A method for the genetically encoded incorporation of FRET pairs into proteins

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

der Georg-August-Universität Göttingen

im Promotionsprogramm Biologie

der Georg-August University School of Science (GAUSS)

vorgelegt von

Christoph Lammers

aus Würselen

Göttingen, 2014

Betreuungsausschuss

Jun.-Prof. Dr. Heinz Neumann, Angewandte Synthetische Biologie, Institut für Mikrobiologie und Genetik

Prof. Dr. Jörg Stülke, Allgemeine Mikrobiolgie, Institut für Mikrobiologie und Genetik

Mitglieder der Prüfungskommission

Referent: Jun.-Prof. Dr. Heinz Neumann, Angewandte Synthetische Biologie, Institut für Mikrobiologie und Genetik

Korreferent: Prof. Dr. Jörg Stülke, Allgemeine Mikrobiolgie, Institut für Mikrobiologie und Genetik

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Ralf Ficner, Molekulare Strukturbiologie, Institut für Mikrobiologie und Genetik

Prof. Dr. Kai Tittmann, Bioanalytik, Albrecht-von-Haller-Institut

Prof. Dr. Rolf Daniel, Genomische und Angewandte Mikrobiologie, Institut für Mikrobiologie und Genetik

Dr. Fabian Commichau, Allgemeine Mikrobiolgie, Institut für Mikrobiologie und Genetik

Tag der mündlichen Prüfung: 15.07.2014

Table of Contents

Table of Contents I

List of Figures VI

List of Tables X

Danksagungen XII

Abbreviations XIV

1 Summary 1

2 Introduction 3

2.1 Proteins Encoded By The Genetic Code 3

2.2 Expansion And Modification of The Genetic Code 5

2.3 Limitations And Optimizations of The Expanded Genetic Code 8

2.4 Applications of The Expanded Genetic Code 13

2.5 The Genetically Encoded Incorporation of FRET Pairs 15

2.6 Specific Aims 17

3 Materials And Methods 18

3.1 Materials 18

3.1.1 Devices And Instruments 18

3.1.2 Chemicals 20

3.1.3 Consumables And Other Materials 20

3.1.4 DNA, Protein And RNA Size Standards 22

3.1.5 Antibiotics 22

3.1.6 Culture Media And Agar Plates 23

3.1.7 Unnatural Amino Acids 23

3.1.8 Enzymes 24

3.1.9 Cell Lines 24

3.1.10 Buffers And Solutions 25

3.1.11 DNA Kit Systems 27

3.1.12 Fluorescent Dyes 27

3.1.13 Antibodies 28

3.2 Methods 29

3.2.1 Microbiological Methods 29

3.2.1.1 Chemical Competent Cells 29

3.2.1.2 Electro Competent Cells 29

3.2.1.3 Transformation of Chemical Competent Cells 29

3.2.1.4 Transformation of Electro Competent Cells 30

3.2.2 Protein Biochemical Methods 30

3.2.2.1 Recombinant Protein Expression 30

3.2.2.2 Cell Lysis 31

3.2.2.3 Discontinuous Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis 31

3.2.2.4 Western Blot 32

3.2.2.5 Measuring Protein Concentration 33

3.2.2.6 Small Scale GST-MBP Expression And Purification 33

3.2.2.7 Small Scale Histone H3 Expression And Purification 34

3.2.2.8 Large Scale Nanobody NbSyn2 Expression And Purification 35

3.2.2.9 Expression And Purification of Isotopically Labeled sfGFP 35

3.2.2.10 Protein Expression For In-Cell NMR Measurements 36

3.2.2.11 Labeling of Proteins With Fluorophores 36

3.2.2.12 Fluorescence Measurement With A Plate Reader 37

3.2.2.13 Fluorescence-Activated Cell Sorting 37

3.2.3 Nucleic Acid Biochemical Methods 38

3.2.3.1 Preparation of Plasmid DNA 38

3.2.3.2 Restriction Enzyme Digestion 38

3.2.3.3 Agarose Gel Electrophoresis 39

3.2.3.4 Extraction of DNA From Agarose Gels 40

3.2.3.5 Separation of Multiple Plasmids ... 40

3.2.3.6 Ligation of Two Double-Stranded DNA Fragments ... 41

3.2.3.7 Ethanol Precipitation ... 41

3.2.3.8 Polymerase Chain Reaction ... 42

3.2.3.9 Purification of PCR Products ... 43

3.2.3.10 QuikChange PCR ... 44

3.2.3.11 Inverse PCR ... 45

3.2.3.12 Measuring Nucleic Acid Concentration ... 47

3.2.3.13 Sequencing of Plasmids ... 48

3.2.3.14 Extraction/Isolation of RNA ... 49

3.2.3.15 Northern Blot ... 50

3.2.3.16 Transcription With T7 RNA Polymerase ... 51

3.2.4 Reporter Assays ... 53

3.2.4.1 Chloramphenicol Reporter Assay ... 53

3.2.4.2 GFP Reporter Assay ... 54

3.2.4.3 Barnase Reporter Assay ... 55

4 Results ... 56

4.1 Abundance And Activity of The aaRS/tRNA Pairs PylS/PylT And MjYRS/MjYT 56

4.1.1 PylS of Methanosarcina barkeri ... 57

4.1.2 PylT of Methanosarcina barkeri ... 62

4.1.3 MjYRS of Methanocaldococcus jannaschii ... 64

4.1.4 MjYT of Methanocaldococcus jannaschii ... 67

4.2 Promoter Libraries ... 71

4.2.1 PylS Library ... 72

4.2.2 MjYRS_AGGA Library ... 75

4.2.3 PylT and MjYT_UCCU Library ... 78

4.3 Defined (Standard) Promoters ... 81

4.4 The Modular Genetic Tool 84

4.4.1 The Concept of The Modular Genetic Tool 84

4.4.2 The Construction of The Modular Genetic Tool 85

4.4.3 Monitoring The Power of The Modular Genetic Tool 90

4.4.4 The Combination of The Modular Genetic Tool and The Promoter Libraries. 92

4.4.5 Alternative Strategies For Screening The Modular Genetic Library Tool 96

4.4.6 The Combination of The Modular Genetic Tool And Defined (Standard) Promoters 100

5 Applications And Side Projects 104

5.1 Evolution of A Quadruplet Decoding MjAzFRS 104

5.2 Labeling of Proteins For FRET 107

5.3 Fluorescently Labeled Nanobodies 111

5.4 Orthogonal Ribosome Assisted In-Cell NMR Spectroscopy 114

6 Discussion 119

6.1 Balancing An Expanded Genetic Code System 119

6.2 Reducing The Metabolic Burden 122

6.3 The Avenue To FRET Measurements 124

6.4 Applications And Side Projects: Nanobodies And In-Cell NMR 125

6.5 Concluding Remarks And Outlook 126

7 References 128

8 Appendix 142

8.1 Plasmids 142

8.2 Oligonucleotides 146

8.3 Construction of Plasmids 152

8.4 Sequence Alignments 158

8.4.1 PylS Promoter Library 158

8.4.2 MjYRS_AGGA Promoter Library 159

8.4.3 PylT Promoter Library 160

8.4.4 MjYT_UCCU Promoter Library 161

8.5 In-Cell sfGFP Fluorescence Measurements 162

8.5.1 Initial Conditions For In-Cell NMR Measurements 162

8.5.2 Optimization For In-Cell NMR Measurements 163

8.6 Curriculum Vitae 164

List of Figures

Figure 2.1: Schematic view for genetic code expansion. 6

Figure 2.2: General selection scheme for the evolution of aminoacyl-tRNA synthetase (aaRS) substrate specificities for unnatural amino acids (UAAs)[44]. 7

Figure 3.1: Schematic view of QuikChange cycles. 44

Figure 3.2: Schematic representation of the restriction enzyme-mediated inverse PCR[135]. 46

Figure 3.3: Schematic view of the chloramphenicol reporter assay. 53

Figure 3.4: Schematic view of the GFP reporter assay. 54

Figure 3.5: Schematic view of the barnase reporter assay. 55

Figure 4.1: Pairwise sequence alignment with PylS from *M. mazei* and *M. barkeri*. 57

Figure 4.2: Crystal structure of the catalytic domain of PylS from *M. mazei*. 58

Figure 4.3: Cm-Assay with pBK PylS His_6-tag variants. 59

Figure 4.4: Detection of PylS His_6-tag (“gap”) via western blot. 59

Figure 4.5: Truncations in the PylS “gap” (“wt”) variant to investigate the effect on the suppression efficiency. 60

Figure 4.6: Cm-Assay with truncation variants of PylS “gap” (“wt”). 61

Figure 4.7: Predicted secondary structure of PylT. 62

Figure 4.8: Establishing northern blots to detect specific tRNAs. 63

Figure 4.9: Detection of PylT in total RNA extracts. 64

Figure 4.10: Crystal structure of tyrosyl-tRNA synthetase from *M. jannaschii* complexed with tRNA(Tyr) and L-tyrosine. 65

Figure 4.11: Cm-Assay with pBK MjYRS His_6-tag variants. 66

Figure 4.12: Detection of MjYRS His_6-tag via western blot. 66

Figure 4.13: Predicted secondary structure of MjYT_CUA. 67

Figure 4.14: Detection of MjYT_CUA in total RNA extracts. 68

Figure 4.15: Sequence alignment of MjYT. 69

Figure 4.16: Refinement of MjYT detection. 70

Figure 4.17: Schematic overview of the synthetic promoter libraries. 71

Figure 4.18: Cm-Assay with single colonies from PylS library. 73

Figure 4.19: Comparative analysis of PylS lib clones via western and northern blot. 74

Figure 4.20: Inverse gradient PCR for the MjYRS promoter library. 75

Figure 4.21: Cm-Assay with single colonies from MjYRS_AGGA library. 76

Figure 4.22: Suppression efficiency of MjYRS_AGGA lib clones. 77

Figure 4.23: Cm-Assay with single colonies from PylT library. 79

Figure 4.24: Cm-Assay with single colonies from MjYT_UCCU library. 79

Figure 4.25: Comparative analysis of defined standard promoters with WT *glnS* and library clone 3E promoter. 83

Figure 4.26: Schematic overview of the modular genetic tool. 85

Figure 4.27: Cm-Assay with final pBK PylS plasmid and intermediate steps from preparation for the modular genetic tool. 87

Figure 4.28: Cm-Assay to detect the stage responsible for PylS activity loss. 87

Figure 4.29: Cm-Assay of the PylS NdeI restore mutants. 88

Figure 4.30: Restriction digests of the final modular genetic tool. 89

Figure 4.31: Comparison of the modular genetic tool to the four plasmid system. 91

Figure 4.32: Cm-Assay for the background reduction of combined PylS and PylT promoter library. 93

Figure 4.33: Restriction digests of the modular genetic tool with wild type and library promoters. 94

Figure 4.34: FACS to identify the best combination of all 4 libraries on the Genetic Tool at once (GT4L). 95

Figure 4.35: Comparison of the modular genetic tool (GTF) to single four library plasmid clones. 99

Figure 4.36: Comparison of the standard modular genetic tool to variants with defined promoters. 101

Figure 4.37: Small scale expression and purification of GST-MBP comparing different genetic tool variants and the 4 plasmid system. 102

Figure 4.38: Growth curve and GFP fluorescence measurement comparing different genetic tool variants and the 4 plasmid system. 103

Figure 5.1: Cm-Assay with the 20 best clones from the MjAzFRS library. 105

Figure 5.2: Incorporation of AzF into GFP and H3 using MjAzFRS library clone 17 (pCLA149). 106

Figure 5.3: Incorporation of BCNK and NorK into myoglobin. 108

Figure 5.4: Incorporation of NorK into histone H3. 109

Figure 5.5: Labeling of histone H3 for (pseudo) FRET. 110

Figure 5.6: Labeling of histone H3 with four different tetrazine conjugated fluorophores. 111

Figure 5.7: Incorporation of BocK and NorK into NbSyn2. 112

Figure 5.8: Fluorescently labeled NbSyn2. 113

Figure 5.9: Incorporation of BocK into NbGFP. 114

Figure 5.10: sfGFP expression for in-cell NMR measurements. 116

Figure 5.11: Arabinose inducible promoter for the orthogonal translational machinery. .. 116

Figure 5.12: Effect of spectinomycin on sfGFP expression. 117

Figure 5.13: sfGFP expression with selected spectinomycin concentrations. 118

Figure 8.1: Sequence alignment of selected PylS promoter library clones. 158

Figure 8.2: Sequence alignment of selected MjYRS_AGGA promoter library clones. 159

Figure 8.3: Sequence alignment of selected PylT promoter library clones. 160

Figure 8.4: Sequence alignment of selected MjYT_UCCU promoter library clones. 161

Figure 8.5: Effect of spectinomycin on in-cell fluorescence of sfGFP. 162

Figure 8.6: Titration of spectinomycin to optimize in-cell fluorescence of sfGFP. 163

List of Tables

Table 3.1: Overview of antibiotics used in growth media/agar plates 22

Table 3.2: Overview of growth media used for culturing of *E. coli*...................................... 23

Table 3.3: Overview of unnatural amino acids used for genetic code expansion 23

Table 3.4: Overview of enzymes used 24

Table 3.5: Overview of cell lines used.. 24

Table 3.6: Overview of commonly used buffers & solutions 25

Table 3.7: Overview of fluorescent dyes used for protein labeling. 27

Table 3.8: Overview of antibodies used for immuno blotting (western & northern blots). . 28

Table 3.9: Composition of polyacrylamide gels for SDS PAGE 31

Table 3.10: Pipetting scheme for standard test and preparative digests 39

Table 3.11: Percentages of agarose gels depending on the size of the DNA (fragments)..... 40

Table 3.12: Pipetting scheme for a standard ligation reaction mix 41

Table 3.13: Standard course of a PCR 43

Table 3.14: Pipetting scheme for the composition of one PCR reaction mix 43

Table 3.15: Pipetting scheme for the composition of one QuikChange PCR reaction mix.... 44

Table 3.16: Standard course of a QC-PCR 45

Table 3.17: Pipetting scheme for the composition of one inverse PCR reaction mix........... 46

Table 3.18: Standard course of a restriction enzyme-mediated inverse PCR....................... 47

Table 3.19: Pipetting scheme for the composition of one sequencing PCR reaction mix 48

Table 3.20: Standard course of a sequencing PCR 48

Table 3.21: Composition of an acid urea polyacrylamide gel for acid urea PAGE 50

Table 3.22: Pipetting scheme for the composition of one transcription reaction mix.......... 52

Table 3.23: Overview of RNA transcripts made by T7 RNA polymerase. 52

Table 4.1: Relative *in vivo* strength of selected promoters. 82

Table 4.2: Overview of all QuikChanges performed in order to construct the modular genetic tool. 86

Table 4.3: Overview of preparative digests performed in order to construct the modular genetic tool. 89

Table 4.4: Cm-Assay for the sorting of the four library plasmids. 97

Table 5.1: Amino acid sequence of the 20 best clones from the MjAzFRS library. 106

Table 8.1: List of plasmids. 142

Table 8.2: List of oligos (primers and probes). 146

Table 8.3: Details of the construction of plasmids made during this thesis. 152

Danksagungen

An dieser Stelle möchte ich einigen Leuten danken, die entweder durch ihre fachliche Kompetenz und/oder ihre Freundschaft entscheidend zum Gelingen dieser Arbeit beigetragen haben:

In erster Linie möchte ich mich bei Herrn **Jun.-Prof. Dr. Heinz Neumann** bedanken, der mir die Möglichkeit gegeben hat, meine Doktorarbeit in seiner Abteilung in dem spannenden Themengebiet der synthetischen Biologie anzufertigen. Ich danke ihm für sein entgegengebrachtes Vertrauen und die kontinuierliche Unterstützung während der gesamten Zeit und hoffe einiges von seinem Ideenreichtum mitgenommen zu haben.

Prof. Dr. Jörg Stülke danke ich für die Übernahme des Korreferats und dafür, dass er auch weit über meine Diplomarbeit hinaus immer ein offenes Ohr hatte und mir mit Rat und Tat freundschaftlich zur Seite stand.

A special “thank you” goes to **Dr. Bryan Wilkins**. I am really thankful that he proofread my thesis although he already read Svenja’s thesis and was himself that busy.

Des Weiteren danke ich sowohl der **AG Neumann**, als auch der **AG Ficner**, für die entspannte Arbeitsatmosphäre, für hilfreiche Diskussionen und auch für manch zünftige (Weihnachts-)Feiern ;).

Ein besonderer Dank gilt auch **Christian Hoffmann**, der mich an so manch einem extrem langen Arbeitstag bis in die Nacht begleitet und mit unzähligen Ideen und Tipps unterstützt hat, aber auch für die netten Abende außerhalb der Arbeit.

Meinen Bachelorstudenten **Miguel Sánchez** und **Julia Motz** bin ich dankbar für ihre erfolgreiche Mithilfe bei einem Teilprojekt dieser Arbeit.

Auch unter der weisen Voraussicht, dass er es wahrscheinlich niemals lesen wird, bedanke ich mich bei meinem ehemaligen Biologielehrer **Herrn Veddeler**, ohne den ich wohl niemals das Biologiestudium angestrebt hätte.

Meinen Freunden/Kommilitonen vor Ort (im Besonderen seien hier **Arne Schmeisky** und **Marcel Kunadt** genannt) danke ich für gesellige Abende und die damit verbundene Ablenkung, wenn mal nicht alles geradeaus lief.

Meinen Freunden zu Hause, besonders den „**Oldenburgern**“, sei für ihre nun teilweise schon weit über ein Jahrzehnt andauernde Freundschaft gedankt und die damit verbundene Unterstützung, die letztlich auch maßgeblich zum Gelingen dieser Arbeit beigetragen hat.

Bei meiner Freundin **Corinna Krüger** möchte ich mich ebenfalls ganz besonders bedanken, für die bereits gemeinsam verbrachte Zeit und für die, die noch folgen wird. Ich danke ihr, dass sie mich so weit es möglich war unterstützt und entlastet hat, mit mir durch die Höhen und Tiefen beim Anfertigen der Arbeit gegangen ist und mir somit das Schreiben um ein Vielfaches erleichtert hat.

Zu guter Letzt, aber definitiv nicht weniger als den zuvor genannten, danke ich meiner Familie und vor allem meinen Eltern **Gert** und **Angelika**. Euer steter Glauben an mich und eure kontinuierliche Unterstützung, die nicht nur finanzieller sondern auch seelischer und moralischer Natur war, hat mir geholfen zu erreichen, wo und wer ich heute bin.

Abbreviations

µg	microgram
µL	microliter
µM	micromolar
AA	amino acid
aaRS/RS	aminoacyl-tRNA synthetase
AG	*Arbeitsgruppe*
Amp	ampicillin
AP	alkaline phosphatase
APS	ammonium persulfate
Ara	arabinose
AzF	4-azido-L-phenylalanine
BCNK	bicyclo[6.1.0]non-4-yn-9-ylmethanol-L-lysine
BME	β-mercaptoethanol
Boc	*tert.*-butyl
BocK	N(ε)-*tert*.-butyl-oxycarbonyl-L-lysine
bp	base pair(s)
BPA	*p*-benzoyl-L-phenylalanine
BSA	bovine serum albumin
CAT	chloramphenicol acetyltransferase
Ch.	chapter
CIP	calf intestinal alkaline phosphatase
Cm	chloramphenicol

ddNTP	dideoxynucleotide triphosphate
DIG	digoxigenin
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
e.g.	*exempli gratia*
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
et al.	*et alii/aliae/alia*
EtOH	ethanol
FACS	fluorescence-activated cell sorting
FRET	Förster/fluorescence resonance energy transfer
g	gram
g	gravitational acceleration
GFP	green fluorescent protein
GST	glutathione *S*-transferase
His	histidine
HRP	horseradish peroxidase
i.e.	*id est*
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
L	liter

LB	lysogeny broth
lib	library
M	molar
M. barkeri (Mb)	*Methanosarcina barkeri*
M. jannaschii (Mj)	*Methanocaldococcus jannaschii*
M. mazei	*Methanosarcina mazei*
MBP	maltose binding protein
MCS	multiple cloning site
MeOH	methanol
mg	milligram
min	minute
MjYRS	*M. jannaschii* tyrosine aminoacyl-tRNA synthetase
MjYT	*M. jannaschii* tyrosine tRNA
mL	milliliter
mM	millimolar
Myc	myelocytomatosis
Nb	nanobody
NB	northern blot
nm	nanometer
nM	nanomolar
NorK	*N*ε-5-norbornene-2-yloxycarbonyl-L-lysine
nt	nucleotide(s)
OD_{600}	optical density at 600 nm wavelength
O-mRNA	orthogonal mRNA

ORBS	orthogonal ribosomal binding site
O-ribosome	orthogonal ribosome
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Pfu	*Pyrococcus furiosus*
PIC	protease inhibitor cocktail
PMSF	phenylmethanesulfonyl fluoride
PVDF	polyvinylidene difluoride
PylS	pyrrolysine aminoacyl-tRNA synthetase
PylT	pyrrolysine tRNA
QC	QuikChange
RNA	ribonucleic acid
rNTP	ribonucleotide triphosphate
rpm	revolutions per minute
RT	room temperature
SAP	shrimp alkaline phosphatase
SDS	sodium dodecyl sulfate
sm	single-molecule
Sm	spectinomycin
SSC	saline-sodium citrate
T7RP	T7 RNA polymerase
Taq	*Thermus aquaticus*
TBE	TRIS-Borate-EDTA buffer

TBS	TRIS buffered saline
TCA	trichloroacetic acid
TCEP	*tris*(2-carboxyethyl)phosphine
TEMED	N,N,N’,N’-tetramethylethylenediamine
Tet	tetracycline
TRIS	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
UAA	unnatural amino acid
UV	ultraviolet
v	volume
w	weight
WB	western blot
WT	wild type

1 Summary

Proteins are composed of 20 canonical amino acids whose unique arrangements predefine a protein's structure and function. Importantly, most proteins are not static conformations but rather very dynamic entities that undergo various structural modifications under different "stimuli". The comprehension of these dynamic processes is necessary to understand how proteins work.

Förster/Fluorescence resonance energy transfer (FRET) became a powerful tool to investigate conformational changes of proteins, and recent advances in technology haven given the capability for studies even on a single-molecule (sm) level. Therefore precise labeling of the proteins with suitable fluorophores is essential, however, remains a challenging task at present. Although chemical bioconjugation of fluorophores and proteins work, more or less reliably, specificity is a drawback for longer polypeptides and full-length proteins.

To overcome the issue of specificity, synthetic biologists have opened new avenues by developing an expansion of the genetic code. This technique requires the introduction of exogenous nonsense suppressor tRNAs and their cognate aminoacyl-tRNA synthetases (aaRS) into the host cell, that have to work completely orthogonal to the endogenous components. This allows the incorporation of additional "unnatural" amino acids (UAAs) into proteins at the genetic level. These UAAs can bear many different functional groups with unique chemical or biophysical properties.

Since we were interested in introducing two fluorophores, site-specifically into a protein, we had to use two tRNA/aaRS pairs, along with the plasmid harboring the gene of interest. This general approach necessitated multiple plasmids with different antibiotic resistances leading to heightened stress in the host cells. Additionally, using two non-cognate tRNA/aaRS pairs displayed toxic side-effects and required balanced levels within cells. Moreover, the two different UAAs drastically decreased the suppression efficiency and, in turn, the expression levels of protein. Therefore, an overall optimization of the system was essential.

Herein, we describe the optimization process. We set out to reduce the number of plasmids used in this system, resulting in a highly modular genetic tool. We designed this system to allow for easy exchange with other tRNA/aaRS pairs to introduce new UAAs. We explored

promoter libraries to fine-tune the expression levels of tRNA/aaRS pairs, which had profound effects on the UAA incorporation efficiency.

Using the above system we achieved higher levels of protein expression with two different UAAs and are currently establishing bioorthogonal labeling strategies for use in smFRET studies.

As a result of the work for the present thesis the following article was submitted to “Chembiochem: a European journal of chemical biology” and will be published soon:

Lammers, C.; Hahn, L. E. & Neumann, H. "Optimized plasmid systems for the incorporation of multiple different unnatural amino acids by evolved orthogonal ribosomes." *Chembiochem,* DOI: 10.1002/cbic.201402033 (2014).

2 Introduction

2.1 Proteins Encoded By The Genetic Code

The three domains of (cellular) life proposed by Woese[1] imply millions of discovered species, as well as those yet undiscovered[2,3]. Despite this huge diversity, all living organisms, whether they are single or multi-celled, have common similarities. They are all built upon simple organic compounds, such as nucleotides and amino acids, the biological “bricks of life”. These monomeric subunits form larger polymeric macromolecules, including DNA/RNA and proteins respectively, assembled in intricate networks of biological processes.

Proteins play a leading part in these highly complex biological systems since they participate in virtually all processes, making them one of the most studied objects in life science. Their variety in structure and function is determined by unique arrangements of the 20 canonical amino acids. Minor mutations, even a single amino acid, can lead to altered conformations that perturb the native function of the protein. The individual amino acid arrangement is dictated by the nucleotide sequence of a protein’s gene, encoded by non-overlapping triplet codons, composed of a combination of the four bases adenine (A), cytosine (C), guanine (G) and thymine (T). This genetic code consists of 64 possible codons from which 61 are assigned to the 20 amino acids and the remaining three to the termination of protein synthesis[4–11].

The assembly of proteins is performed by ribosomes, huge macromolecular complexes composed of ribosomal (r) RNAs and proteins divided into small and large subunits. The process of decoding the mRNA into proteins is called translation and can be separated into the four steps. These are termed initiation, elongation, termination, and recycling[12,13]. Protein synthesis in prokaryotes and eukaryotes is different but primarily they differ only in the initiating steps [14].

Translational events in the prokaryote *Escherichia coli (E. coli)* starts with the complex formation between initiation factors IF1 and IF3, as well as the small (30S) ribosomal subunit[15,16]. This complex is able to bind cellular mRNAs by the interaction of the anti-Shine-Dalgarno sequence on the 16S rRNA and the Shine-Dalgarno sequence (5’-AGGAGG-3’) on the mRNA, located upstream of an AUG start codon[17,18]. The mRNA is guided by this sequence and is thereby positioned in the P site of the ribosome. The recruitment of the GTP-binding protein IF2 and the initiator tRNA, aminoacylated with formylated methionine (fMet-tRNA), prepares the 30S initiation complex for the association with the large (50S) subunit of the ribosome upon GTP hydrolysis. The newly formed 70S initiation complex is poised for the translational elongation phase[13–15].

During elongation, adding amino acids in recurring steps extends the nascent polypeptide chain. Therefore, ternary complexes composed of elongation factor Tu (EF-Tu), GTP and tRNA, which has been previously aminoacylated with the correct amino acid by a cognate aminoacyl-tRNA synthetase (aaRS)[19,20], are delivered into the decoding center within the A site of the ribosome[21]. The ribosome actively monitors the selection of the proper aminoacyl-tRNA, which is dictated by the mRNA codon, and thus maintains the fidelity of the genetic code[22–24]. Whereas the first two codon-anticodon base pairs have to form strict Watson-Crick pairs (A-U and G-C), excluding non-cognate and wobble pairs, wobble pairing at the third position of the codon is tolerated[23,25]. This means, the 5’-base of the tRNA anticodon can form non-Watson-Crick base pairs with the 3’-base of the mRNA codon, permitting some tRNAs to decode more than one codon. The delivery of the appropriate ternary complex into the ribosome’s A site facilitates GTP hydrolysis by EF-TU, followed by its dissociation from the ribosome and the accommodation of the aminoacylated tRNA acceptor stem in the peptidyl transferase center (PTC) within the large subunit for peptide bond formation[13,24]. This causes the transfer of the nascent peptide chain from the P site to the A site tRNA. To remove the deacylated tRNA in the P site and to position a new codon in the A site for the arrival of another aminoacyl-tRNA in the next round, the elongation cycle ends with the translocation of the tRNAs and the mRNA by precisely one codon. This process is catalyzed by the elongation factor EF-G at the cost of GTP hydrolysis[13,26].

The elongation phase terminates when one of the three stop codons (UAA, UAG or UGA) moves into the A site. These codons are recognized by the release factors RF1 (UAA, UAG) and RF2 (UAA, UGA)[27,28] that mediate the cleavage of the nascent polypeptide chain from the tRNA bound in the ribosome’s P site, leading to the release of the newly synthesized protein from the ribosome. Afterwards, the GTP-binding protein RF3 accelerates the dissociation of RF1/RF2 from the ribosome, which is subsequently recycled for the translation of further proteins[13].

2.2 Expansion And Modification of The Genetic Code

The genetic code of all living organisms is near-universally conserved and was long concidered to be restricted to the 20 naturally occurring amino acids. In 1986 two workgroups independently discovered that the nonstandard amino acid selenocysteine (Sec) is directly incorporated into proteins in response to in-frame opal stop codons (UGA), instead of being created by posttranslational modification[29,30]. Afterward, this system was regarded as an expansion of the genetic code and selenocysteine was titled the 21st amino acid[31,32]. Sixteen years later pyrrolysine (Pyl) was found to be the 22nd genetically encoded amino acid, this time in response to the amber stop codon (UAG)[33–35]. Whereas Sec is present in prokaryotes and eukaryotes[36], the distribution of Pyl appears limited to the *Methanosarcinacea* and Gram-positive *Desulfitobacterium hafniense*[37]. Furthermore, both amino acids differ in their aminoacylation mechanism. Sec is made via an enzymatically modified serine that was charged to a special selenocysteinyl-tRNA. In contrast, Pyl is directly paired to pyrrolysyl-tRNA (PylT) by the cognate aminoacyl-tRNA synthetase PylS[37–39].

In addition to Sec and Pyl, even more deviations have been found to the standard genetic code. Genome analyses revealed ten codon reassignments in prokaryotic and eukaryotic nuclear codes which all are a subset of 16 changes occurring in mitochondrial codes[40]. Moreover, some methanogenic archaea compensate the lack of a canonical cysteinyl-tRNA synthetase by a particular pathway using *O*-phosphoserine that is enzymatically converted to cysteine prior to incorporation into a nascent protein[40]. All these modifications show a certain flexibility of the genetic code towards evolutionary novelties, giving the potential for additional genetically encoded nonstandard amino acids that might exist in still-uncharacterized genomes[40]. However, the search for the 23rd amino acid has not yet been successful, making the appearance of further widely spread amino acids improbable[41].

At the end of the last century, scientists began to exploit the degeneracy of the genetic code, in order to artificially expand it for the genetically encoded incorporation of amino acids with new functionalities into proteins. These “unnatural” amino acids (UAAs) bear many different functional groups, such as posttranslational modifications, UV-inducible crosslinkers, spectroscopic probes and chemical handles that can be modified chemically, even in living cells[42]. This required the introduction of exogenous tRNAs and their cognate aaRSs into the host cell, which have to work completely orthogonal to the endogenous components. That means, the introduced tRNAs should not be charged with any canonical amino acid by the host’s aaRSs and, in turn, the orthogonal aaRS should not aminoacylate any endogenous tRNA with UAAs. The anticodon of the orthogonal tRNA is typically complementary to blank (nonsense, frameshift, or otherwise unused) codons, especially the rarely used amber stop codon, allowing the reassignment of the appropriate codon to the amino acid used as a substrate by the orthogonal aaRS. The feasibility of this method was proven in 1998 by Furter who was able to site-specifically incorporate *p*-fluoro-phenylalanine (*p*-F-Phe) into dihydrofolate reductase (DHFR) in *E. coli* expressing a yeast amber suppressor tRNA/phenylalanyl-tRNA synthetase (PheRS) pair[43].

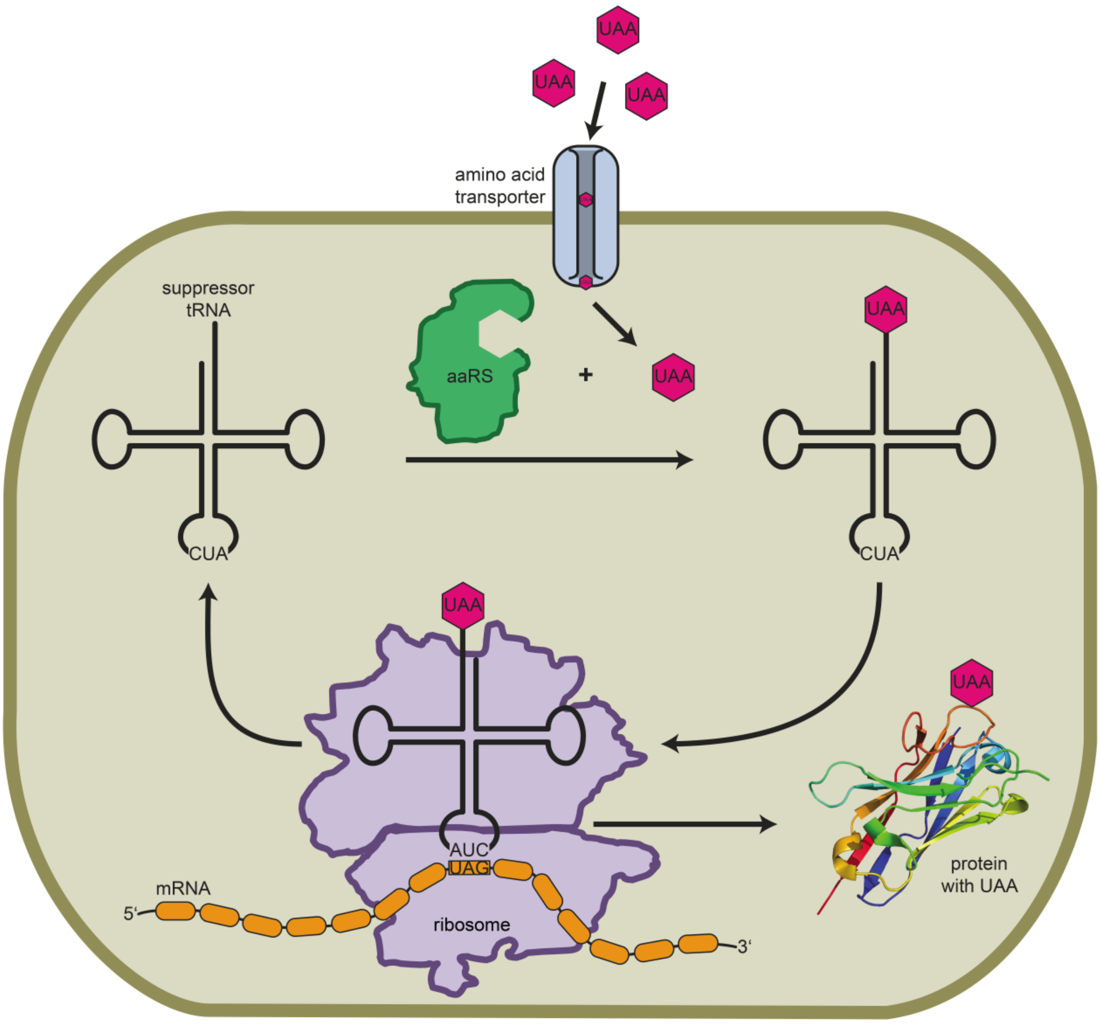


Figure 2.1: Schematic view for genetic code expansion.
Desired unnatural amino acids (UAAs) are taken up by the cell using endogenous transporters. These UAAs are used by evolved aminoacyl-tRNA synthetases (aaRSs) to charge corresponding evolved tRNAs. The tRNAs are then used by the ribosomes to decode (mostly) amber codons introduced in the mRNA, to incorporate the UAAs at predetermined sites on the protein of interest.

Whereas the yeast PheRS accepted the substrate analogue *p*-F-Phe without further modifications, advances in this system necessitated the adaption of aaRS' specificities towards specific UAAs[44]. Furthermore, the original standard amino acid phenylalanine was still a substrate for the PheRS, resulting in non-homogenously labeled protein. This disadvantage was overcome by a large excess (up to 30-fold) of *p*-F-Phe supplementation in the growth medium[43]. The first tRNA/aaRS pair that was truly orthogonal and only recognizing the desired UAA was derived from the tyrosyl pair from *Methanococcus jannaschii* (*M. jannaschii*) in the workgroup of Peter Schultz. This pair was evolved to incorporate *O*-methyl-L-tyrosine into DHFR in response to an amber codon[45] (Figure 2.1). To achieve this, they developed a systematic approach to alter the specificity of a synthetase for a certain UAA. First, all active-site residues interacting with the actual tyrosine substrate were randomly mutated, yielding a large library of synthetase variants which were passed through multiple rounds of stringent positive and negative selection. The positive selection was based on a reporter plasmid containing an antibiotic resistance gene with amber mutations at permissive sites, in this case a chloramphenicol acetyltransferase (CAT). *E. coli* cells simultaneously transformed with this reporter and the aaRS library plasmids were only able to survive in media containing chloramphenicol and the UAA, if they harbored a functional synthetase variant recognizing either the UAA or a natural amino acid (also see Ch. 3.2.4.1). The subsequent negative selection eliminated undesired aaRSs suppressing amber codons with endogenous amino acids in the absence of the UAA. Therefore, active synthetase clones from the first round were combined with a reporter plasmid comprising an amber mutant of the toxic barnase gene. Clones that produce full-length barnase protein using canonical amino acids will die because of its ribonuclease activity, which is toxic to cells without its specific inhibitor barstar[46] (also see Ch. 3.2.4.3). Multiple rounds of these two selections were performed, leading to an orthogonal, highly specific and amber suppressing aaRS variant[44] (Figure 2.2).

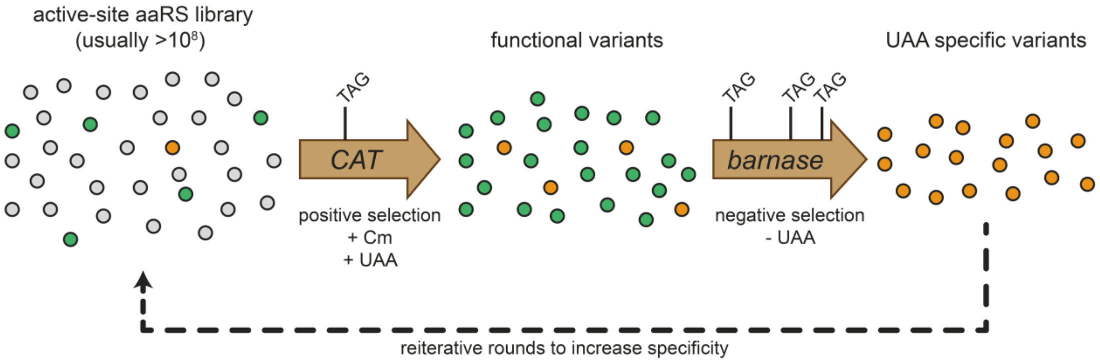


Figure 2.2: General selection scheme for the evolution of aminoacyl-tRNA synthetase (aaRS) substrate specificities for unnatural amino acids (UAAs)[44].
First, non-functional aaRS library variants are removed in a positive selection in the presence of the UAA and chloramphenicol (Cm). Functional variants suppress the amber codon within the chloramphenicol acetyltransferase (CAT) gene with both natural and unnatural amino acids. Synthetases specific for the UAA are isolated in a negative selection in the absence of the UAA, while suppression of the amber codons in the barnase gene leads to cell death.

The success in *E. coli* led to the development of a very similar selection approach in yeast, driven by amber suppression in the transcriptional activator GAL4. The production of full-length GAL4 in the presence of the UAA activates the expression of GAL4-responsive *HIS3*, *URA3*, and *lacZ* reporter genes, allowing for survival on media lacking histidine or uracil. Negative selections to remove unspecific synthetases are based on the conversion of the protoxin 5-fluoroorotic acid (5-FOA) to its toxic product by the *URA3* gene product on media without the UAA[44,47].

The evolution of tRNA synthetases in cells of higher eukaryotes, like mammalians, is more difficult due to technical issues concerning transformation efficiency, slow doubling times, and growth conditions. In this direction, a shuttle approach was applied with tRNA/aaRS pairs, which are orthogonal both in *E. coli* or *Saccharomyces cerevisiae* (*S. cerevisiae*) and in mammalian cells. Pairs evolved for UAAs in the easier to handle bacterium or yeast can then be transferred into mammalian cells while keeping orthogonality[48].

Using the aforementioned techniques a variety of orthogonal tRNA/aaRS pairs from various organisms have been used to add up to 100 unique unnatural amino acids to the genetic code of prokaryotes, including *E. coli* and some mycobacteria, and eukaryotes, like the yeasts *S. cerevisiae* and *Pichia pastoris* and even the multicellular organisms *Caenorhabditis elegans* and *Mus musculus*[42,48]. However, the majority of all genetic code expansion approaches were performed with only four different tRNA/aaRS pairs, each suitable for a particular model organism. First, the already mentioned *M. jannaschii* tyrosyl pair (MjYRS/MjYT) is orthogonal only in *E. coli* and other bacteria. The second and third are two synthetases from *E. coli* decoding for tyrosine (EcTyrRS) and leucine (EcLeuRS) in combination with their cognate tRNAs and can be utilized only in yeast, mammalian and other eukaryotic cells. Lastly, the aforesaid PylS/PylT pair from *Methanosarcina* species provides the advantage of being orthogonal in both bacteria and eukaryotic cells, showing no cross-reactions with endogenous synthetases or tRNAs. Additionally, the natural aaRS substrate specificity must not be destroyed before the evolution for a new UAA, since it decodes pyrrolysine and none of the 20 canonical amino acids[49].

2.3 Limitations And Optimizations of The Expanded Genetic Code

Although the genetically encoded incorporation of UAAs into proteins works, in principle, this system has to cope with a large range of limitations. These restrictions reduce the efficiency or even the feasibility of the system. Not all the desired UAAs exhibit the cell permeability required to cross the cell membrane, impeding the uptake by the cell and thereby the integration into the translational apparatus. Even if the UAA can be imported, the structural nature of the aaRS' active site must allow the evolution of its specificity to the particular amino acid. Furthermore, the incorporation efficiency of UAAs at certain sites in the protein can also be negatively influenced by effects of the mRNA context[50] and the local protein structure[51], as well as further factors like protein folding and stability[48].

One of the main reasons for decreased yields of proteins with UAAs is the competition of reassigned codon suppression with other cell activities, such as binding of a release factor to a stop codon or the erroneous recognition of a frameshift codon by three-base anticodon tRNAs. In *E. coli* suppression of the most commonly used blank codon, the amber codon, is impaired by the interaction with its release factor RF1, leading to truncated proteins. The incorporation of several UAAs into the same protein intensifies this effect. A simple deletion of the RF1-encoding gene, *prfA*, seemed to be unfeasible since this factor has been reported to be essential[52,53].

However, a couple of workgroups successfully reduced the competing effect of RF1 or even managed to perform a *prfA* knockout. For instance, enhanced amber suppression efficiency by overexpressing the C-terminus of the ribosomal protein L11, which has been suggested to play an important role in RF1-mediated translation termination, facilitated the incorporation of up to three acetyllysines into the same green fluorescent protein (GFP)[54]. Rydén and Isaksson found a temperature sensitive RF1 mutant that showed an increased efficiency to several amber tRNA suppressors[55]. This could be utilized to favor UAA incorporation in response to amber codons, but ongoing growth at high temperature would be fatal for the cell and could affect the production of many recombinant proteins adversely.

The Nakamura group was able to compensate for this temperature dependent RF1 growth deficit by engineering RF2 with a single amino acid substitution to an omnipotent release factor (RF2*), capable to terminate translation at all three stop codons[56]. The use of RF2* even permitted a chromosomal RF1/RF2 double knockout, indicating that RF1 is nonessential in *E. coli*. This hypothesis was emphasized by the work of Wang and co-workers who showed the feasibility of a *prfA* deletion strain by "fixing" the RF2-encoding *prfB* gene. The engineering of their RF2 variant (*prfB*^{f}) included the removal of an in-frame UAG autoregulation element and the reversion of a peculiar mutation (A246T) that only occurs in *E. coli* strains derived from K-12, impairing RF2's release activity for the ochre (UAA) codon by a factor of five[57]. The reduced activity of mutated RF2 was supposed to be the reason for the apparent RF1 essentiality. Indeed, later experiments confirmed that previously reported attempts to knock out RF1 failed because they used *E. coli* K-12 strains containing the peculiar A246T mutation. Deletion trials using three common *E. coli* B strains (REL606, BL21, and BL21(DE3)), derived from the second progenitor of most *E. coli* strains and encoding wild type RF2, allowed successful *prfA* knockout. Thus, RF1 was proposed to be nonessential for wild type *E. coli* with nonmutated RF2[58]. The employment of RF1 knockout strains with either wild type background (B strains) or the "fixed" RF2 facilitated the genetically encoded incorporation of up to ten identical UAAs into the same protein.

The workgroups of Yokoyama and Sakamoto accomplished the deletion of *prfA* while leaving RF2 untouched. They used a combination of a bacterial artificial chromosome, harboring the seven essential open reading frames (ORFs) of *E. coli* naturally ending with an UAG codon but replaced by UAA, and an amber suppressor tRNA. This kept the cells viable and permitted the UAG-dependent incorporation of up to ten natural or six unnatural amino acids into the same polypeptide[59]. The role of the amber suppressor tRNA was supposed to alleviate the significantly impaired fitness observed for all mentioned RF1 knockout strains. Its UAG-decoding activity was thought to prevent ribosome stalling at UAG codons and to allow the expression of the nonessential ORFs ending with an amber codon which are then terminated by naturally occurring in-frame "backup" stop codons, opal or ochre, downstream of the suppressed one. Stalled ribosomes have at least three disadvantages that reduce a cell's fitness. First, they cannot finish required protein synthesis. Second, if stalled they are "trapped" to the particular mRNA and cannot be recycled to translate further mRNAs. Third, the tmRNA surveillance system[60,61] is activated, a rescue mechanism responsible for cleavage of the appropriate mRNA and tagging of the premature protein for degradation. Degradation of an overwhelming number of proteins can induce cell death[57,59]. Indeed, Yokoyama and Sakamoto could show in following experiments that an efficient decoding of the amber codon, using a tRNA variant with enhanced suppression activity, instead of reducing the number of UAGs in the genome improved the growth of a *prfA*-deficient strain of *E. coli*[62].

By contrast, the laboratories of Church and Isaacs used multiplex automated genome engineering (MAGE)[63] to create an *E. coli* strain that had all known UAG stop codons replaced with synonymous UAA codons[64]. Afterwards, they deleted RF1 in this strain, reassigning UAG as a sense codon, and compared their approach with those reported by Wang as well as Yokoyama and Sakamoto. They found that their strain was the only one without deleterious effects causing impaired fitness. Furthermore, they proposed that RF1 is essential only for UAG translational termination and not for UAA termination or other essential cellular functions, since RF2 was kept unmodified[65].

Chin and co-workers pursued an alternative strategy, which did not focus on deletion or mutation of the release factors. They utilized orthogonal ribosomes (O-ribosomes) to reduce the competing effect of RF1. Developed by Chin and Rackham, O-ribosomes form the basis of a parallel and independent translational machinery since their anti-Shine-Dalgarno sequence was mutated to exclusively bind to the likewise modified Shine-Dalgarno sequences of orthogonal mRNAs (O-mRNAs). These O-mRNAs, containing an orthogonal ribosomal binding site (ORBS), are in turn, not a substrate for endogenous ribosomes[66]. As part of a parallel and independent system O-ribosomes are unburdened from the maintenance of housekeeping gene expression and therefore evolvable to handle new challenges, whereas mutations of the natural ribosome are potentially lethal to the cell. By mutating merely two positions (U531G and U534A) in a loop of the 16S rRNA in the A site, that was found to be in close proximity to both the anticodon of tRNAs bound to mRNA and the RF1, Chin and co-workers created an evolved O-ribosome (termed ribo-X) which showed greatly enhanced suppression of amber codons placed in O-mRNAs. They hypothesized that this improvement resulted from a decreased functional interaction of ribo-X with RF1[67].

Even in systems optimized for release factor competition an adequate delivery of tRNAs charged with the correct UAA is required. Therefore, the expression of the appropriate tRNAs and their cognate aaRSs have to be carefully balanced, to guarantee an efficient translation but also avoid the overconsumption of biological resources and the generation of toxic side effects by an excessive production of these components. Furthermore, unnecessary metabolic burden for the cell, provoked by, for example, plasmids and corresponding antibiotics, should be prevented. With regard to these issues, several strategies to improve full-length protein yields have been investigated. The simplest solution involved combining the genes encoding the orthogonal tRNA and synthetase on a single plasmid compatible with most expression vectors and strains[68,69]. The Schultz laboratory developed a single-plasmid system (pSup) that was consecutively optimized. By merging derivatives of the *M. jannaschii* tyrosyl pair under the control of enhanced promoters (*proK* instead of *lpp* promoter for the $tRNA_{CUA}$ and a strong mutant *glnS* promoter instead of the wild type variant for the aaRS) on a single vector and raising the tRNA copy number from one to three or six, they observed a 20-fold[48] increase in the yield of an UAA-containing myoglobin protein in *E. coli*[69]. An improved version of pSup, called pSUPAR, comprised an additional copy of the synthetase under the control of an arabinose-inducible *ara* promoter[70]. The replacement of the polycistronic tRNA cluster in pSUPAR with only one copy of an optimized *M. jannaschii* amber suppressor tyrosyl-tRNA, which has been reported to be generally less toxic to *E. coli* and to show increased UAA incorporation efficiency with several aaRS[71], resulted in pEVOL. Compared to the first described pSup, pEVOL led to roughly 250% greater yields of UAA-containing proteins[51]. Based on these insights, they finally designed the suppressor plasmid pUltra which harbored a single copy each of the tRNA cassette from pEVOL and the synthetase MjYRS under the control of a *tacI* promoter. They found pUltra to be better than pEVOL, particularly when tested with weakly expressing mutants or inherently weak suppressor tRNA/aaRS pairs, and achieved the efficient incorporation of up to three UAAs in response to amber stop codons within the same GFP protein.

Since pUltra and pEVOL both have a unique origin of replication and antibiotic resistance marker they are mutually compatible. This was exploited to simultaneously insert two different UAAs into the same protein using amber and ochre codons with pEVOL encoding the *M. jannaschii* amber suppressor pair and pUltra a pyrrolysyl-tRNA/aaRS pair optimized for enhanced ochre suppression. Furthermore, the simple configuration of pUltra enabled the generation of pUltraII, a single plasmid containing both suppression systems for amber and ochre codons. The dual suppressor pUltraII showed similar full-length protein expression levels relative to pUltra and pEVOL in combination, but reduced the metabolic burden for the cell by one plasmid and thereby one antibiotic[72].

Further efforts to find and optimize factors that affect the incorporation efficiency of unnatural amino acids were made. For instance, in one approach MjYRS was mutated to mediate a better recognition of the cognate tRNA's CUA anticodon[73]. Another study described the development of EF-TU mutants that better accept UAAs, particularly those with large side chains or altered backbones[74]. But the simultaneous incorporation of two or more distinct UAAs into the same polypeptide poses a difficult challenge.

Each individual UAA necessitates its own orthogonal tRNA/aaRS pair decoding a unique codon. The use of stop codons to generate proteins with multiple UAAs, as reported by Wan *et al.*[75] or Chatterjee *et al.*[72] (employing pUltra and pEVOL as described above), quickly reaches its limit. Reassigning some of the 61 other triplet codons encoding natural amino acids would be an option but not easy to handle because this requires genome wide replacements with yet unknown consequences. In addition, even if feasible, this would provide a relatively small set of new blank codons, whereas the usage of codons extended to four bases theoretically affords 256 novel blank codons. The combination of a quadruplet and an amber codon has already been successfully utilized to incorporate two unremarkable UAAs into the same model protein[76]. However, the capability of natural ribosomes to decode four-base codons with extended anticodon tRNAs is moderate, probably due to poor tRNA accommodation in the ribosome's decoding center[77]. This inefficiency cannot be circumvented by evolving the endogenous ribosome because this potentially leads to cell death. For this reason, the Chin laboratory further evolved their orthogonal ribo-X to create an O-ribosome, called ribo-Q1, which exhibits enhanced quadruplet and amber decoding on O-mRNAs. Although they randomly mutated 127 nucleotides of the A site's 16S rRNA, ribo-Q1 contains just two mutations (A1196G and A1197G) with respect to ribo-X, allowing an efficient decoding of four-base codons approaching the level of triplet decoding on the natural ribosome[78].

Finally, the number of currently existing mutually orthogonal tRNA/aaRS pairs in bacteria, namely the MjYT/MjYRS pair from *M. jannaschii* and the PylT/PylS pair from *Methanosarcina* species, confines the incorporation of different UAAs into the same polypeptide to two. Since it is not clear if a sufficient number of pairs can be discovered that have been formed by natural evolutionary divergence and which fulfill the absolute required criterion to be orthogonal to both endogenous and already existing exogenous pairs to encode additional unnatural amino acids, alternative strategies need to be invented. One approach was reported by Neumann *et al.* who designed an orthogonal tRNA/aaRS pair *de novo* from an existing one by duplicating the MjYT/MjYRS pair in a logical series of mutagenesis and selection steps[79].

2.4 Applications of The Expanded Genetic Code

The principles of genetic code expansion haven been applied in numerous studies in order to investigate the nature of a protein itself or to give a protein new properties which are not provided by the common 20 amino acids. Site-specifically incorporated unnatural amino acids offer a myriad of new functionalities, such as posttranslational modifications, UV-inducible crosslinkers, spectroscopic probes and chemical handles that can be subsequently altered, *in vitro* and in some cases even *in vivo*[42,48,49]. In the following, a small selection of UAAs and their application is presented, exemplifying the enormous potential of an expanded genetic code for basic and applied scientific research.

To understand how biological processes form life, one has to investigate the participating proteins and their interaction partners. This is sometimes problematic when these interactions are only transient, weak or at particular subcellular locations. Genetically encoded crosslinkers that react with nearby molecules in response to (UV)-light facilitate the covalent trapping of interaction partners for subsequent identification. Several crosslinking systems exist for bacteria, yeast and mammalian cells, including azido derivatives of phenylalanine[47,80], diazirines[81,82] and benzophenones, such as the by far most frequently used *p*-benzoyl-L-phenylalanine (BPA)[47,83]. BPA helped to address questions concerning the mechanism of polypeptide translocation across the cytoplasmic membrane in *E. coli* mediated by the bacterial SecY-SecE-SecG (SecYEG) translocon, a channel-like transmembrane complex, and the SecA ATPase[84,85]. Wilkins *et al.* used BPA to study how histone modifications in yeast influence the complicated but important biological process of chromosome condensation in mitosis[86] and Tagami *et al.* to define conformational changes in the bacterial RNA polymerase of *Thermus thermophilus*[87].

A cells' proteome is dynamically controlled by posttranslational modifications (PTMs) that can change a protein's molecular function, affinity for binding partners, localization and stability. However, for the majority of these PTMs the modifying enzymes are still unknown, making the production of homogeneously and site-specifically modified proteins challenging. This can be achieved by using genetic code expansion, which permits the genetically encoded installation of PTMs into proteins, in order to assess the role of a protein's PTMs. So far, a variety of methods have been reported for incorporating nitrated[88] and sulfated[89] tyrosine, analogues of phosphorylated tyrosine[90], phosphorylated serine[91], mono-[92,93] and dimethylated[94], acetylated[95] and ubiquitinated[96] lysine into recombinant proteins. Employing these methods, the effect of H3 K56 acetylation on nucleosome and chromatin structure and function has been examined, revealing that this modification mediated an increased DNA breathing on the nucleosome[97]. In another case the role of an acetylated lysine in a protein involved in HIV infection, cyclophilin A (CypA), has been identified to regulate key functions of CypA that, among other things, lead to diminished effects of the widely-used immunosuppressive drug cyclosporine[98].

Furthermore, UAAs with spectroscopic properties have been installed into proteins. For example, the phenylalanine derivative *p*-azido-L-phenylalanine (AzF), which can also be used as a crosslinker[47,80] (see above), is suitable for infrared (IR) spectroscopy since its vibration of the azido group absorbs in a clear spectral window at ~2100 cm^{-1}, well separated from intrinsic protein vibrations. This has been exploited to examine fast conformational changes, that the G protein-coupled receptor rhodopsin undergoes after light activation and binding a cytoplasmic G protein, by Fourier transform infrared (FTIR) difference spectroscopy[99,100]. Another IR-labeled phenylalanine, bearing a cyano group in place of the azido moiety (*p*-cyano-L-phenylalanine), absorbing at ∼2200 cm^{-1}, was incorporated into myoglobin to probe metal ion and ligand binding[101]. Unnatural amino acids holding ^{19}F, ^{13}C or ^{15}N isotopes[70,102–104], that can be detected by nuclear magnetic resonance (NMR), enable single site structural investigations of proteins, in contrast to global labeling by growth medium supplementation with ^{13}C-glucose and ^{15}N-ammonium. Li *et al.* could perform NMR supported analysis of protein dynamics inside of cells, “in-cell NMR” (also see Ch. 5.4), since they were able to detect several site-specifically ^{19}F-labeled proteins in living *E. coli*[104].

Genetically encoded UAAs were not only used to determine functionalities of proteins and their related biological processes. Liu and colleagues introduced, among others, *p*-acetyl-L-phenylalanine (pAcF) into different *cis*-regulatory leader-peptide elements. These small molecules served as ON or OFF switches and thereby regulated the transcription of downstream genes[105]. The same amino acid was also employeed to make the therapeutic agent, human growth hormone (hGH), more viable without perturbing its biological activity. To this end, the reactive functional group of pAcF was posttranslationally modified by PEGylation (covalent attachment of polyethylene glycol (PEG)) in a bioorthogonal “click reaction” (for “click reaction” also see next chapter and Ch. 5.1 to 5.3)[48,106].

2.5 The Genetically Encoded Incorporation of FRET Pairs

Highly sensitive fluorophores belong to a class of spectroscopic probes not mentioned in the previous chapter. They can be used to determine protein expression, activity, localization, and function[107]. GFP is one of the most widely studied and exploited fluorophores in biochemistry and cell biology[108]. However, as a fluorescent protein it needs to be correctly folded to be functional. Thus, GFP cannot be site-specifically placed at any position in a protein of interest and is therefore fused to the protein’s termini. In addition, its size would perturb the structure and function of the investigated protein.

Genetically encoded (small) fluorescent UAAs, in theory, would cause minimal structural perturbation and are unlikely to impair a protein’s function and localization. Unfortunately, many fluorophores are not cell-permeable or are simply too large to be a substrate for the aaRSs. Hence, only a small number of fluorescent UAAs have been directly incorporated into proteins [109–112]. Nevertheless, Charbon *et al.* were indeed able to localize the chaperonin protein GroEL under normal and stress conditions in living cells by labeling with a coumarin fluorophore containing UAA, whereas GFP fusions proved to interfere with GroEL function[113]. Summerer *et al.* incorporated a dansyl-UAA into the protein human superoxide dismutase to follow its denaturation in the presence of guanidinium chloride[110].

In order to facilitate the labeling of proteins with a broader range of fluorophores, since hundreds of small organic dyes are commercially available[107], one had to find an alternative strategy. The genetically encoded installation of an UAA containing a bioorthogonal reactive moiety provides one avenue because they allow the subsequent site-specific modification of a protein with almost any probe by bioorthogonal "click chemistry"[114] and hence also any fluorophore that is compatible with the installed UAA.

"Click reactions" between an azide and a terminal alkyne necessitate the use of cytotoxic Cu(I) to catalyze the cycloaddition[115], impeding *in vivo* applications. More recently, Cu(I)-free approaches with cyclooctyne derivatives were developed, which activate the alkyne via ring strain[116,117]. The Ebright laboratory coupled fluorescent probes by Staudinger ligation to genetically encoded AzF to investigate the opening and closing of the bacterial RNA polymerase clamp[118]. Intracellular reduction is one disadvantage of genetically encoded azides. Alternative strategies employ inverse electron-demand Diels-Alder cycloaddition reactions between genetically encoded strained alkenes and alkynes and labeling reagents conjugated to tetrazines and/or azides[119]. The lysine derivatives *N*ε-5-norbornene-2-yloxycarbonyl-L-lysine (NorK) and bicyclo[6.1.0]non-4-yn-9-ylmethanol-L-lysine (BCNK) have been successfully incorporated into proteins by the Chin laboratory, followed by labeling with fluorophores conjugated to tetrazines[120,121].

The simultaneous installation of two or more fluorophores in a single polypeptide permits the investigation of conformational changes, even on a single-molecule (sm) level, using the powerful technique Förster/Fluorescence resonance energy transfer (FRET). FRET relies on the energy transfer from a donor to an acceptor fluorophore in a distance-dependent manner and is capable of detecting distances and their changes in a nanometer scale, both *in vitro* and *in vivo*[122,123]. However, the precise and site-specific labeling of the proteins with suitable fluorophores is essential for FRET experiments but a challenging task and is often the limiting factor. Brustad *et al.* studied T4 lysozyme folding using FRET with site-specific dual-labeling. One label was generated by a genetically encoded pAcF with a hydroxylamine-containing fluorophore and the other from a single cysteine with a maleimide conjugated fluorophore[124]. Nevertheless, this technique is restricted in its applicability to single-cysteine proteins. The incorporation of two distinct and suitable UAAs into the same protein and the subsequent bioorthogonal labeling of these provides an elegant method for the genetically encoded incorporation of FRET pairs into proteins[72,78].

