

**Evolution wachsender, sich reproduzierender Polymere.
Zellfreie Genexpression zum Einbau nicht-kanonischer
Aminosäuren in Proteine sowie zur Analyse des epigenetischen
Escherichia-coli-Pili-Phasenvariationsmechanismus.**

Dissertation
zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät
der Universität des Saarlandes

von

Emanuel Gregor Worst

Saarbrücken
2017

Tag des Kolloquiums: 20.04.2018

Dekan: Univ.-Prof. Dr. Guido Kickelbick

Berichterstatter: Univ.-Prof. Dr. Albrecht Ott
Univ.-Prof. Dr. Dr. Karsten Kruse
Univ.-Prof. Dr. Joachim Rädler

Vorsitz: Univ.-Prof. Dr. Gregor Jung

Akad. Mitarbeiter: Dr. Andreas Tschöpe

Abstrakt

Seit Beginn des Lebens entstehen durch Reproduktion, Variation und Selektion immer neue Spezies. Stetige Erneuerung statt Stagnation ist Teil der Evolution. Ein Modellsystem aus linearen DNA-Molekülen wurde entwickelt, in dem durch spontane Ligation neue Polymere entstehen, die sich Template-basiert vermehren. Bei adäquat hohen Reproduktionsraten wurden unter den länger werdenden Polymeren gewisse Längen selektiert. Dieser stetige, selektive Innovationsprozess zeigt gewisse Parallelen zur biologischen Evolution.

Der Einbau nicht-kanonischer Aminosäuren kann natürliche Proteinfunktionen erweitern. Nicht-kanonische Aminosäuren-Analoga hemmen oft das Expressionswirts-Wachstum, limitieren so den *in-vivo*-Ansatz. Besonders der *in-vivo*-Einbau toxischer Analoga wie L-Canavanin ist schwierig. Jedoch konnte gezeigt werden, dass ein zellfreies *Escherichia coli*-Expressionssystem dieses medizinisch relevante Analogon vollständig einbaut, ein neuer Ansatz, modifizierte Proteine für Folgestudien herzustellen. Weitere nicht-kanonische Aminosäuren wurden auch, jedoch weniger effizient, eingebaut.

Die Regulation des epigenetischen *Escherichia coli*-Pili-Phasenvariationsmechanismus erfolgt transkriptionell durch Modulation des DNA-Methylierungsmusters, involviert Koregulatoren. Es wurde eine synthetische regulatorische DNA im obigen Expressionssystem studiert und theoretisch analysiert. Die DNA wechselwirkte methylierungsabhängig mit sich selbst. Folglich wurde der Mechanismus neu interpretiert.

Abstract

Since the origin of life the emergence of new species has been promoted by reproduction, variation, and selection. Instead of stalling, Evolution innovates ad infinitum. A model system of linear DNA molecules was devised where new template-based proliferating polymers appear by spontaneous ligation. While length increasing polymers were emerging, for sufficiently high reproduction rates particular lengths were constantly being selected. This steady, selective innovation process bears resemblance to biological evolution.

The incorporation of noncanonical amino acids can extend natural protein functionalities. The *in vivo* approach is limited as non-canonical amino acid analogues often act as growth inhibitors for expression hosts. *In vivo* incorporation of toxic analogues like L-canavanine is exceptionally difficult. However, it was shown that an *Escherichia coli* cell-free expression system completely incorporates this medically relevant analogue into a model protein, a new approach producing modified proteins for further studies. Other noncanonical amino acids were incorporated, albeit with lower efficiency.

The regulation of the epigenetic *Escherichia coli* pili phase variation mechanism transcriptionally proceeds by DNA methylation pattern modulation, involving coregulators. A synthetic regulatory DNA was studied in the above expression system and theoretically analyzed. Dependent on the methylation pattern, this DNA self-interacted. Accordingly, the mechanism was reinterpreted.

Publikationsliste

Teile der vorliegenden Dissertation wurden in folgenden Artikeln, die einem *Peer-Review* unterzogen wurden, vorab publiziert:

Emanuel G. Worst*, **Philipp Zimmer***, **Eva Wollrab**, **Karsten Kruse** & **Albrecht Ott**. Unbounded growth patterns of reproducing, competing polymers – similarities to biological evolution. *New Journal of Physics*, **18**(10): 103003 (2016)#

ausgewählt für die Kollektion *Highlights of 2016* des *New Journal of Physics* im Bereich Biophysik,

* gleichwertig beigetragen

Emanuel G. Worst*, **Matthias P. Exner***, **Alessandro De Simone**, **Marc Schenkelberger**, **Vincent Noireaux**, **Nediljko Budisa** & **Albrecht Ott**. Cell-free expression with the toxic amino acid canavanine. *Bioorganic & Medicinal Chemistry Letters*, **25**(17): 3658–3660 (2015)

* gleichwertig beigetragen

Emanuel G. Worst, **Matthias P. Exner**, **Alessandro De Simone**, **Marc Schenkelberger**, **Vincent Noireaux**, **Nediljko Budisa** & **Albrecht Ott**. Incorporation of Noncanonical Amino Acids into Model Proteins Using an *Escherichia coli* Cell-free Transcription-translation System. *Journal of Visualized Experiments*, (14): e54273 (2016)

Zusätzlich wurde an der folgenden Publikation, die einem *Peer-Review* unterzogen wurde, mitgearbeitet:

Marc Schenkelberger, **Siba Shanak**, **Marc Finkler**, **Emanuel G. Worst**, **Vincent Noireaux**, **Volkhard Helms** & **Albrecht Ott**. Expression regulation by a methyl-CpG binding domain in an *E. coli* based, cell-free TX-TL system. *Physical Biology*, **14**(2): 026002 (2017)

Inhaltsverzeichnis

1. Vorwort	1
2. Biochemische, molekularbiologische Grundlagen und Einführung in die zellfreie Proteinsynthese	3
2.1. Aufbau von DNA, RNA und Proteinen	3
2.1.1. Aufbau von DNA und RNA	3
2.1.2. Aufbau der Proteine	7
2.2. Fluss der genetischen Information und die Genexpression	10
2.2.1. Das zentrale Dogma der Molekularbiologie	10
2.2.2. Prokaryotische Organisation von DNA, DNA-Replikation und Genexpression .	10
2.3. Mutationen, Regulation der Genexpression und Epigenetik	16
2.3.1. Mutationen	16
2.3.2. Regulation der Genexpression und Epigenetik	17
2.4. Zellfreie Proteinsynthese	19
2.4.1. Grundstein der modernen zellfreien Proteinsynthese	19
2.4.2. Anwendungsbereiche verschiedener ZFPS-Systeme	20
2.4.3. <i>E.-coli</i> -basierte ZFPS	20
2.5. Fluoreszierende Proteine der GFP-Familie	25
3. Evolution wachsender, sich reproduzierender Polymere	27
3.1. Einführung, Stand der Wissenschaft, Strategie und Zielsetzung	27
3.2. Material und Methoden	29
3.2.1. Herstellung der Proben und deren initiale Zusammensetzung	29
3.2.2. Zyklische Variation der Temperatur, Probenentnahme und -manipulation . . .	30
3.2.3. Gelelektrophoretische Trennung und Quantifizierung der DNA-Menge	31
3.3. Ergebnisse	34
3.3.1. Ergebnisse eigener Vorarbeiten	34
3.3.2. Restriktionsverdau der DNA-Moleküle mit heterogenen Sequenzen	36
3.3.3. AfeI, BstUI und Mungbohnen-Nuklease zur Unterdrückung der Überhang- Ligation	38
3.3.4. Theoretische Analyse	39
3.3.5. Experimentelle Umsetzung der Längenverdopplung	41
3.4. Zusammenfassung, Konklusion und Ausblick	43
4. Zellfreier Einbau nicht-kanonischer Aminosäuren in Modellproteine	45
4.1. Einführung, Stand der Wissenschaft, Strategie und Zielsetzung	45
4.1.1. Grenzen der chemischen Peptidsynthese	45
4.1.2. Einbau nicht-kanonischer Aminosäuren <i>in vivo</i> und <i>in vitro</i>	46
4.1.3. Zielsetzung und Strategie	52
4.2. Material und Methoden	56
4.2.1. Verwendete Chemikalien	56

4.2.2.	Nährmedien und Antibiotika-Selektion	56
4.2.3.	<i>E.-coli</i> -Stämme	57
4.2.4.	Agarosegelelektrophorese	57
4.2.5.	<i>Polymerase chain reaction</i> (PCR)	58
4.2.6.	Restriktionsverdau von DNA-Strängen	59
4.2.7.	Ligation von DNA-Strängen mit klebrigen Enden	60
4.2.8.	Gibson-Assemblierung von DNA-Strängen	61
4.2.9.	Transformation kompetenter <i>E.-coli</i> -Zellen, Plasmid-Amplifikation und Plasmidpräparation	61
4.2.10.	DNA-Sequenzierung	64
4.2.11.	Herstellung der verwendeten DNA-Konstrukte	64
4.2.12.	Konzentrationsbestimmung von DNA	68
4.2.13.	Konzentrationsbestimmung von Proteinen	69
4.2.14.	Herstellung des <i>E.-coli</i> -Zellextraktes, Präparation der zellfreien Reaktionen und Reaktionsbedingungen	70
4.2.15.	SDS-PAGE	72
4.2.16.	His-Tag-Purifikation zellfrei synthetisierter Modellproteine mittels IMAC	73
4.2.17.	Pufferaustausch und Aufkonzentrierung	74
4.2.18.	HPLC-ESI massenspektrometrische und fluoreszenzspektroskopische Analyse der Modellproteine	75
4.3.	Ergebnisse und Diskussion	76
4.3.1.	Optimierung der zellfreien Reaktion	76
4.3.2.	Seitenkettenspezifischer Einbau nicht-kanonischer Aminosäuren in deGFP	76
4.3.3.	Ortsspezifischer Einbau von BocLys in EGFP ^{N149amber}	84
4.4.	Zusammenfassung, Konklusion und Ausblick	91
5.	Zellfreie Analyse des epigenetischen <i>Escherichia coli</i>-Pili-Phasenvariationsmechanismus	95
5.1.	Einführung, Stand der Wissenschaft, Strategie und Zielsetzung	95
5.1.1.	Urinaltraktinfektionen durch <i>E. coli</i> und ihre Folgen	95
5.1.2.	Pyelonephritis-assoziierte Pili von uropathogenen <i>E. coli</i>	95
5.1.3.	Epigenetische Regulation des UPEC-Pili-Phasenvariationsmechanismus	96
5.1.4.	Vorarbeiten in zellfreier <i>E.-coli</i> -ähnlicher Umgebung	99
5.1.5.	Strategie und Zielsetzung der Studie	100
5.2.	Material und Methoden	103
5.2.1.	Herstellung der verwendeten DNA-Konstrukte	103
5.2.2.	Dam-Methylierung der <i>pap</i> -DNA-Konstrukte	107
5.2.3.	Präparation der zellfreien Reaktionen, Reaktionsbedingungen und Fluoreszenz-Intensitäts-Messungen	108
5.2.4.	Vorhersage der <i>pap</i> -DNA-Sekundärstruktur	110
5.3.	Ergebnisse und Diskussion	111
5.3.1.	Epigenetische Regulation des <i>pap</i> -regulatorischen Bereichs ohne zellfreie Lrp- und PapI-Synthese	111
5.3.2.	Epigenetische Regulation des <i>pap</i> -regulatorischen Bereichs bei zusätzlicher Lrp- und PapI-Synthese	114
5.3.3.	Untersuchung der DNA-Wechselwirkung des <i>pap</i> -regulatorischen Bereichs	120
5.4.	Zusammenfassung, Konklusion und Ausblick	123

A. Anhang	127
A.1. Ergänzendes Material zu Kap. 3	127
A.1.1. DNA-Sequenzen der verwendeten DNA-Moleküle	127
A.1.2. Reduzierung der Mungbohnen-Enzym-Aktivität	128
A.1.3. Bestimmung der spontanen Ligations- und autokatalytischen Vermehrungsrate	129
A.2. Ergänzendes Material zu Kap. 4	133
A.2.1. Ergänzungen zur Stopcodon-Suppression mittels der aminoacylierten Suppressor-tRNA	133
A.2.2. DNA-Sequenzen	134
A.3. Ergänzendes Material zu Kap. 5	139
A.3.1. DNA-Sequenzen	139
A.3.2. Kinetik-Messungen zur Analyse der epigenetischen Regulation des <i>pap</i> - regulatorischen Bereichs ohne zellfreie Lrp- und PapI-Synthese	142
A.3.3. Weitere Sekundärstrukturen der synthetischen <i>pap</i> -DNA	143
Literatur	145
Dank und Anerkennung	171

1. Vorwort

Die Darwinsche Evolution [1] verändert seit jeher das Leben auf der Erde: allmählich trieben Mutation und Selektion das irdische Ökosystem, das anfangs eine präbiotische Ursuppe [2]–[4] darstellte, hin zu immer komplexer werdenden, sich reproduzierenden Individuen. Der anfängliche Zustand glich wohl einem promiskuitiven Zustand aus ununterscheidbaren Individuen, da diese unermüdlich Moleküle miteinander austauschten [5]–[10], jedoch führten komplexe dynamische Wechselwirkungen allmählich zur Herausbildung von differenzierbaren Untergruppen, die als Spezies deklariert werden konnten. Offensichtlich bedeutet biologische Vielfalt nicht, dass alle denkbar möglichen sich reproduzierende Individuen entstehen, viel eher bilden sich ständig Nischen, die neue Spezies hervorbringen, die wiederum als Nischen für die weitere Evolution dienen [11], [12], wodurch sich die Biosphäre in gewisser Weise ordnet. Abstrakte [13], [14] oder molekulare Systeme [15], [16] eignen sich zur Analyse grundlegender evolutionärer Prozesse. Üblicherweise selektiert ein in molekularen evolutionären Modellsystemen herrschender konstanter Selektionsdruck eine optimale Evolutionslösung heraus, die mit dem Stillstand der Evolution in einem (Quasi)-Endzustand verbunden ist [17]–[25]. Jedoch streben natürliche Ökosysteme, selbst ohne variierenden externen Selektionsdruck [26]–[28], stetig als eng verflochtenes Kollektiv nach höherer Funktionalität [29], [30]. Bisher sind diese der Evolution zugrundeliegenden dynamischen Kopplungen wenig verstanden. Wie sich andauernde Evolution koexistierender Spezies in molekularen Evolutionssystemen realisieren lässt, bleibt bis dato ungelöst. Die in Kap. 3 durchgeführte Studie beschäftigt sich mit der Erforschung möglicher Konditionen, die es ermöglichen, dass molekulare Evolution unter konstanten äußeren Bedingungen unbegrenzt andauert.

Im Lauf der Evolution wurden die sogenannten 20 kanonischen Aminosäuren selektiert, die bei fast allen Organismen die Bausteine der Proteine bilden (vgl. Kap. 4 aus [31]). Es gibt nur wenige spezielle Organismen, bei denen der kanonische Satz ausnahmsweise um weitere Aminosäuren erweitert wird [32], [33]. Biologische Aktivität wird den Proteinen aufgrund ihrer dreidimensionalen Struktur verliehen [34]. Die dreidimensionale Struktur wiederum wird entscheidend von den Wechselwirkungen zwischen den funktionellen Gruppen der Aminosäuren, die das Protein zusammensetzen, beeinflusst (vgl. Kap. 4 aus [31]). Somit lassen sich zwar mit Hilfe des kanonischen Satzes der Aminosäuren eine große Menge von an das Leben angepasste Proteinfunktionen erzeugen, jedoch müssen diese aufgrund der limitierten Zahl an bei der Proteinsynthese berücksichtigten Aminosäuren keineswegs ideal sein [35]. Diese funktionalen Limitationen aufzuheben und Proteine mit neuen physiko- und biochemischen Eigenschaften zu erzeugen, beabsichtigt der gezielte Einbau nicht-kanonischer Aminosäuren in Proteine [35], [36]. Bei der chemischen Peptidsynthese lassen sich nur kleine Peptide mit einer Länge von bis zu etwa 50 Aminosäuren synthetisieren [37], [38]. Daher hat sich beim Einbau nicht-kanonischer Aminosäuren in Proteine die ribosomale Proteinbiosynthese mittels homo- oder heterologer Genexpression durchgesetzt [39]. Häufig inhibieren nicht-kanonische Aminosäuren-Analoga das Zellwachstum des Expressionswirts [36], was den *in-vivo*-Ansatz einschränkt. Beispielsweise werden sie nicht nur in das Zielprotein eingebaut, sondern von der molekularen Translationsmaschinerie ebenfalls in die wirtseigenen Proteine, beispielsweise in Enzyme, inkorporiert [36], [40]. Insofern gestaltet sich der Einbau nicht-kanonischer Aminosäuren, die starken Einfluss auf die Proteinfaltung haben, als besonders herausfordernd. Jedoch ist der Einbau von solchen Aminosäuren zum Zweck der Beeinflussung

der Proteinstruktur und dem damit verbundenen Potential, außergewöhnliche Proteineigenschaften erzeugen zu können, vielversprechend. Ein Beispiel ist das toxische, nicht-kanonische Arginin-Analogon Canavanin, das in *Escherichia coli* sogar zum unmittelbaren Zelltod führen kann [41], jedoch aus medizinischer Betrachtung Relevanz besitzt, da es anti-virale Eigenschaften zeigt [42], Stimulator für menschliche Autoimmunkrankheiten sein kann [43] und in vielen Studien, u.a. [44]–[49], auf seine anti-Krebs-Eigenschaften hin untersucht wurde. Kap. 4 beschäftigt sich mit dem zellfreien *in-vitro*-Einbau nicht-kanonischer Aminosäuren in Proteine, insbesondere mit dem Einbau von Canavanin, dessen vollständiger *in-vivo*-Einbau bisher nur einmal unter elaborierten Bedingungen nachgewiesen werden konnte [50]. Es wird untersucht, inwieweit sich ein zellfreies System [51], [52], das die Genexpression mit Hilfe der molekularen endogenen *Escherichia-coli*-Transkriptions-Translations-Maschinerie ermöglicht, zu diesem Zweck eignet.

Veränderungen der Zelleigenschaften resultieren nicht nur aus Mutationen, die die DNA-Sequenz bzw. Erbinformation dauerhaft verändern. Der Begriff Epigenetik umfasst alle Veränderungen in der zellulären Regulation der Genexpression, die vererbt und somit stabilisiert werden können, ohne dass dazu die zugrundeliegende DNA-Sequenz verändert wird [53], [54]: Ein epigenetisches Phänomen führt folglich zu einer vererbaren Änderung des Phänotyps einer Zelle, ohne dazu den Genotyp zu verändern. Eine enger gefasste Definition der Epigenetik fordert, dass sich die Veränderung des Phänotyps schalterartig vollzieht, sich demnach entweder vollumfänglich oder gar nicht manifestiert [55]. Diese Definition bezieht somit den in Kap. 5 studierten Phasenvariationsmechanismus uropathogener *Escherichia coli* ein [55], mit dem die Bakterien die Synthese sogenannter Pili regulieren können [56], [57]. Mit Hilfe dieser linearen haar- und polymerartigen Strukturen können die Bakterien an der Epithelzellschicht der ableitenden Harnwege des Wirts adhären [56], [57] und Urinaltraktinfektionen verursachen [58]. Die Phasenvariation wird auf Ebene der Transkription reguliert [59]. Aus diesem Grund stellt das bereits zum Einbau von nicht-kanonischen Aminosäuren verwendete zellfreie *Escherichia-coli*-Transkriptions-Translations-System eine optimale Basis dar, den Phasenvariationsmechanismus zu untersuchen, ein Ansatz der bereits in Vorarbeiten [60] verfolgt wurde und an den die vorliegende Dissertation anknüpft. Insbesondere, im Gegensatz zum *in-vivo*-Ansatz, bei dem verschiedene Umwelteinflüsse die molekulare Zusammensetzung der Zelle ständig verändern können [61], lassen sich die Konzentrationen von involvierten Koregulatoren gezielt variieren.

Auf das Vorwort folgt Kap. 2, das die zum Verständnis der in den Kap. 3–5 durchgeführten Studien notwendigen biochemischen und molekularbiologischen Grundlagen zusammenfasst und in die zellfreie Proteinsynthese einführt. Die durchgeführten Studien werden in sich abgeschlossenen Kapiteln präsentiert. Jedes dieser Kapitel beginnt mit einem einführenden Teil, der den Stand der Wissenschaft zusammenfasst, die Studie motiviert und die Strategie der durchgeführten Experimente erläutert. Es folgen der Material- und Methoden-Teil, die Präsentation und Diskussion der Ergebnisse und eine abschließende Zusammenfassung, Konklusion sowie ein Ausblick.

2. Biochemische, molekularbiologische Grundlagen und Einführung in die zellfreie Proteinsynthese

Kap. 2 resümiert das biochemische, biophysikalische und molekularbiologische Grundlagenwissen, das für das Verständnis der in dieser Dissertation durchgeführten Studien relevant ist. Es führt in die (zellfreie) Proteinsynthese ein.

2.1. Aufbau von DNA, RNA und Proteinen

Um im weiteren Verlauf dieses Kapitels den Vorgang der genetischen Informationsübertragung, der sich Genexpression nennt, zu verstehen, wird zuerst der Aufbau von DNA, RNA und Proteinen beschrieben. Der Aufbau dieser Biomoleküle legt die biomolekularen Wechselwirkungen fest, die in jeder Studie dieser Arbeit von zentraler Bedeutung sind.

2.1.1. Aufbau von DNA und RNA

2.1.1.1. Aufbau der Monomere: DNA- und RNA-Nukleotide

Abschn. 2.1.1.1 fasst das relevante Grundlagenwissen aus den Abschn. 1.2 aus [62] und 3.3.1 aus [63] zusammen.

DNA und RNA sind Nucleinsäuren. Ihr Rückgrat ist ein Polymer, das aus Monomeren, die Nucleotide (Abb. 2.1) heißen, aufgebaut ist. Nucleotide setzen sich aus einer Pentose (Fünffachzucker), einer Stickstoffbase und einer Phosphatgruppe zusammen. Pentosen sind Zucker mit fünf Kohlenstoffatomen, die von 1' bis 5' nummeriert sind und einen Ring bilden (Abb. 2.1).

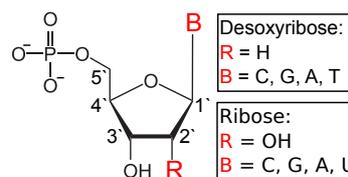


Abbildung 2.1.: Aufbau der Nucleotide, die sich aus einer Pentose, einer Stickstoffbase und einer Phosphatgruppe zusammensetzen. Erstellt in Anlehnung an die Idee von Abb. 3.8 aus [63].

Ist nur die Stickstoffbase über eine N-glykosidische Bindung am 1'-C-Atom der Pentose gebunden, wird von einem Nucleosid gesprochen. Aus dem Nucleosid wird ein Nucleotid, wenn am 5'-C-Atom zusätzlich mindestens eine Phosphatgruppe gebunden ist, die durch Veresterung der dortigen OH-Gruppe der Pentose mit Phosphorsäure entsteht. Des Weiteren befindet sich am 3'-C-Atom der Pentose eine OH-Gruppe, die zusammen mit der Phosphatgruppe am 5'-C-Atom die entscheidenden funktionellen Gruppen für die DNA- und RNA-Polymerisierung darstellt (vgl. Abschn. 2.1.1.2). Ein Unterscheidungsmerkmal zwischen DNA- und RNA-Nucleotiden ist durch die Struktur der Pentose gegeben: Während Ribose der Fünffachzucker der RNA-Nucleotide ist, findet sich 2'-Desoxyribose in DNA-Nucleotiden wieder. Diese beiden Pentosen unterscheiden sich am 2'-C-Atom

voneinander. Bei der Ribose ist an dieser Stelle eine weitere OH-Gruppe gebunden, am 2'-C-Atom der 2'-Desoxyribose ist hingegen ein H-Atom zu finden.

Ein weiteres Unterscheidungsmerkmal zwischen RNA- und DNA-Monomeren bilden die gebundenen Stickstoffbasen, vereinfacht auch Basen genannt (Abb. 2.1, Abb. 2.2B). Im Fall von DNA werden die Basen Adenin (A), Thymin (T), Guanin (G) und Cytosin (C) an die 2'-Desoxyribose gebunden. In RNA wird T durch Uracil (U) ersetzt. Im Gegensatz zu T fehlt U die Methylgruppe an Position 5. Die Basen werden zwei chemischen Klassen zugeteilt: den Purin- und Pyrimidinbasen. Zu den Purinbasen zählen Adenin und Guanin, die zwei überlappende heterozyklische Ringe mit fünf und sechs Gliedern besitzen. Thymin, Cytosin und Uracil sind Pyrimidinbasen, die einen sechsgliedrigen heterozyklischen Ring aufweisen. DNA- und RNA-Nukleotide werden zur Vereinfachung gemäß ihrer Stickstoffbase bezeichnet und abgekürzt (A, G, T/U und C). Während innerhalb der polymerisierten Nucleinsäuren nur die Monophosphate der Nucleoside zu finden sind (vgl. Abschn. 2.1.1.2), besitzen die monomeren Di- oder Triphosphate der Nucleoside im zellulären Stoffwechsel oder als Regulatormoleküle eine herausragende Bedeutung.

