

# Die Bedeutung von Polyprolinmotiven für Signalproteine in *Escherichia coli*

**Dissertation**

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

an der Fakultät für Biologie

Ludwig-Maximilians-Universität München



**vorgelegt von**

Anna Magdalena Motz

München, März 2019

**1. Gutachterin:** Prof. Dr. Kirsten Jung

Department Biologie I, Bereich Mikrobiologie, LMU München

**2. Gutachter:** Prof. Dr. Marc Bramkamp

Department Biologie I, Bereich Mikrobiologie, LMU München

Tag der Abgabe: 18.03.2019

Tag der mündlichen Prüfung: 05.06.2019

**Eidesstattliche Erklärung**

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

Magdalena Motz

München, 18.03.2019

**Statutory Declaration**

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

Magdalena Motz

Munich, 18th March 2019

## Zusammenfassung

Aufgrund seiner außergewöhnlichen stereochemischen Eigenschaften ist Prolin im Vergleich zu allen anderen proteinogenen Aminosäuren ein schlechtes Substrat für die ribosomale Peptidyltransferasereaktion. Bestimmte  $X_{(-2)}X_{(-1)}\text{-PP-}X_{(+1)}$  Motive (polyP) in der entstehenden Peptidsequenz induzieren einen Translationsarrest, welcher nur durch die Beteiligung des bakteriellen Elongationsfaktors EF-P bzw. dessen Orthologe in Eukaryonten und Archaea kompensiert werden kann. Während die zugrunde liegenden Mechanismen bereits gut erforscht sind, existieren nur wenige Erklärungsansätze zur eigentlichen Bedeutung von polyP Motiven. Die Motive wirken auf zwei Ebenen: als zeitliches Element während des Translationsprozesses, durch die Beeinflussung der intramolekularen Translationsrate und räumlich als strukturelles und funktionelles Element der Proteine selbst.

Diese Arbeit beinhaltet eine umfassende Analyse zur Verteilung von EF-abhängigen polyP Motiven in *E. coli*. Die Motive sind in einer großen Bandbreite verschiedener Proteinfamilien vertreten. Dabei wird deutlich, dass polyP Motive im Widerspruch zur Translationseffizienz stehen und daher unter negativer Selektion stehen. Tatsächlich sind die Motive unterrepräsentiert. Eine Akkumulation in bestimmten Proteinregionen weist auf deren Beteiligung an Translationsinitiation, Domänenfaltung und kotranslationaler Translokation von Membranproteinen hin. Der Schwerpunkt einer experimentellen Untersuchung der Rolle von polyP Motiven wurde auf Signalproteine und insbesondere die Histidinkinase EnvZ gelegt. Dabei zeigte sich, dass die Bedeutung der Motive individuell verschieden ist und differenziert bewertet werden muss. Auch innerhalb derselben Proteinklassen ist eine Generalisierung nicht möglich. EnvZ trägt zwei polyP Motive, welche Protein-Protein-Interaktionen vermitteln. Ein stark translationsverzögerndes IPPPL Motiv in der cytoplasmatischen Domäne ist unabdingbar für die Homodimerisierung des Proteins und damit dessen Fähigkeit zur Signaltransduktion an den zugehörigen Antwortregulator OmpR, der die Expression der Zielgene *ompC* und *ompF* reguliert. Das einen moderaten Translationsarrest induzierende periplasmatische VVPPA Motiv ist Voraussetzung für die Bindung des Modulatorproteins MzrA. Eine phylogenetische Analyse zeigt eine Koevolution von MzrA und des polyP Motivs in  $\gamma$ -Proteobakterien auf und untermauert damit diese These. Die stark translationsverzögernden polyP Motive der Histidinkinase EvgS, der Diguanylatcyclase DgcC und der Lon Protease sind von unterschiedlicher Wichtigkeit für die jeweiligen Signalwege betreffende phänotypische Ausprägung. Während eine Substitution des entsprechenden polyP Motivs in EvgS zu einer nichtfunktionellen Proteinvariante führt und vermutlich mit der Signalperzeption interferiert, hat der Austausch der Motive in DgcC und Lon Protease keinen erkennbaren Effekt. Eine durch die starken polyP Motive gewährleistete Limitation der Proteinkopienzahl kann im Fall von EnvZ, EvgS und Lon Protease ausgeschlossen werden. Eine globale Regulation des Translationslevels von Proteinen mit polyP Motiv als Antwort auf verschiedene Stressbedingungen findet in *E. coli* K-12 nur bedingt statt. Unter anderem weisen Zellen im VBNC Zustand (*viable but non-culturable*) kein verändertes Expressionsmuster von Proteinen mit polyP Motiv auf, das darauf hinweisen würde. Die Aktivität von EF-P wird in *E. coli* K-12 durch eine OxyR vermittelte transkriptionelle Regulation von *epmA*, welches für das EF-P modifizierende Enzym  $\beta$ -Lysyltransferase kodiert, geringfügig an oxydative Stressbedingungen angepasst.

Diese Arbeit trägt zum besseren Verständnis der Bedeutung von EF-P abhängigen  $X_{(-2)}X_{(-1)}\text{-PP-}X_{(+1)}$  Motiven in *E. coli* bei und offenbart insbesondere am Beispiel von Rezeptorproteinen deren vielseitigen Einsatz als regulatorisches und funktionelles Element. Aufgrund des universalen Charakters von polyP Motiven sind die Schlussfolgerungen dieser Arbeit auch für ein besseres Verständnis deren Auftretens in Proteomen anderer Organismen relevant.

## Summary

Compared to other proteinogenic amino acids, proline is only a poor substrate for the ribosomal peptidyltransferase reaction due to its special stereochemical properties. Certain  $X_{(-2)}X_{(-1)}\text{-PP-}X_{(+1)}$  motifs (polyP) in the nascent peptide chain induce a translation arrest, which can only be compensated by the participation of the bacterial translation elongation factor EF-P or its archaeal or eukaryotic orthologues, respectively. While the underlying mechanisms are already well described, there exist only few explanatory approaches regarding the significance of polyP motifs. The motifs act on two levels: as temporal element during the translation process, by influencing the intramolecular translation rate and spational, by providing structural and functional elements for the protein itself.

This work provides a comprehensive analysis of the distribution of EF-P dependent polyP motifs in *E. coli*. The motifs are represented in a broad range of different protein families. Their occurrence conflicts with translation efficiency, which sets them under selection pressure. Indeed, the motifs are underrepresented. An accumulation in certain protein regions indicates their participation in translation initiation, domain folding and cotranslational translocation of membrane proteins. The focus of the experimental investigation of the role of polyP motifs was layed on signal proteins and especially the histidine kinase EnvZ. The data emphasises, that the role of polyP motifs is individually different and has to be evaluated differentiated. A generalization is not possible, not even for proteins of the same class. EnvZ harbours two polyP motifs, which participate in protein-protein interaction. The cytoplasmic IPPPL Motiv, which induces a strong translation arrest is prerequisite for the homodimerisation of the protein and therefore for its ability to transfer the signal to its cognate response regulator OmpR, that regulates *ompC* and *ompF* expression. The periplasmic VVPPA is indispensable for the binding of the modulator protein MzrA. A phylogenetic analysis supports this hypothesis by revealing a coevolution of MzrA and the polyP motif in  $\gamma$ -proteobacteria. The strong stalling polyP motifs of the histidine kinase EvgS, the diguanylate cyclase DgcC and the Lon protease are of different importance for the phenotypic characteristics of the respective signaling cascades. While the substitution of the polyP motif in EvgS leads to a nonfunctional protein variant and interferes presumably with signal perception, its replacement in DgcC and Lon protease does not have any detectable effect. A role of the polyP motifs in copy number control can be excluded for EnvZ, EvgS and Lon protease. In *E. coli* K-12, polyP motifs are only marginal involved in regulating the translation level of proteins with polyP motif. Amongst others, the expression pattern of proteins in the VBNC state (viable but non-culturable) does not imply any changes. The activity of EF-P is slightly adjusted in response to oxidative stress conditions by OxyR mediated transcriptional regulation of *epmA*, coding for the EF-P modifying enzyme  $\beta$ -lysyltransferase.

This work contributes to a better understanding of the significance of EF-P dependent  $X_{(-2)}X_{(-1)}\text{-PP-}X_{(+1)}$  motifs in *E. coli* and reveals their broad effects on regulation and function, especially in receptor proteins. Due to the universal character of polyP motifs, the conclusions of this thesis are relevant for a better understanding of their occurrence in proteomes of other organisms.

## Nomenklatur

Die Aminosäuren von Proteinsequenzen sind in aufsteigender Zahlenabfolge nummeriert. Dabei werden artifiziell angefügte N-terminale Peptidsequenzen nicht berücksichtigt.

Die Position von Nukleotiden in Gensequenzen wird analog dazu in Relation zum Translationsstart ATG beschrieben (A ist „1“). Nukleotide, die *upstream* des Startcodons liegen werden dementsprechend mit negativen Zahlen beziffert („-1“ für das Nukleotid, welches *upstream* von ATG liegt). Nukleotide, die *downstream* des Stoppcodons liegen werden in fortfolgender Nummerierung mit „+“ gekennzeichnet („+1“ für das erste Nukleotid *downstream* des Stoppcodons).

Gendeletionen sind durch das Symbol  $\Delta$  beschrieben. Für Ersetzungen wird folgende Schreibweise verwendet:  $\Delta$  „*ersetztes Gen*“ (: „*neues Gen*“).

Basen werden im Ein-, Aminosäuren im Ein- und Dreibuchstabencode abgekürzt. Im Falle einer Substitution wird die native Base, bzw. Aminosäure an erster Stelle genannt, gefolgt von ihrer Position und der Ersetzung. Der Austausch von Prolinmotiven wird als XXPPX/XXXXX gekennzeichnet.

Translationale Fusionen werden entsprechend ihrer Zusammensetzung und ihrer N- bzw. C-terminalen Positionierung benannt (z.B. „Protein-LacZ“ oder „6His-Protein“).

Stämme, Vektoren und Oligonukleotide werden wie im Methodenteil aufgeführt durch die Buchstabenfolgen „St“, „V“ und „Pr“ mit anschließender Kennzahl bezeichnet.

EF-P abhängige Prolinmotive wurden im Rahmen dieser Arbeit als  $X_{(-2)}X_{(-1)}-nP-X_{(+1)}$  definiert, wobei X jede beliebige Aminosäure außer P sein kann und  $n \geq 2$  ist.

## Abkürzungsverzeichnis

aa	Aminosäure
AP	Alkalische Phosphatase
COG	<i>Cluster of Orthologous Groups</i>
DGC	Diguanylatcyclase
PAGFP/PAmCherry	Photoaktivierbares GFP/mCherry
GFP	<i>Green Fluorescent Protein</i>
HK	Histidinkinase
HOG	<i>Hierarchical Orthologous Group</i>
HRP	<i>Horseradish Peroxidase</i>
MES	2-(N-Morpholino)-ethansulfonsäure
MOPS	3-(N-Morpholino)-propansulfonsäure
MU	<i>Miller Units</i>
PAGE	Polyacrylamid-Gelelektrophorese
PBS	<i>Phosphate Buffered Saline</i>
PCR	<i>Polymerase Chain Reaction</i>
PDE	Phosphodiesterase
PGL	<i>Propensity of Gene Loss</i>
PMSF	Phenylmethylsulfonylfluorid
PSEC	<i>Propensity of Stalling Effect Change</i>
qRT-PCR	<i>Quantitative Real-Time PCR</i>
RNS	<i>Reactive Nitrogen Species</i>
ROS	<i>Reactive Oxygen Species</i>
RR	<i>Response Regulator</i>
TCE	Trichlorethanol
TAE	TRIS-Acetat-EDTA-Puffer
TMH	Transmembranhelix
TRIS	Tris(hydroxymethyl)-aminomethan
VBNC	<i>Viable but non-Culturable Cells</i>

## Publikationen, die Bestandteil dieser Arbeit sind

### Kapitel 3.1.1 / 3.1.2.2 / 3.1.6:

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(2):e1005987.

### Kapitel 3.4.3.3:

Vilhena C, Kaganovitch E, Grünberger A, Motz M, Forné I, Kohlheyer D and Jung K (2019). **Importance of pyruvate sensing and transport for the resuscitation of viable but nonculturable *Escherichia coli* K-12.** J Bacteriol. pii: JB.00610-18.

### Kapitel 3.2.4 / 3.3.1 / 3.3.2 / 3.3.5:

Motz M and Jung K (2018). **The role of polyproline motifs in the histidine kinase EnvZ.** PLoS One. 13 (6): e0199782.

## **An bereits veröffentlichten Ergebnissen dieser Arbeit mitwirkende Personen**

### **Kapitel 3.1.1 / 3.1.2.2 / 3.1.5 Abschnitt 2 / 3.1.6 / 4.2 / 4.3:**

Konzept: Magdalena Motz

Entwurf der Experimente: Magdalena Motz, Fei Qi, Jürgen Lassak, Kirsten Jung, Dmitrij Frishman

Entwicklung des Regelsatzes zur Vorhersage der „Stärke“ von polyP Motiven: Magdalena Motz, Jürgen Lassak

Bioinformatische Analyse: Fei Qi

Manuskript: Magdalena Motz, Fei Qi, Jürgen Lassak, Kirsten Jung, Dmitrij Frishman

### **Kapitel 3.4.3.3:**

Konzept: Magdalena Motz, Kirsten Jung

Entwurf der Experimente: Magdalena Motz, Cláudia Vilhena, Ignasi Forné, Kirsten Jung

Versuchsdurchführung: Magdalena Motz, Ignasi Forné

Manuskript (betreffende Textabschnitte): Magdalena Motz, Cláudia Vilhena, Ignasi Forné, Kirsten Jung

### **Kapitel 3.2.4/ 3.3.1 / 3.3.2 / 3.3.5:**

Versuchsdurchführung: Magdalena Motz

Betreuung: Kirsten Jung

Manuskript: Magdalena Motz, Kirsten Jung

## INHALTSVERZEICHNIS

Eidesstattliche Erklärung .....	III
Zusammenfassung.....	IV
Summary .....	V
Nomenklatur .....	VI
Abkürzungsverzeichnis .....	VII
Publikationen und Manuskripte, die Bestandteil dieser Arbeit sind .....	VIII
An bereits veröffentlichten Ergebnissen dieser Arbeit mitwirkende Personen .....	IX
<b>1. EINLEITUNG .....</b>	<b>1</b>
1.1 Strukturelle und dynamische Eigenschaften von Prolin .....	2
1.1.1 Stereochemische Eigenschaften von Prolin.....	2
1.1.2 Prolin als Substrat der ribosomalen Peptidyltransferasereaktion.....	5
1.2 Der bakterielle Translations-Elongationsfaktor EF-P .....	5
1.2.1 Struktur und Funktionsweise .....	6
1.2.2 Posttranslationale Modifikation.....	7
1.2.3 Phänotypische Auswirkung einer Deletion von EF-P oder eIF5A .....	9
1.3 Strukturelle und funktionelle Bedeutung von polyP Motiven .....	9
1.3.1 Experimentelle Untersuchungen zur Bedeutung von polyP Motiven in bakteriellen Proteinen.....	11
1.3.2 Mögliche Funktionen von polyP Motiven in Signalproteinen.....	12
1.3.2.1 Histidinkinasen.....	12
1.3.2.1.1 Der Osmo- und Säurestressrezeptor EnvZ.....	14
1.3.2.1.2 Der pH-Stresssensor EvgS.....	15
1.3.2.2 Andere Signalproteine.....	15
1.3.2.2.1 Die Diguanylatcyclase DgcC.....	15
1.3.2.2.2 Die AAA+ Protease Lon.....	16
1.4 Aufgabenstellung.....	18
<b>2. MATERIAL UND METHODEN .....</b>	<b>19</b>
2.1 Material.....	19
2.2 Stämme, Plasmide und Oligonukleotide.....	22
2.3 Kultivierungsverfahren .....	34
2.4 Molekularbiologische und gentechnologische Methoden.....	36
2.4.1 Isolierung von Plasmiden .....	36
2.4.2 Isolierung von genomischer DNA .....	36

---

2.4.3 Isolierung von RNA .....	36
2.4.4 Reverse Transkription .....	36
2.4.5 Polymerasekettenreaktion (PCR) .....	37
2.4.6 DNA-Sequenzanalyse.....	37
2.4.7 Konstruktion von Plasmiden .....	37
2.4.8 Konstruktion von Stämmen .....	37
2.4.9 Kompetente Zellen und Transformation.....	38
2.4.10 Proteinüberexpression .....	39
2.5 Biochemische Methoden.....	39
2.5.1 Zellaufschluss.....	39
2.5.2 Nickelchelat-Affinitätschromatographie .....	40
2.5.3 <i>Size exclusion chromatography</i> (SEC).....	40
2.5.4 Präparation von Membranvesikeln.....	40
2.5.5 Präparation von Proteinen der äußeren Membran .....	41
2.5.6 Präparation von Sphäroplasten.....	41
2.5.7 Präparation von <i>right-side-out</i> Vesikeln.....	41
2.6 Analytische <i>in vitro</i> Methoden .....	41
2.6.1 DNA Gelelektrophorese .....	41
2.6.2 qRT-PCR.....	42
2.6.3 SDS-PAGE.....	43
2.6.4 Isoelektrische Fokussierung.....	43
2.6.5 Western Blot.....	43
2.6.6 <i>Thermophoretic Shift Assay</i> .....	44
2.6.7 Native PAGE und EMSA .....	44
2.6.8 DNA-Affinitätschromatographie.....	45
2.6.9 Oberflächenplasmonenresonanz-Spektroskopie .....	45
2.6.10 Dynamische Lichtstreuungsdetektoren (DLS).....	46
2.6.11 Flüssigchromatographie-Massenspektrometrie (LC-MS) .....	46
2.6.12 $\beta$ -Galactosidase Assay.....	47
2.7 Analytische <i>in vivo</i> Methoden .....	47
2.7.1 Separierung von intakten, lebenden Zellen aus Langzeitkulturen mittels Percoll.....	47
2.7.2 Zellanfärbung .....	48
2.7.3 Durchflusszytometrie .....	48
2.7.4 <i>Bacterial-two-Hybrid Assay</i> .....	48
2.7.5 Porinexpressionsassay .....	49
2.7.6 Erstellung von Wachstumskurven .....	49
2.8 Phänotypische Charakterisierung.....	49
2.8.1 Säureresistenz Assay .....	49
2.8.2 Persistenzassay .....	50
2.8.3 Generierung von VBNCs .....	50
2.8.4 Visuelle Detektion der Cellulosesynthese von Makrokolonien.....	50

2.9 Mikroskopie.....	50
2.9.1 Probenvorbereitung .....	50
2.9.1.1 Agarpads .....	51
2.9.1.2 Ibidi-Kammern.....	51
2.9.2 Zellfixierung.....	51
2.9.3 Kultivierung von Zellen für die Fluoreszenzmikroskopie.....	51
2.9.4 Immunofluoreszenz.....	52
2.9.5 Weitfeldmikroskopie mit „DMi8“.....	52
2.9.6 Hochauflösende Mikroskopie mit „Delta Vision Elite“.....	52
2.9.7 Photoaktivierte Lokalisationsmikroskopie (PALM).....	53
2.10 Bioinformatische Methoden .....	54
2.10.1 Datenbanken.....	54
2.10.2 Auswahl repräsentativer <i>E. coli</i> Stämme.....	54
2.10.3 Clusterformierung von Proteinen .....	55
2.10.4 Unterscheidung zwischen Kernproteom und akzessorischem Proteom .....	55
2.10.5 Identifikation und Klassifizierung von polyP Motiven .....	56
2.10.6 Untersuchungen zur Akkumulation von polyP Motiven .....	57
2.10.7 Sequenzalignment .....	58
2.10.7.1 Sequenzalignment von EnvZ, EvgS, DgcC und Lon Protease.....	58
2.10.7.2 Multiples Sequenzalignment von Proteinclustern.....	58
2.10.8 Rekonstruktion von phylogenetischen Stammbäumen.....	58
2.10.8.1 Phylogenetische Stammbäume von EnvZ, EvgS, DgcC und Lon Protease .....	58
2.10.8.2 Phylogenetischer Stammbaum von weiteren Proteinclustern aus 43 <i>E. coli</i> Stämmen.....	59
2.10.9 <i>Propensity of Gene Loss</i> (PGL) und <i>Propensity of Stalling Effect Change</i> (PSEC) .....	59
2.10.10 Berechnung von Proteinabundanz und Translationseffizienz.....	60
2.10.11 Vorhersage von Bindestellen von Transkriptionsfaktoren .....	60
2.10.11 Modellierung von dreidimensionalen Proteinstrukturen .....	60
2.10.12 Bildanalyse und graphische Darstellung .....	60
<b>3. ERGEBNISSE .....</b>	<b>61</b>
3.1 Verteilung, Konservierung und Evolution von polyP Motiven in <i>E. coli</i> .....	61
3.1.1 Klassifizierung von polyP Motiven.....	61
3.1.2 Häufigkeit von polyP Motiven in <i>E. coli</i> .....	64
3.1.2.1 Frequenz von polyP Motiven in <i>E. coli</i> K-12 MG1655 .....	64
3.1.2.2 Frequenz von polyP Motiven in 43 ausgewählten <i>E. coli</i> Stämmen .....	65
3.1.3 Antikorrelation zwischen dem Auftreten von polyP Motiven und der Translationseffizienz .....	68
3.1.4 Evolution von polyP Motiven .....	68
3.1.5 Ontologie von Proteinen mit polyP Motiv.....	69

---

3.1.6	Akkumulation von polyP Motiven an strategisch günstigen intramolekularen Loci.....	72
3.1.6.1	Gehäuftes Auftreten von polyP Motiven innerhalb der ersten 50 Aminosäuren eines Proteins...	72
3.1.6.2	Akkumulation von polyP Motiven zwischen Domänen.....	73
3.1.6.3	Akkumulation von polyP Motiven in der Nähe von Transmembranhelices .....	73
3.2	Die Rolle von polyP Motiven für Signalproteine .....	75
3.2.1	Die Bedeutung von HPPPG für die Diguanylatcyclase DgcC.....	75
3.2.1.1	Position und Konservierung der polyP Motive in DgcC.....	75
3.2.1.2	Einfluss des putativ „starken“ HPPPG Motivs auf Lokalisation und Kopienzahl von DgcC .....	77
3.2.1.3	Einfluss des putativ „starken“ HPPPG Motivs auf die Proteinfunktionalität von DgcC.....	80
3.2.2	Die Bedeutung von VGPPG für die Lon Protease.....	83
3.2.2.1	Position und Konservierung der polyP Motive in Lon.....	83
3.2.2.2	Einfluss des putativ „starken“ VGPPG Motivs auf Lokalisation und Kopienzahl von Lon .....	85
3.2.2.3	Einfluss des putativ „starken“ VGPPG Motivs auf die Proteinfunktionalität von Lon.....	85
3.2.3	Die Bedeutung von NYPPD für die Histidinkinase EvgS.....	87
3.2.3.1	Position und Konservierung der polyP Motive in EvgS .....	87
3.2.3.2	Einfluss des putativ „starken“ NYPPD Motivs auf Lokalisation und Kopienzahl von EvgS .....	89
3.2.3.3	Einfluss des putativ „starken“ NYPPD Motivs auf die Proteinfunktionalität von EvgS.....	91
3.2.4	Die Bedeutung von IPPPL für die Histidinkinase EnvZ .....	92
3.2.4.1	Position und Konservierung der polyP Motive in EnvZ .....	92
3.2.4.2	Einfluss des putativ „starken“ IPPPL Motivs auf Lokalisation und Kopienzahl von EnvZ.....	94
3.2.4.3	Einfluss des putativ „starken“ IPPPL Motivs auf die Proteinfunktionalität von EnvZ.....	94
3.3	Die Rolle von EnvZ polyP Motiven in der Proteininteraktion.....	97
3.3.1	Die Beteiligung des IPPPL Motivs an der Homodimerisierung von EnvZ .....	97
3.3.2	Beteiligung von VVPPA und IPPPL an der Bindung des Modulatorproteins MzrA .....	99
3.3.3	Die Bindungsaffinität von MzrA zu EnvZ in Abhängigkeit zur Mediumosmolarität .....	99
3.3.4	Visualisierung von EnvZ und MzrA auf Einzelzellebene .....	101
3.3.4.1	Konstruktion und Charakterisierung von Stämmen mit translationaler MzrA-PAmCherry und EnvZ-PAGFP/mNeonGreen Fusion.....	101
3.3.4.2	Fluoreszenzmikroskopische Visualisierung von MzrA-PAmcherry und EnvZ-PAGF/mNeonGreen.....	103
3.3.4.3	Hochauflösende Lokalisation von MzrA-PAmcherry und EnvZ-PAGFP/-mNeonGreen.....	105
3.3.4.3.1	Größe und Lokalisation von MzrA-PAmcherry und EnvZ-mNeonGreen Proteincluster ...	105
3.3.4.3.2	Kolokalisation von MzrA-PAmCherry und EnvZ-mNeonGreen .....	107
3.3.5	Die phylogenetische Analyse des Auftretens der EnvZ polyP Motive weist auf eine Koevolution mit MzrA hin .....	109
3.4	Regulation der Expression und Aktivität von EF-P.....	112
3.4.1	Hinweise auf eine transkriptionale Regulation unter verschiedenen Stressbedingungen .....	114

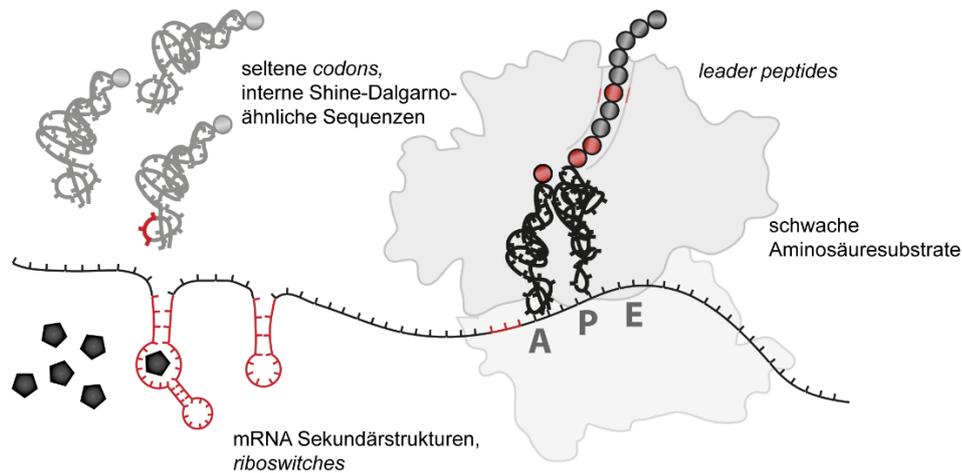
3.4.1.1 Translationslevel eines CadC-LacZ Reporterkonstruktes unter verschiedenen Stressbedingungen .....	115
3.4.1.2 Translationslevel von EF-P und EpmA unter verschiedenen Stressbedingungen.....	115
3.4.1.2.1 Etablierung eines Westernblot-Protokolls zur Immundetektion von EF-P, EpmA und EpmB .....	115
3.4.1.2.2 Auswirkung von oxydativem Stress auf EF-P und EpmA Proteinlevel.....	117
3.4.1.3 Auswirkung von Nährstofflimitation auf die Genexpression von <i>efp</i> , <i>epmA</i> und <i>epmB</i> .....	117
3.4.1.3.1 Beteiligung von OxyR an der Transkription von <i>epmA</i> .....	119
3.4.2 Analyse der Bindeaffinität von OxyR an $P_{epmA}$ .....	119
3.4.2.1 Überexpression und Reinigung von OxyR-Varianten.....	120
3.4.2.2 Bindeaffinität von gereinigtem OxyR zu $P_{epmA}$ .....	122
3.4.2.3 Identifizierung von $P_{epmA}$ bindenden Proteinen mittels DNA-Affinitätschromatographie .....	125
3.4.2.3.1 Bindung von NagC, PurR und PgrR an $P_{epmA}$ .....	125
3.4.2.3.2 Verifizierung des Einflusses von NagC, UlaR und PgrR auf die Expression von <i>efp</i> und <i>epmA</i> .....	127
3.4.3 Expressionsmuster von Proteinen mit polyP Motiv in VBNCs.....	128
3.4.3.1 Etablierung eines Verfahrens zur Separierung von viablen <i>E. coli</i> Zellen aus Langzeitkulturen.....	129
3.4.3.2 Einfluss von EF-P auf das Erreichen des VBNC Zustands .....	131
3.4.3.3 Zusammenhang zwischen der EF-P Aktivität und der Anzahl von Proteinen mit polyP Motiv im VBNC Zustand.....	131
<b>4. DISKUSSION .....</b>	<b>138</b>
4.1 Proteine mit polyP Motiv sind hoch divers.....	138
4.2 polyP Motive stehen unter Selektionsdruck.....	139
4.3 PolyP Motive haben verschiedene kotranslationale Funktionen.....	139
4.4 Die Bedeutung von polyP Motiven in den Signalproteinen EnvZ, EvgS, DgcC und Lon Protease .....	141
4.4.1 PolyP Motive in Signalproteinen sind unabhängig von ihrer „Stärke“ unterschiedlich stark konserviert .....	142
4.4.2 Die native Proteinkopienzahl wird nicht durch polyP Motive limitiert.....	142
4.4.3 PolyP Motive sind für die Proteinfunktionalität von EvgS und EnvZ von essentieller Bedeutung... ..	145
4.4.4 Die polyP Motive von EnvZ vermitteln Protein-Protein-Interaktionen.....	146
4.5 Die Aktivierung von EF-P ist nicht Teil einer zentralen Stressantwort.....	148
<b>5. AUSBLICK.....</b>	<b>151</b>
<b>5. DANKSAGUNG.....</b>	<b>153</b>
<b>6. REFERENZEN .....</b>	<b>154</b>
<b>7. ANHANG.....</b>	<b>177</b>

## 1. EINLEITUNG

Die Proteinbiosynthese gehört zu den Grundvoraussetzungen des Lebens und ist durch das finale Argument im 1958 von Crick formulierten zentralen Dogma der Molekularbiologie beschrieben<sup>1</sup>. Es handelt sich um einen universellen, sehr stark konservierten und hoch effizienten Mechanismus, der im Laufe der Evolution durch Selektion maximal optimiert wurde<sup>2</sup>. Die Translation der durch mRNA kodierten Information in eine Peptidsequenz findet an den Ribosomen statt und gliedert sich in die vier Phasen Initiation, Elongation, Termination und Recycling, welche durch zahlreiche, konzertierende Translationsfaktoren begleitet werden<sup>3,4</sup>. Während der Elongationsphase werden in Bakterien im Durchschnitt 4-22 Codons pro Sekunde prozessiert<sup>5,6</sup>. Jeder Elongationszyklus setzt sich zusammen aus der Bindung der Aminoacyl-tRNA an das zugehörige Codon in der *A-site* des Ribosoms, der Peptidbindungsreaktion zwischen Aminoacyl-tRNA und in der *P-site* lokalisierten Peptidyl-tRNA, sowie der Translokation der beiden tRNAs (von *A-* zu *P-site* und von *P-* zu *E-site*)<sup>3</sup>. Der Prozess wird von verschiedenen Translations-Elongationsfaktoren begleitet. In Prokaryonten unterstützt EF-Tu mit dem assoziierten *nucleotide exchange factor* EF-Ts die Bindung der Aminoacyl-tRNA an das Ribosom, während EF-G die Translokation der tRNA vermittelt<sup>7,8</sup>. Einer Dislokation der Peptidyl-tRNA wird durch die stabilisierende Wirkung von EF-P entgegengewirkt<sup>9</sup>. Obwohl EF-Tu einen ausgleichenden Effekt auf Unterschiede bezüglich der Verfügbarkeit der verschiedenen tRNAs hat und EF-P ungünstige Stereometrien kompensiert, ist die Elongation kinetisch variabel und die individuelle Translationsrate der einzelnen Codons von mehreren Faktoren abhängig. Während in einigen Fällen stereo- und elektrochemische Eigenschaften von funktionell und strukturell essentiellen Aminosäuren zugrunde liegen (z. B. Prolin)<sup>10-12</sup>, bietet die Verzögerung der Peptidyltransferasereaktion auch die Möglichkeit einer dynamischen Regulation der Translation. Diese ist unter anderem Voraussetzung für die Anpassung der Proteinexpression oder die Anreicherung von Zwischenprodukten der Proteinbiosynthese<sup>13,14</sup>. Die kotranslationale Faltung von Proteinen erfordert eine genaue zeitliche Taktung, bei der sich schnelle und langsame Phasen abwechseln und auf die parallel stattfindenden Strukturierungsvorgänge abgestimmt sind<sup>15</sup>. Auch die kotranslationale Translokation von Membranproteinen setzt ein Zeitfenster voraus, währenddessen der Vorgang stattfinden kann. Eine langsame Anlaufphase der Translation in den ersten 50 Aminosäurepositionen wirkt einer Akkumulation von Ribosomen an derselben mRNA entgegen<sup>16,17</sup>. In den letzten Jahrzehnten wurden zahlreiche, translationsverzögernde Elemente beschrieben, die in diesem Zusammenhang stehen (Abb. 1). Dazu zählen seltene Codons, deren zugehörige tRNAs aufgrund ihrer im Vergleich zu anderen tRNAs signifikant schlechteren Verfügbarkeit mit zeitlicher Verzögerung an das Ribosom rekrutiert werden<sup>18,19</sup>. Die Bindungskräfte in Sekundärstrukturen der mRNA, wie zum Beispiel *ribo switches*, verhindern einen konstanten Lesefluss<sup>20-23</sup>. Interne, Shine-Dalgarno-ähnliche Sequenzabschnitte induzieren Translationsstopps<sup>24</sup>. Eine Interaktion der neusynthetisierten Peptidkette mit dem ribosomalen *exit tunnel* hemmt die Translokation und damit das Anknüpfen weiterer Aminosäuren und wird unter anderem durch spezifische *leader peptides*, wie *secM*, *tnaC* oder *ermCL*, vermittelt<sup>25-29</sup>.

Schwerpunkt dieser Arbeit sind konsekutive Proline, deren Eingliederung in die entstehende Peptidkette einen signifikanten Translationsarrest hervorruft<sup>30,31</sup>. Ihre biologische Relevanz ist bislang nur bruchteilhaft verstanden. Der Fokus der Forschung liegt momentan vor allem auf der Aufklärung der dem Translationsarrest zugrunde liegenden molekularen Mechanismen und nur hintergründig bis gar nicht auf dem Sinn und Zweck des Auftretens der Motive selbst. Diese Arbeit soll nun einen Beitrag zu deren Deutung leisten. Auch unter dem Aspekt, dass die

weite Verbreitung von polyP Motiven und die hohe Spezifität des assoziierten Translationselongationsfaktors EF-P für die Entwicklung von neuartigen Antibiotika oder im Fall des eukaryontischen Orthologs eIF5A, in der Etablierung neuer Tumormarker oder Krebstherapeutika von Bedeutung sein könnte.



**Abb. 1:** Schematische Übersicht von translationsverzögernden Elementen und deren Wirkungsort (rot gekennzeichnet).

## 1.1 Strukturelle und dynamische Eigenschaften von Prolin

### 1.1.1 Stereochemische Eigenschaften von Prolin

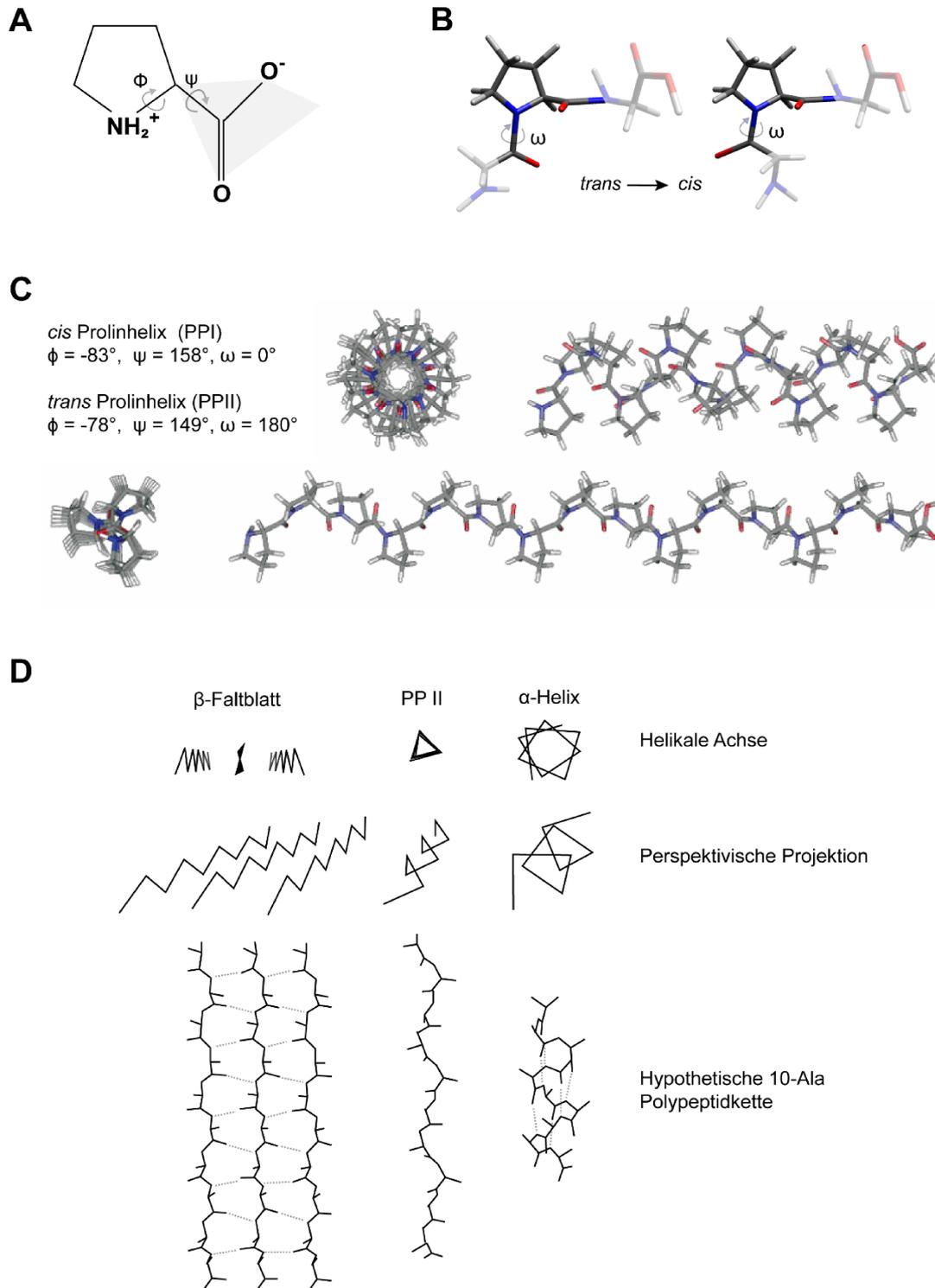
Prolin nimmt aufgrund seiner kovalent an das  $\alpha$ -N gebundenen Seitengruppe eine Sonderstellung unter den 20 proteinogenen Standardamino­säuren ein. Der dihedrale  $\Phi$  Winkel ist zwischen  $-68^\circ$  und  $-75^\circ$  fixiert, während der  $\Psi$ -Winkel eingeschränkt flexibel ist (Abb. 2A)<sup>32</sup>. Auch die Freiheitsgrade von Peptidbindungen werden durch den cyclischen Pyrrolidinring eingeschränkt, wobei das Stickstoffatom durch eine sekundäre Amidbindung an die benachbarte Aminosäure gebunden ist<sup>33-35</sup>. Das als *proton hole*<sup>36</sup> bezeichnete Phänomen des fehlenden Protons der Amingruppe hat mehrere Folgen.

Die alternative Ladungsverteilung führt zu einer verlängerten Peptidbindung (1,36 Å für X-Pro statt 1,33 Å für X-X)<sup>37</sup> und zu einer vergleichsweise niedrigen Aktivierungsenergie für eine *cis/trans* Isomerisierung (13 kcal/mol statt 20 kcal/mol)<sup>38-40</sup>. Die Minima der Konformationsenergie liegen dabei bei  $-55^\circ$  und  $+145^\circ$ <sup>41</sup>. Die Wahrscheinlichkeit, dass die anschließende Peptidbindung in *trans* Konformation vorliegt ist ca. 100 mal geringer, als bei anderen Aminosäuren ( $p = 0,1 - 0,3$  statt 0,001)<sup>42</sup>. Prolin ist die einzige Aminosäure, die sowohl *cis* oder *trans* Konformation einnimmt (Abb. 2B)<sup>33, 32, 43</sup>. Tatsächlich liegen ca. 90% aller Proline in Proteinstrukturen in der *trans* Konformation vor<sup>44, 32</sup>. Eine Sequenz von konsekutiven Prolinen nimmt entweder die Form einer

rechtsgewundenen *cis* Prolinhelix (PPI) ein, oder die einer linksgewundenen *trans* Prolinhelix II (PPII). Eine Mindestanzahl von drei konsekutiven Prolinen oder prolinreiche Sequenzabschnitte, in welchen die Aminosäure an jeder dritten Position steht, führen zur Ausbildung der PPII (Abb 2C und 2D)<sup>45, 46</sup>. Dabei ist die PPII nach  $\alpha$ -Helix und  $\beta$ -Faltblatt die am häufigsten in gefalteten Proteinen anzufindende Sekundärstruktur<sup>47, 48</sup>.

Prolin weist nicht nur selbst eine begrenzte Flexibilität auf, sondern schränkt auch die Konformation benachbarter Aminosäuren in Peptidsequenzen ein<sup>33, 41</sup>. Während die  $\Phi$  und  $\Psi$  Winkel benachbarter Aminosäuren im Normalfall weitgehend unabhängig voneinander sind, limitiert Prolin den zur Verfügung stehenden Raum durch sterische Konflikte zwischen der CH-N Bindung und den NH und CH<sub>2</sub> Gruppen der folgenden Aminosäure<sup>41</sup>. Die flankierenden Aminosäuren haben ihrerseits einen Einfluss auf die *cis/trans* Isomerie des Prolins<sup>33</sup>. Die Interkonversionsraten der X-Pro Peptidbindung variieren in Abhängigkeit der folgenden Aminosäure um das bis zu 10-fache, wobei die Isomerisierungskonstante  $k_{c \rightarrow t}$  umso kleiner ist, je unflexibler die Seitenkette ist (Gly > Ala > Val)<sup>42</sup> und minimale Werte im Fall von aromatischen Aminosäuren einnimmt<sup>49</sup>. Während Aspartat und Asparagin das Gleichgewicht Richtung *cis* Isomer verschieben, begünstigen positiv geladene Aminosäuren die *trans* Konformation<sup>50</sup>. Darüber hinaus wiesen mehrere Studien nach, dass auch weiter entfernt liegende Aminosäuren in einem Abstand von bis zu vier Positionen die Konformation von Prolin beeinflussen<sup>50, 31, 51</sup>. Da die Aminosäuren eines Peptids unter ständiger gegenseitiger Wechselwirkung stehen, lässt sich die endgültige Konformation von Prolin nur durch das Zusammenspiel aller umliegenden Aminosäuren erklären.

Das fehlende Proton der Amingruppe von Prolin und die daraus resultierende Ladungsverteilung in X-Pro Bindungen hat eine erhöhte Elektronegativität der C=O Gruppe der benachbarten Aminosäure zur Folge. Damit steigt deren Tendenz, starke Wasserstoffbrückenbindungen zu bilden<sup>52</sup>. Aus diesem Grund ist Prolin häufig in  $\beta$ - und  $\gamma$ -turns anzufinden<sup>53, 54</sup>. Innerhalb von  $\alpha$ -Helices induziert Prolin einen strukturellen Knick im Winkel von 20-30°<sup>55, 56</sup>, da die X-Pro Peptidbindung keine Wasserstoffbrückenbindung mit benachbarten Carbonylgruppen bilden kann. Folglich steigt in Helices das Potential der vier Aminosäuren entfernt liegenden Carbonylgruppe, alternative Bindungsmuster einzugehen<sup>34</sup>.



**Abb. 2:** Die Rolle von Prolin in der Strukturierung von Proteinen. (A) Darstellung von Prolin als Valenzstrichformel. (B) *Cis/trans*-Isomerisierung der X-Pro Peptidbindung (modifiziert aus <sup>57</sup>). (C) PPI und PPII in Vorder- und Seitenansicht (modifiziert aus <sup>58</sup>). (D) Gegenüberstellung der drei Hauptstrukturen  $\beta$ -Faltblatt, PPII und  $\alpha$ -Helix (modifiziert aus <sup>46</sup>).

### 1.1.2 Prolin als Substrat der ribosomalen Peptidyltransferasereaktion

Aufgrund seiner außergewöhnlichen Rigidität ist Prolin ein schlechtes Substrat für die ribosomale Peptidyltransferasereaktion<sup>9</sup>. Zwei konsekutive Proline haben ein ausreichend großes Potential, die Reaktion um mindestens das 60-fache zu verlangsamen und damit einen Translationsarrest zu induzieren<sup>59, 31, 9</sup>. Während elektrochemische Eigenschaften nur eine untergeordnete Rolle spielen, ist die Stereometrie von Prolin ausschlaggebend für eine erschwerte Positionierung der Peptidyl-PP-tRNA im katalytischen Zentrum des Ribosoms<sup>9</sup>.

Dabei werden vermutlich sowohl der dem nukleophilen Angriff der Peptidbindungsreaktion zugrunde liegende Protonentransfer, als auch die Ausbildung von Ladungen beeinträchtigt<sup>9, 60</sup>. Da sich umliegende Aminosäuren und Prolin gegenseitig in ihrer Konformation beeinflussen, hängt das Ausmaß der Dislozierung von der Sperrigkeit der flankierenden Peptidsequenz ab<sup>30, 61, 59</sup>. Auch die Interaktion der bereits synthetisierten und sich im *exit tunnel* befindenden Peptidsequenz mit ribosomalen Proteinen muss in Betracht gezogen werden. Zum Beispiel schwächen Aminosäuren wie Cys oder Thr *upstream* von drei konsekutiven Proline den Translationsarrest ab, während Arg und His ihn verstärken<sup>61</sup>. Der Einfluss der Aminosäuren ist umso größer, je geringer die Entfernung zu Prolin<sup>30, 61</sup>. Darüberhinaus sind auch intramolekulare Position der konsekutiven Proline und Proteinexpressionsrate von Bedeutung<sup>62, 63</sup>. Vermutlich tragen Zugkräfte einer kotranslationalen Prozessierung des bereits synthetisierten Peptidteils dazu bei, dass konsekutive Proline einen umso kleineren Effekt haben, je weiter entfernt sie vom N-Terminus eines Proteins liegen<sup>64</sup>.

Zur Vereinfachung werden polyP Motive in dieser Arbeit als eine Abfolge konsekutiver Proline mit flankierenden Aminosäuren definiert:  $X_{(-2)}X_{(-1)}nPX_{(+1)}$ . Dabei gilt  $n \geq 2$ .  $X_{(-2)}$ ,  $X_{(-1)}$  und  $X_{(+1)}$  können jede beliebige Aminosäure außer Prolin sein.

## 1.2 Der bakterielle Translations-Elongationsfaktor EF-P

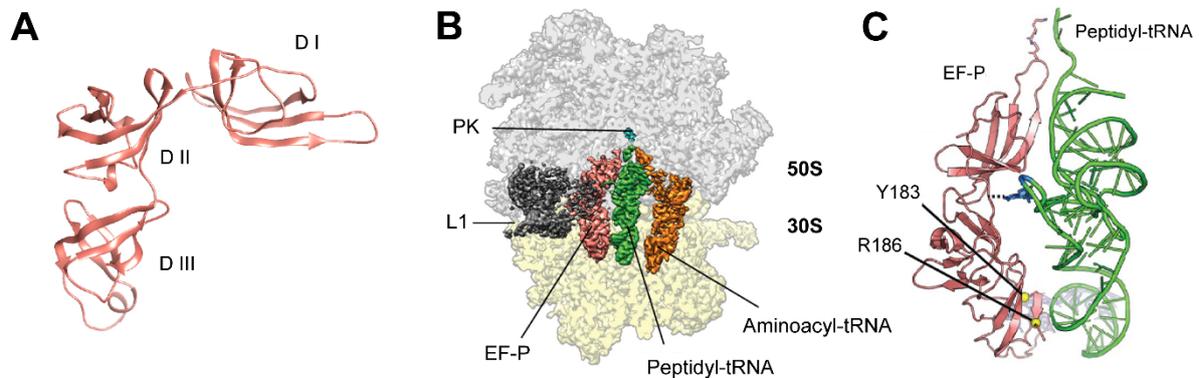
Im Fall eines durch polyP Motive induzierten Translationsarrests bindet der bakterielle Translations-Elongationsfaktor EF-P reversibel zwischen P- und E-*site* des Ribosoms und beschleunigt die Peptidyltransferasereaktion auf ein annähernd reguläres Maß<sup>65, 9, 66-68</sup>. Dennoch ist die Reaktivität im Vergleich zu anderen Peptidbindungen messbar kleiner<sup>31, 69</sup>. Die Reaktion wird von keinem anderen bekannten Translationsfaktor beeinflusst<sup>70</sup>. EF-P ist in Struktur und Funktionalität stark konserviert<sup>70</sup>. Mit Ausnahme einiger symbiontischer Bakterien mit reduzierter Genomgröße, findet sich der Translationsfaktor, bzw. dessen vergleichbar wirkende eukaryontische und archaeale Orthologe eIF5A und aIF5A, in allen lebenden Organismen<sup>71, 45, 72-75</sup>.

### 1.2.1 Struktur und Funktionsweise

EF-P imitiert tRNA in Form und Größe. Der Translations-Elongationsfaktor besteht aus drei  $\beta$ -Fasstrukturen, wobei die Domänen II und III sehr ähnlich strukturiert sind (Abb. 3A) <sup>67, 66, 65</sup>. eIF5A und aIF5A haben im Gegensatz dazu nur zwei Domänen, welche allerdings eine hohe Übereinstimmung zu den entsprechenden EF-P Domänen aufweisen <sup>76, 77</sup>.

EF-P bindet im Fall eines durch polyP Motive induzierten Translationsarrests reversibel zwischen P-*site* und E-*site* im Reaktionszentrum des Ribosoms <sup>9</sup>. Inwiefern die damit einhergehende Stabilisierung der Peptidyl-tRNA die Entropie der ribosomalen Peptidyltransferasereaktion positiv beeinflusst, wurde am Beispiel von *E. coli* aufgeklärt (Abb. 3B) <sup>9</sup>. EF-P berührt beide ribosomalen Untereinheiten. Dabei ist die N-terminale Domäne des Translationsfaktors nahe des Aminoacyl-Akzeptors der Peptidyl-tRNA und die C-terminale Domäne in nächster Nähe der Anticodon-enthaltenden Haarnadelstruktur lokalisiert <sup>78, 9</sup>. Die Domäne II interagiert direkt mit dem ribosomalen Protein L1, welches Translokation und Entlassung der tRNA vermittelt <sup>78</sup>. Für eine Spezifität von EF-P gegenüber Prolin spricht, dass die 9 Nukleotide lange D-*stem* Sequenz der tRNA<sup>Pro</sup>, welche in dieser Form sonst nur in tRNA<sup>Met</sup> vorkommt, *in vitro* eine wichtige Erkennungsstruktur darstellt <sup>79</sup>. Eine konservierte Aminosäure an der Spitze von Domäne I ist in fast allen bislang untersuchten Organismen durch eine einzigartige, posttranslationale Modifikation verlängert <sup>72</sup>. Am Beispiel von *E. coli* konnte gezeigt werden, dass die  $\beta$ -Hydroxylysylierung von K34 eine Überbrückung der Distanz zum Reaktionszentrum ermöglicht und mutmaßlich die Orientierung des Prolinesters in der P-*site* und die Organisation elektrostatischer Interaktionen optimiert <sup>9</sup> (Abb. 3C). Dabei interagiert  $\beta$ -Hydroxylysyl-Lys-34 mit dem CCA Ende der tRNA in der P-*site* und konstruiert eine alternative Konformation der daran gekoppelten Peptidkette <sup>80</sup>. Da EF-P in unmodifizierter Form ebenfalls eine geringe Verbesserung der Translationsrate herbeiführt, wird vermutet, dass der Translationsfaktor auch das Grundgerüst der tRNA stabilisiert <sup>78, 81</sup>. Zahlreiche Beispiele zeigen jedoch, dass die posttranslationale Modifikation Voraussetzung für eine vollständige Aktivität von EF-P und dessen Orthologe ist <sup>82, 83, 30, 84</sup>.

*In vivo* Studien zeigten, dass etwa 30% aller EF-P Moleküle in *E. coli* zeitgleich an Ribosomen gebunden sind <sup>85</sup>. Das Verhältnis zwischen Anzahl an EF-P Molekülen und 70S Ribosomen beträgt in *E. coli* 1:4-10 <sup>86, 87</sup>. Während ein durchschnittlicher Elongationszyklus eine Dauer von 50 ms hat, verbleibt EF-P nur ca. 7 ms am Ribosom <sup>85</sup>. Die Voraussetzung für die Rekrutierung von EF-P an das Ribosom ist dabei eine unbesetzte E-*site* <sup>85</sup>. Da EF-P weit öfter im Komplex mit Ribosomen vorliegt, als es die Abundanz von Pro-Pro Motiven in den zeitgleich vorliegenden mRNA Transkripten implizieren ließe, wird vermutet, dass EF-P generell durch Translationsverzögerungen rekrutiert wird <sup>85</sup>. Analog dazu wurde für EF-P und eIF5A zusätzlich zur Unterstützung der Elongation an polyP Motiven auch eine gewisse Beteiligung an der Peptidbindung von andersartigen, ebenfalls ungünstigen Aminosäuresubstraten und der Translationstermination beschrieben <sup>88, 51, 89</sup>.



**Abb. 3:** Der bakterielle Translations-Elongationsfaktor EF-P erhöht die Effizienz der ribosomalen Peptidyltransferasereaktion im Fall eines durch poyP Motive induzierten Translationsarrestes. (A) Modell der 3D Struktur von EF-P aus *E. coli* mit Domänen I-III (DI-III). (B) Cryoelektronenmikroskopische Rekonstruktion von arretierten Ribosomen (modifiziert aus <sup>80</sup>). (C) Relative Positionierung von EF-P und tRNA<sup>Pro</sup>. Die für Interaktion mit der Peptidyl-tRNA relevanten Aminosäuren Y183 und R186 von EF-P sind hervorgehoben (modifiziert aus <sup>80</sup>).

L1: 50S ribosomales Protein L1 (RplA); PK: Peptidkette

### 1.2.2 Posttranslationale Modifikation

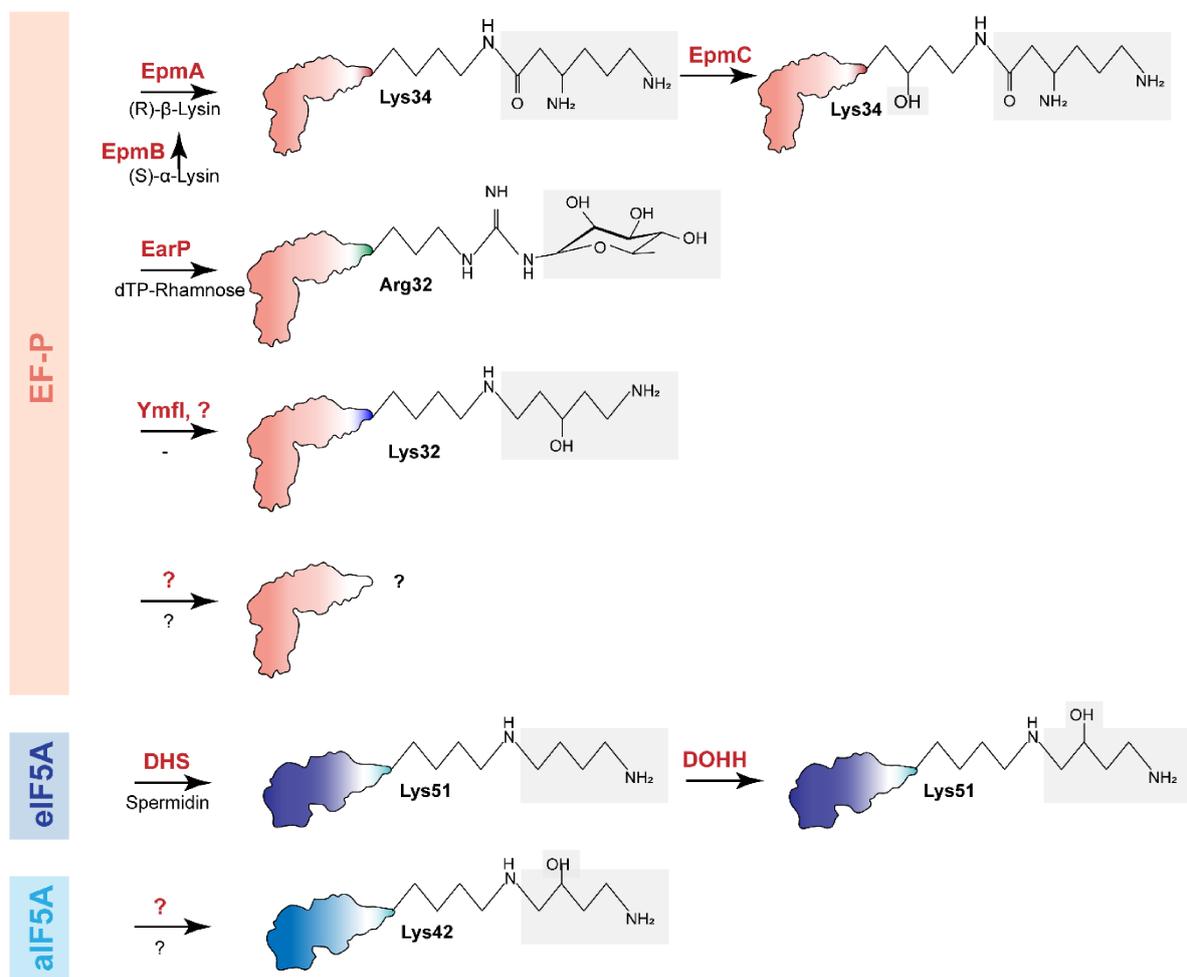
Obwohl die Wirkung von EF-P und eIF5A, bzw. aIF5A auf die Translationseffizienz universell ist und die Proteine die Position einer konservierten Aminosäure gemeinsam haben, unterscheidet sich dennoch die Art von deren Modifikation zwischen den einzelnen Spezies. Interessanterweise handelt es sich dabei immer um sehr ungewöhnliche, in ihrer Form einzigartige Verbindungen (Abb. 4).

In Eukaryonten wird ein konserviertes Lysin durch die Anbindung von Spermidin durch das Enzym Deoxyhypusinsynthase (DHS) und anschließende Hydroxylierung durch die Deoxyhypusinehydroxylase DOHH in Hypusin (N<sup>ε</sup>-(4-Amino-2-Hydroxybutyl)-Lysin) umgewandelt <sup>90, 82, 44</sup>. In Archaea liegt aIF5A ausschließlich in deoxyhypusinylierter Form vor <sup>84</sup>. Der Modifikationsweg ist noch unbekannt.

In Prokaryonten liegen im Gegensatz dazu diverse Modifikationssysteme vor. In 26% aller Bakterien - vor allem  $\gamma$ -Proteobakterien - wird die Aktivität von EF-P durch eine  $\beta$ -Lysin Modifikation des konservierten Lys34 gewährleistet. Die Aminomutase EpmB wandelt  $\alpha$ -Lysin in  $\beta$ -Lysin um, welches durch die Lysyltransferase EpmA an EF-P angefügt wird. Eine Hydroxylierung durch EpmC komplettiert die Modifikation <sup>91, 92</sup>.

In 9% aller Bakterien (darunter *Pseudomonas*, *Bordetella*, *Shewanella*, *Neisseria*) findet eine durch die Glycosyltransferase EarP vermittelte Rhamnosylierung eines konservierten Arginins (Arg32) statt<sup>93-95</sup>. In *Bacillus subtilis* liegt eine 5-Aminopentanol Modifikation an Lys32 vor<sup>96,97</sup>.

Ein großer Teil der Bakterien besitzt keine Orthologen der bereits bekannten Modifikationsenzymen. Die Suche nach neuen Modifikationssystemen ist damit noch nicht abgeschlossen. Es gibt Hinweise darauf, dass Bakterien mit sehr wenigen polyP Motiven nicht auf eine Modifikation angewiesen sind, bzw. alternative Methoden entwickelt haben, um Translationsarreste zu vermeiden. Ein Beispiel dafür ist *Lactobacillus jensenii*, in welchem ValS das einzige Protein mit drei konsekutiven Prolinen ist<sup>98</sup>.



**Abb. 4:** Konservierte Aminosäuren an der Spitze von Domäne I von EF-P, bzw. eIF5A und aIF5A, werden posttranslational durch einzigartige Modifikationen verlängert. Schematische Übersicht der bekannten Modifikationssysteme.

### 1.2.3 Phänotypische Auswirkung einer Deletion von EF-P oder eIF5A

Eine Deletion von *efp*, bzw. *a/eIF5A*, hat mit steigender Komplexität der Organismen und in Korrelation zur Anzahl der polyP Motive einschränkende bis massive Auswirkungen auf die Zellviabilität. Pleiotrope phänotypische Merkmale illustrieren die Bandbreite und Relevanz der EF-P abhängigen Proteine.

EF-P oder das modifizierende Enzym EpmA sind Voraussetzung für die Virulenz von *Shigella flexneri*. Die Abwesenheit eines der beiden Proteine führt zu einer Reduktion der Proteinlevel verschiedener Virulenzregulatoren und damit zu starken Beeinträchtigungen der Fähigkeit, epitheliale Zellen zu invadieren<sup>99</sup>. Die Deletion von *efp* in *E. coli* und *Pseudomonas aeruginosa* hat eine stark reduzierte Wachstumsrate zur Folge<sup>100</sup>. EF-P spielt darüberhinaus eine Rolle in der Virulenz und Stressresistenz von *Salmonella enterica*<sup>101</sup>. In Eukaryonten existieren mehrere Isoformen von eIF5A, welche unterschiedlich stark exprimiert sind. In multizellulären Organismen wie dem menschlichen, ist eIF5A-1 allgegenwärtig, während eIF5A-2 nur in einigen Geweben nachweisbar ist<sup>102, 103</sup>. Hohe Expressionslevel von eIF5A-1 und eIF5A-2 werden mit verschiedenen Krebsarten in Verbindung gebracht und beide Proteine werden als Tumormarker diskutiert<sup>104-108</sup>. Eine homozygote Deletion der genomischen *loci* von eIF5A-1 oder des modifizierenden Enzyms DHS ist in Mäusen, *Caenorhabditis elegans* und *Drosophila melanogaster* lethal<sup>109, 110</sup>.

### 1.3 Strukturelle und funktionelle Bedeutung von polyP Motiven

Prolin findet sich in einem weiten Spektrum von Proteinstrukturen und ist Teil von *loops*, *turns* oder zufälligen *coils*,  $\beta$ -Faltblattstrukturen bis hin zu  $\alpha$ -Helices<sup>33</sup>. Tabelle 1 fasst experimentelle Untersuchungen zu Prolinen in verschiedenen Proteinen und Organismen zusammen. Prolin wird eine Funktion als Helix-Initiator zugeschrieben<sup>33, 111, 112</sup>. Obwohl eine Positionierung von Prolin in Transmembranhelices energetisch ungünstig ist, existieren einige gut beschriebene Beispiele dafür<sup>34</sup>. Nach außen präsentierte Carbonylsauerstoffatome und die X-Pro Peptidbindung bieten eine putative Angriffsfläche für positiv geladene Liganden<sup>113, 114</sup>. Die begrenzte Beweglichkeit von Prolin und die hydrophobe Oberfläche des Pyrrolidinrings reduzieren zudem einen üblicherweise aus der Bindung resultierenden, ungünstigen Entropieverlust<sup>115</sup>. Einer aktiven, durch spezifische Prolin-Isomerasen (PPIasen) vermittelten *cis/trans* Isomerisierung von X-Pro Bindungen wird eine Funktion im Transportmechanismus von Proteinkanälen zugeschrieben<sup>34, 116, 40</sup>. Analog dazu wird eine Beteiligung von Prolin an der stereoisometrischen Aktivierung bestimmter Proteine vermutet<sup>117, 57</sup>. Auch eine Rolle im Schutz gegen unspezifische Proteolyse wurde bereits diskutiert<sup>43</sup>.

Andere Mutmaßungen gehen auf den durch Prolin induzierten Knick in  $\alpha$ -Helices zurück, welcher eine strukturelle Heterogenität und damit die gegenseitige Interaktion von Transmembranhelices ermöglichen könnte<sup>34</sup>. Die PPII ist häufig an den Rändern von  $\alpha$ -Helices, in Linkerregionen oder Interaktionsoberflächen anzufinden<sup>46</sup>. Die unflexible Konformation der Helix und deren oft exponierte Lage hat eine niedrige Bindungsenergie zur Folge. Kurze, prolinreiche Motive wurden bereits als Ziele von verschiedenen proteinbindenden Domänen identifiziert. Dazu zählen zum Beispiel SH3 und WW Domänen<sup>118, 119</sup>. Oft ist die PPII Teil von Protein- oder Nukleinsäurebindestellen<sup>120, 46, 121</sup>.

**Tab. 1:** Übersicht von Proteinen, in welchen die Relevanz ausgewählter Proline mittels SDM (*site directed mutagenesis*) bereits experimentell beschriebenen ist. Art und Position von etwaigen polyP Motiven sind in der Spalte „polyP“ zusammengefasst. P = Lokalisierung der substituierten Aminosäure; TM = transmembran; C = cytosolisch (Referenzen: Strader *et al.*, 1987<sup>122</sup>; Sankara-Ramakrishnan *et al.*, 1989<sup>123</sup>; Vilsen *et al.*, 1989<sup>124</sup>; Wess *et al.*, 1993; Lummis *et al.*, 2005<sup>125</sup>; Lolkema *et al.*, 1988<sup>126</sup>; Vik *et al.*, 1988<sup>127</sup>; Miller *et al.*, 1989<sup>128</sup>)

Spezies	Protein	polyP	SDM	P	Phänotyp	Interpretation	Referenz
<i>Homo sapiens</i>	ATP2A2 Ca <sup>2+</sup> -ATPase	LDPPR (601/602) NKPPR (819/820) RMPPW (925/926)	P308 -> A, G	TM	reduzierte Affinität gegenüber Ca <sup>2+</sup>	Pro gewährleistet Stabilität und damit Positionierung von benachbarten Aminosäuren; eventuelle direkte Beteiligung an Ca <sup>2+</sup> Bindung oder indirekte Beteiligung durch strukturellen "Knick"	Sankara-Ramakrishnan <i>et al.</i> , 1989; Vilsen <i>et al.</i> , 1989
			P802 -> A	TM	reduzierte Affinität gegenüber Ca <sup>2+</sup>		Vilsen <i>et al.</i> , 1989
			P312 -> A, G, L	TM	stark reduzierter Ca <sup>2+</sup> Transport		Vilsen <i>et al.</i> , 1989
	ADRB2 Beta-Adrenorezeptor		P323 -> S	TM	statt 67 kDa Genprodukt: 44 kDa und 50 kDa	Pro an Proteinprozessierung beteiligt (evtl. Faltung)	Strader <i>et al.</i> , 1987
<i>Rattus norvegicus</i>	Chrm3 Acetylcholin Rezeptor 3	TVPPG (126/127)	P242 -> A	TM	35-100-fach reduziertes Expressionslevel	Pro essentiell für Rezeptorexpression	Wess <i>et al.</i> , 1993
			P505 -> A	TM	35-100-fach reduziertes Expressionslevel		Wess <i>et al.</i> , 1993
			P540 -> A	TM	35-100-fach reduziertes Expressionslevel, auf 25% reduzierte Stimulierung von Inositol Hydrolyse	Pro essentiell für Rezeptoraktivität	Wess <i>et al.</i> , 1993
			P201 -> A	TM	drastisch reduzierte Bindeaffinität gegenüber Agonisten und Antagonisten (~450-fach)	Pro essentiell für Ligandenbindung	Wess <i>et al.</i> , 1993
<i>Xenopus laevis</i>	5-HT3 5-Hydroxy-tryptamin Rezeptor	FLPPH (271/272) TRPPA (388/389)	P308 -> G, A, C, V, L, N, P-Analoga	C	Bindeaffinität unverändert, beeinträchtigte Transportaktivität und Konformationsänderung	Pro ausschlaggebend für Transportmechanismus, <i>cis/trans</i> Isomerisierung	Lummis <i>et al.</i> , 2005
<i>Escherichia coli</i>	AtpE F1F0-ATPase Untereinheit C		P43 -> S, A	C	reduziertes Wachstum auf Glukose und geringere Sensitivität gegenüber DCCD Inhibitor	Pro in F1-F0 Bindung involviert	Miller <i>et al.</i> , 1989
	AtpB F1F0-ATPase Untereinheit A	QNPPA (34/35)	P190 -> Q, R	C	reduziertes Wachstum auf Glukose und geringere Sensitivität gegenüber DCCD Inhibitor	Pro in F1-F0 Bindung und Protonentranslokation involviert	Vik <i>et al.</i> , 1988
	LacY <i>lac</i> Permease		P31, 123, 280, 327 -> A, G	TM	Laktosetransport nicht beeinträchtigt	hydropathische Eigenschaften von Pro Voraussetzung für Transporterfunktionalität	Lolkema <i>et al.</i> , 1988; Consler <i>et al.</i> , 1991
			P31, 123, 280, 327 -> I	TM	kein Laktosetransport		Lolkema <i>et al.</i> , 1988; Consler <i>et al.</i> , 1991
P61, 89, 97, 220 -> A, G, L	TM	Laktosetransport nicht beeinträchtigt	Lolkema <i>et al.</i> , 1988; Consler <i>et al.</i> , 1991				
P28 -> A, G, L, S	TM	kein Laktosetransport	Lolkema <i>et al.</i> , 1988; Consler <i>et al.</i> , 1991				

Die Faltung von Proteinen ist ein mehrstufiger Prozess, wobei die langsame Phase zu weiten Teilen nachweislich auf eine enzymatisch katalysierte *cis/trans* Isomerisierung von Prolin zurückgeht<sup>42, 129</sup>. Prolin wird deshalb auch eine für die Dynamik der Bildung von Tertiärstrukturen essentielle Bedeutung zugeschrieben<sup>34, 43</sup>. Aufgrund seiner eingeschränkten Flexibilität spielt Prolin mutmaßlich eine stabilisierende Rolle in Proteinstrukturen<sup>130</sup>.

### 1.3.1 Experimentelle Untersuchungen zur Bedeutung von polyP Motiven in bakteriellen Proteinen

Während die strukturelle und dynamische Rolle einzelner Proline in Proteinsequenzen bereits seit Mitte des 20. Jahrhunderts beschrieben wird, existieren nur wenige Studien zu konsekutiven, in Polyprolinmotiven akkumulierten Prolinen. Diese konzentrieren sich im Wesentlichen auf Proteine aus *Salmonella enterica* und *Escherichia coli* (Tab. 2). Neben den strukturellen und funktionellen Aspekten muss hier unbedingt die Auswirkung auf die Translationseffizienz mit in Betracht gezogen werden. Es erscheint unwahrscheinlich, dass allein die strukturellen Eigenschaften von konsekutiven Prolinen diejenigen von singulären derart überwiegen, dass ein nur mit hohem energetischem Aufwand zu behebender Translationsarrest akzeptiert und im Laufe der Evolution aufrechterhalten worden wäre.

So wurde zum Beispiel für das polyP Motiv des Mg<sup>2+</sup> Transporter MgtB und die Akkumulation von Prolinen im assoziierten leader peptide MgtL eine Funktion in der Regulation der Proteinkopienzahl nachgewiesen<sup>131, 132</sup>.

Der membranintegrierte pH Sensor und Transkriptionsaktivator CadC aus *Escherichia coli* trägt in der unstrukturierten cytoplasmatischen Linkerregion von Transmembran- und DNA-Bindedomäne zwei polyP Motive (SPPPIP und ATPPE)<sup>133, 134</sup>. An seinem Beispiel wurde der durch konsekutive Proline induzierte Translationsarrest und die Rolle von EF-P an dessen Auflösung zum ersten Mal beschrieben<sup>59</sup>. Ein SPPPIP/SAAAIS Austausch in plasmidkodiertem und niedrigexprimiertem CadC führt zu einer Erhöhung des Proteinlevels. Da dadurch das numerische Gleichgewicht zu LysP, der interagierenden und als Cosensor fungierenden Lysinpermease, und damit auch die Expression der Zielgene *cadBA* gestört wird, wurde die Schlussfolgerung gezogen, dass das polyP Motiv in CadC eine Rolle in der Aufrechterhaltung und Feinjustierung einer adäquaten Proteinkopienzahl spielt<sup>59</sup>.

Die universell konservierte Val-tRNA Synthetase ValS trägt drei konsekutive Proline im katalytischen Zentrum, welche nachweislich eine Rolle für eine effiziente Beladung von tRNA mit Valin und für die Vermeidung einer fälschlichen Beladung mit Threonin spielen<sup>71</sup>. Interessanterweise ist dieses Prolinmotiv in allen Orthologen und damit allen Spezies der drei Lebensreiche zu finden, was eine Koevolution von EF-P und dem Prolinmotiv in ValS vermuten lässt<sup>71</sup>.

**Tab. 2:** Übersicht von Proteinen, in welchen die Relevanz ausgewählter polyP Motive mittels SDM (*site directed mutagenesis*) bereits experimentell beschriebenden ist. Art und Position der Motive sind in der Spalte „polyP“ zusammengefasst. P = Lokalisierung der substituierten Aminosäure; TM = transmembran; C = cytosolisch (Referenzen: Choi *et al.*, 2017<sup>131</sup>; Gall *et al.*, 2016<sup>132</sup>; Ude *et al.*, 2013<sup>59</sup>; Starosta *et al.*, 2013<sup>71</sup>)

Spezies	Protein	polyP	SDM	P	Phänotyp	Interpretation	Referenz
<i>Salmonella enterica</i>	<b>MgtB</b> Mg <sup>2+</sup> -Transporter	QVPPA (73/74) LDPPK (555/556)	P555, 556 -> A	C	unveränderte Transporteigenschaften, keine Reduzierung des Translations-levels in Abwesenheit von EF-P	Regulation der Proteinkopienzahl	Choi <i>et al.</i> , 2017
	<b>MgtL</b> <i>leader peptide</i> von <i>mgtA</i>	(MDPEPTP LPRWRIFL FR)	P3 ->S P5 ->L P7 ->S	C	Ausbildung von mRNA stem loops unabhängig von der Prolinkonzentration	Sensor für die Konzentration von tRNA-Pro	Gall <i>et al.</i> , 2016
<i>Escherichia coli</i>	<b>ValS</b> tRNA Synthetase	IPPPN (40-42)	P40 ->G P41 -> G P42 -> G P40-42 -> G	C	Beeinträchtigung der Beladungskapazität	Pro essentiell für katalytische Aktivität	Starosta <i>et al.</i> , 2014
	<b>CadC</b> Transkriptions-aktivator	SPPPIP (120-122/124) ATPPE (146/147)	P120-122/124 -> A	C	verstärkte Expression der Zielgene <i>cadAB</i> , erhöhte Proteinkopienzahl	Feinjustierung Proteinkopienzahl	Ude <i>et al.</i> , 2013

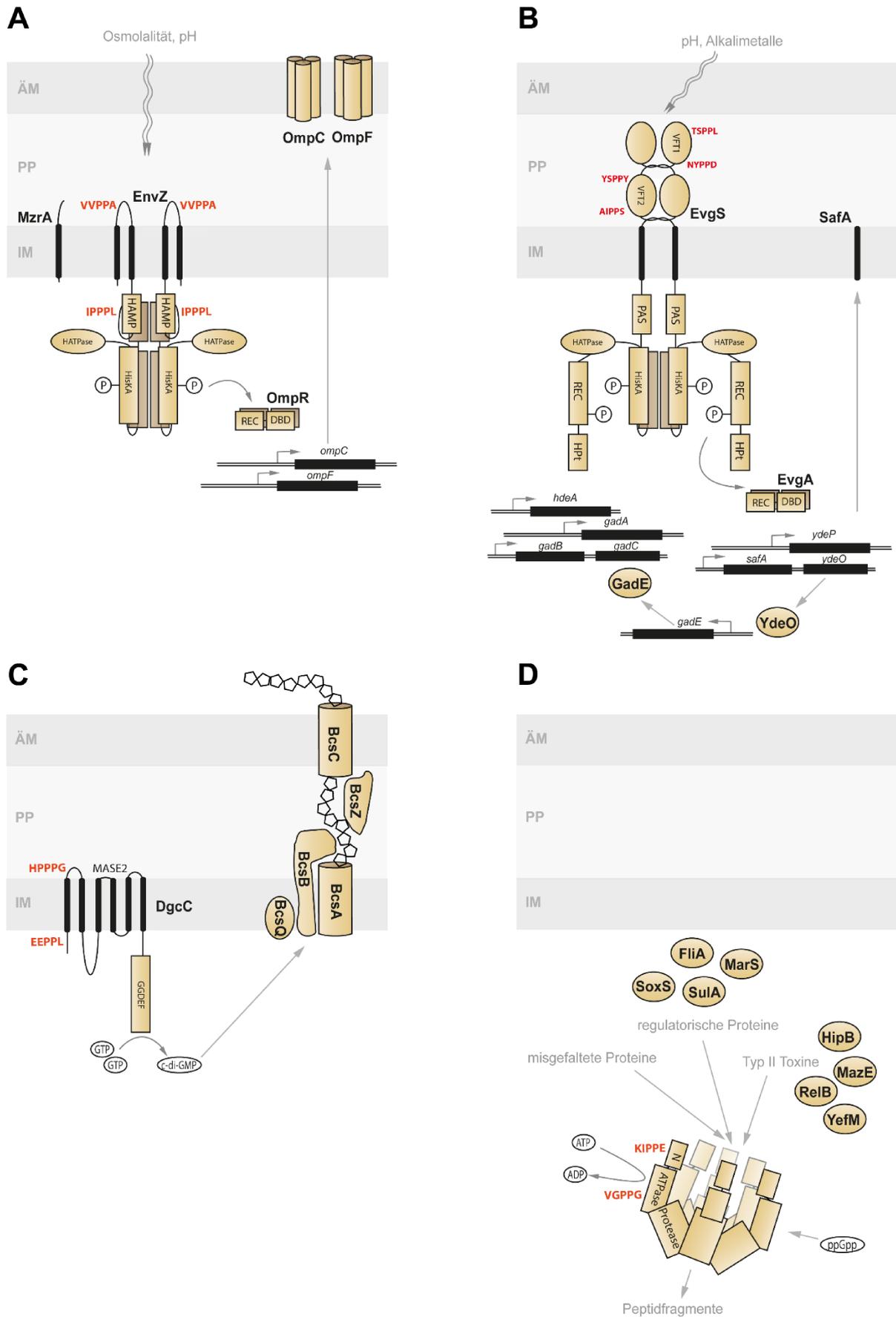
### 1.3.2 Mögliche Funktionen von polyP Motiven in Signalproteinen

Änderungen in der Proteinkopienzahl müssen umso stärker ins Gewicht fallen, je niedriger das jeweilige Expressionslevel ist. Besonders für Proteine, die Teil von komplexen und quervernetzten Signaltransduktionswegen sind, ist ein genau definiertes Verhältnis aus Aktivatoren und Repressoren von großer Bedeutung. Im Rahmen dieser Arbeit wurde die Rolle von polyP Motiven exemplarisch an den *E. coli* K-12 Proteinen EnvZ, EvgS, DgcC und Lon untersucht (Abb. 5). Ausschlaggebend für die Auswahl dieser Proteine war unter anderem die Verfügbarkeit von etablierten Testsystemen zur phänotypischen Charakterisierung von isogenen Mutanten.

#### 1.3.2.1 Histidinkinasen

Auffälligerweise tragen 18 der 30 Histidinkinasen (HKs) in *E. coli* K-12 mindestens ein polyP Motiv, wovon 8 davon potentiell stark translationsverzögernde Motive vorzufinden sind: BaeS, CreC, CpxA, EnvZ, EvgS, QseC, PhoR and RcsD.

Zweikomponentensysteme, die sich typischerweise aus membranintegrierter Histidinkinase und zugehörigem, löslichen Antwortregulator zusammensetzen, sind die Hauptakteure der Signalperzeption und -transduktion in Prokaryonten und gewährleisten eine schnelle und spezifische Anpassung an extrazelluläre Stimuli. Die Information eines von der Histidinkinase detektierten Stimulus wird über eine intramolekulare Konformationsänderung an deren Kinaseeinheit übermittelt. In Folge eines durch Autophosphorylierung an einem konservierten Histidin der Histidinkinase initiierten Phosphotransfers auf ein konserviertes Aspartat des Antwortregulator wird dessen *output* Domäne aktiviert und damit eine zelluläre Antwort angestoßen<sup>135, 136</sup>.



**Abb 5:** Schematische Übersicht der *E. coli* K-12 Signalproteine EnvZ (A), EvgS (B), DgcC (C) und Lon Protease (D), sowie deren assoziierte Signaltransduktionskaskade.

ÄM = äußere Zellmembran; PP = periplasmatischer Raum; IM = innere Zellmembran

Sowohl Histidinkinase, als auch Antwortregulator liegen in homodimerer Form vor<sup>136</sup>. Während die Sensordomäne der Histidinkinasen hoch variabel ist, ist der cytoplasmatischen Transmitterdomäne ein modularer Aufbau aus einem definierten Satz von Domänen gemeinsam, wobei deren Organisation variabel ist<sup>137, 135</sup>. Unterschieden wird zwischen periplasmatisch sensierenden Histidinkinasen, die ihren Stimulus über eine extrazelluläre Rezeptordomäne detektieren, Histidinkinasen, die auf die Erkennung von membranassoziierten Stimuli durch ihre Transmembranregionen spezialisiert sind und Histidinkinasen, die Stimuli ausschließlich im Cytoplasma erfassen<sup>135</sup>. Periplasmatisch sensierende Histidinkinasen setzen sich üblicherweise aus einer periplasmatischen Sensordomäne und einer cytoplasmatischen Transmitterdomäne zusammen. Die Transmitterdomäne beinhaltet eine variable Linkerdomäne und mindestens eine Dimerisierungs- und Phosphotransferdomäne (DHP, bzw. HisKA) mit konserviertem Histidin und die katalytische ATPase Domäne (HATPase). Der Antwortregulator besteht aus Phosphoakzeptordomäne (REC) und einer *output* Domäne für Protein-Protein Interaktionen oder die DNA-Bindung<sup>138, 135</sup>. Der Aufbau von Histidinkinasen kann in abgewandelter Form vorliegen, wie es auch in Hybridkinasen der Fall ist. Diese sind erweitert um die für den Antwortregulator typische Phosphoakzeptordomäne und eine zusätzliche Transmitterdomäne<sup>139</sup>.

Histidinkinasen zeichnen sich durch eine niedrige Proteinkopienzahl aus. Während in der exponentiellen Wachstumsphase (LB-Medium) pro Zelle bis zu 100 Kopien von EnvZ anzufinden sind, sind es im Fall von EvgS sogar nur 5<sup>140, 141</sup>.

#### 1.3.2.1.1 Der Osmo- und Säurestressrezeptor EnvZ

Die Histidinkinase EnvZ (Abb. 5A) bildet zusammen mit dem dazugehörigen Antwortregulator OmpR ein Zweikomponentensystem, das osmotischen und Säurestress sensiert und prozessiert<sup>142–145, 136, 146, 147</sup>. EnvZ ist durch zwei Transmembrandomänen in der inneren Membran verankert. Der cytoplasmatische C-terminale Teil besteht aus den drei Domänen HAMP, HisKA und HATPase<sup>136, 148</sup>. Obwohl davon ausgegangen wird, dass die periplasmatische Domäne in die Reizwahrnehmung involviert ist, zeigten neuere Studien, dass der cytoplasmatische Teil intrazelluläre Auswirkungen von externen osmotischen Schwankungen sensiert<sup>146, 147, 149</sup>. EnvZ ist bifunktional und fungiert unter Bedingungen mit hoher externer Osmolarität als Kinase und bei niedriger Osmolarität als Phosphatase<sup>150</sup>. Die Autophosphorylierung findet an His243 statt<sup>151</sup>. Phosphoryliertes OmpR reguliert die Transkription der beiden für die Hauptporine der äußeren Membran kodierenden Gene *ompC* und *ompF* auf reziproke Art und Weise<sup>152</sup>. Da OmpC und OmpF unterschiedliche Porendurchmesser haben, kann durch eine Verschiebung deren numerischen Verhältnisses die Durchlässigkeit für Moleküle bestimmter Größe limitiert werden. Die Aktivität von EnvZ wird durch das interagierende, ebenfalls in der inneren Membran lokalisierte Protein MzrA moduliert<sup>153</sup>. MzrA verbindet damit außerdem das EnvZ/OmpR System mit den Stressantwortsystemen CpxA/CpxR und  $\sigma^E$ <sup>153, 154</sup>.

EnvZ trägt zwei polyP Motive in der periplasmatischen und HAMP Domäne.

### 1.3.2.1.2 Der pH-Stresssensor EvgS

Die Hybridsensorkinase EvgS (Abb. 5B) bildet zusammen mit dem Antwortregulator EvgA ein Zweikomponentensystem zur Detektion von niedrigen extra- und intrazellulären pH-Werten (~ pH 5,5) und Alkalimetallen und reguliert eine damit verbundene spezifischen Stressantwort<sup>155, 156</sup>. Die Histidinkinase setzt sich zusammen aus einer N-terminalen periplasmatischen Sensordomäne bestehend aus zwei VFT Domänen (*venus fly trap*), einer daran anschließenden Transmembrandomäne und den fünf cytoplasmatischen Domänen PAS, HisKA, HATPase, REC und Hpt<sup>156-160</sup>. Damit zählt EvgS zu den Hybridkinasen. Die Histidinkinase bildet oligomere Cluster in der Zellmembran<sup>161</sup>.

Nach Autophosphorylierung des His721 in der HisKA Domäne wird die Phosphorgruppe an ein Aspartat der REC (*receiver*) Domäne und von dort wiederum auf ein Histidin der Hpt Domäne übertragen<sup>162</sup>. Von dort aus (*output*) wird das Signal final an EvgA übermittelt. EvgA reguliert die Transkription von einem weiten Spektrum an Zielgenen<sup>160, 163</sup>. Die wichtigsten davon betreffen das effektivste der vier Säureresistenzsysteme in *E. coli*, AR2. Dessen Hauptkomponenten sind die beiden Glutamatdecarboxylasen GadA und GadB, welche eine  $\alpha$ -Carboxylgruppe von Glutamat durch ein Proton aus dem Cytoplasma ersetzen und damit das intrazelluläre Protonenlevel reduzieren, sowie der Glutamat/ $\gamma$ -Aminobuttersäure (GABA) Antiporter GadC<sup>164-166</sup>. *GadA* und *gadBC* werden durch das Genprodukt von *gadE* aktiviert. Dessen Transkription steht wiederum unter der direkten und indirekten über YdeO vermittelten Kontrolle von EvgA<sup>166</sup>. Weitere Zielgene von EvgA, welche ebenfalls in die Säurestressantwort involviert sind, sind *emrKY* und *ydeP*<sup>163, 167</sup>.

EvgA reguliert zudem die Transkription von *safA* und *iram*, was EvgAS mit dem Zweikomponentensystem PhoPQ, welches ebenfalls an die Säurestressantwort beteiligt ist, auf verschiedenen Ebenen verknüpft.<sup>155, 168-170</sup> Die Aktivierung von EvgAS ist zudem an RcsB, den Antwortregulator des Dreikomponentensystems RcsC/RcsD/RcsB, gebunden<sup>171, 172</sup>.

EvgS trägt je zwei polyP Motive in den periplasmatischen Sensordomänen VFT1 und VFT2.

### 1.3.2.2 Andere Signalproteine

Um die Rolle von polyP Motiven in einem umfassenderen Kontext zu bewerten, wurden die beiden Proteine DgcC und Lon hinzugezogen. Beide spielen maßgebende Schlüsselrollen in mit Sekundärmetaboliten assoziierten Signalwegen. Funktionsweise und Signalweg unterscheiden sich jedoch grundsätzlich.

#### 1.3.2.2.1 Die Diguanylatcyclase DgcC

Die Diguanylatcyclase DgcC (Abb. 5C) reguliert die c-diGMP vermittelte Cellulosesynthese. Die große Bandbreite an c-diGMP bindenden Proteinen spiegelt die Diversität dieses Signalweges wieder, welcher beinahe jeden molekularen Prozess in Bakterien tangiert, darunter Transkription, mRNA Stabilität, Translation, Protein-Interaktionen oder -Degradation<sup>173</sup>. Das Molekül spielt darüberhinaus eine Schlüsselrolle in dem phänotypischen Übergang von freilebenden Einzelzellen zu koloniebildenden, sessilen Individuen in Biofilmen<sup>174</sup>. Ein komplexes Netzwerk aus verschiedenen Diguanylatcyclasen mit charakteristischer GGDEF Domäne moduliert zusammen

mit den antagonistisch wirkenden Phosphodiesterasen (gekennzeichnet durch EAL oder HD-GYP Domäne) die intrazelluläre, globale und lokale Konzentration des Sekundärmetaboliten c-diGMP.

Die Art und Anzahl der Diguanyltycyclasen und Phosphodiesterasen ist innerhalb der *E. coli* Stämme sehr variabel. Wenige davon sind Teil eines Kernproteoms und bilden ein zentrales, koreguliertes Netzwerk zur Aktivierung der Biofilmproduktion. Dazu zählt unter anderem auch DgcC, was dessen Wichtigkeit unabhängig von den ökologischen Umweltbedingungen unterstreicht<sup>175</sup>. In *E. coli* K-12 finden sich insgesamt 12 Proteine mit GGDEF Domäne und 10 mit EAL Domäne<sup>173</sup>.

Die C-terminale, cytoplasmatische GGDEF Domäne von DgcC weist die für die Enzymklasse der Diguanylatcyclasen charakteristische Aktivität auf, während N-terminal eine aus 6 Transmembrandomänen bestehende MASE2 Rezeptordomäne lokalisiert ist<sup>176, 175</sup>. Aus zwei Molekülen GTP wird mit Hilfe der GGDEF Domäne Bis-(3'-5')-cyklisches dimeres GMP (c-diGMP) katalysiert<sup>177, 174</sup>. Die Homodimerisierung der GGDEF Domäne ist Voraussetzung für die Aktivierung von Diguanylatcyclasen<sup>178, 179</sup>.

Zusammen mit den amyloiden *curli* Fibern bildet das Glycopolysaccharid Cellulose den Hauptbestandteil der extrazellulären *E. coli* Matrix<sup>180, 181</sup>. Die  $\beta$ -1,4 glycosidische Verbindung von UDP-Glucoseeinheiten zu Cellulose findet in *E. coli* durch das Zusammenwirken der Einheiten des membranintegrierten Cellulosesynthase-Komplexes statt. Dieser besteht aus den in den beiden konservierten Operons *yhjR-bcsQABZC* und *bcsEFG* kodierten c-diGMP sensitiven Cellulosesynthase BcsA (C-terminale PilZ Domäne), dem akzessorischen, kokatalytischem Membranprotein BcsB, der periplasmatischen Lyase BcsZ, einem in der äußeren Membran lokalisierten Porin mit periplasmatischem Gerüstmotiv und zusätzlichen für die Struktur und Regulation bedeutenden Untereinheiten, wie zum Beispiel BcsQ<sup>182-186</sup>. In den gängigen *E. coli* Laborstämmen (u.a. K-12, BW25113) verhindert ein Nukleotidmorphismus in *bcsQ* die Transkription des vollständigen *yhjR-bcsQABZC* Operons. Durch eine Wiederherstellung der Ursprungssequenz kann der Cellulosemetabolismus vollständig reaktiviert werden<sup>180</sup>.

DgcC reguliert die Cellulosesynthese, hat aber keinen Effekt auf die *curli* Bildung. In Abwesenheit von DgcC findet keine Cellulosesynthese statt<sup>187</sup>. Da DgcC die intrazelluläre c-diGMP Konzentration lokal reguliert und keinen Einfluss auf das globale Level hat, wird von einer direkten Interaktion mit der Cellulosesynthase ausgegangen<sup>175</sup>.

Die Expression von *dgcC* steht unter der Kontrolle des temperatur- und RpoS-abhängigen Biofilmregulators CsgD und wird damit in der stationären Wachstumsphase aktiviert<sup>175, 188, 189</sup>. Eine zusätzliche Regulation von DgcC durch die Diguanylatzyklase DgcM fügt dem Signaltransduktionsweg eine weitere Komplexitätsstufe hinzu<sup>189</sup>.

DgcC trägt zwei polyP Motive in der MASE2 Domäne.

#### 1.3.2.2.2 Die AAA+ Protease Lon

Die ATP-abhängige Serinprotease Lon (Abb. 5D) gehört zur Familie der AAA+ Proteasen und ist der einzige Vertreter, dessen Orthologe in allen lebenden Organismen anzufinden sind<sup>190, 191</sup>. In *E. coli* degradiert Lon die Hälfte aller missgefalteten Proteine und spielt damit unter anderem eine Rolle in der zellulären Stressantwort auf Umweltbedingungen, die zu Proteindenaturierung und Aggregation führen<sup>192-195</sup>. Darüber hinaus spielt Lon durch die spezifische Proteolyse von intakten, aber instabilen Signalproteinen eine essentielle regulatorische Rolle<sup>196-198</sup>.

*E. coli* Lon gehört zur LonA Familie und zeichnet sich damit durch folgende drei Domänen aus: einer N-terminalen Domäne N, der ATPase Domäne A mit in Nukleotidbindung und Hydrolyse involvierten Walker A und B Motiven und der C-terminalen Proteasedomäne P mit katalytisch aktivem Serin<sup>199-201</sup>. *In vivo* liegt die Protease als Hexamer oder Dodecamer vor, wobei das Homooligomer eine axiale Kammer formt und der Bereich für die Hydrolyse der Peptidbindung im Inneren lokalisiert ist<sup>199, 193</sup>. Die Substrate werden über die aus den N-Domänen gebildete Eingangspore aufgenommen und in ca. 3-15 Aminosäuren lange Peptidfragmente degradiert, welche über die durch die P-Domänen gebildete Ausgangspore entlassen werden<sup>202-205, 199</sup>. Die Bindung und Hydrolyse von ATP induziert eine Konformationsänderung innerhalb des Lon-Komplexes, welche letztendlich die mechanische Entfaltung, Diassemblierung und Translokation der assoziierten Substrate ermöglicht<sup>206-209</sup>. Die Spaltung der Peptidbindung in der durch ein Serin-Lysin-Paar charakterisierten proteolytischen Domäne findet spontan und ohne weiteren Energieaufwand statt<sup>210-212</sup>.

Regulatorische Zielproteine der spezifischen Proteolyse durch Lon sind zahlreich und divers. Sie betreffen unter anderem *quorum sensing*<sup>213, 214, 192</sup>, Zellmotilität<sup>215, 192</sup>, glutamatabhängige Säurestressantwort<sup>216</sup>, das Superoxidregulon<sup>217</sup>, Antibiotikaresistenz<sup>217</sup> und die SOS-DNA-Schadensreparatur<sup>218, 219</sup>. Im Fall einer Aminosäurelimitation werden die freien ribosomalen Proteine durch Lon degradiert, um Aminosäuren zu generieren, die für die Synthese von essentiellen Proteinen eingesetzt werden können<sup>220</sup>.

Lon ist nicht zuletzt an der Regulation des Mengenverhältnisses von TypII Toxin-Antitoxin Systemen beteiligt. Diese basieren auf in einem Operon kodiertem Toxin, welches inhibierend in den zellulären Metabolismus eingreift, und Antitoxin, welches das Toxin bindet und neutralisiert<sup>221, 222</sup>. Antitoxine sind üblicherweise hoch flexibel und zeichnen sich durch eine thermodynamische Instabilität aus, welche für die Bindung des Toxins erforderlich ist, sie auf der anderen Seite aber auch leicht zugänglich für Proteasen macht<sup>223</sup>. Die durch die Toxine gehemmten Zellaktivitäten, wozu unter anderem auch die globale Translationsrate zählt, erlauben letztendlich eine Überdauerung von akuten Stressbedingungen<sup>223, 224</sup>. Toxine sind assoziiert mit phänotypischen Erscheinungen wie Persistenz, Wachstumsinhibition oder Biofilmbildung<sup>225, 224</sup>. Lon ist nachweislich an der Proteolyse der Antitoxine RelB, MazE, HipB und YefM unter Aminosäure- und Stickstofflimitation beteiligt<sup>226, 192, 227-229, 225</sup> und ist die einzige Protease, welche für die durch HipA/HipB vermittelte bakterielle Persistenz benötigt wird<sup>196</sup>.

Da eine Überproduktion den unspezifischen Abbau von intakten Proteinen zur Folge hat, ist eine strikte Regulation der Expression und Aktivierung von Lon unerlässlich<sup>230</sup>. Die Kopienzahl von Lon ist an die jeweils vorherrschende Umweltbedingung angepasst, wobei die Regulation durch verschiedene Mechanismen stattfindet. So wurde unter anderem beschrieben, dass die Expression von Lon durch den RNA Polymerase Sigmafaktor RpoH, welcher eine Stressantwort auf Hitzestress induziert, reguliert wird<sup>231, 232</sup>. Eine Möglichkeit der Aktivierung von Lon ist durch das Polyphosphat (p)ppGpp gegeben<sup>220, 233, 196</sup>.

Lon trägt zwei polyP Motive in der N-terminalen und ATPase Domäne.

#### 1.4 Aufgabenstellung

PolyP Motive sind trotz ihres translationsverzögernden Potentials, dessen Abschwächung mit einem erheblichen energetischen Aufwand einhergeht, in einer großen Bandbreite an Proteinen vertreten. Dies weist auf eine übergeordnete Funktion hin, deren Vorteile die evolutionsbiologisch ungünstigen Eigenschaften überwiegen. Diese Arbeit soll die funktionelle und regulatorische Bedeutung von polyP Motiven am Beispiel des Modellorganismus *E. coli* K-12 ergründen. Der Fokus liegt dabei auf Signalproteinen und insbesondere auf niedrig exprimierten Rezeptorhistidinkinasen. Diese Auswahl begründet sich aus der Beobachtung, dass polyP Motive in dieser Proteinklasse sehr häufig auftreten. Neben funktionellen Aspekten soll geklärt werden, ob die Motive zu einer Limitation der Proteinkopienzahlen beitragen.

Eine umfassende Analyse des Auftretens von polyP Motiven in *E. coli* soll Aufschluss über deren Evolution geben. Die putativ stark translationsverzögernden polyP Motive der niedrigexprimierten Histidinkinasen EnvZ und EvgS aus *E. coli* K-12 sollen im Detail und im Vergleich zu den polyP Motiven der beiden Signalproteine DgcC und Lon Protease untersucht werden. Durch eine Substitution der entsprechenden polyP Motive sollen Rückschlüsse auf deren Beteiligung an der Proteinfunktionalität und -kopienzahl gezogen werden. Desweiteren soll die Dimerisierung von EnvZ und dessen Interaktion mit MzrA in Abhängigkeit zu den darin enthaltenen polyP Motiven ergründet werden. Eine mögliche translatorische Regulation des Expressionslevels von Proteinen mit polyP Motiv durch die Modulierung der EF-P Aktivität bei bestimmten Wachstumsbedingungen soll mittels Expressionsanalysen und der Identifizierung von putativ beteiligten Transkriptionsfaktoren untersucht werden.

## 2. MATERIAL UND METHODEN

### 2.1 Material

Sofern nicht anders gekennzeichnet und in den Tabellen 3, 4 und 5 aufgeführt, wurden Chemikalien und Verbrauchsmittel von den Firmen Bayer (Leverkusen), Cargille (New York, USA), Eppendorf (Hamburg), Gibco/BRL (Eggenstein), Kimberly-Clark (Irving, USA), Menzel (Zürich, Schweiz), Merck (Darmstadt), Riedel-de Häen (Seelze), Roche Diagnostics (Mannheim), Roth (Karlsruhe), Sarstedt AG & Co. (Nümbrecht) und Schott (Mainz) bezogen.

Tabelle 3 fasst alle im Rahmen dieser Arbeit verwendeten Chemikalien zusammen, während Enzyme in Tabelle 4 gelistet sind und Tabelle 5 Antikörper aufführt. Für enzymatische Reaktionen wurden vom Hersteller empfohlene und vertriebene Puffer verwendet, sofern nicht anders beschrieben.

**Tabelle 3:** In dieser Arbeit verwendete Chemikalien wurden von den Herstellern AppliChem (Darmstadt), BD (Franklin Lakes, USA), Biomol (Hamburg), BioRad (München), Biozym (Hess. Oldendorf), Expedon (Over Cambridgeshire, England), Fluka (Buchs, Schweiz), GE Healthcare (Braunschweig), Invitrogen (Karlsruhe), Merck (Darmstadt), Qiagen (Hilden), Roth (Karlsruhe), Remel (Lenexa, USA), Serva (Heidelberg), und Sigma-Aldrich (Deisenhofen) im Reinheitsgrad *pro analysis* bezogen.

Bezeichnung	Hersteller	Bezeichnung	Hersteller
Acetonitril [C <sub>2</sub> H <sub>3</sub> N]	Roth	Kupfersulfat [CuSO <sub>4</sub> ]	Roth
Agar	BD	Lachssperma-DNA	Fluka
Agarose	Serva	Magermilchpulver	Roth
Acrylamid [C <sub>3</sub> H <sub>5</sub> NO]	Serva	Magnesiumchlorid [MgCl <sub>2</sub> ]	Roth
Acrylamid 30%: Bisacrylamid 0,8% (37,1:1)	Roth	Magnesiumsulfat [MgSO <sub>4</sub> ]	Roth
Ammoniumhydrogencarbonat [NH <sub>4</sub> HCO <sub>3</sub> ]	Roth	Manganchlorid [MnCl <sub>2</sub> ]	Merck
Ammoniumchlorid [NH <sub>4</sub> Cl]	Roth	β-Mercaptoethanol [C <sub>2</sub> H <sub>6</sub> OS]	Roth
Ammoniummolybdat [(NH <sub>4</sub> ) <sub>6</sub> Mo <sub>7</sub> O <sub>24</sub> ]	Roth	MES	Roth
Ammoniumpersulfat [(NH <sub>4</sub> ) <sub>2</sub> S <sub>2</sub> O <sub>8</sub> ]	Roth	Midori-Green® Advanced DNA Stain	Biozym
L-Arabinose [C <sub>5</sub> H <sub>10</sub> O <sub>5</sub> ]	Roth	MOPS	Roth
Bisacrylamid [C <sub>7</sub> H <sub>10</sub> N <sub>2</sub> O <sub>2</sub> ]	Serva	Natriumacetat [Na(CH <sub>3</sub> COO)]	Merck
Borsäure [H <sub>3</sub> BO <sub>3</sub> ]	Roth	Natriumcarbonat [Na <sub>2</sub> CO <sub>3</sub> ]	Roth
BSA	AppliChem	Natriumchlorid [NaCl]	Roth
5-Brom-4-Chlor-3-indolyl-β-D-galactopyranosid (X-Gal)	AppliChem	Natriumdodecylsulfat (SDS)	Roth
5-Bromo-4-chloro-3-indolylphosphat (BCIP)	AppliChem	Natriumhydrogencarbonat [NaHCO <sub>3</sub> ]	Roth
Bromphenolblau	Roth	Natriumdihydrogenphosphat [NaH <sub>2</sub> PO <sub>4</sub> ]	Roth
Calciumchlorid [CaCl <sub>2</sub> ]	Roth	Natriumhydroxid [NaOH]	Roth
Calcofluor White	Remel	Natriumthiosulfat [Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub> ]	Fluka
Chloroform [CHCl <sub>3</sub> ]	Roth	Ni-NTA Agarose	Quiagen
Cobaltchlorid [CoCl <sub>2</sub> ]	Roth	Nitroblau-Tetrazolium (NBT)	Biomol
Coomassie Brilliant Blau G250	Roth	o-Nitrophenyl-β-D-galactopyranosid (oNPG)	Roth

Coomassie Instant Blue	Expedon	P-20	Roth
2,6-Diaminopimelinsäure (DAP)	Roth	Paraformaldehyd	Roth
Dikaliumhydrogenphosphat [K <sub>2</sub> HPO <sub>4</sub> ]	Roth	P/C/I (25:24:1)	Roth
Dinatriumhydrogenphosphat [Na <sub>2</sub> HPO <sub>4</sub> ]	Roth	Phenylmethylsulfonylfluorid (PMSF)	Roth
Dithioerythritol (DTT)	Roth	Percoll®	GE Healthcare
Eisensulfat [FeSO <sub>4</sub> ]	Roth	Poly-L-Lysin 0,1% (w/v)	Sigma Aldrich
Essigsäure [CH <sub>3</sub> COOH]	Roth	Propidiumiodid (PI)	Roth
Ethidiumbromid [C <sub>21</sub> H <sub>20</sub> BrN <sub>3</sub> ]	Roth	D(+)-Saccharose [C <sub>12</sub> H <sub>22</sub> O <sub>11</sub> ]	Roth
Ethylendiamintetraacetat (EDTA)	Roth	Salzsäure [HCl]	Roth
Formaldehyd [CH <sub>2</sub> O]	Roth	Silbernitrat [AgNO <sub>3</sub> ]	Roth
Harnstoff [CH <sub>4</sub> N <sub>2</sub> O]	Sigma-Aldrich	SYPRO®-Orange	Invitrogen
Hefeextrakt	BD	SYTOX®	Invitrogen
HEPES	Roth	TEMED	Roth
Hoechst	Sigma_Aldrich	2,2,2-Trichlorethanol [CCl <sub>3</sub> CH <sub>2</sub> OH]	Sigma-Aldrich
D-(+)-Glucose [C <sub>6</sub> H <sub>12</sub> O <sub>6</sub> ]	Roth	Tricin [C <sub>6</sub> H <sub>13</sub> NO <sub>5</sub> ]	Roth
Glycerin [C <sub>3</sub> H <sub>8</sub> O <sub>3</sub> ]	Roth	Trifluoressigsäure [C <sub>2</sub> HF <sub>3</sub> O <sub>2</sub> ]	Roth
Glycin [C <sub>2</sub> H <sub>5</sub> NO <sub>2</sub> ]	Roth	TRIS	Roth
Imidazol	Roth	Triton-X100	BioRad
Isopropyl-β-thiogalactopyranosid (IPTG)	AppliChem	Trypton	Roth
Kaliumchlorid [KCl]	Roth	Tween-20	GE Healthcare
Kaliumdihydrogenphosphat [KH <sub>2</sub> PO <sub>4</sub> ]	Roth	Wasserstoffperoxid (H <sub>2</sub> O <sub>2</sub> )	Sigma-Aldrich
Kaliumsulfat [K <sub>2</sub> SO <sub>4</sub> ]	AppliChem	Xylencyanol	Roth
Kongorot	Merck	Zinksulfat [ZnSO <sub>4</sub> ]	Merck

**Tabelle 4:** In dieser Arbeit verwendete Enzyme wurden von den Firmen New England BioLabs (Frankfurt), Thermo Scientific (Schwerte) und Roth (Karlsruhe) bezogen.

Bezeichnung	Hersteller/Referenz
CIP (calf intestine phosphatase)	New England BioLabs (Frankfurt)
DNase, RNase-frei, 10 U/μl	Thermo Scientific (Schwerte)
OneTaq DNA Polymerase, 5 U/μl	New England BioLabs (Frankfurt)
Q5 Polymerase	New England BioLabs (Frankfurt)
[Restriktionsenzyme]	New England BioLabs (Frankfurt)
RNase	Roth (Karlsruhe)
rSAP (shrimp alkaline phosphatase)	New England BioLabs (Frankfurt)
T4 DNA Ligase	New England BioLabs (Frankfurt)
Ulp Protease (mit His-Tag)	Ingrid Weigl (LMU, München)

**Tabelle 5:** In dieser Arbeit verwendete Antikörper. (AP = alkalische Phosphatase; HRP = *horseradish peroxidase*).

Bezeichnung	Reaktivität	Ursprungsorganismus	Verdünnung	Referenz
$\alpha$ -FLAG-HRP	-	Maus	1:5.000	Sigma Aldrich (Schnelldorf)
$\alpha$ -His	-	Maus	1:5.000	Qiagen (Hilden)
$\alpha$ -EF-P	<i>E. coli</i>	Kaninchen	1:5.000	diese Arbeit
$\alpha$ -EpmA	<i>E. coli</i>	Kaninchen	1:500	diese Arbeit
$\alpha$ -EpmB	<i>E. coli</i>	Kaninchen	1:5.000	diese Arbeit
$\alpha$ -EnvZ	<i>E. coli</i>	Kaninchen	1:4.000	Prof. Dr. Linda Kenney (National University of Singapore and University of Illinois, Chicago)
$\alpha$ -EnvZ	<i>E. coli</i>	Kaninchen	1:4.000	Prof. Dr. Kirsten Jung (LMU, München)
$\alpha$ -OmpR	<i>E. coli</i>	Kaninchen	1:5.000	Prof. Dr. Kirsten Jung (LMU, München)
$\alpha$ -EvgS	<i>E. coli</i>	Kaninchen	1:10.000	Prof. Dr. Utsumi (Kinki University, Nakamachi)
$\alpha$ -Lon	<i>E. coli</i>	Kaninchen	1:5.000	Prof. Dr. Baker (Massachusetts Institute of Technology, Cambridge)
$\alpha$ -rpsB	<i>E. coli</i>	Kaninchen	1:7.500	Fitzgerald Industries (North Acton/USA)
$\alpha$ -mCherry	-	Kaninchen	1:5.000	Thermo Fisher Scientific (Waltham/USA)
$\alpha$ -GFP	-	Maus	1:1.000	Roche (Karlsruhe)
$\alpha$ -mNeonGreen	-	Maus	1:1.000	Chromotek (Martinsried)
$\alpha$ -mCherry-Atto647N	-	Alpaka (Nanobody)	1:200	Chromotek (Martinsried)
$\alpha$ -GFP-Atto488	-	Alpaka (Nanobody)	1:200	Chromotek (Martinsried)
$\alpha$ -Mouse-AP	<i>Mus musculus</i>	Huhn	1:4.000	Rockland (Pottstown/USA)
$\alpha$ -Rabbit-AP	<i>Oryctolagus cuniculus</i>	Ziege	1:4.000	Rockland (Pottstown/USA)
$\alpha$ -Mouse-HRP	<i>Mus musculus</i>	Ziege	1:4.000	Thermo Fisher Scientific (Waltham/USA)
$\alpha$ -Rabbit- HRP	<i>Oryctolagus cuniculus</i>	Ziege	1:4.000	Thermo Fisher Scientific (Waltham/USA)

Die Antikörper  $\alpha$ -EF-P,  $\alpha$ -EpmA und  $\alpha$ -EpmB wurden von der Firma Eurogentec produziert und basieren auf einer Immunisierung von Kaninchen mit gereinigten, rekombinanten, His-getagten Vollängenproteinen (hergestellt von Ingrid Weitl, LMU). Im Rahmen des „Speedy 28-Day“ Programms wurden zwei Tiere mit Injektionen nach 0, 7, 10, 18, 21 und 28 Tagen behandelt. Die antigenspezifischen Antikörper aus einem der beiden Antiseren wurden nach vorheriger Begutachtung im Western Blot von Eurogentec durch eine Affinitätsreinigung („AS-Puri-Small“) konzentriert.

## 2.2 Stämme, Plasmide und Oligonukleotide

Alle im Rahmen dieser Arbeit verwendeten Stämme und Plasmide sind in den Tabellen 6 und 7 aufgeführt. Oligonukleotide wurden von der Firma Sigma Aldrich (München) bezogen und sind in Tabelle 8 aufgelistet.

**Tabelle 6:** Die im Rahmen dieser Arbeit erstellten Stämme wurden wie unter 2.4.8 beschrieben konstruiert.

	Bezeichnung	Genotyp	Konstruktion	Referenz
St1	DH5α λpir	<i>E. coli</i> F <sup>+</sup> Φ80 <i>lacZ</i> ΔM15 Δ( <i>lacZYA-argF</i> ) U169 <i>recA1 endA1 hsdR17</i> (rK <sup>-</sup> , mK <sup>+</sup> ) <i>phoA</i> <i>supE44 λ<sup>-</sup> thi-1 gyrA96 relA1 λpir</i>		Promega
St2	BL21 DE3	<i>E. coli</i> B F <sup>-</sup> <i>dcm</i> (DE3) <i>ompT hsdS</i> (r <sub>B</sub> m <sub>B</sub> <sup>-</sup> ) <i>gal</i>		Studier & Moffatt, 1986 <sup>234</sup>
St3	WM3064	<i>E. coli</i> <i>thrB1004 pro thi rpsL hsdS lacZ</i> ΔM15 RP4-1360 Δ( <i>araBAD</i> )567 Δ <i>dapA1341::[erm pir]</i>		W. Metcalf
St4	BW25113	<i>E. coli</i> Δ( <i>araD-araB</i> )567, Δ <i>lacZ4787::rrnB-3</i> , <i>lambda<sup>-</sup></i> , <i>rph<sup>-1</sup></i> , Δ( <i>rhaD-rhaB</i> )568, <i>hsdR514</i>		Baba <i>et al.</i> , 2006 <sup>235</sup>
St5	BW25113 Δ <i>efp</i>	BW25113; Δ <i>efp</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St6	BW25113 Δ <i>efp</i> Kan <sup>s</sup>	BW25113; Δ <i>efp</i>		Lassak, 2011
St7	BW25113 Δ <i>envZ</i> (JW3367)	BW25113; Δ <i>envZ</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St8	BW25113 Δ <i>evgS</i> (JW2367)	BW25113; Δ <i>evgS</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St9	BW25113 Δ <i>lon</i> (JW0429)	BW25113; Δ <i>lon</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St10	BW25113 Δ <i>oxyR</i> (JW3933)	BW25113; Δ <i>oxyR</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St11	BW25113 Δ <i>epmA</i> (JW4116)	BW25113; Δ <i>epmA</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St12	BW25113 Δ <i>epmB</i> (JW4106)	BW25113; Δ <i>epmB</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St13	BW25113 Δ <i>pgrR</i> (JW2970)	BW25113; Δ <i>pgrR</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St14	BW25113 Δ <i>ulaR</i> (JW4149)	BW25113; Δ <i>ulaR</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St15	BW25113 Δ <i>argP</i> (JW2883)	BW25113; Δ <i>argP</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St16	BW25113 Δ <i>purR</i> (JW1650)	BW25113; Δ <i>purR</i> (:: <i>FRT-kan_casette-FRT</i> )		Baba <i>et al.</i> , 2006 <sup>235</sup>
St17	BW25113 Δ <i>nagC</i> Kan <sup>R</sup>	BW25113; Δ <i>nagC</i> (:: <i>rpsL-neo</i> )	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr13 + Pr14	diese Arbeit
St18	BW25113 EnvZ-PAGFP	BW25113; <i>envZ::PAgfp::pNPTS(backbone)</i>	Einfach homologe Rekombination Plasmid: V33	diese Arbeit

St19	BW25113 $\Delta efp$ EnvZ-PAGFP	BW25113 $\Delta efp$ ; <i>envZ::PAgfp::pNPTS(backbone)</i>	Einfach homologe Rekombination Plasmid: V33	diese Arbeit
St20	BW25113 MzrA-PAmCherry Kan <sup>S</sup>	BW25113; <i>mzrA::PAmCherry</i>	Doppelt homologe Rekombination Plasmid: V36	diese Arbeit
St21	BW25113 $\Delta efp$ MzrA-PAmCherry Kan <sup>S</sup>	BW25113 $\Delta efp$ ; <i>mzrA::PAmCherry</i>	Doppelt homologe Rekombination Plasmid: V386	diese Arbeit
St22	BW25113 EnvZ-PAGFP MzrA-PAmCherry	BW25113; <i>envZ::PAgfp::pNPTS(backbone)</i> <i>mzrA::PAmCherry</i>	Doppelt homologe Rekombination Plasmid: V36 Einfach homologe Rekombination Plasmid: V33	diese Arbeit
St23	BW25113 $\Delta efp$ EnvZ-PAGFP MzrA-PAmCherry	BW25113 $\Delta efp$ ; <i>envZ::pagfp::pNPTS(backbone)</i> <i>mzrA-pamcherry</i>	Doppelt homologe Rekombination Plasmid: V36 Einfach homologe Rekombination Plasmid: V33	diese Arbeit
St24	BW25113 EvgS-PAGFP	BW25113; <i>evgS::PAgfp::pNPTS(backbone)</i>	Einfach homologe Rekombination Plasmid: V34	diese Arbeit
St25	BW25113 $\Delta efp$ EvgS-PAGFP	BW25113 $\Delta efp$ ; <i>evgS::PAgfp::pNPTS(backbone)</i>	Einfach homologe Rekombination Plasmid: V34	diese Arbeit
St26	BW25113 EnvZ-mNeonGreen MzrA-PAmCherry Kan <sup>S</sup>	BW25113; <i>envZ::mNeonGreen</i> <i>mzrA::PAmCherry</i>	Doppelt homologe Rekombination Plasmid: V36 Doppelt homologe Rekombination Plasmid: V35	Dr. Dagmar Koschnitzki (LMU, München)
St27	BW25113 $\Delta efp$ EnvZ-mNeonGreen MzrA-PAmCherry Kan <sup>S</sup>	BW25113 $\Delta efp$ ; <i>envZ::mNeonGreen</i> <i>mzrA::PAmCherry</i>	Doppelt homologe Rekombination Plasmid: V36 Doppelt homologe Rekombination Plasmid: V35	Dr. Dagmar Koschnitzki (LMU, München)
St28	<i>B. subtilis</i> DivVIA-PAGFP	<i>B. subtilis</i> ; <i>divVIA::PAgfp</i>		Prof. Dr. Bramkamp (LMU, München)
St29	<i>B. subtilis</i> DivVIA-PAmCherry	<i>B. subtilis</i> ; <i>divVIA::PAmCherry</i>		Prof. Dr. Bramkamp (LMU, München)
St30	MG1655	<i>E. coli</i> K-12 Referenzstamm; F <sup>-</sup> $\lambda$ <i>ilvG rfb50 rph1</i>		Blattner <i>et al.</i> , 1997 <sup>236</sup>
St31	MG1655 <i>rpsL150</i>	F <sup>-</sup> $\lambda$ <i>ilvG rfb50 rph1 rpsL150</i>		Heermann <i>et al.</i> , 2008 <sup>237</sup>

St32	MG1655 <i>rpsL150</i> $\Delta envZ$ Kan <sup>S</sup>	MG1655 <i>rpsL150</i> ; $\Delta envZ$ (80-1270)	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr1 + Pr2 Flp/FRT Rekombination	diese Arbeit
St33	MG1655 <i>rpsL150</i> EnvZ <sub>IPPL/IAAL</sub>	MG1655 <i>rpsL150</i> ; <i>envZ</i> (C604G, G606A, C607G, G609A, C910G, G612A)	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr1 + Pr2 Flp/FRT rekombination Doppelt homologe Rekombination Plasmid: V30	diese Arbeit
St34	MG1655 <i>rpsL150</i> $\Delta evgS$ Kan <sup>S</sup>	MG1655 <i>rpsL150</i> ; $\Delta evgS$ (81-559)	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr5 + Pr6 Flp/FRT Rekombination	diese Arbeit
St35	MG1655 <i>rpsL150</i> EvgS <sub>NYPPD/NYAAD</sub>	MG1655 <i>rpsL150</i> ; <i>evgS</i> (C520G, C522A, C523G, C525A)	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr5 + Pr6 Flp/FRT Rekombination Doppelt homologe Rekombination Plasmid: V31	diese Arbeit
St36	MG1655 <i>rpsL150</i> $\Delta lon$ Kan <sup>S</sup>	MG1655 <i>rpsL150</i> ; $\Delta lon$ (81-2273)	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr9 + Pr10 Flp/FRT Rekombination	diese Arbeit
St37	MG1655 <i>rpsL150</i> Lon <sub>VGPPG/VGAAG</sub>	MG1655 <i>rpsL150</i> ; <i>lon</i> (C1069G, G1071A, C1072G, G1074A)	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr9 + Pr10 Flp/FRT Rekombination Doppelt homologe Rekombination Plasmid: V32	diese Arbeit
St38	MG1655 <i>rpsL150</i> EvgS- FLAG	MG1655 <i>rpsL150</i> ; <i>evgS</i> :: <i>FLAG</i> :: <i>rpsL-neo</i>	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr7 + Pr8	diese Arbeit
St39	MG1655 <i>rpsL150</i> EvgS <sub>NYPPD/NYAAD</sub> -FLAG	MG1655 <i>rpsL150</i> ; <i>evgS</i> (C520G, C522A, C523G, C525A):: <i>FLA</i> :: <i>rpsL-neo</i>	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr5 + Pr6 Flp/FRT Rekombination Doppelt homologe Rekombination Plasmid: V31 pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr7 + Pr8	diese Arbeit
St40	MG1655 $\Delta efp$	MG1655 <i>rpsL150</i> ; P <sub>tdp</sub> :: <i>lacZ</i> , $\Delta efp$		Dr. Hannah Schramke
St41	EPB273a	MG1655 $\Psi(ompR^+ \text{-} yfp^+) envZ::kan$ <i>attL</i> ::( <i>envZ</i> <sup>-</sup> <i>cat</i> ) $\Phi(ompC^+ \text{-} cfp^+) \Phi(ompF^+ \text{-} yfp^+)$		Prof. Dr. Mark Goulian (University of Pennsylvania)
St42	EPB273a $\Delta envZ$	EPB273a; $\Delta envZ$ (590-634):: <i>FRT-PGK- gb2-neo-FRT</i>	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr3 + Pr4	diese Arbeit
St43	EPB273a EnvZ <sub>IPPL/IAAL</sub>	EPB273a; <i>envZ</i> (C604G, G606A, C607G, G609A, C910G, G612A)	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr3 + Pr4 Flp/FRT Rekombination Doppelt homologe Rekombination Plasmid: V29	diese Arbeit

St44	W3110	<i>thyA36, deoC2, IN(rrnD-rrnE)I</i>		Prof. Dr. Regine Hengge (Humboldt Universität zu Berlin), Serra <i>et al.</i> , 2013 <sup>180</sup>
St45	AR3110	W3110; <i>bcsQ</i> <sup>+</sup>		Prof. Dr. Regine Hengge (Humboldt Universität zu Berlin)
St46	AR3110 AdrA <sub>HPSPG/HPSPG</sub>	AR3110; <i>adrA</i> (C196T)	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr11 + Pr12 Flp/FRT Rekombination Doppelt homologe Rekombination Plasmid: V38	diese Arbeit
St47	AR3110 AdrA-FLAG	AR3110; <i>adrA::FLAG::pNPTS</i> (backbone)	Einfach homologe Rekombination Plasmid: V37	diese Arbeit
St48	AR3110 AdrA <sub>HPSPG/HPSPG</sub> -FLAG	AR3110; <i>adrA</i> (C196T):: <i>FLAG::pNPTS</i> (backbone)	pRED/ET <sup>TM</sup> Rekombination DNA-Fragment: Pr11 + Pr12 Flp/FRT Rekombination Doppelt homologe Rekombination Plasmid: V38 Einfach homologe Rekombination Plasmid: V37	diese Arbeit
St49	AR3110 $\Delta$ <i>csgB</i> AdrA-FLAG	AR3110 AdrA-FLAG; <i>\Delta csgB(::cm)</i>		Prof. Dr. Regine Hengge (Humboldt Universität zu Berlin)
St50	AR3110 $\Delta$ <i>csgB</i> AdrA <sub>HPSPG/HPSPG</sub> -FLAG	AR3110 AdrA <sub>HPSPG/HPSPG</sub> -FLAG; <i>\Delta csgB(::cm)</i>		Prof. Dr. Regine Hengge (Humboldt Universität zu Berlin)
St51	<i>Photothabdus luminescens</i> subsp. <i>laumondi</i> TT01-1°	Wildtyp		Duchaud <i>et al.</i> , 2003 <sup>238</sup>
St52	<i>Salmonella enterica</i> serovar Typhimurium ( <i>S.</i> <i>Typhimurium</i> ) strain 14028	Wildtyp		American Type Culture Collection (ATCC)
St53	MG1655 LysP-mCherry	MG1655; <i>lysP::mCherry</i>		Korinna Burdack (LMU, München)

St54	MG1655 CadB-eGFP	MG1655; <i>cadB::eGFP</i>	Korinna Burdack (LMU, München)
St55	<i>E. coli</i> BTH101	F <sup>-</sup> ; <i>cya-99, araD139, galE15, galK16, rpsL1</i> (Str <sup>r</sup> ), <i>hsdR2, mcrA1, mcrB1</i>	Euromedex, Karimova <i>et al.</i> , 2000 <sup>239</sup>

**Tabelle 7:** In dieser Arbeit verwendete Plasmide. Die im Rahmen dieser Arbeit erstellten Plasmide wurden wie unter 2.4.7 beschrieben konstruiert.