2.6 Specific Aims

This work is intended to apply the principles of genetic code expansion to achieve the efficient incorporation of two different unnatural amino acids into the same polypeptide, allowing us the generation of proteins with a variety of functionalities, including FRET pairs, multiple different NMR and IR probes, combinations of posttranslational modifications and bioorthogonal reactive groups for subsequent chemical modifications.

The eye of a needle for this aim is the efficiency of (multiple) UAA incorporation. Although several trials to optimize this efficiency have been performed by others in the past, none of these approaches provided a systematic study on the correlation of aaRS and tRNA expression levels including the aminoacylation status of the tRNA. Thus, only little insight is given into which of these components is the actual limiting factor. It is therefore necessary to establish assays for the detection of the abundance and activity of each component, in order to identify and optimize the constraints of the genetic code expansion system.

Once the system is optimized, we will start to focus on the production of proteins with FRET pairs to investigate their conformational states and the transitions between them. This can be studied in single-molecule FRET experiments for which the precise labeling of proteins with suitable fluorophores is a limiting factor.

3 Materials And Methods

3.1 Materials

3.1.1 Devices And Instruments

AbiPrism 3100 DNA Sequencer	Applied Biosystems, Darmstadt
Accumax Pipet Help	Accumax, INDIA
Agarose Gel Electrophoresis Chamber	GP-Kuststofftechnik, Kassel
Autoclave HST 4-5-8	Zirbus, Bad Grund
BioPhotometer	Eppendorf, Hamburg
Biovortex V1	Peqlab, Erlangen
Bunsen Burner Fuego Basis	WLD-Tec, Göttingen
Cell Sorter iCyt Synergy	Sony Biotechnology, USA
Centrifuge 5415R	Eppendorf, Hamburg
Centrifuge Allegra 2IR	Beckman Coulter, Krefeld
Centrifuge Avanti J-20 XPIJA-20	Beckman Coulter, Krefeld
Centrifuge HERAEUS Pico 17	Thermo Scientific, Schwerte
Concentrator 5301	Eppendorf, Hamburg
Electroblotter (Semidry)	PeqLab, Erlangen
Electroblotter (Semidry; 20 × 20 cm)	VWR International, Darmstadt
Electroporator Easyject Prima	EquiBio, Willstätt
FLUOstar Omega	BMG Labtech, Ortenberg
Gel Doc 2000	BioRad, München
Gel Shaker Duomax 1030	Heidolph, Schwabach
Gel Shaker Rotamax 120	Heidolph, Schwabach

Hamilton Syringe 50 µL	Hamilton, USA
Hybridizer Oven HB-1000	UVP, USA
Hypercassette 18 × 24 cm	GE Healthcare, München
Incubator Mytron WB 60 k	Mytron, Heiligenstadt
Labcycler	SensoQuest, Göttingen
Magnetic Stirrer MR Hei-Standard	Heidolph, Schwabach
Magnetic Stirrer MR3000	Heidolph, Schwabach
Microfluidizer 110 S	Microfluidics, USA
Optimax X-Ray Film Processor	Protec, Oberstenfeld
pH Meter PT-15	Sartorius, Göttingen
Photometer	Biometra, Göttingen
Pipets Research Plus (10, 100, 1000 µL)	Eppendorf, Hamburg
Power Supply 300V	VWR International, Darmstadt
Power Supply EV231	Consort, BELGIUM
Power Supply MP-250V	Major Science, USA
Rotating Mixer RM5 Assistant 348	Karl Hecht KG, Sondheim/Rhön
Rotor JA-20	Beckman Coulter, Krefeld
Rotor JLA-8.1000	Beckman Coulter, Krefeld
Scale 770	Kern, Balingen-Frommern
Scanner CanoScan 5600F	Canon Deutschland, Krefeld
Sonifier 250	Branson, USA
Special Accuracy Scale BP410S	Sartorius, Göttingen
Tankblotter Criterion (1.3 L)	BioRad, München
Tankblotter Mini Trans-Blot (0.45 L)	BioRad, München

Thermomixer comfort 1.5 mL	Eppendorf, Hamburg
Typhoon 9400 Variable Mode Imager	Amersham (GE Healthcare), München
Unitron Incubators	Infors HT, Einsbach
UV table TFX-20.LC	Vilber Lourmat, Eberhardzell
Vertical Double Gel Systems	PeqLab, Erlangen
Vertical Gel Electrophoresis Chamber H10 Mini	GP-Kunststofftechnik, Kassel
Vortex generator VV3	VWR International, Darmstadt
X-Ray Cassette 18 × 24	Rego X-Ray GmbH, Augsburg

3.1.2 Chemicals

All chemicals were purchased from those companies listed below, unless stated otherwise, and fulfilled the purity grade “pro analysis”.

AppliChem, Darmstadt

BioRad, München

Merck, Darmstadt

Roth ,Karlsruhe

Sigma-Aldrich, Steinheim

VWR International, Darmstadt

3.1.3 Consumables And Other Materials

96-Well Black Microplates	VWR International, Darmstadt
Amersham ECL Plus WB Detection Reagent	GE Healthcare, München
Amersham ECL Prime WB Detection Reagent	GE Healthcare, München
Amersham ECL Select WB Detection Reagent	GE Healthcare, München

Amersham Hyperfilm ECL	GE Healthcare, München
Amicon Ultra-15 Centrifugal Filter Units (NMWL 3 kDa, 10 kDa, 30 kDa, 50 kDa)	Merck, Darmstadt
CDP-Star ECL Substrate	Roche, Mannheim
Centrifugal Ultrafiltration Devices (MWCO 3k, 10k)	Sartorius, Göttingen
Coli Rollers Plating Beads	Novagen, USA
Deep Well Plates, 2 mL - Sterile	Axygen Scientific, USA
Electroporation Cuvettes, 2 mm gap	Peqlab, Erlangen
Eppendorf Tubes (1.5 mL, 2.0 mL)	Eppendorf, Hamburg
Falcon Tubes (15 mL, 50 mL)	Sarstedt, Nümbrecht
Gas Permeable Adhesive Seals	Thermo Scientific, Schwerte
GSH Sepharose Beads	GE Healthcare, München
HisLink Protein Purification Resin	Promega, USA
Hybond-N+ Transfer Membrane, Nylon	GE Healthcare, München
Immobilon-P Transfer Membrane, PVDF	Merck, Darmstadt
Instant Blue	Biozol, Eching
PCR Soft Tubes (0.2 mL)	Biozym, Austria
peqGOLD Universal Agarose	Peqlab, Erlangen
Petri Dishes 145 × 20 mm	Greiner Bio One, Solingen
Petri Dishes 92 × 16 mm	Sarstedt, Nümbrecht
Pipet Tips	Sarstedt, Nümbrecht
Plate 96 DW Square U	VWR International, Darmstadt
Protran Nitrocellulose Transfer Membrane	Whatman, Dassel
Slide-A-Lyzer® Mini Dialysis Units	Thermo Scientific, Schwerte

Spectra/Por Closures, 55 mm	Spectrum Laboratories, Netherlands
Spectra/Por Dialysis Membrane, 50 mm	Spectrum Laboratories, Netherlands
Syringe Filters (0.2 and 0.45 µm)	Sartorius, Göttingen
UV Cuvettes (UVette)	Eppendorf, Hamburg
Whatman filter paper	Whatman, Dassel

3.1.4 DNA, Protein And RNA Size Standards

GeneRuler™ 1 kb DNA Ladder	Thermo Scientific, Schwerte
PageRuler Prestained Protein Ladder	Thermo Scientific, Schwerte
Unstained Protein Molecular Weight Marker	Thermo Scientific, Schwerte
Low Molecular Weight Marker 10-100 nt	USB® Products/Affymetrix, USA
Low Range ssRNA Ladder	NEB, Frankfurt
Oligonucleotide Ladder (20, 35, 45, 58, 70 nt)	Self-made from Sigma-Aldrich primers

3.1.5 Antibiotics

Table 3.1: Overview of antibiotics used in growth media/agar plates

Antibiotic	Work Concentration [µg/mL]	Company
Ampicillin (Amp)	100	AppliChem, Darmstadt
Chloramphenicol (Cm)	50	AppliChem, Darmstadt
Kanamycin (Kan)	50	AppliChem, Darmstadt
Spectinomycin (Sm)	75	Sigma-Aldrich, Steinheim
Tetracycline (Tet)	25	AppliChem, Darmstadt

3.1.6 Culture Media And Agar Plates

For agar plates the corresponding medium was supplemented with 1.5% (w/v) agar-agar. M9 salts, LB and 2YT medium were sterilized by autoclaving at 121 °C for 20 min. Antibiotics and other supplements were added after cooling to at least 55 °C.

Table 3.2: Overview of growth media used for culturing of *E. coli*

LB Medium	2YT Medium	M9 Minimal Medium
10 g tryptone	16 g tryptone	200 mL M9 salts (5x)
5 g yeast extract	10 g yeast extract	2 mL $MgSO_4$ (1 M)
5 g NaCl	5 g NaCl	20 mL glucose (20%; w/v)
Adjust to 1 L ddH_2O	Adjust to 1 L ddH_2O	0.1 mL $CaCl_2$ (1 M)
		Adjust to 1 L ddH_2O
		M9 salts (5x)
		250 mM Na_2HPO_4
		100 mM KH_2PO_4
		50 mM NaCl
		100 mM NH_4Cl

3.1.7 Unnatural Amino Acids

Unnatural amino acids (UAAs) were dissolved with NaOH in ddH_2O (see Table 3.3) just before the addition to the medium.

Table 3.3: Overview of unnatural amino acids used for genetic code expansion

Amino acid	Stock Solution [M]	Work Concentration [mM]	Company
Boc-L-lysine (BocK)	0.5 (in 1 M NaOH)	1	Bachem, Bubendorf (CH)
4-azido-L-phenylalanine (AzF)	0.5 (in 0.5 M NaOH)	1-5	ChemImpex, USA
Norbornene-L-lysine (NorK)	0.1 (in 0.2 M NaOH)	2	AG Hell, Göttingen
Bicyclononynes-L-lysine (BCNK)	0.1 (in 0.2 M NaOH)	2	SynAffix, Nijmegen (NL)

3.1.8 Enzymes

Enzymes were used as recommended by the company’s protocol. The following table gives an overview about the enzymes used in this study.

Table 3.4: Overview of enzymes used

Enyzme	Company
T4 DNA ligase	Thermo Scientific, Schwerte
Shrimp Alkaline Phosphatase (SAP)	
Taq DNA Polymerase	
Phusion DNA Polymerase	
Restriction enzymes: BamHI, DpnI, HindIII, KpnI, MfeI, NcoI, NdeI, NotI, PstI, SacI, SalI, StuI, XhoI	
Proteinase K	
Alkaline Phosphatase, Calf Intestinal (CIP)	NEB, Frankfurt
Restriction enzymes: AgeI, BamHI, BsaI, DpnI, NdeI, NheI, SalI, StuI	
Expand High Fidelity Polymerase Mix	Roche, Mannheim
Pfu Turbo DNA Polymerase	Agilent, Böblingen

3.1.9 Cell Lines

Table 3.5: Overview of cell lines used

Cell line	Genotype	Company
E. coli DH10B	F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15Δ*lac*X74 *rec*A1 *end*A1 *ara*D139Δ (*ara*,*leu*)7697 *gal*U *gal*K λ$^-$ *rps*L *nup*G	Invitrogen, Darmstadt
E. coli BL21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	NEB, Frankfurt
E. coli SHuffle T7 Express	*fhuA2 lacZ::T7 gene1* [lon] *ompT ahpC gal* λatt::pNEB3-r1-*cDsbC* (SpecR, *lacI*^{q}*) ΔtrxB sulA11* R(*mcr-73::miniTn10*--Tet^S)*2 [dcm] R(zgb-210::Tn10* --Tet^S) *endA1 Δgor Δ(mcrC-mrr)114::IS10*	NEB, Frankfurt

3.1.10 Buffers And Solutions

All buffers were prepared with ddH_2O.

Table 3.6: Overview of commonly used buffers & solutions

Buffer/Solution	Ingredients
PIC (1×)	75 µM pefabloc SC 150 nM leupeptin 37.5 µM O-phenanthroline 500 nM pepstatin A Sterilize using syringe filters (0.2 µm)
PBS (1×)	137 mM NaCl 2.7 mM KCl 10 mM Na_2HPO_4 1.8 mM KH_2PO_4 pH 7.4
TBS (1×)	50 mM TRIS-HCl, pH 7.5 150 mM NaCl
$CaCl_2$ Solution for competent cells	60 mM $CaCl_2$ 10 mM Pipes-KOH, pH 7.0 15% glycerol (v/v) Autoclave Store at 4 °C
TBE (1×)	89 mM TRIS base 89 mM boric acid 2 mM EDTA-Na_2
DNA Loading Buffer (1×)	3% glycerol (v/v) 1 mM TRIS base, pH 7.5 1 mM EDTA-Na_2 Bromphenol blue Xylene cyanol
SDS Running Buffer (1×)	25 mM TRIS base 192 mM glycine 0,1% SDS (w/v)

SDS Sample Buffer (1×)	2.5% glycerol (v/v) 12.5 mM TRIS-HCl, pH 6.8 25 mM DTT 0.5% SDS (w/v) 0.025% bromphenol blue (w/v)
WB Transfer Buffer (1×)	1× SDS running buffer 15-20% MeOH (v/v)
PonceauS Solution (0.5%)	0.5% PonceauS (w/v) 5% TCA (w/v)
TE buffer (1×)	10 mM TRIS, pH 8.0 1 mM EDTA-Na_2
Acid Urea Sample Buffer	0.1 M sodium acetate, pH 5.0 8 M urea 0.05% bromphenol blue (w/v) 0.05% xylene cyanol (w/v)
Acid Urea Running Buffer (1×)	0.1 M sodium acetate, pH 5.0
NB Transfer Buffer (1×)	40 mM TRIS-HCl, pH 8.0 2 mM EDTA-Na_2
SSC Buffer (1×)	150 mM NaCl 15 mM trisodium citrate pH 7.0
Prehybridization Solution (1×)	6× SSC 10× Denhardt’s solution (Invitrogen) 0.5% SDS (w/v)
Hybridization Solution (1×)	6× SSC 0.1% SDS (w/v) tRNA probe 1:3000 (v/v) Sterilize using syringe filters (0.2 µm)
Buffer 1 (1×)	100 mM maleic acid 150 mM NaCl pH 7.5 (with NaOH pellets)
Blocking Reagent (10%)	10% blocking reagent ((w/v); Roche) 1× buffer 1 Autoclave and store at 4 °C

Buffer 2 (1×)	10% blocking reagent, 10% (v/v) 1× buffer 1
Buffer 3 (1×)	100 mM TRIS base, pH 9.5 100 mM NaCl

3.1.11 DNA Kit Systems

Kits were used according to manufacturer's protocol, except for XP2 binding buffer from peqGOLD Gel Extraction Kit and CP binding buffer from Cycle Pure Kit. XP2 was used in a w(gel)/v(buffer) ratio of 1:3 instead of 1:1 and CP buffer was used in a v(PCR)/v(buffer) ratio of 1:5 instead of 1:1.

peqGOLD Cycle-Pure Kit (S-Line)	Peqlab, Erlangen
peqGOLD Gel Extraction Kit (S-Line)	Peqlab, Erlangen
peqGOLD Plasmid Miniprep Kit I& II	Peqlab, Erlangen
QIAGEN Plasmid Mini Kit	QIAGEN, Hilden
QIAquick Gel Extraction Kit	QIAGEN, Hilden
QIAquick PCR Purification Kit	QIAGEN, Hilden

3.1.12 Fluorescent Dyes

Table 3.7: Overview of fluorescent dyes used for protein labeling.
All dyes were prepared as stock solutions with 10 mg/mL in DMSO. Abs. Max. = Absorption Maximum; Em. Max = Emission Maxium; Abb. = Abberior

Dye	Abs. Max. [nm]	Em. Max. [nm]	Conjugate	Company
Atto565	563	592	Maleimide	ATTO-TEC, Siegen
Abb. Star635	634	654	Tetrazine	AG Hell, Göttingen[125]
Abb. Star635P	635	655	Tetrazine	AG Hell, Göttingen[125]
KK9046	632	654	Tetrazine	AG Hell, Göttingen[125]
KK114	637	660	Tetrazine	AG Hell, Göttingen[125]

3.1.13 Antibodies

Table 3.8: Overview of antibodies used for immuno blotting (western & northern blots).
His = histidine; GFP = green fluorescent protein; GST = glutathione *S*-transferase; DIG = digoxigenin; Myc = myelocytomatosis; MBP = maltose binding protein; AP = alkaline phosphatase; HRP = horseradish peroxidase

	Antigen	Host	Conjugate	Diluent (w/v)	Product code	Company
Primary Dilution 1:3.000 (v/v)	His	mouse	-	3% BSA-PBS	27-4710-01	GE Healthcare, München
	GFP	mouse	-	3% BSA-PBS	sc-9996	Santa Cruz, Heidelberg
	GST	goat	-	3% BSA-PBS	27-4577-01	GE Healthcare, München
	DIG	sheep	AP	1:10.000 in 1× Buffer 2	11093274910	Roche, Mannheim
	Myc	mouse	-	5% Milk-TBS	sc-40	Santa Cruz, Heidelberg
	MBP	rabbit	-	3% BSA-PBS	#E8030S	NEB, Frankfurt
Secondary Dilution 1:10.000 (v/v)	mouse	goat	HRP	5% Milk-PBS (TBS for Myc)	ab6789 / A8924	Abcam, UK / Sigma-Aldrich, Steinheim
	rabbit	goat	HRP	5% Milk-PBS	ab6721 / R5506	Abcam, UK / Sigma-Aldrich, Steinheim
	goat	rabbit	AP	5% Milk-PBS	A4187	Sigma-Aldrich, Steinheim

3.2 Methods

3.2.1 Microbiological Methods

3.2.1.1 Chemical Competent Cells

Chemical competent cells were prepared from overnight starting cultures by expanding the cells to an OD_{600} of 0.1, in the desired volume needed. After incubation at 37 °C cells were harvested at an OD_{600} of 0.6-0.8 by splitting them into 50 mL aliquots followed by centrifugation (4 °C, 15 min, 4,147 × *g*). All following steps were performed on ice. Each pellet was washed with 10 mL $CaCl_2$ solution (Table 3.6) and centrifuged as before. Afterwards, cells were washed again with 10 mL $CaCl_2$ solution but incubated at 4 °C for 1 h before subsequent final centrifugation. Pellets were resuspended in 2 mL $CaCl_2$ solution and frozen in liquid nitrogen as 200 µL aliquots or used directly for transformations.

3.2.1.2 Electro Competent Cells

For the production of electro competent cells a main culture was grown, split and centrifuged as described in chapter 3.2.1.1. Remaining salts were removed by washing twice with 10 mL 15% glycerol/H_2O (v/v) at 4 °C. After the final centrifugation pellets were resuspended in reflow (remaining supernatant), pooled and directly used for transformations[126].

3.2.1.3 Transformation of Chemical Competent Cells

Chemically competent *E. coli* cells were transformed with plasmids using the heat shock method. 50 µL of chemical competent cells (Ch. 3.2.1.1) were mixed with 150-500 ng DNA and incubated on ice for 10 min. After heat shocking for 2 min at 42 °C, the cells were put back on ice for 3 min. For the recovery, 1 mL LB medium (Table 3.2) was added, followed by incubation (37 °C, 750 rpm) for 1 h. Transformants were plated on agar media containing the appropriate antibiotics (Table 3.1) or used for the inoculation of an overnight culture.

The heat shock method was suitable for the transformation with up to two plasmids at once. If cells needed to be transformed with more plasmids, a two-step variant via the preparation of electro competent cells (Ch. 3.2.1.2) was used.

3.2.1.4 Transformation of Electro Competent Cells

Electro competent cells were transformed with up to 1 µg of plasmid DNA, typically prepared by ethanol precipitation (Ch. 3.2.3.7). A maximum of 10 µL of DNA was added to 100 µL electro competent cells (Ch. 3.2.1.2) and then transferred into precooled electroporation cuvettes. Electro shock was performed at 2.5 kV. For the recovery, 1 mL 2YT medium (Table 3.2) was added and the suspension transferred to an Eppendorf tube, followed by incubation (37 °C, 600 rpm) for 1.5 h. Transformants were plated on agar plates containing appropriate antibiotics (Table 3.1) or used for the inoculation of an overnight culture. In some cases a dilution series was plated to analyze the diversity/efficiency of the transformation[126].

3.2.2 Protein Biochemical Methods

3.2.2.1 Recombinant Protein Expression

Recombinant protein expressions were prepared from cells transformed with the appropriate plasmids. Depending on the type of induction of the protein expression, *E. coli* DH10B or BL21 (DE3) (for T7 driven promoters, Table 3.5) were used.

Overnight precultures were used to inoculate main cultures (20 mL to several liters LB or 2YT medium containing the required antibiotics, Table 3.1) to an OD_{600} of 0.1. Cells were incubated at 37 °C, 200 rpm, to an OD_{600} of 0.8-1.0. Protein expression was induced by supplementing the media with either, 0.2% arabinose (w/v) or 1 mM IPTG (final concentrations). In some cases no addition of an inducer was necessary since the promoter was constitutively active. For the incorporation of UAAs into proteins of interest, the media was supplemented with the appropriate probe by first dissolving (see Table 3.3) and then added to the main culture. This was performed at inoculation when constitutive promoters were being used or in combination with the inducers for inducible promoters.

Cells were typically harvested at 4-6 h after induction by centrifugation, washed with PBS and stored at -20 °C or directly used for cell lysis (Ch. 3.2.2.2).

3.2.2.2 Cell Lysis

Several methods to lyse bacterial cell pellets were used. For whole cell lysate analysis, small pellets from up to 2 mL medium were boiled in 1× SDS sample buffer (Table 3.6) for 10 min at 95 °C. Samples were subsequently centrifuged (16,100 × *g*, RT, 5 min) and directly used for SDS PAGE (Ch. 3.2.2.3). For protein isolation, larger cell pellets (from 50 mL to 500 mL medium) were resuspended in 20-50 mL lysis buffer, which was generally supplemented with protease inhibitors (1× PIC and 1mM PMSF in DMSO; Table 3.6), 1 mg/mL lysozyme, 0.5 mg/mL DNase I and 1 mM DTT (see Ch. 3.2.2.6 to 3.2.2.9 for details). Cells were then incubated on ice for 20-30 min. Disruption of the cells was carried out on ice by either ultrasonic sound, using a sonifier, or by pressure (80 psi) using a pneumatic cell disintegrator (fluidizer). The lysate was clarified by centrifugation (20,000 × *g*, 4 °C, 30 min). Afterwards, the supernatant could be employed for protein purification.

3.2.2.3 Discontinuous Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Table 3.9: Composition of polyacrylamide gels for SDS PAGE

Resolving Gel		**Stacking Gel**
12%	**15%**	**4%**
375 mM TRIS-HCl, pH 8.8	375 mM TRIS-HCl, pH 8.8	125 mM TRIS-HCl, pH 6.8
0.1% SDS (w/v)	0.1% SDS (w/v)	0.1% SDS (w/v)
12% acrylamide/	15% acrylamide/	4% acrylamide/
0.32% bisacrylamide	0.4% bisacrylamide	0.11% bisacrylamide
0.1% APS (w/v)	0.1% APS (w/v)	0.05% APS (w/v)
0.04% TEMED (v/v)	0.04% TEMED (v/v)	0.1% TEMED (v/v)

Discontinuous sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) as described by Laemmli[127] was performed to analyze the size and purity of proteins. Thereby, the strong anionic detergent SDS causes the denaturation of proteins and confers a negative charge to them, simultaneously. The discontinuity between stacking and resolving gel relies on different pore sizes and pH values, as well. The pH gradient is responsible for the stacking of the proteins at the border to the resolving gel. Whereas the stacking gels always have the same concentration of acrylamide, those of the resolving gels were varied depending on the expected protein size. During this study 12 and 15% gels were used. Protein ladders (Ch. 3.1.4) helped to estimate the molecular weights of the separated proteins. Before loading a gel, protein solutions were mixed with 4× sample buffer heated to 95 °C for 10 min followed by a short centrifugation. Small cell pellets were prepared as described in Ch. 3.2.2.2. Electrophoresis was performed at 200 V for 60-120 min in 1× SDS running buffer until the bromphenol blue dye traveled the length of the gel.

After electrophoresis was performed, separated proteins were either visualized by Coomassie Brilliant Blue staining (with Instant Blue, according to manufacturer's manual) or transferred to nitrocellulose, or PVDF, membrane by western blotting (Ch. 3.2.2.4).

3.2.2.4 Western Blot

Immunoblotting was performed to verify the expression of proteins by the direct transfer of proteins from SDS-PAGE (Ch. 3.2.2.3) onto membrane, followed by incubation with specific antibodies against the protein itself, or a protein tag, e.g., His_6-tag.

Semidry blots of SDS gels were performed with nitrocellulose membranes, sandwiched between three layers of Whatman filter papers on each side, which were soaked in 1× WB transfer buffer. This stack was fixed in a semidry blotter and the transfer was carried out with an electric currency of 250 mA constant for 45 min, RT.

Wet blots were performed with PVDF membrane that was first activated with MeOH and then washed with water followed by soaking in 1× WB transfer buffer. The membrane, SDS gel and Whatman filter papers were assembled according to the instruction manual of the blotter. The transfer was applied at 100 V constant for 30-90 min, at 4 °C.

The success of the transfer was determined by Coomassie staining of the gel and PonceauS staining of the membrane. PVDF membrane was first washed with water and MeOH before staining with PonceauS. PonceauS stain was removed from the membranes by washing in water, followed by incubation in TBS or PBS (determined by primary antibody buffer composition). Blocking of the membrane was performed for at least 10 min in buffer conditions similar to the primary antibody of interest (minus the antibody).

Primary antibodies (Table 3.8) were incubated on the membrane for 2 h with shaking, at RT or overnight at 4 °C. The membranes were then washed three times, for 10 min, in the appropriate buffer (TBS or PBS) plus 0.1% Tween20 (v/v). Secondary antibody was applied and allowed to incubate with shaking, for 1 h, at RT. The membrane was then washed as previously described. Remaining detergent was removed by additional washing with water, followed by PBS or TBS.

Chemiluminescence detection of protein was performed dependent upon the enzyme conjugated to the secondary antibody (Table 3.8). The ECL substrates used were, CDP-Star for AP-conjugates and Amersham ECL WB reagents for HRP-conjugates. The substrates were incubated on the membranes for 5 min prior to detection. In an X-ray cassette, emitted light was captured on ECL films for several seconds, to minutes, until the desired band intensities were achieved. The films were developed in an automatic X-ray film processor.

3.2.2.5 Measuring Protein Concentration

Protein concentrations were measured using the Bradford assay developed by Marion Bradford[128]. The reactive reagent is based on Coomassie Brilliant Blue G-250, which turns from its unbound red form to blue upon protein binding. This causes a shift of the absorption spectrum maximum of the dye to 595 nm in a proportional manner. The linear range of the assay depends on the protein and fluctuates between 0 and 2 mg/mL. Therefore, protein samples often needed to be diluted.

In general, 10 µL of protein sample was mixed with 500 µL Bradford reagent, incubated at RT for 5 min and then the absorbance was measured at 595 nm. Water or buffer without protein served as a reference. The amount of protein was estimated by means of a standard curve made with dissolved BSA.

3.2.2.6 Small Scale GST-MBP Expression And Purification

Overnight cultures of *E. coli* DH10B containing the appropriate plasmid combinations (see Figure 4.37) were used to inoculate 50 mL main cultures to OD_{600} of 0.1. The UAA BocK was supplemented to the main cultures at OD_{600} of 0.6 to 0.8 to a final concentration of 1 mM and the *lac* promoter dependent expression of PylS was induced by a final concentration of 1 mM IPTG. Cells with three plasmids were pelleted (4 °C, 15 min, 4,147 × *g*) 280 min after UAA addition and with four plasmids after 330 min. Pellets were washed with 1× PBS and cell counts were normalized by OD_{600}, followed by centrifugation as before. GST-MBP fusion proteins were purified as described by Neumann[78]. Cells were resuspended and lysed in

1 mL lysis buffer (1× PBS supplemented with 1× BugBuster (Novagen), 1× PIC, 1 mM PMSF, lysozyme, 0.5 mg/mL DNaseI) for 2 h on ice. Undigested DNA was sheared by ultrasonic sound using a sonifier and the lysate was subsequently clarified by centrifugation (4 °C, 30 min, 16,100 × *g*). Batch affinity purification of GST containing proteins was performed using 50 µL of glutathione sepharose beads with shaking (1 h, 4 °C). Beads were washed 3 times with 1 mL 1× PBS, before elution by heating for 10 min at 80 °C in 60 µL 1× SDS sample buffer (named Elution). For the investigation of the solubility of GST-MBP fusion proteins, cell debris from clarifying the lysate (named Pellet) were resuspended in 60 µL 1× SDS sample buffer. 15 µL of the unbound proteins in the lysate (named Flowthrough) were mixed with 5 µL 4× SDS sample buffer. All samples were analyzed on a 15% SDS gel (Ch. 3.2.2.3).

3.2.2.7 Small Scale Histone H3 Expression And Purification

Overnight cultures of *E. coli* BL21 containing the appropriate plasmid combinations were used to inoculate 500 mL main cultures to OD_{600} of 0.1. The UAAs BocK, BCNK and NorK were supplemented to the main cultures at OD_{600} of 0.6 to 0.8 to a final concentration of 1 mM and histone H3 expression was induced by a final concentration of 1 mM IPTG 30 min after UAA-addition. Cells were pelleted (4 °C, 30 min, 4,000 × *g*) 4 to 4:30 h after induction and pellets were washed with 1× PBS prior to storage at -20 °C. Frozen pellets were resuspended in 30 mL lysis buffer (1× PBS supplemented with 1× PIC, 1 mM PMSF, 1 mM DTT, lysozyme) and incubated at 37 °C for 30 min with shaking. DNA was sheared by ultrasonic sound using a sonifier (2 min, output: 6, cycle constant, 4 °C) and the lysate was subsequently centrifuged (4 °C, 30 min, 18,000 × *g*). The supernatant was discarded and the inclusion bodies containing pellet was resuspended in 1× PBS supplemented with 1% Triton X-100 and 1 mM DTT. After centrifugation (4 °C, 30 min, 18,000 × *g*), the pellet was washed (1× PBS with 1 mM DTT) and stored at -20 °C after centrifugation as before. Frozen pellets were resuspended and macerated in 1 mL DMSO for 30 min at 37 °C with shaking. Afterwards, proteins were extracted from inclusion bodies with the aid of 25 mL guanidinium solution (6 M guanidinium chloride, 20 mM TRIS (pH 8.0), 2 mM DTT) for 1 h at 37 °C with shaking. After centrifugation as above, the supernatant was transferred into a fresh Falcon tube and His_6-tagged H3 histones were bound in batch to 500 µL of Ni^{2+}-beads with shaking (1 h, 4 °C). Beads were washed with 100 mL wash-buffer (8 M urea, 100 mM NaH_2PO_4, 1 mM DTT, pH 6.2) before proteins were eluted with elution-buffer (7 M urea, 20 mM NaOAc, 200 mM NaCl, 1 mM DTT, pH 4.5). All fractions from elution were analyzed on 15% SDS gels (Ch. 3.2.2.3). Fractions containing the desired histones were pooled and dialyzed overnight against water with 5 mM β-mercaptoethanol (BME; 2 L) followed by two dialyses for 1 to 2 h each. Dialyzed samples were frozen in liquid nitrogen as 500 µL aliquots and stored at -80 °C.

3.2.2.8 Large Scale Nanobody NbSyn2 Expression And Purification

Overnight cultures of *E. coli* BL21 transformed with either pCLA166 + pCLA177 or pCLA174 + pCLA177 were used to inoculate a 500 mL main culture to OD_{600} of 0.1 each. Cultures were incubated at 37 °C for 3 h and subsequently supplemented with the UAA NorK (2 mM final concentration). Nanobody expression was induced by a final concentration of 1 mM IPTG and temperature was shifted to 28 °C for overnight expression (∼16 h). Samples for western blot were taken before and after induction, normalized to OD_{600} of 1.5. Cells were pelleted (4 °C, 30 min, 4,000 × *g*) and stored at -20 °C. The remaining medium of both cultures was pooled and filtered using a vacuum filter system with 0.45 µm pore size filter papers. The filtered medium was supplemented with 5 mM imidazole, 1.5 mM PMSF, 1× PIC and adjusted to pH 7.5. Ni^{2+}-beads (2 mL) were rinsed with washing buffer (50 mM NaH_2PO_4, 300 mM NaCl, 5 mM imidazole, pH 7.5) two times and added to the medium. Binding of His_6-tagged nanobodies was allowed to occur overnight at 4 °C with stirring. Beads were pelleted and two times washed with 50 mL washing buffer at 4 °C for 15 min with shaking. Proteins were eluted six times with 1.5 mL elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole, pH 7.5) at 4 °C for 10 min with shaking. All fractions from elution were analyzed on a 15% SDS gel (Ch. 3.2.2.3). Fractions containing the desired nanobodies were pooled and concentrated to 500 µL using centrifugal filter units.

3.2.2.9 Expression And Purification of Isotopically Labeled sfGFP

An overnight culture of *E. coli* BL21 transformed with pCLA96 and pCLA105, grown in LB medium, was used to inoculate a main culture of 250 mL M9 minimal medium supplemented with ^{15}N labeled $^{15}NH_4Cl$ to OD_{600} of 0.2. Cells were incubated at 37 °C for 7.5 h and pelleted (4 °C, 30 min, 4,000 × *g*) at OD_{600} of ∼1.0. Samples for western blot were taken from overnight and main culture, normalized to OD_{600} of 1.5. Pelleted cells were washed with 1× PBS prior to storage at -20 °C. Frozen cells were resuspended and lysed in 15 mL lysis buffer (Ni^{2+}-wash buffer (20 mM TRIS-HCl, 250 mM NaCl, 10 mM imidazole, pH 8.0) supplemented with 1× PIC, 1 mM PMSF, lysozyme, 0.5 mg/mL DNaseI) for 1 h on ice. Undigested DNA was sheared by ultrasonic sound using a sonifier and the lysate was subsequently clarified by centrifugation (4 °C, 30 min, 16,100 × *g*). His_6-tagged proteins from the lysate were bound in batch to 500 µL of Ni^{2+}-beads with agitation (1 h, 4 °C). Beads were washed with 50 mL cold Ni^{2+}-wash buffer, before elution with four times 1 mL Ni^{2+}-elution buffer (20 mM TRIS-HCl, 250 mM NaCl, 200 mM imidazole, pH 8.0). All fractions from elution were analyzed on a 15% SDS gel (Ch. 3.2.2.3). Fractions containing the desired sfGFP were pooled, concentrated to 400 µL using centrifugal filter units and simultaneously dialyzed against M9 minimal medium without glucose.

3.2.2.10 Protein Expression For In-Cell NMR Measurements

Overnight cultures of *E. coli* BL21 containing the appropriate plasmid combinations, grown in LB medium, were used to inoculate a main culture of 250 mL M9 minimal medium with unlabeled NH_4Cl to OD_{600} of 0.2. Cells were incubated at 37 °C until they reached OD_{600} of 0.6 to 0.8, pelleted (RT, 20 min, 4,000 × *g*) and resuspended in M9 medium supplemented with ^{15}N labeled $^{15}NH_4Cl$ and 150 µg/mL spectinomycin. After additional 4 h of incubation at 37 °C cells were pelleted again (RT, 20 min, 4,000 × *g*) and resuspended in 1 mL unlabeled M9 medium for NMR analysis (performed by AG Zweckstetter).

3.2.2.11 Labeling of Proteins With Fluorophores

In general, the ratio of protein to dye should be 1:10 (w/w) that means 100 µg of protein would be labeled with 10 µg dye. All dyes (Table 3.7) were prepared as stock solutions with 10 mg/mL in DMSO. Here they were used in a ratio of 1:100 (v/v). Labeling reactions were performed at RT for 1.5-2 h or at 4 °C overnight.

For the consecutive double labeling of histone H3, 100 µL of purified histones with BocK, BCNK or NorK incorporated (Ch. 3.2.2.7) were first labeled with 1 µL Abb. Star635 at RT for 1.5 h. An aliquot of 20 µL was kept as “tetrazine only” samples and the rest was dialyzed two times for 30 min against 200 mL water supplemented with 1 mM TCEP using Slide-A-Lyzer® Mini Dialysis Units. Additionally, 100 µL unlabeled histones were dialyzed for “maleimide only” samples. Next, dialyzed unlabeled and single labeled samples were mixed with 1 µL Atto565 at 4 °C overnight. Unbound dyes were removed from all samples by SDS-

PAGE and the gel was scanned with a Typhoon imager to detect labeled proteins prior to Coomassie staining.

The maleimide conjugated dye, Atto565, was excited by the green laser (532 nm) and emitted light passed through a 580 nm band-pass 30 nm (580 BP 30) filter. For all tetrazine conjugated fluorophores the red laser (633 nm) was used in combination with a 670 nm band-pass 30 nm (670 BP 30) filter. Sensitivity was set to normal and the voltage for the photomultiplier tube (PMT) had to be adapted for each blot individually. The resolution for the most scans was 200 µm per pixel, in some cases it was increased to 100 µm per pixel.

3.2.2.12 Fluorescence Measurement With A Plate Reader

A FLUOstar Omega (BMG Labtech) plate reader was used to measure fluorescence from sfGFP in intact cells. Samples from *E. coli* DH10B containing the appropriate plasmid combinations to express sfGFP, from genes with WT sequence or harboring an amber and/or a frameshift codon, were taken with intent to being normalized to OD_{600} of 0.5 by pelleting (3 min, 16,100 × *g*) and resuspending in 1 mL 1× PBS. 200 µL of this cell suspension was transferred into one well of a 96-well black microplate. 200 µL 1× PBS was used as a reference. The fluorescent signals from GFP were measured using the self-written program “GFP_ORBS-REPORTER” (**Plate mode settings:** No. of flashes per well: 10; Scan mode: Orbital averaging; Scan diameter [mm]: 3; **Optic Settings:** Excitation 485 nm; Emission: 520 nm; Gain: variable; **General settings:** Top optic used; Positioning delay [s]: 0.2). Kinetic non-endpoint measurments over time were performed using the program “GFP-NMR” that only differs in the number of measuring points.

3.2.2.13 Fluorescence-Activated Cell Sorting

Fluorescence-activated cell sorting (FACS), established by the workgroup of Herzenberg[129], was done at the MRC Laboratory of Molecular Biology, Cambridge, using the cell sorter iCyt Synergy (Sony Biotechnology). It was used for sorting living *E. coli* based on the fluorescence intensity of sfGFP within the cells. Therefore, 10 mL aliquots normalized to OD_{600} of 1.0 were prepared from 100 mL overnight cultures of *E. coli* DH10B containing the appropriate plasmid combinations to express sfGFP, from genes with WT sequence or harboring an amber and a frameshift codon (see Figure 4.34). Cells were pelleted (4 °C, 15 min, 4,147 × *g*) and three times washed with two-fold filtered (syringe filters (0.2 µm))

1× PBS. Cell suspensions were diluted to 10^5 cells prior to use in the FACS machine, which was finally calibrated to test 30,000 cells per second. Sorted positive clones were temporarily stored in 200 µL SOC medium. For recovery an additional 1 mL SOC medium was added and cells were incubated at 37 °C for 1 h. Recovered cells were transferred into 500 mL LB medium containing the appropriate antibiotics, in order to prepare glycerol stocks (500 µL cell culture with 500 µL glycerol (50%)) and to isolate DNA (Ch. 3.2.3.1).

3.2.3 Nucleic Acid Biochemical Methods

3.2.3.1 Preparation of Plasmid DNA

Plasmid DNA was isolated and purified with the help of kit systems (Ch. 3.1.11) according to manufacturer's manual. In general, bacteria were transformed with the desired plasmid (Ch. 3.2.1.3) and used to inoculate an overnight culture (5 mL LB, 37 °C) containing a selective antibiotic. Up to 4 mL of this culture were harvested by centrifugation (16,100 × *g*, RT, 3 min) and disrupted by alkaline lysis using the kit buffers. The purification of the DNA was performed over silica columns.

3.2.3.2 Restriction Enzyme Digestion

Digestion of DNA was performed by restriction enzymes (Table 3.4) following manufacturer's protocols. These enzymes are endonucleases which operate by recognizing a specific sequence of nucleotides, varying between four and eight base pairs in length, and often palindromic, followed by producing double strand breaks in the DNA[130]. Thereby, some enzymes create overhanging ends (sticky) and some create ends without overhangs (blunt). Categorized into four groups (TypeI-IV) based on different traits they differ amongst others in the position of the cleavage site relative to the recognition sequence.

TypeII endonucleases, which were used in this thesis, cut their substrate near or inside of their recognition sequence. This makes them easier to handle, compared to randomly cutting enzymes, and thereby the most commonly available and used restriction enzymes. Due to the discovery of new enzymes that belonged to this class, but did not fulfill all classical criteria, a number of subfamilies were founded. One of these enzymes, DpnI, is able to recognize and cleave methylated and hemimethylated DNA[131]. Therefore, it was used to remove parental template DNA from PCR reactions (Ch. 3.2.3.8, 3.2.3.10 &

3.2.3.11). In most cases DpnI was supplemented as a fortieth of the PCR volume. If the PCR was purified before (Ch. 3.2.3.9) 10× reaction buffer was added (as a tenth of the volume) as well, otherwise the buffer of the PCR was also used for the digest that usually was performed for 1.5 h at 37 °C.

The majority of the other endonucleases were either used to clone specific DNA fragments into a certain vector backbone (preparative digest) or to conduct test digests (Table 3.10) of prepared plasmid DNA (Ch. 3.2.3.1) from re-transformations or cloning procedures. Digests, especially double digests, were carried out in buffers and relation of enzymes recommended by the manufacturer.

Table 3.10: Pipetting scheme for standard test and preparative digests

Test digest		**Preparative digest**	
1 µL	buffer, 10×	2.5 µL	buffer, 10×
2 µL	DNA	8 µL	DNA (~1 µg)
1 µL	enzyme(s)	2.5 µL	enzyme(s)
6 µL	H_2O	12 µL	H_2O
2 h, 37 °C		3 h, 37 °C	

Digests for cloning (preparative ones) were treated with alkaline phosphatases (0.5-1 µL, 37 °C, 0.5-1 h, Table 3.4) afterwards, to remove 5'- (and 3'- for SAP) phosphate groups from digested vectors to prevent self-ligation.

Phosphatases and restriction enzymes were inactivated by heating (65-80 °C, 15 min) and/or separation on an agarose gel (Ch. 3.2.3.3), which was also used to separate DNA fragments from digest itself.

3.2.3.3 Agarose Gel Electrophoresis

Agarose gel electrophoresis was performed to analyze size and abundance of DNA (fragments) from restriction digests (Ch. 3.2.3.2) and PCRs (Ch. 3.2.3.8, 3.2.3.10 & 3.2.3.11) as well as to purify DNA as necessary.

Depending on the expected size of the DNA, different percentages of agarose gels were prepared (Table 3.11). The appropriate amount of agarose was melted and completely dissolved in 0.5× TBE buffer (Table 3.6) in a microwave, casted into an electrophoresis chamber and supplemented with “gel red” in a concentration of 1:50,000 (v/v) after cooling to 50-60 °C.

DNA samples were mixed with 10× DNA loading buffer (Table 3.6) and loaded onto the gel together with a DNA ladder (Ch. 3.1.4) to estimate the size of the DNA. Electrophoresis was performed at 200 V, 80-100 mA at RT with 0.5× TBE running buffer until the bromphenol blue dye migrated the length of the gel (1-2 h).

Separated DNA bands were visualized by UV light due to the intercalated “gel red” using a gel documentation machine or a UV table. If DNA was needed for further steps, like cloning (Ch. 3.2.3.4), wavelength of the UV light was set to 365 nm instead of 254 nm.

Table 3.11: Percentages of agarose gels depending on the size of the DNA (fragments)

DNA (fragment) size [bp]	[%] agarose in gel (w/v)
< 500	1.5
200 – 3,500	1.0
700 – ≥ 10,000	0.7

3.2.3.4 Extraction of DNA From Agarose Gels

DNA required for cloning was preparatively digested (Ch. 3.2.3.2) and subsequently separated via agarose gel electrophoresis (Ch. 3.2.3.3). Visualized on a UV table (365 nm), those bands of correct size were sliced out of the gel with a scalpel. Afterwards, the DNA was extracted using a gel extraction kit (Ch. 3.1.11) according to manufacturer’s manual. The purified DNA was eluted from the silica columns with max. 30 µL of elution buffer and stored at -20 °C or directly used for ligations (Ch. 3.2.3.6)

3.2.3.5 Separation of Multiple Plasmids

Two methods were used for the separation of multiple plasmids isolated from cells (Ch. 3.2.3.1) transformed with more than one plasmid. First, plasmids were separated on agarose gels (Ch. 3.2.3.3) and DNA/plasmid of expected size extracted from gel (Ch. 3.2.3.4), subsequently. Second, unwanted plasmids were destroyed by restriction enzyme digestion (Ch. 3.2.3.2) prior to transformation of *E. coli* cells. Plasmids of interest were isolated from transformed cells, afterwards.

3.2.3.6 Ligation of Two Double-Stranded DNA Fragments

A double-stranded DNA fragment was typically digested with two different restriction enzymes (Ch. 3.2.3.2) and was ligated with another double-stranded DNA fragment, digested with the same enzymes, using the T4 DNA ligase. Therefore, a 3-5 fold molar excess of the smaller fragment, e.g., a gene for a protein, called the insert is mixed with the bigger fragment, commonly a vector backbone (Table 3.12). For a negative control water replaced the insert DNA. The ligation was carried out for 2 h, at RT, or more frequently at 16 °C overnight. The product of the ligation reaction (5-10 µL) was directly used for transformation (Ch. 3.2.1.3) and the success of the ligation was verified by test digests (Ch. 3.2.3.2) and/or sequencing (Ch. 3.2.3.13).

Table 3.12: Pipetting scheme for a standard ligation reaction mix

	Volume [µL]
T4 Ligase buffer, 10×	1
Vector DNA (big fragment)	2
Insert DNA (small fragment)/H_2O	6.5
T4 DNA ligase	0.5

3.2.3.7 Ethanol Precipitation

Ethanol precipitations were performed to remove salts prior to transformation, if cells needed to be transformed with ligation reactions via electroporation due to required higher efficiencies, and/or to reduce the volume of large scale ligations. For this purpose, the ligation product was mixed in a ratio of 10:1 (v/v) with 3 M sodium acetate solution (pH 5.2) and subsequently supplemented with absolute EtOH equal to double the volume of the whole reaction mixture. After an incubation period of 30 min on ice the DNA was pelleted by centrifugation (16,100 × *g*, 4 °C, 15 min). The pellet was washed two times with 70% EtOH equal to twice the starting volume of the ligation product and centrifuged each time as before. The supernatant was discarded and the DNA pellet was air dried for 5 min at RT. After resuspending the DNA in a small amount of ddH_2O or TE buffer (Table 3.6) an incubation for 10 min at 60 °C followed to dissolve the pellet completely. DNA could then be stored at -20 °C or used for electroporation (Ch. 3.2.1.4).

3.2.3.8 Polymerase Chain Reaction

The polymerase chain reaction (PCR) is a technique used to specifically and exponentially amplify a piece of DNA. Since this method, developed by Kary Mullis[132], relies on thermal cycling with periods of heating and cooling, necessary for DNA melting and enzymatic replication, thermostable enzymes like *Taq* or Phusion DNA polymerases (Table 3.4) are essential. In general, a PCR consists of 6 different steps in which step two to four are repeated in 20-40 cycles.

1. **Initial denaturation:** In this step the reaction is heated up to 94-96 °C to ensure that the template DNA is completely melted which is important for step three, the annealing of the primers. Additionally, it activates DNA polymerases that need heat activation (hot-start PCR)[133].
2. **Cyclic denaturation:** The purpose of this step is the same as for the first one, melting the DNA (template and new synthesized strands) by heating to 94-96 °C to obtain single-stranded DNA but in a cyclic manner.
3. **Cyclic annealing:** The temperature of the reaction is lowered to 50-60 °C, usually 5 °C lower than the melting temperature of the primers, single-stranded oligonucleotides that serve as a starting point for the polymerase. The low temperature allows the primers, which are complementary to the template DNA, flanking the sequence that should be amplified, to hybridize.
4. **Cyclic elongation:** In this step the polymerase synthesizes a new complementary DNA strand starting at a primer by adding dNTPs. Therefore, the temperature is changed to the appropriate optimum for the used polymerase (68-72 °C). The elongation time depends on the length of DNA to be amplified and on the speed of the particular polymerase.
5. **Final elongation:** All remaining single-stranded DNA fragments are fully extended during this step at the same temperature used in step 4.
6. **Stop/cooling:** The reaction mix is cooled down to “stop” enzymatic reactions and for short-term storage.

The standard course for a PCR is shown in Table 3.13 and the composition for one PCR reaction mix is shown in Table 3.14.

Table 3.13: Standard course of a PCR

Step	Temperature [°C]	Duration [s]	
1. Initial denaturation	95	90	
2. Cyclic denaturation	95	30	
3. Cyclic annealing	56	30	30 cycles
4. Cyclic elongation	72	60/1kbp	
5. Final elongation	72	Step 4 + 60	
6. Stop/cooling	4	∞	

Table 3.14: Pipetting scheme for the composition of one PCR reaction mix

	Volume [µL]
Taq buffer with KCl, 10×	5
dNTPs (10 mM each), 50×	1
Primer mix (for + rev), 10 µM	4
$MgCl_2$, 25 mM	4
Template DNA	1
Taq DNA polymerase	1
H_2O	34

Before further usage, PCR products were purified using EtOH precipitation (Ch. 3.2.3.7) or kit systems (Ch. 3.2.3.9).

3.2.3.9 Purification of PCR Products

For the removal of enzymes, unused primers, dNTPs and salts, PCR reactions were purified using EtOH precipitation (Ch. 3.2.3.7) or kit systems (Ch. 3.1.11) according to manufacturer's manual. The cleaned up DNA was eluted from the silica columns with max. 30 µL of elution buffer and stored at -20 °C or directly used, e.g., restriction digests (Ch. 3.2.3.2).

3.2.3.10 QuikChange PCR

A QuikChange (QC) is a site-directed mutagenesis (developed by Agilent, Böblingen former Stratagene) that allows to easily carry out vector modifications (Figure 3.1).

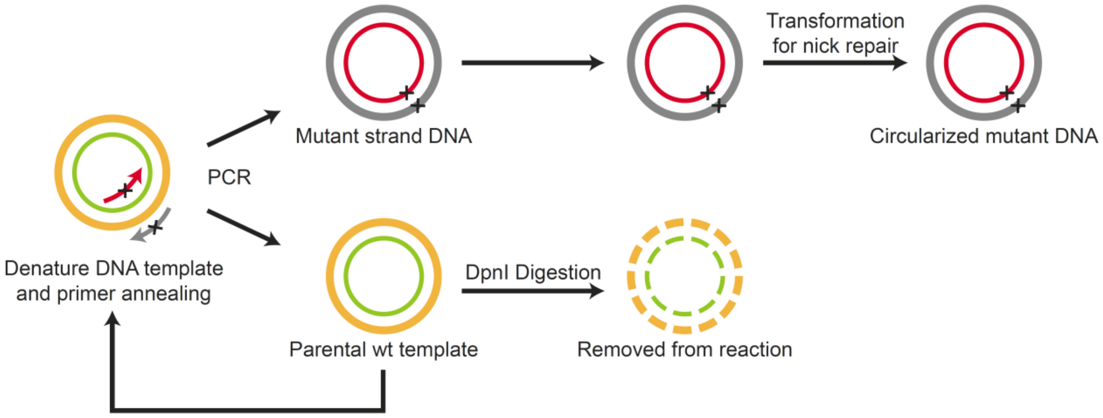

Mutant strand DNA
Transformation for nick repair
Circularized mutant DNA
PCR
Denature DNA template and primer annealing
Parental wt template
DpnI Digestion
Removed from reaction

Figure 3.1: Schematic view of QuikChange cycles.
Figure is adapted from QuikChange II XL site-directed mutagenesis protocol (Agilent). Mutant strand DNA (grey and red rings) is synthesized by performing thermal cycling (PCR) to denature the DNA template (orange and green rings) and to anneal mutagenic primers (grey and red arrows) containing the desired mutations (black crosses). Primers are extended with *Pfu* Turbo DNA polymerase. Parental methylated and hemimethylated template DNA is digested by DpnI. Finally, competent cells are transformed with the mutated molecule for nick repair.

Since these PCR reactions amplify an entire vector a high fidelity proofreading polymerase, *Pfu* Turbo, was required. Primers were designed so that forward and reverse primers both included the desired mutations in a complementary sequence of at least 25 bp. The 3’-ends of each primer were complementary to a minimum of 10-15 bp of the vector backbone to allow annealing to the correct positions. The composition for one QC-PCR reaction mix is shown in Table 3.15.

Table 3.15: Pipetting scheme for the composition of one QuikChange PCR reaction mix

	Volume [µL]
Pfu buffer, 10×	2
dNTPs (10 mM each), 50×	0.4
Pfu Turbo DNA polymerase	0.4
Template DNA	0.4
Primer mix (for + rev), 10 µM	0.8
H_2O	16

The standard course for a QC is shown in Table 3.16. Compared to a common PCR (Ch. 3.2.3.8) fewer cycles were used to avoid the amplification of random mutations. The DNA template was digested with DpnI (0.5 µL) subsequent to PCR (37 °C, 1-1.5 h), as depicted in Figure 3.1, followed by transformation of *E. coli* cells with 5-10 µl of the mixture (Ch. 3.2.1.3). The modified vectors were isolated from those cells (Ch. 3.2.3.1) and the success of the modification determined by sequencing (Ch. 3.2.3.13).

Table 3.16: Standard course of a QC-PCR

Step	Temperature [°C]	Duration [s]	
1. Initial denaturation	96	60	
2. Cyclic denaturation	95	20	
3. Cyclic annealing	56	30	15 cycles
4. Cyclic elongation	72	90/1kbp	
5. Final elongation	72	Step 4 + 60	
6. Stop/cooling	4	∞	

3.2.3.11 Inverse PCR

Inverse PCR is a variation of the conventional PCR (Ch. 3.2.3.8) and was originally used for the rapid *in vitro* amplification of unknown DNA sequences that flank a region of known sequence[134]. The restriction enzyme-mediated inverse PCR incorporates unique restriction enzyme sites at the 5'-ends of inverse tail-to-tail primers[135]. Similar to the QC (Ch. 3.2.3.10) this technique can be used for site-directed mutagenesis but with a higher efficiency. The increase in efficiency is due to the primer design, since only one primer needs to hold the desired mutations, and the generation of sticky-ends by unique restriction enzyme digests which improve ligation efficiency of linear mutant DNA strands. This method amplifies an entire vector like the QC and requires also a high fidelity proofreading polymerase. The Expand High Fidelity Polymerase Mix (Table 3.4) was used for this purpose. The general procedure of the restriction enzyme-mediated inverse PCR is shown in Figure 3.2.

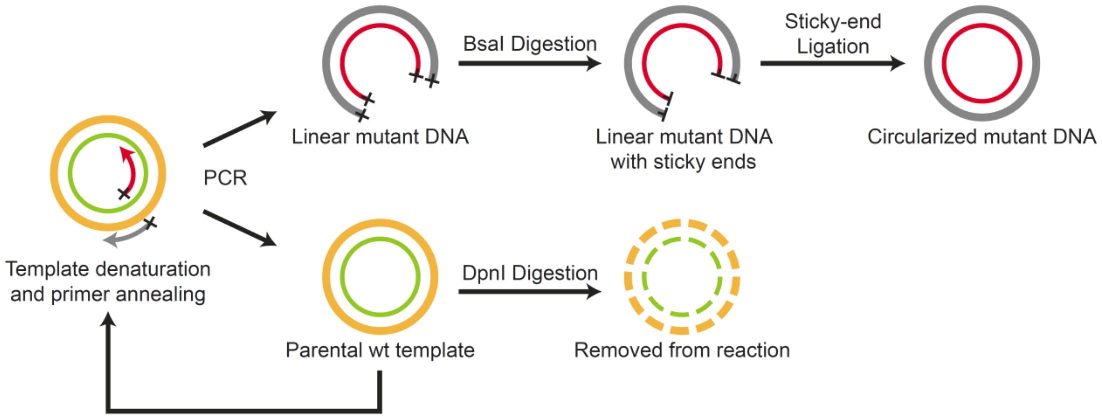

BsaI Digestion
Sticky-end Ligation
Linear mutant DNA
Linear mutant DNA with sticky ends
Circularized mutant DNA
PCR
Template denaturation and primer annealing
DpnI Digestion
Parental wt template
Removed from reaction

Figure 3.2: Schematic representation of the restriction enzyme-mediated inverse PCR[135].
Sense and antisense primers (grey and red arrows) are designed in such a way that they contain a unique restriction site, BsaI (black crosses), and anneal to the opposite ends instead of facing each other as usual. Desired mutations were mostly incorporated into the reverse primer, only. The template DNA (orange and green rings) is amplified via PCR, resulting in linear mutated DNA (grey and red rings), and subsequently eliminated by DpnI digestion. Blunt ends of mutant DNA are made sticky upon digest with BsaI, which improves the ligation efficiency and subsequent circularization of the PCR product.

The composition of one inverse PCR reaction is shown in Table 3.17 and the course in Table 3.18. For the setup of the reaction mix two mixtures were prepared as recommended by Roche's protocol, one with buffer, enzyme and half of the water, the other with the rest. Both mixtures were combined directly before PCR.

Table 3.17: Pipetting scheme for the composition of one inverse PCR reaction mix

	Volume [µL]
Expand HF Buffer (15 mM $MgCl_2$), 10×	5
dNTPs (10 mM each), 50×	1
Template DNA	1
Primer mix (for + rev), 10 µM	4
Expand High Fidelity Polymerase Mix	1
H_2O	38

PCR was followed by purification using kit systems (Ch. 3.2.3.9) and digest of parental template with DpnI from NEB as described in Ch. 3.2.3.2. Afterwards, mutant DNA was digested with BsaI (3 h, 50 °C) to achieve sticky-ends. Therefore, DNA concentration was measured (Ch. 3.2.3.12) and 1 µL of BsaI per 10 µg DNA added to the mixture. Recommended buffer was already supplemented for the DpnI digest. After cutting the recognition and cutting sites for BsaI (5'-GGTCTC-3'), incorporated at the 5'-ends of the primers, are automatically eliminated and not maintained in the final product.

Table 3.18: Standard course of a restriction enzyme-mediated inverse PCR.
A touchdown PCR is performed to avoid nonspecific binding of primers and to increase yields significantly[136,137]. During the first 10 cycles annealing temperature is decreased by 1 °C causing the primers to anneal at highest temperature which is least-permissive for nonspecific binding. Therefore, amplification of nonspecific sequences is outcompeted.

Step	Temperature [°C]	Duration [s]	
1. Initial denaturation	95	60	
2. Cyclic denaturation	95	20	10 cycles (Δ-1 °C)
3. Cyclic annealing	65	30	
4. Cyclic elongation	68	60/1,5kbp	
5. Cyclic denaturation	95	20	25 cycles
6. Cyclic annealing	55	30	
7. Cyclic elongation	68	60/1,5kbp	
8. Final elongation	68	Step 7 + 60	
9. Stop/cooling	4	∞	

Prior to ligation of the sticky-ends using T4 DNA ligase at 16 °C overnight (50 µL digested PCR, 35 µL H2O, 10 µL T4 Ligase buffer (10×), 5 µL T4 DNA ligase) the digested PCR was purified again with PCR purification kits (Ch. 3.2.3.9).

Ligations were usually precipitated with EtOH (Ch. 3.2.3.7) and dissolved DNA used for transformation via electroporation (Ch. 3.2.1.4). The inoculation of big cultures (about 500 mL) followed by the isolation of the modified plasmids allowed the amplification of the DNA for further experiments. A dilution series of these cultures were used to determine the diversity of the PCR. Therefore, 50 µL of the inoculated medium was plated directly on agar plates containing the appropriate antibiotic, representing a 10^4 dilution. Further dilutions up to 10^8 were made by diluting previous ones tenfold with medium and also plated on agar plates.

3.2.3.12 Measuring Nucleic Acid Concentration

DNA and RNA concentrations of aqueous solutions were measured by absorption in UV cuvettes using a photometer. 3 µL of the nucleic acid containing solution was diluted in 100 µL water. Detection occurred at 260 nm (A_{260}) with pure water (100 µL) as a reference. The purity of the nucleic acid was also given by the photometer, calculated by the quotient of A_{260} to A_{280}.

3.2.3.13 Sequencing of Plasmids

DNA sequencing was performed using chain-terminating inhibitors, dideoxynucleotide triphosphates (ddNTPs), developed by Sanger and colleagues[138,139].

This method works similar to a normal PCR (Ch. 3.2.3.8) and needs the same components, i.e., a (single-stranded) DNA template, a primer, a DNA polymerase and normal dNTPs, except the ddNTPs that randomly terminate the extension of the DNA strand due to the lack of the 3'-OH group. The use of ddNTPs each labeled with a different fluorescent dye emitting light of a unique wavelength allows sequencing in a single reaction. At the end of the PCR fragments of all possible lengths are statistically contained in the mixture which can be separated and analyzed by a sequencing machine subsequently.