2.1.1.2. Basenpaarungen und Verknüpfung der Nucleotide zum Biopolymer

Abschn. 2.1.1.2 fasst das relevante Grundlagenwissen aus den Abschn. 1.3.1 aus [62], 3.3.1 und 7.2.1 aus [63] zusammen.

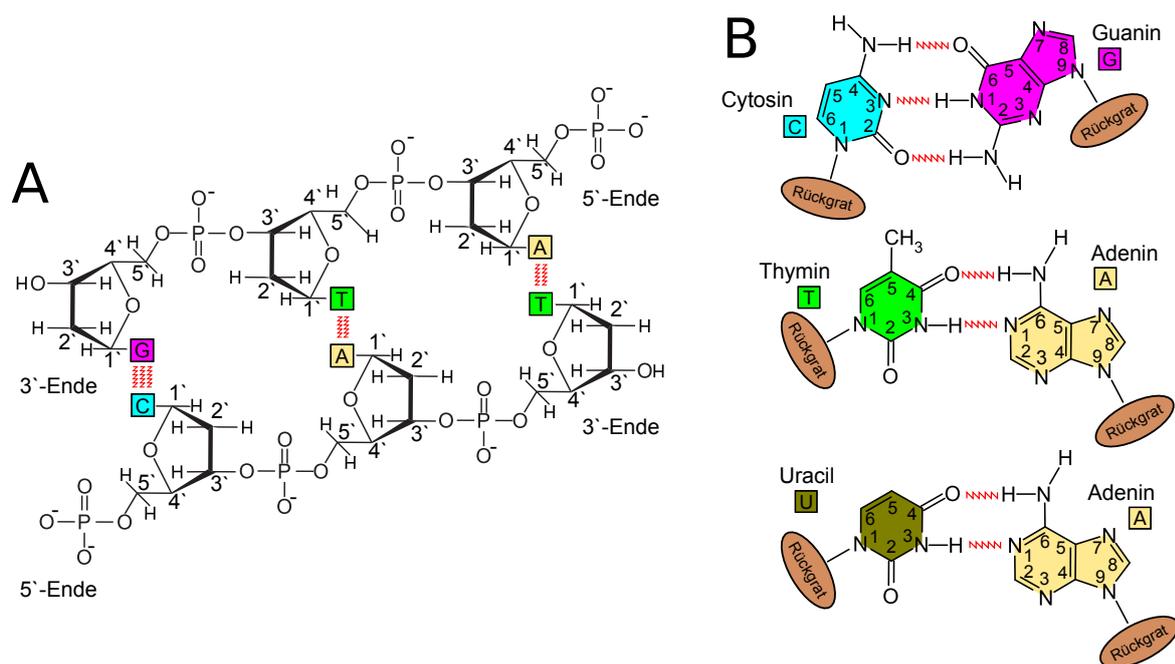


Abbildung 2.2.: Verknüpfung der DNA-Nucleotide zu einzelsträngigen DNA-Polymeren (A), die durch Watson-Crick-Paarungen (B) miteinander wechselwirken können (A,B). RNA-Polymere sind auf die gleiche Weise, wie in (A) dargestellt, aufgebaut, nur dass Ribose Desoxyribose und Uracil Thymin ersetzt. A erstellt entsprechend der Ideen nach Abb. 3.9 und 7.3 aus [63]. B erstellt entsprechend der Idee nach Abb. 7.4 aus [63].

DNA und RNA sind Polymere, die aus monomeren DNA- und RNA-Nucleotiden aufgebaut sind. Zwei aufeinanderfolgende Nucleotide sind dabei über eine Phosphodiesterbindung zwischen der OH-Gruppe

des 3'-C-Atoms des einen Nukleotids und der Phosphatgruppe des 5'-C-Atoms des anderen Monomers verbunden (Abb. 2.2A). Somit setzt sich das DNA/RNA-Rückgrat alternierend aus einer Pentose und einer Phosphatgruppe zusammen. Die entstehenden Polymere erhalten folglich eine Orientierung und enden an einer Seite mit einer Phosphatgruppe (5'-Ende) und einer freien OH-Gruppe an der anderen Seite (3'-Ende). Die Nukleotidsequenz oder vereinfacht Sequenz eines DNA- oder RNA-Einzelstrangs ergibt sich aus der Abfolge seiner Nukleotide, die durch die gebundene Stickstoffbase bezeichnet werden. Durch ihre Sequenz erhalten die DNA und RNA informationstragenden Charakter, der bei der Genexpression eine zentrale Rolle spielt und auf den im späteren Verlauf näher eingegangen wird.

A und T/U sowie G und C wechselwirken über die Ausbildung von Wasserstoffbrückenbindungen (Abb. 2.2B). Sie werden Watson-Crick-Paare genannt. Normalerweise paaren sich nur Purin- und Pyrimidinbasen miteinander, und nur unter speziellen Bedingungen kommt es zu andersartigen Paarungen, wie beispielsweise einer Basenpaarung zwischen zwei Purinbasen. Sofern jedem A ein T und T ein A bzw. jedem G ein C und C ein G zweier sich in entgegengesetzter Richtung gegenüberliegende DNA-Stränge gegenüberstehen, werden diese Stränge als komplementär oder als Stränge mit komplementären Sequenzen bezeichnet. Bei der A/T- beziehungsweise A/U-Paarung werden zwei und bei der C/G-Paarung drei H-Brückenbindungen ausgebildet. Dadurch ist die C/G-Paarung stabiler als die A/T-Paarung.

2.1.1.3. DNA-Doppelhelix und weitere Nukleinsäure-Strukturen

Abschn. 2.1.1.3 fasst das relevante Grundlagenwissen aus den Abschn. 1.3.2 aus [62], 3.3.1 und 7.2.1 aus [63] zusammen.

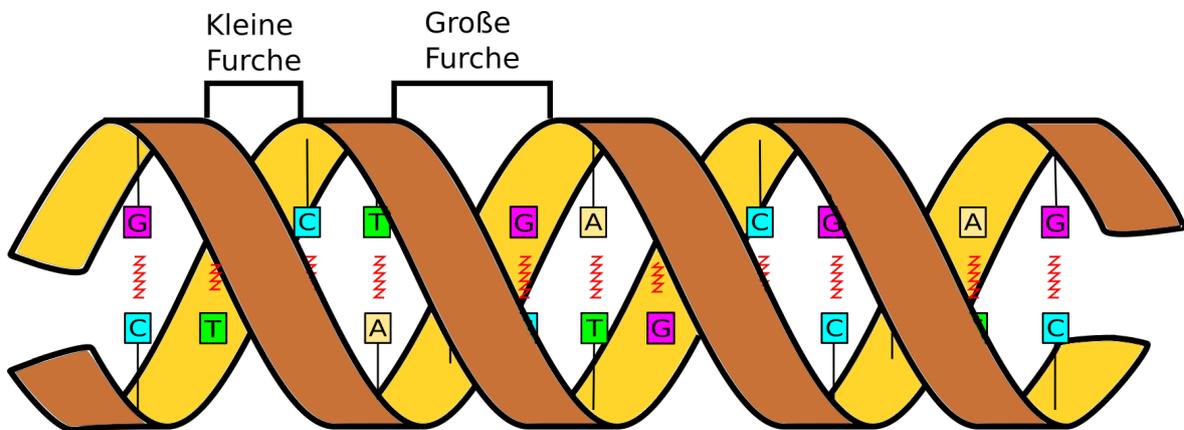


Abbildung 2.3.: Die am häufigsten vorkommende, rechtsgängige Konformation der DNA: die Doppelhelix oder B-Konformation. Erstellt entsprechend der Idee von Abb. 1.6 aus [62].

Die Watson-Crick-Paarungen zwischen komplementären DNA-Strängen ermöglichen ein antiparalleles Anlagern beider Einzelstränge zu einem Doppelstrang (Abb. 2.2A). Meist resultiert unter natürlichen Bedingungen daraus die schraubenförmige, rechtsgängige, doppelhelikale B-Konformation der DNA (Abb. 2.3), die 1953 von Watson und Crick aufgeklärt wurde [64].

In der räumlichen Struktur der B-DNA bilden die alternierenden Pentosen und Phosphatgruppen das außen liegende DNA-Rückgrat, während die Basen in Richtung des Helixinneren zeigen und sich mit den jeweils gegenüberliegenden, komplementären Basen paaren (Abb. 2.3). Da jedes Watson-Crick-Paar normalerweise aus einer Purin- und einer Pyrimidinbase besteht, füllen die gepaarten Basen in

etwa den gleichen Raum aus. Die rechtsgängige Doppelhelix besitzt eine regelmäßige Struktur und hat einen Durchmesser von etwa 2 nm. Innerhalb dieser DNA-Konformation besitzen zwei aufeinanderfolgende Basenpaare (bp) einen Abstand von 0,34 nm. Eine Windung setzt sich aus 10,5 bp zusammen, wodurch sich abgerundet eine Länge pro Windung von etwa 3,4 nm ergibt. Durch das Winden bilden sich die sogenannte große (2,2 nm) und kleine Furche (1,2 nm) aus (Abb. 2.3). Dort liegen die Basen direkt an der Oberfläche. Durch die außen liegenden Basen und die räumliche Struktur der Furchen bilden diese Bereiche oftmals einen Wechselwirkungsbereich für Enzyme und andere Biomoleküle. Die Doppelhelix wird maßgeblich durch hydrophobe π *stacking forces* oder π *stacking interactions* zwischen aufeinanderfolgenden Basen stabilisiert.

DNA existiert auch in der rechtsgängigen, doppelhelikalen A-Konformation, die unter natürlichen Bedingungen seltener als die B-DNA vorkommt und vor allem durch Wasserentzug entsteht. Aufeinanderfolgende Basenpaare der A-DNA besitzen einen Abstand von 0,26 nm. Die A-DNA-Konformation windet sich nach ca. 11 Basen einmal komplett, sodass eine Windung etwa eine Länge von ca. 2,9 nm besitzt. Eine weitere DNA-Konformation ist die linksgängige Z-Konformation, die durch alternierende Purin- und Pyrimidinbasen begünstigt wird, aber durch die räumliche Nähe der Phosphatgruppen zueinander thermodynamisch ungünstig ist.

RNA-Moleküle sind meist einzelsträngig, können allerdings mit sich selbst kurze charakteristische Faltungen (Sekundärstrukturen) ausbilden. Ein Beispiel hierfür ist die sequenzspezifische, charakteristische Struktur der tRNA, auf die später eingegangen wird (vgl. Abschn. 2.2.2.4).

Sekundärstrukturen können auch bei DNA auftreten: DNA-Stammschleifen-Strukturen entstehen durch Wiederholung einer beliebigen Sequenz in kreuzspiegelsymmetrischer Weise im gleichen DNA-Strang.

2.1.1.4. Schmelzen und Renaturierung von DNA

Abschn. 2.1.1.4 fasst das relevante Grundlagenwissen aus den Abschn. 1.4.2 aus [62], 7.2.1 aus [63] und 4.5 aus [65] zusammen.

Da die Bildung des DNA-Doppelstrangs auf reversiblen Wasserstoffbrückenbindungen basiert, kann der Doppelstrang durch Energiezufuhr wieder in zwei Einzelstränge getrennt werden. Dieser Prozess wird als Denaturierung bezeichnet. Geschieht dies über die Zufuhr von Wärme, wird vom Schmelzen der DNA gesprochen. In einer Lösung aus DNA-Doppelsträngen einer gewissen Sequenz wird die Schmelztemperatur T_m als die Temperatur definiert, bei der die Hälfte aller Doppelstränge denaturiert ist. Wird ausgehend von T_m die Temperatur weiter erhöht, steigt der Anteil der Einzelstränge an, und die Konzentration der Doppelstränge nimmt gleichzeitig ab. Bei Abnahme der Temperatur verschiebt sich das Verhältnis zwischen Einzel- und Doppelsträngen in die andere Richtung. T_m hängt vor allem von der DNA-Sequenz, DNA-Länge und Salzkonzentration ab. Da G/C-Paarungen eine Wasserstoffbrückenbindung mehr als A/T-Paarungen besitzen, muss mehr Energie zu ihrer Trennung aufgewendet werden. Insofern steigt T_m mit zunehmender DNA-Länge und dem relativen Anteil an G/C-Paaren (G/C-Anteil). T_m steigt ebenfalls mit zunehmender Konzentration stabilisierender Kationen.

Die erneute Anlagerung zweier Einzelstränge eines zuvor denaturierten Doppelstrangs wird als Renaturierung bezeichnet. Stammen die beiden sich anlagernden (teils) komplementären Einzelstränge aus unterschiedlichen Quellen, wird dieser Vorgang Hybridisierung genannt.

2.1.2. Aufbau der Proteine

Abschn. 2.1.2 fasst das relevante Grundlagenwissen aus den Abschn. 3.3.2-3.3.4 aus [63], 5.1.2 und 5.4.1 aus [62] sowie 3.5 und Kap. 4 aus [31] zusammen.

2.1.2.1. Der Aufbau der Aminosäuren

Wie DNA und RNA setzen sich Proteine aus Monomeren, den Aminosäuren, zusammen. Als proteinogene Aminosäuren werden solche Aminosäuren bezeichnet, die während der Proteinbiosynthese als Bausteine der Proteine verwendet werden. Dies sind ausschließlich sogenannte α -Aminosäuren, deren chemisches Grundgerüst in Abb. 2.4 dargestellt ist, und die sich dadurch auszeichnen, dass die Aminogruppe am benachbarten C-Atom der Carboxylgruppe gebunden ist. Aminosäuren werden durch zwei weitere funktionelle Gruppen charakterisiert: die Carboxylgruppe (COOH-Gruppe) und die Aminogruppe (NH₂-Gruppe).

Bei α -Aminosäuren sind neben diesen beiden Gruppen nur noch eine für die Aminosäure charakteristische Seitenkette, auch Aminosäurerest (R) genannt, und ein Wasserstoffatom an das zentrale C _{α} -Atom gebunden.

Im Folgenden wird sich auf die proteinogenen α -Aminosäuren beschränkt, da sie bei der für diese Arbeit relevanten Proteinbiosynthese zu Proteinen polymerisiert werden.

2.1.2.2. Eigenschaften der Aminosäuren

Wie gut eine Aminosäure dissoziiert, wird durch den pH-Wert der Lösung bestimmt. Im neutralen Milieu liegen Aminosäuren als Zwitterionen vor, bei denen die Aminogruppe protoniert und die Carboxylgruppe deprotoniert ist, sodass sich die Ladungen des Zwitterions zur Nettoladung 0 kompensieren. Im sauren Milieu bleibt die COOH-Gruppe ungeladen, während die NH₂-Gruppe einfach positiv geladen ist. Bei alkalischen pH-Werten kehrt sich dieser Zustand um. Des Weiteren hängt die Nettoladung der Aminosäuren auch von der chemischen Natur ihrer Seitenketten ab. Dies spielt vor allem dann eine Rolle, wenn die Aminosäure über ihre beiden funktionellen Gruppen zum Protein verknüpft wird (vgl. Abschn. 2.1.2.4), da die verknüpften Gruppen nicht mehr ionisierbar sind. Der pH-Wert, bei dem bei einer Aminosäure oder einem Protein saure oder basische Reste in gleicher Anzahl vorliegen, heißt isoelektrischer Punkt.

Die Summenformel eines Moleküls alleine gibt keine Auskunft, wie die beteiligten Atome innerhalb des Moleküls angeordnet sind. Da die biochemischen und physikalischen Eigenschaften vor allem von der Struktur eines Moleküls abhängen, können demnach zwei Moleküle mit derselben Summenformel andere Eigenschaften besitzen. Diese sogenannten isomeren Verbindungen lassen sich in Konstitutions- und Stereoisomere unterteilen. Konstitutionsisomere oder Strukturisomere unterscheiden sich trotz gleicher

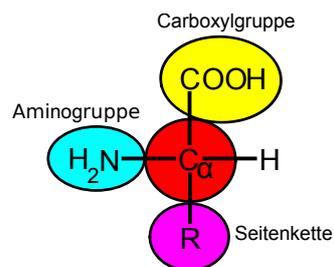


Abbildung 2.4.: Chemisches Grundgerüst der α -Aminosäuren. Am zentralen C _{α} -Atom ist ein H-Atom, eine COOH-, NH₂-Gruppe und eine für die Aminosäuren charakteristische Seitenkette gebunden. Erstellt entsprechend der Idee nach Abb. 4.1 aus [31].

Summenformel in ihrer Struktur, da Atome oder Atomgruppen innerhalb des Moleküls unterschiedlich angeordnet sind. Stereoisomere differieren trotz gleicher Struktur in der räumlichen Anordnung ihrer Atome. Eine Klasse dieser Stereoisomere bilden die Enantiomere. Enantiomere sind Konfigurationsisomere, die auch als chirale Moleküle bezeichnet werden und ein Asymmetriezentrum (Chiralitätszentrum) aufweisen. Mit Ausnahme von Glycin (R=H-Atom) bildet bei allen α -Aminosäuren das C_α -Atom ein solches Zentrum, da dort vier verschiedene Liganden binden (vgl. Abb. 2.4). Enantiomere lassen sich nicht durch Drehung mit ihrem Spiegelbild decken und sind optisch aktiv. Darunter wird die Eigenschaft eines Stoffes, die Polarisationssebene von linear polarisiertem Licht zu drehen, verstanden. Wird die Polarisationssebene im Uhrzeigersinn gedreht, ist die chemische Verbindung rechtsdrehend ((+)-Form). Bei Drehung gegen den Uhrzeigersinn wird der Stoff als linksdrehend ((-)-Form) bezeichnet. Eine äquimolare Mischung zweier entgegengesetzt drehender Enantiomere heißt Racemat, ein optisch inaktives Gemisch. Mit Ausnahme von Glycin existieren von jeder proteinogenen Aminosäure zwei Stereoisomere, eine L- und eine D-Aminosäure. Die meisten biologischen Systeme wechselwirken stereospezifisch mit Biomolekülen und können zwischen verschiedenen Stereoisomeren differenzieren. Beispielsweise werden bei der natürlichen Proteinbiosynthese ausschließlich die L-Aminosäuren von der endogenen, zellulären Translationsmaschinerie erkannt und gebunden.

2.1.2.3. Kanonische und nicht-kanonische Aminosäuren

Bei der natürlichen Proteinbiosynthese werden in der Regel 20 verschiedene L- α -Aminosäuren berücksichtigt. Diese 20 Aminosäuren werden als kanonische Aminosäuren, Standard-Aminosäuren oder natürliche Aminosäuren bezeichnet. Sie unterscheiden sich nur durch ihre Seitenketten, die ihre biochemischen und -physikalischen Eigenschaften sowie jene des Proteins bestimmen (vgl. Abschn. 2.1.2.2, Abschn. 2.1.2.4). Gemäß der chemischen Eigenschaften ihrer Seitenketten lassen sich die kanonischen Aminosäuren in vier Gruppen einteilen. Zu den kanonischen Aminosäuren mit hydrophoben Seitenketten zählen Glycin (Gly, G), Leucin (Leu, L), Isoleucin (Ile, I), Valin (Val, V) und Alanin (Ala, A). Zur Gruppe der kanonischen Aminosäuren mit ungeladenen polaren und aromatischen Seitenketten zählen Phenylalanin (Phe, F), Cystein (Cys, C), Methionin (Met, M), Threonin (Thr, T), Serin (Ser, S), Tryptophan (Trp, W), Prolin (Pro, P), Tyrosin (Tyr, Y), Glutamin (Gln, Q) und Asparagin (Asn, N). Die kanonischen Aminosäuren mit polaren Seitenketten unterteilen sich in solche mit sauren (Glutaminsäure (Glu, E) und Asparaginsäure (Asp, D)) und basischen (Histidin (His, H), Lysin (Lys, K) und Arginin (Arg, R)) Seitenketten. Im Folgenden wird der angegebene Ein- oder Dreibuchstabencode verwendet.

Einige kanonische Aminosäuren werden nach ihrem Einbau in ein Protein chemisch modifiziert. Zu diesen sogenannten posttranslationalen Modifikationen zählen Glykosylierungen, Acylierungen, Phosphorylierungen und Sulfatisierungen.

Neben den 20 universellen kanonischen Aminosäuren wurden bisher noch zwei weitere proteinogene Aminosäuren, Selenocystein (Sec) [32] und Pyrrolysin (Pyl) [33], gefunden. Sie sind in speziellen Organismen Teil des natürlichen Proteinbiosyntheseprozesses und gelten daher als 21. und 22. kanonische Aminosäuren.

Eine gesonderte Rolle nimmt Formylmethionin (fMet) ein, das bei Bakterien die erste Aminosäure eines jeden Proteins darstellt. Es unterscheidet sich vom kanonischen Met dadurch, dass ein Formylrest das H-Atom der Aminogruppe ersetzt.

Als nicht-kanonische Aminosäuren werden alle anderen Aminosäuren bezeichnet, die in Proteine eingebaut werden können [66].

2.1.2.4. Verknüpfung der Aminosäuren zu Proteinen und deren Struktur

Bei der Proteinbiosynthese bilden die kanonischen Aminosäuren die Monomere der Proteine (vgl. Abschn. 2.1.2.3). Die Verknüpfung zweier Aminosäuren vollzieht sich an ihren funktionellen Gruppen. Unter Abspaltung eines H_2O -Moleküls wird die Carboxylgruppe der einen Aminosäure mit der Aminogruppe der anderen Aminosäure verknüpft. Die dabei entstehende Bindung wird als Peptidbindung bezeichnet.

Bei der Polymerisierung werden sukzessive weitere Aminosäuren an das Ende der Polymerkette angehängt. Es entsteht das Peptidrückgrat, das das Atommuster N-C-C repetitiv aufweist. Nur die endständigen funktionellen Gruppen können ionisiert werden, da die inneren funktionellen Gruppen das Peptidrückgrat bilden. Das Aminogruppen- und Carboxylgruppen-Ende wird als N-Terminus bzw. C-Terminus bezeichnet. Dadurch, dass die Peptidbindung zwei mesomere Grenzstrukturen besitzt, von denen eine zu einer Doppelbindung zwischen dem Kohlenstoff- und Stickstoffatom führt, sind Rotationen nur an den $\text{C}_\alpha\text{-N}$ - und C-C_α -Bindungen möglich.

Werden viele Aminosäuren über Peptidbindungen miteinander verknüpft, entsteht ein Polypeptid; ein Protein besteht aus mindestens einem Polypeptid. Die Aminosäuren-Sequenz der Peptidkette wird als Primärstruktur bezeichnet. Die Wechselwirkungen zwischen den Seitenketten (vgl. Abschn. 2.1.2.3) sind entscheidend dafür, dass sich das Polypeptid auf charakteristische Weise dreht und dreidimensional anordnet: es faltet sich und bildet Sekundärstrukturen aus.

Es gibt folgende Beispiele für solche charakteristische Wechselwirkungen. Die unpolaren und aliphatischen Aminosäuren (Ala, Val, Leu, Ile) stabilisieren die Proteinstruktur durch hydrophobe Wechselwirkungen, die sogar wasserfreie Bereiche im Proteininneren schaffen können. Aromatische Seitenketten (Phe, Tyr, Trp) können aufgrund ihrer hydrophoben Wechselwirkung stabil übereinander gestapelt werden. Kovalente und somit sehr stabile Disulfidbrücken entstehen zwischen zwei freien Sulfhydrylgruppen (SH-Gruppe) von Cys. Die Aminosäuren Ser, Thr, Cys, Asn und Gln mit polaren Seitenketten wechselwirken innerhalb der Proteinstruktur über Wasserstoffbrückenbindungen. Die Seitenkette von Prolin ist innerhalb eines Fünfferrings mit dem N-Atom der NH_2 -Gruppe verknüpft, sodass die Bindungsebene zwischen dem C_α -Atom und der NH_2 -Gruppe fixiert ist. Dadurch wird an dieser Stelle das Protein-Rückgrat leicht abgelenkt, und die Polypeptidkette ist nicht mehr frei rotierbar. Die sauren und basischen Aminosäuren (Asp, His, Glu, Lys und Arg) orientieren sich bei Wasserkontakt nach außen und stabilisieren die Proteinstruktur.

Zwei prominente Sekundärstrukturen sind die rechtsgängige α -Helix und das β -Faltblatt. Bei der rechtsgängigen α -Helix zeigt das Peptidrückgrat nach innen und die Aminosäuren-Seitenketten nach außen; stabilisiert wird die recht flexible Sekundärstruktur jeweils durch Wasserstoffbrückenbindungen zwischen einer NH- und der CO-Gruppe der drittnächsten Aminosäure. Im β -Faltblatt faltet sich die Polypeptidkette linear vor und zurück. Die sich gegenüberliegenden Polypeptidabschnitte der eher festen Struktur können antiparallel oder parallel verlaufen (energetisch ungünstiger), wobei die Seitenketten aus der Faltblattstruktur herausragen. In diesem Zusammenhang ist noch die übergeordnete β -Faltblattstruktur namens β -Fass zu nennen, bei der mehrere β -Faltblätter im Kreis angeordnet sind.