	Bezeichnung	Beschreibung	Amplifikation des Inserts	Referenz
V1	pBBR1_MCS2	Kan <sup>R</sup> , pBBR broad host range origin of replication, <i>mob</i> <sup>r</sup>		Kovach <i>et al.</i> , 1995 <sup>240</sup>
V2	pBBR1_MCS2_ <i>lacZ</i>	pBBR1_MCS2; <i>lacZ</i>		Fried <i>et al.</i> , 2012 <sup>241</sup>
V3	pBBR1_MCS2_ <i>lacZ</i> _ P <sub><i>cadC</i></sub>	pBBR1_MCS2_ <i>lacZ</i> ; 500bp <i>cadC</i> promoter region (BamHI/SmaI)	Pr49 + Pr50, template St32	diese Arbeit
V4	pBBR1_MCS2_ <i>lacZ</i> _ P <sub><i>cadC</i></sub> _ <i>envZ</i>	pBBR1_MCS2_ <i>lacZ</i> ; 500bp <i>cadC</i> promoter region (BamHI/SmaI), translational <i>envZ</i> (full length, SmaI/XhoI) - <i>lacZ</i> fusion	Pr49 + Pr50, Pr51 + Pr52, template St32	diese Arbeit
V5	pBBR1_MCS2_ <i>lacZ</i> _ P <sub><i>cadC</i></sub> _ <i>evgS</i>	pBBR1_MCS2_ <i>lacZ</i> ; 500bp <i>cadC</i> promoter region (BamHI/SmaI), translational <i>evgS</i> (full length, SmaI/HindIII) - <i>lacZ</i> fusion	Pr49 + Pr50, Pr53 + Pr54, template St32	diese Arbeit
V6	pBBR1_MCS2_ <i>lacZ</i> _ P <sub><i>cadC</i></sub> _ <i>lon</i>	pBBR1_MCS2_ <i>lacZ</i> ; 500bp <i>cadC</i> promoter region (BamHI/SmaI), translational <i>lon</i> (full length, SmaI/XhoI) - <i>lacZ</i> fusion	Pr49 + Pr50, Pr55 + Pr56, template St32	diese Arbeit
V7	pBBR1_MCS2_ <i>lacZ</i> _ P <sub><i>cadC</i></sub> _ <i>adrA</i> <sub>aa 1-65</sub> (HPPPG)	pBBR1_MCS2_ <i>lacZ</i> ; 500bp <i>cadC</i> promoter region (BamHI/SmaI), translational <i>adrA</i> (1-201, SmaI/XhoI) - <i>lacZ</i> fusion	Pr49 + Pr50, Pr59 + Pr60, template St32	diese Arbeit
V8	pBBR1_MCS2_ <i>lacZ</i> _ P <sub><i>cadC</i></sub> _ <i>adrA</i> <sub>aa 1-65</sub> (HPSPG)	pBBR1_MCS2_ <i>lacZ</i> ; 500bp <i>cadC</i> promoter region (BamHI/SmaI), translational <i>adrA</i> (1-201, C193T, SmaI/XhoI) - <i>lacZ</i> fusion	Pr49 + Pr50, Pr59 + Pr61, template St32	diese Arbeit
V9	pBBR1_MCS2_P <sub><i>cadC</i></sub>	pBBR1_MCS2; 500bp <i>cadC</i> promoter region (BamHI/SmaI)	Pr49 + Pr50, template St32	diese Arbeit
V10	pBBR1_MCS2_P <sub><i>cadC</i></sub> _ <i>evgS</i> -FLAG	pBBR1_MCS2_ <i>lacZ</i> ; 500bp <i>cadC</i> promoter region (BamHI/SmaI), full length <i>evgS</i> -FLAG (SmaI/HindIII)	Pr49 + Pr50, Pr53 + Pr57, template St32	diese Arbeit
V11	pBBR1_MCS2_P <sub><i>cadC</i></sub> _ <i>adrA</i> -FLAG	pBBR1_MCS2_ <i>lacZ</i> ; 500bp <i>cadC</i> promoter region (BamHI/SmaI), full length <i>adrA</i> -FLAG (SmaI/HindIII)	Pr49 + Pr50, Pr59 + Pr66, template St32	diese Arbeit
V12	pBBR1_MCS3	Tet <sup>R</sup> , pBBR broad host range origin of replication, <i>mob</i>		Kovach <i>et al.</i> , 1995 <sup>240</sup>
V13	pBBR1_MCS3_ <i>lacZ</i>	pBBR1_MCS3; <i>lacZ</i>		Fried <i>et al.</i> , 2012 <sup>241</sup>
V14	pBBR1_MCS3_ <i>lacZ</i> _ TF 030_CadC_3Pro_CCG	pBBR1_MCS3_ <i>lacZ</i> ; translational <i>cadC</i> (aa 1-30)-PPP- <i>lacZ</i> fusion		PD Dr. Jürgen Lassak (LMU München)
V15	pBBR1_MCS3_ <i>lacZ</i> _ TF 030_CadC	pBBR1_MCS3_ <i>lacZ</i> ; translational <i>cadC</i> (aa1-30)- <i>lacZ</i> fusion		PD Dr. Jürgen Lassak (LMU München)
V16	pBAD/HisA	Amp <sup>R</sup> , pBBR322 origin, <i>ara</i> <sup>r</sup> , <i>ara</i> operator, 6x His-Tag		Guzman <i>et al.</i> , 1995 <sup>242</sup>
V17	pBAD24/HisA_ <i>lacZ</i>	pBAD/HisA; <i>lacZ</i> under control of P <sub>BAD</sub>		PD Dr. Jürgen Lassak (LMU München)
V18	pBAD/HisA_SP_ <i>lacZ</i>	pBAD/HisA; PPP- <i>lacZ</i> under control of P <sub>BAD</sub>		PD Dr. Jürgen Lassak (LMU München)

V19	pBAD24	Amp <sup>R</sup> , pBBR322 origin, <i>ara</i> <sup>+</sup> , <i>ara</i> operator		Guzman <i>et al.</i> , 1995
V20	pBAD24_EnvZ-FLAG	pBAD24; full length <i>envZ</i> -FLAG (SmaI/HindIII) under control of P <sub>BAD</sub>	Pr51 + Pr58, template St32	diese Arbeit
V21	pBAD24_AdrA-FLAG	pBAD24; full length <i>adrA</i> -FLAG (SmaI/HindIII) under control of P <sub>BAD</sub>	Pr59 + Pr66, template St32	diese Arbeit
V22	pBAD24_AdrA <sup>HPPPG/HPSPG</sup> -FLAG	pBAD24; full length <i>adrA</i> (C196T)-FLAG (SmaI/HindIII) under control of P <sub>BAD</sub>	Pr59 + Pr63/ Pr64 + Pr66, template St32	diese Arbeit
V23	pBAD24_EnvG-FLAG	pBAD24; full length <i>envG</i> -FLAG (SmaI/HindIII) under control of P <sub>BAD</sub>	Pr54 + Pr54, template St32	diese Arbeit
V24	pBAD24_EnvG <sup>S</sup> -PPD/AAD-FLAG	pBAD24; full length <i>envG</i> (C520G,C522A,C523G,C525A) -FLAG (SmaI/HindIII) under control of P <sub>BAD</sub>	Pr53 + Pr20/ Pr21 + Pr54, template St32	diese Arbeit
V25	pFR29	Amp <sup>R</sup> , <i>ompR</i> , <i>envZ</i>		Russo & Silhavy, 1991 <sup>243</sup>
V26	pRedET	Amp <sup>R</sup> , Red/ET expression plasmid		Gene Bridges
V27	pCP 20	Amp <sup>R</sup> , <i>flp</i> , <i>cat</i> , <i>ts-rep</i>		Cherepanov <i>et al.</i> , 1995 <sup>244</sup>
V28	pNPTS138-R6KT	Kan <sup>R</sup> , ori-R6K <i>sacB</i> , suicide plasmid		Lassak <i>et al.</i> , 2010 <sup>245</sup>
V29	pNPTS_EnvZ <sup>IIPPL/IAAAL</sup>	pNPTS138-R6KT; <i>envZ</i> (1-1350; C604G,G606A,C607G,G609A,C910G,G612A)(PstI/BamHI)	Pr27 + Pr16/ Pr17 + Pr28, template St32	diese Arbeit
V30	pNPTS_EnvZ <sup>IIPPL/IAAAL</sup> _500bp+	pNPTS138-R6KT; <i>envZ</i> (-508 to +501; C604G,G606A,C607G,G609A,C910G,G612A)(PstI/BamHI)	Pr29 + Pr16/ Pr17 + Pr30, template St32	diese Arbeit
V31	pNPTS_Evg <sup>S</sup> <sup>NYPPD/NYAAD</sup> _500bp+	pNPTS138-R6KT; <i>evgS</i> (-503 to +476; C520G,C522A,C523G,C525A) (PstI/BamHI)	Pr32 + Pr20/ Pr21 + Pr33, template St32	diese Arbeit
V32	pNPTS_Lon <sup>VGGPG/VGAAG</sup> _500bp+	pNPTS138-R6KT; <i>lon</i> (-449 to +484; C1069G,G1071A,C1072G,G1074A) (EcoRV/EcoRI)	Pr34 + Pr24/ Pr25 + Pr35, template St32	diese Arbeit
V33	pNPTS_EnvZ_PAgfp	pNPTS138-R6KT; <i>envZ</i> (1-1350)(PstI/BamHI)-(GCTGCC)- <i>pagfp</i> (full length)(BamHI/EcoRI)	Pr38 + Pr39 Pr36 + Pr37 template St32, pNPTS_PAgfp	diese Arbeit
V34	pNPTS_EnvG_PAgfp_k	pNPTS138-R6KT; <i>evgS</i> (3070-3591)(PstI/BamHI)-(GCTGCC)- <i>pagfp</i> (full length)(BamHI/EcoRI)	Pr40 + Pr41 Pr36 + Pr37 template St32, pNPTS_PAgfp	diese Arbeit
V35	pNPTS_EnvZ_mNeonGreen_down	pNPTS138-R6KT; <i>envZ</i> (1-1350)-(GCTGCC)- <i>mNeonGreen</i> (full length)- <i>envZ</i> downstream (611 bp)		Dr. Dagmar Koschnitzki
V36	pNPTS_mzrA_PAmCherry_down	pNPTS138-R6KT; <i>mzrA</i> (1-381)(PstI/BamHI)-(GCTGCC)- <i>pamcherry</i> (full length)- <i>mzrA</i> downstream (+1 to +488)(BamHI/EcoRI)	Pr42 + Pr43 Pr44 + Pr45/Pr47 + Pr48 template St32, pNPTS_PAmCherry	diese Arbeit
V37	pNPTS_AdrA-FLAG	pNPTS138-R6KT; <i>adrA</i> (476-1113)-FLAG (PstI/BamHI)	Pr67 + Pr68 template St32	diese Arbeit
V38	pNPTS_AdrA-PPP/PSP	pNPTS138-R6KT; <i>adrA</i> (C196T; -494 to 1048) (PstI/BamHI)	Pr67 + Pr62/Pr63 + Pr65 template St32	diese Arbeit
V39	pKT25	Kan <sup>R</sup> , <i>cyaA/T25</i> (aa 1-224) for C-terminal in frame fusions, under the control of the <i>lac</i> promoter		Euromedex
V40	pKT25_MzrA	pKT25; <i>mzrA</i> (full length)-AA (PstI/BamHI)	Pr75 + Pr76 template St32	diese Arbeit
V41	pKT25_EnvZ	pKT25; <i>envZ</i> (full length)-AA (PstI/BamHI)	PCR: Pr79 + Pr78 template St32	diese Arbeit
V42	pKT25_EnvZ <sup>P/A (p)</sup>	pKT25; <i>envZ</i> (full length, C217G, T219A, C220G, C2222A)-AA (PstI/BamHI)	Pr79 + Pr73/Pr74 + Pr78 template St32	diese Arbeit
V43	pKT25_EnvZ <sup>P/A (c)</sup>	pKT25; <i>envZ</i> (full length, C604G,G606A,C607G,G609A,C910G,G612A)-AA (PstI/BamHI)	Pr79 + Pr16/Pr17 + Pr78 template St32	diese Arbeit
V44	pKT25_EnvZ <sup>P/A (p), P/A (c)</sup>	pKT25; <i>envZ</i> (full length, C217G, T219A, C220G, C2222A, C604G,G606A,C607G,G609A,C910G,G612A)-AA (PstI/BamHI)	Pr79 + Pr73/Pr73 + Pr74/ Pr74 + Pr78 template St32	diese Arbeit

V45	pUT18C	Amp <sup>R</sup> , <i>cyoA/T18</i> (aa 225-339) for C-terminal in frame fusions, under the control of the <i>lac</i> promoter		Euromedex
V46	pUT18_EnvZ	pUT18C; <i>envZ</i> (full length)-A (PstI/BamHI)	Pr77 + Pr78 template St32	diese Arbeit
V47	pUT18_EnvZ <sub>P/A (p)</sub>	pUT18C; <i>envZ</i> (full length, C217G, T219A, C220G, C2222A)-A (PstI/BamHI)	Pr77 + Pr73/Pr74 + Pr78 template St32	diese Arbeit
V48	pUT18_EnvZ <sub>P/A (c)</sub>	pUT18C; <i>envZ</i> (full length, C604G, G606A, C607G, G609A, C910G, G612A)-A (PstI/BamHI)	Pr77 + Pr16/Pr17 + Pr78 template St32	diese Arbeit
V49	pUT18_EnvZ <sub>P/A (p), P/A (c)</sub>	pUT18C; <i>envZ</i> (full length, C217G, T219A, C220G, C2222A, C604G, G606A, C607G, G609A, C910G, G612A)-A (PstI/BamHI)	Pr77 + Pr73/Pr73 + Pr74/ Pr74 + Pr78 template St32	diese Arbeit
V50	pKT25- <i>zip</i>	translational <i>T25-<i>zip</i></i> fusion		Euromedex
V51	pUT18C- <i>zip</i>	translational <i>T18-<i>zip</i></i> fusion		Euromedex
V52	pET_SUMO	Kan <sup>R</sup> , expression plasmid, T7 promoter, SUMO tag		Invitrogen
V53	pET_SUMO_OxyR	pET_SUMO; His-SUMO- <i>oxyR</i> (full length)	Pr69 + Pr72 template St32	diese Arbeit
V54	pET_SUMO_OxyR-C199S	pET_SUMO; His-SUMO- <i>oxyR</i> (G596C) (full length)	Pr69 + Pr70/Pr71 + Pr72 template St32	diese Arbeit

**Tabelle 8:** In dieser Arbeit verwendete Oligonukleotide. Für die Klonierung relevante Nukleotidsequenzen wie Restriktionsschnittstellen oder mit komplementären DNA-Sequenzen überlappende Bereiche sind rot hervorgehoben. Oligonukleotide, die im Rahmen der Untersuchung von DNA-bindenden Transkriptionsfaktoren verwendet wurden, sind durch die Vorsilbe „B“ gekennzeichnet. Oligonukleotide die für die Quantifizierung von Genexpressionsleveln im Rahmen einer qPCR eingesetzt wurden, sind mit der Vorsilbe „Q“ gekennzeichnet.

Nr.	Bezeichnung	Sequenz 5' - 3' (inklusive Modifikation, falls vorhanden)	Restriktionsschnittstelle
Pr1	envZ_rpslKK_fwd	TTCATTTGCCCGTACGTTATTGCTCATCGTCACCTTGCTGTTCCG CCAGCC <b>AATTAACCTCACTAAAGGGCG</b>	
Pr2	envZ_rpslKK_rev	TACCGGCACTGGCAGCCAGGCGCAATGGAAAGCCCGCCCCG CTCGCTGG <b>TAATACGACTCACTATAGGGCTC</b>	
Pr3	envZ_rpsLKK_fwd_EPB	TATCCAGAACCGACCGTTGGTTCGATCTCGAACACGCAGCCTTG CAGGTTG <b>AATTAACCTCACTAAAGGGCG</b>	
Pr4	envZ_rpsLKK_rev_EPB	CTTAACACCAGCCGCATATGGTTAAAGGCACGGGTAACGGA ACGCACCT <b>TAATACGACTCACTATAGGGCTC</b>	
Pr5	EvgS_rpsLKK_fwd	TGTTGTGGTCTTTGGTTCGACCATAAGTTTCGCAGACGAAGATT ACATCGA <b>AATTAACCTCACTAAAGGGCG</b>	
Pr6	EvgS_rpsLKK_rev	AGCTGAGACGGATGCTAATGCCTGATATAAATTTGTAAGAG ATAATTG <b>TAATACGACTCACTATAGGGCTC</b>	
Pr7	evgS_KK_FLAG_fwd	GCTGGACCAGGAGATTGCTGTTTTCTGTCAGAAAAATGACgacta caaggacgacgatgacaagTA <b>AGGCTGGTGATGATGGCGGGATCGTTG</b> <b>TATATTTT</b>	
Pr8	evgS_KK_rev	ATGGCCGCGAAAAGAGCCAGAGAGGGGCGCATGTCGATATAG CGATGTTT <b>TCAGAAGAACTCGTCAAGAAGGGCG</b>	
Pr9	lon_rpslKK_fwd	ATCCCCGTATTGCCGCTGCGCATGTGGTGGTTATCCGCACAT GGTCATA <b>AATTAACCTCACTAAAGGGCG</b>	
Pr10	lon_rpslKK_rev	GTTCAATTTGCAGCGCCAGAGTCAGAATTCCTCAATGCGCTTC ACAGGAT <b>TAATACGACTCACTATAGGGCTC</b>	
Pr11	adrA_rpsLKK_fwd	CTCACCCGAAATATCATACTTCTGCCTTTAGCTCCGTCTCTATA ATTTGG <b>AATTAACCTCACTAAAGGGCG</b>	
Pr12	adrA_rpsLKK_rev	GCCAGACGAACGCCAGCCGACCAACACCAGCCACCACCATA <b>ATACGACTCACTATAGGGCTC</b>	
Pr13	nagC_KK_fwd	ACCAAGACCATCGTTAAACGGTAACGAGGTCGTAACCTCAATAAG AGAAAGT <b>GGCTGGTGATGATGGCGGGATCG</b>	

Pr14	nagC_KK_rev	CATGGACTACCCAGAATATTGACAACAATAAGCGCCACTATAA AAGCACATCAGAAGAACTCGTCAAGAAGGCCG	
Pr15	envZ_upstream_fwd	atgAGGCGATTGCGCTTCTCGCCACGAAG	
Pr16	envZ_upstream_rev	ACTCACGCAGTGCTGCTGCAATAATCCC	
Pr17	envZ_downstream_fwd	AGGGATTATTGCAGCAGCACTGCGTGAGT	
Pr18	envZ_downstream_rev	CCTTCTTTTGTGCTGCCCTGCGCCCGCGT	
Pr19	evgS_upstream_fwd	atgAAGTTTTTACCCTATATTTTTCTTCTC	
Pr20	evgS_upstream_rev_B	ACCTCGTCTGCTGCGTAATTTGCTACCCG	
Pr21	evgS_downstream_fwd_B	ATTACGCAGCAGACGAGGTAATTCATCAA	
Pr22	evgS_downstream_rev	GTCATTTTCTGACAGAAAACAGCAATCTCC	
Pr23	lon_upstream_fwd	atgAATCCTGAGCGTTCTGAACGCATTGAA	
Pr24	lon_upstream_neu_rev	GTTTTACCTACCCTGCTGCCCTACCAGG	
Pr25	lon_downstream_fwd	GCCTGGTAGGGGCAGCAGGGGTAGGTAAAAC	
Pr26	lon_downstream_rev	TTTTGCAGTCACAACCTGCATACCAGACG	
Pr27	envZ_fwd_500bp_up	ATCTGCAGATGAGGCGATTGCGCTTCTC	PstI
Pr28	envZ_rev_500bp_down	ATGGATCCCTTCTTTTGTGCTGCCCTGC	BamHI
Pr29	envZ_up_fwd	ATCTGCAGCAGAGCAACCCGATGCCG	PstI
Pr30	envZ_down_rev	ATGGATCCGAACCGACGCCAACCTTCTC	BamHI
Pr31	envZ_rev	ATGGATCCCTTCTTTTGTGCTGCCCTGCGCCCGCGT	BamHI
Pr32	evgS_fwd_500bp_up	ATCTGCAGGCGGGTGAAAACACTTAAGCCTGA	PstI
Pr33	evgS_rev_500bp_down	ATGGATCCCAATACTCAGGATAAGCCGATTACG	BamHI
Pr34	lon_fwd_500bp_up	ATGATATCCTGATTAGATCCTCAAAGAGCCG	EcoRV
Pr35	lon_rev_500bp_down	ATGAATTCGCTTAGTTTACCGCGTCTTTCAG	EcoRI
Pr36	PAGFP_fwd	ATGGATCCATGGTGAGCAAGGGCGAG	BamHI
Pr37	PAGFP_rev	ATGAATTCCTTACTTGTACAGCTCGTCCATGC	EcoRI
Pr38	envZ_fwd	ATCTGCAGATGAGGCGATTGCGCTTCTCGCCACG	PstI
Pr39	envZ_GS_rev	ATGGATCCGCTGCCCTTCTTTTGTGCTGCCCTG	BamHI
Pr40	evgS_fwd_kurz	ATCTGCAGCTCCGTGAGCAAAAATTCTTCCTTAC	PstI

Pr41	evgS_GS_rev	ATGGATCC GCTGCCGTCATTTTTCTGACAGAAAACAG	BamHI
Pr42	mzrA_fwd	ATCTGCAGATGCAAATACCTCGCATGTCGCTTCG	PstI
Pr43	mzrA_GS_rev	ATGGATCC GCTGCC TCCGAAGCGATGAGAGTTAT	BamHI
Pr44	PAmCherry_fwd	ATGGATCCATGGTGAGCAAGGGCGAG	BamHI
Pr45	PAmCherry_rev_mzrA	GATAACAGGAAGTTACTTGTACAGCTCGTCCATGC	
Pr46	PAmCherry_rev_torS	GCTGATATTGATTTACTTGTACAGCTCGTCCATGC	
Pr47	mzrA_down_fwd_mCher	CTGTACAAGTAACTTCTGTATCCGAAACATATC	
Pr58	mzrA_down_rev	ATGAATTCAGCTCTTCCTGCGCTTCTGCCA	EcoRI
Pr49	cadC_Prom_fwd	ATGGATCCAGCAATTTGCGGGCATATCCCAA	BamHI
Pr50	cadC_Prom_rev	ATCCCGGGAATAGAACTCATTGAAAAGGG	SmaI
Pr51	envZ_fwd	ATCCCGGgatgAGGCGATTGCGCTTCTCGCC	SmaI
Pr52	envZ_rev	ATCTCGAGTTCTTTTGTGTCGTGCCCTGCGC	XhoI
Pr53	evgS_fwd	ATCCCGGgatgAAGTTTTTACCCTATATTTT	SmaI
Pr54	evgS_rev	ATAAGCTTCATTTTTCTGACAGAAAACAGC	HindIII
Pr55	lon_fwd	ATCCCGGgatgAATCTGAGCGTTCTGAAC	SmaI
Pr56	lon_rev	ATCTCGAGTTGCAGTCACAACCTGCATACC	XhoI
Pr57	evgS_rev_FLAG_Stopp	ATAAGCTTTTAcctgtcatcgtcctttagtcGTCATTTTTCTGACAGAAA ACAGCAATCTC	HindIII
Pr58	envZ_rev_FLAG_Stopp	ATAAGCTTTTAcctgtcatcgtcctttagtcCCCTTCTTTTGTGTCGTGCCCT GCGCCC GCGTTAC	HindIII
Pr59	adrA_fwd	ATCCCGGATGTTCCCAAAAATAATGAATGATGAAAAC	SmaI
Pr60	adrA_WT__rev	ATCTCGAGGCCCGCGCGGGTGTGAAAC	XhoI
Pr61	adrA_UTI89_rev	ATCTCGAGGCCCGCGCACGGGTGTGAAAC	XhoI
Pr62	adrA_up_fwd	ATCTGCAGGTTAGCAACCAGACTTAATGGCAGATCAC	PstI
Pr63	adrA_up_rev	CACCAGCCCGGCGACGGGTGTG	
Pr64	adrA_down_fwd	CCGGGCTGGTGGTGGCTGGTGTGGT	
Pr65	adrA_down_rev	ATGGATCCTCAACCACTCACGATAGTACTCATTGTG	BamHI
Pr66	adrA_rev_FLAG_Stopp	ATAAGCTTTTCAcctgtcatcgtcctttagtcGGCCGCCACTTCGGTGGC GTTACGTCCGGC	HindIII
Pr67	adrA_fwd	AAATCTGCAGCTGACGGGCATTACCGTGTGTTTC	PstI

Pr68	adrA_rev_FLAG_Stopp	ATGGATCCTCActgtcatcgtccttgtagcGGCCGCCACTTCGGTGCG GTTACGTCCGGC	BamHI
Pr69	oxyR_up_fwd	ATGAATATTCGTGATCTTGAGTAC	
Pr70	oxyR_up_rev_II	GCCTGATCGCGCAAAGAGTGACCATCTTC	
Pr71	oxyR_down_fwd_II	AGATGGTCACTCTTTGCGCGATC	
Pr72	oxyR_down_rev	TTAAACCGCCTGTTTTAAAACTTTATC	
Pr73	EnvZ_up_rev_VVPPA	GAAAGCTGCTGCCACAACCAACTGCGTGCCGT	
Pr74	EnvZ_down_fwd_VVPPA	GTTGTGGCAGCAGCTTCCGTCGGGAGATCTA	
Pr75	MzrA_fwd_PstI+2	ATCTGCAGAAatgCAAATACCTCGCATGTCGCTTCG	PstI
Pr67	MzrA_rev_BamHI	ATGGATCCttaTCCGAAGCGATGAGAGTTATCCCG	BamHI
Pr77	EnvZ_fwd_PstI+1	ATCTGCAGAatgAGGCGATTGCGCTTCTC	PstI
Pr78	EnvZ_rev_BamHI	ATGGATCCttaCCCTTCTTTTGTCGTGCCCT	BamHI
Pr79	EnvZ_fwd_PstI+2	ATCTGCAGAAatgAGGCGATTGCGCTTCTC	PstI
Pr80	B_epmA_oxyR_fwd	ATTATCACTGAAGATGATTAATTTAATTACTAAACCATCAGAT CGTGCTTTTTTTAGTCA	
Pr81	B_epmA_oxyR	CGACAGCGGCGGAGCCCCCTTC	
Pr82	B_epmA_400bp	CGACAGCGGCGGAGCCCCCTTCTG	
Pr83	B_epmA_-35_fwd	ATAGGTACAACAGTATAGTCTGATGGATAAGTCTGAAATTTGA CGAGATCGATTTTTTTA	
Pr84	B_epmA_-35_rev_Btn	[Btn] TAAAAAATCGATCTCGTCAAATTCAGACTTATCC ATCAGAC TATACTGTTGTACCTAT	
Pr85	B_epmA_fwd_Btn	[Btn] GCGCGCAGCAATTGCAGCACG	
Pr86	B_epmA_rev	AGTTAAATCTCCAGTTTTTGACAAGG	
Pr87	B_epmA_fwd	CCCACAGCCACGTACTTCAGGGTAAG	
Pr88	B_epmA_rev_6FAM	[6FAM] AATTTAGCCCTCTTGCGCACTAAAAAATC	
Pr89	B_epmA_400_fwd_6FAM	[6FAM] GCGCGCAGCAATTGCAGCACG	
Pr90	B_epmA_400_rev_6FAM	[6FAM] TGAATAAAAAAGCACGATCTG	
Pr91	B_katG_fwd_Btn	[Btn] TTCTATACGCTTAACCGTGCTGAAATG	
Pr92	B_katG_rev	CAATGTGCTCCCTCTACAGTGTTACC	
Pr93	B_katG_fwd	GACTTCAAGGGTAAAAAGAGGTG	
Pr94	B_katG_rev_6FAM	[6FAM] TCAAGATCACGAATATTCATTATC	

Pr95	B_katG_fwd_6FAM	[6FAM]TTCTATACGCTTAACCGTGCTGAAATG	
Pr96	B_katG_fwd_150bp_6FAM	[6FAM]CAAAAAAGCTTAATTAAGATC	
Pr97	B_epmA_P_NK_rev_6FAM	[6FAM] TAAATCTCCAGTTTTTGACAAGGGC	
Pr98	B_iolR_400_rev_6FAM	[6FAM]CTGCGGATATCATCCTGCAGTAAGG	
Pr99	B_iolR_400_fwd	CATATCCTTAAAGTAATCAGTGT	
Pr100	B_iolR_150_fwd	GAAGGTAACCACAGCGCAATTT	
Pr101	Q_alaS_fwdII	TCGCTAACCAACAGCTTAACACC	
Pr102	Q_alaS_revII	GCTGAAAGGCGATAGCAATAATC	
Pr103	Q_arcA_fwd	GAAGAACGTCGTAGCGTTGAAAG	
Pr104	Q_arcA_rev	GTCACGTCTACAGTACGGTCGTG	
Pr105	Q_efp_fwd	TAAAGGCCAGGCATTTGCT	
Pr106	Q_efp_rev	GCTGCTCGAAAGTTTCGTTG	
Pr107	Q_epmA_fwd	GTAATGAAGAGATGGGGCGTTAT	
Pr108	Q_epmA_rev	GTCAATTTCCAGATAACGCAAGA	
Pr109	Q_epmB_fwd	ATCGCATGGAGAAAGCAA	
Pr110	Q_epmB_rev	CCGCCTTTGACCAGCAA	
Pr111	Q_oxyR_fwd	ACATTGGTTTGATTCCCACAGTT	
Pr112	Q_oxyR_rev	CATAGATAGCCAGCAACATTGGC	
Pr113	Q_lexA_fwd	TTTGATCTCATCCGTGATCACAT	
Pr114	Q_lexA_rev	CTTCTTCTCTTCCTGCAACAGA	
Pr115	Q_katG_fwd	TAACGTAAGCCTCGATAAAGCG	
Pr116	Q_katG_rev	CAGTTAACATCCAGATCCGGTTC	
Pr117	Q_recN_fwd	TATGCTGGAAGAAGCTACCATCC	
Pr118	Q_recN_rev	CAGTTGCTGCTGTTCTCCAGTA	
Pr119	Q_uspE_fwd	CATACGAAATGACCACCCTGC	
Pr120	Q_uspE_rev	GATCACTTCTGAATGATGGCTT	
Pr121	Q_gadE_fwd	TTCCCTGTCAGAGATCAAAAAAGTAG	
Pr122	Q_gadE_rev	CAAAATAAATAGGCGCTTAGCTTTTA	
Pr123	Q_ydeO_fwd	GTGACGGTGCCTGTTCTCG	

Pr124	Q_ydeO_rev	TGTAATGAAACCTTTGCATGCG	
Pr125	Q_ydeP_fwd	CATTCGACATTTGTGAAAACGG	
Pr126	Q_ydeP_rev	GTAACAGTCGCTGACGGCATC	
Pr127	Q_safA_fwd	CGACCACAGTGAAAAACAAAATC	
Pr128	Q_safA_rev	CTATTGATAATGAAATGTGTAAGATCTCTTG	
Pr129	Q_emrK_fwd	GAATTAGAAGACATGATTAGTACAGATGACG	
Pr130	Q_emrK_rev	CTTTATTGAGTGGGATAGTGGCAT	
Pr131	Q_yfdX_fwd	CACGATGGTCACAGCAATTCT	
Pr132	Q_yfdX_rev	CATGAAACAGAGCCAGGCG	
Pr133	Q_evgsS_fwd	GTGGCATCAGTAGTAACAACCGTG	
Pr134	Q_evgsS_rev	AGCATTAAATACCACGAACCCG	
Pr135	Q_ulaR_fwd	CTCGCACAATTGGGCTTTGTG	
Pr136	Q_ulaR_rev	CTGATGCAGATTCATCGGCG	
Pr137	Q_purR_fwd	GTCGCTGAAGAAACGCGCA	
Pr138	Q_purR_rev	TTCTGGAAGCAATTTTTTCAACTGC	
Pr139	Q_nagC_fwd	GTAACCAAAATTACGCGTCAGCTTATC	
Pr140	Q_nagC_rev	CTTTGCTGCTGAGATCAAACAGAG	
Pr141	Q_argP_fwd	CAGAACATTACAGGCACTGGATGC	
Pr142	Q_argP_rev	CAGTTTTTGGCCCTTGTTCGTC	
Pr143	Q_pgrR_fwd	GTTTGTCTGTCGTTGCAGAGGAG	
Pr144	Q_pgrR_rev	GCCAAGAACAGACAAAAGATGCTC	
Pr145	Q_yfhH_fwd	CTGGCGGATTATCTGCTGCTAC	
Pr146	Q_yfhH_rev	ATCCGGCTGGCTTGCC	

### 2.3 Kultivierungsverfahren

Dem jeweiligen Versuchsaufbau entsprechend und sofern nicht anders beschrieben wurden *E. coli* Stämme bei 37°C in LB Medium<sup>246</sup> [1% (w/v) Bacto-Trypton; 0,5 % (w/v) Hefeextrakt; 171 mM NaCl], M9 Medium<sup>246</sup> [45 mM Na<sub>2</sub>HPO<sub>4</sub>·2H<sub>2</sub>O; 22 mM KH<sub>2</sub>PO<sub>4</sub>; 8,5 mM NaCl; 18,7 mM NH<sub>4</sub>Cl; 2 mM MgSO<sub>4</sub>; 22 mM Glucose; 0,1 mM CaCl<sub>2</sub>], KE-Medium<sup>247</sup> [8,5 mM K<sub>2</sub>HPO<sub>4</sub>; 91,5 mM KH<sub>2</sub>PO<sub>4</sub>; 2,3 mM Na<sub>3</sub>C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>; 8,0 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 6 µM FeSO<sub>4</sub>; 400 µM MgSO<sub>4</sub>; 22 mM Glucose; pH = 5,8], Citrat-gepuffertem Medium [„CBM“: siehe KE-Medium, aber 100 mM Citratpuffer anstatt Phosphatpuffer: 45,6 mM C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>x2H<sub>2</sub>Ox3Na; 54,4 mM C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>; pH = 4,5] oder GTM Medium [1,32 mM K<sub>2</sub>HPO<sub>4</sub>; 0,4% v/v Glycerol; 1% w/v Trypton; 40mM MOPS pH 7,4; 4 mM Tricine pH 7,4; 10 µM FeSO<sub>4</sub>; 9,5 mM NH<sub>4</sub>Cl; 0,276 mM K<sub>2</sub>SO<sub>4</sub>; 0,5 µM CaCl<sub>2</sub>; 0,528 µM MgCl<sub>2</sub>; 50 mM NaCl; Spurenelemente: 3,1 nM (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>; 403,2 nM HBO<sub>3</sub>; 53,8 nM CoCl<sub>2</sub>; 15,6 nM CuSO<sub>4</sub>; 126,4 µM MnCl<sub>2</sub>; 18,1 µM ZnSO<sub>4</sub>] kultiviert. *B. subtilis* Stämme wurden aerob in Reagenzgläsern bei 30°C in LB Medium oder MD Medium<sup>248</sup> [60 mM K<sub>2</sub>HPO<sub>4</sub>; 44 mM KH<sub>2</sub>PO<sub>4</sub>; 4 mM Na<sub>3</sub> Citrat, 111 mM Glucose; 0,2 mM L-Tryptophan; 75 mM Ammonium Eisen(III)-citrat; 150 mM L-Aspartat; 3 mM MgSO<sub>4</sub>; 1 mg/ml Casamino Säuren] kultiviert. Sämtliche Medien wurden vor Verwendung durch Autoklavieren oder Sterilfiltrieren (*Roth*: Rotilab® Spritzenfilter; PVDF; 0,22 µm Porengröße) sterilisiert. Die Kultivierung von Flüssigkulturen erfolgte in erster Linie aerob in Reagenzgläsern oder Schikanekolben. Mikroanaerobe Wachstumsbedingungen wurden durch Kultivierung in luftdicht verschlossenen Hungate-Röhrchen simuliert. Hierfür wurde der Sauerstoff zuvor mittels Gasaustausch durch ein Gemisch aus 2% O<sub>2</sub> und 98% N<sub>2</sub> ersetzt.

Die optische Dichte einer Kultur (OD<sub>600</sub>) wurde mit Hilfe des „Ultraspec 2100 pro“-Photometers (*Amersham Biosciences*) bestimmt. Zur Herstellung von Nährböden wurde LB Medium mit 1,5% (w/v) Agar versetzt. Aus Gründen der Selektion wurden den Medien folgende Antibiotika zugefügt: Kanamycin (25 µg/ml für chromosomal kodierte Resistenzmarker, 50 µg/ml für plasmidkodierte Resistenzmarker), Ampicillin (100 µg/ml), Ofloxacin (100 µg/ml), Chloramphenicol (10 µg/ml), Streptomycin (50 µg/ml) oder Tetracyclin (10 µg/ml). Das Medium zur Kultivierung von WM3064 Zellen wurde durch 300 µM DAP ergänzt.

Zur Simulation verschiedener Stressbedingungen wurden je 10 ml einer *E. coli* Zellkultur wie in Tabelle 9 zusammengefasst, kultiviert. Die Vorkultivierung fand jeweils aerob bei 37°C statt. Die Zellen wurden bei Versuchsbeginn 1:1.000 in frisches Medium überführt.

Bis in die exponentielle Wachstumsphase hinein bei 37°C kultivierte *E. coli* Stämme wurden durch anschließende Langzeitkultivierung bei 4°C in LB Medium in einen Ruhezustand versetzt, aus welchem die Zellen nicht ohne vorherige Induktion rekultivierbar sind (VBNC: *viable but non-culturable*, vgl. 2.8.3). Dauerkulturen wurden in Glycerin-Freezing-Medium angesetzt [65% (v/v) Glycerin; 16 mM KCl; 16 mM NaCl; 0,2 mM MgSO<sub>4</sub>]. Dafür wurden in LB-Medium gewachsene Übernachtskulturen 1:4 verdünnt, in flüssigem Stickstoff schockgefroren und bei -80°C gelagert.

Um den Effekt von osmotischem Stress zu testen, wurde M9 Medium mit NaCl (0,0 bis 0,4 M NaCl) oder Saccharose (0,4 oder 0,8 M) versetzt. Die Osmolalität dieser Medien wurde mit einem Osmometer (*Gonotec*, „Osmomat 030“) gemessen.

**Tabelle 9:** Simulation verschiedener Stressbedingung für das Screening nach Einflussfaktoren der *efp*, *epmA* und *epmB* Expression. I = Induktion der Genexpression mit L-Arabinose

<i>cadC-lacZ</i> Reportergenassay unter Kontrolle des $P_{cadC}$ :		experimentelle Durchführung	
Wachstumsbedingung	Medium	Vorkultivierung	Stress
Kontrolle	M9	3,0 h	-
Kältestress	M9	2,5 h	1,0 h 4°C
Hitzestress	M9	2,0 h	0,25 h 65°C; 0,25 h 37°C
Säure-Stress	M9; pH = 4,5	2,0 h	
Basen-Stress	M9; pH = 10,0	2,75 h	0,25 h
Osmotischer Stress	M9; 0,6 M NaCl	2,0 h	1,0 h
N-Limitation	M9; 0,0 M NH <sub>4</sub> Cl	-	3,0 h
C-Limitation	M9; 0,02% Glucose	-	3,0 h

<i>cadC-lacZ</i> Reportergenassay unter Kontrolle des $P_{BAD}$ :		experimentelle Durchführung	
Wachstumsbedingung	Medium	Vorkultivierung	Stress
Kontrolle	M9	2,0 h - I - 1,0 h	-
Kältestress	M9	2,0 h	4°C: 10 min - I - 30 min
Hitzestress I	M9	2,0 h	42°C: 15 min - I - 1,0 h
Hitzestress II	M9	2,0 h	65°C: 5 min - I - 37°C: 30 min
Säure-Stress	M9; pH = 4,5	2,0 h	15 min - I - 1,0 h
Basen-Stress	M9; pH = 10,0	2,75 h	10 min - I - pH 7,5: 30 min
Osmotischer Stress	M9; 0,6 M NaCl	2,0 h	15 min - I - 1,0 h
N-Limitation	M9; 0,0 M NH <sub>4</sub> Cl	2,0 h	15 min - I - 1,0 h

Proteinexpression EF-P, EpmA und EpmC:		experimentelle Durchführung	
Wachstumsbedingung	Medium	Vorkultivierung	Stress
Kontrolle	LB	3,75 h	-
Osmotischer Stress	LB; 0,6 M NaCl	3,0 h	0,75 h
Säure-Stress	LB; pH = 5,8 (gepuffert)	-	5,75 h
Oxydativer Stress	LB + 30 mM H <sub>2</sub> O <sub>2</sub>	3,0 h	0,75 h
UV-Stress	LB	3,0 h	UV-Licht (253 nm): 20 min; w/o: 40 min

Genexpression <i>efp</i> , <i>epmA</i> und <i>epmC</i> :		experimentelle Durchführung	
Wachstumsbedingung	Medium	Vorkultivierung	Stress
Kontrolle	LB	3,0 h	-
frühe exponentielle Phase	LB	2,0 h	-
stationäre Phase	LB	6,0 h	-
Absterbe-Phase	LB	12,0 h	-
VBNC, Tag 120	LB	4,0 h (OD600 = 1)	120 Tage, 4°C
Oxydativer Stress	LB + 1 mM H <sub>2</sub> O <sub>2</sub>	3,0 h	45 min
DNA-Schädigung	LB + 1 µg/ml Mitomycin c	2,0 h	2, 0 h; UV-Licht 253 nm
Säure-Stress Ia	LB; pH = 5,8	-	3,0 h
Säure-Stress Ib	KE; pH = 5,8	-	3,0 h
Säure-Stress IIa	LB; pH = 4,5	-	3,0 h
Säure-Stress IIb	CBM; pH = 4,5	-	3,0 h
Hitzestress	LB	3,0 h	1,0 h
Osmotischer Stress Ia	M9 + 0,2 M NaCl	-	5,0 h
Osmotischer Stress Ib	M9 + 0,4 M Saccharose	-	5,0 h
Osmotischer Stress II	M9 + 0,4 M NaCl	-	5,0 h
C-Limitation	M9; 0,02% Glucose	-	5,0 h
N-Limitation	M9; 0,0 M NH <sub>4</sub> Cl	-	5,0 h
Kontrolle	M9	5,0 h	-

## 2.4 Molekularbiologische und gentechnologische Methoden

### 2.4.1 Isolierung von Plasmiden

Plasmidpräparationen wurden entsprechend der Herstelleranweisung des Kits „HiYield Plasmid Mini-Kit“ (*Südlaborbedarf Gauting*) durchgeführt. Isolierte Plasmide wurden in H<sub>2</sub>O<sub>dd</sub> gelöst und bei -20°C oder -80°C gelagert.

### 2.4.2 Isolierung von genomischer DNA

Genomische DNA wurde mittels Phenol/Chloroform Extraktion wie folgt isoliert. Zuvor pelletierte Zellen (5 ml) wurden in 1,7 ml TE-Puffer [10 mM Tris; 1mM EDTA; pH 8] mit 0,5% SDS und 180 µg Proteinase K aufgenommen und eine Stunde lang bei 37°C inkubiert. Nach Zugabe von 300 µl NaCl (5 M) und 240 µl CTAB/NaCl-Lösung [0,7 M NaCl; 10% CTAB (w/v)] folgte eine erneute Inkubation für 10 Minuten bei 65°C. Nach Zugabe von 2 ml Chloroform/Isoamylalkohol (24:1) wurden die Proben bei 15.000 x g zentrifugiert. Die wässrige Phase wurde abgenommen und mit dem 0,6-fachen Volumen an Phenol/Chloroform/Isoamylalkohol (25:24:1) versetzt. Nach erneuter Zentrifugation (15.000 x g, 5 Minuten, 4°C) und Phasentrennung wurde die DNA durch Zugabe des 0,6-fachen Volumens an Isopropanol präzipitiert. Nach Zentrifugation (15.000 x g, 30 Minuten, 4°C) wurde das DNA Pellet zweimal mit Ethanol gewaschen (15.000 x g, 10 Minuten, 4°C) und schließlich in H<sub>2</sub>O<sub>dd</sub> gelöst (Endkonzentration ~ 1mg/ml). Gelagert wurde die DNA bei -20 oder -80°C.

### 2.4.3 Isolierung von RNA

Ausgangsmaterial ist ein Zellpellet, das ca. 12 ml einer Flüssigkultur mit OD<sub>600</sub>~0,3 entspricht. Nach Resuspendierung in 500 µl eiskaltem AE-Puffer [20 mM Na(CH<sub>3</sub>COO); 1 mM EDTA; pH 5,2] wurden 500 µl eines 60°C warmen Phenol/Chloroform/Isoamylalkohol-Gemisches (25:24:1) hinzugefügt und der Ansatz für 5 Minuten bei 60°C auf einem Schüttler (1.000 rpm) inkubiert. Nach zweistündiger Abkühlung auf Eis folgte eine Zentrifugation (15.000 x g, 30 Minuten, 4°C). Der Überstand wurde in „phase lock tubes“ (Sigma) überführt und 1/10 Volumen Natriumacetat [3M; pH 5,2], sowie ein äquivalentes Volumen Phenol/Chloroform/Isoamylalkohol hinzugefügt. Nach Abtrennung der wässrigen Phase mittels Zentrifugation (15 Minuten, 4°C, 14.000 rpm) wurde diese in ein frisches Eppendorf-Gefäß überführt. Die Präzipitation erfolgte durch Zugabe von EtOH (70% Endkonzentration) bei -80°C über Nacht. Nach Zentrifugation (15.000 x g, 30 Minuten, 4°C) wurde das Pellet zweimal mit eiskaltem Ethanol (80%) gewaschen und anschließend bei Raumtemperatur getrocknet, in 100-200 µl DEPC-Wasser (1ml DEPC/l; 12h Inkubation bei 37°C vor Autoklavieren) resuspendiert und bei -80°C gelagert.

### 2.4.4 Reverse Transkription

Nach RNA Extraktion wurde DNA mittels DNase (10 Units/µl) (*New England Biolabs*) nach Herstelleranweisung verdaut. Zur cDNA Synthese wurde das Kit „iScriptAdvanced cDNA Synthesis Kit for RT-qPCR“ (*Bio-Rad*) verwendet. cDNA wurde bei -80°C gelagert.

#### 2.4.5 Polymerasekettenreaktion (PCR)

DNA wurde mit Hilfe von OneTaq und Q5 Polymerase (New England Biolabs) amplifiziert. Die PCR wurde in Anlehnung an die Methodenbeschreibungen von Mullis *et al.* (1986)<sup>249</sup> in Thermozyklern der Marke *Eppendorf* (Mastercycler™) durchgeführt. Als *template* wurden gereinigte gDNA, Plasmid-DNA oder lineare DNA-Fragmente verwendet. Zur extensiven Überprüfung von auf LB-Agarplatten kultivierten Organismen mit gentechnologischer Veränderung wurden Bakterienkolonien mit Hilfe einer Pipettenspitze direkt in den vorgelegten Reaktionsansatz überführt. Nukleotidsubstitutionen wurden mittels *overlap PCR* verwirklicht. Dafür wurden *Primer*, die den Bereich der gewünschten Substitution abdecken, so gewählt, dass in einer ersten PCR unabhängig voneinander zwei DNA Fragmente entstehen. Eine Überlappung von ca. 20 bp im zu substituierenden Bereich ist dabei Voraussetzung für die Fusionierung der beiden Fragmente in einem zweiten PCR Ansatz mit *Primern*, die komplementär zu den beiden Enden der finalen Zielsequenz sind.

Die amplifizierte DNA wurde mittels Gelelektrophorese und anschließender Reinigung (vgl. 2.6.1) von unspezifischen Nebenprodukten befreit.

#### 2.4.6 DNA-Sequenzanalyse

Sequenzierungen von Plasmiden oder zuvor mittels PCR amplifizierten DNA Fragmenten wurden auf Basis des von Sanger *et al.* (1977)<sup>250</sup> beschriebenen Kettenabbruchverfahrens durch die „Genomics Service Unit“ des Bereichs Genetik des Biozentrums der LMU unter Leitung von Dr. Andreas Brachmann durchgeführt. Für die Analyse im „ABI 3730 capillary sequencer“ (*Applied Biosystems*, 50 cm Kappilarlänge) kam das Prinzip des „Cycle Sequencing“ (Programm: „Cycle, Clean and Run: Big Dye 3.1“) zur Anwendung. Proben wurden wie von der „Genomics Service Unit“ empfohlen vorbereitet.

#### 2.4.7 Konstruktion von Plasmiden

Standard DNA-Techniken wurden basierend auf den Beschreibungen von Ausubel *et al.* (1987)<sup>251</sup> und Maniatis *et al.* (1989)<sup>252</sup> ausgeführt. Restriktionsenzyme und T4 DNA-Ligase wurden entsprechend der Herstelleranweisungen eingesetzt (*New England Biolabs*). Zur Vermeidung einer Religation von linearisierten Plasmiden wurden die offenen Enden mit alkalischer Phosphatase (CIP oder rSAP von *New England Biolabs*) nach Herstelleranweisung dephosphoryliert. Die präparative Auftrennung von DNA erfolgte mittels Gelelektrophorese, wie in 2.6.1 beschrieben. Alle Plasmide wurden mittels PCR und Sequenzierung auf ihre Korrektheit überprüft.

#### 2.4.8 Konstruktion von Stämmen

Deletionen wurden mit Hilfe des pRED®/ET® Systems von *Gene Bridges* in das *E. coli* Genom eingeführt („Quick & Easy *E. coli* Gene Deletion by Red®/ET® Recombination“ Kit oder „Counter Selection BAC Modification by pRED®/ET®“ Kit). Das Prinzip beruht auf der zielgerichteten Integration einer Kanamycin-Kassette mit zuvor mittels PCR angefügten 50 bp langen Überhängen, die homolog sind zu den entsprechenden genomischen Zielloci.

Die Kanamycin-Kassette des „Quick & Easy *E. coli* Gene Deletion by Red<sup>®</sup>/ET<sup>®</sup> Recombination“ Kit besitzt zwei FRT Stellen, die die Entfernung des Resistenzmarkers mit Hilfe der auf dem Helferplasmid pCP20 kodierten und über einen temperatursensitiven Promotor regulierten FLP Rekombinase erlauben<sup>253</sup>.

Genomische Integrationen wurden mit Hilfe des Plasmides pNPTS138\_R6KT und einfacher oder doppelter homologer Rekombination realisiert<sup>245</sup>. Die zu integrierenden Sequenzen wurden zusammen mit 500 bp langen flankierenden und den Zielloci entsprechenden DNA-Abschnitten in pNPTS138\_R6KT kloniert. Das Plasmid wurde mittels Konjugation über den Donorstamm WM3046 (William Metcalf, University of Illinois) und homologer Rekombination in das Zielgenom integriert. Dafür wurden Übernacht-Kulturen des Donorstamms (WM3064) mit zu übertragendem pNPTS-Konstrukt und Rezeptorstamm 1:500 in 2 ml frisches LB-Medium (bzw. in LB-Medium mit 300  $\mu$ M DAP) überführt und ca 3 Stunden bei 37°C bis in die frühe exponentielle Wachstumsphase kultiviert. Beide Kulturen wurden abzentrifugiert und in je 200  $\mu$ l frischem LB-Medium resuspendiert. 200  $\mu$ l von WM3064 wurden mit 100  $\mu$ l des Rezeptorstammes vermischt und auf eine LB-Agarplatte mit 300  $\mu$ M DAP getropft. Nach vierstündiger Inkubation bei 37°C, während der die Konjugation und homologe Rekombination erfolgten, wurden die Zellen in 2 ml frischem LB Medium ohne DAP resuspendiert und 3 mal gewaschen, bevor sie auf LB-Agarplatten mit entsprechendem Antibiotikum zur Selektion aufgetragen wurden. Das Plasmidgerüst wurde gegebenenfalls, wie von Lassak *et al.* (2010)<sup>245</sup> beschrieben, über eine negative Selektion unter Ausnutzung der auf pNPTS138\_R6KT durch *sacB* kodierten Levansucrase in einer zweiten homologen Rekombination entfernt. Dafür wurden die Stämme für drei Stunden in LB-Medium ohne Antibiotika aerob bei 37°C kultiviert und anschließend über einen Verdünnungsausstrich auf LB-Platten ohne NaCl, aber mit 10% (w/v) D-(+)-Saccharose überführt und über Nacht bei 30°C inkubiert. Kolonien wurden durch paralleles Auftragen auf LB-Platten mit und ohne Antibiotika auf das Vorhandensein der Resistenz überprüft.

Alle Stämme wurden mittels PCR und Sequenzierung auf ihre Korrektheit überprüft.

#### 2.4.9 Kompetente Zellen und Transformation

*E. coli* Stämme wurden mittels Hitzeschock oder Elektroporation mit verschiedenen DNA-Konstrukten transformiert.

Im Zuge der Hitzeschockmethode wurden kompetente Zellen (100  $\mu$ l) mit ca. 20 ng Plasmid-DNA für 30 Minuten bei 4°C präinkubiert. Anschließend wurden die Ansätze 45 Sekunden lang auf 42°C erwärmt und sofort auf Eis abgekühlt. Nach Zugabe von 750  $\mu$ l frischem LB-Medium wurden die Zellen für zwei Stunden inkubiert und schließlich auf LB-Agarplatten ausgestrichen. Die Kompetenz der Zellen wurde wie folgt erzielt. Kompetente *E. coli* WM3964 Zellen wurden mittels CaCl<sub>2</sub>-Methode nach Dagert und Ehrlich (1979)<sup>254</sup> präpariert. Eine Übernachtskultur wurde 1:500 in frisches LB-DAP (300  $\mu$ M DAP) Medium überimpft. Nach aerober Kultivierung bei 37°C bis zu einer OD<sub>600</sub> von 0,3 wurden die Zellen 15 Minuten auf Eis abgekühlt. Durch Zentrifugation (800 x g, 20 Minuten, 4°C) wurden die Zellen dann pelletiert und in 1/5 des Ausgangsvolumens an eiskaltem CaCl<sub>2</sub> (0,1 M) resuspendiert. Nach einer weiteren zwanzigminütigen Inkubation auf Eis wurden die Zellen erneut pelletiert und in 1/5 des vorherigen Volumens CaCl<sub>2</sub> (0,1 M) resuspendiert. Die kompetenten Zellen wurden für höchstens zwei Wochen bei 4°C gelagert. Kompetente *E. coli* DH5 $\alpha$  Zellen wurden auf eine ähnliche Art und Weise nach Inoue *et al.* (1990)<sup>255</sup> hergestellt. Frisches LB-Medium wurde 1:500 mit einer Übernachtskultur

inokuliert und aerob bei 37°C kultiviert. Bei einer OD<sub>600</sub> von 0,3 wurden die Zellen für 10 Minuten auf Eis abgekühlt und anschließend mittels Zentrifugation (4.500 x g, 15 Minuten, 4°C) pelletiert. Nach Resuspendierung in der 0,5-fachen Menge des Ausgangsvolumens an Transformationspuffer [10 mM PIPES pH 6,6; 50 mM CaCl<sub>2</sub>; 15% Glycerin], wurden die Zellen erneut für 20 Minuten auf Eis inkubiert. Nach einer weiteren Zentrifugation (4.500 x g, 15 Minuten, 4°C) wurde das Pellet in 1/10 des vorherigen Volumens Transformationspuffer aufgenommen, zu je 100 µl aliquotiert, in flüssigem Stickstoff schockgefroren und bei -80°C gelagert.

Für die Elektroporation wurden frisch in LB-Medium angezogene Zellen verwendet. Dafür wurden 30 µl einer Übernachtskultur in 1,5 ml LB-Medium überführt und aerob bei 37°C bis zu einer OD<sub>600</sub> von 0,3 kultiviert. Die Zellen wurden mittels Zentrifugation (15.000 x g, 1 Minute, 4°C) pelletiert, drei Mal in eiskaltem 10% Glycerin gewaschen und zuletzt in 100 µl aufgenommen. Nach Zugabe von ca. 20 ng Plasmid-DNA oder 800 ng linearer dsDNA (~1,5 kb) wurden die Transformationsansätze in vorgekühlte Elektroporationsküvetten (0,2 cm Spaltbreite) überführt. Die Elektroporation wurde in einem Gerät der Firma BioRad (*MicroPulser Electroporator*) bei 2,5 kV durchgeführt. Nach Zugabe von 750 µl frischem LB-Medium wurden die Zellen für weitere zwei Stunden bei 37°C kultiviert, bevor sie auf LB-Agarplatten ausgestrichen wurden.

#### 2.4.10 Proteinüberexpression

OxyR Varianten wurden durch Klonierung in den Vektor pET\_SUMO (Invitrogen: „Champion™ pET SUMO Expression System“) N-terminal mit SUMO- und 6xHis-Tag fusioniert und unter die Kontrolle eines T7 Promotors gestellt. Das Plasmid wurde durch Transformation in BL21 DE3 Zellen (Invitrogen), die das Gen für die T7 RNA Polymerase unter Kontrolle des *lacUV5* Promotors tragen, eingebracht. Dies ermöglichte eine induzierbare Überexpression der OxyR Varianten. Dafür wurde eine Übernachtskultur 1:500 in frisches LB-Medium überimpft und bis zu einer OD<sub>600</sub> von 0,4 aerob bei 37°C kultiviert. Es folgte eine Induktion mit 1 mM IPTG, woraufhin die Zellen für weitere vier Stunden inkubiert und anschließend mittels Zentrifugation (4.500 x g, 20 Minuten, 4°C) geerntet wurden.

## 2.5 Biochemische Methoden

### 2.5.1 Zellaufschluss

*E. coli* Zellen wurden nach Kultivierung entweder mittels Hochdruck oder Ultraschall bei 4°C aufgeschlossen. Dazu wurden Bakterienkulturen pelletiert (4.500 x g, 15 Minuten, 4°C) und in dem entsprechendem Aufschlusspuffer resuspendiert. Dabei handelte es sich um Tris-HCl Puffer [20 mM Tris/HCl pH 7,5; 150 mM NaCl; 10% (v/v) Glycerin], HBS-EP Puffer [10 mM HEPES pH 7,4; 150 mM NaCl; 3 mM EDTA; 0,05% P-20] oder Puffer P20 (adaptiert von Teramoto *et al.*, 2013)<sup>256</sup> [20 mM Tris-HCl pH 7,5; 300 mM NaCl; 10% Glycerin; 0,1 mM EDTA; 5 mM MgCl<sub>2</sub>; 0,05% Tween20] mit je 1 mM DTT, 0,5 mM PMSF und 30 ng/ml DNase.

Für den Aufschluss von *E. coli* Zellen mittels Hochdruck (LA Biosystems: „TS“ Modell, French-Press-Prinzip) wurden je 1 g Zellen in 5 ml Aufschlusspuffer resuspendiert und bei 1,5 kbar aufgeschlossen. Mittels Ultraschall (Branson: Digital Sonifier®) wurden Zellen aufgeschlossen, die im jeweiligen Aufschlusspuffer auf eine theoretische OD<sub>600</sub> von 10 konzentriert wurden. Dabei wurde ein maximales Volumen von 10 ml eingesetzt.

Folgende Parameter kamen zur Anwendung: 30% Amplitude; 0,5 Sekunden Puls (*on/off*); 30 Sekunden Gesamtdauer. Der Vorgang wurde drei Mal wiederholt. Zelldebris wurde über eine niedertourige Zentrifugation bei 4.500 x g und 4°C entfernt.

### 2.5.2 Nickelchelate-Affinitätschromatographie

OxyR Varianten wurden nach Überexpression in BL21 DE3 Zellen und Zellaufschluss aus einer zuvor mittels Ultrazentrifugation (250.000 x g, 45 Minuten, 4°C) separierten cytosolischen Zellfraktion (vgl. 2.5.1) gereinigt. Dabei wurde die 6xHis Markierung der OxyR Proteine und deren hohe Affinität zu Nickelchelaten genutzt. Die cytosolischen Proteine (5-10 mg/ml) wurden infolgedessen zusammen mit - zuvor mit H<sub>2</sub>O<sub>ad</sub> gewaschener und mit den entsprechenden Puffern äquilibrierten - Ni<sup>2+</sup>-NTA-Agarose (4 µl/mg Protein) für zwei Stunden bei 4°C auf einem Rollenmischer inkubiert. Anschließend wurden die an die Ni<sup>2+</sup>-NTA-Agarose gebundenen Proteine in einer Säule (CrystalCruz® Chromatography: 1-2,5 cm x 10 cm) aufgefangen und mit dem 5-fachen Volumen an Reinigungspuffer (Tris/HCl-Puffer, HBS-EP Puffer oder Puffer P20, vgl. 2.5.1, mit 20 mM Imidazol) gewaschen. Die Elution erfolgte in 7 Schritten durch Inkubation mit je 1 ml des mit 250 mM versetzten Reinigungspuffers. Imidazol wurde durch eine darauffolgende 4 stündige Dialyse wieder entfernt.

Der über eine SUMO Schnittstelle vom Rest des Proteins getrennte 6x-His-Tag wurde mit Hilfe einer ebenfalls His-markierten Ulp-Protease (produziert und bereitgestellt von Ingrid Weigl, LMU München AG K. Jung) abgespalten. Für den Verdau wurden je 4 mg Protein (~2mg/ml) mit 1 mg Ulp-Protease (~5 mg/ml) für vier Stunden bei 30°C inkubiert. Die His-markierten Peptide wurden wie oben beschrieben durch eine weitere Nickelchelate-Affinitätschromatographie vom nun unmarkierten Zielprotein separiert.

### 2.5.3 Size exclusion chromatography (SEC)

Um zuvor mittels Nickelchelate-Affinitätschromatographie gereinigte Proteine auf ihren Reinheitsgrad zu überprüfen und diesen gegebenenfalls zu erhöhen, wurden sie über eine „Superdex 200 300/10 increase“ Säule (24 ml) mit Hilfe des ÄKTA purifier Systems (GE Healthcare, Software: „UNICORN 5.11“) und des „Fraction Collector Frac-920“ (GE Healthcare) fraktioniert. Der Lauf wurde mit einer Durchflussrate von 0,5 ml/min in Tris/HCl Puffer (vgl. 2.5.1) durchgeführt. Die Proteinkonzentration wurde durch eine Absorptionmessung bei 280 nm bestimmt und die Größe der Moleküle mit Hilfe eines Filtrationsstandards (bereitgestellt von Dr. Yang Wang, LMU München, AG K. Jung) bestimmt.

### 2.5.4 Präparation von Membranvesikeln

*E. coli* Zellen wurden wie in 2.5.1 beschrieben in Tris/HCl Puffer mit 1 mM DTT, 0,5 mM PMSF und 30 ng/ml DNase aufgeschlossen, der Zelldebris wurde abzentrifugiert (4.500 x g, 30 Minuten, 4°C) und Membranproteine von cytosolischen Proteinen durch Ultrazentrifugation separiert (250.000 x g, 45 Minuten, 4°C). Das Membranpellet wurde in 1/10 des vorherigen Volumens Tris/HCl Puffer resuspendiert.

### 2.5.5 Präparation von Proteinen der äußeren Membran

Bakterienpellets wurden für den Zellaufschluss auf eine theoretische OD<sub>600</sub> von 10 in Tris-HCl Puffer (vgl. 2.5.1) resuspendiert. Membranproteine wurden wie in 2.5.4 beschrieben in 1/5 des vorherigen Volumens Na-Phosphatpuffer [10 mM, pH 7,2] mit 2% (w/v) Triton X-100 konzentriert und für eine Stunde bei 37°C unter leichtem Schütteln inkubiert, um die Proteine der cytoplasmatischen Membran zu lösen. Die Suspension wurde anschließend bei 390.000 x g für 30 Minuten bei 4°C zentrifugiert, der Überstand verworfen und das Pellet in Na-Phosphatpuffer ohne Triton X-100 gewaschen. Es folgte eine weitere Zentrifugation (390.000 x g, 30 Minuten, 4°C) und das Pellet wurde in 1/20 des vorherigen Volumens PBS Puffer [8,1 mM Na<sub>2</sub>HPO<sub>4</sub>; 1,47 mM KH<sub>2</sub>PO<sub>4</sub>; 137 mM NaCl; 2,68 mM KCl] resuspendiert. Nach Zugabe von SDS-Probenpuffer [Endkonzentration: 50 mM Tris/HCl pH 6,8; 2% (w/v) SDS; 0,1% (w/v) Bromphenolblau; 10% (v/v) Glycerin, 100 mM DTT] wurde den Proben je 0,25 g/ml Harnstoff beigemischt und es folgte ein fünfminütiges Erhitzen bei 100°C. Die elektrophoretische Auftrennung der Proteine erfolgte in Polyacrylamidgelen mit 4 M Harnstoff (vgl. 2.6.3).

### 2.5.6 Präparation von Sphäroplasten

*E. coli* Zellen wurden nach einer Übernachtskultivierung 1:500 in frisches LB Medium überimpft und aerob bei 37°C bis zu einer OD<sub>600</sub> von 0,5 kultiviert. Die Zellen wurden nach Zentrifugation (4.500 x g, 10 Minuten, 4°C) in 80 ml/g<sub>(Zellen)</sub> Puffer 1 [25% (w/v) Saccharose; 30 mM Tris/HCl pH 8,0] resuspendiert. Nach Zugabe von Lysozym (50 µg/ml) und 10 mM K<sub>2</sub>-EDTA (pH 7,0), wurden die Zellen auf einem Rollenmischer für 45 Minuten bei Raumtemperatur inkubiert. Die Sphäroblastenbildung wurde währenddessen stichprobenartig unter dem Mikroskop überprüft (Leica: „DMi 8“; Software „LAS X“). Für die Fluoreszenzmikroskopie wurden die Sphäroblasten zuvor wie in 2.8.2 beschrieben fixiert, dabei jedoch in Puffer 1 belassen.

### 2.5.7 Präparation von *right-side-out* Vesikeln

100 ml zuvor präparierter Sphäroplasten (vgl. 2.5.7) wurden pelletiert (12.000 x g, 20 Minuten, Raumtemperatur) und in 1 ml Puffer 2 [25% (w/v) Saccharose; 100 mM Tris/HCl pH 7,5; 20 mM MgSO<sub>4</sub>] resuspendiert. Der Ansatz wurde in 15 ml auf 30°C vorgewärmten Puffer 3 [20 mM Tris/HCl pH7,5; 2 mM DTT] überführt und für 15 Minuten bei 30°C inkubiert. Einer niedertourigen Zentrifugation (4.500 x g, 5 Minuten, 4°C) folgte eine weitere Zentrifugation des Überstandes (15.000 x g, 15 Minuten, 4°C) und eine Resuspendierung des Pellets in Puffer 4 [100 mM Tris/HCl pH 7,5; 10% (v/v) Glycerin; 10 mM MgCl<sub>2</sub>].

## 2.6 Analytische *in vitro* Methoden

### 2.6.1 DNA Gelelektrophorese

DNA wurde in Agarosegelen [1,5% (w/v) Agarose; 40 mM Tris-Acetat pH 7,6; 1mM EDTA] mit 0,25 µg/ml Ethidiumbromid oder 0,05 µl/ml Midori Green in Geräten der Firma BioRad („Mini-Sub® DNA Cell GT“ oder „Wide Mini-Sub® Cell GT“ Laufkammer“ bei 100V elektrophoretisch aufgetrennt. Als Größenstandard wurden 5 µl der „2-log ladder“ von New England Biolabs aufgetragen und als Laufpuffer wurde TAE Puffer [40 mM Tris-

Acetat pH 7,6; 1mM EDTA] verwendet. Nach dem Gellauf wurden die in die DNA interkalierten Fluoreszenzfarbstoffe bei UV-Strahlung in einem Transilluminator (BioRad; Software: „Gel Doc™ EZ Imager“) visualisiert.

Die Isolation von DNA aus Agarosegelen erfolgte mit Hilfe des Kits „Hi Yield® Gel/PCR DNA Fragment Extraction“ (Südlabor-Gauting). DNA-Konzentrationen wurden photometrisch mit dem „UV/Vis-Spectrophotometer NanoDrop® ND-100“ (Nanodrop Technologies) bestimmt.

### 2.6.2 qRT-PCR

Ca. 200 bp lange cDNA Fragmente wurden in einem Cycler der Firma BioRad („MyiQ™ Single Color Real-Time PCR Detection System“, Software „Bio-Rad iQ5“, 4titude® 96 well plates, polypropylene, 200 µl capacity) amplifiziert. Reaktionsansätze zu je 25 µl wurden auf Eis vorbereitet und setzten sich folgendermaßen zusammen: 12,5 µl iQ™ SYBR® Green Supermix (BioRad), je 1 µl *Primer* (5 µM), 8,5 µl DEPC-H<sub>2</sub>O, sowie 50 ng *template* DNA (25 ng/µl). Es wurde folgender Temperaturverlauf verwendet: 5 Minuten/95°C, 40 Zyklen aus 10 Sekunden/95°C und 30 Sekunden/60°C. Eine daran anschließende Schmelzkurve wurde folgendermaßen erstellt: 55°C – 95°C mit 1°C Temperaturanstieg/Minute. Der *threshold* wurde so definiert, dass die daraus resultierenden C<sub>t</sub> Werte diejenigen Zyklen markieren, die den Beginn des exponentiellen Anstiegs der reaktionstypischen Fluoreszenz widerspiegeln. Als Negativkontrolle isolierte RNA als *template* (25 ng/µl) eingesetzt. Die qRT-PCR Ergebnisse wurden nur dann akzeptiert, wenn die C<sub>t</sub> Werte der Negativkontrollen größer als 40 waren oder außerhalb des messbaren Bereiches lagen.

Für die Quantifizierung von mRNA wurde RNA wie in 2.4.3 und 2.4.4 beschrieben in cDNA transkribiert und in einer qRT-PCR genspezifisch amplifiziert. Dabei wurden sowohl Ziel- und Referenzgene, als auch Kontroll- und Testwachstumsbedingungen miteinander verglichen. Die Berechnung der relativen Genexpression erfolgte nach Pfaffl *et al.* (2001)<sup>257</sup>. In der unten aufgeführten Formel ist *E* die *Primer*-Effizienz, die für jedes verwendete *Primer*-Paar ermittelt wurde. Sie basiert auf einer Eichgerade aus C<sub>t</sub>-Werten aus qRT-PCR Reaktionen mit definierter Menge an *template* DNA (*E. coli* MG1655 gDNA: 100 ng; 25 ng; 6,25 ng; 1,6 ng; 0,39 ng; 0,1 ng), wobei *m* die Steigung dieser Geraden beschreibt. Die Effizienten aller verwendeten *Primer*-Paare lagen zwischen 1,8 und 2,0.

$$\text{relative Genexpression} = \frac{(E_{\text{Zielgen}})^{\Delta C_{t_{\text{Zielgen}}}}}{(E_{\text{Referenzgen}})^{\Delta C_{t_{\text{Referenzgen}}}}}$$

$$E = 10^{-1/m}$$

$$m = \frac{Ct}{\lg(c_{DNA}[ng/\mu l])}$$

$$\Delta Ct = Ct_{\text{Kontrollbedingung}} - Ct_{\text{Testbedingung}}$$

### 2.6.3 SDS-PAGE

Proteine wurden mittels denaturierender SDS Polyacrylamid-Gelelektrophorese (SDS-PAGE) in Geräten der Firma PeqLab („Perfect double gel system, Twin S/M/L“, 10 cm – 20 cm Länge) nach einem Protokoll von Lämmlí (1970)<sup>258</sup> bei 140V aufgetrennt. Gele setzten sich zusammen aus Sammelgel mit 4,9% (w/v) und Trenngel mit je nach Größe des zu untersuchenden Proteins 7,5% bis 12,5% (w/v) Acrylamid-Bisacrylamid (37,5:1). Zusätzlich wurde dem Trenngel 0,5% (v/v) Trichlorethanol (TCE) hinzugefügt, um Proteine ohne vorherige Anfärbung des Gels unter UV-Strahlung nach vorheriger Quervernetzung im Transilluminator (BioRad; Software: „Gel Doc™ EZ Imager“) zu visualisieren. Der Laufpuffer setzte sich zusammen aus 25 mM Tris, 192 mM Glycin und 0.1% (w/v) SDS (pH 8,3). Proteinproben wurden vor dem Lauf mit SDS-Ladepuffer versetzt [Endkonzentration: 50 mM Tris-HCl pH 6,8; 2% (w/v) SDS; 0,1% (w/v) Bromphenolblau, 10% (v/v) Glycerin; 100 mM DTT] und 5 Minuten bei 95°C denaturiert. Als Größenstandards wurden „Roti®-Mark Standard“ (Carl Roth) und „Page Ruler™ Prestained Protein Ladder“ (Invitrogen) verwendet. Angereicherte Proteine der äußeren Membran wurden zur Vermeidung von Aggregation in einem Harnstoff-Gel nach Lugtenberg *et al.* (1975)<sup>259</sup> aufgetrennt. Im Unterschied zu den oben beschriebenen Gelen beinhaltet das Trenngel in diesem Fall 11,25% Acrylamid-Bisacrylamid (44:1) und 4 M Harnstoff.

Die in den Gelen aufgetrennten Proteine wurden mit Coomassie-Blau (nach Weber und Osborn, 1969<sup>260</sup>), „Coomassie Instant Blue“ (nach Herstelleranweisung: Expedon) oder Silbernitrat (nach Blum *et al.*, 1987)<sup>261</sup> angefärbt, sofern sie nicht für einen Western Blot (vgl. 2.6.5) weiterverwendet wurden.

### 2.6.4 Isoelektrische Fokussierung

Gele der Firma BioRad wurden nach Herstelleranweisung rehydriert und in ein 2D-Elektrophoresesystem von Pharmacia Biotec („Multiphor II“) montiert. Zur Etablierung eines pH-Gradienten fand vor Probenlauf eine Vorfokussierung statt (700 V, 12 mA, 8 W, 20 Minuten). Zuvor mit Probenpuffer (50% v/v Glycerin, 10 mg/ml Bromphenolblau) versetzte Proben wurden auf ein mittig platziertes Filterpapier aufgetragen. Der Probenübergang fand für 20 Minuten bei 500 V, 8 mA und 8 W statt, die isoelektrische Fokussierung für 4 Stunden bei 2000 V, 14 mA und 14 W. Anschließend wurden die Banden für 10 Minuten geschärft (2500 V, 14 mA, 18 W). Die Gele wurden mit Hilfe eines Drahtes von der Kunststoffunterlage abgelöst und für die weitere Analyse durch einen Western Blot (vgl. 2.6.5) vorbereitet.

### 2.6.5 Western Blot

Proteine, die zuvor über eine SDS-PAGE oder isoelektrische Fokussierung aufgetrennt wurden, wurden mittels Nabsblot auf eine Nitrocellulosemembran mit 0,45 µm Porendurchmesser (Sigma Aldrich: „Whatman™, Protran® BA85“) oder auf eine PVDF Membran mit 0,45 µm Porendurchmesser (Sigma Aldrich: „Amersham™ Hybond P 0.45 PVDF blotting membrane“) übertragen. Der Proteintransfer erfolgte für 90 Minuten bei 300 mA in einer „Mini Trans-Blot® Cell“ Kammer (BioRad), bzw. 120 Minuten bei 800 mA in einer „Hoefler TE62 tank transfer unit“ von Amersham.

Die Membran wurde mit in TBS Puffer [20 mM Tris-HCl, pH 7,5; 150 mM NaCl] gelöstem 5% (w/v) Magermilchpulver über Nacht bei 4°C abgesättigt. Der immunologische Nachweis der Proteine erfolgte mit Hilfe von spezifischen Antikörpern (vgl. Tabelle 5). Nach dreimaligem Waschen (je 5 Minuten) mit TBS-TT Puffer [20 mM Tris-HCl, pH 7,5; 500 mM NaCl; 0,05% (v/v) Tween 20; 0,2% Triton X-100] wurde die Membran für 2 Stunden bei Raumtemperatur mit einem primären, in TBS-T Puffer [20 mM Tris-HCl, pH 7,5; 150 mM NaCl; 0,05% (v/v) Tween 20] mit 0,75% Magermilchpulver verdünnten Antikörper inkubiert. Nach dreimaligem Waschen in TBS-TT Puffer (je 5 Minuten) folgte die Inkubation mit einem sekundären, ebenfalls in TBS-T Puffer und 0,75% Magermilchpulver verdünnten Antikörper (1,5 Stunden, Raumtemperatur). Die Membran wurde drei Mal für je 5 Minuten in TBS-TT Puffer und einmal für 10 Minuten in TBS Puffer gewaschen. Das Substrat wurde je nach gewähltem Antikörper-Konjugat gewählt. Antikörper mit alkalischer Phosphatase wurden durch die Inkubation in Natriumcarbonat-Färbelösung [50 mM Na<sub>2</sub>CO<sub>3</sub>; 0,1% NBT; 0,5 mg/ml BCIP] detektiert. Antikörper mit Peroxidase wurden mit Hilfe von „UptiLight™ HRP blot substrate“ (Uptima) nach Herstelleranweisung detektiert und die Lumineszenz in einer Dunkelkammer mit Hilfe einer lichtsensitiven Kamera (PEQLAB, Objektiv: „Spacecom“; Software: „Fusion“) aufgezeichnet.

### 2.6.6 Thermophoretic Shift Assay

Die Stabilität von zuvor gereinigten OxyR Varianten in Abhängigkeit von verschiedenen Pufferzusammensetzungen wurde mittels Thermofluor Screen nach Boivin *et al.* (2013)<sup>262</sup> untersucht. Dafür wurde in einer 96-well Platte (4titude: „96 well plate“, Polypropylene, 200 µl Kapazität) 1 µg des Proteins im zu testenden Puffer verdünnt. Zu einem Reaktionsvolumen von 23 µl wurden 2 µl des Farbstoffs SYPRO Orange (62x, Invitrogen) gegeben. Im Zuge eines Temperaturgradienten von 30-95°C mit einem Temperaturanstieg von 1°C/Minute wurde die Fluoreszenz von SYPRO Orange minütlich gemessen (BioRad: „MyiQ™ Single Color Real-Time PCR Detection System“, Software „Bio-Rad iQ5“).

Im Rahmen dieser Arbeit wurden verschiedene Puffer (Tris-HCl 20-100 mM, PBS) und Pufferzusammensetzungen getestet, wobei der Fokus auf die NaCl-Konzentration (0-500 mM), den pH-Bereich (pH 6,5 – 8), sowie das Vorhandensein von verschiedenen Additiven, wie DTT, EDTA, MgCl<sub>2</sub>, ATP, Urea, Betain, CaCl<sub>2</sub>, Glycin und D-Glucose als putative Stabilisatoren gelegt wurde.

### 2.6.7 Native PAGE und EMSA

Frisch gereinigtes Protein wurde in einem Reaktionsvolumen von 30 µl mit 30 ng 6-FAM markierter doppelsträngiger DNA in Puffer P20 (vgl. Teramoto *et al.*, 2013)<sup>256</sup> [20 mM Tris-HCl pH 7,5; 300 mM NaCl; 10% Glycerin; 0,1 mM EDTA; 5 mM MgCl<sub>2</sub>; 0,05% Tween20; 1 mM DTT; 0,05 mM PMSF], 150 ng Lachsperma-DNA und variierenden Mengen BSA für 30 Minuten bei 30°C inkubiert. Die Bindung des Proteins an die Ziel-DNA wurde in einem Gelretardationsassay, basierend auf den Methodenbeschreibungen von Teramoto *et al.*, (2013)<sup>256</sup>, nachgewiesen. Die Proben wurden dafür in einem 10x10 cm großem Polyacrylamidgel (5,5% (w/v) Acrylamid:Bisacrylamid – ohne SDS, sonst wie in 2.6.3 für Trenngel beschrieben) elektrophoretisch analysiert. Die Gele wurden vor dem eigentlichen Probenlauf 45 Minuten bei 70 V äquilibriert. Nach dem Beladen wurden die Proteine für eine Stunde bei 100 V aufgetrennt. Die Fluoreszenz des 6FAM DNA-Markers wurde bei

488 nm in einem Fluoreszenzscanner (Amersham: „Typhoon TRIO Variable Mode Imager“; Filter: 520 BP40 Cy - 600/88 nm; Software: „Typhoon Scanner Control“) detektiert.

### 2.6.8 DNA-Affinitätschromatographie

DNA-bindende Proteine wurden mit Hilfe von streptavidinummantelten Dynabeads™ M-280 (Thermo Fisher Scientific) und daran gekoppelter dsDNA isoliert. Die Separierung der Beads erfolgte in einer magnetische Halterung (Invitrogen: „DynaMag™-2 Magnet“). Vor Verwendung wurden je 5 mg der Beads drei Mal in DBB Puffer [5 mM Tris/HCl pH 7,5; 0,5 mM EDTA; 1 M NaCl] gewaschen. Die Beads wurden anschließend mit 600 pmol biotinmarkierter, in 1 ml DBB Puffer vorliegender DNA für 4 Stunden bei Raumtemperatur auf einem Roller inkubiert. Zur Entfernung von ungebundener DNA wurden die Beads drei Mal mit DBB Puffer und zwei Mal in PBB Puffer [20 mM Tris/HCl pH 8,9; 1 mM EDTA; 0,05% (v/v) Triton X-100; 10% (v/v) Glycerin; 1 mM DTT; 100 mM NaCl] gewaschen und schließlich in 1 ml PBB Puffer aufgenommen.

*E. coli* Zellen einer 1 l Kultur wurden in PBB Puffer aufgeschlossen und cytosolische Proteine mittels Ultrazentrifugation separiert (vgl. 2.5.1). Zu 20 ml dieser Fraktion (Proteinkonzentration ~8-16 mg/ml) wurden 5 mg der beladenen Beads gegeben und diese für zwei Stunden auf einem Rollenmischer inkubiert. Die Beads wurden anschließend in 1 ml PBB aufgenommen und drei Mal mit PBB Puffer gewaschen. Es folgte eine abschließende Resuspendierung in 50 µl Elutionspuffer [PBB mit 1 M NaCl], ein gründliches Vortexen und sofortiges Abkühlen auf Eis. Nach finaler Separierung der Beads wurde der Überstand in ein frisches Reaktionsgefäß überführt und für die weitere Analyse mittels SDS-PAGE (vgl. 2.6.3) vorbereitet.

### 2.6.9 Oberflächenplasmonenresonanz-Spektroskopie

Frisch gereinigte OxyR Varianten wurden in HBS-EP-Puffer [10 mM HEPES pH 7,4; 150 mM NaCl; 3 mM EDTA; 0,05% P-20] überführt. Biotinmarkierte, doppelsträngige DNA wurde durch eine PCR mit entsprechenden *Primern*, bzw. durch das Annealing komplementärer *Primer* während eines Temperaturgradienten von 95-25°C generiert.

Zur Untersuchung der DNA-Affinität von OxyR Varianten mittels Oberflächenplasmonenresonanz-Spektroskopie wurde ein streptavidinbeschichteter Carboxymethyl-Dextran Sensorchip („SA Sensor Chip Series S“) zuerst äquilibriert (drei Injektionen 1 M NaCl und 50 mM NaOH, 10 µl/min Flussrate) und anschließend mit ca. 200-300 DNA Molekülen pro Flusszelle beladen und in dem Gerät „BiacoreT200“ (GE Healthcare) analysiert. Dabei wurde HBS-EP Puffer als Laufpuffer und ein Gemisch aus 1 M NaCl, 50 mM NaOH und 50% (v/v) Isopropanol als Waschpuffer verwendet. Das Protein wurde in verschiedenen Konzentrationen (0,01 µM – 10 µM), sowie mit einer Kontaktzeit von 180 Sekunden und einer Dissoziationszeit von 300 Sekunden über die Flusszellen geleitet. Die Chipoberfläche wurde nach jedem Zyklus mit 2,5 M NaCl und 0,5% (w/v) SDS mit einer Flussrate von je 60 µl/min regeneriert. Die Auswertung der Sensorgramme erfolgte mit der Software „Biacore Control Software 1.0“ (GE Healthcare), wobei alle Sensorgramme auf die Grundlinie 0 normalisiert wurden. Die Messungen wurden in der Serviceeinheit Bioanalytik (PD Dr. Ralf Heermann) im Bereich Mikrobiologie des Biozentrums der LMU durchgeführt.

### 2.6.10 Dynamische Lichtstreuungsdetektoren (DLS)

Der Oligomerisationszustand von OxyR Varianten, wurde durch die Bestimmung der absoluten Molekülmasse sowie dem Molekülradius mittels DLS (*Dynamic Light Scattering*) untersucht. Die Messungen wurden mit einem DLS-Detektor („Wyatt DynaPro Nanostar“) in der Serviceeinheit Bioanalytik (PD Dr. Ralf Heermann) im Bereich Mikrobiologie des Biozentrums der LMU durchgeführt.

### 2.6.11 Flüssigchromatographie-Massenspektrometrie (LC-MS)

Proteine, die zuvor mittels SDS-PAGE aufgetrennt und mit Coomassie-Blau angefärbt wurden, wurden wie folgt aufbereitet. Die entsprechenden Proteinbanden wurden mit Hilfe eines Skalpellens aus dem Polyacrylamidgel herausgelöst. Für den Verdau der im Gel liegenden Proteine kam ein standardisiertes Protokoll zur Anwendung, das auf Beschreibungen von Sheevchenko *et al.* (2000)<sup>263</sup> und Wilm *et al.* (1996)<sup>264</sup> basiert. Zuerst wurden die Gelstücke zweimal in je 100 µl H<sub>2</sub>O<sub>dd</sub> und dreimal in je 100 ml 25 mM NH<sub>4</sub>HCO<sub>3</sub> gewaschen und durch dreimaliges Waschen in 100 µl Acetonitril dehydriert. Anschließend wurden die Gelstücke für eine Stunde mit 50 µl 10 mM DTT in 25 mM NH<sub>4</sub>HCO<sub>3</sub> inkubiert. Um die reduzierten Cysteine zu carbomidomethylieren, wurden die Stücke für 30 Minuten im Dunkeln mit 50 µl 55 mM Iodoacetamid in 25 mM NH<sub>4</sub>HCO<sub>3</sub> inkubiert. Danach wurden die Gelstücke in 100 µl 25 mM NH<sub>4</sub>HCO<sub>3</sub> gewaschen und in 100 ml Acetonitril rehydriert. Der Verdau erfolgte durch Zugabe von 25 ng/µl Trypsin (gelöst in 25 mM NH<sub>4</sub>HCO<sub>3</sub>) je nach Größe des Gelstückes und einer Inkubation bei 4°C für 45 Minuten. Im Anschluss wurde nicht absorbiertes Trypsin entfernt und Gelstücke mit 25 mM NH<sub>4</sub>HCO<sub>3</sub> bedeckt. Der Verdau erfolgte für 16 Stunden bei 37°C. Zur Extraktion der Peptide wurden die Gelstücke zweimal mit 50 µl einer Acetonitril:H<sub>2</sub>O<sub>dd</sub> 1:1 Lösung mit 0,25% (v/v) Trifluoressigsäure und zwei weitere Male in 50 µl Acetonitril gewaschen. Die Peptide wurden evaporiert und in 30 µl 0,1% (v/v) Trifluoressigsäure resuspendiert.

Proteine zuvor isolierter VBNCs (vgl. 2.8.3) wurden wie folgt für die MS-Analyse aufbereitet. 1x10<sup>9</sup> Zellen wurden pelletiert, in 100 µl PBS Puffer [10 mM Na<sub>2</sub>HPO<sub>4</sub>; 1,8 mM KH<sub>2</sub>PO<sub>4</sub>; 137 mM NaCl; 2,7 mM KCl; pH 7,4] resuspendiert und durch dreimaliges Einfrieren in flüssigem Stickstoff und anschließender fünfminütiger Inkubation im Ultraschall-Wasserbad bei Raumtemperatur aufgeschlossen. 10 µl der Proben wurden einem tryptischen Verdau mit dem „iST Sample Preparation Kit“ von Preomics unterzogen.

5 µl der Peptidlösungen wurden in ein „Ultimate 3000 RSLCnano“ System (Thermo Scientific) injiziert und mit einer analytischen C18 Nanosäule (Aurora, Ionopticks: 25 cm, 75 µm) mittels 120-minütigem Acetonitril-Gradienten (4-40%) in 0,1% Methansäure aufgetrennt. Der Durchlauf aus der HPLC wurde mittels Elektrospray in ein „Q Exactive HF“ System (Thermo Scientific) übertragen. Das Instrument wurde in einem datenabhängigen Modus bedient, um automatisch zwischen „full scan“ MS und Tandemmods („MS/MS“) zu wechseln. „Survey full scan MS“ Spektren (m/z 375-1600) wurden mit einer Auflösung von R=60.000 bei einer m/z von 400 (AGC target von 3x10<sup>6</sup>) aufgezeichnet. Die zehn prominentesten Peptidionen mit Ladungszuständen zwischen drei und fünf wurden fortlaufend isoliert (Zielwert 1x10<sup>5</sup>) und bei 27% normalisierter Kollisionsenergie fragmentiert. Typische Bedingungen der Massenspektrometrie waren: 1,5 kV Sprayspannung; kein Hüll- und Auxiliargasfluss; 250°C aufgeheizte Kapillartemperatur; 33.000 Zählereignisse als Grenzwert für die Ionenauswahl.

Um die Proteine zu identifizieren und mittels LFQ zu quantifizieren, wurde die Software „MaxQuant 1.5.2.8“ mit den folgenden Parametern verwendet: uniprot\_Ecoli\_3AUP00000625\_180323 Datenbank; 10 ppm MS tol; 0,5 Da MS/MS tol; 0,1 Peptide FDR; 5 minimal peptide length; variable Modifikationen; Oxydation (M); fixierte Modifikationen: Carbamidomethyl (C); Peptide für die Proteinquantifizierung: razor and unique; 1 minimal peptides; 2 ratio count. Die identifizierten Proteine wurden als hoch- oder herunterreguliert bezeichnet, falls ihre „MaxQuant LFQ“ Werte  $\log_2$  5-fach erhöht waren und ein p-Wert von 0,05 (t-Test für Mehrfachvergleich) im Vergleich zur Kontrolle nicht überschritten wurde. Die Analyse wurde mit Hilfe der Perseus Plattform<sup>265</sup> durchgeführt und die Ergebnisse durch die Software *Scaffold 4* (Proteomes Software) visualisiert und analysiert.

Die Proteomdaten zur massenspektrometrischen Analyse von VBNCs (vgl. 2.8.3) sind im „ProteomeXchange Consortium“ über die PRIDE<sup>266</sup> Ablage hinterlegt (Daten-Identifikationsnummer PXD009463). Die massenspektrometrische Analyse von verdauten Peptiden wurde von Dr. Ignasi Forné (Zentrallabor für Proteinanalytik, Biomedizinisches Centrum München, LMU) und in Kooperation mit der AG Prof. Dr. Imhoff am Biomedizinischen Centrum der LMU durchgeführt.

### 2.6.12 $\beta$ -Galactosidase Assay

Sofern nicht anders beschrieben, wurden 1 ml bei 37°C kultivierter Zellen, deren  $OD_{600}$  zuvor bestimmt wurde, durch eine Zentrifugation (15.000 x g, 5 Minuten, 4°C) pelletiert und in 1 ml Puffer Z [60 mM  $Na_2HPO_4$ ; 40 mM  $NaH_2PO_4$ ; 10 mM KCl; 1 mM  $MgSO_4$ ; 50 mM  $\beta$ -Mercaptoethanol] resuspendiert. Nach Zugabe von 100  $\mu$ l Chloroform und 50  $\mu$ l 0,1% (w/v) SDS, sowie gründlichem Vortexen wurden die Zellen für 5 Minuten bei 30°C inkubiert. Die enzymatische Aktivität der  $\beta$ -Galactosidase wurde durch Zugabe von 200  $\mu$ l des Substrats O-NPG (4 mg/ml in Puffer Z) bestimmt, wobei die Reaktion nach Gelbfärbung mit 500  $\mu$ l 1 M  $Na_2CO_3$  abgestoppt wurde. Nach Zentrifugation (15.000 x g, 10 Minuten, Raumtemperatur) wurde die Absorption des Überstands bei 420 nm bestimmt (Amersham Biosciences: „Ultraspec 2100 pro“-Photometer). Die  $\beta$ -Galactosidaseaktivität wurde mit folgender Formel berechnet:

$$\text{Aktivität [Miller-Units]} = \frac{A_{420} \times 1000}{\text{Vol[ml]} \times t[\text{min}] \times OD_{600}}$$

## 2.7 Analytische *in vivo* Methoden

### 2.7.1 Separierung von intakten, lebenden Zellen aus Langzeitkulturen mittels Percoll

Das Verfahren wurde im Rahmen dieser Arbeit wie im Ergebnisteil beschrieben entwickelt und optimiert. 100 ml einer VBNC-Kultur ( $OD_{600} = 1$ ) (vgl. 2.8.3) wurden mittels Zentrifugation pelletiert und in 20 ml PBS Puffer [10 mM  $Na_2HPO_4$ ; 1,8 mM  $KH_2PO_4$ ; 137 mM NaCl; 2,7 mM KCl; pH 7,4], sowie 30 ml Percoll® (GE Healthcare Sciences) resuspendiert, sodass die finale Konzentration der kolloidalen Silikalösung bei 60% lag. Die Zellen wurden über eine anschließende Dichtegradientenzentrifugation in mehreren Schritten separiert. Zuerst wurden die Proben bei 11.500 x g und 4°C für eine Stunde zentrifugiert, bis zwei verschiedene Zellschichten deutlich

voneinander unterscheidbar waren. Die untere Zellschicht, welche aus lebenden, intakten Zellen besteht, wurde mit Hilfe einer Pipette gesammelt und in 70 ml PBS überführt. Die Zellen wurden in drei folgenden Zentrifugationsschritten bei 4.500 x g und 4°C für je 45 Minuten gewaschen und final in PBS auf eine theoretische OD<sub>600</sub> von 10 resuspendiert.

### 2.7.2 Zellanfärbung

Zur Bestimmung des Anteils an lebenden Zellen innerhalb einer Kultur wurden der in PBS Puffer verdünnte Farbstoff Propidiumiodid (1 µM) verwendet. Zur Anfärbung wurden 1 ml einer Zellkultur pelletiert und in 1 ml des entsprechend verdünnten Farbstoffes resuspendiert. Nach zehnminütiger Inkubation bei Raumtemperatur folgte ein Waschschriff, bevor die Anfärbung der Zellen mittels Durchflusszytometer (vgl. 2.7.3) oder Mikroskopie (2.9.5) untersucht wurden.

### 2.7.3 Durchflusszytometrie

Zur Erfassung und quantitativen Analyse von zuvor wie in 2.7.2 beschrieben gefärbten Zellen wurde ein Durchflusszytometer der Firma Accuri (Software: „3D Accuri C6“) mit folgenden Einstellungen verwendet: Slow Run; Backflush after each Measurement; 5000 = Count Limitation.

### 2.7.4 Bacterial-two-Hybrid Assay

Um die Interaktion von EnvZ und MzrA zu untersuchen, wurden die entsprechenden Gene C-terminal an die komplementierenden Fragmente T18 und T25 der Adenylatcyclase CyaA in die Vektoren pT18 oder pKT25 kloniert (Euromedex)<sup>267, 268</sup>. Der Reporterstamm BTH101<sup>269</sup> wurde mit Kombinationen aus pT18-EnvZ/pT25-EnvZ oder pT18-EnvZ/pT25-MzrA Varianten transformiert. Eine Komplementation der Adenylatcyclase-Fragmente durch Interaktion der fusionierten Zielproteine führt zur Synthese von cAMP, welches als Signalmolekül fungiert und in BTH101 die Expression des Reportergens *lacZ* aktiviert. Diese wurde mit zwei verschiedenen Methoden quantifiziert: einem β-Galactosidase Assay und einem kolorimetrischen Vergleich von Makrokolonien auf LB-IPTG/X-Gal Platten. Für den β-Galactosidase Assay wurden LB-Übernachtskulturen 1:500 in 1 ml frisches LB-Medium überführt und bei 37°C für zwei Stunden bei 950 rpm auf einem Thermoblock (Eppendorf) inkubiert. Nach der Induktion mit 0,5 mM IPTG (durch Zugabe von 1 ml LB-Medium mit 1 mM IPTG) wurden die Zellen für weitere 4-12 Stunden inkubiert und anschließend geerntet. Die β-Galactosidaseaktivität wurde wie in 2.6.12 beschrieben bestimmt. Alternativ dazu wurden 5 µl einer Zellkultur, die zuvor bei 37°C und 950 rpm auf einem Thermoblock (Eppendorf) bis auf eine OD<sub>600</sub> von 0,2 inkubiert wurde, auf LB-Agarplatten mit 1 mM IPTG und 40 µg/ml X-Gal (5-Brom-4-Chlor-3-Indoxyl-β-D-Galactopyranosid) aufgetropft. Nach 24-stündiger Inkubation bei 30°C wurde die relative β-Galactosidaseaktivität basierend auf der Intensität der Blaufärbung der Kolonien bewertet.

### 2.7.5 Porinexpressionsassay

Der Assay zur Quantifizierung der Expression von *ompC* und *ompF* basiert auf einer von Batchelor *et al.* (2006)<sup>270</sup> publizierten Methode. Der Reporterstamm EPB273a<sup>144</sup>, bzw. dessen Derivat EPB273a EnvZ<sub>IPPL/IAAAL</sub>, wurden über Nacht in M9 Medium vorkultiviert. Aus dieser Kultur wurde frisches M9 Medium inokuliert (1:500) und die Zellen wurden bis zu einer OD<sub>600</sub> von ~0,7 bei 37°C und 950 rpm in einem Thermoblock (Eppendorf) inkubiert. Die Kultur wurde 1:500 in frisches M9 Medium überführt, das je nach Versuchsfragestellung mit NaCl oder Saccharose versetzt war. Aliquots von je 200 µl wurden in 96-well Platten (Corning Incorporated: „Costar® 3610 96 Well Assay Plate, Black, Clear Bottom“) überführt und bei 37°C auf einem Schüttler bis zu einer OD<sub>600</sub> von 0,2 inkubiert. Nach 15 minütiger Inkubation auf Eis wurde die Fluoreszenz von CFP mit einem TECAN Plattenleser („Infinite 200Pro“, Programm: „Tecan i-control“) bestimmt.

### 2.7.6 Erstellung von Wachstumskurven

Eine LB-Übernachtskultur wurde auf eine OD<sub>600</sub> von 0,01 verdünnt und zu je 200 µl in eine 96-well Platte überführt (Corning Incorporated: „Costar® 3610 96 Well Assay Plate, Black, Clear Bottom“). Die Platte wurde kontinuierlich in einem TECAN Plattenleser („Infinite 200Pro“, Programm: „Tecan i-control“) bei 37°C und ständigem Schütteln inkubiert. Alle 30 Minuten erfolgte eine automatische Messung der OD<sub>600</sub>.

## 2.8 Phänotypische Charakterisierung

### 2.8.1 Säureresistenz Assay

Der Assay zur Analyse der Säureresistenz basiert auf einer von Masuda und Church (2002)<sup>271</sup> beschriebenen Methode. Eine Vorkultur, die an zwei aufeinanderfolgenden Tagen jeweils in frischem LB-Medium angezogen wurde, wurde 1:200 in 20 ml LB-Medium verdünnt und bei 37°C aerob bis zu einer OD<sub>600</sub> von 0,7 kultiviert. Nach Einstellung eines pH-Wertes von 3,0 durch Zugabe von LB-Medium mit einem pH von 1,75 wurden die Zellen weiter bei 37°C kultiviert. Unmittelbar nach Einstellung des pH-Wertes ( $t_0$ ), sowie stündlich danach wurden Aliquots von 100 µl entnommen, um durch das Ausplattieren von Verdünnungsreihen auf LB-Agar und anschließende 12-stündige Inkubation bei 37°C die Zahl der noch lebenden und kultivierbaren Zellen zu bestimmen (CFU/ml = *colony forming units*/ml). Nach Abschluss des Versuchs wurde mit einem pH-Meter (inoLab) verifiziert, dass die Änderung des pH-Werts während des Versuchsverlaufs nicht größer als 0,1 war.

Alternativ zu dieser Methode wurden Zellen in Anlehnung an Ma *et al.* (2004)<sup>166</sup> durch Kultivierung in LB-Medium mit einem pH von 5,5 an den Säurestress adaptiert, bevor der pH-Wert durch Zugabe von LB-Medium mit pH 1,75 bei einer OD<sub>600</sub> von 0,7 auf 2,5 eingestellt wurde. Unmittelbar danach, sowie nach 1,5 Stunden Inkubation bei 37°C, wurden Verdünnungsreihen auf LB-Agar ausplattiert, um die CFU/ml zu bestimmen.

### 2.8.2 Persistenzassay

Eine Vorkultur, die an zwei aufeinanderfolgenden Tagen jeweils in frischem LB-Medium angezogen wurde, wurde 1:100 in 50 ml frisches LB-Medium mit 100 µg/ml Ampicillin, respektive 100 µg/ml Ofloxacin, überimpft und bei 37°C aerob kultiviert. Stündlich wurde eine Probe von 100 µl entnommen und Verdünnungsreihen davon auf LB-Agar ausplattiert. Nach 12 Stunden Inkubation bei 37°C wurden die Kolonien gezählt und die CFU/ml (*colony forming units/ml*) bestimmt.

### 2.8.3 Generierung von VBNCs

Zellen wurden wie folgt in einen Ruhezustand (VBNC, *viable but non-culturable*) versetzt, aus welchem eine Kultivierung mit Standardmethoden, wie LB-Agarplatten, *per se* unmöglich ist. Eine aerob bei 37°C kultivierte LB-Übernachtskultur wurde 1:1.000 in frisches LB-Medium verdünnt und die Zellen bis zu einer OD<sub>600</sub> von 1,0 aerob bei 37°C kultiviert. Die Zellen wurden anschließend zentrifugiert (4.500 x g, 20 Minuten, 4°C) und das Pellet in 1,0 l frischem LB-Medium in Glasflaschen (1,0 l; Schott) resuspendiert. Die Kulturen wurden ab diesem Zeitpunkt für die gesamte Dauer des Experiments oder maximal 120 Tage bei 4°C gelagert. In regelmäßigen Abständen wurde der Zustand der Kultur kontrolliert. Die Kultivierbarkeit wurde durch Ausplattieren von Verdünnungsreihen auf LB-Agar festgestellt, die Gesamtzellzahl wurde mit Hilfe von Thoma-Zählkammern (0,0025 mm<sup>2</sup>; 0,1 mm Tiefe) ermittelt und die Viabilität der Zellen mittels Färbung mit Propidiumiodid (1 µM, vgl. 2.7.2) und anschließender Durchflusszytometrie (vgl. 2.7.3) und Fluoreszenzmikroskopie (vgl. 2.9.5) untersucht.

### 2.8.4 Visuelle Detektion der Cellulosesynthese von Makrokolonien

Die Synthese von Cellulose und Curli wurde nach einer Methode der Arbeitsgruppe von Prof. Dr. Hengge<sup>180, 272</sup> untersucht. Frisches LB-Medium wurde 1:500 mit einer Übernachtskultur inokuliert und bei 37°C aerob bei 950 rpm auf einem Thermblock (Eppendorf) bis zu einer OD<sub>600</sub> von ~0,3 kultiviert. 5 µl dieser Kultur wurden auf LB-Agarplatten ohne NaCl getropft, luftdicht mit Parafilm verschlossen und für bis zu 14 Tage bei 30°C inkubiert. Als Ergänzung dazu wurden Makrokolonien zur Visualisierung von Curli auf LB-Agarplatten mit 40 µg/ml des amyloidbindenden Farbstoff Kongorot kultiviert oder, um die Cellulose hervorzuheben, auf LB-Agarplatten mit 0,01% „Calcofluor White M2R“ (remel) angezogen. Die Fluoreszenz wurde dabei in einem Transilluminator (BioRad; Software: „Gel Doc™ EZ Imager“) detektiert.

## 2.9 Mikroskopie

### 2.9.1 Probenvorbereitung

Zellen wurden zur Vermeidung von Fluoreszenzartefakten unter komplettem Ausschluss von Detergenzien in Glasröhrchen inkubiert. Medien und Puffer wurden ausschließlich in mit Wasser gespülten Glasflaschen hergestellt.

### 2.9.1.1 Agarpads

Um die Mobilität von Zellen für die Mikroskopie einzuschränken, wurden 5 µl einer Bakterienkultur auf zuvor auf Glasobjektträgern vorbereitete Agarpads [1% (w/v) Agar in PBS] mit einer Schichtdicke von ~1 mm getropft und mit einem Deckgläschen bedeckt.

### 2.9.1.2 Ibidi-Kammern

Bakterienzellen in Flüssigkultur, bzw. Puffer, wurden mittels Poly-L-Lysin auf die Glasoberfläche eines gekammerten Objektträgers (µ-Slide 8 well, glass bottom, sterilized von ibidi®) adhärirt. Dazu wurden die Wells mit Poly-L-Lysin beschichtet, indem pro Kammer je 300 µl einer 0,02% (w/v) Poly-L-Lysin-Lösung über Nacht bei Raumtemperatur inkubiert wurde. Nicht gebundenes Poly-L-Lysin wurde durch dreimaliges Waschen mit H<sub>2</sub>O<sub>dd</sub> entfernt. Anschließend wurden die Kammern mindestens eine Stunde lang mit 300 µl des versuchsspezifischen Mediums bzw. Puffers äquilibriert. Die Zellen wurden mittels Zentrifugation (2.000 x g, 10 Minuten) in eigens dafür angefertigten Zentrifugationsadaptern aus Schaumstoff an die Oberfläche angeheftet. Überschüssige Zellen wurden durch fünfmaliges Waschen mit dem bereits für das Äquilibrieren verwendeten Medium bzw. Puffer entfernt. Die Kammern wurden für die Mikroskopie mit einem finalen Volumen von 300 µl aufgefüllt.

### 2.9.2 Zellfixierung

Zellen wurden durch Zugabe von 25% (w/v) Paraformaldehyd (gelöst in PBS Puffer pH 7,5) mit einer Endkonzentration von 4% für 30 Minuten bei 37°C inkubiert. Die Reaktion wurde durch die Zugabe von Glycin (15 mM; Stockkonzentration 1,5 M) während einer zehnminütigen Inkubation bei 37°C abgestoppt. Nach Zentrifugation (15.000 x g, 1 Minute) wurde das Zellpellet in PBSG Puffer (PBS mit 20 mM Glucose) resuspendiert.

### 2.9.3 Kultivierung von Zellen für die Fluoreszenzmikroskopie

*E. coli* Stämme wurden über Nacht in LB Medium angezogen und am nächsten Tag 1:500 in frisches M9 Medium überführt, in welchem sie für ca. 4 Stunden bei 37°C aerob bis zu einer OD<sub>600</sub> von 0,4 kultiviert wurden. 1 ml der Kultur wurde wie in 2.9.3 beschrieben mit 1-4% Paraformaldehyd fixiert. Nach Pelletieren und Resuspendierung in 100 µl PBSG Puffer wurden 30 µl in bereits mit Poly-L-Lysin präparierten, gekammerten Objektträgern (vgl. 2.9.1.2) vorliegenden 300 µl PBSG Puffer überführt. Die Zellen wurden wie in 2.9.1.2 beschrieben mittels Zentrifugation an die Glasoberfläche adhärirt.

### 2.9.4 Immunofluoreszenz

Zuvor fixierte und in IBIDI-Kammern immobilisierte Zellen oder Sphäroblasten wurden über Nacht in PBS Puffer mit je nach Versuchsfragestellung entsprechend verdünnten Antikörpern inkubiert. Nicht gebundene Antikörper wurden durch dreimaliges Waschen mit PBSG Puffer entfernt. Die Fluoreszenz wurde im hochauflösenden Mikroskop „Delta Vision Elite“ (GE Healthcare, Applied Precision) untersucht.

Um das Fluoreszenzsignal von EnvZ-PAGFP oder MzrA-PAmCherry Konstrukten zu amplifizieren, wurden diese mit spezifischen, mit Fluoreszenzfarbstoffen markierten Einzeldomänenantikörpern (*Nanobodies*, Firma Chromotek) detektiert. Die Durchlässigkeit der bakteriellen Zellwand von zuvor immobilisierten und fixierten Zellen wurde durch eine fünfminütige Inkubation mit Lysozym (1 mg/ml) bei 30°C erhöht<sup>273</sup>. Es folgte ein Blocken mit 2% (w/v), in PBST gelöstem BSA. Anschließend wurde der 1:200 in PBST mit 2% BSA (w/v) verdünnte Antikörper für eine Stunde bei Raumtemperatur inkubiert. Nach sechsmaligem Waschen mit PBST wurden die Zellen durch Zugabe von 4% Paraformaldehyd ein zweites Mal fixiert. Abschließend wurden die Proben sechs Mal in PBST gewaschen und für die Mikroskopie mit PBSG Puffer bedeckt.

Bakterien wurden alternativ nach einem adaptierten Protokoll von Di Paolo *et al.* (2016)<sup>274</sup> mittels Elektroporation mit Einzeldomänenantikörpern (*Nanobodies*, Firma Chromotek) transformiert. Dafür wurden *E. coli* Stämme bis zu einer OD<sub>600</sub> von 0,5 in 1 ml M9 Medium inkubiert, anschließend geerntet (15.000 x g, 30 Sekunden, 4°C) und drei Mal in eiskaltem 10% (v/v) Glycerin gewaschen. Nach Resuspendierung in 100 µl 10% (v/v) Glycerin wurden *Nanobodies* in einer 1:500 Verdünnung hinzugegeben. Die Proben wurden vorsichtig vermischt und für 10 Minuten auf Eis inkubiert. Es folgte eine Elektroporation in vorgekühlten Küvetten (12 kV/cm). Nach Zugabe von 1 ml auf 30°C erwärmten M9 Medium wurden die Proben in ein Eppendorf-Gefäß überführt und für 30 Minuten bei 30°C in einem Thermomixer (Eppendorf) bei 450 rpm inkubiert. Die Proben wurden daraufhin dreimal mit PBSG Puffer gewaschen und wie in 2.9.3 beschrieben fixiert.

### 2.9.5 Weitfeldmikroskopie mit „DMi8“

Für die Fluoreszenz- und Phasenkontrastmikroskopie mittels Weitfeldmikroskopie wurde ein „DMi8“ Mikroskop (Leica; Software „LAS X“) in Kombination mit der Kamera „DFC365 FX 337“ (Leica) verwendet. Rote Fluoreszenz wurde bei 546 nm angeregt die Emission bei 365 nm mit einem entsprechenden Filter mit 75 nm Bandbreite detektiert. Für grüne Fluoreszenz wurde eine Anregungswellenlänge von 485 nm gewählt, der Emissionsfilter entsprach 510 nm mit 75 nm Bandbreite.

### 2.8.6 Hochauflösende Mikroskopie mit „Delta Vision Elite“

Das hochauflösende Mikroskop „Delta Vision Elite“ (GE Healthcare, Applied Precision) basiert auf dem System der 3D-Dekonvolution („Image Restoration“). Dabei kamen „Standard Insight SSI<sup>TM</sup>“ Vierfarben-Modul, ein X4 Lasermodul und die „CoolSnap HQ2 CCD“ Kamera zur Anwendung. Die Proben wurden in einer Klimakammer bei konstanter Temperatur von 30°C mit dem Objektiv „100x oil PFS U-Plan S-Apo 1.4 NA“ unter Verwendung eines Immersionsöls mit Brechungsindex 1,52 (Cargille) untersucht. Fluoreszenz wurde mit Hilfe von passenden Filtern visualisiert (Grün/GFP: 475/28 nm Anregungsfilter und 523/36 nm Emissionsfilter; Rot/mCherry: 575/25

nm Anregungsfilter und 632/60 nm Emissionsfilter, Rot/Atto647N: 632/22 nm Anregungsfilter und 676/34 nm Emissionsfilter). Photoaktivierbare Fluorophore wie PAGFP oder PAmCherry wurden durch Belichtung bei 390/18 nm mit 100% Intensität für 45 Sekunden aktiviert.

### 2.9.7 Photoaktivierte Lokalisationsmikroskopie (PALM)

Zur Lokalisation von EnvZ-PAGFP, EnvZ-mNeonGreen, sowie MzrA-mCherry mittels 3D PALM/STORM wurde das Mikroskop „Elyra P1“ von Zeiss mit dem Objektiv „Alpha Plan-Apochromat 100x/1.46 oil objective DIC M27“ in Kombination mit der EM-CCD Kamera „iXon DU 897“ (Andor) verwendet. Die Aufnahmen wurden mit Hilfe der Software „ZenBlack“ (Zeiss) verrechnet, wobei eine laterale Verschiebung des Bildausschnitts durch die Verwendung von fluoreszierenden Beads („TetraSpeckTN Microspheres; 0,1  $\mu\text{m}$ , fluorescent blue/green/orange/dark red“ von Thermo Fisher Scientific) korrigiert wurde. Diese wurden während der Probenvorbereitung zusammen mit den Zellen durch Zentrifugation an die Glasoberfläche von gekammerten Objektträgern (ibidi) adhärirt (vgl. 2.9.1.2).

Die PALM-Mikroskopie, sowie die Auswertung der daraus resultierenden Datensätzen wurde von Giacomo Giacomelli aus der Arbeitsgruppe von Prof. Dr. Marc Bramkamp im Bereich Mikrobiologie des Biozentrums der LMU durchgeführt.

PAmCherry wurde mit Hilfe eines Lasers bei 405 nm Lichtwellenlänge (HR Diode, 50 mW Laser) aktiviert und bei 561 nm angeregt. Die Fluoreszenz wurde wie von Bach *et al.* beschrieben unter Verwendung eines passenden Filters detektiert (570 nm)<sup>275</sup>. Dabei kamen folgende Parameter zur Anwendung: peak size: 9 pixel (1 pixel = 100 nm), peak intensity to noise ratio: 6. Zur Stabilisierung der Z-Achse wurde das „Definite Focus“ System verwendet. PAmCherry wurde bei 405 nm unter Verwendung eines linearen Gradienten in zwei Phasen (0,1 – 1% und 1 – 10% Laserintensität) von je 2.000 *frames* aktiviert. Die Laserintensität wurde so gewählt, dass die Wahrscheinlichkeit einer zeitgleichen Konversion von zwei PAmCherry Molekülen in geringer räumlicher Entfernung minimiert wurde. Aktiviertes PAmCherry wurde bei 561 nm angeregt (15%, 50 ms Belichtungszeit, EMCCD gain = 200). Die Fluoreszenzereignisse wurden nach Photonenzahl (60-350 Photonen) und PSF (*half maximum width* = 60 – 160 nm) gefiltert, um unspezifische Hintergrundsignale und das gleichzeitige Auftreten mehrerer PAmCherry Moleküle auszusparen.

mNeonGreen wurde bei 488 nm angeregt (15%, 50 ms Belichtungszeit, EMCCD gain = 200).

Um PAGFP zu detektieren, wurden verschiedene Parameter ausgetestet, was jedoch kein ausreichend hohes Signal zur Folge hatte. Grund dafür ist vermutlich, dass PAGFP auch in Abwesenheit des 405 nm Lasers aktiviert wird und bei Belichtung mit 488 nm Aktivierung und Anregung zusammenfallen<sup>276</sup>.

Für jeden untersuchten *E. coli* Stamm wurden etwa 500 Zellen ausgewertet. Die Koordinatenbasierte-Kolokalisationsanalyse (CBC, *coordinate-based colocalization analysis*) wurde nach Malkusch *et al.* (2012)<sup>277</sup> durchgeführt.

## 2.10 Bioinformatische Methoden

### 2.10.1 Datenbanken

Für die Planung der Experimente wurden DNA Sequenzen von der NCBI Datenbank „Genomes“ heruntergeladen. Informationen bezüglich Proteinsequenz und Vorhersage der molekularen Masse wurden von *UniProt*<sup>278</sup> bezogen. Stoffwechselwege wurden unter anderem bei *Ecocyc*<sup>279</sup> und *KEGG Pathway*<sup>280, 281</sup> recherchiert. Proteinnetzwerke, bzw. indirekte oder direkte Zusammenhänge von mehreren Proteinen, wurden mit Hilfe der Datenbank „STRING“ (Version 10.5) analysiert<sup>282</sup>.

Das Gesamtproteom von *E. coli* K-12 MG1655 wurde von der Datenbank „UniProt“<sup>278</sup> bezogen (UniProt Proteome ID: ECOLI - UP000000625, letzte Modifikation 2.4.2018).

Zur Erstellung eines phylogenetischen Stammbaumes von EnvZ, EvgS und DgcC wurden nichtredundante Orthologe durch eine *BLAST* Suche der *UniProt*<sup>278</sup> Datenbank „Microbial Proteomes“ ermittelt, indem die Gesamtsequenz des jeweiligen Proteins aus *E. coli* K-12 MG1655 (UniProt Proteome ID: ECOLI - UP000000625, letzte Modifikation 2.4.2018) als Vorlage verwendet wurde (Expect value: 1; auto matrix; allow gaps). Sequenzen, die kürzer als 90% der entsprechenden *E. coli* K-12 Proteinsequenzen waren, wurden ausgeschlossen. Orthologe der *E. coli* K-12 Lon Protease wurden analog dazu aus der UniProt Datenbank „Reference Proteomes“ ermittelt.

Nach Orthologen von EnvZ und MzrA in definierten Organismen wurde mit Hilfe der NCBI Datenbank „Protein BLAST“<sup>283</sup> gesucht. Folgende Parameter kamen zur Anwendung: blastp algorithm; expect threshold 10; matrix BLOSUM62.

Teile der Analyse zur Verteilung und Konservierung von polyP Motiven wurden von Dr. Fei Qi (Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitrij Frishman) durchgeführt. Dafür wurden Proteome von 43 ausgewählten *E. coli* Stämmen (33 Wildtypstämme, 10 Laborstämme, darunter *E. coli* K-12 MG1655: Proteome ID ECOLI, letzte Modifikation am 3.3.2009) und Informationen bezüglich orthologer Gruppen mit Hilfe der unabhängigen „OMA“ Software<sup>284</sup> von der „OMA“<sup>285</sup> Datenbank bezogen. OMA Gruppen sind als zueinander orthologe Gensets definiert und schließen damit Genduplikationen aus. Proteindomänen in *E. coli* K-12 MG1655 wurden mit Hilfe der Datenbank CATH<sup>286</sup> und der Datenbank „Gene3D“<sup>287</sup> lokalisiert. Transmembrandomänen wurden separat untersucht. Dazu wurden Transmembranproteine und die Informationen zur Lokalisation der Transmembranabschnitte aus *E. coli* K-12 MG1655 (UniProt Proteome ID: ECOLI - UP000233379) von der *UniProt*<sup>278</sup> Datenbank heruntergeladen.

### 2.10.2 Auswahl repräsentativer *E. coli* Stämme

*E. coli* Stämme wurden nach Clermont *et al.* (2013)<sup>288</sup> in phylogenetische Gruppen klassifiziert. Dafür wurden die Genome der Stämme von der NCBI Datenbank „Genomes“ heruntergeladen und ein *in silico* Annealing der empfohlenen *Primer*-Sequenzen zur Amplifizierung der Markergene *arpA*, *chuA*, *yjaA* und *TspE4* durchgeführt.

**Tab. 10:** Aufstellung und Klassifizierung der im Rahmen dieser Studie verwendeten *E. coli* Stämme in phylogenetische Serogruppen.

taxonomy id	OMA id	Bezeichnung	Phylogenetische Gruppe	Pathogenität
1133853	ECO1E	<i>Escherichia coli</i> O104:H4 (strain 2009EL-2071)	A	ja (EAEC)
511693	ECOB B	<i>Escherichia coli</i> (strain B / BL21)	A	nein
469008	ECOB D	<i>Escherichia coli</i> (strain B / BL21-DE3)	A	nein
413997	ECOB R	<i>Escherichia coli</i> (strain B / REL606)	A	nein
595496	ECOB W	<i>Escherichia coli</i> (strain K12 / MC4100 / BW2952)	A	nein
536056	ECOD 1	<i>Escherichia coli</i> (strain ATCC 33849 / DSM 4235 / NCIB 12045 / K12 / DH1)	A	nein
316385	ECOD H	<i>Escherichia coli</i> (strain K12 / DH10B)	A	nein
316401	ECOH 1	<i>Escherichia coli</i> O78:H11 (strain H10407)	A	ja (ETEC)
331112	ECOH S	<i>Escherichia coli</i> O9:H4 (strain HS)	A	nein
481805	ECOL C	<i>Escherichia coli</i> (strain ATCC 8739 / DSM 1576 / Crooks)	A	nein
511145	ECOL I	<i>Escherichia coli</i> (strain K12 / MG1655)	A	nein
1040638	ECOL X	<i>Escherichia coli</i> O104:H4 LB226692	A	ja (EHEC)
585395	ECO1 0	<i>Escherichia coli</i> O103:H2 (strain 12009)	B1	ja (EHEC)
585396	ECO1 A	<i>Escherichia coli</i> O111:H- (strain 11128)	B1	ja (EHEC)
331111	ECO2 4	<i>Escherichia coli</i> O139:H28 (strain E24377A)	B1	ja (ETEC)
573235	ECO2 6	<i>Escherichia coli</i> O26:H11 (strain 11368)	B1	ja (EHEC)
585055	ECO5 5	<i>Escherichia coli</i> (strain 55989)	B1	ja (EAEC)
585034	ECO8 A	<i>Escherichia coli</i> O8 (strain IA11)	B1	nein
595495	ECOK O	<i>Escherichia coli</i> (strain ATCC 55124 / KO11)	B1	nein
566546	ECOL W	<i>Escherichia coli</i> (strain ATCC 9637 / CCM 2024 / DSM 1116 / NCIMB 8666 / NRRL B-766 / W)	B1	nein
409438	ECOSE	<i>Escherichia coli</i> (strain SE11)	B1	nein
574521	ECO2 7	<i>Escherichia coli</i> O127:H6 (strain E2348/69)	B2	ja (EPEC)
585035	ECO4 5	<i>Escherichia coli</i> O45:K1 (strain S88)	B2	ja (ExPEC)
585057	ECO7 1	<i>Escherichia coli</i> O7:K1 (strain IA139)	B2	ja (ExPEC)
585397	ECO8 1	<i>Escherichia coli</i> O81 (strain ED1a)	B2	nein
685038	ECO8 N	<i>Escherichia coli</i> O83:H1 (strain NRG 857C)	B2	ja (AIEC)
655817	ECOAB	<i>Escherichia coli</i> OR:K5:H- (strain ABU 83972)	B2	ja (ABU)
885275	ECOC 1	<i>Escherichia coli</i> (strain clone D i14)	B2	ja (UPEC)
885276	ECOC 2	<i>Escherichia coli</i> (strain clone D i2)	B2	ja (UPEC)
405955	ECOK 1	<i>Escherichia coli</i> O1:K1	B2	ja (APEC)
714962	ECOK I	<i>Escherichia coli</i> O18:K1:H7 (strain IHE3034)	B2	ja (ExPEC)
362663	ECOL 5	<i>Escherichia coli</i> O6:K15:H31 (strain 536)	B2	ja (UPEC)
199310	ECOL 6	<i>Escherichia coli</i> O6:H1 (strain CFT073 / ATCC 700928)	B2	ja (UPEC)
431946	ECOS 5	<i>Escherichia coli</i> O150:H5 (strain SE15)	B2	nein
869729	ECOUM	<i>Escherichia coli</i> (strain UM146)	B2	ja (AIEC)
364106	ECOUT	<i>Escherichia coli</i> (strain UTI89)	B2	ja (UPEC)
216592	ECO4 4	<i>Escherichia coli</i> O44:H18 (strain 042 )	D	ja (EAEC)
585056	ECOL U	<i>Escherichia coli</i> O17:K52:H18 (strain UMN026)	D	ja (ExPEC)
439855	ECOSM	<i>Escherichia coli</i> (strain SMS-3-5 )	D	nein
155864	ECO5 7	<i>Escherichia coli</i> O157:H7	E	ja (EHEC)
444450	ECOSE	<i>Escherichia coli</i> O157:H7 (strain EC4115)	E	ja (EHEC)
544404	ECO5 T	<i>Escherichia coli</i> O157:H7 (strain TW14359C)	E	ja (EHEC)
701177	ECOCB	<i>Escherichia coli</i> O55:H7 (strain CB9615)	F	ja (EPEC)

### 2.10.3 Clusterformierung von Proteinen

Proteine wurden mittels „CD-HIT“<sup>289</sup> zu Clustern zusammengefasst. Der Grenzwert für eine Sequenzidentität wurde auf 0,8 gesetzt und der Grenzwert für die Sequenzlänge auf 0,7. Die Analyse wurde von Dr. Fei Qi (Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitrij Frishman) durchgeführt.

### 2.10.4 Unterscheidung zwischen Kernproteom und akzessorischem Proteom

Kernproteom und akzessorisches Proteom wurden auf Basis der methodischen Beschreibungen von Lukjancenکو *et al.* (2010)<sup>290</sup> und Friis *et al.* (2010)<sup>291</sup> definiert. Dafür wurden die in orthologe Gruppen eingeteilten Proteine hinsichtlich ihrer Abundanz verglichen. Eine orthologe Gruppe wurde dem Kernproteom zugeordnet, wenn sie in allen untersuchten *E. coli* Stämmen präsent war, andernfalls wurde sie in das akzessorische Proteom klassifiziert. Alle Proteine, die nicht Teil einer orthologen Gruppe sind, wurden zum akzessorischen Proteom gezählt. Die

Analyse wurde von Dr. Fei Qi (Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitrij Frishman) durchgeführt.

### 2.10.5 Identifikation und Klassifizierung von polyP Motiven

Ein polyP-Motiv wurde durch das Auftreten von mindestens zwei konsekutiven Prolinen definiert. Polyprolinmotive,  $X_{(-2)}X_{(-1)}-nP-X_{(+1)}$ , wurden entsprechend ihrem vorhergesagtem Potential, einen Translationsarrest zu induzieren, wie von Qi *et al.* (2018)<sup>292</sup> und in dieser Arbeit beschrieben, in drei Gruppen eingeteilt: stark, moderat und schwach. Diese Vorhersage basiert auf systematischen *in vivo* und *in vitro* Analysen mit *E. coli*<sup>30, 61, 293</sup>. In Anlehnung an die bereits untersuchten Motive, wurde ein Regelsatz für die Vorhersage des Translationsstopps durch alle übrigen Aminosäurekombinationen erstellt.

PolyP Motive im zuvor von der der *UniProt* Datenbank "Proteomes" bezogenen Proteom von *E. coli* K-12 MG1655 wurden mit der Software "Excel" (Microsoft) ausgezählt.

Im Rahmen einer umfassenden Analyse des Auftretens der Motive in insgesamt 43 *E. coli* Stämmen oder spezifischen Proteindomänen wurden PolyP Motive mit Hilfe des Programms „fuzzpro“<sup>294</sup> (EMBOSS package) identifiziert (Dr. Fei Qi, Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitrij Frishman).

Das Arrestpotential der polyP Motive wurde von Dr. Fei Qi (Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitrij Frishman) durch folgende mathematische Modelle beschrieben.

Das Gesamtpotential eines gegebenen Proteins, einen Translationsarrest zu induzieren wurde durch die Anzahl und Art der darin vorkommenden polyP Motive definiert:

$$\text{Stalling\_score}_i = 3 \times N_S \times 2 \times N_M \times 1 \times N_W$$

Dabei beschreiben  $N_S$ ,  $N_M$  und  $N_W$  die Anzahl von starken (strong: S), moderaten (moderate: M) und schwachen (weak: W) polyP Motiven in einem bestimmten Protein.

Die von Starosta *et al.* (2014)<sup>61</sup> erhobenen Daten zur Translationseffizienz eines CadC-XPPP-LacZ Konstruktes wurden mit folgender Formel quantitativ beschrieben:

$$\text{Motif\_score}_i = \left( 1 - \frac{\text{REA}_i}{\text{REA}_{\text{positive\_control}}} \right) \times 100$$

$$\text{REA}_i = \frac{\text{EA}_{i \Delta \text{efp}}}{\text{EA}_{i \text{wild\_type}}}$$

Dabei ist REA die relative  $\beta$ -Galactosidase-Enzymaktivität. EA steht für die Enzymaktivität in Miller-Units. Diese Einordnung wurde durch folgende Formel auf HOGs übertragen:

$$\text{HOG\_Score}_i = \frac{\sum \text{Motif\_score}_i}{\text{Protein\_number}_i}$$

HOGs (*hierarchical orthologous groups*) wurden von der OMA Datenbank<sup>285</sup> erhoben, sind als ein Set von Genen definiert, die auf einen gemeinsamen Vorfahren innerhalb einer taxonomischen Gruppe zurückgehen und beinhalten sowohl Orthologe, als auch Paraloge.  $\sum Motif\_score_i$  ist die Summe der *Scores* aller XPPP Motive im HOG *i*.

### 2.10.6 Untersuchungen zur Akkumulation von polyP Motiven

Die Wahrscheinlichkeit des Auftretens einer bestimmten Aminosäure *X* im Proteom von *E. coli* K12 MG1655 (UniProt Proteome ID: ECOLI - UP000000625) wurde basierend auf der Gesamtanzahl aller Aminosäuren ermittelt. Die Anzahl der bei einer Zufallsverteilung erwarteten Aminosäuredopplungen wurde in Annäherung an die Wahrscheinlichkeit des Auftretens der Motive an jedmöglicher Position im Proteom bestimmt. Dabei wurde miteinander berechnet, dass ein Motiv *XX* nicht an letzter Position eines Proteins lokalisiert sein kann:

$$n(XX)_{erwartet} = \left( \frac{n_{Aminosäure\ X}}{n_{alle\ Aminosäuren}} \right)^2 \times (n_{alle\ Aminosäuren} - n_{Proteine})$$

Darauf basierend wurde das Verhältnis zwischen tatsächlicher und erwarteter Anzahl an *XX* berechnet (Frequenz):

$$Frequenz(XX) = \frac{n(XX)_{tatsächlich}}{n(XX)_{erwartet}}$$

Die Häufigkeit von polyP Motiven in weiteren 42 ausgewählten *E. coli* Stämmen und in spezifischen Proteinstrukturen wurde von Dr. Fei Qi (Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitrij Frishman) wie folgt analysiert. Zur Normalisierung des Auftretens von polyP Motiven in Vollängenproteinen wurden durch eine zufallsbasierte Aneinanderreihung der Aminosäuren eines gegebenen Proteins hypothetische Proteinsequenzen generiert. Dieses Verfahren wurde mit dem Programm „shuffleseq“<sup>294</sup> (EMBOSS package) durchgeführt und wurde pro Protein 1.000 mal wiederholt. Die Werte wurden in einem Vektor zusammengefasst, der das Verhältnis der tatsächlichen zur zufällig generierten Anzahl an polyP Motiven beschreibt. Mit Hilfe des Mann-Whitney-Wilcoxon-Tests wurde die Signifikanz eines Unterschiedes zwischen zwei Vektoren bestimmt. Die Methode wurde für jeden *E. coli* Stamm separat angewandt.

Mit Hilfe des Algorithmus „SPatt“<sup>295</sup> wurde die Häufigkeit von polyP Motiven in spezifischen Proteinstrukturen analysiert. Die Methode basiert auf einem Markov-Kettenmodell der Ordnung *m* („model *Mm*“). Für ein Motiv der Länge *l* ist *m* äquivalent zu *l*-2. Da ein polyP Motiv vorrangig durch zwei Aminosäuren (PP) definiert ist, wurde das Modell *M0* gewählt. Dabei wurde die tatsächliche Anzahl von polyP Motiven mit der erwarteten Anzahl verglichen und ein p-Wert für die Signifikanz einer Abweichung berechnet. Frequenzen jeder einzelnen Aminosäure *X* und deren Dopplung *XX* wurden mit Hilfe des Programms *compseq* des EMBOSS<sup>294</sup> package berechnet.

Die *n*-fache Veränderung des Auftretens, bzw. die Frequenz von polyP Motiven wurde wie im ersten Absatz beschrieben definiert. Für die Bestimmung der Häufigkeit in *E. coli* Proteomen ist  $n(XX)_{erwartet}$  der Mittelwert aus dem zufallsgenerierten Datensatz aus 1.000 Sequenzen. Für die Analyse der Frequenz von polyP Motiven in

spezifischen Proteindomänen, Domänenlinkern oder Transmembranhelices wurde auf den durch den „SPatt“ Algorithmus ermittelten Mittelwert der erwarteten Anzahl an polyP Motiven zurückgegriffen.