The sequencing mix *BigDye* (*BigDye®* Terminator v1.1 Cycle Sequencing Kit, Applied Biosystems, USA) used for sequencing included the DNA polymerase, normal dNTPs and fluorescent labeled ddNTPs as well. The complete composition for a sequencing PCR is shown in Table 3.19 and the course in Table 3.20.

Table 3.19: Pipetting scheme for the composition of one sequencing PCR reaction mix

	Volume [µL]
Seq buffer	1
Seq mix. *BigDye*	1
Primer, 10 µM	0.8
Template DNA, 200-400 ng	1-7.2
H_2O	Adjust to 10 µL

Table 3.20: Standard course of a sequencing PCR

Step	**Temperature [°C]**	**Duration [s]**	
1. Initial denaturation	96	60	
2. Cyclic denaturation	96	10	25 cycles
3. Cyclic annealing	55	15	
4. Cyclic elongation	60	240	
5. Final elongation	60	240	
6. Stop/cooling	4	∞	

The DNA products were purified subsequently to the PCR. Therefore, the mixture was transferred to a 1.5 mL Eppendorf tube, supplemented with 2 mM EDTA-Na_2, 50 mM sodium acetate and 80% EtOH (final concentrations), gently mixed and incubated for 5 min at RT. After a centrifugation step (16,100 × *g*, RT, 15 min) the supernatant was discarded and the pellet washed with 70 µL of 70% EtOH, followed by a second centrifugation for 5 min. The supernatant was discarded again and the pellet was dried in a speedvac/concentrator for 2 min. The DNA pellet was dissolved in 15 µL formamide (99.5%) and given to AG Pieler, Göttingen for sequencing in a capillary sequencing machine.

3.2.3.14 Extraction/Isolation of RNA

Total RNA was isolated from bacterial cells by guanidine thiocyanate-phenol-chloroform extraction using TRIzol Reagent® (Invitrogen, Darmstadt; registered trademark of Molecular Research Center, Inc.).

TRI Reagent is the improved version of the single-step method developed by Chomczynski and Sacchi[140,141]. The guanidine thiocyanate lyses cells and inhibits RNases at the same time. The addition of chloroform (or bromochloropropane) followed by centrifugation splits the mixture into three phases, i.e., an upper aqueous phase containing the RNA, an interphase harboring the DNA and a lower organic phase including the proteins. The RNA is obtained by supplementing isopropanol to the aqueous phase, washed with EtOH and dissolved subsequently.

The isolation was performed according to manufacturer's manual except for the first step. The lysis of a recommended maximum of 10^7 bacterial cells per 1 mL of TRI Reagent did not result in a satisfying yield of RNA. Hence, the pellet of 1 mL bacterial culture with an average OD_{600} of 3 was lysed with 1 mL of TRI Reagent. The washed RNA was dissolved in 30 µL H_2O, finally.

3.2.3.15 Northern Blot

For the identification of specific tRNAs via northern blot total RNA extracts (Ch. 3.2.3.14) were separated by size using acid urea PAGE conforming to the protocol of Köhrer *et al*.[142]

Table 3.21: Composition of an acid urea polyacrylamide gel for acid urea PAGE

Ingredients	Volume [mL]	Final Concentration
H_2O	10	-
30% polyacrylamide (19:1 acrylamid:bisacrylamid)	7.2	6.5%
1 M sodium acetate, pH 5.0	3.3	0.1 M
Urea	16 g $\triangleq$ 11.5 mL	8 M
10% APS (w/v)	1.2	0.36% (w/v)
99.9% TEMED	0.025	0,075% (v/v)

Isolated RNA was mixed 1:1 (v/v) with acid urea sample buffer (Table 3.6) following extraction. Approximately 90 ng of RNA were loaded onto the gel that was subjected to a short pre-electrophoresis for 30 min in the cold room. Electrophoresis was performed at 168 V and 165 mA (~12 V/cm) for 150-210 min in the cold room with 0.1 M sodium acetate pH 5.0 as running buffer until the bromphenol blue dye reached the end of the gel.

As a size standard, low range RNA and oligonucleotide ladders (Ch. 3.1.4) were also loaded onto the gel. *In vitro* RNA transcripts of the correspondent tRNA synthesized by T7 RNA polymerase (Ch. 3.2.3.16) served as size standard and positive control for northern blots.

For the actual blot only the portion of the gel between both dyes containing small RNAs like tRNAs was used. RNAs were transferred to a nylon membrane using a semidry blotter in a similar way as described in Ch. 3.2.2.4. The transfer was carried out at 25 V and 513 mA for 80 min in the cold room with 1× NB transfer buffer (Table 3.6). Afterwards, the gel was soaked in a “gel red” bath to verify the success of the transfer using a gel documentation machine and the RNA was immobilized through covalent linkage to the membrane by heat (2 h, 75 °C in a hybridization oven) or UV light (90 s, 254 nm on a UV table).

For northern hybridization, DNA oligonucleotides were designed (Table 8.2) with a length of 17-23 bp, 5’-labeled with Cy3, Cy5 or DIG and complementary to the anticodon stem and loop domain of the desired tRNA target.

Prior to hybridization a prehybridization in the appropriate solution (30 mL, Table 3.6) was executed for 6-12 h at 42 °C in a hybridizer oven, with constant rotation. Thereafter, the temperature was increased to 68 °C and the solution exchanged with hybridization solution (30 mL, Table 3.6). After incubating the membrane for 12-24 h several washing steps were performed at RT with 6× SSC (twice for 10 min; low stringency, Table 3.6), followed by washes with 4× and 2× SSC (each for 20 min; medium and high stringency).

Northern blots with Cy3 or Cy5 labeled hybridization probes could now directly be utilized for the detection of tRNAs using a Typhoon imager set to the correct laser and filter settings. Cy3 was excited by the green laser (532 nm) and emitted light passed through a 580 nm band-pass 30 nm (580 BP 30) filter. For Cy5 the red laser (633 nm) was used in combination with a 670 nm band-pass 30 nm (670 BP 30) filter. Sensitivity was set to normal and the voltage for the photomultiplier tube (PMT) had to be adapted for each blot individually. The resolution for the most scans was 200 µm per pixel, in some cases it was increased to 100 µm per pixel.

For the detection of DIG-labeled probes further steps were necessary (all carried out at RT). First, the membrane was equilibrated with 30 mL buffer 1 (Table 3.6) for 5-10 min, followed by blocking with 20 mL buffer 2 (Table 3.6) for 30 min and an incubation with anti-DIG-AP-antibody in buffer 2 for 30-60 min (or overnight at 4 °C). Unbound antibodies were removed by washing three times for 10 min with 30 mL buffer 1. The membrane was then equilibrated with 10 mL buffer 3 (Table 3.6) for 10 min and the chemiluminescence reaction started subsequently by adding 1 mL buffer 3 containing 5 µL CDP-Star directly onto the membrane. After 5 min of membrane incubation with the substrate, in an X-ray cassette, emitted light was captured on ECL films for several seconds, to minutes, until the desired band intensities were achieved. The films were developed in an automatic X-ray film processor.

3.2.3.16 Transcription With T7 RNA Polymerase

Positive controls for northern blots (Ch. 3.2.3.15) were made via T7 RNA polymerase (T7RP) transcription. First, a PCR (Ch. 3.2.3.8) was performed to build a template that contained the recognition site for the T7 polymerase, 5’-TAATACGACTCACTATA-3’, and the gene for the appropriate tRNA. The success of the PCR was verified by electrophoresis on an agarose gel (Ch. 3.2.3.3) and the PCR product was purified (Ch. 3.2.3.9).

To destroy remaining RNases, the mixture was treated with proteinase K (1 µL per 50 µL PCR, 30 min, 50-55 °C). Inactivation of the proteinase K was carried out at 95 °C for 10 min.

The transcription was performed at 37 °C for 2 h. The composition for one reaction is shown in Table 3.22; components were received from AG Ficner, Göttingen. The mixture was set up at RT due to the spermidine in the buffer which causes negative effects in the cold[143,144].

Table 3.22: Pipetting scheme for the composition of one transcription reaction mix

	Volume [µL]
$MgCl_2$, 100 mM	4.5
rNTPs (40 mM each)	20 (4×5)
Transcription buffer, 10×	5
Template DNA	18
T7 RNA polymerase (1 mg/mL)	2.5

Transcription was stopped by the addition of 5 µL 0.5 M EDTA-Na_2 and the mixture subsequently centrifuged to remove generated pyrophosphates (16,100 × *g*, RT, 3 min). The supernatant containing the RNA oligos was purified (Ch. 3.2.3.9), and the elution mixed with acid urea sample buffer (Table 3.6) and boiled at 94 °C for 5 min to denature the RNA. Table 3.23 gives an overview of transcripts produced.

Table 3.23: Overview of RNA transcripts made by T7 RNA polymerase.
Transcripts are listed with the primers used for the PCR to produce the T7 polymerase recognition site containing template.

Transcript	Primer	Chapter
T7 PylT	C28 and C29	4.1.2
T7 SerT	C30 and C31	4.1.2
T7 TyrT	C42 and C43	4.1.4

3.2.4 Reporter Assays

3.2.4.1 Chloramphenicol Reporter Assay

The chloramphenicol (Cm) reporter assay (Figure 3.3) was the standard evaluation procedure to classify the suppression efficiency of a given system. The more resistant the cells were to Cm, the more efficient the suppression worked. A system that had insufficient suppression activity led to cell death. The Cm assay could be utilized to evolve synthetases for new UAAs, for example, or to investigate optimizations on components of the genetic code expansion machinery.

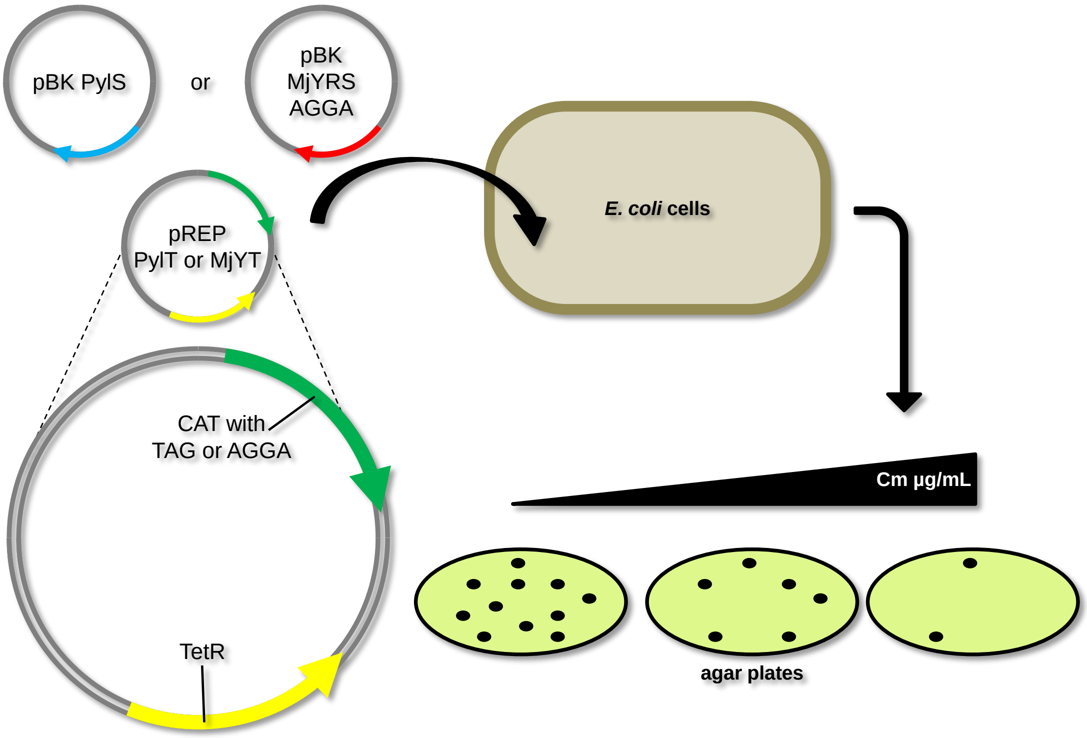


Figure 3.3: Schematic view of the chloramphenicol reporter assay.
A synthetase containing plasmid, derived from pBK PylS[95] (pCLA1) or pBK MjYRS_TAG/AGGA[45,78,83] (pCLA2/pCLA3), was combined with the corresponding pREP[78,95,145] reporter plasmid harboring the gene for PylT (pCLA4) or MjYT (pCLA5/pCLA6; Table 8.1), respectively. In addition to the gene for the tRNA, the plasmid also comprised a gene for a chloramphenicol acetyltransferase (CAT) with either a TAG or an AGGA codon. *E. coli* cells, ususally DH10B, were transformed with both plasmids (Ch. 3.2.1.3 and 3.1.9) and plated on agar plates with an increasing concentration of chloramphenicol, appropriate antibiotics for the plasmids and if required relevant UAAs. The better a plasmid combination worked, the higher the concentration of Cm the cells were able to survive. Resistance genes on plasmids are indicated by the arrows: blue = Kan; red = Amp; green = Cm; yellow = Tet.

3.2.4.2 GFP Reporter Assay

The GFP reporter assay (Figure 3.4) was another option for quantifying the suppression efficiency of a given system. In contrast to the Cm reporter (Ch. 3.2.4.1), weak systems did not cause cell death. The advantage of this reporter was due to the fluorescent trait of GFP, which could be easily and directly measured using a plate reader (Ch. 3.2.2.12). Since all GFP reporter plasmids contained a His$_6$-tagged GFP, detection via western blot (Ch. 3.2.2.4) was also possible.

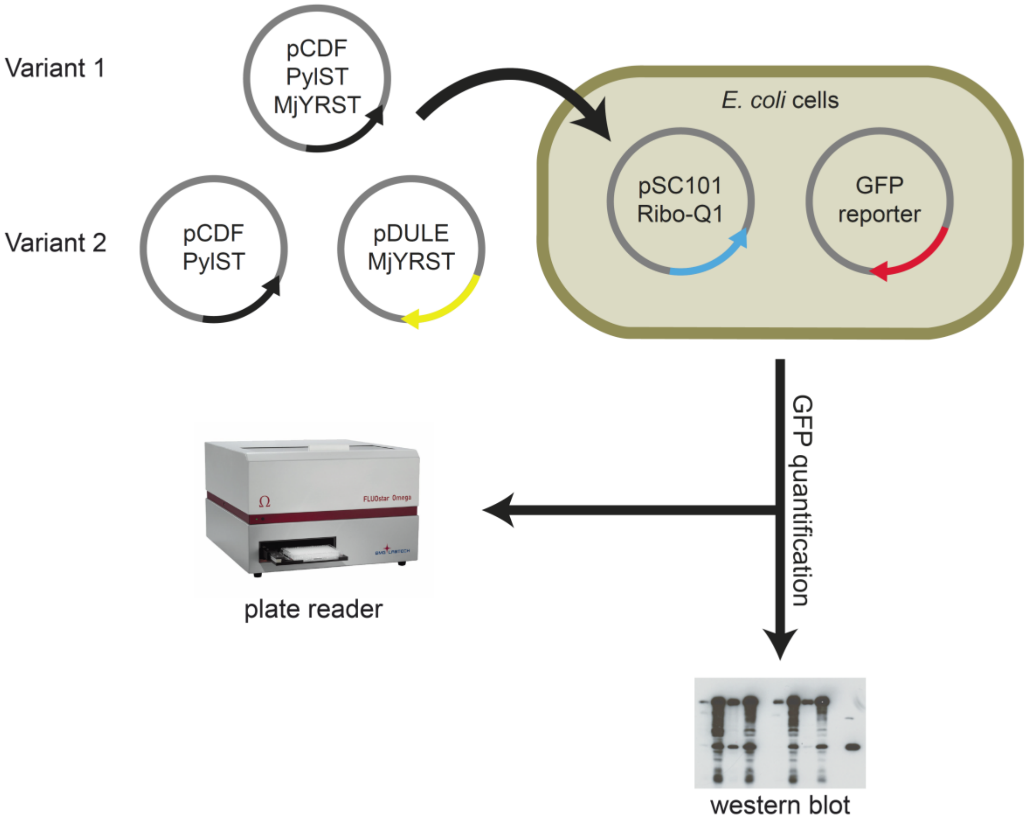


Figure 3.4: Schematic view of the GFP reporter assay.
Electro competent *E. coli* DH10B cells were prepared that already bear a plasmid for the orthogonal ribosome ribo-Q1[78] (pCLA96; Table 8.1) and one out of four GFP reporter plasmids (pCLA105 to pCLA108; Ch. 4.4.3). These cells were transformed with two different sets of plasmids, either Variant 1 or Variant 2. Variant 1 comprised a pCDF Duet-1 (Novagen; pCLA91) based vector, which holds the genes for the tRNA/aaRS pairs PylT/PylS (PylST) as well as MjYT_UCCU/MjYRS_AGGA (MjYRST) (pCLA95). Variant 2 included the same tRNA/aaRS pairs as Variant 1, but split on two plasmids - pCDF PylST[78] (pCLA97) and pDULE MjYRST (this version decodes for UCCU and is based on the CUA version[68,83]; pCLA98). The suppression efficiency of amber and/or frameshift codons was measured by the amount of GFP produced. GFP was either detected using a plate reader (Ch. 3.2.2.12) or by western blot (Ch. 3.2.2.4). The better a plasmid combination worked, the more GFP fluorescence could be detected. Resistance genes on plasmids are indicated by the arrows: black = Sm; blue = Kan; red = Amp; yellow = Tet

3.2.4.3 Barnase Reporter Assay

The barnase assay is used to delete non-orthogonal synthetases that charge their corresponding tRNAs with natural amino acids or non-orthogonal tRNAs that are aminoacylated by endogenous synthetases. This negative selection is based on suppression of amber nonsense mutations in the barnase gene as described by Wang and Schultz[146]. The barnase protein is an extracellular ribonuclease that originates from *Bacillus amyloliquefaciens*. Without its inhibitor, barstar, it is lethal to cells[46]. That means, synthetases that are using natural amino acids in the absence of the UAA will finally suppress the amber codons in the barnase gene and mediate the cell death. Cells having synthetases that are still orthogonal and only using UAAs will be selected for (Figure 3.5).

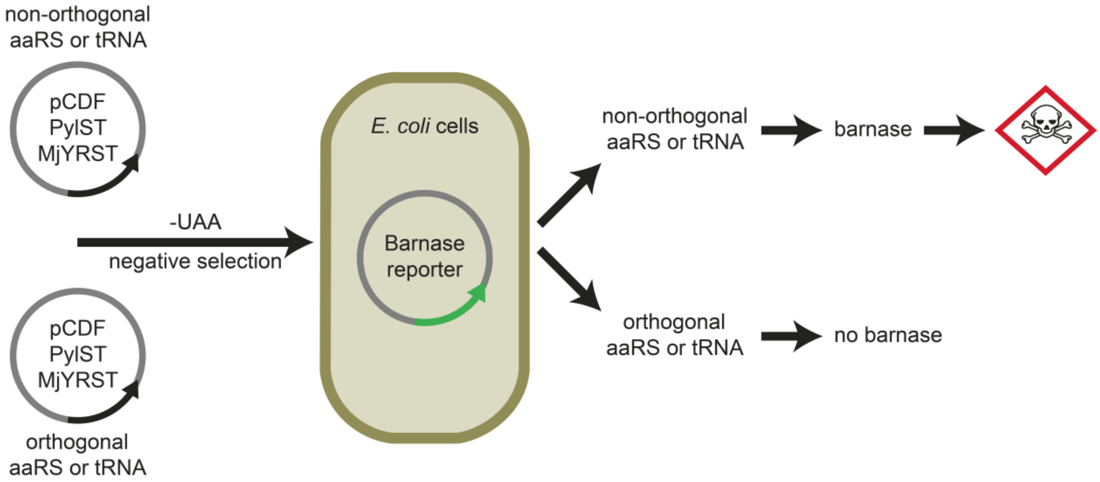

non-orthogonal aaRS or tRNA
pCDF
PylST
MjYRST
-UAA
negative selection
pCDF
PylST
MjYRST
orthogonal aaRS or tRNA
E. coli cells
Barnase reporter
non-orthogonal aaRS or tRNA
barnase
orthogonal aaRS or tRNA
no barnase

Figure 3.5: Schematic view of the barnase reporter assay.
Electro competent *E. coli* DH10B cells containing the barnase reporter plasmid (pCLA114; Table 8.1) were transformed with pCDF four library plasmids. The mixture of the library plasmids comprised undesired mutated aaRS or tRNAs that are no longer orthogonal and desired orthogonal aaRS and tRNAs. Cells were plated on agar plates with appropriate antibiotics in the absence and presence of 0.002% arabinose, to induce the expression of the barnase gene, and 1 mM BocK, as a negative control. If no UAA was added but the barnase expression induced by arabinose, cells with mutated aaRS or tRNAs are supposed to die because of the usage of natural amino acids or the aminoacylation by endogenous synthetases and thereby the suppression of the amber codons within the barnase gene. Thus, only cells with orthogonal aaRS or tRNAs should survive due to no expression of barnase. Resistance genes on plasmids are indicated by the arrows: black = Sm; green = Cm.

4 Results

4.1 Abundance And Activity of The aaRS/tRNA Pairs PylS/PylT And MjYRS/MjYT

The expansion of the genetic code allowed for the site-specific incorporation of synthetic, unnatural amino acids (UAAs) into proteins comprising a diversity of functionalities like posttranslational modifications, UV-activatable crosslinkers or photo-reactive moieties[42]. These modified proteins have been used to investigate biological questions, for example, how acetylated lysines impact chromatin dynamics[97] and gene expression[147]. The UV-activatable crosslinker *p*-benzoyl-L-phenylalanine helped to understand the complicated but important biological process of chromosome condensation in mitosis[86].

Although this technique is a powerful tool for investigating protein function, the incorporation efficiency of UAAs is limited. Several approaches aimed at optimization of the existing systems by addressing suppressor tRNA competition with release factor 1 (RF1) for binding to amber codons in the A-site. This represents one of the major reasons for truncated proteins and is accentuated for protein expression aimed at incorporating multiple identical UAAs. Johnson and colleagues overcame this issue by knocking out the essential RF1 and simultaneously “fixing” RF2[57]. The workgroups of Church and Isaacs reassigned UAG as a sense codon by genome wide replacing with the synonymous UAA codon[64] using multiplex automated genome engineering (MAGE)[63], permitting the deletion of RF1[65].

Other methods included the use of orthogonal ribosomes, which allowed for the formation of a parallel, and independent, translational machinery that was evolved to decode amber and quadruplet codons with a significantly enhanced efficiency[67,78]. Furthermore, a variety of strategies to optimize aaRS and tRNA levels have been pursued that led to improved yields of full-length protein[51,69,72].

Nevertheless, none of the approaches performed to optimize the incorporation efficiency of UAAs into proteins provided a systematic study on the correlation of aaRS and tRNA expression levels, including the loading status of the tRNA. Thus, only little insight is given into which of these components is the actual limiting factor. Here, we describe the establishment of assays for the detection of the abundance and activity of each component (PylS/PylT and MjYRS/MjYT) individually and of both tRNA/aaRS pairs in combination. This work aims to identify and optimize the limiting factors of the genetic code expansion system.

4.1.1 PylS of *Methanosarcina barkeri*

The pyrrolysine synthetase (PylS) of the archaeon *Methanosarcina barkeri* (*M. barkeri*) was the first component being modified for visualization via western blots. Specific antibodies against PylS are not available, necessitating the introduction of well-characterized epitope tags into the protein. We decided to use the small His$_6$-tag for the detection and, possibly, purification of PylS, starting with the incorporation at the N-terminus (plasmid pCLA7) by cloning the appropriate codons into the plasmid pBK PylS[95] (pCLA1; see Table 8.1 and Table 8.3). Although the DNA sequencing (Ch. 3.2.3.13) was positive, for the insertion of the tag, it was not possible to visualize signals on a western blot (Ch. 3.2.2.4) against the His$_6$-tag. Shifting the tag from the N- to C-terminus (pCLA8) also did not result in western blot signals (not shown).

To find other potential positions for the introduction of the His$_6$-tag a sequence alignment with the homologous PylS of *Methanosarcina mazei* (*M. mazei*) was performed (Figure 4.1).

```
sp|Q8PWY1|PYLS_METMA      MDKKPLNTLISATGLWMSRTGTIHKIKHHEVSRSKIYIEMACGDHLVVNNSRSSRTARAL 60
sp|Q6WRH6|PYLS_METBA      MDKKPLDVLISATGLWMSRTGTLHKIKHHEVSRSKIYIEMACGDHLVVNNSRSCRTARAF 60
                          ******:.**************:******************************.*****:

sp|Q8PWY1|PYLS_METMA      RHHKYRKTCKRCRVSDEDLNKFLTKANEDQTSVKVKVVSAPTRTKKAMPKSVARAPKPLE 120
sp|Q6WRH6|PYLS_METBA      RHHKYRKTCKRCRVSDEDINNFLTRSTESKNSVKVRVVSAP-KVKKAMPKSVSRAPKPLE 119
                          ******************:*:***::.*.:.****:***** :.********:*******

sp|Q8PWY1|PYLS_METMA      NTEAAQAQPSGSKFSPAIPVSTQESVSVPASVSTSISSISTGATASALVKGNTNPITSMS 180
sp|Q6WRH6|PYLS_METBA      NSVSAKA-------------STNTSRSVPSPAKSTPNS---------------------S 145
                          *: :*:*             **: * ***:...:: .*                     *

sp|Q8PWY1|PYLS_METMA      APVQASAPALTKSQTDRLEVLLNPKDEISLNSGKPFRELESELLSRRKKDLQQIYAEERE 240
sp|Q6WRH6|PYLS_METBA      VPASAPAPSLTRSQLDRVEALLSPEDKISLNMAKPFRELEPELVTRRKNDFQRLYTNDRE 205
                          .*..*.**:**:** **:*.**.*:*:**** .*******.**::***:*:*::*:::**

sp|Q8PWY1|PYLS_METMA      NYLGKLEREITRFFVDRGFLEIKSPILIPLEYIERMGIDNDTELSKQIFRVDKNFCLRPM 300
sp|Q6WRH6|PYLS_METBA      DYLGKLERDITKFFVDRGFLEIKSPILIPAEYVERMGINNDTELSKQIFRVDKNLCLRPM 265
                          :*******:**:***************** **:*****:***************:*****

sp|Q8PWY1|PYLS_METMA      LAPNLYNYLRKLDRALPDPIKIFEIGPCYRKESDGKEHLEEFTMLNFCQMGSGCTRENLE 360
sp|Q6WRH6|PYLS_METBA      LAPTLYNYLRKLDRILPGPIKIFEVGPCYRKESDGKEHLEEFTMVNFCQMGSGCTRENLE 325
                          ***.********** **.******:*******************:***************

sp|Q8PWY1|PYLS_METMA      SIITDFLNHLGIDFKIVGDSCMVYGDTLDVMHGDLELSSAVVGPIPLDREWGIDKPWIGA 420
sp|Q6WRH6|PYLS_METBA      ALIKEFLDYLEIDFEIVGDSCMVYGDTLDIMHGDLELSSAVVGPVSLDREWGIDKPWIGA 385
                          ::*.:**::* ***:**************:**************:.**************

sp|Q8PWY1|PYLS_METMA      GFGLERLLKVKHDFKNIKRAARSESYYNGISTNL 454
sp|Q6WRH6|PYLS_METBA      GFGLERLLKVMHGFKNIKRASRSESYYNGISTNL 419
                          ********** *.*******:*************
```

Figure 4.1: Pairwise sequence alignment with PylS from *M. mazei* and *M. barkeri*.
Alignment was executed by the program ClustalW[148] using PylS protein sequences for *M. mazei* (METMA, Q8PWY1) and *M. barkeri* (METBA, Q6WRH6) from UniProt database[149]. Similarities between both sequences are highlighted below the alignment, whereas an asterisk (*) indicates identical residues, a colon (:) stands for the conservation between groups of strongly similar properties and a period (.) for weakly similar properties (similarity score: 79.95). The linker between N- and C-terminus seems not to have a stringent sequence dependency, illustrated by the lack of 34 amino acids in the sequence of *M. barkeri* compared to *M. mazei*.

This alignment, as well as alignments with PylS from other organisms (data not shown), revealed a flexible linker region between the tRNA binding domain in the N-terminus and the catalytic domain of the C-terminus[150]. This linker region started around L118 and ended at E183 (according to the sequence of *M. barkeri*). This part of the protein was considered potentially less sensitive for mutagenesis. With the additional help of the crystal structure from *M. mazei* (Figure 4.2) we chose two positions for the incorporation of an internal His_6-tag, flanked by glycines. One tag was directed into the "gap" between S144 and S145, evoked by the lack of 34 amino acids in the sequence of *M. barkeri* compared to *M. mazei* (Figure 4.1; pCLA9) and the other one was guided into a loop which projects out of the structure (Figure 4.2; pCLA10).

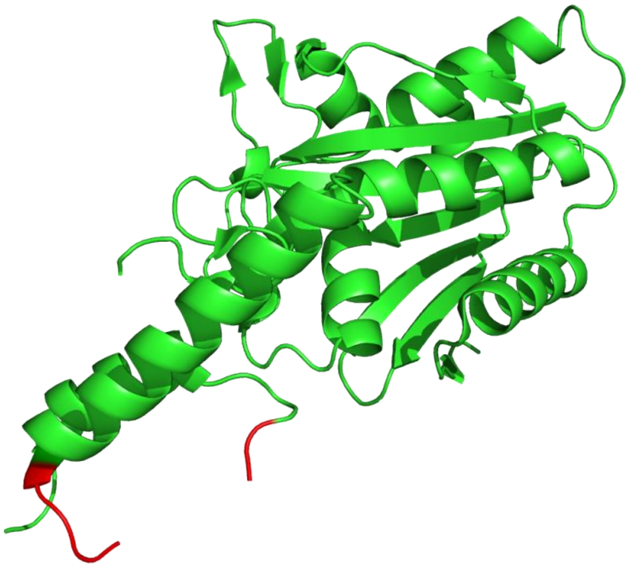


Figure 4.2: Crystal structure of the catalytic domain of PylS from *M. mazei*.
This figure was made with PyMOL[151] using pdb file 2ZIN[152] from the RSCB PDB database[153]. The highlighted red residues indicate the loop D206 to F216 (D171 to F181 in the *M. barkeri* sequence) in the linker region between N- and C-terminus. This loop was chosen to incorporate a His_6-tag because it sticks out of the structure away from the catalytic core.

To investigate if the integration of internal His_6-tags had an influence on the catalytic activity of the enzyme, a chloramphenicol reporter assay (Ch. 3.2.4.1) was performed. Hence, *E. coli* cells were co-transformed with the pREP PylT reporter (pCLA4) and pBK plasmids containing PylS genes with internal tags (pCLA9 and pCLA10), without a His_6-tag (pCLA1) and with a C-terminal tag (pCLA8) (Figure 4.3). The assay showed that cells with wild type (WT) PylS and PylS with a His_6-tag introduced into the "gap" between S144 and S145 exhibit the same growth behavior and were resistant to a Cm concentration of approximately 400 µg/mL. Whereas cells harboring PylS with a C-terminal tag died around 75 µg/mL Cm, cells with the loop version survived Cm up to 150 µg/mL.

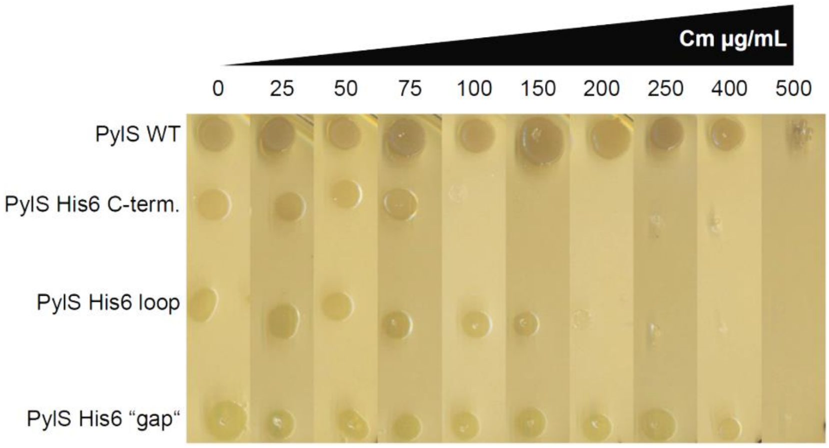


Figure 4.3: Cm-Assay with pBK PylS His$_6$-tag variants.
The catalytic activity of wild type (WT; pCLA1) PylS was compared with different His$_6$-tagged PylS variants (pCLA8 to pCLA10). Therefore a chloramphenicol reporter assay (Ch. 3.2.4.1) was performed and cells were plated on agar plates containing Kan, Tet, increasing Cm and 1 mM BocK. PylS with a tag introduced into the "gap" between S144 and S145 (pCLA9) showed an activity comparable to the WT (pCLA1)since cells were able to grow up to a Cm concentration of 400 µg/mL. A His$_6$-tag at the C-terminus (C-term; pCLA8) allowed the cells to survive a Cm concentration of 75 µg/mL, only. The loop variant (pCLA10) let the cells grow up to 150 µg/mL Cm.

Thus, the PylS "gap" variant (pCLA9) was the most promising mutant and needed to be tested for detectability on western blots. We performed a test expression using cells transformed with the appropriate plasmid and took samples in exponential growth phase, late exponential growth phase and stationary phase (Figure 4.4).

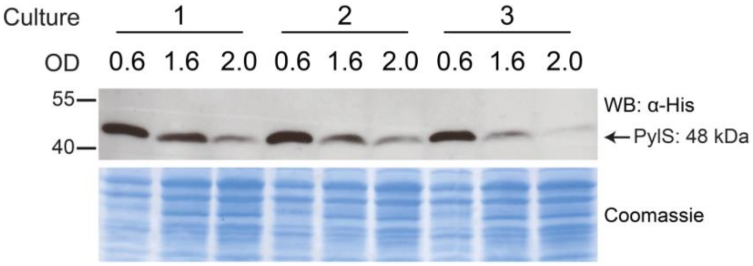


Figure 4.4: Detection of PylS His$_6$-tag ("gap") via western blot.
Three individual cultures with cells harboring plasmid pCLA9 were grown. Samples were taken and normalized to $OD_{600} = 1.5$ in exponential growth phase ($OD_{600} = 0.6$), late exponential growth phase ($OD_{600} = 1.6$) and stationary phase ($OD_{600} = 2.0$). Whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody.

This position allowed for the detection of the His$_6$-tag and thereby the PylS via western blot. As depicted in Figure 4.4, a decrease of the signal at late exponential growth phase and again at stationary phase could be observed, although the loading control (Coomassie stain of the gel) showed even a slight increase in total protein yield. Due to the results from the activity test and from western blot we decided to use the PylS "gap" variant for further experiments.

Due to the fact that no crystal structures of *M. barkeri* PylS are currently available, the only structural insights come from that of the *M. mazei* homologue. This structure only reveals the C-terminal domain[154] because the highly basic, yet hydrophobic, N-terminal domain was unstable and inadequate for crystallization trials[150]. Additionally, the N-terminus was known to be essential for *in vivo* activity but dispensable *in vitro*[155] and its overall function, binding of PylT, was quite recently found by Jiang and Krzycki[150]. As mentioned above, the linker between the N- and C-terminus varies in sequence and length[155]. Furthermore, it has a high content of prolines, which are responsible for the instabilities due to the inability to form hydrogen bonds and the introductions of bends, or kinks, that disrupt helices or sheets. In the case of the *M. barkeri* PylS, the 66 AAs of the linker region include eight prolines, which correlate to circa 12%. In summary, the aaRS PylS exhibits several regions that destabilize the protein and thus potentially decrease the activity.

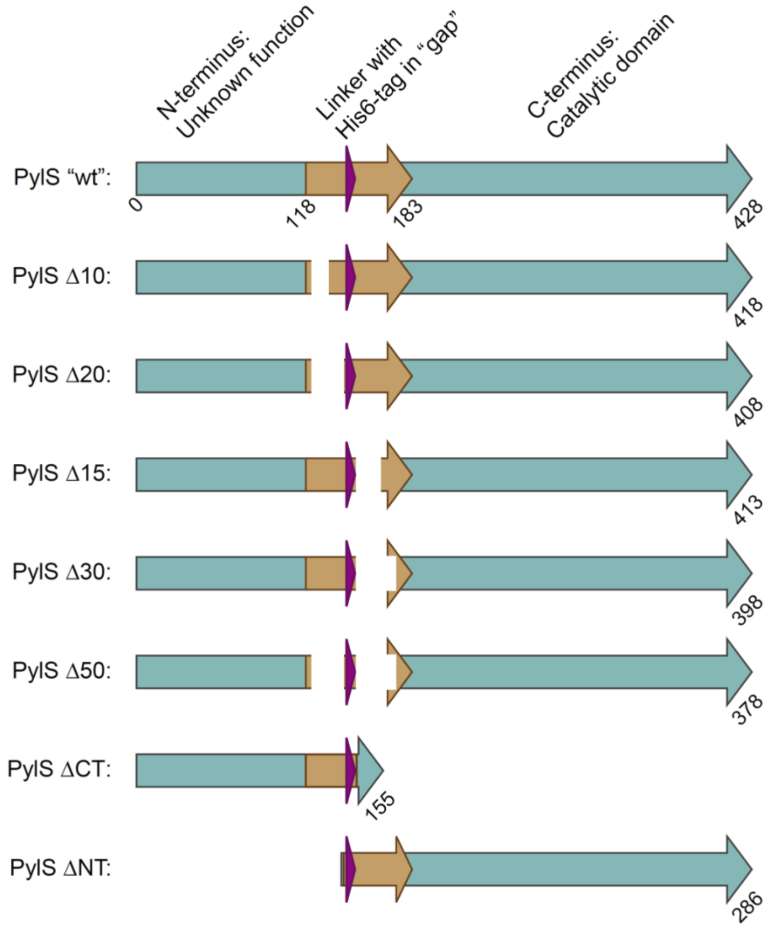

N-terminus:
Unknown function
Linker with
His6-tag in "gap"
C-terminus:
Catalytic domain
PylS "wt":
0
118
183
428
PylS Δ10:
418
PylS Δ20:
408
PylS Δ15:
413
PylS Δ30:
398
PylS Δ50:
378
PylS ΔCT:
155
PylS ΔNT:
286

Figure 4.5: Truncations in the PylS "gap" ("wt") variant to investigate the effect on the suppression efficiency. Truncations (Δ; the number specifies the quantity of AAs deleted; CT = C-terminus; NT = N-terminus) are indicated by the white spaces and were made via QC on pCLA9 resulting in plasmids pCLA11 to pCLA17. The protein PylS is pictured by the turquois arrows with numbers indicating the length in AAs below. The linker region is highlighted by the brown arrow and the internal His_6-tag is accentuated by the violet arrowhead.

In order to determine if these parts of the protein are dispensable and thereby enhance the suppression efficiency, we produced truncations of different lengths inside the linker of the PylS "gap" variant (pCLA9). All of these truncations surrounded the internal His$_6$-tag (Figure 4.5). In addition, the complete N- and C-terminus were deleted to investigate the effect on the suppression efficiency as well.

To find out whether these truncations influence the activity and thereby the suppression efficiency of PylS a chloramphenicol reporter assay was performed (Figure 4.6). Therefore, competent *E. coli* cells harboring the pREP PylT reporter (pCLA4) were transformed with the truncation plasmids (pCLA11 to pCLA17) and compared to PylS "gap" without any changes (pCLA9) as a reference.

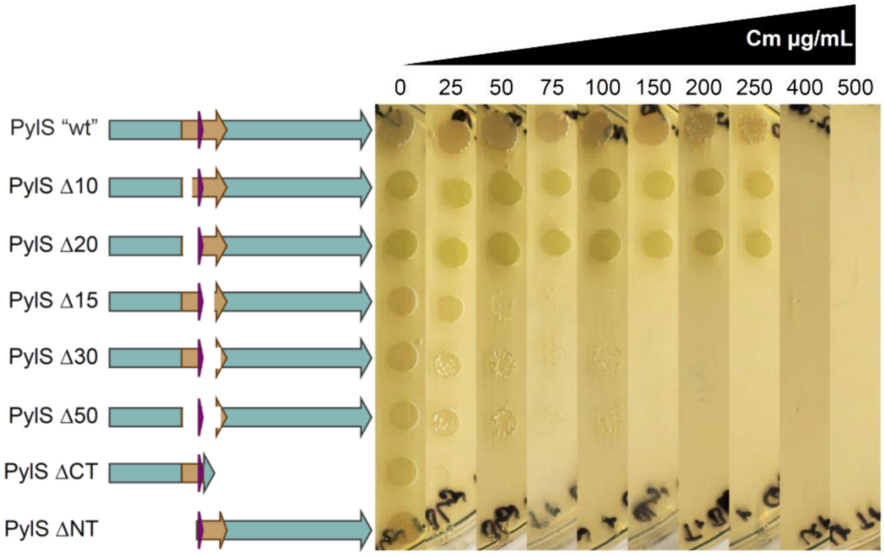


Figure 4.6: Cm-Assay with truncation variants of PylS "gap" ("wt").
The catalytic activity of "wt" PylS "gap" variant (pCLA9) was compared with different truncations made in the linker region between N- and C-terminus (pCLA11 to pCLA17; Figure 4.5). Therefore a chloramphenicol reporter assay (Ch. 3.2.4.1) was performed and cells were plated on agar plates containing Kan, Tet, increasing Cm and 1 mM BocK. Deletions at the beginning of the linker, Δ10 and Δ20, did not influence the suppression efficiency negatively but cells were also not able to survive higher Cm concentrations than cells with the "wt" variant. All three versions let cells grew to a Cm concentration of 250 µg/mL. Cells with the other truncated PylS variants started dying already between Cm 25 and 50 µg/mL. Whereas the smaller truncations in the linker regions still allowed survival of a few cells up to Cm 100 µg/mL, especially the deletion of the C-terminus prevented growth on plates with Cm higher than 25 µg/mL.

We observed that the N-terminal linker truncations, Δ10 and Δ20, revealed the same growth behavior as the "wild type" PylS. The suppression efficiency was neither decreased nor significantly increased. All three versions allowed cell growth to a Cm concentration of 250 µg/mL. Deletions in the C-terminal linker region, as well as the deletion of the whole N- and C-terminus, caused the cells to die between Cm 25 and 50 µg/mL. The removal of the C-terminus caused a severe effect because no growth on agar plates with Cm higher than

25 µg/mL was noticed. The truncations Δ15, Δ30, Δ50 and ΔNT permitted very weak growth up to 100 µg/mL chloramphenicol.

Since truncations of the PylS "gap" variant (pCLA9) did not result in a significant improvement of the suppression efficiency we used the "wt" version of this variant in further experiments.

4.1.2 PylT of *Methanosarcina barkeri*

For the optimization of the aaRS PylS (Ch. 4.1.1) it was also necessary to analyze the cognate tRNA PylT for its abundance and loading status. We chose to approach this assay using Northern blotting (Ch. 3.2.3.15) techniques and therefore hybridization probes had to be designed. As described in Ch. 3.2.3.15, we used DNA oligonucleotides labeled with Cy3, Cy5 or DIG, complementary to the anticodon stem and loop domain of the desired tRNA target. The oligos C25 and C26 (Table 8.2) were designed to anneal to the anticodon loop of PylT as shown in Figure 4.7.

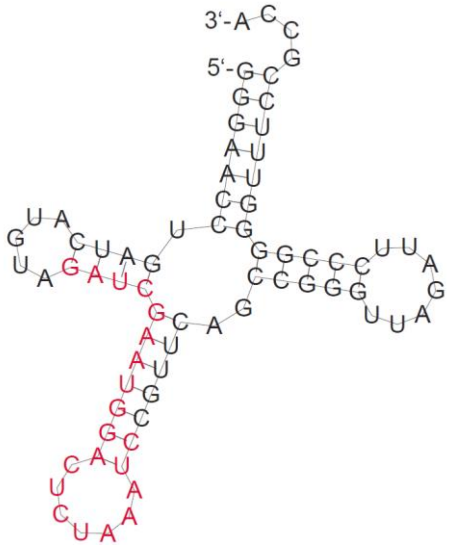


Figure 4.7: Predicted secondary structure of PylT.
The structure was obtained from ViennaRNA Web Services[156]. The region to which the probes C25 and C26 bind is highlighted in red, representing the anticodon stem and a section of the D-loop.

Initial trials for the detection of PylT were carried out with *in vitro* synthesized targets. These were the DNA oligonucleotide C27 that represented the bases 13-42 (of 72) from PylT and a RNA transcript of PylT (T7 PylT; Ch. 3.2.3.16). Both targets could be detected with signal intensities depending on the concentration of the target as depicted in Figure 4.8-A. We first approached this detection using fluorophore labeled probes (Cy3 and Cy5) for hybridization. We then altered our approach to include DIG labeled probes due to increased sensitivity, as shown for T7RP transcribed *E. coli* Ser-tRNA-1 (T7 SerT; Figure 4.8-B).

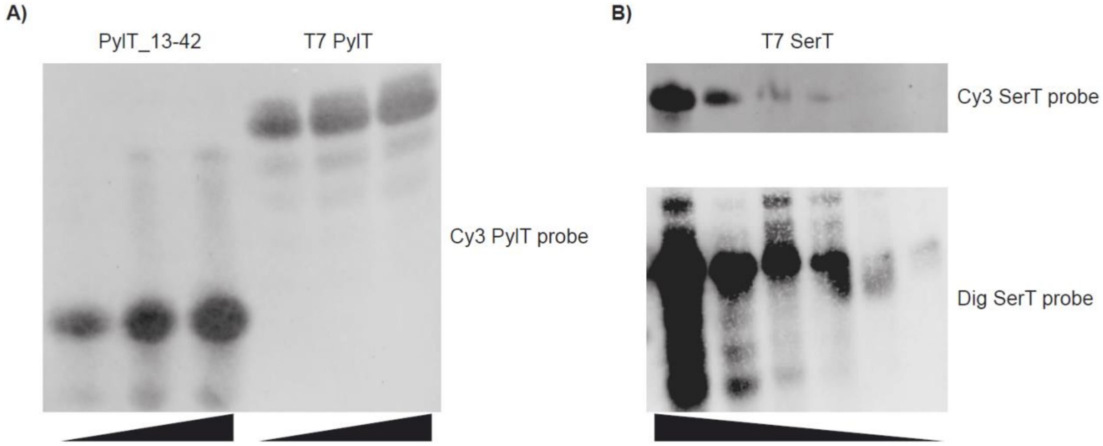


Figure 4.8: Establishing northern blots to detect specific tRNAs.
A) To verify the ability of the hybridization probes C25 and C26 to bind to the specific target tRNA PylT, artificial samples were separated with acid urea PAGE and subsequently blotted (Ch. 3.2.3.15). DNA oligonucleotide C27 (PylT_13-42) was loaded with a total amount of 2.3, 4.6 and 6.9 µg, respectively (left to right, indicated by the black triangle). In contrast 6.0, 12.0 and 18.0 µg of T7 PylT were loaded. A Cy3 labeled PylT probe (C25) was used for hybridization and visualized by a Typhoon imager (Excitation: 532 nm; Filter: 580 BP 30; PMT 400; Resolution: 50 µm per pixel). **B)** Comparison of a Cy3 labeled probe (C32; upper panel) against *E. coli* Ser-tRNA-1 to a DIG labeled probe (C33; lower panel). A total amount of 13.3, 1.33, 0.48, 0.24, 0.048 and 0.024 µg T7 SerT was loaded, respectively (right to left). The Cy3 probe was visualized by a Typhoon imager (Excitation: 532 nm; Filter: 580 BP 30; PMT 460; Resolution: 200 µm per pixel) and the DIG probe by decorating with an anti-DIG-AP-antibody followed by a chemiluminescence reaction using CDP-Star as described in Ch. 3.2.3.15.

The next step was to detect the PylT in total RNA extracts from *E. coli* cells. Hence, different cell cultures harboring plasmids for either PylS only or PylS and PylT (pCLA1 and pCLA4) were grown in the presence, or absence, of BocK. Total RNA was isolated (Ch. 3.2.3.14) and separated with acid urea PAGE. Several standards (Ch. 3.1.4) were used to estimate the size of the isolated RNA compared to the positive control T7 PylT, described above (Figure 4.9-A). The T7 PylT gave a signal around 70 nt which fits to the actual size of the tRNA with 72 nt. Total RNA extracts from cells yielded a smear, as expected, due to the many distinct RNA fragments with different lengths, like rRNAs, mRNAs, tRNAs, riboswitches etc. Subsequent northern blots of separated RNA revealed positive signals for PylT only if the appropriate plasmid was present in the cells (Figure 4.9-B). Decreasing the concentration of loaded RNA samples sufficiently resulted in the emergence of a second band with different migratory behavior. This second band only appeared if RNA was isolated from cells grown in medium supplemented with BocK (Figure 4.9-C). The appearance of the second signal, depending on the addition of the amino acid, suggested that the upper band represented the aminoacylated form of PylT and the lower one the uncharged form.

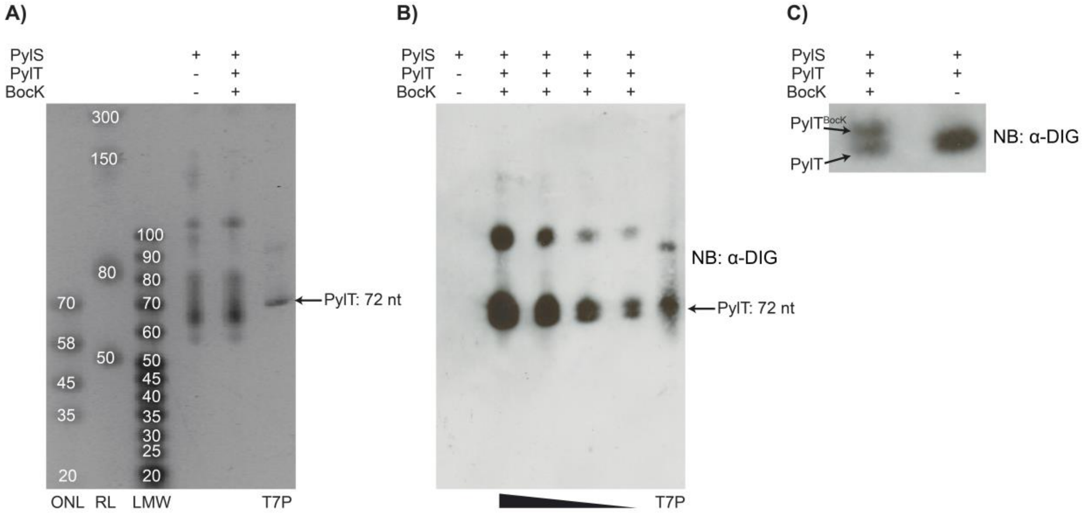


Figure 4.9: Detection of PylT in total RNA extracts.
A) Acid urea PAGE for the size estimation of isolated RNA from cell cultures harboring plasmids for PylS (pCLA1; 1.4 µg loaded onto gel) and PylS/PylT (pCLA1 and pCLA4; 1.45 µg loaded). The latter one was grown in LB medium supplemented with 1 mM BocK. T7 PylT (T7P) with a size of 72 nt served as a positive control (0.5 µg loaded). Three different size standards consisting of DNA and RNA nucleotides were used to consider different running behavior. ONL = Oligonucleotide Ladder; RL = Low Range ssRNA Ladder; LMW = Low Molecular Weight Marker. The gel was soaked in a "gel red" bath (1:10,000 in H_2O (v/v)) and separated DNA/RNA bands were visualized by UV light using a gel documentation machine. **B)** Northern blot of samples described in A). A total amount of 2.8 µg RNA from cells with PylS only was loaded. In contrast 1.45, 0.7, 0.35 and 0.17 µg RNA from cells with PylS and PylT (+ 1 mM BocK) were loaded (left to right, indicated by the black triangle). 0.03 µg of the positive control T7 PylT were used. The DIG labeled probe C26 was utilized for hybridization and chemiluminescence signals recorded as described in Ch. 3.2.3.15. **C)** Northern blot for the comparison of cell cultures grown in the presence and absence of 1 mM Bock in the medium. Approximately 0.045 µg RNA were loaded for each sample. Signals were detected as in B). The upper band ($PylT^{BocK}$) represents the aminoacylated form of PylT, whereas the lower band (PylT) depicts the uncharged form.

4.1.3 MjYRS of *Methanocaldococcus jannaschii*

Analogous to PylS (Ch. 4.1.1) no antibodies against the tyrosine synthetase (MjYRS) of the archaeon *Methanocaldococcus jannaschii* (*M. jannaschii*) were available. Thus, we again took the small His_6-tag as an alternative and cloned it into the N-terminus (pCLA18). We observed again that, similar to the modifications on PylS, no signals on a western blot against the His_6-tag were obtained. Shifting of the tag from N- to C-terminus (pCLA19), we again observed no signal from a western blot (Figure 4.12). Hence, comparable to PylS, we chose an internal position for the His_6-tag based on the crystal structure depicted in Figure 4.10.

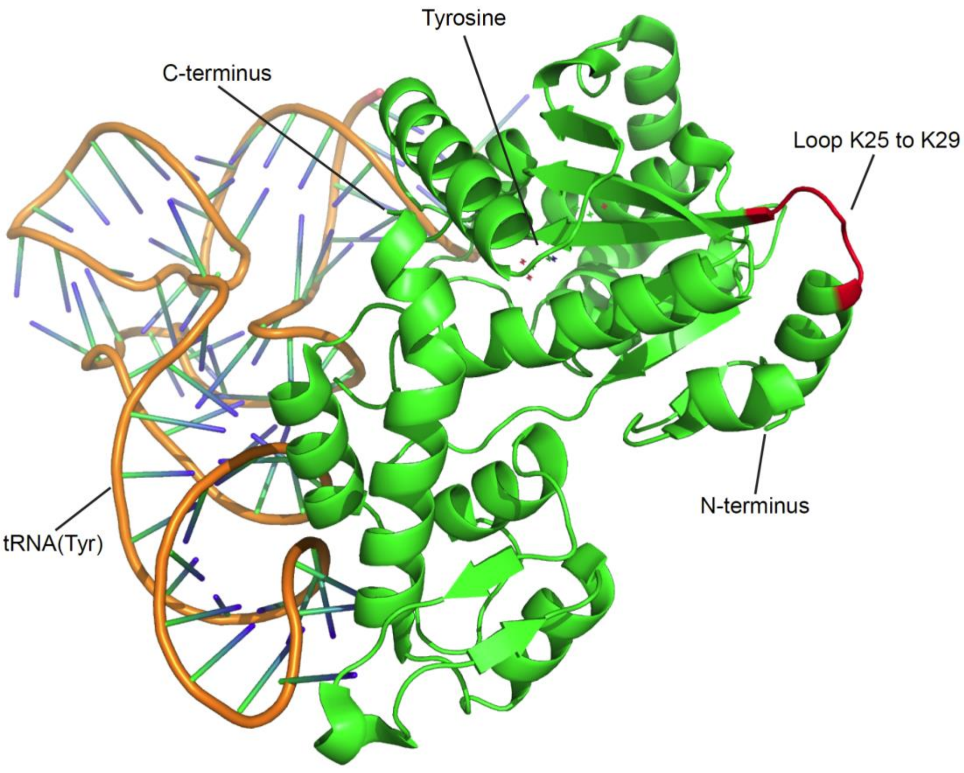


Figure 4.10: Crystal structure of tyrosyl-tRNA synthetase from *M. jannaschii* complexed with tRNA(Tyr) and L-tyrosine.
This figure was made with PyMOL[151] using pdb file 1J1U[73] from the RSCB PDB database[153]. The highlighted red residues indicate the loop K25 to K29. This loop was chosen to incorporate a His_6-tag because it protrudes from the structure away from the catalytic core. The integration into this loop should not interfere with the tRNA sterically.

The His_6-tag, again flanked by glycines, was incorporated into the loop K25 to K29 (pCLA20). As shown in Figure 4.10 the highlighted red loop points away from the catalytic core, indicated by the tyrosine, as well as from the corresponding tRNA. Thus, the mutation should not constrict protein activity what was verified by a chloramphenicol reporter assay (Ch. 3.2.4.1). Therefore *E. coli* cells were transformed with pBK plasmids harboring the genes for MjYRS WT, C-terminal and internal His_6-tag (pCLA2, pCLA19 and pCLA20), respectively, together with the appropriate reporter plasmid (pCLA5).

It seemed that the incorporation of the His_6-tag had no negative influence on the overall catalytic activity. Cells were able to survive all given concentrations of chloramphenicol up to 500 µg/mL, independent of the MjYRS version (Figure 4.11).

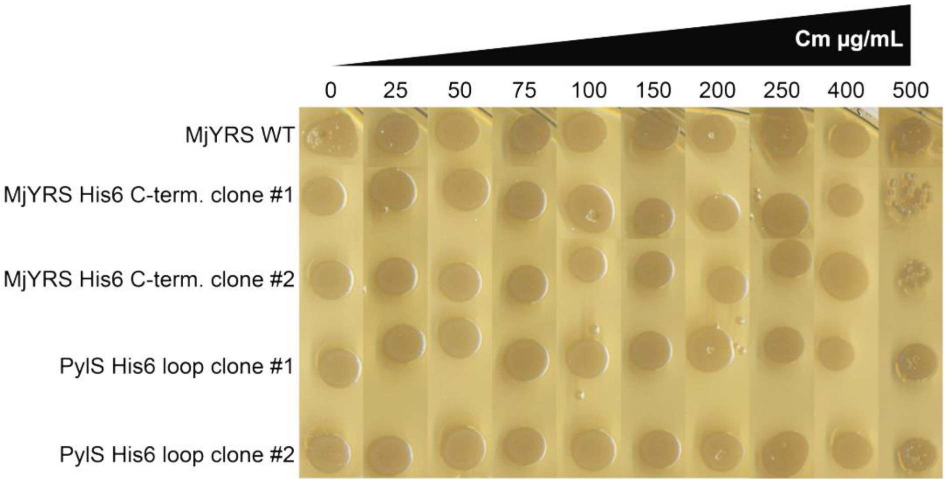

Cm µg/mL
0
25
50
75
100
150
200
250
400
500
MjYRS WT
MjYRS His6 C-term. clone #1
MjYRS His6 C-term. clone #2
PylS His6 loop clone #1
PylS His6 loop clone #2

Figure 4.11: Cm-Assay with pBK MjYRS His$_6$-tag variants.
The catalytic activity of wild type (WT; pCLA2) MjYRS was compared with different His$_6$-tagged MjYRS variants (pCLA19 and pCLA20). Therefore a chloramphenicol reporter assay (Ch. 3.2.4.1) was performed and cells were plated on agar plates containing Amp, Tet and increasing Cm. The His$_6$-tagged versions were plated in duplicates but no significant difference could be observed. Cells with all kinds of His$_6$-tagged MjYRS could grow up 500 µg/mL Cm. If at all, the C-terminal tagged versions show a slightly weaker growth at the highest Cm concentration.

The next step was to investigate if the MjYRS "loop" variant (pCLA20) could be detected via western blot in contrast to the N- and C-terminal tagged variants. We performed a test expression using cells transformed with the appropriate plasmids and took samples in exponential growth phase, late exponential growth phase and stationary phase (Figure 4.12).

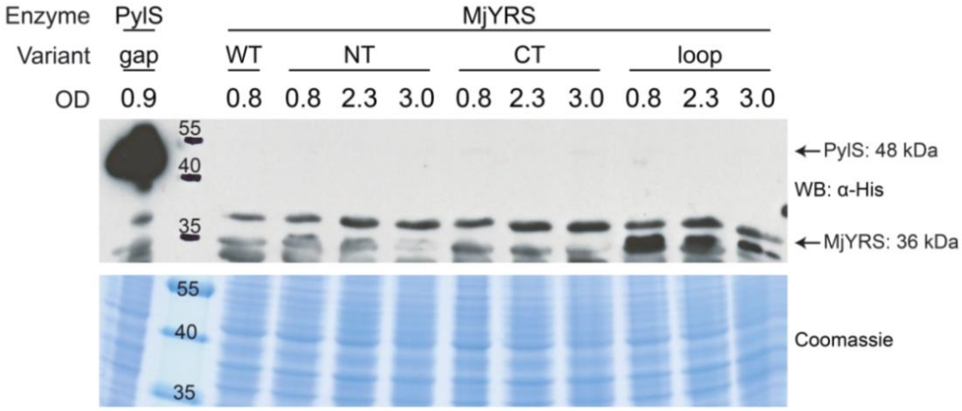

Enzyme
PylS
MjYRS
Variant
gap
WT
NT
CT
loop
OD
0.9
0.8
0.8
2.3
3.0
0.8
2.3
3.0
0.8
2.3
3.0
55
40
35
←PylS: 48 kDa
WB: α-His
←MjYRS: 36 kDa
55
40
35
Coomassie

Figure 4.12: Detection of MjYRS His$_6$-tag via western blot.
Four different cultures were grown with cells harboring the plasmid for WT MjYRS (pCLA2) without a His$_6$-tag, N-terminal (NT) His$_6$-tagged MjYRS (pCLA18), C-terminal tagged MjYRS (CT; pCLA19) and the "loop" variant of MjYRS (pCLA20). Samples were taken in exponential growth phase (OD_{600} = 0.8), late exponential growth phase (OD_{600} = 2.3) and stationary phase (OD_{600} = 3.0). Cells counts were normalized by resuspending the cell pellets in appropriate amounts of SDS sample buffer. Whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody. PylS "gap" from *E. coli*, transformed with the plasmid pCLA9, in exponential growth phase (OD_{600} = 0.9) served as positive control.