Nach Ausbildung von Sekundärstrukturen falten sich Polypeptide durch die Wechselwirkungen ihrer Aminosäuren ein weiteres Mal, sodass sich die für das Protein charakteristische Tertiärstruktur bildet, die ihm seine biologische Funktion verleiht. Zwischen den verschiedenen Sekundärstrukturen liegen weite Bereiche, die von keiner regelmäßigen Struktur geprägt sind. Dort knäult das Protein in Abhängigkeit der Wechselwirkungen zwischen den beteiligten Aminosäuren. Auch diese Regionen tragen zur biologischen Funktion des Proteins bei, indem sie beispielsweise Bewegungen zwischen den verschiedenen Sekundärstrukturen zulassen. Übergeordnet können sich mehrere Polypeptide zu einer übergeordneten Quartärstruktur anlagern. Innerhalb dieser Proteinstruktur besitzt jedes Polypeptid bereits eine eigene Tertiärstruktur.

Proteine können denaturieren, beispielsweise durch hohe Temperaturen, bestimmte Chemikalien, pH-Werte oder Metalle, und verlieren somit ihre Sekundär- und Tertiärstrukturen. Meist ist dieser Prozess irreversibel, und das Protein verliert seine biologische Aktivität ganz oder zumindest teilweise.

2.2. Fluss der genetischen Information und die Genexpression

2.2.1. Das zentrale Dogma der Molekularbiologie

Bereits 1958 beschäftigte sich CRICK mit dem Informationsfluss zwischen den drei Biopolymeren DNA, RNA und Protein und postulierte sein zentrales Dogma, in dem er es für möglich erachtete, dass Information von Nukleinsäuren auf Nukleinsäuren oder von Nukleinsäuren auf Proteine übertragbar ist [67]. Den Informationstransfer von Proteinen auf Proteine oder von Proteinen auf Nukleinsäuren postulierte er als unmöglich [67]. 1970 griff er sein Dogma auf und konkretisierte, dass es darin um den sequentiellen Transfer von Information geht und dass diese Art von Information nicht von Proteinen auf Proteine oder von Proteinen auf Nukleinsäuren übertragbar sei [68].

Er teilte zudem die theoretisch neun möglichen Pfade, auf denen ein Informationsfluss von Polymer zu Polymer stattfinden kann, in drei Gruppen auf [68]. Der allgemeine Transfer (DNA \rightarrow DNA, DNA \rightarrow RNA, RNA \rightarrow Protein) kann bei allen Zellen stattfinden, bezieht kleinere Ausnahmen mit ein, während der spezielle Transfer nur unter außergewöhnlichen Umständen in wenigen Zellen stattfindet [68]. Als mögliche Kandidaten für den speziellen Transfer erachtete er die Informationsübertragung von RNA \rightarrow RNA, RNA \rightarrow DNA und DNA \rightarrow Protein [68]. Die im zentralen Dogma postulierten nie stattfindenden Informationsübertragungen, bei denen Proteine den Ausgangspunkt bilden, definierte er als unbekannte Transfers (Protein \rightarrow Protein, Protein \rightarrow DNA und Protein \rightarrow RNA) [68].

2.2.2. Prokaryotische Organisation von DNA, DNA-Replikation und Genexpression

Zur allgemeinen Übertragung des Informationsflusses (vgl. Abschn. 2.2.1) zählen die zellulären Prozesse der DNA-Replikation (DNA \rightarrow DNA), Transkription (DNA \rightarrow RNA) und Translation (RNA \rightarrow Protein). Der Informationsfluss von DNA \rightarrow RNA \rightarrow Protein wird als Genexpression bezeichnet. Die eukaryotische Replikations- und Genexpressions-Maschinerie ist wesentlich komplexer als die prokaryotische. Da für die vorliegende Dissertation die prokaryotische DNA-Replikation sowie die Genexpression von Bedeutung sind, wird sich im Folgenden auf die Beschreibung dieser Prozesse beschränkt.

2.2.2.1. Organisation der DNA

Abschn. 2.2.2.1 fasst das relevante Grundlagenwissen aus den Abschn. 7.1.1, 7.2.2–7.2.3 aus [63], Abschn. 6.2.1 aus [69] und [70] zusammen.

Das Genom umfasst die Gesamtheit aller Gene einer Zelle. Im Allgemeinen befindet sich der Hauptteil des Genoms im bakteriellen Chromosom, das ringförmig angeordnet ist und in einer einzigen Kopie im Zytoplasma vorliegt. Die ringförmige DNA wird mit Hilfe des Enzyms Topoisomerase II (DNA-Gyrase) um die eigene Achse überspiralisiert und somit dichter gepackt. Die Topoisomerase I hebt diese Überspiralisierung wieder auf. Durch das Wechselspiel der Topoisomerasen kann die DNA bei biologischen Prozessen wie der DNA-Replikation oder Genexpression relaxiert und im Anschluss wieder überspiralisiert werden. Damit nicht bei jeder Aktivität der Topoisomerasen die Ringstruktur komplett relaxiert oder gepackt werden muss, unterteilen stabilisierend DNA-Bindeproteine die Ringstruktur in mehrere überspiralisierte Domänen.

Neben dem Chromosom tragen Bakterien häufig weitere genetische Elemente auf sogenannten Plasmiden. Plasmide bestehen in der Regel aus doppelsträngiger DNA, die ebenfalls ringförmig und überspiralisiert vorliegt. Plasmide können sich autonom, unabhängig vom Chromosom, replizieren, tragen häufig lebenswichtige Gene für die Bakterien und spielen als molekularbiologische Werkzeuge eine wichtige Rolle. In diesem Zusammenhang wird in Abschn. 4.2.9.1 näher auf sie eingegangen.

2.2.2.2. DNA-Replikation

Abschn. 2.2.2.2 fasst das relevante Grundlagenwissen aus Abschn. 7.3 aus [63] zusammen.

Bei der DNA-Replikation duplizieren spezielle Enzyme die Doppelhelix samt der kodierten Information, sodass die Tochterzelle nach der Zellteilung die gleiche genetische Ausstattung besitzt. Die Replikation beginnt am Replikationsursprung, eine 300 bp lange Sequenz, an die ein Initiationsprotein bindet und die Helix öffnet. Im Anschluss bindet die Helikase, bildet die Spitze der sogenannten Replikationsgabel und spreizt die beiden Einzelstränge auf, deren Renaturierung durch das Binden vom Einzelstrangprotein verhindert wird. Die beiden gespreizten Einzelstränge dienen als Matrizen, um jeweils einen komplementären Einzelstrang zu synthetisieren, sodass nach Abschluss der Replikation zwei identische Doppelstränge existieren.

Die DNA-Polymerase III katalysiert das Anfügen eines Nukleotids über dessen 5'-Phosphatgruppe an das 3'-OH-Ende eines anderen Nukleotids. Somit verläuft die Replikation immer in 5' → 3'-Richtung. Deswegen unterscheidet sich der Replikationsmechanismus bei den beiden gespreizten, gegenüberliegenden Matrizen. Der von 5'-Phosphat- zu 3'-OH-Gruppe wachsende Strang (Leitstrang) wird kontinuierlich synthetisiert, da dort stets eine freie OH-Gruppe zur Verfügung steht.

Zum Start der DNA-Replikation des zweiten Strangs (Folgestrang) auf der gegenüberliegenden Matrize wird hingegen ein RNA-Primer, der eine freie OH-Gruppe bereitstellt und von dem aus die Strang-Polymerisierung startet, benötigt. Ein RNA-Primer ist ein kurzer einzelsträngiger RNA-Strang, dessen Polymerisierung von der Primase katalysiert wird. Für die Synthese des Folgestrangs muss die Primase in geeigneten Abständen mehrfach RNA-Primer auf der Matrize synthetisieren, von denen aus jeweils die DNA-Polymerase III Nukleotide an die freien OH-Gruppen bis zum nächsten RNA-Primer anhängt.

Im Anschluss entfernt die DNA-Polymerase I die RNA-Primer und ersetzt sie durch zur Matrize komplementäre DNA-Nukleotide. Zuletzt schließt die DNA-Ligase die dort verbleibenden Einzelstrangbrüche. Der gesamte Replikationsprozess läuft dynamisch und vor allem kooperativ ab. Die beteiligten Enzyme bilden einen gemeinsamen Komplex, der als Replisom bezeichnet wird. Neben den hier aufgezählten wichtigsten Enzymen unterstützen weitere den Replikationsprozess.

Im Allgemeinen verläuft bei Prokaryoten, die ein ringförmiges Chromosom tragen, die DNA-Replikation bidirektional ab: Vom Replikationsursprung aus laufen zwei Replikationsgabeln in entgegengesetzte Richtungen. Die geringe Fehlerrate (10^{-8} - 10^{-9} pro Base) bei der DNA-Replikation beruht auf einem Fehlerkorrekturmechanismus der DNA-Polymerase. Er erkennt falsch eingefügte Nukleotide und ersetzt sie durch das richtige, zur Matrize komplementäre Nukleotid.

2.2.2.3. Transkription

Abschn. 2.2.2.3 fasst das relevante Grundlagenwissen aus [71] und den Abschn. 7.5 und 8.3 aus [63] zusammen.

Bei der Transkription dient DNA als Matrize, um ein komplementäres Stück RNA mit Hilfe des Enzyms RNA-Polymerase zu synthetisieren. Die transkribierte RNA gliedert sich in die *messenger*-RNA (mRNA), Transfer-RNA (tRNA) und ribosomale RNA (rRNA).

Bei der bakteriellen Transkription wird die Synthese von mRNA, tRNA und rRNA nur von einem Typ RNA-Polymerase katalysiert, der sich bei *E. coli* aus fünf Untereinheiten zusammensetzt: je eine Kopie β' (156 kDa), β (151 kDa), σ^{70} (70 kDa) und 2 Kopien α (37 kDa).

Die σ^{70} -Untereinheit initiiert die Transkription, indem sie auf einem Bereich der DNA bindet, der Promotor heißt. Er beinhaltet u.a. zwei kurze Sequenz-Motive, die ungefähr 10 und 35 Basen vor dem Transkriptionsstartpunkt (+1-Stelle) liegen und bei allen Promotoren zu finden sind. Für σ^{70} von *E. coli* hat das -10-Element (Pribnow-Box) die Konsensussequenz 5'-TATAAT-3', die Konsensussequenz des -35-Elements lautet 5'-TTGACAT-3'. Die Konsensussequenz ist als die Sequenz definiert, die, unter Berücksichtigung aller Positionen der Sequenz, an jeder Stelle am wenigsten von einer gegebenen Menge verwandter Sequenzen abweicht [72]. Die Ähnlichkeit einer Promotor-Sequenz zur Konsensussequenz entscheidet über die Bindungsstärke des σ^{70} -Faktors und somit über die Transkriptionseffizienz des Promotors. Der initial an den Promotor bindende RNA-Polymerase-Komplex (Holoenzym) wird als geschlossen bezeichnet, da die DNA am Transkriptionsstartpunkt noch doppelsträngig ist. Erst nach Öffnung der Doppelhelix beginnt der offene Komplex, die RNA-Synthese an der +1-Stelle oder dem Transkriptionsstartpunkt zu katalysieren. Dabei dient der in 3'→5' laufende DNA-Strang als Matrize. Während der Elongation der RNA werden die Ribonukleotidtriphosphate mit der 3'-OH-Gruppe der Ribose des vorhergehenden Nukleotids der wachsenden RNA-Kette über eine Phosphodiesterbindung unter Eliminierung von Pyrophosphat verknüpft (RNA-Synthese-Richtung: 5'→3'). Nach der Initiation der Transkription löst sich σ^{70} ab, und die restlichen Untereinheiten, die das RNA-Polymerase-Kernenzym bilden, fahren mit der RNA-Synthese fort. Die Promotor-Region dient auch der Gen-Regulation auf Transkriptions-Ebene. Beispielsweise regulieren in der Nähe der Promotor-Region spezifisch bindende Moleküle (Repressoren) die Transkriptionsrate nach unten, indem sie mit der RNA-Polymerase um die Bindung konkurrieren. Alle auf diese Weise regulierten Gene bilden ein Operon. Der Locus, an dem der Repressor eine Bindung mit der DNA eingeht, heißt Operator-Region.

Die Transkription wird durch eine DNA-Sequenz, die als Terminator bezeichnet wird, beendet. Die Termination basiert auf der Destabilisierung der Bindung zwischen RNA-Polymerase und DNA. Eine Form der Termination ist die intrinsische Termination. Bei ihr besteht die DNA-Terminatorsequenz meist aus einer sich umgekehrt wiederholenden Sequenz, die durch eine sich nicht repetierende Sequenz getrennt ist und auf die eine A-reiche Sequenz folgt. Dadurch bildet die transkribierte RNA eine Stammschleifen-Sekundärstruktur aus, auf die eine U-reiche Sequenz folgt. Die RNA wechselwirkt in diesem Bereich eher mit sich selbst als mit der DNA, was zur Destabilisierung und zur Auflösung des ternären DNA/RNA-Enzym-Komplexes führt.

2.2.2.4. Translation

Abschn. 2.2.2.4 fasst das relevante Grundlagenwissen aus Abschn. 7.6 aus [63] und Kap. 5 aus [62] zusammen.

Die Translation übersetzt die in der Nukleotidsequenz gespeicherte genetische Information, die während der Transkription von der DNA auf die mRNA übertragen wurde, in die Aminosäuren-Sequenz (Primärstruktur) des zu synthetisierenden Proteins. Dazu transportieren spezifische tRNAs die benötigten Aminosäuren zu den Ribosomen, an denen diese Aminosäuren zur Polypeptidkette verknüpft werden.

Die Übersetzung der Nukleotidbasensequenz in die Primärstruktur des Proteins erfolgt nach den Regeln des genetischen Codes. Ein Codon ist ein Triplett drei aufeinander folgender mRNA-Nukleotidbasen, dem eine spezifische Aminosäure zugeordnet wird. Der genetische Code ist degeneriert: von den insgesamt 64 (4^3) möglichen Codons, kodieren 61 Codons für die 20 kanonischen Aminosäuren. Das bedeutet, dass die meisten kanonischen Aminosäuren von mehreren Codons kodiert werden. Somit kann einem Codon zwar eindeutig eine Aminosäure zugeordnet werden, es ist aber unmöglich, einer Aminosäure eindeutig ein Codon zuzuweisen.

Die drei Stopcodons oder Nonsense-Codons UAA (ocker, *ochre*), UAG (Bernstein, *amber*) und UGA (opalen, *opal*) kodieren für keine Aminosäure, was zum Abbruch der Translation führt. Die Translation beginnt in der Regel mit dem Startcodon AUG, das gleichzeitig für Met kodiert. Der genetische Code gilt als universell, da er nahezu für jeden bekannten Organismus gültig ist. Nur bei sehr wenigen Organismen zeigen sich Abweichungen von diesem Standardcode. Beispielsweise kodiert das UGA-Stopcodon bei eukaryotischen Mitochondrien für Trp. Wie bereits in Abschn. 2.1.2.3 besprochen, gehören Pyl und Sec bei speziellen Organismen zum kanonischen Aminosäuren-Standardrepertoire, werden mittels eines Stopcodons kodiert und während der Translation in Proteine eingebaut [32], [33].

Für die Proteinbiosynthese werden Vermittler, die tRNAs, benötigt, die ein gewisses Codon molekular erkennen und die vom Codon kodierte Aminosäure bereitstellen. tRNAs sind durchschnittlich 80 Nukleotidbasen lang, einzelsträngig und weisen ausgeprägte, Kleeblatt-ähnliche Sekundärstrukturen auf, die aus vier charakteristischen Armen oder Stämmen bestehen (Abb. 2.5). Die Sekundärstruktur ermöglicht die spezifische Wechselwirkung mit den anderen bei der Translation beteiligten Biomolekülen. Der Akzeptorarm bleibt am 3'-Ende einzelsträngig und besitzt immer ein gleich aufgebautes Ende. Auf die ungepaarte Diskriminator-Base folgt bei *E. coli* stets die Sequenz 5'-CCA-3', an die eine Aminosäure kovalent durch das Enzym Aminoacyl-tRNA-Synthetase (aaRS) an die tRNA gebunden wird (s. unten).

Die drei anderen Arme bilden an ihren Enden charakteristische Schleifen (Abb. 2.5): Die TΨC-Schleife trägt ihren Namen aufgrund der Sequenz TΨC (Ψ: Pseudouridin), die bei allen tRNAs zu finden ist; die D-Schleife wird nach der modifizierten Uridinbase Dihydrouridin benannt, der Anticodonarm trägt seine Bezeichnung wegen seines Anticodons. Das Anticodon besteht aus einem Ribonukleotid-Triplett, das zu einem gewissen Codon komplementär ist. Nach der Transkription eines *tRNA*-Gens befindet sich die tRNA in einem längeren Primärtranskript, aus dem sie durch verschiedene Enzyme und Ribozyme herausgeschnitten und prozessiert werden muss, um die beschriebene Sekundärstruktur zu erhalten. Viele tRNAs enthalten zusätzlich innerhalb des Anticodonstamms Introns, die durch spleißende Endonukleasen entfernt werden.

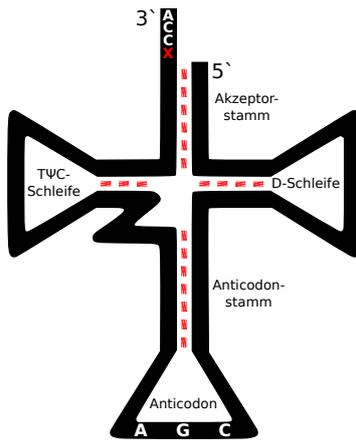


Abbildung 2.5.: Struktur einer *E.-coli*-tRNA mit der charakteristischen D-, Ψ- und Anticodon-Schleife sowie dem Akzeptorstamm. Erstellt in Anlehnung an die Ideen von Abb. 7.36 aus [63] und 5.4 aus [62].

Unterschiedliche Codons, die für die gleiche Aminosäure kodieren, unterscheiden sich häufig nur an der dritten Base. Fehlpaarungen zwischen Codon und Anticodon werden an dieser Stelle häufig toleriert (Wobble-Paarung). Beispielsweise kodieren bei *E. coli* die zwei Codons AAA und AAG für Lys, allerdings besitzt das Bakterium nur eine tRNA^{Lys} mit dem Anticodon UUU [73]. Die mittlere Position der Codons bestimmt hingegen weitgehend die Eigenschaften der Aminosäure, die eingebaut wird (U: hydrophob, A: hydrophil).

Entscheidend für den Einbau einer bestimmten Aminosäure an der richtigen Position in der wachsenden Polypeptidkette ist, dass die tRNAs mit der zum Codon/Anticodon passenden Aminosäure ausgestattet werden. Diesen Beladungsprozess gewährleisten die aaRS-Enzyme mit hoher Spezifität (Fehlerrate 0,01-0,1 % [74]). Die Beladung der tRNA mit einer Aminosäure verläuft in einer zweistufigen Reaktion unter ATP-Verbrauch, bei der die Aminosäure erst aktiviert und anschließend auf das 3'-Ende der tRNA übertragen wird.

Die aaRS-Enzyme werden in Klasse I- und II-Enzyme unterteilt, die eine Aminosäure entweder auf die 2'- oder die 3'-OH-Gruppe der endständigen tRNA-Ribose übertragen. Darüber hinaus entsteht die Bindung zwischen einer Klasse I-aaRS und ihrer entsprechenden tRNA über die kleine Furche in der RNA-Doppelhelix des tRNA-Akzeptorstamms. Eine Klasse II-aaRS bindet ihre tRNAs über die große Furche auf der gegenüberliegenden Seite des Akzeptorstamms. Neben dem Akzeptorstamm erkennt die aaRS die zur Aminosäure passende tRNA an der D-Schleife und insbesondere am Anticodon.

Eine ausreichend hohe Bindungsspezifität zwischen einer aaRS und ihrer Aminosäure entsteht meistens am katalytischen Zentrum, zu dem die eigene kanonische Aminosäure sterisch am besten passt und das die anderen kanonischen Aminosäuren sterisch hindert [74]. Einigen aaRS-Enzymen reicht dieser Mechanismus nicht aus, um eine adäquate Spezifität zu ihrer kanonischen Aminosäure zu generieren [74]. Diese Enzyme haben im Lauf der Evolution Fehlerkorrekturmechanismen entwickelt, die es ihnen erlauben, eine falsch gebundene Aminosäure wieder abzuspalten [74]. Beispielsweise können sowohl Ile als auch das kleinere Val ins katalytische Zentrum der Isoleucyl-tRNA-Synthetase gelangen, in dem dadurch beide Aminosäuren aktiviert und auf die tRNA^{Ile} geladen werden. Das Enzym besitzt ein zweites aktives Zentrum, die *editing*-Stelle, die die Valyl-tRNA^{Ile}-Bindung hydrolysiert. Bei den Fehlerkorrektur-aaRS wirken die beiden aktiven Seiten komplementär zueinander [74]: Zu große,

unpassende Aminosäuren werden am ersten aktiven Zentrum diskriminiert; zu kleine Aminosäuren, die fälschlicherweise im katalytischen Zentrum binden, werden von der *editing*-Stelle hydrolysiert.

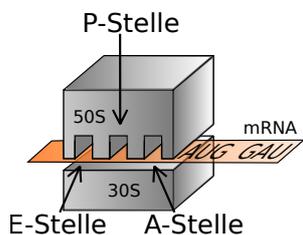


Abbildung 2.6.: Prokaryotisches Ribosom, das sich aus der großen 50S- und der kleineren 30S-Untereinheit zusammensetzt. Er stellt in Anlehnung an die Idee von Abb. 7.38 aus [63].

Auf der beladenen tRNA werden einige Aminosäuren vor dem Einbau in die wachsende Polypeptidkette modifiziert. U.a. besitzen nicht alle Organismen für jede kanonische Aminosäure eine eigene aaRS. Bei vielen Bakterien und Archaea werden beispielsweise jeweils durch eine aaRS die Salze der Glutamin- bzw. Asparaginsäure, Glutamat und Aspartat, jeweils auf die tRNA^{Glu} und tRNA^{Gln} bzw. tRNA^{ASP} und tRNA^{Asn} geladen. Im nächsten Schritt werden die falsch beladenen Glu-tRNA^{Gln} und Asp-tRNA^{Asn} von spezifischen Transamidasen amidiert. Dadurch entstehen die richtig mit Gln bzw. Asn beladenen tRNA^{Gln} und tRNA^{Asn}.

Wichtig für Translationsinitiation bei Bakterien ist die Präsenz von fMet (vgl. Abschn. 2.1.2.3), das aus Met gebildet wird. Für fMet existiert eine eigene Initiator-tRNA_I^{fMet}, die auf normalem Weg zuerst mit Met beladen wird. Anschließend wird die benötigte Formylgruppe durch die Methionyl-tRNA-Formyltransferase von N¹⁰-Formyltetrahydrofolat wie folgt übertragen [75]:



Die tRNA-Entladung und anschließende Verknüpfung der Aminosäuren zum Polypeptid mittels der Regeln des genetischen Codes geschieht an den Ribosomen. Das prokaryotische Ribosom besteht aus einer 50S- und einer kleineren 30S-Untereinheit, die sich zu einem Komplex (70S Ribosom) vereinigen (Abb. 2.6). Jede Untereinheit setzt sich wiederum aus mehreren rRNAs und Proteinen zusammen: Die 30S-Untereinheit besteht aus 21 verschiedenen Proteinen und einer 16S rRNA, die große 50S-Untereinheit hingegen aus 31 Proteinen und zwei rRNAs (23S und 5S rRNA). Bakterienzellen können bis zu 70000 Ribosomen enthalten. An den Ribosomen erfolgt die Bindung der beladenen tRNAs an die mRNA und die Proteinsynthese, bei der drei Regionen des Ribosoms eine entscheidende Rolle spielen: die Peptidyl-, Aminoacyl- und Exit-Stelle (P-, A-, E-Stelle). Eine mRNA wird im Allgemeinen von einem Polysom translatiert. Ein Polysom wird von mehreren Ribosomen gebildet, die sich in einem Abstand von etwa 80 Basen über die mRNA bewegen. Jedes Ribosom verknüpft dabei unabhängig voneinander in etwa 15-20 Aminosäuren pro Sekunde, sodass von jedem Ribosom das gleiche Polypeptid translatiert wird. Bei diesem Prozess ist die kleine 30S-Untereinheit für die mRNA-Bindung, den mRNA-Weitertransport und die tRNA-Erkennung verantwortlich. Die große 50S-Untereinheit beherbergt das Peptidyl-Transferase-Zentrum, in dem die Aminosäuren miteinander verknüpft werden, und bildet einen Austrittskanal für die wachsende Aminosäuren-Kette. tRNAs wechselwirken sowohl mit der großen als auch mit der kleinen Untereinheit.

Die Translation eines Proteins wird mittels drei Initiationsfaktoren initiiert. Im untranslatierten Bereich (*untranslated region* (UTR)) der mRNA, der stromaufwärts des Startcodons liegt, befindet sich die Shine-Dalgarno-Sequenz (5'-AGG-AGG-3'), die an die komplementäre Anti-Shine-Dalgarno-Sequenz der 16S-rRNA der 30S-Untereinheit bindet. Im weiteren Verlauf vereinigen sich die 50S-Untereinheit und 30S-Untereinheit zum 70S-Ribosom, bei dem sich fMet-tRNA_I^{fMet}, über das Anticodon mit dem Startcodon gepaart, an der P-Stelle befindet.