## 2.10.7 Sequenzalignment

### 2.10.7.1 Sequenzalignment von EnvZ, EvgS, DgcC und Lon Protease

Ein paarweises Alignment von *E. coli* K-12 MG1655 EnvZ, EvgS, DgcC und Lon Protease Sequenzen wurde mit einem progressiven Algorithmus der Software „CLC Workbench 7.6“ (CLC Bio Qioagen) mit den folgenden Parametern durchgeführt: gap open cost 10; gap extension cost 1; high accuracy.

### 2.10.7.2 Multiples Sequenzalignment von Proteinclustern

Das Alignment von verschiedenen weiteren Proteinsequenzen zur Untersuchung des Auftretens von polyP Motiven wurde mit „Clustal Omega“<sup>296</sup> durchgeführt (Dr. Fei Qi, Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitrij Frishman). Die Sequenz mit den meisten konsekutiven Prolinen wurde als Referenz verwendet und mit allen anderen Sequenzen verglichen. Identische Sequenzen wurden nur einmal gezählt und die veränderten Sequenzen klassifiziert. Ein polyP Motiv wurde als konserviert bezeichnet, falls alle betrachteten Sequenzen in dieser Region identisch sind.

Die Konservierung eines Motivs, bzw. das Verhältnis zwischen Anzahl an veränderten und unveränderten Motiven, wurde mit folgender Formel beschrieben:

$$Ratio_{i,s} = \frac{N_{i,s} / \sum_i N_{i,s}}{\sum_s N_{i,s} / \sum_{i,s} N_{i,s}}$$

$$Delta\_ratio_i = ratio_{i,conserved} - ratio_{i,changed}$$

Dabei ist  $N_{i,s}$  die Anzahl der Motive  $i$  im Zustand  $s$ . Mit Hilfe des Programms „Needle“ (EMBOSS package)<sup>294</sup> wurden alle möglichen paarweisen Alignments der Proteinsequenzen eines gegebenen Clusters, sowie deren Sequenzübereinstimmungen bestimmt. Die durchschnittliche Sequenzidentität dieses Clusters wurde als das arithmetische Mittel der Sequenzidentitäten aller möglichen paarweisen Alignments definiert.

## 2.10.8 Rekonstruktion von phylogenetischen Stammbäumen

### 2.10.8.1 Phylogenetische Stammbäume von EnvZ, EvgS, DgcC und Lon Protease

Phylogenetische Stammbäume von *E. coli* K-12 MG1655 EnvZ, EvgS, DgcC und Lon Protease wurden auf Basis zuvor erstellter Sequenzalignments mit dem „high-accuracy, distance based, neighbour joining“ Algorithmus der Software „CLC Workbench 7.6“ konstruiert (*default parameters*: 100 bootstrap replicates, Jukes-Cantor distance correction). Die Stammbäume wurden alternativ mit Hilfe des Web-Tools *iTol*<sup>297</sup> visualisiert.

### 2.10.8.2 Phylogenetischer Stammbaum von weiteren Proteinclustern aus 43 *E. coli* Stämmen

Phylogenetische Stammbäume aller polyP enthaltenden Proteine aus 43 *E. coli* Stämmen wurden mit der Software „PhyML“<sup>298</sup> rekonstruiert. Deren Wurzel wurde zentriert definiert. Um Veränderungen im translationsarrestinduzierenden Potentials eines polyP Motivs im Laufe der Evolution nachzuverfolgen, wurden zuerst die externen Knotenpunkte des phylogenetischen Stammbaumes untersucht und der jeweiligen polyP Motivklasse zugeordnet („stark“, „moderat“, „schwach“ und „kein Motiv“). Mit Hilfe des *Maximum Likelihood* Algorithmus<sup>299</sup> wurden dann die internen Knotenpunkte, welche die evolutionären Vorgänger beschreiben, rekonstruiert. Die normalisierte Anzahl von Ereignissen, die zu einer Veränderung eines polyP Motivs führen wurden mit Hilfe der folgenden Formel beschrieben:

$$Number\_normalized_i = \frac{Number_i}{Total\_branch\_length_i}$$

Dabei ist  $Number_i$  die Anzahl an Ereignissen innerhalb eines spezifischen COGs (*cluster of orthologous groups*) und  $Total\_branch\_length_i$  die Gesamtlänge des phylogenetischen Stammbaumes aus dem COG i. COGs, deren Gesamtlänge  $<0,01$  ist, wurden ausgeschlossen, da eine Normalisierung eine überproportional große Anzahl an Ereignissen zur Folge hätte. Die Analyse wurde von Dr. Fei Qi (Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitriy Frishman) durchgeführt.

### 2.10.9 Propensity of Gene Loss (PGL) und Propensity of Stalling Effect Change (PSEC)

Die Neigung zum Genverlust (*propensity of gene loss*: PGL), wird mit einer von Krylov *et al.* (2003)<sup>300</sup> beschriebenen Methode berechnet. Diese basiert auf der Annahme, dass die Wahrscheinlichkeit eines Genverlusts mit der Dauer der Zeit sinkt, während der das Gen hätte verloren gehen können, es jedoch nicht tat. Die PGL ist daher definiert als Verhältnis aus Gesamtlänge der Äste, in welchen das Gen tatsächlich verloren geht und der Gesamtlänge an Ästen, in welcher das Gen hätte verloren gehen können<sup>301, 302</sup>. Dadurch können die Zweige des phylogenetischen Stammbaumes identifiziert werden, in welchen die jeweiligen Gene verloren ( $B_L$ ) oder beibehalten ( $B_P$ ) wurden. Die PGL wird wie folgt berechnet:

$$PGL = \frac{\sum B_L}{\sum B_P + \sum B_L}$$

Analog dazu wurde hier die Neigung zu einer Veränderung in der Klassifizierung von polyP Motiven als PSEC (*propensity of stalling effect change*) definiert. Da im Fall von polyP Motiven sowohl „Verlust“, als auch „Zugewinn“ an putativ translationsarretierendem Potential in Betracht gezogen werden muss, gestaltet sich die Formel etwas komplexer und PSEC ist definiert als die Differenz zwischen *propensity of stalling effect gain* (g) und *loss* (l):

$$PSEC = \frac{\sum B_g}{\sum B_{cg}} - \frac{\sum B_l}{\sum B_{cl}}$$

Dabei sind  $B_g$  und  $B_l$  die Zweiglängen, in denen das Motiv verstärkt (g, *gain*) oder abgeschwächt (l, *loss*) worden ist.  $B_{cg}$  und  $B_{cl}$  sind die Zweiglängen, während der die Änderung potentiell hätte stattfinden können. Damit

beschreibt eine positive PSEC eine Neigung des Motivs, während der Evolution „stärker“ zu werden, während eine negative PSEC den umgekehrten Fall bedeutet.

Die Analyse wurde von Dr. Fei Qi (Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitrij Frishman) durchgeführt.

#### 2.10.10 Berechnung von Proteinabundanz und Translationseffizienz

Daten bezüglich der Abundanz von Proteinen wurden von Wisniewski *et al.* (2014)<sup>303</sup> bezogen. Microarray Daten über Gentranskriptionslevel in *E. coli* MG1655 wurden von der ASAP<sup>304</sup> Datenbank heruntergeladen. Die Translationseffizienz eines gegebenen Proteins wurde durch folgende Formel berechnet:

$$Translation\_efficiency_i = \frac{Protein\_abundance_i}{Transkription\_level_i}$$

Die Analyse wurde von Dr. Fei Qi (Technische Universität München, Department of Genome-Oriented Bioinformatics, Arbeitsgruppe Prof. Dr. Dmitrij Frishman) durchgeführt.

#### 2.10.11 Vorhersage von Bindestellen von Transkriptionsfaktoren

Sequenzen von je 500 bp *upstream* des Startcodons bis zum Stoppcodon des zu untersuchenden Gens wurden mit Hilfe des Programms BPROM<sup>305</sup> von *Softberry* auf potentielle Bindestellen von Transkriptionsfaktoren untersucht. Das Programm erkennt bakterielle Sigma70 Promotorregionen mit 80% Spezifität und Genauigkeit und sagt potentielle Bindestellen partial (8 bp) voraus.

#### 2.10.11 Modellierung von dreidimensionalen Proteinstrukturen

Dreidimensionale Proteinstrukturen von EF-P, EnvZ, und Lon Protease wurden mit Hilfe des Tools „Phyre2“<sup>306</sup> auf Basis von Vollängen-Proteinsequenzen aus *E. coli* K-12 MG1655 (UniProt Proteome ID: ECOLI - UP000000625, letzte Modifikation 2.4.2018) vorhergesagt und modelliert. Bereits veröffentlichte Kristallstrukturen wurden aus dem „PDB“ Archiv der RCSB Proteindatenbank bezogen<sup>307</sup>. Die dreidimensionale Proteinstruktur der periplasmatischen EvgS Domänen wurde 2017 von Sen *et al.*<sup>156</sup> basierend auf der vorhandenen Kristallstruktur des homologen *Bordetella pertussis* Proteins BvgS modelliert und freundlicherweise als Datei zur Verfügung gestellt.

#### 2.10.12 Bildanalyse und graphische Darstellung

Intensitäten von Proteinbanden auf Polyacrylamidgelen und Western Blots wurden mit Hilfe der Software „ImageJ“<sup>308, 309</sup> ausgewertet. Banden im Western Blot wurden normalisiert auf die Intensität einer unveränderten Proteinbande im zugehörigen, mit TCE versetzten SDS-Gel. Proteinstrukturen wurden mit der Software „Chimera“ (UCSF Resource for Biocomputing)<sup>310</sup> dargestellt, während Graphiken und Illustrationen mit den Softwares „Illustrator CC“ (Adobe) und „SigmaPlot 12.5“ (Systat) erstellt wurden.

### 3. ERGEBNISSE

#### 3.1 Verteilung, Konservierung und Evolution von polyP Motiven in *E. coli*

Da trotz der bestehenden Gefahr eines Translationsarrestes zahlreiche Beispiele für eine Akkumulation von Prolinen in Proteinen bekannt sind, stellt sich die Frage, inwiefern das Auftreten von konsekutiven Prolinen einem bestimmten Muster folgt. Bislang konzentrierten sich Untersuchungen dazu hauptsächlich exemplarisch auf eukaryontische Proteindomänen im Zusammenhang mit substratbindenden Eigenschaften<sup>118, 46</sup>, sowie einzelne prokaryontische Proteine, wie CadC und ValS<sup>30, 59, 71</sup>. Für ein besseres globales Verständnis wurde in dieser Arbeit ein Schwerpunkt auf die systematische Analyse der Verteilung von Prolinen im Proteom des Modellorganismus *E. coli* gelegt. Von besonderem Interesse war die mit polyP Motiven verbundene Limitation der Translationseffizienz.

##### 3.1.1 Klassifizierung von polyP Motiven

Eine minimale Einheit aus zwei aufeinanderfolgenden Prolinen, in der vom Ribosom translatierten Peptidsequenz ist ausreichend, um die Effizienz der ribosomalen Peptidyltransferasereaktion signifikant herabzusetzen<sup>59, 311, 31, 75</sup>. *In vitro* und *in vivo* Studien von Peil *et al.* (2013)<sup>30</sup>, Starosta *et al.* (2014)<sup>61</sup> und Woolstenhulme *et al.* (2015)<sup>293</sup> zeigten, inwiefern bestimmte, an die Proline angrenzende Aminosäuren dieses Ereignis verstärken.

Da der Einfluss der benachbarten Aminosäuren umso kleiner wird, je weiter entfernt sie von den Prolinen liegen<sup>61</sup>, wurde die Betrachtung auf zwei Aminosäuren *upstream* und eine Aminosäure *downstream* der Proline begrenzt und ein EF-P abhängiges polyP Motiv als  $X_{(-2)}X_{(-1)}-nP-X_{(+1)}$  definiert (mit  $n \geq 2$ ; im Folgenden auch als polyP bezeichnet). Eine Vielzahl der Motive wurde in den oben genannten Studien bereits experimentell charakterisiert. Um das translationsarrestinduzierende Potential jedes möglichen  $X_{(-2)}X_{(-1)}-nP-X_{(+1)}$  Motivs (mit  $n \geq 2$ ) vorhersagen zu können, wurden diese Daten kombiniert und in Zusammenarbeit mit PD Dr. Jürgen Lassak (LMU München) ein Regelsatz für die Klassifizierung der Motive in „stark“, „moderat“ und „schwach“ aufgestellt (publiziert in Qi *et al.*, 2018<sup>292</sup>). Dazu wurde den zur Verfügung stehenden Datensätzen zuerst der Einfluss jeder Aminosäure X an Position (-2), (-1) oder (+1) entnommen. Prolin flankierende Aminosäuren, die das Proteintranslationslevel des betreffenden Proteins in einem *E. coli*  $\Delta efp$  Stamm um mehr als 70% im Vergleich zum Wildtyp herabsenken, wurden als „stark“ bezeichnet. Eine Verringerung um 30-70% wurde als „moderat“ und alle anderen Fälle als „schwach“ determiniert (Tabelle 11). Darauf basierend wurden Regeln formuliert, die den Gesamteinfluss jeder möglichen Kombination an konsekutiven Prolinen und flankierenden Aminosäuren vorhersagen (Tabelle 12).

Um die Qualität der Vorhersage zu verifizieren, wurde diese mit einer experimentell durch *ribosome profiling* bestimmten Rangfolge der Translationsverzögerung in zahlreichen *E. coli* Proteine mit polyP Motiv von Woolstenhulme *et al.* (2015)<sup>293</sup> verglichen. Der in dieser Studie beschriebene *asymmetry score* steht für die Häufigkeit, mit der Ribosomen während der Translation in einem Wildtyp *E. coli* Stamm im Vergleich zu einem

$\Delta\epsilon_{fp}$  Stamm an bestimmten Proteinmotiven vorzufinden sind. Wenn ein *asymmetry score* von 2 als Grenzwert für die Zuordnung eines Motivs in die Kategorie „stark“ zugewiesen wird, fallen mehr als 75% dieser Motive auch mit Hilfe der hier beschriebenen Vorhersage in diese Klasse, bei einem *asymmetry score* Grenzwert von 3 und 5 sind es bereits 80% bzw. 90% (vgl. Qi *et al.*, 2018<sup>292</sup>).

**Tabelle 11:** Charakterisierung flankierender Aminosäuren bezüglich deren Einflusses auf das arretierende Potential von konsekutiven Prolinen, basierend auf drei experimentellen Studien <sup>293, 61, 30</sup> (Klassifizierung I) und Kombination dieser Datensätze (Klassifizierung II). (Referenzen: Peil *et al.*, 2013<sup>30</sup>; Starosta *et al.*, 2014<sup>61</sup>; Woolstenhulme *et al.*, 2015<sup>293</sup>)

Klassifizierung I	X(-2)	X(-1)	PP	X(+1)	Referenz
stark	K, E, I, V, H, R, P, A	P, G, D	PP	W, N, D, P, K, E, S	Woolstenhulme <i>et al.</i> , 2015
	P, Y, E, K, Q, H, W, R	P, D, A	PP	P, W, D, N, G	Peil <i>et al.</i> , 2013; Starosta <i>et al.</i> , 2014
moderat	G, N, Q, Y, D	A	PP	A, Q, G	Woolstenhulme <i>et al.</i> , 2015
	D, V, G, I, M, F, N	G, S, E	PP	E, S, T, Q	Peil <i>et al.</i> , 2013; Starosta <i>et al.</i> , 2014
schwach	S, M, F, T, L, C	S, E, R, Q, V, I, C, W, K, N, M, H, F, L, T, Y	PP	T, V, H, R, C, I, L, M, Y, F	Woolstenhulme <i>et al.</i> , 2015
	A, L, S, T, C	K, F, V, H, Q, M, I, Y, T, L, W, R, C, N	PP	K, C, V, I, M, A, F, L, Y, R, H	Peil <i>et al.</i> , 2013; Starosta <i>et al.</i> , 2014
Klassifizierung II	X(-2)	X(-1)	PP	X(+1)	Referenz
stark	A, D, E, G, H, I, K, Q, R, V, W, Y, (P)	A, D, G, (P)	PP	D, E, G, N, (P), S, W	kombiniert
moderat	F, M, N	S, E	PP	A, K, Q, T	kombiniert
schwach	C, L, S, T	C, F, H, I, K, L, M, N, Q, R, T, V, W, Y	PP	C, F, H, I, L, M, R, V, Y	kombiniert

**Tabelle 12:** Regelsatz zur Vorhersage der „Stärke“ des durch jedes mögliche  $X_{(-2)}X_{(-1)}nPX_{(+1)}$  Motiv hervorgerufen putativen Translationsarrestes. ( $n \geq 2$ )

Nr.	Regeln zur Klassifizierung von $X_{(-2)}X_{(-1)}nPX_{(+1)}$					
1.	wenn $n \geq 3$ mit $n > 3$ , dann "stark"					
2.	wenn $n \geq 3$ mit $n = 3$ , dann gilt $X_{(-2)}P_{(-1)}PPX_{(+1)}$ und Nr. 2.1-2.3					
2.1	wenn $X_{(-2)} = \text{"stark"}$ , dann "stark"					
	$X_{(-2)}$	$X_{(-1)}$		$X_{(+1)}$	Klassifizierung	Beispiel
	stark	P	PP	stark	stark	KPPPD
	stark	P	PP	moderat	stark	KPPPK
2.2	wenn $X_{(-2)} = \text{"moderat"}$ , dann "stark", wenn $X_{(+1)} = \text{"stark"}$ , sonst "moderat"					
	$X_{(-2)}$	$X_{(-1)}$		$X_{(+1)}$	Klassifizierung	Beispiel
	moderat	P	PP	stark	stark	FPPPD
	moderat	P	PP	moderat	moderat	FPPPK
2.3	wenn $X_{(-2)} = \text{"schwach"}$ , dann $X_{(+1)}$ determinierend					
	$X_{(-2)}$	$X_{(-1)}$		$X_{(+1)}$	Klassifizierung	Beispiel
	schwach	P	PP	stark	stark	SPPPD
	schwach	P	PP	moderat	moderat	SPPPK
3.	wenn $n \geq 2$ mit $n = 2$ , dann gilt $X_{(-2)}X_{(-1)}PPX_{(+1)}$ und Nr. 3.1-3.4					
	wenn $X_{(+1)} = \text{"stark"}$ , dann betrachte $X_{(-1)}$ als $X_{(-2)}$ : $X_{(-2)}PPX_{(+1)}$ und $X_{(-2)}$ ist determinierend					
	$X_{(-2)}$	$X_{(-1)}$		$X_{(+1)}$	Klassifizierung	Beispiel
	stark	-	PP	stark	stark	KPPD
3.1	wenn $X_{(+1)} = \text{"moderat"}$ und $X_{(-2)}$ oder $X_{(-1)} \neq \text{"schwach"}$ , dann "stark", wenn $X_{(-2)}$ und $X_{(-1)} = \text{"stark"}$ , sonst "moderat"					
	$X_{(-2)}$	$X_{(-1)}$		$X_{(+1)}$	Klassifizierung	Beispiel
	moderat	-	PP	stark	moderat	FPPD
	schwach	-	PP	stark	schwach	SPPD
3.2	wenn $X_{(+1)} = \text{"moderat"}$ und $X_{(-2)}$ oder $X_{(-1)} = \text{"schwach"}$ , dann betrachte $X_{(-1)}$ als $X_{(-2)}$ : $X_{(-2)}PPX_{(+1)}$ , "moderat", wenn $X_{(-2)} = \text{"stark"}$ oder "moderat", sonst "schwach"					
	$X_{(-2)}$	$X_{(-1)}$		$X_{(+1)}$	Klassifizierung	Beispiel
	stark	stark	PP	moderat	stark	KGPPK
	stark	moderat	PP	moderat	moderat	KEPPK
3.3	wenn $X_{(+1)} = \text{"schwach"}$ , dann "stark", wenn $X_{(-2)}$ und $X_{(-1)} = \text{"stark"}$ ; "schwach", wenn $X_{(-2)}$ oder $X_{(-1)} = \text{"schwach"}$ , sonst "moderat"					
	$X_{(-2)}$	$X_{(-1)}$		$X_{(+1)}$	Klassifizierung	Beispiel
	stark	-	PP	moderat	moderat	TIPPK
	moderat	-	PP	moderat	moderat	AFPPK
3.4	wenn $X_{(+1)} = \text{"schwach"}$ , dann "stark", wenn $X_{(-2)}$ und $X_{(-1)} = \text{"stark"}$ ; "schwach", wenn $X_{(-2)}$ oder $X_{(-1)} = \text{"schwach"}$ , sonst "moderat"					
	$X_{(-2)}$	$X_{(-1)}$		$X_{(+1)}$	Klassifizierung	Beispiel
	stark	stark	PP	schwach	stark	KGPPC
	stark	moderat	PP	schwach	moderat	KEPPC
	stark	schwach	PP	schwach	schwach	KCPPC
	moderat	stark	PP	schwach	moderat	FGPPC
	moderat	moderat	PP	schwach	moderat	FEPPC
	moderat	schwach	PP	schwach	schwach	FCPPC
	schwach	stark	PP	schwach	schwach	SGPPC
	schwach	moderat	PP	schwach	schwach	SEPPC
schwach	schwach	PP	schwach	schwach	SCPPC	

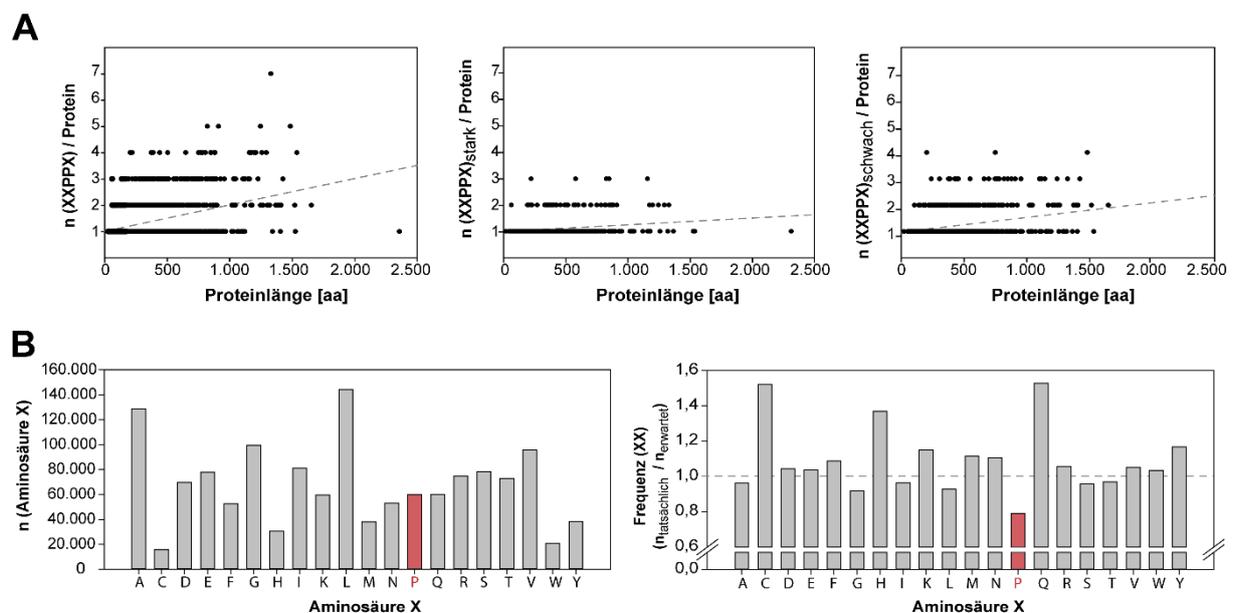
### 3.1.2 Häufigkeit von polyP Motiven in *E. coli*

Die Häufigkeit von polyP Motiven im Proteom von *E. coli* soll Aufschluss geben über die Signifikanz der Motive.

#### 3.1.2.1 Frequenz von polyP Motiven in *E. coli* K-12 MG1655

*E. coli* K-12 MG1655 zählt zu den am besten untersuchten Modellorganismen. Das Proteom ist vollständig entschlüsselt und in der Datenbank UniProt als Referenzproteom vermerkt (UniProt Proteome ID: ECOLI - UP000000625), weshalb dieser Datensatz ausgewählt wurde, um exemplarisch das Vorkommen von polyP Motiven zu untersuchen. Das Proteom von *E. coli* K-12 MG1655 zählt 4.313 Proteine mit insgesamt 2.087  $X_{(-2)}X_{(-1)}-nP-X_{(+1)}$  Motiven. Tabelle S1 zeigt eine vollständige Aufstellung dieser polyP Motive. 478 (22,9%) der polyP Motive sind als „stark“, 446 (21,4%) als „moderat“ und 1.163 (55,7%) als „schwach“ definiert.

Nicht alle denkbaren polyP Varianten treten gleich oft im Proteom von *E. coli* K-12 MG1655 auf. Vielmehr weicht die tatsächliche Verteilung der individuellen  $X_{(-2)}X_{(-1)}-nP-X_{(+1)}$  Motive innerhalb einer Definitionsklasse signifikant von der Normalverteilung ab (Shapiro-Wilk-Test, p-Werte  $< 1e-4$ ). Per Definition fallen 29,4%, 24,9% und 45,7% der theoretisch möglichen  $X_{(-2)}X_{(-1)}-PP-X_{(+1)}$  Kombinationen in die Klassifizierung „stark“, „moderat“ und „schwach“. Der Anteil der tatsächlich auftretenden „starken“ und „moderaten“ an der Gesamtzahl der polyP Motive in *E. coli* K-12 MG1655 ist damit deutlich kleiner.



**Abb. 6:** Auftreten von polyP Motiven im Proteom von *E. coli* K12 MG1655. (A) Anzahl an polyP Motiven je Protein relativ zur Sequenzlänge. (B) Anzahl jeder Aminosäure X und deren Frequenz ( $n_{\text{tatsächlich}}/n_{\text{erwartet}}$ ) als repetitives Element.

Unter Berücksichtigung der 19 betroffenen Phagenproteine (davon sieben Kopien der Transposase InsD: InsD1-6 und InsD8), sind die  $X_{(-2)}X_{(-1)}\text{-nP-X}_{(+1)}$  Motive auf 1.458 Proteine verteilt, womit 33,8% aller *E. coli* K-12 MG1655 Proteine mindestens ein polyP Motiv aufweisen. 1.003 (68,8%) dieser Proteine tragen ein polyP Motiv, während 318 (21,8%) und 137 (9,4%) Proteine zwei bzw. mehr als zwei Motive aufweisen. Die Anzahl von polyP Motiven je Protein ist direkt proportional zur Länge der Aminosäuresequenz (Abb. 6A), wobei eine maximale Anzahl von 7 Motiven in der DNA Translokase FtsK vorzufinden ist. „Starke“ Motive verteilen sich auf 425 Proteine (9,8%), während „moderate“ und „schwache“ Motive in 409 (9,5%) bzw. 952 (22,1%) Proteinen zu finden sind.

Die Frequenz von zwei aufeinanderfolgenden Prolinen ist deutlich geringer als bei einer zufälligen Verteilung der Aminosäuren in den Proteinen zu erwarten wäre (Frequenz =  $n_{\text{tatsächlich}}/n_{\text{erwartet}} = 0,796$ ; Berechnung basierend auf der Wahrscheinlichkeit des Auftretens von Prolin an x-beliebiger Position, Abb. 6B). Im Vergleich dazu kommen Repetitionen der übrigen Aminosäuren häufiger als erwartet oder nur sehr weniger selten vor (Minimalwert Glycin: 0,923).

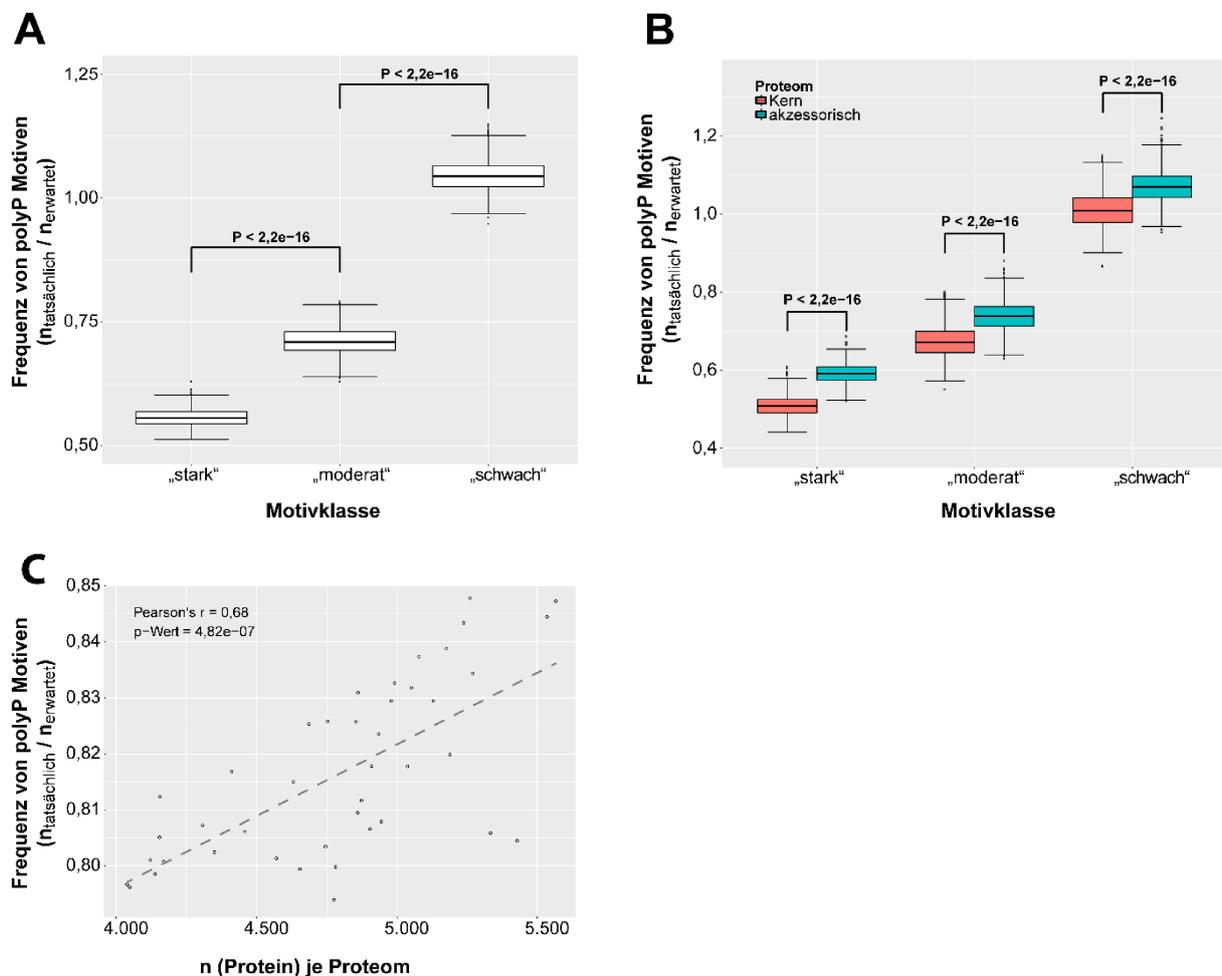
### 3.1.2.2 Frequenz von polyP Motiven in 43 ausgewählten *E. coli* Stämmen

Um einen repräsentativen Datensatz zu generieren, aus dem sich allgemein gültige Vorhersagen ableiten lassen, wurden die Proteome von 43 *E. coli* Stämmen ausgewählt (Kapitel 2.3, Tab. 10) und hinsichtlich der darin vorkommenden polyP Motive analysiert (vgl. Fei *et al.*, 2018<sup>292</sup>). Dabei wurden 11.356 orthologe Gruppen mit insgesamt 195.056 Proteinen erfasst. Zu den *E. coli* Stämmen zählen sowohl Laborstämme, als auch Wildtyp-Isolate, um natürliche Selektionsbedingungen zu berücksichtigen. Die Stämme wurden nach Clermont *et al.* (2013)<sup>288</sup> in phylogenetische Gruppen klassifiziert. Dafür wurden die Genome der Stämme von NCBI Datenbank heruntergeladen und ein *in silico* Annealing der empfohlenen Primer-Sequenzen zur Amplifizierung der Markergene *arpaA*, *chuA*, *yjaA* und *TspE4* durchgeführt. Die ausgewählten Stämme repräsentieren 6 der 7 phylogenetischen *E. coli* Gruppen.

Die Proteome der Stämme zählen 99.386 polyP Motive in 68.710 Proteinen (33,3% aller Proteine). Davon tragen 47.056 (68,5%) ein Motiv, während 15.027 (21,9%) und 6.627 (9,6%) der Proteine zwei bzw. mehr als zwei Motive tragen. 22.253 (22,4%), 21.953 (22,1%) und 55.146 (55,5%) der Motive sind als „stark“, „moderat“ und „schwach“ definiert. Die Verteilung entspricht damit den oben beschriebenen Beobachtungen in *E. coli* K-12 MG1655. Im Vergleich zu einer zufällig generierten Verteilung der Aminosäuren in den Proteinsequenzen sind die polyP Motive unterrepräsentiert. Zwei aufeinanderfolgende Proline weisen in allen untersuchten Stämmen eine signifikant niedrigere Frequenz im Vergleich zum auf 1.000 zufällig generierten Sequenzen basierenden Mittelwert auf (durchschnittliche PP Frequenz der 43 *E. coli* Stämme 0,82; p-Werte < 2,2e-16). Dabei zählt *E. coli* O157:H7 die meisten polyP Motive (0,85) und *E. coli* UM146 die wenigsten Motive (0,79).

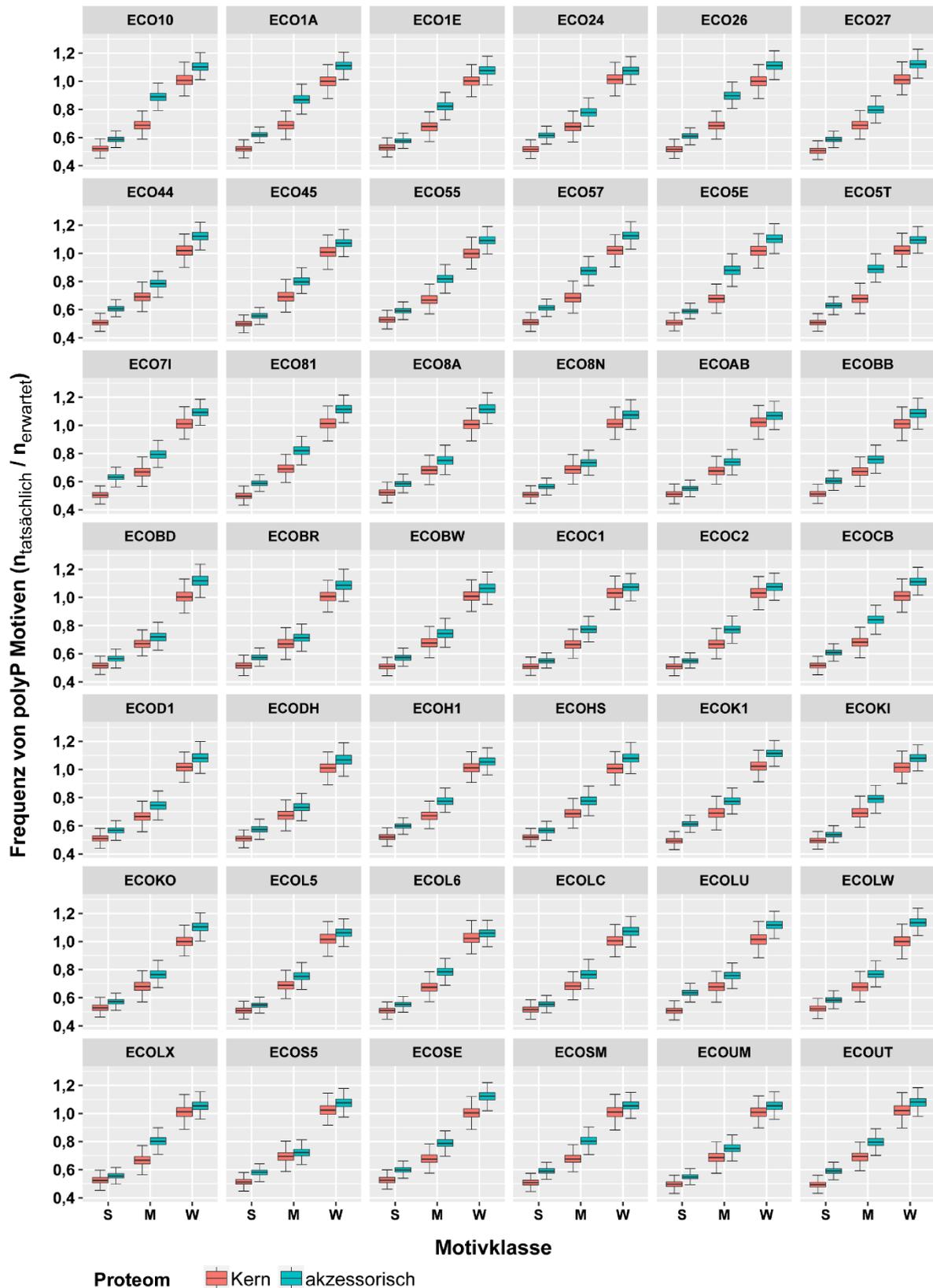
Wird die Klassifizierung der polyP Motive mit in Betracht gezogen, zeigt sich, dass deren Frequenz negativ mit der Motivstärke korreliert. Die Frequenzen von „starken“, „moderaten“ und „schwachen“ Motiven betragen in *E. coli* K-12 MG1655 das 0,55-fache, 0,71-fache und 1,04-fache der erwarteten Frequenz (Abb. 7A). Ein ähnliches Muster ist auch bei den restlichen 42 *E. coli* Stämmen zu erkennen. Damit erhärtet sich die Vermutung eines Selektionsdrucks gegen polyP Motive in Relation zu deren Potential, einen Translationsarrest zu induzieren.

Ein weiterer Punkt, der für diese These spricht, ist die unterschiedliche Häufigkeit von polyP Motiven in Kernproteom und akzessorischem Proteom. Ausgehend von den 43 *E. coli* Stämmen konnten insgesamt 73.745 Proteine dem Kernproteom und 132.615 Proteine dem akzessorischen Proteom zugeordnet werden. Im evolutionsgeschichtlich älteren Kerngenom finden sich signifikant weniger Gene, die für Proteine mit polyP Motiv kodieren, als im akzessorischen Genom (Abb. 7B und 8; Mann-Whitney-Wilcoxon Test, p-Werte  $< 2,2e-16$ ). Dies gilt unabhängig von der „Stärke“ der Motive und es kann nicht völlig ausgeschlossen werden, dass diese Beobachtung auf unterschiedlichen Funktionen und Expressionslevel der kodierten Proteine beruht (Mittelwert  $\log_{10}$ : Translationseffizienz = 1,68 vs. 1,57; Proteinabundanz = 2,05 vs. 1,86). Die Frequenz der polyP Motive des Kernproteoms ist für alle untersuchten Stämme sehr ähnlich, während im akzessorischen Proteom eine größere Diversität festgestellt wurde. Zudem korreliert die Frequenz von polyP Motiven mit der Proteomgröße (Abb. 7C, Pearson's  $r = 0,68$ , p-Wert =  $4,82e-7$ ).



**Abb. 7:** Häufigkeit von polyP Motiven in 43 repräsentativen *E. coli* Stämmen. (A) Frequenz von polyP Motiven in Abhängigkeit von deren Klassifizierung in *E. coli* K-12 MG1655. (B) Häufigkeit von polyP Motiven in Kern- und akzessorischem Proteom von 43 ausgewählten *E. coli* Stämmen. (C) Korrelation zwischen der Häufigkeit von polyP Motiven und der Proteomgröße von 43 ausgewählten *E. coli* Stämmen.

(A und B: Dr. Fei Qi, modifiziert aus Qi *et al.*, 2018<sup>292</sup>; C: Dr. Fei Qi)



**Abb. 8:** Häufigkeit von polyP Motiven in Kern- und akzessorischem Proteom von 43 repräsentativen *E. coli* Stämmen – Einzeldarstellung. (vgl. Tab. 10; Abbildung modifiziert aus Qi *et al.*, 2018<sup>292</sup>)

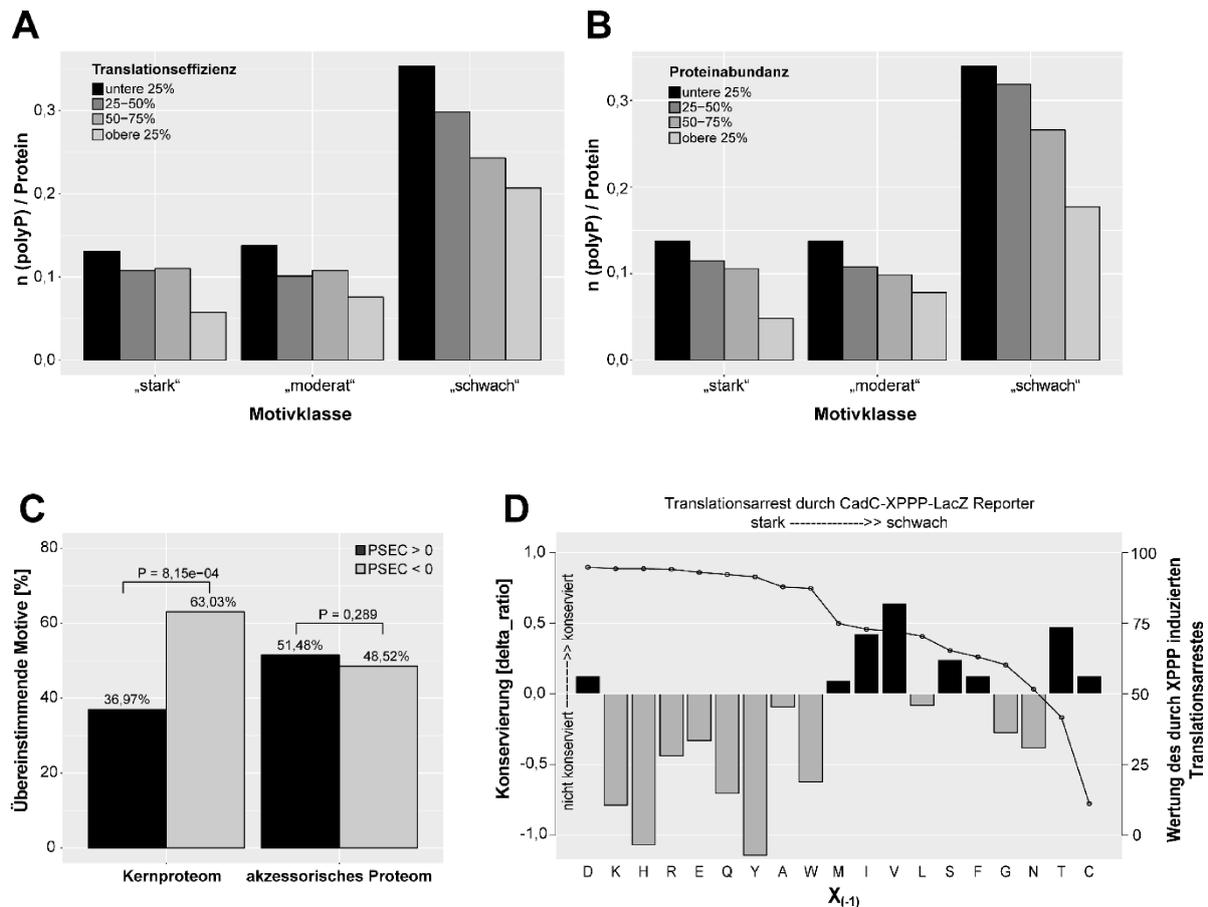
### 3.1.3 Antikorrelation zwischen dem Auftreten von polyP Motiven und der Translationseffizienz

Proteine mit polyP Motiv aus *E. coli* K-12 MG1655 (OMA Datenbank, Proteome ID: ECOLI) weisen eine signifikant niedrigere Translationseffizienz auf, als Proteine ohne Motiv (vgl. Qi *et al.*, 2018<sup>292</sup>; Abb. 9A, Spearman's rho = -0,105; p-Wert = 1,13e-5). Je höher die Translationseffizienz eines Proteins, desto unwahrscheinlicher ist das Auftreten von polyP Motiven. Diese Beobachtung wird besonders deutlich für Proteine mit „starkem“ polyP Motiv. Die Präsenz von polyP Motiven antikorreliert mit zunehmender „Stärke“ außerdem mit der relativen Proteinabundanz (vgl. Qi *et al.*, 2018<sup>292</sup>; Abb. 9B, Spearman's rho = -0,135, p-Wert = 1,47e-8). Dies bedeutet einerseits, dass polyP Motive im Laufe der Evolution zugunsten einer hohen Proteinkopienzahl vermieden wurden, was einen durch die Translationseffizienz begründeten Selektionsdruck gegen polyP Motive impliziert und die niedrige Frequenz erklären würde. Andererseits könnten polyP Motive eine Möglichkeit darstellen, die Kopienzahl einiger Proteine gezielt herabzusetzen, was zum Beispiel für die Aufrechterhaltung eines numerischen Gleichgewichts zwischen wechselwirkenden Komponenten einer Signaltransduktionskaskade von Vorteil wäre. Ebenso denkbar ist, dass die Vorteile, die mit einer bestimmten durch polyP Motive vermittelten strukturellen Konformation oder Funktion einhergehen, die Nachteile hinsichtlich der Translationseffizienz überwiegen und polyP Motive bis zu einem gewissen Grad toleriert werden.

### 3.1.4 Evolution von polyP Motiven

Die Konservierung von polyP Motiven wurde durch ein Sequenzalignment aller Proteine der ausgewählten 43 *E. coli* Stämme mit mindestens einem  $X_{(-2)}X_{(-1)}nPX_{(+1)}$  Motiv analysiert (vgl. Qi *et al.*, 2018<sup>292</sup>). Die Proteine bilden 3.280 orthologe, aus mindestens drei Proteinen bestehende Gruppen. Darin enthalten sind insgesamt 4.980 Sequenzabschnitte mit polyP Motiv, wovon 1.568 eine uneinheitliche Motivstärke aufweisen. Im Detail handelt es sich um 1.923 Einzelereignisse, wovon 955 aus einer Änderung in eine „stärkere“ Klasse resultieren (*gain event*) und 968, bei denen es sich um eine „Abschwächung“ handelt (*loss event*). Daraus wurde ein Index berechnet, der die Neigung eines Motivs sich im Laufe der Evolution zu verändern beschreibt (PSEC, *propensity for stalling effect change*). Im Kerngenom verlieren signifikant mehr polyP Motive an „Stärke“, als dass ein Zugewinn zu verzeichnen ist (Abb. 9C; 63,03% vs. 36,97%; Chi-Quadrat-Test, p-Wert = 8,15e-4). Im Gegensatz dazu treten *gain* und *loss events* im akzessorischen Proteom ähnlich häufig auf (Abb. 9C; 48,52% vs. 51,48%).

Ergänzend wurde in einer zweiten Analyse die Konservierung von  $X_{(-1)}PPP$  Motiven untersucht, deren Auswirkung auf das Translationslevel des betreffenden Proteins bereits experimentell von Starosta *et al.* (2014)<sup>61</sup> unter Verwendung eines CadC-LacZ Reporterkonstruktes untersucht wurde. Darauf basierend wurde ein *score* bestimmt, der das Verhältnis der  $\beta$ -Galactosidaseaktivität zwischen *E. coli* Wildtypstamm und  $\Delta efp$  Stamm für jedes der 20 Motive beschreibt und eine Sortierung bezüglich der „Stärke“ der Motive ermöglicht. Für die bioinformatische Analyse der Konservierung dieser Motive wurden die in der OMA Datenbank hinterlegten *E. coli* Proteome verwendet. Mit Hilfe eines Sequenzalignments der zu orthologen Gruppen zusammengefassten Proteine wurde die Konservierung der polyP Motive bestimmt. Dabei zeigte sich eine Korrelation zwischen der experimentell bestimmten „Stärke“ der XPPP Motive und der Konservierung. Je „stärker“ ein Motiv, desto weniger konserviert ist es (Abb. 9D).



**Abb. 9:** Zusammenhang zwischen dem Auftreten von polyP Motiven und der Translationseffizienz. (A) Anzahl von polyP Motiven in Abhängigkeit zur Translationseffizienz in *E. coli* K-12 MG1655. (B) Anzahl von polyP Motiven in Abhängigkeit zur Proteinabundanz *E. coli* K-12 MG1655. (C) Vergleich des PSEC (*propensity of stalling effect change*) zwischen Kern und akzessorischen Proteom von 43 repräsentativen *E. coli* Stämmen. Ein positiver PSEC beschreibt die Neigung eines Motivs, im Laufe der Evolution durch Aminosäuresubstitutionen an „Stärke“ hinzugewinnen, ein negativer PSEC die Neigung zur Abschwächung. (D) Korrelation zwischen Konservierung eines  $X_{(-1)}$ PPP Motivs und experimentell bestimmter Wertung des durch das selbe Motiv induzierten Translationsarrestes.

(A, B und C: modifiziert aus Qi *et al.*, 2018<sup>292</sup>; D: Dr. Fei Qi)

### 3.1.5 Ontologie von Proteinen mit polyP Motiv

Da die Translationseffizienz durch den Translations-Elongationsfaktor EF-P maßgeblich beeinflusst wird, stellt sich die Frage, ob polyP Motive als regulatorische Elemente fungieren, um die Expressionslevel bestimmter Proteingruppen in Abhängigkeit von EF-P spezifisch an bestimmte Umweltbedingungen anzupassen. Eine derartige Funktion könnte durch ein Muster in der Verteilung von polyP Motiven ersichtlich werden. Proteine aus *E. coli* K-12 MG1655 (UniProt ID: ECOLI - UP000000625) wurden auf eine Akkumulation von „starken“ polyP Motiven in bestimmten Proteinfamilien und Stoffwechsel- bzw. Signaltransduktionswegen untersucht. Darüberhinaus wurde die intrazelluläre Lokalisation der betreffenden Proteine berücksichtigt. Die Motive sind nicht ausschließlich auf Membranproteine oder cytosolische Proteine beschränkt, sondern treten in beiden Gruppen ähnlich häufig auf (38% und 44%). Damit sind polyP Motive überdurchschnittlich oft Teil von Membranproteinen, welche insgesamt nur etwa 25% aller Proteine des Gesamtproteoms von *E. coli* K-12

ausmachen<sup>312</sup>. Die restlichen Proteine mit starken polyP Motiven sind zu jeweils geringen Anteilen in der äußeren Membran (7%), dem Periplasma (2%), in der Zellwand (1%) oder extrazellulär (1%) lokalisiert. Daneben sind 3% der Proteine bakteriophagischen Ursprungs oder können nicht zugeordnet werden (11%).

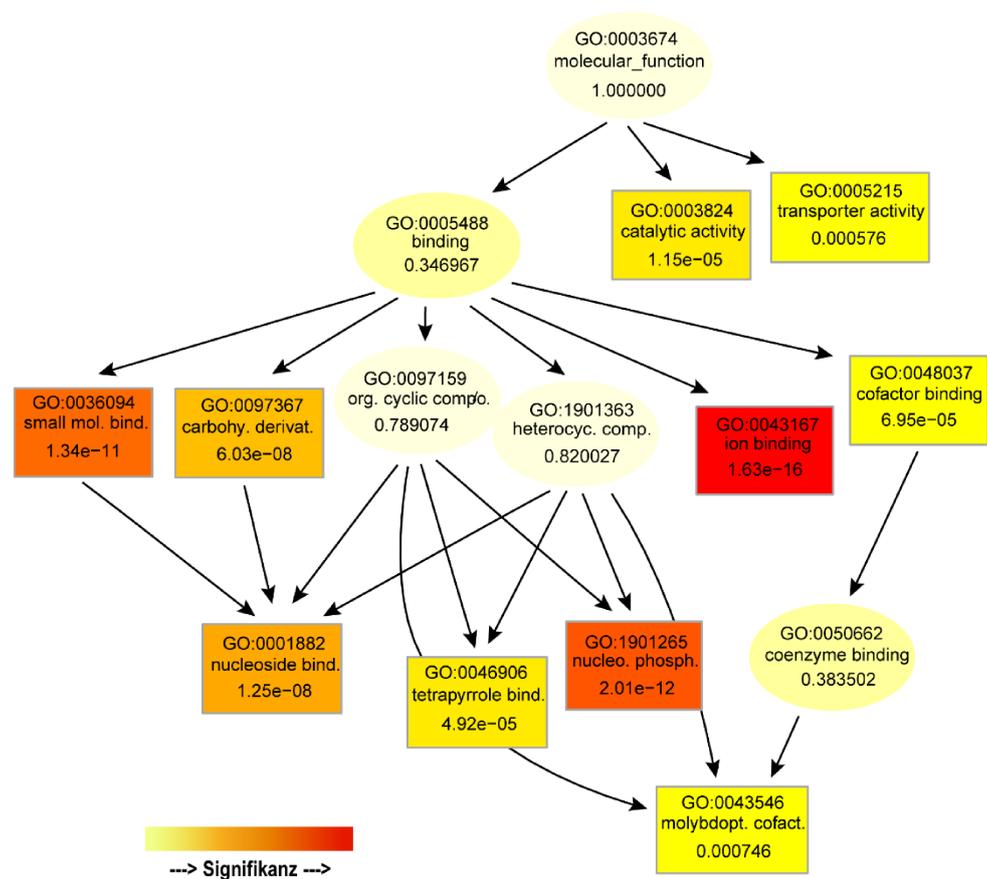
Bei Betrachtung der von UniProt zugewiesenen *gene ontologies* wird deutlich, dass die 425 Proteine eine große Bandbreite hinsichtlich molekularer Funktion und der Zuordnung in Enzymklassen abdecken (Tab. 13). Am häufigsten vertreten sind die Klasse der Hydrolasen und ionenbindenden Proteine.

**Tabelle 13:** Ontologien der 425 Proteine aus *E. coli* K-12 MG1655 mit „starkem“ polyP Motiv. (Mehrfachnennungen)

<b>Molekulare Funktion</b>	<b>Hits</b>	<b>Stoffwechselweg</b>	<b>Hits</b>
<b>Catalytic activity</b>		<i>Amino-acid biosynthesis</i>	8
<i>Hydrolase activity</i>	103	<i>Cell wall/peptidoglycan biosynthesis</i>	8
<i>Transferase activity</i>	77	<i>Glycan metabolism</i>	6
<i>Oxidoreductase activity</i>	41	<i>Carbohydrate metabolism</i>	5
<i>Catalytic activity, acting on a protein</i>	36	<i>Bacterial outer membrane biogenesis</i>	4
<i>Catalytic activity, acting on RNA</i>	24	<i>Amine and polyamine biosynthesis</i>	3
<i>Lyase activity</i>	22	<i>Amine and polyamine degradation</i>	3
<i>Catalytic activity, acting on DNA</i>	21	<i>Aromatic compound metabolism</i>	3
<i>Ligase activity</i>	12	<i>Carbohydrate acid metabolism</i>	3
<i>Isomerase activity</i>	8	<i>Carbohydrate degradation</i>	3
<i>Peptidoglycan murelytic activity</i>	3	<i>Purine metabolism</i>	3
<i>Glycogen debranching enzyme activity</i>	2	<i>Sulfur metabolism</i>	3
<i>Adenylate cyclase activity</i>	1	<i>Carnitine metabolism</i>	2
<b>Peroxidase activity</b>		<i>Amino-acid degradation</i>	2
<i>Catalase activity</i>	1	<i>Cofactor biosynthesis</i>	2
<i>Cytochrome-c peroxidase activity</i>	1	<i>Glycan degradation</i>	2
<b>Signal transducer activity</b>		<i>Lipid metabolism; fatty acid beta-oxidation</i>	2
<i>Histidine phosphotransfer kinase activity</i>	8	<i>Metabolic intermediate biosynthesis</i>	2
<i>Molecular transducer activity</i>	5	<i>Porphyrin-containing compound metabolism</i>	2
<b>Transporter activity</b>		<i>Carbohydrate biosynthesis; gluconeogenesis</i>	1
<i>Transmembrane transporter activity</i>	61	<i>Energy metabolism; nitrogen metabolism</i>	1
<i>Protein transporter activity</i>	3	<i>Glycan biosynthesis; trehalose biosynthesis</i>	1
<b>Binding</b>		<i>Isoprenoid biosynthesis</i>	1
<i>Ion binding</i>	161	<i>Metabolic intermediate (oxalate) degradation</i>	1
<i>Heterocyclic compound binding</i>	156	<i>Organic acid metabolism; propanoate degrad.</i>	1
<i>Organic cyclic compound binding</i>	147	<i>Pyrimidine metabolism</i>	1
<i>Small molecule binding</i>	100	<i>Quinol/quinone metabolism</i>	1
<i>Carbohydrate derivative binding</i>	78	<i>Siderophore (enterobactin) biosynthesis</i>	1
<i>Drug binding</i>	75	<i>Extracellular polysaccharide biosynthesis</i>	1
<i>Protein binding</i>	45		
<i>Cofactor binding</i>	44		
<i>Iron-sulfur cluster binding</i>	19	<b>Enzymklasse</b>	<b>Hits</b>
<i>Sulfur compound binding</i>	9	<i>Hydrolases</i>	69
<i>Amide binding</i>	6	<i>Transferases</i>	68
<i>Ribonucleoprotein complex binding</i>	3	<i>Oxidoreductases</i>	31
<i>Lipid binding</i>	2	<i>Lyases</i>	22
<i>Carbohydrate binding</i>	2	<i>Ligases</i>	8
<i>Toxic substance binding</i>	1	<i>Isomerases</i>	7
<i>Polyamine binding</i>	1		
<i>Phosphopantetheine binding</i>	1		
<i>Virion binding</i>	1		
<i>Carbon monoxide binding</i>	1		
<b>Enzyme regulator activity</b>			
<i>Endopeptidase inhibitor activity</i>	1		
<i>DNA polymerase processivity factor activity</i>	1		
<i>Peptidase regulator activity</i>	1		
<b>Transcription regulator activity</b>			
<i>DNA binding transcription factor activity</i>	13		
<i>Bacterial sigma factor activity</i>	2		

Die Proteine sind Teil unterschiedlicher Stoffwechselwege, was darauf hinweist, dass der Grund für das Auftreten der „starken“ polyP Motive primär nicht in der Regulation spezifischer zellulärer Prozesse liegt. Dennoch handelt es sich nicht um eine Zufallsverteilung, wie eine Analyse der direkten und indirekten Interaktion der Proteine durch die Datenbank STRING ergab. Das errechnete Proteinnetzwerk hat 1477 Verknüpfungen statt 1345, wie erwartet. Damit besteht zumindest teilweise eine biologische Verbindung zwischen den Proteinen (*PPI enrichment p-value*: 0,000531).

Die Analyse der molekularen Funktion von Proteinen mit polyP Motiv wurde auf die Proteome der 43 repräsentativen *E. coli* Stämme ausgeweitet, um auch Proteine spezifischer ökologischer Nischen oder Lebensweisen, insbesondere pathogener Spezies, in Betracht zu ziehen (Abb. 10, Analyse durchgeführt von Dr. Fei Qi). Dabei wurden speziell diese Cluster von orthologen Gruppen untersucht, in welchen die  $X_{(-2)}X_{(-1)}-nP-X_{(+1)}$  Motive konserviert sind. Auch hierbei zeigte sich, dass die Motive häufig in Proteinen vorkommen, die substratbindende Eigenschaften haben.



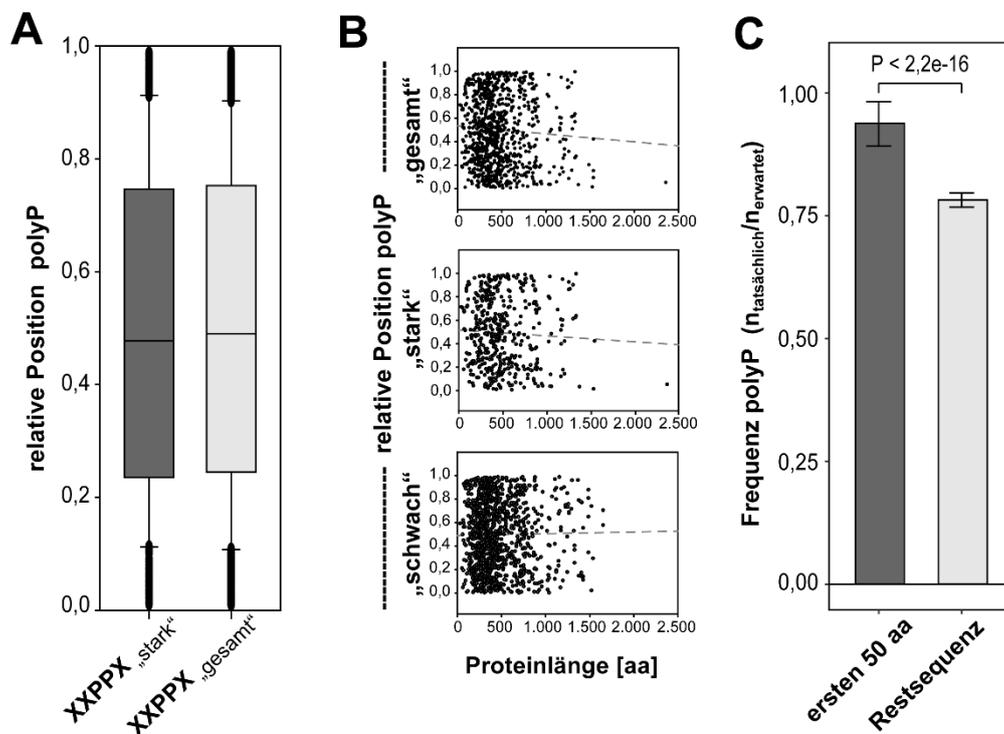
**Abb. 10:** Ontologien von Proteinen aus 43 *E. coli* Stämmen mit konserviertem polyP Motiv der Klassifizierung „stark“. (Dr. Fei Qi)

### 3.1.6 Akkumulation von polyP Motiven an strategisch günstigen intramolekularen Loci

#### 3.1.6.1 Gehäuftes Auftreten von polyP Motiven innerhalb der ersten 50 Aminosäuren eines Proteins

Unabhängig von der Gesamtlänge einer Proteinsequenz handelt es sich bei der relativen intramolekularen Positionierung der polyP Motive in *E. coli* K-12 MG1655 (UniProt ID: ECOLI - UP000000625) um eine Zufallsverteilung. Dasselbe gilt für separat untersuchte „starke“ Motive, die aufgrund ihres Einflusses auf die Translationsgeschwindigkeit mit höherer Wahrscheinlichkeit an regulatorischen kotranslationalen Prozessen beteiligt sind (Abb. 11A und B; Chi-Quadrat-Test, p-Wert < 0,005).

Eine reduzierte Translationseffizienz innerhalb der ersten 50 Aminosäuren eines Proteins wird von Tuller *et al.* (2010)<sup>16</sup> als späte Phase der Translationsinitiation beschrieben und ist durch eine Kombination aus mehreren Faktoren, wie sekundäre mRNA Strukturen, positiv geladene Aminosäuren und seltene Codons begründet. Es ist denkbar, dass auch Proline, welche einen ähnlichen Effekt auf die Translationsgeschwindigkeit haben, daran beteiligt sind. Eine Analyse der Proteine mit polyP Motiv aus *E. coli* K-12 MG1655 (OMA Datenbank, Proteome ID: ECOLI) zeigte nach Normalisierung eine Anhäufung innerhalb der ersten 50 Aminosäuren (Abb. 11C, Mann-Whitney-Wilcoxon Test, p-Wert < 2,2e-16, vgl. Qi *et al.*, 2018<sup>292</sup>). Von den 2.115 untersuchten polyP Motiven sind 325 innerhalb der ersten 50 Aminosäuren lokalisiert und 1.771 in der restlichen Proteinsequenz.



**Abb. 11:** Intramolekulare Verteilung von polyP Motiven in *E. coli* K-12 MG1655. (A) Relative Position zum N-Terminus mit normalisierter Proteinlänge. Darstellung als Boxplot. (B) Relative intramolekulare Verteilung von polyP Motiven in Relation zur Sequenzlänge. (C) Akkumulation von polyP Motiven innerhalb der ersten 50 Aminosäuren einer Proteinsequenz. (C: modifiziert aus Qi *et al.*, 2018<sup>292</sup>)

### 3.1.6.2 Akkumulation von polyP Motiven zwischen Domänen

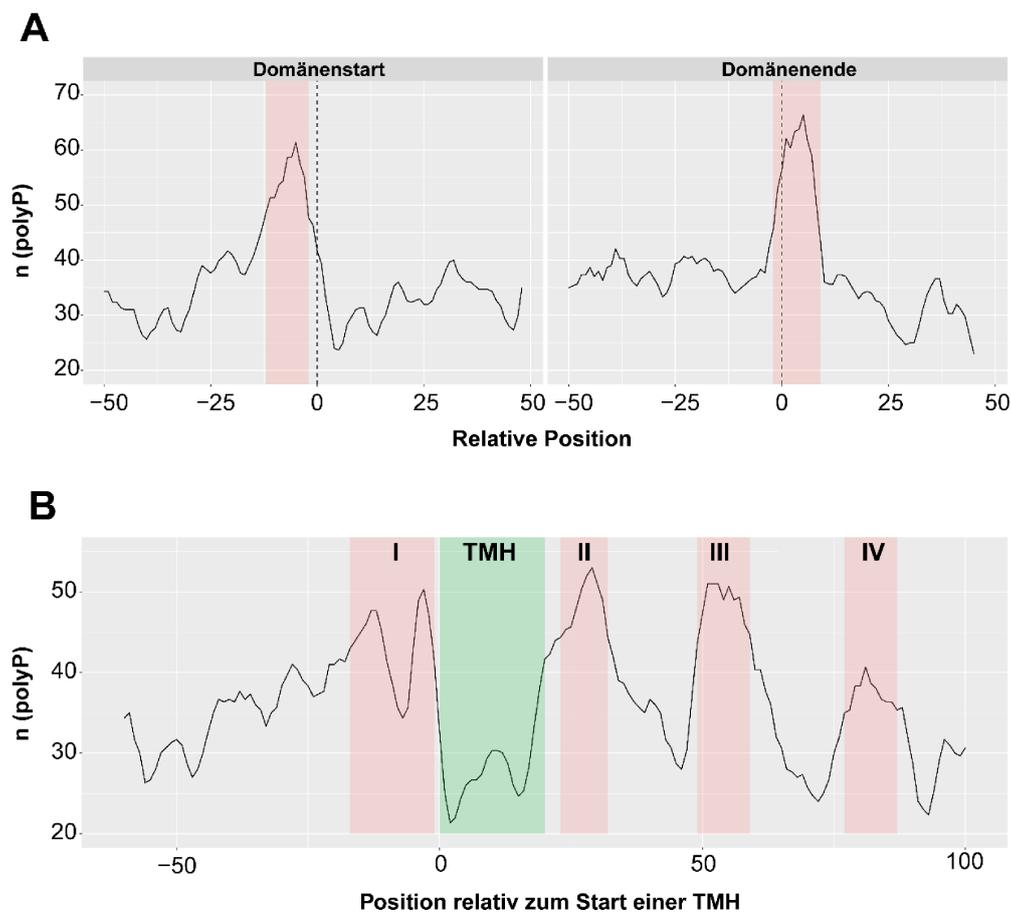
Da es sich bei der kotranslationalen Faltung von Proteinen um einen fein abgestimmten Vorgang handelt, in welchem sich bereits bestehende Strukturen auf die Faltung neu synthetisierter Abschnitte auswirken, ist eine zeitliche Abstimmung des Translationsprozesses unabdinglich<sup>15, 313, 314</sup>. Unter anderem dient der durch seltene Codons herbeigeführte Zeitgewinn nachweislich der kotranslationalen Faltung bereits synthetisierter Domänen, während eine hohe Translationsrate in strukturellen Regionen Fehlfaltungen vorbeugt<sup>315, 316, 314, 317</sup>. Da auch polyP Motive einen Einfluss auf den zeitlichen Ablauf der Translation haben, wurde deren Position in Relation zu Proteindomänen untersucht (vgl. Qi *et al.*, 2018<sup>292</sup>). In der *Gene 3D* Datenbank sind 7.398 strukturelle Domänen und deren Positionen aus 4.080 *E. coli* K-12 MG1655 Proteinen hinterlegt. Darunter sind 1.868 Multidomänenproteine. Interdomänenlinker wurden hier als Sequenzabschnitte zwischen zwei Domänen definiert, die mindestens 5 Aminosäuren lang sind, was der Länge eines  $X_{(-2)}X_{(-1)}\text{-PP-}X_{(+1)}$  Motivs entspricht. Im Fall einer kürzeren Sequenz wurde diese für die folgende Analyse *downstream* verlängert. Im Zuge dessen wurden 3.318 Linker zwischen 5.186 Domänen identifiziert. PolyP Motive kommen signifikant seltener in strukturellen Domänen, als in Linkerregionen vor (p-Wert = 7,86e-80; Frequenz 0,56 vs. p-Wert = 0,912; Frequenz 1,10). Dabei akkumulieren die Motive -12 bis -2 Aminosäuren relativ zum Domänenstart und -2 bis +9 Aminosäuren relativ zum Domänenende (Abb. 12A; p-Werte < 0.05; Frequenzen von 1,19 and 1,23). Eine ähnliche Verteilung wurde für langsam translatierte Codons gezeigt<sup>317, 318, 314, 319</sup>, was vermuten lässt, dass polyP Motive eine vergleichbare Funktion haben.

### 3.1.6.3 Akkumulation von polyP Motiven in der Nähe von Transmembranhelices

Eine zeitliche Koordination der Translation ist auch für die kotranslationale Integration von  $\alpha$ -helikalen Membranproteinen durch Translokons von essentieller Bedeutung<sup>320, 321</sup>. Fluman *et al.*<sup>22</sup> zeigten 2014, dass zwei Shine-Dalgarno-ähnliche Elemente in der *E. coli* mRNA Translationsstopps induzieren, die maßgeblich zur zielgerichteten Translokation von Transmembranproteinen durch *signal recognition particles* (SRP) beitragen. Die erste Translationsverzögerung tritt dabei kurz bevor das neusynthetisierte Peptid aus dem ribosomalen *exit tunnel* hervortritt ein. Nach dem Erscheinen der ersten Transmembranhelix (-5 bis -1 Codons relativ zur nachfolgenden Transmembranhelix) kommt es zum zweiten Stopp. Im Fungus *Emericella nidulans* wurden durch Dessen *et al.* (2000)<sup>322</sup> zwei Translationsarreste im Abstand von 45 und 70 Codons von der Transmembrandomäne nachgewiesen.

Analog dazu wurde hier analysiert, ob auch polyP Motive an diesem Prozess beteiligt sind (vgl. Qi *et al.*, 2018<sup>292</sup>). In der Datenbank *UniProt* sind für *E. coli* K-12 MG1655 912 Transmembranproteine mit insgesamt 5.672 Transmembranhelices annotiert. 39,3% (358) der Proteine tragen polyP Motive. Im Vergleich dazu finden sich signifikant weniger polyP Motive in löslichen Proteinen (32,6%, Qhi-Quadrat-Test, p-Wert = 1,6e-4). Es konnte keine Akkumulation in den von Fluman *et al.*<sup>22</sup> beschriebenen Regionen nachgewiesen werden. Stattdessen wurde festgestellt, dass polyP Motive einerseits wie erwartet nur sehr selten innerhalb von Transmembranhelices vorkommen (signifikant, p-Wert = 1,65e-27, Frequenz = 0,39), andererseits aber in vier distinkten Regionen gehäuft auftreten. Dabei handelt es sich um die vier Sequenzabschnitte von -17 bis -1, 23 bis 32, 49 bis 59 und 77 bis 87 Aminosäuren relativ zum Start der Transmembranhelix (Abb. 12B; Position I-IV).

Unter Ausschluss derjenigen Sequenzabschnitte, die innerhalb einer benachbarten Helix liegen, wurde eine signifikante Anreicherung von polyP Motiven in Position III festgestellt (p-Wert = 0,035; Frequenz = 1,39). Diese Region stimmt mit dem von Dessen *et al.*<sup>322</sup> beschriebenen Translationsstopp überein. Der Translationsstopp tritt in dem Moment ein, in dem die Transmembranhelix, welche durchschnittlich 21 Aminosäuren lang ist, aus dem ribosomalen *exit tunnel*, welcher eine Kapazität für 28 Aminosäuren hat<sup>323</sup>, austritt. Bei dieser Gelegenheit könnte die Helix durch das Translokation in die Membran integriert werden.



**Abb. 12:** Auftreten von polyP Motiven in *E. coli* K-12 MG1655 relativ zu Proteindomänen und Transmembranhelices. (A) PolyP Motive akkumulieren 12 bis -2 Aminosäuren relativ zum Domänenstart und -2 bis +9 Aminosäuren relativ zum Domänenende. Die Domänengrenzen sind als gestrichelte Linie dargestellt, während Regionen mit relativ hohem polyP Vorkommen hellrot hinterlegt sind. (B) Die Frequenz von polyP Motiven relativ zur Startposition von Transmembranhelices ist in vier Regionen (I, II, III und IV) erhöht und hellrot hinterlegt. Die Transmembranhelix ist unter der Annahme einer Sequenzlänge von 21 Aminosäuren grün hinterlegt. Zur besseren Visualisierung wurden die Werte über je drei Datenpunkte geglättet. (modifiziert aus: Qi et al., 2018<sup>292</sup>)

### 3.2 Die Rolle von polyP Motiven für Signalproteine

Proteine, die Teil eines komplexen Signaltransduktionsnetzwerkes sind, müssen sowohl in ihrer Kopienzahl, als auch in ihrer Aktivität genau aufeinander abgestimmt sein, um eine effiziente Weiterleitung eines spezifischen Signals ohne Informationsverlust zu gewährleisten<sup>324, 325</sup>. Unter diesem Gesichtspunkt wurde untersucht, ob polyP Motive, die einen starken Translationsarrest induzieren, eine Rolle für die Kopienzahl von bestimmten Signalproteinen spielen und inwiefern sie an der proteintypischen Funktionalität beteiligt sind, bzw. eine strukturelle Bedeutung haben.

Dabei wurden exemplarisch die putativ stark EF-P abhängigen polyP Motive aus vier *E. coli* K-12 MG155 Proteinen betrachtet: der Diguanylatcyclase DgcC, der ATP-abhängigen, toxindegradierenden Serinprotease Lon und den beiden Histidinkinasen EvgS und EnvZ.

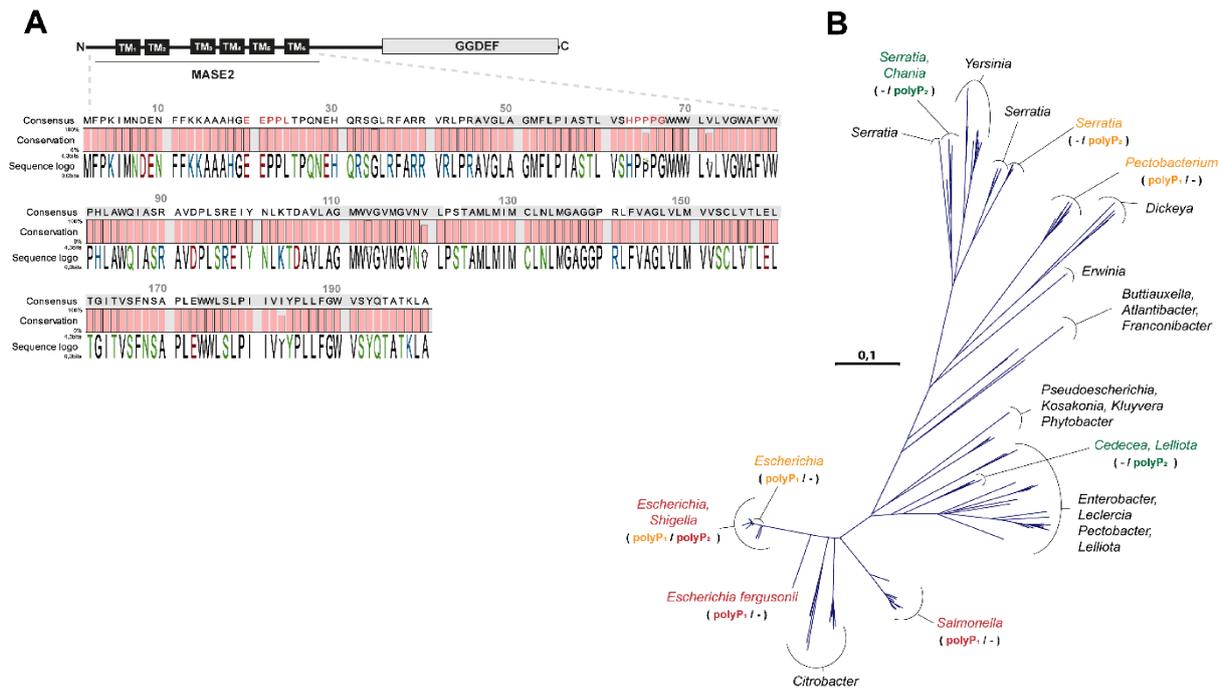
DgcC steht stellvertretend für Proteine mit polyP Motiv, das innerhalb einer Spezies in verschiedener Ausprägung auftritt. Lon repräsentiert hoch konservierte Proteine, die Teil von zentralen Signalnetzwerken sind. Die Histidinkinasen EvgS und EnvZ vermitteln die zelluläre Antwort auf spezifische Umweltreize. Die Tatsache, dass Proteine dieser Klasse überdurchschnittlich oft polyP Motive tragen und sich durch eine niedrige Expression auszeichnen, weist auf eine besondere Bedeutung der Motive hin.

#### 3.2.1 Die Bedeutung von HPPPG für die Diguanylatcyclase DgcC

In DgcC finden sich zwei polyP Motive: das am cytoplasmatischen N-Terminus lokalisierte „moderate“ EEPPL (polyP<sub>1</sub>) und das „starke“ HPPPG (polyP<sub>2</sub>), welches die erste und zweite Transmembrandomäne im periplasmatischen Raum unterteilt (Abb. 13A).

##### 3.2.1.1 Position und Konservierung der polyP Motive in DgcC

Beide polyP Motive sind in der N-terminalen, insgesamt sechs Transmembrandomänen umfassenden MASE2 Domäne lokalisiert<sup>176</sup>. Bedingt durch die hohe Anzahl an Transmembrandomänen existiert bislang weder eine Kristallstruktur des betreffenden Bereichs, noch ist eine Modellierung der Tertiärstruktur möglich. Die beiden Motive sind in *E. coli* unterschiedlich stark konserviert. Während EEPPL (polyP<sub>1</sub>) in allen betrachteten 43 *E. coli* Stämmen (vgl. Kapitel 2.3, Tabelle 10) auftritt, findet sich in 9 *E. coli* Stämmen anstelle des HPPPG Motivs (polyP<sub>2</sub>) HPSPG: den ExPEC O45:K1 (strain S88) und O18:K1:H7 (strain IHE3034), dem EPEC O127:H6 (strain E2348), den AIEC O83:H1 (strain NRG 857C) und UM164, den UPEC O6:K15:H31 (strain 536) und UTI89, sowie dem kommensalen O81 (strain ED1a) und dem APEC O1:K1. Alle Stämme sind Teil der hier insgesamt 15 Stämme umfassenden phylogenetischen Serogruppe B2, was darauf hindeutet, dass es sich bei der zugrunde liegenden Punktmutation C193T im *dgcC* Gen um ein einmaliges Ereignis handelt. Die Konservierung der polyP Motive wurde anhand eines Alignments der homologen DgcC Sequenzen aus *E. coli* visualisiert (Abb. 13A).



**Abb. 13:** Lokalisation der polyP Motive der Adenylatcyclase DgcC aus *E. coli* K-12 MG1655 und deren Konservierung. (A) Schematische Darstellung der Domänen von DgcC, sowie der Konservierung der MASE2 Domäne, sowie der darin enthaltenen EEPLL und HPPPG Motive. Transmembranregionen sind grau hinterlegt. Alignment von orthologen Proteinsequenzen aus 43 repräsentativen *E. coli* Stämmen (vgl. Tab. 10). (B) Rekonstruktion eines phylogenetischen Proteinbaumes aus DgcC Orthologen mit mindestens 44% Sequenzidentität und 90% Sequenzlänge der *E. coli* K-12 MG1655 Variante zur Nachverfolgung der Evolution der DgcC polyP Motive (polyP<sub>1</sub> = EEPLL, polyP<sub>2</sub> = HPPPG). Die farbige Hinterlegung der polyP Motive beschreibt deren Klassifizierung. Grün = „schwach“, orange = „moderat“, rot = „stark“.

Durch die Unterbrechung des Tripleprolin-Motivs durch Serin wird voraussichtlich das translationsarrestinduzierende Potential reduziert und das Motiv fällt in die Klasse der „schwachen“ polyP Motive. P65 zählt damit zu den 2,16% aller Aminosäuren aus DgcC, welche innerhalb der betrachteten 43 *E. coli* Stämme nicht zu 100% konserviert sind. Die Hauptfunktion von DgcC ist die Regulierung der Biofilmbildung durch die Synthese des Sekundärsignalmoleküls c-diGMP, welches unter anderem die Cellulosesynthese durch BcsABC induziert. Alle acht Stämme mit DgcC HPSPG Motiv tragen eine funktionelle Version des *bcsQ* Gens, welches im Gegensatz dazu in Laborstämmen wie *E. coli* K-12 MG1655 eine Punktmutation bzw. ein Stoppcodon nach Aminosäure 5 (TAG → TTG) aufweist. Damit kann ausgeschlossen werden, dass eine daraus resultierende Behinderung der Expression des *downstream* gelegenen *bcsA* Gens, eine definierte Kopienzahl von DgcC obsolet macht.

Ein Alignment von 381 orthologen Proteinsequenzen (Tab. S2) aus  $\gamma$ -Proteobakterien mit mindestens 44% Übereinstimmung und höchstens 10% Sequenzlängenabweichung zu DgcC aus *E. coli* K-12 zeigt, dass EEPLL (polyP<sub>1</sub>) konservierter ist als HPPPG (polyP<sub>2</sub>) (Abb. 13B). Während ein Motiv der Form X<sub>(-2)</sub>X<sub>(-1)</sub>PPX<sub>(+1)</sub> in den Gattungen *Escherichia* und *Shigella* („moderat“: EEPLL), *Pectobacterium* („moderat“: ASPPT) sowie in der Spezies *Citrobacter koseri* („schwach“: SEPPH) auftritt und in der Gattung *Salmonella* sogar bis auf wenige Ausnahmen durch zwei zusätzliche konsekutive Proline verstärkt ist („stark“: VAPPD, hier VAPPDPPD), ist das

$X_{(-1)}$  PPP  $X_{(+1)}$  Motiv ausschließlich in den nahe verwandten Gattungen *Escherichia* und *Shigella* vorzufinden (HPPPG). Interessanterweise ist das polyP Motiv in anderen DgcC Orthologen komplett verschwunden, bzw. findet sich in der entsprechenden Region nur ein einzelnes Prolin. In weiter entfernt verwandten Homologen von Spezies der Gattung *Cedecea*, *Lelliota*, *Serratia*, *Klebsiella* und *Chania* finden sich wiederum Prolincluster an ähnlicher Position, was auf eine unabhängige Entstehungsgeschichte hindeutet.

### 3.2.1.2 Einfluss des putativ „starken“ HPPPG Motivs auf Lokalisation und Kopienzahl von DgcC

Um zu testen, ob das Translationslevel von DgcC durch die natürlich vorkommende HPPPG/HPSPG Substitution beeinflusst wird, wurde ein plasmidbasiertes Reporterkonstrukt aus dem verkürzten (Aminosäure 1-65) nativen *E. coli* K-12 MG1655 *dgcC*, bzw. einer *dgcC* Variante mit P65S Substitution (CCG->TCG), wie sie in einigen anderen *E. coli* Stämmen, wie UTI89, vorkommt, und *lacZ* konstruiert. Die Genexpression wurde analog zu vorhergehenden, bereits veröffentlichten Ergebnissen (*cadC*, *amiB*, *nlpD*, *rzoR*, *tonB*, *uvrB*)<sup>30</sup> unter die Kontrolle des nativen  $P_{cadC}$  Promotors aus *E. coli* K-12 MG1655 gestellt. Die Level des Fusionsproteins wurden indirekt über die Analyse der  $\beta$ -Galactosidaseaktivität von transformierten *E. coli* BW25113 Wildtyp und  $\Delta$ *efp* Zellen bestimmt (Abb. 14A).

Die Translation von DgcC<sub>1-65,HPPPG</sub>-LacZ ist in Abwesenheit von EF-P auf ca. 15% des Wildtyplevels reduziert, während das abgeschwächte polyP Motiv in der Proteinvariante DgcC<sub>1-65,HPSPG</sub>-LacZ die EF-P Abhängigkeit zu einem gewissen Maß aufhebt und Translationslevel von 60% des Wildtyplevels zu beobachten sind. In Anwesenheit von EF-P hat die Art des polyP<sub>2</sub> Motivs (HPPPG oder HPSPG) keine Auswirkung auf das Translationslevel.

Das chromosomal in *E. coli* K-12 MG1655, am nativen Locus mit FLAG-Tag markierte *dgcC* Gen hat eine sehr niedrige Expressionsrate, weshalb das Protein im Western Blot mit spezifischen anti-FLAG-Antikörpern nicht detektierbar ist. Dies war weder im Fall von Gesamtzellextrakten, noch nach Anreicherung von Membranproteinen möglich. Um die Expression von *dgcC* zu erhöhen, wurden die Zellen alternativ auf Agar-Nährboden zwei Tage lang bei 28°C kultiviert (vgl. Sommerfeldt *et al.*, 2009<sup>326</sup>), was jedoch ebenfalls nicht zu ausreichend hohen DgcC-FLAG Kopienzahlen führte.

Aus diesem Grund wurden die *dgcC*-FLAG Varianten unter der Kontrolle des konstitutiv exprimierten *cadC*-Promotors in pBBR1 kloniert. Da auch die Erhöhung der Genkopienzahl für einen Nachweis von DgcC-FLAG nicht ausreichend war, wurde letztendlich auf ein durch L-Arabinose induzierbares Plasmidkonstrukt mit  $P_{BAD}$  Promotor zurückgegriffen. Um den Einfluss der beiden polyP Motive auf die finale Proteinkopienzahl von DgcC in Abhängigkeit von EF-P zu testen, wurden *E. coli* BW25113 Wildtyp und  $\Delta$ *efp* Derivate transformiert. Die Zellen wurden aus einer Übernachtskultur 1:500 in frisches LB Medium überführt, zwei Stunden lang bis zu einer OD<sub>600</sub> von 0,2 bei 37°C aerob inkubiert und die Genexpression mit L-Arabinose induziert (0,1%). Nach weiterer zweistündiger Inkubation wurden die Zellen geerntet, aufgeschlossen, fraktioniert und die beiden DgcC-Varianten (DgcC<sub>HPPPG</sub>-FLAG und DgcC<sub>HPSPG</sub>-FLAG) in einem anschließenden Western Blot relativ quantifiziert (Abb. 14B). Beide Proteinvarianten sind sowohl nach Expression im Wildtypstamm, als auch in der  $\Delta$ *efp* Mutante Teil der Membranfraktion. In der cytosolischen Fraktion finden sich hingegen keine nachweisbaren spezifischen Banden. Daher können verkürzte oder nicht-membranintegrierte, aber lösliche DgcC-Varianten, die aus dem

HPPPG/HPSPG Austausch resultieren, ausgeschlossen werden. Dasselbe gilt für eine maßgebliche Rolle von EF-P in der Membranintegration von DgcC. Im Vergleich zwischen Wildtyp und  $\Delta efp$  Stamm zeigt sich, dass die Translation von DgcC<sub>HPPPG</sub>-FLAG massiv von der Anwesenheit von EF-P abhängt. Während sowohl im Gesamtzelllysate, als auch in der Membranfraktion des Wildtyps deutliche Banden vorliegen, liegt das Level von DgcC<sub>HPPPG</sub>-FLAG im  $\Delta efp$  Stamm unter der Nachweisgrenze. Nach Substitution des „starken“ HPPPG Motivs durch das „schwache“ HPSPG ist das Proteinlevel im Wildtyp nur geringfügig höher als im  $\Delta efp$  Stamm und vergleichbar mit dem Level von DgcC<sub>HPPPG</sub>-FLAG im Wildtyp. Die leichte EF-P Abhängigkeit der DgcC<sub>HPSPG</sub>-FLAG Variante ist vermutlich auf das zusätzliche, hier unveränderte „moderate“ EEPPL Motiv zurückzuführen.

Um eine Proteinakkumulation weitgehend zu vermeiden, wurde die Genexpression durch die Verwendung von niedrigeren L-Arabinosekonzentrationen reduziert.

Während L-Arabinose selbst einen leichten wachstumsinhibierenden Effekt auf *E. coli* BW25113 Zellen hat, führt die Überexpression der plasmidkodierten *dgcC*-FLAG Varianten zu einer erheblichen Beeinträchtigung der Zellteilungsrate in Wildtyp, wie auch in  $\Delta efp$  Zellen (Abb. 14D). Dies unterstreicht die Bedeutung einer kontrollierten, begrenzten Kopienanzahl von DgcC. Gleichzeitig konnte gezeigt werden, dass das niedrigste artifizell erhöhte *dgcC*-Expressionslevel, welches hier zur Anwendung kam, nur geringe pleiotrope Auswirkungen hat. Bereits 0,00002% Arabinose führt zu detektierbaren DgcC-Mengen, welche die Zellteilungsrate jedoch nicht beeinträchtigen (Abb. 14C). Aufgrund einer gewissen Durchlässigkeit des P<sub>BAD</sub>-Promotors kann sogar völlig auf eine Induktion verzichtet werden.

Mit abnehmender Expressionsrate kristallisiert sich heraus, dass die DgcC<sub>HPSPG</sub>-FLAG Proteinlevel in Wildtypzellen etwas höher sind, als die der DgcC<sub>HPPPG</sub>-FLAG Variante (1,51±0,10-fach bei 0,00002% L-Arabinose; 2,77±1,10-fach bei 0,0% L-Arabinose, Abb. 14C). Unter diesen Bedingungen scheint das HPPPG Motiv die Translationseffizienz von DgcC so weit zu beeinträchtigen, dass dies eine Auswirkung auf die Proteinkopienzahl hat.

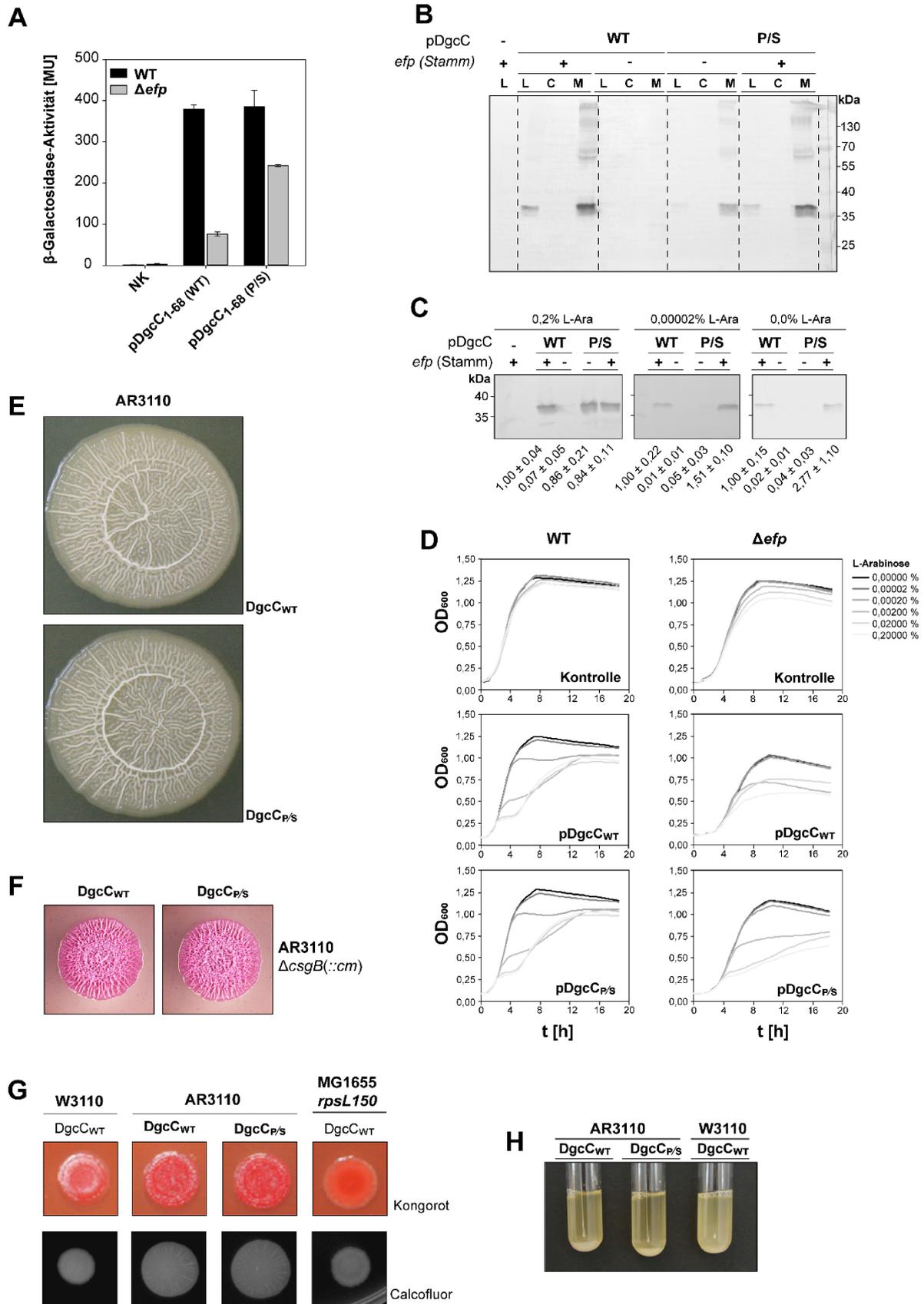


Abb 14: Beschreibung auf folgender Seite.

**Abb 14:** Beurteilung der Auswirkungen einer DgcC HPPPG/HPSPG Substitution (WT oder P/S) auf Translationslevel und Proteinfunktionalität. (A) Quantifizierung des Translationslevels von verkürzten (Aminosäure 1-65), plasmidkodierten *dgcC* Varianten in Abhängigkeit von EF-P mittels  $\beta$ -Galactosidaseassay. (B) Western-Blot zur zellulären Lokalisation der DgcC<sub>HPPPG</sub>-FLAG und DgcC<sub>HPSPG</sub>-FLAG Varianten in Abhängigkeit von EF-P. Die plasmidkodierten *dgcC*-FLAG Varianten stehen unter der Kontrolle des durch 0,1% L-Arabinose induzierten *P<sub>BAD</sub>*. (C) Relative Quantifizierung derselben DgcC<sub>HPPPG</sub>-FLAG und DgcC<sub>HPSPG</sub>-FLAG Varianten in Abhängigkeit zu EF-P und verschiedener Expressionslevel. (D) Wachstumskurven zur Beurteilung des Effekts verschiedener Expressionslevel der *dgcC*-FLAG Varianten auf die Zellteilungsrate in Wildtyp und  $\Delta$ *efp* Hintergrund. (E) Phänotypische Charakterisierung der Celluloseproduktion des Reporterstammes AR3110 nach chromosomaler HPPPG/HPSPG Substitution. Makrokolonien nach sechstägiger Inkubation bei 28°C auf LB-Agarplatten ohne NaCl. (F) Phänotypische Charakterisierung der Celluloseproduktion des Reporterstammes AR3110 nach chromosomaler HPPPG/HPSPG Substitution und Deletion des für die Curlbildung essentiellen Gens *csdB*. Makrokolonien nach viertägiger Inkubation bei 28°C auf Kongorot gefärbten LB-Agarplatten ohne NaCl (Alexandra Possling, AG Prof. Dr. Regine Hengge, Humboldt Universität zu Berlin). (G) Phänotypische Charakterisierung der Celluloseproduktion des Reporterstammes AR3110 nach chromosomaler HPPPG/HPSPG Substitution. Makrokolonien nach viertägiger Inkubation bei 28°C auf Kongorot oder Calcofluor gefärbten LB-Agarplatten ohne NaCl (Material und Methoden: DgcC<sub>WT</sub> = St46, DgcC<sub>P/A</sub> = St47). (H) Sedimentationsassay.

Material und Methoden: (A) NK = V3, pDgcC<sub>1-68(WT)</sub> = V7, pDgcC<sub>1-68(P/S)</sub> = V8, WT = St4,  $\Delta$ *efp* = St6. (B-D) pDgcC<sub>WT</sub> = V21, pDgcC<sub>P/S</sub> = V22, WT (*efp*<sup>+</sup>) = St4,  $\Delta$ *efp* (*efp*<sup>-</sup>) = St6. (E-H) AR3110 DgcC<sub>WT</sub> = St45, AR3110 DgcC<sub>P/A</sub> = St46, AR3110 *csdB::cm* DgcC<sub>WT</sub> = St49, AR3110 *csdB::cm* DgcC<sub>P/S</sub> = St50, W3110 DgcC<sub>WT</sub> = St44, MG1655 *rpsL150* = St31.

Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen, bzw. repräsentieren eine von drei unabhängigen Versuchsdurchführungen.

### 3.2.1.3 Einfluss des putativ „starken“ HPPPG Motivs auf die Proteinfunktionalität von DgcC

Es ist davon auszugehen, dass eine veränderte DgcC Kopienzahl oder eine Beeinträchtigung der Funktionalität des Proteins Auswirkungen auf die assoziierten Signalwege hat. Deshalb wurde getestet, inwiefern diese durch die DgcC HPPPG/HPSPG Substitution beeinflusst werden und ob der Austausch zu einer phänotypischen Veränderung der Bakterienzellen führt, die unter bestimmten, für Stämme mit natürlicher HPPPG/HPSPG Substitution relevanten Umweltbedingungen vorteilhaft ist. Vor diesem Hintergrund wurde die Cellulosesynthese eines *E. coli* AR3110 Stammes nach chromosomaler HPPPG/HPSPG Substitution untersucht.

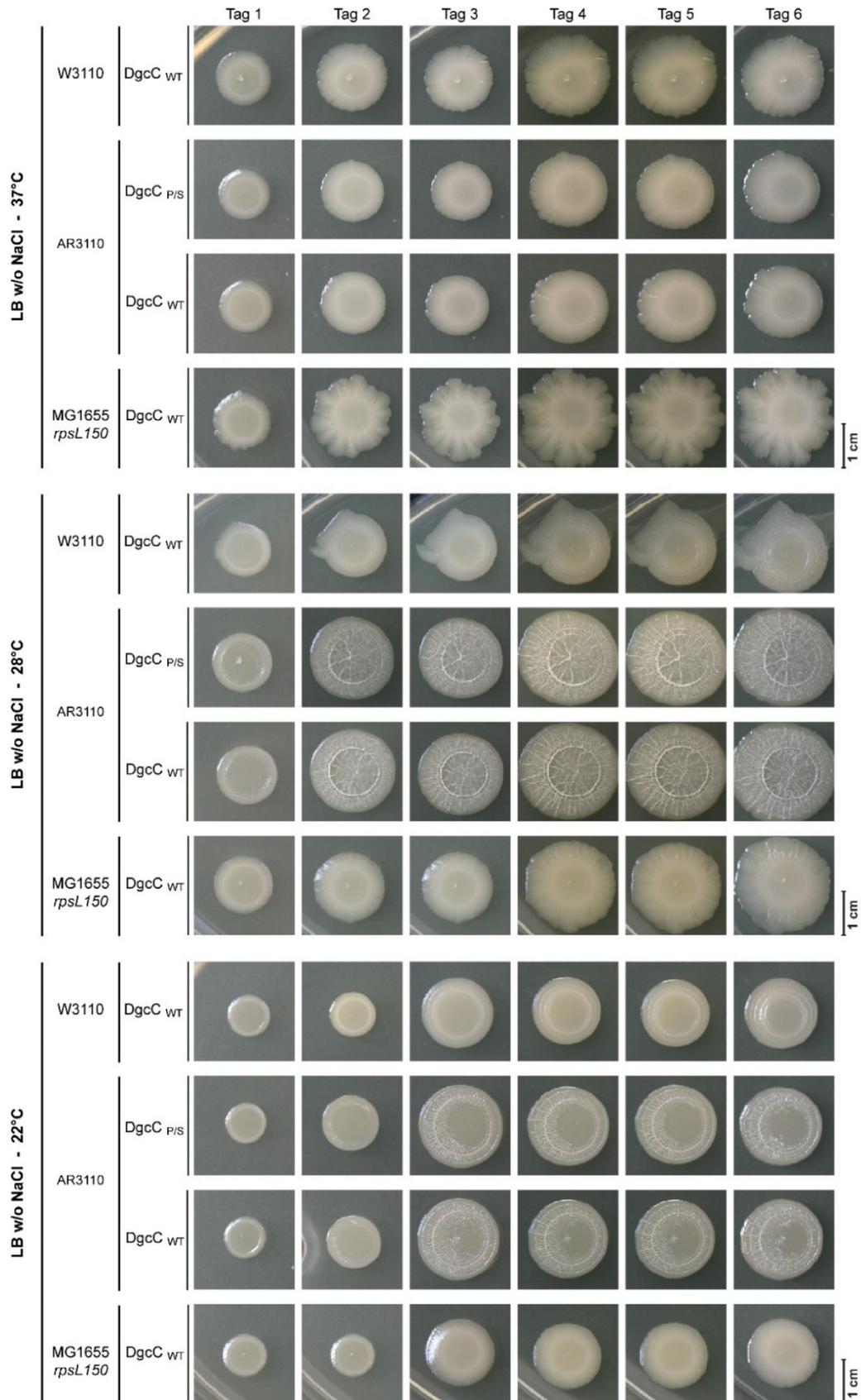
Der Reporterstamm AR3110 (basierend auf W3110) ist durch eine Wiederherstellung des *bcsQ* Gens in der Lage, das Operon *yhjR-bcsQABZC* der für die Cellulosesynthese kodierenden Gene zu transkribieren. Da Cellulose ein Hauptbestandteil der extrazellulären Matrix ist, lässt sich die Aktivität der Cellulose-Synthese durch eine visuelle Betrachtung der Oberfläche von Makrokolonien auf Agarplatten beurteilen. Dafür wurden die Zellen in LB-Medium bis in die frühe exponentielle Wachstumsphase (OD<sub>600</sub> ~ 0,2) bei 37°C vorkultiviert und nach dem Auftropfen auf LB-Agarplatten bei 28°C für bis zu 16 Tage inkubiert. Zur Anfärbung der Cellulose wurden zusätzlich mit Kongorot und Calcofluor versetzte LB-Agarplatten verwendet.

Nach spätestens vier Tagen bildeten sich auf den Kolonien von beiden AR3110 Stämmen (DgcC<sub>HPPPG</sub> und DgcC<sub>HPSPG</sub>) wulstige Erhebungen, in denen sowohl Kongorot, als auch Calcofluor akkumulierten (Abb. 14E und G). Die für Cellulose charakteristischen Erhebungen in der Oberfläche der Makrokolonien konnten in den als Negativkontrolle verwendeten Stämmen W3110 und MG1655 nicht nachgewiesen werden und bildeten sich auch nicht unter anderer, als der für die Cellulosesynthese als optimal beschriebenen Temperatur aus. Über den

gesamten Zeitraum des Experiments ließ sich kein Unterschied zwischen AR3110 Dgc<sub>CHPPP</sub>G und Dgc<sub>CHPSPG</sub> in Art oder Umfang der als Cellulose identifizierten Matrix erkennen (Abb. 15).

Um das Auftreten anderer struktureller Komponenten zu minimieren, wurde in beiden Stämmen zusätzlich das Gen *csgB* deletiert ( $\Delta csgB(::cm)$ ), welches für die Haupteinheit des Curlins kodiert und damit essentiell ist für die Ausbildung von extrazellulären Curli. Doch auch diese Stämme zeigten keinen unterschiedlichen Phänotyp (Abb. 14E). Ergänzend wurden AR3110 Dgc<sub>CHPPP</sub>G und Dgc<sub>CHPSPG</sub> in einem Sedimentationsassay charakterisiert. Dabei wurden Kulturen über Nacht bei 37°C aerob kultiviert und am folgenden Tag hinsichtlich der Sedimentation der Bakterien innerhalb von einer Stunde verglichen. Im Vergleich zu einer *E. coli* K-12 MG1655 Zellkultur, welche sich durch eine homogene Verteilung im Medium und damit ein trübes Erscheinungsbild auszeichnet, kommt es in den beiden cellulosebildenden Stämmen, unabhängig von der Art der DgcC Variante zu einem Absetzen der Zellen. Das Medium erscheint klar (Abb. 14H).

Schlußfolglich schlägt sich die Substitution des „starken“ HPPP Motivs im genetischen Hintergrund von *E. coli* AR3110 nicht in einer detektierbaren phänotypischen Varianz nieder.



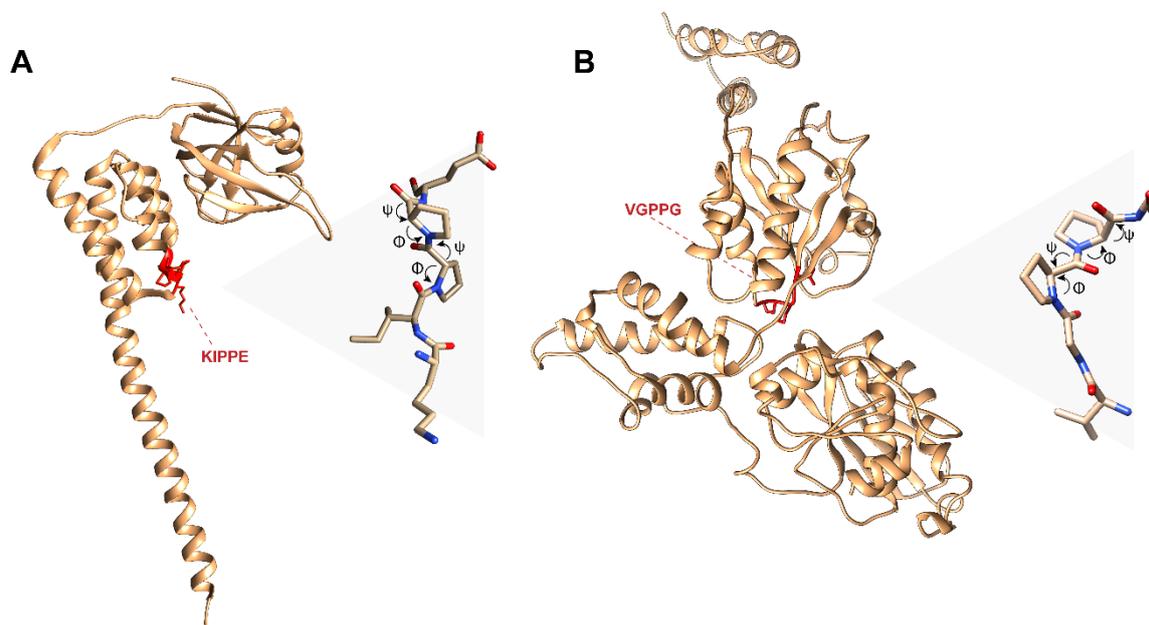
**Abb. 15:** Verlauf der Cellulosesynthese von AR3110 Reporterstämmen mit und ohne chromosomal kodierter HPPPG/HSPSG Substitution (WT oder P/S). Makrokolonien wurden bei 28°C und als Negativkontrolle bei 37°C und 22°C auf LB-Agarplatten ohne NaCl kultiviert. (W3110 DgcC<sub>WT</sub> = St44, AR3110 DgcC<sub>WT</sub> = St45, AR3110 DgcC<sub>P/S</sub> = St46, W3110 DgcC<sub>WT</sub> = St44)

### 3.2.2 Die Bedeutung von VGPPG für die Lon Protease

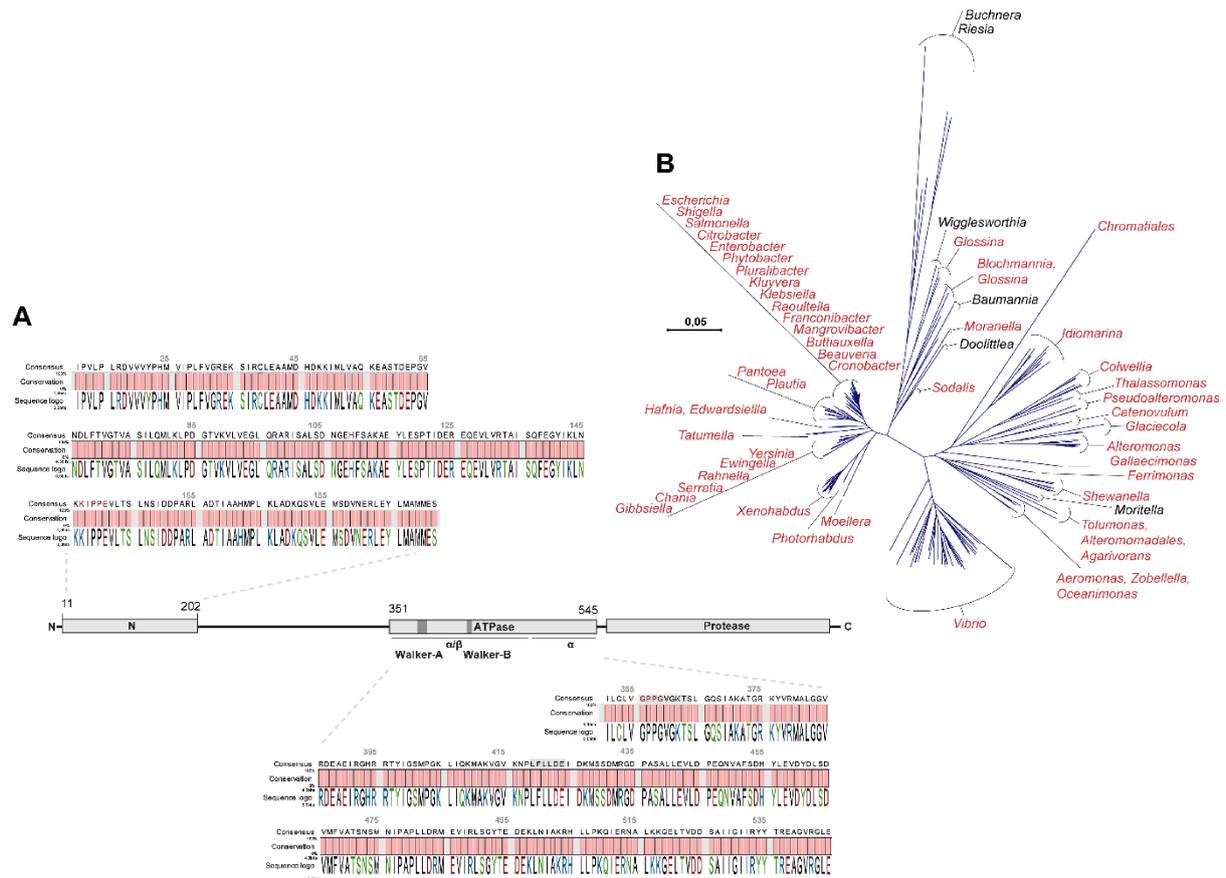
Die *E. coli* K-12 MG1655 Protease Lon trägt zwei „starke“ polyP Motive: KIPPE in der N-terminalen Region und VGPPG in der ATP bindenden Domäne.

#### 3.2.1.1 Position und Konservierung der polyP Motive in Lon

Die Kristallstruktur von Lon ist zum Großteil aufgelöst und deckt das Motiv KIPPE ab (PDB 3LJC<sup>327</sup>). Das Motiv trennt zwei  $\alpha$ -Helices und ist Teil einer Schleife. Der VGPPG beinhaltende Sequenzabschnitt ist bislang noch nicht aufgelöst, weshalb die betreffende 3D-Struktur mittels Phyre2 modelliert wurde. Dabei zeigte sich, dass das Motiv zentral und offen in der ATP bindenden Domäne liegt, was auf eine Beteiligung an der Substratbindung hindeutet (Abb. 16A und B). Lon aus *E. coli* ist zu einem hohen Maß konserviert. Nur 0,26% aller Aminosäurepositionen aus 43 repräsentativen *E. coli* Stämmen sind nicht zu 100% konserviert (Abb. 17A). Die beiden polyP Motive sind unverändert.



**Abb. 16:** Lokalisation der polyP Motive der Protease Lon aus *E. coli* K-12 MG1655. (A) Kristallstruktur (PDB 3LJC, aa 8-245)<sup>327</sup> der N-terminalen Domäne der *E. coli* K-12 Lon Protease mit darin enthaltenem KIPPE Motiv. (B) 3D Modell der Struktur der *E. coli* K-12 Lon ATPase Domäne mit darin enthaltenem VGPPG Motiv (aa 248-771).



**Abb. 17:** Konservierung der polyP Motive der Protease Lon aus *E. coli* K-12 MG1655 (A) Schematische Darstellung der Domänen von Lon, sowie der Konservierung der N-terminalen Domäne und der ATPase Domäne, sowie der darin enthaltenen KIPPE und VGPPG Motive. Walker A und B Motive sind grau hinterlegt. Alignment von orthologen Proteinsequenzen aus 43 repräsentativen *E. coli* Stämmen (vgl. Tab. 10). (B) Rekonstruktion eines phylogenetischen Proteinbaumes aus Lon Orthologen mit mindestens 44% Sequenzidentität und 90% Sequenzlänge der *E. coli* K-12 MG1655 Variante zur Nachverfolgung der Evolution der Lon polyP Motive (polyP<sub>1</sub> = KIPPE, polyP<sub>2</sub> = VGPPG). VGPPG ist mit Ausnahme der Gattungen *Glossina* und *Doolittlea* unverändert. Die farbige Hinterlegung der Gattungen beschreibt die Klassifizierung des polyP<sub>1</sub> Motivs in der N-terminalen Region. Grün = „schwach“, orange = „moderat“, rot = „stark“.

Unter Berücksichtigung von 203 Orthologen (Tab. S3) zu *E. coli* K-12 MG1655 Lon mit mindestens 44% Sequenzidentität und höchstens 10% Abweichung in der Sequenzlänge, zeigte sich, dass die beiden Motive auch über mehrere Gattungen hinweg sehr konserviert sind (Abb. 17B). Aufgrund der hohen Konservierung von Lon in sämtlichen lebenden Organismen wurde das Alignment auf Referenz-Proteome beschränkt. Das Motiv KIPPE ist mit Ausnahme der Gattungen *Doolittlea* (KIPFE), *Baumannia* (KIPLE), *Wigglesworthia* (KIPTE), *Moritella* (KIPAE) *Buchnera* (KIPLE) und *Riesia* (KISPE), in fast allen Orthologen zu finden. In diesen ist eines der beiden konsekutiven Proline substituiert und der translationsarretierende Effekt daher vorraussichtlich nicht gegeben. Im Gegensatz dazu findet sich VGPPG unverändert in allen hier untersuchten homologen Proteinsequenzen, außer in Spezies der Gattungen *Glossina*, *Wigglesworthia* und *Doolittlea* wo das Motiv eine Varianz in den flankierenden Aminosäuren aufweist (MGPPG) oder gar nicht mehr vorhanden ist (VGSPG).

### 3.2.1.2 Einfluss des putativ „starken“ VGPPG Motivs auf Lokalisation und Kopienzahl von Lon

Das an prominenterer Position gelegene und konserviertere „starke“ polyP Motiv VGPPG wurde auf seine Rolle in der Funktionalität und Kopienzahl von Lon untersucht. Eine Analyse der EF-P Abhängigkeit des Translationslevels eines Lon-LacZ Fusionsproteins durch  $\beta$ -Galactosidase-Aktivitätsbestimmung verlief vermutlich aufgrund der proteolytischen Eigenschaften von Lon ergebnislos.

Aufgrund eines ausreichend hohen nativen Proteinlevels kann Lon mittels Western Blot relativ quantifiziert werden (Abb. 18A). Dabei zeigte sich unter der Verwendung von Gesamtzelllysaten von Wildtyp und  $\Delta efp$  Zellen, dass das Translationslevel nicht von der EF-P Präsenz abhängt. Ein Austausch des chromosomal kodierten Lon VGPPG durch VGAAG hat keinen messbaren Effekt auf die Gesamtkopienzahl von Lon, wie eine Analyse von Gesamtzellextrakten und cytosolischer Fraktion der betreffenden *E. coli* K-12 MG1655 Stämme zeigte. Eine Aggregation und das Vorhandensein von unlöslichen Einschlusskörpern können damit ebenfalls ausgeschlossen werden.

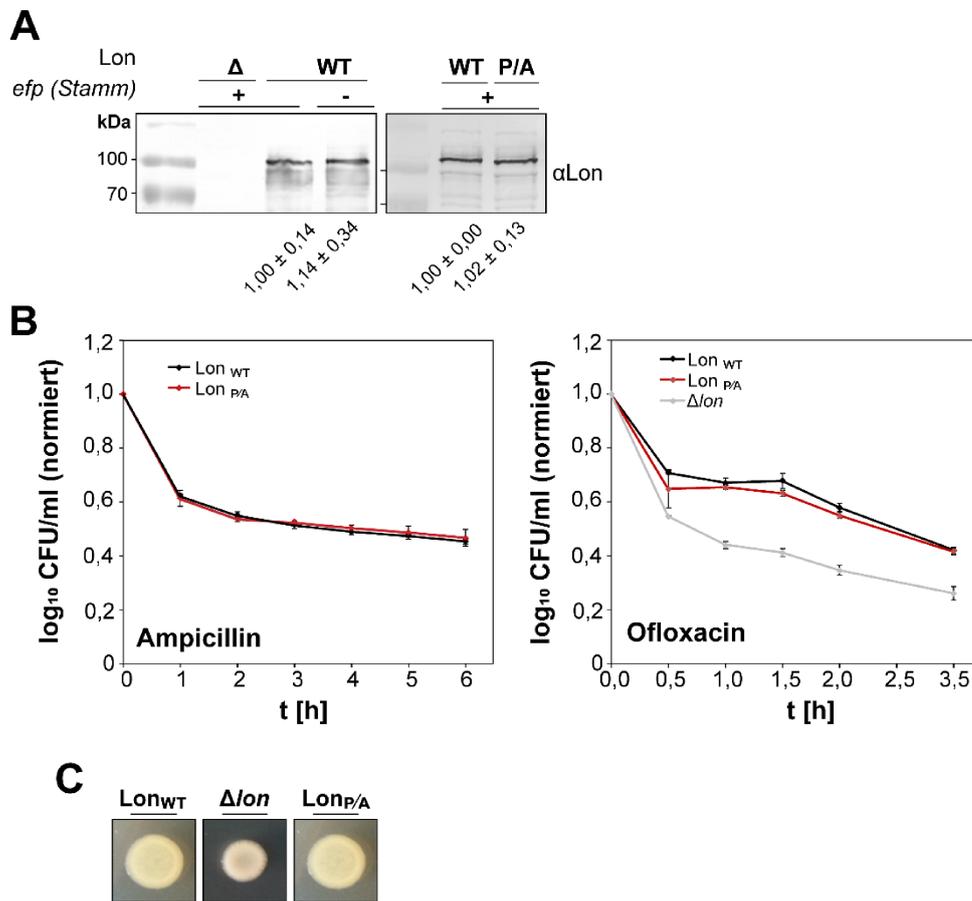
### 3.2.1.3 Einfluss des putativ „starken“ VGPPG Motivs auf die Proteinfunktionalität von Lon

Die Rolle des VGPPG Motivs für die Funktionalität von Lon, bzw. für die damit assoziierten physiologischen Mechanismen, wurde mittels phänotypischer Charakterisierung eines *E. coli* K-12 MG1655 Stammes mit chromosomal kodierter, markerloser VGPPG/VGAAG Substitution untersucht. Lon spielt eine Schlüsselrolle in der Degradierung von Antitoxinen, die Teil von Toxin/Antitoxin Paaren sind und für die Etablierung von persistenten Einzelzellen während anhaltender Stressbedingung sind. Der hier angewendete Persistenz-Assay basiert auf der antibiotikabasierten Selektion von nichtteilenden Zellen. Das  $\beta$ -Lactam-Antibiotikum Ampicillin blockiert die Neusynthese von Murein und damit die Ausbildung einer intakten Zellwand, während das Fluorchinolon Ofloxacin die bakteriellen Topoisomerasen II und IV hemmt und damit die DNA-Replikation inhibiert<sup>328</sup>. Während einer fünfstündigen Inkubation wurden die verbleibenden Zellen durch Ausplattieren auf LB-Agar ausgezählt und damit Rückschlüsse auf deren Eigenschaft, unter den gegebenen Verhältnissen in einen persistenten Zustand überzugehen, gezogen (Abb. 18B).

Dabei zeigte sich, dass die Lon<sub>VGAAG</sub> Variante zu keiner Beeinträchtigung führt. Der entsprechende Stamm verhält sich in diesem Assay unabhängig von der Art des eingesetzten Antibiotikums wie der Wildtyp *E. coli* K-12 MG1655. Im Gegensatz dazu zeigt ein *E. coli* BW25113  $\Delta lon$  Stamm erheblich reduzierte Zellzahlen, was auf eine massive Einschränkung der Fähigkeit, Persister-Zellen auszubilden, hindeutet.

Auf LB-Agarplatten gewachsene Makrokolonien des *E. coli* K-12 MG1655 Stammes mit Lon<sub>VGAAG</sub> Variante zeigen das gleiche Erscheinungsbild, wie der Wildtyp. *E. coli* BW25113  $\Delta lon$  Zellen unterscheiden sich durch ihren, aus einer Überproduktion des Kapselpolysaccharids Cholansäure resultierenden, mukoiden Phänotyp deutlich davon (Abb. 18C)<sup>329</sup>.

Die Resultate weisen auf eine untergeordnete Rolle des VGPPG Motivs in der Funktion von Lon hin. Mögliche phänotypische Effekte liegen unter der Nachweisgrenze der hier verwendeten Assays.



**Abb. 18:** Beurteilung der Auswirkungen einer chromosomal kodierten Lon VGPPG/VGAAG Substitution (WT oder P/A) auf Translationslevel und Proteinfunktionalität. (A) Western-Blot zur relativen Quantifizierung der Lon Proteinlevel in Abhängigkeit des polyP Motivs oder der Präsenz von EF-P. (B) Persistenz-Assay zur phänotypischen Charakterisierung von *E. coli* Stämmen nach chromosomaler Lon VGPPG/VGAAG Substitution. CFU = colony forming units. (C) Erscheinungsbild von Makrokolonien nach Lon VGPPG/VGAAG Substitution und zweitägiger Inkubation bei 37°C auf LB-Agarplatten.

Material und Methoden: (A-C) WT/efp<sup>+</sup> = St4, WT/efp<sup>-</sup> = St6,  $\Delta lon$  = St36, Lon<sub>WT</sub> = St31, Lon<sub>P/A</sub> = St37.

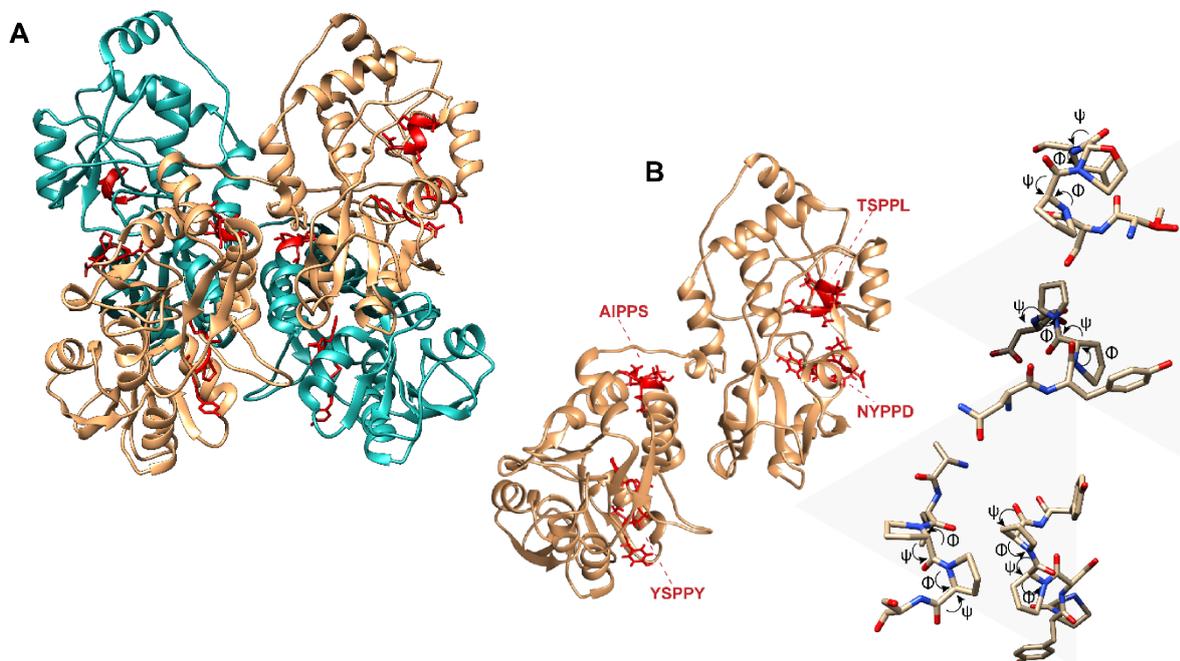
Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen, bzw. repräsentieren eine von drei unabhängigen Versuchsdurchführungen.

### 3.2.3 Die Bedeutung von NYPPD für die Histidinkinase EvgS

Die periplasmatische Sensordomäne der Histidinkinase EvgS beinhaltet zwei VFT-Domänen (*venus fly trap*) und trägt vier polyP Motive unterschiedlicher „Stärke“: das „schwache“ TSPPL (polyP<sub>1</sub>) und das „starke“ NYPPD (polyP<sub>2</sub>) in VFT1, sowie das „moderate“ YSPPY (polyP<sub>3</sub>), welches allerdings eventuell durch ein *upstream* liegendes weiteres Prolin verstärkt wird (PYSPPY) und das „starke“ AIPPS (polyP<sub>4</sub>) in VFT2.