In a similar manner to PylS, as shown in Figure 4.4, only the unconventional internal His_6-tag gave a suitable signal on a western blot against the His_6-tag. The intensity is significantly lower compared to the PylS signal, although both proteins are expressed under the same promoter. As already mentioned for PylS, the signal seemed to decrease at late exponential growth phase and again at stationary phase. The lanes for the samples taken from the culture with the N-terminal tagged MjYRS also exhibit a very weak signal of the correct size, which decreases over time and is not present in those lanes from WT or C-terminal tagged MjYRS. Considering the results from the Cm assay (Figure 4.11), in which no substantial difference between the MjYRS variants could be observed, this could be an indication that the proteins have a very short but distinct half-life. This could mean that there are sufficient molecules for the suppression of the amber codon but not enough left for a proper detection via western blot. The accessibility of the epitope by antibodies should not be the limiting factor because the cell lysates were prepared for western blot with denaturing conditions using SDS.

4.1.4 MjYT of *Methanocaldococcus jannaschii*

The last of the four components analyzed during this study was the corresponding tRNA of the aaRS MjYRS (Ch. 4.1.3) and its derivatives, called MjYT. MjYT, and the corresponding aaRS, were available in two different versions, where one was decoding the amber stop codon UAG (MjYT_CUA; pCLA5) and the other the frameshift codon AGGA (MjYT_UCCU; pCLA6).

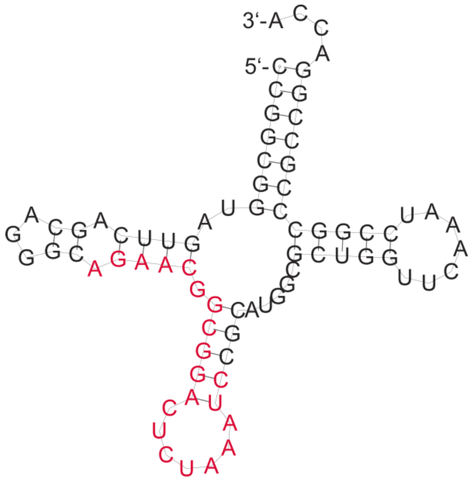


Figure 4.13: Predicted secondary structure of MjYT_CUA.
The structure was obtained from ViennaRNA Web Services[156]. The region to which the probes C40 and C41 bind, that is the anticodon stem and in parts the D-loop, is highlighted in red.

For the detection of these tRNAs on northern blots we developed hybridization probes covering the same regions of the tRNA as shown for PylT (Figure 4.7), namely the anticodon stem and in parts the D-loop. The oligos C40 and C41 (Table 8.2) were designed to anneal to the stop codon tRNA as shown in Figure 4.13.

Next, we extracted RNA from *E. coli* cells (Ch. 3.2.3.14) containing plasmids for either MjYRS_TAG only (pCLA2) or in combination with the cognate tRNA (pCLA2 and pCLA5). Additionally, RNA was isolated from cells transformed with a plasmid that held both components at once but in which the aaRS was decoding for *p*-benzoyl-L-phenylalanine[83] (BPA; pCLA21) instead of tyrosine. Total RNA was separated with acid urea PAGE and its size was estimated by size standards (Ch. 3.1.4) and T7RP transcribed *M. jannaschii* Tyr-tRNA (T7 TyrT) as depicted in Figure 4.14-A.

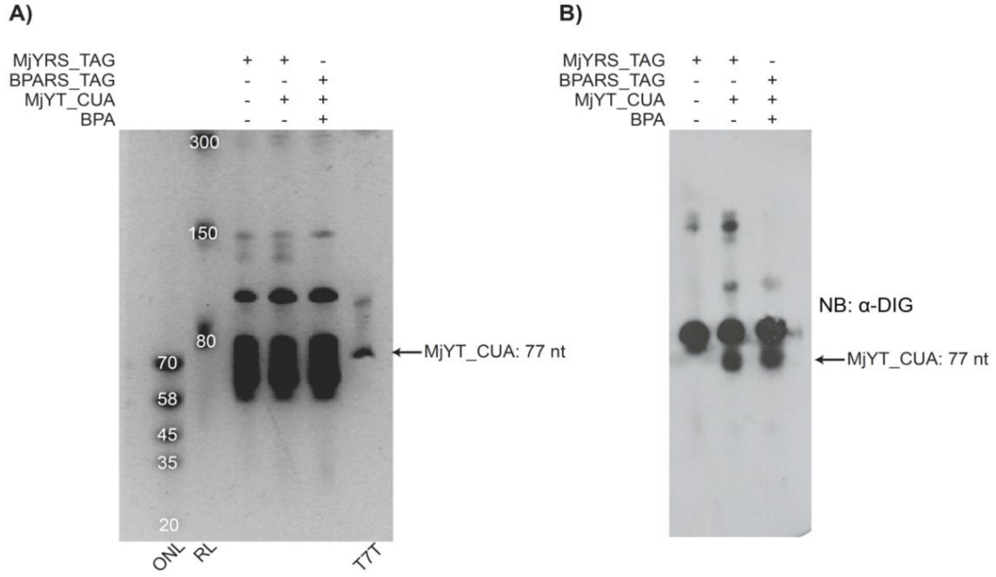
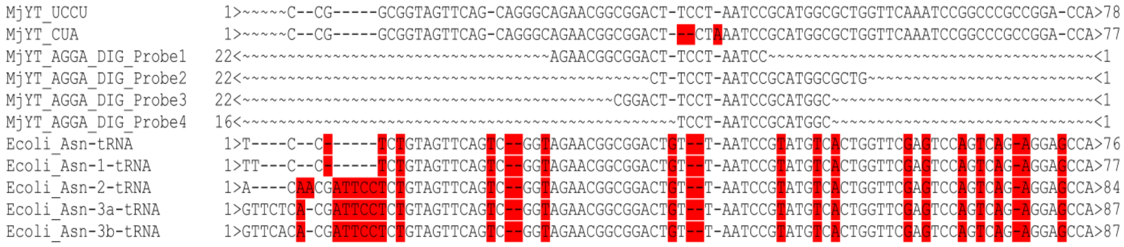


Figure 4.14: Detection of MjYT_CUA in total RNA extracts.
A) Acid urea PAGE for the size estimation of isolated RNA from cell cultures harboring plasmids for MjYRS_TAG (pCLA2; 2.8 µg loaded onto gel), MjYRS_TAG/MjYT_CUA (pCLA2 and pCLA5; 2.7 µg loaded) and BPARS/MjYT_CUA (pCLA21; 3.0 µg loaded). The latter one was grown in LB medium supplemented with 1 mM BPA. T7 TyrT (T7T) with a size of 77 nt served as a positive control (0.8 µg loaded). Two different size standards consisting of DNA and RNA nucleotides were used to consider different running behavior. ONL = Oligonucleotide Ladder; RL = Low Range ssRNA Ladder. The gel was soaked in a “gel red” bath (1:10,000 in H2O (v/v)) and separated DNA/RNA bands were visualized by UV light using a gel documentation machine. **B)** Northern blot of samples described in A). A total amount of 1.4 µg RNA from cells with MjYRS_TAG only was loaded. In contrast 1.35µg RNA from cells with MjYRS_TAG and MjYT_CUA and 1.5 µg RNA from cells with BPARS/MjYT_CUA (+1 mM BPA) were loaded. The DIG labeled probe C41 was utilized for hybridization and chemiluminescence signals recorded as described in Ch. 3.2.3.14.

We observed a signal between 70 and 80 nt in the lane for the T7 TyrT (Figure 4.14-A). The actual size of the tRNA is 77 nt. Total RNA extracts from cells yielded the expected smear. The northern blot of the same samples (Figure 4.14-B) showed two prominent bands. The upper one was present in all lanes and the lower one only if the cells were transformed with the plasmid for the tRNA. Since the lower one only appeared in samples with the tRNA and the size correlated to the signal seen in Figure 4.14-A, we assigned this band as MjYT_CUA.

The appearance of the second upper band was thought to be an unspecific binding of the probe to other RNA fragments, especially other tRNAs, due to its presence in all lanes. Thus, we blasted the sequence of the probe against the genome of *E. coli* (using the program BLASTN 2.2.29+[157], data not shown). It became apparent that the last 13 bases of the probe used for the northern blot were complementary to several kinds of *E. coli* asparagine tRNAs. Hence, new probes binding specifically to the *M. jannaschii* tRNA had to be developed and tested. Figure 4.15 shows an alignment of the two versions of MjYT, CUA and UCCU, together with those asparagine tRNAs to which the probe C41 was hybridizing. In addition to the tRNAs, all different probes tested for a specific binding to MjYT were aligned, those were MjYT_AGGA_DIG_Probe1-4 (C44 to C47). We switched to the frameshift codon version of MjYT since this tRNA was used in further experiments in combination with the UAG decoding tRNA PylT.



```
MjYT_UCCU              1>~~~~~C--CG-----GCGGTAGTTCAG-CAGGGCAGAACGGCGGACT-TCCT-AATCCGCATGGCGCTGGTTCAAATCCGGCCCGCCGGA-CCA>78
MjYT_CUA               1>~~~~~C--CG-----GCGGTAGTTCAG-CAGGGCAGAACGGCGGACT---CTAAATCCGCATGGCGCTGGTTCAAATCCGGCCCGCCGGA-CCA>77
MjYT_AGGA_DIG_Probe1  22<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~AGAACGGCGGACT-TCCT-AATCC~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~<1
MjYT_AGGA_DIG_Probe2  22<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~CT-TCCT-AATCCGCATGGCGCTG~~~~~~~~~~~~~~~~~~~~~~~~~<1
MjYT_AGGA_DIG_Probe3  22<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~CGGACT-TCCT-AATCCGCATGGC~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~<1
MjYT_AGGA_DIG_Probe4  16<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~TCCT-AATCCGCATGGC~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~<1
Ecoli_Asn-tRNA         1>T----C--C------TCTGTAGTTCAGTC--GGTAGAACGGCGGACTGT--T-AATCCGTATGTCACTGGTTCGAGTCCAGTCAG-AGGAGCCA>76
Ecoli_Asn-1-tRNA       1>TT---C--C------TCTGTAGTTCAGTC--GGTAGAACGGCGGACTGT--T-AATCCGTATGTCACTGGTTCGAGTCCAGTCAG-AGGAGCCA>77
Ecoli_Asn-2-tRNA       1>A----CAACGATTCCTCTGTAGTTCAGTC--GGTAGAACGGCGGACTGT--T-AATCCGTATGTCACTGGTTCGAGTCCAGTCAG-AGGAGCCA>84
Ecoli_Asn-3a-tRNA      1>GTTCTCA-CGATTCCTCTGTAGTTCAGTC--GGTAGAACGGCGGACTGT--T-AATCCGTATGTCACTGGTTCGAGTCCAGTCAG-AGGAGCCA>87
Ecoli_Asn-3b-tRNA      1>GTTCACA-CGATTCCTCTGTAGTTCAGTC--GGTAGAACGGCGGACTGT--T-AATCCGTATGTCACTGGTTCGAGTCCAGTCAG-AGGAGCCA>87
```

Figure 4.15: Sequence alignment of MjYT.
The DNA sequences of both MjYT versions (CUA = pCLA5 and UCCU = pCLA6) were compared to northern blot hybridization probes MjYT_AGGA_DIG_Probe1-4 (C44 to C47) and *E. coli* tRNAs Asn, Asn1, Asn2, Asn3a and Asn3b.

Subsequent northern blots revealed that MjYT_AGGA_DIG_Probe3 (C46; Figure 4.16-A) was the only hybridization probe that was able to bind to the tRNA MjYT_UCCU specifically without giving background signals from *E. coli* Asn-tRNAs (Figure 4.16-B). The other probes yielded very weak signals, no signals at all, or exhibited unspecific binding (not shown). Furthermore, Probe3 could discriminate between the two versions of MjYT, only giving signals if the frameshift variant was present in the cells. Comparable to PylT (Figure 4.9-C) we observed two bands at different heights. Since this only happened if the corresponding aaRS was present, the upper signal was assigned as the aminoacylated form of MjYT_UCCU, again.

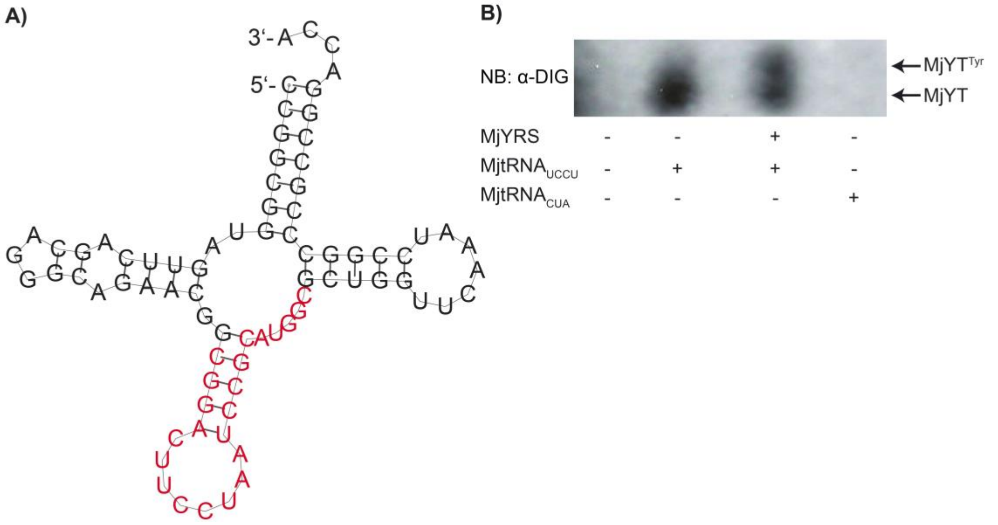


Figure 4.16: Refinement of MjYT detection.
A) Predicted secondary structure of MjYT_UCCU. The structure was obtained from ViennaRNA Web Services[156]. The region to which the probe MjYT_AGGA_DIG_Probe3 (C46) binds is highlighted in red. **B)** Northern blot from the refinement studies using probes MjYT_AGGA_DIG_Probe1-4 (C44 to C47). In this case only MjYT_AGGA_DIG_Probe3 (C46) is shown. Approximately 0.090 µg RNA were loaded for each sample. After hybridization, chemiluminescence signals were recorded as described in Ch. 3.2.3.14. The upper band (MjYT^{Tyr}) represents the aminoacylated form of MjYT_UCCU, whereas the lower band (MjYT) depicts the uncharged form.

4.2 Promoter Libraries

Based on the fact that both tRNA/aaRS pairs could now be detected quantitatively and their activity measured as well (Ch. 4.1), we began to optimize the suppression efficiency of both systems.

One cosideration for optimizations was already discussed in the previous chapter. The standard promoter in front of both synthetases, a *glnS* promoter, was not fully active during the full expression time. More precisely, we discovered a decrease of the expression levels of PylS and MjYRS over time(Figure 4.4 and Figure 4.12). In theory, a higher expression level of a certain synthetase should lead to a higher rate of aminoacylation of the corresponding tRNA and thus, to a more efficient decoding of the appropriate codon. We believed this could be achieved by the introduction of a stronger synthetase promoter. However, a pure overexpression could also cause negative effects due to toxicity, especially if the component is not endogenous. The use of multiple elements in one cell could boost those negative effects.

Therefore, we tried to optimize the expression levels of both aaRS and tRNAs individually and in combination. For this purpose synthetic promoter libraries were used to fine tune the expression of the appropriate genes[158,159].

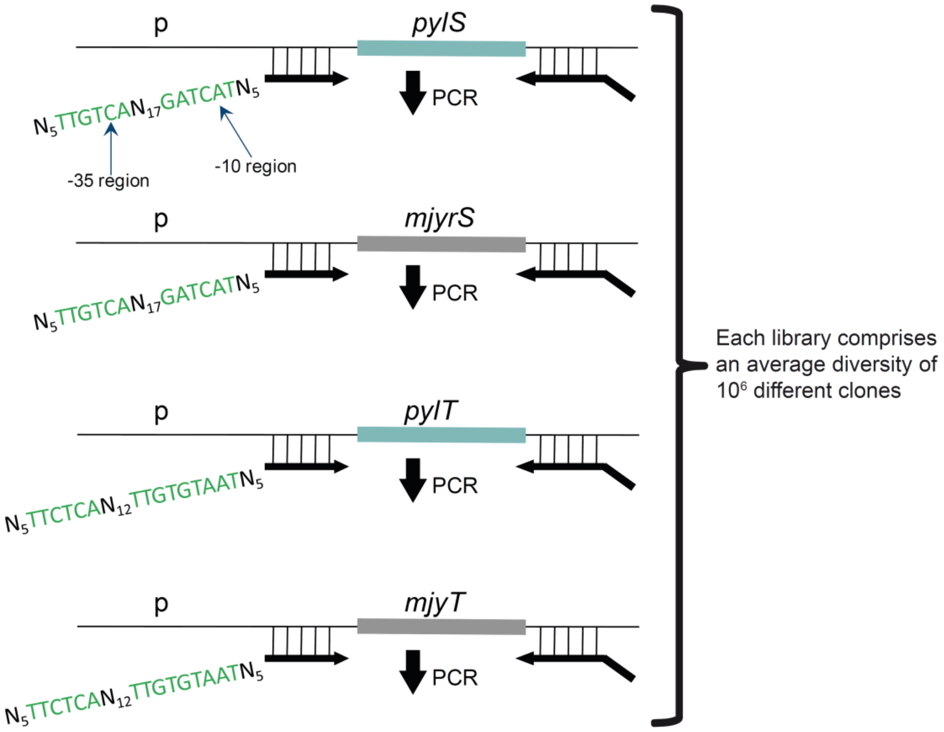

p
pylS
$N_5TTGTCAN_{17}GATCATN_5$
PCR
-35 region
-10 region
p
mjyrS
$N_5TTGTCAN_{17}GATCATN_5$
PCR
p
pylT
$N_5TTCTCAN_{12}TTGTGTAATN_5$
PCR
p
mjyT
$N_5TTCTCAN_{12}TTGTGTAATN_5$
PCR
Each library comprises an average diversity of 10^6 different clones

Figure 4.17: Schematic overview of the synthetic promoter libraries.
The promoters (p) of the genes for *pylS*, *mjyrS_AGGA*, *pylT* and *mjyT_UCCU* were modified randomly via inverse PCR (Ch. 3.2.3.11). The -10 and the -35 box, highlighted in green, remained thereby unaffected. N_x represents the number of exchanged nucleotides, with N standing for the bases A, T, G and C.

The mutagenesis of the *glnS* promoters in front of the aaRS and *lpp* promoters ahead of the tRNAs, was performed using inverse PCR (Ch. 3.2.3.11). As depicted in Figure 4.17, mutations were applied to the nucleotides that surround the predicted -10 and -35 box, with the boxes themselves remaining unaffected. In detail, five bases upstream of the -35 box and further five bases downstream the -10 region were exchanged. Additionally, the 17 nucleotides connecting the two boxes in the *glnS* promoter and the 12 nucleotides in the *lpp* promoter were randomly mutated. Jensen and Hammer observed a range of 400-fold change in activity by randomizing another promoter's spacer sequence between these two conserved boxes without affecting the length[160]. We designed the primers for the PCR in such a way that each modified position could be replaced randomly by any of the four nucleic acid bases. This yielded a theoretical diversity of $1.8*10^{16}$ different *glnS* promoters and $1.8*10^{13}$ *lpp* promoters. The actual diversity achieved in these experiments covered an average of 10^{6} distinct clones.

4.2.1 PylS Library

The *glnS* promoter of the *pylS* gene (pCLA86; this plasmid is described in Ch. 4.4.2) was randomized as described above. We determined a diversity of $5*10^{6}$ different clones for this library. For the separation of inactive clones from active ones the amplified DNA was combined with a reporter plasmid (pCLA4) to select on chloramphenicol (Ch. 3.2.4.1). A dilution of cells was plated that consisted of $5*10^{0}$, $5*10^{1}$, $5*10^{2}$ cells on agar plates with 0 µg/mL Cm and $5*10^{3}$, $5*10^{5}$, $5*10^{7}$ cells on plates with 25, 50 and 100 µg/mL Cm in the absence and presence of BocK. This was done for a clear discrimination of the effects from Cm and to obtain single clones for further investigations, e.g., broader range Cm assays. No growth could be observed on agar plates without Bock.

Cells were collected for the agar plates with BocK and the $5*10^{7}$ cell dilution in order to regain library plasmids from active clones and simultaneously remove reporter plasmids (Ch. 3.2.3.5). We chose the high cell count plates to assure the covering of the library diversity. Plasmid integrity was verified by restriction endonuclease digest (not shown).

As mentioned before, agar plates with low cell counts resulted in single colonies. In order to screen the diversity of the library 90 different clones from plates with 25, 50 and 100 µg/mL Cm were transferred onto fresh agar plates with a broader range of Cm, from 0 to 500 µg/mL, allowing a more precise differentiation of these clones (Figure 4.18).

Referring to Figure 4.18, a tendency of higher resistance to chloramphenicol from left to right could be observed. More cells on the right side of the agar plate survived the Cm concentration of 400 µg/mL in the presence of BocK, but also a Cm concentration of 75 to 100 µg/mL without BocK. This increase in resistance correlated with the amount of chloramphenicol that was existent in the original agar plates. Clones with the reference WT *glnS* promoter already showed a weaker growth at 250 µg/mL.

Next, we addressed changes on the promoters that have positive effects on suppression efficiency. For this purpose, ten out of the 90 single library clones were analyzed by western and northern blots. Since all clones survived a Cm concentration of 250 µg/mL in the presence of BocK, we only selected clones that also survived 400 µg/mL Cm, or those that could not survive low concentrations of Cm without BocK. These were the clones 1D, 2E, 3E, 4F, 6B, 7A, 8G, 10D, 11D and 12H, according to the grid in Figure 4.18. The pBK PylS library plasmid variants were separated from the reporter plasmid (Ch. 3.2.3.5; not shown).

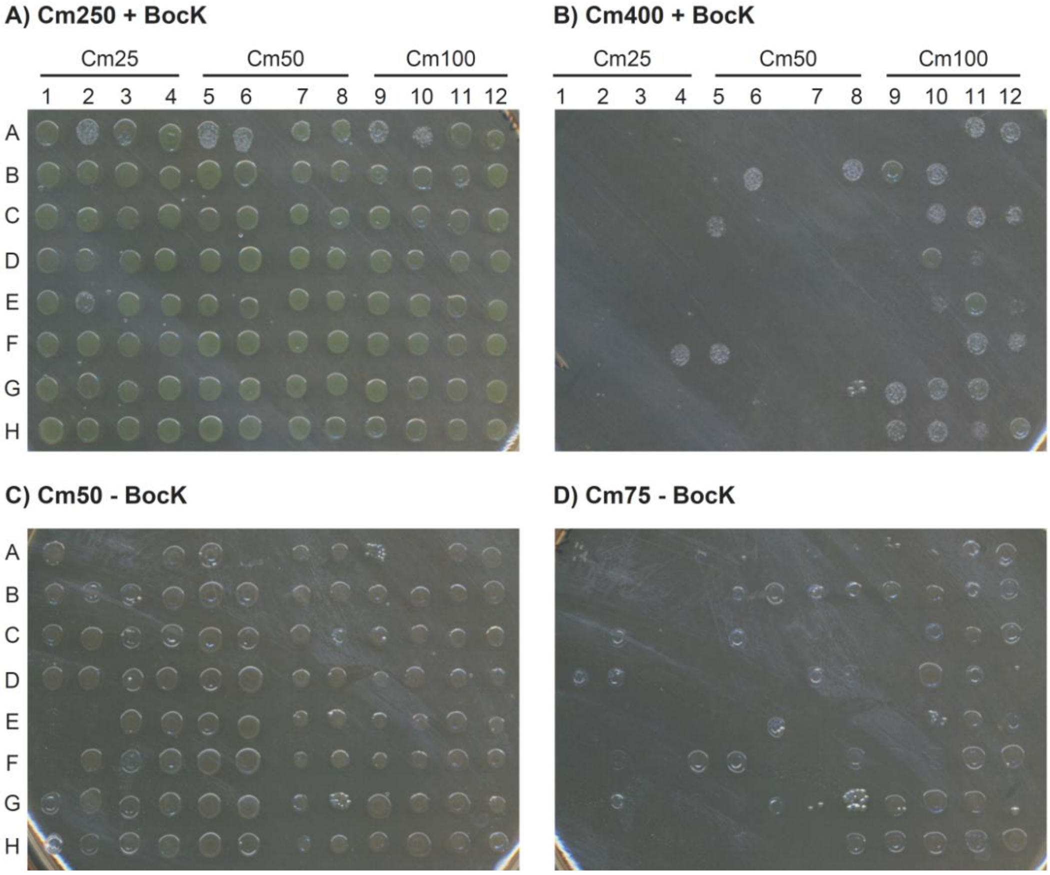


Figure 4.18: Cm-Assay with single colonies from PylS library.
Cells were plated on agar plates containing Kan, Tet, increasing Cm and 1 mM BocK (**A** and **B**) or no BocK (**C** and **D**). A selection of two plates each is shown only. The clones 1A, 5A and 9A contained pCLA9 as a reference and clones 2A, 6A and 10A harbored pCLA86 as a reference. Clones in columns 1-4 originated from the library agar plate with 25 µg/mL Cm and 1 mM Bock, columns 5-8 from the plate with 50 µg/mL Cm and 1 mM BocK and columns 9-12 from the plate with 100 µg/mL Cm and 1 mM BocK.

The purified library plasmids (pCLA22 to pCLA31) were sequenced to assure that all selected clones differed in sequence, but exclusively in the promoter region. Changes in the expression rate of PylS should then be due to mutations in the. The sequencing showed that beside a few silent mutations only those parts of the promoters were mutated which were intended to be. A sequence alignment of all ten library clones revealed that none of them had the same sequence (Figure 8.1).

Lastly, we transformed *E. coli* BL21 cells with the isolated library plasmids in combination with a vector that harbored the genes for the cognate tRNA PylT and a His$_6$-tagged histone H3 with an amber stop codon at position R52 (pCLA32). For suppression of the stop codon during the H3 expression the medium was supplemented with BocK. In the end, two samples of each culture were taken with one being used for comparative western blot analysis of expressed PylS and histone H3. The other was used for northern blot to investigate if the library also had an effect on this level of the translational apparatus (Figure 4.19).

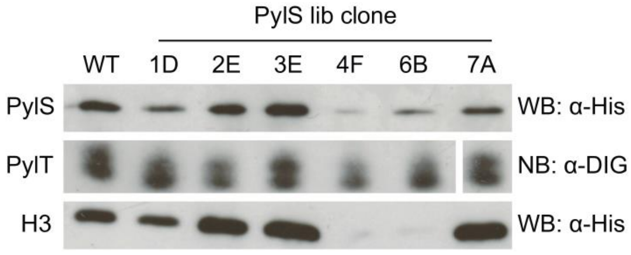


Figure 4.19: Comparative analysis of PylS lib clones via western and northern blot.
For the investigation of the suppression efficiency of the library clones a histone H3 R52TAG (pCLA32) expression was performed. 1 mM Bock was used as UAA and H3 expression was induced with 0.5 mM IPTG at $OD_{600} = 0.8$. A selection of six different library clones (pCLA22 to pCLA27) was compared to WT PylS, which comprised the unchanged *glnS* promoter (pCLA9). All samples were normalized by OD_{600}. For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a nitrocellulose membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody. For the detection of the cognate tRNA PylT total RNA was extracted from a second aliquot (Ch. 3.2.3.14). Approximately 0.090 µg RNA were loaded for each sample. The DIG labeled probe C26 was utilized for hybridization and chemiluminescence signals recorded as described in Ch. 3.2.3.15. The upper band in the second row represents the aminoacylated form of PylT, whereas the lower band depicts the uncharged form.

We found that the PylS promoter library exhibited a variety of different promoter strengths. There were clones, for instance 3E, that had an increased expression efficiency of the synthetase PylS compared to the WT. On the other hand, we also observed clones, such as 4F, which displayed a strong decrease. Furthermore, the more PylS was expressed in the cells, the more PylT in the aminoacylated, slower migrating form, could be observed. Finally, similar to the correlation of synthetase and cognate tRNA, we monitored a clear dependence on the amount of PylS produced and the suppression efficiency, visualized by the amount of histone H3 R52TAG expressed. Hence, it was possible to influence the suppression efficiency of the amber stop codon with the aid of the PylS promoter library with an increase in PylS resulting in higher expression levels of H3.

4.2.2 MjYRS_AGGA Library

The *glns* promoter of the synthetase MjYRS (frameshift codon version = AGGA) was mutated next (pCLA72; this plasmid is described in Ch. 4.4.2.). Unfortunately, the standard touchdown PCR protocol resulted in two PCR products, with the minority having the expected size of approximately 3000 bp but the majority having a size of roughly 1200 bp. This effect was suspected to be due to unspecific primer annealing, despite the touchdown, and needed to be optimized. Therefore, we eliminated the touchdown portion of the PCR, i.e., the first ten cycles, and applied a gradient of increasing annealing temperatures. The gradient covered 54.5 to 66.5 °C, in twelve steps. This modified inverse gradient PCR was set up twice, with one set being supplemented with 5% DMSO, which was thought to reduce the formation of secondary structures of template and primer DNA[161,162].

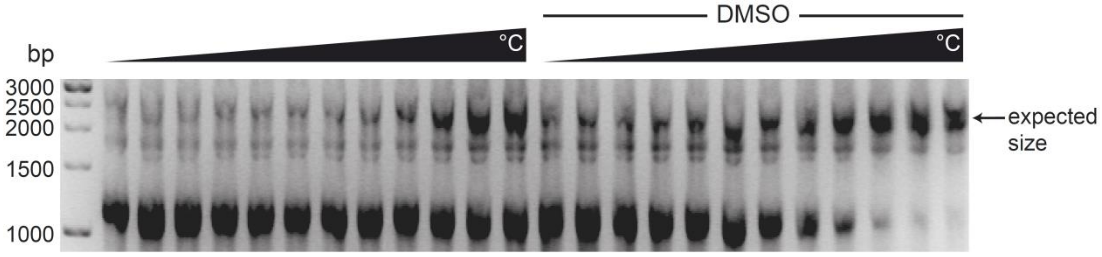


Figure 4.20: Inverse gradient PCR for the MjYRS promoter library.
Two sets were tested with 12 different temperatures each, ranging from 54.5 to 66.5 °C (left to right; indicated by the black triangles). One set was supplemented with 5% DMSO. The expected size of the PCR product is approximately 3000 bp.

An increase of the annealing temperature caused a gain of the desired PCR product, as depicted in Figure 4.20. The addition of DMSO in combination with the high annealing temperature of 66.5 °C finally led to a shift from undesired to the desired PCR product of expected size.

Cells transformed with the product of the PCR using these optimized parameters yielded $1.5*10^6$ different clones. Inactive clones were removed as described above for the PylS library using the Cm reporter (pCLA6). Since MjYRS, in the version described here, decodes for tyrosine, no control plates without the “UAA” could be made. However, a decreasing number of colonies formed on increasing Cm concentrations. The library DNA was again isolated from plates with $5*10^7$ cells and the isolation verified by restriction endonuclease digest (not shown).

The screening of the library diversity was the next step. To this end, we picked 30 different clones each from those single colony plates with 25, 50 and 100 µg/mL of Cm and transferred them onto fresh agar plates with a broader range of Cm, from 0 to 500 µg/mL, (Figure 4.21).

Compared to the PylS library clones no tendency of higher resistance to chloramphenicol from left to right could be noticed. In contrast to PylS, where the reference clones already showed weaker growth at 250 µg/mL Cm, the MjYRS WT promoter clones died more abruptly. None of the selected clones survived the highest concentration of 500 µg/mL.

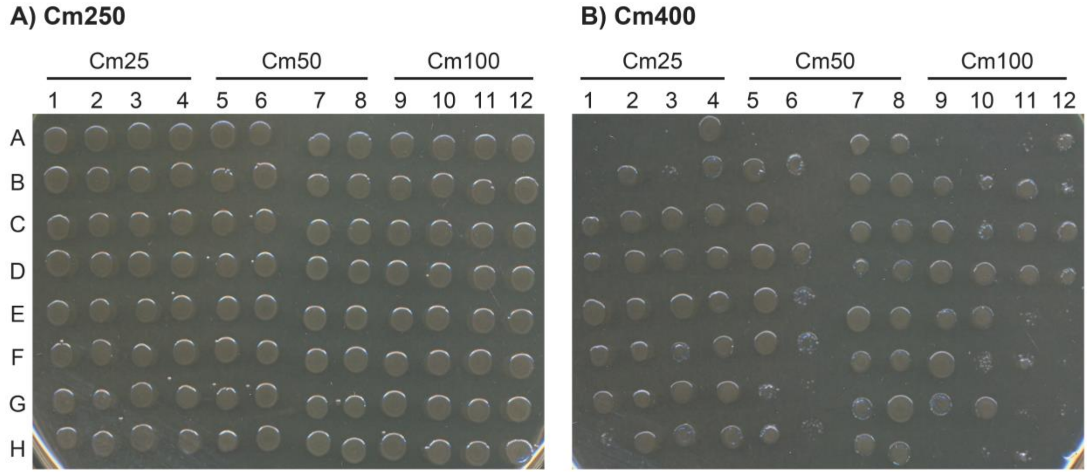


Figure 4.21: Cm-Assay with single colonies from MjYRS_AGGA library.
Cells were plated on agar plates containing Amp, Tet and increasing Cm. A selection of two plates is shown only. The clones 1A, 5A and 9A contained pCLA3 as a reference and clones 2A, 6A and 10A harbored pCLA72 as a reference. Clones in columns 1-4 originated from the library agar plate with 25 µg/mL Cm, columns 5-8 from the plate with 50 µg/mL Cm and columns 9-12 from the plate with 100 µg/mL Cm.

To investigate if the mutagenesis of the *glnS* promoter of the *mjyrS* gene had similar effects on the suppression efficiency as that of the mutant *pylS* promoter, ten clones out of the 90 were chosen. Since all the clones survived a Cm concentration of 250 µg/mL again, we considered how they behaved on the Cm400 plate. The DNA of the colonies 1D, 2F, 3B, 4A, 6D, 7E, 8G, 9F, 10B and 11E, according to the grid in Figure 4.21, was isolated, because these clones represented a range of growth behavior. Thus, they were supposed to be able to repeat the distinct promoter strengths achieved with the PylS library. Therefore, the reporter plasmid was removed as previously described.

Sequencing of the purified library plasmids (pCLA33 to pCLA42) confirmed the intended mutagenesis of the nucleotides surrounding the predicted -10 and -35 box, with the boxes themselves remaining unaffected. A sequence alignment of all ten library clones revealed, that none of them had the same sequence (Figure 8.2).

For the classification of the library promoter strengths myoglobin with a frameshift codon at position S4 (pCLA43) was expressed. For this reason *E. coli* DH10B cells were transformed with the library plasmids pCLA33 to pCLA42 together with pCLA43, a vector that contained the genes for the cognate tRNA MjYT_UCCU and the His_6-tagged myoglobin. Since the synthetase decodes for tyrosine no extra amino acid was added to the medium. Figure 4.22 shows the comparison of the ability to suppress the AGGA codon in case of myoglobin.

Regarding the blot in Figure 4.22 the expression levels of myoglobin fluctuated only in a small range, hardly distinguishable by eye. The quantification using ImageJ revealed a minimal signal intensity of 65% and a maximal intensity of 115% compared to wild type. In contrast to the PylS library (Ch. 4.2.1), the MjYRS_AGGA promoter library seemed to have only a small influence on the suppression efficiency. Correlations between expressed synthetase and the protein of interest were not possible since the aaRS was not detectable on the blots.

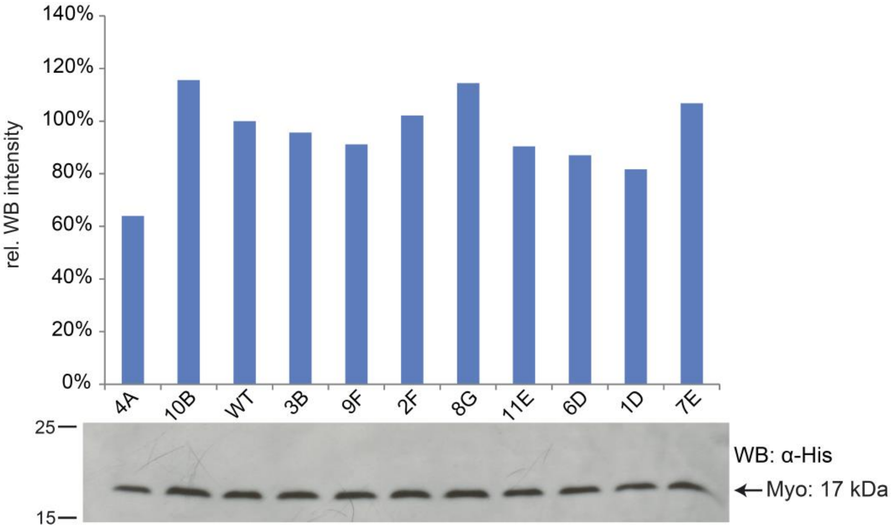


Figure 4.22: Suppression efficiency of MjYRS_AGGA lib clones.
The efficiency was determined by the expression of Myoglobin S4AGGA (pCLA43). Expression was induced with 0.2% arabinose (w/v) at $OD_{600} = 0.8$ and analyzed by western blot (lower panel) and the relative intensities measured by ImageJ (blue bars; upper panel; WT was set to 100%). The ten library clones (pCLA33 to pCLA42) were tested with respect to WT MjYRS_AGGA, which comprised the unchanged *glnS* promoter (pCLA72). For the western blot all samples were normalized by OD_{600}, whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a nitrocellulose membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody.

4.2.3 PylT and MjYT_UCCU Library

The last two promoters that had to be changed were the *lpp* promoters of the *pylT* and *mjYT_UCCU* genes. The modification of the *pylT* promoter (pCLA79; this plasmid is described in Ch. 4.4.2) resulted in a library that held $5*10^5$ different clones. The mutagenesis on the *M. jannaschii* tRNA promoter (pCLA82; this plasmid is described in Ch. 4.4.2) yielded a diversity of $1.2*10^6$ distinct clones.

For the removal of inactive clones via Cm reporter assay the tRNA libraries had to be combined with their cognate aaRS. This is because they were cloned and mutated on the reporter plasmids themselves. The PylT library was used for transformation together with the PylS (pCLA1) and the MjYT_UCCU library together with the MjYRS_AGGA (pCLA3). Cells were plated as before (Ch. 4.2.1) with no extra UAA added to the agar plates for the *M. jannaschii* system and Bock added for the *M. barkeri* pair. The quantity of colonies generally decreased while increasing the Cm concentration. The DNA of both libraries was recovered and checked for correct sequence identity. In the case of PylT a strong contamination with the parental reporter plasmid (pCLA4) was found, which ran higher on an agarose gel (bands at 5625 and 4701 bp) than the mutagenized version (bands at 3928 and 1163 bp) used for making the library (data not shown). Before repeating the whole procedure, the usage of the restriction enzyme NdeI, which only cuts the parental plasmid, helped to clean the PylT library DNA (that still included the inactive clones). Thus, both isolated libraries showed the correct sizes as confirmed by restriction enzyme digest (not shown).

We selected a total number of 96 different clones of each library from the single colony plates with 25, 50 and 100 µg/mL of Cm to transfer them on fresh agar plates for the broad range Cm reporter assay (0 to 500 µg/mL). The results for the PylT library are shown in Figure 4.23 and for the MjYT_UCCU library in Figure 4.24.

The control plates without BocK in Figure 4.23 showed that most of the PylT library clones were not able to survive concentrations of Cm higher than 75 µg/mL and all of them died at 150 µg/mL. In the presence of BocK all clones could resist 250 µg/mL, but only a few colonies in column I withstood 400 µg/mL. These clones originated from the Cm100 single colony plate, a trait already observed for the PylS library, but less distinctive.

Comparable to the MjYRS library, the MjYT_UCCU library clones displayed a wide range of growth efficiencies over the whole plate, independently from the original Cm concentration (Figure 4.24). Most of the clones resisted a Cm concentration of 400 µg/mL, but none of them survived 500 µg/mL. A few colonies became less dense at lower concentrations, e.g., the clones B3 and B6 already at 150 µg/mL (not shown) and the clones H6 and J2 at 250 µg/mL.

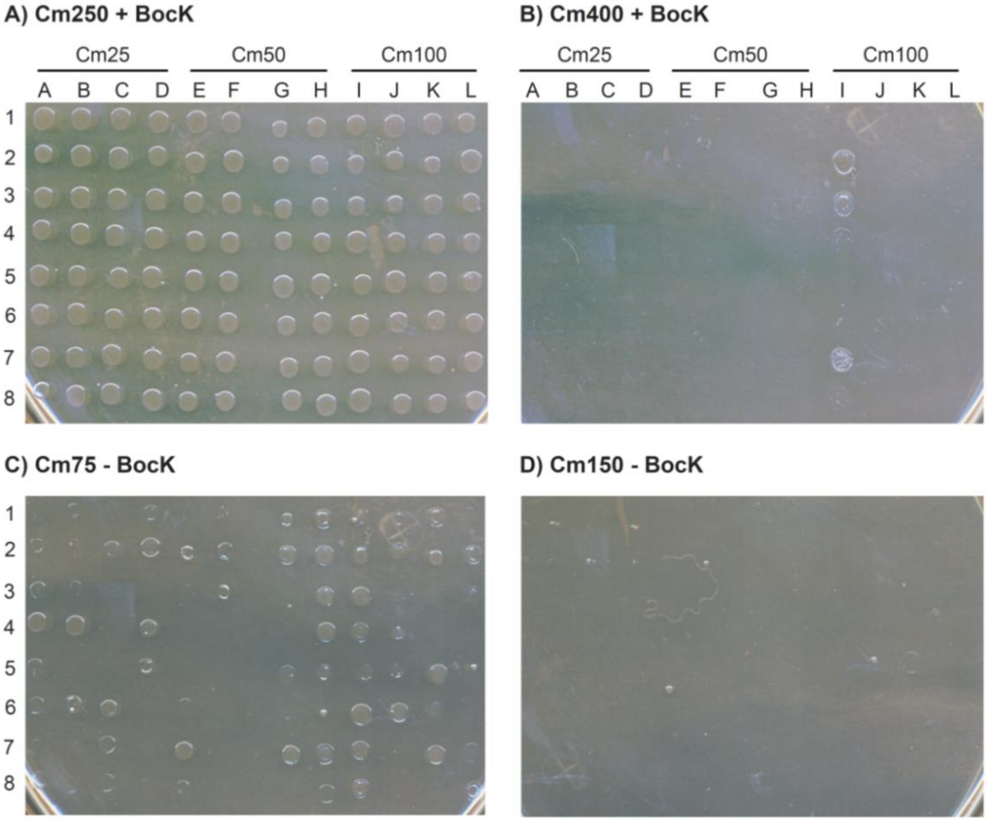

A) Cm250 + BocK
B) Cm400 + BocK
C) Cm75 - BocK
D) Cm150 - BocK
Cm25
Cm50
Cm100
A B C D E F G H I J K L
1
2
3
4
5
6
7
8

Figure 4.23: Cm-Assay with single colonies from PylT library.
Cells were plated on agar plates containing Kan, Tet, increasing Cm and 1 mM BocK (**A** and **B**) or no BocK (**C** and **D**). A selection of two plates each is shown only. Clones in columns A-D originated from the library agar plate with 25 µg/mL Cm and 1 mM Bock, columns E-H from the plate with 50 µg/mL Cm and 1 mM BocK and columns I-L from the plate with 100 µg/mL Cm and 1 mM BocK.

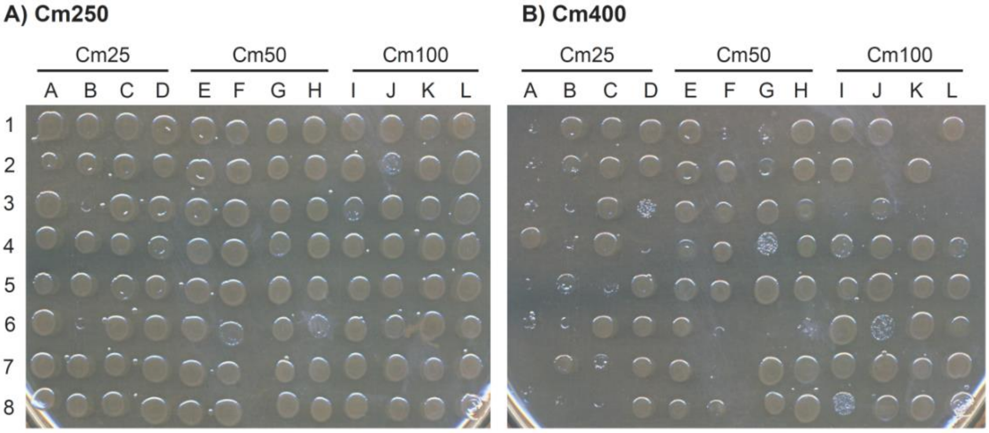

A) Cm250
B) Cm400
Cm25
Cm50
Cm100
A B C D E F G H I J K L
1
2
3
4
5
6
7
8

Figure 4.24: Cm-Assay with single colonies from MjYT_UCCU library.
Cells were plated on agar plates containing Amp, Tet and increasing Cm. A selection of two plates is shown only. Clones in columns A-D originated from the library agar plate with 25 µg/mL Cm, columns E-H from the plate with 50 µg/mL Cm and columns I-L from the plate with 100 µg/mL Cm.

As before, ten out of the 96 clones of each tRNA library were selected to investigate effects on the suppression efficiency, due to changes in the *lpp* promoter, with western and northern blots. The clones A4, C2, E7, F5, H4, I3, I6, J7, K2 and L5, according to the grid in Figure 4.23, were chosen for the PylT library and their DNA separated from the synthetase plasmid (Ch. 3.2.3.5). Full separation was assured by the use of the restriction enzymes NdeI and StuI, which only cut the pBK synthetase plasmids (pCLA1 and pCLA3) but not the modified tRNA containing reporter plasmids (pCLA79 and pCLA82). The same was done for MjYT library clones B3, B8, D4, E1, G4, G7, I2, I8, K6 and L1, according to the grid in Figure 4.24. This procedure failed for PylT clone H4 and for MjYT clone G4, thus only nine clones per library were available (data not shown).

Finally, the sequencing of all purified tRNA library plasmids (pCLA44 to pCLA61) revealed that for both libraries, beside a few minor mutations, only those nucleotides surrounding the predicted -10 and -35 box were altered. The sequence alignment of each library illustrated and confirmed the diversity of the selected clones, with all of them being divergent (Figure 8.3 and Figure 8.4).

In conclusion, promoter libraries for each component of the two tRNA/aaRS pairs originating from the organisms *M. barkeri* and *M. jannaschii* have been prepared. The PylS library revealed a strong dependence of expressed synthetase on the expression of the protein of interest. When more synthetase was expressed it leads to an increased expression level for the desired protein. The MjYRS library did not display such a clear depenence. Individual studies of the tRNA library clones with western and northern blots have not been done, yet. Both libraries were used in combination with the synthetase libraries, as described in a later chapter (Ch. 4.4.4).

4.3 Defined (Standard) Promoters

In chapter 4.2 we found that alterations of the PylS synthetase promoter also effected the suppression efficiency of the amber codon inside a gene of a protein of interest and thereby the level of expression itself. The promoter of library clone 3E was among others the most powerful one and increased the levels of expressed PylS and histone H3 at least twofold (Figure 4.19). Hence, it was known that the promoter was stronger than the wild type *glnS* variant, which was classified as a promoter of moderate strength[163].

Two questions arose then: Is it possible to enhance the performance of the system once more by the usage of promoters characterized as strong, versus weak, and how can the library promoter be categorized in comparison to defined standard promoters like the *E. coli lac* promoter or the lambda (λ) phage promoter? To this end a collection of different promoters was made whose strengths were already evaluated in literature. This set of promoters included endogenous *E. coli* promoters and their derivatives, coliphage promoters, as well as synthetic promoter variants. Their relative strength was mainly specified by so called P_{bla}-units[164]. These units represent the ratio of RNA transcribed from a certain gene under the control of a promoter of interest with the transcripts of the β-lactamase gene (*bla*) under its endogenous promoter (P_{bla}). For instance, the promoter P_{A1} of the phage T7, with 76 ± 9 P_{bla}-units, was the strongest one found by Deuschle[164] and thereby 76-fold more efficient than P_{bla} with 1 P_{bla}-unit. To avoid the need of a DE3 cell background, what means to be independent of the T7 RNA polymerase, no promoters of phage T7 were chosen. A list of the selected promoters together with their P_{bla}-units, if available, is shown in Table 4.1.

The sequences of the compiled promoters were cloned in front of the *pylS* gene (pCLA85; this plasmid is described in Ch. 4.4.2) while exchanging the *glnS* promoter at the same time.

We tested the capacity of these defined standard promoters concerning the suppression efficiency in the same way as we did for the PylS promoter library (Ch. 4.2.1) by expressing a His_6-tagged histone H3 with an amber stop codon at position R52. Thus, *E. coli* BL21 cells were transformed with the relevant pBK plasmids, including WT *glnS* (pCLA9) and library 3E promoter (pCLA24) as references, again in combination with the PylT and H3 R52TAG genes containing pCDF plasmid (pCLA32). The expression was performed with BocK and the addition of IPTG started the synthesis of H3 as well as of PylS if provided with a P_{lac} derived promoter. All other promoters are constitutive and do not need an inducer. The comparative western blot is depicted in Figure 4.25.

Table 4.1: Relative *in vivo* strength of selected promoters.
The strength of the promoters (P) from various origins is given in P_{bla}-units. The sequences were obtained from the mentioned references. No P_{bla}-units were found for $P_{lac1\text{-}6\ Mt5}$ and P_{cp25}. $P_{lac1\text{-}6\ Mt5}$ was described to have a relative transcription efficiency of 112% and is consequently 56-fold stronger than P_{lac} that exhibited only 2%[165]. The synthetic promoter Pcp25 was merely characterized as "quite strong" without showing the data[160]. The strength of P_{λ} found by Deuschle[164] was corrected upwards by Knaus[166].

Origin	Promoter (P)	Relative strength [P_{bla}-units]	Reference
Phage T5	P_{H207}	55 ± 4	Deuschle[164]
	$P_{D/E20}$	56 ± 8	Deuschle[164]
Phage lambda (λ)	P_{λ}	37 ± 7 / 53 ± 8	Deuschle[164]/ Knaus[166]
	$P_{\lambda con/N25DSR}$	65 ± 9	Knaus[166]
E. coli	P_{lac}	5.7 ± 0.5	Deuschle[164]
	P_{tacI}	17 ± 2	Deuschle[164]
	$P_{lac1\text{-}6\ Mt5}$	"56-fold of P_{lac}"	Liu[165]
Synthetic	P_{con}	4 ± 0.2	Deuschle[164]
	P_{cp25}	"quite strong"	Jensen[160]

Our initial observations suggested that the expression of the synthetase PylS worked for all the different promoters tested except P_{cp25}, $P_{\lambda con/N25DSR}$ and P_{H207} (Figure 4.25). For P_{cp25} no signal could be detected, but the $P_{\lambda con/N25DSR}$ promoter samples revealed a relatively intense band smaller than the expected size of 48 kDa probably due to degradation. The P_{H207} lane showed a very weak signal of the same reduced size.

The various promoters were in control of a broad spectrum of produced PylS, ranging from small amounts, in the case of the WT promoter, up to a multiple fold increases of these signals, as in the case of the phage T5 promoter $P_{D/E20}$. The yield of PylS correlated with most of the relative promoter strengths given in Table 4.1. We noticed the lowest abundances for the *glnS* WT promoter, which was characterized as moderate[163] before, and the library promoter P_{lib3E} followed by the consensus sequence based P_{con}, which had the fewest P_{bla}-units of all. P_{lac} was the next strongest candidate referring to P_{bla}-units, what also coincided with the signal intensities on the blot. The modified *lac* promoter variant $P_{lac1\text{-}6\ Mt5}$ seemed to be more powerful than its precursor, but not to that extent as stated by Liu[165]. The phage promoters appeared to be the overall strongest ones, as expected by P_{bla}-units. In previously performed experiments (not shown) $P_{\lambda con/N25DSR}$ already turned out to be the most powerful promoter, according to the 65 P_{bla}-units, assuming that the mass of degraded PylS protein had a high transcription rate as a cause. However, despite its strength this promoter was not successful because no full-length PylS was produced in the end. Regarding the blot, the promoter hybrid P_{tacI} revealed higher PylS intensities as anticipated with respect to P_{bla}-units, arranging between P_{λ} and $P_{D/E20}$.

We then observed a strong correlation of produced PylS and expressed histone H3 R52TAG, similar to the PylS library clones (Figure 4.19). This means, that with an increasing level of PylS the level of histone increased as well. Nevertheless, there seemed to be a certain limit since even a moderate increase of PylS (compare P_{con}, lane 1, and P_{glnS}, lane 10) was already beneficial, while yields of histone H3 were not further improved, if more PylS was expressed. In contrast to the current result P_{lac} has been shown in previous trials (not shown) to mediate histone H3 amounts similar to P_{con}.

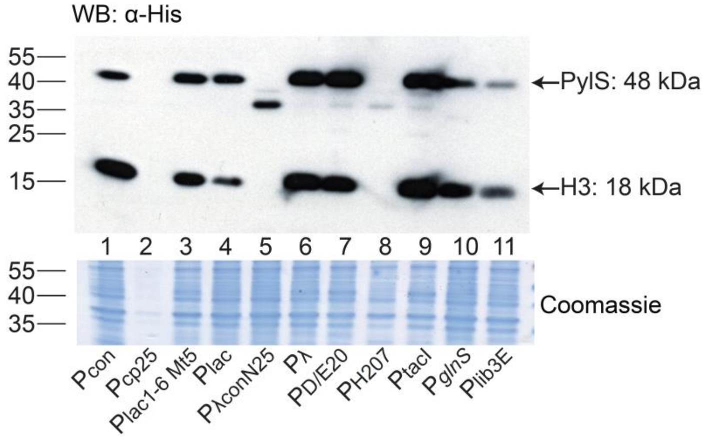


Figure 4.25: Comparative analysis of defined standard promoters with WT *glnS* and library clone 3E promoter. The suppression efficiency in histone H3 R52TAG (pCLA32) expression of systems under the control of defined standard promoters (pCLA62 to pCLA70) was compared to systems where the aaRS PylS was expressed under the control of the WT *glnS* (pCLA9) or library clone 3E promoter (pCLA24). 1 mM Bock was used as UAA and H3 expression was induced with 0.5 mM IPTG at OD_{600} = 0.8 as well as PylS that was dependent on P_{lac} promoter and its derivatives. All samples were normalized by OD_{600}. For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody.

Referring to the two questions at the beginning of this chapter, the library promoter 3E and also the WT *glnS* promoter could be classified as weak, if compared to the power of the other standard promoters. Furthermore, it was possible to significantly improve the suppression efficiency of the given system by using strong promoters. However, we considered P_{con} and P_{lac} to use in later experiments (Ch. 4.4.6) because the former one provided an enhancement of the system comparable to the strongest promoters and the latter one offered the advantage of an inducible promoter, what could be essential if a product is (very) toxic to the host cell.

4.4 The Modular Genetic Tool

4.4.1 The Concept of The Modular Genetic Tool

In chapters 4.2 and 4.3 we tried to optimize the suppression efficiency of amber stop codons or frameshift codons by the use of promoter libraries. A change of the accordant synthetase expression level was intended to have an effect on the expression rate of a protein of interest that contains the nonsense or quadruplet codon in its gene. Several promoters were found, especially for PylS, which improved the suppression efficiency.

Whereas the optimization trials mentioned above focused on the incorporation of a single UAA, the simultaneous insertion of two or more UAAs was the greater goal. For the production of proteins containing multiple UAAs there requires an equivalent number of tRNA/aaRS pairs. These pairs are typically harbored on a single plasmid, where one would need two plasmids to accommodate for the suppression of two unique amino acids. In conjunction with the plasmid for the protein of interest, a minimum of three different plasmids is needed to produce a protein with two UAAs. Three plasmids also imply the use of three distinct antibiotics leading to increased stress levels in host cells.

Moreover, the incorporation of two different UAAs decreases the expression efficiency once more. Neumann and colleagues showed an approach to overcome this deficit by the application of an evolved orthogonal ribosome that efficiently decodes a series of quadruplet codons as well as the amber codon[78]. The disadvantage of this method is the necessity for yet another plasmid that contains the orthogonal ribosomal RNA. In summary, although the expression efficiency is improved, the total count of plasmids and antibiotic stress is increased in host cells.

Hence, another possibility for the optimization of the genetic code expansion system is given by reducing the number of plasmids used. This should be achieved by combining both tRNA/aaRS pairs on a single vector based on pCDF Duet-1 as depicted in Figure 4.26-A. Furthermore, this plasmid should provide the opportunity to exchange each component for other tRNA/aaRS pairs, individually, to introduce new UAAs. Finally, even the replacement of each promoter and terminator sequence of all four components should be available, resulting in a highly modular genetic tool (Figure 4.26-B). The Schultz laboratory pursued a similar approach by combining an amber codon suppressing MjTyr pair and an ochre suppressing pyrrolysine pair on an evolved suppressor plasmid[72]. They were able to efficiently incorporate BocK and *p*-acetyl-L-phenylalanine (pAcF) into GFP albeit GFP levels did not exceed those if both suppression systems were split on two vectors.

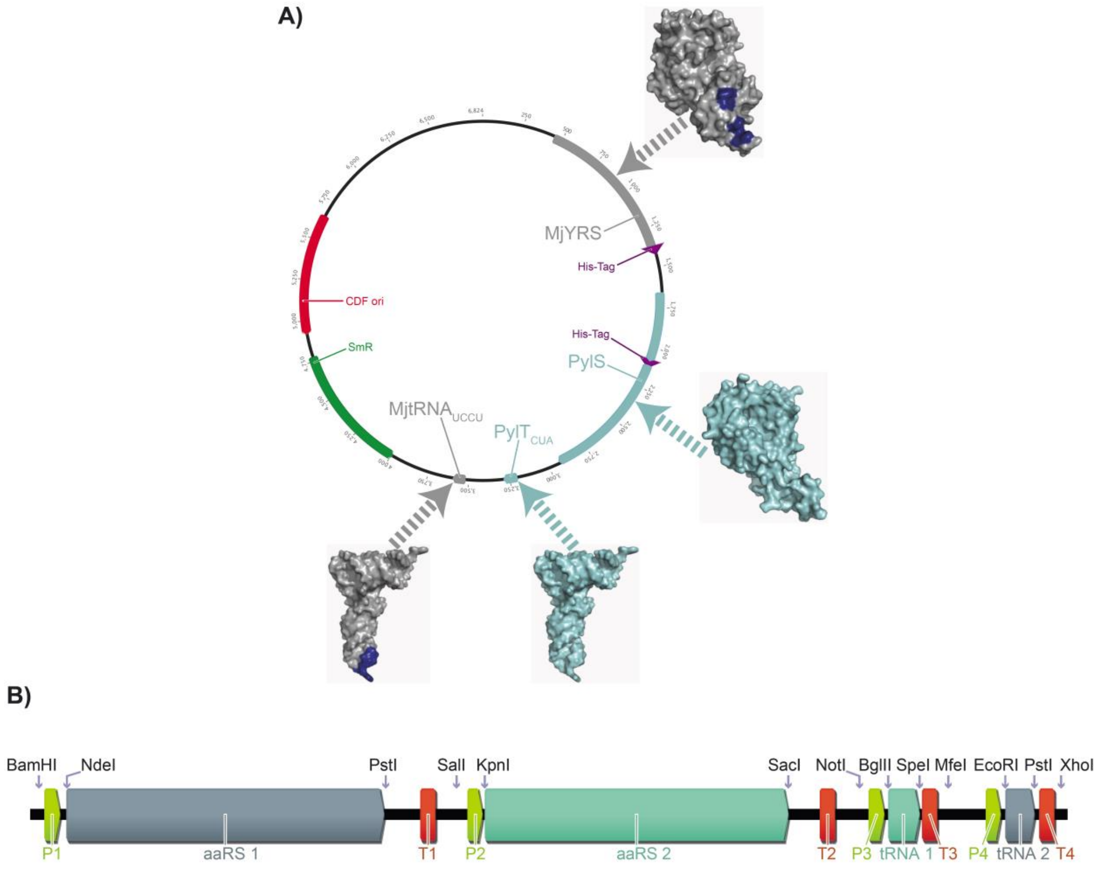

A)
MjYRS
His-Tag
His-Tag
PylS
CDF ori
SmR
MjtRNA UCCU
PylT CUA
B)
BamHI
NdeI
PstI
SalI
KpnI
SacI
NotI
BglII
SpeI
MfeI
EcoRI
PstI
XhoI
P1
aaRS 1
T1
P2
aaRS 2
T2
P3
tRNA 1
T3
P4
tRNA 2
T4

Figure 4.26: Schematic overview of the modular genetic tool.
A) The genes *mjyrS_AGGA*, *pylS*, *pylT* and *mjyT_UCCU* will be combined on a single vector based on pCDF Duet-1 (Novagen). This will reduce the number of plasmids required for the incorporation of two UAAs and the needed antibiotics at the same time by one. **B)** For the construction of the modular genetic tool that allows the replacement of every single component, including their promoters (P1-4; light green) and terminators (T1-4, orange), a series of unique restriction sites is mandatory.