Bei der Elongation, die sich unter Einfluss von verschiedenen Elongationsfaktoren vollzieht, binden weitere mit Aminosäuren beladene tRNAs an der A-Stelle, die miteinander zum Polypeptid verknüpft

werden. Die dabei entstehende Polypeptidkette verlässt den Austrittskanal. Während dieses Prozesses wird die Methionyl-Gruppe vom N-terminalen fMet sofort durch das Enzym Peptid-Deformylase abgespalten, sodass wieder Met entsteht [76]. Allerdings wird das N-Terminale Met häufig enzymatisch durch das Enzym Met-Aminopeptidase abgespalten [77]–[79]. Diese Abspaltung ist umso wahrscheinlicher, je kleiner die Seitenkette der zweiten Aminosäure nach dem initialen Met ist [77]–[79]. Während der Elongation ist es der Elongationsfaktor EF-Tu, der die beladenen tRNAs erkennt und sie zur A-Stelle bringt, an der die Codon-Anticodon-Paarung auf Passgenauigkeit geprüft wird. Ein weiterer Elongationsfaktor heißt EF-G, der zur Translokation des Ribosoms um drei Ribonukleotide auf der mRNA führt. Die Translokation ermöglicht die Bindung der nächsten beladenen tRNA an der nun frei gewordenen A-Stelle. Die unbeladene tRNA wird an der E-Stelle entlassen.

Die Translation wird durch ein Stopcodon und mehrere Terminationsfaktoren (*release factors*, RF) beendet. Erreicht ein Stopcodon die A-Stelle, wird es von Terminationsfaktoren identifiziert. Bei *E. coli* werden das amber- und opal-Stopcodon von RF1 bzw. RF2 spezifisch erkannt. Das am häufigsten vorkommende UAA-Stopcodon erkennen sie hingegen beide. Die Termination wird durch das Binden von RF1 oder RF2 an der P-Stelle eingeleitet. Durch eine Konformationsänderung des Ribosoms und durch weitere spezifische Wechselwirkungen löst sich die Polypeptidkette vom Ribosom. Weitere RF führen zur Dissoziation der beiden ribosomalen Untereinheiten und setzen die mRNA und tRNA wieder frei.

Bleibt die A-Stelle im Ribosom eine längere Zeit unbesetzt, wird die Translation vorzeitig mit Hilfe einer *transfer-messenger*-RNA (tmRNA) terminiert, die bei *E. coli* ca. 350 Nukleotide lang ist. Die tmRNA hat einen Aminosäuren-Akzeptorstamm und zwei den D- und T Ψ C-Schleifen ähnelnde Sekundärstrukturen, aber anstelle der Anticodon-Schleife einen Verbindungsstamm, der eine kurze mRNA-Sequenz enthält. Das erste Codon dieser mRNA-Sequenz kodiert für Ala. Es folgt eine Nukleotidsequenz, die für Ala-Asn-Asp-Glu-Asn-Tyr-Ala-Leu-Ala-Ala kodiert. Ein Stopcodon bildet das Ende dieser mRNA-Sequenz. Die mit Ala beladene tmRNA bindet an der A-Stelle des Ribosoms, Ala wird an die letzte Aminosäure der wachsenden Polypeptidkette gehängt. Die mRNA-Sequenz der tmRNA wird bis zum Abbruch beim Stopcodon translatiert und dort terminiert. Durch das Anfügen von Ala-Ala-Asn-Asp-Glu-Asn-Tyr-Ala-Leu-Ala-Ala an das Polypeptid wird die Elongationsblockade aufgehoben. Zugleich ist die angefügte Aminosäuren-Sequenz ein Abbausignal, sodass das unfertige Protein abgebaut wird.

Viele der bei der Translation beschriebenen Prozesse, in die die verschiedenen Faktoren involviert sind, laufen unter Hydrolyse von Guanosintriphosphat (GTP) ab.

2.3. Mutationen, Regulation der Genexpression und Epigenetik

Die Eigenschaften und Funktionen einer Zelle werden im Wesentlichen durch die Mengen und Arten der in der Zelle vorhandenen Proteine bestimmt (vgl. Abschn. 3.3.3 aus [63]). Dementsprechend können Zellen ihre Eigenschaften durch Veränderung ihrer Proteinzusammensetzung verändern.

2.3.1. Mutationen

Abschn. 2.3.1 fasst das relevante Grundlagenwissen aus Kap. 10 aus [62], den Abschn. 3.3.3, 7.6.1 und 10.1 aus [63] zusammen.

Eine Möglichkeit, biologische Protein-Funktionen und somit die Eigenschaften einer Zelle zu modifizieren, ergibt sich durch die Substitution bestimmter Aminosäuren, einer Verkürzung oder Verlängerung des Proteins, da dies zu veränderten Wechselwirkungen zwischen den ins Protein eingebauten Aminosäuren und somit der Protein-Struktur führt. Solche Modifizierungen können durch Mutationen entstehen. Eine Mutation ist eine Veränderung der Erbinformation, d.h. der Nukleotidsequenz der DNA. Mutationen reichen vom Austausch, Einfügen oder Entfernen einzelner Nukleotide bis hin zu einer Insertion, Deletion oder Translokation ganzer Nukleotid-Sequenzen. Eine Zelle, deren genomische Nukleotidsequenz im Vergleich zur Mutterzelle mutiert ist, wird als Mutante bezeichnet; die Mutante besitzt einen anderen Genotyp als ihre Mutterzelle.

Zwei Zellen mit unterschiedlichen Genotypen müssen nicht zwangsläufig unterschiedliche Merkmale aufweisen. Ist beispielsweise die letzten Base der für Val kodierenden Sequenz GTA in einem Gen mutiert, kodiert es nach den Regeln des genetischen Codes immer noch für Val. Mutiert hingegen innerhalb eines Gens die für Arg kodierende Sequenz AGA zur für das Stopcodon kodierenden Sequenz TAG, wird das zu translatierende Protein drastisch verkürzt. Es verliert wahrscheinlich seine Funktion, und die Zelleigenschaften können sich je nach Wichtigkeit des Proteins stark verändern. Ähnliches gilt beim Einfügen oder der Deletion einer Base inmitten eines Gens. Solche Mutationen führen zu einer Verschiebung des offenen Leserahmens und einer veränderten Primärstruktur des Proteins. Mutationen, die keine Veränderung der Zelleigenschaften nach sich ziehen oder solche, die sich negativ auf die Zelle auswirken, treten häufiger auf als solche, die einer Zelle einen Wettbewerbsvorteil bei der Reproduktion und dem Überleben bringen. Letztere setzen sich evolutionär hingegen eher durch und manifestieren sich. Ein Austausch einer oder weniger Aminosäuren in einem bestimmten Bereich des Proteins kann dazu führen, dass es eine neue chemische Eigenschaft gewinnt, die sich auf die Mutante überträgt.

Mutationen entstehen spontan aufgrund chemischer oder physikalischer Einwirkungen auf die DNA (Strahlung, chemische Substanzen, etc.), die die Basenstruktur der DNA verändern. Während der DNA-Replikation macht die DNA-Polymerase häufig Fehler an Positionen mit solch einer veränderten Basenstruktur. Dazu zählen der Einbau von nicht-komplementären Nukleotiden, das Überspringen oder zusätzliche Einfügen von DNA-Monomeren im Tochterstrang.

Die bisher beschriebenen Mutationen werden als Genmutationen bezeichnet, da sie sich lokal auf einzelne Gene des Chromosoms beziehen. Bei eukaryotischen Zellen, deren Genom aus mehreren Chromosomen besteht, existieren weitere Mutationsformen. Bei einer Chromosomenmutation finden größere strukturelle Umbauten eines Chromosoms statt. Eine Genommutation umfasst darüber hinaus die Veränderung der Chromosomenzahl einer Zelle.

Mutationen, die sich evolutionär durchsetzen, verändern Zelleigenschaften dauerhaft, da sie vererbt werden. Eine einmal manifestierte Mutation lässt sich nur durch eine zufällige Rückmutation korrigieren.

2.3.2. Regulation der Genexpression und Epigenetik

Der erste Abs. von Abschn. 2.3.2 fasst das relevante Grundlagenwissen aus Kap. 11 aus [62] und Kap. 8 aus [63] zusammen. Eine Zelle benötigt nicht in jeder Situation jedes von der DNA kodierte Protein und muss vor allem kurzfristig auf variierende Umweltbedingungen reagieren. Daher besitzen Zellen komplexe und dynamische Regulationsmechanismen, mit denen sie in der Lage sind, ihre Proteinzusammensetzung im Zuge der sich ändernden Bedingungen anzupassen. Die Regulation kann

prinzipiell auf jeder Stufe der Genexpression stattfinden, angefangen bei der Transkription bis hin zur posttranslationalen Modifizierung des anhand des Gens synthetisierten Proteins. Beispielsweise wird bei der Regulation auf Transkriptionsebene die Transkriptionsrate der RNA-Polymerase durch bestimmte Regulatormoleküle, die Repressoren heißen, abgeschwächt. In der extremsten Form wird somit die Transkription eines Gens deaktiviert. Während Mutationen zufällig den Phänotyp einer Zelle dauerhaft verändern (vgl. Abschn. 2.3.1), verändern die im letzten Absatz beschriebenen Regulationsmechanismen die Merkmale einer Zelle nur so lange, wie ein bestimmter äußerlicher Reiz dies erforderlich macht.

Eine Sonderrolle nimmt die epigenetische Regulation ein, ein Begriff, dessen Definition in der Mitte des letzten Jahrhunderts vor allem von WADDINGTON geprägt [80]–[82] wurde und seitdem im stetigen Wandel ist [53], [54]. Eine minimalistische Definition von Epigenetik, auf die sich die wissenschaftliche Gemeinschaft heute weitgehend geeinigt hat, ist, dass die Epigenetik alle Veränderungen in der Regulation der Genexpression umfasst, die nach der Zellteilung auf die Tochterzelle vererbt und somit stabilisiert werden, ohne dass die Erbinformation (DNA-Sequenz) selbst verändert wird [53], [54]: Ein epigenetisches Phänomen führt folglich zu einer vererbaren Änderung des Phänotyps einer Zelle, ohne den Genotyp zu verändern. Eine erweiterte Definition fordert darüber hinaus, dass sich die Veränderung des Phänotyps schalterartig vollzieht, also entweder die Veränderung komplett oder gar nicht ausgeprägt ist und sie sich auch dann vererbt, wenn der verursachende Reiz nicht mehr existiert [55]. Diese Definition bezieht somit den in dieser Dissertation untersuchten Pili-Phasenvariationsmechanismus von uropathogenen *E.-coli*-Bakterien (vgl. Kap. 5) ein [55].

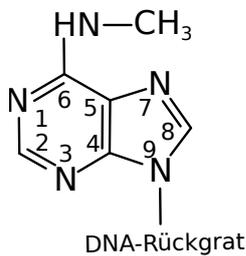


Abbildung 2.7.: Strukturformel von N⁶-Methyladenosin

Dieser Abs. fasst das relevante Grundlagenwissen aus [54] zusammen. Eine zentrale Rolle in der Epigenetik spielt die DNA-Methylierung in eukaryotischen Zellen, die durch die Enzymklasse der DNA-Methyltransferasen (DNMTs) katalysiert wird. Als Kofaktor benötigen diese Enzyme S-Adenosylmethionin (SAM). Bei der Methylierung katalysieren die DNMTs die Übertragung der Methylgruppe (CH₃) eines SAM-Moleküls auf eine Stickstoffbase der DNA. Dies geschieht bei gleichzeitiger Hydrolyse von SAM zu S-Adosylhomocystein (SAH). Auch nach Methylierung kommt es weiterhin zur Watson-Crick-Paarung zwischen komplementären Nucleotiden, sodass die DNA-Sequenz bei der Replikation durch diese Modifikation nicht verändert wird. Allerdings ist direkt nach der DNA-Replikation im neu synthetisierten Doppelstrang nur einer der beiden Einzelstränge methyliert. Dies ist der Einzelstrang, der der DNA-Polymerase als Matrize gedient hat. Dieser Zustand wird als hemimethyliert bezeichnet. Im Anschluss sorgen die DNMTs dafür, dass auch der unmethylierte Einzelstrang methyliert wird.

Die vorliegende Arbeit beschäftigt sich mit der bakteriellen DNA-Methylierung, die durch die DNA-Adenin-Methylase (DAM) katalysiert wird [83]. DAM katalysiert den Transfer einer Methylgruppe auf die 6-Aminogruppe des Adenins innerhalb von GATC-Tetrameren (6meA, Abb. 2.7), die das Enzym spezifisch erkennt [83]. Diese Erkennungssequenz ist im Schnitt alle 243 Nucleotide im *E.-coli*-Genom vorzufinden, wobei sie häufiger in kodierenden als in nicht-kodierenden oder nicht-translatierten Bereichen auftritt [83]. In der Regel sind die Tetramere methyliert, wobei sie auch in hemi- und unmethylierter Form vorliegen können [83]. DAM und spezifische DNA-Bindeproteine scheinen gerade an unmethylierten GATC-Stellen um die Bindung an die DNA zu konkurrieren, sodass davon ausgegangen wird, dass diese DNA-Bereiche bei der Genexpression-Regulation von Bedeutung sind [83].

2.4. Zellfreie Proteinsynthese

Zellfreie Proteinsynthese-Systeme besitzen die notwendige molekulare Maschinerie, um mittels Template-mRNA oder -DNA Proteine zu synthetisieren [84]. Die Grundlage vieler solcher Systeme bildet der sogenannte rohe Extrakt, der aus den Zellen extrahiert und für die zellfreie Proteinsynthese aufbereitet wird [84]. Die zellfreie Genexpression bildet in den Kap. 4 und 5 die experimentelle Umgebung der durchgeführten Studien. Es wird dort ein auf NOIREAUX *et al.* [51], [52] zurückgehendes, zellfreies Expressionssystem verwendet, das auf einem zellfreien *E.-coli*-Extrakt basiert.

2.4.1. Grundstein der modernen zellfreien Proteinsynthese

Die (frühe) Entwicklungsgeschichte der zellfreien Proteinsynthese (ZFPS) wurde mehrfach zusammengefasst, beispielsweise in [85], [86], die bei der Literaturrecherche zur Erstellung von Abschn. 2.4.1 herangezogen wurden.

Bereits um 1950 inkorporierten unter anderem [87]–[89] radioaktiv markierte Aminosäuren mittels eukaryotischer Gewebehomogenate in Proteine und wiesen somit nach, dass zur Proteinsynthese keine intakten Zellen notwendig sind. Dieser Nachweis gelang ebenfalls in prokaryotischen Zellsysaten, beispielsweise auf der Grundlage von Staphylokokken [90]. Diese frühen ZFPS-Systeme waren jedoch auf die Translation endogener mRNA beschränkt, die bereits vor dem Zellaufschluss an die Ribosomen gebunden waren [85].

1961 fanden NIRENBERG und MATTHAEI heraus, dass die Präinkubation ihres *E.-coli*-Extraktes bei physiologischen Temperaturen die anschließende Translation exogener RNA ermöglicht [91]. NIRENBERG mutmaßte, dass die Präinkubation die endogene mRNA depletiert, somit die zellfreie Translation exogener mRNA ermöglicht wird und teilweise in proportionalem Zusammenhang mit der Menge der exogenen mRNA steht [92]. Jedoch fand [93] Anhaltspunkte, dass der positive Effekt der Präinkubation, auch *Run-off* genannt, auf die zellfreie Proteinsyntheserate nicht nur ausschließlich darauf zurückzuführen sein kann, dass die Ribosomen von endogener mRNA befreit werden. [93] spekulierte, dass der *Run-off* Faktoren, die die Proteinsynthese begünstigen, aktiviert oder Substanzen, die die Proteinsynthese inhibieren, deaktivieren könnte.

Später ließ sich exogene mRNA auch in eukaryotischen Weizenkeimextrakten translatieren [94], [95]. Darüber hinaus konnte gezeigt werden, dass Weizenkeimextrakte wenig endogene mRNA enthalten und somit eine effiziente Translation exogener RNA ermöglichen [96], [97]. Hingegen gestaltete sich die Translation exogener mRNA in Kaninchen-Retikulocyten-Extrakten erst als effizient, nachdem die in hoher Konzentration enthaltene endogene mRNA enzymatisch vor der Proteinsynthese degradiert wurde [98].

Ebenfalls 1961 zeigten NIRENBERG und MATTHAEI, dass DNase-Aktivität die ZFPS anfangs gar nicht, aber nach einiger Zeit stark reduzierte [99]. Sie schlossen auf die *de-novo*-Synthese von mRNA durch im zellfreien System vorhandene DNA [99]. Diese Kopplung zwischen zellfreier Transkription und Translation wurde im weiteren Verlauf intensiv studiert, und es wurde versucht, das molekulare Zusammenspiel zwischen DNA, RNA und Protein mittels der ZFPS besser zu verstehen [100], [101]. ZUBAY *et al.* entwickelten ein gekoppeltes *E.-coli*-basiertes ZFPS, das im Vergleich zu früheren Systemen höhere Proteinausbeuten, insbesondere bei Verwendung von *E.-coli*-DNA, generierte [102], [103]. Dies erreichten sie durch Optimierung der Zusammensetzung der zellfreien Reaktion und der

Präinkubation des Extraktes bei 37 °C mit einem Mix, der vor allem aus den 20 kanonischen Aminosäuren, Phosphoenolpyruvat, der Phosphoenolpyruvat-Kinase, Adenosintriphosphat, Magnesium und Puffersubstanzen bestand [102], [103].

Weitere Optimierungsversuche bei der bakteriellen ZFPS zielten darauf ab, die hohe Nuklease-Aktivität abzuschwächen, indem beispielsweise zur Extrakt-Herstellung Bakterienstämme genommen wurden, denen RNase-Gene fehlten [104]. Ein Durchbruch in der ZFPS konnte erzielt werden, als das zu exprimierende Gen unter Kontrolle des T7-Promotors gestellt und die Synthese-stärkere T7-RNA-Polymerase aus dem T7-Bakteriophagen, die hochspezifisch auf ihrem Promotor bindet, in ein *E.-coli*-ZFPS integriert wurde [105]. Das Zurückgreifen auf die T7-Transkription in *E.-coli*-ZFPS-Systemen stellt auch bei der moderneren ZFPS den Standard dar [61], [106]. Die Integration von Bakteriophagen-Transkriptions-Systemen ermöglichte schließlich auch, effiziente eukaryotische ZFPS-Systeme zu entwickeln, in denen die Transkription und Translation kombiniert wurden [107].

2.4.2. Anwendungsbereiche verschiedener ZFPS-Systeme

Aus Zellen verschiedener Organismen konnten rohe Extrakte zum Zweck der ZFPS hergestellt werden [108]. Die Zahl an Organismen, aus denen Extrakte im kommerziellen Maßstab hergestellt werden können, ist jedoch gering [86]. Die Wahl eines bestimmten Organismus bringt Vor- oder Nachteile für die spätere ZFPS mit sich und hängt davon ab, für welche Art der ZFPS das resultierende System verwendet werden soll [108]. Auf Grundlage der Übersichtsartikel [86], [108], [109] zusammengefasst, lassen sich in aller Regel die größten Proteinausbeuten unter geringsten Kosten mit *E.-coli*-basierten ZFPS-Systemen erzielen, die darüber hinaus bereits im industriellen Maßstab betrieben werden können (100 Liter Reaktion und Ausbeuten von 700 mg/l in 10 h [110]). Eukaryotische ZFPS-Systeme bieten typischerweise mehr Möglichkeiten der posttranslationalen Proteinmodifikation; sie gewährleisten die richtige Faltung eukaryotischer Proteine, aber sie verursachen höhere Kosten bei geringen Proteinausbeuten (s. jeweils Tab. 1 in [86], [108], [109]).

Der Einbau nicht-kanonischer Aminosäuren in Proteine zielt u.a. darauf ab, Proteine mit neuen bio- und physikochemischen Funktionen herzustellen (vgl. Abschn. 4.1), die beispielsweise therapeutisch nutzbar gemacht werden können [108], [111]. Diese Anforderungen erfüllen aktuell vor allem die *E.-coli*-basierten ZFPS-Systeme aufgrund der einfachen und kostengünstigen Extrakt-Herstellung, ihrer preiswerten Versorgung mit zu ergänzenden Komponenten, ihren hohen Proteinausbeuten und der Möglichkeit, mit ihnen bereits Proteine im industriellen Maßstab zu produzieren (s. jeweils Tab. 1 in [86], [108], [109]).

2.4.3. *E.-coli*-basierte ZFPS

Im Folgenden wird auf die *E.-coli*-basierte ZFPS eingegangen, da diese für den Kontext der vorliegenden Arbeit relevant ist. Die *E.-coli*-basierte ZFPS existiert in verschiedenen Formaten. Die wichtigsten werden nachstehend präsentiert.

2.4.3.1. Extrakt-basierte ZFPS im Batch-Modus

Die Grundlage eines *E.-coli*-Extrakt-basierten ZFPS-Systems bildet der rohe *E.-coli*-Extrakt, der die recycelten Komponenten der molekularen Maschinerie der bei der Extrakt-Herstellung aufgeschlossenen *E.-coli*-Zellen beinhaltet und der durch Zentrifugation weitgehend von Zellwandresten und der chromosomalen Bakterien-DNA gereinigt ist [61]. Ein Vergleich (Abb. 1 aus [112], [113]) verschiedener Protokolle zur Extrakt-Herstellung zeigt, dass sich diese Protokolle in gemeinsame, teils optionale, Schritte unterteilen lassen, die sich jedoch in der detaillierten Ausführung unterscheiden: Das Zellmaterial von zuvor aufgeschlossenen Zellen (2) einer Zellkultur (1) wird zentrifugiert (3), das daraus resultierende Zelllysate wird dem sogenannten *Run-off* unterzogen (4), optional wird es anschließend dialysiert (5) und abschließend zentrifugiert (6). Der Extrakt wird typischerweise mit Aminosäuren, RNA-Nukleotiden, tRNA, Kofaktoren, Puffersubstanzen, einer Bakteriophagen RNA-Polymerase und der kodierenden DNA ergänzt [61]. Entscheidend ist zudem, dass die zur ZFPS benötigte Energie regeneriert wird [61].

Selbst noch um die Jahrtausendwende brach die ZFPS im Batch-Modus meist unter einer halben Stunde ab, und die Proteinausbeuten beliefen sich auf ca. 100 µg [114]. Auf die wichtigsten Weiterentwicklungen, die zu den heutigen effizienten Systemen geführt haben, wird nun eingegangen.

2.4.3.1.1. Energieregeneration und Vermeidung der Anhäufung von inhibitorischem anorganischem Phosphat

Adenosintriphosphat (ATP) ist der universelle molekulare Energielieferant für die lebende Zelle: Die Zelle hydrolysiert ATP zu Adenosindiphosphat (ADP), nutzt die frei werdende Energie, unter anderem für die Transkription oder Translation, und ist in der Lage, ADP über komplexe Stoffwechselvorgänge zu ATP zu regenerieren (vgl. Kap. 5 aus [63]). Eine Form der ATP-Regeneration ist die Glykolyse, bei der aus Glukose über eine Kette von enzymatischen Reaktionen Pyruvat als Zwischenprodukt entsteht, das entweder danach anaerob in verschiedene Produkte fermentiert oder aerob über den Zitronensäurezyklus in CO₂ und H₂O umgewandelt wird (vgl. Kap. 5 aus [63]).

In ZFPS-Systemen gestaltete sich die ATP-Regeneration anfangs hingegen einfach. Ein Beispiel für ein solches ATP-Regenerationssystem ist das Phosphoenolpyruvat-Pyruvatkinase-System (PEP-PK-System), das beispielsweise beim ZUBAY-System Anwendung fand [103]. Es entspricht dem letzten Schritt in der Glykolysekette, bei der die Pyruvatkinase die Übertragung energiereicher Phosphate von PEP auf ADP zur Bildung von ATP katalysiert. SWARTZ und KIM entdeckten, dass sich durch Integration des ersten Zwischenprodukts der Glykolyse, dem Glukose-6-Phosphat (G6P), anstelle von PEP oder Pyruvat, die ATP-Regeneration als effizient und günstiger gestaltet [115]. Sie vermuteten, dass somit prinzipiell jedes weitere Zwischenprodukt der Glykolyse als initiale Energiequelle zur ATP-Regeneration taugt, da alle bei der Glykolyse beteiligten Enzyme im Zellextrakt vorhanden zu sein schienen [115]. Tatsächlich wurden Glukose und Glukose-6-Phosphat [116], Fruktose-1,6-bisphosphat (FBP) [117] und 3-Phosphoglycerat (3-PGA) [118] als effiziente Energiequellen zur ATP-Regeneration nutzbar gemacht. Ebenso konnte mit Hilfe von Acetylphosphat (AP) und Acetat-Kinase (AK) [119] sowie Kreatinphosphat (KP) und Kreatinkinase (KK) [120] ATP regeneriert werden.