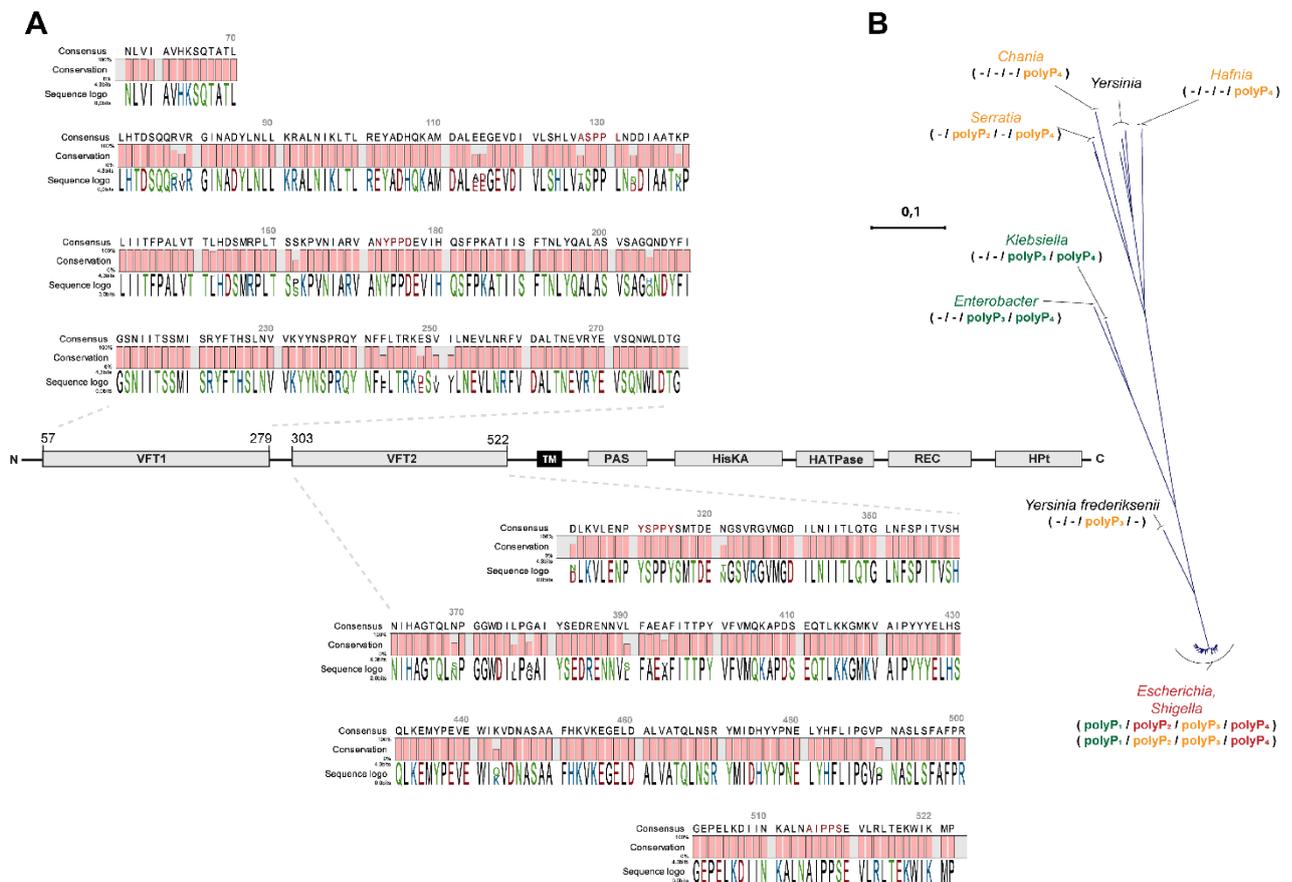
#### 3.2.1.1 Position und Konservierung der polyP Motive in EvgS

Die Kristallstruktur der periplasmatischen EvgS Domänen aus *E. coli* K-12 legt die Position der darin enthaltenen polyP Motive offen<sup>156</sup> (Abb. 19A und B). TSPPL, NYPPD und YSPPD formen Schleifenstrukturen. TSPPL und NYPPD liegen in relativ geringer räumlicher Entfernung (Abstand von P129 zu P175: 7,0 Å; S128 zu P175: 2,8 Å) in zwei gegenüberliegenden unterschiedlichen globulären Einheiten. YSPPD liegt in einer zentralen Region von VFT2 und wird umrahmt von zwei  $\alpha$ -Helices und einer  $\beta$ -Faltblattstruktur. EvgS liegt als Dimer vor. Die Strukturen der beiden periplasmatischen Domänen sind dabei komplementär und passen ineinander. Dabei formieren die VFT Domänen einen Kern aus insgesamt sechs  $\alpha$ -Helices. In dieser Struktur liegen sich die AIPPS Motive der beiden EvgS Moleküle gegenüber, wobei der Abstand zwischen den beiden S509 6,35 Å beträgt und die Proline selbst voneinander abgewandt positioniert sind (P508: 17,7 Å).



**Abb. 19:** Lokalisation der polyP Motive der Histidinkinase EvgS aus *E. coli* K-12 MG1655 (A) Kristallstruktur<sup>156</sup> der periplasmatischen EvgS Domänen mit darin enthaltenen polyP Motiven. Darstellung als Dimer (aa 26-519). (B) Kristallstruktur<sup>156</sup> der periplasmatischen EvgS Domänen aus *E. coli* K-12 mit darin enthaltenen polyP Motiven. Darstellung als Monomer (aa 26-519).

Die Motive sind innerhalb von 43 repräsentativen *E. coli* Stämmen unterschiedlich stark konserviert (Abb. 20A). Mit Ausnahme des *E. coli* Stammes O78:H11 (strain H10407) tragen Stämme der phylogenetischen Gruppen A und B1 ausschließlich TSPPL (polyP<sub>1</sub>), während sich in allen untersuchten Stämmen der phylogenetischen Gruppen B2, D, E und F an analoger Position das „moderate“ Motiv ASPPL vorfindet. Diese Zweiteilung spricht für ein einmaliges Substitutionsereignis, welches zur Änderung der Motivstärke führte. NYPPD (polyP<sub>2</sub>) ist zu 100% konserviert und in allen EvgS Homologen anzufinden. Dasselbe gilt für PYSPPY (polyP<sub>3</sub>) und AIPPS (polyP<sub>4</sub>). Insgesamt sind 4,76% aller Aminosäuren nicht zu 100% in allen 43 *E. coli* Stämmen konserviert.



**Abb. 20:** Konservierung der polyP Motive der Histidinkinase EvgS aus *E. coli* K-12 MG1655. (A) Schematische Darstellung der Domänen von EvgS, sowie der darin enthaltenen polyP Motive TSPPL, NYPPD (P)YSPPY und AIPPS. Alignment von orthologen Proteinsequenzen aus 43 repräsentativen *E. coli* Stämmen und Darstellungen der Konservierung von VFT1 und VFT2 Domänen (vgl. Tab. 10). (B) Rekonstruktion eines phylogenetischen Proteinbaumes aus Orthologen mit mindestens 44% Sequenzidentität und 90% Sequenzlänge zu EvgS aus *E. coli* K-12 MG1655 zur Nachverfolgung der Evolution der polyP Motive (polyP<sub>1</sub> = TSPPL, polyP<sub>2</sub> = NYPPD, polyP<sub>3</sub> = (P)YSPPY, polyP<sub>4</sub> = AIPPS). Die farbige Hinterlegung der polyP Motive beschreibt deren Klassifizierung. Grün = „schwach“, orange = „moderat“, rot = „stark“.

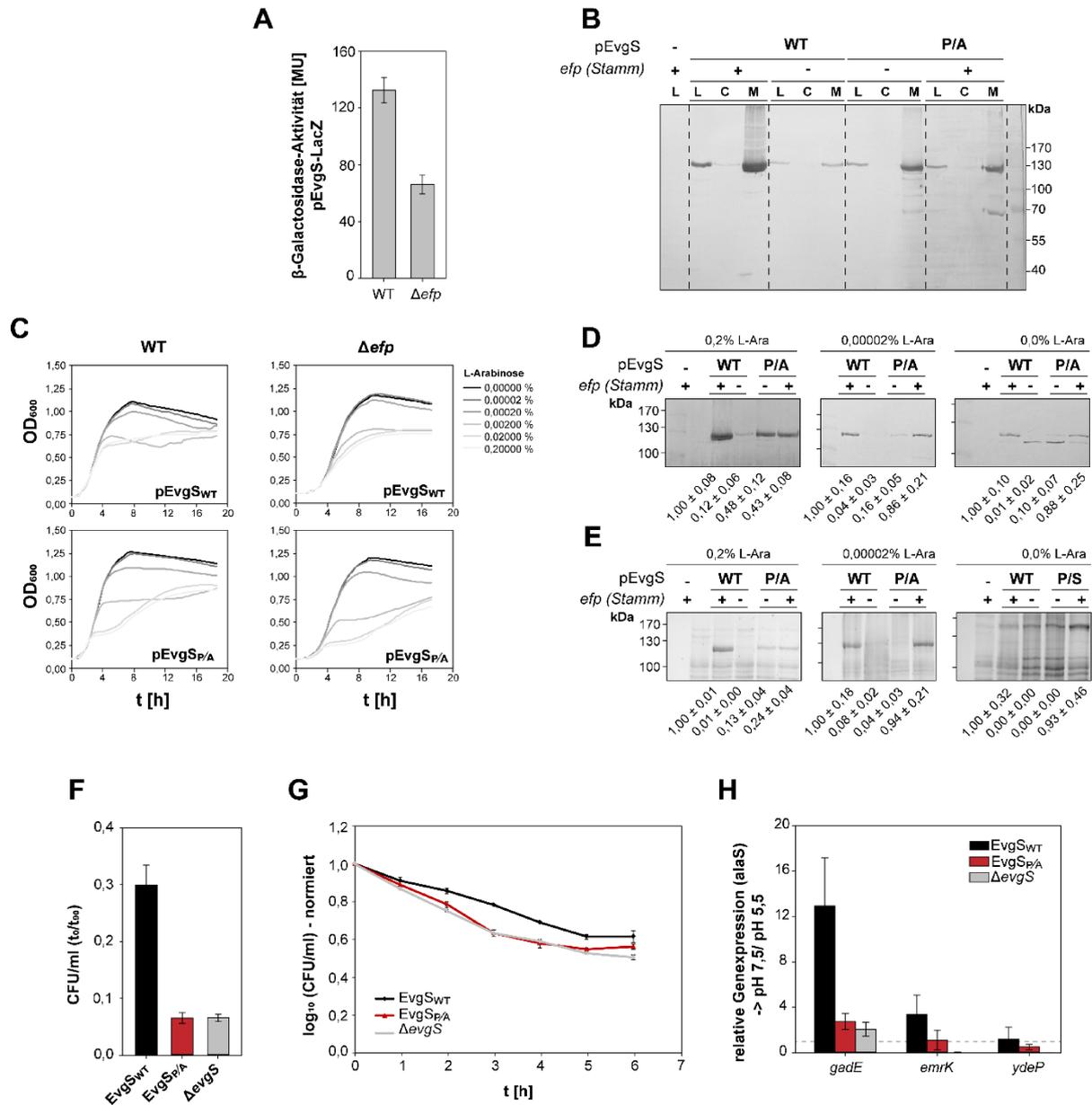
Ein Alignment von 183 Orthologen (Tab. S4) mit mindestens 44% Sequenzidentität und höchstens 10% Sequenzlängenabweichung von EvgS aus *E. coli* K-12 MG1655 verdeutlicht die Evolution der Motive (Abb. 20B). Die Proteinsequenzen der einzelnen Gattungen (ausgenommen *Escherichia* und *Shigella*) unterscheiden sich sehr stark voneinander. TSPPL (polyP<sub>1</sub>), bzw. ASPPL finden sich ausschließlich in Spezies der Gattungen *Escherichia* und *Shigella*, während in weiter entfernt verwandten Gattungen keine konsekutiven Proline an analoger Position auftreten. NYPPD (polyP<sub>2</sub>) findet sich ebenfalls in *Escherichia* und *Shigella*, wobei zwei konsekutive Proline mit abweichenden, flankierenden Aminosäuren ebenfalls in *Serratia* (DFPPK/E, „moderat“) auftreten. (P)YSPPY (polyP<sub>3</sub>) ist stärker konserviert und existiert unverändert in EvgS Orthologen der Gattungen *Escherichia*, *Shigella* und der Spezies *Yersinia frederiksenii* (PYSPPF, „moderat“). In den Gattungen *Enterobacter* und *Klebsiella* (SFHPPF, „schwach“), sowie *Serratia*, *Chania*, *Yersinia* und *Hafnia* (PYYAPL/F, „kein Effekt“) liegt das polyP Motiv in abgeschwächter Form oder gar nicht vor. Ein polyP Motiv an analoger Position zu AIPPS (polyP<sub>4</sub>) in EvgS aus *E. coli* K-12 MG1655 ist in allen Gattungen außer *Yersinia* (NIPQK, kein Effekt) anzufinden. Eine abgeschwächte Form findet sich in *Serratia* (DIPPK, „moderat“) und *Hafnia* (NIPPK, „moderat“). Die EvgS polyP Motive sind umso konservierter, je stärker der vorhergesagte Translationsarrest ist.

### 3.2.1.2 Einfluss des putativ „starken“ NYPPD Motivs auf Lokalisation und Kopienzahl von EvgS

Das Translationslevel eines plasmidbasierten und unter der Kontrolle des konsekutiven *cadC* Promotor stehenden EvgS-LacZ Reporterkonstruktes ist in Abwesenheit von EF-P auf ca. 60% des Wildtyplevels reduziert (Abb. 21A). Die Translation von EvgS ist demnach von EF-P abhängig.

Die Detektion von nativen EvgS Proteinleveln im Western Blot mit spezifischen  $\alpha$ -EvgS Antikörpern ist aufgrund dessen niedriger Expression auch nach Anreicherung von membrangebundenen Proteinen nicht möglich. Eine FLAG-Markierung stellt eine einfache Möglichkeit dar, ein Protein mit einer bekannten Struktur zu ergänzen, die mit hoher Affinität von Antikörpern gebunden wird. Jedoch konnte auch ein C-terminal an EvgS fusionierter FLAG-Tag im Westernblot mittels spezifischem  $\alpha$ -FLAG Antikörper nicht eindeutig nachgewiesen werden.

Aus diesem Grund wurde die Expression von *evgS* artifiziell erhöht. Da ein unter der Kontrolle des konsekutiven *cadC*-Promotor stehendes *evgS-FLAG* Konstrukt nicht zu einer ausreichend hohen Expressionsrate führte, wurde auf den L-Arabinose-induzierbaren P<sub>BAD</sub>-Promotor zurückgegriffen. Dabei wurde neben dem nativen EvgS Protein eine Variante getestet, in welcher eines der beiden „starken“ polyP Motive, NYPPD, durch NYAAD substituiert wurde. Beide Proteine, EvgS<sub>NYPPD</sub>-FLAG und EvgS<sub>NYAAD</sub>-FLAG, können nach Induktion mit 0,1% L-Arabinose im Western Blot detektiert werden (Abb. 21B). Dabei stellt sich eine deutliche EF-P-Abhängigkeit des Proteinlevels dar. Die spezifische Bande von EvgS<sub>NYPPD</sub>-FLAG ist im Wildtyphintergrund deutlich stärker, als es bei einer  $\Delta$ *efp* Mutante der Fall ist. Beide EvgS-Varianten sind sowohl in Wildtyp-, als auch in  $\Delta$ *efp* Zellen Teil der Membranfraktion, während keine löslichen Derivate nachgewiesen werden konnten.



**Abb. 21:** Beurteilung der Auswirkungen einer EvgS NYPPD/NYAAD Substitution (WT oder P/A) auf Translationslevel und Proteinfunktionalität. (A) Quantifizierung des Translationslevels eines plasmidkodierten EvgS-LacZ Reporters in Abhängigkeit von EF-P mittels  $\beta$ -Galactosidaseassay. (B) Western-Blot zur relativen Quantifizierung von plasmidkodierten, unter der Kontrolle des  $P_{BAD}$  Promotors stehenden EvgS<sub>NYPPD</sub>-FLAG und EvgS<sub>NYAAD</sub>-FLAG in Abhängigkeit von EF-P. (C) Wachstumskurven zur Beurteilung des Effekts verschiedener Expressionslevel der EvgS-FLAG Varianten auf die Zellteilungsrate in Wildtyp und  $\Delta$ efp Hintergrund. (D) SDS-PAGE von angereicherten Membranproteinen zur relativen Quantifizierung der Proteinlevel von plasmidkodierten EvgS<sub>NYPPD</sub>-FLAG und EvgS<sub>NYAAD</sub>-FLAG Varianten in Abhängigkeit von EF-P bei verschiedenen hohen Expressionsleveln. (E) Auf (D) basierender Western-Blot. (F) Phänotypische Charakterisierung einer chromosomal kodierten EvgS<sub>NYAAD</sub> Variante. Anteil der verbleibende CFUs nach 90-minütiger Inkubation ( $t_{90}$ ) bei pH 2 nach vorheriger Adaptation an moderate Säurestressbedingungen von pH 5,5 ( $t_0$ ). (G) Stichprobenartige Analyse der Anzahl verbleibender CFUs während der Kultivierung bei pH 3,1 über einen Zeitraum von sechs Stunden. (H) qPCR Analyse des Expressionslevels der durch EvgS/EvgA regulierten Gene *gadE*, *emrK* und *ydeP* nach zweistündiger Inkubation bei moderatem Säurestress (pH 5,5) im Vergleich zu pH 7,5. Vergleich von Stämmen mit EvgS<sub>NYPPD</sub> oder EvgS<sub>NYAAD</sub> Variante.

Material und Methoden: (A) pEvgS-LacZ = V5, WT = St4,  $\Delta$ efp = St6. (B-E) pEvgS<sub>NYPPD</sub>-FLAG = V23, pEvgS<sub>NYAAD</sub>-FLAG = V24, WT = St4,  $\Delta$ efp = St6. (F-H) EvgS<sub>WT</sub> = St31, EvgS<sub>P/A</sub> = St35,  $\Delta$ evgS = St34.

Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen, bzw. repräsentieren eine von drei unabhängigen Versuchsdurchführungen.

Eine artifiziell erhöhte Kopienzahl der EvgS-Varianten wirkt sich auf die Teilungsrate der transformierten Zelllinien aus, wie Wachstumskurven mit unterschiedlich induzierten Kulturen zeigten (Abb. 21C). Dabei nähern sich die Werte von mit 0,00002% L-Arabinose induzierten Kulturen, denen von nicht-induzierten an. Aus diesem Grund entsprechen die dazugehörigen Westernblot-Ergebnisse am wahrscheinlichsten dem tatsächlich vorherrschenden Zustand eines nativ exprimierten *evgS* Gens. Aufgrund der Durchlässigkeit des  $P_{BAD}$ -Promotors kann auch hier völlig auf den Induktor verzichtet werden. Die EvgS-Varianten können bei diesem Expressionslevel nach Anreicherung von Membranproteinen im Western Blot detektiert werden (Abb. 21E). Da die EvgS-spezifische Bande nach SDS-PAGE an sehr prominenter Position zu finden ist, ist es möglich, die Proteinbande bereits im Gel mittels TCE-Färbung zu identifizieren und quantifizieren (Abb. 21D).

Nach Induktion mit 0,2% L-Arabinose ist das Proteinlevel von EvgS<sub>NYAAD</sub>-FLAG in Wildtyp und  $\Delta$ *efp* Stämmen gleich. Im Gegensatz dazu ist das Proteinlevel der EvgS<sub>NYPPD</sub>-FLAG Variante in Abwesenheit von EF-P stark reduziert (basierend auf Proteinfärbung im Gel: 0,12±0,06-fach). Die Entfernung des „starken“ polyP Motivs führt demnach zu einer EF-P unabhängigen Translation. Bei derartig hohem Expressionslevel ist allerdings auch zu vermerken, dass die EvgS<sub>NYAAD</sub>-FLAG Variante in beiden Stämmen in niedrigerer Kopienzahl translatiert wird, als die native EvgS<sub>NYPPD</sub>-FLAG Variante in Wildtypzellen (basierend auf Proteinfärbung im Gel: 0,40±0,12-fach, bzw. 0,43±0,08-fach in  $\Delta$ *efp* Mutante, bzw. Wildtyp). Bei niedrigerem Expressionslevel und damit einer geringeren Akkumulation bereits translaterter Proteine, ergibt sich ein leicht abweichendes Bild. Hier wird deutlich, dass EvgS<sub>NYAAD</sub> in Abwesenheit von EF-P in geringerer Kopienzahl auftritt, als in Wildtypzellen (basierend auf Proteinfärbung im Gel: 0,16±0,05-fach statt 0,86±0,21 bzw. 0,10±0,07-fach statt 0,88±0,25 bei 0,00002%, bzw. 0,0% L-Arabinose). Grund dafür sind die verbliebenen, ebenfalls putativ EF-P abhängigen drei polyP-Motive (TSPPL, PYSPPY und AIPPS). Eine Anhebung des Translationslevels in Folge der NYPPD/NYAAD Substitution konnte in keinem der Fälle beobachtet werden.

### 3.2.1.3 Einfluss des putativ „starken“ NYPPD Motivs auf die Proteinfunktionalität von EvgS

Um zu klären, welche Rolle das „starke“ NYPPD Motiv auf die Funktionalität von EvgS hat, wurde das Motiv markerlos auf chromosomaler Ebene durch NYAAD substituiert und der resultierende *E. coli* K-12 MG1655 EvgS<sub>NYAAD</sub> Stamm phänotypisch charakterisiert. Dabei wurde der Fokus auf die durch das Zweikomponentensystem EvgAS regulierte Antwort auf moderaten Säurestress gelegt.

Der zur Anwendung gekommene Säurestress-Assay ist ein gut beschriebenes Testsystem zur Bewertung von EvgS-Mutanten<sup>271, 169, 161</sup>. Zellen wurden dafür bis in die frühe exponentielle Wachstumsphase ( $OD_{600} \sim 0,2$ ) in LB-Medium angezogen und der pH-Wert bei Start des Experimentes ( $t_0$ ) auf 3,0 eingestellt. Die über die Zeit verbleibenden koloniebildenden Individuen (CFU, *colony forming units*) spiegeln die Säureresistenz der Zellen wieder. Die Zahl der koloniebildenden *E. coli* K-12 MG1655 EvgS<sub>NYAAD</sub> Zellen war über einen Zeitraum von 5 Stunden vergleichbar mit der einer  $\Delta$ *evgS* Mutante. Wildtypzellen mit nativem EvgS wiesen einen signifikant höheren Anteil an stressresistenten Zellen auf (Abb. 21G). Ergänzend wurde die Fähigkeit der Säureadaptation des EvgS<sub>NYAAD</sub> exprimierenden Stammes getestet. Dafür wurden Zellen in LB-Medium mit einem pH-Wert von 5,0 vorkultiviert, bevor der pH auf 3,0 eingestellt wurde. Die Zahl der überlebenden und koloniebildenden Zellen wurde nach drei Stunden Inkubationszeit ausgezählt (Abb. 21F). Auch hier bestätigte sich, dass der NYPPD/NYAAD Austausch in EvgS denselben Phänotyp induziert, wie es bei einer  $\Delta$ *evgS* Mutante der Fall ist.

Zellen des Wildtyp-Stammes sind signifikant resistenter gegenüber der getesteten Stressbedingung und bei Versuchsende wurden etwa 6-mal mehr Kolonien gezählt. Daraus lässt sich schlussfolgern, dass  $EvgS_{NYAAD}$  nicht funktionell ist.

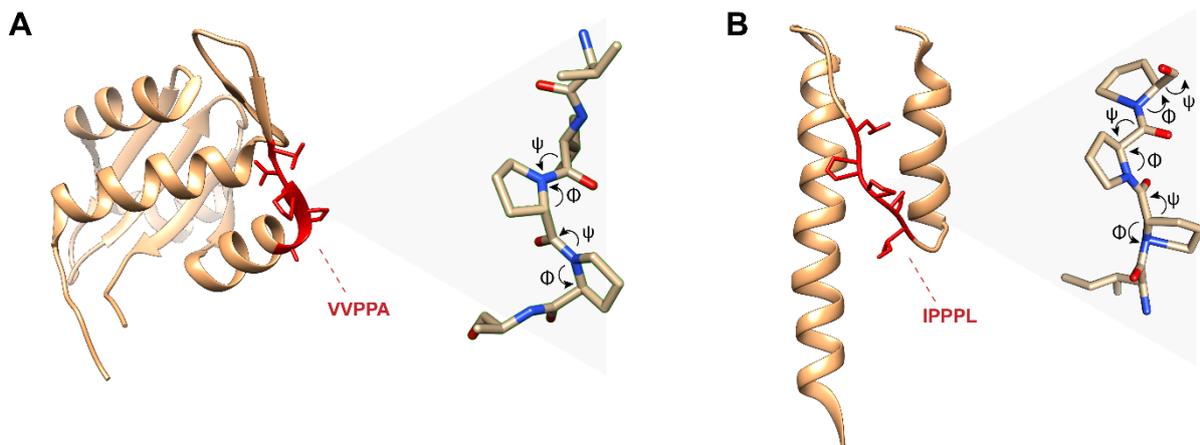
Darüberhinaus zeigte eine qPCR Analyse, dass NYPPD auch für die Expression der durch *EvgA* regulierten Gene *gadE*, *emrK* und *ydeP* bei moderatem Säurestress von Bedeutung ist (Abb. 21H). Die Expressionslevel von *gadE* und *emrK* sind auf ca. 18%, bzw. 25% des Wildtyplevels reduziert und vergleichbar mit denen der  $\Delta evgS$  Mutante.

### 3.2.4 Die Bedeutung von IPPPL für die Histidinkinase EnvZ

EnvZ trägt zwei polyP Motive unterschiedlicher „Stärke“: das in der periplasmatischen Domäne lokalisierte „moderate“ VVPPA (polyP<sub>1</sub>) und das „starke“ IPPPL (polyP<sub>2</sub>) in der HAMP Domäne.

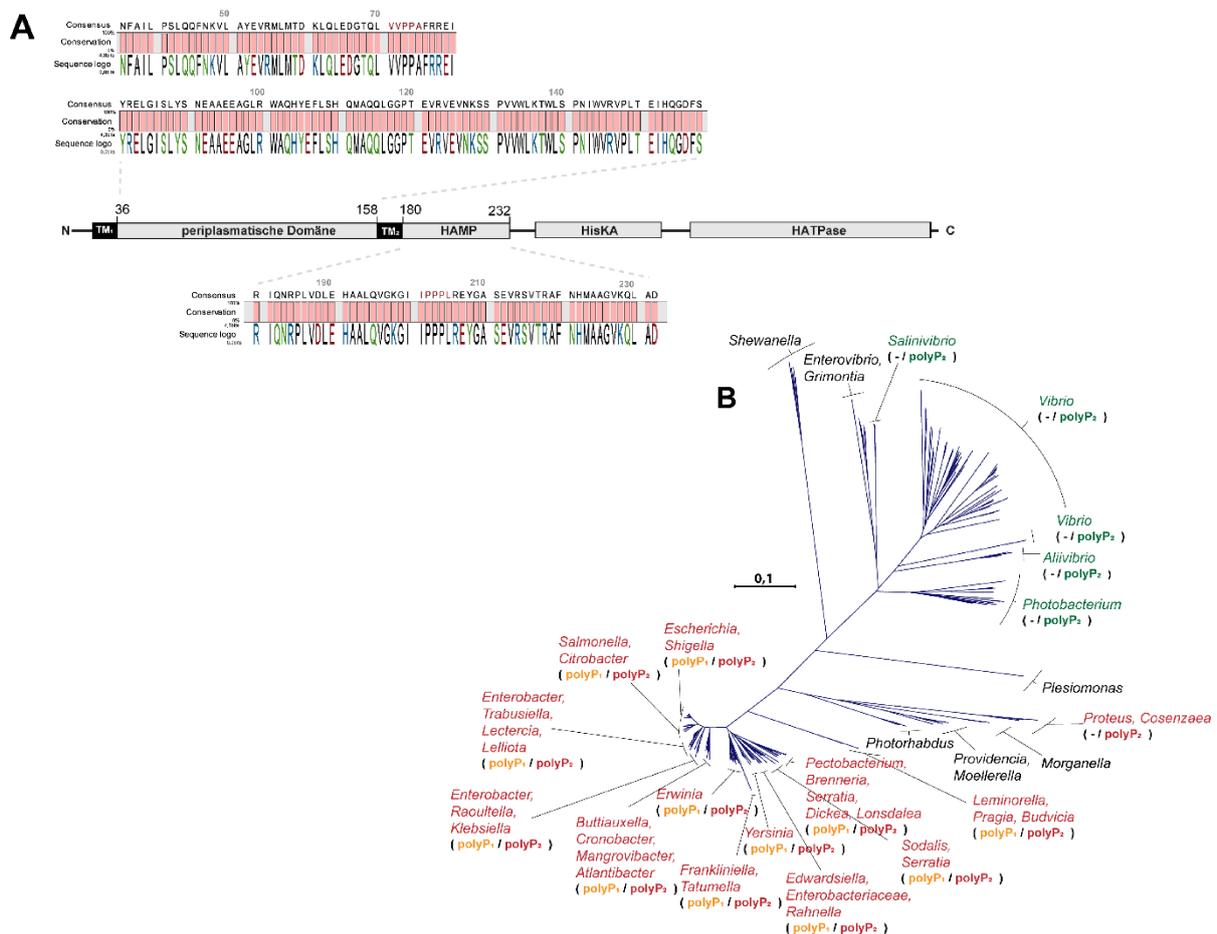
#### 3.2.1.1 Position und Konservierung der polyP Motive in EnvZ

Die periplasmatische Domäne wurde bereits kristallisiert (PDB 5XGA)<sup>330</sup> und die 3D Proteinstruktur zeigt, dass VVPPA nach außen orientiert positioniert ist und ein  $\beta$ -Faltblatt von einer  $\alpha$ -Helix trennt (Abb. 22A). Die EnvZ HAMP Domäne wurde mit Hilfe des Tools Phyre2 modelliert (Abb. 22B). Die daraus resultierende Strukturvorhersage ergab, dass IPPPL in einer unstrukturierten Verbindungssequenz zwischen den beiden für HAMP Domänen charakteristischen antiparallelen  $\alpha$ -Helices liegt.



**Abb. 22:** Lokalisation der polyP Motive der Histidinkinase EnvZ aus *E. coli* K-12 MG1655. (C) Kristallstruktur (PDB 5XGA, aa 36 - 125)<sup>330</sup> der periplasmatischen EnvZ Domäne aus *E. coli* K-12 und Markierung des darin enthaltenen VVPPA Motivs. (D) Modell der 3D Struktur der cytoplasmatischen HAMP Domäne (aa 183 – 238) von EnvZ aus *E. coli* K-12 und Markierung des darin enthaltenen IPPPL Motivs.

In Homologen aus 43 repräsentativen *E. coli* Stämmen sind beide Motive unverändert, wobei insgesamt nur 1,56% der EnvZ Aminosäuren nicht zu 100% konserviert sind (Abb. 23A). Ein Alignment von EnvZ aus 793 Orthologen (Tab. S5) mit mindestens 44% Identität und höchstens 10% Sequenzlängenabweichung ergab, dass IPPPL stärker konserviert ist, als VVPPA (Abb. 23B). In insgesamt 32 Gattungen liegt in der zu VVPPA analogen Position ein polyP Motiv vor (mit leichter Abweichung von VVPPA: wie EVPPA, AVPPA und PVPPA) – unter anderem in *Escherichia*, *Salmonella*, *Enterobacter*, *Klebsiella*, *Cronibacter*, *Erwinia*, *Edwardsiella*, *Yersinia*, *Sodalis*, *Brenneria* und *Budvicia*. In allen EnvZ Orthologen dieser Gattungen ist zugleich auch ein zweites polyP Motiv in der HAMP Domäne vorzufinden. Die Gattungen *Vibrio*, *Aliivibrio*, *Salinivibrio*, *Photobacterium*, *Proteus* und *Cosenzae* tragen dagegen nur ein zur Position von IPPPL analoges polyP Motiv. Die „Stärke“ dieses XPPPX Motivs ist dabei variabel, so dass die Erhaltung der Proline selbst vermutlich eine wichtigere Rolle spielt, als deren translationsverzögernder Effekt. Der phylogenetische Stammbaum lässt offen, ob die XPPPX Motive auf einen gemeinsamen Vorfahren zurückgehen, oder ob es sich um eine unabhängige Entwicklung handelt.



**Abb. 23:** Konservierung der polyP Motive der Histidinkinase EnvZ aus *E. coli* K-12 MG1655. (A) Schematische Darstellung der Domänen von EnvZ, sowie der darin enthaltenen polyP Motive VVPPA und IPPPL. Alignment von orthologen Proteinsequenzen aus 43 repräsentativen *E. coli* Stämmen und Darstellung der periplasmatischen und der HAMP Domäne (vgl. Tab. 10). (B) Rekonstruktion eines phylogenetischen Proteinbaumes aus Orthologen mit mindestens 44% Sequenzidentität und 90% Sequenzlänge zu EnvZ aus *E. coli* K-12 MG1655 zur Nachverfolgung der Evolution der EnvZ polyP Motive (polyP<sub>1</sub> = VVPPA, polyP<sub>2</sub> = IPPPL). Die farbige Hinterlegung der polyP Motive beschreibt deren Klassifizierung. Grün = „schwach“, orange = „moderat“, rot = „stark“.

### 3.2.1.2 Einfluss des putativ „starken“ IPPPL Motivs auf Lokalisation und Kopienzahl von EnvZ

Eine indirekte Bestimmung des Translationslevels eines plasmidbasierten und unter der Kontrolle des *cadC* Promotors stehenden EnvZ-LacZ Reporterkonstruktes durch Messung der  $\beta$ -Galactosidaseaktivität von transformierten Wildtyp *E. coli* BW25113 und BW25113  $\Delta$ *efp* Zellen ergab, dass das Proteinlevel von EnvZ in Abwesenheit von EF-P auf etwa 75% reduziert ist (Abb 24A). Damit konnte die vorhergesagte EF-P Abhängigkeit der EnvZ Translation bestätigt werden.

Um zu untersuchen, welche Rolle das „starke“ IPPPL Motiv auf das Translationslevel und die Lokalisation von EnvZ hat, wurde das Motiv auf chromosomaler Ebene durch IAAAL ersetzt. Die relative Kopienzahl von EnvZ wurde mittels Western Blot und spezifischen  $\alpha$ -EnvZ- Antikörpern quantifiziert (Abb. 24C). Das Proteinlevel der EnvZ<sub>IAAAL</sub> Variante unterschied sich nicht von dem des nativen EnvZ des Wildtypstamms. Die EnvZ<sub>IAAAL</sub> Variante konnte ebenso wie das native EnvZ nach Fraktionierung in der Membranfraktion festgestellt werden. Lösliche Derivate der EnvZ<sub>IAAAL</sub> Variante können ausgeschlossen werden. Die Proteinbande der EnvZ<sub>IAAAL</sub> Variante läuft in der SDS-PAGE etwa 1.250 Da niedriger, als die des Wildtyp EnvZ<sub>IPPPL</sub>. Vermutlich spiegelt dies ein aus dem fehlenden polyP Motiv resultierendes, verändertes Laufverhalten des Proteins im Polyacrylamidgel wider.

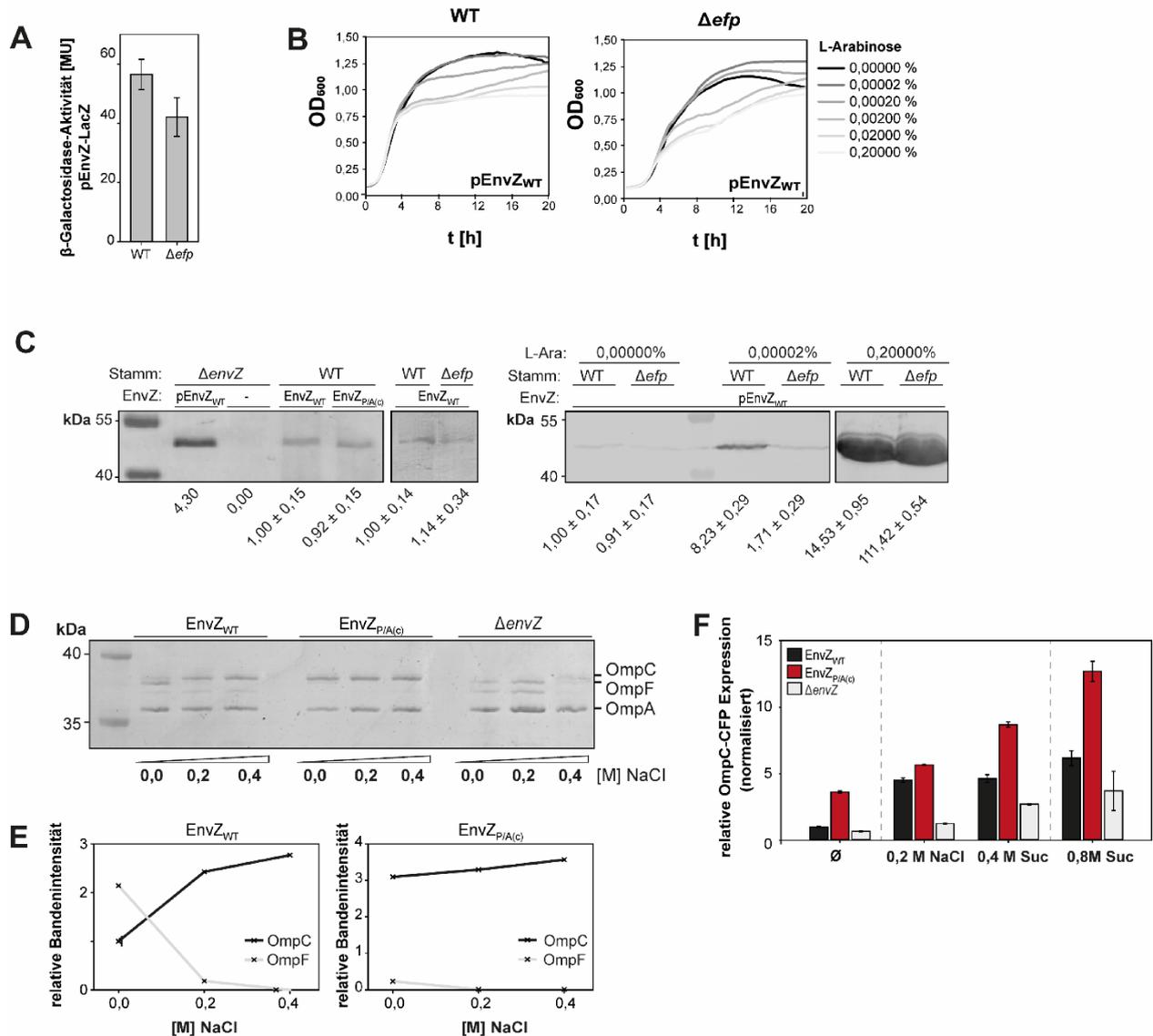
Eine EF-P Abhängigkeit des Proteinlevels der Wildtyp EnvZ<sub>IPPPL</sub> Variante wurde erst nach plasmidbasierter Expression von *envZ* unter der Kontrolle des *P<sub>BAD</sub>* Promotors sichtbar. Je höher die Expression, desto deutlicher der Unterschied zwischen Wildtyp und  $\Delta$ *efp* Zellen. Allerdings beeinträchtigt eine erhöhte Genexpression, wie im Fall von *evgS* und *dgcC* die Teilungsrate von Wildtyp und  $\Delta$ *efp* *E. coli* Zellen gleichermaßen (Abb. 24B).

### 3.2.1.3 Einfluss des putativ „starken“ IPPPL Motivs auf die Proteinfunktionalität von EnvZ

Um die Bedeutung des IPPPL Motivs für die Signalaktivität von EnvZ zu untersuchen, wurde die Fähigkeit einer *E. coli* K-12 MG1655 Mutante, in welcher das Motiv in EnvZ auf chromosomaler Ebene durch IAAAL ausgetauscht wurde, auf osmotischen Stress durch die OmpR vermittelte Regulation der Expression von *ompC* und *ompF* zu reagieren, beurteilt.

In einem ersten Ansatz wurden die Porine von Wildtyp und isogener EnvZ<sub>IAAAL</sub>, sowie  $\Delta$ *envZ* Mutante nach Exposition mit steigenden NaCl-Konzentrationen isoliert und anschließend in einem Coomassieblau-gefärbten SDS-Harnstoffgel detektiert und quantifiziert (Abb. 24D und E). Wie erwartet steigen bzw. sinken die OmpC und OmpF Proteinlevel in Wildtypzellen als Reaktion der steigenden Mediumosmolarität. Im Gegensatz dazu sind in der  $\Delta$ *envZ* Mutante stressunabhängige, konstant reduzierte OmpC Level zu verzeichnen. Die EnvZ<sub>IAAAL</sub> Variante zeigt unter Normalbedingungen hingegen hohen OmpC Leveln niedrige OmpF Level. Mit steigender Osmolarität findet nur ein minimaler Anstieg, bzw. Abfall der OmpC und OmpF Level statt. Daraus lässt sich schließen, dass die Kinaseaktivität der EnvZ<sub>IAAAL</sub> Variante im „ON“ Status fixiert ist.

In einem ergänzenden Versuchsansatz wurde die Fluoreszenz des Reporterstammes EPB273a mit verschiedenen chromosomal kodierten EnvZ-Varianten, bzw. mit *envZ* Deletion untersucht (Abb. 24F).



**Abb. 24:** Beurteilung der Auswirkungen einer EnvZ IPPL/IAAL Substitution (WT oder P/A) auf Translationslevel und Proteinfunktionalität. (A) Quantifizierung des Translationslevels eines plasmidkodierten EnvZ-LacZ Reporters in Abhängigkeit von EF-P mittels  $\beta$ -Galactosidaseassay. (B) Wachstumskurven zur Beurteilung des Effekts verschiedener Expressionslevel von EnvZ-FLAG auf die Zellteilungsrate in Wildtyp und  $\Delta$ efp Hintergrund. (C) Western-Blot zur relativen Quantifizierung der Proteinlevel von chromosomal kodiertem EnvZ<sub>WT</sub> und EnvZ<sub>P/A</sub> oder plasmidkodiertem, unter der Kontrolle des  $P_{BAD}$  stehenden EnvZ-FLAG in Abhängigkeit von EF-P. (D) Reinigung und Visualisierung von Proteinen der äußeren Membran auf einem Coomassieblau-gefärbten Harnstoffgel. Vergleich der Proteinlevel von *E. coli* K-12 Wildtyp und isogener EnvZ<sub>IAAAL</sub> bzw.  $\Delta$ envZ Mutante nach Exposition mit verschiedenen hohen NaCl-Konzentrationen in M9 Medium. (E) Quantifizierung der Bandenintensitäten aus (D). (F) Auf OmpC-CFP basierender Reporterassay zur Untersuchung der Veränderung des OmpC Expressionslevels als Reaktion auf Osmostress. Vergleich der Fluoreszenz von EPB273a Stämmen mit chromosomal kodierter EnvZ<sub>IPPL</sub>, bzw. EnvZ<sub>IAAAL</sub> Variante oder envZ Deletion. Der Zusatz von 0,2 M NaCl oder 0,4 M Saccharose entspricht einem Anstieg der Medium-Osmolalität von 0,200 auf 0,460 Osmol/kg, sowie 1,080 Osmol/kg nach Zugabe von 0,8 M Saccharose.

Material und Methoden: (A) pEnvZ-LacZ = V4, WT = St4,  $\Delta$ efp = St6. (B) pEnvZ<sub>WT</sub> = V20, WT = St4,  $\Delta$ efp = St6. (C) EnvZ<sub>WT</sub> = St31, EnvZ<sub>P/A</sub> = St33,  $\Delta$ envZ = St32,  $\Delta$ efp = St6 - Überexpression envZ: pEnvZ<sub>WT</sub> = V20 in WT = St4 oder  $\Delta$ efp = St6 St32 - Kontrolle: pEnvZ<sub>WT</sub> = V25 in  $\Delta$ envZ = St32. (D+E) EnvZ<sub>WT</sub> = St31, EnvZ<sub>P/A(c)</sub> = St33,  $\Delta$ envZ = St32. (F) EnvZ<sub>WT</sub> = St41, EnvZ<sub>P/A(c)</sub> = St43,  $\Delta$ envZ = St42.

Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen, bzw. repräsentieren eine von drei unabhängigen Versuchsdurchführungen. Die Ergebnisse wurden 2018 publiziert<sup>331</sup>.

Der Assay basiert auf einer chromosomalen *ompC-cfp* Fusion, durch welche das relative Expressionslevel von OmpC als Antwort auf unterschiedliche osmotische Stressbedingungen beobachtet werden kann. Wie bereits beschrieben<sup>144</sup>, führt die Kultivierung von Wildtypzellen mit 0,2 M NaCl zu erhöhten OmpC-CFP Leveln, welche durch eine weitere Erhöhung der Osmolarität noch gesteigert werden kann. Die  $\Delta envZ$  Mutante produziert nur geringe OmpC-Level und reagiert minimal auf osmotischen Stress. Im Gegensatz dazu sind die OmpC-CFP Level der EnvZ<sub>I<sup>AAA</sup></sub> Mutante im Vergleich zum Wildtyp unter allen getesteten Bedingungen signifikant erhöht.

Folglich kann geschlossen werden, dass IPPPL eine essentielle Rolle in der Signalperzeption und -weiterleitung durch EnvZ spielt.

### 3.3 Die Rolle von EnvZ polyP Motiven in der Proteininteraktion

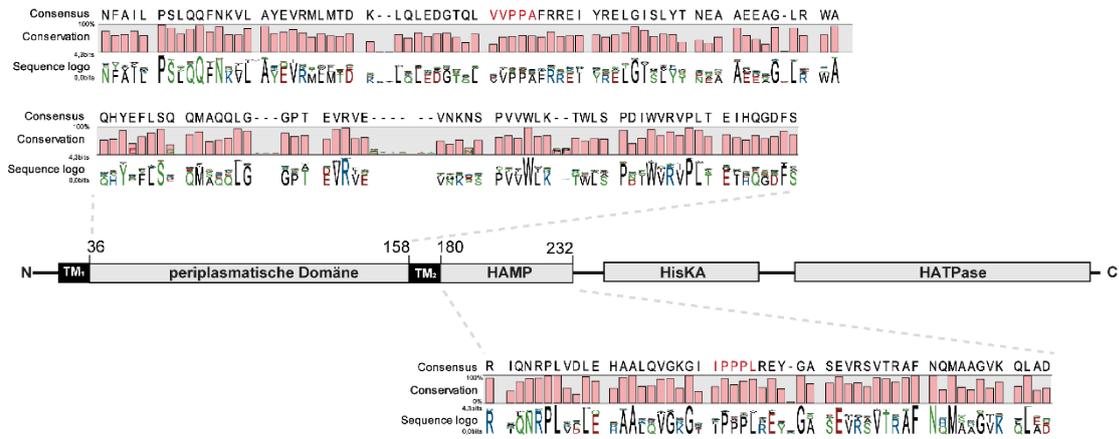
Ein Alignment von 62 EnvZ Orthologen mit mindestens 44% Sequenzidentität zu EnvZ aus *E. coli* K-12 MG1655 aus Organismen externer Knotenpunkte eines kürzlich veröffentlichten phylogenetischen Stammbaums von Gammaproteobakterien<sup>332</sup> gibt Aufschluss über das Maß der Konservierung der polyP Motive (Abb. 25A). Dabei zeigte sich, dass das periplasmatische polyP<sub>1</sub> Motiv (Val<sub>71</sub> 30%, Val<sub>72</sub> 56%, Pro<sub>73</sub> 59%, Pro<sub>74</sub> 62%, Ala<sub>75</sub> 60%) weniger stark konserviert ist, als das cytoplasmatische polyP<sub>2</sub> Motiv (Ile<sub>201</sub> 60%, Pro<sub>202</sub> 89%, Pro<sub>203</sub> 68%, Pro<sub>204</sub> 81%, Leu<sub>205</sub> 87%). Bei Letzterem ist das flankierende Isoleucin in vielen Fällen durch Phenylalanin ersetzt, was generell zu einer Abschwächung des vorhergesagten Translationsarrests führt.

#### 3.3.1 Die Beteiligung des IPPPL Motivs an der Homodimerisierung von EnvZ

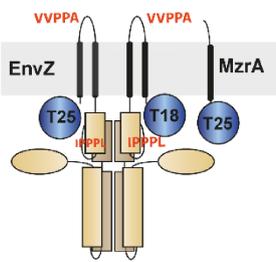
Die EnvZ HAMP Domäne bildet eine typische homodimere, sich aus vier  $\alpha$ -Helices zusammensetzende, parallele *coiled-coil* Struktur<sup>333–335</sup>. HAMP Domänen sind entscheidend für die Signaltransduktion von Rezeptorproteinen<sup>336</sup>. Ein polyP Motiv wie im Fall von EnvZ tritt in dieser Domäne nur in Einzelfällen auf<sup>337, 336</sup>. Der Frage, ob IPPPL ein stabilisierendes, strukturformendes Element oder durch eine direkte Beteiligung an der Proteininteraktion wichtig für die Dimerisierung von EnvZ ist, wurde mit Hilfe eines *bacterial-two-hybrid* Assays nachgegangen (Abb. 25C und D). Dabei wurde die Dimerisierungskapazität der geteilten Adenylatcyclase CyaA (T18 und T25) von *Bordetella pertussis*<sup>267, 338</sup> in Abhängigkeit der C-terminal und translational an die beiden Fragmente fusionierten EnvZ Varianten bestimmt (Abb. 25B). Die Reassemblierung hat eine funktionelle Adenylatcyclase zur Folge, welche die cAMP Signalkaskade induziert und damit die Transkription des *lac* Operons im mit verschiedenen Plasmidkombinationen transformierten *E. coli* Reporterstamm BTH101 aktiviert. Für die relative Quantifizierung des Expressionslevels wurde die  $\beta$ -Galactosidaseaktivität bestimmt. Alternativ dazu wurde die Blaufärbung von auf X-Gal/IPTG kultivierten Makrokolonien verglichen.

Als Positivkontrolle fungierte die Dimerisierung des Transkriptionsfaktors GCN4 (Zip)<sup>339</sup> aus *Saccharomyces cerevisiae* und als Negativkontrolle wurde gezeigt, dass T18 und T25 allein nicht interagieren. Eine Kombination von zwei EnvZ Wildtypproteinen führte zu hohen  $\beta$ -Galactosidasewerten, womit eine Dimerisierung des nativen Proteins bestätigt wurde. Eine Substitution des IPPPL Motivs durch IAAAL in einem der beiden EnvZ Konstrukte führt hingegen zu signifikant kleineren Werten. Ein dramatischer Abfall der Dimerisierung ist im Fall eines in beiden Interaktionspartnern ausgetauschten IPPPL Motivs zu verzeichnen. Ein Austausch des periplasmatischen VVPPA Motifs in VVAAA hat hingegen kaum einen Effekt auf die Dimerisierung von EnvZ. Die Beobachtungen gelten auch bei einem Austausch der Motive in beiden Interaktionspartnern (Abb. 25E und F).

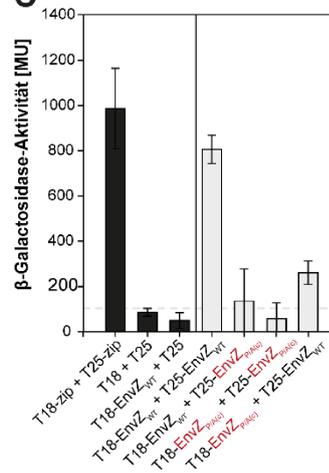
**A**



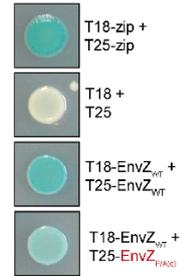
**B**



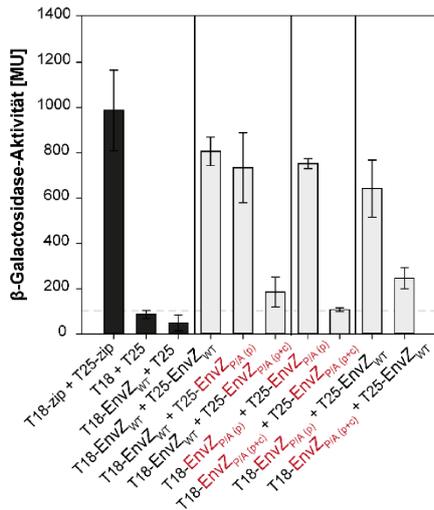
**C**



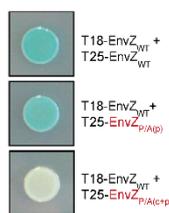
**D**



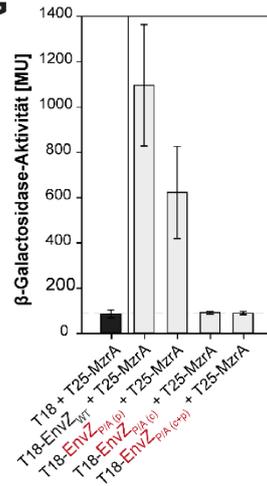
**E**



**F**



**G**



**H**

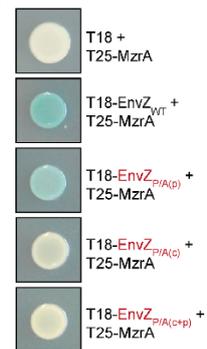


Abb. 25: Beschreibung auf folgender Seite.

**Abb. 25:** Die Rolle von periplasmatischem VVPPA (polyP<sub>p</sub>) und cytoplasmatischem IPPPL (polyP<sub>c</sub>) Motiv der *E. coli* K-12 MG1655 Histidinkinase EnvZ für Homodimerisierung und Interaktion mit dem Modulatorprotein MzrA. P/A = Prolin-Alanin-Substitution innerhalb des betreffenden polyP Motivs. (A) Aminosäurekonservierung der periplasmatischen und HAMP Domäne aus orthologen EnvZ Sequenzen aus 63 verschiedenen Vertretern von  $\gamma$ -Proteobakterien, die dieses Protein aufweisen. Auswahl basierend auf einem kürzlich veröffentlichten *tree of life* <sup>332</sup>. (B) Prinzip des *bacterial-two-hybrid* Assays zur Untersuchung der EnvZ-EnvZ Dimerisierung oder EnvZ-MzrA Interaktion. (C -H) *Bacterial-two-hybrid* Assay zur Determinierung der EnvZ Dimerisierung, bzw. der EnvZ-MzrA Interaktion in Abhängigkeit der beiden polyP Motive mittels  $\beta$ -Galactosidaseassay oder Vergleich der Blaufärbung von auf IPTG/X-Gal gewachsenen Makrokolonien.

Material und Methoden: T18 = V45, T25 = V39, T18-zip = V50, T25-zip = V51, T25-MzrA = V40, T18-EnvZ<sub>WT</sub> = V46, T18-EnvZ<sub>P/A (p)</sub> = V47, T18-EnvZ<sub>P/A (c)</sub> = V48, T18-EnvZ<sub>P/A (p+c)</sub> = V49, T25-EnvZ<sub>WT</sub> = V41, T25-EnvZ<sub>P/A (p)</sub> = V42, T25-EnvZ<sub>P/A (c)</sub> = V43, T25-EnvZ<sub>P/A (p+c)</sub> = V44, in: *E. coli* BTH101 (St53).

Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen, bzw. repräsentieren eine von drei unabhängigen Versuchsdurchführungen. Die Ergebnisse von wurden 2018 publiziert <sup>331</sup>.

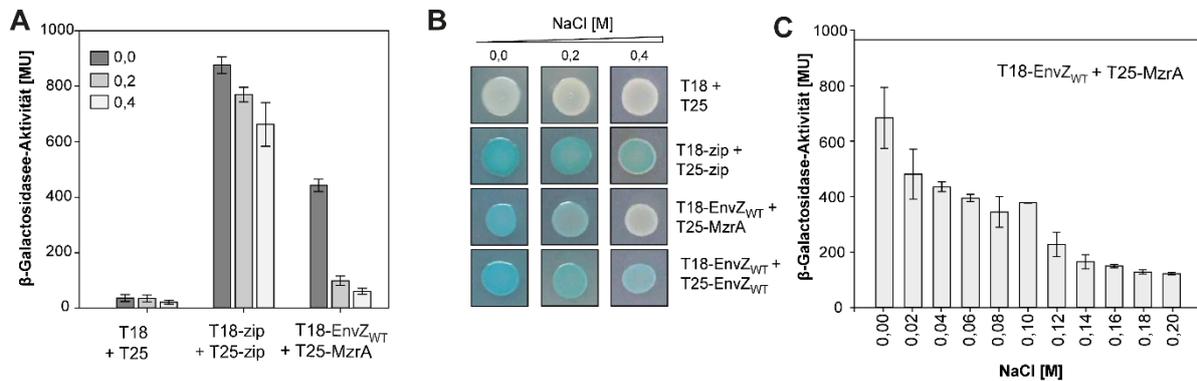
### 3.3.2 Beteiligung von VVPPA und IPPPL an der Bindung des Modulatorproteins MzrA

Aufgrund seiner exponierten Lage im Periplasma wurde das EnvZ VVPPA Motiv auf seine Beteiligung an der Interaktion zwischen EnvZ und dem membranintegrierten Modulator MzrA hin untersucht. Es ist bereits bekannt, dass zwei Aminosäuresubstitutionen in der periplasmatischen Domäne von MzrA ausreichen, um dessen Bindeaffinität gegenüber EnvZ signifikant herabzusetzen <sup>153</sup>. Ein *bacterial-two-hybrid* Assay mit C-terminal an das Adenylatcyclase-Fragment T25 fusioniertem MzrA und verschiedenen EnvZ Varianten unterstreicht die Wichtigkeit des EnvZ VVPPA Motivs (Abb. 25G und H). MzrA und das native EnvZ interagieren wie erwartet und führen den hohen  $\beta$ -Galactosidaseaktivitäten des Reporterstammes BTH101 nach zu urteilen zu einer stabilen Komplementation der Adenylatcyclase-Fragmente T18 und T25. Eine Substitution des VVPPA Motivs durch VVAAA führt dagegen zu einem Abfall der  $\beta$ -Galactosidaseaktivitäten auf 60%, was auf eine reduzierte Affinität der EnvZ<sub>VVAAA</sub> Variante gegenüber MzrA hindeutet. Die IPPPL/IAAAL Substitution des in der HAMP Domäne lokalisierten polyP Motivs führt zu einer kompletten Aufhebung der EnvZ-MzrA Interaktion und  $\beta$ -Galactosidasewerten, die nur etwa 10% der Kontrolle entsprechen. Es bleibt zu diskutieren, ob die 13 cytoplasmatischen MzrA Aminosäuren, welche für eine direkte Interaktion mit der EnvZ HAMP Domäne in Frage kommen, eine Rolle spielen, oder ob eine Dimerisierung von EnvZ Voraussetzung für die Interaktion mit MzrA ist.

### 3.3.3 Die Bindungsaffinität von MzrA zu EnvZ in Abhängigkeit zur Mediumosmolarität

Die Affinität von EnvZ und MzrA ist stark abhängig von der Osmolarität, wie eine weitere *bacterial-two-hybrid* Analyse mit den oben beschriebenen Komponenten zeigt (Abb. 26). Die Zugabe von bis zu 0,4 M NaCl senkt die Interaktion der Positivkontrolle T18-Zip/T25-Zip, sowie die Dimerisierung von T18-EnvZ/T25-EnvZ nur leicht und mit linearem Charakter herab. Möglicherweise ist dies eine Folge der Verdrängung von Wassermolekülen auf den Proteinoberflächen und einer damit einhergehenden Veränderung der Proteinstabilität. Im Gegensatz dazu fällt

die T18-EnvZ/T25-MzrA Interaktion exponentiell ab. Bereits der Zusatz von 0,2 M NaCl hat eine Reduzierung der  $\beta$ -Galactosidaseaktivität auf 22% zur Folge, während eine weitere Erhöhung der Osmolarität durch 0,4 M NaCl zu Werten von 13,6% des Normallevels führt (Abb. 26A und B). Um den fraglichen Bereich zwischen 0,0 M und 0,2 M NaCl besser aufzulösen, wurden in einem folgenden Versuchsansatz 11 Konditionen mit um je 0,02 M abweichenden NaCl Konzentrationen getestet (Abb. 26C). Ein entscheidender Abfall der Bindestärke zwischen MzrA und EnvZ findet demnach zwischen dem Zusatz von 0,0 M und 0,02 M NaCl statt. Dies lässt Rückschlüsse auf die Funktion von MzrA als Modulator der Aktivität von EnvZ zu. Anscheinend bindet MzrA unter Normalbedingungen an EnvZ, während die Bindung unter osmotischen Stressbedingungen aufgelöst wird, was der Funktionsweise eines Repressors entsprechen würde.



**Abb. 26:** *Bacterial-two-hybrid* Assay zur Analyse der EnvZ-MzrA Interaktion in Abhängigkeit zur Mediumosmolarität. (A + B) Vergleich der MzrA-EnvZ Interaktion mittels  $\beta$ -Galactosidaseassay und Vergleich der Blaufärbung von auf IPTG/X-Gal gewachsenen Makrokolonien. (C) Detaillierte Untersuchung der relevanten Mediumosmolaritäten.

Material und Methoden: T18 = V45; T25 = V39, T18-zip = V50, T25-zip = V51, T25-MzrA = V40, T18-EnvZ<sub>WT</sub> = V46, in: *E. coli* BTH101 (St53).

Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen, bzw. repräsentieren eine von drei unabhängigen Versuchsdurchführungen.

### 3.3.4 Visualisierung von EnvZ und MzrA auf Einzelzellebene

Eine Analyse der Lokalisation von EnvZ und MzrA auf Einzelzellebene gibt neben Informationen zu Lage und Oligomerisierung der beiden Proteine auch Aufschluss darüber, inwiefern eine korrekte Membranintegration von EnvZ von einer zeitlichen Taktung dessen Translation, bzw. der Präsenz von EF-P, abhängt. Zwar konnte gezeigt werden, dass EnvZ auch in Abwesenheit von EF-P in der Membran verankert ist, unklar ist jedoch, ob dies kotranslational geschieht und ob ein definiertes Zeitfenster Voraussetzung für das Zusammenfinden von Interaktionspartnern ist. Ein ebenfalls mögliches Szenario ist, dass EnvZ in Abwesenheit von EF-P nur unvollständig und unter Verlust der für Proteininteraktionen relevanten Strukturen in die Membran integriert wird.

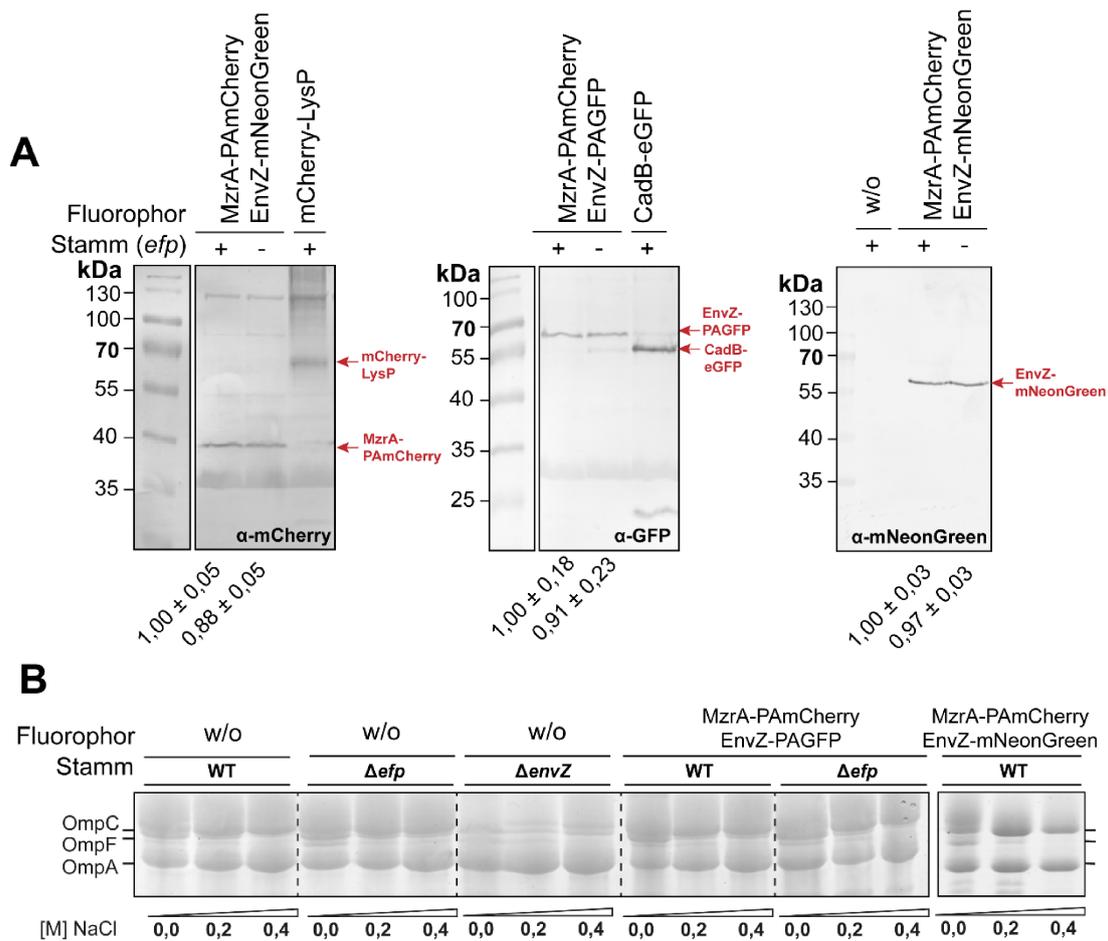
#### 3.3.4.1 Konstruktion und Charakterisierung von Stämmen mit translationaler MzrA-PAmCherry und EnvZ-PAGFP/mNeonGreen Fusion

Um EnvZ und MzrA im *E. coli* K-12 MG1655 Wildtyp und in einer  $\Delta efp$  Mutante auf Einzelzellebene hinsichtlich Lokalisierung und Clusterbildung zu untersuchen, wurden die beiden Proteine mittels chromosomal am nativen Locus angefügten Fluorophoren C-terminal markiert (Linkeraminosäuresequenz: GSGS). Obwohl der C-Terminus von MzrA nachweislich im periplasmatischen Raum liegt, zeigt ein von Gerken *et al.* (2010)<sup>153</sup> publiziertes Beispiel (MzrA-PhoA), dass die Voraussetzung für die Funktionalität von daran fusionierten Proteinen gegeben ist. PAmCherry weist im Gegensatz zu GFP Derivaten keine Cytosine auf und ist damit resistent gegenüber den oxydierenden Bedingungen des Periplasmas<sup>340</sup>. Eine effektive Translokation und nachfolgende Faltung findet nachweislich über verschiedene Zweige des Sec Transportweges statt<sup>341</sup>. Die hier gewählte Konstellation von cytoplasmatisch und periplasmatisch fusionierten Fluorophoren hat den Vorteil, dass die Fluoreszenzproteine sich nicht räumlich behindern. Die Verwendung der Fluoreszenzproteine PAGFP und PAmCherry erlaubt grundsätzlich eine photoaktivierte Lokalisationsmikroskopie (PALM) mit sehr hoher Auflösung<sup>275</sup>. Im Laufe der hier durchgeführten Experimente kristallisierte sich heraus, dass PAGFP aufgrund der Autofluoreszenz von *E. coli* in einem ähnlichen Wellenlängenbereich ungeeignet ist, weshalb auf das auch ohne vorherige Aktivierung fluoreszierende mNeonGreen zurückgegriffen wurde.

Mittels Westernblot-Analyse wurde verifiziert, dass die fusionierten Fluoreszenzproteine keine Auswirkung auf die relativen Proteinlevel von EnvZ und MzrA haben (Abb. 27A). EnvZ-PAGFP und EnvZ-mNeonGreen werden in Wildtyp *E. coli* Zellen und  $\Delta efp$  Zellen gleich stark translatiert. Dies entspricht der Beobachtung von nativen EnvZ Kopienzahlen. Das Proteinlevel von MzrA-PAmCherry ist in  $\Delta efp$  Zellen etwas kleiner, als in Wildtypzellen (~0,9-fach). Untersucht wurden Stämme, in welchen *mzrA* und *envZ* zugleich markiert sind. Als Kontrolle wurden die bereits charakterisierten Konstrukte mCherry-Lys und CadB-eGFP detektiert.

Die Funktionalität der Konstrukte wurde überprüft, indem die Fähigkeit der Stämme, auf steigende NaCl-Konzentrationen des Mediums mit einer Regulation der Expression der Porine *ompC* und *ompF* zu reagieren, getestet wurde (Abb. 27B). Nach Kultivierung der Stämme in entsprechend modifiziertem M9-Medium wurden die Proteine der äußeren Membran isoliert, mittels Harnstoff-SDS-PAGE aufgetrennt und mit Coomassie-Blau angefärbt. Als Referenz dienen *E. coli* BW25113 Wildtyp und  $\Delta efp$  Mutante. Beide Stämme zeigten dasselbe Expressionsmuster. Unter normalen Wachstumsbedingungen ist die Kopienzahl von OmpF höher als die von OmpC. Mit steigender Osmolarität (durch Zugabe von 0,2 und 0,4 M NaCl) sinkt das Level von OmpF und steigt

das Level von OmpC. Eine chromosomal kodierte *envZ-PAGFP*, bzw. *envZ-mNeonGreen* und *mzrA-PAmCherry* Fusion hat weder im Wildtyp, noch im  $\Delta efp$  Hintergrund einen Einfluss darauf.



**Abb. 27:** Verifizierung der Expression und Funktionalität von EnvZ-PAGFP, bzw. EnvZ-mNeonGreen und MzrA-PAmCherry. (A) Vergleich der Translationslevel der chromosomal kodierten Konstrukte mittels Western Blot mit spezifischen Antikörpern. Untersuchung von zuvor isolierten Membranfraktionen. (B) Untersuchung der Fähigkeit der Stämme, auf variierende Mediumosmolaritäten mit einer Veränderung der EnvZ/OmpR regulierten Expression der für die Porine OmpC und OmpF kodierenden Gene zu reagieren. Isolierung von Proteinen der äußeren Membran und Harnstoff-SDS-PAGE mit anschließender Coomassie-Blaufärbung.

Material und Methoden: WT = St4,  $\Delta efp$  = St6,  $\Delta envZ$  = St7, MzrA-PAmCherry/EnvZ-PAGFP/*efp*<sup>+</sup> = St22, MzrA-PAmCherry/EnvZ-PAGFP/*efp*<sup>-</sup> = St23, MzrA-PAmCherry/EnvZ-mNeonGreen/*efp*<sup>+</sup> = St26, MzrA-PAmCherry/EnvZ-mNeonGreen/*efp*<sup>-</sup> = St27, mCherry-LysP/*efp*<sup>+</sup> = St53, CadB-eGFP/*efp*<sup>+</sup> = St54.

Die gezeigten Resultate repräsentieren eine von drei unabhängigen Versuchsdurchführungen.

### 3.3.4.2 Fluoreszenzmikroskopische Visualisierung von MzrA-PAmcherry und EnvZ-PAGF/mNeonGreen

Bevor die Stämme mittels PALM untersucht wurden, wurde die Fluoreszenz der konstruierten Fusionsproteine durch Epifluoreszenzmikroskopie überprüft (Abb. 28). Dafür wurden *E. coli* Wildtyp und  $\Delta efp$  Zellen mit chromosomal kodierter MzrA-PAmCherry und EnvZ-mNeonGreen, bzw. EnvZ-PAGFP Fusion bis in die frühe exponentielle Phase kultiviert, anschließend fixiert und auf eine mit PBSG bedeckte Glasoberfläche immobilisiert.

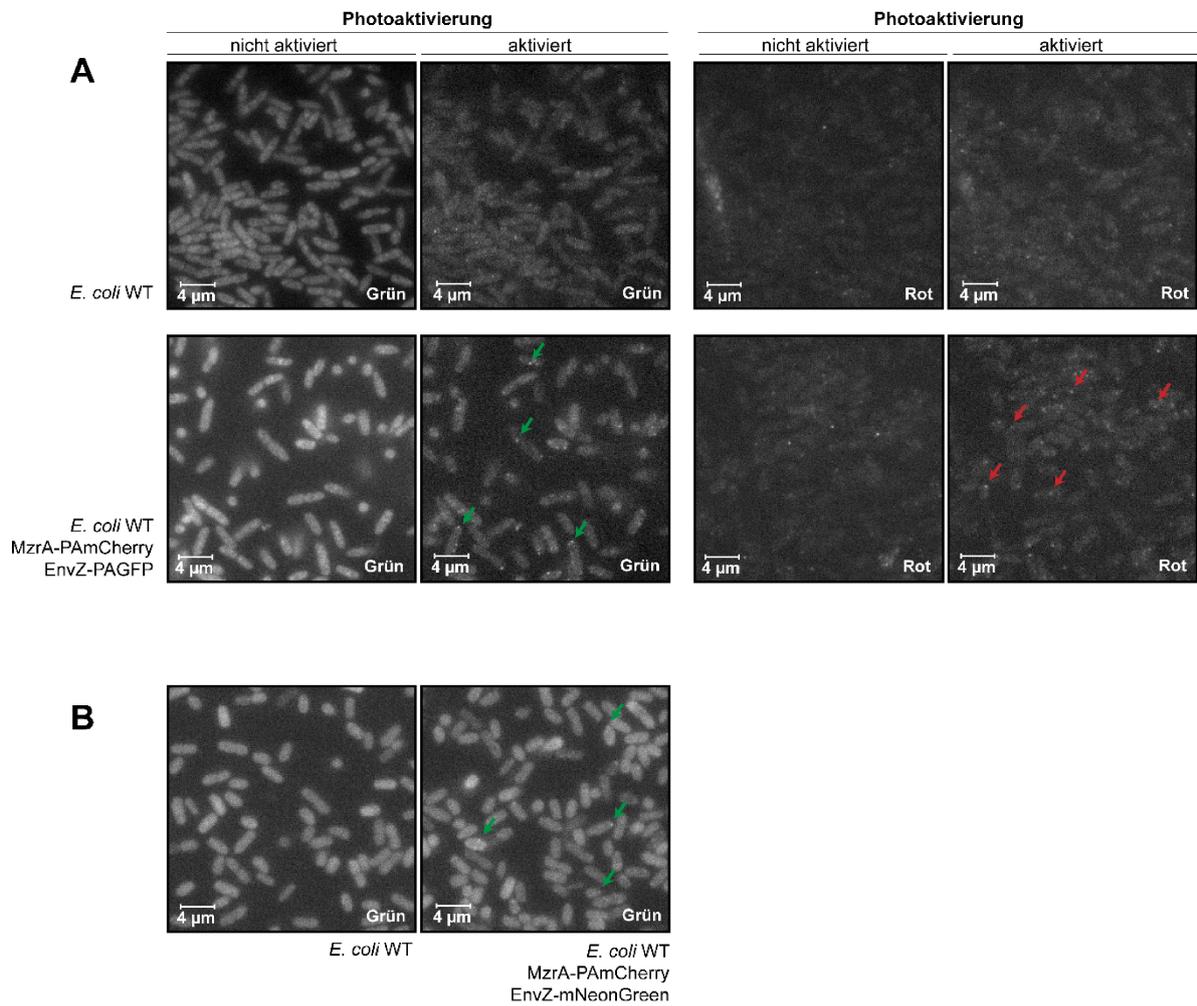
Durch Prüfung verschiedener Werte der Parameter zur Anregung und Detektion der photoaktivierbaren Fluorophore wurden geeignete Bildaufnahmebedingungen herausgearbeitet, um die Proteine zu detektieren. Als Positivkontrolle wurde das bereits bekannte Fluoreszenzverhalten der Proteine DivIVA-PAmCherry und DivIVA-PAGFP aus *B. subtilis*, welche am Teilungsseptum lokalisiert sind, herbeigezogen<sup>342</sup>.

Eine Aktivierung der Fluorophore wurde durch eine 45 sekündige Exposition mit Licht der Wellenlänge 400 nm simuliert. Direkt im Anschluss konnten in einigen der zentral im Bildausschnitt liegenden Zellen ein bis zwei grün, bzw. rot fluoreszierende Foci detektiert werden (Abb. 28A). Aufgrund der geringen Proteinkopienzahlen konnte das Signal nur bei sehr intensiver und langer Belichtung detektiert werden, welche gleichzeitig zu einem hohen Hintergrundsignal führte. Dabei ist das Signal von MzrA-PAmCherry deutlich stärker, als das von EnvZ-PAGFP, für welches längere Belichtungszeiten verwendet werden müssen. Der verwendete Wellenlängenbereich zur Anregung und Detektion von PAmCherry und PAGFP induziert hier auch im nicht fluoreszenzmarkierten Wildtyp unspezifische, punktförmig konzentrierte Signale. Diese erschweren eine eindeutige Identifizierung der Fluorophore, wobei der Hintergrund im Fall von PAGFP deutlich höher ist. Durch die Belichtung der Zellen mit einer Glühlampe (Farbtemperatur 2700°K) während der Kultivierung konnten unspezifische Signale geringfügig reduziert werden.

Im Vergleich zu EnvZ-PAGFP konnten für EnvZ-mNeonGreen deutlich höhere und spezifische Fluoreszenzintensitäten beobachtet werden (Abb. 28B). Allerdings limitiert mNeonGreen aufgrund seiner konstanten Fluoreszenz die durch PALM Mikroskopie maximal erreichbare Auflösung.

Um das Signal zur Detektion der Fluorophore zu amplifizieren, wurden ergänzend spezifische  $\alpha$ -mCherry und  $\alpha$ -mNeonGreen Nanobodies mit gekoppeltem fluoreszierendem Farbstoff verwendet ( $\alpha$ -GFP-Atto488 und  $\alpha$ -mCherry-Atto647N). Durch ihre geringe Größe sollten in der Theorie mehrere Nanobodies intrazellulär an ein Fluoreszenzprotein binden. Weder eine Vorbehandlung der Zellen mit Lysozym, die Verwendung von zuvor präparierten Sphäroblasten und *right-side-out* Vesikeln, noch das Einbringen der Nanobodies in intakte Zellen mittels Elektroporation führten zu einer reproduzierbaren und gleichmäßigen intrazellulären Verteilung der Nanobodies.

Obwohl die oben beschriebenen fluoreszenzmikroskopischen Untersuchungen keine eindeutigen Aussagen zur Lokalisation von MzrA-PAmCherry und EnvZ-PAGFP, bzw. EnvZ-mNeonGreen zulassen, konnte die Fluoreszenz der Fusionsproteine bestätigt werden, womit die Voraussetzung für eine Analyse mittels PALM gegeben ist.



**Abb. 28:** Überprüfung der Fluoreszenz von MzrA-PAmCherry und EnvZ-PAGFP, bzw. EnvZ-mNeonGreen Konstrukten mittels Epifluoreszenzmikroskopie (Delta Vision Elite). (A) Photoaktivierung der chromosomal kodierten Fluorophore MzrA-PAmCherry („rot“, roter Pfeil) und EnvZ-PAGFP („grün“, grüner Pfeil) in intakten *E. coli* BW25113 Stämmen. (B) Fluoreszenz von chromosomal kodiertem EnvZ-mNeonGreen („grün“, grüner Pfeil) in *E. coli* BW25113.

Material und Methoden: *E. coli* WT = St4, *E. coli* MzrA-PAmcherry/EnvZ-PAGFP = St22, *E. coli* MzrA-PAmcherry/EnvZ-mNeonGreen = St26.

### 3.3.4.3 Hochauflösende Lokalisation von MzrA-PAmcherry und EnvZ-PAGFP/-mNeonGreen

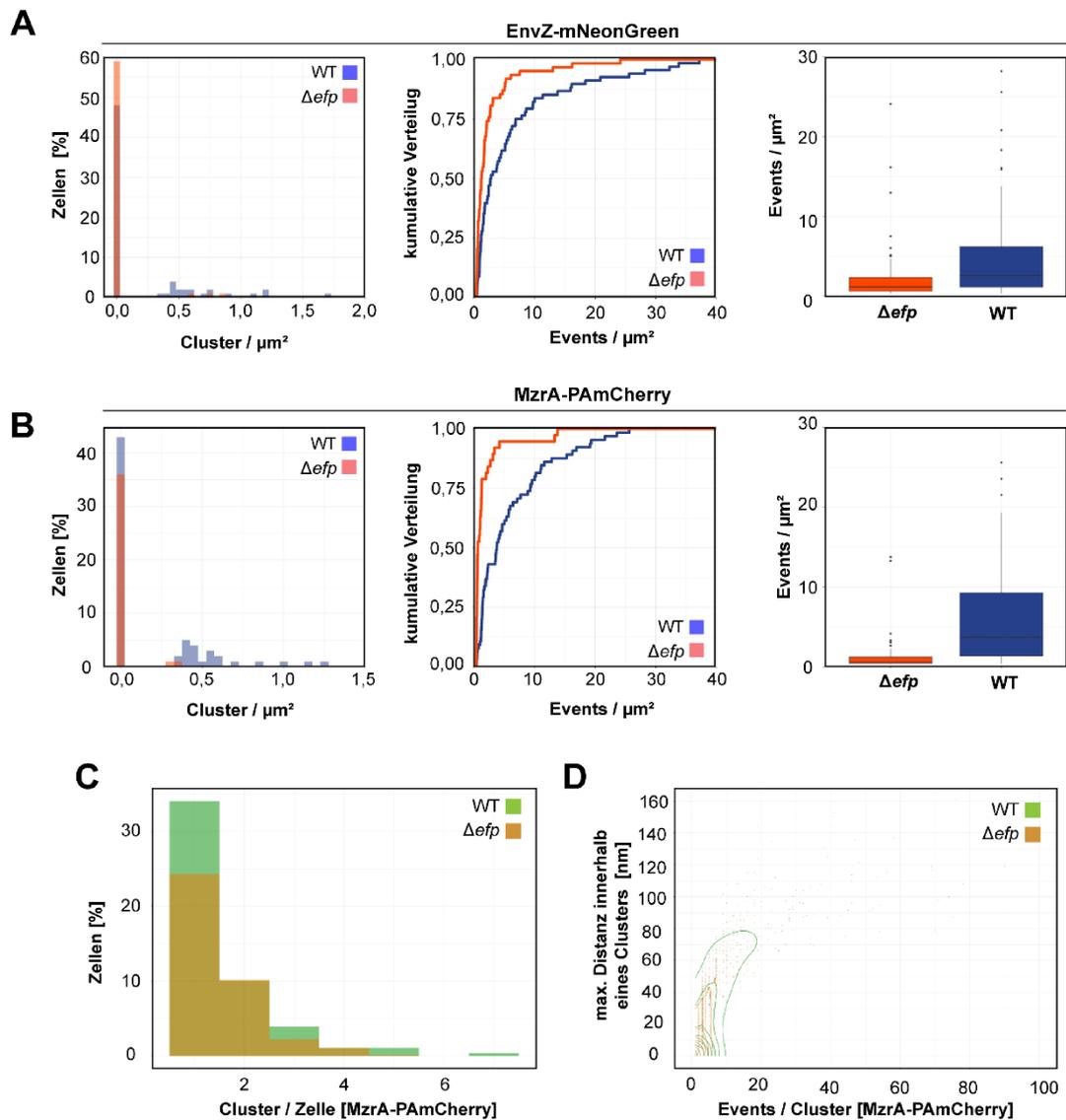
Die intrazelluläre Lokalisation von MzrA-PAmCherry und EnvZ-mNeonGreen wurde in Zusammenarbeit mit Giacomo Giacomelli (AG Prof. Dr. M. Bramkamp, Bereich Mikrobiologie, LMU München) durch PALM Mikroskopie untersucht. *E. coli* Wildtyp und  $\Delta efp$  Zellen mit zugleich chromosomal kodierter MzrA-PAmCherry und EnvZ-mNeonGreen Fusion wurden wie oben beschrieben bis in die frühe exponentielle Phase kultiviert, fixiert und immobilisiert.

MzrA-PAmCherry und EnvZ-mNeonGreen konnten sowohl in *E. coli* BW25113 Wildtyp, als auch  $\Delta efp$  Mutante identifiziert und lokalisiert werden. Damit wurde das nativ exprimierte EnvZ Vollängenprotein entgegen der gängigen Erwartung<sup>343</sup> zum ersten Mal auf Einzelzellebene visualisiert. Bis *dato* wurden verkürzte und plasmidkodierte EnvZ Derivate verwendet<sup>146</sup>. Im Fall von MzrA existieren bislang keinerlei Daten zur intrazellulären Lokalisation der einzelnen Moleküle, sodass die hier vorliegenden Ergebnisse erstmals einen Einblick gewähren.

#### 3.3.4.3.1 Größe und Lokalisation von MzrA-PAmcherry und EnvZ-mNeonGreen Proteincluster

Obwohl die konstante Fluoreszenz von mNeonGreen die Auflösung limitiert und somit keine Unterscheidung von Einzelmolekülen in Proteinclustern zulässt, konnten dennoch wichtige Informationen zur intrazellulären Verteilung und Lokalisation von EnvZ-mNeonGreen gesammelt werden. EnvZ-mNeonGreen formt mehrere Cluster pro Zelle, wobei sich deren Größe und Anzahl im Wildtyp und  $\Delta efp$  Hintergrund unterscheidet (Abb. 29A). Dabei sind die Cluster heterogen verteilt und ihr Vorkommen ist auf einen Teil der Zellen beschränkt. Der Anteil an Wildtypzellen, in welchem Cluster auftreten ist größer als der der  $\Delta efp$  Zellen. Während Wildtypzellen durchschnittlich 2,65 Fluoreszenzereignisse pro  $\mu\text{m}^2$  aufweisen, sind es in Abwesenheit von EF-P nur 1,12 Fluoreszenzereignisse pro  $\mu\text{m}^2$ . Da die per Western Blot bestimmten Werte zur relativen Kopienzahl von EnvZ-mNeonGreen keine Abweichungen in Abhängigkeit von EF-P aufweisen, muss für den Anteil an EnvZ-mNeonGreen Molekülen, welcher hier nicht als Fluoreszenzereignis erfasst wurde, eine veränderte Faltung und damit einhergehende Fluoreszenzeigenschaft von mNeonGreen in Abwesenheit von EF-P in Betracht gezogen werden.

Auch MzrA-PAmCherry formt mehrere Cluster pro Zelle. Nur 50% der Zellen weisen Cluster auf und auch hier ist deren Anzahl in Wildtypzellen höher, als in  $\Delta efp$  Zellen (Abb. 29B). Dabei wurde in Wildtypzellen eine deutlich höhere Dichte an Fluoreszenzereignissen festgestellt, als im  $\Delta efp$  Hintergrund (3,66 events/ $\mu\text{m}^2$  vs. 0,56 events/ $\mu\text{m}^2$ ). Auch die Anzahl der Fluoreszenzereignisse pro Zelle ist im Wildtyp 29,95% größer, als in der  $\Delta efp$  Mutante (13,15 vs. 10,12 Fluoreszenzereignisse/Zelle). Da annähernd davon ausgegangen werden kann, dass jedes Molekül aufgrund des Ausbleicheffekts nur einmal detektiert wird, kann direkt auf die Anzahl der MzrA-PAmCherry Moleküle rückgeschlossen werden. Der Unterschied zwischen den beiden Stämmen entspricht der Westernblot-basierten relativen Quantifizierung. Ein Großteil der Zellen trägt ein bis zwei Cluster, die maximale Anzahl liegt bei sieben. Die Cluster setzen sich im Mittel aus 10 Fluoreszenzereignissen zusammen und haben einen Durchmesser von etwa 40 nm (Abb. 29C und D).



**Abb 29:** Detektion von MzrA-PAmCherry und EnvZ-mNeonGreen in doppelmarkierten *E. coli* BW25113 Wildtyp und  $\Delta efp$  Stamm mittels PALM. (A) Anzahl und Verteilung der Fluoreszenzcluster pro Zelle, kumulative Verteilung und mittlere Dichte der mit EnvZ-mNeonGreen assoziierten Fluoreszenzereignisse (*events*). (B) Anzahl und Verteilung der Fluoreszenzcluster pro Zelle, kumulative Verteilung und mittlere Dichte der mit MzrA-PAmCherry assoziierten Fluoreszenzereignisse (*events*). (C) MzrA-PAmCherry Cluster pro Zelle. (D) Fluoreszenzereignisse (*events*) pro MzrA-PAmCherry Cluster.

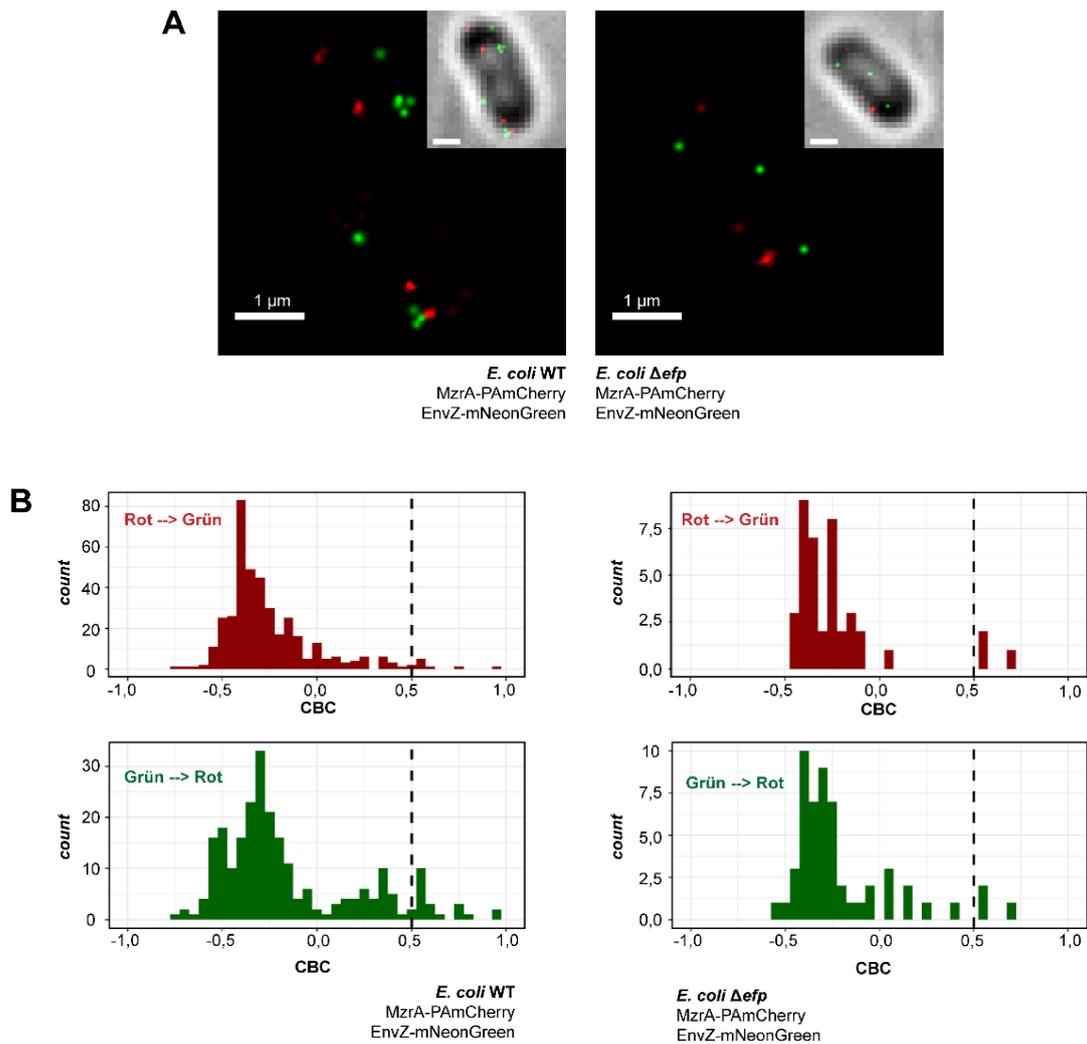
Material und Methoden: *E. coli* BW25113 Wildtyp MzrA-PAmCherry/EnvZ-mNeonGreen („WT“) = St26, *E. coli* BW25113  $\Delta efp$  MzrA-PAmCherry/EnvZ-mNeonGreen („ $\Delta efp$ “) = St27.

Analyse und Darstellung durch Giacomo Giacomelli (AG Prof. Dr. M. Bramkamp, Bereich Mikrobiologie, LMU München). Die Daten beruhen auf der Auswertung der Fluoreszenz von ca. 500 Zellen.

### 3.3.4.3.2 Kolokalisation von MzrA-PAmCherry und EnvZ-mNeonGreen

Cluster von EnvZ-mNeonGreen und MzrA-PAmCherry sind zufällig in der Zellmembran verteilt. Es kann weder im Wildtyp, noch im  $\Delta\text{efp}$  Hintergrund eine Akkumulation in bestimmten Regionen festgestellt werden (Abb. 30A). Die vorangegangenen *in vivo* Versuche zeigten, dass EnvZ und MzrA im Wildtyp unter den hier angewendeten Wachstumsbedingungen interagieren. Durch die Berechnung der relativen Position der fluoreszierenden Moleküle zueinander nach einer von Malkusch *et al.* (2012)<sup>277</sup> beschriebenen Methode wurde die Kolokalisation von EnvZ-mNeonGreen und MzrA-PAmCherry in Abhängigkeit von EF-P auf Einzelzellebene untersucht (Abb 30B).

Entgegen der Erwartung stellte sich heraus, dass EnvZ-mNeonGreen und MzrA-PAmCherry in Wildtypzellen nicht kolokalisieren. Nur ein Bruchteil der grün fluoreszierenden Cluster liegt in ausreichend geringer Distanz zu roten und umgekehrt. Dasselbe gilt für die  $\Delta\text{efp}$  Mutante. Grund dafür könnte die Fixierung der Zellpräparate sein, welche die Proteinstrukturen möglicherweise so verändert, dass die Interaktion aufgelöst wird, bevor die Proteine durch Quervernetzung vollständig immobilisiert sind. Wie bereits gezeigt, ist die Interaktion zwischen EnvZ und MzrA dynamisch. Denkbar wäre daher auch, dass eine temporäre Bindung zum Zeitpunkt der Zellfixierung nicht sichtbar ist. Die hier angewendeten Versuchsbedingungen lassen keine verlässlichen Aussagen zur Kolokalisation von EnvZ und MzrA auf Einzelzellebene zu.



**Abb 30:** Lokalisation von MzrA-PAmCherry („rot“) und EnvZ-mNeonGreen („grün“) in doppelmarkierten *E. coli* BW25113 Wildtyp und  $\Delta$ efp Stämmen mittels PALM. (A) Visualisierung der Proteincluster am Beispiel einzelner Zellen. (B) Koordinatenbasierte Kolokalisationsanalyse (CBC)<sup>277</sup>: Werte von 0 bedeuten keine Kolokalisation, während eine vollständige Kolokalisation durch den Wert 1 beschrieben wird. Nur bis zu etwa 30% aller erfassten Fluoreszenzereignisse (*events*) haben einen Abstand von weniger als 300 nm zum nächsten andersfarbigen Event. (Wildtyp: „grün“ -> „rot“ 27.63307% - „rot“ -> „grün“ 46.90476%;  $\Delta$ efp: „grün“ -> „rot“ 18.43003% - „rot“ -> „grün“ 28.77698%).

Material und Methoden: *E. coli* BW25113 Wildtyp MzrA-PAmCherry/EnvZ-mNeonGreen („WT“) = St26, *E. coli* BW25113  $\Delta$ efp MzrA-PAmCherry/EnvZ-mNeonGreen („ $\Delta$ efp“) = St27.

Analyse und Darstellung durch Giacomo Giacomelli (AG Prof. Dr. M. Bramkamp, Bereich Mikrobiologie, LMU München). Die Daten beruhen auf der Auswertung der Fluoreszenz von ca. 500 Zellen.

### 3.3.5 Die phylogenetische Analyse des Auftretens der EnvZ polyP Motive weist auf eine Koevolution mit MzrA hin

Unter der Annahme, dass das periplasmatische EnvZ polyP Motiv VVPPA auch in anderen Organismen als *E. coli* die Interaktion mit MzrA unterstützt, wurde untersucht, ob das Auftreten von MzrA mit der Präsenz dieses Motivs oder mindestens zwei konsekutiver Proline an entsprechender Position korreliert. Obwohl das Motiv in der unstrukturierten Linkerregion der Domäne liegt und damit ungeeignet für die Erstellung von üblichen Homologiemodellen ist<sup>344</sup>, wurde aufgrund seiner hohen Konservierung in EnvZ Orthologen aus verschiedenen Gammaproteobakterien und seiner Rolle in der Signalaktivität des Rezeptors ein detaillierterer Blick auf seine Evolution geworfen.

In *E. coli* existieren keine Paraloge zu EnvZ. Um die Verteilung von EnvZ polyP Motiven zu visualisieren, wurde der in Kapitel 3.3.4 gezeigte phylogenetische Protein-Stammbaum aus EnvZ Orthologen mit einer Sequenzübereinstimmung von mindestens 44% und einer Längenabweichung von höchstens 10% zum *E. coli* K12 EnvZ verwendet. Mittels NCBI BLAST wurde innerhalb der ausgewählten Proteome nach Orthologen von MzrA aus *E. coli* K-12 MG1655 gesucht, wobei eine Sequenzübereinstimmung von mindestens 31% empirisch als Grenzwert für die Definition einer Proteinsequenz als „MzrA“ festgelegt wurde. Unter Verwendung eines Farbschemas wurde das Auftreten des periplasmatischen polyP<sub>p</sub> (analog zu VVPPA in *E. coli* K12 MG1655) und des cytoplasmatischen polyP<sub>c</sub> (analog zu IPPPL in *E. coli* K12 MG1655) Motivs in den betreffenden Organismen markiert (Abb. 31A). Ein cytoplasmatisches polyP<sub>c</sub> Motiv mit drei konsekutiven Prolinen ist weit verbreitet. Dies unterstreicht die Wichtigkeit eines Triple-Prolins in der EnvZ HAMP Domäne und dessen hohen Konservierungsgrad. Im Gegensatz dazu findet sich das periplasmatische polyP<sub>p</sub> Motiv ausschließlich in näher verwandten EnvZ Orthologen, und insbesondere in Repräsentanten der *Enterobacteriaceae*. Interessanterweise existieren MzrA Orthologe exklusiv in Organismen, in welchen EnvZ dieses polyP Motiv trägt. Ausnahmen sind *Pragia*, *Leminorella* und *Budvicia*, welche ein periplasmatisches polyP Motiv aufweisen, jedoch nicht MzrA. Dies spricht für eine Adaptation von EnvZ an seinen Modulator MzrA.

Um die Entwicklung der beiden EnvZ polyP Motive im Laufe der Evolution nachzuverfolgen, wurde auf einen Spezies-Stammbaum von Gammaproteobakterien zurückgegriffen<sup>332</sup>. Die Proteome der Organismen der Knotenpunkte wurden mit Hilfe der BLAST Homologiensuche auf das Vorhandensein von EnvZ Orthologen hin untersucht, welche entsprechend ihrer Sequenzübereinstimmung farblich markiert wurden (Abb. 31B). Gleichzeitig wurde das Vorhandensein von MzrA Orthologen mit mindestens 31% Sequenzübereinstimmung zu MzrA aus *E. coli* K-12 MG1655, sowie der beiden polyP Motive dargestellt. Dabei wird verdeutlicht, dass beide polyP Motive spät während der Evolution entstanden und das periplasmatische Motiv zusammen mit MzrA in Erscheinung tritt.

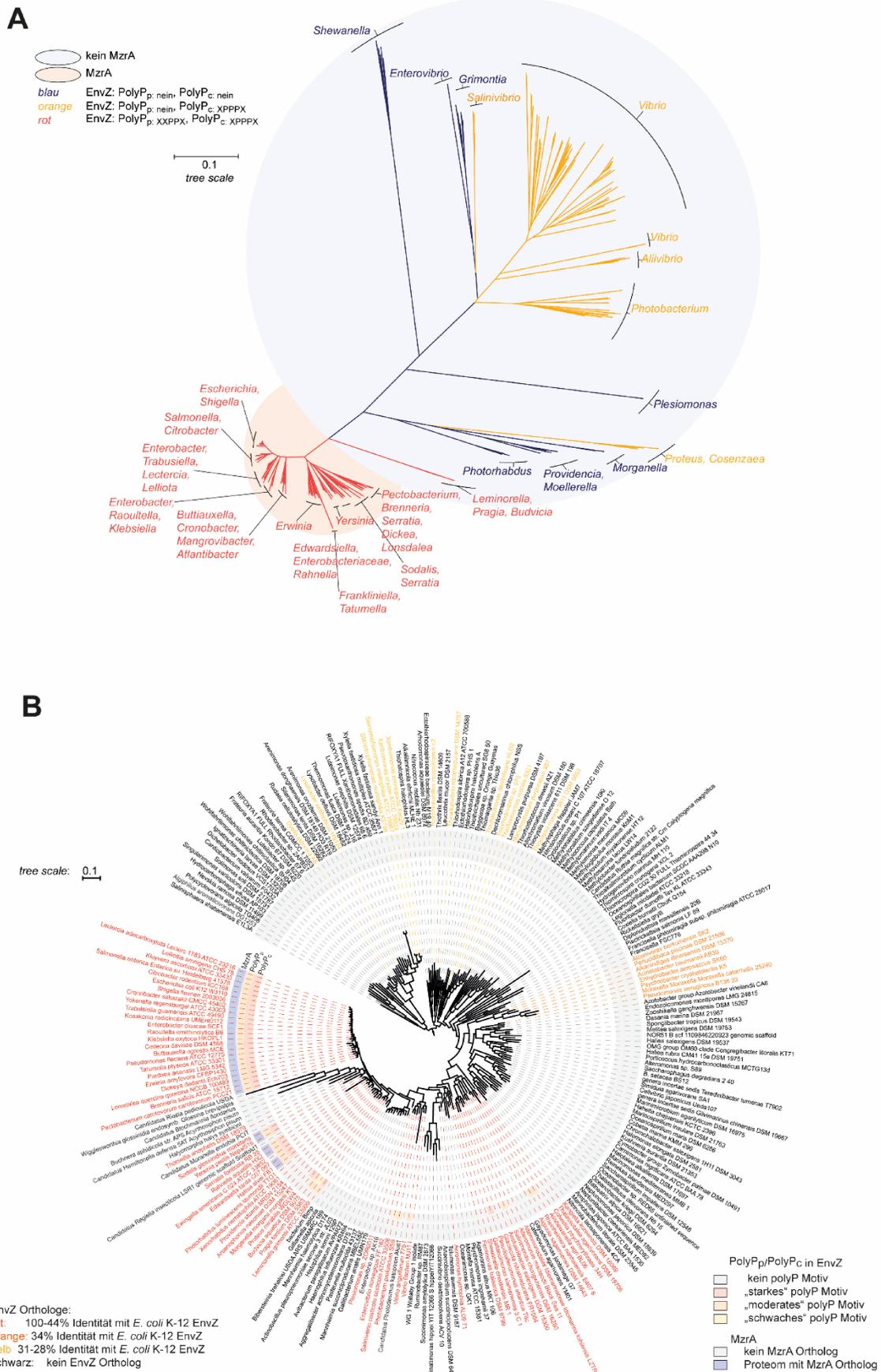


Abb. 31: Beschreibung auf folgender Seite.

**Abb. 31:** Koevolution von MzrA und der beiden EnvZ polyP Motive. (A) Phylogenetischer Protein-Stammbaum, basierend auf einem Sequenzalignment von orthologen EnvZ Sequenzen (>44% Sequenzidentität zu *E. coli* K-12 MG1655 EnvZ). Stämme, die zugleich Orthologe des *E. coli* K-12 MG1655 Proteins MzrA (< 31% Sequenzidentität) tragen, sind durch einen roten Hintergrund hervorgehoben. Die Art der EnvZ polyP Motive ist durch die Schriftfarbe, sowie die Farbe der Äste gekennzeichnet. (B) Speziesbaum der Gammaproteobakterien aus einem von Hug *et al.* (2016)<sup>332</sup> veröffentlichten „tree of life“. EnvZ Orthologe der verschiedenen bakteriellen Spezies sind durch die Schriftfarbigkeit unterteilt nach dem Grad der Sequenzübereinstimmung zu EnvZ aus *E. coli* K-12 MG1655. Die Anwesenheit von MzrA Orthologen ist durch eine blaue Hinterlegung gekennzeichnet. Die polyP Motive sind durch eine ihrer vorhergesagten „Stärke“ entsprechenden farbigen Hinterlegung dargestellt.

Die Daten wurden 2018 publiziert <sup>331</sup>.

### 3.4 Regulation der Expression und Aktivität von EF-P

PolyP Motive stellen durch ihre Abhängigkeit von einem spezifischen Translations-Elongationsfaktor (EF-P) ein potentiell Instrument dar, um das Translationslevel von Proteinen gezielt zu manipulieren. Eine an bestimmte Umweltbedingungen angepasste Expression oder Aktivierung von EF-P könnte die Kopienzahl von Proteinen bestimmter Signalwege oder Stoffwechselwege zugunsten einer Adaptation der Bakterienzelle beeinflussen.

*Efp* und *epmB* sind benachbarte, entgegengesetzt ausgerichtete Gene mit einem Abstand von 42 bp, deren von Aoki *et al.* (1991)<sup>68</sup> und Huerta *et al.* (2003)<sup>345</sup> vorhergesagte Promotorregionen sich überschneiden und in das jeweils andere Gen hineinragen. *EpmA* liegt 325 bp vom ebenfalls entgegengesetzt ausgerichteten Gen *frdA* entfernt. Im Gegensatz zu *efp* und *epmB* wurde der Promotor von *epmA* experimentell bestimmt (Mendoza-Vargas *et al.* 2009<sup>346</sup>).

Die Genexpression von *efp*, *epmA* und *epmB* steht unter der Kontrolle des Sigmafaktor70<sup>345, 346</sup>. Dies erlaubt die Vorhersage der zugehörigen Promotorsequenzen und putativen Bindestellen für Transkriptionsfaktoren mit Hilfe des Tools BPROM (Tab. 14). Die ermittelten Positionen des Transkriptionsstarts entsprechen im Fall von *efp* und *epmB* exakt den bereits publizierten Vorhersagen (Huerta *et al.*, 2003<sup>345</sup>), weichen im Fall von *epmA* jedoch stark davon ab (Vorhersage durch Mendoza-Vargas *et al.*, 2009<sup>346</sup>: 25 bp *upstream* von ATG). Dabei ist der hohe *score* von 7,02 der BPROM Vorhersage für einen *epmA* Transkriptionsstart im Abstand von 142 bp vom Startcodon des Gens hervorzuheben, der die Wahrscheinlichkeit dieses Ergebnisses bekräftigt.

**Tab. 14:** Putative Bindestellen von Transkriptionsfaktoren in den Promotorregionen von *efp*, *epmA* und *epmB*, wie sie in einer Analyse mittels BPROM vorhergesagt werden. Die „Position“ beschreibt, wieviele Nukleotide die erste Base des genannten DNA-Motivs *upstream* vom Startcodon des betreffenden Gens entfernt liegt. TSS = *transcriptional start site*.

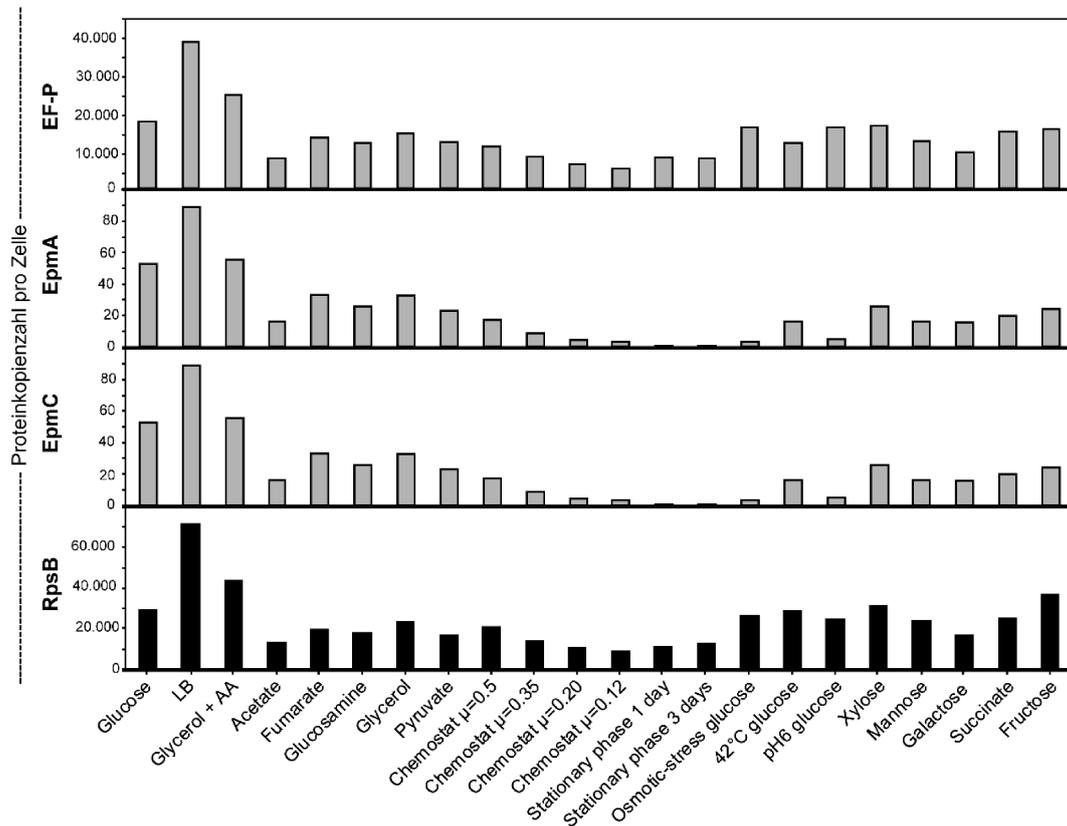
<i>efp</i>				<i>epmA</i>			
Promotor #1	Position	Sequenz	Score	Promotor #1	Position	Sequenz	Score
TSS	-44		<b>3,73</b>	TSS	-142		<b>7,07</b>
-10 Box	-59	GGTTACAAT	74	-10 Box	-159	GGCTAAATT	58
-35 Box	-80	TTCTCT	28	-35 Box	-176	TTTTTA	26
<b>TF Bindestellen:</b>				<b>TF Bindestellen:</b>			
Crp	-89	TAACCAAT	6	Lrp	-179	ATTTTTTT	11
<b><i>epmB</i></b>				LexA	-178	TTTTTTTA	16
Promotor #1	Position	Sequenz	Score	RpoD	-165	CAAGAGGG	7
TSS	-45		<b>3,32</b>	RpoD	-124	CTAAACCA	6
-10 Box	-60	TGCTATAGT	64	OxyR	-138	GATTAATT	9
-35 Box	-83	TTAAG	35	Ihf	-136	TTAATTTA	12
<b>TF Bindestellen:</b>				Hns	-135	TAATTTAA	9
SoxS	-28	ATTTGTTA	9	PhoB	-132	TTTAATTA	11
RpoD	-40	AGGCCCTC	7	Promotor #2	Position	Sequenz	Score
				TSS	-492		<b>2,23</b>
				-10 Box	-508	TGTCATGAT	51
				-35 Box	-528	GTGAAA	27
				<b>TF Bindestellen:</b>			
				SoxS	-537	TACTGTAT	9
				RpoS	-508	TGTCATGA	7

Die Analyse zeigt für *efp* und *epmB* einige wenige putative Bindestellen für Transkriptionsfaktoren auf. Im Fall der mittels BPROM für *efp* identifizierten Bindestellen handelt es sich um die partielle Consensussequenz der Bindestelle von Crp, einem durch cAMP aktivierten globalen Regulator von über 300 Genen, die überwiegend an der Katabolisierung von sekundären Kohlenstoffquellen beteiligt sind<sup>347, 348</sup>. In der Promotorsequenz von *epmB* befinden sich mögliche Bindestellen für RpoD und SoxS. Der Sigmafaktor70 (RpoD) spielt zusammen mit der RNA Polymerase Sigma70 vor allem während der exponentiellen Wachstumsphase eine große Rolle und ist assoziiert mit Genen, die essentiell für ein normales Wachstum sind<sup>349, 350</sup>. SoxS ist an der Aktivierung von Genen, die Teil der Stressantwort gegen Super- und Stickoxide sind, beteiligt<sup>351</sup>.

Die Länge des nichtkodierenden Bereichs zwischen *epmA* und dessen Nachbaroperon *frdABCD* deutet bereits auf das Vorhandensein von regulatorischen Elementen hin. Dem 93 bp langen Promotor des *frdABCD* Operons werden Bindestellen für NarL, Fnr und DcuR zugeordnet<sup>352-355</sup>. Auffälligerweise finden sich mit Hilfe von BPROM *upstream* von *epmA* mögliche Bindestellen für insgesamt 7 Transkriptionsfaktoren: Lrp, LexA, RpoD, OxyR, Ihf, Hns und PhoB. Die Bindestellen für den Sigmafaktor70 (RpoD), sowie für Lrp weisen auf eine globale Regulation des Promotors hin. Lrp reguliert die Transkription von mehr als 10% und während der stationären Phase sogar bis zu dreiviertel aller *E. coli* Gene.<sup>356</sup> Dazu zählen Gene die mit Aminosäurekatabolismus und – synthese assoziiert sind, sowie Gene für Stofftransport, Pilisynthese oder Kohlenstoffmetabolismus<sup>357-359</sup>. Hns und IHF sind in erster Linie an der Stabilisierung der DNA-Architektur beteiligt. Im Gegensatz zu diesen universalen Mechanismen regulieren die Transkriptionsfaktoren PhoB, LexA und OxyR die Expression ihrer Zielgene spezifisch als Antwort auf bestimmte Umweltreize. PhoB agiert in Abhängigkeit von extrazellulärem Phosphatkonzentrationen und spielt eine Rolle in der Virulenz pathogener *E. coli* Spezies<sup>360-362</sup>. LexA reprimiert im Rahmen der SOS Schadensantwort die Transkription von Genen, die mit der DNA-Reparatur und -Replikation in Verbindung stehen<sup>363, 364</sup>. Der oxydative Stressregulator OxyR aktiviert oder reprimiert Gene, die an der zellulären Antwort gegen oxydativen Stress, bzw. erhöhten Konzentrationen von Wasserstoffperoxid beteiligt sind, sowie bei nitrosativem Stress die Gene des Regulons der S-Nitrosylierung<sup>365-367</sup>. In einer zweiten, jedoch mit einer geringeren Wahrscheinlichkeit vorhergesagten Promotorregion (*score* 2,23) befinden sich Bindestellen für LexA und RpoS. Der alternative Sigmafaktor  $\sigma$ S (RpoS) ist Hauptregulator in der allgemeinen Stressantwort von *E. coli*<sup>350</sup>. Mit OxyR, LexA, PhoB und RpoS ist die Expression von *epmA* potentiell mit extremen, sowie spezifischen Stressbedingungen assoziiert. Eine Absenkung der *epmA* Expression könnte dazu führen, dass in Folge die Anzahl an aktivierten EF-P Molekülen derart sinkt, dass dies eine Auswirkung auf die Translationsrate von Proteinen mit polyP Motiv hat. Demzufolge würde das Translationslevel dieser Proteine ebenfalls sinken. Eine verminderte Translation von Proteinen, die keinen positiven Beitrag zur Beseitigung der Stresssituation leisten, könnte von Vorteil sein. Es ist aber ebenso denkbar, dass die Transkriptionsfaktoren die Expression von *epmA* erhöhen. Durch eine höhere Anzahl von aktivierten EF-P Molekülen könnte die Translationseffizienz von Proteinen mit polyP Motiv, die zu einer Adaptation an die veränderte Umweltbedingung beitragen, möglicherweise gesteigert werden.

### 3.4.1 Hinweise auf eine transkriptionale Regulation unter verschiedenen Stressbedingungen

Um festzustellen, ob die Aktivität von EF-P tatsächlich variabel ist, wurden Transkription und Translation von EpmA, aber auch EpmB und EF-P unter verschiedenen Wachstumsbedingungen untersucht. Eine quantitative, massenspektrometrische Untersuchung des Proteoms von *E. coli* durch Schmidt *et al.* (2016)<sup>87</sup> zeigt, dass die Kopienzahl der drei Proteine durch die Verwendung verschiedener Kohlenstoffquellen nur wenig beeinflusst werden (Abb.32). Ein ähnliches Muster zeigt sich für das ribosomale Protein RpsB. Hitzestress und Osmotress scheinen sich auf EpmA und EpmB Level im Vergleich zu RpsB moderat auszuwirken. Im Rahmen dieser Arbeit wurde der Einfluss weiterer Stressbedingungen auf das Transkriptions- und Translationslevel von EF-P, EpmA und EpmB untersucht. Die dafür angewandten Kultivierungsmethoden sind im Kapitel 2.3 (Kultivierungsverfahren) aufgeführt.



**Abb. 32:** Zusammenfassung der von Schmidt *et al.* (2016)<sup>87</sup> veröffentlichten Daten zur Kopienzahl der *E. coli* K-12 Proteine EF-P, EpmA, EpmB und RpsB unter verschiedenen Wachstumsbedingungen.

### 3.4.1.1 Translationslevel eines CadC-LacZ Reporterkonstruktes unter verschiedenen Stressbedingungen

Im natürlichen Habitat ist *E. coli* zahlreichen Stressfaktoren ausgesetzt. Dazu zählen extreme pH-Werte, Temperaturschwankungen, hohe Osmolaritäten und Limitation der Verfügbarkeit von Stickstoff und Kohlenstoff. In einem ersten Test wurde daher unabhängig von den oben ermittelten putativen Bindestellen für Transkriptionsfaktoren untersucht, ob diese Bedingungen einen Einfluss auf das Translationslevel von Proteinen mit polyP Motiv haben, was einen indirekten Rückschluss auf die Aktivität von EF-P erlaubt. Stellvertretend wurde für diese Analyse ein plasmidkodiertes CadC-PPP-LacZ Reporterkonstrukt mit PPP-Motiv gewählt und die relative Abundanz des Proteins durch die Bestimmung der  $\beta$ -Galactosidaseaktivität ermittelt.

In einem ersten Ansatz wurde das Konstrukt unter der Kontrolle des konstitutiven Promotors  $P_{cadC}$  in einem Wildtyp *E. coli* BW25113 Stamm und in einem  $\Delta efp$  Stamm untersucht. Die Zellen wurden aus einer Übernachtskultur in M9 Medium (Start- $OD_{600} = 0,1$ ) überführt und bis in die exponentielle Wachstumsphase kultiviert, bevor die Stressoren durch einen Austausch des Mediums eingesetzt wurden. Ein Kontrollversuch mit einem CadC-LacZ Reporterkonstrukt ohne polyP-Motiv bestätigte, dass das Motiv zu einer EF-P Abhängigkeit des Translationslevels führt (Abb. 33A). Obwohl nach Kultivierung in den gewählten Wachstumsbedingungen leichte Schwankungen in der aus CadC-PPP-LacZ resultierenden  $\beta$ -Galactosidaseaktivität zu erkennen sind, konnte unter keiner der getesteten Bedingungen eine signifikante Abweichung beobachtet werden (Abb. 33B).

Ergänzend wurde das CadC-PPP-LacZ Reporterkonstrukt unter Kontrolle des arabinoseinduzierbaren PBAD-Promotors untersucht. Damit wird gewährleistet, dass die Translation des Reporterkonstruktes während der definierten Zeitspanne des Stresseinflusses stattfindet. Geringe Unterschiede lassen sich auf diese Weise deutlicher auflösen und bereits davor translatierte Proteine fallen nicht ins Gewicht. Eine Induktion mit 0,33% L-Arabinose und eine 30 - 60 minütige Inkubation erwiesen sich als geeignete Parameter. Allerdings konnte auch mit dieser Methode bei keiner der getesteten Bedingungen eine signifikante Änderung des Translationslevels des EF-P abhängigen Reporterkonstruktes festgestellt werden (Abb. 33C).

PolyP Motive werden bei keiner der hier getesteten Stressbedingungen translational reguliert. Eine Feinabstimmung lässt sich dennoch nicht völlig ausschließen.

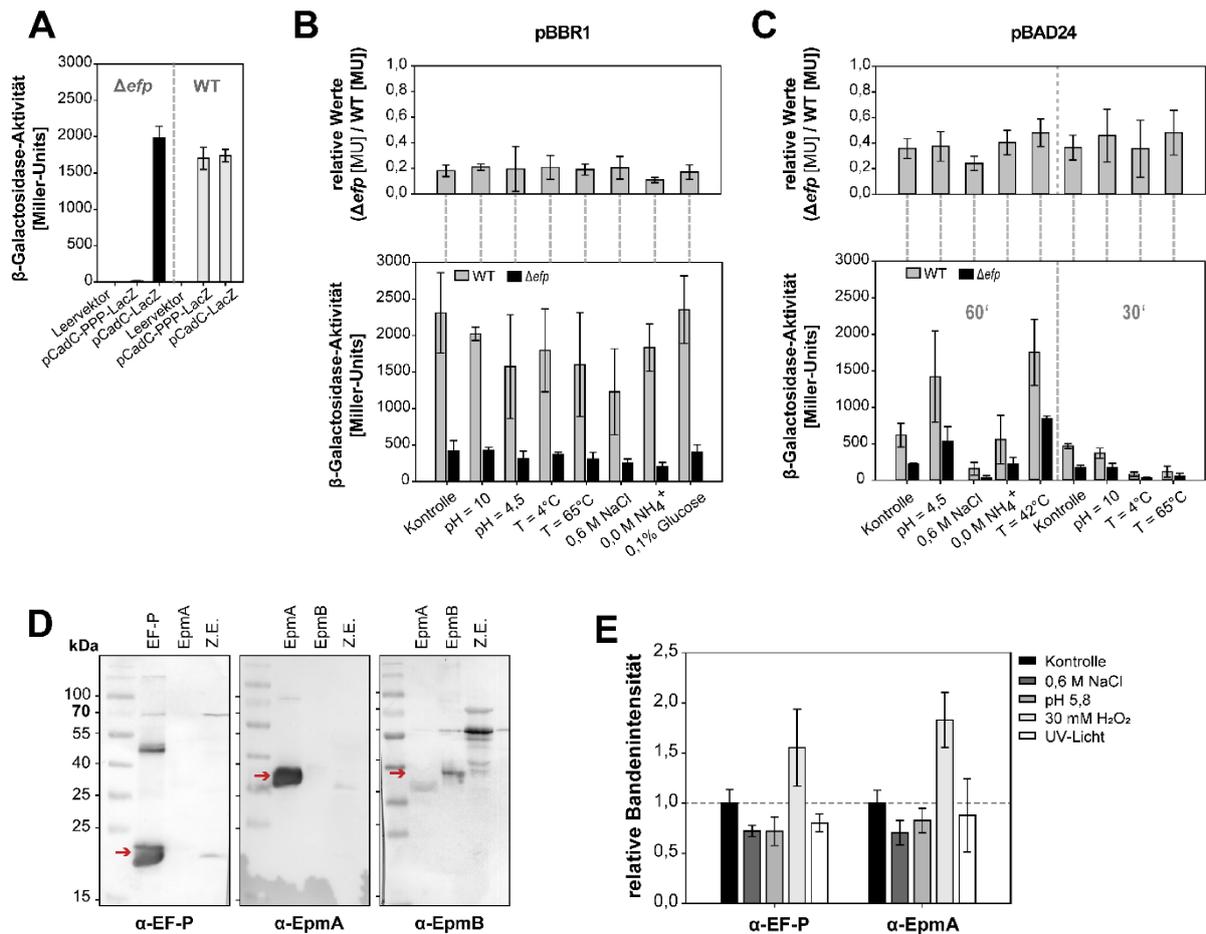
### 3.4.1.2 Translationslevel von EF-P und EpmA unter verschiedenen Stressbedingungen

Mittels Western Blot Analyse wurden die Proteinlevel von EF-P und EpmA aus *E. coli* BW25113 unter verschiedenen Stressbedingungen direkt quantitativ miteinander verglichen.

#### 3.4.1.2.1 Etablierung eines Westernblot-Protokolls zur Immundetektion von EF-P, EpmA und EpmB

Durch die Immunisierung von Kaninchen mit gereinigten, His-markierten EF-P, EpmA und EpmB wurden spezifische Antikörper generiert. Die Antikörper wurden durch eine Affinitätsreinigung der Antiseren konzentriert. Das in den Methoden (Kapitel 2.6.5) beschriebene Standard-Westernblot-Protokoll erwies sich als

geeignet, um die EF-P und EpmA im Western Blot zu analysieren. Bei einer Verdünnung von 1:5.000 konnten die jeweiligen Zielproteine durch  $\alpha$ -EF-P und  $\alpha$ -EpmA mit ausreichender Sensitivität und Spezifität detektiert werden. Dies gilt sowohl für gereinigte Proteine, als auch für *E. coli* Gesamtzellextrakte (Abb. 33D). Der Antikörper  $\alpha$ -EpmB zeigte eine starke unspezifische Reaktivität gegenüber anderen *E. coli* Proteinen, die eine eindeutige Detektion von EpmB im Western Blot unmöglich machte.



**Abb. 33:** Auswirkungen verschiedener Stresskonditionen auf die Translationslevel von EF-P und des modifizierenden Enzyms EpmA in *E. coli* BW25113. Kultivierungsbedingungen wie in Kapitel 3.2, Tabelle 9 beschrieben. (A)  $\beta$ -Galactosidase-Assay zur Verifizierung der EF-P Abhängigkeit eines CadC-PPP-LacZ Reporterkonstruktes. (B) Quantifizierung des Translationslevels eines unter der Kontrolle des konstitutiven  $P_{cadC}$  Promotors stehenden CadC-PPP-LacZ Reporterkonstruktes nach Exposition mit verschiedenen Stressbedingungen. Unten: absolute  $\beta$ -Galactosidase-Aktivität, oben: Verhältnis der  $\beta$ -Galactosidase-Aktivitäten zwischen WT und  $\Delta efp$ . (C) Quantifizierung des Translationslevels eines unter der Kontrolle des induzierbaren  $P_{BAD}$  Promotors stehenden CadC-PPP-LacZ Reporterkonstruktes nach Exposition mit verschiedenen Stressbedingungen. (D) Prüfung von Spezifität und Sensitivität von für diese Arbeit generierten spezifischen Antikörpern gegen EF-P, EpmA und EpmB im Western Blot. („EF-P“ und „EpmA“: gereinigte, His-markierte Proteine, „Z.E.“ = *E. coli* BW25113 Zellextrakt) (E) Relative Quantifizierung der nativen EF-P und EpmA Proteinlevel aus *E. coli* BW25113 Gesamtzellextrakten nach Exposition mit verschiedenen Stressbedingungen mittels Western Blot. Relative Quantifizierung der Bandenintensitäten.

Material und Methoden: (A) Leervektor = V13, pCadC-PPP-LacZ = V14, pCadC-LacZ = V15, WT = St4,  $\Delta efp$  = St6. (B) pCadC-PPP-LacZ = V14, WT = St4,  $\Delta efp$  = St6. (C) pCadC-PPP-LacZ = V18, WT = St4,  $\Delta efp$  = St6. (D+E) St4.

Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen.

#### 3.4.1.2.2 Auswirkung von oxydativem Stress auf EF-P und EpmA Proteinlevel

Um die Ergebnisse des Reporterassays zu bestätigen und den Einfluss weiterer Stressfaktoren, insbesondere der durch die Promotoranalyse plausibel erscheinenden oxydativen Stressbedingung und DNA-Schädigung, zu testen, wurde *E. coli* K-12 MG1655 in LB Medium angezogen und die Stressbedingungen durch verschiedene Kultivierungsmethoden und UV-Licht simuliert (Kapitel 2.3). Nach Ernte und Zellaufschluss wurden die Proteinlevel von EF-P und EpmA im Western Blot analysiert. Es konnten keine signifikanten Unterschiede im Vergleich zur Kontrolle festgestellt werden, außer für den oxydativen Stress. Die im Western Blot detektierten Banden von EF-P und EpmA weisen hier eine 1,5-fach, bzw. 1,6-fach stärkere Intensität auf, als die Kontrollen (Abb. 33E, zweiseitiger t-Test, Signifikanzniveau 5%).

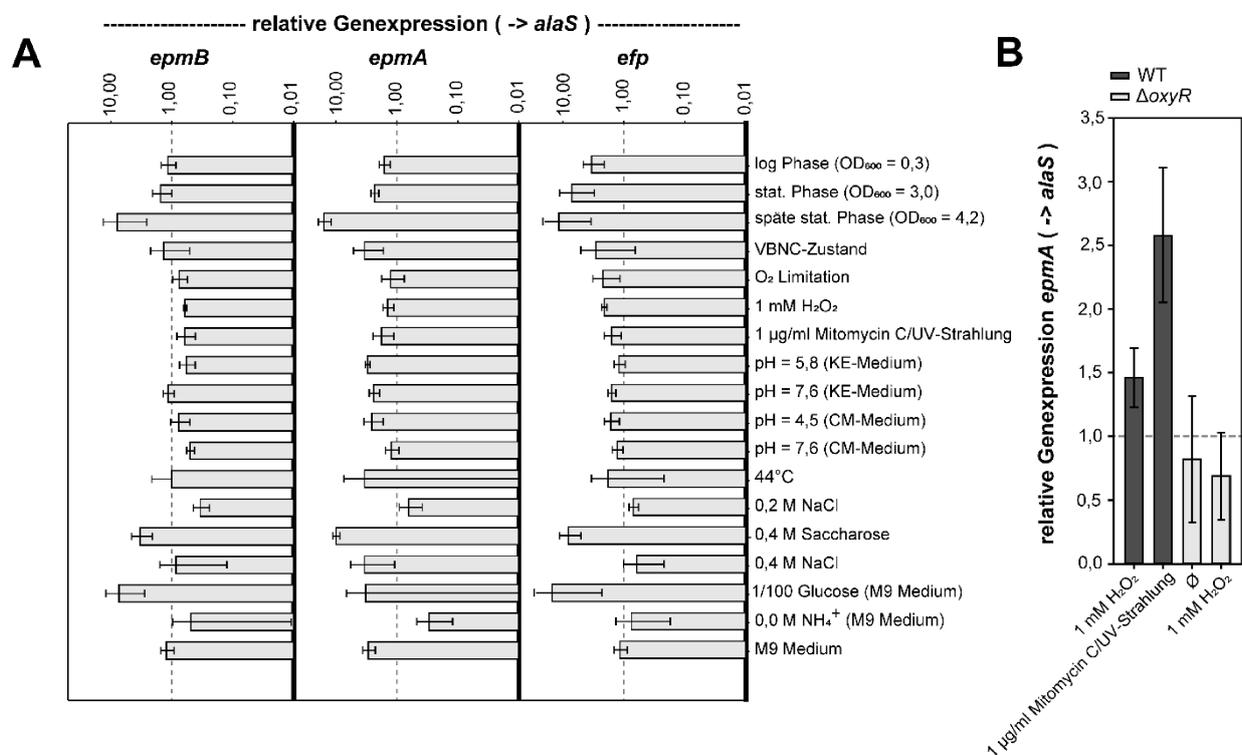
#### 3.4.1.3 Auswirkung von Nährstofflimitation auf die Genexpression von *efp*, *epmA* und *epmB*

Um eine mögliche Anpassung der Expression von *efp*, *epmA* und *epmB* auf transkriptioneller Ebene festzustellen, wurden mRNA-Level von *E. coli* K-12 MG1655 nach Exposition mit verschiedenen Stressfaktoren mittels quantitativer *real-time* PCR (qRT-PCR) relativ zu in LB-Medium aerob bei 37°C bis zu einer OD<sub>600</sub> von 0,7 kultivierten Zellen quantifiziert. Die Werte wurden normalisiert auf das Expressionslevel der Aminoacyl-tRNA-Synthetase *alaS* (Abb. 34A). Die Kultivierung erfolgte im Einzelnen bei den bereits zuvor untersuchten Bedingungen Kohlenstoff- und Stickstofflimitation, hoher Osmolarität, pH-Stress, Hitzestress, oxydativem Stress und DNA-Schädigung. Zusätzlich wurden verschiedene Wachstumsphasen und die anaerobe Atmung betrachtet.