4.4.2 The Construction of The Modular Genetic Tool

The construction of the modular genetic tool started with the incorporation and/or the exchange of the numerous unique restriction sites according to Figure 4.26-B. To this end, a series of twelve different QuikChanges (Ch. 3.2.3.10) was performed that mutated the basic pBK synthetase plasmids pCLA9 (PylS His$_6$-tagged “gap” variant) and pCLA3 (MjYRS_AGGA) as well as the pREP plasmids for the tRNAs PylT (pCLA4) and MjYT_UCCU (pCLA6). Table 4.2 shows a list of all applied QCs including plasmids that were used for mutagenesis. Every cloning step was checked for success via restriction digest using the appropriate enzymes (data not shown) and sequencing data revealed that all twelve QCs were successfully mutated.

Table 4.2: Overview of all QuikChanges performed in order to construct the modular genetic tool.
Twelve different QCs were made in total. The number of the QC is aligned with the used template, the purpose of the QC and the resulting plasmid. Except for pCLA77 and pCLA80 all produced plasmids were sequenced. Details of the construction (primers and sequencing primers) can be found in Table 8.3.

	QC #	Template	Purpose	Product
MjYRS_AGGA	1	pCLA3	C-terminal His$_6$-tag on MjYRS_AGGA	pCLA71
	2	pCLA71	SalI restriction site behind MjYRS terminator	pCLA72
PylS	3	pCLA9	NdeI restriction site in front of PylS start codon changed to KpnI	pCLA73
	4	pCLA73	PstI restriction site behind PylS stop codon changed to SacI	pCLA74
	5	pCLA74	SalI restriction site in front of PylS promoter	pCLA75
	6	pCLA75	NotI restriction site behind PylS terminator	pCLA76
PylT	7	pCLA4	deletion of genes for T7 RNA Polymerase, AraC and GFP	pCLA77
	8	pCLA77	NotI restriction site in front of PylT promoter	pCLA78
	9	pCLA78	MfeI restriction site behind PylT terminator	pCLA79
MjYT_UCCU	10	pCLA6	deletion of genes for T7 RNA Polymerase, AraC and GFP	pCLA80
	11	pCLA80	MfeI restriction site in front of MjYT_UCCU promoter	pCLA81
	12	pCLA81	XhoI restriction site behind MjYT_UCCU terminator	pCLA82

The extent of the mutagenesis in the preparation of the modular genetic tool holds a certain amount of risk because one or more mutations could influence the efficiency of the system negatively. Thus, chloramphenicol reporter assays were made to test the final plasmids pCLA72, pCLA76, pCLA79 and pCLA82 as well as the intermediate steps. Most of the plasmids revealed the same suppression efficiency as the basic plasmid (data not shown), except for PylS (Figure 4.27). Independent of the plasmid used for transformation, the cells were not able to survive Cm concentrations higher than 25 µg/mL. However, cells that harbored the unchanged basic plasmid (pCLA9) were still able to grow on agar plates with 400 µg/mL chloramphenicol.

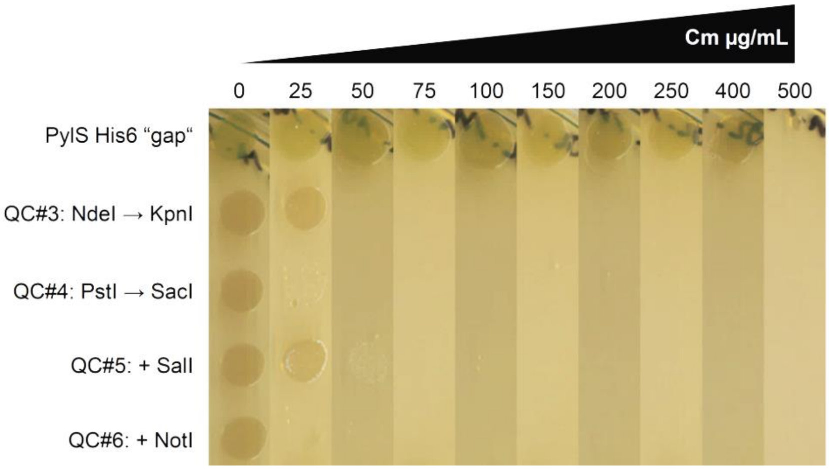

Cm µg/mL
0
25
50
75
100
150
200
250
400
500
PylS His6 "gap"
QC#3: NdeI → KpnI
QC#4: PstI → SacI
QC#5: + SalI
QC#6: + NotI

Figure 4.27: Cm-Assay with final pBK PylS plasmid and intermediate steps from preparation for the modular genetic tool.
The catalytic activity of PylS His_6 "gap" on the basic plasmid (pCLA9) was compared to PylS on the mutated plasmids pCLA73 to pCLA76 from QC#3-6 (Table 4.2). Therefore a chloramphenicol reporter assay (Ch. 3.2.4.1) was performed and cells were plated on agar plates containing Kan, Tet, increasing Cm and 1 mM BocK. Only cells with the basic plasmid were able to survive Cm concentrations higher than 25 µg/mL.

Since the activity already decreased after the first QC that exchanged the NdeI restriction with KpnI and the resulting plasmid served as precursor for the following QCs, we needed to determine if the catalytic activity of PylS itself was disturbed or if there was another cause for this effect. For this reason QC#5 and #6 (Table 4.2) were repeated using the basic plasmid (pCLA9) as template this time (QC#5B: pCLA83 and #6B: pCLA84), in order to incorporate the corresponding sites in a different order. Additionally, two further clones of QC#3 were tested in a subsequent Cm reporter assay (Figure 4.28).

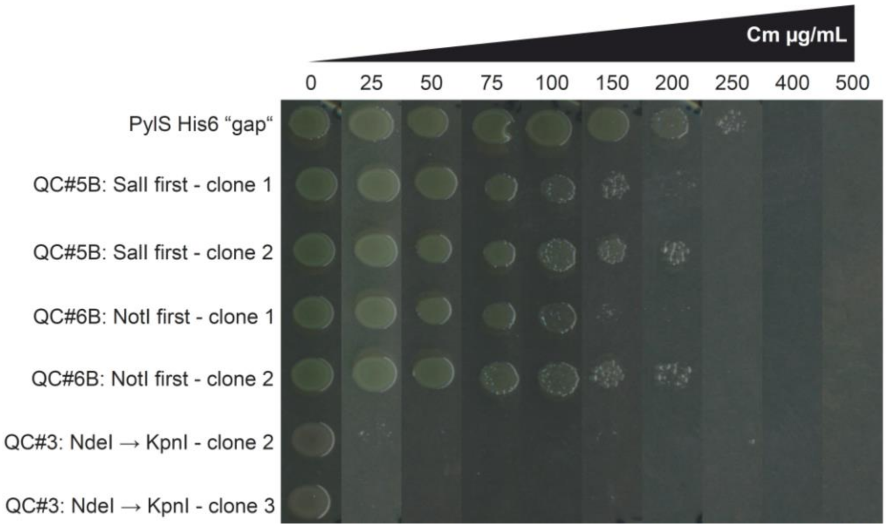

Cm µg/mL
0
25
50
75
100
150
200
250
400
500
PylS His6 "gap"
QC#5B: SalI first - clone 1
QC#5B: SalI first - clone 2
QC#6B: NotI first - clone 1
QC#6B: NotI first - clone 2
QC#3: NdeI → KpnI - clone 2
QC#3: NdeI → KpnI - clone 3

Figure 4.28: Cm-Assay to detect the stage responsible for PylS activity loss.
The catalytic activity of PylS His_6 "gap" on the basic plasmid (pCLA9) was compared to PylS on the mutated plasmids pCLA73, pCLA83 and pCLA84 from QC#3, #5B and #6B (two clones each). Therefore a chloramphenicol reporter assay (Ch. 3.2.4.1) was performed and cells were plated on agar plates containing Kan, Tet, increasing Cm and 1 mM BocK. Only cells with the plasmid from QC#3 were not able to survive Cm concentrations of 25 µg/mL.

It could be stated that the incorporation of the restriction sites SalI and NotI alone did not perturb the suppression efficiency of PylS as depicted in Figure 4.28. In contrast, the two other clones of QC#3 tested, showed the same phenotype as before. Hence, the replacement of the NdeI site upstream of the PylS start codon by KpnI seemed to be the reason for the PylS activity loss.

It was necessary to exchange, or at least to destroy, the NdeI site at the PylS gene for the construction of the modular genetic tool because the gene for MjYRS_AGGA also possesses an NdeI site. We then asked two questions. First, is it possible to restore the PylS activity on pCLA76 by mutating the KpnI site back to NdeI? Second, is it possible to restore the activity and destroy the NdeI site simultaneously? A series of six different QCs was applied to pCLA76, one to restore the NdeI site and the other ones to restore the site while introducing point mutations (pCLA85 to pCLA90). The NdeI-restore mutants were again screened via Cm reporter assay (Figure 4.29).

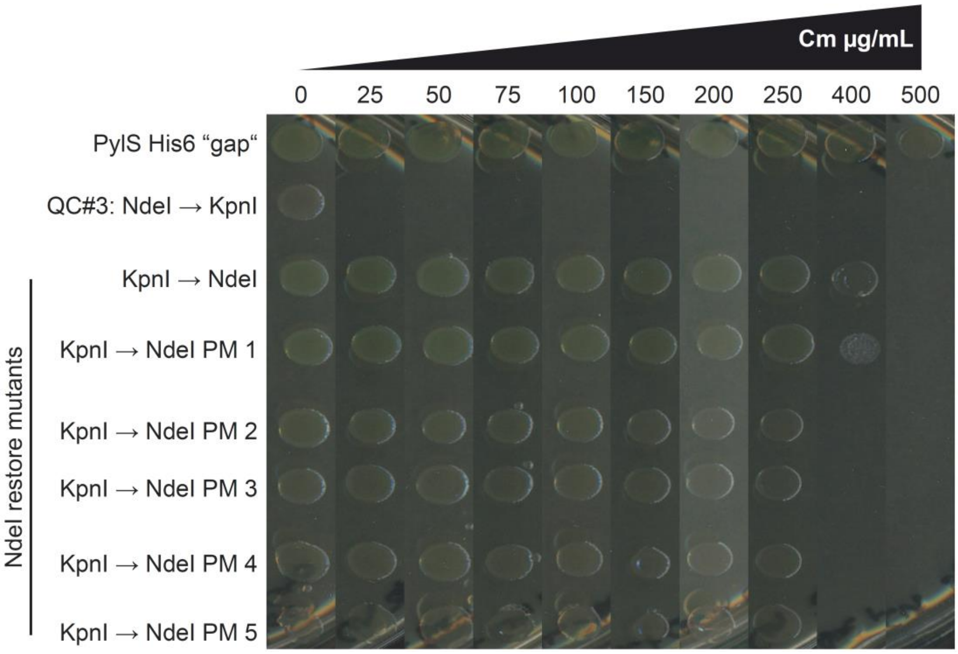


Figure 4.29: Cm-Assay of the PylS NdeI restore mutants.
The catalytic activity of PylS His$_6$ “gap” on the basic plasmid (pCLA9) was compared to PylS on the mutated plasmids pCLA73 (QC#3), pCLA85 (NdeI restore) and pCLA86 to pCLA90 (NdeI restore Point Mutations (PM) 1-5). Therefore a chloramphenicol reporter assay (Ch. 3.2.4.1) was performed and cells were plated on agar plates containing Kan, Tet, increasing Cm and 1 mM BocK. Only cells with the plasmid from QC#3 were not able to survive Cm concentrations of 25 µg/mL.

The activity of PylS could be almost restored by reinstalling the NdeI site in front of the PylS start codon. The PylS activity of the point mutation variants of NdeI was similar to the fully reconstituted NdeI site, as shown in Figure 4.29. Moreover, the point mutated NdeI sites were not digested by the restriction enzyme NdeI any more (data not shown). Due to the highest PylS activity of all point mutation clones, pCLA86 (NdeI-restore_PM1) was chosen for further experiments.

Since all needed restriction sites were cloned into the basic pBK and pREP vectors, according to Figure 4.26-B and Table 4.2, we performed digests in a preparative scale (Ch. 3.2.3.2; Table 4.3) to cut out the genes for both tRNA/aaRS pairs and combined it on a pCDF Duet-1 backbone (Novagen; pCLA91).

Table 4.3: Overview of preparative digests performed in order to construct the modular genetic tool.
Plasmids for inserts and backbones were digested with the same enzymes. The insert is equal to the component mentioned in the table. The plasmid pCLA86 had to be cut by three restriction enzymes to assure the full separation of backbone and insert on a subsequent agarose gel. Cutting with SalI and NotI only would yield bands with sizes of 1555 and 1707 bp, which are difficult to excise from gel precisely.

Plasmid		Enzymes	Component	Product
Insert	Backbone			
pCLA72	pCLA91	BamHI + SalI	MjYRS_AGGA	pCLA92
pCLA86	pCLA92	SalI + NotI (+ XhoI)	PylS	pCLA93
pCLA79	pCLA93	NotI + MfeI	PylT	pCLA94
pCLA82	pCLA94	MfeI + XhoI	MjYT_UCCU	pCLA95

All four components could be successfully transferred into the pCDF Duet-1 vector. Plasmid integrity was verified by restriction enzyme digests (Figure 4.30) and sequencing.

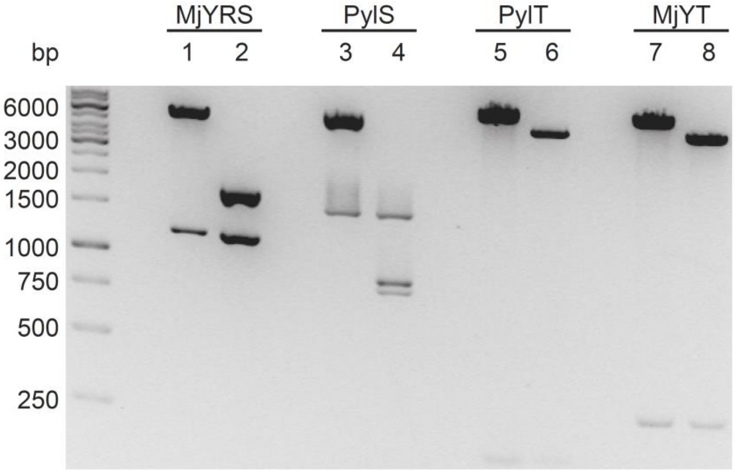


Figure 4.30: Restriction digests of the final modular genetic tool.
The final plasmid pCLA95 that contains all four components (digests 1, 3, 5 and 7) was compared to the insert plasmids (Table 4.3) pCLA72 (digest 2), pCLA86 (digest 4), pCLA79 (digest 6) and pCLA82 (digest 8). Enzymes and sizes: BamHI and SalI (1 = 5618 and 1206; 2 = 1820 and 1206); SalI and NotI (3 =5269 and 1555; 4 = 1555, 890 and 817); NotI and MfeI (5 = 6612 and 212; 6 = 4879 and 212); MfeI and XhoI (7 = 6506 and 318; 8 = 4776 and 318).

4.4.3 Monitoring The Power of The Modular Genetic Tool

In order to determine if the modular genetic tool had benefits compared to the classical system using four different plasmids, it had to be tested using a reporter system. Therefore, we established a second reporter in parallel to the Cm assay – the GFP reporter assay (Ch. 3.2.4.2). This assay exists in two variants, with variant 1 consisting of three plasmids, namely the modular genetic tool (pCLA95), a plasmid for the orthogonal ribosome ribo-Q1[78] (pCLA96) and finally one out of four different GFP reporter plasmids (pCLA105 to pCLA108). The second variant only differed by the separation of the four components of the genetic tool onto two plasmids, namely pCDF PylST[78] (pCLA97) and pDULE MjYRST (this version decodes for UCCU and is based on the CUA version[68,83]; pCLA98).

For the construction of the GFP reporter a plasmid made by Wang and others, called pO-*gst-malE*[67] (pCLA99), was taken as a basis. This vector contains a mutated Shine-Dalgarno sequence that is only bound by orthogonal ribosomes (ORBS = orthogonal ribosomal binding site), in this case the ribo-Q1 provided on pCLA96. As a first step the gene for the fusion protein GST-MBP was removed, except for the first five triplet codons to assure that the binding of the orthogonal ribosome is not perturbed in any way. Simultaneously, a small multiple cloning site (MCS) was introduced, including the unique restriction sites for NcoI, BamHI, NotI and HindIII. The latter one was embedded in the final ochre stop codon of the open reading frame. The plasmid (pCLA100) was controlled by restriction digest and sequencing. Next, we amplified the actual reporter genes encoding for superfolder GFP[167] from pBAD sfGFP vectors, made by Stokes *et al.*[168], including four different genes for a wild type version, a gene with an amber codon at position N150, a frameshift codon at position D134 or both substitutions on one gene at the same time. All versions contained a C-terminal His$_6$-tag and the former restriction site, XhoI, at the end of the gene, was changed to HindIII during the PCR, allowing the cloning into the ORBS reporter pCLA100 subsequent to restriction digest (pCLA105 to pCLA108).

Four sets of electro competent cells were prepared (Ch. 3.2.1.2), with all of them containing the plasmid for the orthogonal ribosome ribo-Q1 (pCLA96) and one out of the four different ORBS GFP reporter plasmids (pCLA105 to pCLA108). For the comparison of the three plasmid system, variant 1, and the four plasmid system, variant 2, these competent cells were transformed with appropriate plasmids and grown in medium supplemented with BocK for the suppression of the amber codon. The suppression efficiency of amber and/or frameshift codons was measured by the amount of GFP produced. We expected that the better a plasmid combination worked, the higher the concentration of GFP would be. In

Figure 4.31 a comparison of the modular genetic tool to the four plasmid system is shown. Additionally, a variant of the genetic tool was tested harboring a gene for MjYT_UCCU with a mismatch in the anticodon stem (C_{29}-A_{44} instead of C_{29}-G_{44}, pCLA109).

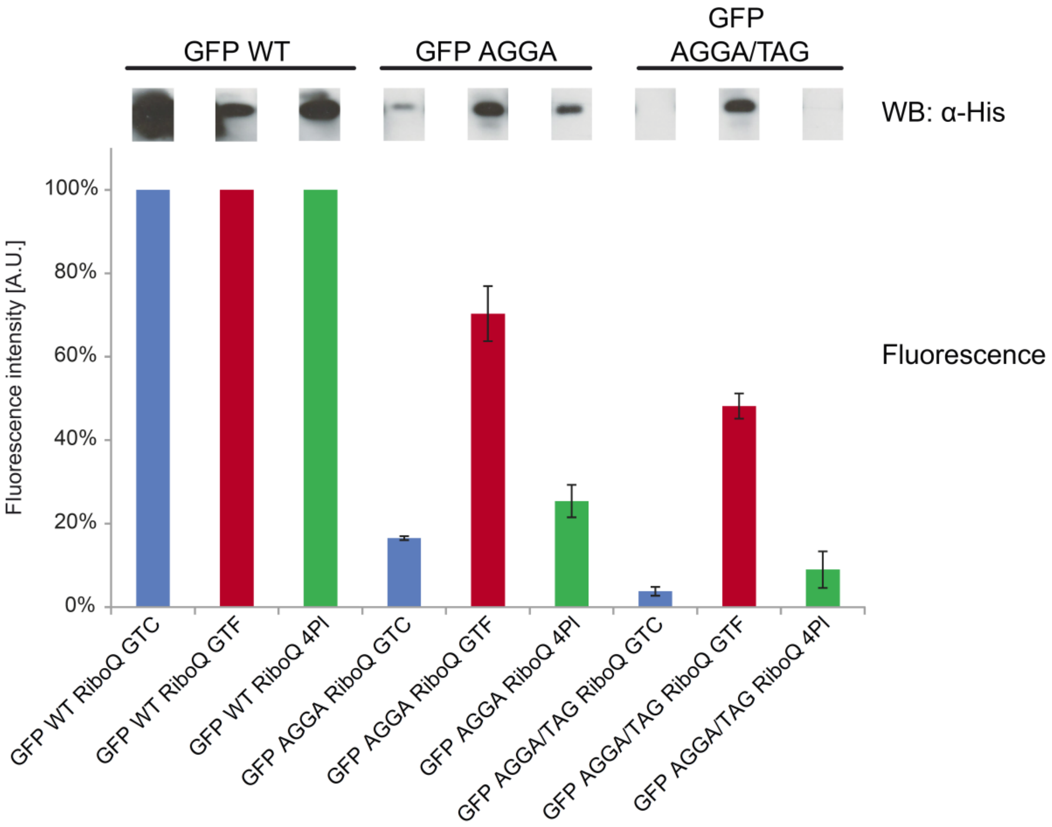


Figure 4.31: Comparison of the modular genetic tool to the four plasmid system.
The expression efficiencies of WT GFP (pCLA105), GFP with a frameshift codon (pCLA108) and GFP with a frameshift and an amber codon (pCLA107) were investigated via fluorescence measurement in a plate reader (lower panel; Ch. 3.2.2.12) and western blot (upper panel; Ch. 3.2.2.4). Three systems were compared: GTC = Genetic Tool Control (mismatch in MjYT_UCCU; pCLA109), GTF = Genetic Tool Final (pCLA95), 4PI = 4 Plasmid system (pCLA97 and pCLA98). All orthogonal mRNA translations were driven by ribo-Q1 (pCLA96). A final concentration of 1 mM Bock was added to the medium. Two distinct samples were taken of each condition tested at different time points (GTC and GTF after 7 and 15 h; 4PI after 72 h (pre-culture), 20 and 26.5 h (main culture) because they grew much slower than variant 1-cells), with one being normalized to OD_{600} of 1.5 for usage with western blots and one being normalized to OD_{600} of 0.5 for the detection of GFP in a plate reader (Ch. 3.2.2.12). For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody. Fluorescence measurements were always performed in duplicates. The mean values of GFP WT were set to 100% for each system individually.

The employment of the modular genetic tool resulted in higher expression levels of GFP compared to the four plasmid system (Figure 4.31). Although the general translation efficiency was similar as shown for GFP WT, the decoding of an AGGA codon performed up to three fold more efficiently. The combination of an AGGA and a UAG codon yielded five times more protein. Results could be confirmed by western blot. Cells with the control mismatch tRNA seemed to be almost completely inactive.

Using the modular genetic tool increased the expression level of the reporter protein GFP significantly. Beside this, we observed that the viability of cells transformed with the genetic tool, and thereby with one plasmid less than those cells transformed with the four plasmid system, was three to five times better (for details see Ch. 4.4.6). Concerning these findings the modular genetic tool already demonstrated a powerful tool that optimized the established system.

4.4.4 The Combination of The Modular Genetic Tool and The Promoter Libraries

The promoter libraries explained in Ch. 4.2, especially for the aaRS PylS, already showed improvements for the incorporation of UAAs into proteins. The positive effect by using the modular genetic tool instead of the four plasmid system was illustrated in Ch. 4.4.3. Consequently, the next step was to test the combination of both optimization trials to obtain an even more efficient genetic tool with enhanced promoters to balance the simultaneous expression of both tRNA/aaRS pairs in the cell.

To this end, large scale preparative digests with 10 to 30 µg of DNA, from backbone and insert vectors, were performed to cover the diversities of the distinct libraries. The same restriction enzymes described in Table 4.3 were used because the library plasmids were based upon the plasmids (pCLA72, pCLA86, pCLA79 and pCLA82) used to construct the modular genetic tool. To avoid contaminations with wild type promoters, the libraries were cloned into a fresh pCDF Duet-1 (pCLA91) vector instead of exchanging the existing WT versions on the final plasmid (pCLA95). The ligation and transformation efficiency was checked by dilution series for each cloning step.

First, we cloned the PylS and PylT promoter libraries into pCDF Duet. The transformations of *E. coli* DH10B via electroporation yielded with 10^6 to 10^7 different clones. However, the cloning of the PylT library revealed a strong background, defined by the number of colonies counted on agar plates with clones from the negative control (ligation of vector backbone and water instead of insert DNA). Since these background clones were thought to be due to re-ligation of the digested backbone vector, we presumed to eliminate them using a modified chloramphenicol reporter harboring the gene for the chloramphenicol acetyltransferase with an amber codon but no tRNA (pCLA110). That means clones should only be able to grow on Cm if both libraries were successfully cloned into the pCDF vector. Hence, electro competent cells containing the modified reporter were prepared and transformed with up to 10 µg of the combined pCDF PylS-PylT library plasmids. Afterwards, cells were selected on chloramphenicol (Figure 4.32).

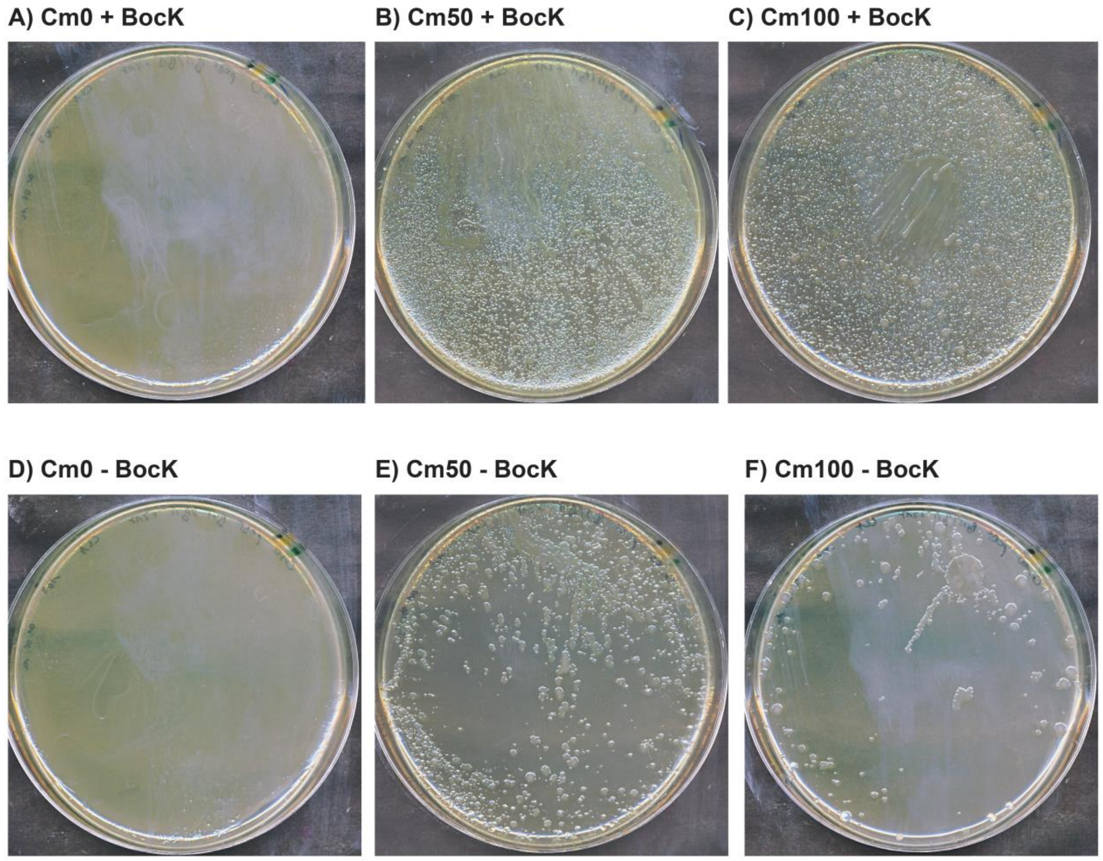


Figure 4.32: Cm-Assay for the background reduction of combined PylS and PylT promoter library.
Cells harboring the delta tRNA Cm reporter (pCLA110) and the pCDF PylS-PylT promoter library plasmids were plated (2 mL cell culture with an OD_{600} of 2.6 each) on agar plates containing Sm, Tet, increasing Cm and 1 mM BocK (**A** to **C**) or no BocK (**D** to **F**).

Assuming that cells without any PylT did not survive the Cm assay, the background reduction for the combination of the PylS and PylT library conferred amino acid dependency, since only a fraction of cells survived without the UAA BocK, as depicted in Figure 4.32. The cells from the agar plate with a Cm concentration of 50 µg/mL and BocK were scraped off for plasmid isolation, as already described in Ch. 4.2.1. The purified pCDF PylS-PylT library plasmids were subsequently used for the cloning of the MjYRS_AGGA and MjYT_UCCU libraries. Both yielded, on average, 10^6 different clones. Thus, all four libraries described in Ch. 4.2 were successfully cloned into pCDF Duet, confirmed by restriction digests in comparison to the wild type modular genetic tool (pCLA95; Figure 4.33).

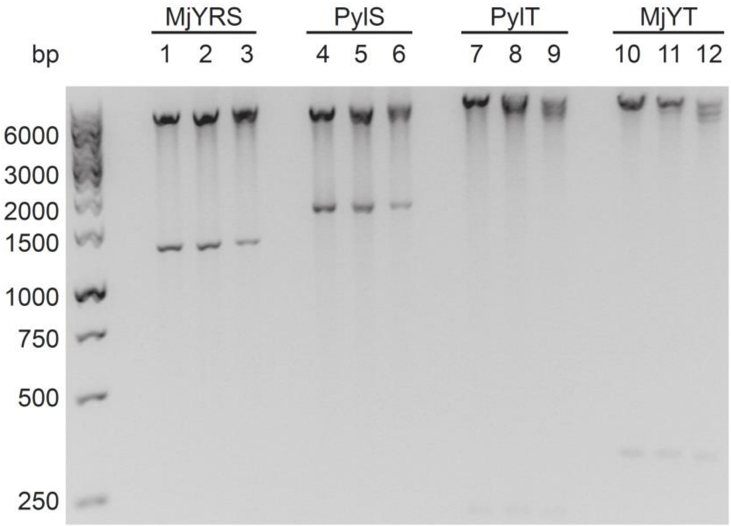


Figure 4.33: Restriction digests of the modular genetic tool with wild type and library promoters.
The final modular genetic tool with wild type *glnS* and *lpp* promoters (pCLA95; digests 1, 4, 7 and 10) was compared to two sets of the tool containing the promoter libraries of all four components (sets differed in diversity only: set1 had $5*10^5$ clones (digests 2, 5, 8 and 11); set2 had $3*10^6$ clones (digests 3, 6, 9 and 12). Enzymes and sizes: BamHI and SalI (13 = 5618 and 1206); SalI and NotI (4-6 =5269 and 1555); NotI and MfeI (7-9 = 6612 and 212); MfeI and XhoI (10-12 = 6506 and 318).

The quantification of the suppression efficiency for the modular genetic library tool was performed by GFP reporter assay (Ch. 3.2.4.2), as previously described for the WT promoter version (Ch. 4.4.3). Therefore, previously prepared competent cells that held the plasmids pCLA96 and pCLA107 for the orthogonal ribosome ribo-Q1 and the ORBS GFP reporter comprising the frameshift codon AGGA and the amber codon UAG, respectively, were transformed with the pCDF PylS-PylT-MjYRS-MjYT library plasmids. The enormous variety of different promoter sequences made it impossible to investigate the GFP expression levels separately via plate reader or western blot. The method of choice used for this purpose was the fluorescence-activated cell sorting (FACS) which was established by the workgroup of Herzenberg[129]. FACS (Ch. 3.2.2.13) was performed at the MRC Laboratory of Molecular Biology, Cambridge. Cells transformed with plasmids pCLA96 and pCLA105, expressing GFP WT, were used as a positive control. In order to set a threshold for the suppression of the AGGA and TAG codon in the GFP gene (pCLA107) that had to be exceeded by more efficient library clones, cells were prepared that contained the modular genetic tool with wild type promoters (pCLA95) as a reference. The results of the FACS are shown in Figure 4.34 in summary.

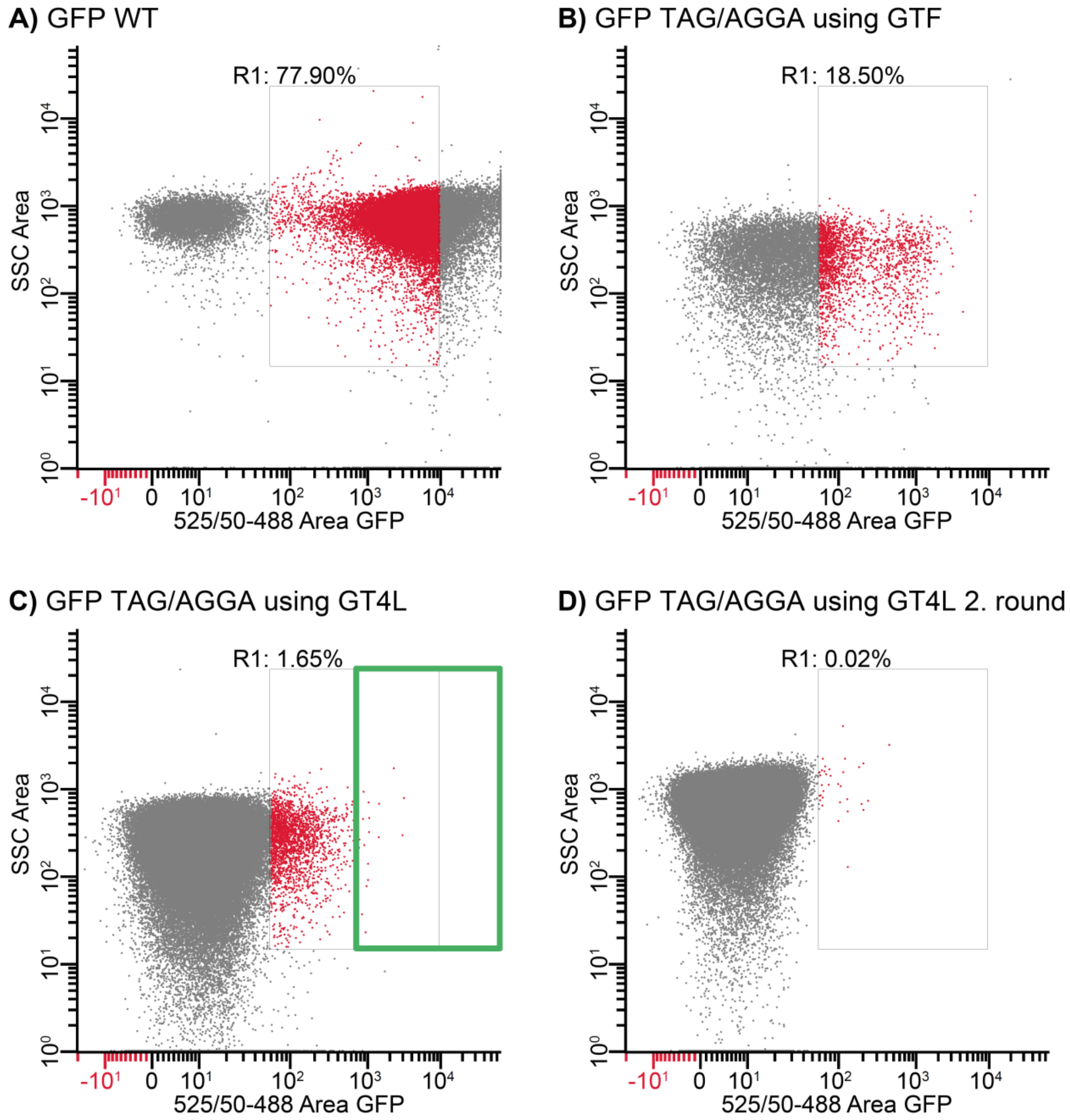


Figure 4.34: FACS to identify the best combination of all 4 libraries on the Genetic Tool at once (GT4L).
A) sfGFP WT (pCLA105) expression to set up the screening area. Positive fluorescence signals are found on the right side of the plot. Each dot illustrates a measurement of a single cell/event by the FACS machine. Cells/events chosen to be collected are found within the rectangle and colored red. R1 represents the relative amount of cells/events within the rectangle compared to the total amount of cells/events tested. **B)** sfGFP harboring a TAG and an AGGA codon in the gene (pCLA107) was expressed using the GTF (Genetic Tool Final (pCLA95) – standard *glnS* (aaRS)/*lpp* (tRNAs) promoter) to identify maximum fluorescence/expression level using the standard system. **C)** Screening for high fluorescence signals (green rectangle) resulting from sfGFP expression using the GT4L plasmids. **D)** Second round of screening cells with GT4L plasmids using isolated cells from C). It seemed the machine sorted out mainly false positive ones, thus no improved GFP signals could be detected.

We did not observe clones from the four library plasmids that displayed enhanced GFP expression compared to the standard WT promoter genetic tool (pCLA95). In Figure 4.34-A the majority of the events, which ideally represent the fluorescence measurements of single cells, are arranged on the right side of the plot. This represents an intense fluorescence signal due to high expression levels of sfGFP. In order to exclude false positive events by virtue of more than one cell measured at the same time, the screening and collecting area (rectangles) was not set to the most right. This setting still included about 78% of all cells tested. In contrast, the expression of the GFP gene with an amber and a frameshift codon by the aid of the standard modular genetic tool (pCLA95) reduced the number of positive clones to roughly 19% (Figure 4.34-B). The majority of these clones was less active, or inactive, indicated by the shift to the left side of the plot. The employment of the four library plasmids (Figure 4.34-C) had an even more dramatic effect on the fluorescence intensities. A minority, of less than 2% of the cells tested, was active and located within the rectangle. To obtain the clones that were able to produce more GFP than those with the standard genetic tool, the screening area was shifted to the right (green rectangle). In the end, performing a second round of FACS with those collected cells (Figure 4.34-D) revealed no improvements.

4.4.5 Alternative Strategies For Screening The Modular Genetic Library Tool

Although FACS is an established method for sorting cells based on their fluorescent characteristic, it was not possible to find any better library composition that exceeded the performance of the standard modular genetic tool. Moreover, a lot of clones seemed to be completely inactive (Figure 4.34). Considering these facts, the search for alternative strategies to screen the four library plasmids led back to the already established chloramphenicol reporter system, which should eliminate inactive clones. To keep the expression system comparable to the GFP reporter used in the last chapters, we made a new Cm reporter including an orthogonal ribosomal binding site based on a reporter produced by Wang[67] (pCLA111). This plasmid contained the gene for the CAT only but not for a tRNA. We exchanged the codons for T6 and D111 in the CAT gene for TAG and AGGA, respectively (pCLA112 and pCLA113).

Similar to the GFP reporter previously prepared competent cells harboring the ORBS CAT reporter plasmid (pCLA113) and the orthogonal ribosome plasmid (pCLA96) were transformed with the four library plasmids. Three library plasmid variants, missing the library for MjYT, and the standard genetic tool (pCLA95) were transformed as negative and positive control, respectively. In addition, we investigated differences between a CAT gene with TAG alone and together with the AGGA codon. After a low stringency selection on chloramphenicol colonies were counted to determine the reliability of the reporter (Table 4.4).

Table 4.4: Cm-Assay for the sorting of the four library plasmids.
Competent cells harboring plasmids for the orthogonal ribosome ribo-Q1 (pCLA96) and either the ORBS CAT reporter with T6TAG alone (pCLA112) or together with D111AGGA (pCLA113) were transformed with the standard modular genetic tool (pCLA95; GTF), the four library plasmids (GT4L) or the three library variants (GT3L; MjYT library missing). In case of the library plasmids up to $5*10^7$ cells were plated and up to 1000 cells were plated for the standard genetic tool on agar plates containing Sm, Tet, increasing Cm in the absence and presence of 1 mM BocK. Plates were incubated at 37 C for two days due to slow growth of the cells. MSC = “many single colonies”

	GTF (standard)				GT4L		GT3L
Cm [µg/mL]	**T6 +BocK**	**T6 -BocK**	**T6/D111 +BocK**	**T6/D111 -BocK**	**T6/D111 +BocK**	**T6/D111 -BocK**	**T6/D111 +BocK**
0	≈960	872	1088	≈864	Lawn	Lawn	Lawn
10	≈788	0	507	0	Lawn	Lawn	37
25	420	0	30	0	MSC	MSC	12
50	2	0	1	0	≈1000	>1000	3
75	0	0	0	0	78	≈500	1
100	0	0	0	0	25	75	0
150	0	0	0	0	9	8	0

The orthogonal CAT reporters (pCLA112 and pCLA113) both worked as demonstrated in Table 4.4 by the quantity of colonies for the standard modular genetic tool. Cells were able to survive treatment with Cm only in the presence of the UAA BocK. The additional AGGA codon in the CAT gene reduced the growth of the bacteria 10-fold concerning the plates with 25 µg/mL Cm. The reporter itself seemed not to be that strong than the pREP reporters used earlier since the cutoff was around 50 µg/mL instead of 250 to 400 µg/mL. The majority of cells lacking the MjYT library died upon Cm treatment, confirming that no other endogenous tRNA/aaRS pairs suppressed the AGGA codon. Considering the four library plasmids, no dependence on the UAA could be observed. Approximately the same numbers of colonies survived on plates with and without BocK.

These findings led us to the assumption that even if inactive clones were eliminated using the chloramphenicol reporter assay, collateral mutations during the mutagenesis of the promoter regions could have caused the use of natural amino acids by the synthetases, especially PylS. Therefore, the four library plasmids from the agar plate with 25 µg/mL Cm and BocK were separated from the orthogonal CAT reporter and ribosome plasmids by destroying the unwanted plasmids with the restriction enzyme SphI and removing them by transformation of *E. coli* DH10B with the intact library plasmids. In order to get rid of PylS synthetases which used natural amino acids as substrate to charge PylT, a negative selection was performed based on suppression of amber nonsense mutations in the barnase gene as described by Wang and Schultz[146]. Without its inhibitor, barstar, it is lethal to cells[46]. That means, synthetases that are using natural amino acids in the absence of the UAA will finally suppress the amber codons in the barnase gene and mediate the cell death. Cells with synthetases that are still orthogonal and only using UAAs will be filtered (Ch. 3.2.4.3). To this end, electro competent cells containing the plasmid with the amber mutated barnase gene (pCLA114) were transformed with the purified four library plasmids as well as the standard modular genetic tool (pCLA95) as a comparison. A total number of $5*10^7$ transformed cells were plated on agar plates with the appropriate antibiotics in the absence and presence of 0.002% arabinose, to induce the expression of the barnase gene, and BocK, as a negative control.

We observed nearly the same growth behavior on all plates regardless of conditions (data not shown). A small tendency of less growth in the presence of UAA BocK could be stated. Thus, we were not able to solve the problem of the amino acid independent suppression of amber codons by using the barnase assay.

A second alternative strategy was then used. This method relied on the direct monitoring of GFP expression via western blot. 96 colonies were picked for this purpose from the remaining agar plates of the ORBS CAT reporter assay mentioned at the beginning of this chapter – including two clones for four library plasmids from plates without BocK and two clones from the three library plasmid plate. These 96 single clones were put again on agar plates with increasing Cm for re-phenotyping. Finally, a selection of 20 clones for the four library plasmids including the two clones from plates without BocK and the two clones for the three library plasmids were further analyzed. Isolated library plasmids were investigated by restriction digest (not shown) and sequencing. The digest revealed that only half of the clones had the correct banding pattern, indicating the absence of the MjYRS library. However, the sequencing confirmed that the majority of mutations were located in the promoter regions, as desired. Next, similar to the procedure described in Ch. 4.4.3 competent cells were prepared that held plasmids for the orthogonal ribosome ribo-Q1 (pCLA96) and the ORBS GFP reporter comprising the frameshift codon AGGA and the amber codon UAG (pCLA107), respectively, were transformed with all 22 plasmids to perform the GFP reporter assay (Ch. 3.2.4.2). The standard modular genetic tool (pCLA95) served as a reference for the GFP expression level. The comparative western blot is shown in Figure 4.35.

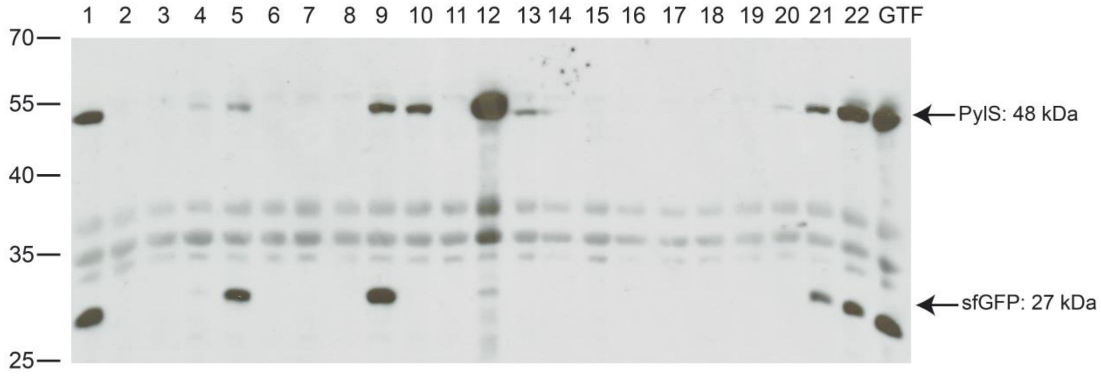

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 GTF
70
55
40
35
25
PylS: 48 kDa
sfGFP: 27 kDa

Figure 4.35: Comparison of the modular genetic tool (GTF) to single four library plasmid clones.
The expression efficiencies of GFP with a frameshift and an amber codon (pCLA107) were investigated via western blot (Ch. 3.2.2.4). The modular genetic tool (GTF; pCLA95) was compared to 22 different library clones. Clone number 5 and 15 originated from four library plasmid plates without 1 mM Bock. Clone number 10 and 20 were picked from three library plasmid plates. The rest of the clones came from four library plasmid agar plates with BocK and different concentrations of Cm. All orthogonal mRNA translations were driven by ribo-Q1 (pCLA96) with 1 mM of BocK to suppress the UAG codon. Samples taken at the end were normalized to OD_{600} of 1.5. For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody.

Figure 4.35 clearly shows that most of the selected clones did not express any sfGFP. Those clones that did were, in the end, not better than the wild type modular genetic tool (GTF; pCLA95). For clone number 10 and 20 no GFP was expected due to the lack of MjYT. We noticed again the correlation of produced PylS and protein of interest observed in earlier chapters (Ch. 4.2.1 and 4.3), with more PylS leading to more GFP in this case. However, it seemed that a strong expression level of PylS caused a negative effect as depicted by clone number 12. This clone revealed the highest levels of expressed PylS but only a very small amount of GFP.

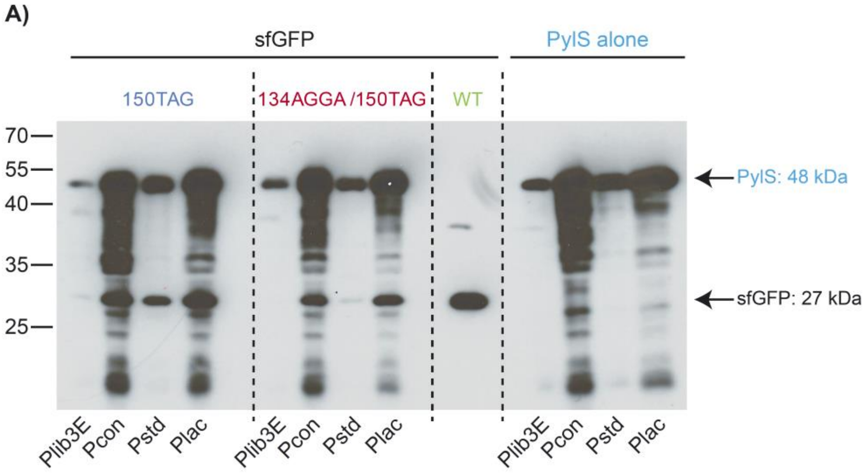
In summary, it was not possible to isolate a clone out of the four component library plasmids that exceeded the standard modular genetic tool’s suppression efficiency for a UAG and AGGA codon at the same time.

4.4.6 The Combination of The Modular Genetic Tool And Defined (Standard) Promoters

Since the combination of the modular genetic tool and the promoter libraries brought no significant improvements in the end (Ch. 4.4.4 and 4.4.5), we thought to incorporate the defined standard promoters introduced in Ch. 4.3 into the tool. The most promising ones were selected, i.e., P_{lib3E}, P_{con} and P_{lac}, and cloned into pCLA95 exchanging the *glnS* promoter thereby.

The effect of the promoter exchanges was tested with the GFP reporter assay in the same way as described in Ch. 4.4.3. GFP expression levels were determined by western blots (Ch. 3.2.2.4) and by in-cell fluorescence of sfGFP in a plate reader (Ch. 3.2.2.12). In a first attempt we found a strong degradation of PylS on western blot (not shown). To exclude that bands at the height of GFP were due to PylS degradation, background controls with the synthetase containing pCDF vectors alone were also analyzed (Figure 4.36).

Both the western blot and the fluorescence measurements in Figure 4.36 revealed that the change of the *glnS* wild type promoter to P_{con} and P_{lac} drastically increased the GFP expression as well for the amber codon only version as for the combined amber and frameshift codon version. In detail, P_{con} was at least 1.5 to 2-fold better and P_{lac} 2.5 to 3-fold compared to the standard promoter in terms of fluorescence measured (Figure 4.36-B). Interestingly, P_{lib3E}, which was shown to be significantly stronger than P_{std} in earlier chapters (Ch. 4.2.1), was found to be weaker after changing the plasmid from pBK to pCDF. It could be excluded that signals at the height of GFP detected on the western blot in Figure 4.36-A relied on PylS degradation because no comparable signals were found in the samples with PylS alone.



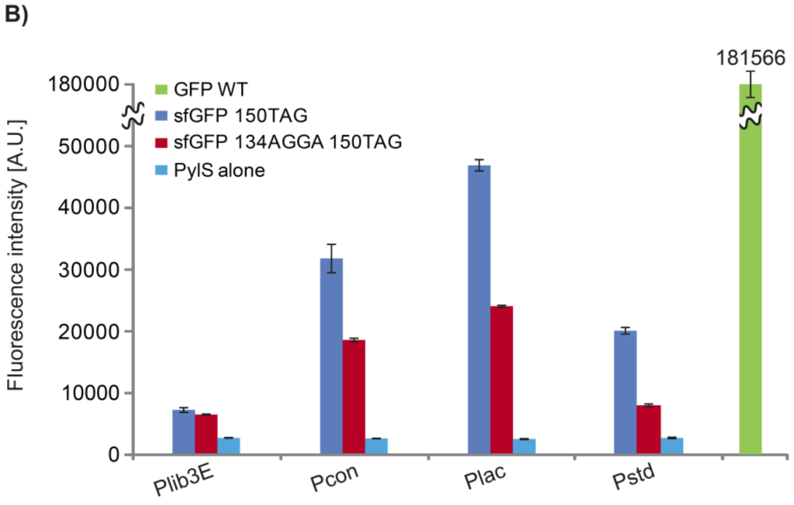


Figure 4.36: Comparison of the standard modular genetic tool to variants with defined promoters.
The expression efficiencies of WT GFP (pCLA105), GFP with an amber codon (pCLA106) and GFP with a frameshift and an amber codon (pCLA107) were investigated **A)** via western blot (Ch. 3.2.2.4) and **B)** via fluorescence measurement in a plate reader (Ch. 3.2.2.12). Four variants of the modular genetic tool were compared, only differing in the promoter in front of the *pylS* gene: P_{std} = wild type *glnS* promoter on pCLA95; P_{lib3E} = PylS library promoter 3E on pCLA115; P_{con} = consensus promoter on pCLA116; P_{lac} = *lac* promoter on pCLA117. All orthogonal mRNA translations were driven by ribo-Q1 (pCLA96). *E. coli* DH10B transformed with pCLA95, pCLA115, pCLA116 and pCLA117 only, were used as a background control. A final concentration of 1 mM Bock was added to the growth medium for the suppression of the amber codon. In case of the *lac* promoter, expression was induced with 1 mM IPTG Two distinct samples were taken of each condition tested, with one being normalized to OD_{600} of 1.5 for usage with western blots and one being normalized to OD_{600} of 0.5 for the detection of GFP in a plate reader (Ch. 3.2.2.12). For the western blot whole cell extracts were separated with SDS PAGE (10 µL loaded for each sample, except for GFP WT with 5 µL; Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody. Fluorescence measurements were always performed in duplicates.

As a proof of principle, to demonstrate that the effects on the suppression efficiencies observed for GFP due to the changes in the PylS promoter were also valid for other proteins, we did a final experiment comparing P_{std}, P_{con}, P_{lac} and the previously used four plasmid system (with standard *glnS* and *lpp* promoters). The GFP reporter plasmid was exchanged with a variant of its original vector pO-*gst-malE*[67] which harbored an AGGA codon at position Y17 and a TAG codon at position N234 in the *malE* gene (pCLA118). We kept the conditions for the expression the same, except for the scale. A 50 mL small scale expression was performed and the GST-MBP fusion proteins were purified from cell lysates (Ch. 3.2.2.6) to visualize differences in protein levels by SDS-PAGE instead of a western blot (Figure 4.37).

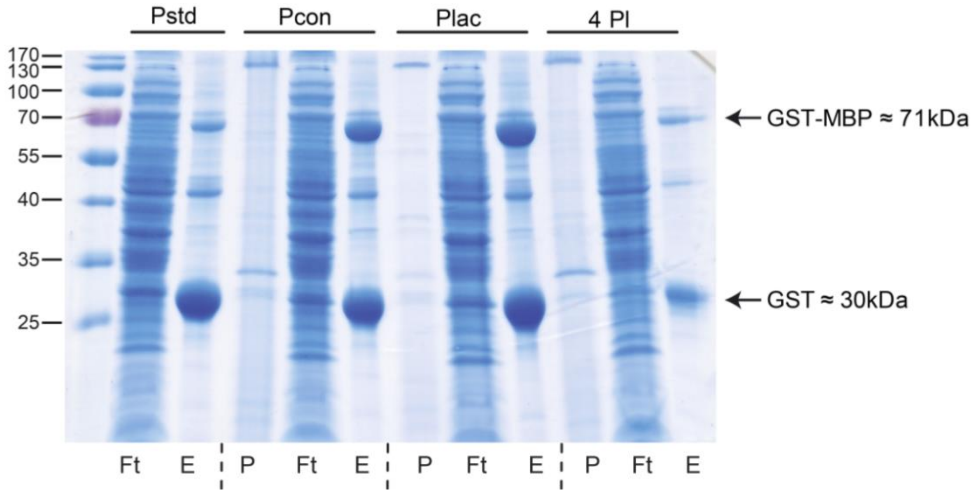


Figure 4.37: Small scale expression and purification of GST-MBP comparing different genetic tool variants and the 4 plasmid system.
The expression efficiency of *gst-malE* harboring an AGGA codon at position Y17 and a TAG codon at position N234 in the *malE* gene (pCLA118) was compared in four different systems: P_{std} = wild type *glnS* promoter on pCLA95; P_{con} = consensus promoter on pCLA116; P_{lac} = *lac* promoter on pCLA117; 4Pl = 4 Plasmid system (pCLA97 and pCLA98). All orthogonal mRNA translations were driven by ribo-Q1 (pCLA96). A final concentration of 1 mM Bock was added to the growth medium for the suppression of the amber codon. Expression was performed in a 50 mL scale and cell densities normalized by OD_{600} after pelleting. GST-MBP fusion proteins were purified from cell lysates using glutathione sepharose beads (Ch. 3.2.2.6).Proteins were separated with SDS PAGE (15 µL loaded for each sample; Ch. 3.2.2.3). Ft = Flowthrough; E = Elution; P = Pellet.

The small scale expression and purification experiment shown in Figure 4.37 validated our previous experiments. The four plasmid system offered the lowest expression and thereby suppression efficiency of all the conditions tested. The use of the standard modular genetic tool increased the full-length GST-MBP level substantially. However, when the *glnS* promoter of the *pylS* gene is exchanged for P_{con} or P_{lac} the yield of full-length GST-MBP was again boosted enormously.

Beside the fact that cells with the four plasmid system had the lowest expression efficiencies, another observation already showed that their growth rate was at least three to five times lower compared to the three plasmid system with the modular genetic tool variants (Ch. 4.4.3). To visualize these differences we made a growth curve with the three modular genetic tool variants P_{std}, P_{con}, P_{lac} and the four plasmid system. For this reason, competent cells that harbored the plasmids for the orthogonal ribosome ribo-Q1 (pCLA96) and the ORBS GFP reporter, comprising the frameshift codon AGGA and the amber codon UAG (pCLA107), respectively, were transformed with the appropriate plasmids in order to monitor the growth rate and the production of GFP, simultaneously, by using a plate reader. The results of the growth assay and the fluorescence measurements are plotted in one graph (Figure 4.38).

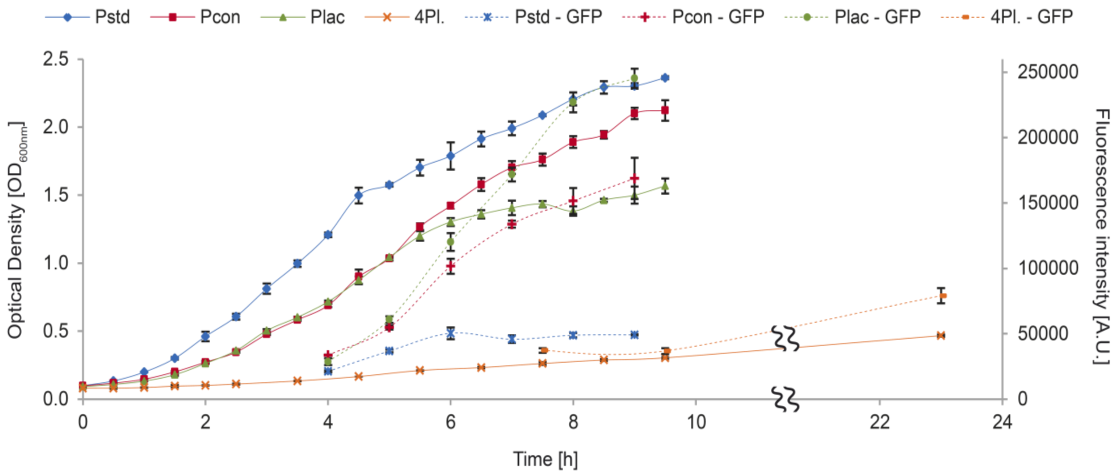


Figure 4.38: Growth curve and GFP fluorescence measurement comparing different genetic tool variants and the 4 plasmid system.

The growth rates and the expression efficiencies of sfGFP harboring an AGGA codon at position D134 and a UAG codon at position N150 (pCLA107) was compared in four different systems: P_{std} = wild type *glnS* promoter on pCLA95; P_{con} = consensus promoter on pCLA116; P_{lac} = *lac* promoter on pCLA117; 4Pl = 4 Plasmid system (pCLA97 and pCLA98). All orthogonal mRNA translations were driven by ribo-Q1 (pCLA96). Expression was performed in a 50 mL scale. Full-length GFP expression was initiated by the addition of the UAA BocK to a final concentration of 1 mM after 3 hours for P_{std}, P_{con} and P_{lac} and after 3.5 hours for 4Pl due to the slow growth. A final concentration of 1 mM IPTG was supplemented for P_{lac} as well. Samples for fluorescence measurements of P_{std}, P_{con} and P_{lac} were taken every hour upon UAA addition corresponding to 4, 5, 6, 7, 8 and 9 hours and normalized to OD_{600} of 0.5 for the detection of GFP in a plate reader (Ch. 3.2.2.12). Samples for 4Pl were taken after 4, 6 and 20 hours upon UAA addition corresponding to 7.5, 9.5 and 23.5 hours. OD_{600} was measured every 30 min for P_{std}, P_{con} and P_{lac} over a period of 9.5 hours. OD_{600} for 4Pl was measured every 30 min for the first 2.5 hours, every 60 min for the next 7 hours and finally one sample was taken after overnight incubation at 23.5 hours. All samples were taken in triplicates from three individual cultures.

The growth assay depicted in Figure 4.38 confirmed the observations made before. The viability of cells using the four plasmid system is several magnitudes lower than the cells with one of the three plasmid systems P_{std}, P_{con} or P_{lac}. The fluorescence intensity was indeed higher than for P_{std} but this was only achieved after overnight incubation. When comparing the three plasmid variants, P_{std} had the fastest growth rates but also the lowest fluorescence intensities with a maximum already reached after 2 hours. Cells with the P_{con} modular genetic tool were almost as viable as P_{std}, with a medium to high expression of GFP. P_{lac} mediated a medium growth rate but produced, by far, the highest fluorescence intensities.To summarize, the use of the modular genetic tool improved the suppression efficiencies of amber and frameshift codons in protein expressions performed with the aid of the orthogonal ribosome ribo-Q1. The introduction of P_{con} and P_{lac} further increased this effect, with having one constitutively active and one inducible promoter for different purposes. Furthermore, the viability of the cells could be drastically enhanced by the reduction of plasmids from four to three, leading to better results in shorter expression times.