SWARTZ und KIM fanden weiter heraus, dass bis zu 70 % der eingesetzten PEP-Menge enzymatisch durch endogene Phosphatasen in Pyruvat und anorganisches Phosphat degradiert wurden [121]. Dadurch fehlte PEP nicht nur als Energiequelle, sondern es häufte sich auch anorganisches Phosphat an, das die Proteinsynthese inhibierte [121]. Zusätzlich beobachteten sie die Degradation der Aminosäuren Arg, Trp und Cys [114]. Durch die wiederholte Hinzugabe der drei oben genannten Aminosäuren

und PEP in Abständen von 20 min stieg die Proteinsynthesedauer von 20 min auf 80 min, wurde der zellfreien Reaktion zusätzlich Magnesiumacetat nach 60 min addiert, so verlängerte sich die Proteinsynthesedauer auf 2 h [114]. Allerdings stieg mit der höheren PEP-Konzentration auch die Konzentration an anorganischem Phosphat [114]. Dadurch, dass sich die Proteinsynthesedauer durch die Hinzugabe von Magnesiumacetat nochmals verlängern ließ, schlossen SWARTZ und KIM, dass Magnesium Komplexe mit anorganischem Phosphat bildet und dass die reduzierte Magnesiumkonzentration die Proteinsynthese limitiert [114]. Die Wichtigkeit von Magnesium für die ZFPS zeigte sich auch im System von KIM *et al.* [122]: Sank die Magnesiumkonzentration unter 12 mM, kam die ZFPS zum Erliegen, wurde die Konzentration stattdessen zwischen 15 mM-18 mM durch Hinzugabe von Magnesiumionen gehalten, verlängerte sich die Proteinsynthesedauer und -ausbeute. Dass die Anhäufung von anorganischem Phosphat die Konzentration an freiem Magnesium reduziert, wurde auch in diesem System bestätigt [122]. Im KIGAWA-System optimiert eine Magnesium-Konzentration zwischen 10 mM-14 mM die Proteinsynthese [123].

Eine weitere Strategie, die Anhäufung von anorganischem Phosphat zu vermeiden, lag darin, die *E. coli*-Zellen vor der Extrakterstellung in einem Medium wachsen zu lassen, das anorganisches Phosphat und Glukose in hoher Konzentration enthielt (2xYTPG Medium) [124]. Somit wurde die Expression endogener Phosphatasen während des Zellwachstums unterdrückt [124].

SWARTZ und KIM entwickelten einen mehrstufigen ATP-Regenerationsprozess, der anorganisches Phosphat recycelt und somit dessen Anhäufung reduziert [121]. Dem ZFPS-System wurde Pyruvat und das exogene Enzym Pyruvatoxidase (PO) hinzugegeben [121]. Anorganisches Phosphat wurde durch die PO unter Einbeziehung von O_2 und H_2O zu AP, CO_2 und H_2O_2 umgewandelt [121]. Im zweiten Schritt katalysierte die endogene, im Extrakt enthaltene AK die Übertragung von Phosphat auf ADP, und es entstanden ATP und Acetat [121]. Das dadurch anfallende anorganische Phosphat gelangte erneut in den Kreislauf [121]. Die endogene Katalase-Aktivität baute H_2O_2 ausreichend zu O_2 und H_2O ab [121]. Durch das mehrstufige ATP-Regenerationssystem hielt die Proteinsynthese 2 h an, wurden dem ZFPS-System regelmäßig Aminosäuren und Pyruvat hinzugegeben, verlängerte sich die Synthesedauer auf 6 h [121]. Obwohl Pyruvat wesentlich günstiger als PEP war, trieb zusätzlich benötigte PO die Kosten für dieses System in die Höhe [121]. Daher entwickelten SWARTZ und KIM dieses System weiter und griffen auf gleich mehrere endogene, im ZFPS-System enthaltene *E. coli*-Enzyme zur ATP-Regeneration zurück [115]. Zuerst katalysierte die Pyruvatdehydrogenase (PDH) mit den zur zellfreien Reaktion hinzugegebenen Kofaktoren NAD und Koenzym A (CoA) die Kondensation von Pyruvat und CoA zu Acetyl-CoA, das danach in AP durch die Phosphotransacetylase umgewandelt wurde [115]. Die Phosphorylierung von ADP durch die endogene AK verlief wie zuvor [121]. Wurde PEP anstelle von Pyruvat eingesetzt, ließ sich zusätzliches ATP durch PK-Aktivität regenerieren, und das Pyruvat wurde für den PDH-AK-Stoffwechselweg als Nebenprodukt bereitgestellt [115]. Der *E. coli*-PEP-Synthetase-Inhibitor Oxalat [125] wurde zur weiteren Pyruvatstabilisierung verwendet [115].

Im System von KIM *et al.* wurde zur ATP-Regeneration der KP- und Glukose-Metabolismus gekoppelt: Das durch den KP-Stoffwechsel anfallende anorganische Phosphat wird in den Glukose-Metabolismus gespeist, sodass kein inhibitorisches anorganisches Phosphat anfällt und beide Systeme parallel ATP regenerieren [126].

SITARAMAN *et al.* konnten durch den Einsatz von 3-PGA im Vergleich zu einer mit PEP versorgten zellfreien Reaktion die Proteinausbeute in etwa verdoppeln und die Proteinsynthesedauer ca. verdreifachen, ohne dass Kofaktoren oder weitere Reagenzien der zellfreien Reaktion hinzugegeben werden

mussten [118]. Es wurde vermutet [118], dass 3-PGA weniger anfällig für die enzymatische Degradation durch endogene Phosphatasen als beispielsweise die sonst verwendeten Energiequellen G6P, AP oder KP ist. Es wurde spekuliert [118], dass die Umsetzung von 3-PGA in PEP mit einer für die ATP-Regeneration angepassten Rate abläuft, sodass es sich nicht akkumuliert, sondern direkt nach Bildung zur ATP-Regeneration genutzt wird und daher weniger der Phosphatasen-Aktivität unterliegt [118]. Das 3-PGA-ATP-Regenerationssystem wurde neuerdings mit dem endogenen *E.-coli*-Maltose-Metabolismus gekoppelt, in dem anfallendes anorganisches Phosphat recycelt wird, um Maltose über zwei enzymatische Reaktionen in G6P umzuwandeln, das im Anschluss im Glykolyse-Metabolismus verstoffwechselt wird und ATP regeneriert [127].

Neben den hier vorgestellten ATP-Regenerationssystemen existieren weitere Systeme, wie beispielsweise das Cytomim-System, das versucht, die zellfreie Umgebung an möglichst reale zytoplasmatische *E.-coli*-Bedingungen anzupassen [128], [129]. Unter derartigen Bedingungen ist sogar die oxidative Phosphorylierung möglich [130].

2.4.3.1.2. Aminosäurenstabilisierung

Die Degradierung der Aminosäuren Arg, Trp, Cys und Ser in ZFPS-Systemen und die damit verbundene Reduzierung der Proteinsynthese-Effizienz stellte ein Problem in der zellfreien Proteinsynthese dar [114], [129]. Die anfängliche Strategie, die Aminosäuredegradation zu kompensieren, bestand in der Erhöhung der initialen Aminosäurenkonzentration [115], [129] oder der Aminosäuren-Supplementierung der Reaktion während der ZFPS [114], [115], [129].

Weiterführende Strategien zielten darauf ab, die degradierenden Enzyme aus dem rohen Zellextrakt zu entfernen [131] oder sie bei der ZFPS zu inhibieren [124]. Die durch die Glutamacysteinligase hervorgerufene Cys-Degradation, bei der Cys und Glutamat zu γ -Glutamylcystein ligiert werden, konnte mit dem Inhibitor L-Buthioninsulfoximin reduziert werden [124]. Wurde der Extrakt aus dem mutierten *E.-coli*-Stamm KC1 gewonnen, in dessen Genom die Gene, die für die Arg-Decarboxylase (*speA*), Tryptophanase (*tnaA*) und Serin-Deaminasen (*sdaA* und *sdaB*) kodieren, deletiert wurden, konnten Arg, Trp und Ser stabilisiert werden [131]. Später wurde im *E.-coli*-Stamm KC1 außerdem das Gen, das für die Glutamacysteinligase kodiert, deletiert [132]. Wurde der rohe Zellextrakt aus dem daraus resultierenden Stamm KC6 erzeugt, konnte zusätzlich Cys stabilisiert werden [132].

2.4.3.1.3. Stabilisierung von linearer DNA

Der *E.-coli*-Enzym-Komplex ReBCD degradiert lineare DNA [133]. Die verkürzte Version GamS des λ -Phagen Proteins Gam inhibiert die *E.-coli*-RecBCD-Nuklease-Aktivität, indem es den Komplex daran hindert, an die Enden des DNA-Doppelstrangs zu binden [134]. Durch die enzymatische Aktivität von GamS in einem zellfreien Genexpressionssystem konnte lineare Template-DNA länger stabilisiert und somit die darauf befindlichen Gene effizienter exprimiert werden [118]. Außerdem wurde ein Extrakt aus einem mutierten *E.-coli*-Stamm erzeugt, in dem das Gen, das für die Endonuklease I (*endA*) kodiert, entfernt und das *recBCD*-Operon durch das Rekombinationssystem des λ -Phagens ersetzt wurde [135]. In diesem ZFPS-System konnte das Gen, das für ein Modellprotein kodiert, von linearer DNA in vergleichbar hoher Ausbeute wie von Plasmid-DNA exprimiert werden [135].

2.4.3.2. Extrakt-basierte ZFPS, die Stoffaustausch mit der Umgebung ermöglichen

Ende der 1980er Jahre verfolgten SPIRIN *et al.* eine zum Batch-Modus alternative Strategie, die ZFPS effizienter zu gestalten [136], [137]. Es wurde eine Extrakt-basierte zellfreie Reaktion kontinuierlich

mit einer Nährlösung bestehend aus Aminosäuren, ATP und GTP [136] oder Aminosäuren, sämtlichen NTPs und PEP [137] versorgt. Während die Reaktionsprodukte, insbesondere die synthetisierten Proteine, aus der Kammer abgeführt wurden, wurde die molekulare Proteinsynthese-Maschinerie über eine Membran in der Reaktionskammer zurückgehalten [136], [137]. Die Synthesedauer konnte durchschnittlich von 1 h bei einer vergleichbaren Batch-Reaktion auf bis zu 20 h-40 h verlängert werden, und auch die Proteinausbeute stieg deutlich an [136].

Bei einem passiveren Verfahren wurde die zellfreie Reaktion gegen eine Nährlösung dialysiert [138]. Hier hielt die Dialysemembran die synthetisierten Proteine zurück, und es wurden nur kleine Moleküle zwischen der Reaktion und der Nährlösung per Diffusion ausgetauscht [138]. Auf diesem Weg konnte die Synthesedauer auf 14 h bei einer Proteinausbeute von 1,2 mg/ml verlängert werden [138]. Es existieren weitere ähnliche ZFPS-Verfahren, die ebenfalls auf Stoffaustausch zwischen der zellfreien Reaktion und der Umgebung aufbauen, beispielsweise in [120], wo über eine Dauer von 21 h 6 mg/ml eines Proteins synthetisiert werden konnten.

2.4.3.3. PURE-System

Ein ZFPS-System, das nicht auf einem rohen Zellextrakt basiert, ist das sogenannte *protein synthesis using recombinant elements* (PURE)-System, das 2001 von UEDA *et al.* entwickelt wurde [139], [140]. Es setzt sich aus der minimalistischen *E.-coli*-Proteinsynthese-Maschinerie zusammen, die zuvor *in vivo* überexprimiert und mittels His-Tag-Purifikation extrahiert und aufgereinigt wurde (vgl. Abschn. 4.2.16) [139]. Der erste Versuch, ein gekoppeltes ZFPS-System aus der aufgereinigten Transkriptions-Translations-Maschinerie zu rekonstruieren, geht auf KUNG *et al.* zurück [141].

Obwohl das PURE-System im Vergleich zu Extrakt-basierten Systemen frei von inhibitorischen Substanzen wie beispielsweise Nukleasen, Proteasen oder Phosphatasen ist [139], müssen neben der Isolation der Ribosomen mindestens zehn Translationsfaktoren, die T7-RNA-Polymerase, 20 aaRS, die Kreatinkinase, die Myokinase und die Nukleosid-Diphosphatase *in vivo* überexprimiert und aufgereinigt werden [140]. Insofern ist die ZFPS mit dem PURE-System zeitaufwendig und teuer [109].

2.4.3.4. Verwendetes zellfreies Expressions-System

Wie bereits in Abschn. 2.4.1 erläutert, stellte die Integration des T7-Transkriptionssystems einen Durchbruch in der ZFPS dar, da wesentlich höhere Proteinausbeuten im Vergleich zum endogenen *E.-coli*-Transkriptionssystem erreicht werden konnten [105]. Auf diesem Erfolg aufbauend, wurden ZFPS-Systeme, die mittels eines Bakteriophagen-Transkriptions-Systems DNA exprimieren, zwecks längerer Synthesedauer und höheren Proteinausbeuten weiter optimiert, u.a. in [112]–[114], [116], [118], [121]–[125], [130], [132], [138], [142]–[144] (vgl. Abschn. 2.4.3.1). Somit hat sich die Bakteriophagen-Transkription in der ZFPS heute als Standard etabliert [61], [106]. Solche Systeme sind demnach zwar in Bezug auf die Proteinausbeute effizient, aber kaum in der Lage, *E.-coli*-endogene Regulationsmechanismen zu studieren [52].

Einen alternativen Weg haben NOIREAUX *et al.* eingeschlagen, indem sie ein Extrakt-basiertes ZFPS-System entwickelten, in dem die endogene *E.-coli*-Transkriptions-Translations-Maschinerie ausreicht,

um ähnlich hohe Proteinausbeuten wie in ZFPS-Systemen zu generieren, die auf Bakteriophagen-Transkription zurückgreifen [51], [52]. Mit Hilfe dieses Systems lassen sich somit endogene *E.-coli*-Gen-Regulations-Mechanismen und die Expression von endogenen *E.-coli*-Genen in zellfreier Umgebung studieren [52].

2.5. Fluoreszierende Proteine der GFP-Familie

Das grün fluoreszierende Protein (GFP), das 1962 von SHIMOMURA *et al.* in der Qualle *Aequorea victoria* entdeckt wurde [145], oder eines seiner Varianten werden häufig zur Analyse der Organisation lebendiger Systeme verwendet [146]. Mit Hilfe von fluoreszierenden Proteinen lässt sich beispielsweise die Aktivität eines Promotors charakterisieren [146]. Die fluoreszierenden Proteine GFP, EGFP, deGFP und mCherry wurden in der vorliegenden Dissertation als Reportermoleküle oder Modellproteine verwendet und werden im Folgenden kurz vorgestellt.

Die posttranslationale autokatalytische GFP-Fluorophor-Bildung ist das Resultat der Zyklisierung der drei Aminosäuren Ser, Tyr und Gly an den Positionen 65, 66, und 67 und der anschließenden Dehydrierung (C_{α} - C_{β} bei Tyr66) mit molekularem Sauerstoff, wodurch sich die Masse des Proteins um 20 Da verringert [146]. Dadurch, dass das Fluorophor sowohl als protoniert (neutral), als auch unprotoniert (anionisch) vorliegen kann, lässt sich GFP mit Licht einer Wellenlänge von ca. 400 nm (protoniert) und 480 nm (unprotoniert) anregen [146]. Die angeregte anionische Form emittiert grünes Licht mit einer Wellenlänge von etwa 510 nm; wird GFP hingegen in protoniertem Zustand angeregt, wird es meist in die anionische Form überführt und emittiert wiederum Licht einer Wellenlänge von ca. 510 nm [146]. GFP besteht aus β -Faltblättern, die über Schleifen miteinander verbunden sind und eine Fass-Struktur bilden [147].

Beim *enhanced GFP* (EGFP, Phe64Leu [148] und Ser65Thr [149]) ist der molare Extinktionskoeffizient bei einer Anregungswellenlänge von 488 nm in etwa sechs Mal höher als der von GFP, während sich die Quantenausbeute von EGFP im Vergleich zum Wild-Typen von 79 % auf 60 % reduziert [150]. Beim EGFP wird wegen der kleinen Seitenketten der zweiten Aminosäure Val das N-Terminale Met enzymatisch entfernt (vgl. Abschn. 2.2.2.4), sodass die Nummerierung der EGFP-Sequenz mit Val beginnt. EGFP-Del6-229 (deGFP) ist eine verkürzte Version von EGFP. Das *degfp*-Gen geht auf Publikation [51] zurück, in der auf den Arbeiten von [151] aufgebaut wird. Bei deGFP befindet sich das nicht entfernte Start-Met an der fünften Position in Bezug auf EGFP, und alle Aminosäuren hinter der Position 229 fehlen. Somit stimmen die Proteinsequenzen zwischen den Aminosäuren an den Positionen 6-229 überein. Die Position N149 (bzw. N150, sofern das abgespaltene Met mitgezählt wird) an der EGFP-Oberfläche im Übergang eines β -Faltblatts zu einer Schleife, hat sich als geeignete Position zur Platzierung eines Stopcodons erwiesen, da der Einbau einer nicht-kanonischen Aminosäure dort keinen Einfluss auf die globale Proteinstruktur zeigt [152]–[154]. Insofern eignet sich EGFP^{N149amber} als Reporter-Molekül in Stopcodon-Experimenten, da sich das Fluorophor nur bei erfolgter Stopcodon-Suppression bildet und in diesem Fall das Protein fluoresziert [152]–[154].

Die GFP-ähnlichen Chromophore von rot fluoreszierenden Proteinen des Typs DsRed entstehen durch einen zusätzlichen Oxidationsschritt (C_{α} -N-Bindung an Position 65) mit molekularem Sauerstoff [146]. Das GFP-ähnliche DsRed stammt aus der Scheibenanemone *Discosoma sp.* und geht auf [155] zurück. Es konnten optimierte DsRed-Varianten erzeugt werden [156]. Durch auf den Arbeiten von [156] aufbauenden Modifizierungen konnte mCherry erzeugt werden [157]. Es besitzt ein Anregungsmaximum

bei einer Wellenlänge von 587 nm, das Emissionsmaximum liegt bei einer Wellenlänge von 610 nm [157].

3. Evolution wachsender, sich reproduzierender Polymere

Teile des Kap. 3 wurden vorab in [158] mit den Koautoren Philipp Zimmer, Eva Wollrab sowie den Projekt-betreuenden Professoren Karsten Kruse und Albrecht Ott veröffentlicht. Die initiale Idee für diesen Teil der Dissertation wurde im Rahmen des übergeordneten Projekts „Origin of Life“ des Lehrstuhls für Biologische Experimentalphysik von Albrecht Ott, Karsten Kruse und Eva Wollrab entwickelt und von Philipp Zimmer theoretisch [159] sowie in dieser Dissertation experimentell aufgegriffen. Die in Kap. 3 vorgestellte Studie wurde bereits während der unveröffentlichten Diplomarbeit (*Enzymbasierte Selbstreplikation von DNA Fragmenten*, Universität des Saarlandes, Lehrstuhl für Biologische Experimentalphysik, 2012) begonnen und im Verlauf der Promotion abgeschlossen. Die wesentlichen, in der Diplomarbeit erzielten Ergebnisse, werden als Motivation der fortführenden Experimente während der Promotionsphase zu Beginn des Ergebnisteils (Abschn. 3.3) präsentiert.

3.1. Einführung, Stand der Wissenschaft, Strategie und Zielsetzung

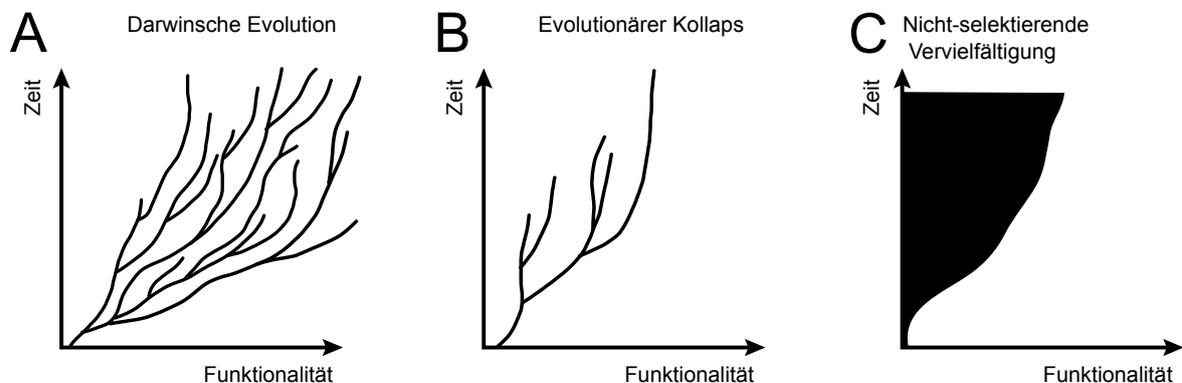


Abbildung 3.1.: Schematische Darstellung möglicher evolutionärer Szenarien anhand des Lebensbaums. A: Darwinsche Evolution, die die Koexistenz mehrerer Spezies zu jedem Zeitpunkt ermöglicht. Hier evolvieren die Spezies als Ensemble in Richtung höherer Funktionalität. B: Evolutionskollaps, bei dem eine Spezies alle anderen verdrängt und die Evolution stoppt. C: Nicht-selektierende Vervielfältigung durch die zwar ständig neue und komplexere Individuen entstehen, diese aber keine abgrenzbaren Spezies bilden. Abb. 3.1 modifiziert aus [158] übernommen.

Das Leben auf der Erde ist eng mit der Darwinschen Evolution [1] verflochten, die das irdische Ökosystem von einer präbiotischen Ursuppe [2]–[4] in Richtung immer komplexer werdender, sich reproduzierender Individuen durch Mutation und Selektion vorantrieb. Das frühe Leben ähnelte einem promiskuitiven Zustand, indem es zwischen den Individuen vermutlich zum regen Austausch von Molekülen kam [5]–[10]. Dann erst formierten sich allmählich diese Individuen aufgrund von komplexen

dynamischen Wechselwirkungen zu unterscheidbaren Untergruppen, die als Spezies klassifiziert werden konnten. Dieser Organisationsprozess führte zu einer gewissen Ordnung innerhalb der Biosphäre (Abb. 3.1A). Daher ist biologische Vielfalt nicht mit der bloßen Entstehung aller denkbaren, sich selbst reproduzierenden Individuen (Abb. 3.1C) gleichzusetzen. Viel eher formten sich fortwährend ökologische Nischen, durch die neue Spezies entstanden, die ihrerseits wiederum als neue Nischen für die weitere Evolution dienten [11], [12]. Sogar ohne externen Selektionsdruck [26]–[28] evolviert das Ökosystem in diesem Sinne als eng verflochtenes Kollektiv zu höherer Funktionalität [29], [30]. Bisher sind diese, der Evolution zugrundeliegenden dynamischen Kopplungen, wenig verstanden.

Abstrakte [13], [14] oder molekulare Systeme [15], [16], in denen Spezies üblicherweise durch DNA oder RNA verkörpert werden, eignen sich zur Analyse grundlegender evolutionärer Prozesse. In derartigen Systemen führen typischerweise Mutationen der Nukleotidsequenzen zur evolutionären Variation, und die molekulare Reproduktion der Nukleotidsequenzen wird mittels eines Template-basierten Prozesses realisiert [14]–[16], [160]–[168]. Durch Energiezufuhr oder zyklisches Ändern der physikalischen Parameter, beispielsweise der Temperatur oder dem Wechsel zwischen Feucht- und Trockenperioden, wird ein solcher molekularer Reproduktionsprozess ermöglicht [169], [170]. Die Template-basierte Reproduktion läuft typischerweise wie folgt ab: Zwei komplementäre Nukleotide hybridisieren zuerst auf dem Template, werden danach kovalent verbunden [171], [172] und lösen sich im letzten Schritt als neu entstandenes Nukleotid vom Template ab, das sich somit kreuzkatalytisch reproduziert, ein Mechanismus, der zum exponentiellen Wachstum der Template-Konzentration führen kann [173]. Die Ligase-Kettenreaktion (LCR) [174], [175] ist eine hoch effiziente Variante dieses Reproduktionsprozesses, bei der zwei DNA-Stränge auf einem komplementären Template hybridisieren und durch das Enzym Taq-DNA-Ligase kovalent miteinander verbunden werden.

In molekularen Evolutionssystemen treten Variationen üblicherweise während der Reproduktion auf, wenn diese beispielsweise unvollständig abläuft [17], [18], falsche Nukleotide eingesetzt werden [176] oder es zu Fehlpaarungen bei der Hybridisierung kommt [177]. Ist das Verhältnis der Variations- zur Reproduktionsrate klein genug, können durch externen Selektionsdruck Spezies selektiert werden [19]. Üblicherweise treibt ein solcher konstanter Druck molekulare evolutionäre Modellsysteme zur Selektion einer optimalen Evolutionslösung, die mit dem Stoppen der weiteren Evolution verbunden ist (Abb. 3.1B) [17]–[25]. Diese Eigenschaft machen sich Verfahren der Molekularen Evolution wie SELEX gezielt zu Nutze, um bestmögliche Liganden zu selektieren [178]–[180]. Hingegen läuft die Evolution in natürlichen Ökosystemen bei gleichzeitiger Wahrung der Koexistenz der Spezies permanent weiter, sogar bei fehlendem externen Druck [26]–[28]. Die Umsetzung einer andauernden Evolution koexistierender Spezies in molekularen Evolutionssystemen bleibt bis dato ungelöst. Ein Lösungsansatz besteht in der Verlängerung von Nukleotidsträngen als grundlegender Prozess nicht nur, um die molekulare Funktionalität zu erhöhen, sondern auch der Evolution ein offenes Ende zu lassen [181].