Das Level von *epmA*-Transkripten ist in der frühen exponentiellen Phase leicht erhöht (2,9-fach), während alle drei Gene in der späten stationären Phase um das 0,3 bis 0,8-fache weniger stark exprimiert sind. Nach einer längeren Kultivierung über 12 Stunden und dem Erreichen der Absterbephase, während der es bereits zu Abbauprozessen kommt, steigen die mRNA Level aller drei Gene deutlich an. Dies gilt jedoch nicht für 120 Tage alte Zellkulturen, die bereits vor Erreichen der stationären Phase bei einer OD<sub>600</sub> von 1,0 in einen Ruhezustand versetzt wurden (VBNC). In diesen Zellen ähneln die Transkriptionslevel denen der frühen exponentiellen Phase (0,1 bis 0,5-fache Expression). Dies lässt vermuten, dass die für die Absterbephase typischen Faktoren wie Nährstofflimitation, hohe Konzentrationen an sezernierten Stoffwechselprodukten und die aus Degradationsprozessen resultierende Akkumulation von Toxinen und reaktiven Sauerstoffspezies<sup>368, 369</sup> den entscheidenden Ausschlag geben. Unter mikroanaeroben Wachstumsbedingungen und bei Verfügbarkeit von Stickstoff findet der Energiestoffwechsel von *E. coli* hauptsächlich über die dissimilatorische Nitratreduktion statt, wobei Stickstoffmonoxyd als Nebenprodukt entsteht<sup>370, 371</sup>. Reaktive Spezies treten sowohl in dieser Wachstumsbedingung, als auch in der Absterbephase auf, möglicherweise ist diese Gemeinsamkeit der Grund für die erhöhte Genexpression von *efp*, *epmA* und *epmB* (3,2-, 10,1- und 8,7-fach). Sowohl oxydativer, als auch nitrosativer Stress werden in *E. coli* durch den Transkriptionsfaktor OxyR beantwortet<sup>365, 367</sup>.

Das niedrige Expressionslevel des Kontrollgens *katG* (6,1±3,80-fach) bei Konfrontation mit 1 mM Wasserstoffperoxid verdeutlicht, dass die hier gewählte oxydative Stressbedingung sehr mild war. Unter diesen Umständen ist keine Veränderung der Transkriptionslevel von *efp*, *epmA* oder *epmB* zu beobachten. Die Induzierung der DNA-Schadensreparatur durch Mitomycin C (*recN*: 342,8± 53,0-fach) hatte keinen Einfluss auf die Transkription von *efp*, *epmA* oder *epmB*.

Ebenso wenig konnten die Transkriptionslevel der drei Gene durch Säurestress, Hitzestress, Osmostress oder Stickstofflimitation verändert werden. Eine Limitation der Kohlenstoffquelle führte hingegen zu einer leichten Erhöhung des Levels an *efp*-Transkripten ( $4,5 \pm 0,5$ -fach) und *epmA*-Transkripten ( $16,4 \pm 4,1$ -fach). Dies ist konsistent mit den Resultaten zur stationären Wachstumsphase. Unter beiden Bedingungen herrscht Nährstoffknappheit. Zusammenfassend ergibt sich eine mäßige Erhöhung der Transkription von *efp*, *epmA* und *epmB* unter der Voraussetzung von Nährstofflimitation und der Präsenz von Degradationsprodukten und Radikalen.



**Abb. 34:** Auswirkungen verschiedener Stresskonditionen auf die Genexpressionslevel von *efp*, *epmA* und *epmB* in *E. coli* BW25113. Kultivierungsbedingungen wie in Kapitel 3.2 beschrieben. (A) Quantifizierung der relativen Genexpression von *efp*, *epmA* und *epmB* aus *E. coli* BW25113 in Abhängigkeit verschiedener Stressbedingungen mittels qPCR. Sofern nicht anders gekennzeichnet, wurden die Zellen in LB-Medium kultiviert. Normalisierung der Genexpression auf die Abundanz von *alaS* kodierender mRNA und normale Wachstumsbedingungen (LB, 37°C, aerob, OD<sub>600</sub> = 0,7). Als Negativkontrolle wurde bestätigt, dass in Stämmen mit Deletion einer der drei untersuchten Gene keine entsprechenden mRNA Level nachgewiesen werden können ( $\Delta efp$ /St6,  $\Delta epmA$ /St11,  $\Delta epmB$ /St12). (B) Quantifizierung der relativen Genexpression von *epmA* unter oxydativen Stressbedingungen und in Abhängigkeit von OxyR mittels qPCR.

Material und Methoden: (A) St4 (B) WT = St4,  $\Delta oxyR$  = St10.

Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen.

### 3.4.1.3.1 Beteiligung von OxyR an der Transkription von *epmA*

Um zu prüfen, ob die Genexpression von *epmA* wie durch BPROM vorhergesagt und durch die vorangegangene Analyse impliziert von dem Transkriptionsfaktoren OxyR und LexA reguliert wird, wurden damit assoziierte Stressbedingungen näher analysiert. Dafür wurden ein Wildtyp *E. coli* Stamm (BW25113), sowie eine  $\Delta oxyR$  Mutante (JW3933) in 20 ml LB Medium bei einer Ausgangsverdünnung von 1:2.000 bis in die frühe exponentielle Phase kultiviert. Nach zwei Stunden wurden die Ansätze entweder mit einer Endkonzentration von 1 mM H<sub>2</sub>O<sub>2</sub> oder 1 µg/ml Mitomycin C versetzt. Nach weiterer zweistündiger Kultivierung unter den veränderten Wachstumsbedingungen wurden die Zellen geerntet, die RNA extrahiert und in cDNA transkribiert. Mittels qPCR wurden die relativen Level der mRNA-Transkripte von *epmA*, sowie der als Positivkontrolle fungierenden Gene *katG* und *recN* quantifiziert (Abb. 34B). Als Referenz diente eine unbehandelte Kultur. Die oxydativen Stressbedingungen konnten erfolgreich induziert werden, wie die 4250±1800-fache Erhöhung der Expression von *katG* im Wildtypstamm verdeutlichen. Auch das Markergen für DNA-Schäden, *recN*, zeigte unter der entsprechenden Bedingung im Wildtypstamm eine 17,8±1,9-fach höhere Expression, als in der Kontrolle. Wie zu erwarten, konnte in der  $\Delta oxyR$  Mutante JW3933 kein Transkript von *oxyR* nachgewiesen werden.

Durch Mitomycin C und UV-Strahlung induzierte DNA-Schädigungen führen zu einem leicht erhöhten Transkriptionslevel von *epmA* (3,0±0,7-fach). Unter oxydativem Stress (1 mM H<sub>2</sub>O<sub>2</sub>) ist das Level von *epmA*-Transkripten im *E. coli* Wildtyp signifikant auf das 1,74±0,31-fache erhöht (zweiseitiger t-Test, Signifikanzniveau 2,5%). Dies ist in JW3933 hingegen nicht der Fall (0,7±0,4-fach), was darauf hindeutet, dass die Anhebung des *epmA* Transkriptionslevels abhängig von OxyR ist.

### 3.4.2 Analyse der Bindeaffinität von OxyR an P<sub>*epmA*</sub>

Der Schwerpunkt der weiteren Analyse wurde auf OxyR gelegt. OxyR bindet mit mittlerer Affinität als Tetramer an vier aufeinanderfolgende, jeweils vier Nukleotide lange Erkennungssequenzen<sup>372, 373</sup>. Bei der Konsensussequenz der OxyR-Bindestelle handelt es sich um eine sogenannte *degenerate recognition site*, was bedeutet, dass die tatsächlichen Bindestellen nicht hundertprozentig übereinstimmen, jedoch ohne dass diese Abweichungen eine Auswirkung auf die Stärke der DNA-Bindung hat. Ausschlaggebend ist vielmehr das Verteilungsmuster der Nukleotide im betreffenden DNA-Sequenzabschnitt<sup>373</sup>. Die durch BPROM vorhergesagte Bindestelle von OxyR liegt in einer Region *upstream* von *epmA*, die Teile der OxyR Konsensussequenz ATAGnnnnnnCTATnnnnnnATAGnnnnnnCTAT enthält. Sieben der jeweils vier Nukleotide langen Erkennungssequenzen der OxyR-Untereinheiten tauchen dabei in unregelmäßigen Abständen von 6-19 Nukleotiden auf. Interessanterweise findet sich *upstream* in einiger Entfernung eine DNA-Region, die mit größerer Genauigkeit mit der Konsensussequenz übereinstimmt: ATAGnnnnnnGTATnnnnnATGGnnnnnnCTGA. OxyR ist bifunktional und agiert sowohl als Repressor als auch als Aktivator, wobei die Zielpromotoren zum Teil übereinstimmen. Die Aktivität wird über den Oxydationszustand zweier Cysteine (C199 und C208) vermittelt. Unter oxydierenden Bedingungen bilden gegenüberliegende Cysteine im OxyR Dimer intermolekulare (C199) und intramolekulare (C208) Disulfidbrückenbindung aus, die eine Konformationsänderung zur Folge haben<sup>374, 375</sup>. So finden sich im Promotor von *oxyRS* leicht alternierende Bindestellen für OxyR im oxydierten

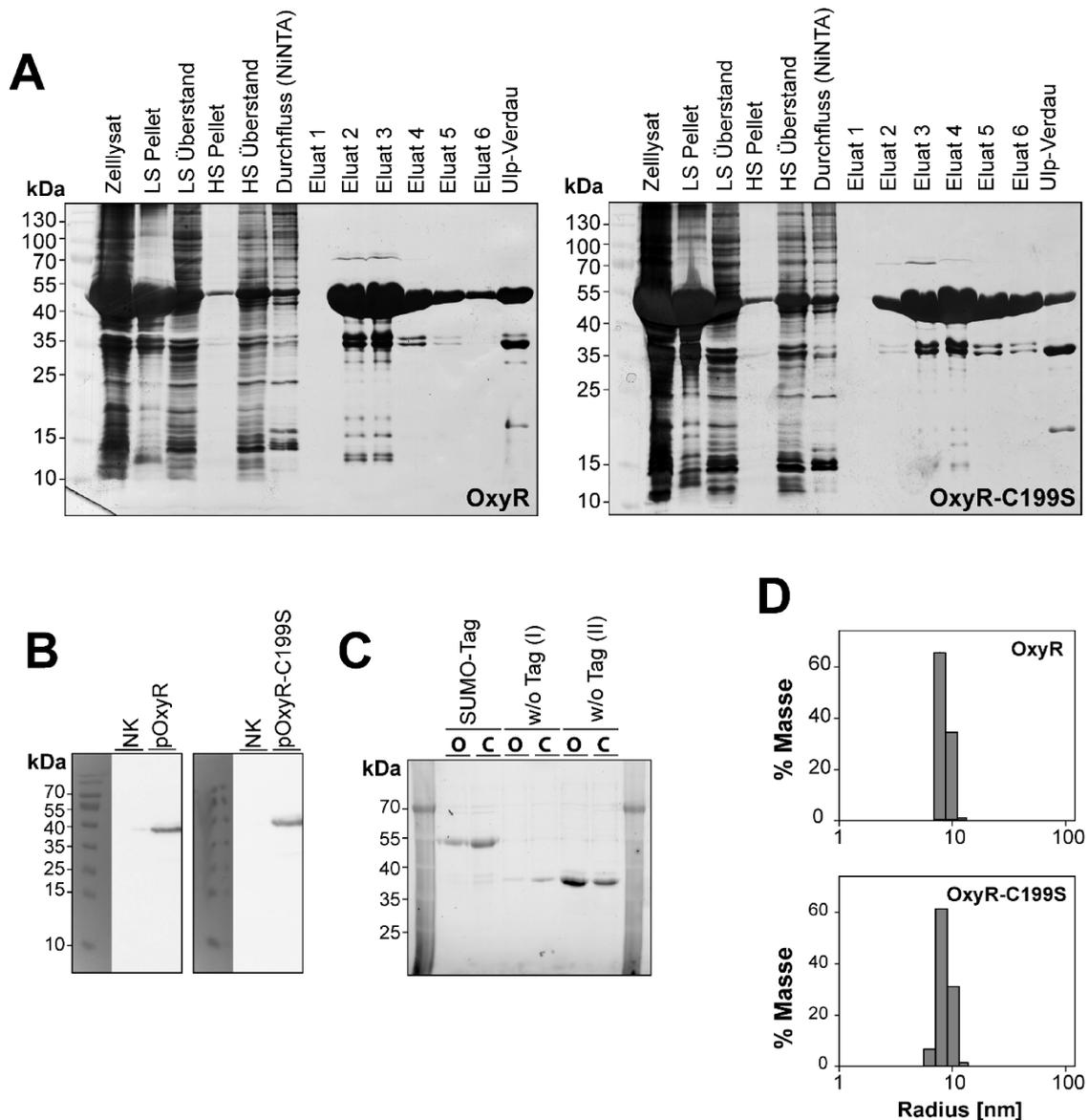
und im reduzierten Zustand, während der Promotor von *katG* hauptsächlich von oxydiertem OxyR gebunden wird<sup>372</sup>. Eine C199S Substitution fixiert das Protein im reduzierten Zustand<sup>372, 376, 375</sup>.

Um die Bindestelle für OxyR in der Promotorregion von *epmA* zu verifizieren, wurde OxyR gereinigt und dessen Bindeaffinität zu der betreffenden DNA-Sequenz *in vitro* getestet. Dafür wurden plasmidkodierte und unter der Kontrolle des durch IPTG induzierbaren T7 Promoters stehende OxyR Varianten mit His-Tag in *E. coli* BL21 DE3 exprimiert und anschließend über eine Nickelchelate-Affinitätschromatographie gereinigt. Bei den OxyR Varianten handelte es sich zum einen um das Wildtypprotein aus *E. coli* K-12 MG1655, zum anderen wurde eine Variante mit C199S Substitution generiert. Während die C199S Variante in der für den reduzierten Zustand typischen Konformation fixiert ist, liegt das native Protein durch die hier angewendete Reinigungsmethode im oxydierten Zustand vor. Dies ermöglicht Rückschlüsse auf die Situation unter normalen Wachstumsbedingungen und bei oxydativem Stress.

### 3.4.2.1 Überexpression und Reinigung von OxyR-Varianten

Die Expression der OxyR-SUMO/His Varianten wurde mittels Western Blot mit spezifischen  $\alpha$ -His Antikörpern überprüft (Abb. 35B). Die Trennung von cytosolischen und membranassoziierten Proteinen zeigte, dass sowohl Wildtyp OxyR-SUMO/His, als auch die OxyR-C199S-SUMO/His Proteinvariante zu über 90% in Lösung vorliegen (Abb. 35A). Nach Reinigung mittels Nickelchelate-Affinitätschromatographie und Abspaltung der His-Markierung unter Verwendung der Ulp-Protease wurde die Reinheit der Proteine im SDS-Polyacrylamidgel überprüft (Abb. 35A und C). Bereits im zweiten Elutionsschritt konnten die Proteine mit einer Reinheit von über 90% und in hoher Konzentration (ca. 2 mg/ml) gewonnen werden. Durch eine säulenchromatographische Größenauftrennung der eluierten Proteine und einem Vergleich mit einem Größenstandard konnte verifiziert werden, dass der Hauptanteil der Moleküle eine Größe von 134 kDa aufweist, was der Größe, des *in vivo* als Tetramer vorliegenden OxyR-Oligomers entspricht.

Durch ein Screening mittels ThermoFluor-Assay wurde ein Puffer für eine adäquate Proteinstabilität identifiziert, der auf den methodischen Beschreibungen von Teramoto *et al.* (2013)<sup>256</sup> basiert (20 mM Tris-HCl pH 7,5; 300 mM NaCl; 10% Glycerin; 0,1 mM EDTA; 5 mM MgCl<sub>2</sub>; 0,05% Tween20; 1 mM DTT; 0,5 mM PMSF). Um ein Umpuffern nach Überexpression und Reinigung der OxyR-Varianten zu vermeiden, wurden die OxyR-Varianten für folgende Affinitätsuntersuchungen direkt in diesem Puffer gereinigt. Eine Analyse der gereinigten Proteine mittels Dynamic Light Scattering (DLS) bestätigte den hohen Reinheitsgrad der monodispersen Lösung und konnte das Vorliegen von Aggregaten, bzw. maßgeblichen Anteilen von degradierten oder unspezifischen Proteinen ausschließen. Die beiden OxyR-Varianten weisen einen Durchmesser von 86 Å (OxyR) und 84 Å (OxyR-C199S) auf (Abb. 35D), was mit den Abmessungen der Kristallstruktur des OxyR-Oligomers (PDB-ID: 1I69) in etwa übereinstimmt.



**Abb. 35:** Reinigung von OxyR und der Variante OxyR-C199S nach Expression in *E. coli* BL21 DE3. (A) Übersicht der Reinigungsschritte nach Zellaufschluss im mit Silbernitrat gefärbten SDS-Gel. Die Spuren „Zellysat“ bis „HS Überstand“ enthalten äquivalent konzentrierte Proben von je 20  $\mu$ l aus in 50 ml resuspendierten Zellen einer 1 l Kultur. Die Spuren „Eluat 1-6“ und „Ulp-Verdau“ enthalten je 10  $\mu$ l der Eluationsschritte bzw. des Verdau mit einer Proteinkonzentration von  $\sim$  2 mg/ml. LS: Probe nach *low speed* Zentrifugation zur Separierung von Zelldebris, HS: Probe nach *high speed* Zentrifugation zur Trennung von löslichen und membrangebundenen Proteinen. (B) Verifizierung der Expression der auf pET\_SUMO Plasmiden kodierte und His-markierte *oxyR*-Varianten nach Induktion und einstündiger Kultivierung bei 37°C mittels Western Blot und spezifischen  $\alpha$ -His Antikörpern. Pro Spur wurden je 150  $\mu$ g Protein aus Gesamtzellextrakten verwendet. NK: nichttransformierte BL21 DE3 Zellen. (C) SDS-Gelelektrophorese zur Verifizierung der Abtrennung des SUMO/His Tags nach Reinigung von OxyR und OxyR-C199S. O: OxyR, C: OxyR-C199S, „SUMO-Tag“: Proteine vor Ulp-Verdau, „w/o Tag“: Proteine nach Ulp-Verdau und Reinigung. Es wurden unterschiedliche Proteinmengen aufgetragen (I und II). (D) Dynamic Light Scattering zur Überprüfung des Reinheitsgrades und der Dispersität der Proteinlösungen.

Material und Methoden: pOxyR = V53, pOxyR-C199S = V54, in: St2.

### 3.4.2.2 Bindeaffinität von gereinigtem OxyR zu $P_{epmA}$

Die Bindeaffinität der gereinigten OxyR-Varianten zu DNA-Sequenzabschnitten aus der putativen *epmA* Promotorregion wurde mit zwei verschiedenen Methoden untersucht: einem Gelretardationsassay, welcher die unterschiedlichen Laufeigenschaften von freier und gebundener 6FAM-markierter DNA während der Polyacrylamid-Gelelektrophorese ausnutzt und der direkten Detektion der Protein-DNA-Interaktion mittels Oberflächen-Plasmonen-Resonanz (SPR, *surface plasmon resonance*), welche auf Lichtbrechungsindizes von immobilisierter dsDNA mit oder ohne gebundenem Protein basiert. Dabei wurden bekannte DNA-Binderegionen als Positivkontrolle und Promotoren, die mit hoher Wahrscheinlichkeit nicht von OxyR reguliert werden, als Negativkontrolle herbeigezogen, sowie verschiedene DNA-Fragmentlängen getestet (Abb. 36).

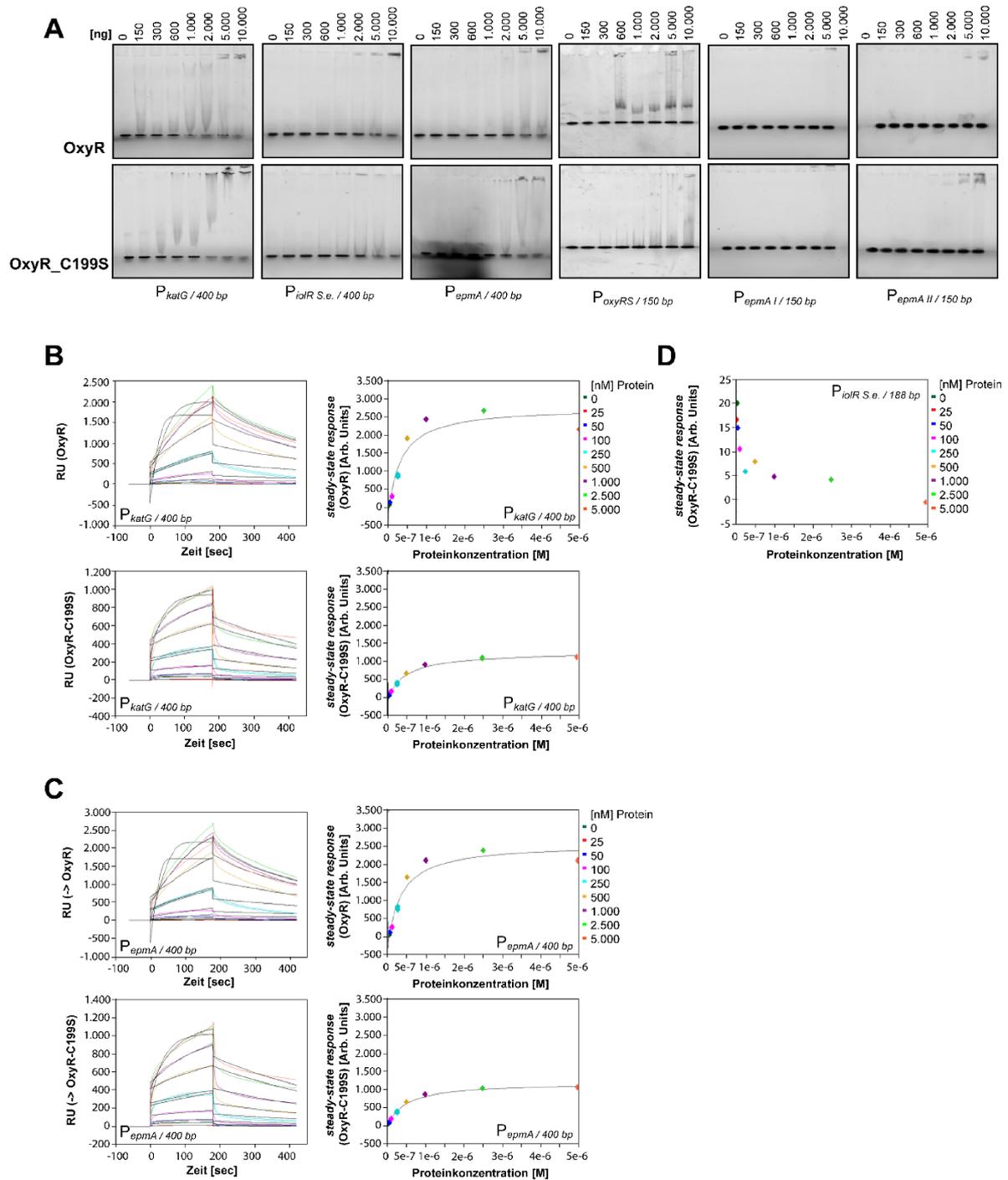
Die Versuchsaufbau des Gelretardationsassays basierte auf den methodischen Beschreibungen zur Identifizierung der *E. coli* OxyR Bindestelle von Tartaglia *et al.* (1992)<sup>373</sup>. Es konnte bestätigt werden, dass OxyR wie erwartet an ein 400 bp langes DNA-Fragment bindet, das im Genom von *E. coli* K-12 MG1655 direkt *upstream* von *katG* liegt und die bekannte OxyR-Konsensussequenz enthält. Je höher die Proteinkonzentration, desto weniger ungebundene DNA. Auch das die im reduzierten Zustand vorliegende Konformation simulierende OxyR-C199S zeigt eine gewisse, jedoch geringere Affinität gegenüber dieser DNA-Sequenz (Abb. 37A). Im Gegensatz dazu kann dies im Fall der Negativkontrolle - ein 400 bp langes Fragment aus der Promotorregion des *Salmonella enterica* Gens *iolR* - eindeutig ausgeschlossen werden. Beide OxyR Varianten zeigen im Gegensatz dazu eine gewisse Bindeaffinität zur 400 bp langen *epmA* Promotorsequenz. Die im Gel ersichtliche Bandenverschiebung ist deutlich erkennbar, aber schwächer als im Fall der *katG* DNA-Sequenz. Um die Binderegion schärfer einzugrenzen, wurde auf 150 bp lange DNA-Fragmente zurückgegriffen. Beide OxyR Varianten zeigen eine gewisse Bindeaffinität gegenüber einem 150 bp langen Fragment aus der *oxyRS* Promotorsequenz mit bekannter OxyR Bindestelle. Es besteht ebenfalls eine geringe Bindung an eine *epmA* Promotorsequenz ( $P_{epmA II}$ ) mit durch BPROM vorhergesagter OxyR Bindestelle.

Eine Analyse mittels Oberflächen-Plasmonen-Resonanzspektroskopie bestätigte diese Ergebnisse größtenteils. Beide OxyR-Varianten zeigen eine hohe Affinität zu 400 bp langen Promotorsequenzen von *katG* und *epmA* (Abb. 38B und C). Allerdings zeichnet sich die Bindung in diesem Assay durch eine sehr hohe Sensitivität und geringe Spezifität aus, was sich durch die in allen Fällen sehr hohen Assoziations- und Dissoziationskoeffizienten widerspiegelt. Deshalb wurde auf die Berechnung eines  $K_d$ -Wertes verzichtet. Als eindeutige Negativkontrolle fungierte eine 188 bp lange Promotorregion von *iolR* aus *Salmonella enterica* (Abb. 37D).

Die Ergebnisse untermauern die These einer geringen, aber nachweisbaren Affinität von OxyR bezüglich der putativen *epmA* Promotorregion und weisen damit auf eine regulatorische Rolle von OxyR in der Genexpression von *epmA* hin. Dies gilt sowohl für die für oxydierende, als auch die für reduzierende Bedingungen typische Konformation des Proteins.



**Abb. 36:** Verifizierte und putative OxyR Erkennungssequenzen in den Promotor-Regionen von *katG*, *oxyRS*, *epmA* und *iolR* (Gene grau hinterlegt). Die Konsensussequenzen sind durch rot, die zu testenden möglichen Bindestellen durch grün hervorgehobene Basen markiert. Die für Gelretardations-Assay und Oberflächen-Plasmonen-Resonanzspektroskopie verwendeten Sequenzabschnitte sind durch gestrichelte Linien dargestellt. Die zur Amplifikation der jeweiligen Sequenzen aus *E. coli* (St31) und *S. enterica* (St52) verwendeten *Primer* sind oberhalb davon aufgeführt. Violett: 400 bp lange Sequenzabschnitte, grün: 150 bp lange Sequenzabschnitte, blau: 188 bp lange Sequenzabschnitte.



**Abb. 37:** Analyse der Bindeaffinität von OxyR an die putative Promotor-DNA von *epmA*. (A) Gelretardationsassay zum Nachweis der Bindung von OxyR oder OxyR-C199S an Promotor-DNA-Fragmente der Länge 400 bp und 150 bp. (B) SPR Analyse zum Nachweis der Bindung von OxyR oder OxyR-C199S an 400 bp lange DNA-Fragmente aus dem *katG* Promotor (Positiv-Kontrolle). Kinetischer Screen (Modell: 1:1 Bindung) und Affinitäts-Screen. (C) SPR Analyse zum Nachweis der Bindung von OxyR oder OxyR-C199S an 400 bp lange DNA-Fragmente aus dem *epmA* Promotor. Kinetischer Screen (Modell: 1:1 Bindung) und Affinitäts-Screen. (D) SPR Analyse mit 400 bp langen DNA-Fragmenten aus dem *ioIR* Promotor und OxyR-C199S (Negativ-Kontrolle).

### 3.4.2.3 Identifizierung von $P_{epmA}$ bindenden Proteinen mittels DNA-Affinitätschromatographie

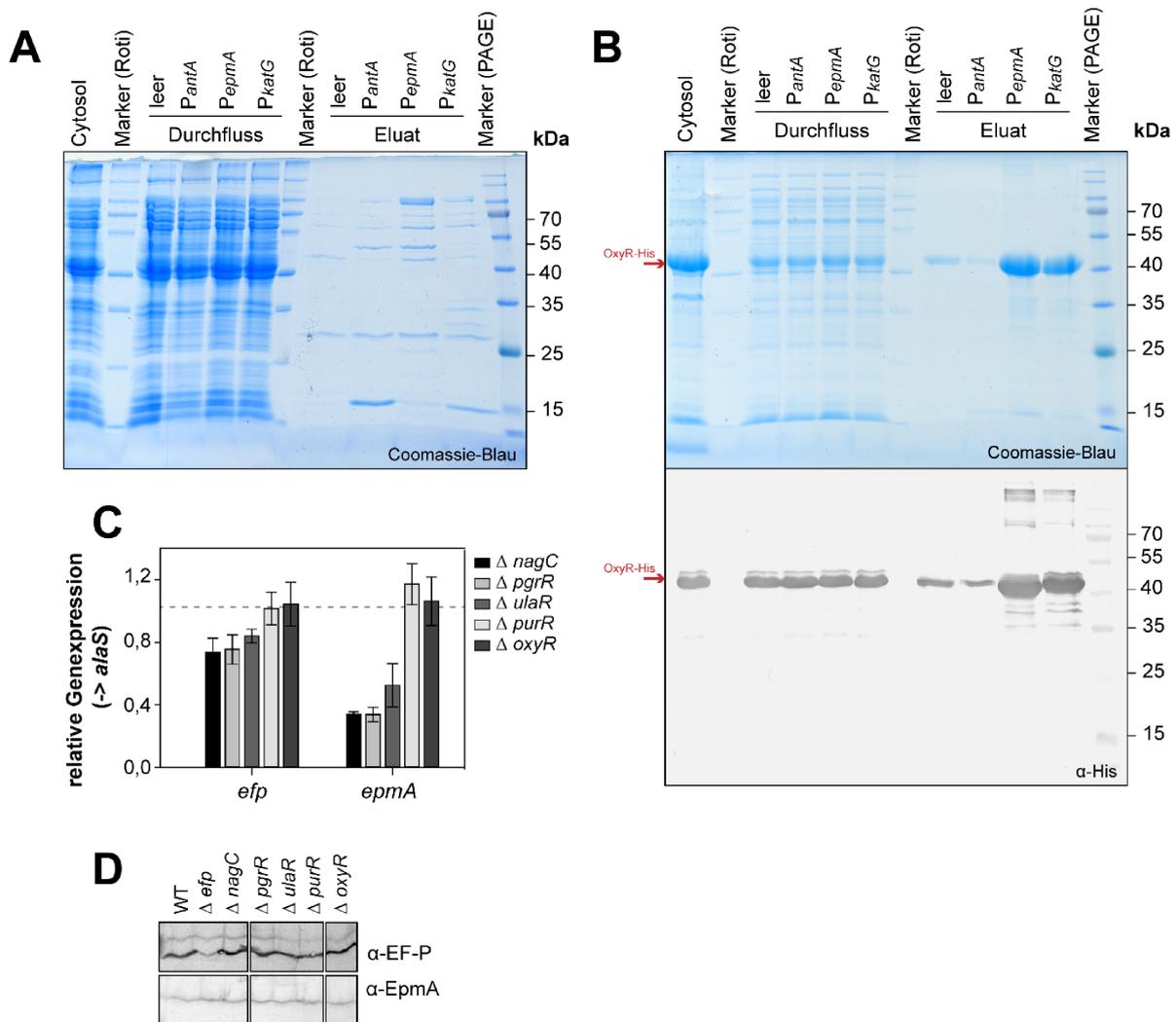
Um die vorherigen Ergebnisse zu bestätigen und weitere Proteine, die an die Promotorregion von *epmA* binden, zu identifizieren, wurde eine DNA-Affinitätschromatographie durchgeführt. Dafür wurden 400 bp lange, doppelsträngige DNA-Fragmente aus genomischer DNA amplifiziert, mittels Biotin-Streptavidin-Bindung an magnetische Beads gekoppelt und in einer zytosolischen Zellfraktion von *E. coli* inkubiert. Lösliche Proteine, die an die jeweilige DNA binden, konnten auf diese Weise isoliert werden, sowie in einer SDS-PAGE und anschließender massenspektrometrischen Analyse identifiziert werden.

#### 3.4.2.3.1 Bindung von NagC, PurR und PgrR an $P_{epmA}$

Bei den getesteten DNA-Fragmenten handelt es sich um Sequenzabschnitte, die direkt *upstream* an das Startcodon der betreffenden Gene anschließen. Als Positivkontrolle wurde die Promotorregion von *katG* (*E. coli* BW25113) mit der bekannten OxyR Bindestelle und als Negativkontrolle ein entsprechendes Fragment der Promotorregion des *Photobacterium luminescens* (TT01) Gens *antA* verwendet. AntA ist an der Synthese von Antraquinon beteiligt und, soweit bekannt, nicht mit OxyR assoziiert<sup>377</sup>.

Zuerst wurden die DNA-Fragmente mit einer cytosolischen Zellfraktion von *E. coli* BW25113 (Proteinkonzentration 16 mg/ml) aus der frühen exponentiellen Wachstumsphase inkubiert. Die gebundenen Proteine wurden nach einer Elution mittels SDS-PAGE aufgetrennt (Abb. 38A). Im coomassieblaugefärbten Gel zeigen sich sowohl Proteinbanden, die nur in einem der drei Eluate auftauchen (im Fall der *katG* Promotor-DNA z. B. bei 35 kDa), unterschiedlich prominente Banden (z. B. bei 50 kDa in allen drei Eluat) oder Banden, die in allen drei Proben gleich stark sind (z. B. bei 30 kDa). Unspezifisch an die leeren Beads bindende Proteine treten nur in vernachlässigbarer Anzahl auf.

Mittels Massenspektrometrie wurden die eluierten Proteine identifiziert und relativ quantifiziert. Die Ergebnisse sind in Tabelle 14 gelistet. Neben Proteinen, die an der DNA-Modifikation beteiligt sind (z. B. Topoisomerase TopA), und sehr abundanten Proteinen, die durch die Waschschrte nicht vollständig entfernt werden konnten (z.B. ribosomale Proteine wie RpsC), wurde auch eine Vielzahl von Transkriptionsfaktoren identifiziert. So konnte unter anderem die Bindung von OxyR an die *katG* Promotor-DNA bestätigt werden. Der Transkriptionsfaktor bindet darüber hinaus in signifikanter Weise die Promotorregion von *epmA*. Die relative Anzahl der eluierten Proteine ist zwar ungefähr 9-mal geringer als es für  $P_{katG}$  der Fall ist, jedoch 10-mal höher, als bei der Negativkontroll-DNA  $P_{antA}$ , für welche die Anzahl an OxyR-spezifischen Peptiden am unteren Limit der Nachweisgrenze liegt. Transkriptionsfaktoren, die präferentiell an die Promotorregion von *epmA* binden, sind NagC, ArgP, MraZ, PurR und PgrR.



**Abb. 38:** Verifizierung der Bindung von OxyR an  $P_{epmA}$  und Identifizierung weiterer  $P_{epmA}$  bindender Transkriptionsfaktoren. (A) DNA-Affinitätschromatographische Identifizierung von  $P_{epmA}$  bindenden Proteinen aus *E. coli* MG1655. Als Positivkontrolle wurden DNA-Sequenzen aus  $P_{katG}$  und als Negativkontrolle aus  $P_{antA}$  verwendet. Analyse der DNA-bindenden und nichtbindenden Proteine mittels SDS-PAGE und Anfärbung mit Coomassie-Blau. „Cytosol“: lösliche Proteine vor Inkubation mit DNA-gekoppelten *beads*, „Durchfluss“: nicht gebundene Proteine, „Eluat“: an DNA-gekoppelte *beads* bindende Proteine nach Eluation. (B) DNA-Affinitätschromatographische Identifizierung von  $P_{epmA}$  bindenden Proteinen aus OxyR-His überexprimierenden *E. coli* BL21 DE3. Oben: Coomassie-Blaufärbung des SDS-Gels, unten: Westernblot-Analyse desselben mittels spezifischer  $\alpha$ -His Antikörper zur Verifizierung von OxyR-His. (C) qPCR Analyse der Genexpression von *epmA* in *E. coli* BW25113 Stämmen mit Deletion ausgewählter, putativ  $P_{epmA}$  bindender Transkriptionsfaktoren. (D) Westernblot-Analyse des Proteinlevels von EF-P und EpmA in Abwesenheit dieser Transkriptionsfaktoren mit spezifischen  $\alpha$ -EF-P oder  $\alpha$ -EpmA Antikörpern.

Material und Methoden: (A) St31. (B) St2, V53. (C und D) WT = St4,  $\Delta$ *efp* = St5,  $\Delta$ *nagC* = St17,  $\Delta$ *pgrR* = St13,  $\Delta$ *ulaR* = St14,  $\Delta$ *purR* = StSt16,  $\Delta$ *oxyR* = St10.

Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen, bzw. repräsentieren eine von drei unabhängigen Versuchsdurchführungen.

**Tab. 14:** Massenspektrometrische Identifizierung der mittels DNA-Affinitätschromatographie ermittelten Transkriptionsfaktoren, die von Beads mit gekoppelter Promotor-DNA von *antA*, *epmA* und *katG*, bzw. unbeladenen Beads eluiert wurden.

Accession Number	MW	Total Unique Peptide Count:				Accession Number	MW	Total Unique Peptide Count:			
		empty	P <sub>antA</sub>	P <sub>epmA</sub>	P <sub>katG</sub>			empty	P <sub>antA</sub>	P <sub>epmA</sub>	P <sub>katG</sub>
HNS_ECOLI	16 kDa	0	170	48	78	DEOR_ECOLI	29 kDa		8	1	14
STPA_ECOLI	15 kDa	0	117	19	53	LACI_ECOLI	39 kDa		2	1	2
LRP_ECOLI	19 kDa		15	16	35	AAER_ECOLI	35 kDa		1	1	3
<b>NAGC_ECOLI</b>	<b>45 kDa</b>		<b>0</b>	<b>15</b>		YAFC_ECOLI	34 kDa		11	0	1
NSRR_ECOLI	16 kDa	0	28	12	21	DICA_ECOLI	16 kDa		3	0	0
OXYR_ECOLI	34 kDa	0	1	11	87	MURR_ECOLI	31 kDa			0	2
CYTR_ECOLI	38 kDa		8	10	7	YCIT_ECOLI	28 kDa		16		
<b>ARGP_ECOLI</b>	<b>33 kDa</b>		<b>4</b>	<b>9</b>		CPXR_ECOLI	26 kDa		6		0
NADR_ECOLI	47 kDa		6	8	4	PAA_X_ECOLI	35 kDa		6		
<b>MRAZ_ECOLI</b>	<b>17 kDa</b>		<b>1</b>	<b>8</b>	<b>2</b>	YGBI_ECOLI	27 kDa		5		62
<b>PGRR_ECOLI</b>	<b>33 kDa</b>		<b>2</b>	<b>7</b>	<b>0</b>	CRA_ECOLI	38 kDa		5		4
RCSB_ECOLI	24 kDa		11	5	4	OMPR_ECOLI	27 kDa		5		1
INTS_ECOLI	44 kDa	0	1	4	0	YIEP_ECOLI	26 kDa		5		
YFHH_ECOLI	31 kDa		31	3		GALR_ECOLI	37 kDa		4		1
GNTR_ECOLI	36 kDa		6	3	6	DCTR_ECOLI	20 kDa		3		
ALLR_ECOLI	29 kDa		5	3	3	FUCR_ECOLI	27 kDa		2		1
<b>PURR_ECOLI</b>	<b>38 kDa</b>		<b>1</b>	<b>2</b>		NHAR_ECOLI	34 kDa		2		
ULAR_ECOLI	28 kDa		0	2	2	YJHU_ECOLI	36 kDa		1		6
CYSB_ECOLI	36 kDa		14	1	10	KDGR_ECOLI	30 kDa				3

Die Bindung von OxyR an die Promotorregion von *epmA* wurde in einem zweiten Versuchsansatz unter der Annahme, dass die zuvor beobachtete OxyR-DNA Bindung noch keine Sättigung erreicht hat, verifiziert. Diesmal wurden die an Beads gekoppelten DNA-Fragmente mit cytosolischen Proteinen eines mit pET\_SUMO\_OxyR transformierten und damit OxyR-His überexprimierenden *E. coli* BL21 Stammes (Proteinkonzentration 10,33 mg/ml) inkubiert (Abb. 38B). Tatsächlich ist OxyR das prominenteste der eluierten Proteine. Die entsprechende Bande ist nach SDS-PAGE und Anfärbung des Gels mit Coomassie-Blau im Fall von P<sub>katG</sub> und P<sub>epmA</sub> ca. 100-fach stärker als es bei P<sub>antA</sub> und den leeren Beads der Fall ist. Unspezifisch bindende Proteine sind vermutlich eine Folge der überproportional hohen Konzentration von OxyR-His. Durch einen Western Blot mit spezifischen anti-His Antikörpern wurde bestätigt, dass es sich bei der Proteinbande tatsächlich um OxyR-His handelt.

Ein Screening nach unter verschiedenen Stressbedingungen wie Phosphat- (6,7 mM), Stickstoff- (1,8 mM) oder Sauerstofflimitation spezifisch bindenden Transkriptionsfaktoren lieferte keine über die obigen Versuche hinausgehenden Ergebnisse.

#### 3.4.2.3.2 Verifizierung des Einflusses von NagC, UlaR und PgrR auf die Expression von *efp* und *epmA*

Der Einfluss von Transkriptionsfaktoren, welche nachweislich an die Promotor-DNA von *epmA* binden, wurde mittels quantitativer *realtime* PCR überprüft. Dazu wurden Stämme untersucht, in welchen die für die Transkriptionsfaktoren kodierenden Gene durch die Insertion einer Kanamycin-Kassette deletiert sind.

Das Expressionslevel von *epmA* ist in Abwesenheit von *nagC*, *ulaR* und *pgrR* signifikant um das 0,3-, 0,5- und 0,3-fache reduziert (Abb. 38C). Alle drei Transkriptionsfaktoren besitzen ein *helix-turn-helix* Bindemotiv und regulieren wichtige metabolische Prozesse<sup>378-381</sup>. Während NagC am Kohlenhydratmetabolismus beteiligt ist, kontrolliert UlaR den L-Ascorbatmetabolismus<sup>382, 383</sup>. PgrR spielt eine Rolle in der Peptidoglykandegradierung<sup>384</sup>. Die Veränderung des Expressionslevels von *epmA* in Folge einer Deletion der Gene, die für diese Transkriptionsfaktoren kodieren, ist allerdings nicht groß genug, um Auswirkungen auf das EpmA Proteinlevel im Western Blot auflösen zu können (Abb. 38D).

### 3.4.3 Expressionsmuster von Proteinen mit polyP Motiv in VBNCs

Wie es für die meisten Bakterien der Fall ist, wird auch *E. coli* aus praktischen Gründen im Labor standardmäßig unter möglichst optimalen Wachstumsbedingungen kultiviert und die exponentielle Wachstumsphase als Referenzzustand herbeigezogen. Dabei gerät jedoch die Tatsache in den Hintergrund, dass Bakterien in ihrem natürlichen Habitat die meiste Zeit einer Substratlimitation ausgesetzt sind und sich in einem katabolischen Ruhezustand befinden, bis sich eine Gelegenheit zur Zellteilung ergibt<sup>385, 386</sup>. Aus diesem Grund können Freilandisolate verschiedener Bakteriengattungen oftmals nicht im Labor kultiviert werden. Der Verlust der Kultivierbarkeit (VBNC, *viable but non-culturable cells*), kann nur durch das Vorhandensein bestimmter Reize beendet werden<sup>387</sup>. Dieses Phänomen wurde zuerst in *E. coli* und *V. cholerae* erkannt<sup>386, 388</sup>. Durch verschiedene externe Stimuli, wie Langzeitkultivierung bei niedriger Temperatur, Nährstofflimitation, hohe Osmolarität oder Strahlung können Bakterienzellen künstlich in einen vorübergehend unkultivierbaren Zustand (VBNC) versetzt werden<sup>389</sup>.

Der Wachstumsstillstand von VBNCs manifestiert sich vor allem in einer verlangsamten Ribosomenbiogenese und damit einhergehenden Proteinbiosynthese, der vermehrten Synthese von Membranphospholipiden, einer Kondensierung des Nukleoids und schließlich einem reduzierten Gesamtvolumen<sup>387</sup>. Dieser Phänotyp wird maßgeblich durch den Sigmafaktor S (RpoS) und den Sekundärmetaboliten (p)ppGpp hergestellt<sup>387</sup>. Weitere lebensnotwendige Prozesse, wie die Aufrechterhaltung des elektrochemischen Protonengradienten, werden beibehalten<sup>390, 391</sup>.

Das pathogene Bakterium *Pseudomonas aeruginosa* liegt in infiziertem Lungengewebe in einem Ruhezustand vor, der eine Überdauerung der konstanten Sauerstofflimitation erlaubt. Eine Simulierung der anaeroben Bedingungen im Labor führt zu erhöhten Proteinleveln von GreA, S10 und EF-P (2,02-, 3,50- und 2,35-fach)<sup>392</sup>. Daraus wurde der Rückschluss gezogen, dass die durch Substratlimitation ins Stocken geratene Transkription und Translation mit Hilfe einer Anhebung der Kopienzahl bestimmter Transkriptions- sowie Translationselongationsfaktoren ausgeglichen werden kann<sup>387</sup>.

Entgegen dieser Hypothese wäre jedoch auch denkbar, dass eine Absenkung der Anzahl an aktiven EF-P Molekülen die Translation von Proteinen, die nicht relevant zur Überbrückung des Ruhezustandes sind, reduziert, womit wertvolle Ressourcen eingespart werden könnten. Wie bereits gezeigt, ist *efp* in *E. coli* und in zahlreichen anderen Bakterien nicht essentiell, bzw. führt eine Deletion des Gens zwar zu Wachstumsdefiziten, ist aber nicht lethal<sup>100, 99</sup>. Daraus lässt sich ableiten, dass keines der von EF-P (im Gegensatz zu eIF5A, bzw. aIF5A) abhängigen Proteine eine essentielle, lebensnotwendige Rolle spielt.

### 3.4.3.1 Etablierung eines Verfahrens zur Separierung von viablen *E. coli* Zellen aus Langzeitkulturen

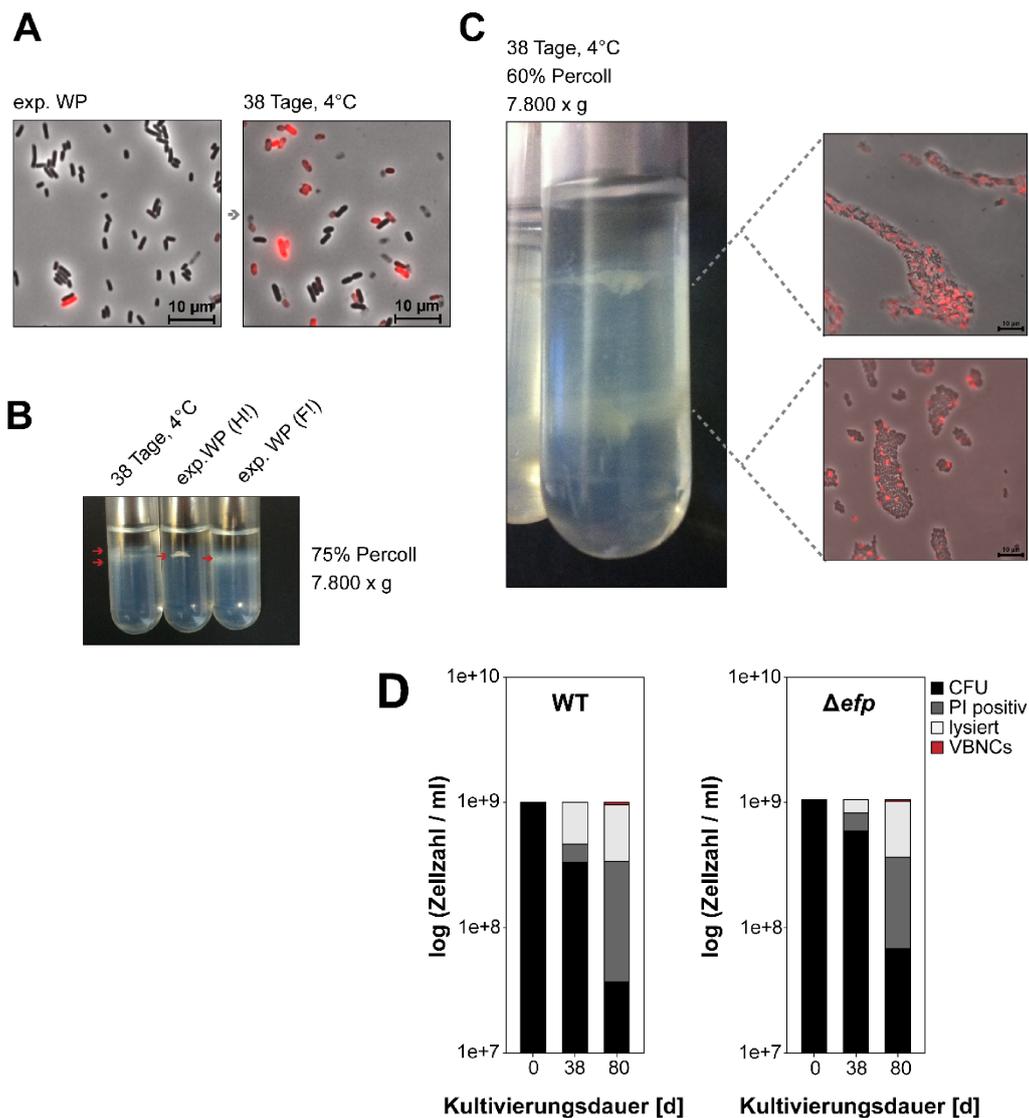
Um zu klären, ob und welche Rolle EF-P während der Dormanz von *E. coli* spielt, wurden die Zellen durch Langzeitkultivierung bei 4°C in einen vorübergehend unkultivierbaren Zustand (VBNC) versetzt.

Ähnlich wie bei der Persistenz besteht für das Auftreten von VBNCs eine stochastische Wahrscheinlichkeit, während ein Großteil der Bakterien abstirbt<sup>393</sup>. Um VBNCs separat untersuchen zu können, ist es nötig, sie von der restlichen Kultur abzutrennen. Zu diesem Zweck wurde im Rahmen dieser Arbeit eine auf einer Dichtegradienten-Zentrifugation basierende Methode entwickelt.

Bereits 1981 beschrieb Killik<sup>394</sup>, dass sich die kolloidalen, polyvinylpyrrolidonummantelten Silikatpartikel der von Pertoft *et al.* (1987)<sup>395</sup> entwickelten Percoll-Lösung dazu eignen, um verschiedene sich in ihrer Form unterscheidende Phänotypen derselben Zellen zu separieren. Dabei handelte es sich um *Dyctiostelium discoideum* Sporen und Stängelzellen. Eine ähnliche Methode wurde für eine erfolgreiche Separierung von *Caulobacter crescentis* Schwärmer- und sessilen Zellen angewandt<sup>396</sup>. Es stellte sich daher die Frage, ob die Unterschiede in der Partikelform und -größe von intakten, viablen und abgestorbenen oder bereits degradierten *E. coli* Zellen groß genug sind, um eine ähnliche Trennung durchzuführen.

Mit Hilfe einer Testreihe wurde nach einer geeigneten Percoll-Verdünnung und Zentrifugationsmethode geforscht, welche alle Voraussetzungen für einen geeigneten Gradienten erfüllt. Dabei kamen *E. coli* K-12 MG1655 Kulturen, welche bis zu einer OD<sub>600</sub> von 1 bei 37°C kultiviert und anschließend durch fünfminütiges Aufkochen bei 95°C, bzw. Hochdruck aufgeschlossen oder zur Induktion des VBNC Zustands anschließend 31 Tage bei 4°C inkubiert wurden, zur Anwendung. Letztere setzt sich zusammen aus intakten, lebenden und abgestorbenen Zellen (Abb. 39A). Zur Ermittlung geeigneter Parameter wurde deren Separierung in verschiedenen konzentrierten Percoll-Lösungen (25-100% in PBS-Puffer pH 8,0) untersucht. Der Dichtegradient wurde durch eine einstündige Zentrifugation erstellt, wobei Geschwindigkeiten von 180 – 7.800 x g getestet wurden.

In allen Fällen bildete sich mindestens eine distinkte Zellschicht aus, in einigen Ansätzen konnte im Fall der bei 4°C inkubierten Zellen eine zweite Zellschicht festgestellt werden (Abb. 39B). Eine optimale Trennung fand in einer 60% Percoll-Lösung bei 7.800 x g statt. Zellen aus beiden Schichten wurden mit einer Pipette entnommen, mit Propidiumiodid zur Unterscheidung von abgestorbenen Zellen angefärbt und mikroskopisch untersucht (Abb. 39C). Die im oberen Bereich des Dichtegradienten liegende Zellschicht besteht vollständig aus abgestorbenen Zellen und Zellfragmenten. Im Gegensatz dazu finden sich in der darunter liegenden Schicht über 95% lebende Zellen. Zwar finden sich vereinzelt auch abgestorbene Zellen, diese sind aber möglicherweise eine Folge der harschen Zentrifugationsbedingung. Die Separierung von viablen Zellen und den eine geringere Dichte aufweisenden abgestorbenen, sowie fragmentierten Zellen ist mittels der hier beschriebenen, auf Percoll basierenden Dichtegradientenzentrifugation möglich.



**Abb. 39:** Analyse des Anteils an intakten und potentiell lebenden im Vergleich zu abgestorbenen und fragmentierten *E. coli* K-12 MG1655 Zellen aus Langzeitkulturen nach Induzierung des VBNC Zustandes. (A) Zustand einer *E. coli* K-12 MG1655 Zellkultur nach Kälte-Induzierung des VBNC Zustandes und 38-tägiger Inkubation im Vergleich zu Zellen in der exponentiellen Wachstumsphase (exp. WP). Anfärbung von abgestorbenen Zellen mit Propidiumiodid und Fluoreszenzmikroskopie (DMI8) von auf Agar-Pads immobilisierten Zellen. (B) Vergleichende Dichtegradientenzentrifugation von Zellen nach Induktion des VBNC Zustandes und 38-tägiger Inkubation, sowie hitzeinaktivierten (HI) oder aufgeschlossenen (FI) Zellen der späten exponentiellen Wachstumsphase. Beispielparameter. (C) Verifizierung des Zustands der nach Dichtegradientenzentrifugation in zwei distinkten Schichten enthaltenen Zellen (38 Tage Inkubation bei 4°C). Anfärbung von abgestorbenen Zellen mit Propidiumiodid und Fluoreszenzmikroskopie. (D) Nachverfolgung der Veränderungen der Zellzahlen von lysierten („lysiert“), abgestorbenen („PI positiv“), lebenden und kultivierbaren („CFU“), sowie lebenden und unkultivierbaren („VBNC“) Zellen in *E. coli* K-12 MG1655 Kulturen nach 38 und 80 Tagen Inkubation bei 4°C.

Material und Methoden: WT = St31  $\Delta$ efp = St40.

### 3.4.3.2 Einfluss von EF-P auf das Erreichen des VBNC Zustands

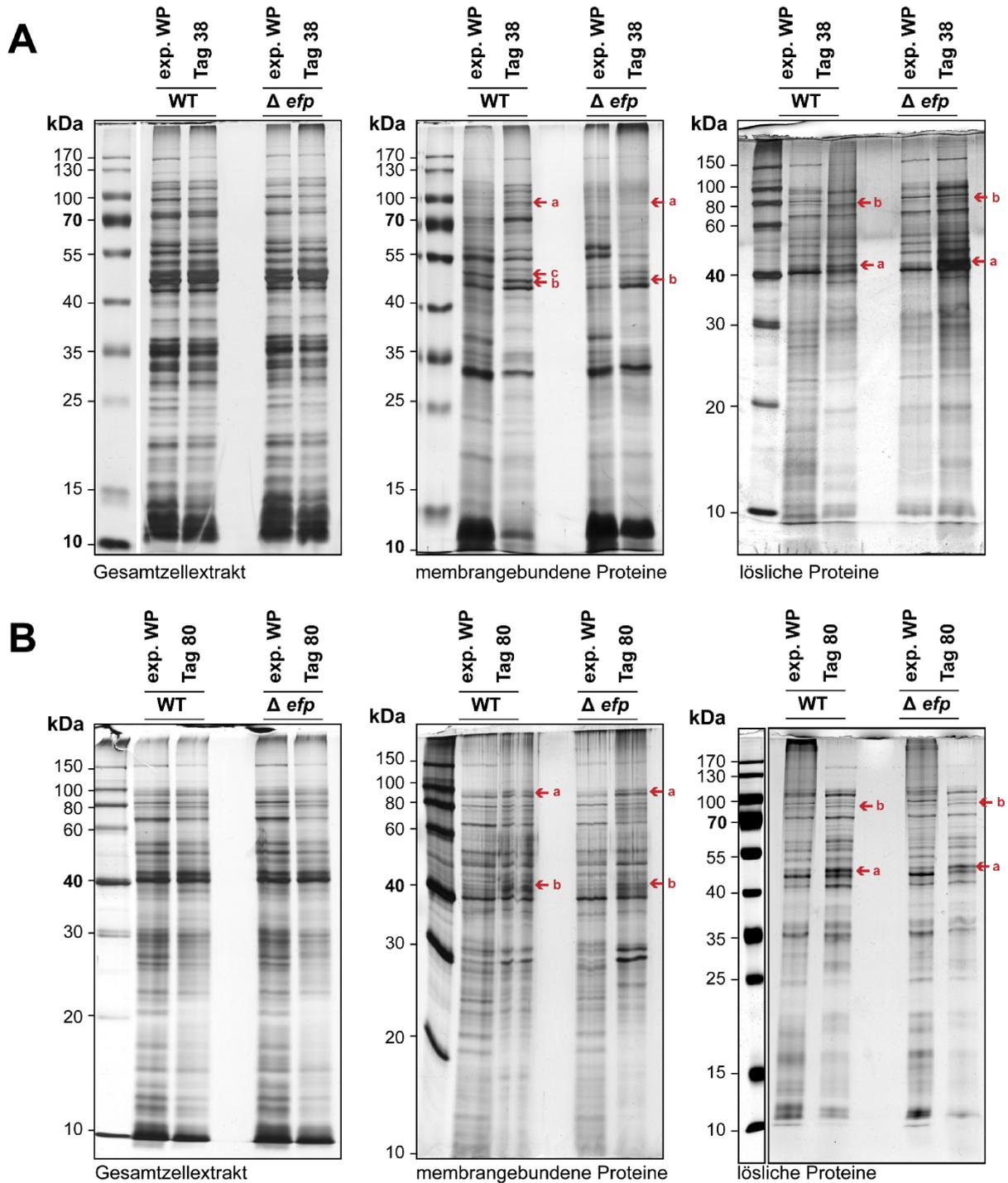
Um die Rolle von EF-P, bzw. von EF-P abhängigen polyP Motiven in der Etablierung von VBNCs zu testen wurden ein *E. coli* K-12 MG1655 Wildtypstamm, sowie eine  $\Delta\text{efp}$  Mutante untersucht.

Der VBNC Zustand wurde wie im Methodenteil beschrieben mittels Kälte (4°C) und 120-tägiger Langzeitkultivierung induziert. Eine Kontamination wurde durch eine Sequenzierung der Stämme bei Versuchsende ausgeschlossen. Durch die Verfügbarkeit aller Nährstoffe ist während der gesamten Versuchsdauer eine funktionelle Proteinbiosynthese gewährleistet. Die niedrige Temperatur verlangsamt jedoch alle enzymatischen Reaktionen und stellt eine enorme Stresssituation für das Bakterium dar<sup>389</sup>. Während des gesamten Zeitraums der Inkubation bei 4°C wurden regelmäßig Proben entnommen und die Kultivierbarkeit der Bakterien überprüft. Dabei ließ sich kein Unterschied zwischen Wildtyp und  $\Delta\text{efp}$  Mutante feststellen. Nach 38 Tagen war die Anzahl der koloniebildenden Bakterien in beiden Kulturen auf 50% gesunken, nach 80 Tagen auf 5% und nach 120 Tagen auf 0,001%. Durch eine Bestimmung der Gesamtzellzahl und Propidiumiodid anfärbung mit anschließender fluoreszenzmikroskopischer und durchflusszytometrischer Analyse wurden die Zellen auf ihren physiologischen Zustand hin untersucht (Abb. 39 D). Nach 38 Tagen waren noch 50% und nach 80 Tagen noch 35% der Bakterienzellen intakt und damit zählbar. Davon waren nach 38 Tagen nur noch 70% kultivierbar, während es nach 80 Tagen 10% waren. Im Gegensatz dazu fanden sich in einer *E. coli* Kultur in der exponentiellen Wachstumsphase über 99% viable und kultivierbare Zellen. Unter Berücksichtigung eines Fehlers im Auszählen der Zellzahl von 10%, wurde die Anzahl der weder kultivierbaren, noch durch Propidiumiodid anfärbaren Zellen berechnet. Demnach befanden sich nach 38 Tagen 0% und nach 80 Tagen 1% der Ausgangspopulation in einem viablen, aber unkultivierbaren Zustand (Abb 39D).

Da die Menge an abgestorbenen Zellen und Zellfragmenten nach 120 Tagen trotz der konservierenden Bedingungen überproportional hoch war, war eine Bestimmung der Gesamtzellzahl und Anfärbung mit Propidiumiodid in diesem Zustand weder sinnvoll noch zielführend. Parallel in dieser Arbeitsgruppe durchgeführte und von Vilhena *et al.* (2018)<sup>397</sup> publizierte Untersuchungen zeigten, dass nach 120 Tagen 8% der verbleibenden intakten Zellen in einem VBNC Zustand vorliegen.

### 3.4.3.3 Zusammenhang zwischen der EF-P Aktivität und der Anzahl von Proteinen mit polyP Motiv im VBNC Zustand

Um die Proteinbiosynthese von viablen Wildtyp und  $\Delta\text{efp}$  Zellen zu untersuchen, wurden diese nach 38 und 80 Tagen mittels Percoll-Dichtegradientenzentrifugation isoliert und membranständige, sowie lösliche Proteine durch eine SDS-PAGE analysiert (Abb. 40A und B). Bereits nach 38 Tagen zeigte sich ein deutlich von Zellen aus der exponentiellen Wachstumsphase abweichendes Muster. Die Proteine besonders prominenter Banden wurden massenspektrometrisch identifiziert. Folgende Membranproteine waren stärker präsent, als in der exponentiellen Wachstumsphase: das Chaperon ClpB und die die NADH Dehydrogenase Ndh. Im Wildtyp ist eine der ATP-Synthaseuntereinheit AtpD zugeordnete Proteinbande nach 38 Tagen schwächer, als in der Vergleichsprobe aus der exponentiellen Wachstumsphase. Die cytosolische Tryptophanase TnaA und die Alanyl-tRNA-Synthetase AlaS wiesen bereits nach 38 Tagen erhöhte, bzw. reduzierte Level auf.



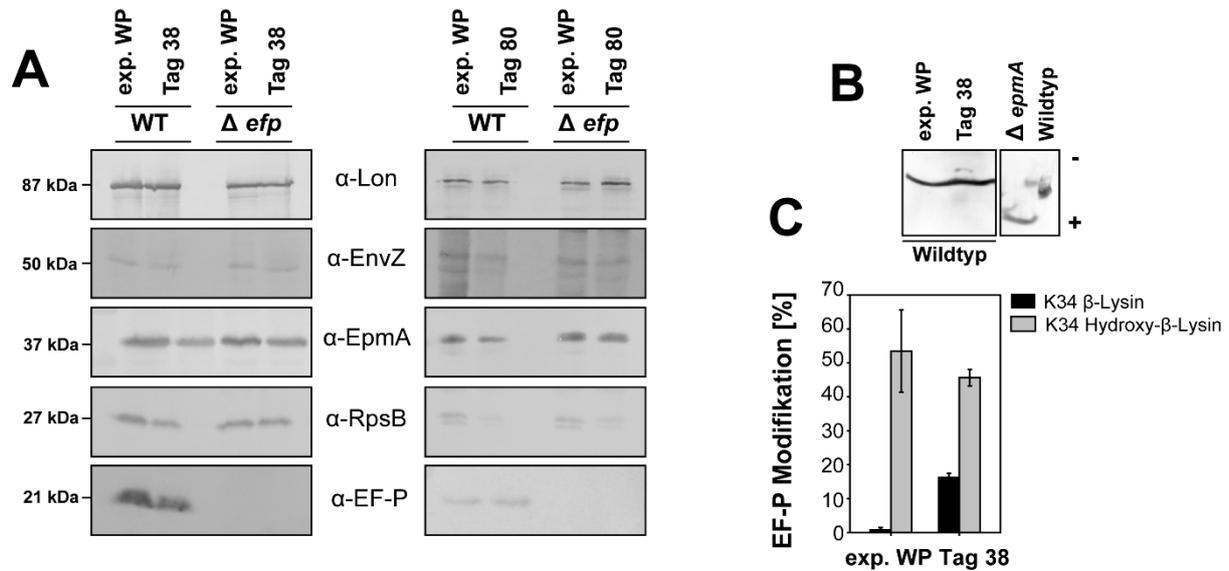
**Abb. 40:** Charakterisierung des Proteoms von lebenden *E. coli* K-12 MG1655 Wildtyp (St31) und  $\Delta$ *efp* Zellen (St40) nach 38 (A), bzw. 80 (B) Tagen Langzeitkultivierung bei 4°C. SDS-PAGE mit anschließender Silbernitrat-Färbung zur Analyse der gesamten, löslichen und membrangebundenen Proteine. Rote Pfeile markieren Proteinbanden, die sich von der entsprechenden Bandenintensität der Kontrollprobe, Zellen aus der mittleren exponentiellen Wachstumsphase (exp. WP), unterscheiden. Membrangebundene Proteine: a = ClpB, b = Ndh, c = AtpD; lösliche Proteine: a = TnaA, b = AlaS.

Mittels Western Blot wurden ausgewählte Proteine der nach 38 und 80 Tagen entnommenen Zellen immunologisch detektiert und relativ quantifiziert (Abb. 41A). Als Referenz wurde das ribosomale Protein RpsB (S2) herangezogen, welches nach 38 Tagen keine Veränderung zur exponentiellen Wachstumsphase aufweist und

nach Tag 80 in geringfügig kleineren Kopienzahlen vorliegt. Die Proteine EnvZ und Lon tragen polyP Motive, welche laut Vorhersage einen „starken“ Translationsarrest induzieren und damit potentiell abhängig von EF-P sind. Es konnte jedoch weder nach 38 noch nach 80 Tagen eine Veränderung deren Proteinlevels festgestellt werden. Auch das Proteinlevel von EF-P blieb konstant. Eine leichte Absenkung der Kopienzahl von EpmA lies sich hingegen in Wildtyp und  $\Delta efp$  Zellen nach 38 und 80 Tagen beobachten.

Um zu überprüfen, inwiefern das nach 38 Tagen reduzierte EpmA-Level eine Auswirkung auf den Modifikationszustand von EF-P hat, wurde das Protein nach einem tryptischen Verdau massenspektrometrisch untersucht (Abb. 41B). Als Referenz dienten die EF-P Peptidsequenzen „VPLFVQIGE“, „FWHFMNNE“ und „FWHFM[+16]NNE“, wobei letztere eine Methioninoxydierung beinhaltet, welche typisch für die hier angewendete Probenaufbereitung ist. Relativ dazu wurden je zwei Peptide unterschiedlicher Länge mit Lysin- (128,1 Da) bzw. Oxylysinmodifikation (144,1 Da) an Position K34 quantifiziert: „FVKPGK[+128.1]GQAFAR“ und „ASEFVKPGK[+128.1]GQAFAR“, bzw. „FVKPGK[+144.1]GQAFAR“ und „ASEFVKPGK[+144.1]GQAFAR“. Unter der Annahme, dass die Anzahl der Referenzpeptide konstant bleibt, ist nach 38-tägiger Inkubation im Vergleich zur exponentiellen Wachstumsphase eine  $20,77 \pm 1,73$ -fache Anhebung des Levels der  $\beta$ -Lysilylierung zu verzeichnen. Im Vergleich dazu bleibt das Level der Hydroxylysinmodifikation gleich ( $0,85 \pm 0,20$ -fache Veränderung). Diese Modifikation begründet sich sowohl durch EpmA vermittelte  $\beta$ -Lysilylierung (128,1 Da), als auch durch die EpmC vermittelte Hydroxylierung (15,995 Da). Bezeichnenderweise konnten keine Peptide mit unmodifiziertem K34 identifiziert werden. Dabei ist allerdings nicht auszuschließen, dass unmodifizierte Peptide weniger stabil, damit zugänglicher für den tryptischen Verdau sind und deshalb nicht erfasst wurden. Die Ergebnisse widerlegen demnach eine Absenkung der EpmA Kopienzahl im Zuge der Langzeitkultivierung. Möglicherweise ist eine erhöhte Zahl an Peptiden mit K34 Lysinmodifikation im Vergleich zu Referenzpeptiden eine Folge der bei 4°C verlangsamten Proteindegradation. Ein erhöhtes Level an Peptiden mit K34 Lysinmodifikation könnte andererseits auch auf eine geringere Aktivität von EpmC hindeuten.

Der Modifikationszustand von EF-P nach Tag 38 wurde ergänzend mittels isoelektrischer Fokussierung und anschließendem Western Blot untersucht. (Abb. 41C) Als Referenz dienten zwei bis in die exponentielle Wachstumsphase kultivierte Stämme: ein Wildtyp *E. coli* Stamm und eine  $\Delta epmA$  Mutante, in welcher EF-P nur in unmodifizierter Form vorliegt. Dabei zeigte sich, dass sowohl in der exponentiellen Wachstumsphase, als auch nach Tag 38 ausschließlich EF-P Moleküle mit identischer Aminosäurezusammensetzung vorliegen, welche der einer gereinigten EF-P Variante mit K34-Hydroxy- $\beta$ -lysinmodifikation entspricht. Eine Auflösung zwischen Proteinen mit  $\beta$ -Lysin- und solchen mit Hydroxy- $\beta$ -lysinmodifikation ist mit dieser Methode nicht möglich. Alles deutet darauf hin, dass EF-P unter beiden Bedingungen zu annähernd 100% modifiziert ist.



**Abb. 41:** Charakterisierung der Level von ausgewählten Proteinen in lebenden *E. coli* K-12 MG1655 Wildtyp und  $\Delta efp$  Zellen nach Langzeitkultivierung bei 4°C. (B) Western-Blot Analyse mit spezifischen Antikörpern zur Analyse ausgewählter Proteinlevel nach 38 und 80 Tagen. (C) Isoelektrische Fokussierung von Proteinen aus einem Gesamtzellextrakt von *E. coli* K-12 MG1655 Wildtypzellen und anschließende Detektion im Western Blot mit spezifischen  $\alpha$ -EF-P Antikörpern zur Ermittlung des Modifikationszustandes von EF-P nach 38 Tagen. (D) Massenspektrometrische Bestimmung des Modifikationslevels von EF-P nach 38 Tagen. Prozentualer Anteil von EF-P Peptiden mit K34 Modifikation im Vergleich zu Referenzpeptiden aus einem anderem Proteinabschnitt (vergleiche Text).

Material und Methoden: (A und C) WT = St31,  $\Delta efp$  = St40. (B) WT= St4,  $\Delta epmA$  = St11.

Die gezeigten Resultate repräsentieren jeweils die Mittelwerte und Standardabweichungen von Daten aus drei voneinander unabhängig durchgeführten Versuchen, bzw. repräsentieren eine von drei unabhängigen Versuchsdurchführungen.

Nach 120-tägiger Inkubation wurde das gesamte Proteom des Wildtyp *E. coli* Stamms mittels Massenspektrometrie untersucht. Dabei wurden 2.399 verschiedene Proteine erfasst (Tab. 15). Der vollständige Datensatz ist im „ProteomeXChange Consortium“ über das Archiv „PRIDE“ (<http://www.ebi.ac.uk/pride>) mit der Datenidentifikationsnummer PXD009463 hinterlegt.

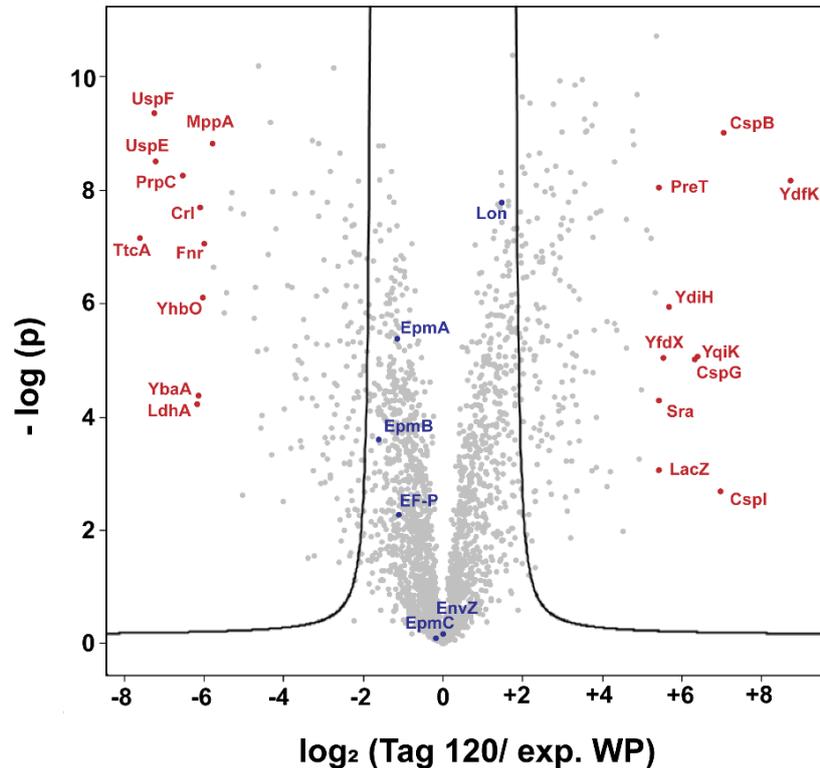
Im Vergleich zu Zellen aus der exponentiellen Wachstumsphase sind die Level einzelner Proteine um bis zu 418,65-fach erhöht (YdfK) und um bis zu 0,0051-fach reduziert (TtcA). In insgesamt 1.455 Proteinen konnten signifikant abweichende Proteinzahlen festgestellt werden (t-Test, Mehrfachvergleich, p-Wert < 0,05), wobei 565 Proteine erhöhte und 658 Proteine reduzierte Level im Vergleich zur exponentiellen Wachstumsphase aufwiesen (Abb. 42). Dabei bestätigten sich die bereits nach 38 und 80 Tagen festgestellten veränderten Kopienzahlen von ClpB (6,13-fach), TnaA (14,16-fach) und AlaS (0,53-fach). Die Level von EF-P, EpmA, EpmB und EpmC sind auf das je 0,48-, 0,45-, 0,33- und 0,94-fache reduziert (nicht signifikant) und gleichen damit den Leveln von anderen für die Proteinbiosynthese wichtigen Proteinen. Dazu zählen die Alanyl-tRNA-Synthetase AlaS (0,53-fach), 21 der ribosomalen 30S-Untereinheit Proteinen RpsA-U (0,11-fache bis 0,49-fache Proteinlevel im Vergleich zur exponentiellen Wachstumsphase), oder die Translationselongationsfaktoren TufB (0,59-fach) und Tsf (0,73-fach). Das Level des mit der stationären Wachstumsphase assoziierten RpsV Proteins der 30S-Untereinheit ist dagegen um das 42,26-fache erhöht.

Unter den 656 Proteinen mit signifikant erhöhter Kopienzahl finden sich 268 (40,9%) Proteine mit polyP Motiv, davon je 62 (10,9%) mit starken, 64 (9,7%) mit moderaten und 142 (21,6%) mit schwachem Motiv. Im Vergleich dazu finden sich in den 799 Proteinen mit signifikant reduzierter Kopienzahl 286 Proteine mit insgesamt 68 (8,5%) starken, 79 (9,9%) moderaten und 149 (18,6%) schwachen polyP Motiven. Es kann keine signifikante Abweichung von einer Zufallsverteilung festgestellt werden (zweiseitiger t-Test, 95% Konfidenzintervall).

Die relativen Level von 52 bzw. 47 Proteinen sind im Vergleich zur exponentiellen Wachstumsphase um je das mindestens 10-fache erhöht bzw. reduziert (Tab 16). Bei den Proteinen mit derartig erhöhter Kopienzahl handelt es sich in erster Linie um solche, die mit Stress und insbesondere Kältestress assoziiert sind, wie BhsA, YdfK, CspB, CspI, CspG, BluF und BluR. Dazu zählt außerdem eine auffällig hohe Zahl an Proteinen, die Teil des Glycerophospholipidmetabolismus (GlpA, GlpB, GlpD, GlpT, GlpQ, SraA, YtfE, BhsA) und des Zuckermetabolismus (LacZ, TreC, SrlD, FucO) sind. Deutlich reduzierte Proteinlevel finden sich im Fall von mit andersartigem Stress assoziierten Proteinen (SodC, HslJ, HdeA) und den ribosomalen Proteinen RpmG und RpmH, sowie Transportern wie YojI und PlaP. Proteine, die im Vergleich zur exponentiellen Wachstumsphase weniger abundant sind haben unterschiedlichen Funktionen und sind Teil eines breiten Spektrums an Stoffwechselwegen. Interessanterweise sind Level von Proteinen des Pyruvatmetabolismus (PflA 5,6-fach; PckA 4,6-fach) und des Pyruvattransporters BtsT (7,68-fach) in VBNCs erhöht, was im Rahmen einer darauf aufbauenden Analyse zur Bedeutung der Pyruvat-Wahrnehmung für das Beenden des VBNC Zustands detaillierter untersucht wurde (vgl. Vilhena *et al.*, 2018<sup>397</sup>).

**Tab. 15:** Verteilung von polyP Motiven in allen Proteinen mit signifikant veränderten Proteinleveln nach 120-tägiger Inkubation bei 4°C im Vergleich zur exponentiellen Wachstumsphase.

	Gesamtanzahl Proteine			Anzahl Proteine mit polyP Motiven			Anzahl Proteine mit "starken" polyP Motiven			Anzahl Proteine mit "moderaten" polyP Motiven			Anzahl Proteine mit "schwachen" polyP Motiven		
	absolut	relativ	n polyP	absolut	relativ	n polyP	absolut	relativ	n polyP	absolut	relativ	n polyP	absolut	relativ	n polyP
Gesamtproteom	4313	100%	2087	1458	33,8%	2087	425	9,8%	478	409	9,5%	446	952	22,1%	1163
VBNC Tag 120; erhöhte Proteinlevel	656	100%	307	229	34,9%	307	62	9,5%	69	64	9,7%	67	142	21,6%	171
VBNC Tag 120; reduzierte Proteinlevel	799	100%	346	286	35,8%	346	68	8,5%	81	79	9,9%	86	149	18,6%	179
	Anzahl Proteine mit n=1 polyP Motiven			Anzahl Proteine mit n>1 polyP Motiven											
	absolut	relativ	n polyP	absolut	relativ	n polyP									
Gesamtproteom	1003	23,3%	1003	455	10,5%	1084									
VBNC Tag 120; erhöhte Proteinlevel	167	25,5%	167	62	9,5%	140									
VBNC Tag 120; reduzierte Proteinlevel	195	24,4%	195	91	11,4%	197									



**Abb 42:** Massenspektrometrische Analyse des Proteoms von lebenden *E. coli* K-12 MG1655 Wildtypzellen nach 120-tägiger Inkubation bei 4°C im Vergleich zu Zellen aus der mittleren exponentiellen Wachstumsphase. Die n-fache Änderung der Proteinlevel wird als das  $\log_2$ -Verhältnis der relativen Proteinabundanz der beiden Proben dargestellt. Die Signifikanz dieser Änderung wird durch den negativen Logarithmus der p-Werte eines t-Tests (Mehrfachvergleich) beschrieben. Exemplarische Proteinlevel, die im Vergleich zur exponentiellen Wachstumsphase nach 120 Tagen um das über 10-fache reduziert, bzw. erhöht sind, sind in rot markiert. Die für diese Arbeit relevanten Proteinlevel von EF-P, EpmA, EpmC und der Signalproteine EnvZ und Lon Protease sind blau hervorgehoben, wobei deren Veränderung nicht signifikant ist.

Die Daten basieren auf drei voneinander unabhängig durchgeführten Versuchen. Das Diagramm wurde erstellt von Dr. Ignasi Forné (Biomedizinisches Centrum, LMU München) und 2018 publiziert<sup>397</sup>.

**Tab. 16:** Detaillierte Übersicht über die massenspektrometrisch identifizierten Proteine, deren Level im Vergleich zur exponentiellen Wachstumsphase nach 120-tägiger Inkubation bei 4°C um das über 10-fache erhöht, bzw. reduziert sind. Die Anzahl der in den Proteinen auftretenden polyP Motive, sowie deren Klassifizierung (W = schwach, M = moderat, S = stark) sind im Einzelnen aufgeführt.

Tab 16: Beschreibung auf vorheriger Seite.

	Protein	UniProt ID	polyP	Proteinlevel TI20/exp	p-Wert	Protein	UniProt ID	polyP	Proteinlevel TI20/exp	p-Wert
<b>YdfK</b>	Cold shock protein	P76154		418,65	6,86E-09	<b>EcnB</b>			418,65	6,86E-09
<b>CspB</b>	Cold shock-like protein	P36995		131,98	9,74E-10	<b>YahO</b>	Enterickin B Uncharacterized protein	P0ADB7 P75694	0,10 0,09	3,08E-02 4,02E-06
<b>CspI</b>	Cold shock-like protein	P0A986		125,12	2,14E-03	<b>TatE</b>	Sec-independent protein translocase protein	P0A843	0,09	1,91E-05
<b>YqkK</b>	Inner membrane protein	P77306		83,66	8,76E-06	<b>Bfr</b>	Bacterioferritin	P0ABD3	0,09	2,14E-06
<b>CspG</b>	Cold shock-like protein	P0A978		79,85	9,71E-06	<b>DdpA</b>	Probable D,D-dipeptide-binding periplasmic protein	P76128	0,09	3,37E-04
<b>YdhH</b>	Uncharacterized protein	P64476		51,33	1,17E-06	<b>DusB</b>	rRNA-dihydropyrimidine synthase B	P0ABT5	0,09	1,46E-05
<b>YdhI</b>	Protein	P76520		46,50	8,99E-06	<b>SodC</b>	Superoxide dismutase [Cu-Zn]	P0AGD1	0,08	4,60E-05
<b>PreT</b>	NAD-dependent dihydropyrimidine dehydrogenase subunit	P76440		42,57	9,10E-09	<b>NarG</b>	Respiratory nitrate reductase 1, alpha chain	P09152	0,08	2,64E-04
<b>RpsV</b>	Stationary-phase-induced ribosome-associated protein	P68191	1 x W	42,26	5,08E-05	<b>Yde1</b>	Uncharacterized protein	P31130	0,08	1,00E-05
<b>LacZ</b>	Beta-galactosidase	P00722	2 x W	42,16	8,98E-04	<b>YohF</b>	Uncharacterized oxidoreductase	P33368	0,08	2,03E-05
<b>YdfE</b>	Iron-sulfur cluster repair protein	P69506	1 x S	41,01	1,89E-11	<b>Yjbj</b>	UHP0337 protein	P68206	0,08	1,15E-08
<b>Oxc</b>	Oxalyl-CoA decarboxylase	P0AF10		39,87	6,61E-07	<b>HslJ</b>	Heat shock protein	P52644	0,08	5,63E-06
<b>GlpA</b>	Anaerobic glycerol-3-phosphate dehydrogenase subunit A	P0A9C0	1 x W	36,96	7,14E-08	<b>WbbL</b>	Rhamnosyltransferase	P36667	0,08	1,84E-05
<b>YdeN</b>	Uncharacterized sulfatase	P77318	1 x W	34,76	5,14E-08	<b>Fis</b>	DNA-binding protein	P0A6R3	0,07	5,24E-07
<b>GlpB</b>	Anaerobic glycerol-3-phosphate dehydrogenase subunit B	P13033	1 x W	31,46	3,26E-05	<b>RpmG</b>	50S ribosomal protein	P0A7N9	0,07	3,53E-04
<b>FlaA</b>	RNA polymerase sigma factor	P0AEM6		30,52	5,45E-04	<b>LstB</b>	Autoinducer 2-binding protein	P76142	0,07	4,79E-07
<b>GlpK</b>	Glycerol kinase	P0A6F3	1 x W	29,29	2,11E-10	<b>CsID</b>	Protein YcE	P76621	0,06	7,11E-05
<b>BluF</b>	Blue light- and temperature-regulated antirepressor	P75990		27,82	1,57E-09	<b>HdeA</b>	Acid stress chaperone	P0AES9	0,06	3,08E-03
<b>GlpD</b>	Aerobic glycerol-3-phosphate dehydrogenase	P13035		27,32	9,00E-10	<b>AdhP</b>	Alkohol dehydrogenase, propanol-prefering	P39451	0,05	4,87E-08
<b>GlpT</b>	Glycerol-3-phosphate transporter	P08194	3 x W	24,28	3,30E-08	<b>Ggt</b>	Glutathione hydroxylase proenzyme	P18956	0,05	1,08E-08
<b>ArgB</b>	Acetylglutamate kinase	P0A6C8		24,07	6,08E-06	<b>AstD</b>	N-succinylglutamate 5-semialdehyde dehydrogenase	P76217	0,05	4,59E-04
<b>XdhD</b>	Probable hypoxanthine oxidase	Q46814	1 x W	22,83	1,06E-02	<b>NrdH</b>	Glutaredoxin-like protein	P21363	0,05	6,53E-10
<b>PreA</b>	NAD-dependent dihydropyrimidine dehydrogenase subunit	P25889		20,71	2,96E-06	<b>YceH</b>	Low-affinity putrescine importer	P0AA65	0,05	1,39E-07
<b>GarR</b>	2-hydroxy-3-oxopropionate reductase	P0ABQ2		19,95	3,15E-09	<b>Plap</b>	Phosphate starvation-inducible protein	P0AA47	0,05	6,31E-06
<b>ScpB</b>	Methylmalonyl-CoA decarboxylase	P52045		18,62	2,29E-06	<b>PsIF</b>	ABC transporter ATP-binding/permease protein	P0AFM4	0,04	9,20E-05
<b>FucO</b>	Lactaldehyde reductase	P0A9S1	1 x M, 1 x W	18,35	9,30E-04	<b>YojI</b>	Aldehyde dehydrogenase	P33941	0,04	1,55E-04
<b>YfjB</b>	Putative lipoprotein	P07021		17,26	2,07E-06	<b>AldB</b>	2-methylsuccinate lyase	P37685	0,04	6,52E-11
<b>BdcA</b>	Cyclic-di-GMP-binding biofilm dispersal mediator protein	P39333	1 x S	17,24	1,70E-03	<b>PrpB</b>	Outer membrane protein	P77541	0,04	7,01E-06
<b>GarL</b>	5-oxo-4-deoxy-D-glucarate aldolase	P23522		15,43	2,35E-08	<b>FecB</b>	Fe(3+) dicitrate-binding periplasmic protein	P15028	0,04	5,31E-07
<b>YgeY</b>	Uncharacterized protein	P65807		14,46	7,77E-04	<b>OmpW</b>	Probable hydrolase	P0A915	0,04	1,85E-06
<b>HcxB</b>	Hydroxycarboxylate dehydrogenase B	P30178		14,41	2,49E-08	<b>YcaC</b>	50S ribosomal protein	P21367	0,03	2,62E-08
<b>YgeV</b>	Uncharacterized sigma-54-dependent transcriptional regulator	Q46802	1 x S	14,35	1,04E-04	<b>RpmH</b>	Uncharacterized protein	P0A7P5	0,03	2,41E-03
<b>TnaA</b>	Tryptophanase	P0A853	2 x W	14,16	3,07E-10	<b>YdcJ</b>	Protein	P76097	0,03	1,10E-08
<b>YgeX</b>	Diaminopropionate ammonia-lyase	P66899		13,97	2,92E-06	<b>YciF</b>	Uncharacterized zinc-type alcohol dehydrogenase-like protein	P21362	0,03	2,16E-08
<b>CheW</b>	Chemotaxis protein	P0A964		13,64	6,85E-08	<b>YbdR</b>	Uncharacterized protein	P77316	0,02	6,38E-07
<b>BluR</b>	HTH-type transcriptional repressor	P75989		13,52	1,63E-05	<b>YbgS</b>	Protein	P0AAV6	0,02	1,47E-06
<b>StrD</b>	Sorbitol-6-phosphate 2-dehydrogenase	P05707		13,41	9,00E-04	<b>PhnB</b>	Periplasmic murein peptide-binding protein	P16681	0,02	2,27E-07
<b>GlpQ</b>	Glycerophosphodiester phosphodiesterase, periplasmic	P09394		13,26	7,38E-10	<b>MppA</b>	Fumarate and nitrate reduction regulatory protein	P77348	0,02	1,50E-09
<b>RibB</b>	3,4-dihydroxy-2-butanone 4-phosphate synthase	P0A710		12,82	3,12E-08	<b>Fnr</b>	Protein/mucic acid dehydrogenase 2	P0A9E5	0,02	8,89E-08
<b>TrcC</b>	Trehalose-6-phosphate hydrolase	P28904	1 x W	12,17	9,17E-10	<b>YhbO</b>	Sigma factor-binding protein	P45470	0,02	7,85E-07
<b>HmpP</b>	Flavochemoprotein	P24232		12,07	9,57E-10	<b>Ch</b>	Uncharacterized protein	P24251	0,01	2,07E-08
<b>GlpC</b>	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	P0A996		12,03	1,60E-05	<b>YbaA</b>	D-lactate dehydrogenase	P0AAO6	0,01	4,25E-05
<b>LldD</b>	L-lactate dehydrogenase	P33232	1 x W	11,74	5,62E-10	<b>LdhA</b>	2-methylcitrate synthase	P52643	0,01	6,06E-05
<b>GudD</b>	Glucate dehydratase	P0AES2		11,72	8,02E-09	<b>PrpC</b>	Universal stress protein E	P31660	0,01	5,54E-09
<b>NanA</b>	N-acetylneuraminate lyase	P0A6L4		11,26	1,15E-10	<b>UspE</b>	Universal stress protein F	P0AAC0	0,01	3,24E-09
<b>GlpF</b>	Glycerol uptake facilitator protein	P0AER0		10,79	2,69E-04	<b>UspF</b>	tRNA-cytidine (32) 2-sulfuryltransferase	P37903	0,01	4,51E-10
<b>PsyE</b>	Thiosulfate sulfurtransferase	P23857		10,67	8,34E-07	<b>TtcA</b>		P76055	0,01	7,00E-08
<b>Year</b>	Uncharacterized protein	P64488	1 x M, 1 x W	10,59	1,24E-04					
<b>FliM</b>	Flagellar motor switch protein	P06974	1 x W	10,32	2,05E-03					
<b>CheZ</b>	Protein phosphatase	P0A9H9		10,20	8,28E-06					
<b>YdfO</b>	Uncharacterized protein	P76156		10,12	6,09E-05					
<b>BhsA</b>	Multiple stress resistance protein	P0AB40		10,08	3,28E-06					

## 4. DISKUSSION

Prolin unterscheidet sich von allen anderen proteinogenen Standard-Aminosäuren durch besondere strukturelle Eigenschaften, die mit einer hohen Rigidität von X-Pro Peptidbindungen einhergehen. Dies gilt insbesondere für konsekutive Proline.  $X_{(-2)}X_{(-1)}-nP-X_{(+1)}$  Motive wirken nicht nur posttranslational auf räumlicher Ebene als strukturgebendes Element in Proteinen, sondern auch kotranslational durch eine zeitliche Beeinflussung der ribosomalen Peptidyltransferasereaktion und einer erheblichen Beeinträchtigung der Translationsrate. Beide Wirkebenen sind untrennbar miteinander verbunden.

Der durch konsekutive Proline induzierte Translationsarrest wird durch die Art der flankierenden Aminosäuren moduliert. In Anbetracht der Tatsache, dass diese den verzögernden Effekt oft verstärken, erscheint es unwahrscheinlich, dass konsekutive Proline allein von struktureller Bedeutung sind. Vielmehr muss analog zu bereits gut beschriebenen Elementen wie seltenen Codons oder *leader peptides* eine kotranslational regulatorische Funktion berücksichtigt werden.

Bereits publizierte Fallbeispiele einiger ausgewählter Proteine unterstreichen die biologische Signifikanz von konsekutiven Prolinen, sind jedoch nicht ausreichend für ein weitgreifendes Verständnis deren Auftretens<sup>59, 71, 131</sup>. Globale proteomische Analysen berücksichtigten bislang nur einzelne polyP Motive<sup>45, 398</sup>. Diese Arbeit umfasst eine breit angelegte Untersuchung des Auftretens und der Evolution von allen putativ arretierenden Polyprolinmotiven in *E. coli* und liefert damit einen umfangreichen Datensatz, der allgemeingültige Aussagen ermöglicht. Gleichzeitig wurden spezifische Fragestellungen bezüglich der Rolle von polyP Motiven in ausgewählten *E. coli* K-12 Rezeptorproteinen beantwortet.

### 4.1 Proteine mit polyP Motiv sind hoch divers

In *E. coli* treten polyP Motive in einer großen Bandbreite von Proteinfamilien unterschiedlichster Funktion auf. Eine 2014 von Mandal *et al.*<sup>45</sup> publizierte proteomweite Analyse des Auftretens von PPP und PPG Motiven in Organismen aus sechs taxonomische Gruppen – Archaea, Eubacteria, Protista, Fungi, Plantae und Animalia – zeigte, dass diese Beobachtung universell zutreffend ist. Die Funktionen von Proteinen mit diesen Motiven sind sehr divers und betreffen nahezu alle wichtigen zellulären Prozesse. Dazu zählen DNA/RNA-Bindung, Signaltransduktion, Replikation und Transkription, Zellzyklus und Morphologie, aber auch die Bildung von strukturellen Komponenten wie dem Cytoskelett und der extrazellulären Matrix<sup>45</sup>.

Im Rahmen dieser Arbeit konnte in *E. coli* darüberhinaus bei Betrachtung der katalytischen Aktivität und der substratbindenden Eigenschaften eine leichte Akkumulation von polyP Motiven in der Enzymklasse der Hydrolasen und Transferasen, sowie in Proteinen, die niedermolekulare Moleküle binden, festgestellt werden. Zudem sind in Membranproteinen mehr polyP Motive anzufinden, als es eine Zufallsverteilung vermuten ließe. Dies könnte auf eine besondere Funktion in diesen Proteinen hindeuten. Das Proteom von *E. coli* beinhaltet ca. 25% putativ membranintegrierte Proteine<sup>312</sup>, auf welche sich mehr als die Hälfte aller „starken“ polyP Motive verteilen. Dabei muss berücksichtigt werden, dass das Auftreten konsekutiver Proline in Transmembrandomänen

aufgrund der strukturellen Eigenschaften *per se* unwahrscheinlich ist, wenn die Proline nicht wie im Fall der *cis/trans*-Isomerisierung von Xaa-Pro Peptidbindungen in Transportproteinen eine übergeordnete Funktion haben.

## 4.2 polyP Motive stehen unter Selektionsdruck

Die Effizienz ist ausschlaggebendes Merkmal für die Art aller an der Translation beteiligten Komponenten<sup>399, 400</sup>. Hier konnte am Beispiel von *E. coli* gezeigt werden, dass polyP Motive aufgrund ihrer der Translationseffizienz entgegenwirkenden Eigenschaften unter negativem Selektionsdruck stehen.

Das Auftreten von polyP Motiven antikorreliert mit Translationseffizienz und Proteinabundanz. Motive, die einen starken Translationsarrest hervorrufen sind signifikant unterrepräsentiert. Je „stärker“ ein Motiv, desto niedriger ist seine Frequenz, insbesondere mit steigender Translationseffizienz der Proteine. Gleichzeitig gilt: je „stärker“ ein Motiv ist, desto weniger konserviert ist es.

Hinzu kommt, dass evolutionsgeschichtlich ältere Proteine weniger polyP Motive tragen und diese wahrscheinlicher durch eine „schwächere“ Variante substituieren, als jüngere. Dies wird durch die Erkenntnis bestätigt, dass auch evolutionsgeschichtlich „jüngere“ und komplexere Organismen – Animalia und Plantae im Vergleich zu Archaea und Eubakteria - eine höhere Frequenz an PPP und PPG Motiven aufweisen<sup>45</sup>. Während eine Deletion von *efp* Eubakterien in den meisten Fällen beeinträchtigt, ist eine Deletion von eIF5A in Eukaryonten aufgrund des hohen Anteils an betroffenen Proteinen lethal<sup>99–101, 109, 110</sup>.

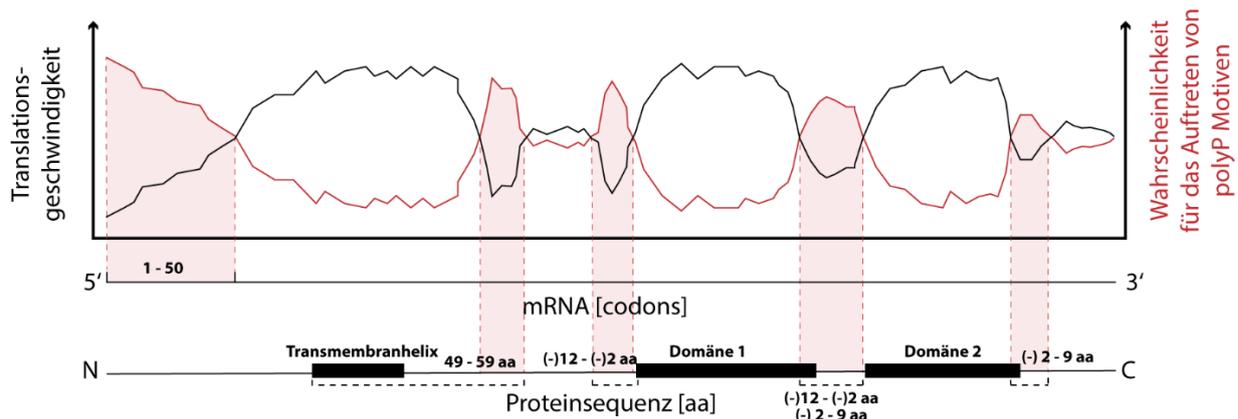
Je geringer die Translationsrate eines Proteins, bzw. die Anzahl an translatierten Proteinkopien innerhalb einer gegebenen Zeit, desto kleiner ist der Einfluss von einzelnen translationsverzögernden Elementen auf dessen Gesamtkopienzahl pro Zelle. Eine Auswirkung auf die finale Proteinabundanz kann nur bei genügend kurzen zeitlichen Abständen zwischen den Translationszyklen, bzw. einer Akkumulation von Polysomen stattfinden. Dass die arretierende Wirkung von polyP Motiven mit der Translationsinitiationsrate eines Proteins sinkt, wurde bereits von Hersch *et al.* (2014) nachgewiesen<sup>63</sup>. Eine geringe Proteinabundanz wäre unter diesem Aspekt keine Folge von polyP Motiven, sondern der Grund, wieso diese im Laufe der Evolution nicht verloren gegangen sind.

Der auf Translationseffizienz basierende Evolutionsdruck formt die Proteinsequenz und ist daher vermutlich treibende Kraft für den Verlust von polyP Motiven. Gleichzeitig bedeutet dies, dass konsekutive Proline trotz der damit verbundenen negativen Auswirkungen auf die Translationseffizienz bestimmte übergeordnete Vorteile bieten, auf welche nicht ohne weiteres verzichtet werden kann.

## 4.3 PolyP Motive haben verschiedene kotranslationale Funktionen

Trotz des übergeordneten Bestrebens nach einer größtmöglichen Effizienz, handelt es sich bei der Translationselongation um einen nonuniformen Prozess<sup>401, 402</sup>. Im Rahmen dieser Arbeit wurden am Beispiel von *E. coli* mehrere Bereiche innerhalb von Proteinsequenzen identifiziert, welche während der Translation typischerweise mit verlangsamter Peptidbindungsrate synthetisiert werden und in welchen auch polyP Motive vermehrt auftreten. Dabei handelt sich um die zur späten Phase der Translationsinitiation zählenden ersten 50

Aminosäuren eines Proteins, um Interdomänenregionen und Regionen, die mit der Translokation von Membranproteinen assoziiert sind (Abb. 43). Das Phänomen lässt sich auch in dem Gram-positiven Bakterium *B. subtilis* beobachten (vgl. Qi *et al.*, 2018<sup>292</sup>), sodass davon ausgegangen werden kann, dass die polyP vermittelte Translationsregulation in Bakterien universal ist.



**Abb 43:** Schematische Darstellung des Zusammenhangs zwischen Translationsgeschwindigkeit und dem Auftreten von polyP Motiven (hypothetische Werte). PolyP Motive sind generell unterrepräsentiert, treten jedoch gehäuft in Proteinbereichen auf, in welchen eine reduzierte Translationsgeschwindigkeit von Vorteil ist. Dabei handelt es sich um die ersten 50, mit der Translationsinitiation assoziierten Aminosäuren, einen Bereich 49-59 Aminosäuren *downstream* von Transmembranhelix-Starts und Interdomänenregionen.

Die Akkumulation von polyP Motiven in den ersten 50 Aminosäuren eines Proteins gleicht dem vermehrten Auftreten von anderen, bereits beschriebenen regulatorischen Elementen, die einen langsamen Start der Translation gewährleisten<sup>403, 16</sup>. Dieser zur späten Phase der Translationsinitiation zählende Proteinbereich zeichnet sich durch ein hohes Vorkommen an seltenen Codons, positiv geladenen Aminosäuren und starken mRNA Strukturen aus und dient vermutlich der Vermeidung von Ribosomenstaus<sup>404</sup>. Die translationsverzögernden Eigenschaften von polyP Motiven machen eine ähnliche Funktion wahrscheinlich, könnten aber auch der Grund dafür sein, wieso sie in diesem Sequenzabschnitt mit hoher Wahrscheinlichkeit toleriert werden.

PolyP Motive akkumulieren in nichtstrukturierten Bereichen eines Proteins und in Interdomänenregionen. Eine Anpassung der Translationsrate steht hier vermutlich mit kotranslationalen Faltungsprozessen bestimmter Domänenstrukturen in Verbindung. Ein Alterieren von schnellen und langsamen Phasen der Peptidbiosynthese ist Voraussetzung für die Faltung bestimmter Domänenstrukturen<sup>405</sup>. Während der Faltung eines Proteins beeinflussen sich dessen strukturelle Elemente gegenseitig, weshalb eine zeitliche Koordination deren Auftretens unabdingbar für eine korrekte Faltung ist<sup>406, 407</sup>. Eine schnelle Translation der mRNA Abschnitte, welche für

strukturelle Domänen kodieren vermeidet missgefaltete Intermediate, während Translationspausen zwischen den Domänen eine unabhängige Faltung der einzelnen Domänen vereinfacht<sup>408, 406, 409</sup>.

PolyP Motive akkumulieren in einem Bereich von 49-59 Aminosäuren *downstream* des Transmembranhelixstarts. Unter der Annahme, dass die meisten Transmembranhelices 21 Aminosäuren lang sind und der ribosomale *exit tunnel* etwa 28 Aminosäuren einschließt, findet der durch polyP Motive induzierte Translationsarrest vermutlich dann statt, wenn die Transmembranhelix aus dem *exit tunnel* austritt und durch das Translokons in die Membran integriert wird<sup>323, 321, 410</sup>. Die Leitung von  $\alpha$ -helicalen Transmembranproteinen an die Translokons durch *signal recognition particles* (SRP) und deren Insertion in die Membran ist ein weiterer, parallel zur Translation stattfindender Prozess, welcher durch die Gewährleistung eines gewissen Zeitfensters vereinfacht wird<sup>320, 321, 18, 19, 411, 322, 412, 22</sup>. Der hier identifizierte Locus stimmt nicht mit der von Fluman *et al.* (2014)<sup>22</sup> beschriebenen Position von Shine-Dalgarno-ähnlichen Sequenzabschnitten in *E. coli* mRNAs überein. Allerdings identifizierten Dessen *et al.* (2000)<sup>322</sup> in *E. nidulans* 45-70 Aminosäuren *downstream* von Transmembranhelices einen Cluster aus seltenen Codons, welcher dem hier beschriebenen Bereich entspricht<sup>322, 22</sup>.

Der intramolekulare Translationsrhythmus wird im Allgemeinen durch verschiedene Elemente reguliert. Zum einen wird die Konformation der entstehenden Peptidkette direkt durch das Ribosom beeinflusst, indem positiv geladene Aminosäuregruppen von der negativ geladenen Ribosomenoberfläche angezogen, während negativ geladene abgestoßen werden<sup>413</sup>. Zum anderen wird die Translationsrate durch die kodierende Sequenz selbst determiniert. Ausschlaggebend dafür sind tRNA-Level, synonyme Codons mit unterschiedlicher Translationsrate und vermutlich Shine-Dalgarno-ähnliche Sequenzen in der mRNA<sup>408, 22, 414, 313, 415-417</sup>. Darüberhinaus haben bereits synthetisierte Peptidsequenzen das Potential, die Translationsrate zu beeinflussen – entweder durch eine direkte Interaktion mit dem ribosomalen *exit tunnel* oder durch die für die Peptidyltransferase relevante Substrateignung der jeweiligen Aminosäuren<sup>418, 12, 311, 293</sup>. Im Gegensatz zu den oben beschriebenen Elementen hat der Einbau von polyP Motiven jedoch einen zusätzlichen Begleiteffekt. Konsekutive Proline determinieren eine spezifische, unflexible Proteinstruktur, die aus diesem Grund nur in wenigen Domänen oder Sequenzabschnitten sinnvoll und nicht störend ist<sup>33</sup>. Translationsarrest und Struktur bedingen sich hier gegenseitig. Falls die primäre Funktion eines polyP Motivs die Manipulation der Translationsrate ist, so muss das Motiv an einer Position innerhalb eines Proteins lokalisiert sein, an dem es dessen Tertiärstruktur nur marginal beeinflusst oder gleichzeitig einen positiven Beitrag dazu leistet.

#### **4.4 Die Bedeutung von polyP Motiven in den Signalproteinen EnvZ, EvgS, DgcC und Lon Protease**

Die hier näher untersuchten Signalproteine DgcC, Lon Protease, EvgS und EnvZ tragen jeweils mehrere polyP Motive unterschiedlicher „Stärke“. Deren intramolekulare Lokalisation ist sehr divers. Keines der Motive befindet sich in einer der für Translationsinitiation, Domänenfaltung oder Membranintegration in Frage kommenden Positionen. Die vier Proteine sind Initiationspunkte von Signaltransduktionswegen, die eine Adaptation der Bakterienzelle an definierte Umweltbedingungen vermitteln. Unter diesem Aspekt wurde die Rolle der EF-P abhängigen polyP Motive individuell bewertet.

Der Fokus wurde hier auf die Rolle von putativ stark translationsverzögernde polyP Motive gelegt. Dabei handelt es sich im Einzelnen um VGPPG (Lon), HPPPG (DgcC), NYPPD (EvgS) und IPPPL (EnvZ). Auf Basis der gewonnenen Erkenntnisse und zur Komplettierung des Datensatzes wurde auch das zweite polyP Motiv (VVPPA) der Histidinkinase EnvZ genauer untersucht.

#### **4.4.1 PolyP Motive in Signalproteinen sind unabhängig von ihrer „Stärke“ unterschiedlich stark konserviert**

Die polyP Motive der vier ausgewählten Proteine sind unabhängig von ihrer vorhergesagten „Stärke“ unterschiedlich stark konserviert. Der Trend, dass putativ stark translationsverzögernde Motive weniger konserviert sind, lässt sich aufgrund der kleinen Stichprobengröße nicht erkennen. Dies impliziert aber andererseits, dass die Motive eine übergeordnete Funktion haben, welche ihren individuellen Verbleib entgegen dem negativen Selektionsdruck erklärt.

Das VGPPG Motiv der Lon Protease ist sehr stark konserviert. Alle untersuchten homologen Proteine mit 44% Sequenzidentität zur *E. coli* K-12 Protease weisen zwei konservative Proline an analoger Position auf. Dies könnte auf die Wichtigkeit des Motivs hinweisen. Das „starke“ HPPPG Motiv der Diguanylatzyklase DgcC ist ausschließlich in Stämmen der Spezies *E. coli* zu finden. In Stämmen der phylogenetischen Serogruppe B2 ist das Motiv durch eine Prolin-Serin Substitution abgeschwächt (HPSPG) und gleichzeitig an einen engen genomischen Kontext gebunden. Die starken polyP Motive NYPPD und IPPPL in EvgS und EnvZ sind neben einer 100% Konservierung in *E. coli* dagegen auch in anderen Organismen zu finden und demnach Grundelemente dieser Histidinkinasen.

#### **4.4.2 Die native Proteinkopienzahl wird nicht durch polyP Motive limitiert**

Unabhängig von der Frage, ob die Kopienanzahl eines Proteins mit polyP Motiv durch das Level an aktiven EF-P Proteinen moduliert werden kann, tragen polyP Motive das Potential, die Translationsrate durch eine Verminderung der Translationseffizienz konstant niedrig zu halten. Dies würde einen zusätzlichen Kontrollmechanismus zu Genexpression darstellen, um eine definierte Anzahl von Proteinen pro Zelle nicht zu überschreiten. Besonders für Signalproteine, die Teil eines komplexen und fein abgestimmten Signaltransduktionsnetzwerks sind, könnte dies von Bedeutung sein. Im Fall von niedrigexprimierten Proteinen, könnten bereits geringe Änderungen in der Kopienzahl signifikante, phänotypische Effekte haben. Obwohl die für Histidinkinase und zugehörigen Antwortregulator kodierenden Gene üblicherweise in einem Operon organisiert sind und damit unter der Kontrolle desselben Promotors stehen, ist die Proteinkopienzahl der beiden Komponenten sehr verschieden. Die Stöchiometrie von EnvZ – OmpR liegt bei 1:35 (ca. 100 Kopien EnvZ/Zelle und 3.500 Kopien OmpR/Zelle)<sup>141</sup>, die von EvgS – EvgA bei etwa 1:6 (ca. 5 Kopien EvgS/Zelle<sup>140</sup> und ca. 30 Kopien EvgA/Zelle<sup>419, 87</sup>). Eine differenzierte Proteinbiosynthese von Proteinen deselben Operons wird in Bakterien und Archaea durch eine unterschiedliche Translationseffizienz der einzelnen Einheiten erzielt<sup>420</sup>. PolyP Motive wären neben den bereits beschriebenen Instrumenten wie seltenen Codons oder mRNA Sekundärstrukturen eine Möglichkeit, die Translationseffizienz von EvgS und EnvZ zu reduzieren.

Im Rahmen dieser Arbeit wurde daher der Einfluss von polyP Motiven auf die Proteinkopienzahl der ausgewählten Proteine EnvZ, EvgS, Lon Protease und DgcC untersucht.

Der Einfluss von putativ stark translationsverzögernden polyP Motiven auf das native Proteinlevel von EnvZ und Lon wurde unter Verwendung von Wildtyp und  $\Delta efp$  *E. coli* Stämmen mit chromosomal kodierten Proteinvarianten mit Pro-Ala substituierten polyP Motiven untersucht. Hervorzuheben ist, dass die Gegenüberstellung im Gegensatz zu vorangegangenen Studien auf den nativen Genexpressionsleveln basiert, um eine unbeabsichtigte Verfälschung der Ergebnisse durch die Manipulation der Transkriptionsrate zu vermeiden.

Die Lon Protease steht stellvertretend für Proteine mit hohem Translationslevel. Hier konnte gezeigt werden, dass das native Proteinlevel der Lon Protease nicht abhängig von EF-P ist. Auch ein chromosomaler Austausch des putativ „starken“ VGPPG Motivs in VGAAG beeinflusst es nicht. Das native Proteinlevel der niedrigexprimierten Histidinkinase EnvZ ist ebenfalls nicht EF-P abhängig. Eine chromosomal exprimierte Proteinvariante, in welcher das putativ stark translationsverzögernde IPPPL Motiv durch IAAAL substituiert ist, weist dieselbe Kopienzahl auf, wie das native Protein. Die Motive haben im Kontext der chromosomal kodierten und unter der Kontrolle ihres nativen Promotors stehenden Genvarianten keinen Einfluss auf die Translationsrate.

Wird die Genexpression hingegen durch die Verwendung von plasmidkodierten *envZ* Varianten artifiziell erhöht, lässt sich die EF-P Abhängigkeit des Proteinlevels feststellen. Sofern allerdings ein kritischer Punkt erreicht ist, und es zu einer starken Anreicherung des Proteins kommt, gleichen sich die Proteinlevel wieder an. Damit wird zum einen die Wichtigkeit der Verwendung von chromosomal kodierten *envZ* Varianten unterstrichen und zum anderen verdeutlicht, dass das zelluläre Proteinlevel von verschiedenen Parametern beeinflusst wird und ein Resultat der *turn over* Rate ist, die sich aus Neusynthese und Degradierung, bzw. Proteinstabilität zusammensetzt<sup>421</sup>. Durch die Induktion der Überexpression eines Gens zu Versuchsbeginn wird das Gleichgewicht in Richtung Neusynthese verschoben. Eine erhöhte Transkriptkonzentration führt dazu, dass die Translationsrate erhöht wird und verzögernde Elemente wie polyP Motive die maximal erreichbare Anzahl an Translationsprozessen innerhalb eines bestimmten Zeitabschnitts limitieren. Demnach hat IPPPL translationsverzögernde Eigenschaften, die jedoch keine Auswirkung auf die native Kopienzahl von 50 EnvZ Molekülen pro Zelle<sup>141</sup> haben.

Am Beispiel des membranintegrierten pH-Sensors und Transkriptionsaktivators CadC wurde in vorangegangenen Studien eine Relevanz von polyP Motiven für die Aufrechterhaltung eines für die Signaltransduktion notwendigen Gleichgewichts demonstriert<sup>59</sup>. In diesem Fall wurde nachgewiesen, dass der Austausch des stark translationsverzögernden PIPPP Motivs zu einer Erhöhung der Kopienzahl von CadC führte, welches unter nativer Promotorkontrolle auf einem *low copy* Plasmid und damit in annähernd originaler Expressionsrate kodiert ist<sup>59</sup>. Eine damit einhergehende Erhöhung der Expressionsrate der Zielgene *cadAB* wurde mit der, nun unverhältnismäßig kleineren, Anzahl des unveränderten inhibierenden Cosensors LysP in Verbindung gebracht. Für EnvZ, das ebenfalls mit einer membranintegrierten Komponente interagiert (MzrA)<sup>154</sup>, konnte ein derartiges Konzept nicht nachgewiesen werden.

EvgS und DgcC können im nativen Expressionlevel nicht mittels Western Blot detektiert werden. Im Fall von EvgS liegt dies augenscheinlich an der geringen Kopienzahl von etwa 5 Molekülen pro Zelle<sup>140</sup>. Zu DgcC liegen diesbezüglich keine Daten vor. Eine Erhöhung des Expressionslevels durch die Verwendung von plasmidkodierten, induzierbaren *dgcC* und *evgS* Varianten ermöglichte die Untersuchung der Relevanz der „starken“ polyP Motive NYPPD und HPPPG. Wie bereits für EnvZ gezeigt, lassen sich Unterschiede bezüglich der EF-P Abhängigkeit der Proteinlevel bei moderaten Expressionsleveln besser auflösen, als bei sehr hohen.

Beide Proteine sind in der Wildtypform stark abhängig von EF-P und nur schwach EF-P abhängig nach NYPPD/NYAD, bzw. HPPPG/HSPPG Substitution. Interessanterweise ist das Proteinlevel der EvgS<sub>NYAAD</sub> Variante im Wildtyphintergrund geringer, als das des nativen Proteins. Das polyP Motiv scheint einen positiven Effekt auf den Translationsprozess zu haben. Möglicherweise behindert sein Fehlen die kotranslationale Prozessierung des Proteins, wodurch es zu Verzögerungen kommt. Die Ergebnisse bestätigen die Vorhersage, dass HPPPG und NYPPD in DgcC und EvgS die Translationsgeschwindigkeit reduzieren. Da allerdings keine Ergebnisse zu chromosomal kodierten Varianten vorliegen, können keine Aussagen bezüglich der Auswirkungen der Motive auf die tatsächliche native Proteinkopienzahl gemacht werden.

Im Fall der Diguanylatcyclase DgcC konnte bei moderatem Expressionslevel eine leichte Erhöhung der Kopienzahl nach Substitution des „starken“ HPPPG Motivs festgestellt werden. Für EnvZ, EvgS und Lon Protease wurde jedoch unter keiner getesteten Bedingung – weder für natives noch für erhöhtes Genexpressionslevel – eine Erhöhung der Proteinkopienzahl als Folge eines Austauschs der „starken“ polyP Motive im genetischen Hintergrund des Wildtyps festgestellt. Der negative Einfluss der polyP Motive auf die Translationseffizienz dieser Proteine muss daher durch die Anwesenheit von EF-P relativiert sein.

Da EF-P den durch polyP Motive induzierten Translationsarrest nicht vollständig nivellieren kann<sup>31, 69, 293</sup>, wird deren translationsverzögernde Wirkung hier vermutlich auch durch andere Komponenten überlagert, welche ebenfalls einen Einfluss auf die Synthesekapazität eines Proteins haben. Dazu zählen Gentranskription, mRNA Prozessierung und Halbwertszeit, Primär- und Sekundärstruktur der mRNA, Verfügbarkeit an ribosomalen Untereinheiten, Anzahl an beladenen tRNAs und Translationsinitiation<sup>422</sup>. Die Architektur von Signaltransduktionsnetzwerken beinhaltet oftmals *feed back loops* und damit die Selbstregulation der Expressionslevel von einzelnen Komponenten. Im Fall des Zweikomponentensystems EvgAS ist bekannt, dass der Antwortregulator EvgA die Transkription des *evgAS* Operons reguliert<sup>271, 423</sup>. Eine *feed back* Inhibition von Diguanylatcyclasen, wie DgcC findet über deren Produkt c-diGMP statt<sup>174</sup>.

Mehrere Studien stellten für verschiedene metabolische Stoffwechselwege und Signaltransduktionswege die Hypothese auf, dass die Architektur des zugrunde liegenden Netzwerkes eine größere Rolle für dessen Stabilität spielt, als die jeweilige Anzahl der einzelnen Komponenten<sup>424-426</sup>. Goulian *et al.* bewiesen 2003 ein mathematisches Modell, wonach Schwankungen in der Expression von EnvZ keine Auswirkungen auf die Phosphorylierung des Antwortregulators OmpR haben, solange die Kopienanzahl von OmpR deutlich höher ist, als die von EnvZ<sup>427</sup>. Zwei-Komponenten-Systeme wie EnvZ/OmpR und EvgS/EvgA weisen demnach eine gewisse Robustheit gegenüber einer stöchiometrischen Varianz der beteiligten Komponenten auf. Eine Regulation der Proteinkopienzahl der Rezeptorproteine hätte demnach keine große Auswirkung auf den assoziierten Signalweg.

Hier konnte gezeigt werden, dass polyP Motive in Proteinen mit niedrigem Translationslevel keine oder nur geringe Auswirkungen auf die native Proteinkopienzahl haben. Damit würde ein geringerer Selektionsdruck zugunsten der Translationseffizienz einhergehen, was wiederum die Etablierung von polyP Motiven in diesen Proteinen erklären würde. Vor diesem Hintergrund muss das vorhergesagte translationsarretierende Potential des VGPPG Motivs der Lon Protease neu bewertet werden.

#### 4.4.3 PolyP Motive sind für die Proteinfunktionalität von EvgS und EnvZ von essentieller Bedeutung

Zur Bewertung der Rolle von polyP Motive für die Funktionalität der untersuchten Proteine EnvZ, EvgS, Lon Protease und DgcC wurden isogene Mutanten mit chromosomal kodierten Proteinvarianten erstellt. Auch hier wurde Wert darauf gelegt, das native Expressionslevel der Proteine nicht zu verfälschen.

Trotz seiner starken Konservierung ist das VGPPG Motiv der Lon Protease nicht essentiell für die phänotypische Ausprägung von *E. coli* K-12 und dessen Fähigkeit, in einen Zustand der Persistenz überzugehen. Die Proline sind Teil des universell konservierten Walker A Motivs (Consensussequenz: GXXXXGKT/S, hier: GPPGVGKS) der Protease Domäne<sup>428</sup>. Der charakteristische P-loop dieser Domäne vermittelt die Bindung von Phosphat, bzw. ATP und initiiert damit die Nukleotidhydrolyse<sup>429, 430</sup>. Ausschlaggebend für die Phosphatbindung ist Lysin<sup>431, 432</sup>. Da ein großer Teil der Proteine der AAA+ Proteasefamilie ein Walker A Motiv ohne Prolin trägt, aber ähnliche Substratbindeigenschaften zeigt, hat Prolin vermutlich primär eine andere Funktion<sup>429</sup>. Diese könnte beispielsweise in der Stabilisierung der loop-Struktur liegen. Dabei ist das VGPPG Motiv der Lon Protease vermutlich nur von geringer Bedeutung für die Proteinfunktionalität.

Das HPPPG Motiv der Diguanylatcyclase DgcC hat ebenfalls keine erkennbare Relevanz für die Funktionalität des Proteins, bzw. den assoziierten Signalweg im genetischen Kontext eines celluloseproduzierenden Reporterstammes. Pathogene *E. coli* Stämme kolonisieren zum Teil sehr diverse ökologische Nischen, was sich auch in deren Genomen bzw. Proteomen widerspiegelt. Das Pangenom von *E. coli* beinhaltet mindestens 35 Gene, die für putative DGCs (GGDEF Domäne) und PDEs (EAL Domäne) kodieren. Abgesehen von abweichenden Proteinvarianten, fehlen *E. coli* K-12 im Vergleich zu pathogenen Stämmen vier PDE Gene<sup>173</sup>. DgcC zählt zu den 8 konservierten Genen und spielt demnach eine Schlüsselrolle in dem Signaltransduktionsnetzwerk. Das HPPPG Motiv von DgcC aus *E. coli* K-12 ist in der MASE2 Domäne lokalisiert, für welche eine Rezeptorfunktion für ein bislang nicht identifiziertes Signal vorhergesagt wird<sup>176</sup>. Möglicherweise spielt das jeweilige Habitat und die Architektur des Signaltransduktionsnetzwerkes der unterschiedlichen Stämme eine Rolle in der Ausprägung des polyP Motivs dieser Domäne. Es bleibt daher offen, welche Auswirkungen ein stark translationsverzögerndes HPPPG Motiv, das wie hier gezeigt im Vergleich zu HPSPG mit einer Reduzierung der Proteinkopienzahl einhergeht, in diesem Kontext hätte.

Die individuelle Funktion der polyP Motive unterscheidet sich selbst in Vertretern derselben Proteinklasse, wie das Beispiel der beiden Histidinkinasen EvgS und EnvZ zeigt (Abb. 44).

Das putativ stark translationsverzögernde NYPPD Motiv aus EvgS ist essentiell für die Funktionalität des Zweikomponentensystems EvgSA. Eine chromosomal kodierte EvgS<sub>NAYYD</sub> Variante führt zu einer verminderten Resistenz gegenüber moderatem Säurestress. Der Phänotyp des entsprechenden *E. coli* K-12 Stammes gleicht dem einer  $\Delta$ evgS Mutante. Das Motiv ist Teil der periplasmatischen VFT1 Domäne. Obwohl das Substratmolekül noch nicht experimentell nachgewiesen wurde, deutet vieles darauf hin, dass die Domäne direkt an der Signalerkennung beteiligt ist<sup>433</sup>. Eine partielle Deletion der VFT1 hat nachweislich den Verlust der Aktivierbarkeit der Histidinkinase bei niedrigen pH Werten zu Folge<sup>156, 158</sup>. Einen ähnlichen Effekt könnte die NYPPD/NYAAD Substitution haben. Dies weist auf die Beteiligung des Motivs an der Substratbindung hin; entweder direkt oder indirekt durch die Stabilisierung von beteiligten Strukturen.

Das cytoplasmatische IPPPL Motiv in EnvZ determiniert die enzymatische Aktivität der Histidinkinase. EnvZ ist bifunktional und agiert den extra- und intrazellulären Stimuli entsprechend als Phosphatase oder als Kinase. Bei hohen Osmolaritäten überwiegt die Kinaseaktivität<sup>150</sup>. Das enzymatische Reaktionsgleichgewicht einer EnvZ<sub>I<sub>AAA</sub>L</sub> Variante ist im Vergleich zum Wildtypprotein stark auf die Seite der Kinaseaktivität verschoben. Die Substitution des cytoplasmatischen polyP Motivs interferiert möglicherweise mit der homodimeren, parallelen *coiled-coil* Struktur der HAMP Domäne und behindert damit deren konformelle Struktur, was den Rezeptor letztendlich im „ON“ Kinasezustand „einfriert“. Verschiedene Studien untermauern diese These. Die Konformation der HAMP Domäne ist abhängig von der Signalperzeption und vermittelt die Signalweiterleitung an die cytoplasmatischen enzymatischen Domänen<sup>434</sup>. Der genaue Mechanismus wurde am Beispiel des Tar Rezeptors aufgeklärt. Während eine kompakte Anordnung der Helixbündel mit dem „OFF“ Zustand assoziiert ist, hat eine lockere Anlagerung der Helices den „ON“ Status zur Folge<sup>435</sup>. Aminosäuresubstitutionen innerhalb der HAMP Domäne eines chimären Tar-EnvZ Rezeptors arretieren die enzymatische Aktivität des Proteins im Kinasezustand und führen zu einer konstitutiven, von der osmotischen Bedingung unabhängigen Expression von OmpC<sup>436</sup>. Dieser Phänotyp ist auch charakteristisch für eine A193V Substitution innerhalb der EnvZ HAMP Domäne und tritt ebenfalls auf, wenn das Tar Fragment N-terminal direkt *up-* oder *downstream* der drei konsekutiven Proline fusioniert wird<sup>436, 333</sup>. Deshalb wurde vorgeschlagen, dass diese Aminosäuresubstitutionen die ohnehin dynamische HAMP Domäne destabilisieren und/oder deren Fähigkeit, externe Stimuli an die HisKA Domäne zu übertragen, entgegenwirken<sup>333</sup>. Einen ähnlichen Effekt hat vermutlich die hier analysierte Substitution des IPPPL Motivs. Dass die Funktionalität von EnvZ primär von den strukturellen Eigenschaften dieses Motivs und nicht von dessen translationsverzögernden Eigenschaften abhängt, wird dadurch untermauert, dass eine Deletion von *efp* keinen Einfluss auf die Signaltransduktion und das Expressionsmuster der beiden Zielgene *ompC* und *ompF* hat.