5 Applications And Side Projects

5.1 Evolution of A Quadruplet Decoding MjAzFRS

A large variety of UAAs with different functionalities can be incorporated into proteins facilitated by the expansion of the genetic code. All previous experiments described in Ch. 4.1 to 4.4 were performed using the efficient substrate analogue BocK[169] and/or tyrosine. Both amino acids do not provide any special functions because the Boc-protected lysine is not useful for further modifications and the tyrosine can also be incorporated using endogenous codons and tRNA/aaRS pairs. In contrast, aryl-azide, a photocrosslinking agent that was used first in a biological context by Fleet and colleagues who photochemically labeled antibodies with hapten[170], represents one notably interesting group for incorporation into proteins. Furthermore, this group has been used to irreversibly inactivate enzymes[171] and to probe protein-protein interactions[172,173].

Chin and others enabled the site-specific incorporation of the azide containing UAA *p*-azido-L-phenylalanine (AzF) in good yield with high fidelity by evolving the *M. jannaschii* tyrosyl-tRNA synthetase (MjYRS) to an orthogonal synthetase that incorporates AzF (MjAzFRS)[80]. Once incorporated into proteins the azide can be easily modified via bioorthogonal "click chemistry" even in cells and living organisms[114]. Whereas the paradigm of all click reactions, the Cu(I)-catalyzed cycloaddition between an azide and a terminal alkyne[115], is problematic due to its requirement for the cytotoxic copper[42], the Cu(I)-free alternative with cyclooctyne derivatives employs ring strain for alkyne activation[116,117]. Chin and co-workers encoded AzF in combination with *N6*-[(2-propynyloxy)carbonyl]-L-lysine (CAK) in GST-calmodulin forming a redox-insensitive, crosslink by bioorthogonal cycloaddition that can be used to specifically constrain protein structure on the nanometer scale[78]. The Ebright laboratory coupled fluorescent probes by Staudinger ligation to genetically encoded AzF to investigate the opening and closing of the bacterial RNA polymerase clamp[118].

The original tRNA/aaRS pair was designed to incorporate AzF in response to the amber codon UAG. In order to combine the *M. jannaschii* derived pair with similar amber decoding *M. barkeri* tRNA/aaRS pairs on the modular genetic tool we needed to change the codon specificity. In collaboration with the bachelor students Miguel Sánchez[174] and Julia Motz[175] the tRNA anticodon recognizing region of MjAzFRS was evolved to accept the expanded anticodon of the MjYT_UCCU tRNA as described by Neumann *et al*[78,79].

To this end, the open reading frame (ORF) of pBK PylS (pCLA1) was replaced by MjAzFRS (pCLA119) and subsequently used to create a library of 3*10^6 mutant clones by randomizing the codons for residues Y230, C231, P232, F261, H283 and D286 (on pCLA120) in the tRNA anticodon recognizing region.

We analyzed 96 clones after an AzF dependent low stringency selection on chloramphenicol via broad range Cm reporter assay (0 to 500 µg/mL). The 20 most promising clones (pCLA121 to pCLA140) were transferred to new Cm containing agar plates for re-phenotyping (Figure 5.1) and sequenced to monitor codon changes (Table 5.1).

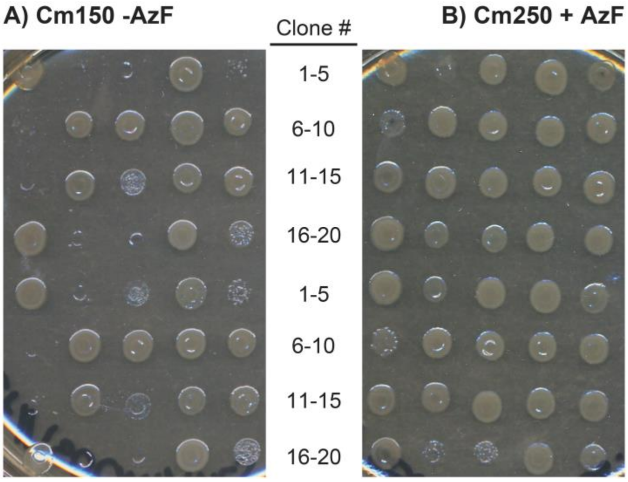


Figure 5.1: Cm-Assay with the 20 best clones from the MjAzFRS library.
Cells were plated on agar plates containing Kan, Tet, increasing Cm and 1 mM AzF (B) or no AzF (A). A selection of two plates is shown only. Clones were plated in duplicates. Numbers ascend from left to right.

The clones with number 2, 3, 5, 6, 11, 17 and 18 (marked orange in Table 5.1) showed a growth clearly dependent on the presence of the amino acid AzF allowing them to survive a chloramphenicol concentration up to 250 µg/mL. Without AzF they were not able to resist Cm at concentrations higher than 150 µg/mL. The other clones grew on plates with a high Cm concentration even in the absence of AzF.

The amino acid sequence of the selected 20 clones revealed the common mutations Y230K, C231K and P232K, but divergent mutations at positions F261, H283 and D286. This observation was also made by Neumann and colleagues[78] who suggested that amino acids 230, 231 and 232 grant affinity and specificity for the anticodon, and that 261, 283 and 286 may link the identity of the anticodon to the amino acid identity.

Table 5.1: Amino acid sequence of the 20 best clones from the MjAzFRS library.
The mutated six amino acids are compared to the wild type amino acid sequence. Clones that showed an unnatural amino acid (AzF) dependent growth on Cm containing agar plates are marked in orange.

Clone #	Amino acid					
WT	**Y230**	**C231**	**P232**	**F261**	**H283**	**D286**
1	K	K	K	T	G	C
2	P	P	A	M	F	A
3	Y	C	P	G	G	A
4	K	K	M	T	G	L
5	R	K	E	F	S	W
6	K	K	K	P	R	Y
7	K	K	K	F	L	Q
8	K	K	K	F	S	F
9	T	P	P	Q	G	G
10	K	K	K	S	P	S
11	R	A	H	F	V	M
12	P	K	L	S	G	A
13	K	K	K	G	A	T
14	R	K	R	F	P	H
15	K	K	K	F	A	T
16	K	K	K	W	G	Y
17	K	K	K	W	G	Y
18	K	K	K	F	frameshift @ K280	
19	K	K	K	S	P	Y
20	K	K	K	S	P	Y

To confirm the ability of the synthetases to suppress AGGA codons by charging the corresponding tRNA with AzF, we performed test expressions of GFP (pCLA108) and histone H3 (pCLA143) in combination with the orthogonal ribosome ribo-Q1 (pCLA96), depicted in Figure 5.2. In order to include the genes for the evolved MjAzFRS synthetases on one plasmid with the appropriate tRNA MjYT_UCCU, these were cloned into pDULE (pCLA144 to pCLA150).

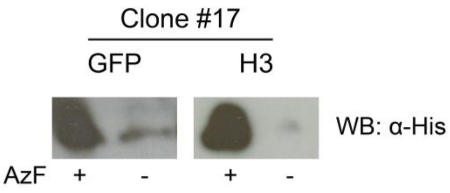


Figure 5.2: Incorporation of AzF into GFP and H3 using MjAzFRS library clone 17 (pCLA149).
Expressions of GFP 134AGGA (pCLA108) and H3 T6AGGA (pCLA143) were performed in the absence and presence of 1 mM AzF. All orthogonal mRNA translations were driven by ribo-Q1 (pCLA96). For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody.

The MjAzFRS library clone 17 exhibited the best amino acid dependent expression and thereby AGGA codon suppressing efficiency. The gene for this mutant synthetase was finally cloned into the modular genetic tool (pCLA155).

5.2 Labeling of Proteins For FRET

Förster/Fluorescence resonance energy transfer (FRET) is a powerful tool to investigate conformational changes of biomolecules like proteins and DNA, even on a single-molecule (sm) level[122,123]. Uemura *et al.* used FRET to study the transit of tRNAs on single translating ribosomes at codon resolution in real-time[176]. smFRET experiments aided the understanding of the mechanism of Rrp44 (Dis3), a key catalytic subunit of the yeast exosome[177], and to obtain an experimentally derived model of the synaptotagmin 1-SNARE fusion complex[178].

However, the precise and site-specific labeling of the proteins with suitable fluorophores is essential for FRET experiments but remains a challenging task and often the limiting factor. A combination of chemical ligations of fluorophores and proteins using maleimide conjugations and bioorthogonal labeling has been successfully used to study T4 lysozyme folding[124]. Nevertheless, this technique is restricted in its applicability to single-cysteine proteins. The incorporation of two distinct and suitable UAAs into the same protein[72,78] and the subsequent bioorthogonal labeling of these provides an elegant method for the site-specifically installation of FRET pairs.

Genetically encoded UAAs with azide function are among those suitable, where they can be utilized as crosslinkers or be altered with bioorthogonal “click chemistry”. Disadvantages of incorporating azides are that they suffer from intracellular reduction and, in the case of fluorescent labeling, the endangered design of a fluorogenic labeling scheme[119]. Alternative approaches are based on inverse electron-demand Diels-Alder cycloaddition reactions between genetically encoded strained alkenes and alkynes[119], like norbornenes[120], bicyclononynes[121] and *trans*-cyclooctenes[121], and tetrazines and/or azides. Besides not having the disadvantages mentioned above, these kinds of bioorthogonal reactions have reported rate constants that are orders of magnitude faster than other established methods[179–182]. Furthermore, fluorophores conjugated to strongly quenching tetrazines can become intrinsically fluorogenic[119,183], a trait difficult to attain if the tetrazines are genetically incorporated into proteins[119,184], instead.

The lysine derivatives *N*ε-5-norbornene-2-yloxycarbonyl-L-lysine (NorK) and bicyclo[6.1.0]non-4-yn-9-ylmethanol-L-lysine (BCNK) have been successfully incorporated into proteins by the Chin laboratory, followed by labeling with fluorophores conjugated to tetrazines[120,121]. Whereas the NorK could be incorporated using the wild type PylS/PylT pair, BCNK required a mutant PylS (BCNRS) with the three amino acid substitutions Y271M, L274G, and C313A in the enzyme active site.

We performed an expression of an amber codon containing myoglobin (S4TAG; pCLA156) to test the incorporation of theses amino acids (Figure 5.3). The necessary mutations needed for PylS to accept BCNK as a substrate were previously introduced by Heinz Neumann (pCLA158).

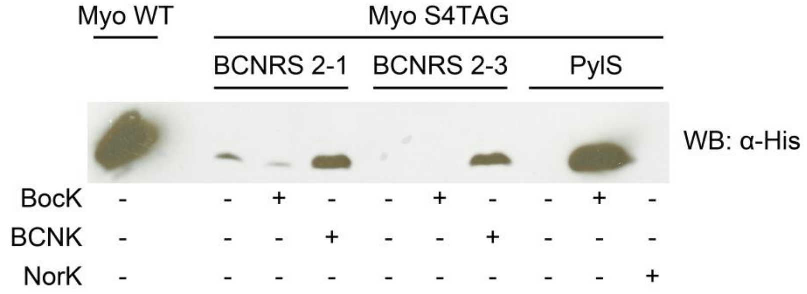


Figure 5.3: Incorporation of BCNK and NorK into myoglobin.
Expressions of wild type (Myo WT; pCLA157) and amber codon containing myoglobin (Myo S4TAG; pCLA156) were induced with arabinose (0.2% final concentration) and tested in the absence of UAAs and presence of 1 mM BocK, BCNK or NorK, respectively. Two different clones of the BCNK accepting PylS (pCLA158) were tested. For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody.

We observed specific incorporation of BCNK in response to the amber codon for both BCNRS synthetases. Although published, the wild type PylS was only tolerating the substrate analogue BocK but not the desired lysine derivative NorK.

To find another potential synthetase that is able to mediate the suppression of the amber codon with NorK without creating a new library, we tested a selection of existing PylS derivatives. These included the wild type PylS (pCLA1), the BCNK employing BCNRS 2-1 (pCLA158), an evolved synthetase for photocaged lysine (pcKRS; pCLA159), an azido-photocaged lysine using variant (AzpcKRS; pCLA160) and two synthetases optimized for the incorporation of acetyllysine (AcK), namely AcKRS1 (pCLA161) and AcKRS3 (pCLA162). We chose an amber codon containing variant of histone H3 (K56TAG; pCLA163) as the target protein for the incorporation of NorK (Figure 5.4).

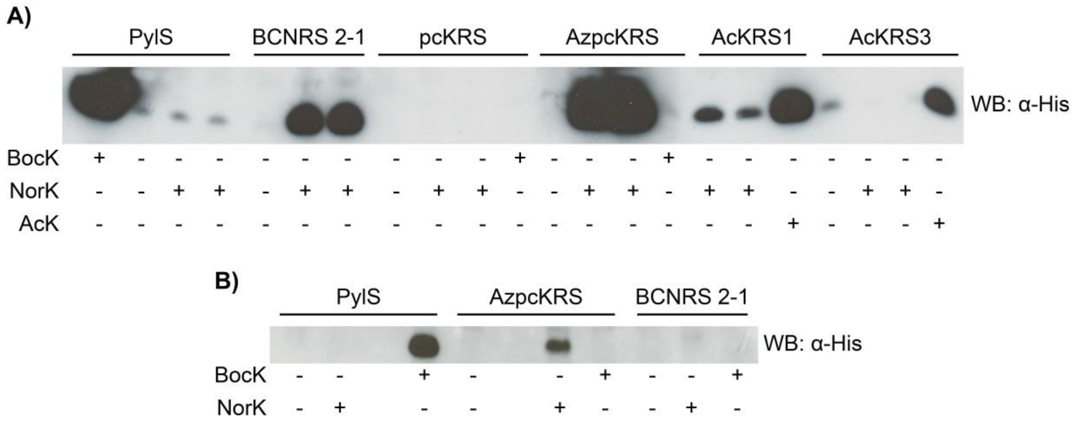

A)
PylS
BCNRS 2-1
pcKRS
AzpcKRS
AcKRS1
AcKRS3
WB: α-His
BocK + - - - - - - - - - + - - - + - - - - - - -
NorK - - + + - + + - + + - - + + - + + - - + + -
AcK - - - - - - - - - - - - - - - - - + - - - +
B)
PylS
AzpcKRS
BCNRS 2-1
WB: α-His
BocK - - + - - + - - +
NorK - + - - + - - + -

Figure 5.4: Incorporation of NorK into histone H3.
A) Expression of amber codon containing histone H3 (K56TAG; pCLA163) in *E. coli* BL21 was induced with IPTG (1 mM final concentration) and tested in the absence of UAAs and presence of 1 mM BocK, NorK or AcK, respectively. Six different PylS derivatives were tested, including wild type (pCLA1), BCNRS 2-1 (pCLA158), pcKRS (pCLA159), AzpcKRS (pCLA160), AcKRS1 (pCLA161) and AcKRS3 (pCLA162). For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody. **B)** Same samples as in A) but five times less volume loaded onto gel.

The wild type PylS did not recognize NorK as a substrate, as already shown before (Figure 5.3). This means that no full-length histone could be detected (Figure 5.4-A). The synthetases for the photocaged lysine (pcKRS) as well as the acetyllysine (AcKRS3) were also not able to bind NorK. AcKRS1 showed some NorK dependent expression of H3 but clearly less than H3 with the actual corresponding amino acid AcK. BCNRS 2-1 seemed to incorporate NorK as efficient as the AcKRSs their substrate AcK. The azido-photocaged lysine aaRS revealed good yields of NorK dependent full-length histone H3. Expression levels were hard to compare with BocK dependent H3 expression using PylS due to overexposure on the western blot. A second blot with less sample loaded (Figure 5.4-B) verified that AzpcKRS is able to encode NorK with an efficiency approximately threefold lower than BocK encoded by wild type PylS.

As a preparation for future FRET experiments we incorporated the amino acids BocK, BCNK and NorK into the single-cysteine protein histone H3 (V35TAG, F78C, C110A; pCLA164), in order to test the combination of tetrazine conjugated fluorophores with maleimide conjugations for in-gel pseudo FRET measurements. Therefore, the H3 histones were purified from *E. coli* inclusion bodies (Ch. 3.2.2.7) yielding 1.4 mg protein with BocK, 1.1 mg with NorK and 0.4 mg with BCNK from 0.5 L culture each. Afterwards, the histones were consecutively labeled with fluorophores coupled to tetrazines and maleimides (Ch. 3.2.2.11). Labeling was verified and visualized using a Typhoon imager (Figure 5.5).

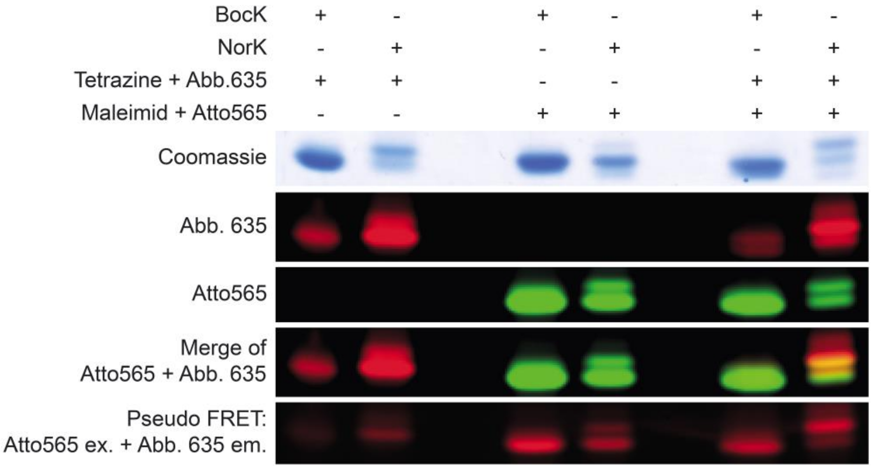

BocK + - + - + -
NorK - + - + - +
Tetrazine + Abb.635 + + - - + +
Maleimid + Atto565 - - + + + +
Coomassie
Abb. 635
Atto565
Merge of Atto565 + Abb. 635
Pseudo FRET: Atto565 ex. + Abb. 635 em.

Figure 5.5: Labeling of histone H3 for (pseudo) FRET.
NorK and BocK were incorporated into histone H3 (V35TAG, F78C, C110A; pCLA164) with BocK serving as a negative control. Histone H3-NorK was labeled with the fluorophore Abberior Star635 (Abb. 635) conjugated with a tetrazine. BocK showed only weak interactions respective to protein level. For (pseudo) FRET an Atto565 dye conjugated with a maleimide was bound to the one cysteine. The SDS-PAGE gel was scanned using a Typhoon imager (Abb.635 – excitation (ex.): 633 nm (red laser) and emission (em.): Filter 670 nm BP30 / Atto565 - excitation: 532 nm (green laser) and emission: Filter 580 nm BP30).

With respect to the amount of protein loaded onto the gel only little background was observed. The tetrazine conjugated fluorophore Abberior Star635 could be specifically “clicked” to the strained alkene containing protein. The labeling of the protein, especially the double-labeling, caused a small shift of 1-2 kDa in the gel. This fits with the sizes of the dyes that have a molecular mass of approximately 1 kDa each. The most shifted band belonged to H3 with NorK incorporated and double labeled with both dyes and showed a clear signal for pseudo FRET. That means the emission of the excited Atto565 fluorophore provoked the excitation of Abberior Star635 whose emission was finally detected. We obtained comparable results for labeling experiments with histones comprising BCNK instead of NorK (data not shown).

Finally, three additional tetrazine conjugated fluorescent dyes were tested for labeling using the same protein purifications as above (Figure 5.6).

The Typhoon scan showed only weak background signals for all four dyes tested as demonstrated by histones with non-reactive BocK incorporated. The labeling efficiencies for histones with NorK and BCNK were similar with Abberior Star635 and KK114 giving the brightest signals. It seemed that the four dyes did not have the same concentration, although properly dissolved with regard to their indicated weight, leading to distinct labeling efficiencies and thereby signal intensities.

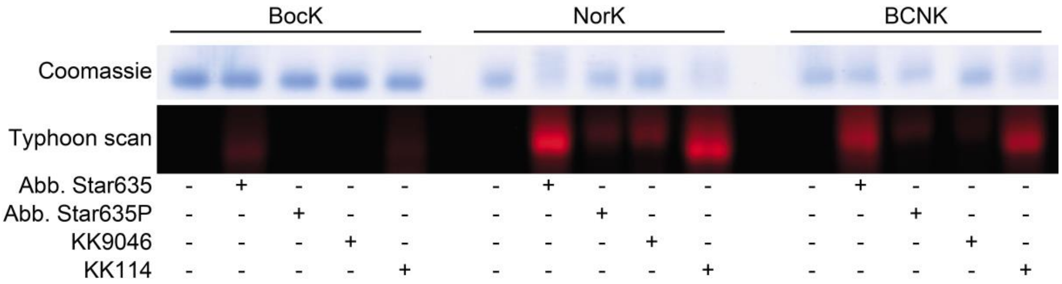

BocK
NorK
BCNK
Coomassie
Typhoon scan
Abb. Star635 - + - - - - + - - - - + - - -
Abb. Star635P - - + - - - - + - - - - + - -
KK9046 - - - + - - - - + - - - - + -
KK114 - - - - + - - - - + - - - - +

Figure 5.6: Labeling of histone H3 with four different tetrazine conjugated fluorophores.
NorK, BCNK and BocK were incorporated into histone H3 (V35TAG, F78C, C110A; pCLA164) with BocK serving as a negative control. Histones were labeled with the tetrazine conjugated fluorophores Abberior (Abb.) Star635, Abberior Star635P, KK9046 and KK114[125]. BocK showed only weak interactions respective to protein level. The SDS-PAGE gel was scanned using a Typhoon imager (Excitation: 633 nm (red laser) and emission: Filter 670 nm BP30).

In conclusion, the strained alkyne BCNK and alkene NorK could be genetically encoded into proteins that then could be site-specifically labeled with different fluorophores. We cloned the genes for both corresponding synthetases into the modular genetic tool substituting the wild type PylS gene (pCLA165 and pCLA166). The UAA NorK can be combined with AzF, encoded by AGGA, whereas BCNK would cross-react with the azide. To potentially optimize the incorporation efficiency of both UAAs the wild type *glnS* promoter in front of both synthetases was replaced by the promising alternatives P_{con}, P_{lac} (see Ch. 4.3 and 4.4.6; pCLA171 to pCLA174) and will ultimately be tested in future experiments.

5.3 Fluorescently Labeled Nanobodies

In 1993 Hamers and co-workers discovered antibodies in the camel serum that have naturally evolved to be devoid of light chains[185]. Later, these very small single-domain antibodies, about 15 kDa[186], were shown to be adaptable to other desired antigens by immunizing dromedaries followed by a screening process that led to extremely stable, highly soluble minimum sized antigen binders. These binders, termed nanobodies (Nb) by the Ablynx company, could be recombinantly expressed in *E. coli* and were proven to react specifically and with high affinity to the antigens[187].

Due to their advantageous properties nanobodies have been used in a broad range of applications as research and diagnostic tools as well as therapeutics[188]. Saerens *et al.* found Nbs to sense conformational changes on different prostate-specific antigen isoforms which could be helpful to discriminate the stages of prostate cancer[189]. Nanobodies served as tools for molecular tumor imaging[190,191] and were further developed for in vivo imaging of specific immune cell types[186]. The Leonhardt laboratory engineered an anti-GFP nanobody that allows the purification of GFP-fusion proteins and their interacting partners for biochemical studies on, e.g., DNA binding, enzymatic activity and complex formation[192].

The Rizzoli group (ENI, Göttingen) is interested in synaptic vesicle function, with an emphasis on synaptic vesicle recycling, and therefore combines fluorescently labeled nanobodies and high resolution imaging provided by stimulated emission depletion (STED[193,194]) microscopy. Chemical ligations of fluorophores and Nbs prevent a quantitative analysis due to multiple but erratic binding. Here, we show a collaborative approach for the site-specific incorporation of fluorophores, suitable for STED[125], by applying the principles of genetic code expansion as described in Ch. 5.2.

First, a set of α-synuclein targeting Nbs[195] (NbSyn2), including wild type and the amber codon containing mutants S44TAG, K70TAG, R72TAG, K93TAG, K115TAG and S153TAG (provided by Felipe Opazo from AG Rizzoli),were expressed in the absence and presence of BocK and NorK to test the overall incorporation of UAAs into nanobodies (Figure 5.7).

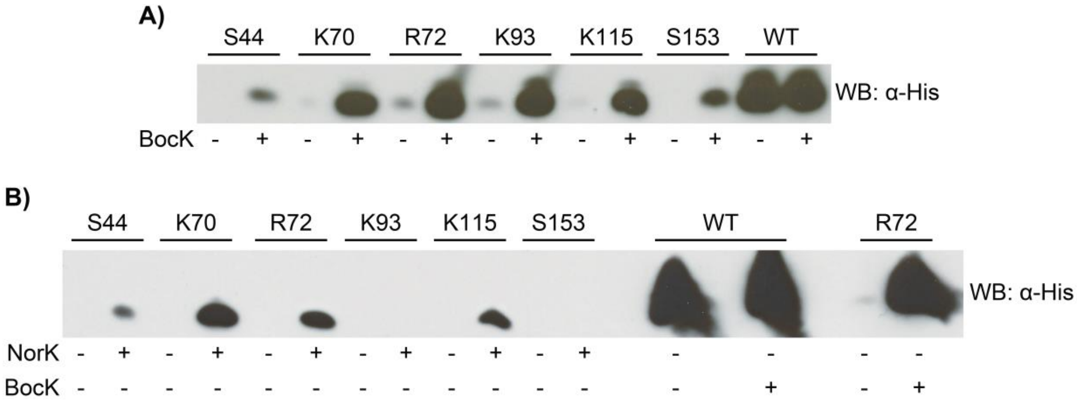

A)
S44
K70
R72
K93
K115
S153
WT
WB: α-His
BocK - + - + - + - + - + - + - +
B)
S44
K70
R72
K93
K115
S153
WT
R72
WB: α-His
NorK - + - + - + - + - + - + - - - -
BocK - - - - - - - - - - - - - + - +

Figure 5.7: Incorporation of BocK and NorK into NbSyn2.
A) Expression of wild type (WT) and amber codon containing nanobodies NbSyn2 (pCLA175 to pCLA181) in *E. coli* BL21 was induced with IPTG (1 mM final concentration) after two hours of incubation at 37 °C. BocK (1 mM final concentration) was added subsequently and temperature was shifted to 28 °C for overnight expression (~16 h). PylS and PylT were used as aaRS and tRNA pair (pCLA97). For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody. **B)** Same experiment as in A) but NorK was used as UAA instead of BocK together with the appropriate aaRS AzpcKRS (pCLA166). WT and R72 samples on the right side originate from A).

We found UAA dependent expression of NbSyn2 with BocK being incorporated with higher efficiencies than NorK yielding more full-length protein, as seen before (Figure 5.4-B). The clones with amber codons at positions K70 and R72 revealed the best expression levels. Both mutations are located in the same loop opposite the target binding pocket, according to the crystal structure of the nanobody (pdb file 2X6M[195], not shown).

We finally chose the NbSyn2 K70TAG clone for large scale expression and subsequent labeling with the Abberior Star635 fluorophore. Since the nanobody was cloned behind a secretion signal its production occurred in the periplasm, where the oxidizing environment is optimal for the required disulfide formation[188]. Secreted proteins were purified from the medium (1 L; Ch. 3.2.2.8) yielding 0.11 mg full-length protein and then labeled with the tetrazine conjugated fluorophore (Ch. 3.2.2.11). Labeling was verified and visualized using a Typhoon imager (Figure 5.8).

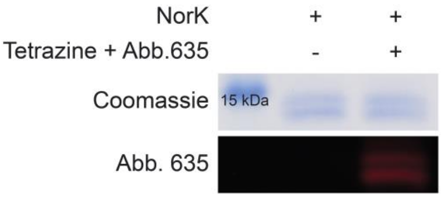


Figure 5.8: Fluorescently labeled NbSyn2.
NorK was incorporated into nanobody NbSyn2 (K70TAG; pCLA177). NbSyn2-NorK was labeled with the fluorophore Abberior Star635 (Abb. 635) conjugated with a tetrazine. The SDS-PAGE gel was scanned using a Typhoon imager (Excitation: 633 nm (red laser) and emission: Filter 670 nm BP30).

NbSyn2 could be successfully labeled with the tetrazine conjugated fluorophore Abberior Star635 and was given to Felipe Opazo for the removal of unbound dye and further experiments.

Due to the low yield of purified NbSyn2 from medium, we tested a second nanobody targeting GFP [192] (NbGFP). NbGFP is expressed in good yields and exhibits excellent binding to its target. Additionally, the cell line was changed from *E. coli* BL21 to *E. coli* SHuffle. SHuffle cells were engineered to assist disulfide bond formation in the cytoplasm, e.g., in nanobodies, rendering the transport of proteins into the periplasm unnecessary (see NbSyn2). This is achieved by the simultaneous deletion of the reductases trxB and gor and the constitutive expression of a chromosomal copy of the disulfide bond isomerase DsbC which promotes the correction of mis-oxidized proteins into their correct form[196–198]. This cell line also possesses a disadvantage in that a chromosomal copy of the resistance gene for spectinomycin, which is also present on the pCDF plasmids used for the modular genetic tools.

To this end, we cloned the resistance genes for chloramphenicol and ampicillin from the Duet vectors pACYC and pET, respectively, into the modular genetic tools pCLA116 and pCLA171, replacing the spectinomycin resistance genes (pCLA184 to pCLA187).

In order to test the genetic code expansion system in the SHuffle cell line, we prepared competent cells harboring the optimized modular genetic tool with an ampicillin resistance gene and encoding BocK in response to an amber codon (pCLA185). These cells were transformed with plasmids for NbGFP containing amber codons at positions Q13 and R76 (pCLA189 and pCLA192; provided by Felipe Opazo from AG Rizzoli). Nanobodies were expressed in the absence and presence of BocK (Figure 5.9).

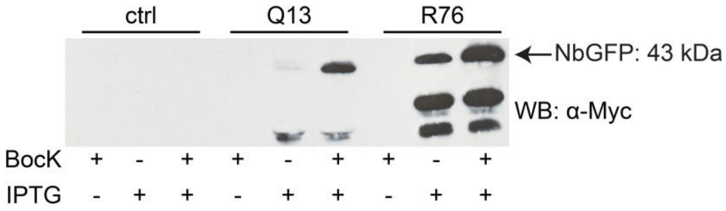


Figure 5.9: Incorporation of BocK into NbGFP.
Expression of amber codon containing nanobodies NbGFP (pCLA189 and pCLA192) in *E. coli* SHuffle was induced with IPTG (1 mM final concentration) after 2.5 hours of incubation at 30 °C. BocK (1 mM final concentration) was added subsequently and temperature was shifted to 28 °C for overnight expression (~16 h). PylS and PylT were used as aaRS and tRNA pair (pCLA185). Control cells (ctrl) contained only pCLA185 but no NbGFP plasmid. For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-Myc-antibody was used as primary antibody.

The adapted modular genetic tool containing the ampicillin resistance gene allowed the incorporation of the UAA BocK into NbGFP. We found UAA dependent expression with clone R76TAG showing more full-length protein. Since the Myc-tag is cloned to the C-terminus of the protein only full-length nanobodies should be detected by the antibody. Therefore, the intense signals of smaller sizes are probably due to strong degradation.

In summary, we were able to encode UAAs in both nanobody proteins, NbSyn2 and NbGFP. NbSyn2 was successfully labeled with a fluorophore suitable for STED, opening an avenue for highly specific binding to its target followed by quantifiable high resolution microscopy. First experiments with NbGFP expression in *E. coli* SHuffle cells looked promising concerning high yield expression. The use of this cell line will simplify the purification of the nanobodies enormously.

5.4 Orthogonal Ribosome Assisted In-Cell NMR Spectroscopy

Living organisms are highly complex biological systems with intricate networks of biological processes, simultaneously exercised by an enormous number of macromolecules like proteins. The complexity is not only determined by the quantity of proteins expressed, but also by the combinatorial interactions between them. Deciphering these interactions and putting them into a meaningful context is required to fully understand how life works.

The *in vivo* investigation of proteins in their natural cellular environment within cells assures the perpetuation of structural conformations, which are of biological importance but not reproducible *in vitro*. Techniques like X-ray crystallography and high-resolution electron microscopy are usually confined to artificial, dilute, and isolated *in vitro* experimental setups because of their necessity for pure samples and crystalline or vitrified specimen[199]. Among several attempts to find new *in vivo* techniques for structural biology and cellular imaging[200], in-cell nuclear magnetic resonance (NMR) technologies have been developed[201] allowing the direct monitoring of conformational changes at the atomic level of proteins with “NMR-active” atomic nuclei. Since protein purification is unnecessary, in-cell NMR spectroscopy can be applied to studies of proteins that are difficult to purify. Proteolytically labile proteins would be advantageous for this method because they are better protected from the proteolytic machinery whose tight regulation in cells is lost in lysates [202]. Furthermore, in contrast to lysates, less protein amount is needed in intact cells due to higher local concentrations of interacting partners, thus increasing the likelihood of detecting weaker interactions[203]. Both qualitative structural[204] and quantitative dynamic[205,206] data of proteins in living prokaryotic cells has been collected using high resolution NMR measurements.

The characterization of only one protein at a time was a limitation of the first developed in-cell NMR techniques[201] because the simultaneous overexpression of more than one protein inside the cell led to extremely intricate NMR spectra. In order to study ubiquitin interactions, the Shekhtman laboratory used sequential protein expression in a time-controlled manner to overcome this issue[203]. Here, we show an approach in collaboration with the Zweckstetter group (MPI BPC, Göttingen) that also regulates protein expression in a time-controlled fashion with the aid of an orthogonal ribosome evolved in the Chin laboratory[78]. The idea was to block protein synthesis by the endogenous ribosome using spectinomycin but to maintain translation of the desired protein by the mutant ribosome, which is spectinomycin resistant by virtue of a C1192U mutation in 16S rRNA[66].

First trials of in-cell NMR experiments were performed in *E. coli* BL21 transformed with plasmids for the orthogonal ribosome ribo-Q1 (pCLA96) and an orthogonal mRNA coding for sfGFP (pCLA105). Labeling experiments were grown in M9 minimal medium containing ^{15}N labeled ammonium chloride for isotopic labeling of GFP (Ch. 3.2.2.10). ^{15}N labeling schemes have been shown to exhibit negligible background levels[207] compared to other labeling schemes using, e.g., ^{13}C-isotopes. Purified ^{15}N labeled GFP (1.534 mg protein out of 250mL culture; Ch 3.2.2.9) and cells grown without the GFP mRNA were used as references. Despite a clear difference in GFP expression levels detected by western blot (Figure 5.10), in-cell NMR measurements, carried out by Martin Schwalbe from AG Zweckstetter, did not result in distinguishable NMR spectra for the GFP control or the control cells.

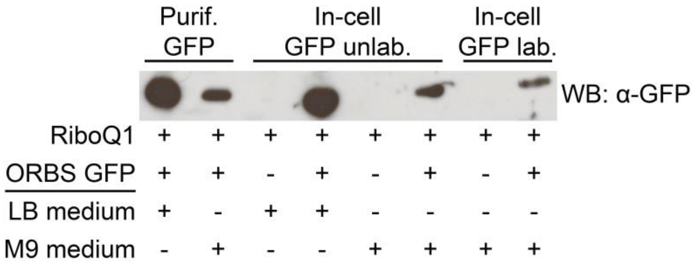


Figure 5.10: sfGFP expression for in-cell NMR measurements.
Expression of sfGFP in *E. coli* BL21 using an orthogonal mRNA (ORBS GFP; pCLA105) and the orthogonal ribosome ribo-Q1 (pCLA96). Samples were taken and normalized to $OD_{600} = 1.5$ from cells used for the purification of ^{15}N labeled GFP (Purif. GFP; Ch 3.2.2.9) and for in-cell NMR measurements, each with samples from overnight cultures grown in LB medium and main cultures grown in M9 minimal medium. Cells for in-cell NMR were grown in M9 medium with unlabeled NH_4Cl (unlab.) to $OD_{600} = 0.6$ to 0.8 and subsequently transferred to M9 medium supplemented with ^{15}N labeled (lab.) $^{15}NH_4Cl$ and 150 µg/mL spectinomycin (Ch. 3.2.2.10) to block protein synthesis by the endogenous ribosome. For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-GFP-antibody was used as primary antibody.

Although expression of GFP in M9 minimal medium yielded considerably less protein compared to expressions in LB medium, it was still sufficient to be clearly detected by western blot. However, using this method it was not possible to discriminate between the labeled and unlabeled form of GFP, with the latter being produced constitutively. To this end, we generated an arabinose inducible variant of the plasmid containing the orthogonal mRNA coding for sfGFP (pCLA193) allowing the simultaneous blocking of the endogenous ribosome and the induction of orthogonal ribosome mediated ^{15}N labeled GFP expression (Figure 5.11).

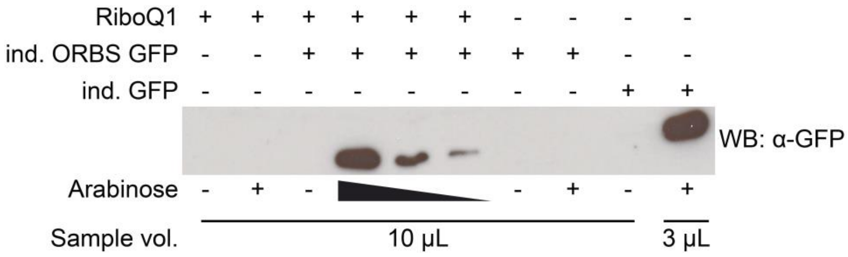

RiboQ1 + + + + + + - - - -
ind. ORBS GFP - - + + + + + + - -
ind. GFP - - - - - - - - + +
WB: α-GFP
Arabinose - + - - + - +
Sample vol.
10 µL
3 µL

Figure 5.11: Arabinose inducible promoter for the orthogonal translational machinery.
Arabinose induced expression of sfGFP in *E. coli* BL21 using an orthogonal mRNA (ind. ORBS GFP; pCLA193) in combination with the orthogonal ribosome ribo-Q1 (pCLA96) and a non-orthogonal mRNA (ind. GFP; pCLA101). Black triangle indicates decreasing concentrations of arabinose (0.2%, 0.04% and 0.008%). For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-GFP-antibody was used as primary antibody.

We observed GFP expression that was dependent on the arabinose concentration only in the presence of the orthogonal ribosome. The expression efficiency was approximately eightfold lower in comparison to the original non-orthogonal pBAD vector, as depicted in Figure 5.11 and as confirmed by in-cell fluorescence measurement (not shown). In contrast, compared to the non-inducible orthogonal GFP encoding plasmid (pCLA105) we found a 2.5-fold decrement only (data not shown).

Both the non-orthogonal and the orthogonal inducible variants were used to investigate the effect of spectinomycin on cells and ribosomes, respectively, since the applied concentration of 150 µg/mL was supposed to be too harsh in order to produce acceptable amounts of ^{15}N labeled GFP. Therefore, in-cell fluorescence of sfGFP was monitored over time in the absence and presence of 150 µg/mL spectinomycin (Figure 8.5) and samples were periodically taken for western blot analysis (Figure 5.12).

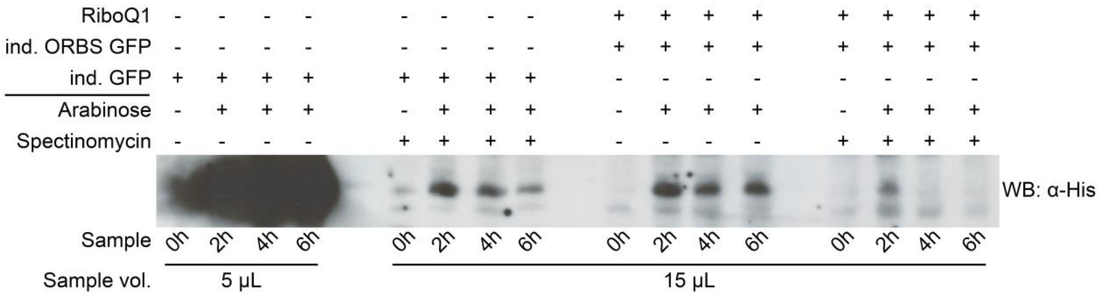

RiboQ1 - - - - - - - - + + + + + + + +
ind. ORBS GFP - - - - - - - - + + + + + + + +
ind. GFP + + + + + + + + - - - - - - - -
Arabinose - + + + - + + + - + + + - + + +
Spectinomycin - - - - + + + + - - - - + + + +
WB: α-His
Sample 0h 2h 4h 6h 0h 2h 4h 6h 0h 2h 4h 6h 0h 2h 4h 6h
Sample vol. 5 µL 15 µL

Figure 5.12: Effect of spectinomycin on sfGFP expression.
Arabinose induced expression of sfGFP in *E. coli* BL21 using an orthogonal mRNA (ind. ORBS GFP; pCLA193; 0.2% arabinose) in combination with the orthogonal ribosome ribo-Q1 (pCLA96) and a non-orthogonal mRNA (ind. GFP; pCLA101; 0.02% arabinose). Expressions were performed in the absence and presence of 150 µg/mL spectinomycin. For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody.

The in-cell fluorescence measurements (Figure 8.5) as well as the western blot (Figure 5.12) verified the expected effect of spectinomycin on non-orthogonal translational components with GFP levels being tremendously reduced. The combination of an inducible promoter and an orthogonal ribosomal binding site has been shown to produce less GFP than the non-orthogonal or non-inducible systems (Figure 5.11). However, it seems that the use of 150 µg/mL spectinomycin also had a strong negative influence on the orthogonal ribosome and thereby on the GFP expression efficiency.

In order to find a spectinomycin concentration that offers sufficient blocking of the endogenous ribosome while the orthogonal one is still mostly active, we repeated the in-cell fluorescence experiment with spectinomycin being titrated (Figure 8.6). A concentration of 10 µg/mL spectinomycin already decreased the final GFP fluorescence of cells with the non-orthogonal system by approximately 30%. At 40 µg/mL spectinomycin only 35% of the fluorescence intensity from cells with unblocked ribosomes remained. Further increase of spectinomycin yielded only a moderate decrease in GFP levels to a maximum of approximately 17%. Cells transformed with plasmids for the orthogonal components revealed that at 40 µg/mL spectinomycin there was still a fluorescence intensity of circa 56% compared to cells grown in medium without spectinomycin. Further supplements of spectinomycin led to a maximum loss of fluorescence of 60%. Thus, a spectinomycin concentration of 40 µg/mL seemed to be sufficient for blocking endogenous ribosomes but keeping the orthogonal ones active enough. This was confirmed in a final experiment that analyzed GFP expression levels of cells grown in medium with the selected spectinomycin concentrations of 0, 20 and 40 µg/mL by western blot (Figure 5.13). After normalizing cells according to in-cell fluorescence in the absence of spectinomycin (data not shown) only those containing the orthogonal mRNA coding for GFP and the orthogonal ribosome revealed GFP expression past treatment with 40 µg/mL spectinomycin.

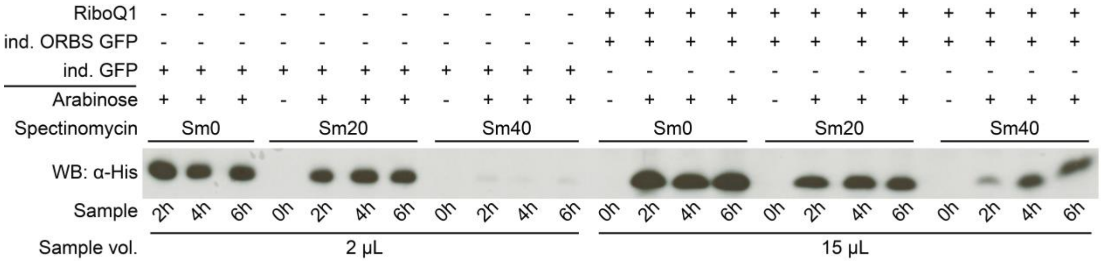

RiboQ1 - - - - - - - - - - - + + + + + + + + + + + +
ind. ORBS GFP - - - - - - - - - - - + + + + + + + + + + + +
ind. GFP + + + + + + + + + + + - - - - - - - - - - - -
Arabinose + + + - + + + - + + + - + + + - + + + - + + +
Spectinomycin
Sm0
Sm20
Sm40
Sm0
Sm20
Sm40
WB: α-His
Sample
2h 4h 6h 0h 2h 4h 6h 0h 2h 4h 6h 0h 2h 4h 6h 0h 2h 4h 6h 0h 2h 4h 6h
Sample vol.
2 µL
15 µL

Figure 5.13: sfGFP expression with selected spectinomycin concentrations.
Arabinose induced expression of sfGFP in *E. coli* BL21 using an orthogonal mRNA (ind. ORBS GFP; pCLA193; 0.2% arabinose) in combination with the orthogonal ribosome ribo-Q1 (pCLA96) and a non-orthogonal mRNA (ind. GFP; pCLA101; 0.001% arabinose). Expressions were performed in the absence and presence of 20 or 40 µg/mL spectinomycin. Cells using orthogonal and non-orthogonal components were normalized by in-cell fluorescence of GFP (data not shown). For the western blot whole cell extracts were separated with SDS PAGE (Ch. 3.2.2.3) and blotted onto a PVDF membrane (Ch. 3.2.2.4). Anti-His-antibody was used as primary antibody.

The size of sfGFP, roughly 28 kDa, could present another reason that the previously described in-cell NMR experiment (Figure 5.10) did not work. The Zweckstetter group suggested testing the small peptide -NYHLENEV**A**RLKKLVG-, representing the amino acids 16 to 31 of the yeast transcriptional activator GCN4, that has been used for NMR experiments before[208]. The peptide was expressed and ^{15}N labeled as described for sfGFP but under the control of the arabinose inducible promoter (pCLA194) and with optimized conditions using only 40 µg/mL spectinomycin. However, we were again not able to detect the isotopically labeled peptide by NMR in living cells.

6 Discussion

Studies on biological processes usually involve the examination of its participating proteins, which may be difficult, or impossible in some cases, using standard methods. In recent years, the expansion of the genetic code has become a powerful tool for biological studies. This technique provides the ability to incorporate unnatural amino acids, site-specifically into proteins, establishing new side chain possibilities and protein properties. The ongoing progress of this approach facilitated the expansion of the genetic code in several prokaryotic and eukaryotic organisms. However, the creation of proteins with (multiple) UAAs is significantly limited by the UAA incorporation efficiency. Efforts made to optimize this drawback are lacking in systematic studies that correlate aaRS and tRNA expression and aminoacylation levels with the efficiency of UAA incorporation[42,48,49,77]. This thesis focused on these issues by the establishment of assays to detect the abundance and activity of each component. We ultimately determined that the synthetase PylS was responsible for restraining the performance of the system. Furthermore, we lowered the metabolic burden for the cell by balancing the expression levels of distinct components and by reducing the number of plasmids used. This enhanced cell viability and also provided a 5-10 fold increase in yields of proteins containing UAAs.

6.1 Balancing An Expanded Genetic Code System

One could suppose that the low efficiency of UAA incorporation could be easily adjusted by simply overexpressing the required components, i.e., the tRNA/aaRS pairs, and/or by increasing the UAA concentration in the media. However, the overexpression of endogenous components could also provoke toxic ancillary effects and exhaust biosynthetic energy. Moreover, UAAs, especially those that are not commercially obtainable, have to be used as economically as possible.

In order to determine optimal expression and aminoacylation levels of the tRNA/aaRS pairs, which guarantee sufficient suppression of the blank codons without distressing the cells, we developed assays for their detection. The absence of specific antibodies against the aaRS PylS and MjYRS led us introduce the well-characterized small His_6-tag into these proteins but neither the addition to the N- nor the C-terminus allowed for clear detection of the synthetases by western blot using antibodies against the His_6-tag. Since the cell lysates were treated under denaturing conditions for the blot, using SDS and heat, the accessibility of the terminal epitopes by antibodies should not be restricted. This was proven by the Schultz group which obtained a relatively weak but distinct signal of a C-terminal His_6-tagged BPARS, a mutant MjYRS that decodes for BPA, in cell lysates by western blot analysis[69]. The detection of His_6-tagged PylS also failed for them, until they used a variant that was codon-optimized for expression in *E. coli*[72]. Considering the crystal structures of both aaRS (Figure 4.2 and 4.10), their C-termini are exposed to the protein surfaces and in case of MjYRS not directed towards the tRNA. This holds also true for the N-terminus of MjYRS, but concerning PylS no statement can be made because structural data in combination with PylT and for the N-terminal domain is missing. However, unless there are unknown interactions with other components and as long as the proteins are properly folded, the His_6-epitope should be accessible even under native conditions. Srinivasan *et al.* were therefore able to purify PylS using an N-terminal His_6-tag[33]. An observation made by Chatterjee *et al.*[72] and, independently by ourselves, is that PylS is mostly found in the insoluble fraction, probably due to misfolding in *E. coli*. Improper folded proteins are prime targets for proteolysis[209] and could thereby represent one explanation for the poor detectability of (tagged) PylS. We tried to solve this problem by using chaperones, reducing the temperature during expression and deleting parts of the proline-rich linker (Figure 4.1, 4.5 and 4.6) without success (data not shown).

In addition to these findings, we discovered that MjYRS accepted a His_6-tag at the C-terminus without losing catalytic activity (Figure 4.11), whereas PylS revealed a grave loss of activity upon addition of this tag (Figure 4.3). By contrast, shifting the His_6-tag into the flexible linker region of PylS, between the N-terminus tRNA binding domain and the C-terminus catalytic domain, proved beneficial because we were then able to detect the non-codon-optimized protein by western blot without severely impairing its enzymatic functionality (Figure 4.3 and 4.4). The introduction of an unconventional internal His_6-tag also helped us to determine MjYRS expression levels by western blot, which were many times lower than PylS levels, although under the control of the same *glnS* promoter (Figure 4.12). When expressed from the same *tacI* promoter Chatterjee *et al.* observed MjYRS to be stronger expressed than PylS[72]. This demonstrates how important different promoters can be for the regulation of certain protein levels and, in the case of these two aaRSs, very likely for the suppression efficiency of the blank codons.

When we quantified PylS and MjYRS expression levels at different stages of growth in rich medium under the control of a *glnS* promoter, we found both proteins exhibited diminished expression levels when the cultures reached high cell densities (Figure 4.4 and 4.12). This was consistent with previous reports stating that, in *E. coli*, the enzyme level of the naturally regulated glutaminyl-tRNA synthetase increases with increasing growth rate[210], because cultures that approach saturation exhibit decreased growth rates. Thus, the aaRSs may become limiting in their ability to charge their cognate tRNAs with UAAs, and thereby for suppression of blank codons in recombinant proteins, during prolonged expressions. This, in turn, renders the wild type *glnS* promoter amendable for this purpose. Indeed, the Schultz laboratory achieved a fivefold improved suppression efficiency using a mutant *glnS* promoter (*glnS'*) that led to a twofold increase in BPARS expression[69]. Additional improvements were made by further raising the expression levels of different MjYRS derivatives using a dual *glnS'*/*araBAD* promoter system or a single *tacI*-driven aaRS expression cassette[51,70,72].

We decided to use promoter libraries for optimization, which have been shown to provide a broad range of promoter activities[158–160,211]. We determined, that for the PylS, there was a direct positive dependency upon promoter strength for enzyme level, tRNA aminoacylation and suppression efficiency (Figure 4.19). Whereas Jensen and Hammer directly measured the effect of their synthetic promoter library in *Lactococcus lactis* and *E. coli* by β-galactosidase activity, obtaining up to 400-fold variation in promoter strength[160], we performed indirect measurements and observed a fivefold variation, only. This was determined by the ability of the PylS to suppress an amber codon in the genes of CAT or histone H3, with the most efficient promoter being two to threefold more efficient than the wild type *glnS* variant. In the case of the MjYRS the promoter libraries did not reveal a similar tendency (Figure 4.22), leading to the assumption that the PylS constrains the performance of the current system for the incorporation of two distinct UAAs. The replacement of the wild type *glnS* promoter in front of the PylS gene with some well-defined standard promoters of different strength, including the most promising candidate used by the Schultz group, i.e., *tacI*[72], permitted us to enhance the performance of the system once more (Figure 4.25). Furthermore, we could characterize the *glnS* promoter and also its strongest library variant, 3E, as relatively weak. However, we discovered modest increases in PylS levels already being advantageous, yielding higher amounts of UAA containing proteins. Promoters with stronger activities, like *tacI*, did not prove more useful and likely consumed biosynthetic resources, particularly if more than one component is controlled by these regulatory elements. To this end, we performed subsequent experiments using the constitutive promoter P_{con} or the IPTG inducible P_{lac} to carry out PylS expression.

In order to optimize the expression levels of the appropriate tRNAs, PylT and MjYT_UCCU, the original *lpp* promoters were also exchanged by promoter libraries. DIG labeled hybridization probes were created for the simultaneous detection of their abundance and loading status by northern blot analysis. We obtained specific signals for both tRNAs and were able to distinguish their aminoacylated form from their free form by means of a mobility shift in acid urea gels (Figure 4.9 and 4.16), as demonstrated by Polycarpo *et al.*[39,212] or Köhrer and Rajbhandary[142]. However, the effect of the tRNA promoter libraries on the suppression efficiency has not yet been investigated by western and northern blot. They were merely used in combination with the two aaRS promoter libraries to construct a four component suppressor plasmid (see Ch. 6.2). According to the results of the Schultz laboratory, where they replaced the weak *lpp* promoter of MjYT_CUA by the stronger *proK* promoter[69], one could expect improved suppression efficiencies by increasing the tRNA expression levels. While they initially observed a positive correlation between suppression efficiency and increased copies of tRNA, up to six tRNA copies per plasmid organized in two clusters under the control of *proK*, they later stated that these increases caused toxicity effects to rise to significant levels, leading to considerably slowed growth rates of *E. coli*. They could show that a single copy of an optimized *M. jannaschii* amber suppressor tyrosyl-tRNA[71] exceeded the performance of multiple non-optimized tRNAs, while the cell growth was less affected[51].

In summary, it was shown by us, and others, that raising expression levels of both aaRS and tRNA to a certain limit provides benefits to the genetic code expansion system. Expression levels beyond these limits are either producing toxic side effects or do not further enhance suppression efficiencies and are probably exhausting biosynthetic energy. Instead of increasing the component's abundance, additional opportunities to optimize the system could focus on the functionalities of the components. This is exemplified by the work of Kobayashi *et al.* who created a MjYRS mutant that aminoacylates its cognate tRNA 65 times more efficiently than its wild type variant, mainly due to improved recognition of the amber anticodon[73].

6.2 Reducing The Metabolic Burden

The efficient incorporation of two distinct unnatural amino acids, which can be altered by bioorthogonal "click chemistry", is an important intermediate step for the creation of proteins that can be double labeled with fluorophores for smFRET studies. Developed by the Chin laboratory, the original orthogonal ribosome assisted system for the dual suppression of an amber and a frameshift codon consisted of six exogenous components (two cognate pairs of tRNA/aaRS, an O-mRNA and the O-ribosome) distributed on four different plasmids[78]. This number of plasmids, harboring potentially toxic components and corresponding antibiotics for their maintenance, impairs cell growth rates significantly (Figure 4.38) and thus limits the possibilities of this system. By merging a tRNA/aaRS pair from two vectors into a single one, the reduction of required plasmids has already been shown to be advantageous for the incorporation of a single UAA into proteins[68,69]. Chatterjee *et al.* demonstrated the feasibility of a dual nonsense (amber and ochre) suppressor plasmid that was used to efficiently install two UAAs into GFP. This system did not take benefits from an O-ribosome; hence, only five exogenous components were present, distributed on two plasmids. However, protein yields were not improved compared to a three plasmid system[72].

In contrast, we intended to optimize the six component system introduced by Chin and co-workers to profit from the evolved O-ribosome ribo-Q1 that facilitates UAA incorporation in response to both amber and frameshift codons, with the latter one providing 256 new possibilities to encode UAAs. We therefore constructed a modular genetic tool that also works as a dual suppressor plasmid by expressing both required tRNA/aaRS pairs, in order to reduce the metabolic burden provoked by the four plasmids of the original system. Furthermore, this tool permits the exchange of individual parts (including tRNAs, aaRS and their promoters and terminators) by a series of unique restriction sites to adapt the system for new UAAs (Figure 4.26). Compared to the established four plasmid system, we indeed observed significantly enhanced growth rates (Figure 4.38) and improved yields of UAA containing proteins (Figure 4.31) by employing our three plasmid system.

Trials to optimize our first version of the tool by replacing the original *glnS* and *lpp* promters of the aaRS and tRNAs with promoter libraries, to balance the expression levels of each individual component, failed so far. It seemed that the replacement generated a multitude of less active or even inactive clones, which were difficult to separate from more efficient clones using FACS (Figure 4.34). The observed higher GFP expression levels from putatively optimized clones were possibly due to the simultaneous measurement of more than one cell by virtue of the small size of *E. coli* cells. In some cases a strong overexpression of PylS was found, while only tiny amounts of UAA containing GFP were produced (Figure 4.35). This is consistent with the aforementioned hypothesis that overexpressing one component is likely to produce toxicity and consume biosynthetic energy.

By contrast, when we exchanged the *glnS* promoter of the PylS on the modular genetic tool with P_{con} or P_{lac}, we observed enhanced amber suppression efficiency, on its own and in combination with a frameshift codon. P_{con} was at least 1.5 to 2-fold more efficient and P_{lac} 2.5 to 3-fold compared to the standard *glnS* promoter (Figure 4.36). In comparison with the original four plasmid system, these second optimized versions of our tool provided a 5-10 fold increase in yields of proteins containing UAAs, while the improved growth rates, due to the reduced metabolic burden, were maintained (Figure 4.37 and 4.38).

6.3 The Avenue To FRET Measurements

For a start to demonstrate the feasibility that the existing system can be optimized, all previously described optimizations were performed with the wild type PylS and MjYRS using the commercially obtainable but unremarkable UAAs BocK and tyrosine as a substrate. The site-specific labeling of proteins with fluorophores by bioorthogonal "click chemistry" requires the incorporation of suitable UAAs and in some cases the evolution of the aaRS' specificity towards them. Chin *et al.* evolved MjYRS to incorporate one of these suitable UAAs, the azide containing UAA *p*-azido-L-phenylalanine (AzF), in response to an amber codon[80]. Later on, Neumann *et al.* changed the decoding property of this mutant aaRS (MjAzFRS) from UAG to AGGA by randomizing the tRNA anticodon recognizing region to accept the expanded anticodon of the MjYT_UCCU tRNA[78,79]. Since the amber codon is exclusive for the *M. barkeri* PylS/PylT pair in our optimized dual suppressor system, we also needed to evolve MjAzFRS (provided by the Chin laboratory) to incorporate AzF in response to AGGA. Using the published procedure[78,79], we successfully obtained a MjAzFRS mutant that suppressed AGGA codons in GFP and histone H3 with high specificity for AzF (Figure 5.2). Sequencing of this, and other clones, revealed the common amino acid substitutions Y230K, C231K and P232K, but divergent mutations at positions F261, H283 and D286. This observation was also made by Neumann *et al.* who suggested that amino acids 230, 231 and 232 grant affinity and specificity for the anticodon, and that 261, 283 and 286 may link the identity of the anticodon to the amino acid identity[78].

Furthermore, we were able to incorporate the strained alkyne bicyclo[6.1.0]non-4-yn-9-ylmethanol-L-lysine (BCNK; Figure 5.3) and alkene Nε-5-norbornene-2-yloxycarbonyl-L-lysine (NorK; Figure 5.4) into proteins in response to amber codons using evolved variants of PylS. Whereas BCNK incorporation was successful by means of issued PylS mutations[121], we could not produce NorK containing proteins with the aid of wild type PylS as published[120], but by the use of an alternative PylS mutant originally designed to recognize an azido-photocaged lysine as substrate. The poor substrate specificity of wild type PylS for NorK was also found by Plass *et al.*, who incorporated this amino acid and other strained alkene and alkyne UAAs using a rationally designed PylS double mutant[119].

However, both BCNK and NorK contain chemical handles that can react with tetrazines and/or azides, allowing the labeling of proteins with fluorophores conjugated to these functional groups even in living cells[119–121]. To test this, as a preparation for future FRET experiments, we successfully labeled histone H3 proteins with different tetrazine conjugated fluorophores (Figure 5.5 and 5.6), which have been reported to exhibit superior performance in STED microscopy[125]. We even double labeled H3 by using a second maleimide conjugated fluorophore, permitting us to perform in-gel (pseudo) FRET measurements (Figure 5.5). Moreover, in collaboration with the Rizzoli group (ENI, Göttingen) we successfully expressed, and site-specifically labeled, the α-synuclein targeting nanobody NbSyn2[195] with Abberior Star635, which is one of the tetrazine conjugated fluorophores (Figure 5.8). In contrast to classical chemical ligations of fluorophores and nanobodies, which occur protein wide and erratically, this approach will facilitate quantitative analyses in their research on synaptic vesicle function by STED microscopy.

In summary, we demonstrated the incorporation of the UAAs AzF, BCNK and NorK into different proteins in response to the frameshift codon AGGA or the amber codon UAG, respectively. Thereby, functional groups for bioorthogonal conjugations were introduced into these proteins, allowing us to site-specifically label them with fluorophores.