In dieser Studie wurden mögliche Bedingungen erforscht, die ein unbegrenztes Andauern molekularer Evolution unter gleichbleibenden äußeren Bedingungen ermöglichen. Zu diesem Zweck wurde ein System betrachtet, in dem molekulare Spezies durch Polymere mit unterschiedlichen Sequenzen repräsentiert werden, deren Länge und nicht die Sequenz die Spezies-definierende Eigenschaft darstellt. Aus diesem Grund entstehen Variationen durch Veränderung der Polymerlänge einer existierenden Spezies. Experimentell wurden diese linearen Polymere durch lineare DNA-Moleküle verkörpert, die sich mittels LCR reproduzieren. Variationen wurden durch spontane Ligationen realisiert.

3.2. Material und Methoden

Die Methoden zur Herstellung (Abschn. 3.2.1), Inkubation, Entnahme und Manipulation (Abschn. 3.2.2) der Proben sowie zur gelelektrophoretischen Trennung der DNA-Fragmente (Abschn. 3.2.3) wurden während der Diplomarbeit etabliert. Die initiale Idee, die Menge der gelelektrophoretisch getrennten, fluoreszierenden Proben-DNA mit Hilfe einer Kalibrierungsmessung der Fluoreszenz-Intensität einer Verdünnungsreihe eines DNA-Standards und dem Bildverarbeitungsprogramm ImageJ [182], [183] zu quantifizieren (Abschn. 3.2.3), wurde ebenfalls während der Diplomarbeit entwickelt.

3.2.1. Herstellung der Proben und deren initiale Zusammensetzung

Alle verwendeten DNA-Einzelstränge wurden von *Metabion international AG* (Martinsried) synthetisiert, HPLC-purifiziert und sind am 5'-Ende phosphoryliert, sodass die Taq-DNA-Ligase (*New England Biolabs GmbH*, Frankfurt am Main) die Einzelstränge ligieren konnte. Die zugehörigen DNA-Sequenzen sind in Anh. A.1.1 zu finden.

Für jedes Experiment wurde eine Mastermix-Lösung in verschließbaren DNase- und RNase-freien 0,2 ml-PCR-Reaktionsgefäßen (*Biozym Scientific GmbH*, Hessisch Oldendorf) hergestellt, in der 1x Taq-DNA-Ligase-Reaktionspufferbedingungen (*New England Biolabs GmbH*, Frankfurt am Main) herrschten. Alle Bestandteile der Mastermix-Lösung wurden durch Auf- und Abpipettieren und anschließendes kurzes Vortexen (*Vortex 3, IKA®-Werke GmbH*, Staufen) gemischt. Im Anschluss wurde die Mastermix-Lösung zu jeweils gleichen Volumina auf neue PCR Reaktionsgefäße verteilt. Jede gesplittete Reaktion wird im Folgenden als Probe bezeichnet. Alle Schritte wurden auf Eis ausgeführt, um die enzymatische Aktivität der Taq-DNA-Ligase zu unterdrücken bis die Proben den Temperaturzyklen ausgesetzt werden.

Im Experiment, dessen Ergebnisse in Abb. 3.4A präsentiert werden, betrug die initiale Konzentration der beiden komplementären 20 b langen Einzelstränge Σ und Σ^c jeweils 500 nM, die Taq-DNA Ligase war 1,7 units/ μ l konzentriert. Im Experiment, dessen Ergebnisse Abb. 3.4B zeigt, wurde die Konzentration dieser Stränge jeweils halbiert, und die Taq-DNA-Ligase-Konzentration erhöhte sich auf 8 units/ μ l. In diesen beiden Experimenten betrug das Probenvolumen jeweils 10 μ l.

Im Experiment, dessen Ergebnisse in Abschn. 3.3.2 präsentiert und diskutiert werden, betrug die initiale Konzentration der beiden komplementären 10 b langen Einzelstränge σ und σ^c jeweils 1000 nM, die Taq-DNA-Ligase befand sich mit einer Konzentration von 4 units/ μ l in den Proben, die ein Volumen von 10 μ l aufwiesen.

Im Experiment, dessen Ergebnisse in Abschn. 3.3.3 präsentiert und diskutiert werden, betrug die initiale Konzentration der beiden komplementären 10 b langen Einzelstränge σ und σ^c jeweils 200 nM. Die Taq-DNA-Ligase besaß eine Konzentration von 4 units/ μ l in den Proben. Das Probenvolumen belief sich auf 10 μ l.

Im Experiment, dessen Ergebnisse in Abschn. 3.3.5 präsentiert und diskutiert werden, war jeder der insgesamt 16 10 b langen und paarweise komplementären DNA-Einzelstränge 17,5 nM konzentriert. Die Konzentration der acht 20 b langen und paarweise komplementären DNA-Einzelstränge betrug jeweils 3,75 nM, die Konzentration der Taq-DNA-Ligase belief sich auf 0,8 units/ μ l. Das Volumen der Proben umfasste 20 μ l.

3.2.2. Zyklische Variation der Temperatur, Probenentnahme und -manipulation

Die Proben wurden mit dem *MultiGene™ 96-Well Gradient Thermal Cycler* (Labnet International Inc., Edison, USA) mit zyklisch variierenden Temperaturen inkubiert. Die Temperaturen wurden an die Schmelztemperaturen der verwendeten DNA-Stränge angepasst, die Schmelztemperaturen mit Hilfe von DINAMelt berechnet [184]–[186].

Im Experiment, dessen Ergebnisse Abb. 3.4A präsentiert, wurden die Proben mit dem Inkubationsprogramm, das Tab. 3.1 zeigt, zyklisch inkubiert. Initial wurde die DNA in allen Proben für 1,5 min bei einer Temperatur von 84 °C denaturiert. Zu Beginn wurde je eine Probe nach zwei, vier, sechs, acht, zehn und zwölf Zyklen entnommen, danach erfolgte nach jedem Zyklus eine weitere Probenentnahme.

Temperatur [°C]	60	66	60	84
Zeit [min]	15	1	15	1

Tabelle 3.1.: Temperatur-Zeit-Profil eines Zyklus des Experiments, dessen Ergebnisse in Abb. 3.4A illustriert werden.

Im Experiment, dessen Ergebnisse Abb. 3.4B zeigt, wurden die Proben nach initialer Denaturierung aller DNA-Stränge bei einer Temperatur von 84 °C und einer Dauer von 1,5 min ebenfalls mit dem Inkubationsprogramm, das Tab. 3.1 zeigt, für fünf Zyklen inkubiert. Darauf folgte eine einzelne fünfminütige Inkubationsphase bei einer Temperatur von 30 °C. Ab dem sechsten Zyklus wurde das Inkubationsprogramm verwendet, das in Tab. 3.2 zusammengefasst wird. Es wurde vor Beginn eines jeden dieser Zyklen 1 unit des Enzyms Mungbohnen-Nuklease (*Mung Bean Nuclease*, New England Biolabs GmbH, Frankfurt am Main) in die Proben gegeben. Das Enzym wurde während des Experiments unter 1x Mungbohnen-Reaktionspufferbedingungen (*New England Biolabs GmbH*, Frankfurt am Main) bei 4 °C gelagert.

Temperatur [°C]	30	60	66	60	84	30
Zeit [min]	30	15	1	15	1	5

Tabelle 3.2.: Temperatur-Zeit-Profil eines Zyklus, der im Experiment, dessen Ergebnisse in Abb. 3.4B präsentiert werden, das in Tab. 3.1 illustrierte Temperatur-Zeit-Profil der ersten fünf Zyklen, wie im Text beschrieben, ablöste.

Im Experiment, dessen Ergebnisse in Abschn. 3.3.2 präsentiert und diskutiert werden, wurde nach initialer 1,5-minütiger DNA-Denaturierung bei 60 °C in zyklischer Weise das Temperatur-Zeit-Profil aus Tab. 3.3 verwendet. Die Restriktionsenzyme AfeI und BstUI (*New England Biolabs GmbH*, Frankfurt am Main) wurden nach acht (je 1 unit), zwölf (je 1 unit) und 20 Zyklen (je 0,75 units) in die Proben gegeben. Da AfeI nicht thermostabil ist, wurden zusätzlich je 0,5 units AfeI nach 15 und 18 Zyklen in die Proben pipettiert. Aus dem gleichen Grund wurden ab dem 22. Zyklus je 0,75 units von AfeI alle zwei Zyklen in die Proben bis zum Ende des Experiments gegeben. 0,75 units von BstUI wurden den Proben nach 32 Zyklen hinzugefügt. Die beiden Restriktionsenzyme wurden während des Experiments unter 1x Reaktionspuffer-Bedingungen (*New England Biolabs GmbH*, Frankfurt am Main) bei 4 °C gelagert. Nach 20 Zyklen wurden die Temperaturen des Inkubationsprogrammes von 34 °C auf 60 °C bzw. von 60 °C auf 76 °C nach einem einzelnen 1,5-minütigen DNA-Denaturierungsschritt bei 76 °C erhöht. Die Proben wurden nach 12, 20, 28 und 36 Zyklen entnommen.

Im Experiment, dessen Ergebnisse in Abschn. 3.3.3 präsentiert und diskutiert werden, wurde nach initialer 1,5-minütiger DNA-Denaturierung bei 60 °C in zyklischer Weise ebenfalls das Temperatur-

Temperatur [°C]	34	60
Zeit [min]	13,5	1

Tabelle 3.3.: Temperatur-Zeit-Profil eines Zyklus, der zur Proben-Inkubation während der ersten 20 Zyklen im Experiment verwendet wurde, das in Abschn. 3.3.2 präsentiert und diskutiert wird. Nach 20 Zyklen wurden die Temperaturen von 34 °C auf 60 °C bzw. von 60 °C auf 76 °C nach einem einzelnen 1,5-minütigen DNA-Denaturierungsschritt bei 76 °C erhöht.

Zeit-Profil aus Tab. 3.3 benutzt. Die Temperaturen des Inkubationsprogrammes wurden in gleicher Weise wie im in Abschn. 3.3.2 besprochenen Experiment erhöht, nur dass die Erhöhung zwei Zyklen später, also erst nach 22 Zyklen, erfolgte. Die Restriktionsenzyme AfeI und BstUI wurden nach acht (je 1 unit), zwölf (je 1 unit), 15 (je 0,5 units), 18 Zyklen (je 0,5 units) und 22 Zyklen (je 0,5 units) in die Proben gegeben. Jeweils 0,25 units der Mungbohnen-Nuklease wurden außerdem nach 15 und 18 Zyklen hinzugegeben.

Im Experiment, präsentiert und diskutiert in Abschn. 3.3.5, wurde nach initialer einminütiger DNA-Denaturierung bei 80 °C in zyklischer Weise das Temperatur-Zeit-Profil aus Tab. 3.4 angewendet. Nach 24 Zyklen wurden 80 fmol der 40 bp-Verknüpfer in die Proben gegeben. Nach 29 Zyklen wurden 40 fmol der 80 bp-Verknüpfer in die Proben pipettiert. Die Verknüpfer wurden während des Experiments unter Reaktionspufferbedingungen gelagert. Jeweils eine Probe wurde nach 6, 12, 18 und 24 Zyklen entnommen. Anschließend erfolgte die Probenentnahme nach jedem Zyklus.

Temperatur [°C]	28 / 55 / 74	80 / 90 / 94
Zeit [min]	2	1

Tabelle 3.4.: Temperatur-Zeit-Profil der Zyklen, die zur Proben-Inkubation im Experiment verwendet wurden, das in Abschn. 3.3.5 präsentiert und diskutiert wird. Der erste Wechsel auf die jeweils höheren Temperaturen erfolgte nach 24 Zyklen (von 28 °C auf 55 °C bzw. von 80 °C auf 90 °C). Der letzte Wechsel fand nach 29 Zyklen statt (von 55 °C auf 74 °C bzw. von 90 °C auf 94 °C).

Während der Hinzugabe der verschiedenen Enzyme oder DNA-Moleküle wurde das Inkubationsprogramm pausiert, die Probenmanipulation erfolgte auf Eis. Nach Abschluss der Molekül-Hinzugabe wurden die Proben zuerst kurz (2s-3s) auf mittlerer Drehzahl gevortext (*Vortex 3, IKA®-Werke GmbH*, Staufen) und dann etwa 2s-3s abzentrifugiert (*Sprout Minizentrifuge, Biozym Scientific GmbH*, Hesisch Oldendorf). Danach wurde das Inkubationsprogramm wie angegeben fortgesetzt.

3.2.3. Gelelektrophoretische Trennung und Quantifizierung der DNA-Menge

Sofort nach ihrer Entnahme wurde den Proben 6x konzentrierter Probenpuffer bzw. Ladepuffer (*Fisher Scientific GmbH*, Schwerte) hinzupipettiert, um die weitere Enzym-Aktivität zu inhibieren und um die Reaktionen gelelektrophoretisch trennen zu können. Dieser Probenpuffer setzt sich wie folgt zusammen: 10 mM Tris-HCl (pH 7.6), 0,03 % Bromphenolblau, 0,03 % Xylencyanol FF, 60 % Glycerin und 60 mM EDTA. Die Proben wurden bis zur Gelelektrophorese bei -20 °C gelagert. Das hinzupipettierte Volumen des 6x Probenpuffers betrug stets ein Sechstel des Gesamtvolumens aus Probe und Probenpuffer. Die doppelsträngigen DNA-Moleküle wurden mittels nativer, vertikaler Polyacrylamid-Gelelektrophorese (PAGE) unter einfachen Tris-Borat-EDTA-Pufferbedingungen (89 mM Tris, 89 mM Borsäure, 2 mM EDTA bei pH 8,3, autoklaviert (Varioklav® Dampfsterilisator, Typ 75 S, *H+P Labortechnik AG*, Oberschleißheim)) getrennt. Es wurden 15-20 %-ige Polyacrylamidgele verwendet. Die PAGE lief gekühlt ab (≈ 4 °C), indem der Tris-Borat-EDTA-Puffer (TBE-Puffer) vorgekühlt wurde

3. Evolution wachsender, sich reproduzierender Polymere

und die vertikale Gelelektrophoresekammer während der Gelelektrophorese mit Eis umhüllt war. Typischerweise dauerte die PAGE zwischen 60 min-90 min bei einer angelegten Spannung von 140 V. Die getrennte DNA wurde mittels Ethidiumbromid (EthBr) gefärbt (1x TBE-Puffer mit $\approx 1 \mu\text{g/ml}$). Die fluoreszierende EthBr-gefärbte DNA wurde mit dem *GelVue UV Transilluminator* sichtbar gemacht (*Syngene Europe*, Cambridge, UK).

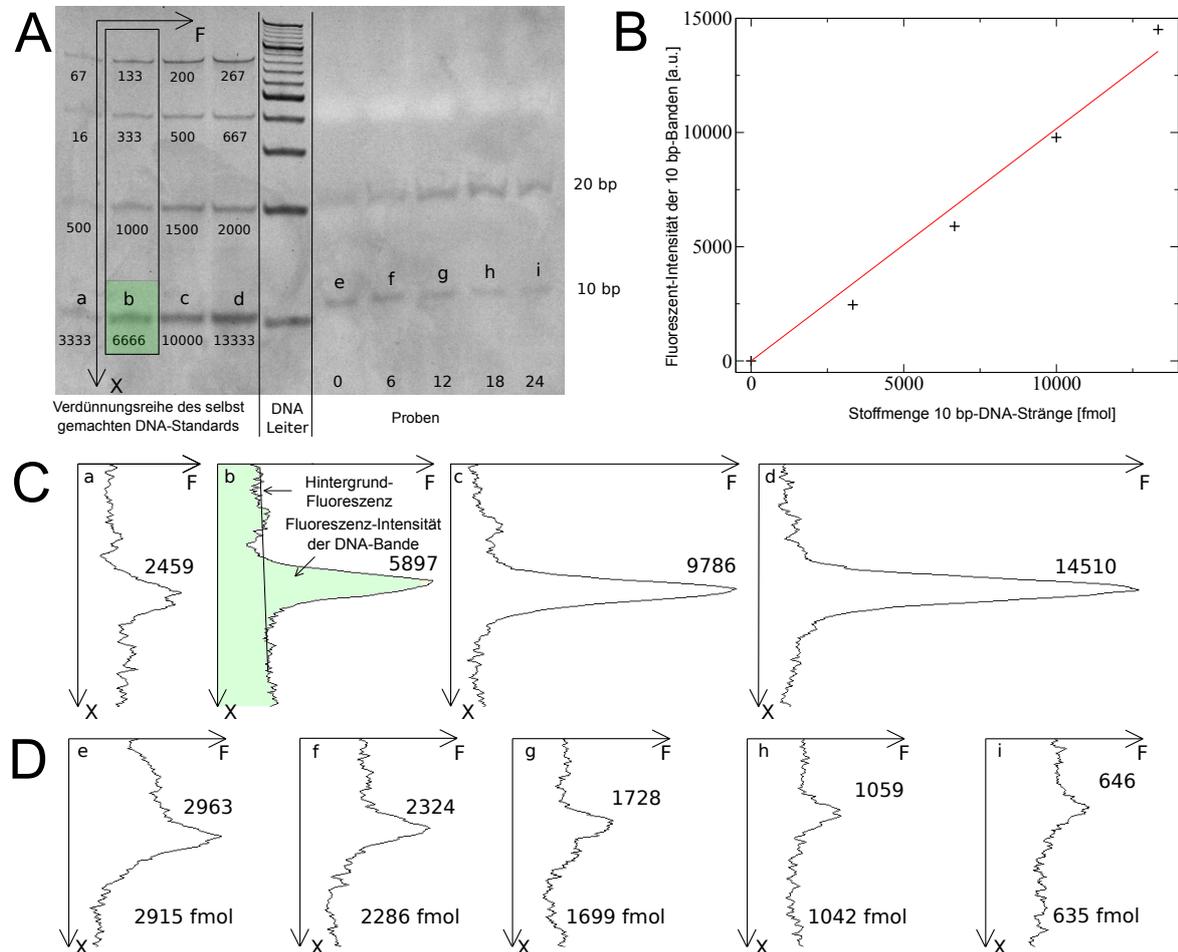


Abbildung 3.2.: Illustration der Quantifizierung der DNA-Stoffmenge mit ImageJ [182], [183]. A: Beispiel eines Polyacrylamidgels. Von links ausgehend zeigen die ersten vier Gelbahnen die aufgetrennten 10 bp, 20 bp, 40 bp und 80 bp langen DNA-Stränge des eigens hergestellten DNA-Standards (Sequenzen in Anh. A.1.1). Die Zahlen entsprechen der Stoffmenge in fmol der sich in den DNA-Banden befindlichen DNA-Stränge. In der fünften Bahn wurden die 15 DNA-Stränge mit Längen von jeweils 10 bp, 20 bp, ..., 150 bp der *O'RangeRuler-10 bp-DNA-Leiter* (*O'RangeRuler 10 bp DNA Ladder*, verwendungsbereit, *Thermo Fisher Scientific*, Waltham, USA) getrennt. In den restlichen Bahnen wurden die DNA-Stränge der Proben getrennt. Die Zahlen geben die Zyklen an, nach denen die jeweiligen Proben entnommen wurden. B: Kalibrierungsgerade, die aus den 10 bp-Strängen der Verdünnungsreihe des DNA-Standards entstand. Die Gerade ist ein Fit mit der Steigung $1,017 \text{ fmol}^{-1} \pm 0,045 \text{ fmol}^{-1}$. C und D: Ausschnitte aus den Plots der Fluoreszenz-Intensität um die 10 bp-Stränge der Verdünnungsreihe des DNA-Standards und die 10 bp-Stränge der Proben herum. Die Buchstaben ordnen den DNA-Banden (A) ihre Plots (C, D) zu. Abb. 3.2 ist modifiziert aus [158] übernommen.

Neben den Proben wurde auf jedem Gel die DNA der *O'RangeRuler 10 bp DNA Ladder* (verwendungsbereit, *Thermo Fisher Scientific*, Waltham, USA) bei jeder Gelelektrophorese getrennt. Dieser DNA-Standard besitzt 15 DNA-Stränge mit Längen von jeweils 10 bp, 20 bp, ..., 150 bp. Im Experiment, dessen Ergebnisse Abb. 3.4A präsentiert, wurden zur Quantifizierung der Proben-DNA-Stränge auf jedem

Gel die DNA-Stränge einer Verdünnungsreihe der *GeneRuler-Low-Range-DNA-Quantifizierungsleiter* (*GeneRuler Low Range DNA Ladder* verwendungsbereit, 0,1 µg/µl, *Thermo Fisher Scientific*, Waltham, USA) gelelektrophoretisch getrennt. Im Experiment, mit dem sich Abschn. 3.3.5 beschäftigt, wurde ein eigens hergestellter DNA-Standard zur Quantifizierung der Proben-DNA-Stränge benutzt, der aus den in diesem Experiment verwendeten 10 bp, 20 bp, 40 bp und 80 bp langen DNA-Strängen (Sequenzen in Anh. A.1.1) bestand.

Auf jedem Gel wurden die DNA-Stränge einer Verdünnungsreihe dieses Standards mittels des Polyacrylamidgels aufgetrennt (Abb. 3.2A). Die Gel-Bilder wurden mit dem Bildverarbeitungsprogramm ImageJ [182], [183] analysiert, um die DNA-Stoffmenge zu quantifizieren. Dazu mussten die Gelbilder in Graustufen-Bilder umgewandelt werden, wenn nötig, wurde der Kontrast und die Helligkeit des gesamten Bildes angepasst. Diese Standard-Methode zur Quantifizierung der Stoffmenge von auf Gelen getrennten Biomolekülen wird exemplarisch anhand von Abb. 3.2 erklärt, die ein Exempel eines Polyacrylamidgels zeigt (Abb. 3.2A). Zuerst wurden die Gelbahnen markiert, in denen die DNA-Stränge der Verdünnungsreihe des DNA-Standards und der Proben getrennt wurden, beispielhaft anhand der zweiten Gel-Bahn von links auf Abb. 3.2A illustriert. Die Fluoreszenz-Intensitäten der markierten Bereiche wurden geplottet. Abb. 3.2C und Abb. 3.2D zeigen Ausschnitte der Plots im Bereich der 10 bp-Stränge des Standards bzw. der 10 bp-DNA-Stränge der Proben. Nach Abzug der Hintergrund-Fluoreszenz des Gels war die Fläche eines Peaks proportional zur DNA-Stoffmenge bzw. -Masse der zugehörigen DNA-Bande (Abb. 3.2B,C). Zur Quantifizierung der Stoffmenge der Proben-DNA-Stränge in den Experimenten, deren Ergebnisse in Abb. 3.4A und 3.8 präsentiert werden, wurden die Kalibrierungsgeraden der Kalibrierungs-DNA-Stränge gleicher Länge verwendet (beispielsweise Abb. 3.2B). Für alle anderen DNA-Längen wurden interpolierte Werte verwendet. Für die anderen Experimente wurden die Fluoreszenz-Intensitäten aller DNA-Banden auf die Fluoreszenz-Intensität der Bande des kürzesten DNA-Moleküls der Nullprobe normiert. Von diesen Verhältnissen wurde die relative DNA-Stoffmenge einer jeden DNA-Bande abgeleitet.

3.3. Ergebnisse

3.3.1. Ergebnisse eigener Vorarbeiten

Abschn. 3.3.1 ist eine überarbeitete Fassung der für die Dissertation relevanten Ergebnisse der eigenen unveröffentlichten Diplomarbeit (*Enzymbasierte Selbstreplikation von DNA Fragmenten*, Universität des Saarlandes, Lehrstuhl für Biologische Experimentalphysik, 2012). Abschn. 3.3.1 soll dem grundlegenden Verständnis der auf der Diplomarbeit aufbauenden Studie, die während der Promotionsphase durchgeführt wurde, dienen. Er soll die im Anschluss präsentierten und diskutierten Experimente motivieren.