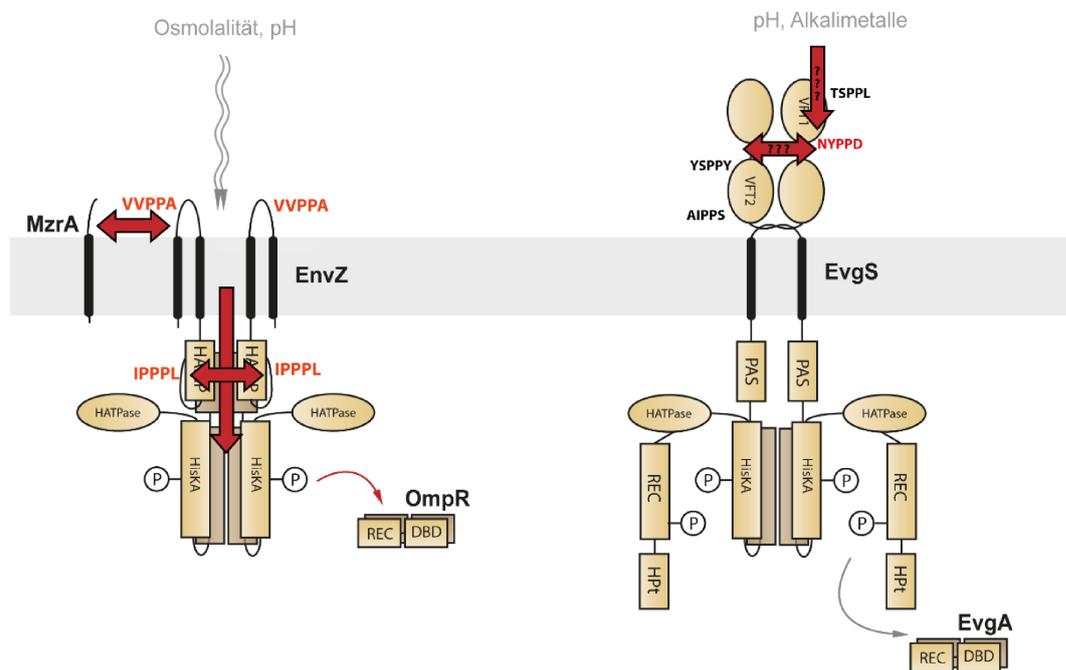
Das hohe Auftreten von polyP Motiven in Histidinkinasen weist auf deren große Bedeutung für diese Proteinklasse hin. Im Rahmen dieser Arbeit konnte am Beispiel zweier wichtiger Vertreter die Notwendigkeit der Motive für die Proteinfunktionalität gezeigt werden. Die Aktivität von EnvZ und EvgS hängt maßgeblich von der Präsenz der darin auftretenden „starken“ polyP Motive ab. Inwiefern die Funktionalität anderer Histidinkinasen von dem Auftreten von polyP Motiven abhängt, muss aufgrund der Lokalisierung in unterschiedlichen Domänen individuell untersucht werden.

#### 4.4.4 Die polyP Motive von EnvZ vermitteln Protein-Protein-Interaktionen

Die Wichtigkeit des polyP IPPPL Motivs in der EnvZ HAMP Domäne für die Proteinfunktionalität hängt mit dessen Beteiligung an der Proteindimerisierung zusammen. Die Homodimerisierung von EnvZ wird in erster Linie über die HisKA Domäne vermittelt, die gesamte cytoplasmatische Berührungsfläche besteht jedoch sowohl aus HisKA, als auch aus HAMP Domäne<sup>437, 438</sup>. *In vitro* Studien zeigten, dass die HAMP Domäne entbehrlich für die Dimerisierung einer trunkierten HisKA Domäne ist<sup>439</sup> und dass deren primäre Funktion in der Transduktion des extrazellulären Stimulus zu den cytoplasmatischen, enzymatischen Domänen von EnvZ liegt<sup>440, 441</sup>.

Das membranintegrierte Modulatorprotein MzrA interagiert mit EnvZ bekanntermaßen im periplasmatischen Raum<sup>153</sup>. Eine Substitution der Proline des cytoplasmatischen IPPPL Motivs in der EnvZ HAMP Domäne verhindert die EnvZ-MzrA Interaktion jedoch vollständig. Nach einem Konzept von Williams und Steward (1999)

ist die HAMP Helix I parallel zur Membranoberfläche angeordnet, anstatt ins Cytoplasma hineinzuragen, was eine physische Interaktion zwischen MzrA und dem polyP Motiv erlauben würde<sup>442</sup>. Auf der anderen Seite hat eine Deletion der ersten fünf cytosolischen Aminosäuren von MzrA keine Auswirkungen auf das Zusammenspiel zwischen EnvZ und MzrA und eine kürzlich veröffentlichte biophysikalische Studie zeigte, dass die Helix I von EnvZ nicht mit einer *bicelle* Oberfläche interagiert, wie es bei Tar und NarX der Fall ist<sup>153, 443</sup>. Daher kann davon ausgegangen werden, dass die Verringerung der EnvZ-MzrA Interaktion in Folge einer Substitution des cytoplasmatischen EnvZ polyP Motivs wahrscheinlicher eine Konsequenz der veränderten Konformation der EnvZ<sub>IAAAL</sub> Variante ist.



**Abb. 44:** Übersicht der Funktionen von polyP Motiven in den Histidinkinasen EnvZ (links) und EvgS (rechts). Das periplasmatische EnvZ VVPPA Motiv ist nachweislich an der Interaktion mit dem Modulator MzrA beteiligt. Das periplasmatische EnvZ IPPPL Motiv ist für die Homodimerisierung erforderlich, sowie an der Signalweiterleitung und damit der Determinierung der enzymatischen Aktivität von EnvZ beteiligt. Das periplasmatische EvgS NYPPD ist ebenfalls Voraussetzung für die Funktionalität des Proteins. Inwiefern das Motiv an der Signalperzeption oder der Homodimerisierung der VFT Domänen beteiligt ist, bleibt noch zu klären.

Die Ergebnisse dieser Arbeit demonstrieren, dass das periplasmatische EnvZ VVPPA Motiv an der Interaktion mit MzrA involviert ist (Abb. 44). Ob die auf Seiten von MzrA an der Interaktion beteiligten Aminosäuren D51 und I78<sup>153</sup> das polyP Motiv direkt binden, bleibt noch zu untersuchen. Die Wichtigkeit des periplasmatischen EnvZ polyP Motivs für die Interaktion mit MzrA wird von einer phylogenetischen Analyse gestützt, welche eine Koevolution zwischen MzrA und dem Auftreten eines periplasmatischen polyP Motivs in EnvZ aufzeigt. Damit zählen die beiden EnvZ polyP Motive zu der Vielzahl an prolinenthaltenden Peptidsequenzen, deren Beteiligung an der Protein-Proteininteraktion bereits beschrieben sind. Konsekutive Proline sind häufig exponiert in Proteinstrukturen und ihre stabile Konformation erlaubt eine Bindung mit nur kleinen entropischen

Veränderungen. Daher erleichtern prolinreiche Regionen eine schnelle und unspezifische Bindung, im Gegensatz zu der langsamen, aber spezifischen Bindung von globulären Interaktionspartnern<sup>444</sup>. In Übereinstimmung mit dieser Idee wurde die Bindeaffinität zahlreicher Proteine, wie den humanen Speichelproteinen oder dem actinbindenden Protein von *Dictyostelium discoideum* in früheren Studien bereits mit deren prolinreichen Sequenzmotiven in Verbindung gebracht<sup>115, 445</sup>.

Eine Beteiligung der EvgS polyP Motive an inter- und intramolekularen Wechselwirkungen der periplasmatischen VFT Domänen ist aufgrund der geringen Distanzen der Aminosäuremotive innerhalb des Homodimers ebenfalls denkbar (Abb. 44). Es bleibt noch nachzuweisen, ob die reduzierte Funktionalität der EvgS<sub>NYPPD</sub> Variante Folge einer veränderten oder gelockerten Domänenanordnung aufgrund fehlender Bindungen ist.

Der Modulator MzrA ist nachweislich an der Justierung der enzymatischen Aktivität von EnvZ und der damit einhergehenden Regulation der Porinexpression unter verschiedenen osmotischen Bedingungen beteiligt, spielt jedoch keine Rolle bei der pH-abhängigen oder Procain-abhängigen Signalantwort<sup>154</sup>. Hier konnte erstmals gezeigt werden, dass diese Spezifität durch die Bindeeigenschaften von MzrA an EnvZ unter verschiedenen osmotischen Bedingungen vermittelt wird. Hohe Natriumchloridkonzentrationen verändern die Bindeaffinität von MzrA an EnvZ.

Die Interaktion und Kollokalisierung von MzrA und EnvZ auf Einzelzellebene konnte mittels hochauflösender Fluoreszenzmikroskopie nicht bestätigt werden. Die dafür nötige Probenpräparation setzt eine Zellfixierung voraus, was vermuten lässt, dass es sich bei der EnvZ-MzrA Interaktion um eine transiente Bindung handelt, welche auf diese Weise nicht visualisiert werden kann. Aus diesem Grund sollten zukünftige mikroskopische Analysen *in vivo* durchgeführt werden. Nichtsdestotrotz stellt der im Rahmen dieser Arbeit verwendete Versuchsaufbau einen wichtigen Beitrag zur Etablierung der erst kürzlich für Prokaryonten beschriebenen Methoden zur intrazellulären Lokalisierung zweier Fluorophore mittels PALM dar<sup>275</sup>. Darüberhinaus wurden sowohl EnvZ, als auch MzrA zum ersten Mal bei nativem Expressionslevel auf Einzelzellebene visualisiert. Es konnte gezeigt werden, dass beide Proteine distinkte Cluster in der Zellmembran formieren, wobei deren Anzahl pro Zelle bei etwa 2-3 liegt. Dies stimmt mit der von Sommer *et al.* (2013) gezeigten Clusterbildung von EvgS überein<sup>161</sup>. MzrA liegt wie die Histidinkinasen als Dimer vor; eine Oligomerisierung zu Clustern erscheint daher plausibel<sup>153</sup>.

#### **4.5 Die Aktivierung von EF-P ist nicht Teil einer zentralen Stressantwort**

In ihrer natürlichen Umweltumgebung sind Bakterien sich ständig verändernden Wachstumsbedingungen ausgesetzt. Um auch ungünstige Stresssituationen zu bewältigen, haben sie zahlreiche Strategien für eine schnelle Adaptation entwickelt. Während eine spezifische Signaltransduktion als Antwort auf ausgewählte Umweltreize durch ein als Regulon bezeichnetes Set aus Regulator und ausgewählten Zielgenen vermittelt wird, umfasst die globale Stressantwort ein weitverzweigtes Signaltransduktionsnetzwerk, welches umfassende Anpassungen auf metabolischer und transkriptionaler Ebene ermöglicht und während verschiedener Stresssituationen aktiviert wird<sup>446</sup>. Die translationsverzögernden Eigenschaften von konsekutiven Prolinen machen sie potentiell zu einem wirkungsvollen Instrument zur Regulation der Proteinexpression. Sofern polyP Motive konstante Zielelemente

einer Translationsregulation sind, muss EF-P als Variable fungieren – entweder durch eine flexible Kopienzahl oder eine angepasste Aktivierung. Eine translatorische Regulation der Expression von Proteinen mit polyP Motiv durch EF-P als Reaktion auf spezifische Umweltreize ist aufgrund der hohen Diversität der potentiellen Zielproteine sehr wahrscheinlich eine globale Stressantwort.

In *E. coli* spielt der alternative Sigmafaktor RpoS eine ausschlaggebende Rolle sowohl für Stressadaptation, als auch für Stressresistenz gegenüber einem breiten Spektrum an ungünstigen Umweltbedingungen. RpoS wird durch diverse Adaptoren, die unter verschiedenen Stresssituationen induziert werden, hinsichtlich Transkription, Translation, Degradierung, und Aktivität beeinflusst<sup>447</sup>. Im Komplex mit der RNA Polymerase RNAP reguliert RpoS etwa ein Zehntel des gesamten *E. coli* Proteoms<sup>448</sup>. Dies entspricht in etwa der Anzahl der Proteine mit „stark“ EF-P abhängigem polyP Motiv. Stressbedingungen, welche zur Induktion von RpoS führen und auch im natürlichen Habitat von *E. coli* anzufinden sind - wie DNA Schädigung<sup>449</sup>, Stickstoff- und Kohlenstofflimitation<sup>450, 447</sup>, Temperaturen außerhalb des Toleranzbereiches<sup>451, 452</sup>, extreme pH-Werte<sup>453, 454</sup>, hohe Osmolarität<sup>455</sup> und oxydativer Stress<sup>456</sup> – wurden hier auf deren Wirkung auf die Expression und Aktivität von EF-P getestet. Es wurden jedoch nur wenige Anhaltspunkte für eine Beteiligung von EF-P an einer globalen Stressantwort gefunden.

Eine leichte Erhöhung des Transkriptionslevels von *efp*, sowie *epmA* und *epmB* konnte in der späten stationären und Absterbephase, unter anaeroben Wachstumsbedingungen und bei Kohlenstofflimitation festgestellt werden. Diese Stressbedingungen haben eine Limitation von Ressourcen gemeinsam, während welcher eine Reduzierung der Proteinbiosynthese von nicht essentiellen Proteinen von Vorteil sein könnte. Gleichzeitig zeichnen sie sich durch eine Akkumulation von reaktiven Sauerstoff- und Stickstoffspezies (ROS und RNS) aus. Dementsprechend konnte eine transkriptionale Regulation von *epmA* durch OxyR nachgewiesen werden. Unter durch H<sub>2</sub>O<sub>2</sub> induziertem oxydativen Stress kommt es zu einer geringfügigen Erhöhung des mRNA Levels von *epmA*. Die Interaktion von OxyR mit der Promotor-DNA von *epmA* konnte mittels DNA-Affinitätschromatographie verifiziert werden. Darüberhinaus deuten Expressionsanalysen darauf hin, dass die Transkriptionsfaktoren NagC und UlaR an der Transkription von *epmA* beteiligt sind. Demnach findet unter Bedingungen, die eine generelle Stressantwort induzieren, eine Feinjustierung der Anzahl von EF-P modifizierendem EpmA statt.

Aufgrund der nur sehr geringen Veränderungen der *epmA* Transkription handelt es sich wahrscheinlich nicht um eine Stressantwort, sondern um eine Reaktion auf die allgemein veränderte Proteintranslation unter den getesteten Bedingungen. Im Detail handelt es sich dabei um limitierte Kohlenstoffquellen, oxydativen oder nitrosativen Stress (OxyR), Glucosemetabolismus (NagC) und anaerobe Wachstumsbedingungen (UlaR). Analog dazu wurde bereits eine moderate (p)ppGP vermittelte Anpassung der *efp* Transkription in Folge von Aminosäurelimitation und einer damit einhergehenden reduzierten Proteinbiosynthese beschrieben<sup>457</sup>.

Ein metabolischer Zustand, während dem eine Reduzierung des Translationslevels von nicht essentiellen Proteinen besonders denkbar ist, existiert in VBNCs – lebenden, aber nicht kultivierbaren Zellen (*viable but non-culturable*), welche sich in einem metabolischen Ruhezustand befinden. Durch Langzeitkultivierung bei konstantem Kältestress geht ein geringer Prozentsatz der Individuen einer Zellkultur in einen Überdauerungszustand über. Der Phänotyp gleicht dabei dem der Persistenz<sup>458</sup>. Die Bakterien sind unempfindlich gegenüber Antibiotika, was insbesondere im Fall von Pathogenen ein enormes Gesundheitsrisiko darstellt<sup>389</sup>. Während des Wachstumsstillstands ist die allgemeine Genexpression sehr niedrig und die Proteinbiosynthese auf das Nötigste beschränkt<sup>388, 387</sup>. Zusätzlich ist die Anzahl an Ribosomen, welche um Aminosäuresubstrate konkurrieren, stark

reduziert<sup>388, 387</sup>. Daraus ergibt sich theoretisch auch eine geringere Notwendigkeit von EF-P. In *Pseudomonas aeruginosa* wurde dagegen ein leicht EF-P Level nachgewiesen, dem eine puffernde Wirkung in einer substratlimitierten Umwelt zugeschrieben wurde<sup>387</sup>. Da der VBNC Zustand in dieser Arbeit mit Hilfe von Kälte herbeigeführt wurde und eine Substratlimitation ausgeschlossen werden kann, muss die Rolle von EF-P hier neu bewertet werden.

Dank einer im Rahmen dieser Arbeit etablierten Methode zur Separierung von VBNCs nach Langzeitkultivierung und einer anschließenden massenspektrometrischen Analyse von über 50% des Proteoms gelang es, umfassende Aussagen über deren proteinogene Zusammensetzung zu treffen. Die Proteinlevel von EF-P, EpmA, EpmB und EpmC sind in VBNCs im Vergleich zur exponentiellen Wachstumsphase nicht signifikant verändert. Trotzdem konnte nach 38 Tagen Inkubation bei 4°C ein Abfall der für EpmC typischen Hydroxylierung der EF-P K34  $\beta$ -Lysin Modifikation festgestellt werden. EpmC spielt für die Aktivität von EF-P jedoch nur eine untergeordnete Rolle<sup>92, 30</sup>. Insofern bleibt die Bedeutung dieser Beobachtung unklar. Es konnte kein allgemeiner Zusammenhang zwischen dem Auftreten von polyP Motiven und der relativen Kopienzahl der betreffenden Proteine im VBNC Zustand festgestellt werden.

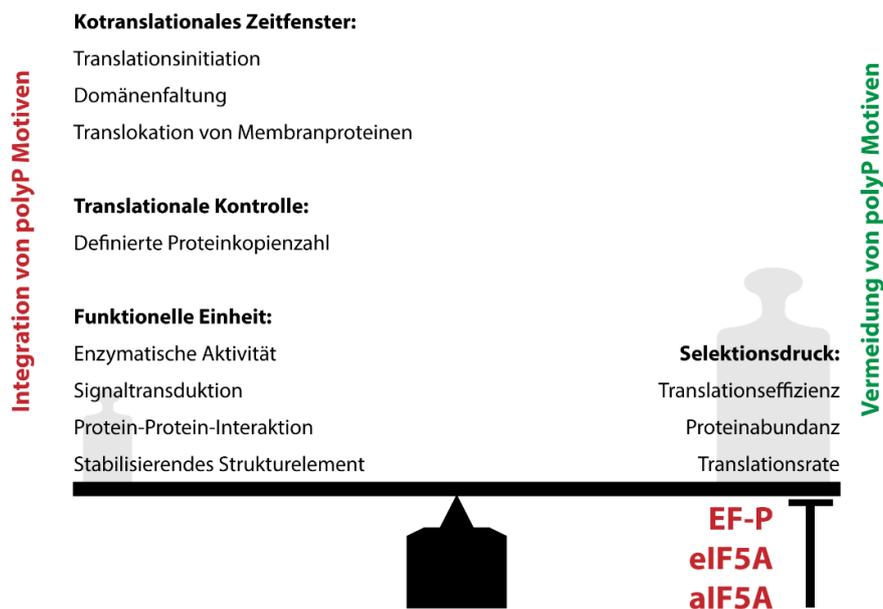
Darüberhinaus lieferten die Daten aber wichtige Informationen zur Charakterisierung der Proteinexpression in *E. coli* VBNCs und leisten damit einen Beitrag zum besseren Verständnis dieses Zustandes. VBNCs unterscheiden sich in vielerlei Hinsicht von Zellen aus der exponentiellen Wachstumsphase. Zu den Charakteristika zählen unter anderem Veränderungen in Morphologie, Zellwand und -membranzusammensetzung, Metabolismus und Resistenzen<sup>389</sup>. Die Resultate der hier gezeigten Analyse stimmen damit überein. Neben einigen Kälteschockproteinen (*cold shock*) sind im metabolischen Ruhezustand vermehrt Proteinlevel einzelner Stoffwechselwege im Vergleich zum Gesamtproteom und der exponentiellen Wachstumsphase signifikant erhöht. Dabei handelt es sich um Proteine des Glycerophospholipidmetabolismus und des Zuckermetabolismus. Eine Anpassung der Membranfluidität durch den Einbau alternativer Glycerophospholipide ist eine typische Reaktion der Bakterienzelle auf Kältestress<sup>459</sup>. Im Ruhezustand konzentriert sich der energetische Haushalt auf wenige, aber essentielle Komponenten. In diesem Fall handelt es sich um den Zuckermetabolismus. Es konnte bestätigt werden, dass die Kopienzahl von ribosomalen Proteinen reduziert ist. Die Wiederaufnahme der Zellteilung bei Wiederherstellung günstiger Wachstumsbedingungen wird nachweislich durch Pyruvat angestoßen<sup>388, 460</sup>. Hier konnte gezeigt werden, dass die Level von Proteinen des Pyruvatmetabolismus und des Pyruvattransporters BtsT in VBNCs im Vergleich zur exponentiellen Wachstumsphase erhöht sind. Diese Beobachtung trug zur Aufklärung der Bedeutung der Pyruvataufnahme für das Beenden des VBNC Zustandes bei (vgl. Vilhena *et al.*, 2018<sup>397</sup>).

Die hier durchgeführten Versuche betreffen den Laborstamm *E. coli* K-12. Es ist nicht auszuschließen, dass EF-P in pathogenen Stämmen insbesondere während der Infektion, welche mit einer starken phänotypischen Veränderung einhergeht, eine größere regulatorische Rolle spielt. So wurde beispielsweise beschrieben, dass EF-P und EpmA essentiell für die Virulenz von *Shigella flexneri* sind<sup>99</sup>. Andere Studien zeigten hingegen eine Herunterregulierung von EF-P in *Salmonella enterica* während der Infektion<sup>461</sup>. Inwiefern es sich dabei um eine gezielte Manipulation der Translationslevel von mit der Pathogenität assoziierten Proteinen handelt oder ob dies in Folge einer allgemeinen Anpassung der Proteinbiosynthese geschieht, bleibt noch zu klären.

## 5. Ausblick

Diese Arbeit leistet einen Beitrag zum besseren Verständnis der Rolle von polyP Motiven, insbesondere in Signalproteinen aus *E. coli* und verdeutlicht deren weites Wirkspektrum.

Der durch polyP Motive induzierte Translationsarrest stellt aufgrund der Beeinträchtigung der Translationseffizienz prinzipiell einen selektiven Nachteil dar. Theoretisch reicht die Existenz eines einzigen essentiellen Proteins mit stark translationsverzögerndem polyP Motiv aus, um die Anwesenheit eines unterstützenden Translationsfaktors zu begründen. Tatsächlich sind EF-P und dessen Orthologe *eIF5A* an das Auftreten einer minimalen translationsarretierenden, aber essentiellen Einheit geknüpft – dem Triprolinmotiv der Valyl-tRNA-Synthetase Vals<sup>71</sup>. Dieses ist universell konserviert, in allen bislang untersuchten Organismen zu finden und macht die Anwesenheit des Translations-Elongationsfaktors erforderlich. Damit geht vermutlich ein gewisser Toleranzbereich für das Auftreten weiterer polyP Motive einher (Abb. 45).



**Abb. 45:** Modell für das Auftreten und die Funktion von polyP Motiven. PolyP Motive stehen aufgrund ihrer translationsverzögernden Eigenschaften, welche der Translationseffizienz, sowie einer hohen Proteinabundanz und Translationsrate entgegenwirken, unter negativer Selektion. Durch die Existenz von EF-P und dessen Orthologe wird der arretierende Effekt von polyP Motiven abgeschwächt, so dass deren Integration in Proteinsequenzen weniger stark ins Gewicht fällt. PolyP Motive wirken sowohl auf kotranslationaler Ebene, als auch im reifen Protein und haben vielfältige Funktionen.

Im Rahmen dieser Arbeit wurden in Ergänzung zu den bereits bekannten Funktionen zahlreiche und vielfältige weitere Wirkweisen von polyP Motiven beschrieben. Diese sind ebenso divers, wie die Art der betreffenden Proteine und polyP Motive stellen sich als universelles Instrument dar, das sowohl kotranslational, als auch funktional im reifen Protein wirkt. PolyP Motive agieren während der Translation als regulatorisches Element, welches die Translationsgeschwindigkeit von strategisch relevanten Proteinbereichen ähnlich wie andere bereits beschriebene cis-agierende Faktoren, wie seltene codons oder sekundäre mRNA Strukturen, moduliert. Hier konnte eine Assoziation mit Translationsinitiation, Domänenfaltung und Translokation von Membranproteinen gezeigt werden. Darüberhinaus wurde eine Rolle von polyP Motiven in der Funktionalität und Signaltransduktion, sowie der Dimerisierungskapazität und Protein-Protein-Interaktion von Rezeptorhistidinkinasen nachgewiesen. Die Ergebnisse dieser Arbeit liefern Anhaltspunkte für eine generelle Bewertung von polyP Motiven. Aufgrund der zahlreichen möglichen Einsatzgebiete muss deren spezifische Bedeutung letztendlich aber in jedem Protein individuell betrachtet werden.

EnvZ und MzrA wurden bei nativer Expression erfolgreich in *E. coli* K-12 auf Einzelzellebene visualisiert und detektiert. Die damit einhergehenden Erkenntnisse können für eine darauf aufbauende *in vivo* Analyse herangezogen werden. Sollte es gelingen, die Kolokalisation der beiden Proteine in lebenden Zellen nachzuweisen, kann im nächsten Schritt die Dynamik der EnvZ-MzrA Interaktion untersucht werden.

## 5. DANKSAGUNG

Herzlichen Dank an Frau Prof. Kirsten Jung für das spannende Thema und die Möglichkeit, Teil ihrer Arbeitsgruppe zu sein, für ihr entgegengebrachtes Vertrauen, sowie die langjährige Unterstützung und Förderung.

Für ihr Interesse an meiner Arbeit möchte ich dem gesamten Prüfungskomitee und insbesondere dem Zweitgutachter Herrn Prof. Marc Bramkamp danken.

Folgenden Personen möchte ich sehr für die erfolgreiche Zusammenarbeit danken:

Prof. Marc Bramkamp und Giacomo Giacomelli, die mir eine große Hilfe bei der Planung und Durchführung der PALM Versuche zur Lokalisierung von EnvZ und MzrA auf Einzelzellebene waren.

Prof. Dmitrij Frishman und Dr. Fei Qi, die einen großen Beitrag zur Analyse der Evolution von polyP Motiven in *E. coli* geleistet und durch ihren Enthusiasmus und ihre analytische Denkweise viele neue Impulse gegeben haben.

Prof. Ralf Heermann danke ich für die kompetente Unterstützung bei der Durchführung der SPR-Analysen.

Dr. Ignasi Forné, der die massenspektrometrischen Analysen durchgeführt hat und einen großen Beitrag zur Probenvorbereitung geleistet hat.

Ich möchte allen gegenwärtigen und ehemaligen Mitgliedern der AGs K. Jung, H. Jung und R. Heermann für die schöne Zeit danken, insbesondere meinen Büro-Kollegen Sophie Buchner, Stefan Behr, Yang Wang, Elisabeth Hoyer und Bruno Pinheiro.

Besonderer Dank auch an Dr. Jae Yen Shin und Dr. Jürgen Lassak für ihre exzellenten Ratschläge und die Einführung in neue Methoden. Vielen Dank an Ingrid Weitzl und Sabine Peschek für ihre hervorragende technische Assistenz.

Nichtzuletzt bin ich meiner Familie - vor allem meinen Eltern und meiner Schwester - unendlich dankbar für ihren bedingungslosen Rückhalt!

## 6. REFERENZEN

1. Crick FH (1958). On protein synthesis. *Symp Soc Exp Biol* **12**, 138–163.
2. Noller HF (2012). Evolution of protein synthesis from an RNA world. *Cold Spring Harb Perspect Biol* **4**, a003681.
3. Rodnina MV (2018). Translation in Prokaryotes. *Cold Spring Harb Perspect Biol.* **4**, a032664.
4. Schmeing TM & Ramakrishnan V (2009). What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461**, 1234–1242.
5. Mitarai N, Sneppen K & Pedersen S (2008). Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization. *J Mol Biol* **382**, 236–245.
6. Sørensen MA & Pedersen S (1991). Absolute in vivo translation rates of individual codons in *Escherichia coli*. *J Mol Biol* **222**, 265–280.
7. Rodnina MV, Fricke R, Kuhn L & Wintermeyer W (1995). Codon-dependent conformational change of elongation factor Tu preceding GTP hydrolysis on the ribosome. *EMBO J* **14**, 2613–2619.
8. Ticu C, Nechifor R, Nguyen B, Desrosiers M & Wilson KS (2009). Conformational changes in switch I of EF-G drive its directional cycling on and off the ribosome. *EMBO J* **28**, 2053–2065.
9. Doerfel LK, Wohlgemuth I, Kubyshkin V, Starosta AL, Wilson DN, Budisa N & Rodnina MV (2015). Entropic Contribution of Elongation Factor P to Proline Positioning at the Catalytic Center of the Ribosome. *J Am Chem Soc* **137**, 12997–13006.
10. Muto H & Ito K (2008). Peptidyl-prolyl-tRNA at the ribosomal P-site reacts poorly with puromycin. *Biochem Biophys Res Commun* **366**, 1043–1047.
11. Wohlgemuth I, Brenner S, Beringer M & Rodnina MV (2008). Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. *J Biol Chem* **283**, 32229–32235.
12. Pavlov MY, Watts RE, Tan Z, Cornish VW, Ehrenberg M & Forster AC (2008). Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proc Natl Acad Sci U S A* **106**, 50–54.
13. Grundy FJ & Henkin TM (2006). From ribosome to riboswitch: control of gene expression in bacteria by RNA structural rearrangements. *Crit Rev Biochem Mol Biol* **41**, 329–338.
14. Korolev SA, Zverkov OA, Seliverstov AV & Lyubetsky VA (2016). Ribosome reinitiation at leader peptides increases translation of bacterial proteins. *Biol Direct* **11**.
15. Marin M (2008). Folding at the rhythm of the rare codon beat. *Biotechnol J* **3**, 1047–1057.
16. Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, Pan T, Dahan O, Furman I & Pilpel Y (2010). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* **141**, 344–354.
17. Tuller T, Veksler-Lublinsky I, Gazit N, Kupiec M, Ruppin E & Ziv-Ukelson M (2011). Composite effects of gene determinants on the translation speed and density of ribosomes. *Genome Biol* **12**, R110.
18. Pechmann S, Chartron JW & Frydman J (2014). Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP in vivo. *Nat Struct Mol Biol* **21**, 1100–1105.
19. Morgunov AS & Babu MM (2014). Optimizing membrane-protein biogenesis through nonoptimal-codon usage. *Nat Struct Mol Biol* **21**, 1023–1025.
20. Mao Y, Liu H, Liu Y & Tao S (2014). Deciphering the rules by which dynamics of mRNA secondary structure affect translation efficiency in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **42**, 4813–4822.

21. McCown PJ, Corbino KA, Stav S, Sherlock ME & Breaker RR (2017). Riboswitch diversity and distribution. *RNA* **23**, 995–1011.
22. Fluman N, Navon S, Bibi E & Pilpel Y (2014). mRNA-programmed translation pauses in the targeting of *E. coli* membrane proteins. *eLife* **3**.
23. Sherwood AV & Henkin TM (2016). Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. *Annu Rev Microbiol* **70**, 361–374.
24. Diwan GD & Agashe D (2016). The Frequency of Internal Shine-Dalgarno-like Motifs in Prokaryotes. *Genome Biol Evol* **8**, 1722–1733.
25. Payne SH, Bonissone S, Wu S, Brown RN, Ivankov DN, Frishman D, Pasa-Tolić L, Smith RD & Pevzner PA (2012). Unexpected diversity of signal peptides in prokaryotes. *mBio* **3**, pii: e00339-12.
26. Yap M-N & Bernstein HD (2011). The translational regulatory function of SecM requires the precise timing of membrane targeting. *Mol Microbiol* **81**, 540–553.
27. Ito K & Chiba S (2013). Arrest peptides: cis-acting modulators of translation. *Annu Rev Biochem* **82**, 171–202.
28. Mason N, Ciufò LF & Brown JD (2000). Elongation arrest is a physiologically important function of signal recognition particle. *EMBO J* **19**, 4164–4174.
29. Wilson DN, Arenz S & Beckmann R (2016). Translation regulation via nascent polypeptide-mediated ribosome stalling. *Curr Opin Struct Biol* **37**, 123–133.
30. Peil L, Starosta AL, Lassak J, Atkinson GC, Virumae K, Spitzer M, Tenson T, Jung K, Remme J & Wilson DN (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.
31. Tanner DR, Cariello DA, Woolstenhulme CJ, Broadbent MA & Buskirk AR (2009). Genetic identification of nascent peptides that induce ribosome stalling. *J Biol Chem* **284**, 34809–34818.
32. Yaron A & Naider F (1993). Proline-dependent structural and biological properties of peptides and proteins. *Crit Rev Biochem Mol Biol* **28**, 31–81.
33. MacArthur MW & Thornton JM (1991). Influence of proline residues on protein conformation. *J Mol Biol* **218**, 397–412.
34. Williams KA & Deber CM (1991). Proline residues in transmembrane helices: structural or dynamic role? *Biochemistry* **30**, 8919–8923.
35. Venkatachalam CM & Ramachandran GN (1969). Conformation of polypeptide chains. *Annu Rev Biochem* **38**, 45–82.
36. Dunker AK (1982). A proton motive force transducer and its role in proton pumps, proton engines, tobacco mosaic virus assembly and hemoglobin allostereism. *J Theor Biol* **97**, 95–127.
37. Kartha G, Ashida T & Kakudo M (1974). The crystal structure of t -amyloxy carbonyl- L -prolyl- L -prolyl- L -proline. *Acta Crystallogr B Struct Sci* **30**, 1861–1866.
38. Schulz GE & Schirmer RH (1979) Principles of Protein Structure (Prinzipien der Proteinstruktur). Springer Advanced Texts in Chemistry, Edit. C. R. Cantor. Springer-Verlag, Berlin -Heidelberg -New York. (Buch)
39. Owens NW, Braun C, O'Neil JD, Marat K & Schweizer F (2007). Effects of glycosylation of (2S,4R)-4-hydroxyproline on the conformation, kinetics, and thermodynamics of prolyl amide isomerization. *J Am Chem Soc* **129**, 11670–11671.
40. Fischer S, Dunbrack RL & Karplus M (1994). Cis-Trans Imide Isomerization of the Proline Dipeptide. *J Am Chem Soc* **116**, 11931–11937.
41. Schimmel PR & Flory PJ (1968). Conformational energies and configurational statistics of copolypeptides containing L-proline. *J Mol Biol* **34**, 105–120.

42. Brandts JF, Halvorson HR & Brennan M (1975). Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* **14**, 4953–4963.
43. Vanhoof G, Goossens F, Meester I de, Hendriks D & Scharpe S (1995). Proline motifs in peptides and their biological processing. *FASEB J* **9**, 736–744.
44. Dever TE, Gutierrez E & Shin B-S (2014). The hypusine-containing translation factor eIF5A. *Crit Rev Biochem Mol Biol* **49**, 413–425.
45. Mandal A, Mandal S & Park MH (2014). Genome-wide analyses and functional classification of proline repeat-rich proteins: potential role of eIF5A in eukaryotic evolution. *PLoS ONE* **9**, e111800.
46. Adzhubei AA, Sternberg MJE & Makarov AA (2013). Polyproline-II helix in proteins: structure and function. *J Mol Biol* **425**, 2100–2132.
47. Adzhubei AA & Sternberg MJ (1993). Left-handed polyproline II helices commonly occur in globular proteins. *J Mol Biol* **229**, 472–493.
48. Stapley BJ & Creamer TP (1999). A survey of left-handed polyproline II helices. *Protein Sci* **8**, 587–595.
49. Grathwohl C & Wüthrich K (1981). Nmr studies of the rates of proline cis - trans isomerization in oligopeptides. *Biopolymers* **20**, 2623–2633.
50. Dyson HJ, Rance M, Houghten RA, Lerner RA & Wright PE (1988). Folding of immunogenic peptide fragments of proteins in water solution. *J Mol Biol* **201**, 161–200.
51. Hersch SJ, Wang M, Zou SB, Moon K-M, Foster LJ, Ibba M & Navarre WW (2013). Divergent protein motifs direct elongation factor P-mediated translational regulation in *Salmonella enterica* and *Escherichia coli*. *mBio* **4**, e00180-13.
52. Veis A & Nawrot CF (1970). Basicity differences among peptide bonds. *J Am Chem Soc* **92**, 3910–3914.
53. Smith JA & Pease LG (1980). Reverse turns in peptides and proteins. *CRC Crit Rev Biochem* **8**, 315–399.
54. Deber CM, Glibowicka M & Woolley GA (1990). Conformations of proline residues in membrane environments. *Biopolymers* **29**, 149–157.
55. Barlow DJ & Thornton JM (1988). Helix geometry in proteins. *J Mol Biol* **201**, 601–619.
56. Richardson JS & Richardson DC (1989). Principles and Patterns of Protein Conformation. In: Prediction of Protein Structure and the Principles of Protein Conformation. (Fasman GD, ed), pp. 1–98. Boston, MA. Springer Verlag. (Buch)
57. Lu KP, Finn G, Lee TH & Nicholson LK (2007). Prolyl cis-trans isomerization as a molecular timer. *Nat Chem Biol* **3**, 619–629.
58. Counterman AE & Clemmer DE (2004). Anhydrous Polyproline Helices and Globules. *J Phys Chem B* **108**, 4885–4898.
59. Ude S, Lassak J, Starosta AL, Kraxenberger T, Wilson DN & Jung K (2013). Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* **339**, 82–85.
60. Polikanov YS, Steitz TA & Innis CA (2014). A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. *Nat Struct Mol Biol* **21**, 787–793.
61. Starosta AL, Lassak J, Peil L, Atkinson GC, Virumae K, Tenson T, Remme J, Jung K & Wilson DN (2014). Translational stalling at polyproline stretches is modulated by the sequence context upstream of the stall site. *Nucleic Acids Res* **42**, 10711–10719.
62. Lassak J, Wilson DN & Jung K (2016). Stall no more at polyproline stretches with the translation elongation factors EF-P and IF-5A. *Mol Microbiol* **99**, 219–235.
63. Hersch SJ, Elgamil S, Katz A, Ibba M & Navarre WW (2014). Translation initiation rate determines the impact of ribosome stalling on bacterial protein synthesis. *J Biol Chem* **289**, 28160–28171.

64. Cymer F & Heijne G von (2013). Cotranslational folding of membrane proteins probed by arrest-peptide-mediated force measurements. *Proc Natl Acad Sci U S A* **110**, 14640–14645.
65. Hanawa-Suetsugu K, Sekine S-i, Sakai H, Hori-Takemoto C, Terada T, Unzai S, Tame JRH, Kuramitsu S, Shirouzu M & Yokoyama S (2004). Crystal structure of elongation factor P from *Thermus thermophilus* HB8. *Proc Natl Acad Sci U S A* **101**, 9595–9600.
66. Choi S & Choe J (2011). Crystal structure of elongation factor P from *Pseudomonas aeruginosa* at 1.75 Å resolution. *Proteins* **79**, 1688–1693.
67. Kristensen O & Laurberg M (2002). Expression, refolding and crystallization of *Aquifex aeolicus* elongation factor P. *Acta Crystallogr D Biol Crystallogr* **58**, 1039–1041.
68. Aoki H, Adams SL, Chung DG, Yaguchi M, Chuang SE & Ganoza MC (1991). Cloning, sequencing and overexpression of the gene for prokaryotic factor EF-P involved in peptide bond synthesis. *Nucleic Acids Res* **19**, 6215–6220.
69. Chevance FFV, Le Guyon S & Hughes KT (2014). The effects of codon context on in vivo translation speed. *PLoS genetics* **10**, e1004392.
70. Doerfel LK & Rodnina MV (2013). Elongation factor P: Function and effects on bacterial fitness. *Biopolymers* **99**, 837–845.
71. Starosta AL, Lassak J, Peil L, Atkinson GC, Woolstenhulme CJ, Virumäe K, Buskirk A, Tenson T, Remme J, Jung K & Wilson DN (2014). A conserved proline triplet in Val-tRNA synthetase and the origin of elongation factor P. *Cell reports* **9**, 476–483.
72. Rossi D, Kuroshu R, Zanelli CF & Valentini SR (2014). eIF5A and EF-P: two unique translation factors are now traveling the same road. *Wiley Interdiscip Rev RNA*. **5**, 209–222.
73. Bartig D, Lemkemeier K, Frank J, Lottspeich F & Klink F (1992). The archaeobacterial hypusine-containing protein. Structural features suggest common ancestry with eukaryotic translation initiation factor 5A. *Eur J Biochem* **204**, 751–758.
74. Saini P, Eyler DE, Green R & Dever TE (2009). Hypusine-containing protein eIF5A promotes translation elongation. *Nature* **459**, 118–121.
75. Gutierrez E, Shin B-S, Woolstenhulme CJ, Kim J-R, Saini P, Buskirk AR & Dever TE (2013). eIF5A promotes translation of polyproline motifs. *Mol Cell* **51**, 35–45.
76. Dias CAO, Garcia W, Zanelli CF & Valentini SR (2013). eIF5A dimerizes not only in vitro but also in vivo and its molecular envelope is similar to the EF-P monomer. *Amino acids* **44**, 631–644.
77. Aoki H, Xu J, Emili A, Chosay JG, Golshani A & Ganoza MC (2008). Interactions of elongation factor EF-P with the *Escherichia coli* ribosome. *FEBS J* **275**, 671–681.
78. Blaha G, Stanley RE & Steitz TA (2009). Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. *Science* **325**, 966–970.
79. Katoh T, Wohlgemuth I, Nagano M, Rodnina MV & Suga H (2016). Essential structural elements in tRNA(Pro) for EF-P-mediated alleviation of translation stalling. *Nat Commun* **7**, 11657.
80. Huter P, Arenz S, Bock LV, Graf M, Frister JO, Heuer A, Peil L, Starosta AL, Wohlgemuth I, Peske F, Nováček J, Berninghausen O, Grubmüller H, Tenson T, Beckmann R, Rodnina MV, Vaiana AC & Wilson DN (2017). Structural basis for polyproline-mediated ribosome stalling and rescue by the translation elongation factor EF-P. *Mol Cell* **68**, 515–527.e6.
81. Doerfel LK, Wohlgemuth I, Kothe C, Peske F, Urlaub H & Rodnina MV (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* **339**, 85–88.
82. Park MH (2006). The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A (eIF5A). *J Biochem* **139**, 161–169.
83. Park J-H, Johansson HE, Aoki H, Huang BX, Kim H-Y, Ganoza MC & Park MH (2012). Post-translational modification by  $\beta$ -lysylation is required for activity of *Escherichia coli* elongation factor P (EF-P). *J Biol Chem* **287**, 2579–2590.

84. Prunetti L, Graf M, Blaby IK, Peil L, Makkay AM, Starosta AL, Papke RT, Oshima T, Wilson DN & Crécy-Lagard V de (2016). Deciphering the translation initiation factor 5A modification pathway in halophilic Archaea. *Archaea* **2016**, 7316725.
85. Mohapatra S, Choi H, Ge X, Sanyal S & Weisshaar JC (2017). Spatial distribution and ribosome-binding dynamics of EF-P in live *Escherichia coli*. *mBio* **8**.
86. An G, Glick BR, Friesen JD & Ganoza MC (1980). Identification and quantitation of elongation factor EF-P in *Escherichia coli* cell-free extracts. *Can J Biochem* **58**, 1312–1314.
87. Schmidt A, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, Knoops K, Bauer M, Aebersold R & Heinemann M (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nat Biotechnol* **34**, 104–110.
88. Schuller AP, Wu CC-C, Dever TE, Buskirk AR & Green R (2017). eIF5A functions globally in translation elongation and termination. *Mol Cell* **66**, 194–205.e5.
89. Pelechano V & Alepuz P (2017). eIF5A facilitates translation termination globally and promotes the elongation of many non polyproline-specific tripeptide sequences. *Nucleic Acids Res* **45**, 7326–7338.
90. Park MH, Cooper HL & Folk JE (1981). Identification of hypusine, an unusual amino acid, in a protein from human lymphocytes and of spermidine as its biosynthetic precursor. *Proc Natl Acad Sci U S A* **78**, 2869–2873.
91. Yanagisawa T, Sumida T, Ishii R, Takemoto C & Yokoyama S (2010). A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. *Nat Struct Mol Biol* **17**, 1136–1143.
92. Peil L, Starosta AL, Virumäe K, Atkinson GC, Tenson T, Remme J & Wilson DN (2012). Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. *Nat Chem Biol* **8**, 695–697.
93. Lassak J, Keilhauer EC, Furst M, Wuichet K, Godeke J, Starosta AL, Chen J-M, Sogaard-Andersen L, Rohr J, Wilson DN, Haussler S, Mann M & Jung K (2015). Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat Chem Biol* **11**, 266–270.
94. Sengoku T, Suzuki T, Dohmae N, Watanabe C, Honma T, Hikida Y, Yamaguchi Y, Takahashi H, Yokoyama S & Yanagisawa T (2018). Structural basis of protein arginine rhamnosylation by glycosyltransferase EarP. *Nat Chem Biol* **14**, 368–374.
95. Krafczyk R, Macošek J, Jagtap PKA, Gast D, Wunder S, Mitra P, Jha AK, Rohr J, Hoffmann-Röder A, Jung K, Hennig J & Lassak J (2017). Structural Basis for EarP-Mediated Arginine Glycosylation of Translation Elongation Factor EF-P. *mBio* **8**.
96. Witzky A, Hummels KR, Tollerson R2, Rajkovic A, Jones LA, Kearns DB & Ibba M (2018). EF-P posttranslational modification has variable impact on polyproline translation in *Bacillus subtilis*. *mBio* **9**.
97. Rajkovic A, Hummels KR, Witzky A, Erickson S, Gafken PR, Whitelegge JP, Faull KF, Kearns DB & Ibba M (2016). Translation control of swarming proficiency in *Bacillus subtilis* by 5-amino-pentanolyated elongation factor P. *J Biol Chem* **291**, 10976–10985.
98. Rajkovic A & Ibba M (2017). Elongation factor P and the control of translation elongation. *Annu Rev Microbiol* **71**, 117–131.
99. Marman HE, Mey AR & Payne SM (2014). Elongation factor P and modifying enzyme PoxA are necessary for virulence of *Shigella flexneri*. *Infect Immun* **82**, 3612–3621.
100. Balibar CJ, Iwanowicz D & Dean CR (2013). Elongation factor P is dispensable in *Escherichia coli* and *Pseudomonas aeruginosa*. *Curr Microbiol* **67**, 293–299.
101. Zou SB, Roy H, Ibba M & Navarre WW (2011). Elongation factor P mediates a novel post-transcriptional regulatory pathway critical for bacterial virulence. *Virulence* **2**, 147–151.
102. Caraglia M, Park MH, Wolff EC, Marra M & Abbruzzese A (2013). eIF5A isoforms and cancer: two brothers for two functions? *Amino acids* **44**, 103–109.

103. Clement PMJ, Henderson CA, Jenkins ZA, Smit-McBride Z, Wolff EC, Hershey JWB, Park MH & Johansson HE (2003). Identification and characterization of eukaryotic initiation factor 5A-2. *Eur J Biochem* **270**, 4254–4263.
104. Mathews MB & Hershey JWB (2015). The translation factor eIF5A and human cancer. *Biochim Biophys Acta* **1849**, 836–844.
105. Guan XY, Sham JS, Tang TC, Fang Y, Huo KK & Yang JM (2001). Isolation of a novel candidate oncogene within a frequently amplified region at 3q26 in ovarian cancer. *Cancer research* **61**, 3806–3809.
106. Kolligs FT (2012). An alternative way for epithelial-to-mesenchymal transition in colorectal cancer via EIF5A2? *Gut* **61**, 473–474.
107. Cracchiolo BM, Heller DS, Clement PMJ, Wolff EC, Park M-H & Hanauke-Abel HM (2004). Eukaryotic initiation factor 5A-1 (eIF5A-1) as a diagnostic marker for aberrant proliferation in intraepithelial neoplasia of the vulva. *Gynecol Oncol* **94**, 217–222.
108. Yang S-S, Gao Y, Wang D-Y, Xia B-R, Liu Y-D, Qin Y, Ning X-M, Li G-Y, Hao L-X, Xiao M & Zhang Y-Y (2016). Overexpression of eukaryotic initiation factor 5A2 (EIF5A2) is associated with cancer progression and poor prognosis in patients with early-stage cervical cancer. *Histopathology* **69**, 276–287.
109. Pällmann N, Braig M, Sievert H, Preukschas M, Hermans-Borgmeyer I, Schweizer M, Nagel CH, Neumann M, Wild P, Haralambieva E, Hagel C, Bokemeyer C, Hauber J & Balabanov S (2015). Biological relevance and therapeutic potential of the hypusine modification system. *J Biol Chem* **290**, 18343–18360.
110. Park MH, Nishimura K, Zanelli CF & Valentini SR (2010). Functional significance of eIF5A and its hypusine modification in eukaryotes. *Amino acids* **38**, 491–500.
111. Richardson JS & Richardson DC (1988). Amino acid preferences for specific locations at the ends of alpha helices. *Science* **240**, 1648–1652.
112. Argos P & Palau J (1982). Amino acid distribution in protein secondary structures. *Int J Pept Protein Res* **19**, 380–393.
113. Brandl CJ & Deber CM (1986). Hypothesis about the function of membrane-buried proline residues in transport proteins. *Proc Natl Acad Sci U S A* **83**, 917–921.
114. Sansom MS (1992). Proline residues in transmembrane helices of channel and transport proteins: a molecular modelling study. *Protein engineering* **5**, 53–60.
115. Murray NJ, Williamson MP, Lilley TH & Haslam E (1994). Study of the interaction between salivary proline-rich proteins and a polyphenol by 1H-NMR spectroscopy. *Eur J Biochem* **219**, 923–935.
116. Kawakami T, Akizawa Y, Ishikawa T, Shimamoto T, Tsuda M & Tsuchiya T (1988). Amino acid substitutions and alteration in cation specificity in the melibiose carrier of *Escherichia coli*. *J Biol Chem* **263**, 14276–14280.
117. Rothschild KJ, He YW, Gray D, Roepe PD, Pelletier SL, Brown RS & Herzfeld J (1989). Fourier transform infrared evidence for proline structural changes during the bacteriorhodopsin photocycle. *Proc Natl Acad Sci U S A* **86**, 9832–9835.
118. Kay BK, Williamson MP & Sudol M (2000). The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J* **14**, 231–241.
119. Zarrinpar A, Bhattacharyya RP & Lim WA (2003). The structure and function of proline recognition domains. *Sci STKE* **2003**, RE8.
120. Adzhubei AA, Eisenmenger F, Tumanyan VG, Zinke M, Brodzinski S & Esipova NG (1987). Third type of secondary structure: noncooperative mobile conformation. Protein Data Bank analysis. *Biochem Biophys Res Commun* **146**, 934–938.
121. Jha AK, Colubri A, Zaman MH, Koide S, Sosnick TR & Freed KF (2005). Helix, sheet, and polyproline II frequencies and strong nearest neighbor effects in a restricted coil library. *Biochemistry* **44**, 9691–9702.
122. Strader CD, Sigal IS, Register RB, Candelore MR, Rands E & Dixon RA (1987). Identification of residues required for ligand binding to the beta-adrenergic receptor. *Proc Natl Acad Sci U S A* **84**, 4384–4388.

123. Sankara-Ramakrishnan R & Vishveshwara S (1989). A hydrogen bonded chain in bacteriorhodopsin by computer modelling approach. *J Biomol Struct Dyn* **7**, 187–205.
124. Vilsen B, Andersen JP, Clarke DM & MacLennan DH (1989). Functional consequences of proline mutations in the cytoplasmic and transmembrane sectors of the Ca<sup>2+</sup>-ATPase of sarcoplasmic reticulum. *J Biol Chem* **264**, 21024–21030.
125. Lummis SCR, Beene DL, Lee LW, Lester HA, Broadhurst RW & Dougherty DA (2005). Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature* **438**, 248–252.
126. Lolkema JS, Püttner IB & Kaback HR (1988). Site-directed mutagenesis of Pro327 in the lac permease of *Escherichia coli*. *Biochemistry* **27**, 8307–8310.
127. Vik SB, Cain BD, Chun KT & Simoni RD (1988). Mutagenesis of the alpha subunit of the F<sub>1</sub>F<sub>0</sub>-ATPase from *Escherichia coli*. Mutations at Glu-196, Pro-190, and Ser-199. *J Biol Chem* **263**, 6599–6605.
128. Miller MJ, Fraga D, Paule CR & Fillingame RH (1989). Mutations in the conserved proline 43 residue of the uncE protein (subunit c) of *Escherichia coli* F<sub>1</sub>F<sub>0</sub>-ATPase alter the coupling of F<sub>1</sub> to F<sub>0</sub>. *J Biol Chem* **264**, 305–311.
129. Seckler R & Jaenicke R (1992). Protein folding and protein refolding. *FASEB J* **6**, 2545–2552.
130. Woolfson DN & Williams DH (1990). The influence of proline residues on alpha-helical structure. *FEBS letters* **277**, 185–188.
131. Choi E, Choi S, Nam D, Park S, Han Y, Lee J-S & Lee E-J (2017). Elongation factor P restricts *Salmonella*'s growth by controlling translation of a Mg<sup>2+</sup> transporter gene during infection. *Scientific reports* **7**, 42098.
132. Gall AR, Datsenko KA, Figueroa-Bossi N, Bossi L, Masuda I, Hou Y-M & Csonka LN (2016). Mg<sup>2+</sup> regulates transcription of *mgtA* in *Salmonella typhimurium* via translation of proline codons during synthesis of the MgtL peptide. *Proc Natl Acad Sci U S A* **113**, 15096–15101.
133. Watson N, Duniak DS, Rosey EL, Slonczewski JL & Olson ER (1992). Identification of elements involved in transcriptional regulation of the *Escherichia coli cad* operon by external pH. *J Biol Chem* **174**, 530–540.
134. Buchner S, Schlundt A, Lassak J, Sattler M & Jung K (2015). Structural and functional analysis of the signal-transducing linker in the pH-responsive one-component system CadC of *Escherichia coli*. *J Biol Chem* **427**, 2548–2561.
135. Mascher T, Helmann JD & Uden G (2006). Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol Mol Biol Rev* **70**, 910–938.
136. Dutta R, Qin L & Inouye M (1999). Histidine kinases: diversity of domain organization. *Mol Microbiol* **34**, 633–640.
137. Bhate MP, Molnar KS, Goulian M & DeGrado WF (2015). Signal transduction in histidine kinases: insights from new structures. *Structure* **23**, 981–994.
138. Galperin MY (2006). Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J Bacteriol* **188**, 4169–4182.
139. Appleby JL, Parkinson JS & Bourret RB (1996). Signal transduction via the multi-step phosphorelay: not necessarily a road less traveled. *Cell* **86**, 845–848.
140. Taniguchi Y, Choi PJ, Li G-W, Chen H, Babu M, Hearn J, Emili A & Xie XS (2010). Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533–538.
141. Cai SJ & Inouye M (2002). EnvZ-OmpR interaction and osmoregulation in *Escherichia coli*. *J Biol Chem* **277**, 24155–24161.
142. Mizuno T, Wurtzel ET & Inouye M (1982). Osmoregulation of gene expression. II. DNA sequence of the *envZ* gene of the *ompB* operon of *Escherichia coli* and characterization of its gene product. *J Biol Chem* **257**, 13692–13698.
143. Jung K, Hamann K & Revermann A (2001). K<sup>+</sup> stimulates specifically the autokinase activity of purified and reconstituted EnvZ of *Escherichia coli*. *J Biol Chem* **276**, 40896–40902.

144. Libby EA, Ekici S & Goulian M (2010). Imaging OmpR binding to native chromosomal loci in *Escherichia coli*. *J Bacteriol* **192**, 4045–4053.
145. King ST & Kenney LJ (2007). Application of fluorescence resonance energy transfer to examine EnvZ/OmpR interactions. *Methods Enzymol* **422**, 352–360.
146. Foo YH, Gao Y, Zhang H & Kenney LJ (2015). Cytoplasmic sensing by the inner membrane histidine kinase EnvZ. *Prog Biophys Mol Biol* **118**, 119–129.
147. Wang LC, Morgan LK, Godakumbura P, Kenney LJ & Anand GS (2012). The inner membrane histidine kinase EnvZ senses osmolality via helix-coil transitions in the cytoplasm. *EMBO J* **31**, 2648–2659.
148. Ferris HU, Coles M, Lupas AN & Hartmann MD (2014). Crystallographic snapshot of the *Escherichia coli* EnvZ histidine kinase in an active conformation. *J Struct Biol* **186**, 376–379.
149. Ghosh M, Wang LC, Ramesh R, Morgan LK, Kenney LJ & Anand GS (2017). Lipid-Mediated Regulation of Embedded Receptor Kinases via Parallel Allosteric Relays. *Biophys J* **112**, 643–654.
150. Dutta R & Inouye M (1996). Reverse phosphotransfer from OmpR to EnvZ in a kinase-/phosphatase+ mutant of EnvZ (EnvZ.N347D), a bifunctional signal transducer of *Escherichia coli*. *J Biol Chem* **271**, 1424–1429.
151. Forst S, Delgado J & Inouye M (1989). Phosphorylation of OmpR by the osmosensor EnvZ modulates expression of the ompF and ompC genes in *Escherichia coli*. *Proc Natl Acad Sci U S A* **86**, 6052–6056.
152. Nikaido H (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* **67**, 593–656.
153. Gerken H & Misra R (2010). MzrA-EnvZ interactions in the periplasm influence the EnvZ/OmpR two-component regulon. *J Bacteriol* **192**, 6271–6278.
154. Gerken H, Charlson ES, Cicirelli EM, Kenney LJ & Misra R (2009). MzrA: a novel modulator of the EnvZ/OmpR two-component regulon. *Mol Microbiol* **72**, 1408–1422.
155. Itou J, Eguchi Y & Utsumi R (2009). Molecular mechanism of transcriptional cascade initiated by the EvgS/EvgA system in *Escherichia coli* K-12. *Biosci Biotechnol Biochem* **73**, 870–878.
156. Sen H, Aggarwal N, Ishionwu C, Hussain N, Parmar C, Jamshad M, Bavro VN & Lund PA (2017). Structural and functional analysis of the *Escherichia coli* acid-sensing histidine kinase EvgS. *J Bacteriol* **199**.
157. Perraud AL, Kimmel B, Weiss V & Gross R (1998). Specificity of the BvgAS and EvgAS phosphorelay is mediated by the C-terminal Hpt domains of the sensor proteins. *Mol Microbiol* **27**, 875–887.
158. Johnson MD, Bell J, Clarke K, Chandler R, Pathak P, Xia Y, Marshall RL, Weinstock GM, Loman NJ, Winn PJ & Lund PA (2014). Characterization of mutations in the PAS domain of the EvgS sensor kinase selected by laboratory evolution for acid resistance in *Escherichia coli*. *Mol Microbiol* **93**, 911–927.
159. Utsumi R, Katayama S, Taniguchi M, Horie T, Ikeda M, Igaki S, Nakagawa H, Miwa A, Tanabe H & Noda M (1994). Newly identified genes involved in the signal transduction of *Escherichia coli* K-12. *Gene* **140**, 73–77.
160. Perraud AL, Rippe K, Bantscheff M, Glocker M, Lucassen M, Jung K, Sebald W, Weiss V & Gross R (2000). Dimerization of signalling modules of the EvgAS and BvgAS phosphorelay systems. *Biochim Biophys Acta* **1478**, 341–354.
161. Sommer E, Koler M, Frank V, Sourjik V & Vaknin A (2013). The sensory histidine kinases TorS and EvgS tend to form clusters in *Escherichia coli* cells. *PLoS ONE* **8**, e77708.
162. Kinoshita-Kikuta E, Kinoshita E, Eguchi Y, Yanagihara S, Edahiro K, Inoue Y, Taniguchi M, Yoshida M, Yamamoto K, Takahashi H, Sawasaki T, Utsumi R & Koike T (2015). Functional characterization of the receiver domain for phosphorelay control in hybrid sensor kinases. *PLoS ONE* **10**, e0132598.
163. Masuda N & Church GM (2003). Regulatory network of acid resistance genes in *Escherichia coli*. *Mol Microbiol* **48**, 699–712.

164. Smith DK, Kassam T, Singh B & Elliott JF (1992). *Escherichia coli* has two homologous glutamate decarboxylase genes that map to distinct loci. *J Bacteriol* **174**, 5820–5826.
165. Castanie-Cornet MP, Penfound TA, Smith D, Elliott JF & Foster JW (1999). Control of acid resistance in *Escherichia coli*. *J Bacteriol* **181**, 3525–3535.
166. Ma Z, Masuda N & Foster JW (2004). Characterization of EvgAS-YdeO-GadE branched regulatory circuit governing glutamate-dependent acid resistance in *Escherichia coli*. *J Bacteriol* **186**, 7378–7389.
167. Kato A, Ohnishi H, Yamamoto K, Furuta E, Tanabe H & Utsumi R (2000). Transcription of emrKY is regulated by the EvgA-EvgS two-component system in *Escherichia coli* K-12. *Biosci Biotechnol Biochem* **64**, 1203–1209.
168. Eguchi Y, Itou J, Yamane M, Demizu R, Yamato F, Okada A, Mori H, Kato A & Utsumi R (2007). B1500, a small membrane protein, connects the two-component systems EvgS/EvgA and PhoQ/PhoP in *Escherichia coli*. *Proc Natl Acad Sci U S A* **104**, 18712–18717.
169. Eguchi Y, Ishii E, Hata K & Utsumi R (2011). Regulation of acid resistance by connectors of two-component signal transduction systems in *Escherichia coli*. *J Bacteriol* **193**, 1222–1228.
170. Eguchi Y, Okada T, Minagawa S, Oshima T, Mori H, Yamamoto K, Ishihama A & Utsumi R (2004). Signal transduction cascade between EvgA/EvgS and PhoP/PhoQ two-component systems of *Escherichia coli*. *J Bacteriol* **186**, 3006–3014.
171. Castanié-Cornet M-P, Cam K, Bastiat B, Cros A, Bordes P & Gutierrez C (2010). Acid stress response in *Escherichia coli*: mechanism of regulation of gadA transcription by RcsB and GadE. *Nucleic acids research* **38**, 3546–3554.
172. Johnson MD, Burton NA, Gutiérrez B, Painter K & Lund PA (2011). RcsB is required for inducible acid resistance in *Escherichia coli* and acts at gadE-dependent and -independent promoters. *J Bacteriol* **193**, 3653–3656.
173. Povolotsky TL & Hengge R (2016). Genome-based comparison of cyclic di-GMP signaling in pathogenic and commensal *Escherichia coli* strains. *J Bacteriol* **198**, 111–126.
174. Schirmer T & Jenal U (2009). Structural and mechanistic determinants of c-di-GMP signalling. *Nat Rev Microbiol* **7**, 724–735.
175. Sarenko O, Klauck G, Wilke FM, Pfiffer V, Richter AM, Herbst S, Kaefer V & Hengge R (2017). More than enzymes that make or break cyclic di-GMP-local signaling in the interactome of GGDEF/EAL domain proteins of *Escherichia coli*. *mBio* **8**.
176. Nikolskaya AN, Mulkidjanian AY, Beech IB & Galperin MY (2003). MASE1 and MASE2: two novel integral membrane sensory domains. *J Mol Microbiol Biotechnol* **5**, 11–16.
177. Hengge R (2009). Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* **7**, 263–273.
178. Paul R, Abel S, Wassmann P, Beck A, Heerklotz H & Jenal U (2007). Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization. *J Biol Chem* **282**, 29170–29177.
179. Wassmann P, Chan C, Paul R, Beck A, Heerklotz H, Jenal U & Schirmer T (2007). Structure of BeF3<sup>-</sup>-modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure* **15**, 915–927.
180. Serra DO, Richter AM & Hengge R (2013). Cellulose as an architectural element in spatially structured *Escherichia coli* biofilms. *J Bacteriol* **195**, 5540–5554.
181. Serra DO & Hengge R (2017). Experimental Detection and Visualization of the Extracellular Matrix in Macrocolony Biofilms. *Methods Mol Biol* **1657**, 133–145.
182. Krasteva PV, Bernal-Bayard J, Travier L, Martin FA, Kaminski P-A, Karimova G, Fronzes R & Ghigo J-M (2017). Insights into the structure and assembly of a bacterial cellulose secretion system. *Nature communications* **8**.
183. Römling U & Galperin MY (2015). Bacterial cellulose biosynthesis: diversity of operons, subunits, products and functions. *Trends Microbiol* **23**, 545–557.

184. Solano C, García B, Valle J, Berasain C, Ghigo J-M, Gamazo C & Lasa I (2002). Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. *Mol Microbiol* **43**, 793–808.
185. Morgan JLW, Strumillo J & Zimmer J (2013). Crystallographic snapshot of cellulose synthesis and membrane translocation. *Nature* **493**, 181–186.
186. Whitney JC & Howell PL (2012). Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. *Trends Microbiol* **21**, 63–72.
187. Zogaj X, Nimtz M, Rohde M, Bokranz W & Römling U (2001). The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. *Mol Microbiol* **39**, 1452–1463.
188. Bokranz W, Wang X, Tschäpe H & Römling U (2005). Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. *J Med Microbiol* **54**, 1171–1182.
189. Weber H, Pesavento C, Possling A, Tischendorf G & Hengge R (2006). Cyclic-di-GMP-mediated signalling within the sigma network of *Escherichia coli*. *Mol Microbiol* **62**, 1014–1034.
190. Nishii W, Kukimoto-Niino M, Terada T, Shirouzu M, Muramatsu T, Kojima M, Kihara H & Yokoyama S (2015). A redox switch shapes the Lon protease exit pore to facultatively regulate proteolysis. *Nat Chem Biol* **11**, 46–51.
191. Amerik A, Antonov VK, Gorbalenya AE, Kotova SA, Rotanova TV & Shimbarevich EV (1991). Site-directed mutagenesis of La protease. A catalytically active serine residue. *FEBS letters* **287**, 211–214.
192. van Melderen L & Aertsen A (2009). Regulation and quality control by Lon-dependent proteolysis. *Res Microbiol* **160**, 645–651.
193. Vieux EF, Wohlever ML, Chen JZ, Sauer RT & Baker TA (2013). Distinct quaternary structures of the AAA+ Lon protease control substrate degradation. *Proc Natl Acad Sci U S A* **110**, E2002–8.
194. Fredriksson A, Ballesteros M, Dukan S & Nyström T (2005). Defense against protein carbonylation by DnaK/DnaJ and proteases of the heat shock regulon. *J Bacteriol* **187**, 4207–4213.
195. Rosen R, Biran D, Gur E, Becher D, Hecker M & Ron EZ (2002). Protein aggregation in *Escherichia coli*: role of proteases. *FEMS Microbiol Lett* **207**, 9–12.
196. Gerdes K & Maisonneuve E (2012). Bacterial persistence and toxin-antitoxin loci. *Annu Rev Microbiol* **66**, 103–123.
197. Bissonnette SA, Rivera-Rivera I, Sauer RT & Baker TA (2010). The IbpA and IbpB small heat-shock proteins are substrates of the AAA+ Lon protease. *Mol Microbiol* **75**, 1539–1549.
198. Gottesman S, Halpern E & Trisler P (1981). Role of sulA and sulB in filamentation by lon mutants of *Escherichia coli* K-12. *J Bacteriol* **148**, 265–273.
199. Rotanova TV, Botos I, Melnikov EE, Rasulova F, Gustchina A, Maurizi MR & Wlodawer A (2006). Slicing a protease: structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains. *Protein Sci* **15**, 1815–1828.
200. Neuwald AF, Aravind L, Spouge JL & Koonin EV (1999). AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome research* **9**, 27–43.
201. Rotanova TV, Melnikov EE, Khalatova AG, Makhovskaya OV, Botos I, Wlodawer A & Gustchina A (2004). Classification of ATP-dependent proteases Lon and comparison of the active sites of their proteolytic domains. *Eur J Biochem* **271**, 4865–4871.
202. Nishii W, Suzuki T, Nakada M, Kim Y-T, Muramatsu T & Takahashi K (2005). Cleavage mechanism of ATP-dependent Lon protease toward ribosomal S2 protein. *FEBS letters* **579**, 6846–6850.
203. van Melderen L, Thi MH, Lecchi P, Gottesman S, Couturier M & Maurizi MR (1996). ATP-dependent degradation of CcdA by Lon protease. Effects of secondary structure and heterologous subunit interactions. *J Biol Chem* **271**, 27730–27738.

204. Cha S-S, An YJ, Lee CR, Lee HS, Kim Y-G, Kim SJ, Kwon KK, Donatis GM de, Lee J-H, Maurizi MR & Kang SG (2010). Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *The EMBO journal* **29**, 3520–3530.
205. Sauer RT & Baker TA (2011). AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **80**, 587–612.
206. Ogura T & Wilkinson AJ (2001). AAA+ superfamily ATPases: common structure--diverse function. *Genes Cells* **6**, 575–597.
207. Iyer LM, Leipe DD, Koonin EV & Aravind L (2004). Evolutionary history and higher order classification of AAA+ ATPases. *J Struct Biol* **146**, 11–31.
208. Martin A, Baker TA & Sauer RT (2005). Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. *Nature* **437**, 1115–1120.
209. Patterson J, Vineyard D, Thomas-Wohlever J, Behshad R, Burke M & Lee I (2004). Correlation of an adenine-specific conformational change with the ATP-dependent peptidase activity of *Escherichia coli* Lon. *Biochemistry* **43**, 7432–7442.
210. Maurizi MR (1992). Proteases and protein degradation in *Escherichia coli*. *Experientia* **48**, 178–201.
211. Starkova NN, Koroleva EP, Rumsh LD, Ginodman LM & Rotanova TV (1998). Mutations in the proteolytic domain of *Escherichia coli* protease Lon impair the ATPase activity of the enzyme. *FEBS letters* **422**, 218–220.
212. Botos I, Melnikov EE, Cherry S, Tropea JE, Khalatova AG, Rasulova F, Dauter Z, Maurizi MR, Rotanova TV, Wlodawer A & Gustchina A (2004). The catalytic domain of *Escherichia coli* Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J Biol Chem* **279**, 8140–8148.
213. Bertani I, Rampioni G, Leoni L & Venturi V (2007). The *Pseudomonas putida* Lon protease is involved in N-acyl homoserine lactone quorum sensing regulation. *BMC Microbiol* **7**, 71.
214. Takaya A, Tabuchi F, Tsuchiya H, Isogai E & Yamamoto T (2008). Negative regulation of quorum-sensing systems in *Pseudomonas aeruginosa* by ATP-dependent Lon protease. *J Bacteriol* **190**, 4181–4188.
215. Barembruch C & Hengge R (2007). Cellular levels and activity of the flagellar sigma factor FliA of *Escherichia coli* are controlled by FlgM-modulated proteolysis. *Mol Microbiol* **65**, 76–89.
216. Heuveling J, Possling A & Hengge R (2008). A role for Lon protease in the control of the acid resistance genes of *Escherichia coli*. *Mol Microbiol* **69**, 534–547.
217. Griffith KL, Shah IM & Wolf RE (2004). Proteolytic degradation of *Escherichia coli* transcription activators SoxS and MarA as the mechanism for reversing the induction of the superoxide (SoxRS) and multiple antibiotic resistance (Mar) regulons. *Mol Microbiol* **51**, 1801–1816.
218. Mizusawa S & Gottesman S (1983). Protein degradation in *Escherichia coli*: the lon gene controls the stability of sulA protein. *Proc Natl Acad Sci U S A* **80**, 358–362.
219. Gonzalez M & Woodgate R (2002). The "tale" of UmuD and its role in SOS mutagenesis. *BioEssays* **24**, 141–148.
220. Kuroda A, Nomura K, Ohtomo R, Kato J, Ikeda T, Takiguchi N, Ohtake H & Kornberg A (2001). Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in *E. coli*. *Science* **293**, 705–708.
221. Harms A, Brodersen DE, Mitarai N & Gerdes K (2018). Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. *Mol Cell* **70**, 768–784.
222. Gerdes K, Christensen SK & Løbner-Olesen A (2005). Prokaryotic toxin-antitoxin stress response loci. *Nature reviews. Microbiology* **3**, 371–382.
223. Brzozowska I & Zielenkiewicz U (2013). Regulation of toxin-antitoxin systems by proteolysis. *Plasmid* **70**, 33–41.
224. Wang X & Wood TK (2011). Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Appl Environ Microbiol* **77**, 5577–5583.

225. Hansen S, Vulić M, Min J, Yen T-J, Schumacher MA, Brennan RG & Lewis K (2012). Regulation of the *Escherichia coli* HipBA toxin-antitoxin system by proteolysis. *PLoS ONE* **7**, e39185.
226. Tsilibaris V, Maenhaut-Michel G & van Melderen L (2006). Biological roles of the Lon ATP-dependent protease. *Res Microbiol* **157**, 701–713.
227. Christensen SK, Mikkelsen M, Pedersen K & Gerdes K (2001). RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc Natl Acad Sci U S A* **98**, 14328–14333.
228. Christensen SK, Pedersen K, Hansen FG & Gerdes K (2003). Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. *J Mol Biol* **332**, 809–819.
229. Christensen-Dalsgaard M, Jørgensen MG & Gerdes K (2010). Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Mol Microbiol* **75**, 333–348.
230. Goff SA & Goldberg AL (1987). An increased content of protease La, the lon gene product, increases protein degradation and blocks growth in *Escherichia coli*. *J Biol Chem* **262**, 4508–4515.
231. Goff SA, Casson LP & Goldberg AL (1984). Heat shock regulatory gene htpR influences rates of protein degradation and expression of the lon gene in *Escherichia coli*. *Proc Natl Acad Sci U S A* **81**, 6647–6651.
232. Phillips TA, VanBogelen RA & Neidhardt FC (1984). lon gene product of *Escherichia coli* is a heat-shock protein. *J Bacteriol* **159**, 283–287.
233. Maisonneuve E, Castro-Camargo M & Gerdes K (2013). (p)ppGpp controls bacterial persistence by stochastic induction of toxin-antitoxin activity. *Cell* **154**, 1140–1150.
234. Studier FW & Moffatt BA (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113–130.
235. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL & Mori H (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006.0008.
236. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B & Shao Y (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462.
237. Heermann R, Zeppenfeld T & Jung K (2008). Simple generation of site-directed point mutations in the *Escherichia coli* chromosome using Red(R)/ET(R) Recombination. *Microb Cell Fact* **7**, 14.
238. Duchaud E, Rusniok C, Frangeul L, Buchrieser C, Givaudan A, Taourit S, Bocs S, Boursaux-Eude C, Chandler M, Charles J-F, Dassa E, Derose R, Derzelle S, Freyssinet G, Gaudriault S, Medigue C, Lanois A, Powell K, Siguier P, Vincent R, Wingate V, Zouine M, Glaser P, Boemare N, Danchin A & Kunst F (2003). The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. *Nature biotechnology* **21**, 1307–1313.
239. Karimova G, Ullmann A & Ladant D (2000). A bacterial two-hybrid system that exploits a cAMP signaling cascade in *Escherichia coli*. *Methods Enzymol* **328**, 59–73.
240. Kovach ME, Elzer PH, Steven Hill D, Robertson GT, Farris MA, Roop RM & Peterson KM (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**, 175–176.
241. Fried L, Lassak J & Jung K (2012). A comprehensive toolbox for the rapid construction of lacZ fusion reporters. *J Microbiol Methods* **91**, 537–543.
242. Guzman LM, Belin D, Carson MJ & Beckwith J (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**, 4121–4130.
243. Russo FD & Silhavy TJ (1991). EnvZ controls the concentration of phosphorylated OmpR to mediate osmoregulation of the porin genes. *J Mol Biol* **222**, 567–580.
244. Cherepanov PP & Wackernagel W (1995). Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9–14.

245. Lassak J, Henche A-L, Binnenkade L & Thormann KM (2010). ArcS, the cognate sensor kinase in an atypical Arc system of *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* **76**, 3263–3274.
246. Miller J (1972). Experiments in molecular genetics. *Cold Spring Harbor Laboratory Press, New York*.
247. Epstein W & Kim BS (1971). Potassium transport loci in *Escherichia coli* K-12. *J Bacteriol* **108**, 639–644.
248. van Baarle S, Bramkamp M & Rénia L (2010). The MinCDJ system in *Bacillus subtilis* prevents minicell formation by promoting divisome disassembly. *PLoS ONE* **5**, e9850.
249. Mullis K, Faloona F, Scharf S, Saiki R, Horn G & Erlich H (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* **51 Pt 1**, 263–273.
250. Sanger F, Nicklen S & Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463–5467.
251. Ausubel, M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (1987) Current protocols in molecular biology. Wiley-Blackwell, New York. (Buch)
252. Maniatis, T., Fritsch, E., Sambrook, J. (1989) Molecular cloning. Cold spring harbor laboratory press, New York. (Buch)
253. Datsenko KA & Wanner BL (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640–6645.
254. Dagert M & Ehrlich SD (1979). Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene* **6**, 23–28.
255. Inoue H, Nojima H & Okayama H (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23–28.
256. Teramoto H, Inui M & Yukawa H (2013). OxyR acts as a transcriptional repressor of hydrogen peroxide-inducible antioxidant genes in *Corynebacterium glutamicum*. *R. FEBS J* **280**, 3298–3312.
257. Pfaffl MW (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**, e45.
258. Laemmli UK (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
259. Lugtenberg B, Meijers J, Peters R, van der Hoek P & van Alphen L (1975). Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. *FEBS letters* **58**, 254–258.
260. Weber K & Osborn M (1969). The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. *J Biol Chem* **244**, 4406–4412.
261. Blum, H., Beier, H., Gross, H.J (1987). Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis*, 93–99.
262. Boivin S, Kozak S & Meijers R (2013). Optimization of protein purification and characterization using ThermoFluor screens. *Protein Expr Purif* **91**, 192–206.
263. Shevchenko A, Chernushevich I, Wilm M & Mann M (2000). De Novo peptide sequencing by nanoelectrospray tandem mass spectrometry using triple quadrupole and quadrupole/time-of-flight instruments. *Methods Mol Biol* **146**, 1–16.
264. Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T & Mann M (1996). Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **379**, 466–469.
265. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M & Cox J (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature methods* **13**, 731–740.
266. Vizcaíno JA, Csordas A, del-Toro N, Dienes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu Q-W, Wang R & Hermjakob H (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* **44**, D447–56.

267. Karimova G, Pidoux J, Ullmann A & Ladant D (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A* **95**, 5752–5756.
268. Karimova G, Dautin N & Ladant D (2005). Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. *J Bacteriol* **187**, 2233–2243.
269. Karimova G, Ladant D, Ullmann A, Selig L, Legrain P Bacterial two-hybrid system for protein-protein interaction screening, new strains for use therein, and their applications: US Application.
270. Batchelor E & Goulian M (2006). Imaging OmpR localization in *Escherichia coli*. *Mol Microbiol* **59**, 1767–1778.
271. Masuda N & Church GM (2002). *Escherichia coli* gene expression responsive to levels of the response regulator EvgA. *J Bacteriol* **184**, 6225–6234.
272. Richter AM, Povolotsky TL, Wieler LH & Hengge R (2014). Cyclic-di-GMP signalling and biofilm-related properties of the Shiga toxin-producing 2011 German outbreak *Escherichia coli* O104:H4. *EMBO molecular medicine* **6**, 1622–1637.
273. Primo ED, Otero LH, Ruiz F, Klinke S & Giordano W (2018). The disruptive effect of lysozyme on the bacterial cell wall explored by an in-silico structural outlook. *Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology* **46**, 83–90.
274. Di Paolo D, Afanar O, Armitage JP & Berry RM (2016). Single-molecule imaging of electroporated dye-labelled CheY in live *Escherichia coli*. *Philos Trans R Soc Lond B Biol Sci* **371**.
275. Bach JN, Giacomelli G & Bramkamp M (2017). Sample preparation and choice of fluorophores for single and dual color photo-activated localization microscopy (PALM) with bacterial cells. *Methods Mol Biol* **1563**, 129–141.
276. Durisic N, Laparra-Cuervo L, Sandoval-Álvarez A, Borbely JS & Lakadamyali M (2014). Single-molecule evaluation of fluorescent protein photoactivation efficiency using an in vivo nanotemplate. *Nature methods* **11**, 156–162.
277. Malkusch S, Endesfelder U, Mondry J, Gelléri M, Verveer PJ & Heilemann M (2012). Coordinate-based colocalization analysis of single-molecule localization microscopy data. *Histochem Cell Biol* **137**, 1–10.
278. UniProt Consortium T (2018). UniProt: the universal protein knowledgebase. *Nucleic Acids Res* **46**, 2699.
279. Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martinez C, Caspi R, Fulcher C, Gama-Castro S, Kothari A, Krummenacker M, Latendresse M, Muniz-Rascado L, Ong Q, Paley S, Peralta-Gil M, Subhraveti P, Velazquez-Ramirez DA, Weaver D, Collado-Vides J, Paulsen I & Karp PD (2017). The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res* **45**, D543-D550.
280. Kanehisa M & Goto S (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27–30.
281. Kanehisa M, Furumichi M, Tanabe M, Sato Y & Morishima K (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**, D353-D361.
282. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ & Mering C von (2014). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **43**, D447-D452.
283. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
284. Altenhoff AM, Gil M, Gonnet GH & Dessimoz C (2013). Inferring hierarchical orthologous groups from orthologous gene pairs. *PLoS ONE* **8**, e53786.
285. Altenhoff AM, Skunca N, Glover N, Train C-M, Sueki A, Pilizota I, Gori K, Tomiczek B, Muller S, Redestig H, Gonnet GH & Dessimoz C (2015). The OMA orthology database in 2015: function predictions, better plant support, synteny view and other improvements. *Nucleic Acids Res* **43**, D240-9.
286. Sillitoe I, Lewis TE, Cuff A, Das S, Ashford P, Dawson NL, Furnham N, Laskowski RA, Lee D, Lees JG, Lehtinen S, Studer RA, Thornton J & Orengo CA (2015). CATH: comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Res* **43**, D376-81.

287. Lees J, Yeats C, Perkins J, Sillitoe I, Rentzsch R, Dessailly BH & Orengo C (2012). Gene3D: a domain-based resource for comparative genomics, functional annotation and protein network analysis. *Nucleic Acids Res* **40**, D465-71.
288. Clermont O, Christenson JK, Denamur E & Gordon DM (2013). The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* **5**, 58–65.
289. Li W & Godzik A (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659.
290. Lukjancenko O, Wassenaar TM & Ussery DW (2010). Comparison of 61 sequenced *Escherichia coli* genomes. *Microb Ecol* **60**, 708–720.
291. Friis C, Wassenaar TM, Javed MA, Snipen L, Lagesen K, Hallin PF, Newell DG, Toszeghy M, Ridley A, Manning G & Ussery DW (2010). Genomic characterization of *Campylobacter jejuni* strain M1. *PLoS ONE* **5**, e12253.
292. Qi F, Motz M, Jung K, Lassak J & Frishman D (2018). Evolutionary analysis of polyproline motifs in *Escherichia coli* reveals their regulatory role in translation. *PLoS computational biology* **14**, e1005987.
293. Woolstenhulme CJ, Guydosh NR, Green R & Buskirk AR (2015). High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. *Cell reports* **11**, 13–21.
294. Rice P, Longden I & Bleasby A (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276–277.
295. Nuel G. (2011). Significance Score of Motifs in Biological Sequences In: Mahdavi M A., editor. In: Bioinformatics—Trends and Methodologies. pp. 173–94. (Buch)
296. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD & Higgins DG (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539.
297. Letunic I & 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**, W242-5.
298. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W & Gascuel O (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307–321.
299. Paradis E, Claude J & Strimmer K (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**, 289–290.
300. Krylov DM, Wolf YI, Rogozin IB & Koonin EV (2003). Gene loss, protein sequence divergence, gene dispensability, expression level, and interactivity are correlated in eukaryotic evolution. *Genome Res* **13**, 2229–2235.
301. Borenstein E, Shlomi T, Ruppin E & Sharan R (2007). Gene loss rate: a probabilistic measure for the conservation of eukaryotic genes. *Nucleic Acids Res* **35**, e7.
302. Albalat R & Cañestro C (2016). Evolution by gene loss. *Nat Rev Genet* **17**, 379–391.
303. Wisniewski JR & Rakus D (2014). Quantitative analysis of the *Escherichia coli* proteome. *Data in brief* **1**, 7–11.
304. Glasner JD, Liss P, Plunkett G3, Darling A, Prasad T, Rusch M, Byrnes A, Gilson M, Biehl B, Blattner FR & Perna NT (2003). ASAP, a systematic annotation package for community analysis of genomes. *Nucleic Acids Res* **31**, 147–151.
305. Solovyev V (2011). Automatic annotation of microbial genomes and metagenomic sequences. In: Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies (Ed. R.W. Li), Nova Science Publishers, p.61-78.(Buch)
306. Kelley LA, Mezulis S, Yates CM, Wass MN & Sternberg MJE (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* **10**, 845–858.
307. Berman HM (2000). The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242.
308. Rasband WS (1997-2018) ImageJ. U. S. National Institutes of Health, Bethesda, Maryland, USA.

309. Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET & Eliceiri KW (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC bioinformatics* **18**, 529.
310. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004). UCSF Chimera - a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.
311. Woolstenhulme CJ, Parajuli S, Healey DW, Valverde DP, Petersen EN, Starosta AL, Guydosh NR, Johnson WE, Wilson DN & Buskirk AR (2013). Nascent peptides that block protein synthesis in bacteria. *Proc Natl Acad Sci U S A* **110**, E878-87.
312. Corbin RW, Paliy O, Yang F, Shabanowitz J, Platt M, Lyons CE, Root K, McAuliffe J, Jordan MI, Kustu S, Soupene E & Hunt DF (2003). Toward a protein profile of *Escherichia coli*: comparison to its transcription profile. *Proc Natl Acad Sci U S A* **100**, 9232–9237.
313. Kim SJ, Yoon JS, Shishido H, Yang Z, Rooney LA, Barral JM & Skach WR (2015). Protein folding. Translational tuning optimizes nascent protein folding in cells. *Science* **348**, 444–448.
314. Thanaraj TA & Argos P (1996). Ribosome-mediated translational pause and protein domain organization. *Protein Sci* **5**, 1594–1612.
315. O'Brien EP, Vendruscolo M & Dobson CM (2014). Kinetic modelling indicates that fast-translating codons can coordinate cotranslational protein folding by avoiding misfolded intermediates. *Nature communications* **5**, 2988.
316. Zhang G, Hubalewska M & Ignatova Z (2009). Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat Struct Mol Biol* **16**, 274–280.
317. Komar AA (2009). A pause for thought along the co-translational folding pathway. *Trends Biochem Sci* **34**, 16–24.
318. Komar AA & Jaenicke R (1995). Kinetics of translation of  $\gamma$ B crystallin and its circularly permuted variant in an in vitro cell-free system: possible relations to codon distribution and protein folding. *FEBS letters* **376**, 195–198.
319. Purvis IJ, Bettany AJE, Santiago TC, Coggins JR, Duncan K, Eason R & Brown AJP (1987). The efficiency of folding of some proteins is increased by controlled rates of translation in vivo. *J Mol Biol* **193**, 413–417.
320. Rapoport TA (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* **450**, 663–669.
321. Cymer F, Heijne G von & White SH (2015). Mechanisms of integral membrane protein insertion and folding. *J Mol Biol* **427**, 999–1022.
322. Dessen P & Képès F (2000). The PAUSE software for analysis of translational control over protein targeting: application to *E. nidulans* membrane proteins. *Gene* **244**, 89–96.
323. Bornemann T, Jöckel J, Rodnina MV & Wintermeyer W (2008). Signal sequence-independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel. *Nat Struct Mol Biol* **15**, 494–499.
324. Azeloglu EU & Iyengar R (2015). Signaling Networks: Information Flow, Computation, and Decision Making. *Cold Spring Harb Perspect Biol* **7**.
325. Steffen M, Petti A, Aach J, D'haeseleer P & Church G (2002). Automated modelling of signal transduction networks. *BMC bioinformatics* **3**, 34.
326. Sommerfeldt N, Possling A, Becker G, Pesavento C, Tschowri N & Hengge R (2009). Gene expression patterns and differential input into curli fimbriae regulation of all GGDEF/EAL domain proteins in *Escherichia coli*. *Microbiology* **155**, 1318–1331.
327. Li M, Gustchina A, Rasulova FS, Melnikov EE, Maurizi MR, Rotanova TV, Dauter Z & Wlodawer A (2010). Structure of the N-terminal fragment of *Escherichia coli* Lon protease. *Acta Crystallogr D Biol Crystallogr* **66**, 865–873.
328. Whitehouse Station, ed (2013) The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals. Royal Society of chemistry. (Buch)

329. Torres-Cabassa AS & Gottesman S (1987). Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis. *J Bacteriol* **169**, 981–989.
330. Hwang E, Cheong H-K, Kim S-Y, Kwon O, Blain KY, Choe S, Yeo KJ, Jung YW, Jeon YH & Cheong C (2017). Crystal structure of the EnvZ periplasmic domain with CHAPS. *FEBS letters* **591**, 1419–1428.
331. Motz M & Jung K (2018). The role of polyproline motifs in the histidine kinase EnvZ. *PLoS ONE* **13**, e0199782.
332. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, HERNSDORF AW, Amano Y, Ise K, Suzuki Y, Dudek N, Relman DA, Finstad KM, Amundson R, Thomas BC & Banfield JF (2016). A new view of the tree of life. *Nat Microbiol* **1**, 16048.
333. Kishii R, Falzon L, Yoshida T, Kobayashi H & Inouye M (2007). Structural and functional studies of the HAMP domain of EnvZ, an osmosensing transmembrane histidine kinase in *Escherichia coli*. *J Biol Chem* **282**, 26401–26408.
334. Parkinson JS (2010). Signaling mechanisms of HAMP domains in chemoreceptors and sensor kinases. *Annu Rev Microbiol* **64**, 101–122.
335. Sukomon N, Widom J, Borbat PP, Freed JH & Crane BR (2017). Stability and Conformation of a Chemoreceptor HAMP Domain Chimera Correlates with Signaling Properties. *Biophys J* **112**, 1383–1395.
336. Dunin-Horkawicz S & Lupas AN (2010). Comprehensive analysis of HAMP domains: implications for transmembrane signal transduction. *J Mol Biol* **397**, 1156–1174.
337. Aravind L & Ponting CP (1999). The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. *FEMS Microbiol Lett* **176**, 111–116.
338. Ladant D & Ullmann A (1999). *Bordetella pertussis* adenylate cyclase: a toxin with multiple talents. *Trends Microbiol* **7**, 172–176.
339. Mittal N, Guimaraes JC, Gross T, Schmidt A, Vina-Vilaseca A, Nedialkova DD, Aeschmann F, Leidel SA, Spang A & Zavolan M (2017). The Gcn4 transcription factor reduces protein synthesis capacity and extends yeast lifespan. *Nat Commun* **8**, 457.
340. Dammeyer T & Tinnefeld P (2012). Engineered fluorescent proteins illuminate the bacterial periplasm. *Comput Struct Biotechnol J* **3**, e201210013.
341. Dinh T & Bernhardt TG (2011). Using superfolder green fluorescent protein for periplasmic protein localization studies. *J Bacteriol* **193**, 4984–4987.
342. Bach JN, Albrecht N & Bramkamp M (2014). Imaging DivIVA dynamics using photo-convertible and activatable fluorophores in *Bacillus subtilis*. *Front Microbiol* **5**, 59.
343. Foo YH, Spahn C, Zhang H, Heilemann M & Kenney LJ (2015). Single cell super-resolution imaging of *E. coli* OmpR during environmental stress. *Integr Biol (Camb)* **7**, 1297–1308.
344. Brown CJ, Johnson AK & Daughdrill GW (2010). Comparing models of evolution for ordered and disordered proteins. *Mol Biol Evol* **27**, 609–621.
345. Huerta AM & Collado-Vides J (2003). Sigma70 promoters in *Escherichia coli*: specific transcription in dense regions of overlapping promoter-like signals. *J Mol Biol* **333**, 261–278.
346. Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, Taboada B, Jimenez-Jacinto V, Salgado H, Juárez K, Contreras-Moreira B, Huerta AM, Collado-Vides J & Morett E (2009). Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS ONE* **4**, e7526.
347. Zheng D, Constantinidou C, Hobman JL & Minchin SD (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Res* **32**, 5874–5893.
348. Aiba H (1985). Transcription of the *Escherichia coli* adenylate cyclase gene is negatively regulated by cAMP-cAMP receptor protein. *J Biol Chem* **260**, 3063–3070.

349. Jishage M, Iwata A, Ueda S & Ishihama A (1996). Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. *J Bacteriol* **178**, 5447–5451.
350. Maciag A, Peano C, Pietrelli A, Egli T, Bellis G de & Landini P (2011). In vitro transcription profiling of the  $\sigma$ S subunit of bacterial RNA polymerase: re-definition of the  $\sigma$ S regulon and identification of  $\sigma$ S-specific promoter sequence elements. *Nucleic Acids Res* **39**, 5338–5355.
351. Hidalgo E, Ding H & Demple B (1997). Redox signal transduction via iron-sulfur clusters in the SoxR transcription activator. *Trends Biochem Sci* **22**, 207–210.
352. Jones HM & Gunsalus RP (1985). Transcription of the *Escherichia coli* fumarate reductase genes (*frdABCD*) and their coordinate regulation by oxygen, nitrate, and fumarate. *J Bacteriol* **164**, 1100–1109.
353. Li J, Kustu S & Stewart V (1994). In vitro interaction of nitrate-responsive regulatory protein NarL with DNA target sequences in the *fdnG*, *narG*, *narK* and *frdA* operon control regions of *Escherichia coli* K-12. *J Mol Biol* **241**, 150–165.
354. Abo-Amer AE, Munn J, Jackson K, Aktas M, Golby P, Kelly DJ & Andrews SC (2004). DNA interaction and phosphotransfer of the C4-dicarboxylate-responsive DcuS-DcuR two-component regulatory system from *Escherichia coli*. *J Bacteriol* **186**, 1879–1889.
355. Jones HM & Gunsalus RP (1987). Regulation of *Escherichia coli* fumarate reductase (*frdABCD*) operon expression by respiratory electron acceptors and the *fir* gene product. *J Bacteriol* **169**, 3340–3349.
356. Tani TH, Khodursky A, Blumenthal RM, Brown PO & Matthews RG (2002). Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proc Natl Acad Sci U S A* **99**, 13471–13476.
357. Ernsting BR, Atkinson MR, Ninfa AJ & Matthews RG (1992). Characterization of the regulon controlled by the leucine-responsive regulatory protein in *Escherichia coli*. *J Bacteriol* **174**, 1109–1118.
358. Brinkman AB, Ettema TJG, Vos WM de & van der Oost J (2003). The Lrp family of transcriptional regulators. *Mol Microbiol* **48**, 287–294.
359. Calvo JM & Matthews RG (1994). The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol Rev* **58**, 466–490.
360. Wanner BL (1993). Gene regulation by phosphate in enteric bacteria. *J Cell Biochem* **51**, 47–54.
361. Baek JH & Lee SY (2006). Novel gene members in the Pho regulon of *Escherichia coli*. *FEMS Microbiol Lett* **264**, 104–109.
362. Crépin S, Chekabab S-M, Le Bihan G, Bertrand N, Dozois CM & Harel J (2011). The Pho regulon and the pathogenesis of *Escherichia coli*. *Vet Microbiol* **153**, 82–88.
363. Amábile-Cuevas CF & Demple B (1991). Molecular characterization of the *soxRS* genes of *Escherichia coli*: two genes control a superoxide stress regulon. *Nucleic Acids Res* **19**, 4479–4484.
364. Semchyshyn H, Bagnyukova T & Lushchak V (2005). Involvement of *soxRS* regulon in response of *Escherichia coli* to oxidative stress induced by hydrogen peroxide. *Biochemistry (Mosc)* **70**, 1238–1244.
365. Storz G, Tartaglia LA & Ames BN (1990). The OxyR regulon. *Antonie van Leeuwenhoek* **58**, 157–161.
366. Zheng M, Wang X, Templeton LJ, Smulski DR, LaRossa RA & Storz G (2001). DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *J Bacteriol* **183**, 4562–4570.
367. Seth D, Hausladen A, Wang Y-J & Stamler JS (2012). Endogenous protein S-Nitrosylation in *E. coli*: regulation by OxyR. *Science* **336**, 470–473.
368. Navarro Llorens JM, Tormo A & Martínez-García E (2010). Stationary phase in gram-negative bacteria. *FEMS Microbiol Rev* **34**, 476–495.
369. Dukan S & Nyström T (1999). Oxidative stress defense and deterioration of growth-arrested *Escherichia coli* cells. *The J Biol Chem* **274**, 26027–26032.

370. Moreno-Vivián C, Cabello P, Martínez-Luque M, Blasco R & Castillo F (1999). Prokaryotic Nitrate Reduction: Molecular Properties and Functional Distinction among Bacterial Nitrate Reductases. *J Bacteriol* **181**, 6573–6584.
371. Zumft WG (1997). Cell biology and molecular basis of denitrification. *Microbiol Mol Biol Rev* **61**, 533–616.
372. Toledano MB, Kullik I, Trinh F, Baird PT, Schneider TD & Storz G (1994). Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. *Cell* **78**, 897–909.
373. Tartaglia LA, Gimeno CJ, Storz G & Ames BN (1992). Multidegenerate DNA recognition by the OxyR transcriptional regulator. *J Biol Chem* **267**, 2038–2045.
374. Kullik I, Toledano MB, Tartaglia LA & Storz G (1995). Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for oxidation and transcriptional activation. *J Bacteriol* **177**, 1275–1284.
375. Choi H, Kim S, Mukhopadhyay P, Cho S, Woo J, Storz G & Ryu SE (2001). Structural basis of the redox switch in the OxyR transcription factor. *Cell* **105**, 103–113.
376. Zheng M, Aslund F & Storz G (1998). Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science* **279**, 1718–1721.
377. Heinrich AK, Glaeser A, Tobias NJ, Heermann R & Bode HB (2016). Heterogeneous regulation of bacterial natural product biosynthesis via a novel transcription factor. *Heliyon* **2**, e00197.
378. Plumbridge J (2001). DNA binding sites for the Mlc and NagC proteins: regulation of nagE, encoding the N-acetylglucosamine-specific transporter in *Escherichia coli*. *Nucleic Acids Res* **29**, 506–514.
379. Beck von Bodman S, Hayman GT & Farrand SK (1992). Opine catabolism and conjugal transfer of the nopaline Ti plasmid pTiC58 are coordinately regulated by a single repressor. *Proc Natl Acad Sci U S A* **89**, 643–647.
380. Campos E, Baldoma L, Aguilar J & Badia J (2004). Regulation of expression of the divergent ulaG and ulaABCDEF operons involved in LaAscorbate dissimilation in *Escherichia coli*. *J Bacteriol* **186**, 1720–1728.
381. Santiago AS, Santos CA, Mendes JS, Toledo MAS, Beloti LL, Souza AA & Souza AP (2015). Characterization of the LysR-type transcriptional regulator YcjZ-like from *Xylella fastidiosa* overexpressed in *Escherichia coli*. *Protein Expr Purif* **113**, 72–78.
382. Campos E, Aguilar J, Baldoma L & Badia J (2002). The Gene *yjfQ* encodes the repressor of the *yjfr-X* regulon (*ula*), which is involved in l-ascorbate metabolism in *Escherichia coli*. *J Bacteriol* **184**, 6065–6068.
383. Pannetier C, Domínguez-Ramírez L & Plumbridge J (2008). Different regions of Mlc and NagC, homologous transcriptional repressors controlling expression of the glucose and N-acetylglucosamine phosphotransferase systems in *Escherichia coli*, are required for inducer signal recognition. *Mol Microbiol* **67**, 364–377.
384. Shimada T, Yamazaki K & Ishihama A (2013). Novel regulator PgrR for switch control of peptidoglycan recycling in *Escherichia coli*. *Genes Cells* **18**, 123–134.
385. Kell DB, Kaprelyants AS, Weichart DH, Harwood CR & Barer MR (1998). Viability and activity in readily culturable bacteria: a review and discussion of the practical issues. *Antonie van Leeuwenhoek* **73**, 169–187.
386. Xu HS, Roberts N, Singleton FL, Attwell RW, Grimes DJ & Colwell RR (1982). Survival and viability of nonculturable *Escherichia coli* and *Vibrio cholerae* in the estuarine and marine environment. *Microb Ecol* **8**, 313–323.
387. Bergkessel M, Basta DW & Newman DK (2016). The physiology of growth arrest: uniting molecular and environmental microbiology. *Nat Rev Microbiol* **14**, 549–562.
388. Ramamurthy T, Ghosh A, Pazhani GP & Shinoda S (2014). Current Perspectives on Viable but Non-Culturable (VBNC) Pathogenic Bacteria. *Front Public Health* **2**.
389. Li L, Mendis N, Trigui H, Oliver JD & Faucher SP (2014). The importance of the viable but non-culturable state in human bacterial pathogens. *Front Microbiol* **5**, 258.

390. Nyström T & Gustavsson N (1998). Maintenance energy requirement: what is required for stasis survival of *Escherichia coli*? *Biochim Biophys Acta* **1365**, 225–231.
391. Koch AL (1997). Microbial physiology and ecology of slow growth. *Microbiol Mol Biol Rev* **61**, 305–318.
392. Babin BM, Bergkessel M, Sweredoski MJ, Moradian A, Hess S, Newman DK & Tirrell DA (2016). SutA is a bacterial transcription factor expressed during slow growth in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **113**, E597-605.
393. Khan MMT, Pyle BH & Camper AK (2010). Specific and rapid enumeration of viable but nonculturable and viable-culturable gram-negative bacteria by using flow cytometry. *Appl Environ Microbiol* **76**, 5088–5096.
394. Killick KA (1981). Purification of *Dictyostelium discoideum* spores by centrifugation in percoll density gradients with retention of morphological and biochemical integrity. *Analytical Biochemistry* **114**, 46–52.
395. Pertoft H, Laurent TC, Låås T & Kågedal L (1978). Density gradients prepared from colloidal silica particles coated by polyvinylpyrrolidone (Percoll). *Analytical Biochemistry* **88**, 271–282.
396. Fumeaux C, Radhakrishnan SK, Ardisson S, Théraulaz L, Frandi A, Martins D, Nesper J, Abel S, Jenal U & Viollier PH (2014). Cell cycle transition from S-phase to G1 in *Caulobacter* is mediated by ancestral virulence regulators. *Nature communications* **5**, 4081.
397. Vilhena C, Kaganovitch E, Grünberger A, Motz M, Forné I, Kohlheyer D & Jung K (2019). Importance of pyruvate sensing and transport for the resuscitation of viable but nonculturable *Escherichia coli* K-12. *J Bacteriol* **111**; e00610-18 .
398. Morgan AA & Rubenstein E (2013). Proline: The distribution, frequency, positioning, and common functional roles of proline and polyproline sequences in the human proteome. *PLoS ONE* **8**, e53785.
399. Ikemura T (1981). Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *J Mol Biol* **151**, 389–409.
400. dos Reis M, Savva R & Wernisch L (2004). Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res* **32**, 5036–5044.
401. Varenne S, Buc J, Lloubes R & Lazdunski C (1984). Translation is a non-uniform process. Effect of tRNA availability on the rate of elongation of nascent polypeptide chains. *J Mol Biol* **180**, 549–576.
402. Gorochoowski TE, Ignatova Z, Bovenberg RAL & Roubos JA (2015). Trade-offs between tRNA abundance and mRNA secondary structure support smoothing of translation elongation rate. *Nucleic Acids Res* **43**, 3022–3032.
403. Quax TEF, Claassens NJ, Söll D & van der Oost J (2015). Codon bias as a means to fine-tune gene expression. *Mol Cell* **59**, 149–161.
404. Tuller T & Zur H (2015). Multiple roles of the coding sequence 5' end in gene expression regulation. *Nucleic Acids Res* **43**, 13–28.
405. Braselmann E, Chaney JL & Clark PL (2013). Folding the proteome. *Trends Biochem Sci* **38**, 337–344.
406. Yu C-H, Dang Y, Zhou Z, Wu C, Zhao F, Sachs MS & Liu Y (2015). Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. *Mol Cell* **59**, 744–754.
407. Hardesty B, Tsalkova T & Kramer G (1999). Co-translational folding. *Curr Opin Struct Biol* **9**, 111–114.
408. O'Brien EP, Ciryam P, Vendruscolo M & Dobson CM (2014). Understanding the influence of codon translation rates on cotranslational protein folding. *Acc Chem Res* **47**, 1536–1544.
409. Buhr F, Jha S, Thommen M, Mittelstaet J, Kutz F, Schwalbe H, Rodnina MV & Komar AA (2016). Synonymous codons direct cotranslational folding toward different protein conformations. *Mol Cell* **61**, 341–351.
410. Tu L, Khanna P & Deutsch C (2014). Transmembrane segments form tertiary hairpins in the folding vestibule of the ribosome. *J Mol Biol* **426**, 185–198.

411. Nørholm MHH, Light S, Virkki MTI, Elofsson A, Heijne G von & Daley DO (2012). Manipulating the genetic code for membrane protein production: what have we learnt so far? *Biochim Biophys Acta* **1818**, 1091–1096.
412. Képès F (1996). The "+70 pause": hypothesis of a translational control of membrane protein assembly. *J Mol Biol* **262**, 77–86.
413. Kaiser CM, Goldman DH, Chodera JD, Tinoco I & Bustamante C (2011). The Ribosome Modulates Nascent Protein Folding. *Science* **334**, 1723–1727.
414. Rosenblum G, Chen C, Kaur J, Cui X, Zhang H, Asahara H, Chong S, Smilansky Z, Goldman YE & Cooperman BS (2013). Quantifying elongation rhythm during full-length protein synthesis. *J Am Chem Soc* **135**, 11322–11329.
415. Spencer PS, Siller E, Anderson JF & Barral JM (2012). Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. *J Mol Biol* **422**, 328–335.
416. Qi F & Frishman D (2017). Melting temperature highlights functionally important RNA structure and sequence elements in yeast mRNA coding regions. *Nucleic Acids Res* **45**, 6109–6118.
417. Mohammad F, Woolstenhulme CJ, Green R & Buskirk AR (2016). Clarifying the translational pausing landscape in bacteria by ribosome profiling. *Cell reports* **14**, 686–694.
418. Charneski CA & Hurst LD (2013). Positively charged residues are the major determinants of ribosomal velocity. *PLoS biology* **11**, e1001508.
419. Masuda T, Saito N, Tomita M & Ishihama Y (2009). Unbiased quantitation of *Escherichia coli* membrane proteome using phase transfer surfactants. *Mol Cell Proteomics* **8**, 2770–2777.
420. Quax TEF, Wolf YI, Koehorst JJ, Wurtzel O, van der Oost R, Ran W, Blombach F, Makarova KS, Brouns SJJ, Forster AC, Wagner EGH, Sorek R, Koonin EV & van der Oost J (2013). Differential translation tunes uneven production of operon-encoded proteins. *Cell reports* **4**, 938–944.
421. Claydon AJ & Beynon R (2012). Proteome Dynamics: Revisiting Turnover with a Global Perspective. *Mol Cell Proteomics* **11**, 1551–1565.
422. Li G-W, Burkhardt D, Gross C & Weissman JS (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624–635.
423. Eguchi Y, Oshima T, Mori H, Aono R, Yamamoto K, Ishihama A & Utsumi R (2003). Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*. *Microbiology* **149**, 2819–2828.
424. Alon U, Surette MG, Barkai N & Leibler S (1999). Robustness in bacterial chemotaxis. *Nature* **397**, 168–171.
425. Hart Y, Madar D, Yuan J, Bren A, Mayo AE, Rabinowitz JD & Alon U (2011). Robust control of nitrogen assimilation by a bifunctional enzyme in *E. coli*. *Mol Cell* **41**, 117–127.
426. Dassow G von, Meir E, Munro EM & Odell GM (2000). The segment polarity network is a robust developmental module. *Nature* **406**, 188–192.
427. Batchelor E & Goulian M (2003). Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. *Proc Natl Acad Sci U S A* **100**, 691–696.
428. Walker JE, Saraste M, Runswick MJ & Gay NJ (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1**, 945–951.
429. Ramakrishnan C, Dani VS & Ramasarma T (2002). A conformational analysis of Walker motif A GXXXXGKT (S) in nucleotide-binding and other proteins. *Protein Eng* **15**, 783–798.
430. Yedidi RS, Wendler P & Enenkel C (2017). AAA-ATPases in Protein Degradation. *Front Mol Biosci* **4**, 42.
431. Wendler P, Ciniawsky S, Kock M & Kube S (2012). Structure and function of the AAA+ nucleotide binding pocket. *Biochim Biophys Acta* **1823**, 2–14.

432. Saraste M, Sibbald PR & Wittinghofer A (1990). The P-loop — a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* **15**, 430–434.
433. Eguchi Y & Utsumi R (2014). Alkali metals in addition to acidic pH activate the EvgS histidine kinase sensor in *Escherichia coli*. *J Bacteriol* **196**, 3140–3149.
434. Hulko M, Berndt F, Gruber M, Linder JU, Truffault V, Schultz A, Martin J, Schultz JE, Lupas AN & Coles M (2006). The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell* **126**, 929–940.
435. Ames P, Hunter S & Parkinson JS (2016). Evidence for a Helix-Clutch Mechanism of Transmembrane Signaling in a Bacterial Chemoreceptor. *J Mol Biol* **428**, 3776–3788.
436. Zhu Y & Inouye M (2003). Analysis of the role of the EnvZ linker region in signal transduction using a chimeric Tar/EnvZ receptor protein, Tez1. *J Biol Chem* **278**, 22812–22819.
437. Hidaka Y, Park H & Inouye M (1997). Demonstration of dimer formation of the cytoplasmic domain of a transmembrane osmosensor protein, EnvZ, of *Escherichia coli* using Ni-histidine tag affinity chromatography. *FEBS letters* **400**, 238–242.
438. Heininger A, Yusuf R, Lawrence RJ & Draheim RR (2016). Identification of transmembrane helix 1 (TM1) surfaces important for EnvZ dimerisation and signal output. *Biochim Biophys Acta* **1858**, 1868–1875.
439. Ashenberg O, Rozen-Gagnon K, Laub MT & Keating AE (2011). Determinants of homodimerization specificity in histidine kinases. *J Mol Biol* **413**, 222–235.
440. Park H, Saha SK & Inouye M (1998). Two-domain reconstitution of a functional protein histidine kinase. *Proc Natl Acad Sci U S A* **95**, 6728–6732.
441. Qin L, Cai S, Zhu Y & Inouye M (2003). Cysteine-scanning analysis of the dimerization domain of EnvZ, an osmosensing histidine kinase. *J Bacteriol* **185**, 3429–3435.
442. Williams SB & Stewart V (1999). Functional similarities among two-component sensors and methyl-accepting chemotaxis proteins suggest a role for linker region amphipathic helices in transmembrane signal transduction. *Mol Microbiol* **33**, 1093–1102.
443. Unnerståle S, Måler L & Draheim RR (2011). Structural characterization of AS1-membrane interactions from a subset of HAMP domains. *Biochim Biophys Acta* **1808**, 2403–2412.
444. Williamson MP (1994). The structure and function of proline-rich regions in proteins. *Biochem J* **297 ( Pt 2)**, 249–260.
445. Noegel AA, Gerisch G, Lottspeich F & Schleicher M (1990). A protein with homology to the C-terminal repeat sequence of Octopus rhodopsin and synaptophysin is a member of a multigene family in *Dictyostelium discoideum*. *FEBS letters* **266**, 118–122.
446. Battesti A, Majdalani N & Gottesman S (2011). The RpoS-mediated general stress response in *Escherichia coli*. *Annu Rev Microbiol* **65**, 189–213.
447. Lange R & Hengge-Aronis R (1994). The cellular concentration of the sigma S subunit of RNA polymerase in *Escherichia coli* is controlled at the levels of transcription, translation, and protein stability. *Genes & development* **8**, 1600–1612.
448. Weber H, Polen T, Heuveling J, Wendisch VF & Hengge R (2005). Genome-wide analysis of the general stress response network in *Escherichia coli*: sigmaS-dependent genes, promoters, and sigma factor selectivity. *J Bacteriol* **187**, 1591–1603.
449. Merrikh H, Ferrazzoli AE, Bougdour A, Olivier-Mason A & Lovett ST (2009). A DNA damage response in *Escherichia coli* involving the alternative sigma factor, RpoS. *Proc Natl Acad Sci U S A* **106**, 611–616.
450. Mandel MJ & Silhavy TJ (2005). Starvation for different nutrients in *Escherichia coli* results in differential modulation of RpoS levels and stability. *J Bacteriol* **187**, 434–442.
451. Sledjeski DD, Gupta A & Gottesman S (1996). The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *The EMBO journal* **15**, 3993–4000.

452. Muffler A, Barth M, Marschall C & Hengge-Aronis R (1997). Heat shock regulation of sigmaS turnover: a role for DnaK and relationship between stress responses mediated by sigmaS and sigma32 in *Escherichia coli*. *J Bacteriol* **179**, 445–452.
453. Waterman SR & Small PL (1996). Identification of sigma S-dependent genes associated with the stationary-phase acid-resistance phenotype of *Shigella flexneri*. *Mol Microbiol* **21**, 925–940.
454. Nair S & Finkel SE (2004). Dps protects cells against multiple stresses during stationary phase. *J Bacteriol* **186**, 4192–4198.
455. Muffler A, Traulsen DD, Lange R & Hengge-Aronis R (1996). Posttranscriptional osmotic regulation of the sigma(s) subunit of RNA polymerase in *Escherichia coli*. *J Bacteriol* **178**, 1607–1613.
456. Sak BD, Eisenstark A & Touati D (1989). Exonuclease III and the catalase hydroperoxidase II in *Escherichia coli* are both regulated by the *katF* gene product. *Proc Natl Acad Sci U S A* **86**, 3271–3275.
457. Traxler MF, Summers SM, Nguyen H-T, Zacharia VM, Hightower GA, Smith JT & Conway T (2008). The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol Microbiol* **68**, 1128–1148.
458. Kim J-S, Chowdhury N, Yamasaki R & Wood TK (2018). Viable but non-culturable and persistence describe the same bacterial stress state. *Environ Microbio*.
459. Mansilla MC, Cybulski LE, Albanesi D & Mendoza D de (2004). Control of membrane lipid fluidity by molecular thermosensors. *J Bacteriol* **186**, 6681–6688.
460. Morishige Y, Fujimori K & Amano F (2013). Differential resuscitative effect of pyruvate and its analogues on VBNC (viable but non-culturable) *Salmonella*. *Microbes Environ* **28**, 180–186.
461. Nam D, Choi E, Shin D & Lee E-J (2016). tRNA<sup>Pro</sup>-mediated downregulation of elongation factor P is required for *mgtCBR* expression during *Salmonella* infection. *Mol Microbiol* **102**, 221–232.

## 7. ANHANG

S. 178

**Tabelle S1:** Übersicht aller polyP Motive in *E. coli* K-12 MG1655.

S. 192

**Tabelle S2:** 381 nichtredundante Orthologe aus der Uni Prot Datenbank „Microbial Proteomes“ mit mindestens 44% Übereinstimmung und höchstens 10% Sequenzlängenabweichung zu DgcC aus *E. coli* K-12.

S. 195

**Tabelle S3:** 203 nichtredundante Orthologe aus der Uni Prot Datenbank „Reference Proteomes“ mit mindestens 44% Übereinstimmung und höchstens 10% Sequenzlängenabweichung zu Lon aus *E. coli* K-12.

S. 197

**Tabelle S4:** 183 nichtredundante Orthologe aus der Uni Prot Datenbank „Microbial Proteomes“ mit mindestens 44% Übereinstimmung und höchstens 10% Sequenzlängenabweichung zu EvgS aus *E. coli* K-12.

S. 199

**Tabelle S5:** 793 nichtredundante Orthologe aus der Uni Prot Datenbank „Microbial Proteomes“ mit mindestens 44% Übereinstimmung und höchstens 10% Sequenzlängenabweichung zu EnvZ aus *E. coli* K-12.

Tabelle S1: Übersicht aller polyP Motive in *E. coli* K-12 MG1655.