Each of the corresponding aaRS was cloned into our optimized genetic tool substituting the wild type MjYRS and PylS genes. In order to produce proteins with two distinct UAAs that can be double labeled with fluorophores for smFRET studies, one could combine NorK and AzF but BCNK would cross-react with the azide of AzF. Initial trials to express proteins containing NorK and AzF yielded only tiny amounts of full-length protein, barely detectable by western blot (data not shown). This could have been due to the poor incorporation efficiency of NorK, which therefore have to be optimized for future experiments.

6.4 Applications And Side Projects: Nanobodies And In-Cell NMR

Improving proteins yields is generally important and desirable for most studies, including the production of fluorescently labeled nanobodies and the orthogonal ribosome assisted in-cell NMR spectroscopy. Since both applications/side projects were not part of the main focus of this thesis their optimization will only be briefly discussed.

We demonstrated that our optimized genetic tool was conducive to increase expression levels of proteins containing UAAs. However, the purification of the NorK containing nanobody NbSyn2 resulted only in very small amounts of protein (0.11 mg full-length protein out of 1 L medium). This was probably due to the poor incorporation efficiency of NorK but also due to the intricate but necessary production of NbSyn2 in the cell's periplasm[188] and its release to the medium from which it was purified. Besides enhancing protein yields by optimizing the incorporation of NorK or using a more efficiently incorporated UAA, one could simplify the production and purification of the nanobody. Our first approach using the specialized *E. coli* SHuffle cell line turned out to be useful for the high yield expression of proteins that contain disulfide bonds (Figure 5.9).

The production of uniquely ^{15}N labeled proteins for in-cell NMR experiments was based on the spectinomycin resistance of the used orthogonal ribosome[66]. We found that the initially applied concentration of 150 µg/mL Sm was even too harsh for the O-ribosome, preventing an efficient protein synthesis (Figure 5.12). The reduction to 40 µg/mL Sm showed a sufficient blocking of the endogenous ribosomes while the orthogonal ones were still active (Figure 5.13). However, a detection of isotopically labeled polypeptides by NMR in living cells was not possible. Maybe the amount of ^{15}N labeled protein was too low to be detected despite optimized conditions. Additional mutations in the O-ribosome may increase the resistance to spectinomycin and thereby improve protein yields. We observed that the expression of the protein from the modified pBAD vector containing an ORBS was approximately eightfold lower in comparison to the original pBAD vector (Figure 5.11). This could perhaps be overcome by the use of a stronger inducible promoter, for example the IPTG driven *lac* or *tacI* promoter, or by optimizing the expression level of the O-ribosome.

6.5 Concluding Remarks And Outlook

Since synthetic biologists initiated the expansion of the genetic code more than a decade ago, it has become a powerful tool for basic and applied scientific research. Numerous efforts have been made to expand the catalogue of useful UAAs in several organisms, including prokaryotes and eukaryotes. We expect future developments to further extend the existing system by evolving new tRNA/aaRS pairs for the incorporation of novel UAAs in response to additional blank codons even in higher eukaryotes. Optimizing the translational apparatus for a more efficient production of proteins with multiple different UAAs will possibly allow its reprogramming to create fully synthetic unnatural polymers.

In this study, we described a part of these optimizations by significantly improving the reliability and yield of UAA containing proteins produced with the aid of orthogonal ribosomes. Our analysis of the limiting factors revealed bottlenecks of the existing system and will be helpful for future studies. For instance, the integration of some components into the host's chromosome for a further reduction of the metabolic burden, if one can find a solution to compensate the loss in copy numbers compared to plasmids. Optimized expression levels of O-ribosomes will also prove beneficial for the performance of the system. One approach, which addressed both concerns, was recently published by the Chin laboratory. That work combined the orthogonal ribosomal RNA of ribo-Q1 and the O-mRNA of the protein of interest on a single vector, reducing the number of required plasmids by one[213]. Furthermore, they evolved a series of quadruplet-decoding PylT variants that showed enhanced UAA specific decoding of their corresponding codons on ribo-Q1 in combination with PylS or some of its derivatives in comparison with quadruplet-decoding pairs derived from MjYT/MjYRS. Both strategies together led to improved incorporation efficiencies of two distinct UAAs into the same protein and thus to the efficient genetically encoded installation of FRET pairs[213]. Moreover, the combination of their recent work and the work presented in this study will most likely lead to an extremely more efficient genetic code expansion system.

7 References

[1] Woese, C. R.; Kandler, O. & Wheelis, M. L. "Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya." *Proc. Natl. Acad. Sci. U. S. A.* **87,** 4576–9 (1990).

[2] May, R. M. "How many species are there on Earth?" *Science* **241,** 1441–9 (1988).

[3] Mora, C.; Tittensor, D. P.; Adl, S.; Simpson, A. G. B. & Worm, B. "How many species are there on Earth and in the ocean?" *PLoS Biol.* **9,** e1001127 (2011).

[4] Tsugita, A. & Fraenkel-Conrat, H. "The amino acid composition and C-terminal sequence of a chemically evoked mutant of TMV." *Proc. Natl. Acad. Sci. U. S. A.* **46,** 636–42 (1960).

[5] Crick, F. H.; Barnett, L.; Brenner, S. & Watts-Tobin, R. J. "General nature of the genetic code for proteins." *Nature* **192,** 1227–32 (1961).

[6] Nirenberg, M. W.; Matthaei, J. H. & Jones, O. W. "An intermediate in the biosynthesis of polyphenylalanine directed by synthetic template RNA." *Proc. Natl. Acad. Sci. U. S. A.* **48,** 104–9 (1962).

[7] Nirenberg, M. & Leder, P. "RNA Codewords and Protein Synthesis: The Effect of Trinucleotides upon the Binding of sRNA to Ribosomes" *Science (80-.).* **145,** 1399–1407 (1964).

[8] Khorana, H. G. "Polynucleotide synthesis and the genetic code." *Harvey Lect.* **62,** 79–105 (1966).

[9] Khorana, H. G. *et al.* "Polynucleotide synthesis and the genetic code." *Cold Spring Harb. Symp. Quant. Biol.* **31,** 39–49 (1966).

[10] Crick, F. H. C. "The origin of the genetic code" *J. Mol. Biol.* **38,** 367–379 (1968).

[11] Garen, A. "Sense and nonsense in the genetic code. Three exceptional triplets can serve as both chain-terminating signals and amino acid codons." *Science* **160,** 149–59 (1968).

[12] Ramakrishnan, V. "Ribosome structure and the mechanism of translation." *Cell* **108,** 557–72 (2002).

[13] Schmeing, T. M. & Ramakrishnan, V. "What recent ribosome structures have revealed about the mechanism of translation." *Nature* **461,** 1234–42 (2009).

[14] Kozak, M. "Initiation of translation in prokaryotes and eukaryotes." *Gene* **234,** 187–208 (1999).

[15] Laursen, B. S.; Sørensen, H. P.; Mortensen, K. K. & Sperling-Petersen, H. U. "Initiation of protein synthesis in bacteria." *Microbiol. Mol. Biol. Rev.* **69,** 101–23 (2005).

[16] Simonetti, a *et al.* "A structural view of translation initiation in bacteria." *Cell. Mol. Life Sci.* **66,** 423–36 (2009).

[17] Shine, J. & Dalgarno, L. "The 3’-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites." *Proc. Natl. Acad. Sci. U. S. A.* **71,** 1342–6 (1974).

[18] Steitz, J. A. & Jakes, K. "How ribosomes select initiator regions in mRNA: base pair formation between the 3’ terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli." *Proc. Natl. Acad. Sci. U. S. A.* **72,** 4734–8 (1975).

[19] Woese, C. R.; Olsen, G. J.; Ibba, M. & Söll, D. "Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process." *Microbiol. Mol. Biol. Rev.* **64,** 202–36 (2000).

[20] O’Donoghue, P. & Luthey-Schulten, Z. "On the evolution of structure in aminoacyl-tRNA synthetases." *Microbiol. Mol. Biol. Rev.* **67,** 550–73 (2003).

[21] LaRiviere, F. J.; Wolfson, A. D. & Uhlenbeck, O. C. "Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation." *Science* **294,** 165–8 (2001).

[22] Ogle, J. M.; Murphy, F. V; Tarry, M. J. & Ramakrishnan, V. "Selection of tRNA by the ribosome requires a transition from an open to a closed form." *Cell* **111,** 721–32 (2002).

[23] Ogle, J. M.; Carter, A. P. & Ramakrishnan, V. "Insights into the decoding mechanism from recent ribosome structures." *Trends Biochem. Sci.* **28,** 259–66 (2003).

[24] Rodnina, M. V; Gromadski, K. B.; Kothe, U. & Wieden, H.-J. "Recognition and selection of tRNA in translation." *FEBS Lett.* **579,** 938–42 (2005).

[25] Crick, F. H. "Codon--anticodon pairing: the wobble hypothesis." *J. Mol. Biol.* **19,** 548–55 (1966).

[26] Rodnina, M. V & Wintermeyer, W. "The ribosome as a molecular machine: the mechanism of tRNA-mRNA movement in translocation." *Biochem. Soc. Trans.* **39,** 658–62 (2011).

[27] Capecchi, M. R. "Polypeptide chain termination in vitro: isolation of a release factor." *Proc. Natl. Acad. Sci. U. S. A.* **58,** 1144–51 (1967).

[28] Scolnick, E.; Tompkins, R.; Caskey, T. & Nirenberg, M. "Release factors differing in specificity for terminator codons." *Proc. Natl. Acad. Sci. U. S. A.* **61,** 768–74 (1968).

[29] Zinoni, F.; Birkmann, A.; Stadtman, T. C. & Böck, A. "Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from Escherichia coli." *Proc. Natl. Acad. Sci. U. S. A.* **83,** 4650–4 (1986).

[30] Chambers, I. *et al.* "The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the “termination” codon, TGA." *EMBO J.* **5,** 1221–7 (1986).

[31] Böck, A. *et al.* "Selenocysteine: the 21st amino acid." *Mol. Microbiol.* **5,** 515–20 (1991).

[32] Böck, A.; Forchhammer, K.; Heider, J. & Baron, C. "Selenoprotein synthesis: an expansion of the genetic code." *Trends Biochem. Sci.* **16,** 463–7 (1991).

[33] Srinivasan, G.; James, C. M. & Krzycki, J. a. "Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA." *Science* **296,** 1459–62 (2002).

[34] Hao, B. *et al.* "A new UAG-encoded residue in the structure of a methanogen methyltransferase." *Science* **296,** 1462–6 (2002).

[35] Atkins, J. F. & Gesteland, R. "Biochemistry. The 22nd amino acid." *Science* **296,** 1409–10 (2002).

[36] Driscoll, D. M. & Copeland, P. R. "Mechanism and regulation of selenoprotein synthesis." *Annu. Rev. Nutr.* **23,** 17–40 (2003).

[37] Krzycki, J. A. "The direct genetic encoding of pyrrolysine." *Curr. Opin. Microbiol.* **8,** 706–12 (2005).

[38] Blight, S. K. *et al.* "Direct charging of tRNA(CUA) with pyrrolysine in vitro and in vivo." *Nature* **431,** 333–5 (2004).

[39] Polycarpo, C. *et al.* "An aminoacyl-tRNA synthetase that specifically activates pyrrolysine." *Proc. Natl. Acad. Sci. U. S. A.* **101,** 12450–4 (2004).

[40] Ambrogelly, A.; Palioura, S. & Söll, D. "Natural expansion of the genetic code." *Nat. Chem. Biol.* **3,** 29–35 (2007).

[41] Lobanov, A. V; Kryukov, G. V; Hatfield, D. L. & Gladyshev, V. N. "Is there a twenty third amino acid in the genetic code?" *Trends Genet.* **22,** 357–60 (2006).

[42] Neumann, H. "Rewiring translation - Genetic code expansion and its applications." *FEBS Lett.* **586,** 2057–64 (2012).

[43] Furter, R. "Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli." *Protein Sci.* **7,** 419–26 (1998).

[44] Cropp, T. A. & Schultz, P. G. "An expanding genetic code." *Trends Genet.* **20,** 625–30 (2004).

[45] Wang, L.; Brock, A.; Herberich, B. & Schultz, P. G. "Expanding the genetic code of Escherichia coli." *Science* **292,** 498–500 (2001).

[46] Hartley, R. W. "Barnase and barstar. Expression of its cloned inhibitor permits expression of a cloned ribonuclease." *J. Mol. Biol.* **202,** 913–5 (1988).

[47] Chin, J. W. *et al.* "An expanded eukaryotic genetic code." *Science* **301,** 964–7 (2003).

[48] Liu, C. C. & Schultz, P. G. "Adding new chemistries to the genetic code." *Annu. Rev. Biochem.* **79,** 413–44 (2010).

[49] Davis, L. & Chin, J. W. "Designer proteins: applications of genetic code expansion in cell biology." *Nat. Rev. Mol. Cell Biol.* **13,** 168–82 (2012).

[50] Bossi, L. "Context effects: translation of UAG codon by suppressor tRNA is affected by the sequence following UAG in the message." *J. Mol. Biol.* **164,** 73–87 (1983).

[51] Young, T. S.; Ahmad, I.; Yin, J. A. & Schultz, P. G. "An enhanced system for unnatural amino acid mutagenesis in E. coli." *J. Mol. Biol.* **395,** 361–74 (2010).

[52] Gerdes, S. Y. *et al.* "Experimental determination and system level analysis of essential genes in Escherichia coli MG1655." *J. Bacteriol.* **185,** 5673–84 (2003).

[53] Yamazaki, Y.; Niki, H. & Kato, J. "Profiling of Escherichia coli Chromosome database." *Methods Mol. Biol.* **416,** 385–9 (2008).

[54] Huang, Y. *et al.* "A convenient method for genetic incorporation of multiple noncanonical amino acids into one protein in Escherichia coli." *Mol. Biosyst.* **6,** 683–6 (2010).

[55] Rydén, S. M. & Isaksson, L. A. "A temperature-sensitive mutant of Escherichia coli that shows enhanced misreading of UAG/A and increased efficiency for some tRNA nonsense suppressors." *Mol. Gen. Genet.* **193,** 38–45 (1984).

[56] Ito, K.; Uno, M. & Nakamura, Y. "Single amino acid substitution in prokaryote polypeptide release factor 2 permits it to terminate translation at all three stop codons." *Proc. Natl. Acad. Sci. U. S. A.* **95,** 8165–9 (1998).

[57] Johnson, D. B. F. *et al.* "RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites." *Nat. Chem. Biol.* **7,** 779–86 (2011).

[58] Johnson, D. B. F. *et al.* "Release factor one is nonessential in Escherichia coli." *ACS Chem. Biol.* **7,** 1337–44 (2012).

[59] Mukai, T. *et al.* "Codon reassignment in the Escherichia coli genetic code." *Nucleic Acids Res.* **38,** 8188–95 (2010).

[60] Li, X.; Yokota, T.; Ito, K.; Nakamura, Y. & Aiba, H. "Reduced action of polypeptide release factors induces mRNA cleavage and tmRNA tagging at stop codons in Escherichia coli." *Mol. Microbiol.* **63,** 116–26 (2007).

[61] Moore, S. D. & Sauer, R. T. "The tmRNA system for translational surveillance and ribosome rescue." *Annu. Rev. Biochem.* **76,** 101–24 (2007).

[62] Ohtake, K. *et al.* "Efficient decoding of the UAG triplet as a full-fledged sense codon enhances the growth of a prfA-deficient strain of Escherichia coli." *J. Bacteriol.* **194,** 2606–13 (2012).

[63] Wang, H. H. *et al.* "Programming cells by multiplex genome engineering and accelerated evolution." *Nature* **460,** 894–8 (2009).

[64] Isaacs, F. J. *et al.* "Precise manipulation of chromosomes in vivo enables genome-wide codon replacement." *Science* **333,** 348–53 (2011).

[65] Lajoie, M. J. *et al.* "Genomically recoded organisms expand biological functions." *Science* **342,** 357–60 (2013).

[66] Rackham, O. & Chin, J. W. "A network of orthogonal ribosome x mRNA pairs." *Nat. Chem. Biol.* **1,** 159–66 (2005).

[67] Wang, K.; Neumann, H.; Peak-Chew, S. Y. & Chin, J. W. "Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion." *Nat. Biotechnol.* **25,** 770–7 (2007).

[68] Farrell, I. S.; Toroney, R.; Hazen, J. L.; Mehl, R. A. & Chin, J. W. "Photo-cross-linking interacting proteins with a genetically encoded benzophenone." *Nat. Methods* **2,** 377–84 (2005).

[69] Ryu, Y. & Schultz, P. G. "Efficient incorporation of unnatural amino acids into proteins in Escherichia coli." *Nat. Methods* **3,** 263–5 (2006).

[70] Cellitti, S. E. *et al.* "In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy." *J. Am. Chem. Soc.* **130,** 9268–81 (2008).

[71] Guo, J.; Melançon, C. E.; Lee, H. S.; Groff, D. & Schultz, P. G. "Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids." *Angew. Chem. Int. Ed. Engl.* **48,** 9148–51 (2009).

[72] Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H. & Schultz, P. G. "A versatile platform for single- and multiple-unnatural amino acid mutagenesis in Escherichia coli." *Biochemistry* **52,** 1828–37 (2013).

[73] Kobayashi, T. *et al.* "Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion." *Nat. Struct. Biol.* **10,** 425–32 (2003).

[74] Doi, Y.; Ohtsuki, T.; Shimizu, Y.; Ueda, T. & Sisido, M. "Elongation factor Tu mutants expand amino acid tolerance of protein biosynthesis system." *J. Am. Chem. Soc.* **129,** 14458–62 (2007).

[75] Wan, W. *et al.* "A facile system for genetic incorporation of two different noncanonical amino acids into one protein in Escherichia coli." *Angew. Chem. Int. Ed. Engl.* **49,** 3211–4 (2010).

[76] Anderson, J. C. *et al.* "An expanded genetic code with a functional quadruplet codon." *Proc. Natl. Acad. Sci. U. S. A.* **101,** 7566–71 (2004).

[77] Wang, K.; Schmied, W. H. & Chin, J. W. "Reprogramming the genetic code: from triplet to quadruplet codes." *Angew. Chem. Int. Ed. Engl.* **51,** 2288–97 (2012).

[78] Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M. & Chin, J. W. "Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome." *Nature* **464,** 441–4 (2010).

[79] Neumann, H.; Slusarczyk, A. L. & Chin, J. W. "De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs." *J. Am. Chem. Soc.* **132,** 2142–4 (2010).

[80] Chin, J. W. *et al.* "Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli." *J. Am. Chem. Soc.* **124,** 9026–7 (2002).

[81] Ai, H.; Shen, W.; Sagi, A.; Chen, P. R. & Schultz, P. G. "Probing protein-protein interactions with a genetically encoded photo-crosslinking amino acid." *Chembiochem* **12,** 1854–7 (2011).

[82] Chou, C.; Uprety, R.; Davis, L.; Chin, J. W. & Deiters, A. "Genetically encoding an aliphatic diazirine for protein photocrosslinking" *Chem. Sci.* **2,** 480 (2011).

[83] Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L. & Schultz, P. G. "Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli." *Proc. Natl. Acad. Sci. U. S. A.* **99,** 11020–4 (2002).

[84] Mori, H. & Ito, K. "Different modes of SecY-SecA interactions revealed by site-directed in vivo photo-cross-linking." *Proc. Natl. Acad. Sci. U. S. A.* **103,** 16159–64 (2006).

[85] Das, S. & Oliver, D. B. "Mapping of the SecA·SecY and SecA·SecG interfaces by site-directed in vivo photocross-linking." *J. Biol. Chem.* **286,** 12371–80 (2011).

[86] Wilkins, B. J. *et al.* "A Cascade of Histone Modifications Induces Chromatin Condensation in Mitosis" *Science (80-.).* **343,** 77–80 (2014).

[87] Tagami, S. *et al.* "Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein." *Nature* **468,** 978–82 (2010).

[88] Neumann, H.; Hazen, J. L.; Weinstein, J.; Mehl, R. a & Chin, J. W. "Genetically encoding protein oxidative damage." *J. Am. Chem. Soc.* **130,** 4028–33 (2008).

[89] Liu, C. C. & Schultz, P. G. "Recombinant expression of selectively sulfated proteins in Escherichia coli." *Nat. Biotechnol.* **24,** 1436–40 (2006).

[90] Xie, J.; Supekova, L. & Schultz, P. G. "A genetically encoded metabolically stable analogue of phosphotyrosine in Escherichia coli." *ACS Chem. Biol.* **2,** 474–8 (2007).

[91] Park, H.-S. *et al.* "Expanding the genetic code of Escherichia coli with phosphoserine." *Science* **333,** 1151–4 (2011).

[92] Nguyen, D. P.; Garcia Alai, M. M.; Kapadnis, P. B.; Neumann, H. & Chin, J. W. "Genetically encoding N(epsilon)-methyl-L-lysine in recombinant histones." *J. Am. Chem. Soc.* **131,** 14194–5 (2009).

[93] Ai, H.-W.; Lee, J. W. & Schultz, P. G. "A method to site-specifically introduce methyllysine into proteins in E. coli." *Chem. Commun. (Camb).* **46,** 5506–8 (2010).

[94] Nguyen, D. P.; Garcia Alai, M. M.; Virdee, S. & Chin, J. W. "Genetically directing ε-N, N-dimethyl-L-lysine in recombinant histones." *Chem. Biol.* **17,** 1072–6 (2010).

[95] Neumann, H.; Peak-Chew, S. Y. & Chin, J. W. "Genetically encoding N(epsilon)-acetyllysine in recombinant proteins." *Nat. Chem. Biol.* **4,** 232–4 (2008).

[96] Virdee, S. *et al.* "Traceless and site-specific ubiquitination of recombinant proteins." *J. Am. Chem. Soc.* **133,** 10708–11 (2011).

[97] Neumann, H. *et al.* "A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation." *Mol. Cell* **36,** 153–63 (2009).

[98] Lammers, M.; Neumann, H.; Chin, J. W. & James, L. C. "Acetylation regulates cyclophilin A catalysis, immunosuppression and HIV isomerization." *Nat. Chem. Biol.* **6,** 331–7 (2010).

[99] Ye, S.; Huber, T.; Vogel, R. & Sakmar, T. P. "FTIR analysis of GPCR activation using azido probes." *Nat. Chem. Biol.* **5,** 397–9 (2009).

[100] Ye, S. *et al.* "Tracking G-protein-coupled receptor activation using genetically encoded infrared probes." *Nature* **464,** 1386–9 (2010).

[101] Schultz, K. C. *et al.* "A genetically encoded infrared probe." *J. Am. Chem. Soc.* **128,** 13984–5 (2006).

[102] Deiters, A.; Geierstanger, B. H. & Schultz, P. G. "Site-specific in vivo labeling of proteins for NMR studies." *Chembiochem* **6,** 55–8 (2005).

[103] Jackson, J. C.; Hammill, J. T. & Mehl, R. a. "Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity." *J. Am. Chem. Soc.* **129,** 1160–6 (2007).

[104] Li, C. *et al.* "Protein (19)F NMR in Escherichia coli." *J. Am. Chem. Soc.* **132,** 321–7 (2010).

[105] Liu, C. C.; Qi, L.; Yanofsky, C. & Arkin, A. P. "Regulation of transcription by unnatural amino acids." *Nat. Biotechnol.* **29,** 164–8 (2011).

[106] Deiters, A.; Cropp, T. A.; Summerer, D.; Mukherji, M. & Schultz, P. G. "Site-specific PEGylation of proteins containing unnatural amino acids." *Bioorg. Med. Chem. Lett.* **14,** 5743–5 (2004).

[107] Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H. & Tsien, R. Y. "The fluorescent toolbox for assessing protein location and function." *Science* **312,** 217–24 (2006).

[108] Tsien, R. Y. "The green fluorescent protein." *Annu. Rev. Biochem.* **67,** 509–44 (1998).

[109] Wang, L.; Brock, A. & Schultz, P. G. "Adding L-3-(2-Naphthyl)alanine to the genetic code of E. coli." *J. Am. Chem. Soc.* **124,** 1836–7 (2002).

[110] Summerer, D. *et al.* "A genetically encoded fluorescent amino acid." *Proc. Natl. Acad. Sci. U. S. A.* **103,** 9785–9 (2006).

[111] Wang, J.; Xie, J. & Schultz, P. G. "A genetically encoded fluorescent amino acid." *J. Am. Chem. Soc.* **128,** 8738–9 (2006).

[112] Lee, H. S.; Guo, J.; Lemke, E. a; Dimla, R. D. & Schultz, P. G. "Genetic incorporation of a small, environmentally sensitive, fluorescent probe into proteins in Saccharomyces cerevisiae." *J. Am. Chem. Soc.* **131,** 12921–3 (2009).

[113] Charbon, G. *et al.* "Localization of GroEL determined by in vivo incorporation of a fluorescent amino acid." *Bioorg. Med. Chem. Lett.* **21,** 6067–70 (2011).

[114] Baskin, J. M. & Bertozzi, C. R. "Bioorthogonal Click Chemistry: Covalent Labeling in Living Systems" *QSAR Comb. Sci.* **26,** 1211–1219 (2007).

[115] Kolb, H. C.; Finn, M. G. & Sharpless, K. B. "Click Chemistry: Diverse Chemical Function from a Few Good Reactions." *Angew. Chem. Int. Ed. Engl.* **40,** 2004–2021 (2001).

[116] Agard, N. J.; Prescher, J. A. & Bertozzi, C. R. "A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems." *J. Am. Chem. Soc.* **126,** 15046–7 (2004).

[117] Baskin, J. M. *et al.* "Copper-free click chemistry for dynamic in vivo imaging." *Proc. Natl. Acad. Sci. U. S. A.* **104,** 16793–7 (2007).

[118] Chakraborty, A. *et al.* "Opening and closing of the bacterial RNA polymerase clamp." *Science* **337,** 591–5 (2012).

[119] Plass, T. *et al.* "Amino acids for Diels-Alder reactions in living cells." *Angew. Chem. Int. Ed. Engl.* **51,** 4166–70 (2012).

[120] Lang, K. *et al.* "Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction." *Nat. Chem.* **4,** 298–304 (2012).

[121] Lang, K. *et al.* "Genetic Encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels-Alder reactions." *J. Am. Chem. Soc.* **134,** 10317–20 (2012).

[122] Weiss, S. "Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy." *Nat. Struct. Biol.* **7,** 724–9 (2000).

[123] Selvin, P. R. "The renaissance of fluorescence resonance energy transfer." *Nat. Struct. Biol.* **7,** 730–4 (2000).

[124] Brustad, E. M.; Lemke, E. A.; Schultz, P. G. & Deniz, A. A. "A general and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer." *J. Am. Chem. Soc.* **130,** 17664–5 (2008).

[125] Wurm, C. A. *et al.* "Novel red fluorophores with superior performance in STED microscopy" *Opt. Nanoscopy* **1,** 7 (2012).

[126] Dower, W. J.; Miller, J. F. & Ragsdale, C. W. "High efficiency transformation of E. coli by high voltage electroporation." *Nucleic Acids Res.* **16,** 6127–45 (1988).

[127] Laemmli, U. K. "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." *Nature* **227,** 680–5 (1970).

[128] Bradford, M. M. "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." *Anal. Biochem.* **72,** 248–54 (1976).

[129] Julius, M. H.; Masuda, T. & Herzenberg, L. A. "Demonstration that antigen-binding cells are precursors of antibody-producing cells after purification with a fluorescence-activated cell sorter." *Proc. Natl. Acad. Sci. U. S. A.* **69,** 1934–8 (1972).

[130] Kessler, C. & Manta, V. "Specificity of restriction endonucleases and DNA modification methyltransferases a review (Edition 3)." *Gene* **92,** 1–248 (1990).

[131] Pingoud, A. & Jeltsch, A. "Structure and function of type II restriction endonucleases." *Nucleic Acids Res.* **29,** 3705–27 (2001).

[132] Mullis, K. B. & Faloona, F. A. "Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction." *Methods Enzymol.* **155,** 335–50 (1987).

[133] Sharkey, D. J.; Scalice, E. R.; Christy, K. G.; Atwood, S. M. & Daiss, J. L. "Antibodies as thermolabile switches: high temperature triggering for the polymerase chain reaction." *Biotechnology. (N. Y).* **12,** 506–9 (1994).

[134] Ochman, H.; Gerber, A. S. & Hartl, D. L. "Genetic applications of an inverse polymerase chain reaction." *Genetics* **120,** 621–3 (1988).

[135] Williams, M.; Louw, A. I. & Birkholtz, L.-M. "Deletion mutagenesis of large areas in Plasmodium falciparum genes: a comparative study." *Malar. J.* **6,** 64 (2007).

[136] Don, R. H.; Cox, P. T.; Wainwright, B. J.; Baker, K. & Mattick, J. S. "“Touchdown” PCR to circumvent spurious priming during gene amplification." *Nucleic Acids Res.* **19,** 4008 (1991).

[137] Hecker, K. H. & Roux, K. H. "High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR." *Biotechniques* **20,** 478–85 (1996).

[138] Sanger, F. & Coulson, A. R. "A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase." *J. Mol. Biol.* **94,** 441–8 (1975).

[139] Sanger, F.; Nicklen, S. & Coulson, A. R. "DNA sequencing with chain-terminating inhibitors." *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–7 (1977).

[140] Chomczynski, P. & Sacchi, N. "Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction." *Anal. Biochem.* **162,** 156–9 (1987).

[141] Chomczynski, P. "A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples." *Biotechniques* **15,** 532–4, 536–7 (1993).

[142] Köhrer, C. & Rajbhandary, U. L. "The many applications of acid urea polyacrylamide gel electrophoresis to studies of tRNAs and aminoacyl-tRNA synthetases." *Methods* **44,** 129–38 (2008).

[143] Pelta, J.; Livolant, F. & Sikorav, J. L. "DNA aggregation induced by polyamines and cobalthexamine." *J. Biol. Chem.* **271,** 5656–62 (1996).

[144] Limsuwun, K. & Jones, P. G. "Spermidine acetyltransferase is required to prevent spermidine toxicity at low temperatures in Escherichia coli." *J. Bacteriol.* **182,** 5373–80 (2000).

[145] Santoro, S. W.; Wang, L.; Herberich, B.; King, D. S. & Schultz, P. G. "An efficient system for the evolution of aminoacyl-tRNA synthetase specificity." *Nat. Biotechnol.* **20,** 1044–8 (2002).

[146] Wang, L. & Schultz, P. G. "A general approach for the generation of orthogonal tRNAs." *Chem. Biol.* **8,** 883–90 (2001).

[147] Tropberger, P. *et al.* "Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer." *Cell* **152,** 859–72 (2013).

[148] Larkin, M. A. *et al.* "Clustal W and Clustal X version 2.0." *Bioinformatics* **23,** 2947–8 (2007).

[149] Consortium, T. U. "Update on activities at the Universal Protein Resource (UniProt) in 2013." *Nucleic Acids Res.* **41,** D43–7 (2013).

[150] Jiang, R. & Krzycki, J. A. "PylSn and the homologous N-terminal domain of pyrrolysyl-tRNA synthetase bind the tRNA that is essential for the genetic encoding of pyrrolysine." *J. Biol. Chem.* **287,** 32738–46 (2012).

[151] DeLano, W. L. "The PyMOL Molecular Graphics System, Version 1.1" *Schrödinger LLC* http://www.pymol.org (2002).

[152] Yanagisawa, T. *et al.* "Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification." *Chem. Biol.* **15,** 1187–97 (2008).

[153] Berman, H. M. *et al.* "The Protein Data Bank." *Nucleic Acids Res.* **28,** 235–42 (2000).

[154] Yanagisawa, T.; Ishii, R.; Fukunaga, R.; Nureki, O. & Yokoyama, S. "Crystallization and preliminary X-ray crystallographic analysis of the catalytic domain of pyrrolysyl-tRNA synthetase from the methanogenic archaeon Methanosarcina mazei." *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **62,** 1031–3 (2006).

[155] Herring, S. *et al.* "The amino-terminal domain of pyrrolysyl-tRNA synthetase is dispensable in vitro but required for in vivo activity." *FEBS Lett.* **581,** 3197–203 (2007).

[156] Gruber, A. R.; Lorenz, R.; Bernhart, S. H.; Neuböck, R. & Hofacker, I. L. "The Vienna RNA websuite." *Nucleic Acids Res.* **36,** W70–4 (2008).

[157] Altschul, S. F. *et al.* "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucleic Acids Res.* **25,** 3389–402 (1997).

[158] Alper, H.; Fischer, C.; Nevoigt, E. & Stephanopoulos, G. "Tuning genetic control through promoter engineering." *Proc. Natl. Acad. Sci. U. S. A.* **102,** 12678–83 (2005).

[159] Hammer, K.; Mijakovic, I. & Jensen, P. R. "Synthetic promoter libraries--tuning of gene expression." *Trends Biotechnol.* **24,** 53–5 (2006).

[160] Jensen, P. R. & Hammer, K. "The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters." *Appl. Environ. Microbiol.* **64,** 82–7 (1998).

[161] Winship, P. R. "An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide." *Nucleic Acids Res.* **17,** 1266 (1989).

[162] Hung, T.; Mak, K. & Fong, K. "A specificity enhancer for polymerase chain reaction." *Nucleic Acids Res.* **18,** 4953 (1990).

[163] Berezhnoy, A. Y. & Shckorbatov, Y. G. "Dependence of the E. coli promoter strength and physical parameters upon the nucleotide sequence." *J. Zhejiang Univ. Sci. B* **6,** 1063–8 (2005).

[164] Deuschle, U.; Kammerer, W.; Gentz, R. & Bujard, H. "Promoters of Escherichia coli: a hierarchy of in vivo strength indicates alternate structures." *EMBO J.* **5,** 2987–94 (1986).

[165] Liu, M.; Tolstorukov, M.; Zhurkin, V.; Garges, S. & Adhya, S. "A mutant spacer sequence between -35 and -10 elements makes the Plac promoter hyperactive and cAMP receptor protein-independent." *Proc. Natl. Acad. Sci. U. S. A.* **101,** 6911–6 (2004).

[166] Knaus, R. & Bujard, H. "PL of coliphage lambda: an alternative solution for an efficient promoter." *EMBO J.* **7,** 2919–23 (1988).

[167] Pédelacq, J.-D.; Cabantous, S.; Tran, T.; Terwilliger, T. C. & Waldo, G. S. "Engineering and characterization of a superfolder green fluorescent protein." *Nat. Biotechnol.* **24,** 79–88 (2006).

[168] Stokes, A. L. *et al.* "Enhancing the utility of unnatural amino acid synthetases by manipulating broad substrate specificity." *Mol. Biosyst.* **5,** 1032–8 (2009).

[169] Mukai, T. *et al.* "Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases." *Biochem. Biophys. Res. Commun.* **371,** 818–22 (2008).

[170] Fleet, G. W. J.; Porter, R. R. & Knowles, J. R. "Affinity Labelling of Antibodies with Aryl Nitrene as Reactive Group" *Nature* **224,** 511–512 (1969).

[171] Escher, E.; Jost, R.; Zuber, H. & Schwyzer, R. "p-Azido-L-Phenylalanine Peptides. I: Synthesis of Peptide Ligands for Chymotrypsin and Aminopeptidases." *Isr. J. Chem.* **12,** 129–138 (1974).

[172] Krieg, U. C.; Walter, P. & Johnson, A. E. "Photocrosslinking of the signal sequence of nascent preprolactin to the 54-kilodalton polypeptide of the signal recognition particle." *Proc. Natl. Acad. Sci. U. S. A.* **83,** 8604–8 (1986).

[173] Chen, Y.; Ebright, Y. W. & Ebright, R. H. "Identification of the target of a transcription activator protein by protein-protein photocrosslinking." *Science* **265,** 90–2 (1994).

[174] Sánchez, M. A. L. "Generation of tRNA Synthetases Capable of Incorporating Azido-phenylalanine For The Genetic Code Expansion of Escherichia coli." (2012).

[175] Motz, J. "Einbau von Azido-Phenylalanin und Acetyl- Lysin durch orthogonale evolvierte Translationskomponenten in Histon H3" (2012).

[176] Uemura, S. *et al.* "Real-time tRNA transit on single translating ribosomes at codon resolution." *Nature* **464,** 1012–7 (2010).

[177] Lee, G.; Bratkowski, M. A.; Ding, F.; Ke, A. & Ha, T. "Elastic coupling between RNA degradation and unwinding by an exoribonuclease." *Science* **336,** 1726–9 (2012).

[178] Choi, U. B. *et al.* "Single-molecule FRET-derived model of the synaptotagmin 1-SNARE fusion complex." *Nat. Struct. Mol. Biol.* **17,** 318–24 (2010).

[179] Blackman, M. L.; Royzen, M. & Fox, J. M. "Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity." *J. Am. Chem. Soc.* **130,** 13518–9 (2008).

[180] Devaraj, N. K.; Weissleder, R. & Hilderbrand, S. a. "Tetrazine-based cycloadditions: application to pretargeted live cell imaging." *Bioconjug. Chem.* **19,** 2297–9 (2008).

[181] Taylor, M. T.; Blackman, M. L.; Dmitrenko, O. & Fox, J. M. "Design and synthesis of highly reactive dienophiles for the tetrazine-trans-cyclooctene ligation." *J. Am. Chem. Soc.* **133,** 9646–9 (2011).

[182] Devaraj, N. K. & Weissleder, R. "Biomedical applications of tetrazine cycloadditions." *Acc. Chem. Res.* **44,** 816–27 (2011).

[183] Devaraj, N. K.; Hilderbrand, S.; Upadhyay, R.; Mazitschek, R. & Weissleder, R. "Bioorthogonal turn-on probes for imaging small molecules inside living cells." *Angew. Chem. Int. Ed. Engl.* **49,** 2869–72 (2010).

[184] Seitchik, J. L. *et al.* "Genetically encoded tetrazine amino acid directs rapid site-specific in vivo bioorthogonal ligation with trans-cyclooctenes." *J. Am. Chem. Soc.* **134,** 2898–901 (2012).

[185] Hamers-Casterman, C. *et al.* "Naturally occurring antibodies devoid of light chains." *Nature* **363,** 446–8 (1993).

[186] De Groeve, K. *et al.* "Nanobodies as tools for in vivo imaging of specific immune cell types." *J. Nucl. Med.* **51,** 782–9 (2010).

[187] Arbabi Ghahroudi, M.; Desmyter, A.; Wyns, L.; Hamers, R. & Muyldermans, S. "Selection and identification of single domain antibody fragments from camel heavy-chain antibodies." *FEBS Lett.* **414,** 521–6 (1997).

[188] Muyldermans, S. "Nanobodies: natural single-domain antibodies." *Annu. Rev. Biochem.* **82,** 775–97 (2013).

[189] Saerens, D. *et al.* "Single domain antibodies derived from dromedary lymph node and peripheral blood lymphocytes sensing conformational variants of prostate-specific antigen." *J. Biol. Chem.* **279,** 51965–72 (2004).

[190] Cortez-Retamozo, V. *et al.* "Efficient cancer therapy with a nanobody-based conjugate." *Cancer Res.* **64,** 2853–7 (2004).

[191] Huang, L. *et al.* "SPECT imaging with 99mTc-labeled EGFR-specific nanobody for in vivo monitoring of EGFR expression." *Mol. Imaging Biol.* **10,** 167–75 (2008).

[192] Rothbauer, U. *et al.* "A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins." *Mol. Cell. Proteomics* **7,** 282–9 (2008).

[193] Hell, S. W. & Wichmann, J. "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy." *Opt. Lett.* **19,** 780–2 (1994).

[194] Klar, T. A. & Hell, S. W. "Subdiffraction resolution in far-field fluorescence microscopy." *Opt. Lett.* **24,** 954–6 (1999).

[195] De Genst, E. J. *et al.* "Structure and properties of a complex of α-synuclein and a single-domain camelid antibody." *J. Mol. Biol.* **402,** 326–43 (2010).

[196] Bessette, P. H.; Aslund, F.; Beckwith, J. & Georgiou, G. "Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm." *Proc. Natl. Acad. Sci. U. S. A.* **96,** 13703–8 (1999).

[197] Qiu, J.; Swartz, J. R. & Georgiou, G. "Expression of active human tissue-type plasminogen activator in Escherichia coli." *Appl. Environ. Microbiol.* **64,** 4891–6 (1998).

[198] Levy, R.; Weiss, R.; Chen, G.; Iverson, B. L. & Georgiou, G. "Production of correctly folded Fab antibody fragment in the cytoplasm of Escherichia coli trxB gor mutants via the coexpression of molecular chaperones." *Protein Expr. Purif.* **23,** 338–47 (2001).

[199] Selenko, P. & Wagner, G. "NMR mapping of protein interactions in living cells." *Nat. Methods* **3,** 80–1 (2006).

[200] Sali, A.; Glaeser, R.; Earnest, T. & Baumeister, W. "From words to literature in structural proteomics." *Nature* **422,** 216–25 (2003).

[201] Serber, Z. & Dötsch, V. "In-cell NMR spectroscopy." *Biochemistry* **40,** 14317–23 (2001).

[202] Groll, M.; Bochtler, M.; Brandstetter, H.; Clausen, T. & Huber, R. "Molecular machines for protein degradation." *Chembiochem* **6,** 222–56 (2005).

[203] Burz, D. S.; Dutta, K.; Cowburn, D. & Shekhtman, A. "Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR)." *Nat. Methods* **3,** 91–3 (2006).

[204] Dedmon, M. M.; Patel, C. N.; Young, G. B. & Pielak, G. J. "FlgM gains structure in living cells." *Proc. Natl. Acad. Sci. U. S. A.* **99,** 12681–4 (2002).

[205] Bryant, J. E.; Lecomte, J. T. J.; Lee, A. L.; Young, G. B. & Pielak, G. J. "Protein dynamics in living cells." *Biochemistry* **44,** 9275–9 (2005).

[206] Li, C. *et al.* "Differential dynamical effects of macromolecular crowding on an intrinsically disordered protein and a globular protein: implications for in-cell NMR spectroscopy." *J. Am. Chem. Soc.* **130,** 6310–1 (2008).

[207] Serber, Z.; Corsini, L.; Durst, F. & Dötsch, V. "In-cell NMR spectroscopy." *Methods Enzymol.* **394,** 17–41 (2005).

[208] Steinmetz, M. O. *et al.* "Molecular basis of coiled-coil formation." *Proc. Natl. Acad. Sci. U. S. A.* **104,** 7062–7 (2007).

[209] Gottesman, S. "Proteases and their targets in Escherichia coli." *Annu. Rev. Genet.* **30,** 465–506 (1996).

[210] Cheung, A. Y.; Watson, L. & Söll, D. "Two control systems modulate the level of glutaminyl-tRNA synthetase in Escherichia coli." *J. Bacteriol.* **161,** 212–8 (1985).

[211] Miksch, G.; Bettenworth, F.; Friehs, K. & Flaschel, E. "The sequence upstream of the -10 consensus sequence modulates the strength and induction time of stationary-phase promoters in Escherichia coli." *Appl. Microbiol. Biotechnol.* **69,** 312–20 (2005).

[212] Polycarpo, C. R. *et al.* "Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase." *FEBS Lett.* **580,** 6695–700 (2006).

[213] Wang, K. *et al.* "Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET." *Nat. Chem.* **6,** 393–403 (2014).

8 Appendix

8.1 Plasmids

Table 8.1: List of plasmids.
Details for the construction of plasmids made during this thesis (method, primers, template etc.) are listed in Table 8.3. Plasmids labeled with “Cambridge” were kindly provided by the Chin laboratory (MRC Laboratory of Molecular Biology, Cambridge). The remaining plasmids were made by other group members of AG Neumann as indicated.

No.	Name	Resistance	Source
pCLA1	pBK PylS	Kan	Cambridge
pCLA2	pBK MjYRS_TAG	Amp	Cambridge
pCLA3	pBK MjYRS_AGGA (2D12)	Amp	Cambridge
pCLA4	pREP PylT	Tet	Cambridge
pCLA5	pREP YC JYCUA (MjYT_CUA)	Tet	Cambridge
pCLA6	pREP YC JYCUA AGGA2 (MjYT_UCCU)	Tet	Cambridge
pCLA7	pBK His$_6$ PylS	Kan	Thesis
pCLA8	pBK H6C PylS	Kan	Thesis
pCLA9	pBK H6 34AA lack to mazei (gap) PylS	Kan	Thesis
pCLA10	pBK H6 Loop206-216 in mazei PylS	Kan	Thesis
pCLA11	pBK H6 PylS linker delta 10	Kan	Thesis
pCLA12	pBK H6 PylS linker delta 20	Kan	Thesis
pCLA13	pBK H6 PylS linker delta 15	Kan	Thesis
pCLA14	pBK H6 PylS linker delta 30	Kan	Thesis
pCLA15	pBK H6 PylS linker delta 50	Kan	Thesis
pCLA16	pBK H6 PylS linker delta C-terminus	Kan	Thesis
pCLA17	pBK H6 PylS linker delta N-terminus	Kan	Thesis
pCLA18	pBK His$_6$ MjYRS_TAG	Amp	Thesis
pCLA19	pBK H6C MjYRS_TAG	Amp	Thesis
pCLA20	pBK H6 Loop25-29 MjYRS_TAG	Amp	Thesis
pCLA21	pDULE BPA (BPA aaRS + MjYT_CUA)	Tet	Cambridge
pCLA22	pBK H6 gap PylS_Prom Lib 1D	Kan	Thesis
pCLA23	pBK H6 gap PylS_Prom Lib 2E	Kan	Thesis
pCLA24	pBK H6 gap PylS_Prom Lib 3E	Kan	Thesis
pCLA25	pBK H6 gap PylS_Prom Lib 4F	Kan	Thesis
pCLA26	pBK H6 gap PylS_Prom Lib 6B	Kan	Thesis
pCLA27	pBK H6 gap PylS_Prom Lib 7A	Kan	Thesis
pCLA28	pBK H6 gap PylS_Prom Lib 8G	Kan	Thesis
pCLA29	pBK H6 gap PylS_Prom Lib 10D	Kan	Thesis
pCLA30	pBK H6 gap PylS_Prom Lib 11D	Kan	Thesis
pCLA31	pBK H6 gap PylS_Prom Lib 12H	Kan	Thesis
pCLA32	pCDFDuet-PylT H3 R52TAG	Sm	Liljan Hahn
pCLA33	pBK H6C MjYRS_AGGA_Prom Lib 1D	Amp	Thesis
pCLA34	pBK H6C MjYRS_AGGA_Prom Lib 2F	Amp	Thesis
pCLA35	pBK H6C MjYRS_AGGA_Prom Lib 3B	Amp	Thesis
pCLA36	pBK H6C MjYRS_AGGA_Prom Lib 4A	Amp	Thesis
pCLA37	pBK H6C MjYRS_AGGA_Prom Lib 6D	Amp	Thesis
pCLA38	pBK H6C MjYRS_AGGA_Prom Lib 7E	Amp	Thesis
pCLA39	pBK H6C MjYRS_AGGA_Prom Lib 8G	Amp	Thesis
pCLA40	pBK H6C MjYRS_AGGA_Prom Lib 9F	Amp	Thesis
pCLA41	pBK H6C MjYRS_AGGA_Prom Lib 10B	Amp	Thesis

pCLA42	pBK H6C MjYRS_AGGA_Prom Lib 11E	Amp	Thesis
pCLA43	pMyo4AGGA MjtRNAYuccu-3	Tet	Cambridge
pCLA44	pREP PylT short_Prom Lib A4	Tet	Thesis
pCLA45	pREP PylT short_Prom Lib C2	Tet	Thesis
pCLA46	pREP PylT short_Prom Lib E7	Tet	Thesis
pCLA47	pREP PylT short_Prom Lib F5	Tet	Thesis
pCLA48	pREP PylT short_Prom Lib I3	Tet	Thesis
pCLA49	pREP PylT short_Prom Lib I6	Tet	Thesis
pCLA50	pREP PylT short_Prom Lib J7	Tet	Thesis
pCLA51	pREP PylT short_Prom Lib K2	Tet	Thesis
pCLA52	pREP PylT short_Prom Lib L5	Tet	Thesis
pCLA53	pREP MjYT_UCCU short_Prom Lib B3	Tet	Thesis
pCLA54	pREP MjYT_UCCU short_Prom Lib B8	Tet	Thesis
pCLA55	pREP MjYT_UCCU short_Prom Lib D4	Tet	Thesis
pCLA56	pREP MjYT_UCCU short_Prom Lib E1	Tet	Thesis
pCLA57	pREP MjYT_UCCU short_Prom Lib G7	Tet	Thesis
pCLA58	pREP MjYT_UCCU short_Prom Lib I2	Tet	Thesis
pCLA59	pREP MjYT_UCCU short_Prom Lib I8	Tet	Thesis
pCLA60	pREP MjYT_UCCU short_Prom Lib K6	Tet	Thesis
pCLA61	pREP MjYT_UCCU short_Prom Lib L1	Tet	Thesis
pCLA62	pBK H6 gap PylS_con Prom	Kan	Thesis
pCLA63	pBK H6 gap PylS_cp25 Prom	Kan	Thesis
pCLA64	pBK H6 gap PylS_lac1-6 Mt5 Prom	Kan	Thesis
pCLA65	pBK H6 gap PylS_lacP1 Prom	Kan	Thesis
pCLA66	pBK H6 gap PylS_lambda ConN25 Prom	Kan	Thesis
pCLA67	pBK H6 gap PylS_lambda Prom	Kan	Thesis
pCLA68	pBK H6 gap PylS_T5 D/E20 Prom	Kan	Thesis
pCLA69	pBK H6 gap PylS_T5 H207 Prom	Kan	Thesis
pCLA70	pBK H6 gap PylS_tacI Prom	Kan	Thesis
pCLA71	pBK H6C MjYRS_AGGA (2D12)	Amp	Thesis
pCLA72	pBK H6C MjYRS_AGGA_SalI-Term	Amp	Thesis
pCLA73	pBK H6 gap PylS_NdeI to KpnI	Kan	Thesis
pCLA74	pBK H6 gap PylS-KpnI_PstI to SacI	Kan	Thesis
pCLA75	pBK H6 gap PylS-KpnI-SacI_SalI	Kan	Thesis
pCLA76	pBK H6 gap PylS-KpnI-SacI-SalI_NotI	Kan	Thesis
pCLA77	pREP PylT short	Tet	Thesis
pCLA78	pREP PylT short_NotI	Tet	Thesis
pCLA79	pREP PylT short-NotI_MfeI	Tet	Thesis
pCLA80	pREP MjYT_UCCU short	Tet	Thesis
pCLA81	pREP MjYT_UCCU short_MfeI	Tet	Thesis
pCLA82	pREP MjYT_UCCU short-MfeI_XhoI	Tet	Thesis
pCLA83	pBK H6 gap PylS_SalI	Kan	Thesis
pCLA84	pBK H6 gap PylS_NotI	Kan	Thesis
pCLA85	pBK H6 gap PylS_NdeI-restore	Kan	Thesis
pCLA86	pBK H6 gap PylS_NdeI-restore_PM1	Kan	Thesis
pCLA87	pBK H6 gap PylS_NdeI-restore_PM2	Kan	Thesis
pCLA88	pBK H6 gap PylS_NdeI-restore_PM3	Kan	Thesis
pCLA89	pBK H6 gap PylS_NdeI-restore_PM4	Kan	Thesis
pCLA90	pBK H6 gap PylS_NdeI-restore_PM5	Kan	Thesis
pCLA91	pCDF Duet-1	Sm	Novagen
pCLA92	pCDF_H6C MjYRS_AGGA	Sm	Thesis

pCLA93	pCDF_H6C MjYRS_H6 gap PylS	Sm	Thesis
pCLA94	pCDF_H6C MjYRS_H6 gap PylS_PylT	Sm	Thesis
pCLA95	pCDF_H6C MjYRS_H6 gap PylS_PylT_MjYT_UCCU	Sm	Thesis
pCLA96	pSC101*-ribo-Q1	Kan	Cambridge
pCLA97	pCDF PylST	Sm	Cambridge
pCLA98	pDULE MjYRST (AGGA)	Tet	Cambridge
pCLA99	pO-*gst-malE*	Amp	Cambridge
pCLA100	pO-delta-*gst-malE*_MCS	Amp	Thesis
pCLA101	pBAD sfGFP WT	Amp	Cambridge
pCLA102	pBAD sfGFP 150TAG	Amp	Cambridge
pCLA103	pBAD sfGFP 134AGGA 150TAG-1	Amp	Cambridge
pCLA104	pBAD sfGFP 134AGGA-1	Amp	Cambridge
pCLA105	pO-sfGFP WT	Amp	Thesis
pCLA106	pO-sfGFP 150TAG	Amp	Thesis
pCLA107	pO-sfGFP 134AGGA 150TAG-1	Amp	Thesis
pCLA108	pO-sfGFP 134AGGA-1	Amp	Thesis
pCLA109	pCDF_MjYRS_PylS_PylT_MjYT_UCCU-mismatch	Sm	Thesis
pCLA110	pREP YC JYCUA delta tRNA	Tet	Cambridge
pCLA111	pO-CAT	Tet	Cambridge
pCLA112	pO-CAT T6TAG	Tet	Thesis
pCLA113	pO-CAT T6TAG D111AGGA	Tet	Thesis
pCLA114	pYOBB2 PylT-1	Cm	Cambridge
pCLA115	pCDF_GTF_PylS_Prom Lib 3E	Sm	Thesis
pCLA116	pCDF_GTF_PylS_con Prom	Sm	Thesis
pCLA117	pCDF_GTF_PylS_lacP1 Prom	Sm	Thesis
pCLA118	G9GSTMBP Y17AGGA N234TAG-1	Amp	Cambridge
pCLA119	pDULE AzFRS (TAG)	Tet	Cambridge
pCLA120	pBK AzFRS (TAG)	Kan	Thesis
pCLA121	pBK AzFRS_AGGA Lib1	Kan	Thesis
pCLA122	pBK AzFRS_AGGA Lib2	Kan	Thesis
pCLA123	pBK AzFRS_AGGA Lib3	Kan	Thesis
pCLA124	pBK AzFRS_AGGA Lib4	Kan	Thesis
pCLA125	pBK AzFRS_AGGA Lib5	Kan	Thesis
pCLA126	pBK AzFRS_AGGA Lib6	Kan	Thesis
pCLA127	pBK AzFRS_AGGA Lib7	Kan	Thesis
pCLA128	pBK AzFRS_AGGA Lib8	Kan	Thesis
pCLA129	pBK AzFRS_AGGA Lib9	Kan	Thesis
pCLA130	pBK AzFRS_AGGA Lib10	Kan	Thesis
pCLA131	pBK AzFRS_AGGA Lib11	Kan	Thesis
pCLA132	pBK AzFRS_AGGA Lib12	Kan	Thesis
pCLA133	pBK AzFRS_AGGA Lib13	Kan	Thesis
pCLA134	pBK AzFRS_AGGA Lib14	Kan	Thesis
pCLA135	pBK AzFRS_AGGA Lib15	Kan	Thesis
pCLA136	pBK AzFRS_AGGA Lib16	Kan	Thesis
pCLA137	pBK AzFRS_AGGA Lib17	Kan	Thesis
pCLA138	pBK AzFRS_AGGA Lib18	Kan	Thesis
pCLA139	pBK AzFRS_AGGA Lib19	Kan	Thesis
pCLA140	pBK AzFRS_AGGA Lib20	Kan	Thesis
pCLA141	pCDF Duet-PylT H3opt Q76C (pCHR029)	Sm	C. Hoffmann
pCLA142	pCDF Duet-PylT H3opt Q76C_T6AGGA	Sm	Thesis
pCLA143	pO-H3opt Q76C_T6AGGA	Amp	Thesis

pCLA144	pDULE AzFRS_AGGA Lib2	Tet	Thesis
pCLA145	pDULE AzFRS_AGGA Lib3	Tet	Thesis
pCLA146	pDULE AzFRS_AGGA Lib5	Tet	Thesis
pCLA147	pDULE AzFRS_AGGA Lib6	Tet	Thesis
pCLA148	pDULE AzFRS_AGGA Lib11	Tet	Thesis
pCLA149	pDULE AzFRS_AGGA Lib17	Tet	Thesis
pCLA150	pDULE AzFRS_AGGA Lib18	Tet	Thesis
pCLA151	pBK AzFRS_BamHI_SalI (TAG)	Kan	K. Halder
pCLA152	pCDF_AzFRS_BamHI_SalI (TAG)	Sm	Thesis
pCLA153	pCDF_AzFRS_BamHI_SalI (TAG)_PylT_MjYT_UCCU	Sm	Thesis
pCLA154	pCDF_PylT_MjYT_UCCU_AzFRS (AGGA)	Sm	Thesis
pCLA155	pCDF_PylT_MjYT_UCCU_AzFRS (AGGA)_PylS	Sm	Thesis
pCLA156	pMyo4TAG PylT	Tet	Cambridge
pCLA157	pMyo 4K PylT-1 (amber reverted to AAA (Lys))	Tet	Cambridge
pCLA158	pBK BCNRS	Kan	H. Neumann
pCLA159	pBK pcKRS-1	Amp	Cambridge
pCLA160	pBK AzpcK2-3	Amp	Cambridge
pCLA161	pBK AcKRS1	Kan	Cambridge
pCLA162	pBK AcKRS3	Kan	Cambridge
pCLA163	pCDF PylT1 H3K56TAG-1	Sm	Cambridge
pCLA164	pCDF Duet-PylT H3opt V35TAG F78C C110A	Sm	Liljan Hahn
pCLA165	pCDF_PylT_MjYT_UCCU_AzFRS (AGGA)_BCNRS	Sm	Thesis
pCLA166	pCDF_PylT_MjYT_UCCU_AzFRS (AGGA)_AzpcKRS	Sm	Thesis
pCLA167	pBK BCNRS_con Prom	Kan	Thesis
pCLA168	pBK BCNRS_lacP1 Prom	Kan	Thesis
pCLA169	pBK AzpcKRS_con Prom	Kan	Thesis
pCLA170	pBK AzpcKRS_lacP1 Prom	Kan	Thesis
pCLA171	pCDF_PylT_MjYT_AzFRS_BCNRS_con Prom	Sm	Thesis
pCLA172	pCDF_PylT_MjYT_AzFRS_BCNRS_lacP1 Prom	Sm	Thesis
pCLA173	pCDF_PylT_MjYT_AzFRS_AzpcKRS_con Prom	Sm	Thesis
pCLA174	pCDF_PylT_MjYT_AzFRS_AzpcKRS_lacP1 Prom	Sm	Thesis
pCLA175	pHEN6_NbSyn2_Thrombin_His$_6$-tag	Amp	AG Rizzoli
pCLA176	pHEN6_NbSyn2_Thrombin_His$_6$-tag_S44TAG	Amp	AG Rizzoli
pCLA177	pHEN6_NbSyn2_Thrombin_His$_6$-tag_K70TAG	Amp	AG Rizzoli
pCLA178	pHEN6_NbSyn2_Thrombin_His$_6$-tag_R72TAG	Amp	AG Rizzoli
pCLA179	pHEN6_NbSyn2_Thrombin_His$_6$-tag_K93TAG	Amp	AG Rizzoli
pCLA180	pHEN6_NbSyn2_Thrombin_His$_6$-tag_K115TAG	Amp	AG Rizzoli
pCLA181	pHEN6_NbSyn2_Thrombin_His$_6$-tag_K153TAG	Amp	AG Rizzoli
pCLA182	pACYC Duet	Cm	Novagen
pCLA183	pET Duet	Amp	Novagen
pCLA184	pCDF_GTF_PylS_con Prom_CatR	Cm	Thesis
pCLA185	pCDF_GTF_PylS_con Prom_AmpR	Amp	Thesis
pCLA186	pCDF_PylT_MjYT_AzFRS_BCNRS_con Prom_CatR	Cm	Thesis
pCLA187	pCDF_PylT_MjYT_AzFRS_BCNRS_con Prom_AmpR	Amp	Thesis
pCLA188	K63_GFP_VHH_optimized_Nanobody	Kan	AG Rizzoli
pCLA189	K63_GFP_VHH_optimized_Nanobody_Q13TAG	Kan	AG Rizzoli
pCLA190	K63_GFP_VHH_optimized_Nanobody_S17TAG	Kan	AG Rizzoli
pCLA191	K63_GFP_VHH_optimized_Nanobody_R19TAG	Kan	AG Rizzoli
pCLA192	K63_GFP_VHH_optimized_Nanobody_R76TAG	Kan	AG Rizzoli
pCLA193	pBAD_ORBS_sfGFP WT	Amp	Thesis
pCLA194	pBAD_ORBS_GCN4p 16-31	Amp	Thesis

8.2 Oligonucleotides

Oligonucleotides like primers and probes for RNA hybridization were ordered from Sigma-Aldrich (Steinheim) and dissolved with ddH$_2$O as indicated on manufacturer's technical datasheet to a final concentration of 10µM.

Table 8.2: List of oligos (primers and probes).
For/f = forward; rev/r = reverse.