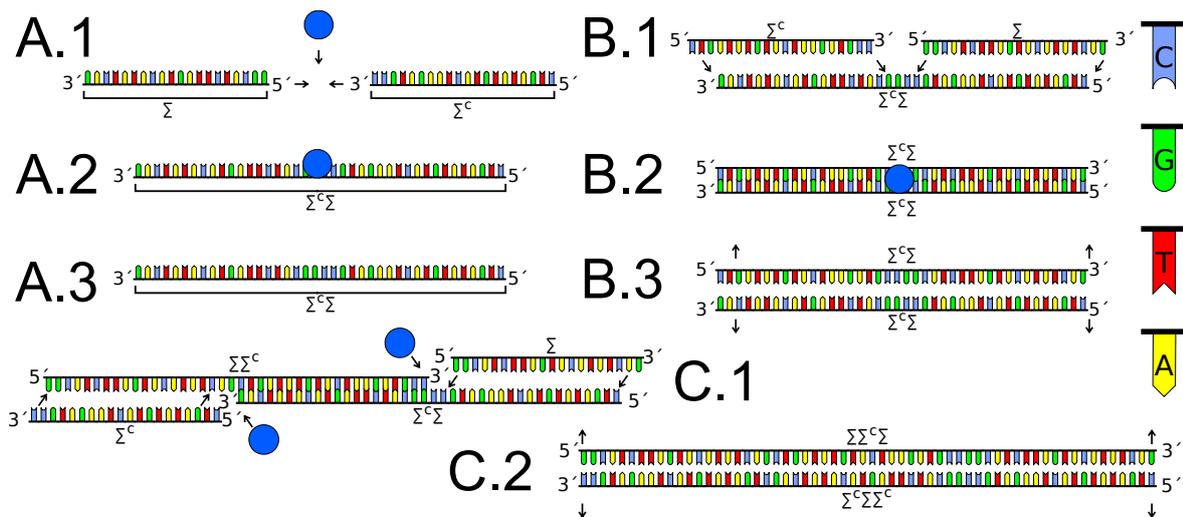


Abbildung 3.3.: Schematische Darstellung verschiedener Ligationenprozesse. A: Freie Ligation, bei der die reaktiven Enden zweier komplementärer DNA-Einzelstränge mit den Sequenzen Σ und Σ^c und die Taq-DNA-Ligase (blauer Kreis) spontan aufeinander treffen (A.1). Das Enzym verbindet die beiden Einzelstränge kovalent (A.2). Nach Ablösen des Enzyms ist ein längerer DNA-Strang entstanden (A.3). B: Template-basierte Ligation, bei der zwei DNA-Einzelstränge auf dem komplementären Template hybridisieren (B.1). Durch die Hybridisierung werden die reaktiven Enden der zu ligierenden DNA-Stränge in räumliche Nähe gebracht und von der Taq-DNA-Ligase kovalent verbunden (B.2). Ablösen des Enzyms ermöglicht es, dem neu entstandenen Molekül selbst als Template zu fungieren (B.3). C: Überhang-Ligation bei der DNA-Stränge unvollständig hybridisieren und einzelsträngige Überhänge erzeugen, die als Template dienen (C.1). Dadurch entstehen DNA-Stränge verschiedener Längen, die ihrerseits als Templates dienen (C.2). Abb. 3.3 ist modifiziert aus [158] übernommen.

Zur experimentellen Erforschung möglicher Bedingungen, die ein unbegrenztes Andauern molekularer Evolution ermöglichen, wurden, wie eingangs erwähnt, lineare DNA-Moleküle verwendet, die durch ihre DNA-Sequenz miteinander in Wechselwirkung treten können. Die Taq-DNA-Ligase [187] wurde dem experimentellen System hinzugegeben. Dieses Enzym katalysiert die kovalente Bindung der reaktiven 5'- und 3'-Enden zweier aneinandergrenzender einzelsträngiger DNA-Moleküle. Dieser Prozess kann entweder beim zufälligen Aufeinandertreffen zweier freier DNA-Enden und des Enzyms selten ablaufen und zu einer Variation führen (Abb. 3.3A) oder häufig in einer Template-basierten Art und Weise für die Reproduktion stattfinden (Abb. 3.3B).

Es wurden 20 b lange DNA-Einzelstränge mit den komplementären Sequenzen Σ und Σ^c (Mengenverhältnis 1:1) verwendet. Die Sequenzen wurden zufällig unter der Bedingung ausgewählt, dass die Schmelztemperatur der Stränge im Bereich der höchsten enzymatischen Aktivität der Taq-DNA-Ligase

lag und dass die Stränge keine Sekundärstrukturen bildeten. Die Proben wurden einer zyklischen Variation der Temperatur ausgesetzt. Die Inkubation der Proben leicht oberhalb der Schmelztemperatur ermöglicht beides, die freie Ligation wie auch die Template-basierte Reproduktion. Die zyklischen Temperatursprünge auf die hohen Temperaturen ermöglichten die Denaturierung der DNA-Doppelstränge, sodass die gebildeten freien Einzelstränge der Reproduktion als auch der freien Ligation als Ressourcen zur Verfügung standen und die Produkthemmung minimiert wurde. Die hohen Temperaturen reichten aus, um auch die im Verlauf des Experiments entstandenen, längeren DNA-Doppelstränge zu denaturieren. Diese Änderung der physischen Parameter war Teil des Reproduktionsprozesses, veränderten aber nicht den externen Selektionsdruck.

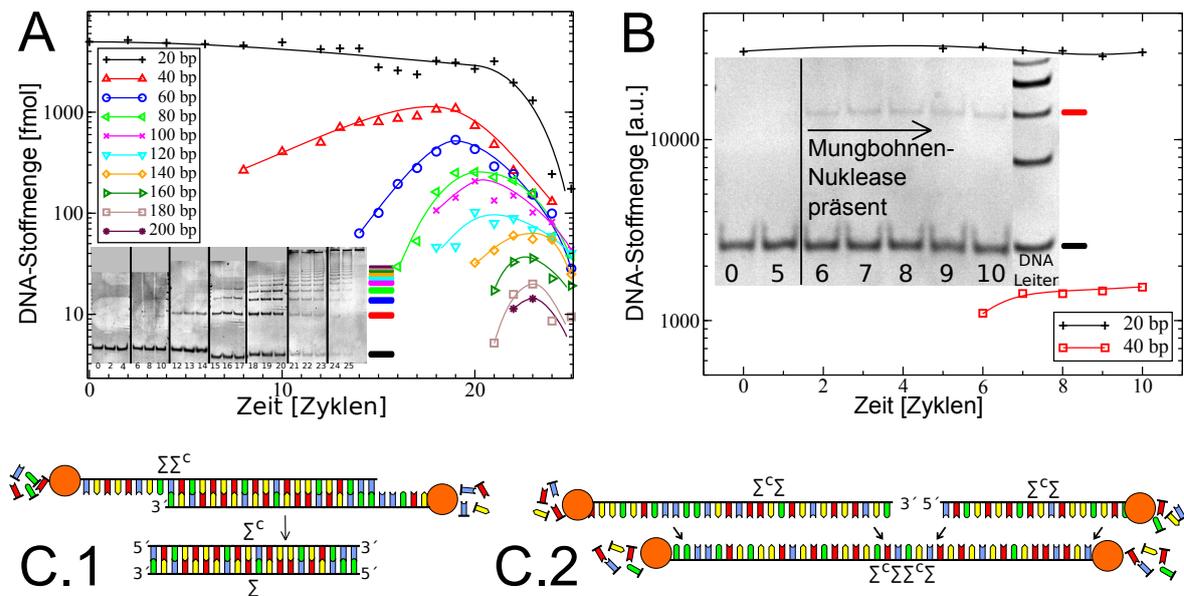


Abbildung 3.4.: Verlängerung der DNA-Stränge im Verlauf der Zeit und enzymatische Unterdrückung der Überhang-Ligation durch die Mungbohnen-Nuklease. A: Zeitlicher Stoffmengenverlauf der DNA-Doppelstränge in Präsenz der Taq-DNA-Ligase. Per freier Ligation wurden aus den initial vorhandenen 20 b langen Einzelsträngen 40 b lange Einzelstränge erzeugt, die als Template für die autokatalytische Bildung weiterer 40 b-Stränge dienen. Im weiteren Verlauf entstanden weitere Doppelstränge der Länge $20 \cdot N$ ($N = 1, 2, \dots, 5$ und darüber), die mittels PAGE detektiert werden konnten (Einbettung). B: Experiment zum Test der Hypothese, ob sich die in (A) beobachtete Wachstumsdynamik durch Abwesenheit des Überhang-Ligations-Mechanismus ändert. Die Mungbohnen-Nuklease (ab dem sechsten Zyklus im System) entfernte einzelsträngige Überhänge (C.1) und stoppte die Strang-Verlängerung komplett. Einerseits unterdrückte die Exonuklease die Überhang-Ligation erfolgreich, andererseits attackierte sie auch die spontan gebildeten längeren DNA-Templates (C.2) so schnell, dass deren Template-basierte Reproduktion inhibiert wurde, die Population der 40 bp-Stränge nicht mehr zunahm (eingefügtes Polyacrylamidgel) und somit auch keine längeren Stränge oberhalb einer Länge von 40 bp mehr in detektierbarer Menge entstanden. Die Linien dienen der optischen Führung. Abb. 3.3 ist modifiziert aus [158] übernommen.

Abb. 3.4A zeigt die Ergebnisse dieses Experiments. Mit der Zeit erzeugte das betrachtete System längere DNA-Stränge. Initial entstanden per freier Ligation zweier 20 b-Einzelstränge 40 b lange Einzelstränge (Abb. 3.3A), die dann im weiteren Verlauf als Templates für die autokatalytische Bildung weiterer 40 b-Stränge fungierten (Abb. 3.3B). Nach und nach konnten weitere Doppelstränge der Länge $20 \cdot N$ ($N = 1, 2, \dots, 5$ und darüber) mittels PAGE detektiert werden. Diese Wachstumsdynamik kann durch den Mechanismus der Überhang-Ligation erklärt werden (Abb. 3.3C). Um diese Hypothese zu testen, wurde untersucht, ob sich die beschriebene Wachstumsdynamik des Systems durch Abwesenheit der Überhang-Ligation ändert.

Zu diesem Zweck wurde im folgenden Experiment die Mungbohnen-Nuklease [188], [189], die einzelsträngige DNA Sequenz-unspezifisch und daher insbesondere einzelsträngige Verlängerungen an doppelsträngigen DNA-Molekülen degradiert, in das System integriert. Während der Inkubation der 20 b-Einzelstränge (Σ und Σ^c) wurde die Mungbohnen-Nuklease, kurz bevor die 40 b-Stränge nachgewiesen werden konnten, in die Proben gegeben. Dies führte dazu, dass keine Doppelsträngige, die länger als 40 bp lang waren, beobachtet werden konnten (Abb. 3.4B). Die Entstehung längerer DNA-Stränge konnte durch zwei mit der Mungbohnen-Nuklease-Aktivität assoziierte Mechanismen unterdrückt werden: Einerseits wurden die einzelsträngigen Überhänge durch das Enzym entfernt und somit dem Überhang-Ligations-Mechanismus die Basis entzogen. Andererseits degradierte die Mungbohnen-Nuklease spontan gebildete längere, einzelsträngige Templates zu schnell und verhinderte somit deren Reproduktion. Die Reduzierung der enzymatischen Aktivität dieser Exonuklease ließ neben den 20 bp- und 40 bp-Strängen zusätzlich 60 bp lange DNA-Moleküle entstehen (Anh. A.1.2). Diese Ergebnisse unterstützten die Hypothese, dass die Entstehung längerer Stränge auf die Überhang-Ligation zurückzuführen ist.

Die in der Diplomarbeit durchgeführte Studie ließ die Frage offen, ob und wie es möglich wäre, unbegrenztes Längenwachstum der DNA-Stränge ohne den Mechanismus der Überhang-Ligation zu erzeugen und ob sich ohne diesen Überhangmechanismus die Wachstumsdynamik des Systems ändern würde. Es wurde dort abschließend spekuliert, dass der gezielte Restriktionsverdau der im Verlauf des Experiments entstehenden heterogenen DNA-Moleküle ($\Sigma\Sigma^c$ und $\Sigma^c\Sigma$, Abb. 3.3C), die vornehmlich Überhänge ausbilden, unbegrenztes Längenwachstum der DNA-Fragmente in Abwesenheit der Überhang-Ligation ermöglichen könnte. Diese Idee wurde in der vorliegenden Dissertation aufgegriffen und realisiert.

3.3.2. Restriktionsverdau der DNA-Moleküle mit heterogenen Sequenzen

In den zuvor präsentierten und diskutierten Experimenten bildeten sich vor allem dann einzelsträngige Überhänge aus, wenn zwei DNA-Stränge mit heterogenen Sequenzen $\Sigma\Sigma^c$ und $\Sigma^c\Sigma$ (Abb. 3.3C) mit einem Versatz hybridisierten. Ein alternativer Ansatz zur Sequenz-unspezifischen Degradation einzelsträngiger DNA durch die Mungbohnen-Nuklease besteht darin, die Doppelstrang-DNA mit heterogenen Sequenzen enzymatisch zu verdauen, bevor sich aus diesen Strängen Überhänge bilden können (Abb. 3.5A,B). Die Restriktionsenzyme AfeI und BstUI schneiden bzw. verdauen doppelsträngige DNA an komplementären Erkennungssequenzen: Während AfeI DNA mit der Sequenz 5'-AGC[↓]GCT-3' (Abb. 3.5A) schneidet, verdaut BstUI DNA mit der Sequenz 5'-CG[↓]CG-3' (Abb. 3.5B); der Pfeil gibt jeweils die Schnittstelle an.

Im folgenden Experiment wurden initial 10 b lange DNA-Stränge verwendet, die die komplementären Sequenzen $\sigma = \text{CGT TAT TAG C}$ und $\sigma^c = \text{GCT AAT AAC G}$ (in 5' \rightarrow 3'-Richtung dargestellt, Mengenverhältnis 1:1) besaßen (Abb. 3.5A-C). Bildeten sich per Ligation im Verlauf des Experiments DNA-Stränge mit den heterogenen Sequenzen $\sigma\sigma^c$ und $\sigma^c\sigma$, besaßen sie die Erkennungssequenzen der beiden Restriktionsenzyme und konnten enzymatisch degradiert werden (Abb. 3.5A,B). DNA-Doppelstränge mit der homogenen Sequenz $\sigma\sigma$ wurden von den Enzymen hingegen nicht gespalten, da sich die jeweiligen AfeI- und BstUI-Erkennungssequenzen nicht trugen (Abb. 3.5C).

Abb. 3.5D zeigt die Ergebnisse des Experiments. Wie bei den vorherigen Experimenten bildeten sich aus den initialen DNA-Molekülen (10 b lang) die nächst längeren DNA-Stränge (20 b lang) per freier Ligation. Allerdings konnten keine 30 bp-Stränge detektiert werden. Die nächst längeren DNA-Moleküle

besaßen eine Länge von 40 bp. Im weiteren Verlauf des Experiments entstanden längere Doppelstränge der Länge $10 \cdot N$ ($N = 5, 6, \dots, 8$ und darüber).

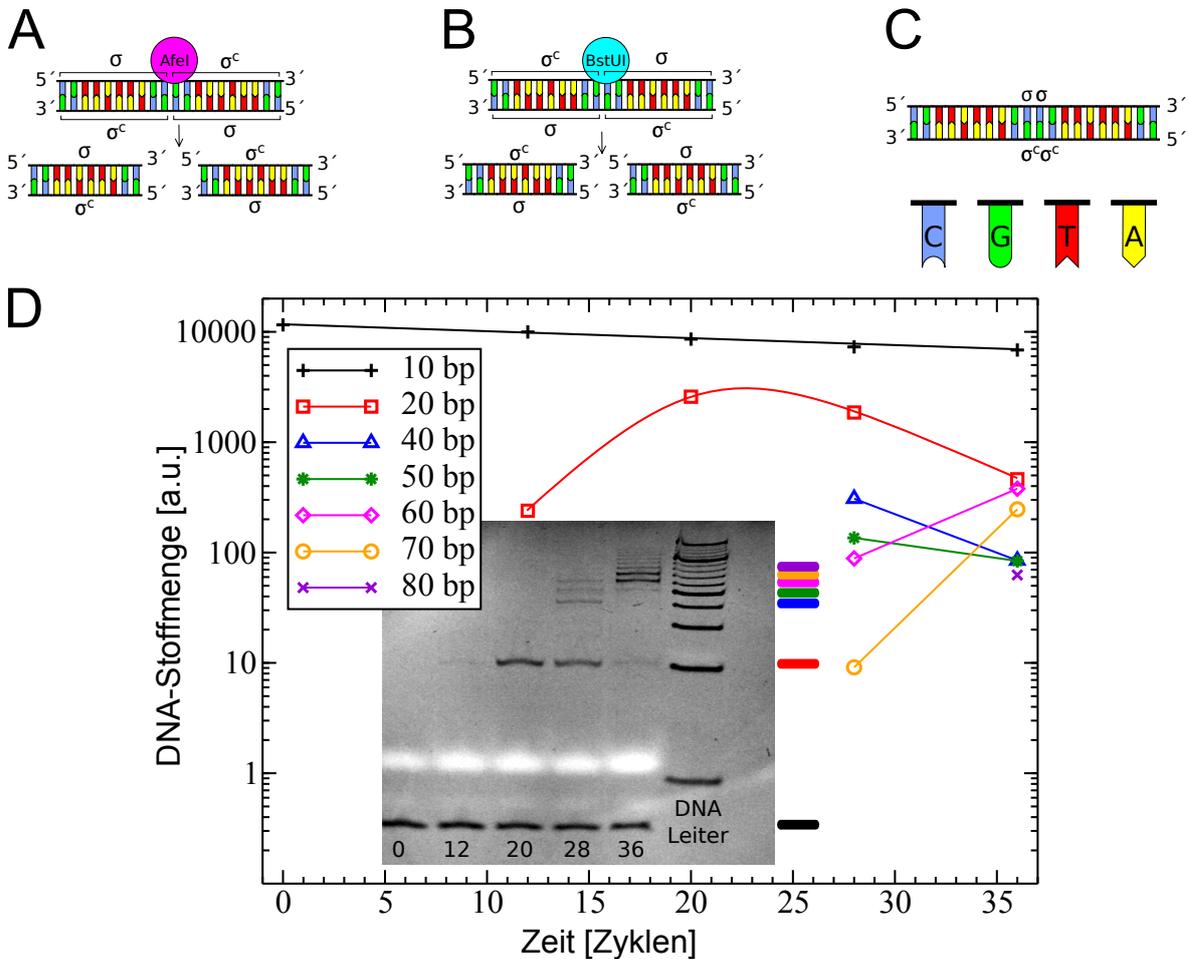


Abbildung 3.5.: Restriktionsverdau der doppelsträngigen DNA-Moleküle mit den heterogenen Sequenzen $\sigma\sigma^c$ und $\sigma^c\sigma$, um die Überhang-Ligation im Experiment zu unterdrücken. A-C: Schematische Darstellung des Restriktionsverdaus der doppelsträngigen DNA mit den heterogenen Sequenzen $\sigma\sigma^c$ und $\sigma^c\sigma$ durch die Enzyme AfeI (violetter Kreis, A) bzw. BstUI (türkisfarbener Kreis, B). Nach dem Restriktionsverdau sind die verdauten DNA-Stränge erneut ligierbar. DNA-Stränge mit den homogenen Sequenzen $\sigma\sigma$ werden nicht von den Restriktionsenzymen AfeI und BstUI verdaut (C). Zeitlicher relativer Stoffmengenverlauf der DNA-Doppelstränge in Präsenz von AfeI und BstUI, die nach dem achten Zyklus das erste Mal hinzugegeben wurden. Eingefügt ist das zugehörige Polyacrylamidgel zu sehen. Die Zyklenzahl bei Probenentnahme ist darauf angegeben. Rechts zu sehen ist die Auftrennung der 10 bp, 20 bp, ..., 150 bp-DNA-Stränge der DNA-Leiter (Abschn. 3.2.3). Die Linien dienen der optischen Führung. Mit der Zeit bildeten sich 20 bp-Stränge und 40 bp-Stränge, ohne dass 30 bp-Stränge detektiert wurden. Oberhalb einer Länge von 40 bp konnten $10 \cdot N$ bp lange Doppelstränge ($N = 5, 6, \dots, 8$ und darüber) beobachtet werden. Abb. 3.5 ist modifiziert aus [158] übernommen.

Es ist anzunehmen, dass die Überhang-Ligation erneut die Wachstumsdynamik dominierte, da auch die nicht gespaltenen DNA-Moleküle mit den homogenen Sequenzen $\sigma\sigma$ und $\sigma^c\sigma^c$ eine geringe Wahrscheinlichkeit zur Überhang-Bildung besaßen.

3.3.3. AfeI, BstUI und Mungbohnen-Nuklease zur Unterdrückung der Überhang-Ligation

Im nächsten Experiment wurde die enzymatische Aktivität der Restriktionsenzyme AfeI und BstUI sowie der Mungbohnen-Nuklease gebündelt, um die Überhang-Ligation komplett zu unterdrücken. AfeI und BstUI sollten, wie auch schon im letzten Experiment (Abschn. 3.3.2), die DNA-Doppelstränge mit heterogenen Sequenzen enzymatisch degradieren, sodass diese keine Überhänge bilden können. Sofern sich, wie angenommen, Überhänge aufgrund unvollständig hybridisierter DNA-Stränge mit den homogenen Sequenzen $\sigma\sigma$ und $\sigma^c\sigma^c$ bilden sollten, würden diese durch die zusätzliche Hinzugabe der Mungbohnen-Nuklease entfernt werden.

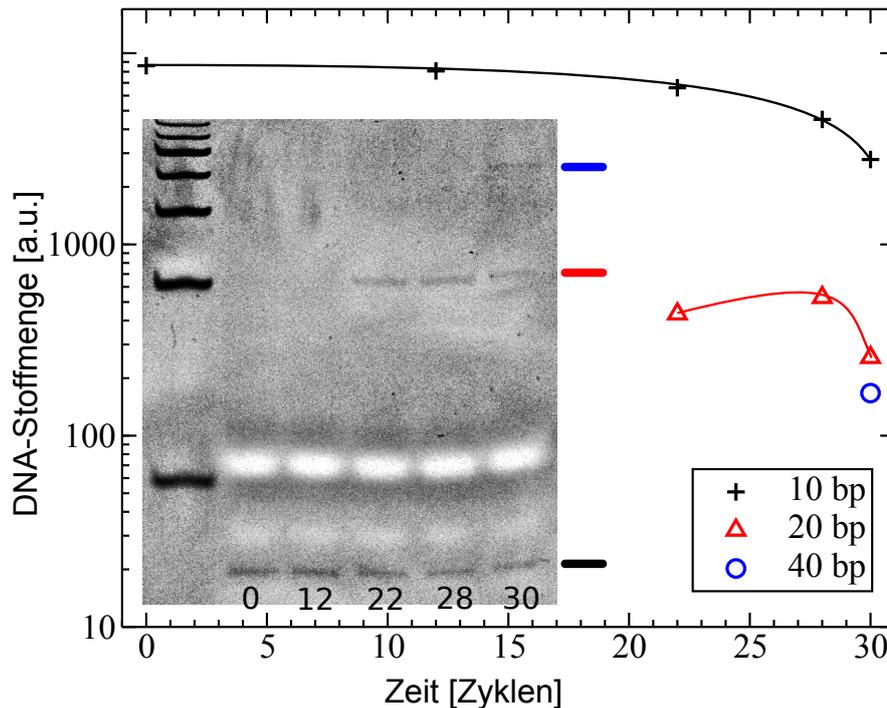


Abbildung 3.6.: Zeitlicher Stoffmengen-Verlauf der DNA-Moleküle bei kombinierter enzymatischer Aktivität von AfeI, BstUI und der Mungbohnen-Nuklease. AfeI und BstUI wurden initial nach acht Zyklen und die Mungbohnen-Nuklease das erste Mal nach 15 Zyklen hinzugegeben. Links zu sehen ist die Auftrennung der 10 bp, 20 bp, ..., 150 bp-DNA-Stränge der DNA-Leiter (Abschn. 3.2.3). Die Linien dienen der optischen Führung. Die Zyklenzahl bei Probenentnahme ist auf dem Gel angegeben. Initial befanden sich die 10 b langen Stränge mit den Sequenzen σ und σ^c im System. Es entstand ein Regime, in dem neben den 10 bp-Strängen noch 20 bp- und 40 bp-Stränge detektiert werden konnten. Abb. 3.6 ist modifiziert aus [158] übernommen.

Abb. 3.6 zeigt die Ergebnisse des Experiments. Mit der Zeit konnte die Entstehung von 20 bp- und 40 bp-Strängen neben den sich initial im System befindlichen 10 bp-Strängen beobachtet werden. Die Population der 30 bp langen DNA-Moleküle blieb wie im vorangegangenen Experiment unterhalb der Detektionsgrenze. Demnach gelang es tatsächlich, durch die zusätzliche Hinzugabe der Mungbohnen-Nuklease die Überhang-Ligation komplett zu unterdrücken. Allerdings entstand eine zu kleine Population an 40 bp langen DNA-Molekülen, als dass daraus detektierbare Mengen noch längerer DNA-Moleküle hätten entstehen können.

Die bisherigen experimentellen Ansätze ließen nach wie vor die Frage offen, ob ein unbegrenztes Längenwachstum der DNA-Moleküle in Abwesenheit der Überhang-Ligation möglich sei und ob sich

ohne diesen Mechanismus die Wachstumsdynamik des Systems ändern würde. Um eine Antwort auf diese Frage zu erhalten, wurde im Folgenden die Dynamik des Längenwachstums der DNA-Stränge theoretisch analysiert. Dieser Ansatz ermöglicht es, die Dynamik des Längenwachstums nur in Präsenz der spontanen und autokatalytischen Strang-Verlängerung ohne den Mechanismus der Überhang-Ligation zu untersuchen.

3.3.4. Theoretische Analyse

Die theoretische Analyse, deren für weitere Experimente relevanten Ergebnisse in Abschn. 3.3.4 zusammenfasst werden, wurde von Philipp Zimmer durchgeführt. Diese und weitere Ergebnisse seiner Studie können in seiner Dissertation [159] gefunden werden.