Anzahl Motive				Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive			
S	M	W	Σ			Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4
2	2	3	7	spP46889	FTSK_ECOLI	DNA translocase FtsK	APPPF	TPPS	-	DVPPA	PVPPQ	ARPL	AVPPV	QTPPV	-
2	0	3	5	spP19319	NARZ_ECOLI	Respiratory nitrate reductase 2 alpha chain	RVPPG	QIPPV	-	-	-	DLPPA	NRPPR	YRPPI	-
2	0	3	5	spP37624	RBBA_ECOLI	Ribosome-associated ATPase	EIPPV	EAPPV	-	-	-	PVPPV	VIPPY	EQPLV	-
2	0	3	5	spP76655	YQIG_ECOLI	Putative outer membrane usher protein YqjG	NVPPG	IQPPV	-	-	-	MLPPD	SHPLV	WTPEE	-
0	1	4	5	spP09831	GLTB_ECOLI	Glutamate synthase [NADPH] large chain	-	-	-	LVPPA	-	IRPPV	NNPFF	SPPPH	TNPPI
3	0	1	4	spP37650	BCSC_ECOLI	Cellulose synthase operon protein C	GAPPE	QPPPS	SQPPS	-	-	DLPPQ	-	-	-
3	1	0	4	spP00936	CYAA_ECOLI	Adenylyate cyclase	YVPPG	AAPPD	QDPPK	NMPPA	-	-	-	-	-
3	1	0	4	spP09549	DEDD_ECOLI	Cell division protein DedD	VAPPK	EAPPA	TQPPE	VEPPK	-	-	-	-	
2	1	1	4	spP11454	ENTF_ECOLI	Enterobactin synthase component F	PVPPG	TWPEE	-	LPPPA	-	-	TLPPH	-	-
2	1	1	4	spP30855	EVGS_ECOLI	Sensor protein EvgS	AIPPS	NYPPD	-	YSPPY	-	-	TSPLL	-	-
2	0	2	4	spP30130	FIMD_ECOLI	Outer membrane usher protein FimD	YIPPE	TVPPG	-	-	-	-	ELPPG	QLPPE	-
2	0	2	4	spP52647	NIFJ_ECOLI	Probable pyruvate-flavodoxin oxidoreductase	VVPEE	SRPPS	-	-	-	-	ALPPD	NRPPV	-
2	1	1	4	spP27431	ROXA_ECOLI	50S ribosomal protein L16 3-hydroxylase	IAPPE	YIPPP	-	PEPPY	-	-	DVPPR	-	-
2	1	1	4	spP76393	YEGJ_ECOLI	Uncharacterized protein YegJ	VAPPA	KPPPR	-	TPPPA	-	-	FTPPE	-	-
2	0	2	4	spP76616	YGAA_ECOLI	Putative uncharacterized protein YgaQ	IPPPG	GEPPN	-	-	-	-	ITPLL	QIPPV	-
1	1	2	4	spP31553	CAIT_ECOLI	L-carnitine/gamma-butyrobetaine antiporter	NEPPE	-	-	EEPLL	-	-	FFPLL	STPFF	-
1	0	3	4	spP21189	DPO2_ECOLI	DNA polymerase II	VRPPE	-	-	-	-	-	EVPPH	NVPPH	YRPPL
1	2	1	4	spP30015	LHR_ECOLI	Probable ATP-dependent helicase Ihr	TIPPP	-	-	VNPPA	RNPPD	-	LTPLH	-	-
1	2	1	4	spP0QD69	PHNEI_ECOLI	1 Putative cryptic phosphonate transport system permease protein PhnE1	IAPPK	-	-	FAPPV	FFPPD	-	STPPA	-	-
1	2	1	4	spP39321	TAMB_ECOLI	Translocation and assembly module subunit TamB	SGPPD	-	-	EIPPA	KMPPS	-	TVPPM	-	-
1	0	3	4	spP77196	YFUC_ECOLI	Putative outer membrane usher protein YfucU	WDPPS	-	-	-	-	-	MLPPN	SHPPV	TDPPV
1	2	1	4	spP76223	YNIJ_ECOLI	Protein YniJ	RPPPD	-	-	-	-	-	DPPPS	PQPPQ	-
0	1	3	4	spP30131	HYPF_ECOLI	Carbamoyltransferase HypF	-	-	-	GKPPA	-	-	ALPPG	HCPLL	NIPPV
0	0	4	4	spP0A0F2	PARC_ECOLI	DNA topoisomerase 4 subunit A	-	-	-	-	-	-	VLPPQ	TLPPG	DIPPH
0	1	3	4	spP76577	PBPC_ECOLI	Penicillin-binding protein 1C	-	-	-	SQPPV	-	-	RLPPA	RLPPR	SCPYP
0	0	4	4	spP0AAA1	YAGU_ECOLI	Inner membrane protein YagU	-	-	-	-	-	-	FLPPR	LNPPY	LTPPL
0	2	2	4	spP77504	YBPA_ECOLI	Uncharacterized ABC transporter permease YhbP	-	-	-	NWPPK	TNPPG	-	ALPPA	QLPPE	-
0	1	3	4	spP75666	YDBA_ECOLI	Putative exported protein YdbA	-	-	-	PVPPV	-	-	FLPPT	FTPPD	FTPPK
0	1	3	4	spP64426	YDDW_ECOLI	UPF0748 lipoprotein YdW	-	-	-	SKPPA	-	-	STPPE	VTPPA	DWPPV
0	2	2	4	spP33346	YEHJ_ECOLI	Uncharacterized protein YehJ	-	-	-	SAPPK	TEPPA	-	IMPPF	VSPPW	-
3	0	0	3	spP18956	GET_ECOLI	Glutathione hydrolase proenzyme	MPPPS	PAPPV	AAPPA	-	-	-	-	-	-
3	0	0	3	spP33341	YEHB_ECOLI	Outer membrane usher protein YehB	EVPPG	YVPEE	EVPPS	-	-	-	-	-	-
2	1	0	3	spP36683	ACNB_ECOLI	Aconitate hydratase B	RVPPG	VAPPT	-	KNPPA	-	-	-	-	-
2	0	1	3	spP37653	BESA_ECOLI	Cellulose synthase catalytic subunit [UDP-forming]	IAPPT	YGPPT	-	-	-	-	LIPPV	-	-
2	1	0	3	spP0A0E82	CPXA_ECOLI	Sensor histidine kinase CpxA	WAPPG	NDPPN	-	NFPPG	-	-	-	-	-
2	1	0	3	spP54901	CSIE_ECOLI	Stationary phase-inducible protein CsiE	PVPPD	LAPPS	-	FSPLL	-	-	-	-	-
2	0	1	3	spP76541	EULI_ECOLI	Ethanolamine utilization protein EulI	VAPPL	VPPPS	-	-	-	-	KLPPH	-	-
2	0	1	3	spP15286	FLK_ECOLI	Flagellar regulator Flk	GPPPG	QPPPG	-	-	-	-	FLPPM	-	-
2	0	1	3	spP0A0A13	FTSH_ECOLI	ATP-dependent zinc metalloprotease FtsH	GEPEE	VGPBG	-	VRPPA	-	-	-	-	-
2	0	1	3	spP77650	HCAJ_ECOLI	3-phenylpropionate dioxygenase ferredoxin-NAD(+)-reductase component	APPPV	LPPPV	-	-	-	-	ERPLL	-	-
2	0	1	3	spP0A0B7	HFLI_ECOLI	Modulator of FtsH protease HflK	QCPDD	ARPPE	-	-	-	-	RLPPA	-	-
2	0	1	3	spP43377	NUDL_ECOLI	Uncharacterized Nudix hydrolase NudL	AIPPS	IPPPD	-	-	-	-	VLPPV	-	-
2	1	0	3	spP02918	PBPA_ECOLI	Penicillin-binding protein 1A	TPPPG	QIPPE	-	NSPPQ	-	-	-	-	-
2	0	1	3	spP75864	RLMKL_ECOLI	18S ribosomal RNA large subunit methyltransferase K.L	YAPPK	IDPPT	-	-	-	-	SNPPY	-	-
2	0	1	3	spP52043	SCPC_ECOLI	Propionyl-CoA: succinate CoA transferase	RIPPE	GAPPR	-	-	-	-	VIPPP	-	-
2	0	1	3	spP13482	TREA_ECOLI	Periplasmic chaperone	YVPEE	PQPPD	-	-	-	-	SOPPP	-	-
2	0	1	3	spP75780	YBQJ_ECOLI	Uncharacterized outer membrane usher protein YbgQ	KVPPG	WAPPS	-	-	-	-	QFPLL	-	-
2	0	1	3	spP31126	YDEI_ECOLI	Uncharacterized MFS-type transporter YdeE	IGPLL	IAPPE	-	-	-	-	SLPPS	-	-
2	1	0	3	spP22787	YIFB_ECOLI	Uncharacterized protein YifB	LPPPG	IGPPG	-	NAPPI	-	-	-	-	-
1	0	2	3	spP31224	ACRB_ECOLI	Multidrug efflux pump subunit AcrB	IAPPA	-	-	-	-	-	GLPPK	GTPPV	-
1	0	2	3	spP37652	BCSB_ECOLI	Cyclic di-GMP-binding protein	APPPG	-	-	-	-	-	NLPPD	TMPPV	-
1	0	2	3	spP0A0D99	BRNQ_ECOLI	Branched-chain amino acid transport system 2 carrier protein	IAPPM	-	-	-	-	-	IFPPM	IYPPC	-
1	0	2	3	spP21179	CATE_ECOLI	Catalase HPH	TPPPD	-	-	-	-	-	ETPPG	PTPPG	-
1	1	1	3	spP52095	DCLZ_ECOLI	Constitutive lysine decarboxylase	PYPPG	-	-	WQPPQ	-	-	ITPPF	-	-
1	0	2	3	spP21169	DCOR_ECOLI	Constitutive ornithine decarboxylase	PYPPG	-	-	-	-	-	LLPPF	FIPPV	-
1	0	2	3	spP00882	DPO1_ECOLI	DNA polymerase I	GVPEE	-	-	-	-	-	AFPLL	HRPPM	-
1	1	1	3	spP10443	DPO3A_ECOLI	DNA polymerase III subunit alpha	LIPPD	-	-	GEPLL	-	-	ILPPD	-	-
1	2	0	3	spP06710	DPO3X_ECOLI	DNA polymerase III subunit tau	EEPEE	-	-	QVPPQ	TIPPT	-	-	-	-
1	0	2	3	spP0C839	EPTC_ECOLI	Phosphoethanolamine transferase EptC	HVPPG	-	-	-	-	-	ALPLL	DTPPH	-
1	0	2	3	spP38134	ETK_ECOLI	Tyrosine-protein kinase etk	QVPPN	-	-	-	-	-	NTPPG	DTPPM	-
1	1	1	3	spP52612	FLII_ECOLI	Flagellin-specific ATP synthase	GYPSS	-	-	GEPPA	-	-	ITPPF	-	-
1	0	2	3	spP15067	GLGX_ECOLI	Glycogen debranching enzyme	DAPPR	-	-	-	-	-	AIPPF	NFPLL	-
1	1	1	3	spP21517	MALZ_ECOLI	Maltodextrin glucosylase	EDPPG	-	-	RMPPA	-	-	PVPPF	-	-
1	0	2	3	spP76398	MDTB_ECOLI	Multidrug resistance protein MdtB	ADPPI	-	-	-	-	-	VLPPS	PNPPV	-
1	1	1	3	spP76080	PAAD_ECOLI	Putative 1,2-phenylacetyl-CoA epoxidase, subunit D	IAPPG	-	-	ISPPA	-	-	HLPPE	-	-
1	1	1	3	spP69434	PGAA_ECOLI	Poly-beta-1,6-N-acetyl-D-glucosamine export protein	IPPPV	-	-	YSPLL	-	-	TSPPF	-	-
1	1	1	3	spP16687	PHNI_ECOLI	Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnI	IDPPG	-	-	GEPPP	-	-	RTPPV	-	-
1	0	2	3	spP24230	RECG_ECOLI	ATP-dependent DNA helicase RecG	RPPPT	-	-	-	-	-	ELPPG	LLPPE	-
1	1	1	3	spP37634	RLMJ_ECOLI	Ribosomal RNA large subunit methyltransferase J	IDPPY	-	-	INPPW	-	-	KLPPV	-	-
1	0	2	3	spP0A0812	RUVB_ECOLI	Holliday junction ATP-dependent DNA helicase RuvB	FGPPD	-	-	-	-	-	ITPPE	DLPFF	-
1	1	1	3	spP58041	RZOD_ECOLI	Prophage outer membrane lipoprotein RzoD	QPPPD	-	-	RPPPA	-	-	VKPPR	-	-
1	0	2	3	spP33018	SFGH2_ECOLI	1S-formylglutathione hydrolase YegG	TQPPW	-	-	-	-	-	LLPPR	TTPPV	-
1	1	1	3	spP37025	THPR_ECOLI	RNA 2',3'-cyclic phosphodiesterase	IPPPG	-	-	HFPPE	-	-	RQPPR	-	-
1	2	0	3	spP02929	TONB_ECOLI	Protein TonB	QPPPE	-	-	LEPPQ	PEPPK	-	-	-	-
1	1	1	3	spP06612	TOPI_ECOLI	DNA topoisomerase I	VAPPK	-	-	TKPPA	-	-	ALPPK	-	-
1	1	1	3	spP0A0AB4	UBID_ECOLI	3-oxopentyl-4-hydroxybenzoate carboxy-lyase	GRPPD	-	-	PEPPK	-	-	YLPEE	-	-
1	1	1	3	spP76387	WZC_ECOLI	Tyrosine-protein kinase wzc	QVPPN	-	-	NKPPA	-	-	DTPPI	-	-
1	0	2	3	spP77221	YAHG_ECOLI	Uncharacterized protein YahG	AGPPI	-	-	-	-	-	WTPPG	VHPPF	-
1	1	1	3	spP75822	YBTT_ECOLI	Putative NAD(P)-binding protein YbtT	QAPPH	-	-	SVPPT	-	-	LTPPR	-	-
1	0	2	3	spP76065	YDAU_ECOLI	Uncharacterized protein YdaU	AIPPG	-	-	-	-	-	LNPPR	TDPLL	-
1	0	2	3	spP52645	YDBH_ECOLI	Uncharacterized protein YdbH	LVPPD	-	-	-	-	-	VLPPS	NQPPV	-
1	0	2	3	spP76186	YDHK_ECOLI	Uncharacterized transporter YdhK	QAPPG	-	-	-	-	-	BSPGG	TNPPV	-
1	0	2	3	spP33342	YEHJ_ECOLI	Probable fibrillar chaperone YehC	DIPPN	-	-	-	-	-	SLPPE	LTPPV	-
1	0	2	3	spP77656	YDFK_ECOLI	Uncharacterized protein YdfK	IPPPV	-	-	-	-	-	GLPPK	PTPPA	-
1	0	2	3	spQ46811	YGFK_ECOLI	Putative oxidoreductase YgfK	NVPPPE	-	-	-	-	-	GCPPH	KQPPM	-
1	2	0	3	spP39407	YJII_ECOLI	Uncharacterized protein YjiI	EKPPG	-	-	IYPPK	VIPPA	-	-	-	-
1	1	1	3	spP09154	YMFS_ECOLI	Uncharacterized protein YmfS	VIPPD	-	-	VFPPS	-	-	PTPPA	-	-

Anzahl Motive					Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive			
S	M	W	Σ	Motiv 1			Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4	
0	1	2	3	spp069741	MBHT_ECOLI	Hydrogenase-2 small chain	-	-	-	GKPK	-	-	GCPCN	QRPPV	-	-
0	1	2	3	spp060075	MEPA_ECOLI	Penicillin-insensitive murein endopeptidase	-	-	-	FEPPK	-	-	PLPSP	TPPPL	-	-
0	1	2	3	spp130099	METH_ECOLI	Methionine synthase	-	-	-	GFPEE	-	-	RTPPV	YTPPV	-	-
0	0	3	3	spp30958	MFD_ECOLI	Transcription-repair-coupling factor	-	-	-	-	-	-	PLPL	ATPPA	LLPQ	-
0	3	0	3	spp0A065	MRDA_ECOLI	Peptidoglycan D,D-transpeptidase MrdA	-	-	-	EVPPQ	HEPPV	VYPPA	-	-	-	-
0	2	1	3	spp76342	MSRP_ECOLI	Protein-methionine-sulfoxide reductase catalytic subunit MsrP	-	-	-	DRPPA	ERPPT	-	ALPPQ	-	-	-
0	0	3	3	spp07648	RECC_ECOLI	RecBCD enzyme subunit RecC	-	-	-	-	-	-	ALPPV	TCPPG	RFPPL	-
0	0	3	3	spp16919	RHSD_ECOLI	Protein RhsD	-	-	-	-	-	-	PLPPY	ALPPD	SLPPV	-
0	0	3	3	spp75782	RLMF_ECOLI	Ribosomal RNA large subunit methyltransferase F	-	-	-	-	-	-	NLPL	CNPPF	LCPPV	-
0	0	3	3	spp36929	RSMB_ECOLI	Ribosomal RNA small subunit methyltransferase B	-	-	-	-	-	-	ILPL	QRPPM	RIPPH	-
0	2	1	3	spp0A0676	SBCD_ECOLI	Nuclease SbcD subunit D	-	-	-	NFPPA	QEPVV	-	GSPPS	-	-	-
0	2	1	3	spp27253	SCPA_ECOLI	Methylmalonyl-CoA mutase	-	-	-	IYPPK	VIPPQ	-	GLPPY	-	-	-
0	2	1	3	spp21507	SRMB_ECOLI	ATP-dependent RNA helicase SrmB	-	-	-	AIPPA	GVPPQ	-	SGPPR	-	-	-
0	2	1	3	spp75853	SSUA_ECOLI	Putative aliphatic sulfonates-binding protein	-	-	-	VEPPK	HRPPT	-	DIPPI	-	-	-
0	1	2	3	spp37610	TAUD_ECOLI	Alpha-ketoglutarate-dependent taurine dioxygenase	-	-	-	DNPPD	-	-	ETPPA	KNPPL	-	-
0	1	2	3	spp33225	TORA_ECOLI	Trimethylamine-N-oxide reductase 1	-	-	-	VVPPQ	-	-	KLPPV	SIPPV	-	-
0	1	2	3	spp23841	XAPR_ECOLI	HTH-type transcriptional regulator XapR	-	-	-	TEPPT	-	-	SQPL	TMPVV	-	-
0	0	3	3	spp0A0C13	XYLR_ECOLI	Xylose operon regulatory protein	-	-	-	-	-	-	LVPPV	SYPPV	TLPPQ	-
0	1	2	3	spp36682	YACH_ECOLI	Uncharacterized protein YacH	-	-	-	LPPPA	-	-	AYPPV	SAPPV	-	-
0	1	2	3	spp77328	YBBY_ECOLI	Putative purine permease YbbY	-	-	-	VVPPY	-	-	ALPPV	QLPPF	-	-
0	1	2	3	spp77779	YBFO_ECOLI	Putative uncharacterized protein YbfO	-	-	-	NKPPY	-	-	QLPPG	VFPVV	-	-
0	1	2	3	spp75783	YBIO_ECOLI	Moderate conductance mechanosensitive channel YbiO	-	-	-	IVPPT	-	-	ATPPA	PLPPG	-	-
0	2	1	3	spp22525	YCBB_ECOLI	Probable L,D-transpeptidase YcbB	-	-	-	NVPPY	VNPPW	-	ATPPL	-	-	-
0	2	1	3	spp76176	YDGD_ECOLI	Uncharacterized serine protease YdgD	-	-	-	IVPPA	RNPPS	-	LTPPK	-	-	-
0	0	3	3	spp33349	YEHM_ECOLI	Uncharacterized protein YehM	-	-	-	-	-	-	HCPQQ	PQPL	WLPPR	-
0	1	2	3	spp77775	YFCH_ECOLI	Epimerase family protein YfcH	-	-	-	EEPPH	-	-	DTPPS	MLPPF	-	-
0	2	1	3	spp46892	YGBN_ECOLI	Inner membrane permease YgbN	-	-	-	INPPG	LMPPS	-	AVPPH	-	-	-
0	1	2	3	spp46798	YGER_ECOLI	Uncharacterized lipoprotein YgeR	-	-	-	ISPPY	-	-	SWPPV	YLPPQ	-	-
0	0	3	3	spp52062	YGGW_ECOLI	Oxygen-independent coproporphyrinogen-III oxidase-like protein YggW	-	-	-	-	-	-	KLPL	LNPPH	SRPPV	-
0	0	3	3	spp45537	YHFK_ECOLI	Uncharacterized protein YhfK	-	-	-	-	-	-	ALPL	ALPPE	MFPMM	-
2	0	0	2	spp58042	RZOR_ECOLI	Prophage outer membrane lipoprotein RzoR	PPPPP	QPPPD	-	-	-	-	-	-	-	-
2	0	0	2	spp76518	ACOCT_ECOL	1 Acetyl-CoA:oxalate CoA-transferase	VEPPG	DAPPV	-	-	-	-	-	-	-	-
2	0	0	2	spp31552	CAIC_ECOLI	Crotonobetaine/carnitine-CoA ligase	YAPPL	VQPPS	-	-	-	-	-	-	-	-
2	0	0	2	spp00864	CAPP_ECOLI	Phosphoenolpyruvate carboxylase	LPPPE	SQPPG	-	-	-	-	-	-	-	-
2	0	0	2	spp38038	CYSI_ECOLI	Sulfite reductase [NADPH] flavoprotein alpha-component	GEPPY	QVPPS	-	-	-	-	-	-	-	-
2	0	0	2	spp29680	DCUP_ECOLI	Uroporphyrinogen decarboxylase	YAPPA	DVPPY	-	-	-	-	-	-	-	-
2	0	0	2	spp45523	FKBA_ECOLI	FKBP type peptidyl-prolyl cis-trans isomerase FkpA	VIPPE	GPPN	-	-	-	-	-	-	-	-
2	0	0	2	spp46846	GNTX_ECOLI	DNA utilization protein YghH	YAPPL	KPPPW	-	-	-	-	-	-	-	-
2	0	0	2	spp0A0AK1	HYCB_ECOLI	Formate hydrogenase subunit 2	PAPPA	KAPPA	-	-	-	-	-	-	-	-
2	0	0	2	spp0A09M0	LON_ECOLI	Lon protease	VGPPG	KIPPE	-	-	-	-	-	-	-	-
2	0	0	2	spp41052	MLTB_ECOLI	Membrane-bound lytic murein transglycosylase B	GVPPY	VKPPS	-	-	-	-	-	-	-	-
2	0	0	2	spp17802	MUTY_ECOLI	Adenine DNA glycosylase	AQPPS	QRPPS	-	-	-	-	-	-	-	-
2	0	0	2	spp0A0E4	TAMA_ECOLI	Translocation and assembly module subunit TamA	KVPPD	RPPPK	-	-	-	-	-	-	-	-
2	0	0	2	spp07913	TDH_ECOLI	L-threonine 3-dehydrogenase	GAPPA	GIPPS	-	-	-	-	-	-	-	-
2	0	0	2	spp76192	YDHV_ECOLI	Uncharacterized oxidoreductase YdhV	EVPPG	EIPPG	-	-	-	-	-	-	-	-
1	1	0	2	spp25516	ACNA_ECOLI	Aconitate hydratase A	VVPPG	-	ASPPL	-	-	-	-	-	-	-
1	0	1	2	spp24177	ACRD_ECOLI	Probable aminoglycoside efflux pump	LAPPA	-	-	-	-	-	-	SSPPA	-	-
1	0	1	2	spp24181	ACRF_ECOLI	Multidrug export protein AcrF	IAPPA	-	-	-	-	-	-	KLPPK	-	-
1	0	1	2	spp31440	ADEQ_ECOLI	Adenine permease AdeQ	SIPPD	-	-	-	-	-	-	IVPPY	-	-
1	0	1	2	spp28629	ADIA_ECOLI	Biolegradative arginine decarboxylase	PYPYG	-	-	-	-	-	-	LLPPL	-	-
1	0	1	2	spp25550	ASLB_ECOLI	Anaerobic sulfatase-maturing enzyme homolog AslB	DGPPY	-	-	-	-	-	-	HLPPY	-	-
1	1	0	2	spp0A0B8	ATMA_ECOLI	Magnesium-transporting ATPase, P-type 1	VVPPS	-	LDPPK	-	-	-	-	-	-	-
1	0	1	2	spp30847	BAES_ECOLI	Signal transduction histidine-protein kinase BaeS	PIPPD	-	-	-	-	-	-	GMPPH	-	-
1	0	1	2	spp0A0C9	BETT_ECOLI	High-affinity choline transport protein	MQPPY	-	-	-	-	-	-	SLPPE	-	-
1	1	0	2	spp46896	CASI_ECOLI	CRISPR-associated endonuclease Cas1	PAPPE	-	IQPPA	-	-	-	-	-	-	-
1	0	1	2	spp38036	CAS3_ECOLI	CRISPR-associated endonuclease/helicase Cas3	DIPPD	-	-	-	-	-	-	QLPPR	-	-
1	1	0	2	spp36659	CBPA_ECOLI	Curved DNA-binding protein	TIPPG	-	VMPPK	-	-	-	-	-	-	-
1	0	1	2	spp76175	CLCB_ECOLI	Voltage-gated Cl <sup>-</sup> type chloride channel CkC	LAPPW	-	-	-	-	-	-	TAPPL	-	-
1	1	0	2	spp0A0E2	CSGG_ECOLI	Curli production assembly/transport component CsgG	SVPPY	-	TAPPK	-	-	-	-	-	-	-
1	1	0	2	spp36649	CUEO_ECOLI	Blue copper oxidase CueO	IIPPG	-	GKPPA	-	-	-	-	-	-	-
1	0	1	2	spp38054	CUSA_ECOLI	Cation efflux system protein CusA	KIPPE	-	-	-	-	-	-	WVPPY	-	-
1	0	1	2	spp22255	CYSQ_ECOLI	3(2),5-bisphosphate nucleotidase CysQ	EDPPG	-	-	-	-	-	-	ARPLL	-	-
1	0	1	2	spp0A0E7	DAPE_ECOLI	Succinyl-diaminopelate desuccinylase	VVPPG	-	-	-	-	-	-	INPPF	-	-
1	0	1	2	spp24169	DCOS_ECOLI	Inducible ornithine decarboxylase	PYPYG	-	-	-	-	-	-	LRPPF	-	-
1	1	0	2	spp0A0A1	DGCC_ECOLI	Probable diguanylate cyclase DgcC	HPPPY	-	EEPLL	-	-	-	-	-	-	-
1	0	1	2	spp37349	DHAM_ECOLI	Protein-lysine deacetylase	VAPPP	-	-	-	-	-	-	PVPPV	-	-
1	0	1	2	spp0A0B5	DNAG_ECOLI	DNA primase	FAPPG	-	-	-	-	-	-	LVPLL	-	-
1	0	1	2	spp76129	DOSP_ECOLI	Oxygen sensor protein DosP	HVPPS	-	-	-	-	-	-	SSPPF	-	-
1	0	1	2	spp75858	ELFG_ECOLI	Uncharacterized fibrial-like protein ElfG	GVPPY	-	-	-	-	-	-	VIPPM	-	-
1	0	1	2	spp52600	EMRY_ECOLI	Probable multidrug resistance protein EmrY	NYPPE	-	-	-	-	-	-	AKPPF	-	-
1	0	1	2	spp24077	ENTS_ECOLI	Enterobactin exporter EntS	LPPPP	-	-	-	-	-	-	QTPPQ	-	-
1	1	0	2	spp0A0E4	ENVZ_ECOLI	Osmolarity sensor protein EnvZ	IPPLL	-	VVPPA	-	-	-	-	-	-	-
1	0	1	2	spp39176	ERFK_ECOLI	Probable L,D-transpeptidase ErfK/SrfK	YPPPD	-	-	-	-	-	-	PLPPE	-	-
1	1	0	2	spp0A0W3	ETTA_ECOLI	Energy-dependent translational throttle protein EtnA	FIPPG	-	VVPPK	-	-	-	-	-	-	-
1	0	1	2	spp0A0A7	EXUT_ECOLI	Hexuronate transporter	IAPPL	-	-	-	-	-	-	YLPLL	-	-
1	0	1	2	spp77399	FADJ_ECOLI	Fatty acid oxidation complex subunit alpha	DAPPA	-	-	-	-	-	-	GPPPF	-	-
1	0	1	2	spp07821	FHUC_ECOLI	Iron(3+)-hydroxamate import ATP-binding protein Fhuc	HQPPS	-	-	-	-	-	-	QLPPA	-	-
1	0	1	2	spp76298	FLHA_ECOLI	Flagellar biosynthesis protein FlhA	GAPPV	-	-	-	-	-	-	FLPPV	-	-
1	1	0	2	spp0A0Y7	FLHC_ECOLI	Flagellar transcriptional regulator FlhC	CQPPS	-	SPPPK	-	-	-	-	-	-	-
1	1	0	2	spp0A0C5	FLIP_ECOLI	Flagellar biosynthetic protein FlhP	SAPPN	-	MVPPA	-	-	-	-	-	-	-
1	0	1	2	spp06136	FTSQ_ECOLI	Cell division protein FtsQ	SVPPY	-	-	-	-	-	-	PLPPE	-	-
1	0	1	2	spp39829	GARD_ECOLI	Galactarate dehydratase (L-threo-forming)	EAPPL	-	-	-	-	-	-	PLPLL	-	-
1	0	1	2	spp75883	GFCC_ECOLI	Uncharacterized protein GfcC	EPPPG	-	-	-	-	-	-	SNPLL	-	-
1	0	1	2	spp27305	GLUQ_ECOLI	Glutamyl-Q tRNA(Asp) synthetase	IDPPY	-	-	-	-	-	-	MLPPY	-	-
1	0	1	2	spp25552	GPPA_ECOLI	Guanosine-5'-triphosphate,3'-diphosphate pyrophosphatase	DIPPS	-	-	-	-	-	-	AVPPR	-	-
1	1	0	2	spp75797	GSIB_ECOLI	Glutathione-binding protein GsiB	VVPPS	-	NWPPY	-	-	-	-	-	-	-
1	1	0	2	spp45763	GSPL_ECOLI	Putative type II secretion system protein L	LIPPE	-	QFPEE	-	-	-	-	-	-	-
1	0	1	2	spp37651	GUN_ECOLI	Endoglycanase	GYPPE	-	-	-	-	-	-	YLPPY	-	-
1	1	0	2	spp47140	HCAF_ECOLI	3-phenylpropionate/cinnamic acid dioxygenase subunit beta	EEPPS	-	VQPPY	-	-	-	-	-	-	-
1	0	1	2	spp45753	HOFM_ECOLI	DNA utilization protein HofM	LPPPG	-	-	-	-	-	-	RQPLL	-	-
1	1	0	2	spp45750	HOPF_ECOLI	DNA utilization protein HofP	FKPPE	-	CEPPQ	-	-	-	-	-	-	-
1	1	0	2	spp28631	HOLB_ECOLI	DNA polymerase III subunit delta'	APPPE	-	EEPPA	-	-	-	-	-	-	-
1	1	0	2	spp0A0H5	HSLU_ECOLI	ATP-dependent protease ATPase subunit HslU	MAPPG	-	LIPPA	-	-	-	-	-	-	-
1	0	1	2	spp33129	HTRE_ECOLI	Outer membrane usher protein HtrE	VVPPG	-	-	-	-	-	-	MLPPT	-	-
1	0	1	2	spp77668	HYFL_ECOLI	Hydrogenase-4 component 1	QIPPS	-	-	-	-	-	-	GCPCY	-	-
1	0	1	2	spp39344	IDNT_ECOLI	Gnt-II system L-idonate transporter	KAPPE	-	-	-	-	-	-	FLPPH	-	-
1	1	0	2	spp63183	KUP_ECOLI	Low affinity potassium transport system protein Kup	EIPPN	-	KSPPV	-	-	-	-	-	-	-
1	0	1	2	spp0A0H3	LDCL_ECOLI	Inducible lysine decarboxylase	PYPYG	-	-	-	-	-	-	ILPLL	-	-
1	1	0	2	spp39398	LGOT_ECOLI	Probable L-galactonate transporter	PVPPD	-	VSPPI	-	-	-	-	-	-	-
1	0	1	2	spp15977	MALQ_ECOLI	4-alpha-glucanotransferase	GAPPD	-	-	-	-	-	-	GLPPM	-	-
1	0	1	2	spp31121	MARB_ECOLI	Multiple antibiotic resistance protein MarB	VVPPS	-	-</							

Anzahl Motive					Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive				
S	M	W	Σ	Motiv 1			Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4		
1	0	1	2	spi16256	PANF_EC01	Sodium/pantothenate symporter	WKPE	-	-	-	-	-	VLPFF	-	-	-	-
1	1	0	2	spi42588	PAT_EC01	Putrescine aminotransferase	LPPG	-	-	-	-	-	IEPLL	-	-	-	-
1	1	0	2	spi00ABF1	PCNB_EC01	Poly(A) polymerase I	SAPPD	-	-	-	-	-	DIPPA	-	-	-	-
1	1	0	2	spi21514	PDEL_EC01	Cyclic di-GMP phosphodiesterase PdeL	IIPPD	-	-	-	-	-	YSPPV	-	-	-	-
1	0	1	2	spi06999	PFKB_EC01	ATP-dependent 6-phosphofruktokinase isozyme 2	VPPPV	-	-	-	-	-	-	SLPPG	-	-	-
1	0	1	2	spi16692	PHNP_EC01	Phosphoribosyl 1,2-cyclic phosphate phosphodiesterase	YGGPD	-	-	-	-	-	-	SHPPR	-	-	-
1	1	0	2	spi07001	PNTA_EC01	NAD(P) transhydrogenase subunit alpha	PAPPI	-	-	-	-	-	KVPPA	-	-	-	-
1	0	1	2	spi31135	POTH_EC01	Putrescine transport system permease protein PotH	SKPPG	-	-	-	-	-	-	AIPPY	-	-	-
1	1	0	2	spi00ADR8	PPNN_EC01	Pyrimidine/purine nucleotide 5'-monophosphate nucleosidase	AEPPN	-	-	-	-	-	INPPE	-	-	-	-
1	1	0	2	spi43671	PQIB_EC01	Paraquat-inducible protein B	LAPPD	-	-	-	-	-	DSPPL	-	-	-	-
1	0	1	2	spi37344	PSPF_EC01	Psp operon transcriptional activator	RRPPE	-	-	-	-	-	-	QLPPL	-	-	-
1	0	1	2	spi32154	PTFLB_EC01	1 Fructose-like PTS system EHB component	TIPPS	-	-	-	-	-	-	LMPPV	-	-	-
1	0	1	2	spi09546	PUTA_EC01	Bifunctional protein PutA	GVPPG	-	-	-	-	-	-	ATPPA	-	-	-
1	1	0	2	spi00AAZ4	RARA_EC01	Replication-associated recombination protein A	WGPPG	-	-	-	-	-	YFPPE	-	-	-	-
1	0	1	2	spi08394	RECB_EC01	RecBCD enzyme subunit RecB	APPPD	-	-	-	-	-	-	GCPLL	-	-	-
1	1	0	2	spi00AG30	RHO_EC01	Transcription termination factor Rho	VAPPK	-	-	-	-	-	IRPPK	-	-	-	-
1	1	0	2	spi06992	RSMA_EC01	Ribosomal RNA small subunit methyltransferase A	EVPPS	-	-	-	-	-	TPPPK	-	-	-	-
1	0	1	2	spi75892	RUTG_EC01	Putative pyrimidine permease RutG	VPPPE	-	-	-	-	-	-	LMPPV	-	-	-
1	0	1	2	spi00AGH5	SACP_EC01	Putrescine export system permease protein SapC	KRPPG	-	-	-	-	-	-	LLPSS	-	-	-
1	0	1	2	spi14081	SELB_EC01	Selenocysteine-specific elongation factor	DVPPE	-	-	-	-	-	-	LNPPR	-	-	-
1	0	1	2	spi80645	SSUD_EC01	Alkanesulfonate monooxygenase	GEPPP	-	-	-	-	-	-	PYPLL	-	-	-
1	0	1	2	spi008956	TIRK_EC01	Type I restriction enzyme EcoKI R protein	QDPPI	-	-	-	-	-	-	NIPPC	-	-	-
1	0	1	2	spi76155	TFAQ_EC01	Prophage tail fiber assembly protein homolog TfaQ	IAPPD	-	-	-	-	-	-	YIPPH	-	-	-
1	0	1	2	spi77163	TFAR_EC01	Prophage tail fiber assembly protein homolog TfaR	IAPPD	-	-	-	-	-	-	YIPPH	-	-	-
1	1	0	2	spi07464	THGA_EC01	Galactoside-acetyltransferase	DIPPN	-	-	-	-	-	VEPPV	-	-	-	-
1	0	1	2	spi52097	TILS_EC01	tRNA(Ile)-lysidine synthase	GVPPW	-	-	-	-	-	-	IRPPR	-	-	-
1	0	1	2	spi62601	TREF_EC01	Cytoplasmic trehalase	GRPPN	-	-	-	-	-	-	SQPPV	-	-	-
1	0	1	2	spi00A877	TRPA_EC01	Tryptophan synthase alpha chain	AAPPL	-	-	-	-	-	-	ICPPN	-	-	-
1	0	1	2	spi000895	TRPE_EC01	Anthraniolate synthase component 1	AAPPL	-	-	-	-	-	-	RFPPL	-	-	-
1	0	1	2	spi07023	TYRA_EC01	T-protein	GVPPD	-	-	-	-	-	-	KLPLL	-	-	-
1	0	1	2	spi00A6A0	UBIB_EC01	Probable protein kinase UbiB	WVPPD	-	-	-	-	-	-	LFPPH	-	-	-
1	1	0	2	spi12295	UNG_EC01	Uracil-DNA glycosylase	AIPPS	-	-	-	-	-	IVPPQ	-	-	-	-
1	0	1	2	spi00ADN0	VIAA_EC01	Protein ViaA	RVPPP	-	-	-	-	-	-	LLPPE	-	-	-
1	0	1	2	spi36881	YADI_EC01	Putative phosphotransferase enzyme IIA component YadiI	GAPPY	-	-	-	-	-	-	KNPPF	-	-	-
1	0	1	2	spi77324	YAGS_EC01	Putative xanthine dehydrogenase YagS FAD-binding subunit	YHPPG	-	-	-	-	-	-	LLPPL	-	-	-
1	0	1	2	spi46890	YBAE_EC01	Uncharacterized protein YbaE	SAPPG	-	-	-	-	-	-	ITPPL	-	-	-
1	1	0	2	spi40876	YCBF_EC01	Uncharacterized fibrial chaperone YcbF	AAPPL	-	-	-	-	-	VVPPK	-	-	-	-
1	0	1	2	spi75954	YCF5_EC01	Probable L,D-transpeptidase Ycf5	YPPPG	-	-	-	-	-	-	PLPPA	-	-	-
1	0	1	2	spi75961	YCFZ_EC01	Inner membrane protein YcfZ	NAPPG	-	-	-	-	-	-	STPPL	-	-	-
1	0	1	2	spi76010	YCGR_EC01	Flagellar brake protein YcgR	VPPPT	-	-	-	-	-	-	LHPPY	-	-	-
1	0	1	2	spi77156	YDCU_EC01	Inner membrane ABC transporter permease protein YdcU	VGPPG	-	-	-	-	-	-	RLPSS	-	-	-
1	0	1	2	spi77748	YDJU_EC01	Uncharacterized protein YdjI	CVPPE	-	-	-	-	-	-	ICPPE	-	-	-
1	0	1	2	spi76464	YFAS_EC01	Putative UPF0192 protein YfaS	YAPPA	-	-	-	-	-	-	TMPPV	-	-	-
1	0	1	2	spi67729	YFEO_EC01	Putative ion-transport protein YfeO	PVPPS	-	-	-	-	-	-	QQPPH	-	-	-
1	0	1	2	spi52060	YGGU_EC01	UPF0235 protein YggU	QIPPE	-	-	-	-	-	-	TAPPV	-	-	-
1	1	0	2	spi28638	YHDJ_EC01	DNA adenine methyltransferase YhdJ	ADPPY	-	-	-	-	-	KNPPQ	-	-	-	-
1	0	1	2	spi36677	YHDN_EC01	Uncharacterized protein YhdN	HVPPE	-	-	-	-	-	-	CLPPE	-	-	-
1	0	1	2	spi45549	YHFW_EC01	Uncharacterized protein YhfW	APPPQ	-	-	-	-	-	-	PLPPL	-	-	-
1	0	1	2	spi00AGH1	YHHU_EC01	Inner membrane transport permease YhhJ	NIPPN	-	-	-	-	-	-	FLPPE	-	-	-
1	1	0	2	spi37197	YHJA_EC01	Probable cytochrome c peroxidase	GGPPL	-	-	-	-	-	TMPPT	-	-	-	-
1	0	1	2	spi37675	YIAN_EC01	2,3-diketo-L-gulonate TRAP transporter large permease protein Yian	IIPPS	-	-	-	-	-	-	ITPPI	-	-	-
1	1	0	2	spi00A8C8	YIDD_EC01	Putative membrane protein insertion efficiency factor	PVPPG	-	-	-	-	-	MAPPL	-	-	-	-
1	1	0	2	spi00ADN6	YIFL_EC01	Uncharacterized lipoprotein YifL	APPPT	-	-	-	-	-	YFPPA	-	-	-	-
1	0	1	2	spi32688	YJBG_EC01	Uncharacterized protein YjgG	GPPPS	-	-	-	-	-	-	GNPPL	-	-	-
1	1	0	2	spi032528	YJPD_EC01	Uncharacterized lipoprotein Yjpd	DAPPL	-	-	-	-	-	MPPPF	-	-	-	-
1	1	0	2	spi046871	YQJH_EC01	NADPH-dependent ferric-chelate reductase	PRPPS	-	-	-	-	-	FVPPT	-	-	-	-
1	0	1	2	spi64612	ZAPE_EC01	Cell division protein ZapE	NIPPD	-	-	-	-	-	-	STPPA	-	-	-
0	0	2	2	spi00A9Q5	ACCD_EC01	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	-	-	-	-	-	-	-	KLPPG	VVPPV	-	-
0	0	2	2	spi00CK95	ACFD_EC01	Putative lipoprotein AcfD homolog	-	-	-	-	-	-	-	DTPPV	NRPPR	-	-
0	2	0	2	spi21515	ACPH_EC01	Acyl carrier protein phosphodiesterase	-	-	-	-	-	-	DSPPR	SFPD	-	-	-
0	0	2	2	spi23872	AES_EC01	Acetyl esterase	-	-	-	-	-	-	-	DLPPW	EVPPC	-	-
0	2	0	2	spi33224	AIDB_EC01	Putative acyl-CoA dehydrogenase AibB	-	-	-	-	-	-	ASPPM	VNPPE	-	-	-
0	0	2	2	spi77434	ALAC_EC01	Glutamate-pyruvate aminotransferase AlaC	-	-	-	-	-	-	-	ATPPH	RLPPY	-	-
0	1	1	2	spi37685	ALDB_EC01	Aldehyde dehydrogenase B	-	-	-	-	-	-	NNPPS	-	-	-	-
0	0	2	2	spi77671	ALLB_EC01	Alkanoate	-	-	-	-	-	-	-	CSPP1	PCPPE	-	-
0	1	1	2	spi00081	AMPC_EC01	Beta-lactamase	-	-	-	-	-	-	NVPPA	-	-	-	-
0	1	1	2	spi00AE0	ARAC_EC01	Arabinose operon regulatory protein	-	-	-	-	-	-	LFPPG	-	-	-	-
0	0	2	2	spi00A8S1	ARGP_EC01	HTH-type transcriptional regulator ArgP	-	-	-	-	-	-	-	DLPPG	TVPPR	-	-
0	0	2	2	spi00A6D3	AROA_EC01	3-phosphohikimate 1-carboxyvinyltransferase	-	-	-	-	-	-	-	ITPPE	NYPLA	-	-
0	0	2	2	spi00A962	ASPGI_EC01	1L-asparaginase 1	-	-	-	-	-	-	-	NLPLL	NTPPA	-	-
0	0	2	2	spi00065	ATOC_EC01	Regulatory protein AtoC	-	-	-	-	-	-	-	ILPPL	DLPPQ	-	-
0	0	2	2	spi00072	BGAL_EC01	Beta-galactosidase	-	-	-	-	-	-	-	AHPFF	VNPPF	-	-
0	0	2	2	spi12999	BIOC_EC01	Malonyl-[acyl-carrier protein]-methyltransferase	-	-	-	-	-	-	-	FLPPD	LSPPM	-	-
0	2	0	2	spi12998	BIOF_EC01	8-amino-7-oxononanoate synthase	-	-	-	-	-	-	IRPPT	SMPPA	-	-	-
0	0	2	2	spi00661	BTUD_EC01	Vitamin B12 import ATP-binding protein Btud	-	-	-	-	-	-	-	LTPPN	QTPPF	-	-
0	0	1	2	spi37028	BTUF_EC01	Vitamin B12-binding protein	-	-	-	-	-	-	DYPPQ	-	-	-	-
0	0	2	2	spi23890	CADC_EC01	Transcriptional activator CadC	-	-	-	-	-	-	-	SPPPI	ATPPE	-	-
0	0	2	2	spi77808	CINAL_EC01	1 NMN amidohydrolase-like protein YfaY	-	-	-	-	-	-	-	QLPPG	PQPPV	-	-
0	1	1	2	spi00AE74	CITT_EC01	Citrate/succinate antiporter	-	-	-	-	-	-	GMPPQ	-	-	-	-
0	1	1	2	spi76113	CURA_EC01	NADPH-dependent curcumin reductase	-	-	-	-	-	-	YSPPV	-	-	-	-
0	0	2	2	spi77463	DDPC_EC01	Probable D,D-dipeptide transport system permease protein DdpC	-	-	-	-	-	-	-	SLPPL	LLPSS	-	-
0	0	2	2	spi00A6P5	DER_EC01	GTPase Der	-	-	-	-	-	-	-	HQPLL	YNPPI	-	-
0	0	2	2	spi15877	DHG_EC01	Quinoprotein glucose dehydrogenase	-	-	-	-	-	-	-	FTPPS	TSPPI	-	-
0	0	2	2	spi000393	DHNA_EC01	NADH dehydrogenase	-	-	-	-	-	-	-	ALPPR	FVPPR	-	-
0	0	2	2	spi27296	DING_EC01	ATP-dependent DNA helicase DniG	-	-	-	-	-	-	-	AFPP1	TIPPL	-	-
0	0	2	2	spi006149	DLD_EC01	Quinone-dependent D-lactate dehydrogenase	-	-	-	-	-	-	-	HLPPE	HLPPR	-	-
0	2	0	2	spi00AEG1	DPCC_EC01	Dipeptide transport system permease protein DppC	-	-	-	-	-	-	AQPPT	LAPPA	-	-	-
0	0	2	2	spi75901	EFEU_EC01	Putative inactive ferrous iron permease EfeU	-	-	-	-	-	-	-	ALPPR	IWPLL	-	-
0	0	2	2	spi19636	EUTC_EC01	Ethanolamine ammonia-lyase light chain	-	-	-	-	-	-	-	ILPPL	GTPPV	-	-
0	1	1	2	spi00AA15	FABF_EC01	3-oxoacyl-[acyl-carrier-protein] synthase 2	-	-	-	-	-	-	AVPPT	-	-	-	-
0	0	2	2	spi21177	FADB_EC01	Fatty acid oxidation complex subunit alpha	-	-	-	-	-	-	-	GPPFF	YPPPV	-	-
0	0	2	2	spi047146	FADE_EC01	Acyl-coenzyme A dehydrogenase	-	-	-	-	-	-	-	DLPPE	VMPPM	-	-
0	1	1	2	spi76503	FADI_EC01	3-keatoacyl-CoA thiolase FadI	-	-	-	-	-	-	PVPPA	-	-	-	-
0	0	2	2	spi13024	FDHE_EC01	Protein FdhE	-	-	-	-	-	-	-	GKPLL	MIPPL	-	-
0	0	2	2	spi19323	FHLA_EC01	Formate hydrogenly											

Anzahl Motive				Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive			
S	M	W	Σ			Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4
0	0	2	2	spi19932	HYAF_ECOLI	Hydrogenase-1 operon protein Hyaf	-	-	-	-	-	-	-	-	
0	1	1	2	spi77437	HYFF_ECOLI	Hydrogenase-4 component F	-	-	-	AKPPQ	-	-	-	-	
0	1	1	2	spi00AB80	ILVE_ECOLI	Branched-chain-amino-acid aminotransferase	-	-	-	VNPPA	-	-	-	-	
0	0	2	2	spi00CF91	INSL1_ECOLI	1 Putative transposase InsL for insertion sequence element IS186A	-	-	-	-	-	-	-	-	
0	0	2	2	spi00CF92	INSL2_ECOLI	1 Putative transposase InsL for insertion sequence element IS186B	-	-	-	-	-	-	-	-	
0	0	2	2	spi00CF93	INSL3_ECOLI	1 Putative transposase InsL for insertion sequence element IS186C	-	-	-	-	-	-	-	-	
0	0	2	2	spi75987	IRAM_ECOLI	Anti-adaptor protein Iram	-	-	-	-	-	-	-	-	
0	0	2	2	spi00A754	KEFF_ECOLI	Glutathione-regulated potassium-efflux system ancillary protein Keff	-	-	-	-	-	-	-	-	
0	0	2	2	spi30125	LEU3_ECOLI	3-isopropylmalate dehydrogenase	-	-	-	-	-	-	-	-	
0	0	2	2	spi25772	LIGB_ECOLI	DNA ligase B	-	-	-	-	-	-	-	-	
0	0	2	2	spi33231	LLDP_ECOLI	L-lactate permease	-	-	-	-	-	-	-	-	
0	0	2	2	spi00ACV0	LPXL_ECOLI	Lipid A biosynthesis lauroyltransferase	-	-	-	-	-	-	-	-	
0	0	2	2	spi24205	LPXM_ECOLI	Lipid A biosynthesis myristoyltransferase	-	-	-	-	-	-	-	-	
0	0	2	2	spi68187	MALK_ECOLI	Maltose/maltodextrin import ATP-binding protein MalK	-	-	-	-	-	-	-	-	
0	0	2	2	spi69739	MBHS_ECOLI	Hydrogenase-1 small chain	-	-	-	-	-	-	-	-	
0	0	2	2	spi00AEZ1	METF_ECOLI	5,10-methylenetetrahydrofolate reductase	-	-	-	-	-	-	-	-	
0	1	1	2	spi77397	MHPA_ECOLI	3-(3-hydroxy-phenyl)propionate/3-hydroxycinnamic acid hydroxylase	-	-	-	LAPPK	-	-	-	-	
0	0	2	2	spi00ABR9	MHPB_ECOLI	2,3-dihydroxyphenylpropionate/2,3-dihydroxycinnamic acid 1,2-dioxygenase	-	-	-	-	-	-	-	-	
0	1	1	2	spi00ACV6	MPAA_ECOLI	Protein MpaA	-	-	-	AFPPG	-	-	-	-	
0	1	1	2	spi16926	MREC_ECOLI	Cell shape-determining protein MreC	-	-	-	AQPPA	-	-	-	-	
0	0	2	2	spi23367	MUTL_ECOLI	DNA mismatch repair protein MutL	-	-	-	-	-	-	-	-	
0	0	2	2	spi009152	NARG_ECOLI	Respiratory nitrate reductase 1 alpha chain	-	-	-	-	-	-	-	-	
0	1	1	2	spi76007	NHAP2_ECOLI	1 K(+)-H(+) antiporter NhaP2	-	-	-	QSPPV	-	-	-	-	
0	1	1	2	spi77567	NHOA_ECOLI	N-hydroxyarylamine-acetyltransferase	-	-	-	SNPPA	-	-	-	-	
0	2	0	2	spi00ADA3	NLPD_ECOLI	Murein hydrolase activator NlpD	-	-	-	SNPPA	TPPPK	-	-	-	
0	1	1	2	spi37596	NORW_ECOLI	Náric oxide reductase FNR-NAD(+) reductase	-	-	-	LMPEE	-	-	-	-	
0	1	1	2	spi28903	NRDD_ECOLI	Anaerobic ribonucleoside-triphosphate reductase	-	-	-	IEPPK	-	-	-	-	
0	1	1	2	spi32710	NRFE_ECOLI	Probable cytochrome c-type biogenesis protein NrFe	-	-	-	IFPPA	-	-	-	-	
0	0	2	2	spi33021	NUPX_ECOLI	Putative nucleoside permease NupX	-	-	-	-	-	-	-	-	
0	1	1	2	spi40120	OPGD_ECOLI	Glucans biosynthesis protein D	-	-	-	YFPPA	-	-	-	-	
0	1	1	2	spi62517	OPGH_ECOLI	Glucans biosynthesis glucosyltransferase H	-	-	-	YSPPQ	-	-	-	-	
0	0	2	2	spi23843	OPPA_ECOLI	Periplasmic oligopeptide-binding protein	-	-	-	-	-	-	-	-	
0	1	1	2	spi76027	OPPD_ECOLI	Oligopeptide transport ATP-binding protein OppD	-	-	-	GNPPN	-	-	-	-	
0	2	0	2	spi77737	OPPF_ECOLI	Oligopeptide transport ATP-binding protein OppF	-	-	-	INPPS	WQPPK	-	-	-	
0	1	1	2	spi75793	PFLF_ECOLI	Putative formate acetyltransferase 3	-	-	-	HQPPQ	-	-	-	-	
0	2	0	2	spi36938	PGM_ECOLI	Phosphoglucomutase	-	-	-	HNPEE	YNPPN	-	-	-	
0	0	2	2	spi46806	PHYDA_ECOLI	1 D-phenylhydantoïnase	-	-	-	-	-	-	-	-	
0	1	1	2	spi43676	PITB_ECOLI	Probable low-affinity inorganic phosphate transporter 2	-	-	-	MEPPL	-	-	-	-	
0	0	2	2	spi00A7A7	PLSB_ECOLI	Glycerol-3-phosphate acyltransferase	-	-	-	-	-	-	-	-	
0	1	1	2	spi26647	PLSC_ECOLI	1-acyl-sn-glycerol-3-phosphate acyltransferase	-	-	-	VQPPY	-	-	-	-	
0	1	1	2	spi31134	POTG_ECOLI	Putrescine transport ATP-binding protein PotG	-	-	-	EEPPA	-	-	-	-	
0	1	1	2	spi007102	PPA_ECOLI	Periplasmic AppA protein	-	-	-	PHPPQ	-	-	-	-	
0	0	2	2	spi77495	PRPE_ECOLI	Propionate-CoA ligase	-	-	-	-	-	-	-	-	
0	0	2	2	spi77743	PRPR_ECOLI	Propionate catabolism operon regulatory protein	-	-	-	-	-	-	-	-	
0	1	1	2	spi33024	PSUT_ECOLI	Putative pseudouridine transporter	-	-	-	YFPPG	-	-	-	-	
0	0	2	2	spi00A9M8	PTA_ECOLI	Phosphate acetyltransferase	-	-	-	-	-	-	-	-	
0	0	2	2	spi00A7D1	PTH_ECOLI	Peptidyl-tRNA hydrolase	-	-	-	-	-	-	-	-	
0	0	2	2	spi77212	RCLA_ECOLI	Probable pyridine nucleotide-disulfide oxidoreductase RclA	-	-	-	-	-	-	-	-	
0	1	1	2	spi15043	RECQ_ECOLI	ATP-dependent DNA helicase RecQ	-	-	-	LDPPK	-	-	-	-	
0	0	2	2	spi16916	RHSA_ECOLI	Protein RhsA	-	-	-	-	-	-	-	-	
0	0	2	2	spi16917	RHSB_ECOLI	Protein RhsB	-	-	-	-	-	-	-	-	
0	0	2	2	spi16918	RHSC_ECOLI	Protein RhsC	-	-	-	-	-	-	-	-	
0	2	0	2	spi000452	RIR1_ECOLI	Ribonucleoside-diphosphate reductase 1 subunit alpha	-	-	-	FEPPA	IEPPR	-	-	-	
0	1	1	2	spi00A7J7	RLI1_ECOLI	50S ribosomal protein L11	-	-	-	PSPPV	-	-	-	-	
0	1	1	2	spi75876	RLM1_ECOLI	Ribosomal RNA large subunit methyltransferase I	-	-	-	MDPPK	-	-	-	-	
0	1	1	2	spi33643	RLUD_ECOLI	Ribosomal large subunit pseudouridine synthase D	-	-	-	PRPPK	-	-	-	-	
0	0	2	2	spi00AC10	ROB_ECOLI	Right origin-binding protein	-	-	-	-	-	-	-	-	
0	0	2	2	spi00ABG7	RODA_ECOLI	Peptidoglycan glycosyltransferase MrdB	-	-	-	-	-	-	-	-	
0	0	2	2	spi00AFX9	RSEB_ECOLI	Sigma-E factor regulatory protein RseB	-	-	-	-	-	-	-	-	
0	0	2	2	spi76182	RSXD_ECOLI	Ion-translocating oxidoreductase complex subunit D	-	-	-	-	-	-	-	-	
0	1	1	2	spi47622	SAPA_ECOLI	Probable ABC transporter periplasmic-binding protein SapaA	-	-	-	ESPPH	-	-	-	-	
0	0	2	2	spi77249	SFMC_ECOLI	Probable fibrial chaperone SfmC	-	-	-	-	-	-	-	-	
0	1	1	2	spi75715	SFMH_ECOLI	Uncharacterized fibrial-like protein SfmH	-	-	-	QFPPQ	-	-	-	-	
0	1	1	2	spi39366	SGCX_ECOLI	Putative aminopeptidase SgcX	-	-	-	YNPPQ	-	-	-	-	
0	1	1	2	spi30852	SMF_ECOLI	Protein Smf	-	-	-	FYPPQ	-	-	-	-	
0	2	0	2	spi00AGD7	SRP54_ECOLI	1 Signal recognition particle protein	-	-	-	AQPPA	MMPPG	-	-	-	
0	1	1	2	spi00ACA3	SSPA_ECOLI	Stringent starvation protein A	-	-	-	DNPPQ	-	-	-	-	
0	1	1	2	spi00AFZ3	SSPB_ECOLI	Stringent starvation protein B	-	-	-	DEPPQ	-	-	-	-	
0	1	1	2	spi76072	STFR_ECOLI	Prophage site tail fiber protein homolog StfR	-	-	-	GFPPS	-	-	-	-	
0	0	2	2	spi008312	SYFA_ECOLI	Phenylalanine-tRNA ligase alpha subunit	-	-	-	-	-	-	-	-	
0	0	2	2	spi005719	TISK_ECOLI	Type-1 restriction enzyme EcoKI specificity protein	-	-	-	-	-	-	-	-	
0	1	1	2	spi47539	TAUC_ECOLI	Taurine transport system permease protein TauC	-	-	-	LNPPQ	-	-	-	-	
0	0	2	2	spi31549	THIP_ECOLI	Thiamine transport system permease protein ThiP	-	-	-	-	-	-	-	-	
0	0	2	2	spi00ABV6	TOLR_ECOLI	Tol-Pal system protein TolR	-	-	-	-	-	-	-	-	
0	1	1	2	spi14294	TOP3_ECOLI	DNA topoisomerase 3	-	-	-	KSPPA	-	-	-	-	
0	0	2	2	spi007649	TRUA_ECOLI	tRNA pseudouridine synthase A	-	-	-	-	-	-	-	-	
0	0	2	2	spi39414	TTDT_ECOLI	L-tartrate/succinate antiporter	-	-	-	-	-	-	-	-	
0	1	1	2	spi00AG80	UGPB_ECOLI	sn-glycerol-3-phosphate-binding periplasmic protein UgpB	-	-	-	EQPPK	-	-	-	-	
0	1	1	2	spi00AGM7	URAA_ECOLI	Uracil permease	-	-	-	LFPPA	-	-	-	-	
0	1	1	2	spi75970	VXIS_ECOLI	Prophage excisionase-like protein	-	-	-	SNPPS	-	-	-	-	
0	0	2	2	spi71237	WCAC_ECOLI	Putative colanic acid biosynthesis glycosyl transferase WcaC	-	-	-	-	-	-	-	-	
0	1	1	2	spi00AGM9	XANP_ECOLI	Xanthine permease XanP	-	-	-	IFPPE	-	-	-	-	
0	2	0	2	spi46799	XDHA_ECOLI	Putative xanthine dehydrogenase molybdenum-binding subunit XdhA	-	-	-	TNPPA	GEPPY	-	-	-	
0	0	2	2	spi30143	YAAJ_ECOLI	Uncharacterized transporter YaaJ	-	-	-	-	-	-	-	-	
0	0	2	2	spi00AFN6	YADH_ECOLI	Inner membrane transport permease Yadh	-	-	-	-	-	-	-	-	
0	1	1	2	spi33128	YADV_ECOLI	Probable fibrial chaperone Yadv	-	-	-	EVPPK	-	-	-	-	
0	0	2	2	spi00A8U2	YAFD_ECOLI	UPF0294 protein YafD	-	-	-	-	-	-	-	-	
0	0	2	2	spi75682	YAGE_ECOLI	Putative 2-dehydro-3-deoxy-D-gluconate aldolase YagE	-	-	-	-	-	-	-	-	
0	0	2	2	spi77596	YAGF_ECOLI	D-xylonate dehydratase YagF	-	-	-	-	-	-	-	-	
0	2	0	2	spi77301	YBAP_ECOLI	Uncharacterized protein YbaP	-	-	-	NAPPQ	AMPPG	-	-	-	
0	1	1	2	spi75749	YBGP_ECOLI	Uncharacterized fibrial chaperone YbgP	-	-	-	EIPPA	-	-	-	-	
0	0	2	2	spi30177	YBIB_ECOLI	Uncharacterized protein YbiB	-	-	-	-	-	-	-	-	
0	1	1	2	spi37443	YCAL_ECOLI	Uncharacterized protein YcaI	-	-	-	MQPPA	-	-	-	-	
0	2	0	2	spi00A8X4	YCCY_ECOLI	UPF0319 protein YccY	-	-	-	AIPPA	ISPPL	-	-	-	
0	0	2	2	spi29013	YCGB_ECOLI	Uncharacterized protein YcgB	-	-	-	-	-	-	-	-	
0	1	1	2	spi76108	YDCS_ECOLI	Bifunctional polyhydroxybutyrate synthase / ABC transporter protein	-	-	-	AEPPT	-	-	-	-	
0	0	2	2	spi00AA70	YEDA_ECOLI	Uncharacterized inner membrane transporter YedA	-	-	-	-	-	-	-	-	

Anzahl Motive					Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive				
S	M	W	Σ	Motiv 1			Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4		
0	1	1	2	spQ46844	YGHT_ECOLI	Uncharacterized ATP-binding protein YghT	-	-	-	VFPPN	-	-	ITPPL	-	-	-	
0	1	1	2	spQ46861	YGIQ_ECOLI	UPF0313 protein YgiQ	-	-	-	INPPA	-	-	VQPPR	-	-	-	
0	0	2	2	spQ46863	YGIS_ECOLI	Probable deoxycholate-binding periplasmic protein YgiS	-	-	-	-	-	-	LTPPE	EKPPF	-	-	
0	1	1	2	spP0AA73	YHBE_ECOLI	Uncharacterized inner membrane transporter YhbE	-	-	-	MEPPT	-	-	RLPPL	-	-	-	
0	0	2	2	spP0ADW6	YHCC_ECOLI	Protein YhcC	-	-	-	-	-	-	HTPPE	WLPTT	-	-	
0	0	2	2	spP46474	YHDP_ECOLI	Uncharacterized protein YhdP	-	-	-	-	-	-	TLPPL	NMPPM	-	-	
0	0	2	2	spP45767	YHDX_ECOLI	Putative amino-acid ABC transporter permease protein YhdX	-	-	-	-	-	-	IIPPL	NIPPL	-	-	
0	0	2	2	spP45768	YHDY_ECOLI	Inner membrane amino-acid ABC transporter permease protein YhdY	-	-	-	-	-	-	LIPPL	SHPPR	-	-	
0	1	1	2	spP37645	YHJG_ECOLI	AsmA family protein YhjG	-	-	-	GNPPD	-	-	DTPPF	-	-	-	
0	1	1	2	spP37648	YHJJ_ECOLI	Protein YhjJ	-	-	-	DIPPA	-	-	PMPPV	-	-	-	
0	0	2	2	spP37691	YIBQ_ECOLI	Uncharacterized protein YibQ	-	-	-	-	-	-	NLPPD	STPPK	-	-	
0	2	0	2	spP60872	YIDE_ECOLI	Putative transport protein YidE	-	-	-	FMPPS	TDPPA	-	-	-	-	-	
0	0	2	2	spP32139	YIHR_ECOLI	Uncharacterized protein YihR	-	-	-	-	-	-	FLPPS	SCPPN	-	-	
0	0	2	2	spP32681	YJAH_ECOLI	Uncharacterized protein YjaH	-	-	-	-	-	-	SLPPF	SKPPV	-	-	
0	0	2	2	spP32717	YJCS_ECOLI	Putative alkylaryl-sulfatase YjCS	-	-	-	-	-	-	KLPPA	VTPPA	-	-	
0	0	2	2	spP39359	YJHH_ECOLI	Probable 2-dehydro-3-deoxy-D-pentonate aklolase YjH	-	-	-	-	-	-	IIPPV	CLPPI	-	-	
0	0	2	2	spP39389	YJIR_ECOLI	Uncharacterized HTH-type transcriptional regulator YjIR	-	-	-	-	-	-	AQPPV	PVPPM	-	-	
0	1	1	2	spP75968	YMFEC	Uncharacterized protein YmfE	-	-	-	FEPPK	-	-	YVPMI	-	-	-	
0	0	2	2	spP76150	YNEK_ECOLI	Uncharacterized protein YneK	-	-	-	-	-	-	ETPPK	VRPPI	-	-	
0	1	1	2	spP77739	YNTA_ECOLI	Putative kinase YniA	-	-	-	EQPPQ	-	-	YLPPR	-	-	-	
0	0	2	2	spP76222	YNTA_ECOLI	Uncharacterized protein YniA	-	-	-	-	-	-	GLPPL	VTPPT	-	-	
0	1	1	2	spP76224	YNJC_ECOLI	Inner membrane ABC transporter permease protein YnjC	-	-	-	GNPPT	-	-	FTPPM	-	-	-	
0	2	0	2	spP76559	YYPG_ECOLI	Uncharacterized protein YypG	-	-	-	SVPPA	DEPPL	-	-	-	-	-	
0	1	1	2	spP46570	YQGC_ECOLI	Uncharacterized protein YagC	-	-	-	VSPRR	-	-	PRPPV	-	-	-	
0	1	1	2	spP77616	YQIH_ECOLI	Uncharacterized fibrial chaperone YqiH	-	-	-	EIPPK	-	-	TVPPM	-	-	-	
0	1	1	2	spP39310	YTFB_ECOLI	Uncharacterized protein YtfB	-	-	-	SQPPT	-	-	PLPPM	-	-	-	
0	1	1	2	spP75757	ZITB_ECOLI	Zinc transporter ZitB	-	-	-	RRPPT	-	-	VIPPH	-	-	-	
1	0	0	1	spP75616	YAAAX_ECOLI	Uncharacterized protein YaaX	HGPPPPP	-	-	-	-	-	-	-	-	-	
1	0	0	1	spP46481	AAEB_ECOLI	p-hydroxybenzoic acid efflux pump subunit AaeB	QAPPQ	-	-	-	-	-	-	-	-	-	
1	0	0	1	spP11071	ACEK_ECOLI	Isocitrate dehydrogenase kinase/phosphatase	IPPPR	-	-	-	-	-	-	-	-	-	
1	0	0	1	spP31466	ADEP_ECOLI	Adenine permease AdeP	SAPPS	-	-	-	-	-	-	-	-	-	
1	0	0	1	spP0ACR0	ALLS_ECOLI	HTH-type transcriptional activator AIs	MRPPS	-	-	-	-	-	-	-	-	-	
1	0	0	1	spP26365	AMIB_ECOLI	N-acetylmuramoyl-L-alanine amidase AmiB	VPPPPPPP	-	-	-	-	-	-	-	-	-	
1	0	0	1	spP0AD70	AMPH_ECOLI	D-alanyl-D-alanine-carboxypeptidase/endopeptidase AmpH	YAPPG	-	-	-	-	-	-	-	-	-	
1	0	0	1	spP12008	AROC_ECOLI	Chorismate synthase	GVPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP76460	ATOE_ECOLI	Putative short-chain fatty acid transporter	DAPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0ABB0	ATPA_ECOLI	ATP synthase subunit alpha	RRPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0ABB4	ATPB_ECOLI	ATP synthase subunit beta	NEPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spC1P605	AZUC_ECOLI	Uncharacterized protein AzuC	DVPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP37657	BCSE_ECOLI	Cyclic di-GMP binding protein BcsE	GAPPL	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP66948	BEPA_ECOLI	Beta-barrel assembly-enhancing protease	SRPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP06864	BGA2_ECOLI	Evolved beta-galactosidase subunit alpha	IAPPV	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP05804	BGLR_ECOLI	Beta-glucuronidase	TIPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP56976	BLR_ECOLI	Divisome-associated membrane protein Blr	YQPPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP39297	B SMA_ECOLI	Lipoprotein BsmA	PAPPV	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP76345	C56H_ECOLI	Cytochrome b561 homolog 1	VPPPP	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP75925	C56L_ECOLI	Cytochrome b561 homolog 2	SPPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0ABM9	CCMH_ECOLI	Cytochrome c-type biogenesis protein CcmH	YDPLL	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0AE60	CEDA_ECOLI	Cell division activator Ceda	PAPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0ABH9	CLPA_ECOLI	ATP-dependent Clp protease ATP-binding subunit ClpA	GAPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP63284	CLPB_ECOLI	Chaperone protein ClpB	GAPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0A8Q6	CLPS_ECOLI	ATP-dependent Clp protease adapter protein ClpS	LKPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0A6H8	CLSA_ECOLI	Cardiolipin synthase A	LPPPP	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0A8A4	CLSB_ECOLI	Cardiolipin synthase B	VDPPA	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP08401	CREC_ECOLI	Sensor protein CreC	GPPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0DM85	CRFC_ECOLI	Clamp-binding protein CrfC	LPPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP52107	CSGC_ECOLI	Curli assembly protein CsgC	QWPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP76621	CSID_ECOLI	Protein CsiD	AAPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0AEA5	CYOE_ECOLI	Protoheme IX farnesyltransferase	AAPPV	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP33013	DACD_ECOLI	D-alanyl-D-alanine carboxypeptidase DacD	PQPPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0A9P6	DEAD_ECOLI	ATP-dependent RNA helicase DeaD	IVPPD	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP18775	DMSA_ECOLI	Dimethyl sulfoxide reductase DmsA	SWPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP03004	DNAA_ECOLI	Chromosomal replication initiator protein DnaA	IETPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0A8J2	DNAT_ECOLI	Primosomal protein 1	QIPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0AA89	DOSC_ECOLI	Diguanylate cyclase DosC	ADPPI	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP37313	DPPF_ECOLI	Dipeptide transport ATP-binding protein DppF	NPPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP36655	DSBD_ECOLI	Thioldisulfide interchange protein DsbD	CYPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP77263	ECPE_ECOLI	Probable fibrial chaperone EcpE	AKPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP75857	ELFC_ECOLI	Probable outer membrane usher protein Elic	YVPEE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP37661	EPTB_ECOLI	Kdo(2)-lipid A phosphoethanolamine 7'-transferase	LAPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP69902	FCTA_ECOLI	Formyl-CoA:oxalate CoA-transferase	DGPPPL	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP32176	FDOG_ECOLI	Formate dehydrogenase-major subunit	AAPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP05825	FEP A_ECOLI	Ferrienterobactin receptor	WVPPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP75780	FIU_ECOLI	Catechololate siderophore receptor Fiu	QPPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP76299	FLHB_ECOLI	Flagellar biosynthetic protein FlhB	EAPPL	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP77733	FOCB_ECOLI	Probable formate transporter 2	APPPF	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP25437	FRMA_ECOLI	S-(hydroxymethyl)glutathione dehydrogenase	VAPPK	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0AD68	FTSL_ECOLI	Peptidoglycan D,D-transpeptidase FtsI	VDPPV	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0AEN8	FUCM_ECOLI	L-fucose mutarotase	YAPPL	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP25748	GALS_ECOLI	HTH-type transcriptional regulator GalS	IIPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP23522	GARL_ECOLI	5-keto-4-deoxy-D-glucarate aklolase	RYPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0ACL0	GLPR_ECOLI	Glycerol-3-phosphate regulon repressor	APPPV	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP09832	GLTD_ECOLI	Glutamate synthase [NADPH] small chain	VDPPK	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spQ46851	GPR_ECOLI	L-glycerol-3-phosphate reductase	GPPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP04425	GSHB_ECOLI	Glutathione synthetase	KDPPF	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP75796	GSIA_ECOLI	Glutathione import ATP-binding protein GsiA	QAPPI	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP30178	HCBX_ECOLI	Hydroxycarboxylate dehydrogenase B	PVPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP06987	HIS7_ECOLI	Histidine biosynthesis bifunctional protein HisB	SEPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP08142	ILVB_ECOLI	Acetolactate synthase isozyme 1 large subunit	MVPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0OCF53	INS D1_ECOLI	I Transposase InsD for insertion element IS2A	AVPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0OCF54	INS D2_ECOLI	I Transposase InsD for insertion element IS2D	AVPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0OCF55	INS D3_ECOLI	I Transposase InsD for insertion element IS2F	AVPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0OCF56	INS D4_ECOLI	I Transposase InsD for insertion element IS2H	AVPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0OCF57	INS D5_ECOLI	I Transposase InsD for insertion element IS2I	AVPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0OCF58	INS D6_ECOLI	I Transposase InsD for insertion element IS2K	AVPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP0OCF60	INS D8_ECOLI	I Putative transposase InsD for insertion element IS2E	AVPPS	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP39347	INTB_ECOLI	Putative protein IntB	TIPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP75969	INTE_ECOLI	Prophage integrase IntE	EVPPG	-	-	-	-	-	-	-	-	-	-
1	0	0	1	spP62615	ISPE_ECOLI	4-diphosphotyridyl-2-C-methyl-D-erythritol kinase	VDPPE	-	-	-	-	-	-	-	-	-	-
1	0	0	1	sp													

Anzahl Motive					Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive			
S	M	W	Σ	Motiv 1			Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4	
1	0	0	1	1	spiP17109	MEND_ECOLI	2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP54745	MNGA_ECOLI	PTS system 2-alpha-mannosyl-D-glycerate-specific EIIBC component	-	-	-	-	-	-	-		
1	0	0	1	1	spiP32173	MOBA_ECOLI	Molybdenum cofactor guanylyltransferase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76096	MOKB_ECOLI	Regulatory protein MokB	-	-	-	-	-	-	-		
1	0	0	1	1	spiP09348	MOTA_ECOLI	Motility protein A	-	-	-	-	-	-	-		
1	0	0	1	1	spiP33940	MQO_ECOLI	Malate:quinone oxidoreductase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP77338	MSCK_ECOLI	Mechanosensitive channel MscK	-	-	-	-	-	-	-		
1	0	0	1	1	spiP09424	MTLD_ECOLI	Mannitol-1-phosphate 5-dehydrogenase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0F10	MTLR_ECOLI	Mannitol operon repressor	-	-	-	-	-	-	-		
1	0	0	1	1	spiP22634	MURI_ECOLI	Glutamate racemase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP39370	NANS_ECOLI	Probable 9-acetyl-N-acetylneuraminic acid deacetylase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0A0	NAPF_ECOLI	Ferredoxin-type protein NapF	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0A3	NAPG_ECOLI	Ferredoxin-type protein NapG	-	-	-	-	-	-	-		
1	0	0	1	1	spiP68739	NFI_ECOLI	Endonuclease V	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0G2	NHAR_ECOLI	Transcriptional activator protein NhaR	-	-	-	-	-	-	-		
1	0	0	1	1	spiP23898	NLPC_ECOLI	Probable endopeptidase NlpC	-	-	-	-	-	-	-		
1	0	0	1	1	spiP32664	NUDC_ECOLI	NADH pyrophosphatase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP39401	OPGB_ECOLI	Phosphoglycerol transferase I	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0CQ4	OXYR_ECOLI	Hydrogen peroxide-inducible genes activator	-	-	-	-	-	-	-		
1	0	0	1	1	spiP02919	PBPB_ECOLI	Penicillin-binding protein 1B	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76261	PDED_ECOLI	Probable cyclic di-GMP phosphodiesterase PdeD	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76446	PDEN_ECOLI	Probable cyclic di-GMP phosphodiesterase PdeN	-	-	-	-	-	-	-		
1	0	0	1	1	spiP08400	PHOR_ECOLI	Phosphate regulon sensor protein PhoR	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0A7A5	PIMT_ECOLI	Protein-L-aspartate-methyltransferase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0FK9	POTD_ECOLI	Spermidine/putrescine-binding periplasmic protein	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0AF1	POTE_ECOLI	Putrescine transporter PotE	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0B75	PROB_ECOLI	Glutamate 5-kinase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP07654	PSTA_ECOLI	Phosphate transport system permease protein PstA	-	-	-	-	-	-	-		
1	0	0	1	1	spiP69797	PTNAB_ECOLI	1 PTS system mannose-specific EIIB component	-	-	-	-	-	-	-		
1	0	0	1	1	spiP42910	PTPC1_ECOLI	1 N-acetylgalactosamine permease IIC component 1	-	-	-	-	-	-	-		
1	0	0	1	1	spiP42905	PTPC2_ECOLI	1 Putative N-acetylgalactosamine permease IIC component 2	-	-	-	-	-	-	-		
1	0	0	1	1	spiP07117	PUTP_ECOLI	Sodium/proline symporter	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76037	PUUP_ECOLI	Putrescine importer PuuP	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0E1	PYRD_ECOLI	Dihydroorotate dehydrogenase (quinone)	-	-	-	-	-	-	-		
1	0	0	1	1	spiP28304	QOR1_ECOLI	Quinone oxidoreductase 1	-	-	-	-	-	-	-		
1	0	0	1	1	spiP40719	QSEC_ECOLI	Sensor protein QseC	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76062	RACR_ECOLI	Prophage repressor RacR	-	-	-	-	-	-	-		
1	0	0	1	1	spiP31473	RAVA_ECOLI	ATPase RavA	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A08V0	RBN_ECOLI	Ribonuclease BN	-	-	-	-	-	-	-		
1	0	0	1	1	spiP39838	RCSD_ECOLI	Phosphotransferase RcsD	-	-	-	-	-	-	-		
1	0	0	1	1	spiP10100	RLPA_ECOLI	Endolytic peptidoglycan transglycosylase RlpA	-	-	-	-	-	-	-		
1	0	0	1	1	spiP30850	RNB_ECOLI	Exoribonuclease 2	-	-	-	-	-	-	-		
1	0	0	1	1	spiP21499	RNR_ECOLI	Ribonuclease R	-	-	-	-	-	-	-		
1	0	0	1	1	spiP24255	RP54_ECOLI	RNA polymerase sigma-54 factor	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0GB6	RPOE_ECOLI	ECF RNA polymerase sigma-E factor	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0DX9	RSMD_ECOLI	Ribosomal RNA small subunit methyltransferase D	-	-	-	-	-	-	-		
1	0	0	1	1	spiP68567	RSMJ_ECOLI	Ribosomal RNA small subunit methyltransferase J	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0FR0	RSSA_ECOLI	NTE family protein RssA	-	-	-	-	-	-	-		
1	0	0	1	1	spiP77611	RSXC_ECOLI	Ion-translocating oxidoreductase complex subunit C	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0G74	RUSA_ECOLI	Crossover junction endodeoxyribonuclease RusA	-	-	-	-	-	-	-		
1	0	0	1	1	spiP31675	SETA_ECOLI	Sugar efflux transporter A	-	-	-	-	-	-	-		
1	0	0	1	1	spiP33026	SETB_ECOLI	Sugar efflux transporter B	-	-	-	-	-	-	-		
1	0	0	1	1	spiP31436	SETC_ECOLI	Sugar efflux transporter C	-	-	-	-	-	-	-		
1	0	0	1	1	spiP77468	SFMD_ECOLI	Outer membrane usher protein SfmD	-	-	-	-	-	-	-		
1	0	0	1	1	spiP33595	SGRR_ECOLI	HTH-type transcriptional regulator SgrR	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0840	SURE_ECOLI	5/3'-nucleotidase SurE	-	-	-	-	-	-	-		
1	0	0	1	1	spiP000957	SYA_ECOLI	Alanine-tRNA ligase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP000956	SYI_ECOLI	Isoleucine-tRNA ligase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP07118	SYV_ECOLI	Valine-tRNA ligase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0867	TALA_ECOLI	Translokase A	-	-	-	-	-	-	-		
1	0	0	1	1	spiP19934	TOLA_ECOLI	Tol-Pal system protein TolA	-	-	-	-	-	-	-		
1	0	0	1	1	spiP33226	TORC_ECOLI	Cytochrome c-type protein TorC	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0858	TPIS_ECOLI	Triosephosphate isomerase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP23003	TRMA_ECOLI	tRNA/tmRNA (uracil-C(5))-methyltransferase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0GJ7	TRML_ECOLI	tRNA (cytidine(34)-2'-	-	-	-	-	-	-	-		
1	0	0	1	1	spiP43672	UUP_ECOLI	ABC transporter ATP-binding protein uup	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A08F8	UVRB_ECOLI	UvrABC system protein B	-	-	-	-	-	-	-		
1	0	0	1	1	spiP24215	UXUA_ECOLI	Mannonate dehydratase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP37750	WBBJ_ECOLI	Putative lipopolysaccharide biosynthesis-acetyl transferase WbbJ	-	-	-	-	-	-	-		
1	0	0	1	1	spiP71244	WCAM_ECOLI	Colanic acid biosynthesis protein WcaM	-	-	-	-	-	-	-		
1	0	0	1	1	spiP56258	WECF_ECOLI	TDP-N-acetylglucosamine:lipid II N-acetylglucosaminyltransferase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP27835	WZYE_ECOLI	Probable ECA polymerase	-	-	-	-	-	-	-		
1	0	0	1	1	spiP75971	Y1142_ECOLI	1 Putative uncharacterized protein b1142	-	-	-	-	-	-	-		
1	0	0	1	1	spiP37049	YAEI_ECOLI	Phosphodiesterase YaeI	-	-	-	-	-	-	-		
1	0	0	1	1	spiP77489	YAGR_ECOLI	Putative xanthine dehydrogenase YagR molybdenum-binding subunit	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A08D3	YAIL_ECOLI	UPF0178 protein YaiI	-	-	-	-	-	-	-		
1	0	0	1	1	spiP47536	YAIP_ECOLI	Uncharacterized glycosyltransferase YaiP	-	-	-	-	-	-	-		
1	0	0	1	1	spiP77400	YBAT_ECOLI	Inner membrane transport protein YbaT	-	-	-	-	-	-	-		
1	0	0	1	1	spiP68661	YBCO_ECOLI	Putative nuclease YbcO	-	-	-	-	-	-	-		
1	0	0	1	1	spiP77174	YBDM_ECOLI	Uncharacterized protein YbdM	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0898	YBEY_ECOLI	Endonuclease YbeY	-	-	-	-	-	-	-		
1	0	0	1	1	spiP12994	YBHB_ECOLI	UPF0098 protein YbhB	-	-	-	-	-	-	-		
1	0	0	1	1	spiP75763	YBHI_ECOLI	Inner membrane protein YbhI	-	-	-	-	-	-	-		
1	0	0	1	1	spiP75788	YBIR_ECOLI	Inner membrane protein YbiR	-	-	-	-	-	-	-		
1	0	0	1	1	spiP75791	YBIU_ECOLI	Uncharacterized protein YbiU	-	-	-	-	-	-	-		
1	0	0	1	1	spiP75810	YBJJ_ECOLI	Inner membrane protein YbjJ	-	-	-	-	-	-	-		
1	0	0	1	1	spiP43674	YCAL_ECOLI	Uncharacterized metalloprotease YcaL	-	-	-	-	-	-	-		
1	0	0	1	1	spiP75836	YCAN_ECOLI	Uncharacterized HTH-type transcriptional regulator YcaN	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0B43	YCGI_ECOLI	Protein YcgI	-	-	-	-	-	-	-		
1	0	0	1	1	spiP30192	YCHG_ECOLI	Putative uncharacterized protein YchG	-	-	-	-	-	-	-		
1	0	0	1	1	spiP45848	YCIQ_ECOLI	Uncharacterized protein YcIQ	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76046	YCIJ_ECOLI	Uncharacterized protein YcIJ	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76097	YDCJ_ECOLI	Uncharacterized protein YdcJ	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76103	YDCO_ECOLI	Inner membrane protein YdcO	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76123	YDDK_ECOLI	Putative uncharacterized protein YddK	-	-	-	-	-	-	-		
1	0	0	1	1	spiP31130	YDEI_ECOLI	Uncharacterized protein YdeI	-	-	-	-	-	-	-		
1	0	0	1	1	spiP0A0FS7	YDIK_ECOLI	Putative transport protein YdiK	-	-	-	-	-	-	-		
1	0	0	1	1	spiP77649	YDIU_ECOLI	UPF0061 protein YdiU	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76206	YDIY_ECOLI	Uncharacterized protein YdiY	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76347	YEEI_ECOLI	Uncharacterized protein YeeI	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76369	YEEY_ECOLI	Uncharacterized HTH-type transcriptional regulator YeeY	-	-	-	-	-	-	-		
1	0	0	1	1	spiP33030	YEIR_ECOLI	Zinc-binding GTPase YeiR	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76481	YFBK_ECOLI	Uncharacterized protein YfbK	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76498	YFCO_ECOLI	Uncharacterized protein YfcO	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76501	YFCR_ECOLI	Uncharacterized fibrial-like protein YfcR	-	-	-	-	-	-	-		
1	0	0	1	1	spiP76569	YFGD_ECOLI	Uncharacterized protein YfgD	-	-	-	-	-	-	-		
1	0	0	1	1	spiP52125	YFJJ_ECOLI	Uncharacterized protein YfjJ	-	-	-	-	-	-	-		

Anzahl Motive					Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive			
S	M	W	Σ	Motiv 1			Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4	
1	0	0	1	1	spi76628	YGAY_ECOLI	Putative uncharacterized transporter YgaY	AAPPF	-	-	-	-	-	-		
1	0	0	1	1	spi55138	YGCE_ECOLI	Uncharacterized sugar kinase YgcE	CVPPG	-	-	-	-	-	-		
1	0	0	1	1	spiQ46905	YGCO_ECOLI	Ferredoxin-like protein YgcO	RYPPS	-	-	-	-	-	-		
1	0	0	1	1	spiP11664	YGGC_ECOLI	Uncharacterized protein YggC	CAPPG	-	-	-	-	-	-		
1	0	0	1	1	spiP45766	YHDW_ECOLI	Putative amino-acid ABC transporter-binding protein YhdW	YAPPV	-	-	-	-	-	-		
1	0	0	1	1	spiP64624	YHEO_ECOLI	Uncharacterized protein YheO	FVPPV	-	-	-	-	-	-		
1	0	0	1	1	spiP37621	YHHS_ECOLI	Uncharacterized MFS-type transporter YhhS	KRPPE	-	-	-	-	-	-		
1	0	0	1	1	spiP37626	YHIL_ECOLI	Uncharacterized protein YhhI	RIPPE	-	-	-	-	-	-		
1	0	0	1	1	spiP37641	YHJC_ECOLI	Uncharacterized HTH-type transcriptional regulator YhjC	DIPPG	-	-	-	-	-	-		
1	0	0	1	1	spiP37682	YIAU_ECOLI	Uncharacterized HTH-type transcriptional regulator YiaU	YAPPK	-	-	-	-	-	-		
1	0	0	1	1	spiP31455	YICJ_ECOLI	Inner membrane symporter YicJ	EAPPT	-	-	-	-	-	-		
1	0	0	1	1	spiP31446	YIDL_ECOLI	Inner membrane protein YidI	APPTT	-	-	-	-	-	-		
1	0	0	1	1	spiP31453	YIDP_ECOLI	Uncharacterized HTH-type transcriptional regulator YidP	PAPPA	-	-	-	-	-	-		
1	0	0	1	1	spiP31455	YIDR_ECOLI	Uncharacterized protein YidR	PAPPR	-	-	-	-	-	-		
1	0	0	1	1	spiP56259	YIFN_ECOLI	Putative uncharacterized protein YifN	RIPPE	-	-	-	-	-	-		
1	0	0	1	1	spiP77536	YKGF_ECOLI	Uncharacterized electron transport protein YkgF	QDPPS	-	-	-	-	-	-		
1	0	0	1	1	spiP77129	YLB_ECOLI	Uncharacterized protein YlbE	AGPPM	-	-	-	-	-	-		
1	0	0	1	1	spiP76227	YNJH_ECOLI	Uncharacterized protein YnjH	NVPPV	-	-	-	-	-	-		
1	0	0	1	1	spiP76280	YOBB_ECOLI	Uncharacterized protein YobB	PAPPD	-	-	-	-	-	-		
1	0	0	1	1	spiP77585	YPDE_ECOLI	Aminopeptidase YpdE	IAPPK	-	-	-	-	-	-		
1	0	0	1	1	spiP52143	YPJA_ECOLI	Uncharacterized outer membrane protein YpjA	YAPPG	-	-	-	-	-	-		
1	0	0	1	1	spiP63340	YQEG_ECOLI	Inner membrane transport protein YqeG	SIPPS	-	-	-	-	-	-		
1	0	0	1	1	spiQ46858	YQHG_ECOLI	Uncharacterized protein YqhG	AQPPS	-	-	-	-	-	-		
1	0	0	1	1	spiP69506	YTFE_ECOLI	Iron-sulfur cluster repair protein YtfE	TPPPE	-	-	-	-	-	-		
0	0	1	1	1	spiP04395	3MG2_ECOLI	DNA-3-methyladenine glycosylase 2	-	-	-	-	-	-	WQPPY		
0	0	1	1	1	spiP0A028	5FCL_ECOLI	5-formyltetrahydrofolate cyclo-ligase	-	-	-	-	-	-	TYPPV		
0	0	1	1	1	spiP46478	AAEX_ECOLI	Protein AaeX	-	-	-	-	-	-	SFFPI		
0	0	1	1	1	spiP31119	AAS_ECOLI	Bifunctional protein Aas	-	-	-	-	-	-	ILPPT		
0	1	0	1	1	spiP00509	AAT_ECOLI	Aspartate aminotransferase	-	-	SNPPA	-	-	-	-		
0	0	1	1	1	spiP46133	ABGT_ECOLI	p-aminobenzoyl-glutamate transport protein	-	-	-	-	-	-	IMPPM		
0	0	1	1	1	spiP75747	ABRB_ECOLI	Putative regulator AbrB	-	-	-	-	-	-	WPPPV		
0	0	1	1	1	spiP24182	ACCC_ECOLI	Biotin carboxylase	-	-	-	-	-	-	TVPPY		
0	1	0	1	1	spiP77580	ACDH_ECOLI	Acetaldehyde dehydrogenase	-	-	AEPLP	-	-	-	-		
0	0	1	1	1	spiP37623	ACPT_ECOLI	4-phosphopantetheinyl transferase AcpT	-	-	-	-	-	-	PLPPG		
0	0	1	1	1	spiP26646	ACUI_ECOLI	Probable acrylyl-CoA reductase Acui	-	-	-	-	-	-	MTPPE		
0	0	1	1	1	spiP31441	ADEC_ECOLI	Adenine deaminase	-	-	-	-	-	-	SAPPY		
0	0	1	1	1	spiP39451	ADHP_ECOLI	Alcohol dehydrogenase, propanol-preferring	-	-	-	-	-	-	GLPPE		
0	0	1	1	1	spiP60061	ADIC_ECOLI	Arginine/gmatine antiporter	-	-	-	-	-	-	LFPPI		
0	0	1	1	1	spiP42907	AGAS_ECOLI	Putative D-galactosamine-6-phosphate deaminase AgaS	-	-	-	-	-	-	ILPPS		
0	0	1	1	1	spiP19926	AGP_ECOLI	Glucose-1-phosphatase	-	-	-	-	-	-	CPPPY		
0	0	1	1	1	spiP08660	AK3_ECOLI	Lysine-sensitive aspartokinase 3	-	-	-	-	-	-	ENPLP		
0	0	1	1	1	spiP77256	AKRMG_ECOLI	1 NADH-specific methylglyoxal reductase	-	-	-	-	-	-	SVPPF		
0	1	0	1	1	spiP25553	ALDA_ECOLI	Lactaldehyde dehydrogenase	-	-	YYPTT	-	-	-	-		
0	1	0	1	1	spiP0A991	ALFI_ECOLI	Fructose-bisphosphate aldolase class 1	-	-	NRPPA	-	-	-	-		
0	0	1	1	1	spiP75713	ALLE_ECOLI	(S)-ureidoglycine aminohydrolase	-	-	-	-	-	-	YCPGG		
0	1	0	1	1	spiP63883	AMIC_ECOLI	N-acetylmuramoyl-L-alanine amidase AmiC	-	-	QVPPA	-	-	-	-		
0	0	1	1	1	spiP75820	AMID_ECOLI	N-acetylmuramoyl-L-alanine amidase AmiD	-	-	-	-	-	-	AVPPR		
0	0	1	1	1	spiP46883	AMO_ECOLI	Primary amine oxidase	-	-	-	-	-	-	LLPPD		
0	1	0	1	1	spiP68767	AMPA_ECOLI	Cytosolic aminopeptidase	-	-	NMPPN	-	-	-	-		
0	0	1	1	1	spiP13016	AMPD_ECOLI	1,6-anhydro-N-acetylmuramyl-L-alanine amidase AmpD	-	-	-	-	-	-	SLPPG		
0	0	1	1	1	spiP0A016	AMPG_ECOLI	Protein AmpG	-	-	-	-	-	-	YTPPF		
0	0	1	1	1	spiP25718	AMY1_ECOLI	Periplasmic alpha-amylase	-	-	-	-	-	-	NTPPQ		
0	0	1	1	1	spiP26612	AMY2_ECOLI	Cytoplasmic alpha-amylase	-	-	-	-	-	-	WLPPA		
0	1	0	1	1	spiP0A0F08	APBC_ECOLI	Iron-sulfur cluster carrier protein	-	-	DMPPG	-	-	-	-		
0	0	1	1	1	spiP0A0E22	APHA_ECOLI	Class B acid phosphatase	-	-	-	-	-	-	GRPPM		
0	0	1	1	1	spiP26458	APPB_ECOLI	Cytochrome bd-II ubiquinol oxidase subunit 2	-	-	-	-	-	-	LVPPV		
0	1	0	1	1	spiP05052	APPY_ECOLI	HTH-type transcriptional regulator AppY	-	-	EEPLP	-	-	-	-		
0	0	1	1	1	spiP08204	ARAB_ECOLI	Ribulokinase	-	-	-	-	-	-	GLPPA		
0	1	0	1	1	spiP02924	ARAF_ECOLI	L-arabinose-binding periplasmic protein	-	-	VEPPK	-	-	-	-		
0	0	1	1	1	spiP0A0E3	ARCB_ECOLI	Aerobic respiration control sensor protein ArcB	-	-	-	-	-	-	DLPLP		
0	0	1	1	1	spiP37306	ARCC_ECOLI	Carbamate kinase	-	-	-	-	-	-	QMPPV		
0	1	0	1	1	spiP77624	ARCM_ECOLI	Carbamate kinase-like protein Yahi	-	-	HFPPG	-	-	-	-		
0	0	1	1	1	spiP77398	ARNA_ECOLI	Bifunctional polymyxin resistance protein ArnA	-	-	-	-	-	-	HFPPF		
0	1	0	1	1	spiP77757	ARNC_ECOLI	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	-	-	QNPPV	-	-	-	-		
0	0	1	1	1	spiP76474	ARNF_ECOLI	Probable 4-amino-4-deoxy-L-arabinose-phosphoundecaprenol flippase subunit	-	-	-	-	-	-	HLPPM		
0	1	0	1	1	spiP76473	ARNT_ECOLI	Undecaprenyl-phosphate 4-amino-4-deoxy-L-arabinose arabinosyl transferase	-	-	AIPPA	-	-	-	-		
0	0	1	1	1	spiP07639	AROB_ECOLI	3-dehydroquinate synthase	-	-	-	-	-	-	TLPPR		
0	0	1	1	1	spiP0A0B91	AROG_ECOLI	Phospho-2-dehydro-3-deoxyheptone aldolase, Phe-sensitive	-	-	-	-	-	-	LLPPV		
0	0	1	1	1	spiP0A0D7	AROK_ECOLI	Shikimate kinase 1	-	-	-	-	-	-	ETPPR		
0	0	1	1	1	spiP0A0E28	AROM_ECOLI	Protein AroM	-	-	-	-	-	-	ILPPL		
0	0	1	1	1	spiP23325	ARPA_ECOLI	Ankyrin repeat protein A	-	-	-	-	-	-	ITPPL		
0	0	1	1	1	spiP0A0B96	ARSC_ECOLI	Arsenate reductase	-	-	-	-	-	-	ETPPT		
0	0	1	1	1	spiP30859	ARTL_ECOLI	Putative ABC transporter arginine-binding protein 2	-	-	-	-	-	-	SYPPF		
0	0	1	1	1	spiP30860	ARTI_ECOLI	ABC transporter arginine-binding protein 1	-	-	-	-	-	-	TYPPF		
0	1	0	1	1	spiP24242	ASCG_ECOLI	HTH-type transcriptional regulator AscG	-	-	FSPPK	-	-	-	-		
0	0	1	1	1	spiP76216	ASTB_ECOLI	N-succinylarginine dihydrolase	-	-	-	-	-	-	VIPPH		
0	0	1	1	1	spiP76217	ASTD_ECOLI	N-succinylglutamate 5-semialdehyde dehydrogenase	-	-	-	-	-	-	GLPPG		
0	1	0	1	1	spiP76461	ATOB_ECOLI	Acetyl-CoA acetyltransferase	-	-	GVPPA	-	-	-	-		
0	1	0	1	1	spiP0A0B98	ATP6_ECOLI	ATP synthase subunit a	-	-	QNPPA	-	-	-	-		
0	0	1	1	1	spiP0A0A0	ATPF_ECOLI	ATP synthase subunit b	-	-	-	-	-	-	VWPPD		
0	0	1	1	1	spiP69228	BAER_ECOLI	Transcriptional regulatory protein BaeR	-	-	-	-	-	-	QTPPD		
0	1	0	1	1	spiP0A903	BAMC_ECOLI	Outer membrane protein assembly factor BamC	-	-	IRPPA	-	-	-	-		
0	1	0	1	1	spiP0A0C02	BAMD_ECOLI	Outer membrane protein assembly factor BamD	-	-	DNPPN	-	-	-	-		
0	0	1	1	1	spiP0A0E3	BARA_ECOLI	Signal transduction histidine-protein kinase BarA	-	-	-	-	-	-	ALPPA		
0	0	1	1	1	spiP28246	BCR_ECOLI	Bicyclomyacin resistance protein	-	-	-	-	-	-	TLPPE		
0	1	0	1	1	spiP37659	BCSG_ECOLI	Cellulose biosynthesis protein BcsG	-	-	TAPPT	-	-	-	-		
0	1	0	1	1	spiP39334	BDCR_ECOLI	HTH-type transcriptional repressor BdcR	-	-	INPPS	-	-	-	-		
0	0	1	1	1	spiP33363	BGLX_ECOLI	Periplasmic beta-glucosidase	-	-	-	-	-	-	YMPPY		
0	0	1	1	1	spiP12995	BIOA_ECOLI	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase	-	-	-	-	-	-	LMPPY		
0	0	1	1	1	spiP12996	BIOB_ECOLI	Biotin synthase	-	-	-	-	-	-	PTPPE		
0	0	1	1	1	spiP13000	BIOD1_ECOLI	1 ATP-dependent dethiobiotin synthetase BioD 1	-	-	-	-	-	-	VTPPG		
0	0	1	1	1	spiP0A901	BLC_ECOLI	Outer membrane lipoprotein Blc	-	-	-	-	-	-	TPPPR		
0	1	0	1	1	spiP0A0B2	BOLA_ECOLI	DNA-binding transcriptional regulator BolA	-	-	ASPPC	-	-	-	-		
0	0	1	1	1	spiP39396	BTST_ECOLI	Pyruvate/proton symporter BtsT	-	-	-	-	-	-	NTPPA		
0	0	1	1	1	spiP0A0B5	C561_ECOLI	Cytochrome b561	-	-	-	-	-	-	PTPPI		
0	0	1	1	1	spiP31551	CAID_ECOLI	Carnitinyl-CoA dehydratase	-	-	-	-	-	-	ILPPA		
0	0	1	1	1	spiP00968	CARB_ECOLI	Carbamoyl-phosphate synthase large chain	-	-	-	-	-	-	VIPPY		
0	0	1	1	1	spiP45956	CAS2_ECOLI	CRISPR-associated endonuclease Cas2	-	-	-	-	-	-	NVPPR		
0	0	1	1	1	spiP06961	CCA_ECOLI	Multifunctional CCA protein	-	-	-	-	-	-	LTPEE		
0	0	1	1	1	spiP0A0B1	CCMC_ECOLI	Heme exporter protein C	-	-	-	-	-	-	AIPPR		
0	0	1	1	1	spiP69490	CCME_ECOLI	Cytochrome c-type biogenesis protein CcmE	-	-	-	-	-	-	YTPPE		
0	0	1	1	1	spiP33927	CCMF_ECOLI	Cytochrome c-type biogenesis protein CcmF	-	-	-	-	-	-	FHPPL		
0	0	1	1	1	spiP0A0F6	CDD_ECOLI	Cytidine deaminase	-	-	-	-	-	-	TLPPL		
0	0	1	1	1	spiP0A0G1	CDSA_ECOLI	Phosphatidate cytidyltransferase	-	-	-	-	-	-	LLPPV		
0	0	1	1	1	spiP07363	CHEA_ECOLI	Chemotaxis protein CheA	-	-	-	-	-	-	STPPV		
0	1	0	1	1	spiP07330	CHEB_ECOLI	Chemotaxis response regulator protein-glutamate methyltransferase	-	-	HMPPG	-	-	-	-		
0	0	1	1	1	spiP77390	CITC_ECOLI	[Citrate [pro-3S]-lyase] ligase	-	-	-	-	-	-	SAPPI		

Anzahl Motive					Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive							
S	M	W	Σ	Motiv 1			Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4					
0	0	1	1	spiP0A6G5	CITX_ECOLI	Apo-citrate lyase phosphoribosyl-dephospho-CoA transferase	-	-	-	-	-	-	-	SLPPR	-	-	-	-		
0	1	0	1	spiP0A6H1	CLPX_ECOLI	ATP-dependent Clp protease ATP-binding subunit ClpX	-	-	-	-	-	-	-	AVPPQ	-	-	-	-		
0	0	1	1	spiP0A8Q0	COABC_ECOLI	1 Coenzyme A biosynthesis bifunctional protein CoaBC	-	-	-	-	-	-	-	PTPPF	-	-	-	-		
0	0	1	1	spiP25524	CODA_ECOLI	Cytosine deaminase	-	-	-	-	-	-	-	VIPPF	-	-	-	-		
0	0	1	1	spiP0A8A2	CODB_ECOLI	Cytosine permease	-	-	-	-	-	-	-	AIPPV	-	-	-	-		
0	0	1	1	spiP75952	COMR_ECOLI	HTH-type transcriptional repressor ComR	-	-	-	-	-	-	-	EIPPH	-	-	-	-		
0	0	1	1	spiQ59385	COPA_ECOLI	Copper-exporting P-type ATPase	-	-	-	-	-	-	-	LTPTT	-	-	-	-		
0	0	1	1	spiP0A8I4	CORA_ECOLI	Magnesium transport protein CorA	-	-	-	-	-	-	-	FLPPT	-	-	-	-		
0	1	0	1	spiP08331	CPDB_ECOLI	2,3-cyclic-nucleotide 2'-phosphodiesterase/3'-nucleotidase	-	-	-	-	-	-	-	VVPPQ	-	-	-	-		
0	0	1	1	spiP0A8E5	CPXP_ECOLI	Periplasmic protein CpxP	-	-	-	-	-	-	-	-	EQPPV	-	-	-	-	
0	0	1	1	spiP0AACP1	CRA_ECOLI	Catabolite repressor/activator	-	-	-	-	-	-	-	-	SLPPE	-	-	-	-	
0	0	1	1	spiP13518	CSRD_ECOLI	RNase E specificity factor CsrD	-	-	-	-	-	-	-	-	EWPPR	-	-	-	-	
0	0	1	1	spiP0A9G4	CUER_ECOLI	HTH-type transcriptional regulator CueR	-	-	-	-	-	-	-	-	VTPPP	-	-	-	-	
0	1	0	1	spiP77239	CUSB_ECOLI	Cation efflux system protein CusB	-	-	-	-	-	-	-	-	AEPPA	-	-	-	-	
0	0	1	1	spiP0A8K2	CYDB_ECOLI	Cytochrome bd-1 ubiquinol oxidase subunit 2	-	-	-	-	-	-	-	-	FVPLP	-	-	-	-	
0	0	1	1	spiP23886	CYDC_ECOLI	ATP-binding/permease protein CydC	-	-	-	-	-	-	-	-	LMPLL	-	-	-	-	
0	0	1	1	spiP27111	CYNR_ECOLI	HTH-type transcriptional regulator CynR	-	-	-	-	-	-	-	-	LAPPL	-	-	-	-	
0	0	1	1	spiP0A9D4	CYSE_ECOLI	Serine acetyltransferase	-	-	-	-	-	-	-	-	PVPPH	-	-	-	-	
0	1	0	1	spiP17846	CYSI_ECOLI	Sulfite reductase [NADPH] hemoprotein beta-component	-	-	-	-	-	-	-	-	VIPPO	-	-	-	-	
0	0	1	1	spiP16700	CYSP_ECOLI	Thiosulfate-binding protein	-	-	-	-	-	-	-	-	LNPPF	-	-	-	-	
0	0	1	1	spiP0A6J3	CYSZ_ECOLI	Sulfate transporter CysZ	-	-	-	-	-	-	-	-	ATPPD	-	-	-	-	
0	1	0	1	spiP24228	DACB_ECOLI	D-alanyl-D-alanine carboxypeptidase DacB	-	-	-	-	-	-	-	-	SAPPA	-	-	-	-	
0	0	1	1	spiP11557	DAMX_ECOLI	Cell division protein DamX	-	-	-	-	-	-	-	-	SLPPI	-	-	-	-	
0	1	0	1	spiP0A9D8	DAPD_ECOLI	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	-	-	-	-	-	-	-	-	VVPPA	-	-	-	-	
0	1	0	1	spiP0A6K1	DAPF_ECOLI	Diaminopimelate epimerase	-	-	-	-	-	-	-	-	VEPPY	-	-	-	-	
0	0	1	1	spiP28248	DCD_ECOLI	dCTP deaminase	-	-	-	-	-	-	-	-	PRPPV	-	-	-	-	
0	1	0	1	spiP0AED9	DCM_ECOLI	DNA-cytosine methyltransferase	-	-	-	-	-	-	-	-	PKPPA	-	-	-	-	
0	0	1	1	spiP24171	DCP_ECOLI	Dipeptidyl carboxypeptidase	-	-	-	-	-	-	-	-	-	AIPPR	-	-	-	-
0	0	1	1	spiP0A8N9	DCUB_ECOLI	Anaerobic C4-dicarboxylate transporter DcuB	-	-	-	-	-	-	-	-	GKPPV	-	-	-	-	
0	0	1	1	spiP45428	DCUD_ECOLI	Putative cryptic C4-dicarboxylate transporter DcuD	-	-	-	-	-	-	-	-	NVPLL	-	-	-	-	
0	1	0	1	spiP77308	DDPB_ECOLI	Probable D,D-dipeptide transport system permease protein DdpB	-	-	-	-	-	-	-	-	LDPPT	-	-	-	-	
0	1	0	1	spiP77790	DDPX_ECOLI	D-alanyl-D-alanine dipeptidase	-	-	-	-	-	-	-	-	SVPPA	-	-	-	-	
0	1	0	1	spiP0A6K6	DEOB_ECOLI	Phosphopentomutase	-	-	-	-	-	-	-	-	VEPPA	-	-	-	-	
0	0	1	1	spiQ6BF17	DGOD_ECOLI	D-galactonate dehydratase	-	-	-	-	-	-	-	-	-	RLPPR	-	-	-	-
0	0	1	1	spiP31459	DGOK_ECOLI	2-dehydro-3-deoxygalactonokinase	-	-	-	-	-	-	-	-	GLPPQ	-	-	-	-	
0	0	1	1	spiP0A8A76	DGOT_ECOLI	D-galactonate transporter	-	-	-	-	-	-	-	-	YQPPR	-	-	-	-	
0	0	1	1	spiP15723	DGTP_ECOLI	Deoxyguanosinetriphosphate triphosphohydrolase	-	-	-	-	-	-	-	-	GNPPF	-	-	-	-	
0	0	1	1	spiP76016	DHAR_ECOLI	PTS-dependent dihydroxyacetone kinase operon regulatory protein	-	-	-	-	-	-	-	-	TIPPL	-	-	-	-	
0	0	1	1	spiP06966	DICA_ECOLI	HTH-type transcriptional regulator DicA	-	-	-	-	-	-	-	-	PTPPV	-	-	-	-	
0	0	1	1	spiP31680	DJLA_ECOLI	Co-chaperone protein DjaA	-	-	-	-	-	-	-	-	GLPPE	-	-	-	-	
0	1	0	1	spiP37672	DLGD_ECOLI	2,3-diketo-L-gulonate reductase	-	-	-	-	-	-	-	-	VMPPW	-	-	-	-	
0	0	1	1	spiP42628	DLST_ECOLI	Probable serine transporter	-	-	-	-	-	-	-	-	SLPPL	-	-	-	-	
0	0	1	1	spiP0A8E8	DMA_ECOLI	DNA adenine methylase	-	-	-	-	-	-	-	-	CDPPY	-	-	-	-	
0	0	1	1	spiP69853	DMSD_ECOLI	Tat proofreading chaperone DmsD	-	-	-	-	-	-	-	-	PSPPW	-	-	-	-	
0	0	1	1	spiP0A8C0	DNAB_ECOLI	Replicative DNA helicase	-	-	-	-	-	-	-	-	KVPPH	-	-	-	-	
0	1	0	1	spiP0A6Y8	DNAK_ECOLI	Chaperone protein DnaK	-	-	-	-	-	-	-	-	MAPPQ	-	-	-	-	
0	0	1	1	spiP77510	DPIB_ECOLI	Sensor histidine kinase DpiB	-	-	-	-	-	-	-	-	QLPPG	-	-	-	-	
0	1	0	1	spiP23847	DPPA_ECOLI	Periplasmic dipeptide transport protein	-	-	-	-	-	-	-	-	LIPPT	-	-	-	-	
0	1	0	1	spiP0A8G6	DSBC_ECOLI	Thiol-disulfide interchange protein DsbC	-	-	-	-	-	-	-	-	YQPPK	-	-	-	-	
0	0	1	1	spiP08555	DSDX_ECOLI	D-serine transporter DsdX	-	-	-	-	-	-	-	-	-	VVPPH	-	-	-	-
0	1	0	1	spiP36837	DTPB_ECOLI	Dipeptide and tripeptide permease B	-	-	-	-	-	-	-	-	CYPPK	-	-	-	-	
0	0	1	1	spiP32695	DUSA_ECOLI	tRNA-dihydrouridine(20/20a) synthase	-	-	-	-	-	-	-	-	-	EIPPL	-	-	-	-
0	0	1	1	spiP0A8T5	DUSB_ECOLI	tRNA-dihydrouridine synthase B	-	-	-	-	-	-	-	-	LLPPL	-	-	-	-	
0	0	1	1	spiP77488	DXS_ECOLI	1-deoxy-D-xylulose-5-phosphate synthase	-	-	-	-	-	-	-	-	GVPPI	-	-	-	-	
0	0	1	1	spiP31125	EAMA_ECOLI	Probable amino-acid metabolite efflux pump	-	-	-	-	-	-	-	-	NMPLL	-	-	-	-	
0	0	1	1	spiP06846	EBGR_ECOLI	HTH-type transcriptional regulator EbgR	-	-	-	-	-	-	-	-	TFPPL	-	-	-	-	
0	0	1	1	spiP0A8F6	EDD_ECOLI	Phosphogluconate dehydratase	-	-	-	-	-	-	-	-	LMPLL	-	-	-	-	
0	0	1	1	spiP31545	EFEB_ECOLI	Deferrochelate/oxidase EfeB	-	-	-	-	-	-	-	-	RLPPL	-	-	-	-	
0	1	0	1	spiP0A8B4	EFE0_ECOLI	Iron uptake system component EfeO	-	-	-	-	-	-	-	-	AFPPS	-	-	-	-	
0	0	1	1	spiP0A6N4	EFP_ECOLI	Elongation factor P	-	-	-	-	-	-	-	-	-	VTPPN	-	-	-	-
0	0	1	1	spiP75856	ELFD_ECOLI	Probable fibrillar chaperone protein EIfD	-	-	-	-	-	-	-	-	ITPPL	-	-	-	-	
0	1	0	1	spiP0C960	EMTA_ECOLI	Endo-type membrane-bound lytic murein transglycosylase A	-	-	-	-	-	-	-	-	TNPPW	-	-	-	-	
0	1	0	1	spiP25736	END1_ECOLI	Endonuclease-1	-	-	-	-	-	-	-	-	AEPPA	-	-	-	-	
0	1	0	1	spiP10378	ENTE_ECOLI	Enterobactin synthase component E	-	-	-	-	-	-	-	-	LVPPA	-	-	-	-	
0	1	0	1	spiP30845	EPTA_ECOLI	Phosphoethanolamine transferase EptA	-	-	-	-	-	-	-	-	RYPPQ	-	-	-	-	
0	0	1	1	spiP77218	EUTD_ECOLI	Ethanolamine utilization protein EutD	-	-	-	-	-	-	-	-	-	KTPPD	-	-	-	-
0	0	1	1	spiP77445	EUTE_ECOLI	Ethanolamine utilization protein EutE	-	-	-	-	-	-	-	-	GNPPV	-	-	-	-	
0	0	1	1	spiP76553	EUTG_ECOLI	Ethanolamine utilization protein EutG	-	-	-	-	-	-	-	-	SVPPV	-	-	-	-	
0	1	0	1	spiP77277	EUTI_ECOLI	Ethanolamine utilization protein EutJ	-	-	-	-	-	-	-	-	SFPPG	-	-	-	-	
0	0	1	1	spiP0A8E0	EXOX_ECOLI	Exodeoxyribonuclease 10	-	-	-	-	-	-	-	-	-	QTPPG	-	-	-	-
0	0	1	1	spiP42593	FADH_ECOLI	2,4-dienoyl-CoA reductase	-	-	-	-	-	-	-	-	RTPLP	-	-	-	-	
0	0	1	1	spiP38135	FADK_ECOLI	Medium-chain fatty-acid-CoA ligase	-	-	-	-	-	-	-	-	TLPPG	-	-	-	-	
0	1	0	1	spiP0A8V6	FADR_ECOLI	Fatty acid metabolism regulator protein	-	-	-	-	-	-	-	-	RFPPG	-	-	-	-	
0	0	1	1	spiP0A8A3	FDNH_ECOLI	Formate dehydrogenase, nitrate-inducible, iron-sulfur subunit	-	-	-	-	-	-	-	-	-	ITPPS	-	-	-	-
0	1	0	1	spiQ47208	FDRA_ECOLI	Protein FdrA	-	-	-	-	-	-	-	-	SKPPA	-	-	-	-	
0	0	1	1	spiP15029	FECD_ECOLI	Fe(3+) dicitrate transport system permease protein FecD	-	-	-	-	-	-	-	-	-	IHPPL	-	-	-	-
0	0	1	1	spiP0A8E6	FEPB_ECOLI	Ferrienterobactin-binding periplasmic protein	-	-	-	-	-	-	-	-	KLPPQ	-	-	-	-	
0	0	1	1	spiP26266	FEPE_ECOLI	Ferrienterobactin transport protein FepE	-	-	-	-	-	-	-	-	VTPPE	-	-	-	-	
0	0	1	1	spiP13039	FES_ECOLI	Enterochelin esterase	-	-	-	-	-	-	-	-	QLPPA	-	-	-	-	
0	0	1	1	spiP07822	FHUD_ECOLI	Iron(3+)-hydroxamate-binding protein FluH	-	-	-	-	-	-	-	-	SEPLL	-	-	-	-	
0	0	1	1	spiP39405	FHFU_ECOLI	Ferric iron reductase protein FhfU	-	-	-	-	-	-	-	-	MVPLL	-	-	-	-	
0	1	0	1	spiP39264	FIML_ECOLI	Fimbrin-like protein FimI	-	-	-	-	-	-	-	-	NRPPA	-	-	-	-	
0	0	1	1	spiP0A9L3	FKBB_ECOLI	FKBP-type 22 kDa peptidyl-prolyl cis-trans isomerase	-	-	-	-	-	-	-	-	-	SIPPF	-	-	-	-
0	0	1	1	spiP0A8W9	FLGB_ECOLI	Flagellar basal body rod protein FlgB	-	-	-	-	-	-	-	-	LTPTT	-	-	-	-	
0	0	1	1	spiP75937	FLGE_ECOLI	Flagellar hook protein FlgE	-	-	-	-	-	-	-	-	GTPTT	-	-	-	-	
0	0	1	1	spiP0A6S3	FLGI_ECOLI	Flagellar P-ring protein	-	-	-	-	-	-	-	-	SLPPF	-	-	-	-	
0	0	1	1	spiP76297	FLHE_ECOLI	Flagellar protein FlhE	-	-	-	-	-	-	-	-	LIPPL	-	-	-	-	
0	1	0	1	spiP31068	FLIH_ECOLI	Flagellar assembly protein FlhI	-	-	-	-	-	-	-	-	LAPPO	-	-	-	-	
0	0	1	1	spiP52614																



Anzahl Motive					Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive				
S	M	W	Σ	Motiv 1			Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4		
0	0	1	1	spi28306	MLTG_ECOLI	Endolytic murein transglycosylase	-	-	-	-	-	-	-	GLPPG	-	-	-
0	0	1	1	spi54746	MNGB_ECOLI	Mannosylglycerate hydrolase	-	-	-	-	-	-	-	-	KMPVV	-	-
0	0	1	1	spi25522	MNME_ECOLI	tRNA modification GTPase MnmE	-	-	-	-	-	-	-	-	ATPPG	-	-
0	0	1	1	spi30748	MOAD_ECOLI	Molybdopterin synthase sulfur carrier subunit	-	-	-	-	-	-	-	-	FTPPV	-	-
0	0	1	1	spi0A0F01	MODB_ECOLI	Molybdenum transport system permease protein ModB	-	-	-	-	-	-	-	-	VLLPV	-	-
0	1	0	1	spi09833	MODC_ECOLI	Molybdenum import ATP-binding protein ModC	-	-	-	-	-	LQPPQ	-	-	-	-	-
0	1	0	1	spi33345	MOLR_ECOLI	Putative molybdate metabolism regulator	-	-	-	-	-	VFPPE	-	-	-	-	-
0	0	1	1	spi37773	MPL_ECOLI	ligase Mpl	-	-	-	-	-	-	-	-	VYPPM	-	-
0	0	1	1	spi77348	MPPA_ECOLI	Periplasmic murein peptide-binding protein	-	-	-	-	-	-	-	-	YTPPG	-	-
0	0	1	1	spi0A0CR9	MPRA_ECOLI	Transcriptional repressor MprA	-	-	-	-	-	-	-	-	VLLPQ	-	-
0	0	1	1	spi0A0A9X4	MREB_ECOLI	Rod shape-determining protein MreB	-	-	-	-	-	-	-	-	QCPEE	-	-
0	0	1	1	spi0A742	MSCL_ECOLI	Large-conductance mechanosensitive channel	-	-	-	-	-	-	-	-	IMPLL	-	-
0	0	1	1	spi39285	MSCM_ECOLI	Minicconductance mechanosensitive channel MscM	-	-	-	-	-	-	-	-	PFPFF	-	-
0	0	1	1	spi0A744	MSRA_ECOLI	Peptide methionine sulfoxide reductase MsrA	-	-	-	-	-	-	-	-	CLPPE	-	-
0	0	1	1	spi0A746	MSRB_ECOLI	Peptide methionine sulfoxide reductase MsrB	-	-	-	-	-	-	-	-	TEPFF	-	-
0	0	1	1	spi25738	MSYB_ECOLI	Acidic protein MsyB	-	-	-	-	-	-	-	-	LEPLL	-	-
0	0	1	1	spi40874	MTOX_ECOLI	N-methyl-L-tryptophan oxidase	-	-	-	-	-	-	-	-	HMPPH	-	-
0	1	0	1	spi0A0AD2	MTR_ECOLI	Tryptophan-specific transport protein	-	-	-	-	-	FAPPV	-	-	-	-	-
0	0	1	1	spi14900	MURD_ECOLI	UDP-N-acetylmuramoylalanine-D-glutamate ligase	-	-	-	-	-	-	-	-	MTPPG	-	-
0	0	1	1	spi0A0F16	MURJ_ECOLI	Lipid II flippase MurJ	-	-	-	-	-	-	-	-	FNPPV	-	-
0	0	1	1	spi76535	MURQ_ECOLI	N-acetylmuramic acid 6-phosphate esterase	-	-	-	-	-	-	-	-	ECPPT	-	-
0	0	1	1	spi06722	MUTH_ECOLI	DNA mismatch repair protein MthH	-	-	-	-	-	-	-	-	LSPPE	-	-
0	0	1	1	spi23909	MUTS_ECOLI	DNA mismatch repair protein MutS	-	-	-	-	-	-	-	-	DTPPV	-	-
0	1	0	1	spi08337	MUTT_ECOLI	8-oxo-dGTP diphosphatase	-	-	-	-	-	DFPPA	-	-	-	-	-
0	1	0	1	spi11458	NADA_ECOLI	Quinolinate synthase A	-	-	-	-	-	PFPFK	-	-	-	-	-
0	0	1	1	spi10902	NADB_ECOLI	L-aspartate oxidase	-	-	-	-	-	-	-	-	-	-	-
0	0	1	1	spi0A752	NADD_ECOLI	Nicotinate-nucleotide adenyltransferase	-	-	-	-	-	-	-	-	TLPPW	-	-
0	0	1	1	spi18843	NADE_ECOLI	NH(3)-dependent NAD(+) synthetase	-	-	-	-	-	-	-	-	NVPPH	-	-
0	0	1	1	spi0A0F18	NAGA_ECOLI	N-acetylglucosamine-6-phosphate deacetylase	-	-	-	-	-	-	-	-	RRPPI	-	-
0	1	0	1	spi33937	NAPA_ECOLI	Periplasmic nitrate reductase	-	-	-	-	-	NRPPQ	-	-	ELPPE	-	-
0	0	1	1	spi0A0BL3	NAPB_ECOLI	Periplasmic nitrate reductase, electron transfer subunit	-	-	-	-	-	-	-	-	NQPPM	-	-
0	0	1	1	spi11349	NARH_ECOLI	Respiratory nitrate reductase 1 beta chain	-	-	-	-	-	-	-	-	YVPLL	-	-
0	1	0	1	spi27896	NARQ_ECOLI	Nitrate/nitrite sensor protein NarQ	-	-	-	-	-	DSPLL	-	-	-	-	-
0	0	1	1	spi0A0FA2	NARX_ECOLI	Nitrate/nitrite sensor protein NarX	-	-	-	-	-	-	-	-	QLPPR	-	-
0	0	1	1	spi19318	NARY_ECOLI	Respiratory nitrate reductase 2 beta chain	-	-	-	-	-	-	-	-	YVPLL	-	-
0	0	1	1	spi33920	NDPA_ECOLI	Nucleoid-associated protein YejK	-	-	-	-	-	-	-	-	GTPPN	-	-
0	0	1	1	spi0A0DL1	NEPI_ECOLI	Purine ribonucleoside efflux pump NepI	-	-	-	-	-	-	-	-	LVPPR	-	-
0	0	1	1	spi63020	NFUA_ECOLI	Fe/S biogenesis protein NfuA	-	-	-	-	-	-	-	-	YCPDD	-	-
0	0	1	1	spi33591	NIKB_ECOLI	Nickel transport system permease protein NkB	-	-	-	-	-	-	-	-	NLPPT	-	-
0	1	0	1	spi33594	NIKE_ECOLI	Nickel import ATP-binding protein NtKE	-	-	-	-	-	KRPPQ	-	-	-	-	-
0	1	0	1	spi76241	NIMR_ECOLI	HTH-type transcriptional regulator NimR	-	-	-	-	-	MVPPQ	-	-	-	-	-
0	0	1	1	spi76242	NIMT_ECOLI	2-nitroimidazole transporter	-	-	-	-	-	-	-	-	CGPPL	-	-
0	0	1	1	spi08201	NIRB_ECOLI	Nitrite reductase (NADH) large subunit	-	-	-	-	-	-	-	-	WIPPI	-	-
0	0	1	1	spi37013	NORR_ECOLI	Anaerobic nitric oxide reductase transcription regulator NorR	-	-	-	-	-	-	-	-	SVPLL	-	-
0	1	0	1	spi0A0AG6	NQOR_ECOLI	NAD(P)H dehydrogenase (quinone)	-	-	-	-	-	TMPPQ	-	-	-	-	-
0	0	1	1	spi32711	NRFF_ECOLI	Formate-dependent nitrite reductase complex subunit NrfF	-	-	-	-	-	-	-	-	YNPLL	-	-
0	0	1	1	spi0A0B5	NTRB_ECOLI	Sensory histidine kinase/phosphatase NtrB	-	-	-	-	-	-	-	-	GIPPH	-	-
0	0	1	1	spi0A0B8	NTRC_ECOLI	DNA-binding transcriptional regulator NtrC	-	-	-	-	-	-	-	-	HLPLL	-	-
0	0	1	1	spi0A0C7	NUOB_ECOLI	NADH-quinone oxidoreductase subunit B	-	-	-	-	-	-	-	-	GCPPR	-	-
0	1	0	1	spi33599	NUOCD_ECOLI	1 NADH-quinone oxidoreductase subunit C/D	-	-	-	-	-	TPPPK	-	-	-	-	-
0	0	1	1	spi31979	NUOF_ECOLI	NADH-quinone oxidoreductase subunit F	-	-	-	-	-	-	-	-	SKPPF	-	-
0	0	1	1	spi0A0FD4	NUOH_ECOLI	NADH-quinone oxidoreductase subunit H	-	-	-	-	-	-	-	-	LLPFF	-	-
0	0	1	1	spi0A0FD6	NUOI_ECOLI	NADH-quinone oxidoreductase subunit I	-	-	-	-	-	-	-	-	YLPPR	-	-
0	0	1	1	spi33607	NUOL_ECOLI	NADH-quinone oxidoreductase subunit L	-	-	-	-	-	-	-	-	IVPLL	-	-
0	0	1	1	spi0A0FF4	NUPG_ECOLI	Nucleoside permease NupG	-	-	-	-	-	-	-	-	DFPPI	-	-
0	1	0	1	spi0A0FG6	ODO2_ECOLI	component of 2-oxoglutarate dehydrogenase complex	-	-	-	-	-	INPPQ	-	-	-	-	-
0	0	1	1	spi0A0A15	OMPW_ECOLI	Outer membrane protein W	-	-	-	-	-	-	-	-	HLPPT	-	-
0	0	1	1	spi27298	OPDA_ECOLI	Oligopeptidase A	-	-	-	-	-	-	-	-	ELPFF	-	-
0	0	1	1	spi75920	OPGC_ECOLI	Glucans biosynthesis protein C	-	-	-	-	-	-	-	-	VYPPI	-	-
0	0	1	1	spi0A0FH2	OPPB_ECOLI	Oligopeptide transport system permease protein	-	-	-	-	-	-	-	-	TLPPE	-	-
0	0	1	1	spi76083	PAAH_ECOLI	3-hydroxyadipyl-CoA dehydrogenase	-	-	-	-	-	-	-	-	VCPQQ	-	-
0	1	0	1	spi77455	PAAZ_ECOLI	Bifunctional protein PaaZ	-	-	-	-	-	FFPPT	-	-	-	-	-
0	0	1	1	spi20083	PARE_ECOLI	DNA topoisomerase 4 subunit B	-	-	-	-	-	-	-	-	ALPPL	-	-
0	0	1	1	spi22259	PCKA_ECOLI	Phosphoenolpyruvate carboxylase (ATP)	-	-	-	-	-	-	-	-	VLPVV	-	-
0	0	1	1	spi77473	PDEB_ECOLI	Probable cyclic di-GMP phosphodiesterase PdeB	-	-	-	-	-	-	-	-	ESPPD	-	-
0	1	0	1	spi32701	PDEC_ECOLI	Probable cyclic di-GMP phosphodiesterase PdeC	-	-	-	-	-	AMPPQ	-	-	-	-	-
0	1	0	1	spi75995	PDEG_ECOLI	Probable cyclic di-GMP phosphodiesterase PdeG	-	-	-	-	-	NPPPL	-	-	-	-	-
0	0	1	1	spi37646	PDEH_ECOLI	Cyclic di-GMP phosphodiesterase PdeH	-	-	-	-	-	-	-	-	RLPPD	-	-
0	0	1	1	spi77334	PDER_ECOLI	Cyclic di-GMP phosphodiesterase PdeR	-	-	-	-	-	-	-	-	LIPPL	-	-
0	0	1	1	spi0A0CL9	PDHR_ECOLI	Pyruvate dehydrogenase complex repressor	-	-	-	-	-	-	-	-	KLPEE	-	-
0	0	1	1	spi05459	PDXB_ECOLI	Erythronate-4-phosphate dehydrogenase	-	-	-	-	-	-	-	-	CDPPR	-	-
0	1	0	1	spi25906	PDXL_ECOLI	Pyridoxine 4-dehydrogenase	-	-	-	-	-	FGPPR	-	-	-	-	-
0	0	1	1	spi40191	PDXK_ECOLI	Pyridoxine/pyridoxal/pyridoxamine kinase	-	-	-	-	-	-	-	-	ILPPL	-	-
0	1	0	1	spi77150	PDXY_ECOLI	Pyridoxal kinase PdxY	-	-	-	-	-	VMPPS	-	-	-	-	-
0	0	1	1	spiQ57083	PERR_ECOLI	HTH-type transcriptional regulator PerR	-	-	-	-	-	-	-	-	MTPPN	-	-
0	1	0	1	spi0A0A9N4	PFLA_ECOLI	Pyruvate formate-lyase 1-activating enzyme	-	-	-	-	-	VKPPK	-	-	-	-	-
0	1	0	1	spi75906	PGAB_ECOLI	Poly-beta-1,6-N-acetyl-D-glucosamine N-deacetylase	-	-	-	-	-	FPPQ	-	-	-	-	-
0	1	0	1	spi77366	PGMB_ECOLI	Beta-phosphoglucomutase	-	-	-	-	-	GVPPQ	-	-	-	-	-
0	0	1	1	spi0A0A924	PGPB_ECOLI	Phosphatidylglycerophosphatase B	-	-	-	-	-	-	-	-	LTPPA	-	-
0	0	1	1	spi16688	PHNJ_ECOLI	Alpha-D-ribose 1-methylphosphonate 5-phosphate C-P lyase	-	-	-	-	-	-	-	-	AVPPF	-	-
0	0	1	1	spi23836	PHOP_ECOLI	Transcriptional regulatory protein PhoP	-	-	-	-	-	-	-	-	SLPFF	-	-
0	0	1	1	spi000914	PHR_ECOLI	Deoxyribodipyrimidine photo-lyase	-	-	-	-	-	-	-	-	ILPPG	-	-
0	0	1	1	spi77170	PINQ_ECOLI	Serine recombinase PinQ	-	-	-	-	-	-	-	-	GRPPV	-	-
0	0	1	1	spi0A0D10	PINR_ECOLI	Serine recombinase PinR	-	-	-	-	-	-	-	-	GRPPV	-	-
0	0	1	1	spi0A0AF7	PITA_ECOLI	Low-affinity inorganic phosphate transporter 1	-	-	-	-	-	-	-	-	KKPPF	-	-
0	0	1	1	spi76594	PKA_ECOLI	Protein lysine acetyltransferase Pka	-	-	-	-	-	-	-	-	ALPPL	-	-
0	0	1	1	spi67080	PLPHP_ECOLI	1 Pyridoxal phosphate homeostasis protein	-	-	-	-	-	-	-	-	ELPPL	-	-
0	0	1	1	spi05055	PNP_ECOLI	Polyribonucleotide nucleotidyltransferase	-	-	-	-	-	-	-	-	NFPPI	-	-
0	0	1	1	spi33644	POLOX_ECOLI	1 Polyphenol oxidase	-	-	-	-	-	-	-	-	SLPPI	-	-
0	1	0	1	spi31133	POTF_ECOLI	Putrescine-binding periplasmic protein	-	-	-	-	-	IVYPA	-	-	-	-	-
0	1	0	1	spi07003	POXB_ECOLI	Pyruvate dehydrogenase [ubiquinone]	-	-	-	-	-	AIPPQ	-	-	-	-	-
0	1	0	1	spi08372	PPDC_ECOLI	Prepilin peptidase-dependent protein C	-	-	-	-	-	ISPPA	-	-	-	-	-
0	0	1	1	spi0A0FL3	PPIA_ECOLI	Peptidyl-prolyl-cis-trans isomerase A	-	-	-	-	-	-	-	-	PNPPI	-	-
0	0	1	1	spi0A0B11	PPK1_ECOLI	Polyphosphatase kinase	-	-	-	-	-	-	-	-	NLPPE	-	-
0	0	1	1	spi0A0FL6	PPX_ECOLI	Exopolyphosphatase	-	-	-	-	-	-	-	-	TTPTT	-	-
0	0	1	1	spi17888	PRIA_ECOLI	Primosomal protein N'	-	-	-	-	-	-	-	-	QLPPW	-	-
0	1	0	1	spi23862	PRIC_ECOLI	Primosomal replication protein N''	-	-	-	-	-	SAPPK	-	-	-	-	-
0	0	1	1	spi39199	PRMB_ECOLI	50S ribosomal protein L3 glutamine methyltransferase	-	-	-	-	-	-	-	-	TNPPY	-	-
0																	





Anzahl Motive					Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive					
S	M	W	Σ	Motiv 1			Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4			
0	0	1	1	1	1	YCEA_ECOLI	UPF0176 protein YceA	-	-	-	-	-	-	ALPPE	-	-	-	-
0	0	1	1	1	1	YCFT_ECOLI	Inner membrane protein Ycft	-	-	-	-	-	-	FSPPS	-	-	-	-
0	0	1	1	1	1	YCGV_ECOLI	Uncharacterized protein YcgV	-	-	-	-	-	-	AQPPI	-	-	-	-
0	1	0	1	1	1	YCGY_ECOLI	Uncharacterized protein YcgY	-	-	-	-	YFPPG	-	-	-	-	-	-
0	0	1	1	1	1	YCHO_ECOLI	Uncharacterized protein YchO	-	-	-	-	-	-	ATPPW	-	-	-	-
0	0	1	1	1	1	YCJR_ECOLI	Probable ketose 3-epimerase	-	-	-	-	-	-	RLPPM	-	-	-	-
0	1	0	1	1	1	YCJS_ECOLI	Uncharacterized oxidoreductase Ycjs	-	-	-	-	EKPPA	-	-	-	-	-	-
0	1	0	1	1	1	YCIY_ECOLI	Uncharacterized protein YcY	-	-	-	-	YFPPK	-	-	-	-	-	-
0	0	1	1	1	1	YDAS_ECOLI	Uncharacterized protein YdaS	-	-	-	-	-	-	VSPPS	-	-	-	-
0	0	1	1	1	1	YDBD_ECOLI	Uncharacterized protein YdbD	-	-	-	-	-	-	HTPPA	-	-	-	-
0	0	1	1	1	1	YDCP_ECOLI	Uncharacterized protease YdcP	-	-	-	-	-	-	DIPPI	-	-	-	-
0	0	1	1	1	1	YDCR_ECOLI	Uncharacterized HTH-type transcriptional regulator YdcR	-	-	-	-	-	-	NLPPG	-	-	-	-
0	0	1	1	1	1	YDCT_ECOLI	Uncharacterized ABC transporter ATP-binding protein YdcT	-	-	-	-	-	-	NLPPW	-	-	-	-
0	1	0	1	1	1	YDCV_ECOLI	Inner membrane ABC transporter permease protein YdcV	-	-	-	-	SFPPO	-	-	-	-	-	-
0	1	0	1	1	1	YDDE_ECOLI	Uncharacterized isomerase YdIE	-	-	-	-	FEPLL	-	-	-	-	-	-
0	1	0	1	1	1	YDDH_ECOLI	Uncharacterized protein YdIH	-	-	-	-	FEPLR	-	-	-	-	-	-
0	1	0	1	1	1	YDEM_ECOLI	Anaerobic sulfatase-maturing enzyme homolog YdeM	-	-	-	-	SVPPPT	-	-	-	-	-	-
0	0	1	1	1	1	YDEN_ECOLI	Uncharacterized sulfatase YdeN	-	-	-	-	-	-	SQPLL	-	-	-	-
0	0	1	1	1	1	YDEP_ECOLI	Protein YdeP	-	-	-	-	-	-	FTPPH	-	-	-	-
0	1	0	1	1	1	YDET_ECOLI	Putative uncharacterized protein ydet	-	-	-	-	VKPPPT	-	-	-	-	-	-
0	0	1	1	1	1	YDEU_ECOLI	Uncharacterized protein YdeU	-	-	-	-	-	-	NNPPV	-	-	-	-
0	1	0	1	1	1	YDFW_ECOLI	Uncharacterized protein YdfW	-	-	-	-	VKPPK	-	-	-	-	-	-
0	0	1	1	1	1	YDHI_ECOLI	Uncharacterized protein YdIH	-	-	-	-	-	-	YFPFF	-	-	-	-
0	0	1	1	1	1	YDHP_ECOLI	Inner membrane transport protein YdhP	-	-	-	-	-	-	VVPPL	-	-	-	-
0	0	1	1	1	1	YDHS_ECOLI	Uncharacterized protein YdhS	-	-	-	-	-	-	EIPPI	-	-	-	-
0	0	1	1	1	1	YDHW_ECOLI	Uncharacterized protein YdhW	-	-	-	-	-	-	LSPPY	-	-	-	-
0	0	1	1	1	1	YDHX_ECOLI	Uncharacterized ferredoxin-like protein YdhX	-	-	-	-	-	-	GFFPI	-	-	-	-
0	0	1	1	1	1	YDIF_ECOLI	Acetate CoA-transferase YdIF	-	-	-	-	-	-	VKPPR	-	-	-	-
0	0	1	1	1	1	YDIM_ECOLI	Inner membrane transport protein YdIM	-	-	-	-	-	-	TFPHI	-	-	-	-
0	1	0	1	1	1	YDIQ_ECOLI	Putative electron transfer flavoprotein subunit YdIQ	-	-	-	-	RVPPQ	-	-	-	-	-	-
0	0	1	1	1	1	YDJE_ECOLI	Inner membrane metabolite transport protein YdJE	-	-	-	-	-	-	RLPLL	-	-	-	-
0	1	0	1	1	1	YDJJ_ECOLI	Uncharacterized zinc-type alcohol dehydrogenase-like protein YdJJ	-	-	-	-	FIPPK	-	-	-	-	-	-
0	0	1	1	1	1	YDJK_ECOLI	Putative metabolite transport protein YdJK	-	-	-	-	-	-	PLPPV	-	-	-	-
0	0	1	1	1	1	YEAK_ECOLI	Uncharacterized protein YeaK	-	-	-	-	-	-	AIPPF	-	-	-	-
0	0	1	1	1	1	YEAL_ECOLI	UPF0756 membrane protein YeaL	-	-	-	-	-	-	TLPPS	-	-	-	-
0	1	0	1	1	1	YEAR_ECOLI	Uncharacterized protein YeaR	-	-	-	-	VFPPE	-	-	-	-	-	-
0	1	0	1	1	1	YEAY_ECOLI	Uncharacterized lipoprotein YeaY	-	-	-	-	IMPPQ	-	-	-	-	-	-
0	0	1	1	1	1	YEBQ_ECOLI	Uncharacterized transporter YebQ	-	-	-	-	-	-	FLPPN	-	-	-	-
0	0	1	1	1	1	YEBT_ECOLI	Uncharacterized protein YebT	-	-	-	-	-	-	ATPPG	-	-	-	-
0	1	0	1	1	1	YEBZ_ECOLI	Inner membrane protein YebZ	-	-	-	-	GFPPT	-	-	-	-	-	-
0	0	1	1	1	1	YECS_ECOLI	L-cystine transport system permease protein YecS	-	-	-	-	-	-	ALPPL	-	-	-	-
0	0	1	1	1	1	YEDK_ECOLI	Putative S	-	-	-	-	-	-	DKPPL	-	-	-	-
0	0	1	1	1	1	YEDL_ECOLI	Uncharacterized N-acetyltransferase YedL	-	-	-	-	-	-	QSPPE	-	-	-	-
0	0	1	1	1	1	YEEE_ECOLI	UPF0394 inner membrane protein YeeE	-	-	-	-	-	-	TLPPR	-	-	-	-
0	1	0	1	1	1	YEEZ_ECOLI	Protein YeeZ	-	-	-	-	LEPPQ	-	-	-	-	-	-
0	0	1	1	1	1	YEGH_ECOLI	UPF0053 protein YegH	-	-	-	-	-	-	KLPPK	-	-	-	-
0	0	1	1	1	1	Y EGL_ECOLI	Uncharacterized protein YegL	-	-	-	-	-	-	FFPPI	-	-	-	-
0	0	1	1	1	1	YEGS_ECOLI	Lipid kinase YegS	-	-	-	-	-	-	RLPPD	-	-	-	-
0	0	1	1	1	1	YEGT_ECOLI	Putative nucleoside transporter YegT	-	-	-	-	-	-	DTPPK	-	-	-	-
0	1	0	1	1	1	YEHE_ECOLI	Uncharacterized protein YehE	-	-	-	-	ISPPT	-	-	-	-	-	-
0	0	1	1	1	1	YEHW_ECOLI	Glycine betaine uptake system permease protein YehW	-	-	-	-	-	-	TFPPV	-	-	-	-
0	0	1	1	1	1	YEJE_ECOLI	Inner membrane ABC transporter permease protein YejE	-	-	-	-	-	-	PSPPS	-	-	-	-
0	1	0	1	1	1	YEJF_ECOLI	Uncharacterized ABC transporter ATP-binding protein YejF	-	-	-	-	PSPPV	-	-	-	-	-	-
0	0	1	1	1	1	YFAQ_ECOLI	Uncharacterized protein YfaQ	-	-	-	-	-	-	STPPE	-	-	-	-
0	0	1	1	1	1	YFAT_ECOLI	Uncharacterized protein YfaT	-	-	-	-	-	-	YLPPE	-	-	-	-
0	1	0	1	1	1	YFBL_ECOLI	Uncharacterized protein YfbL	-	-	-	-	EPPPF	-	-	-	-	-	-
0	0	1	1	1	1	YFBN_ECOLI	Uncharacterized protein YfbN	-	-	-	-	-	-	LLPPV	-	-	-	-
0	1	0	1	1	1	YFBP_ECOLI	Uncharacterized protein YfbP	-	-	-	-	PFPFN	-	-	-	-	-	-
0	0	1	1	1	1	YFBS_ECOLI	Uncharacterized transporter YfbS	-	-	-	-	-	-	ATPPN	-	-	-	-
0	1	0	1	1	1	YFCF_ECOLI	Glutathione S-transferase YfcF	-	-	-	-	FAPPT	-	-	-	-	-	-
0	0	1	1	1	1	YFCN_ECOLI	UPF0115 protein YfcN	-	-	-	-	-	-	WLPEE	-	-	-	-
0	0	1	1	1	1	YFCQ_ECOLI	Uncharacterized fibrial-like protein YfcQ	-	-	-	-	-	-	LSPPT	-	-	-	-
0	0	1	1	1	1	YFDM_ECOLI	Putative uncharacterized protein YfdM	-	-	-	-	-	-	GNNPV	-	-	-	-
0	0	1	1	1	1	YFDR_ECOLI	Uncharacterized protein YfdR	-	-	-	-	-	-	GLPPV	-	-	-	-
0	0	1	1	1	1	YFDS_ECOLI	Uncharacterized protein YfdS	-	-	-	-	-	-	ELPPA	-	-	-	-
0	0	1	1	1	1	YFDV_ECOLI	Uncharacterized transporter YfdV	-	-	-	-	-	-	ALPPA	-	-	-	-
0	1	0	1	1	1	YFIR_ECOLI	Protein Yfir	-	-	-	-	SGPPK	-	-	-	-	-	-
0	0	1	1	1	1	YFIK_ECOLI	Uncharacterized protein YfiK	-	-	-	-	-	-	HLPPG	-	-	-	-
0	0	1	1	1	1	YFIQ_ECOLI	UPF0380 protein YfiQ	-	-	-	-	-	-	LPPPV	-	-	-	-
0	0	1	1	1	1	YFJR_ECOLI	Uncharacterized HTH-type transcriptional regulator YfJR	-	-	-	-	-	-	DSPPV	-	-	-	-
0	0	1	1	1	1	YFJS_ECOLI	Lipoprotein YfJS	-	-	-	-	-	-	LTPPR	-	-	-	-
0	1	0	1	1	1	YFJU_ECOLI	Putative arsenate reductase-like protein	-	-	-	-	DMPPT	-	-	-	-	-	-
0	0	1	1	1	1	YGCN_ECOLI	Probable electron transfer flavoprotein-quinone oxidoreductase YgcN	-	-	-	-	-	-	PVPPV	-	-	-	-
0	1	0	1	1	1	YGCQ_ECOLI	Putative electron transfer flavoprotein subunit YgcQ	-	-	-	-	LFPPG	-	-	-	-	-	-
0	0	1	1	1	1	YGFI_ECOLI	Uncharacterized HTH-type transcriptional regulator YgfI	-	-	-	-	-	-	ELPPL	-	-	-	-
0	0	1	1	1	1	YGFT_ECOLI	Uncharacterized protein YgfT	-	-	-	-	-	-	GIPPF	-	-	-	-
0	0	1	1	1	1	YGfZ_ECOLI	rRNA-modifying protein YgfZ	-	-	-	-	-	-	PFPFR	-	-	-	-
0	1	0	1	1	1	YGGM_ECOLI	Uncharacterized protein YggM	-	-	-	-	GEPPPT	-	-	-	-	-	-
0	0	1	1	1	1	YGHF_ECOLI	Putative type II secretion system C-type protein YghF	-	-	-	-	-	-	FCPPY	-	-	-	-
0	0	1	1	1	1	YGHS_ECOLI	Uncharacterized ATP-binding protein YghS	-	-	-	-	-	-	STPPG	-	-	-	-
0	0	1	1	1	1	YGHU_ECOLI	Disulfide-bond oxidoreductase YghU	-	-	-	-	-	-	HNPPV	-	-	-	-
0	0	1	1	1	1	YGID_ECOLI	4,5-DOPA dioxygenase extradiol	-	-	-	-	-	-	ETPPT	-	-	-	-
0	0	1	1	1	1	YGJI_ECOLI	Inner membrane transporter YgjI	-	-	-	-	-	-	MLPPL	-	-	-	-
0	0	1	1	1	1	YHAB_ECOLI	Uncharacterized protein YhaB	-	-	-	-	-	-	SIPPM	-	-	-	-
0	1	0	1	1	1	YHAV_ECOLI	Ribonuclease toxin YhaV	-	-	-	-	GHPPA	-	-	-	-	-	-
0	0	1	1	1	1	YHBW_ECOLI	Luciferase-like monooxygenase	-	-	-	-	-	-	LPPLI	-	-	-	-
0	1	0	1	1	1	YHCA_ECOLI	Uncharacterized fibrial chaperone YhcA	-	-	-	-	GIPPK	-	-	-	-	-	-
0	1	0	1	1	1	YHCD_ECOLI	Uncharacterized outer membrane usher protein YhcD	-	-	-	-	FSPPA	-	-	-	-	-	-
0	0	1	1	1	1	YHCG_ECOLI	Putative nuclease YhcG	-	-	-	-	-	-	ENPPI	-	-	-	-
0	0	1	1	1	1	YHET_ECOLI	Putative esterase YheT	-	-	-	-	-	-	SLPPQ	-	-	-	-
0	0	1	1	1	1	YHFS_ECOLI	Uncharacterized protein YhfS	-	-	-	-	-	-	EIPPL	-	-	-	-
0	1	0	1	1	1	YHGF_ECOLI	Protein YhgF	-	-	-	-	DEPPK	-	-	-	-	-	-
0	0	1	1	1	1	YHHM_ECOLI	Uncharacterized protein YhhM	-	-	-	-	-	-	SKPPL	-	-	-	-
0	0	1	1	1	1	YHHT_ECOLI	Putative transport protein YhhT	-	-	-	-	-	-	AVPPM	-	-	-	-
0	0	1	1	1	1	YHHW_ECOLI	Quercetin 2,3-dioxygenase	-	-	-	-	-	-	DLPPV	-	-	-	-
0	0	1	1	1	1	YHHZ_ECOLI	Uncharacterized protein YhhZ	-	-	-	-	-	-	IVPPM	-	-	-	-
0	1	0	1	1	1	YHJE_ECOLI	Inner membrane metabolite transport protein YhjE	-	-	-	-	NAPPR	-	-	-	-	-	-
0	1	0	1	1	1	YIAA_ECOLI	Inner membrane protein YiaA	-	-	-	-	INPPK	-	-	-	-		

Anzahl Motive					Protein-ID	Beschreibung	Art starke Motive			Art moderate Motive			Art schwache Motive			
S	M	W	Σ	Motiv 1			Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 1	Motiv 2	Motiv 3	Motiv 4	
0	0	1	1	spi32136	YIHQ_ECOLI	Putative sulfoquinovose importer	-	-	-	-	-	-	RNPL	-	-	-
0	0	1	1	spi32144	YIHW_ECOLI	Uncharacterized HTH-type transcriptional regulator YihW	-	-	-	-	-	-	ELPPA	-	-	-
0	0	1	1	spi32151	YIIG_ECOLI	Uncharacterized protein YigG	-	-	-	-	-	-	TLPPA	-	-	-
0	1	0	1	spi32668	YIUF_ECOLI	Uncharacterized protein YifF	-	-	-	AVPPT	-	-	-	-	-	-
0	1	0	1	spi32689	YJBH_ECOLI	Uncharacterized lipoprotein YjbH	-	-	-	TNPPQ	-	-	-	-	-	-
0	0	1	1	spi32700	YJCB_ECOLI	Uncharacterized protein YjcB	-	-	-	-	-	-	-	QTPPQ	-	-
0	0	1	1	spi32703	YJCE_ECOLI	Uncharacterized Na <sup>(+)</sup> /H <sup>(+)</sup> exchanger YjcE	-	-	-	-	-	-	-	FIPPL	-	-
0	0	1	1	spi33222	YJFC_ECOLI	Putative acid-amine ligase Yjfc	-	-	-	-	-	-	-	AIPPL	-	-
0	0	1	1	spi39298	YJFP_ECOLI	Esterase Yjfp	-	-	-	-	-	-	-	LFPPA	-	-
0	1	0	1	spiP0A8X0	YJGA_ECOLI	UPF0307 protein YigA	-	-	-	NKPPK	-	-	-	-	-	-
0	0	1	1	spi39332	YJGH_ECOLI	RutC family protein YjgH	-	-	-	-	-	-	-	SAPPY	-	-
0	1	0	1	spi39336	YJGL_ECOLI	Uncharacterized protein YjgL	-	-	-	LNPPE	-	-	-	-	-	-
0	1	0	1	spi39342	YJGR_ECOLI	Uncharacterized protein YjgR	-	-	-	NNPPA	-	-	-	-	-	-
0	0	1	1	spi39351	YJGZ_ECOLI	Uncharacterized protein YjgZ	-	-	-	-	-	-	-	MLPPG	-	-
0	1	0	1	spi39352	YJHB_ECOLI	Putative metabolite transport protein YjhB	-	-	-	VNPPQ	-	-	-	-	-	-
0	1	0	1	spi39354	YJHD_ECOLI	Putative uncharacterized protein YjhD	-	-	-	PEPPR	-	-	-	-	-	-
0	0	1	1	spi39357	YJHF_ECOLI	Uncharacterized permease Yjhf	-	-	-	-	-	-	-	FLPPH	-	-
0	0	1	1	spi39374	YJIC_ECOLI	Uncharacterized protein YjcC	-	-	-	-	-	-	-	YLPL	-	-
0	0	1	1	spi39376	YJIE_ECOLI	HTH-type transcriptional regulator Yjie	-	-	-	-	-	-	-	QMPL	-	-
0	0	1	1	spi39381	YJII_ECOLI	Uncharacterized protein YjiI	-	-	-	-	-	-	-	WLPPF	-	-
0	1	0	1	spi39384	YJIM_ECOLI	Putative dehydratase subunit Yjim	-	-	-	LNPPA	-	-	-	-	-	-
0	0	1	1	spi39385	YJIN_ECOLI	Uncharacterized protein Yjin	-	-	-	-	-	-	-	FLPPN	-	-
0	0	1	1	spi39393	YJIV_ECOLI	Putative uncharacterized protein Yjiv	-	-	-	-	-	-	-	FQPL	-	-
0	0	1	1	spi39410	YJII_ECOLI	Toxin YjiI	-	-	-	-	-	-	-	SEPPF	-	-
0	0	1	1	spi39408	YJIV_ECOLI	Uncharacterized metal-dependent hydrolase Yjiv	-	-	-	-	-	-	-	DPPPF	-	-
0	0	1	1	spi75677	YKFF_ECOLI	UPF0401 protein YkfF	-	-	-	-	-	-	-	LLPPG	-	-
0	0	1	1	spiP0AAS5	YLBK_ECOLI	Uncharacterized protein YlbK	-	-	-	-	-	-	-	ITPPA	-	-
0	0	1	1	spi75982	YMFQ_ECOLI	Uncharacterized protein YmfQ	-	-	-	-	-	-	-	LLPPG	-	-
0	1	0	1	spi75979	YMFY_ECOLI	Uncharacterized protein YmfY	-	-	-	IYPPA	-	-	-	-	-	-
0	0	1	1	spiP0AEB5	YNAI_ECOLI	Low conductance mechanosensitive channel YnaI	-	-	-	-	-	-	-	ITPPE	-	-
0	0	1	1	spi76093	YNBD_ECOLI	Uncharacterized protein YnbD	-	-	-	-	-	-	-	KLPPA	-	-
0	1	0	1	spi76115	YNCD_ECOLI	Probable TonB-dependent receptor YncD	-	-	-	QQPPT	-	-	-	-	-	-
0	0	1	1	spi76117	YNCG_ECOLI	Uncharacterized GST-like protein YncG	-	-	-	-	-	-	-	LAPPV	-	-
0	0	1	1	spi77559	YNFL_ECOLI	Uncharacterized HTH-type transcriptional regulator YnfL	-	-	-	-	-	-	-	SQPLL	-	-
0	0	1	1	spi78067	YNJE_ECOLI	Thiosulfate sulfurtransferase Ynje	-	-	-	-	-	-	-	GTPPK	-	-
0	0	1	1	spi76257	YOAA_ECOLI	Probable ATP-dependent DNA helicase YoaA	-	-	-	-	-	-	-	SLPPA	-	-
0	0	1	1	spiP0AECO	YOAE_ECOLI	UPF0053 inner membrane protein YoeE	-	-	-	-	-	-	-	KLPPK	-	-
0	1	0	1	spi64493	YOAF_ECOLI	Uncharacterized protein YofA	-	-	-	PKPPQ	-	-	-	-	-	-
0	0	1	1	spi33366	YOHD_ECOLI	Inner membrane protein YohD	-	-	-	-	-	-	-	QLPPK	-	-
0	0	1	1	spiP0AE39	YPDB_ECOLI	Transcriptional regulatory protein YpdB	-	-	-	-	-	-	-	KLPPS	-	-
0	1	0	1	spi77396	YPDC_ECOLI	Uncharacterized HTH-type transcriptional regulator YpdC	-	-	-	NMPPG	-	-	-	-	-	-
0	0	1	1	spiQ2EET2	YPFN_ECOLI	UPF0370 protein YpfN	-	-	-	-	-	-	-	ELPPH	-	-
0	0	1	1	spi76584	YPHB_ECOLI	Uncharacterized protein YphB	-	-	-	-	-	-	-	ITPPA	-	-
0	1	0	1	spi76585	YPHG_ECOLI	Uncharacterized protein YphG	-	-	-	VFPDD	-	-	-	-	-	-
0	0	1	1	spi64432	YPIJ_ECOLI	Inner membrane protein YpiJ	-	-	-	-	-	-	-	EMPLL	-	-
0	0	1	1	spiP0AE42	YQAE_ECOLI	UPF0057 membrane protein YqaE	-	-	-	-	-	-	-	ILPPL	-	-
0	0	1	1	spi77051	YQCE_ECOLI	Inner membrane protein YqeC	-	-	-	-	-	-	-	TYPPL	-	-
0	1	0	1	spiQ46808	YQEB_ECOLI	Uncharacterized protein YqeB	-	-	-	LFPSS	-	-	-	-	-	-
0	1	0	1	spiQ46942	YQEL_ECOLI	Uncharacterized protein YqeI	-	-	-	NAPPA	-	-	-	-	-	-
0	0	1	1	spiP0A8W5	YQGE_ECOLI	UPF0301 protein YagE	-	-	-	-	-	-	-	HTPPS	-	-
0	0	1	1	spiQ46856	YQHD_ECOLI	Alcohol dehydrogenase YqhD	-	-	-	-	-	-	-	TLPPR	-	-
0	0	1	1	spiP0A8Z7	YQIA_ECOLI	Esterase YqiA	-	-	-	-	-	-	-	QLPPY	-	-
0	0	1	1	spiQ46868	YQIC_ECOLI	Uncharacterized protein YqiC	-	-	-	-	-	-	-	TLPPT	-	-
0	1	0	1	spiQ42620	YQIG_ECOLI	Glutathionyl-hydroquinone reductase YqgG	-	-	-	YYPPA	-	-	-	-	-	-
0	0	1	1	spiP42914	YRAI_ECOLI	Probable fibrial chaperone YraI	-	-	-	-	-	-	-	ITPPL	-	-
0	0	1	1	spiP45394	YRBG_ECOLI	Inner membrane protein YrbG	-	-	-	-	-	-	-	GIPPL	-	-
0	0	1	1	spi39314	YTFE_ECOLI	Inner membrane protein YtfE	-	-	-	-	-	-	-	QMPLL	-	-
0	0	1	1	spiP0ACN2	YTFH_ECOLI	Uncharacterized HTH-type transcriptional regulator YtfH	-	-	-	-	-	-	-	VVPPH	-	-
0	0	1	1	spi39187	YTFI_ECOLI	Uncharacterized protein YtfI	-	-	-	-	-	-	-	RVPPH	-	-
0	1	0	1	spiP39328	YTFJ_ECOLI	Inner membrane ABC transporter permease protein YtfJ	-	-	-	GFPPE	-	-	-	-	-	-
0	0	1	1	spiP0DN74	YTIA_ECOLI	Uncharacterized protein YtiA	-	-	-	-	-	-	-	TCPPG	-	-
0	1	0	1	spiP0DPC5	YTID_ECOLI	Protein YtiD	-	-	-	NFPPE	-	-	-	-	-	-
0	0	1	1	spiP0ADS2	ZAPA_ECOLI	Cell division protein Zapa	-	-	-	-	-	-	-	NCPDP	-	-
0	0	1	1	spiP37617	ZNTA_ECOLI	Zinc/cadmium/lead-transporting P-type ATPase	-	-	-	-	-	-	-	LVPPL	-	-
0	0	1	1	spiP64423	ZNTB_ECOLI	Zinc transport protein ZntB	-	-	-	-	-	-	-	QIPPR	-	-
0	1	0	1	spiP0AAA9	ZRAP_ECOLI	Zinc resistance-associated protein	-	-	-	ANPPD	-	-	-	-	-	-
0	1	0	1	spiP0AC51	ZUR_ECOLI	Zinc uptake regulation protein	-	-	-	AKPPT	-	-	-	-	-	-

**Tabelle S2:** 381 nichtredundante Orthologe aus der Uni Prot Datenbank „Microbial Proteomes“ mit mindestens 44% Übereinstimmung und höchstens 10% Sequenzlängenabweichung zu DgcC aus *E. coli* K-12.