No.	Name	Sequence 5' → 3'
C1	H6PylSf	TGCATCACCATCACCACCATGATAAAAAACCGCTG
C2	H6PylSr	TCATGGTGGTGATGGTGATGCATATGGGATTCCTC
C3	pBKseqf	CTCGGGTTGTCAGCCTGTC
C4	His$_6$cPylSf	TGCACCATCACCATCATCACTAACTGCAGTTTCAAAC
C5	His$_6$cPylSr	TAGTGATGATGGTGATGGTGCAGGTTCGTGCTAATG
C6	pBKseqr	CAGAGATCATGTAGGCCTG
C7	H6Lack34PylSf	GCGGGCACCATCACCATCATCACGGGAGCGTTCCGGCGTC
C8	H6Lack34PylSr	CTCCCGTGATGATGGTGATGGTGCCCGCTGTTCGGGGTG
C9	H6LoopPylSf	AAGGGCACCATCACCATCATCACGGGATGGCGAAACCGTTTC
C10	H6LoopPylSr	ATCCCGTGATGATGGTGATGGTGCCCTTTATCTTCCGGAGAC
C11	PylS_H6L_d10_f	GAAAATAGCGTGAGCGTTCCGAGCCCGGCGAAAAG
C12	PylS_H6L_d10_r	GGCTCGGAACGCTCACGCTATTTTCCAGCGGTTTC
C13	PylS_H6L_d20_f	AAATAGCGTGAACAGCGGGCACCATCACCATCATC
C14	PylS_H6L_d20_r	GCCCGCTGTTCACGCTATTTTCCAGCGGTTTCGGCGCAC
C15	PylS_H6L_d15_f	CACGGGAGCGTTCGTGTGGAAGCGCTGCTGTCTCCGGAAG
C16	PylS_H6L_d15_r	CTTCCACACGAACGCTCCCGTGATGATGGTGATGGTG
C17	PylS_H6L_d30_f	CACGGGAGCGTTATGGCGAAACCGTTTCGTGAACTG
C18	PylS_H6L_d30_r	GTTTCGCCATAACGCTCCCGTGATGATGGTGATGGTG
C19	PylS_H6L_d50_f	TGAACAGCGGGCACCATCACCATCATCACGGGAGCGTTATGGCGAAACCGTTTCGTGAAC
C20	PylS_H6L_d50_r	ATAACGCTCCCGTGATGATGGTGATGGTGCCCGCTGTTCACGCTATTTTCCAGCGGTTTC
C21	PylS_H6L_dCT_f	CACGGGAGCGTTTAACTGCAGTTTCAAACGCTAAATTG
C22	PylS_H6L_dCT_r	ACTGCAGTTAAACGCTCCCGTGATGATGGTGATGGTG
C23	PylS_H6L_dNT_f	ATCCCATATGAACAGCGGGCACCATCACCATCATC

C24	PylS_H6L_dNT_r	GTGCCCGCTGTTCATATGGGATTCCTCAAAGCGTAAAC
C25	Mbarkeri_PylT_Probe	[CY3]GATTTAGAGTCCATTCGATC
C26	Mbarkeri_PylT_DIG_Probe	[DIG]GATTTAGAGTCCATTCGATC
C27	PylT_13-42	ATGTAGATCGAATGGACTCTAAATCCGTTC
C28	T7P_PylT_for	GAAATTAATACGACTCACTATAGGGAACCTGATCATGTAG
C29	T7P_PylT_rev	TGGCGGAAACCCCGGGAATCTAAC
C30	T7P_EcSerT_for	GAAATTAATACGACTCACTATAGGGAGTGTGGCCGAGCGGTTG
C31	T7P_EcSerT_rev	TGGCGGGAGCGCAGAGATTCGAAC
C32	Ecoli_serT_Probe	[CY3]GTTTTCAAGACCGGTGCCTT
C33	Ecoli_serT_DIG_Probe	[DIG]GTTTTCAAGACCGGTGCCTT
C34	H6JYRSf	TGCATCACCATCACCACCATGACGAATTTGAAATG
C35	H6JYRSr	TCATGGTGGTGATGGTGATGCATATGGGATTCCTC
C36	His$_6$cJYRSf	TACACCATCACCATCATCACTAACTGCAGTTTCAAAC
C37	His$_6$cJYRSr	TAGTGATGATGGTGATGGTGTAATCTCTTTCTAATTG
C38	H6LoopJYRSf	ATGGGCACCATCACCATCATCACGGGGAAAAATCTGCTTACATAG
C39	H6LoopJYRSr	TCCCCGTGATGATGGTGATGGTGCCCATCTTTTTTTAAAACCTCTC
C40	MjYT_CUA_Probe	[CY5]GATTTAGAGTCCGCCGTTCT
C41	MjYT_CUA_DIG_Probe	[DIG]GATTTAGAGTCCGCCGTTCT
C42	T7P_MjTyrT_for	GAAATTAATACGACTCACTATAGGGGCGGTAGTTCAGCAGGGCAG
C43	T7P_MjTyrT_rev	TGGTGGGGCGGCCGGATTTG
C44	MjYT_AGGA_DIG_Probe1	[DIG]GGATTAGGAAGTCCGCCGTTCT
C45	MjYT_AGGA_DIG_Probe2	[DIG]CAGCGCCATGCGGATTAGGAAG
C46	MjYT_AGGA_DIG_Probe3	[DIG]GCCATGCGGATTAGGAAGTCCG
C47	MjYT_AGGA_DIG_Probe4	[DIG]GCCATGCGGATTAGGA
C48	PylS_glnS_lib_f	GCGCAGGAAAGGTCTCATATACGTTGTTTACGCTTTGAGGAATC
C49	PylS_glnS_lib_r	GCGCAGAGTAGGTCTCATATAACNNNNNATGATCNNNNNNNNNNNNNNNNNTGACAANNNNNGTCGACGGATCCTTCAACTCAGCAAAAG
C50	pBK_PylS_SalI_seq_f	GGTTGTAACACTGGCAGAG
C51	MJYRS_glnS_lib_r	GCGCAGAGTAGGTCTCATATAACNNNNNATGATCNNNNNNNNNNNNNNNNNTGACAANNNNNGAGGATCCTCTGACGCTCAGTGGAAC
C52	pBK_MjYRS-lib_seqf	TAAAAGGATCTAGGTGAAG
C53	pBK_NotI_SalI_seq_r	CTCGTCAGGGGGGCGGAGC
C54	PylT_lpp_lib_f	GCGCAGGAAAGGTCTCAACGCTAGATCTGGGAACCTGATCATG
C55	PylT_lpp_lib_r	GCGCAGAGTAGGTCTCAGCGTTANNNNNATTACACAANNNNNNNNNNNNTGAGAANNNNNTTTTGATGGGGCGCCACTTATTTTTG
C56	MJYT_lpp_lib_f	GCGCAGGAAAGGTCTCAACGCTGAATTCCCGGCGGTAGTTC
C57	pREP_tRNA_seq_r	GGGAGAGCGTCTGGCGAAAG

C58	pBK_conProm_f	GCGCAGGAAAGGTCTCACGGTTATAATGGTACCATAAGGAGGTGGATCCGGCAGTTTACGCTTTGAGGAATC
C59	pBK_conProm_r	GCGCAGAGTAGGTCTCAACCGCCAAGCTTAAAAATGTCAACAACGACGGTGAATTGTCGACGGATCCTTCAACTC
C60	pBK_CP25Prom_f	GCGCAGGAAAGGTCTCAGAGGGGGCTGGTATAATCACATAGTACTGTTGTTTACGCTTTGAGGAATC
C61	pBK_CP25Prom_r	GCGCAGAGTAGGTCTCACCTCACTACATGTCAAGAATAAACTGCCAAAGGTCGACGGATCCTTCAACTC
C62	pBK_lac1-6Mt5Prom_f	GCGCAGGAAAGGTCTCATGTTTATGTTGTGTGGAATTGTGAGCGGATAACAATGTTTACGCTTTGAGGAATC
C63	pBK_lac1-6Mt5Prom_r	GCGCAGAGTAGGTCTCAAACATAAATTGCTTAAAGTGTAAAGCCTGGGGTGCCTAAGTCGACGGATCCTTCAACTC
C64	pBK_lacP1Prom_f	GCGCAGGAAAGGTCTCACTGGTATGTTGTGTGGAATTGTGAGCGGATAACAATGTTTACGCTTTGAGGAATC
C65	pBK_lacP1Prom_r	GCGCAGAGTAGGTCTCACCAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAGTCGACGGATCCTTCAACTC
C66	pBK_lambda_ConN25Prom_f	GCGCAGGAAAGGTCTCACGGTTATAATGAGCACATAAATTTGAGAGAGGAGTTGTTTACGCTTTGAGGAATC
C67	pBK_lambdaProm_f	GCGCAGGAAAGGTCTCACGGTGATACTGAGCACATCAGCAGGACGCACTGACCGTTTACGCTTTGAGGAATC
C68	pBK_lambdaProm_r	GCGCAGAGTAGGTCTCAACCGCCAGTGGTATTTATGTCAACACCGCCAGAGATAAGTCGACGGATCCTTCAACTC
C69	pBK_T5DE20Prom_f	GCGCAGGAAAGGTCTCAGCTTTAAGATGTACCCAGTTCGATGAGAGCGATAACGTTTACGCTTTGAGGAATC
C70	pBK_T5DE20Prom_r	GCGCAGAGTAGGTCTCAAAGCCTATCGGCTAGGGTGTCAAACTATTTTTGCAGTTGTCGACGGATCCTTCAACTC
C71	pBK_T5H207Prom_f	GCGCAGGAAAGGTCTCACTCGTATAATATACTTCATAAATTGATAAACAAAAAGTTTACGCTTTGAGGAATC
C72	pBK_T5H207Prom_r	GCGCAGAGTAGGTCTCACGAGAATTTGAAGCGTTTAGCAAATGAATTTTTTAAAAGTCGACGGATCCTTCAACTC
C73	pBK_tacIProm_f	GCGCAGGAAAGGTCTCACTCGTATAATGTGTGGAATTGTGAGCGGATAACAATGTTTACGCTTTGAGGAATC
C74	pBK_tacIProm_r	GCGCAGAGTAGGTCTCACGAGCCGATGATTAATTGTCAACAGCTCATTTCAGAAGTCGACGGATCCTTCAACTC
C75	JYRS_2D12_SalI_f	GAATTAATTGTCGACCCGCTTCGCAACATGTGAG
C76	JYRS_2D12_SalI_r	GCGAAGCGGGTCGACAATTAATTCGCGAAGAAAAG

C77 PylS_H6L_NdeI_KpnI_f	GAGGAATCCGGTACCATGGATAAAAAACCGCTGGATG
C78 PylS_H6L_NdeI_KpnI_r	TTTATCCATGGTACCGGATTCCTCAAAGCGTAAAC
C79 PylS_H6L_PstI_SacI_f	AACCTGTAAGAGCTCTTTCAAACGCTAAATTGCCTG
C80 PylS_H6L_PstI_SacI_r	CGTTTGAAAGAGCTCTTACAGGTTCGTGCTAATG
C81 PylS_H6L_SalI_f	GAAGGATCCGTCGACTCGGGTTGTCAGCCTGTCCCGCTTATAAG
C82 PylS_H6L_SalI_r	ACAACCCGAGTCGACGGATCCTTCAACTCAGCAAAAG
C83 PylS_H6L_NotI_f	ATTAATTGGCGGCCGCCCGCTTCGCACATGTGAGCAAAAG
C84 PylS_H6L_NotI_r	CGAAGCGGGCGGCCGCCAATTAATTCGCGAAGAAAAG
C85 pREP_PylT_short_ORI_f	TACTCTCGCATGGCCAGATGGTAAGCCCTCCCGTATC
C86 pREP_PylT_short_ORI_r	TTACCATCTGGCCATGCGAGAGTAGGGAACTGCCAG
C87 pREP_PylT_NotI_f	AACGATCGGCGGCCGCAAAAATAAGTGGCGCCCCATC
C88 pREP_PylT_NotI_r	TATTTTTGCGGCCGCCGATCGTTCGCTCAAAGAAG
C89 pREP_tRNA_seq_f	GAGGCCGCTCATGGCGTTC
C90 pREP_PylT_MfeI_f	GCCGTCGTTCAATTGTTACAACGTCGTGACTGGGAAAAC
C91 pREP_PylT_MfeI_r	ACGTTGTAACAATTGAACGACGGCCAGTGCCAAG
C92 pREP_YCJYcua_AGGA2_MfeI_f	GACTGTTGTCAATTGAACTCAGAATAAGAAATGAG
C93 pREP_YCJYcua_AGGA2_MfeI_r	TTCTGAGTTCAATTGACAACAGTCCGCACCGCTGCCGGTAG
C94 pREP_YCJYcua_AGGA2_XhoI_f	GTCGTTTTACTCGAGCAACGTCGTGACTGGGAAAAC
C95 pREP_YCJYcua_AGGA2_XhoI_r	ACGACGTTGCTCGAGTAAAACGACGGCCAGTGCCAAG
C96 PylS_NdeI-Restore_f	GAGGAATCCCATATGATGGATAAAAAACCGCTGGATG
C97 PylS_NdeI-Restore_r	TTTATCCATCATATGGGATTCCTCAAAGCGTAAAC
C98 PylS_NdeI-Restore_PM1_f	GAGGAATCCTATATGATGGATAAAAAACCGCTGGATG
C99 PylS_NdeI-Restore_PM1_r	TTTATCCATCATATAGGATTCCTCAAAGCGTAAAC
C100 PylS_NdeI-Restore_PM2_f	GAGGAATCCCGTATGATGGATAAAAAACCGCTGGATG
C101 PylS_NdeI-Restore_PM2_r	TTTATCCATCATACGGGATTCCTCAAAGCGTAAAC
C102 PylS_NdeI-Restore_PM3_f	GAGGAATCCCATACGATGGATAAAAAACCGCTGGATG
C103 PylS_NdeI-Restore_PM3_r	TTTATCCATCGTATGGGATTCCTCAAAGCGTAAAC
C104 PylS_NdeI-Restore_PM4_f	GAGGAATCCCATATAATGGATAAAAAACCGCTGGATG

C105 PylS_NdeI-Restore_PM4_r	TTTATCCATTATATGGGATTCCTCAAAGCGTAAAC
C106 PylS_NdeI-Restore_PM5_f	GAGGAATCCTATATAATGGATAAAAAACCGCTGGATG
C107 PylS_NdeI-Restore_PM5_r	TTTATCCATTATATAGGATTCCTCAAAGCGTAAAC
C108 T7seqf	TAATACGACTCACTATAGG
C109 pCDF43_PylS_seqfII	CCAATGGATTTAAAAAATG
C110 pCDF_4comp_seq1f	ATAGATTGGCTTTAAAAAC
C111 pCDF_4comp_seq2f	AGATTTTAGAGCCAATTAG
C112 pCDF_4comp_seq3f	AGATATCAACAACTTTCTG
C113 pCDF_4comp_seq4f	CCTGTGCCTGCGTCCGATG
C114 pCDF_4comp_seq5f	CGGCATTAGCACGAACCTG
C115 ORBS_rep_dGST_uRS_f	ACTAGCCATGGATCCAGGCGGCCGCTAAGCTTAATTAGCTG
C116 ORBS_rep_dGST_uRS_r	CTTAGCGGCCGCCTGGATCCATGGCTAGTATAGGGGACATTTG
C117 ORBS_rep_seq_f	CACTCATTAGGCACCCCAG
C118 ORBS_rep_seq_r	TCACCGTCATCACCGAAAC
C119 ORBS_rep_sfGFP_NcoI_f	ATCACCATGGTTAGCAAAGGTGAAGAACTGTTTAC
C120 ORBS_rep_sfGFP_HindIII_r	TGATAAGCTTAATGGTGATGATGATGGTGGCTG
C121 pCAT_T6TAG_f	GAAAAAAATCTAGGGATATACCACCGTTGATATATC
C122 pCAT_T6TAG_r	TGGTATATCCCTAGATTTTTTTCTCCATTTGCGGAG
C123 pCAT_D111AGGA_f	TGAATACCACAGGAGATTTCCGGCAGTTTCTACAC
C124 pCAT_D111AGGA_r	GCCGGAAATCTCCTGTGGTATTCACTCCAGAGCGATG
C125 pCATseq1750rv	CAGCCTTTTTTCTCCTGCCAC
C126 pCATseq1100rv	ATCTGCAGAATTCGCCCTTC
C127 pCATseq350fw	CCAAAACAGCCAAGCTGGAC
C128 CNPheRS_NdeI_f	ATCACATATGGACGAATTTGAAATGATAAAGAGAAAC
C129 CNPheRS_PstI_r	TGATCTGCAGTTATAATCTCTTTCTAATTGGCTCTAAAATC
C130 MjYRS23012NNK_for	GTAGGTCTCGGCTAAGATAAAGAAAGCANNKNNKNNKGCTGGAGTTGTTGAAGGAAATCCAATAATGGAGATAG
C131 MjYRS2836NNK_for	GTAGGTCTCATGAGGAGTTAGAGAGTTTATTTAAAAATAAGGAATTGNNKCCAATGNNKTTAAAAAATGCTGTAGCTGAAGAACTTATAAAG
C132 MjYRS261MNN_rev	GTAGGTCTCCCTCATAGCTATTAACTGTCAAATCTCCACCMNNTTTTTCTGGCCTTTTTATGGTTAAAGGATATTC
C133 H3opt_T6AGGA_f	TACCAAACAGAGGAGCGCGTAAAAGCACCGGCGGCAAAG
C134 H3opt_T6AGGA_r	TTTTACGCGCTCCTCTGTTTGGTACGGCCCTGAAAATAC
C135 ORBS_rep_H3_NcoI_f	ATCACCATGGGCAGCAGCCATCACCATCATCACCACAG
C136 ORBS_rep_H3optT6AGGA_HindIII_r	TGATAAGCTTACGCACGTTCGCCACGAATACGAC

C137 pACYC_AgeI_f	ATCAACCGGTAAACCAGCAATAGACATAAGCGGCTATTTAAC
C138 pACYC_NheI_r	TGATGCTAGCGCAGAATAAATAAATCCTGGTGTCCCTG
C139 pET_AgeI_f	ATCAACCGGTGGATCTTCACCTAGATCCTTTTAAATTAAAAATG
C140 pET_NheI_upstream_r	TGATGCTAGCTGTTTATTTTTCTAAATACATTCAAATATG
C141 pCDF_resi_seqf	CTGAAACCTCAGGCATTTG
C142 pCDF_resi_seqr	ACTTTGTATGTGTCCGCAG
C143 ORBS_pBAD_GFP_f	TGGGTTCATATCCCTCCGCAAATGGTTAGCAAAGGTG
C144 ORBS_pBAD_GFP_r	CATTTGCGGAGGGATATGAACCCAAAAAAACGGGTATG
C145 pBADseqf	ATTAGCGGATCCTACCTGAC
C146 ORBS_pBAD_GCN4_f	AATGAACTACCACCTGGAAAACGAAGTTGCGCGTCTGAAAAAACTGGTTGGTTAAAGCTCGAGATCTG
C147 ORBS_pBAD_GCN4_r	TTTAACCAACCAGTTTTTTCAGACGCGCAACTTCGTTTTCCAGGTGGTAGTTCATTTGCGGAGGGATATG

8.3 Construction of Plasmids

Table 8.3: Details of the construction of plasmids made during this thesis.
Plasmids that are listed in Table 8.1 and made during this thesis are aligned with the method used, primers (listed in Table 8.2), the template and the purpose of the mutagenesis. The corresponding sequencing primers (Seq. Primer) can also be found in the list. Methods: QC = Quik Change (Ch. 3.2.3.10); Inv = Inverse PCR (Ch. 3.2.3.11); Sub = Subcloning (Ch. 3.2.3.1 to 3.2.3.6).

Product	Method	Primer	Template	Purpose	Seq. Primer
pCLA7	QC	C1 + C2	pCLA1	N-terminal His$_6$-tag on PylS	C3
pCLA8	QC	C4 + C5	pCLA1	C-terminal His$_6$-tag on PylS	C3/C6
pCLA9	QC	C7 + C8	pCLA1	Internal His$_6$-tag in the “gap” between S144 and S145 on PylS	C3/C6
pCLA10	QC	C9 + C10	pCLA1	Internal His$_6$-tag in the loop D171 to F181 on PylS	C3/C6
pCLA11	QC	C11 + C12	pCLA9	deletion of 10 AAs of the linker region (in front of internal His$_6$-tag)	C3/C6
pCLA12	QC	C13 + C14	pCLA9	deletion of 20 AAs of the linker region (in front of internal His$_6$-tag)	C3/C6
pCLA13	QC	C15 + C16	pCLA9	deletion of 15 AAs of the linker region (behind internal His$_6$-tag)	C3/C6
pCLA14	QC	C17 + C18	pCLA9	deletion of 30 AAs of the linker region (behind internal His$_6$-tag)	C3/C6
pCLA15	QC	C19 + C20	pCLA9	deletion of 50 AAs of the linker region (in front of and behind internal His$_6$-tag)	C3/C6
pCLA16	QC	C21 + C22	pCLA9	deletion of C-terminal part beginning behind internal His$_6$-tag	C3/C6
pCLA17	QC	C23 + C24	pCLA9	deletion of N-terminal part up to internal His-Tag	C3/C6
pCLA18	QC	C34 + C35	pCLA2	N-terminal His$_6$-tag on MjYRS_TAG	C3
pCLA19	QC	C36 + C37	pCLA2	C-terminal His$_6$-tag on MjYRS_TAG	C3/C6
pCLA20	QC	C38 + C39	pCLA2	Internal His$_6$-tag in the loop K25 to K29 on MjYRS_TAG	C3/C6
pCLA22	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA23	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA24	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA25	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA26	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA27	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA28	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA29	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA30	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA31	Inv	C48 + C49	pCLA86	PylS promoter library	C6/C50
pCLA33	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53
pCLA34	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53
pCLA35	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53

pCLA36	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53
pCLA37	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53
pCLA38	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53
pCLA39	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53
pCLA40	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53
pCLA41	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53
pCLA42	Inv	C48 + C51	pCLA72	MjYRS_AGGA promoter library	C52/C53
pCLA44	Inv	C54 + C55	pCLA79	PylT promoter library	C57
pCLA45	Inv	C54 + C55	pCLA79	PylT promoter library	C57
pCLA46	Inv	C54 + C55	pCLA79	PylT promoter library	C57
pCLA47	Inv	C54 + C55	pCLA79	PylT promoter library	C57
pCLA48	Inv	C54 + C55	pCLA79	PylT promoter library	C57
pCLA49	Inv	C54 + C55	pCLA79	PylT promoter library	C57
pCLA50	Inv	C54 + C55	pCLA79	PylT promoter library	C57
pCLA51	Inv	C54 + C55	pCLA79	PylT promoter library	C57
pCLA52	Inv	C54 + C55	pCLA79	PylT promoter library	C57
pCLA53	Inv	C56 + C55	pCLA82	MjYT_UCCU promoter library	C57
pCLA54	Inv	C56 + C55	pCLA82	MjYT_UCCU promoter library	C57
pCLA55	Inv	C56 + C55	pCLA82	MjYT_UCCU promoter library	C57
pCLA56	Inv	C56 + C55	pCLA82	MjYT_UCCU promoter library	C57
pCLA57	Inv	C56 + C55	pCLA82	MjYT_UCCU promoter library	C57
pCLA58	Inv	C56 + C55	pCLA82	MjYT_UCCU promoter library	C57
pCLA59	Inv	C56 + C55	pCLA82	MjYT_UCCU promoter library	C57
pCLA60	Inv	C56 + C55	pCLA82	MjYT_UCCU promoter library	C57
pCLA61	Inv	C56 + C55	pCLA82	MjYT_UCCU promoter library	C57
pCLA62	Inv	C58 + C59	pCLA85	*glnS* promoter in front of *pylS* exchanged for P_{con}	C50/C53
pCLA63	Inv	C60 + C61	pCLA85	*glnS* promoter in front of *pylS* exchanged for P_{cp25}	C50/C53
pCLA64	Inv	C62 + C63	pCLA85	*glnS* promoter in front of *pylS* exchanged for $P_{lac1\text{-}6\ Mt5}$	C50/C53
pCLA65	Inv	C64 + C65	pCLA85	*glnS* promoter in front of *pylS* exchanged for P_{lac}	C50/C53
pCLA66	Inv	C66 + C68	pCLA85	*glnS* promoter in front of *pylS* exchanged for $P_{\lambda con/N25DSR}$	C50/C53
pCLA67	Inv	C67 + C68	pCLA85	*glnS* promoter in front of *pylS* exchanged for P_{λ}	C50/C53
pCLA68	Inv	C69 + C70	pCLA85	*glnS* promoter in front of *pylS* exchanged for $P_{D/E20}$	C50/C53
pCLA69	Inv	C71 + C72	pCLA85	*glnS* promoter in front of *pylS* exchanged for P_{H207}	C50/C53
pCLA70	Inv	C73 + C74	pCLA85	*glnS* promoter in front of *pylS* exchanged for P_{tacI}	C50/C53
pCLA71	QC	C36 + C37	pCLA3	C-terminal His$_6$-tag on MjYRS_AGGA	C3/C6
pCLA72	QC	C75 + C76	pCLA71	SalI restriction site behind MjYRS terminator	C3/C53
pCLA73	QC	C77 + C78	pCLA9	NdeI restriction site in front of PylS start codon changed to KpnI	C3/C6
pCLA74	QC	C79 + C80	pCLA73	PstI restriction site behind PylS stop codon changed to SacI	C50/C53

pCLA75	QC	C81 + C82	pCLA74	SalI restriction site in front of PylS promoter	C6/C50
pCLA76	QC	C83 + C84	pCLA75	NotI restriction site behind PylS terminator	C50/C53
pCLA77	QC	C85 + C86	pCLA4	deletion of genes for T7 RNA Polymerase, AraC and GFP	digest only
pCLA78	QC	C87 + C88	pCLA77	NotI restriction site in front of PylT promoter	C89
pCLA79	QC	C90 + C91	pCLA78	MfeI restriction site behind PylT terminator	C89
pCLA80	QC	C85 + C86	pCLA6	deletion of genes for T7 RNA Polymerase, AraC and GFP	digest only
pCLA81	QC	C92 + C93	pCLA80	MfeI restriction site in front of MjYT_UCCU promoter	C57
pCLA82	QC	C94 + C95	pCLA81	XhoI restriction site behind MjYT_UCCU terminator	C57/C89
pCLA83	QC	C81 + C82	pCLA9	SalI restriction site in front of PylS promoter	C6/C50
pCLA84	QC	C83 + C84	pCLA9	NotI restriction site behind PylS terminator	C50/C53
pCLA85	QC	C96 + C97	pCLA76	NdeI site on pCLA76 restored	C50/C53
pCLA86	QC	C98 + C99	pCLA76	point mutation in restored NdeI site	C50/C53
pCLA87	QC	C100 + C101	pCLA76	point mutation in restored NdeI site	C50/C53
pCLA88	QC	C102 + C103	pCLA76	point mutation in restored NdeI site	C50/C53
pCLA89	QC	C104 + C105	pCLA76	point mutation in restored NdeI site	C50/C53
pCLA90	QC	C106 + C107	pCLA76	point mutation in restored NdeI site	C50/C53
pCLA92	Sub	-	pCLA72	MjYRS_AGGA gene cloned with BamHI + SalI into pCLA91	C79/C89/ C108-C114
pCLA93	Sub	-	pCLA86	PylS gene cloned with SalI + NotI (+ XhoI) into pCLA92	C79/C89/ C108-C114
pCLA94	Sub	-	pCLA79	PylT gene cloned with NotI + MfeI into pCLA93	C79/C89/ C108-C114
pCLA95	Sub	-	pCLA82	MjYT_UCCU gene cloned with MfeI + XhoI into pCLA94	C79/C89/ C108-C114
pCLA100	QC	C115 + C116	pCLA99	deletion of the GST-MBP gene and incorporation of a MCS	C117/C118
pCLA105	Sub	C119 + C120	pCLA101	sfGFP WT gene cloned with NcoI + HindIII into pCLA100	C117/C118
pCLA106	Sub	C119 + C120	pCLA102	sfGFP 150TAG gene cloned with NcoI + HindIII into pCLA100	C117/C118
pCLA107	Sub	C119 + C120	pCLA103	sfGFP 134AGGA 150TAG gene cloned with NcoI + HindIII into pCLA100	C117/C118
pCLA108	Sub	C119 + C120	pCLA104	sfGFP 134AGGA gene cloned with NcoI + HindIII into pCLA100	C117/C118
pCLA112	QC	C121 + C122	pCLA111	Codon for T6 in the CAT gene exchanged for TAG	C125-C127

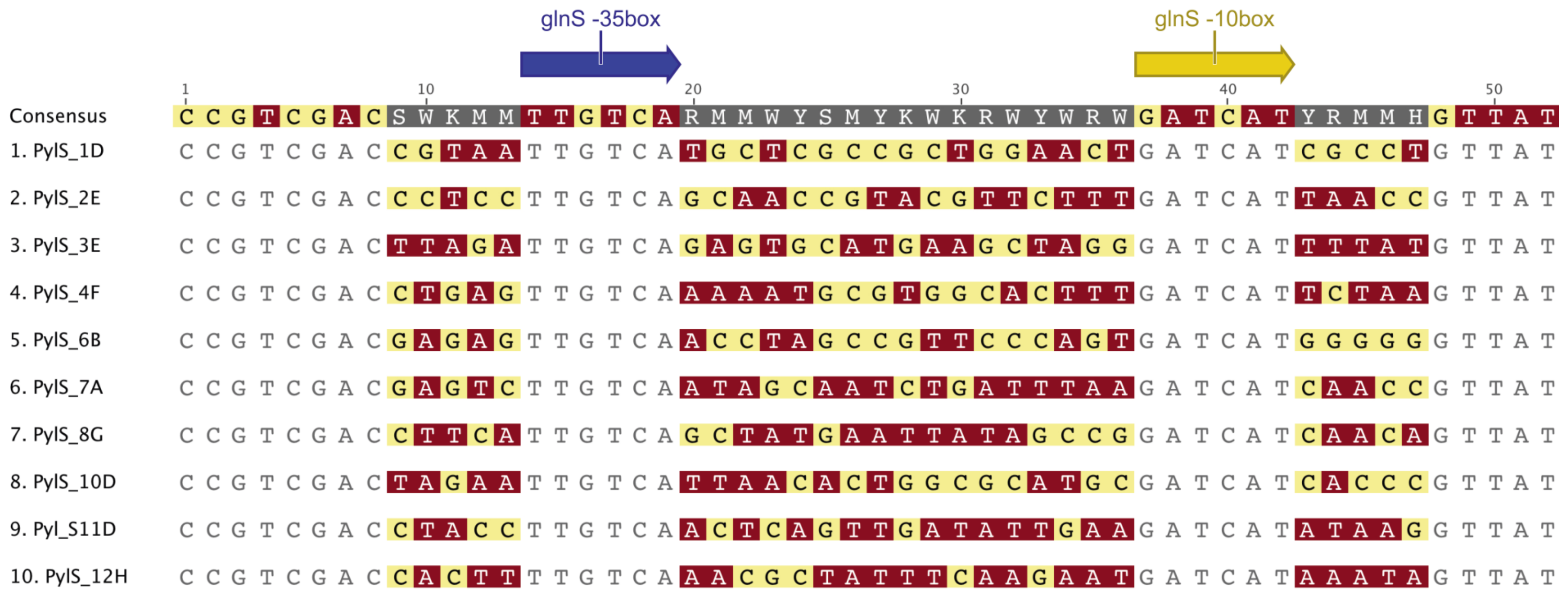
pCLA113	QC	C123 + C124	pCLA112	Codon for D111 in the CAT gene exchanged for AGGA	C125-C127
pCLA115	Sub	-	pCLA24	*pylS* with P_{glnS} exchanged for *pylS* with P_{lib3E} by cloning into pCLA95 with SalI + NotI (+ XhoI)	C111
pCLA116	Sub	-	pCLA62	*pylS* with P_{glnS} exchanged for *pylS* with P_{con} by cloning into pCLA95 with SalI + NotI (+ XhoI)	C111
pCLA117	Sub	-	pCLA65	*pylS* with P_{glnS} exchanged for *pylS* with P_{lac} by cloning into pCLA95 with SalI + NotI (+ XhoI)	C111
pCLA120	Sub	C128 + C129	pCLA119	MjAzFRS gene cloned with NdeI + PstI into pCLA1	C3/C6
pCLA121	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA122	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA123	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA124	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA125	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA126	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA127	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA128	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA129	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA130	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA131	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA132	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA133	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA134	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA135	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA136	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA137	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA138	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA139	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6

pCLA140	Inv	C130 and C131 + C132	pCLA120	MjAzFRS library in two rounds of inverse PCR	C6
pCLA142	QC	C133 + C134	pCLA141	Codon for T6 in the H3opt gene exchanged for AGGA	C108
pCLA143	Sub	C135 + C136	pCLA142	H3opt T6AGGA gene cloned with NcoI + HindIII into pCLA100	C117
pCLA144	Sub	-	pCLA122	MjYRS gene on pCLA98 replaced by MjAzFRS gene (clone #2) with NdeI + StuI	digest only
pCLA145	Sub	-	pCLA123	MjYRS gene on pCLA98 replaced by MjAzFRS gene (clone #3) with NdeI + StuI	digest only
pCLA146	Sub	-	pCLA125	MjYRS gene on pCLA98 replaced by MjAzFRS gene (clone #5) with NdeI + StuI	digest only
pCLA147	Sub	-	pCLA126	MjYRS gene on pCLA98 replaced by MjAzFRS gene (clone #6) with NdeI + StuI	digest only
pCLA148	Sub	-	pCLA131	MjYRS gene on pCLA98 replaced by MjAzFRS gene (clone #11) with NdeI + StuI	digest only
pCLA149	Sub	-	pCLA137	MjYRS gene on pCLA98 replaced by MjAzFRS gene (clone #17) with NdeI + StuI	digest only
pCLA150	Sub	-	pCLA138	MjYRS gene on pCLA98 replaced by MjAzFRS gene (clone #18) with NdeI + StuI	digest only
pCLA152	Sub	-	pCLA151	MjAzFRS (TAG) gene cloned into pCLA91 with BamHI + SalI	C3/C6
pCLA153	Sub	-	pCLA95	tRNA cassette (PylT and MjYT_UCCU) cloned into pCLA152 with NotI + XhoI	C89/C109
pCLA154	Sub	-	pCLA137	MjAzFRS (TAG) gene on pCLA153 replaced by MjAzFRS (AGGA, clone #17) with NdeI + StuI	C3/C6
pCLA155	Sub	-	pCLA95	PylS gene cloned into pCLA154 with SalI + NotI	C111-C113
pCLA165	Sub	C81 + C84	pCLA158	BCNRS gene cloned into pCLA154 with SalI + NotI	C111-C113
pCLA166	Sub	C81 + C84	pCLA160	AzpcKRS gene cloned into pCLA154 with SalI + NotI	C111-C113
pCLA167	Sub	-	pCLA158	PylS gene on pCLA62 replaced by BCNRS gene with NdeI + StuI	C6/C50
pCLA168	Sub	-	pCLA158	PylS gene on pCLA65 replaced by BCNRS gene with NdeI + StuI	C6/C50
pCLA169	Sub	-	pCLA160	PylS gene on pCLA62 replaced by AzpcKRS gene with NdeI + StuI	C6/C50
pCLA170	Sub	-	pCLA160	PylS gene on pCLA65 replaced by AzpcKRS gene with NdeI + StuI	C6/C50
pCLA171	Sub	-	pCLA167	BCNRS gene with P_{con} cloned into pCLA154 with SalI + NotI (+ XhoI)	C111-C114

pCLA172	Sub	-	pCLA168	BCNRS gene with P_{lac} cloned into pCLA154 with SalI + NotI (+ XhoI)	C111-C114
pCLA173	Sub	-	pCLA169	AzpcKRS gene with P_{con} cloned into pCLA154 with SalI + NotI (+ XhoI)	C111-C114
pCLA174	Sub	-	pCLA170	AzpcKRS gene with P_{lac} cloned into pCLA154 with SalI + NotI (+ XhoI)	C111-C114
pCLA184	Sub	C137 + C138	pCLA182	SmR gene on pCLA116 replaced by CatR gene with AgeI + NheI	C111/ C141/C142
pCLA185	Sub	C139 + C140	pCLA183	SmR gene on pCLA116 replaced by AmpR gene with AgeI + NheI	C111/ C141/C142
pCLA186	Sub	C137 + C138	pCLA182	SmR gene on pCLA171 replaced by CatR gene with AgeI + NheI	C111/ C141/C142
pCLA187	Sub	C139 + C140	pCLA183	SmR gene on pCLA171 replaced by AmpR gene with AgeI + NheI	C111/ C141/C142
pCLA193	QC	C143 + C144	pCLA101	Embedding ORBS while “killing” the WT Shine-Dalgarno sequence	C145
pCLA194	QC	C146 + C147	pCLA193	sfGFP gene exchanged by sequence for GCN4 16-31	C145

8.4 Sequence Alignments

8.4.1 PylS Promoter Library



Consensus	CCGTCGACSWKMMTTGTCARMMWYSMYKWKRWYWRWGATCATYRMMHGTTAT
1. PylS_1D	CCGTCGACCGTAATTGTCATGCTCGCCGCTGGAACTGATCATCGCCTGTTAT
2. PylS_2E	CCGTCGACCCTCCTTGTCAGCAACCGTACGTTCTTTGATCATTAACCGTTAT
3. PylS_3E	CCGTCGACTTAGATTGTCAGAGTGCATGAAGCTAGGGATCATTTTATGTTAT
4. PylS_4F	CCGTCGACCTGAGTTGTCAAAAATGCGTGGCACTTTGATCATTCTAAGTTAT
5. PylS_6B	CCGTCGACGAGAGTTGTCAACCTAGCCGTTCCCAGTGATCATGGGGGGTTAT
6. PylS_7A	CCGTCGACGAGTCTTGTCAATAGCAATCTGATTTAAGATCATCAACCGTTAT
7. PylS_8G	CCGTCGACCTTCATTGTCAGCTATGAATTATAGCCGGATCATCAACAGTTAT
8. PylS_10D	CCGTCGACTAGAATTGTCATTAACACTGGCGCATGCGATCATCACCCGTTAT
9. Pyl_S11D	CCGTCGACCTACCTTGTCAACTCAGTTGATATTGAAGATCATATAAGGTTAT
10. PylS_12H	CCGTCGACCACTTTTGTCAAACGCTATTTCAAGAATGATCATAAATAGTTAT

Figure 8.1: Sequence alignment of selected PylS promoter library clones.
The alignment shows the mutated promoter region of the ten selected library clones. Mutated positions are highlighted in red (bases A and T) or yellow (bases C and G). Mutagenesis only occurred at intended positions. All ten clones differed in sequence.

8.4.2 MjYRS_AGGA Promoter Library

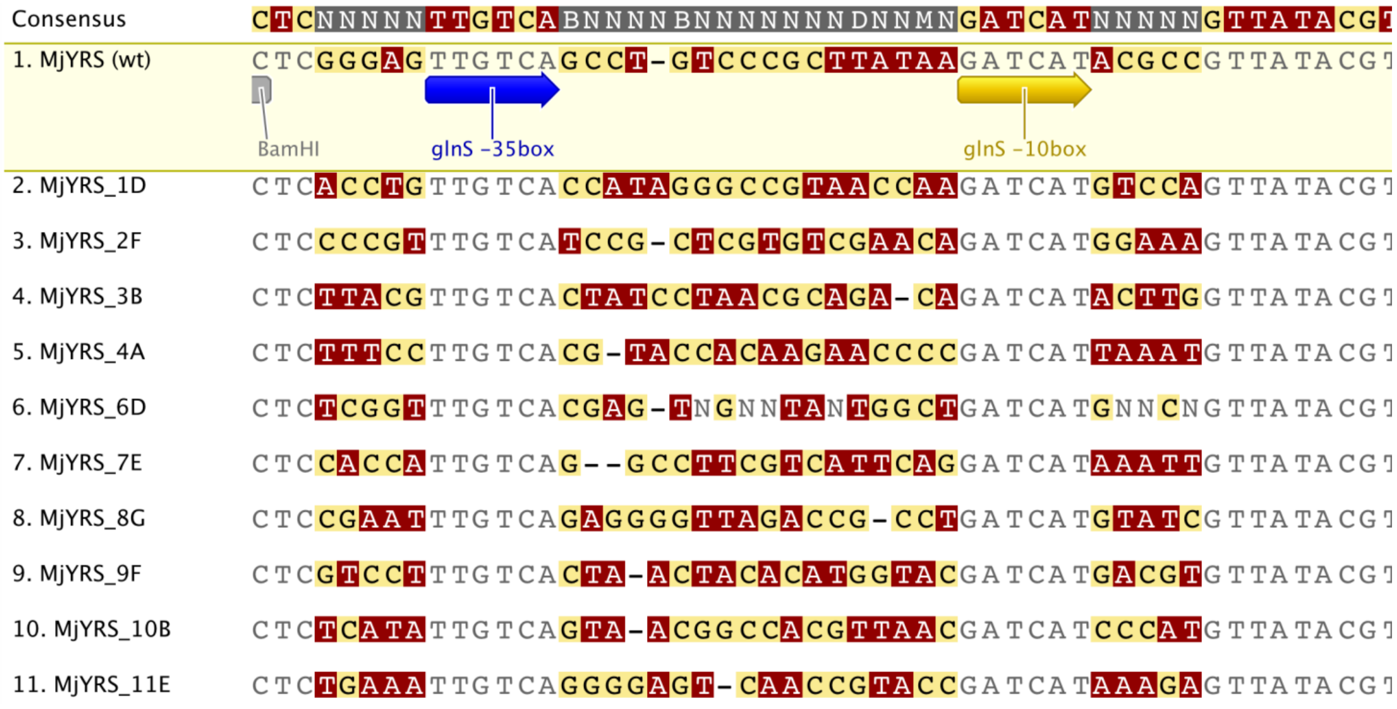

Consensus CTCNNNNNTTGTCABNNNNBNNNNNNNDNNMNGATCATNNNNNGTTATACG
1. MjYRS (wt) CTCGGGAGTTGTCAGCCT-GTCCCGCTTATAAGATCATACGCCGTTATACG
BamHI
glnS -35box
glnS -10box
2. MjYRS_1D CTCACCTGTTGTCACCATAGGGCCGTAACCAAGATCATGTCCAGTTATACG
3. MjYRS_2F CTCCCCGTTTGTCATCCG-CTCGTGTCGAACAGATCATGGAAAGTTATACG
4. MjYRS_3B CTCTTACGTTGTCACTATCCTAACGCAGA-CAGATCATACTTGGTTATACG
5. MjYRS_4A CTCTTTCCTTGTCACG-TACCACAAGAACCCCGATCATTAAATGTTATACG
6. MjYRS_6D CTCTCGGTTTGTCACGAG-TNGNNTANTGGCTGATCATGNNCNGTTATACG
7. MjYRS_7E CTCCACCATTGTCAG--GCCTTCGTCATTCAGGATCATAAATTGTTATACG
8. MjYRS_8G CTCCGAATTTGTCAGAGGGGTTAGACCG-CCTGATCATGTATCGTTATACG
9. MjYRS_9F CTCGTCCTTTGTCACTA-ACTACACATGGTACGATCATGACGTGTTATACG
10. MjYRS_10B CTCTCATATTGTCAGTA-ACGGCCACGTTAACGATCATCCCATGTTATACG
11. MjYRS_11E CTCTGAAATTGTCAGGGGAGT-CAACCGTACCGATCATAAAGAGTTATACG

Figure 8.2: Sequence alignment of selected MjYRS_AGGA promoter library clones.
The alignment shows the mutated promoter region of the ten selected library clones compared to the WT *glnS* promoter sequence. Mutated positions are highlighted in red (bases A and T) or yellow (bases C and G). Mutagenesis only occurred at intended positions. All eleven clones differed in sequence.

8.4.3 PylT Promoter Library

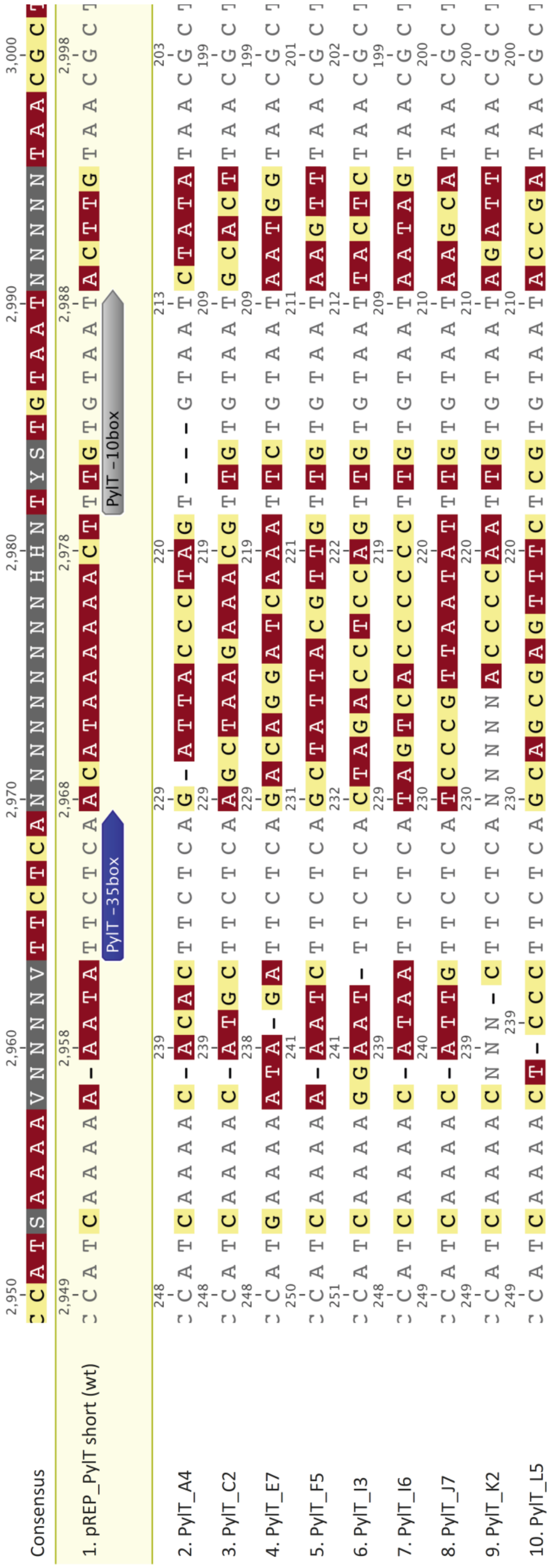

Consensus
1. pREP_PylT short (wt)
2. PylT_A4
3. PylT_C2
4. PylT_E7
5. PylT_F5
6. PylT_I3
7. PylT_I6
8. PylT_J7
9. PylT_K2
10. PylT_L5
PylT -35box
PylT -10box
2,950
2,960
2,970
2,980
2,990
3,000

Figure 8.3: Sequence alignment of selected PylT promoter library clones.
The alignment shows the mutated promoter region of the nine selected library clones compared to the WT *lpp* promoter sequence. Mutated positions are highlighted in red (bases A and T) or yellow (bases C and G). Mutagenesis mainly occurred at intended positions. All ten clones differed in sequence.

8.4.4 MjYT_UCCU Promoter Library

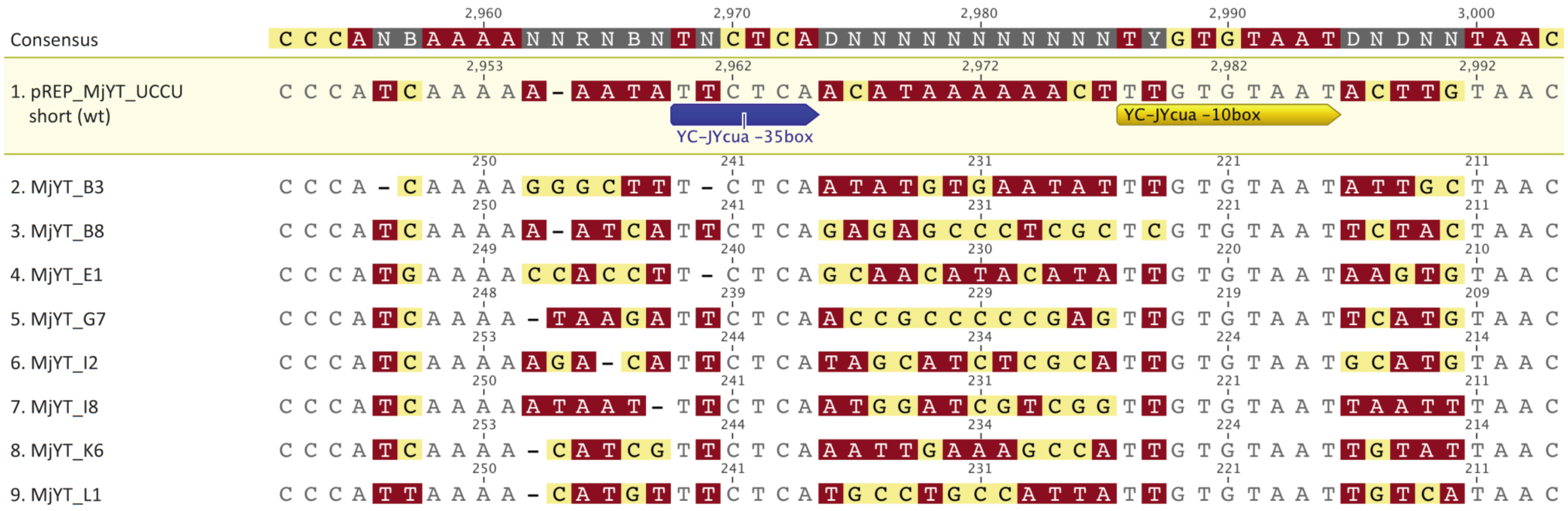

2,960
2,970
2,980
2,990
3,000
Consensus
CCCANBAAAANNRNBNTNCTCADNNNNNNNNNNNTYGTGTAATDNDNNTAAC
2,953
2,962
2,972
2,982
2,992
1. pREP_MjYT_UCCU short (wt)
CCCATCAAAAA-AATATTCTCAACATAAAAAACTTTGTGTAATACTTGTAAC
YC-JYcua -35box
YC-JYcua -10box
2. MjYT_B3
CCCA-CAAAAGGGCTTT-CTCAATATGTGAATATTTGTGTAATATTGCTAAC
3. MjYT_B8
CCCATCAAAAA-ATCATTCTCAGAGAGCCCTCGCTCGTGTAATTCTACTAAC
4. MjYT_E1
CCCATGAAAACCACCTT-CTCAGCAACATACATATTGTGTAATAAGTGTAAC
5. MjYT_G7
CCCATCAAAA-TAAGATTCTCAACCGCCCCCGAGTTGTGTAATTCATGTAAC
6. MjYT_I2
CCCATCAAAAAGA-CATTCTCATAGCATCTCGCATTGTGTAATGCATGTAAC
7. MjYT_I8
CCCATCAAAAATAAT-TTCTCAATGGATCGTCGGTTGTGTAATTAATTTAAC
8. MjYT_K6
CCCATCAAAA-CATCGTTCTCAAATTGAAAGCCATTGTGTAATTGTATTAAC
9. MjYT_L1
CCCATTAAAA-CATGTTTCTCATGCCTGCCATTATTGTGTAATTGTCATAAC

Figure 8.4: Sequence alignment of selected MjYT_UCCU promoter library clones.
The alignment shows the mutated promoter region of eight of the nine selected library clones compared to the WT *lpp* promoter sequence. Sequencing of clone D4 failed. Mutated positions are highlighted in red (bases A and T) or yellow (bases C and G). Mutagenesis mainly occurred at intended positions. All nine clones differed in sequence.

8.5 In-Cell sfGFP Fluorescence Measurements

8.5.1 Initial Conditions For In-Cell NMR Measurements

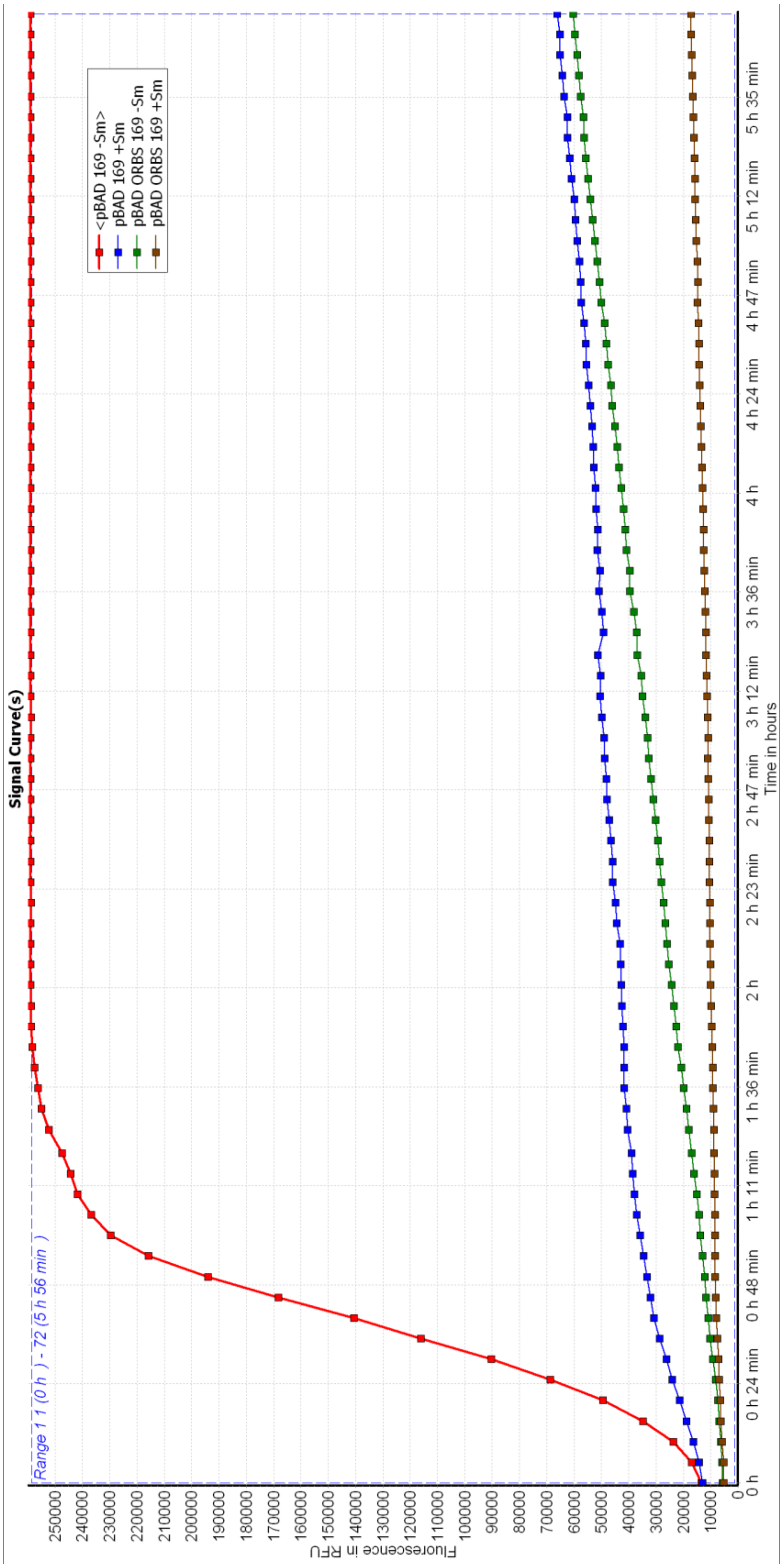


Figure 8.5: Effect of spectinomycin on in-cell fluorescence of sfGFP. Arabinose induced expression of sfGFP in *E. coli* BL21 using an orthogonal mRNA (pBAD ORBS 169; pCLA193; 0.2% arabinose) in combination with the orthogonal ribosome ribo-Q1 (pCLA96) and a non-orthogonal mRNA (pBAD 169; pCLA101; 0.02% arabinose). Expressions were performed in the absence (-Sm) and presence (+Sm) of 150 µg/mL spectinomycin. In-cell fluorescence was measured every 5 min over 6 h using a plate reader (Ch. 3.2.2.12).

8.5.2 Optimization For In-Cell NMR Measurements

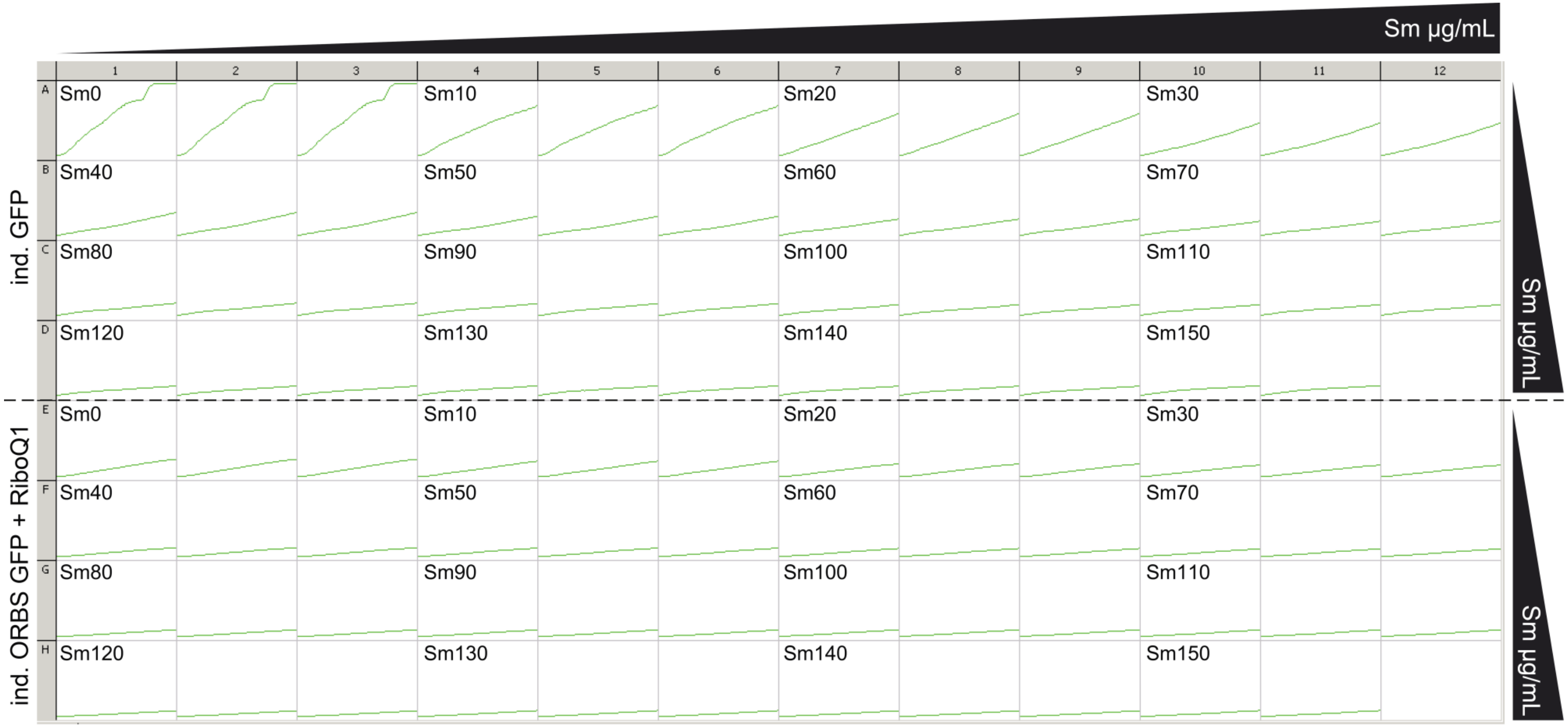


Figure 8.6: Titration of spectinomycin to optimize in-cell fluorescence of sfGFP.
Arabinose induced expression of sfGFP in *E. coli* BL21 using an orthogonal mRNA (ind. ORBS GFP; pCLA193; 0.2% arabinose) in combination with the orthogonal ribosome ribo-Q1 (pCLA96) and a non-orthogonal mRNA (ind. GFP; pCLA101; 0.01% arabinose). Expressions were performed with spectinomycin concentrations ranging from 10 to 150 µg/mL. In-cell fluorescence was measured in triplicates every 5 min over 6.5 h using a plate reader (Ch. 3.2.2.12).

8.6 Curriculum Vitae

Name: Christoph Lammers
Geburtsdatum, -ort: 02. März 1984 in Würselen
Familienstand: ledig
Staatsangehörigkeit: deutsch
Anschrift: Annastraße 66 / App. 65, 37075 Göttingen
Email: clammer1@gwdg.de

PROMOTION

Seit 2010	Promotion an der Georg-August-Universität, Göttingen Abteilung für Angewandte Synthetische Biologie

---

STUDIUM DER BIOLOGIE

09/2009	Abschluss des Studiums als **Diplom-Biologe** Thema der Diplomarbeit: „Erstellung eines Wikis zu den Genen von *Bacillus subtilis*” (Abteilung für Allgemeine Mikrobiologie bei Prof. Dr. Stülke)
2004 - 2009	Studium der Biologie an der Georg-August-Universität, Göttingen **Hauptfach:** Mikrobiologie **Nebenfächer:** Biochemie, organische Chemie

---

WEHRDIENST

2003 - 2004	Soldat im Grundwehrdienst (Sanitäter) 4./Sanitätsregiment 12, Fürstenau

---

SCHULISCHE AUSBILDUNG

1996 - 2003	Ubbo-Emmius-Gymnasium, Leer Abschluss: Allgemeine Hochschulreife
1994 - 1996	Orientierungsstufe, Collhusen
1990 - 1994	Grundschule, Ihrhove