Uni Prot ID	Spezies	Uni Prot ID	Spezies
A0A0M7MNT6	Achromobacter sp.	A0A0E2L8W1	Escherichia coli (strain UMEA 3162-1)
H5V4M8	Atlantibacter hermannii NBRC 105704	Q1RFF9	Escherichia coli (strain UT189 / UPEC)
A0A085GHX3	Buttiauxella agrestis ATCC 33320	I2SYM0	Escherichia coli 1.2264
A0A1B7IUX3	Buttiauxella brennerae ATCC 51605	I2RKB5	Escherichia coli 1.2741
A0A2N0CM18	Cedecea lapagei	U9ZG70	Escherichia coli 110957
W0LCU8	Chania multitudinisensens RB-25	A0A125X7X6	Escherichia coli 1-110-08_S3_C1
A0A2S4S0L6	Citrobacter amalonaticus	U9XI27	Escherichia coli 113290
A0A0F6TTE8	Citrobacter amalonaticus Y19	U9XQW3	Escherichia coli 113303
A0A1R0G039	Citrobacter braaki	A0A0809Q07	Escherichia coli 1-250-04_S3_C2
A0A223JNB6	Citrobacter farmeri	A0A0E1LTV6	Escherichia coli 1303
A0A023V8J8	Citrobacter freundii CFNIH1	A0A074IAE9	Escherichia coli 1-392-07_S4_C1
A0A2I8S399	Citrobacter freundii complex sp. CFNIH3	A0A080G8V7	Escherichia coli 1-392-07_S4_C3
A0A2L0T811	Citrobacter freundii complex sp. CFNIH4	I2XGY0	Escherichia coli 2.3916
A0A2I8TPY9	Citrobacter freundii complex sp. CFNIH9	A0A029IIP9	Escherichia coli 2-005-03_S4_C2
A0A064E7J0	Citrobacter freundii MGH 56	A0A029IB30	Escherichia coli 2-005-03_S4_C3
A0A2L1NN16	Citrobacter koseri (Citrobacter diversus)	A0A070V6E4	Escherichia coli 2-177-06_S3_C2
A8AK79	Citrobacter koseri (strain ATCC BAA-895 / CDC 4225-83 / SGSC4696)	A0A070SL33	Escherichia coli 2-210-07_S3_C3
A0A0A1RWCS	Citrobacter pasteurii	E3XHH0	Escherichia coli 2362-75
D2TKW9	Citrobacter rodentium (strain ICC168) (Citrobacter freundii biotype 4280)	A0A073G0D2	Escherichia coli 2-427-07_S4_C3
A0A0V9JZZ0	Citrobacter sp. 50677481	A0A069XPZ0	Escherichia coli 3-373-03_S4_C2
A0A1V4NW82	Citrobacter sp. A316	I2UQL2	Escherichia coli 4.0522
R8WYU8	Citrobacter sp. KTE151	I2WR70	Escherichia coli 4.0967
A0A2N4V4E8	Citrobacter sp. L55	A0A073VAM2	Escherichia coli 5-366-08_S1_C1
A0A0J1K9N3	Citrobacter sp. MGH100	A0A073HQ47	Escherichia coli 5-366-08_S1_C3
A0A0J1Q8Q7	Citrobacter sp. MGH105	I4SYQ8	Escherichia coli 541-15
A0A0J1LMR3	Citrobacter sp. MGH106	I2W4W4	Escherichia coli 9.0111
A0A1H0W368	Citrobacter sp. NLAE-zl-C269	V0STV8	Escherichia coli 907672
A0A090TYK1	Citrobacter werkmanii NBRC 105721	U9Z8Y4	Escherichia coli 907713
D4BD88	Citrobacter youngae ATCC 29220	V0VC44	Escherichia coli 908519
A0A2K8QJW8	Dickeya fangzhongdai	V0XFW1	Escherichia coli 908525
C6CCW8	Dickeya paradisiaca (strain Ech703) (Dickeya dadantii (strain Ech703))	V0Z9C9	Escherichia coli 908573
A0A2K8W491	Dickeya solani RNS 08.23.3.1.A	V0AT47	Escherichia coli 909945-2
D2BVJ3	Dickeya zeae (strain Ech586) (Dickeya dadantii (strain Ech586))	I2S4E5	Escherichia coli 97.0246
A0A0A8FA82	Dickeya zeae EC1	V6FPJ6	Escherichia coli 99.0741
G2S965	Enterobacter asburiae (strain LF7a)	V8F9L4	Escherichia coli ATCC BAA-2209
A0A2J7SAT8	Enterobacter bugandensis	D6HSR3	Escherichia coli B088
D2ZFW1	Enterobacter cancerogenus ATCC 35316	D6I6J4	Escherichia coli B185
A0A205KVR2	Enterobacter cloacae complex 'Hoffmann cluster IV'	D6I6V9	Escherichia coli B354
A0A0F1B5P1	Enterobacter cloacae complex sp. 35699	J7RNL8	Escherichia coli chi7122
A0A0F4BSI8	Enterobacter cloacae complex sp. CIDeIMsCOL9	F4NMD6	Escherichia coli D9
A0A2S4SRJ0	Enterobacter cloacae complex sp. ECNIH11	H4I747	Escherichia coli DEC1B
A0A2S4SFX6	Enterobacter cloacae complex sp. ECNIH12	H4IMM5	Escherichia coli DEC1C
A0A0XGZ0I4	Enterobacter cloacae complex sp. GN02468	H4J413	Escherichia coli DEC1D
A0A0X4EVA6	Enterobacter cloacae complex sp. GN04363	H4KCP2	Escherichia coli DEC2C
A0A167IA20	Enterobacter cloacae complex sp. GN04787	H4L8B7	Escherichia coli DEC2E
A0A0W2I6D5	Enterobacter cloacae complex sp. SMART_763	H4UF70	Escherichia coli DEC6A
G8LP83	Enterobacter cloacae EcWSU1	W1X4W7	Escherichia coli DORA_A_5_14_21
W0BTA0	Enterobacter cloacae P101	A0A1X3IQR6	Escherichia coli E1114
A0A0A0HCF8	Enterobacter cloacae subsp. cloacae (strain ATCC 13047 / NCDC 279-56)	A0A0A0FAB2	Escherichia coli G3/10
V3E2G1	Enterobacter cloacae UCICRE 12	A0A1X3JM05	Escherichia coli H386
A0A0M2C8K0	Enterobacter cloacae UCICRE 3	A0A1X3L7G3	Escherichia coli H420
FSRV64	Enterobacter hormaechei ATCC 49162	A0A1X3K7R7	Escherichia coli H461
A0A0W2YLC6	Enterobacter hormaechei subsp. oharae	F4VA55	Escherichia coli H591
A0A0E1J673	Enterobacter hormaechei subsp. steigerwaltii	A0A1X3KPR7	Escherichia coli H605
A0A242VK56	Enterobacter kobei	F4SIZ8	Escherichia coli H736
A0A229HU80	Enterobacter mori	V2RZH4	Escherichia coli HVH 50 (4-2593475)
A0A198GTV8	Enterobacter soli ATCC BAA-2102	T6DWO0	Escherichia coli HVH 70 (4-2963531)
A4W760	Enterobacter sp. (strain 638)	T6MMP7	Escherichia coli HVH 87 (4-5977630)
A0A0V9EA58	Enterobacter sp. 50588862	W1FWH7	Escherichia coli ISC11
A0A0I1VVV3	Enterobacter sp. BIDMC92	L2V8D2	Escherichia coli KTE10
A0A2I6K7K0	Enterobacter sp. Crenshaw	S1HS83	Escherichia coli KTE100
W7NU01	Enterobacter sp. DC4	S1J6B0	Escherichia coli KTE107
A0A2T3CK83	Enterobacter sp. FS01	S1TT88	Escherichia coli KTE108
A0A0I0J7K2	Enterobacter sp. GN02600	L4W214	Escherichia coli KTE112
A0A2H6KWD5	Enterobacter sp. KINAN-G	L4JGL0	Escherichia coli KTE146
A0A1ESMR89	Enterobacter sp. ku-bf2	S1QAM8	Escherichia coli KTE182
A0A0E2MN81	Enterobacter sp. MGH 14	L3BS22	Escherichia coli KTE193
V3PYE3	Enterobacter sp. MGH 24	S0W027	Escherichia coli KTE20
A0A1U6IN86	Enterobacter sp. NFR05	S0YK45	Escherichia coli KTE37
A0A202D9L2	Enterobacter xiangfangensis	S0ZVB6	Escherichia coli KTE38
A0A2I8Q7Y7	Enterobacteriaceae bacterium ENNIH1	S1DF15	Escherichia coli KTE64
A0A2I8QVX8	Enterobacteriaceae bacterium ENNIH2	L3PDL5	Escherichia coli KTE66
A0A0L7T2H6	Erwinia iniecta	S1FG87	Escherichia coli KTE73
A0A1X3JT17	Escherichia albertii B156	L3QCX0	Escherichia coli KTE75
W0ARL8	Escherichia albertii KF1	V8ISZ6	Escherichia coli LAU-EC10
B7L540	Escherichia coli (strain 55989 / EAEC)	A0A1X3BFB0	Escherichia coli M056
E0U0J9	Escherichia coli (strain ATCC 9637 / DSM 1116 / NCIMB 8666 / NRRL B-766 / W)	A0A0K9TC73	Escherichia coli M114
A0A140NEW1	Escherichia coli (strain B / BL21-DE3)	F4SUJ9	Escherichia coli M605
P0AAP1	Escherichia coli (strain K12)	F4TAP0	Escherichia coli M718
D8A4E9	Escherichia coli (strain MS 21-1)	E9YXN2	Escherichia coli M863
B6HZI5	Escherichia coli (strain SE11)	M9HD29	Escherichia coli MP021561.2
B1LIR9	Escherichia coli (strain SMS-3-5 / SECEC)	D7Y387	Escherichia coli MS 115-1

Uni Prot ID	Spezies	Uni Prot ID	Spezies
E9TLK3	Escherichia coli MS 117-3	A0A1M5LV19	Pantoea sesami
D8E176	Escherichia coli MS 119-7	Q6D772	Pectobacterium atrosepticum (strain SCRI 1043 / ATCC BAA-672)
E1IZQ8	Escherichia coli MS 124-1	A0A1V2R8K4	Pectobacterium carotovorum subsp. actinidiae
E1IPL4	Escherichia coli MS 145-7	A0A086EA40	Pectobacterium carotovorum subsp. brasiliense
D7XC99	Escherichia coli MS 198-1	C6DCM6	Pectobacterium carotovorum subsp. carotovorum (strain PC1)
D7ZJJ4	Escherichia coli MS 69-1	J7KX68	Pectobacterium carotovorum subsp. carotovorum PCC21
D7XKE9	Escherichia coli MS 84-1	A0A0U5XZT3	Pectobacterium carotovorum subsp. carotovorum UGC32
E6BHL1	Escherichia coli MS 85-1	A0A094UGI5	Pectobacterium carotovorum subsp. odoriferum
A0A0H2YW49	Escherichia coli O1:K1 / APEC	D0K124	Pectobacterium parmentieri (strain WPP163)
A0A2S1UXT8	Escherichia coli O103 str. RM8385	A0A093SW48	Pectobacterium polaris
CSU271	Escherichia coli O103:H2 (strain 12009 / EHEC)	A0A0H3H4V9	Pectobacterium sp. (strain SCC3193)
A0A0E0Y395	Escherichia coli O104:H4 (strain 2011C-3493)	A0A2T3VV94	Pectobacterium sp. SS95
CSUIG9	Escherichia coli O111:H- (strain 11128 / EHEC)	A0A1D7Z453	Pectobacterium wasabiae CFBP 3304
K4WSQ1	Escherichia coli O111:H11 str. CVM9455	A0A1V3X5H3	Pedobacter himalayensis
K4W6F3	Escherichia coli O111:H8 str. CVM9634	A0A0G3QHA8	Phytobacter ursingii
A0A027TXX5	Escherichia coli O111:NM str. K6722	A0A090V143	Pseudoescherichia vulneris NBRC 102420
A0A028DMW4	Escherichia coli O118:H16 str. 2009C-4446	B5EW86	Salmonella agona (strain SL483)
A0A2S1UHK2	Escherichia coli O121 str. RM8352	A0A2T8N0W6	Salmonella anatum
A0A0E2TIN4	Escherichia coli O121:H19 str. 2010C-3609	A0A0K0H7T8	Salmonella bongori (strain ATCC 43975 / DSM 13772 / NCTC 12419)
B7UJK9	Escherichia coli O127:H6 (strain E2348/69 / EPEC)	S5MS16	Salmonella bongori N268-08
A0A070FGS5	Escherichia coli O128:H2 str. 2011C-3317	A0A248KD46	Salmonella bongori serovar 66:z41:- str. SA19983605
A7ZID4	Escherichia coli O139:H28 (strain E24377A / ETEC)	Q57SH9	Salmonella choleraesuis (strain SC-B67)
A0A023YT24	Escherichia coli O145:H28 str. RM12581	A0A2S1H5J3	Salmonella diarizonae
A0A025CFX5	Escherichia coli O145:NM str. 2010C-3526	A0A1X2TCT3	Salmonella enterica subsp. arizonae serovar 50:r:z
A0A1Z3UU30	Escherichia coli O157	A0A2T9I5G5	Salmonella enterica subsp. enterica serovar Agona
A0A0H3JCC4	Escherichia coli O157:H7	G5LJS2	Salmonella enterica subsp. enterica serovar Alachua str. R6-377
Q8XEB2	Escherichia coli O157:H7	A0A2T9MKJ3	Salmonella enterica subsp. enterica serovar Bareilly
A0A0H3PRW9	Escherichia coli O157:H7 (strain EC869)	A0A0T7RPM9	Salmonella enterica subsp. enterica serovar Bovismorbificans
A0A0F6FIS2	Escherichia coli O157:H7 str. S552	A0A2T9LJ88	Salmonella enterica subsp. enterica serovar Braenderup
B7N8T7	Escherichia coli O17:K52:H18 (strain UMN026 / ExPEC)	A0A2T9Q8N8	Salmonella enterica subsp. enterica serovar Cerro
A0A026V5A8	Escherichia coli O174:H8 str. 04-3038	A0A2R4DCX9	Salmonella enterica subsp. enterica serovar Concord
A0A192CJ0	Escherichia coli O25b:H4	A0A2T9JAL8	Salmonella enterica subsp. enterica serovar Copenhagen
W8ZF56	Escherichia coli O25b:H4-ST131	V7ISU6	Salmonella enterica subsp. enterica serovar Cubana str. 76814
A0A2S1VEN7	Escherichia coli O26 str. RM10386	M7RRZ6	Salmonella enterica subsp. enterica serovar Dublin str. UC16
A0A0A8UHS6	Escherichia coli O26:H11	S4H36	Salmonella enterica subsp. enterica serovar Enteritidis str. 2009K0958
D3GU78	Escherichia coli O44:H18 (strain O42 / EAEC)	A0A0G2NK67	Salmonella enterica subsp. enterica serovar Gallinarum str. SG9
B7MD44	Escherichia coli O45:K1 (strain S88 / ExPEC)	A0A2T8XC90	Salmonella enterica subsp. enterica serovar Gaminara
D3QIF2	Escherichia coli O55:H7 (strain CB9615 / EPEC)	A0A2T9PXC7	Salmonella enterica subsp. enterica serovar Give
P0A AP2	Escherichia coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC)	A0A265B7I2	Salmonella enterica subsp. enterica serovar Heidelberg
A0A028AH34	Escherichia coli O69:H11 str. 08-4661	A0A1Z3Q0Z5	Salmonella enterica subsp. enterica serovar India str. SA20085604
A0A0H3ML36	Escherichia coli O7:K1 (strain IA139 / ExPEC)	G4CSB9	Salmonella enterica subsp. enterica serovar Infantis str. SARB27
A0A0E0V179	Escherichia coli O7:K1 str. CE10	A0A2T9CP19	Salmonella enterica subsp. enterica serovar Javiana
E3PFK2	Escherichia coli O78:H11 (strain H10407 / ETEC)	A0A0L9F6W6	Salmonella enterica subsp. enterica serovar Johannesburg
B7MP66	Escherichia coli O81 (strain ED1a)	A0A2T9LHJ8	Salmonella enterica subsp. enterica serovar Kentucky
A0A0H3E8D0	Escherichia coli O83:H1 (strain NRG 857C / AIEC)	A0A2T9KHH1	Salmonella enterica subsp. enterica serovar Kiambu
A0A0G3K0D6	Escherichia coli PCN033	A0A2C9P300	Salmonella enterica subsp. enterica serovar Macclesfield str. S-1643
A0A1X3LMY2	Escherichia coli PA054	A0A2T9M776	Salmonella enterica subsp. enterica serovar Meleagridis
A0A1X3LTL6	Escherichia coli TA249	A0A2T8PVC2	Salmonella enterica subsp. enterica serovar Mississippi
A0A1X3J6E3	Escherichia coli TA447	G5FYI6	Salmonella enterica subsp. enterica serovar Mississippi str. A4-633
E9XPS0	Escherichia coli TW10509	A0A2S1BTR6	Salmonella enterica subsp. enterica serovar Montevideo str. CDC 2010K-0257
T9AJZ3	Escherichia coli UMEA 3200-1	V1XG82	Salmonella enterica subsp. enterica serovar Muenchen str. baal594
T9ES07	Escherichia coli UMEA 3212-1	A0A1X2RJY1	Salmonella enterica subsp. enterica serovar Rough O:d:1,7
T9UHL3	Escherichia coli UMEA 3718-1	A0A249V1R5	Salmonella enterica subsp. enterica serovar Saintpaul
Q0FC79	Escherichia coli UMN118	A0A1X2Y8R1	Salmonella enterica subsp. enterica serovar Saphra
A0A0E0TU64	Escherichia coli UMNK88	A0A2T1CCN7	Salmonella enterica subsp. enterica serovar Schwarzengrund
I0VQX9	Escherichia coli Wuz2	A0A1X2USH9	Salmonella enterica subsp. enterica serovar Seftenburg
A0A0F6BZK3	Escherichia coli Xuzhou21	A0A2S1HWK5	Salmonella enterica subsp. enterica serovar Senftenberg str. 361154004
A0A2S7SQ55	Escherichia fergusonii	A0A2R4IFW5	Salmonella enterica subsp. enterica serovar Tennessee
B7LMK0	Escherichia fergusonii (strain ATCC 35469 / DSM 13698 / CDC 0568-73)	A0A0U1GR99	Salmonella enterica subsp. enterica serovar Typhimurium str. DT104
F0UMY7	Escherichia fergusonii ECD227	G5RBE6	Salmonella enterica subsp. enterica serovar Uganda str. R8-3404
A0A2T3UG66	Escherichia sp. 4726-5	G5S761	Salmonella enterica subsp. enterica serovar Wandswoth str. A4-580
A0A2S4QM72	Escherichia sp. ESNIH1	A0A2T9HUA4	Salmonella enterica subsp. enterica serovar Weslaco
S0TZN6	Escherichia sp. KTE114	V1HC15	Salmonella enterica subsp. indica serovar 6,14,25:z10:(1,2),7 str. 1121
S1M3T4	Escherichia sp. KTE159	A0A241SGH5	Salmonella enterica subsp. salamae serovar 55:k:z39 str. 1315K
S0XB25	Escherichia sp. KTE31	A0A0V2D4J4	Salmonella enterica subsp. salamae serovar 56:z10:e,n,x str. 1369-73
S1CA28	Escherichia sp. KTE52	A0A1R2MWY9	Salmonella enteritidis
A0A2T3BRI0	Escherichia sp. MOD1-EC5451	B5RSX6	Salmonella gallinarum (strain 287/91 / NCTC 13346)
A0A0I8VSM0	Franconibacter pulveris	A0A1I6YEC5	Salmonella houtenae
A0A155F3G1	Klebsiella aerogenes	A0A2T9N8U6	Salmonella infantis
A0A171CYE7	Klebsiella oxytoca	A0A2T7HL15	Salmonella muenster
A0A2U2HWS3	Kluyvera intestini	A0A0R9N0L2	Salmonella newport
A0A2T3XBM2	Kluyvera sp. Nf5	A0A0H3BXP5	Salmonella newport (strain SL254)
A0A1H1X0V7	Kosakonia sp. S29	A0A0H2WR45	Salmonella paratyphi A (strain ATCC 9150 / SARB42)
A0A21I2P98	Leclercia adecarboxylata	COQ7Q8	Salmonella paratyphi C (strain RKS4594)
A0A0I6Q5J9	Leclercia sp. LK8	A0A2T8W4J1	Salmonella rubislaw
A0A2I8NYV3	Leclercia sp. LSNIH1	A0A0N1QU07	Salmonella schwarzengrund (strain CVM19633)
A0A2S4XJ16	Leclercia sp. LSNIH4	A0A1F2K6U0	Salmonella sp. HMSC13B08
A0A066PES8	Lelliottia amnigena CHS 78	A0A2T9P1K7	Salmonella thompson
A0A2L2JB75	Lelliottia nimpipressuralis	A0A0F6AKX3	Salmonella typhimurium (strain 14028s / SGSC 2262)
A0A1J5W4I8	Lelliottia nimpipressuralis	E8XJ75	Salmonella typhimurium (strain 4/74)
A0A2S5CS31	Lelliottia sp. 7254-16	H9L4B6	Salmonella typhimurium (strain LT2 / SGSC1412 / ATCC 700720)
A0A2U3TIS2	Lelliottia sp. WB101	A0A0H3N9U4	Salmonella typhimurium (strain SL1344)

Uni Prot ID	Spezies	Uni Prot ID	Spezies
U2NM58	<i>Serratia fonticola</i> AU-AP2C		
A0A1C0R780	<i>Serratia marcescens</i> subsp. <i>marcescens</i>		
A0A086G4X1	<i>Serratia nematodiphila</i> DZ0503SBS1		
A0A1S8CP63	<i>Serratia oryzae</i>		
S0ADJ9	<i>Serratia plymuthica</i> 4Rx13		
A0A1B1KML2	<i>Serratia plymuthica</i> PRI-2C		
S4YQ92	<i>Serratia plymuthica</i> S13		
A0A1W5DFY9	<i>Serratia proteamaculans</i>		
A8GB07	<i>Serratia proteamaculans</i> (strain 568)		
A0A1C7W5H1	<i>Serratia</i> sp. 14-2641		
A0A087KVG4	<i>Serratia</i> sp. Ag1		
A0A083ZV18	<i>Serratia</i> sp. DD3		
A0A0X2NUC6	<i>Serratia</i> sp. LCN16		
A0A2T3WMN9	<i>Serratia</i> sp. Nf2		
A0A2G7ZRT4	<i>Serratia</i> sp. OLAL2		
A0A1B3F8V0	<i>Serratia</i> sp. YD25		
I6EU12	<i>Shigella boydii</i> 4444-74		
F3WEG2	<i>Shigella boydii</i> 5216-82		
E7SUV1	<i>Shigella boydii</i> ATCC 9905		
B2U3Y7	<i>Shigella boydii</i> serotype 18 (strain CDC 3083-94 / B5512)		
Q32SL3	<i>Shigella boydii</i> serotype 4 (strain Sb227)		
E2XC03	<i>Shigella dysenteriae</i> 1617		
I6G7N8	<i>Shigella dysenteriae</i> 225-75		
E7SNJ8	<i>Shigella dysenteriae</i> CDC 74-1112		
Q32ID4	<i>Shigella dysenteriae</i> serotype 1 (strain Sd197)		
A0A090NBS4	<i>Shigella dysenteriae</i> WRSd3		
I6C2Z9	<i>Shigella flexneri</i> 2850-71		
A0A226K605	<i>Shigella flexneri</i> 2a str. 301		
A0A127GGY3	<i>Shigella flexneri</i> 4c		
A0A0F6MA83	<i>Shigella flexneri</i> 5a str. M90T		
E7T6V8	<i>Shigella flexneri</i> CDC 796-83		
FSNQD0	<i>Shigella flexneri</i> K-227		
Q0T7K3	<i>Shigella flexneri</i> serotype 5b (strain 8401)		
D2A8D5	<i>Shigella flexneri</i> serotype X (strain 2002017)		
FSMY23	<i>Shigella flexneri</i> VA-6		
A0A200LH78	<i>Shigella sonnei</i>		
A0A1E2VIN2	<i>Shigella</i> sp. FC2928		
A0A1E3MZZ6	<i>Shigella</i> sp. FC569		
A0A142GZS8	<i>Shigella</i> sp. PAMC 28760		
A0A0M3E341	<i>Vibrio parahaemolyticus</i>		
A0A0T9UD15	<i>Yersinia aldovae</i>		
A0A2G4U259	<i>Yersinia bercovieri</i>		
A0A2K7ZF98	<i>Yersinia enterocolitica</i> LC20		
A1JPK6	<i>Yersinia enterocolitica</i> serotype O:8 / biotype 1B (strain NCTC 13174 / 8081)		
A0A0T9UU77	<i>Yersinia frederiksenii</i>		
A0A209A887	<i>Yersinia intermedia</i>		
A0A0T9L7D5	<i>Yersinia kristensenii</i>		
A0A2R4NS15	<i>Yersinia massiliensis</i>		
A0A0U113H5	<i>Yersinia mollaretii</i>		
A0A0T9KHZ6	<i>Yersinia nurmii</i>		
A0A0U1HUS8	<i>Yersinia rohdei</i>		
A0A085USN6	<i>Yersinia ruckeri</i>		
A0A2I2B7Z4	<i>Yokenella regensburgei</i>		

**Tabelle S3:** 203 nichtredundante Orthologe aus der Uni Prot Datenbank „Reference Proteomes“ mit mindestens 44% Übereinstimmung und höchstens 10% Sequenzlängenabweichung zu Lon aus *E. coli* K-12.

Uni Prot ID	Spezies	Uni Prot ID	Spezies
A0A2H9U7T8	<i>Aeromonas cavemicola</i>	A0A1S8YF12	<i>Izhakiella australiensis</i>
A0KJU3	<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i> (strain ATCC 7966 / DSM 30187 / KCTC 2358)	A0A0H3FS34	<i>Klebsiella aerogenes</i> (strain ATCC 13048 / DSM 30053 / NBRC 13534 / NCIMB 10102)
RIH7V1	<i>Aeromonas molluscorum</i> 848	A0A0H3H9I7	<i>Klebsiella oxytoca</i> (strain ATCC 8724 / DSM 4798 / JCM 20051 / NRRL B-199 / KCTC 1686)
R9PNM3	<i>Agarivorans albus</i> MKT 106	W1HVX8	<i>Klebsiella pneumoniae</i> IS39
A0A2N0WRV0	Alteromonadales bacterium alter-6D02	W1E6L4	<i>Klebsiella pneumoniae</i> IS46
A0A075P131	<i>Alteromonas australica</i>	W1H174	<i>Klebsiella pneumoniae</i> ISC21
A0A1E7Z7Y8	<i>Alteromonas confluentis</i>	A6T5I2	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i> (strain ATCC 700721 / MGH 78578)
K0CZS8	<i>Alteromonas nacleodii</i> (strain Black Sea 11)	A0A0H3GJ60	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i> (strain HS11286)
F5ZC46	<i>Alteromonas naphthalenivorans</i>	A0A0L0AN15	<i>Klebsiella</i> sp. RIT-PI-d
A0A2A6AQZ9	<i>Alteromonas pelagiomontana</i>	A0A085I687	<i>Kluyvera ascorbata</i> ATCC 33433
A0A1J0TFZ2	<i>Alteromonas</i> sp. RW2A1	A0A071LTY2	<i>Mangrovibacter</i> sp. MFB070
HSUYG5	<i>Atlantibacter hermannii</i> NBRC 105704	A0A0N1KUJ3	<i>Moellerella wisconsensis</i> ATCC 35017
A0A08YHN6	bacteria symbiont Bfo I of <i>Frankliniella occidentalis</i>	F7XY00	<i>Moranelia endobia</i> (strain PCTT)
A0A2AST0E4	bacterium symbiont of <i>Melanocetus johnsonii</i>	A0A090L1P1	<i>Moritella viscosa</i>
Q1LTK1	<i>Baumannia cicadellinicola</i> subsp. <i>Homolodisca</i> coagulata	H2FU93	<i>Oceanimonas</i> sp. (strain GK1)
A0A0A2WDM7	<i>Beauveria bassiana</i> D1-5	A0A1W6B196	<i>Pantoea alhagi</i>
A0A0E3Y936	<i>Blochmannia endosymbiont</i> of <i>Camponotus (Colobopsis) obliquus</i>	D4GLB9	<i>Pantoea ananatis</i> (strain LMG20103)
A0A0E3Y759	<i>Blochmannia endosymbiont</i> of <i>Polyrhachis (Hedomyza) turneri</i>	A0A1V9DID9	<i>Pantoea latae</i>
Q7VRB9	<i>Blochmannia floridiana</i>	U2LA G3	<i>Pantoea</i> sp. AS-PWVM4
P57549	<i>Buchnera aphidicola</i> subsp. <i>Acyrtosiphon pisum</i> (strain APS)	A0A0A0Z874	<i>Pantoea</i> sp. PSNIH1
Q89A99	<i>Buchnera aphidicola</i> subsp. <i>Baizongia pistaciae</i> (strain Bp)	H3R9B0	<i>Pantoea stewartii</i> subsp. <i>stewartii</i> DC283
Q057E4	<i>Buchnera aphidicola</i> subsp. <i>Cinara cedri</i> (strain Cc)	A0A0U3THJ9	<i>Pantoea vagans</i>
A0A172WDX8	<i>Buchnera aphidicola</i> subsp. <i>Schlechtendalia chinensis</i>	E1SF98	<i>Pantoea vagans</i> (strain C9-1) ( <i>Pantoea agglomerans</i> (strain C9-1))
A0A085G137	<i>Buttiauxella agrestis</i> ATCC 33320	A0A1X1DC65	<i>Pantoea wallisii</i>
A0A0K2BK68	Candidatus <i>Baumannia cicadellinicola</i>	C7BH88	<i>Photorhabdus asymbiotica</i> subsp. <i>asymbiotica</i> (strain ATCC 43949 / 3105-77)
A0A143W575	Candidatus <i>Doolittlea endobia</i>	Q7N0L5	<i>Photorhabdus luminescens</i> subsp. <i>laumondii</i> (strain DSM 15139 / CIP 105565 / TT01)
S3EHH3	Candidatus <i>Photodesmus katoptron</i> Akat1	A0A0CG3QF89	<i>Phytobacter ursingii</i>
EDW5H0	Candidatus <i>Regiella insecticola</i> LSR1	U3U670	<i>Plautia stali</i> symbiont
A0A0C1S030	Candidatus <i>Riesia pediculuschaeffi</i> PTSU	A0A089PRJ4	<i>Pluralibacter gergoviae</i> ( <i>Enterobacter gergoviae</i> )
W7QIZ6	<i>Catenovulum agarivorans</i> DS-2	A0A090V5P5	<i>Pseudesccherichia vulneris</i> NBRC 102420
A0A0I8GSM3	<i>Catenovulum maritimum</i>	A0A111QQC4	<i>Pseudoalteromonas denitrificans</i> DSM 6059
W0LJC6	<i>Chania multitudineis</i> RB-25	L8DE29	<i>Pseudoalteromonas luteoviolacea</i> B = ATCC 29581
A0A2N6CI08	Chromatiales bacterium	A0A269PLC2	<i>Pseudoalteromonas</i> sp. NBT06-2
A8AK14	<i>Citrobacter koseri</i> (strain ATCC BAA-895 / CDC 4225-83 / SC5C4696)	H2IZL5	<i>Rahnella aquatilis</i> (strain ATCC 33071 / DSM 4594 / JCM 1683 / CIP 78.65)
A0A0V9YAO	<i>Citrobacter</i> sp. 50677481	A0A0B5IHM0	<i>Raoultella ornithinolytica</i> ( <i>Klebsiella ornithinolytica</i> )
A0A090TY91	<i>Citrobacter werkmanii</i> NBRC 105721	D4G7X0	<i>Riesia pedicicola</i> (strain USDA)
A0A1H7GNW3	<i>Colwellia chukchiensis</i>	A9MM21	<i>Salmonella arizonae</i> (strain ATCC BAA-731 / CDC346-86 / RSK2980)
Q47XM0	<i>Colwellia psycherythraea</i> (strain 34H / ATCC BAA-681) ( <i>Vibrio psycherythraeus</i> )	Q8ZR89	<i>Salmonella typhimurium</i> (strain LT2 / SC5C1412 / ATCC 700720)
A0A2N1UGC6	<i>Colwellia</i> sp. 12G3	A0A0F7HCJ2	<i>Serratia fonticola</i>
A0A0S2JHT5	<i>Colwellia</i> sp. MT41	D4DW99	<i>Serratia odorifera</i> DSM 4582
A0A1D8RSJ1	<i>Colwellia</i> sp. PAMC 20917	A0A087L0A5	<i>Serratia</i> sp. Ag1
A0A1V0KU14	<i>Colwellia</i> sp. PAMC 21821	A0A084A5H2	<i>Serratia</i> sp. DD3
A7MF16	<i>Cronobacter sakazakii</i> (strain ATCC BAA-894) ( <i>Enterobacter sakazakii</i> )	L0MDH5	<i>Serratia</i> sp. FG94
A0A1D8MZZ2	<i>Edwardsiella hoshinae</i>	A0A0Q4MDN6	<i>Serratia</i> sp. Leaf50
Q8LPU9	<i>Enterobacter cloacae</i> EcW5U1	Q12LA3	<i>Shewanella denitrificans</i> (strain OS217 / ATCC BAA-1090 / DSM 15013)
V5A670	<i>Enterobacter cloacae</i> S611	Q07ZY0	<i>Shewanella frigidimarina</i> (strain NCIMB 400)
E3G6W6	<i>Enterobacter lignolyticus</i> (strain SCF1)	A8FT11	<i>Shewanella sediminis</i> (strain HAW-EB3)
A0A28SINE9	<i>Enterobacter</i> sp. CC120223-11	A0A2N7DK28	<i>Shewanella</i> sp. 10N.286.48.B5
A0A2I0FSW0	Enterobacterales bacterium CwR94	A0A2I0DWN8	<i>Shewanella</i> sp. GutDb-MelDb
L0M5Y0	Enterobacteriaceae bacterium (strain FGI 57)	A0A0Q0LZQ0	<i>Shewanella</i> sp. P1-14-1
D8MNY4	<i>Erwinia billingiae</i> (strain Eb661)	Q83M54	<i>Shigella flexneri</i>
A0A0U5E7P5	<i>Erwinia gerundensis</i>	I2BBP9	<i>Shimwellia blattae</i> (strain ATCC 29907 / DSM 4481 / CDC 9005-74)
A0A0L7TGE9	<i>Erwinia imiecta</i>	Q2NV77	<i>Sodalis glossinidius</i> (strain morsitans)
A0A0I4NT00	<i>Erwinia mallochora</i>	A0A0957EU2	<i>Tatumella morbirosei</i>
A0A0Q4MFZ9	<i>Erwinia</i> sp. Leaf53	A0A085JAS0	<i>Tatumella ptyseos</i> ATCC 33301
A0A2G8EMM9	<i>Erwinia</i> sp. OLSSP12	A0A0D8CY54	<i>Thalassomonas actinurum</i>
B2VHU3	<i>Erwinia tasmaniensis</i> (strain DSM 17950 / CIP 109463 / E1199)	A0A099LFP5	<i>Thalassotalea</i> sp. ND16A
A0A0M2K4M9	<i>Erwinia tracheiphila</i>	C4LDB5	<i>Tolomonas auensis</i> (strain DSM 9187 / TA4)
P0A9M0	<i>Escherichia coli</i> (strain K12)	A0A1M5X4G9	<i>Vibrio aerogenes</i> CECT 7868
A0A152UUM3	<i>Escherichia coli</i> O157:H7	F7YLE8	<i>Vibrio anguillarum</i> (strain ATCC 68554 / 775) ( <i>Listonella anguillarum</i> )
A0A085G788	<i>Ewingella americana</i> ATCC 33852	U3CGP8	<i>Vibrio azureus</i> NBRC 104587
A0A1G8XXZ7	<i>Ferrimonas sediminum</i>	A0A177XZB3	<i>Vibrio bivalvicida</i>
A0A0I8YFJ3	<i>Franconibacter pulveris</i>	A0A193KAF0	<i>Vibrio breoganii</i>
K2J071	<i>Gallaecimonas xiamenensis</i> 3-C-1	A7MV83	<i>Vibrio campbellii</i> (strain ATCC BAA-1116 / BB120)
A0A250AYV1	<i>Gibbsiella quercinecans</i>	E3BHX7	<i>Vibrio caribbeanicus</i> ATCC BAA-2122
G4QK64	<i>Glaciecola nitratireducens</i> (strain JCM 12485 / KCTC 12276 / FR1064)	A0A1T4PQL0	<i>Vibrio cincinnatiensis</i> DSM 19608
K6YSM0	<i>Glaciecola pallidula</i> DSM 14239 = ACAM 615	A0A1B1V7Q7	<i>Vibrio coralliculticus</i>
H5T7Z9	<i>Glaciecola punicea</i> ACAM 611	A0A2J8GUG6	<i>Vibrio diazotrophicus</i>
A0A1A9UK58	<i>Glossina austeni</i> (Savannah tsetse fly)	Q5E6Q3	<i>Vibrio fischeri</i> (strain ATCC 700601 / ES114)
A0A1A9X6F9	<i>Glossina fuscipes fuscipes</i> (Riverine tsetse fly)	F0LPI3	<i>Vibrio furnissii</i> (strain DSM 14383 / NCTC 11218)
A0A1B0C1S6	<i>Glossina palpalis gambiensis</i>	A0A0F4NT13	<i>Vibrio galathea</i>
A0A097R554	<i>Hafnia alvei</i> FB1	A0A1E5D151	<i>Vibrio genomsp. F6 str. FF-238</i>
A0A094HNL7	<i>Idiomarina atlantica</i>	A0A1Y6IPI3	<i>Vibrio mangrovi</i>
A0A1G6DNA9	<i>Idiomarina indica</i>	C9P5T0	<i>Vibrio metschnikovii</i> CIP 69.14
Q5QXP0	<i>Idiomarina lohiensis</i> (strain ATCC BAA-735 / DSM 15497 / L2-TR)	A0A0C3ICZ8	<i>Vibrio mytili</i>
A0A1I6GMQ9	<i>Idiomarina maritima</i>	A0A0M0HHR0	<i>Vibrio nereis</i>
A0A241R4S8	<i>Idiomarina piscisalsi</i>	U4JZR1	<i>Vibrio nigripulchritudo</i>
A0A1Y6EM56	<i>Idiomarina planktonica</i>	A0A1R4B1P9	<i>Vibrio palustris</i>
A0A094UJ7	<i>Idiomarina salinarum</i>	A0A1Q9HJU3	<i>Vibrio panuliri</i>
A0A0W1SDH7	<i>Idiomarina</i> sp. H105	P74956	<i>Vibrio parahaemolyticus</i> serotype O3:K6 (strain RIMD 2210633)
A0A0K6GY82	<i>Idiomarina woesei</i>	A0A090PBH5	<i>Vibrio ponticus</i>
K2KB65	<i>Idiomarina xiamenensis</i> 10-D-4	U2ZE85	<i>Vibrio proteolyticus</i> NBRC 13287

Uni Prot ID	Spezies	Uni Prot ID	Spezies
A0A1E5E3Q1	<i>Vibrio rumoiensis</i> IS-45		
A0A0A5JJ9	<i>Vibrio sinoaloensis</i>		
F9RHK4	<i>Vibrio</i> sp. (strain N418)		
A0A2N7GC62	<i>Vibrio</i> sp. 10N.261.55.A7		
A0A2N7DC76	<i>Vibrio</i> sp. 10N.286.49.B3		
A0A2N7D1D3	<i>Vibrio</i> sp. 10N.286.49.C2		
A0A090R9Z6	<i>Vibrio</i> sp. C7		
A0A2A9F432	<i>Vibrio</i> sp. ES.051		
A0A0W8JHQ2	<i>Vibrio</i> sp. MEBIC08052		
A0A2C6MLV9	<i>Vibrio</i> sp. PID17_43		
A0A1X1MPJ7	<i>Vibrio</i> sp. qd031		
A0A1A6L020	<i>Vibrio</i> sp. UCD-FRSSP16_10		
A0A2N0XR58	<i>Vibrio</i> sp. vnigr-6D03		
A0A2N8ZDK0	<i>Vibrio tapetis</i> subsp. tapetis		
B7VHZ8	<i>Vibrio tasmaniensis</i> (strain LCP32) ( <i>Vibrio splendidus</i> (strain Mel32))		
Q7MMG5	<i>Vibrio vulnificus</i> (strain YJ016)		
A0A1G7ZBV4	<i>Vibrio xiamenensis</i>		
A0A0M0IRI5	<i>Vibrio xuii</i>		
Q8D348	<i>Wigglesworthia glossinidia brevipalpis</i>		
A0A1Y2STX4	<i>Xenorhabdus beddingii</i>		
D3VOC8	<i>Xenorhabdus bovienii</i> (strain SS-2004)		
W1IUD9	<i>Xenorhabdus cabanillasii</i> JM26		
A0A068QPK8	<i>Xenorhabdus doucetiae</i>		
A0A2G0Q8Q9	<i>Xenorhabdus hominickii</i>		
A0A1N6MTN5	<i>Xenorhabdus innexi</i>		
A0A0J5ISY6	<i>Xenorhabdus khoisanae</i>		
A0A1H3Q9F7	<i>Xenorhabdus mauleonii</i>		
D3VL94	<i>Xenorhabdus nematophila</i> (strain ATCC 19061 / DSM 3370 / LMG 1036 / NCIB 9965 / AN6)		
A0A068R1X1	<i>Xenorhabdus poinarii</i> G6		
A0A2D0KVL6	<i>Xenorhabdus stockiae</i>		
W1J054	<i>Xenorhabdus szentirmai</i> DSM 16338		
A0A1Q5U5J1	<i>Xenorhabdus thuongxuanensis</i>		
A0A1Y2SK35	<i>Xenorhabdus vietnamensis</i>		
C4S7R0	<i>Yersinia mollaretii</i> ATCC 43969		
Q0WCC8	<i>Yersinia pestis</i>		
A0A085USA8	<i>Yersinia ruckeri</i>		
A0A2N5EQP6	Yersiniaceae bacterium 2016Iso3		
G9Z494	<i>Yokenella regensburgei</i> ATCC 43003		
A0A231MY60	<i>Zobellella denitrificans</i>		

**Tabelle S4:** 183 nichtredundante Orthologe aus der Uni Prot Datenbank „Microbial Proteomes“ mit mindestens 44% Übereinstimmung und höchstens 10% Sequenzlängenabweichung zu EvgS aus *E. coli* K-12.

Uni Prot ID	Spezies	Uni Prot ID	Spezies
A0A0M7P199	Achromobacter sp.	V8K1X7	Escherichia coli LAU-EC10
W0LJ93	Chania multidisensitens RB-25	A0A1X3I4M0	Escherichia coli M056
A0A2J7P3D5	Enterobacter cancerogenus	A0A0K9TD41	Escherichia coli M114
A0A2I6E5V3	Enterobacter cancerogenus	F4T1G8	Escherichia coli M605
B7LBS3	Escherichia coli (strain 55989 / EAEC)	F4TH46	Escherichia coli M718
E0IXQ2	Escherichia coli (strain ATCC 9637 / CCM 2024 / DSM 1116 / NCIMB 8666 / NRRL B-766 / W)	E9YQL9	Escherichia coli M863
A0A140N502	Escherichia coli (strain B / BL21-DE3)	M9FUQ0	Escherichia coli MP021561.2
P30855	Escherichia coli (strain K12)	D7Y3V7	Escherichia coli MS 115-1
D8A5X9	Escherichia coli (strain MS 21-1)	E9TL35	Escherichia coli MS 117-3
B6I6S1	Escherichia coli (strain SE11)	D8E2V6	Escherichia coli MS 119-7
B1LMG6	Escherichia coli (strain SMS-3-5 / SECEC)	E1J8U4	Escherichia coli MS 124-1
A0A0E2L3T1	Escherichia coli (strain UMEA 3162-1)	E1IV22	Escherichia coli MS 145-7
Q1R8Z6	Escherichia coli (strain UT189 / UPEC)	D7XF46	Escherichia coli MS 198-1
I2SY04	Escherichia coli 1.2264	D7ZGT9	Escherichia coli MS 69-1
I2RIC9	Escherichia coli 1.2741	D7XJQ1	Escherichia coli MS 84-1
U9Y7B4	Escherichia coli 110957	E6BH36	Escherichia coli MS 85-1
A0A125X114	Escherichia coli 1-110-08_S3_C1	A0A0H2Z166	Escherichia coli O1:K1 / APEC
U9Y8B8	Escherichia coli 113290	A0A2S1V5F4	Escherichia coli O103 str. RM8385
U9XLP8	Escherichia coli 113303	C8U7Q6	Escherichia coli O103:H2 (strain 12009 / EHEC)
A0A080IRY5	Escherichia coli 1-250-04_S3_C2	A0A0E0XZS1	Escherichia coli O104:H4 (strain 2011C-3493)
A0A0E1LZS9	Escherichia coli 1303	C8UDV4	Escherichia coli O111:H- (strain 11128 / EHEC)
A0A074IF92	Escherichia coli 1-392-07_S4_C1	K4VT98	Escherichia coli O111:H11 str. CVM9455
A0A080FP10	Escherichia coli 1-392-07_S4_C3	K4WFX5	Escherichia coli O111:H8 str. CVM9634
I2XDG8	Escherichia coli 2.3916	A0A027TYX3	Escherichia coli O111:NM str. K6722
A0A029IQR5	Escherichia coli 2-205-03_S4_C2	A0A028EE26	Escherichia coli O118:H16 str. 2009C-4446
A0A029HU22	Escherichia coli 2-205-03_S3_C3	A0A2S1URS0	Escherichia coli O121 str. RM8352
A0A070VBK2	Escherichia coli 2-177-06_S3_C2	A0A0E2TYN7	Escherichia coli O121:H19 str. 2010C-3609
A0A070SXZ8	Escherichia coli 2-210-07_S3_C3	B7UG80	Escherichia coli O127:H6 (strain E2348/69 / EPEC)
E3XQF7	Escherichia coli 2362-75	A0A070FHQ2	Escherichia coli O128:H2 str. 2011C-3317
A0A073G5I0	Escherichia coli 2-427-07_S4_C3	A7ZPH6	Escherichia coli O139:H28 (strain E24377A / ETEC)
A0A069XSM9	Escherichia coli 3-373-03_S4_C2	A0A023Z0J5	Escherichia coli O145:H28 str. RM12581
I2USV4	Escherichia coli 4.0522	A0A025CY49	Escherichia coli O145:NM str. 2010C-3526
I2WVK0	Escherichia coli 4.0967	A0A1Z3V1C2	Escherichia coli O157
A0A0E1SYJ1	Escherichia coli 53638	P58402	Escherichia coli O157:H7
A0A073UA2	Escherichia coli 5-366-08_S1_C1	A0A0H3PXV8	Escherichia coli O157:H7 (strain EC869)
A0A073HBYS	Escherichia coli 5-366-08_S1_C3	A0A0F6FA92	Escherichia coli O157:H7 str. S552
I2WGW5	Escherichia coli 9.0111	B7N5X0	Escherichia coli O17:K52:H18 (strain UMN026 / EPEC)
V0SE95	Escherichia coli 907672	A0A026UKE8	Escherichia coli O174:H8 str. 04-3038
U9Z1Z2	Escherichia coli 907713	A0A192C9B9	Escherichia coli O25b:H4
V0XQ83	Escherichia coli 908525	W8ZUA5	Escherichia coli O25b:H4-ST131
V0Z9D8	Escherichia coli 908573	A0A2S1VN18	Escherichia coli O26 str. RM10386
U9ZWT8	Escherichia coli 909945-2	A0A0A8UBP0	Escherichia coli O26:H11
I2S203	Escherichia coli 97.0246	B7MH30	Escherichia coli O45:K1 (strain S88 / EPEC)
V6FX22	Escherichia coli 99.0741	D3QLW5	Escherichia coli O55:H7 (strain CB9613 / EPEC)
V8FTJ8	Escherichia coli ATCC BAA-2209	A0A0H2V999	Escherichia coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC)
D6HZF1	Escherichia coli B088	A0A028ATU2	Escherichia coli O69:H11 str. 08-4661
D6I9U8	Escherichia coli B185	A0A0H3MSK3	Escherichia coli O7:K1 (strain IA139 / EPEC)
D6ID98	Escherichia coli B354	A0A0E0V8E8	Escherichia coli O7:K1 str. CE10
J7QNQ3	Escherichia coli ch7122	E3PEX7	Escherichia coli O78:H11 (strain H10407 / ETEC)
F4NCX4	Escherichia coli D9	B7MY29	Escherichia coli O81 (strain ED1a)
H4IDT6	Escherichia coli DEC1B	A0A0H3EMQ3	Escherichia coli O83:H1 (strain NRG 857C / AIEC)
H4IUU6	Escherichia coli DEC1C	A0A0G3K7R9	Escherichia coli PCN033
H4JAT8	Escherichia coli DEC1D	A0A1X3LGI8	Escherichia coli TA054
H4KJN8	Escherichia coli DEC2C	A0A1X3M0I3	Escherichia coli TA249
H4LF03	Escherichia coli DEC2E	A0A1X3I2T7	Escherichia coli TA447
H4ULZ3	Escherichia coli DEC6A	E9XKN5	Escherichia coli TW10509
W1WW76	Escherichia coli DORA_A_5_14_21	T8ZZW8	Escherichia coli UMEA 3200-1
A0A1X3IL12	Escherichia coli E1114	T9DUY4	Escherichia coli UMEA 3212-1
A0A0A0FDL8	Escherichia coli G3/10	T9SAN6	Escherichia coli UMEA 3718-1
A0A1X3FB1	Escherichia coli H386	G0FEV9	Escherichia coli UMNFI8
A0A1X3KZY5	Escherichia coli H420	A0A0E0U1M0	Escherichia coli UMNK88
A0A1X3KE80	Escherichia coli H461	I0VT02	Escherichia coli W26
F4VEI3	Escherichia coli H591	A0A0F6C766	Escherichia coli Xuzhou21
F4SH15	Escherichia coli H736	A0A085H3Y0	Hafnia alvei ATCC 13337
V2QXK7	Escherichia coli HVH 50 (4-2593475)	A0A097R8F0	Hafnia alvei FB1
T6HL25	Escherichia coli HVH 70 (4-2963531)	A0A181WQT7	Klebsiella oxytoca
T6MBO0	Escherichia coli HVH 87 (4-5977630)	A0A0L0ARCI	Klebsiella sp. RT-P1-d
L2UWT8	Escherichia coli KTE10	A0A1Q5VC96	Serratia fonticola
S1H550	Escherichia coli KTE100	A0A1C7WQ24	Serratia sp. 14-2641
S1I2V8	Escherichia coli KTE107	I6DXX3	Shigella boydii 4444-74
S1HRU7	Escherichia coli KTE108	F3WKA7	Shigella boydii 5216-82
L4V9E6	Escherichia coli KTE112	E7SY87	Shigella boydii ATCC 9905
L4J4C0	Escherichia coli KTE146	B2TW99	Shigella boydii serotype 18 (strain CDC 3083-94 / BS512)
S1P767	Escherichia coli KTE182	Q31YA1	Shigella boydii serotype 4 (strain Sb227)
L3C6K8	Escherichia coli KTE193	B3XOG4	Shigella dysenteriae 1012
SOVXP7	Escherichia coli KTE20	F3V7Z2	Shigella dysenteriae 155-74
SOYIC9	Escherichia coli KTE37	F3VB78	Shigella dysenteriae 155-74
SOZ6Z9	Escherichia coli KTE38	E2XDV0	Shigella dysenteriae 1617
S1DE60	Escherichia coli KTE64	A0A0A7A2V2	Shigella dysenteriae 1617
L3P598	Escherichia coli KTE66	I6FPU8	Shigella dysenteriae 225-75
S1F2B8	Escherichia coli KTE73	E7SP95	Shigella dysenteriae CDC 74-1112
L3Q243	Escherichia coli KTE75	Q328A1	Shigella dysenteriae serotype 1 (strain Sd197)

Uni Prot ID	Spezies
A0A090N9Z1	<i>Shigella dysenteriae</i> WRSd3
K0X0P3	<i>Shigella flexneri</i> 1485-80
I6BH47	<i>Shigella flexneri</i> 2850-71
A0A226KD62	<i>Shigella flexneri</i> 2a str. 301
A0A127GTQ3	<i>Shigella flexneri</i> 4c
A0A0F6MF18	<i>Shigella flexneri</i> 5a str. M90T
E7TDE3	<i>Shigella flexneri</i> CDC 796-83
FSNYD8	<i>Shigella flexneri</i> K-227
I6CIJ3	<i>Shigella flexneri</i> K-315
Q0T2C7	<i>Shigella flexneri</i> serotype 5b (strain 8401)
D2AG60	<i>Shigella flexneri</i> serotype X (strain 2002017)
FSN480	<i>Shigella flexneri</i> VA-6
Q3YZG0	<i>Shigella sonnei</i> (strain Ss046)
A0A142H921	<i>Shigella</i> sp. PAMC 28760
A0A209AX16	<i>Yersinia frederiksenii</i>
A0A209A2J9	<i>Yersinia frederiksenii</i>
A0A0T9N341	<i>Yersinia intermedia</i>
A0A2R4NMP0	<i>Yersinia massiliensis</i>
A0A0U1HYP4	<i>Yersinia rohdei</i>

Uni Prot ID	Spezies
-------------	---------

**Tabelle S5:** 793 nichtredundante Orthologe aus der Uni Prot Datenbank „Microbial Proteomes“ mit mindestens 44% Übereinstimmung und höchstens 10% Sequenzlängenabweichung zu EnvZ aus *E. coli* K-12.

Uni Prot ID	Spezies	Uni Prot ID	Spezies
A0A0M7GA Y0	<i>Achromobacter</i> sp.	A0A0X4EQ51	<i>Enterobacter cloacae</i> complex sp. GN04363
A0A0M9J588	<i>Achromobacter</i> sp.	A0A16IZRV5	<i>Enterobacter cloacae</i> complex sp. GN04787
A0A1B9PA19	<i>Aliivibrio fischeri</i>	A0A162TYJ9	<i>Enterobacter cloacae</i> complex sp. GN06232
A0A1ESALM3	<i>Aliivibrio fischeri</i>	A0A1C7ZNR3	<i>Enterobacter cloacae</i> complex sp. SMART_763
A0A1B9NVK1	<i>Aliivibrio logei</i>	G8LJB0	<i>Enterobacter cloacae</i> EcWSU1
B6EPQ4	<i>Aliivibrio salmonicida</i> (strain LFI1238)	W0BJJ3	<i>Enterobacter cloacae</i> P101
A0A1B9PQT3	<i>Aliivibrio</i> sp. IS175	V5AVK0	<i>Enterobacter cloacae</i> S611
A0A090IIC3	<i>Aliivibrio wodanis</i>	A0A0F0T8X8	<i>Enterobacter cloacae</i> subsp. cloacae
H5V3D5	<i>Atlantibacter hermannii</i> NBRC 105704	A0A246P6C8	<i>Enterobacter cloacae</i> subsp. cloacae
A0A0J9B555	bacteria symbiont Bfo1 of <i>Frankliniella occidentalis</i>	A0A0H3CUG0	<i>Enterobacter cloacae</i> subsp. cloacae (strain ATCC 13047 / NCDC 27956)
A0A0I8YM61	bacteria symbiont Bfo2 of <i>Frankliniella occidentalis</i>	A0A0M1R9K0	<i>Enterobacter cloacae</i> subsp. dissolvens SDM
A0A1Q4P9N5	BMT49_24555	V3E041	<i>Enterobacter cloacae</i> UCICRE 12
A0A0G4JUL3	<i>Brenneria goodwinii</i>	A0A0M2BWS8	<i>Enterobacter cloacae</i> UCICRE 3
G7LV48	<i>Brenneria</i> sp. EniD312	A0A246L8K8	<i>Enterobacter hormaechei</i>
A0A2C6DGK3	<i>Budvicia aquatica</i>	F5S2X8	<i>Enterobacter hormaechei</i> ATCC 49162
A0A085QW4	<i>Buttiauxella agrestis</i> ATCC 33320	A0A0W1YL96	<i>Enterobacter hormaechei</i> subsp. ohanae
A0A1B7IHA3	<i>Buttiauxella brenneriae</i> ATCC 51605	A0A0EIJHG2	<i>Enterobacter hormaechei</i> subsp. steigerwaltii
A0A1B7IKM1	<i>Buttiauxella ferruginae</i> ATCC 51602	A0A0F0YAB7	<i>Enterobacter kobei</i>
A0A1B7IQ07	<i>Buttiauxella gaviniae</i> ATCC 51604	E3GB28	<i>Enterobacter lignolyticus</i> (strain SCF1)
A0A1B7IUG8	<i>Buttiauxella noackiae</i> ATCC 51607	A0A198GZX1	<i>Enterobacter soli</i> ATCC BAA2102
W0HH31	<i>Candidatus Sodalis pierantonius</i> str. SOPE	A4WFJ2	<i>Enterobacter</i> sp. (strain 638)
S3ILA2	<i>Cedecea davisae</i> DSM 4568	A0A0V9DYB1	<i>Enterobacter</i> sp. 50588862
A0A089Q5D7	<i>Cedecea neteri</i>	J0MGX4	<i>Enterobacter</i> sp. Ag1
A0A1X7IMP0	<i>Cedecea</i> sp. NFXS7	A0A0I1VEP2	<i>Enterobacter</i> sp. BIDMC92
W0LQJ4	<i>Chania multitudinisentis</i> RB25	W7NNK9	<i>Enterobacter</i> sp. DC4
A0A083ZGK7	<i>Citrobacter amalonaticus</i>	A0A127KX40	<i>Enterobacter</i> sp. FY07
A0A0F6TXL3	<i>Citrobacter amalonaticus</i> Y19	A0A0I0NTS3	<i>Enterobacter</i> sp. GN02600
A0A1R0FW10	<i>Citrobacter braakii</i>	A0A1E9Y2J7	<i>Enterobacter</i> sp. HMSC055A11
A0A023V4M6	<i>Citrobacter freundii</i> CFNIH1	A0A1E5MPC8	<i>Enterobacter</i> sp. kubf2
A0A064D5S0	<i>Citrobacter freundii</i> MGH 56	A0A0E2M624	<i>Enterobacter</i> sp. MGH 14
A8AQV8	<i>Citrobacter koseri</i> (strain ATCC BAA895 / CDC 422583 / SGSC4696)	V3P7Q5	<i>Enterobacter</i> sp. MGH 24
A0A0A1RGI3	<i>Citrobacter pasteurii</i>	A0A1K1ZDN2	<i>Enterobacter</i> sp. NFX03
D2TLN5	<i>Citrobacter rodentium</i> (strain ICC168)	A0A1U6HXL5	<i>Enterobacter</i> sp. NFR05
A0A0V9J7B1	<i>Citrobacter</i> sp. S0677481	R9VRP7	<i>Enterobacter</i> sp. R4368
A0A1V4NX48	<i>Citrobacter</i> sp. A316	A0A175SSH7	<i>Enterobacter</i> sp. siangfangensis
R8WM42	<i>Citrobacter</i> sp. KTEI51	L0M0Z1	<i>Enterobacteriaceae</i> bacterium (strain FGI 57)
A0A0I1J5B2	<i>Citrobacter</i> sp. MGH100	E5YM27	<i>Enterobacteriaceae</i> bacterium 9_2_54FAA
A0A0I1KZP9	<i>Citrobacter</i> sp. MGH103	A0A1E5BMW0	<i>Enterovibrio calviensis</i> FF85
A0A0I1LZX0	<i>Citrobacter</i> sp. MGH105	A0A135IB75	<i>Enterovibrio coralii</i>
A0A0I1JL93	<i>Citrobacter</i> sp. MGH106	A0A1T4UH36	<i>Enterovibrio nigricans</i> DSM 22720
A0A0I1MPX6	<i>Citrobacter</i> sp. MGH109	A0A1I5PFU2	<i>Enterovibrio norvegicus</i> DSM 15893
A0A1X4HL03	<i>Citrobacter werkmanii</i>	A0A1E5CF52	<i>Enterovibrio norvegicus</i> FF454
A0A090U1N4	<i>Citrobacter werkmanii</i> NBRC 105721	A0A1C3EJ81	<i>Enterovibrio pacificus</i>
D4BI29	<i>Citrobacter youngae</i> ATCC 29220	D4I336	<i>Erwinia amylovora</i> (strain CFBP1430)
A0A198GF5	<i>Cosenzaea myofaciens</i> ATCC 19692	D8MY26	<i>Erwinia billingiae</i> (strain Eb661)
K8A019	<i>Cronobacter condimentii</i> 1330	A0A0U5L125	<i>Erwinia gerundensis</i>
K8B4P8	<i>Cronobacter dublinensis</i> 1210	A0A0L7TEGI	<i>Erwinia iniecta</i>
A0A0K2P6X2	<i>Cronobacter dublinensis</i> subsp. dublinensis LMG 23823	A0A014NAW6	<i>Erwinia mallotivora</i>
K8C3R2	<i>Cronobacter malonaticus</i> 507	V5ZC79	<i>Erwinia piflorinigra</i> CFBP 5888
A0A0U4H387	<i>Cronobacter malonaticus</i> LMG 23826	E3DIQ6	<i>Erwinia</i> sp. (strain Ejp617)
K8A9J3	<i>Cronobacter mytjensii</i> 530	A0A0Q4N4Z9	<i>Erwinia</i> sp. Leaf53
A0A0K2P8H5	<i>Cronobacter mytjensii</i> ATCC 51329	B2VJW0	<i>Erwinia tasmaniensis</i> (strain DSM 17950 / CIP 109463 / E1/99)
A7MG08	<i>Cronobacter sakazakii</i> (strain ATCC BAA894)	A0A0M2KDY6	<i>Erwinia tracheiphila</i>
K8D7W7	<i>Cronobacter sakazakii</i> 696	A0A0A3YWU4	<i>Erwinia typographi</i>
K8CQ24	<i>Cronobacter sakazakii</i> 701	B1E155	<i>Escherichia albertii</i> (strain TW07627)
C9Y288	<i>Cronobacter turicensis</i> (strain DSM 18703 / LMG 23827 / z3032)	A0A1X3PC0	<i>Escherichia albertii</i> B156
K8BCC7	<i>Cronobacter turicensis</i> 564	W0AVU2	<i>Escherichia albertii</i> KF1
K8DCS3	<i>Cronobacter universalis</i> NCTC 9529	A0A1V2GIM7	<i>Escherichia coli</i>
A0A0A3ZJA5	<i>Dickeya chrysanthemi</i>	A0A228XD68	<i>Escherichia coli</i>
C6CH25	<i>Dickeya chrysanthemi</i> (strain Ech1591)	B7L4T4	<i>Escherichia coli</i> (strain 55989 / EAEC)
E0SJT2	<i>Dickeya dadantii</i> (strain 3937)	E0JZW9	<i>Escherichia coli</i> (strain ATCC 9637 / NRRL B766 / W)
C6CZ75	<i>Dickeya paradisiaca</i> (strain Ech703)	A0A140N5C3	<i>Escherichia coli</i> (strain B / BL21DE3)
A0A089V5H0	<i>Dickeya solani</i>	P0AEJ4	<i>Escherichia coli</i> (strain K12)
U6ZD14	<i>Dickeya solani</i> D s04321	D8A6U3	<i>Escherichia coli</i> (strain MS 211)
D2BY53	<i>Dickeya zeae</i> (strain Ech586)	B6L2W6	<i>Escherichia coli</i> (strain SE11)
A0A0A8F11	<i>Dickeya zeae</i> EC1	B1LHK4	<i>Escherichia coli</i> (strain SMS35 / SECEC)
A0A076LMK1	<i>Edwardsiella anguillarum</i> ET080813	A0A0E2L0C4	<i>Escherichia coli</i> (strain UMEA 31621)
A0A1D8MU19	<i>Edwardsiella hoshinae</i>	Q1R5N1	<i>Escherichia coli</i> (strain UT189 / UPEC)
CSBGS5	<i>Edwardsiella ictaluri</i> (strain 93146)	I2SV14	<i>Escherichia coli</i> I.2264
A0A034T3E3	<i>Edwardsiella piscicida</i>	I2RE62	<i>Escherichia coli</i> I.2741
A0A1D7VCF6	<i>Edwardsiella piscicida</i>	U9YGA2	<i>Escherichia coli</i> I10957
D0ZG11	<i>Edwardsiella tarda</i> (strain EIB202)	A0A125WX41	<i>Escherichia coli</i> I11008_S3_C1
A0A0H3DWZ4	<i>Edwardsiella tarda</i> (strain FL660)	U9YGB8	<i>Escherichia coli</i> I13290
M0Q802	<i>Edwardsiella tarda</i> ATCC 1594	U9XQ01	<i>Escherichia coli</i> I13303
D4FA97	<i>Edwardsiella tarda</i> ATCC 23685	A0A0I7I764	<i>Escherichia coli</i> I17605_S3_C2
A0A0X8Q9M0	<i>Enterobacter agglomerans</i>	A0A080IFC4	<i>Escherichia coli</i> I25004_S3_C2
G2S2H5	<i>Enterobacter asburiae</i> (strain LF7a)	A0A0E1M2I3	<i>Escherichia coli</i> I303
D2ZIR2	<i>Enterobacter cancerogenus</i> ATCC 35316	A0A074HRU0	<i>Escherichia coli</i> I39207_S4_C1
A0A1S2AN28	<i>Enterobacter cloacae</i> complex Hoffmann cluster IV	A0A080FNE8	<i>Escherichia coli</i> I39207_S4_C3
A0A0F1AV05	<i>Enterobacter cloacae</i> complex sp. 35699	I2XS94	<i>Escherichia coli</i> I2.3916
A0A0F4BN61	<i>Enterobacter cloacae</i> complex sp. CIDEIMsCOL9	A0A029IU3	<i>Escherichia coli</i> I200503_S4_C2
A0A0X6Z051	<i>Enterobacter cloacae</i> complex sp. GN02468	A0A029HU1	<i>Escherichia coli</i> I200503_S4_C3

Uni Prot ID	Spezies	Uni Prot ID	Spezies
A0A062Y5U5	<i>Escherichia coli</i> 201108_S1_C1	D7ZAG5	<i>Escherichia coli</i> MS 691
A0A070UVW7	<i>Escherichia coli</i> 217706_S3_C2	D7XWB6	<i>Escherichia coli</i> MS 841
A0A070SJ20	<i>Escherichia coli</i> 221007_S3_C3	E6BIF9	<i>Escherichia coli</i> MS 851
E3XSW3	<i>Escherichia coli</i> 236275	A0A0H2Z4E0	<i>Escherichia coli</i> O1:K1 / APEC
A0A073FQR1	<i>Escherichia coli</i> 242707_S4_C3	C8TY20	<i>Escherichia coli</i> O103:H2 (strain 12009 / EHEC)
A0A073GJ49	<i>Escherichia coli</i> 326703_S4_C1	A0A0E0XUU4	<i>Escherichia coli</i> O104:H4 (strain 2011C3493)
A0A069XU89	<i>Escherichia coli</i> 337303_S4_C2	C8UHN5	<i>Escherichia coli</i> O111:H (strain 11128 / EHEC)
I2UJ66	<i>Escherichia coli</i> 4.0522	K4XBG9	<i>Escherichia coli</i> O111:H11 str. CVM9455
I2X1A2	<i>Escherichia coli</i> 4.0967	K4VG05	<i>Escherichia coli</i> O111:H8 str. CVM9634
A0A0E1SZY3	<i>Escherichia coli</i> 53638	A0A027U142	<i>Escherichia coli</i> O111:NM str. K6722
A0A073UPJ2	<i>Escherichia coli</i> 536608_S1_C1	A0A028E932	<i>Escherichia coli</i> O118:H16 str. 2009C4446
A0A073H4V8	<i>Escherichia coli</i> 536608_S1_C3	A0A0E2TQV4	<i>Escherichia coli</i> O121:H19 str. 2010C3609
I4T7Q2	<i>Escherichia coli</i> 54115	B7UKA9	<i>Escherichia coli</i> O127:H6 (strain E2348/69 / EPEC)
A0A080EU35	<i>Escherichia coli</i> 653708_S1_C3	A0A070FCT8	<i>Escherichia coli</i> O128:H2 str. 2011C3317
I2WEM0	<i>Escherichia coli</i> 9.0111	A7ZST2	<i>Escherichia coli</i> O139:H28 (strain E24377A / ETEC)
VUSG68	<i>Escherichia coli</i> 907672	A0A070DFQ0	<i>Escherichia coli</i> O145:H25 str. 073858
U9ZTJ2	<i>Escherichia coli</i> 907713	A0A023Z435	<i>Escherichia coli</i> O145:H28 str. RM12581
V0VJD1	<i>Escherichia coli</i> 908519	A0A025CP53	<i>Escherichia coli</i> O145:NM str. 2010C3526
V0XFK0	<i>Escherichia coli</i> 908525	A0A070DSA7	<i>Escherichia coli</i> O146:H21 str. 2010C3325
V0YI39	<i>Escherichia coli</i> 908573	A0A123UYQ7	<i>Escherichia coli</i> O157
V0AZG6	<i>Escherichia coli</i> 9099452	Q8X731	<i>Escherichia coli</i> O157:H7
I2RSC8	<i>Escherichia coli</i> 97.0246	A0A0H3PTN3	<i>Escherichia coli</i> O157:H7 (strain EC869)
V6FW35	<i>Escherichia coli</i> 99.0741	A0A0F6FD96	<i>Escherichia coli</i> O157:H7 str. SS52
V8F190	<i>Escherichia coli</i> ATCC BAA 2209	B7NE09	<i>Escherichia coli</i> O17:K52:H18 (strain UMN026 / ExPEC)
D6I282	<i>Escherichia coli</i> B088	A0A026V6R9	<i>Escherichia coli</i> O174:H8 str. 043038
D6IEI7	<i>Escherichia coli</i> B185	A0A192CCQ6	<i>Escherichia coli</i> O25b:H4
D6JFY0	<i>Escherichia coli</i> B354	W8ZXF6	<i>Escherichia coli</i> O25b:H4ST131
I7QS12	<i>Escherichia coli</i> ch7122	A0A0A8UF11	<i>Escherichia coli</i> O26:H11
F4NNB8	<i>Escherichia coli</i> D9	D3GZM2	<i>Escherichia coli</i> O44:H18 (strain 042 / EAEC)
H4IH68	<i>Escherichia coli</i> DEC1B	B7MDN0	<i>Escherichia coli</i> O45:K1 (strain S88 / ExPEC)
H4Y47	<i>Escherichia coli</i> DEC1C	D3QTM1	<i>Escherichia coli</i> O55:H7 (strain CB9615 / EPEC)
H4JE07	<i>Escherichia coli</i> DEC1D	A0A0H2VD53	<i>Escherichia coli</i> O6:H1 (strain CFT073 / ATCC 700928 / UPEC)
H4KN01	<i>Escherichia coli</i> DEC2C	A0A028ANP5	<i>Escherichia coli</i> O69:H11 str. 084661
H4LJ76	<i>Escherichia coli</i> DEC2E	A0A0H3MLR3	<i>Escherichia coli</i> O7:K1 (strain IA139 / ExPEC)
H4UQB6	<i>Escherichia coli</i> DEC6A	A0A0E0VBQ6	<i>Escherichia coli</i> O7:K1 str. CE10
A0A1X3IGW1	<i>Escherichia coli</i> E1114	E3PLL1	<i>Escherichia coli</i> O78:H11 (strain H10407 / ETEC)
A0A0A0FHD4	<i>Escherichia coli</i> G3/10	B7N137	<i>Escherichia coli</i> O81 (strain ED1a)
A0A1X3JBD9	<i>Escherichia coli</i> H386	A0A0H3EQK3	<i>Escherichia coli</i> O83:H1 (strain NRG 857C / AIEC)
A0A1X3KX55	<i>Escherichia coli</i> H420	A0A0G3K9V0	<i>Escherichia coli</i> PCN033
A0A1X3K300	<i>Escherichia coli</i> H461	A0A1X3LCR1	<i>Escherichia coli</i> TA054
F4VK40	<i>Escherichia coli</i> H591	A0A1X3LPX9	<i>Escherichia coli</i> TA249
F4SQD5	<i>Escherichia coli</i> H736	F4V7A2	<i>Escherichia coli</i> TA280
T5U9R6	<i>Escherichia coli</i> HHV 25 (45851939)	A0A1X3IUS2	<i>Escherichia coli</i> TA447
T6A1N7	<i>Escherichia coli</i> HHV 41 (42677849)	E9XTB3	<i>Escherichia coli</i> TW10509
V2RTW4	<i>Escherichia coli</i> HHV 50 (42593475)	A0A064T8M5	<i>Escherichia coli</i> UCI 65
T6GKE7	<i>Escherichia coli</i> HHV 70 (42963531)	T9A1C3	<i>Escherichia coli</i> UMEA 32001
T6LSR9	<i>Escherichia coli</i> HHV 87 (45977630)	T9CLF2	<i>Escherichia coli</i> UMEA 32121
W1G424	<i>Escherichia coli</i> ISC11	V2SRN2	<i>Escherichia coli</i> UMEA 33231
W1HIJ6	<i>Escherichia coli</i> ISC56	T9S324	<i>Escherichia coli</i> UMEA 37181
W1ETV2	<i>Escherichia coli</i> ISC7	G0F772	<i>Escherichia coli</i> UMN18
T8KJC4	<i>Escherichia coli</i> KOEJ671 (186a)	A0A0E0U4G6	<i>Escherichia coli</i> UMNK88
L2VFB5	<i>Escherichia coli</i> KTE10	I0VV64	<i>Escherichia coli</i> W26
S1GIC6	<i>Escherichia coli</i> KTE100	A0A0F6C9Y9	<i>Escherichia coli</i> Xuzhou21
S1HTD3	<i>Escherichia coli</i> KTE107	B7LSA7	<i>Escherichia fergusonii</i> (strain ATCC 35469 / DSM 13698 / CDC 056873)
S1IND4	<i>Escherichia coli</i> KTE108	F0J54	<i>Escherichia fergusonii</i> ECD227
L4V4H0	<i>Escherichia coli</i> KTE112	L2VK96	<i>Escherichia</i> sp. KTE11
L4YQ3	<i>Escherichia coli</i> KTE146	S0UAH5	<i>Escherichia</i> sp. KTE114
S1PJM3	<i>Escherichia coli</i> KTE182	S1LF37	<i>Escherichia</i> sp. KTE159
L3C2K4	<i>Escherichia coli</i> KTE193	S0XEF7	<i>Escherichia</i> sp. KTE31
S0VHU5	<i>Escherichia coli</i> KTE20	S1CHW0	<i>Escherichia</i> sp. KTE52
L2XA22	<i>Escherichia coli</i> KTE21	S1FPD3	<i>Escherichia</i> sp. KTE96
R8XC97	<i>Escherichia coli</i> KTE33	A0A090V5C2	<i>Escherichia vulneris</i> NBRC 102420
S0XP38	<i>Escherichia coli</i> KTE37	A0A085G574	<i>Ewingella americana</i> ATCC 33852
S0YW90	<i>Escherichia coli</i> KTE38	A0A0I8Y5H1	<i>Franconibacter pulveris</i>
S1DSL6	<i>Escherichia coli</i> KTE64	A0A128FD34	<i>Grimontia celer</i>
L3NVW9	<i>Escherichia coli</i> KTE66	D0I469	<i>Grimontia hollisae</i> CIP 101886
S1EAH5	<i>Escherichia coli</i> KTE73	R1DM18	<i>Grimontia indica</i>
L3PZC6	<i>Escherichia coli</i> KTE75	A0A128F7W5	<i>Grimontia marina</i>
L3QB37	<i>Escherichia coli</i> KTE76	A0A0F5Z8R4	<i>Grimontia</i> sp. AD028
V8JN44	<i>Escherichia coli</i> LAUEC10	A0A085H2B6	<i>Hafnia alvei</i> ATCC 13337
A0A1X3I115	<i>Escherichia coli</i> M056	G9Y419	<i>Hafnia alvei</i> ATCC 51873
A0A0K9TEJ1	<i>Escherichia coli</i> M114	A0A097R778	<i>Hafnia alvei</i> FB1
F4T4U0	<i>Escherichia coli</i> M605	A0A158YKY6	<i>Izhakiella australiensis</i>
F4TLE4	<i>Escherichia coli</i> M718	A0A0H3FMC5	<i>Klebsiella aerogenes</i> (strain ATCC 13048 / NCTC 10006)
E9YVD1	<i>Escherichia coli</i> M863	A0A0J2H344	<i>Klebsiella michiganensis</i>
M9G8W2	<i>Escherichia coli</i> MP021561.2	H3N490	<i>Klebsiella michiganensis</i>
D7Y617	<i>Escherichia coli</i> MS 1151	A0A0H3H307	<i>Klebsiella oxytoca</i> (strain ATCC 8724 / KCTC 1686)
E9TFR0	<i>Escherichia coli</i> MS 1173	A0A068HJH8	<i>Klebsiella oxytoca</i> KONIH1
D8E859	<i>Escherichia coli</i> MS 1197	B5XTT3	<i>Klebsiella pneumoniae</i> (strain 342)
E1J1V8	<i>Escherichia coli</i> MS 1241	A0A0E1C7X0	<i>Klebsiella pneumoniae</i> 30660/NJST258_1
E1INZ6	<i>Escherichia coli</i> MS 1457	W8VD91	<i>Klebsiella pneumoniae</i> 30684/NJST258_2
D7X8S5	<i>Escherichia coli</i> MS 1981	V0B100	<i>Klebsiella pneumoniae</i> 909957

Uni Prot ID	Spezies	Uni Prot ID	Spezies
W1HUC9	<i>Klebsiella pneumoniae</i> IS39	J7L4A8	<i>Pectobacterium carotovorum</i> subsp. <i>carotovorum</i> PCC21
W1DT00	<i>Klebsiella pneumoniae</i> IS43	A0A015XPM4	<i>Pectobacterium carotovorum</i> subsp. <i>carotovorum</i> UGC32
W1DXC8	<i>Klebsiella pneumoniae</i> IS46	A0A094UCS9	<i>Pectobacterium carotovorum</i> subsp. <i>odoriferum</i>
W1ELZ7	<i>Klebsiella pneumoniae</i> IS53	D0KBD1	<i>Pectobacterium parmentieri</i> (strain WPP163)
A0A1Y0PTW2	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>	A0A0H3IB70	<i>Pectobacterium</i> sp. (strain SCC3193)
A6TF27	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i> (strain ATCC 700721 / MGH 78578)	A0A1D7YYU2	<i>Pectobacterium wasabiae</i> CFBP 3304
A0A0H3H416	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i> (strain HS11286)	A0A1V3W986	<i>Pedobacter himalayensis</i>
J2LNU3	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i> DSM 30104	A0A0D8MND1	<i>Photobacterium angustum</i>
C8TCE2	<i>Klebsiella pneumoniae</i> subsp. <i>rhinoscleromatis</i> ATCC 13884	A0A0D8QZT0	<i>Photobacterium angustum</i>
R4Y8F9	<i>Klebsiella pneumoniae</i> subsp. <i>rhinoscleromatis</i> SB3432	Q1ZKA9	<i>Photobacterium angustum</i> (strain S14 / CCUG 15956)
A0A060VBK2	<i>Klebsiella quasipneumoniae</i> subsp. <i>quasipneumoniae</i>	A0A090QV42	<i>Photobacterium aphoticum</i>
A0A1F2M0L0	<i>Klebsiella</i> sp. HMSC16C06	A0A011JYV0	<i>Photobacterium aquae</i>
A0A1D8K0Q9	<i>Klebsiella</i> sp. LTGPAF6F	A0A1A6TIU9	<i>Photobacterium aquimaris</i>
J5YCA5	<i>Klebsiella</i> sp. OBRCT7	A0A1B8HWH9	<i>Photobacterium aquimaris</i>
A0A0L0AV70	<i>Klebsiella</i> sp. RITPld	A0A1Q9H284	<i>Photobacterium damsela</i> subsp. <i>piscicida</i>
A0A0B7GAX7	<i>Klebsiella varicola</i>	A0A0D8LH30	<i>Photobacterium damsela</i> subsp. <i>damsela</i>
A0A085HE2	<i>Kluyvera ascorbata</i> ATCC 33433	A0A1C3DIZ9	<i>Photobacterium damsela</i> subsp. <i>damsela</i>
A0A1B7K7D2	<i>Kluyvera georgiana</i> ATCC 51603	D0Z1B7	<i>Photobacterium damsela</i> subsp. <i>damsela</i> CIP 102761
A0A0G3QFY4	<i>Kluyvera intermedia</i>	A0A0C5WJ68	<i>Photobacterium gaetbulicola</i> Gung47
A0A1U7E1D5	<i>Kosakonia cowanii</i>	A0A066RNN7	<i>Photobacterium galathea</i>
A0A1T5H074	<i>Kosakonia oryzae</i>	A0A01JTX9	<i>Photobacterium ganghwense</i>
A0A1C4C5X9	<i>Kosakonia oryzendophytica</i>	A0A0F5VDJ9	<i>Photobacterium halotolerans</i>
A0A1C4CYH2	<i>Kosakonia oryziphila</i>	A0A0D8Q3K0	<i>Photobacterium ilopiscarium</i>
A0A1V0L7Y6	<i>Kosakonia radicincitans</i> DSM 16656	A0A178KKK5	<i>Photobacterium jeanii</i>
W6J586	<i>Kosakonia sacchari</i> SP1	A0A0H3T111	<i>Photobacterium kishitanii</i>
A0A1V2Z3Y0	<i>Leclercia adecarboxylata</i>	A0A0D8N462	<i>Photobacterium leiognathi</i>
A0A016M930	<i>Leclercia</i> sp. LK8	A0A0U1P8Y4	<i>Photobacterium leiognathi</i> Irivu.4.1
A0A1A9FGW0	<i>Lelliottia amnigena</i>	A0A0M9FB15	<i>Photobacterium leiognathi</i> subsp. <i>mandapamensis</i>
A0A066P7U9	<i>Lelliottia amnigena</i> CHS 78	F2PBH8	<i>Photobacterium leiognathi</i> subsp. <i>mandapamensis</i> svers.1.1.
A0A1J5WCP9	<i>Lelliottia nimpresuralis</i>	L8J8Z0	<i>Photobacterium marinum</i>
A0A085HCR4	<i>Leminorella grimonitii</i> ATCC 3399	A0A1A6UKS8	<i>Photobacterium phosphoreum</i>
A0A1X3RIZ1	<i>Lonsdalea britannica</i>	A0A0B7J7B4	<i>Photobacterium phosphoreum</i> ANT2200
A0A1X3RU95	<i>Lonsdalea ibérica</i>	A0A1T5I5L2	<i>Photobacterium piscicola</i>
A0A1X3RGV7	<i>Lonsdalea populi</i>	Q6LVQ1	<i>Photobacterium profundum</i> (strain SS9)
A0A1H4FA36	<i>Lonsdalea quercina</i>	Q1Z2L1	<i>Photobacterium profundum</i> 3TCK
A0A1B7L4L5	<i>Mangrovibacter phragmitis</i>	A0A1Q9GGH8	<i>Photobacterium proteolyticum</i>
A0A071LT05	<i>Mangrovibacter</i> sp. MF070	A0A135ZQP7	<i>Photobacterium sanguinicaneri</i>
A0A0H2QFV1	<i>Moellerella wisconsensis</i>	Q2BYF8	<i>Photobacterium</i> sp. SKA 34
A0A0N0Z902	<i>Moellerella wisconsensis</i> ATCC 35017	A0A0I8VD92	<i>Photobacterium swingsii</i>
M7D0Z7	<i>Morganella morganii</i> SC01	A0A1T4S1Q5	<i>Photobacterium tonni</i>
J7TBT7	<i>Morganella morganii</i> subsp. <i>morganii</i> KT	C7BCS5	<i>Photorhabdus asymbiotica</i> subsp. <i>asymbiotica</i> (strain ATCC 43949 / 310577)
A0A1B8H308	<i>Morganella psychrotolerans</i>	A0A1C0U8J0	<i>Photorhabdus asymbiotica</i> subsp. <i>australis</i>
A0A1B7KIW3	<i>Obesumbacterium proteus</i> ATCC 12841	A0A0M9D1C1	<i>Photorhabdus heterorhabditis</i>
A0A059IBZ9	<i>Pantoea agglomerans</i> Eh318	A0A089V9J8	<i>Photorhabdus luminescens</i> subsp. <i>laumondii</i> (strain DSM 15139 / CIP 105565 / TT01)
A0A1W6B3G1	<i>Pantoea alhagi</i>	U7R0P4	<i>Photorhabdus temperata</i> J3
A0A221NBX4	<i>Pantoea ananas</i>	W3V724	<i>Photorhabdus temperata</i> subsp. <i>khanii</i> NC19
A0A0H3L808	<i>Pantoea ananatis</i> (strain AJ13355)	TOQB7	<i>Photorhabdus temperata</i> subsp. <i>temperata</i> M1021
D4GD23	<i>Pantoea ananatis</i> (strain LMG 20103)	A0A081RZZ1	<i>Photorhabdus temperata</i> subsp. <i>temperata</i> Meg1
A0A0F5FCB7	<i>Pantoea anthophila</i>	A0A0F7LLE7	<i>Photorhabdus temperata</i> subsp. <i>thracensis</i>
A0A1X1CC00	<i>Pantoea calida</i>	U3U262	<i>Plautia stali</i> symbiont
A0A1X1BY95	<i>Pantoea conspicua</i>	A0A1A9AUC5	<i>Plesiomonas shigelloides</i>
A0A1X1EZJ6	<i>Pantoea cypripedii</i>	R8ASQ3	<i>Plesiomonas shigelloides</i> 30273
A0A1X1DW51	<i>Pantoea deleyi</i>	A0A089PGP3	<i>Pluralibacter gergoviae</i>
A0A147GZ28	<i>Pantoea dispersa</i>	A0A0G3CFV7	<i>Pragia fontium</i>
A0A1X1DU53	<i>Pantoea gaviniae</i>	A0A1B7JT40	<i>Proteus hauseri</i> ATCC 700826
A0A1V9DP43	<i>Pantoea latae</i>	V6MFV6	<i>Proteus hauseri</i> ZM44
A0A0B1RTK3	<i>Pantoea rodasii</i>	B4EZJ4	<i>Proteus mirabilis</i> (strain H14320)
A0A1X1CTS4	<i>Pantoea rwandensis</i>	A0A1F2HHL5	<i>Proteus</i> sp. HMSC10D02
A0A1M5WNE8	<i>Pantoea sesami</i>	A0A1F2K2H2	<i>Proteus</i> sp. HMSC14B05
E6WH04	<i>Pantoea</i> sp. (strain AT9b)	A0A0G4QGY1	<i>Proteus vulgaris</i>
A0A0F5XX76	<i>Pantoea</i> sp. 3.5.1	B6X8H1	<i>Providencia alcalifaciens</i> DSM 30120
E0M0P1	<i>Pantoea</i> sp. aB	X6QFK6	<i>Providencia alcalifaciens</i> F902004
U2M5C8	<i>Pantoea</i> sp. ASPVM4	W3YJZ2	<i>Providencia alcalifaciens</i> PAL3
A0A0F3LTM1	<i>Pantoea</i> sp. BL1	K8W161	<i>Providencia burhodogranariae</i> DSM 19968
A0A0L7DJ51	<i>Pantoea</i> sp. CFSAN033090	A0A1B7JS59	<i>Providencia heimbachae</i> ATCC 35613
J3D257	<i>Pantoea</i> sp. GM01	K8W1T4	<i>Providencia rettgeri</i> (strain Dme11)
A0A176UQ79	<i>Pantoea</i> sp. OXW06B1	D4C2F1	<i>Providencia rettgeri</i> DSM 1131
A0A0A0Z197	<i>Pantoea</i> sp. PSNIH1	D1P0V3	<i>Providencia rustigianii</i> DSM 4541
A0A0A1B4U0	<i>Pantoea</i> sp. PSNIH2	K3WU50	<i>Providencia sneebii</i> DSM 19967
A0A246NB57	<i>Pantoea</i> sp. VS1	A0A140NLB5	<i>Providencia stuartii</i> (strain MRSN 2154)
J2UPM2	<i>Pantoea</i> sp. YR343	B2PUL0	<i>Providencia stuartii</i> ATCC 25827
H3RH88	<i>Pantoea stewartii</i> subsp. <i>stewartii</i> DC283	H2IQG1	<i>Rahnella aquatilis</i> (strain ATCC 33071 / DSM 4594 / CIP 78.65)
A0A1W9EPQ4	<i>Pantoea vagans</i>	H8NP73	<i>Rahnella aquatilis</i> HX2
E1SGV0	<i>Pantoea vagans</i> (strain C91)	A0A0H3F4P9	<i>Rahnella</i> sp. (strain Y9602)
A0A1X1D9W9	<i>Pantoea wallisii</i>	A0A0B51QA0	<i>Raoultella ornithinolytica</i>
Q6CZP4	<i>Pectobacterium atrosepticum</i> (strainSCRI 1043 / ATCC BAA672)	A0A1Y6GAU5	<i>Raoultella ornithinolytica</i>
A0A093S2T5	<i>Pectobacterium betavasculorum</i>	A0A1V2BMT6	<i>Raoultella terrigena</i>
A0A1V2R1X8	<i>Pectobacterium carotovorum</i> subsp. <i>actinidiae</i>	A0A1X0WF38	<i>Rouxella badensis</i>
A0A086ESB8	<i>Pectobacterium carotovorum</i> subsp. <i>brasiliense</i>	A0A1X0W571	<i>Rouxella silvae</i>
A0A0M2F3V2	<i>Pectobacterium carotovorum</i> subsp. <i>brasiliense</i>	A0A1V3H1K1	<i>Salinivibrio shamensis</i>
A0A093SUK6	<i>Pectobacterium carotovorum</i> subsp. <i>carotovorum</i>	A0A1N6HK19	<i>Salinivibrio</i> sp. ES.052
C6DGE9	<i>Pectobacterium carotovorum</i> subsp. <i>carotovorum</i> (strain PCI)	A0A1V3I5E3	<i>Salinivibrio</i> sp. IB872

Uni Prot ID	Spezies	Uni Prot ID	Spezies
A0A1V3GM31	Salinivibrio sp. IC202	A0A087L3L3	Serratia sp. Ag1
A0A0F5AQD6	Salinivibrio sp. KP1	A0A084A021	Serratia sp. DD3
A0A1V3HC74	Salinivibrio sp. MA351	A0A0Q4MCX5	Serratia sp. Leaf50
A0A1V3GR88	Salinivibrio sp. ML290	A0A0Q4NDP1	Serratia sp. Leaf51
B5F8L7	Salmonella agona (strain SL483)	I0QWY2	Serratia sp. M24T3
A9MMC4	Salmonella arizonae (strain ATCC BAA731 / CDC34686 / RSK2980)	A0A1B3FHR3	Serratia sp. YD25
A0A0K0HF17	Salmonella bongori (strain ATCC 43975 / DSM 13772 / NCTC 12419)	A0A068Z5H1	Serratia symbiotica
SSN1A0	Salmonella bongori N26808	E9CMX7	Serratia symbiotica str. Tucson
Q57IX3	Salmonella choleraesuis (strain SCB67)	A0A1S2TYG9	Shewanella algae
A0A1J4QKQ0	Salmonella enterica I	Q12IB2	Shewanella denitrificans (strain OS217 / ATCC BAA1090 / DSM 15013)
A0A221ZGX8	Salmonella enterica I	A0A119CZX5	Shewanella frigidimarina
A0A1M3XSZ1	Salmonella enterica subsp. arizonae serovar 18z4,z23: str. CVM N6509	Q089I5	Shewanella frigidimarina (strain NCIMB 400)
A0A1X2TR86	Salmonella enterica subsp. arizonae serovar 41z4,z23:	E6KQ38	Shewanella putrefaciens (strain 200)
A0A1X2TD23	Salmonella enterica subsp. arizonae serovar 50zrz	A4YC70	Shewanella putrefaciens (strain CN32 / ATCC BAA453)
A0A089HHD8	Salmonella enterica subsp. arizonae serovar 62z36: str. RKS2983	A8FPK7	Shewanella sediminis (strain HAWEB3)
A0A0V2HU3	Salmonella enterica subsp. arizonae serovar 63g,z51: str. So 20/20	I6DMC7	Shigella boydii 444474
A0A241R7W5	Salmonella enterica subsp. diarizonae serovar 50kz: str. MZ0080	F3WNU9	Shigella boydii 521682
A0A241RUZ0	Salmonella enterica subsp. diarizonae serovar 65z:z str. SA20044251	E7STW9	Shigella boydii ATCC 9905
G5LEH4	Salmonella enterica subsp. enterica serovar Adelaide str. A4669	B2U3L2	Shigella boydii serotype 18 (strain CDC 308394 / BS512)
G5LUR5	Salmonella enterica subsp. enterica serovar Alachua str. R6377	Q31VN0	Shigella boydii serotype 4 (strain Sb227)
A0A1S1AR07	Salmonella enterica subsp. enterica serovar Bareilly	B3WVX2	Shigella dysenteriae 1012
A0A0U0WIK6	Salmonella enterica subsp. enterica serovar Bovismorbificans	F3VBR0	Shigella dysenteriae 15574
V7V0D0	Salmonella enterica subsp. enterica serovar Cero str. CFSAN001590	E28X33	Shigella dysenteriae 1617
V7WA1	Salmonella enterica subsp. enterica serovar Cubana str. 76814	I6FK69	Shigella dysenteriae 22575
S5HPZ5	Salmonella enterica subsp. enterica serovar Cubana str. CFSAN002050	E7SI46	Shigella dysenteriae CDC 741112
M7RMV6	Salmonella enterica subsp. enterica serovar Dublin str. UC16	Q32AL8	Shigella dysenteriae serotype 1 (strain Sd197)
S4HWX6	Salmonella enterica subsp. enterica serovar Enteritidis str. 2009K0958	A0A090NKT1	Shigella dysenteriae WR5d3
A0A0G2NTL0	Salmonella enterica subsp. enterica serovar Gallinarum str. S09	K0WPG0	Shigella flexneri 148580
G5MPB8	Salmonella enterica subsp. enterica serovar Give str. S5487	I6B812	Shigella flexneri 285071
A0A1Z3Q804	Salmonella enterica subsp. enterica serovar India str. SA20085604	A0A226KF07	Shigella flexneri 2a str. 301
G4C7G4	Salmonella enterica subsp. enterica serovar Infantis str. SARB27	A0A127GQI0	Shigella flexneri 4c
G5NJ55	Salmonella enterica subsp. enterica serovar Inveness str. R83668	A0A0F6MHV0	Shigella flexneri 5a str. M90T
A0A0L9F432	Salmonella enterica subsp. enterica serovar Johannesburg	E77945	Shigella flexneri CDC 79683
G5PTV2	Salmonella enterica subsp. enterica serovar Mississippi str. A4633	F5P329	Shigella flexneri K227
A0A1CZS59	Salmonella enterica subsp. enterica serovar Montevideo str. 6094582	I6CB19	Shigella flexneri K315
G5Q8U0	Salmonella enterica subsp. enterica serovar Montevideo str. S5403	Q05ZR4	Shigella flexneri serotype 5b (strain 8401)
V1WVN4	Salmonella enterica subsp. enterica serovar Muenchen str. baa1594	D2A9K3	Shigella flexneri serotype X (strain 2002017)
A0A1X2RSR1	Salmonella enterica subsp. enterica serovar Rough Oad:1,7	F5NA53	Shigella flexneri VA6
G5QQV5	Salmonella enterica subsp. enterica serovar Rubislaw str. A4653	Q3YWM1	Shigella sonnei (strain Ss046)
A0A1X2YSY6	Salmonella enterica subsp. enterica serovar Saphra	A0A1E3N6I1	Shigella sp. FC569
A0A1X2U5Y5	Salmonella enterica subsp. enterica serovar Seftenburg	A0A142H7Q9	Shigella sp. PAMC 28760
G5R6R0	Salmonella enterica subsp. enterica serovar Senftenberg str. A4543	I2B4A7	Shimwellia blattae (strain ATCC 29907 / DSM 4481 / JCM 1650 / NBRC 105725 / CDC 900574)
X2KH59	Salmonella enterica subsp. enterica serovar Tennessee str. TXSC_TXSC0819	Q2NQI2	Sodalis glossinidius (strain morsitans)
A0A0U1IAA2	Salmonella enterica subsp. enterica serovar Typhimurium str. DT104	W0HNQ5	Sodalis praecaptivus
G5RLS5	Salmonella enterica subsp. enterica serovar Uganda str. R83404	A0A1J7C9Z6	Sodalis sp. TMEI
G5S221	Salmonella enterica subsp. enterica serovar Urbana str. R82977	A0A095TSA2	Tatumella morbirosei
G5SHS2	Salmonella enterica subsp. enterica serovar Wandsworth str. A4580	A0A083JP57	Tatumella ptyseos ATCC 33301
V1HG64	Salmonella enterica subsp. indica serovar 6,14,25z:10,1,2,7 str. 1121	A0A085A3E4	Trabulsiella guamensis ATCC 49490
A0A0F5BBE1	Salmonella enterica subsp. salamae	A0A0L0GMJ3	Trabulsiella odontotermis
A0A241SSV6	Salmonella enterica subsp. salamae serovar 55kz:39 str. 1315K	A0A1M5YV71	Vibrio aerogenes CECT 7868
A0A0V2DD59	Salmonella enterica subsp. salamae serovar 56z:10e,n,x str. 136973	C2HY47	Vibrio albensis VL426
A0A1R2FQV4	Salmonella enteritidis	A0A0H0Y7W8	Vibrio alginolyticus
B5R7L3	Salmonella gallinarum (strain 287/91 / NCTC 13346)	A0A0L7YJ86	Vibrio alginolyticus
A0A1J6YBY3	Salmonella houtenae	A0A1W6VHN4	Vibrio alginolyticus
A0A0R9N8A2	Salmonella Newport	D0W W58	Vibrio alginolyticus 40B
A0A0H3BUZ9	Salmonella Newport (strain SL254)	A0A191W0W9	Vibrio anguillarum
A0A0H2WUJ3	Salmonella paratyphi A (strain ATCC 9150 / SARB42)	F7YP48	Vibrio anguillarum (strain ATCC 68554 / 775)
C0Q0H6	Salmonella paratyphi C (strain RKS4594)	A0A1M3U253	Vibrio antiquarius (strain Ex25)
A0A1F2JLQ5	Salmonella sp. HMSC13B08	A0A1C3IRK4	Vibrio atlanticus
P41406	Salmonella typhi	U3CBE2	Vibrio azureus NBRC 104587
A0A0D6IGZ4	Salmonella typhimurium	A0A1J4QYG4	Vibrio barjuae
A0A0F6B7V8	Salmonella typhimurium (strain 14028s / SGSC 2262)	A0A177Y5F2	Vibrio bivalvicida
EXXEK4	Salmonella typhimurium (strain 4/74)	E8LY90	Vibrio brasiliensis LMG 20546
P08982	Salmonella typhimurium (strain LT2 / SGSC1412 / ATCC 700720)	A7MSS2	Vibrio campbellii (strain ATCC BAA1116 / BB120)
A0A0H3NIL4	Salmonella typhimurium (strain SL1344)	L8XNJ8	Vibrio campbellii CAIM 51
U2N160	Serratia fonticola AUP2C	A0A0A6QKS3	Vibrio caribbeanicus
U2LQP9	Serratia fonticola AUP3(3)	E3BIY6	Vibrio caribbeanicus ATCC BAA2122
A0A084YMD5	Serratia grimesii	A0A1C3JBY8	Vibrio celticus
A0A069CX27	Serratia liquefaciens FK01	A0A0X1L458	Vibrio cholerae (strain MO10)
L0MP79	Serratia marcescens FG94	A0A0K9UQV9	Vibrio cholerae 274080
A0A1C0KPG5	Serratia marcescens subsp. marcescens	A0A0H3Q4L1	Vibrio cholerae B33
A0A0N2ACK3	Serratia marcescens subsp. marcescens Db11	D7HFY9	Vibrio cholerae RC385
A0A086G8V3	Serratia nematodiphila DZ0503SBS1	Q9KNL9	Vibrio cholerae serotype O1 (strain ATCC 39315 / El Tor Inaba N16961)
D4DYP0	Serratia odorifera DSM 4582	A0A0H3AHU5	Vibrio cholerae serotype O1 (strain ATCC 39541 / Classical Ogawa 395 / O395)
S0ALC9	Serratia plymuthica 4Rxl3	C3LSE0	Vibrio cholerae serotype O1 (strain M662)
A0A1B1KXQ1	Serratia plymuthica PR12C	A0A151L1L2	Vibrio cidiicii
S4YTT4	Serratia plymuthica S13	A0A1T4RRQ1	Vibrio cincinnatiensis DSM 19608
A0A1W5DT66	Serratia proteamaculans	A0A1B1VER7	Vibrio corallilyticus
A8GKS3	Serratia proteamaculans (strain 568)	A0A0T7DMX2	Vibrio crassostreae
A0A126VFT0	Serratia rubidaea	A0A1D2YD97	Vibrio crassostreae 9ZC13
V3TRQ7	Serratia sp. (strain ATCC 39006)	A0A1D2XL38	Vibrio crassostreae ZF91
A0A1C7WCQ4	Serratia sp. 142641	A0A1E5DJN7	Vibrio cyclophilicus FF160

Uni Prot ID	Spezies	Uni Prot ID	Spezies
A0A0T7DWM4	Vibrio diabolicus	A0A0H2MXE3	Vibrio sp. VPAP30
A0A178JAL0	Vibrio europaeus	A0A1B9Q184	Vibrio sp. ZF57
Q5E8N6	Vibrio fischeri (strain ATCC 700601 / ES114)	A0A1N6M2L3	Vibrio spartinae
BSFFE6	Vibrio fischeri (strain MJ11)	A0A0P6Z6B2	Vibrio splendidus
S7ISB4	Vibrio fluvialis PG41	A0A1A6LAB2	Vibrio splendidus
A0A066UWF8	Vibrio fortis	A0A1R3EXW1	Vibrio splendidus
A0A0Q2V1G1	Vibrio furnissii	A0A1V9JH7	Vibrio splendidus
F0LRG6	Vibrio furnissii (strain DSM 14383 / NCTC 11218)	A3UPD0	Vibrio splendidus (strain 12B01)
A0A0F4NMP7	Vibrio galathea	A0A1E5FWW9	Vibrio splendidus 12E03
A0A1Z2SFN8	Vibrio gazogenes	F9SJB9	Vibrio splendidus ATCC 33789
A0A1M5D3S0	Vibrio gazogenes DSM 21264	A0A1E5BVB3	Vibrio splendidus FP6
A0A1B9QVB9	Vibrio genomosp. F10	A0A1E5B7S3	Vibrio splendidus ZS139
A0A1E5BD52	Vibrio genomosp. F10 str. ZF129	A0A1A6KNX2	Vibrio tasmaniensis
A0A1ESDAE3	Vibrio genomosp. F6 str. FF238	B7VHH7	Vibrio tasmaniensis (strain LQP32)
D0XS02	Vibrio harveyi (strain 1DA3)	A0A1E5E5W39	Vibrio tasmaniensis 1F187
A0A0M0HX27	Vibrio hepatarius	A0A1E5EHV4	Vibrio tasmaniensis 1F267
F9S5F4	Vibrio ichthyenteri ATCC 700023	A0A1200DH3	Vibrio toranzoniae
A0A1ESDMS2	Vibrio kanaloae 5S149	A0A0A5I2P0	Vibrio tubiashii
A0A0L1L6F9	Vibrio lentus	F9T8K1	Vibrio tubiashii ATCC 19109
A0A1B9QAD2	Vibrio lentus	Q7MPX4	Vibrio vulnificus (strain YJ016)
A0A1Y6I171	Vibrio mangrovi	A0A0H0Y1G2	Vibrio vulnificus CladeAyb158
A0A1A7V1D8	Vibrio mediterranei	A0A0M0IPG9	Vibrio xuii
A0A0N8U119	Vibrio metoecus	A5KYP2	Vibrionales bacterium (strain SWAT3)
A0A0Q0TKU1	Vibrio metoecus	A0A0T9T1G7	Yersinia aldovae
C9P1T1	Vibrio metschnikovii CIP 69.14	A0A0T9R2H0	Yersinia bercovieri
D0HIW1	Vibrio mimicus VM223	A0A0T7NUR9	Yersinia enterocolitica
D2YL71	Vibrio mimicus VM573	A0A0T9S1N5	Yersinia enterocolitica
D2YF94	Vibrio mimicus VM603	A0A0T9TYK2	Yersinia enterocolitica
A0A1B1E8M7	Vibrio natriegens	A0A1W9FF69	Yersinia enterocolitica
A0A099LQZ4	Vibrio navamensis	A1JIS0	Yersinia enterocolitica serotype O:8 / biotype 1B (strain NCTC 13174 / 8081)
A0A0F4NY85	Vibrio neptunius	A0A0H3P102	Yersinia enterocolitica subsp. palearctica serotype O:3 (strain DSM 13030 / CIP 106945 / Y11)
A0A0M0HLA1	Vibrio nereis	A0A094UPY2	Yersinia frederiksenii ATCC 33641
U4EFA5	Vibrio nigripulchritudo FTn2	A0A0B6FKN7	Yersinia frederiksenii Y225
A0A1E5FM58	Vibrio ordalii 12B09	A0A0H5M0L7	Yersinia intermedia
C9QDL4	Vibrio orientalis CIP 10289	A0A0T9MAT5	Yersinia intermedia
A0A0AC1VTM2	Vibrio owensii 476661	A0A0T9U1Q6	Yersinia intermedia
A0A0C1VSS1	Vibrio owensii CAIM 185	A0A0U1EW47	Yersinia intermedia
A0A1R4B775	Vibrio palustris	A0A0T9LIA0	Yersinia kristensenii
A0A1Q9HAB8	Vibrio panuliri	A0A0T9R043	Yersinia massiliensis
A0A0D1DRJ0	Vibrio parahaemolyticus 49	A0A0U1KEX3	Yersinia mollaretii
A0A0F6LVZ3	Vibrio parahaemolyticus 863	C4SCQ6	Yersinia mollaretii ATCC 43969
A0A0D1EPL2	Vibrio parahaemolyticus 901128	A0A0T9KRZ2	Yersinia nurmii
X7RS41	Vibrio parahaemolyticus M0605	A0A0T9PJ50	Yersinia pekanenii
SS1YF5	Vibrio parahaemolyticus O1:Kuk str. FDA_R31	Q7CFW9	Yersinia pestis
A6B2M0	Vibrio parahaemolyticus serotype O3:K6 (strain A Q3810)	A0A0M1VD06	Yersinia pestis biovar Orientalis str. IP275
Q87TB5	Vibrio parahaemolyticus serotype O3:K6 (strain RIMD 2210633)	A0A0E1NVC9	Yersinia pestis bv. Antiqua (strain Antiqua)
A0A0D1DNQ5	Vibrio parahaemolyticus VP766	A0A0H2YNH3	Yersinia pestis bv. Antiqua (strain Nepal516)
A0A090P6G0	Vibrio ponticus	Q664K6	Yersinia pseudotuberculosis serotype I (strain IP32953)
U2ZZY5	Vibrio proteolyticus NBRC 13287	A0A0U1QWU0	Yersinia pseudotuberculosis serotype O:1b (strain IP 31758)
A0A1M7Z0B4	Vibrio quintilis	A0A0H3A WS8	Yersinia pseudotuberculosis serotype O:3 (strain YPIII)
A0A0C2P5H2	Vibrio renipiscarius	A0A0U1HQW9	Yersinia rohdei
A0A241TRV9	Vibrio rotiferianus	A0A085U721	Yersinia ruckeri
A0A1R4LLN0	Vibrio ruber DSM 16370	A0A0T9PCW5	Yersinia similis
A0A1E5E613	Vibrio rumoiensis 1S45	A0A0E8XNW0	Yersinia wautersii
F9RQW6	Vibrio scophthalmi LMG 19158	C9Z0P7	Yokenella regensburgei ATCC 43003
A6D0J4	Vibrio shilonii AK1		
A0A0A5HTA3	Vibrio sinaloensis		
E8M506	Vibrio sinaloensis DSM 21326		
H2IBT8	Vibrio sp. (strain EJY3)		
F9RCU2	Vibrio sp. (strain N418)		
A0A1R3F9J6	Vibrio sp. 10N.222.47.A.9		
A0A1R3FTE6	Vibrio sp. 10N.261.45.E1		
B8K8X5	Vibrio sp. 16		
A0A220TEZ9	Vibrio sp. 252189		
A8TG65	Vibrio sp. AND4		
A0A086W1F0	Vibrio sp. B183		
A0A084T3D3	Vibrio sp. ER1A		
K5T1M4	Vibrio sp. HENC01		
K5VF17	Vibrio sp. HENC02		
K5TE71	Vibrio sp. HENC03		
A0A165SCE5	Vibrio sp. HI00D65		
A0A0T7EEG0	Vibrio sp. J229		
A0A0T7EJ12	Vibrio sp. J231		
A0A0T7DDT6	Vibrio sp. J24		
A0A1R4K0H4	Vibrio sp. JB196		
A0A034TH77	Vibrio sp. JCM 18904		
A0A034TXM8	Vibrio sp. JCM 18905		
A0A061Q2X0	Vibrio sp. JCM 19052		
A0A061PN34	Vibrio sp. JCM 19053		
A0A0W8JK89	Vibrio sp. MEBIC08052		
D0IEB0	Vibrio sp. RCS86		
A0A0F2H4I6	Vibrio sp. S2345		