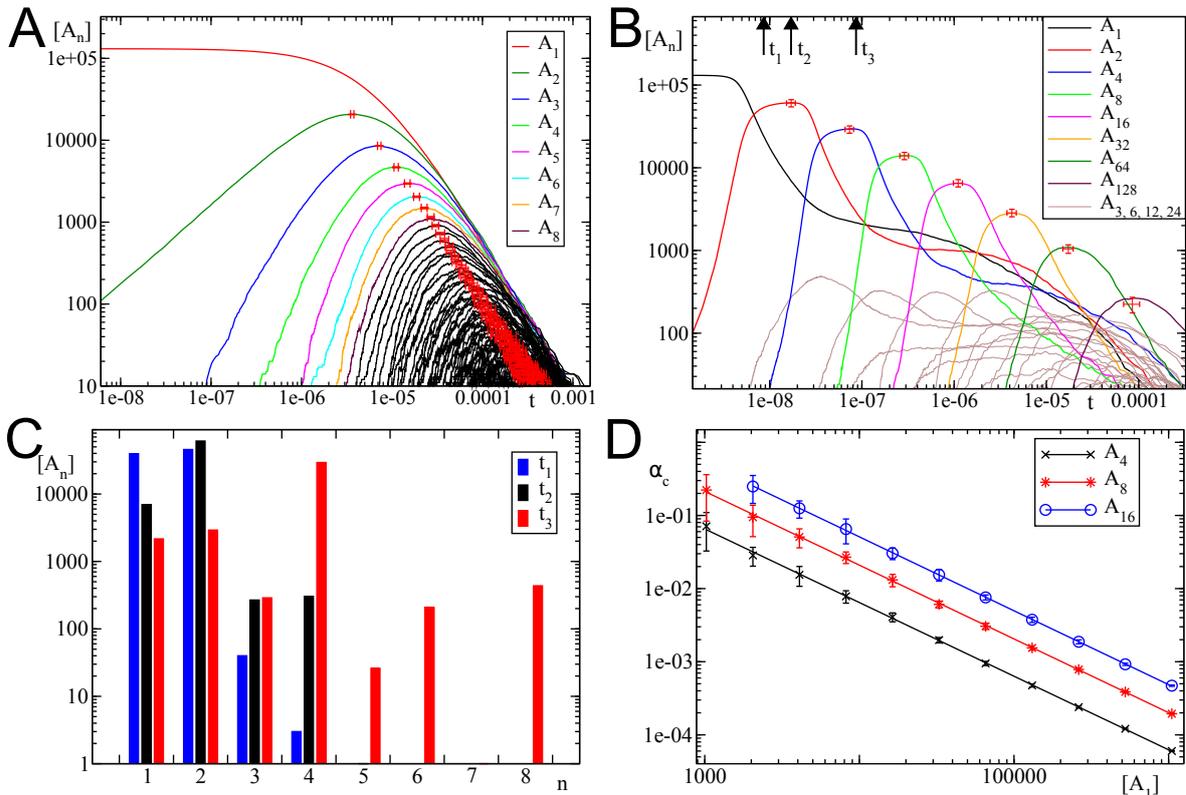


Abbildung 3.7.: Verlangern der Polymere im Verlauf der Zeit und Analyse des stochastischen Modells. A: Simulation des stochastischen Modells mit $\alpha=0,00001$. Wie im DNA-System (Abb. 3.4A) entstanden im Wesentlichen alle möglichen Polymerlängen im Verlauf der Zeit. B: Simulation des stochastischen Modells mit $\alpha=0,1$. Es entstanden zwar mit der Zeit mehrere unterschiedlich lange Polymere, allerdings dominierten die Polymere, deren Längen aus Verdopplungsschritten hervorgingen, zahlenmäßig das System. C: Konzentrationen aller vorhandenen Polymere zu den Zeitpunkten t_1 , t_2 und t_3 , die in (B) markiert sind. t_1 und t_3 markieren Zeitpunkte zu Beginn bzw. gegen Ende der Wachstumsphase von A_4 , t_2 liegt zwischen diesen Zeitpunkten. D: Kritischer Wert $\alpha_{c,n}$, bei dem Polymere bis zu der Länge n selektiert wurden, als Funktion der initialen Konzentration A_1 . Simulationen und deren Analyse von Philipp Zimmer durchgeführt. Abb. 3.7 ist modifiziert aus [158] übernommen.

Im theoretischen System wurde ein Mix aus Polymeren der Länge A_n betrachtet. A_n stellt eine Kette, die sich aus n Monomeren zusammensetzt ($n=1, 2, \dots$), dar. Zwei Polymere der Länge n und m können durch die Reaktion $A_n + A_m \rightarrow A_{n+m}$ miteinander zum Polymer A_{n+m} der Länge $n+m$ verknüpft werden. Diese Reaktion verläuft entweder spontan gemäß der Massenwirkungskinetik mit der

Rate β ab oder sie verläuft autokatalytisch mit einer Rate $\alpha \cdot [A_{n+m}]$, die linear mit Konzentration der Polymere A_{n+m} ansteigt. $[A_{n+m}]$ bezeichnet die Konzentration der Polymere A_{n+m} . Sequenzeffekte und Komplementarität von DNA wurden im theoretischen Modell vernachlässigt. Dies ermöglicht, die beiden Verknüpfungsmechanismen ohne Rücksicht auf molekulare Details und somit auf allgemeingültige Weise zu untersuchen. Die Reaktionskinetik der durchschnittlichen Polymer-Konzentrationen ist durch folgende chemische Ratengleichung gegeben:

$$\begin{aligned} \frac{d}{dt}[A_{n+m}] = & (\alpha[A_{n+m}] + \beta) [A_n][A_m] \\ & - \sum_k (1 + \delta_{n+m,k}) (\alpha[A_{n+m+k}] + \beta) [A_{n+m}][A_k]. \end{aligned} \quad (3.1)$$

Die Konstante α ist ein Maß für die Stärke der Autokatalyse. Gleichung 3.1 impliziert, dass die Polymere im System, das ein festes Volumen besitzt, gut gemischt sind, sodass räumliche Konzentrationsgradienten vernachlässigt werden können. Die Konzentration $[A_n]$ der Polymere A_n wird durch die Zahl der Moleküle ausgedrückt, die im Folgenden dieselbe Nomenklatur wie die Konzentration $[A_n]$ besitzt.

Da chemische Reaktionen stochastisch ablaufen, wurde die Dynamik der Verknüpfungen mit Hilfe einer Mastergleichung beschrieben. Diese wurde mittels einer stochastischen Simulation in Form eines Gillespie-Algorithmus [190], [191] untersucht. Die Analyse wurde auf mehreren Rechnern parallelisiert [192]. Anfänglich befanden sich nur Monomere im zu analysierenden System, d.h. $[A_n] = 0$ für $n \geq 2$. Die Zahl der initial vorhandenen Monomere lag zwischen 2^9 und 2^{20} , üblicherweise wurde 2^{17} ausgewählt. Die Simulationen wurden abgebrochen, wenn die Polymere des Systems eine gewisse Länge, typischerweise 2^9 , erreicht hatten. Die Zeit wurde so gewählt, dass $\beta = 1$ galt.

Je nach Größe des Wertes für α konnten zwei unterschiedliche Regime bei der Wachstumsdynamik der Polymere beobachtet werden (Abb. 3.7). Ist der Wert für α klein genug, entstanden prinzipiell Polymere aller Längen im Verlauf der Zeit (Abb. 3.7A). War der Wert für α jedoch groß genug, so evolvierte das System aus Polymeren nach folgendem Mechanismus: Zuerst bildeten sich aus den Monomeren A_1 per spontaner Verknüpfung Dimere A_2 . Aufgrund der starken Autokatalyse reproduzierten sich die Polymere A_2 effizient und reduzierten somit die Konzentration der Monomere A_1 erheblich. Der schnelle Umsatz von Monomeren zu Dimeren reduzierte die Wahrscheinlichkeit einer spontanen Verknüpfung von Dimeren und Monomeren zu Trimeren, vor allem aber die Reproduktionsrate der Trimere. Obwohl die Trimere vor den Tetrameren entstanden, waren die Tetramere nach den Dimeren die nächst längeren, in großer Anzahl beobachtbaren Polymere. Für genügend große Werte von α wiederholte sich dieser Prozess auch bei den Tetrameren. Auf diese Weise wurden zwar Polymere aller Längen gebildet, aber nur solche Polymere, die aus dem beschriebenen Längenverdopplungsprozess hervorgingen, dominierten zahlenmäßig das System (Abb. 3.7B).

Die Trimere waren nicht nur, wie eben bereits erwähnt, zeitlich vor den Tetrameren zu finden, sondern die Tetramere erreichten ihre maximale Konzentration in Präsenz von Pentameren, Hexameren und Oktameren (Abb. 3.7C), sodass verschieden lange Polymere koexistierten.

Der Längenverdopplungsprozess setzte genau dann ein, wenn die Reproduktionsrate der 2^k langen Polymere einen kritischen Wert überschritt. Für jedes k existierte ein kritischer Wert $\alpha_{c,n}$, über dem der Längenverdopplungsprozess bis zu einer Polymerlänge von $n = 2^k$ das System im Wesentlichen dominierte. Ein Ergebnis der stochastischen Simulation war, dass diese Werte algebraisch von der initialen Konzentration der Monomere $[A_1]_{\text{init}}$ abhingen: $\alpha_{c,n} \propto [A_1]_{\text{init}}^\gamma$ mit $\gamma = -1,01(1)$ für alle $n = 2^k$ (Abb. 3.7D). Wie stark die aus dem Längenverdopplungsprozess hervorgehenden Polymere das System

zahlenmäßig dominierten, hing von der Größe des Wertes für α ab. Je größer der Wert von α war, desto größer waren auch die Konzentrationsunterschiede zwischen den selektierten längen-verdoppelten Polymeren und den Polymeren anderer Längen. Für sehr große α -Werte entstanden sogar ausgedehnte Phasen, in denen nur Polymere einer einzigen Länge existierten¹.

3.3.5. Experimentelle Umsetzung der Längenverdopplung

Die theoretische Analyse zeigte, dass es zur Selektion von Polymeren bestimmter Längen nötig ist, ein kritisches Verhältnis zwischen autokatalytischer Reproduktion und nicht-autokatalytischer Polymerbildung zu überschreiten.

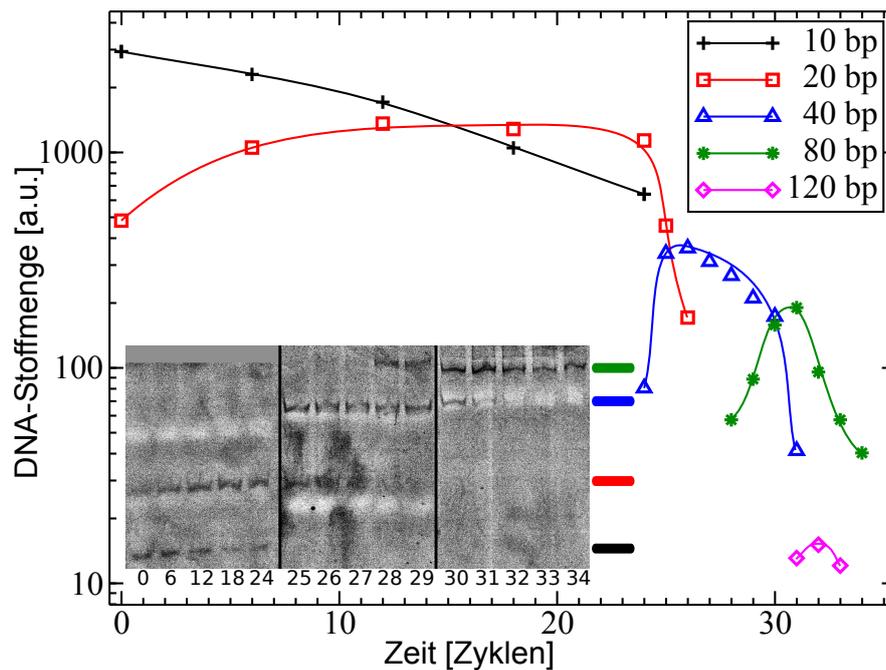


Abbildung 3.8.: Verlängerung der DNA-Stränge im Verlauf der Zeit. Initial bestand das System aus einem Mix aus 10 b langen DNA-Strängen mit 16 unterschiedlichen und paarweise komplementären Sequenzen und der Taq-DNA-Ligase. Mit Hilfe von „Verknüpfer“-Molekülen (Abschn. 3.2 und Anh.A.1.1) wurde die autokatalytische Reproduktion initiiert, und die Template-basierte Reproduktionsrate $\alpha \cdot [A_n]$ konnte durch eine Erhöhung der Menge an hinzugegebenen „Verknüpfern“ verstärkt werden. Die 20 b langen „Verknüpfer“ befanden sich von Beginn an im System, nach 24 und 29 Zyklen wurden die 40 b und 80 b langen Verknüpfer in das System gebracht. Eingefügt zu sehen sind die zugehörigen Polyacrylamidgels. Die Linien dienen der optischen Führung. Die Zyklenzahl bei Probenentnahme ist unterhalb der Gele angegeben. Die durch die „Verknüpfer“ verstärkte Template-basierte Reproduktionsrate führte zu einem Regime, in dem solche DNA-Stränge selektiert wurden, die aus einem Längenverdopplungsprozess resultierten, d.h. 10 bp-, 20 bp, 40 bp und 80 bp-Stränge dominierten zahlenmäßig das System. Abb. 3.8 ist modifiziert aus [158] übernommen.

Für weitere Experimente bedeutete dies, dass der Einfluss der Überhang-Ligation drastisch reduziert werden musste, da dieser Mechanismus in den bisherigen Experimenten die Wachstumsdynamik der DNA-Stränge dominiert hatte. Darüber hinaus mussten die Raten im experimentellen System kontrolliert werden. Um beides zu erreichen, wurde das experimentelle System wie folgt modifiziert: Es wurden 16 10 b lange DNA-Stränge mit paarweise komplementären Sequenzen verwendet, die unter

¹Weitere Details zur theoretischen Analyse können [158], [159] entnommen werden.

den experimentellen Bedingungen weder kreuzhybridisierten noch mit sich selbst hybridisierten. Die verwendeten Sequenzen wurden aus einer Menge zufällig generierter Sequenzen unter der Bedingung, dass die Schmelztemperaturen [185], [186] ähnlich waren, ausgewählt². Diese DNA-Stränge können bis zu einer Länge von 80 bp ohne Wiederholungen eines Sequenzmotivs verknüpft werden. Auf diese Weise sollte das unvollständige Hybridisieren von DNA-Strängen, die Ausbildung von Überhängen und die damit verbundene Überhang-Ligation entscheidend unterdrückt werden.

Um ausschließlich DNA-Stränge mit solchen Sequenzen zu erhalten, die höchstens eine Kopie einer jeden 10 b-Sequenz besitzen, wurde die initiale DNA-Lösung mit einer kleinen Menge an 20 b-langen Einzelsträngen inokuliert. Die Sequenzen dieser 20 b-Stränge waren Kombinationen aus den Sequenzen der initialen 10 b-Stränge, sodass die resultierenden 20 b-Stränge paarweise komplementär waren. Diese „Verknüpfer“ (Abschn. 3.2 und Anh.A.1.1) initiierten die autokatalytische Reproduktion. Die Template-basierte Reproduktionsrate $\alpha \cdot [A_n]$ konnte somit durch eine Erhöhung der Menge an hinzugegebenen Verknüpfern verstärkt werden. Durch ein festes experimentelles β (Anh. A.1.3) konnte demnach das Verhältnis $\omega_n = \alpha \cdot [A_n] / \beta_n$ (Anh. A.1.3) zwischen autokatalytischer Vermehrungsrate und spontaner Ligationsrate eingestellt werden. Nach dem gleichen Prinzip, das bei den 20 b langen „Verknüpfern“ angewendet wurde, wurden 40 b und 80 b lange „Verknüpfers“ in das System gebracht, wenn die jeweiligen Vorgänger-Stränge ihren Bereich der maximalen Konzentration erreicht hatten.

Durch die beschriebenen Modifikationen entstand auch im Experiment ein Regime, in dem solche DNA-Stränge selektiert wurden, die aus einem Längenverdopplungsprozess resultierten, folglich dominierten 10 bp-, 20 bp-, 40 bp- und 80 bp-Stränge zahlenmäßig das System (Abb. 3.8). Neben diesen DNA-Doppelsträngen blieben nur doppelsträngige DNA-Moleküle einer Länge von 120 bp oberhalb der Detektionsgrenze. Eine Bestimmung des Verhältnisses ω_n zwischen autokatalytischer Vermehrungsrate und nicht-autokatalytischer Produktion ergab für dieses Längenverdopplungsregime (Abb. 3.8) einen über alle DNA-Längen und Zyklen gemittelten Wert für $\omega \approx 510$. Im Gegensatz dazu wurde für das Regime, in dem die Überhang-Ligation die Wachstumsdynamik dominierte (Abb. 3.4A), nur ein Wert von $\omega \approx 40$ gefunden.

²Die Sequenzanalyse und -auswahl wurde von Philipp Zimmer durchgeführt. Details können in [158], [159] nachgelesen werden.

3.4. Zusammenfassung, Konklusion und Ausblick

Grundlegende Aspekte evolutionärer Prozesse können mit Hilfe molekularer [15], [16] oder abstrakter Systeme [13], [14] erforscht werden. So geschehen in Kap. 3, das sich mit der Dynamik wachsender, sich reproduzierender linearer Polymere beschäftigt. Experimentell wurden diese Polymere durch lineare DNA-Moleküle verkörpert, die mittels der Taq-DNA-Ligase zu längeren Strängen ligiert wurden. Diese Verlängerung wurde durch zwei Ligationsmechanismen realisiert: die spontane Ligation, die neue Moleküle produzierte, und die Template-basierte Reproduktion, durch die sich existierende Moleküle vermehren konnten. Da die Länge der Moleküle die definierende Eigenschaft eines Polymers darstellt, bezieht sich „neu“ und „existierend“ auf die Länge der Moleküle.

Das betrachtete System ermöglichte konzeptionell unbegrenztes Längenwachstum der Polymere. Darüber hinaus konnten zwei unterschiedliche Regime während der Elongation der Polymere beobachtet werden. Im Fall, dass die spontane Ligation die Wachstumsdynamik dominierte, entstanden Polymere aller möglichen Längen. Dieses Regime bildete sich dann aus, wenn die dominierende Überhang-Ligation schnell Polymere neuer Längen produzierte. Während dieses Regime bereits in der Studie der eigenen, vorangegangenen Diplomarbeit beobachtet wurde, blieb die Frage offen, ob unbegrenztes Längenwachstum der Polymere in Abwesenheit der Überhang-Ligation realisierbar ist und ob sich die Wachstumsdynamik des Systems ohne diesen Mechanismus ändert. Auf der Diplomarbeit aufbauende Versuche, die Überhang-Ligation durch enzymatische Entfernung aller Überhangbildender DNA-Konfigurationen zu unterdrücken, ließen entweder die Überhang-Ligation nach einer gewissen Zeit erneut dominieren oder führten zu einem Stop der Evolution. Inspiriert von der theoretischen Analyse von Philipp Zimmer, wurde durch den Einsatz von „Verknüpfer“-Molekülen und der damit verbundenen Verstärkung der Template-basierten Reproduktion ein Regime erzeugt, in dem die Template-basierte Reproduktion dominierte. In diesem Regime entstanden nur Polymere, deren Längen eine Teilmenge aller möglichen Längen darstellten. Dieser Prozess kann als dynamische Längen-Selektion bezeichnet werden.

Die meisten Studien zur molekularen Evolution selektieren üblicherweise eine optimale Molekül-Konfiguration [17]–[25], die mit dem Stillstand der Evolution in einem (Quasi)-Endzustand verbunden ist (Abb. 3.1B). Es ist ein beschränkter Konfigurationsraum, der den Grund für dieses evolutionäre Szenario darstellt, das im Kontrast zur Evolution natürlicher Ökosysteme steht, die selbst ohne variierenden externen Selektionsdruck [26]–[28] stetig als eng verflochtenes Kollektiv nach höherer Funktionalität [29], [30] streben. Hingegen bietet unbegrenztes Längenwachstum von Polymeren durch Ligationen einen konzeptionellen Ausweg aus dieser evolutionären Sackgasse. Im Kontext der biologischen Evolution interpretiert, produziert polymeres Längenwachstum jedoch im einfachsten Fall eine gegenüber der nicht-selektierenden Vervielfältigung ähnliche Dynamik (Abb. 3.1C). Die Analogie zur biologischen Evolution ergibt sich durch die Betrachtung der Länge der Polymere als phänotypische Eigenschaft. Variationen ergeben sich demnach durch Polymer-Verlängerung, Reproduktion wird durch die Template-basierte Ligation realisiert.

Erst die dynamische Längen-Selektion ermöglicht es, das betrachtete System als ein einfaches evolutionäres System zu interpretieren, das Ähnlichkeiten zur Darwinschen Evolution besitzt (Abb. 3.1A). Dieses Szenario weist die grundlegende Dynamik biologischer Systeme auf, bei der fortwährend ökologische Nischen entstehen, aus denen Spezies mit neuen gewissen phänotypischen Eigenschaften hervorgehen, die ihrerseits wiederum neue Nischen für die weitere Evolution bilden [11], [12]. Beispielsweise wird die initiale Nische aus Monomeren bzw. DNA-Strängen der kleinsten Länge A_n gebildet. Durch eine Variation (Fluktuation) entsteht ein Dimer bzw. DNA-Strang der Länge A_{2n} und somit eine neue

Nische, in der die Dimere bzw. DNA-Stränge der Länge A_{2n} die Monomere bzw. DNA-Stränge der Länge A_n konsumieren. Durch weitere derartige Fluktuationen bilden sich neue Polymerlängen und Nischen, in denen die längeren Polymere die kleineren konsumieren. Unabhängig von den molekularen Details des Experiments zeigt das theoretische Modell von Philipp Zimmer, dass somit das Zusammenspiel von Template-basierter Reproduktion und spontaner Ligation den Phasenraum nicht nur erweitern, sondern dabei auch gewisse Zustände selektieren kann. Das Experiment ist im betrachteten Rahmen eine mögliche Realisierung dieses Modells.

Des Weiteren erweisen sich molekulare evolutionäre Systeme als geeignet, den Ursprung des Lebens zu studieren [181]. Im betrachteten System ermöglicht die Taq-DNA-Ligase die Verlängerung der Moleküle. Um dynamische Verlängerungsprozesse im Kontext des Ursprungs des Lebens studieren zu können, ist die Erforschung weiterer nicht-enzymatischer Ligations-Mechanismen erforderlich [181]. Darüber hinaus sollte das in Kap. 3 präsentierte, molekulare evolutionäre System in anschließenden Arbeiten erweitert werden. Neben der Länge kann die Miteinbeziehung der Sequenzen der Polymere als weiteres phänotypisches Merkmal in das System erfolgen. Um Evolution in Zukunft auf komplexere Weise studieren zu können, müssten die beteiligten Moleküle Funktionalität entwickeln können.

4. Zellfreier Einbau nicht-kanonischer Aminosäuren in Modellproteine

Teile des Kap. 4 wurden vorab in [193], [194] mit den Koautoren Matthias P. Exner, Alessandro De Simone, Marc Schenkelberger sowie den Projekt-betreuenden Professoren Vincent Noireaux, Nediljko Budisa und Albrecht Ott veröffentlicht. Die initiale Idee, nicht-kanonische Aminosäuren mittels des verwendeten zellfreien *E.-coli*-Expressionssystems in Proteine einzubauen, wurde mit Albrecht Ott entwickelt. Die massen- und fluoreszenzspektroskopischen Messungen der in diesem Teil der Dissertation zellfrei synthetisierten Proteine wurden von Matthias P. Exner durchgeführt. Die in [193], [194] veröffentlichten Daten wurden mit Matthias P. Exner gemeinsam analysiert und interpretiert. Vergleichende *in-vivo*-Experimente wurden von Alessandro De Simone durchgeführt, analysiert und interpretiert. Marc Schenkelberger arbeitete allgemein in die zellfreie Genexpression ein.

4.1. Einführung, Stand der Wissenschaft, Strategie und Zielsetzung

Im Lauf der Evolution wurden 20 kanonische Aminosäuren selektiert, die universell vom genetischen Code kodiert und während der natürlichen Proteinbiosynthese an den Ribosomen zu Proteinen translatiert werden. Nur bei speziellen Organismen wird dieser kanonische Satz an Aminosäuren durch Selenocystein (Sec) [32] und Pyrrolysin (Pyl) [33] ergänzt. Ihre biologische Funktion erlangen Proteine durch die dreidimensionale Struktur [34], die maßgeblich aus den Wechselwirkungen der funktionellen Gruppen der konstituierenden Aminosäuren entsteht (vgl. Kap. 4 aus[31]). So existiert zwar ein an das Leben angepasstes, umfangreiches Spektrum an Proteinfunktionen, das aber aufgrund der begrenzten Anzahl der kanonischen Aminosäuren limitiert ist und keineswegs ideal sein muss [35]. Der Einbau nicht-kanonischer Aminosäuren in Proteine zielt darauf ab, diese funktionellen Beschränkungen aufzuheben und Proteine mit neuen physiko- und biochemischen Eigenschaften zu erzeugen [35], [36].

4.1.1. Grenzen der chemischen Peptidsynthese

Im Gegensatz zur ribosomalen Proteinbiosynthese, bei der ein ganzer Satz an Biomolekülen die korrekte Verknüpfung zweier Aminosäuren dirigiert, kommt es bei der chemischen Peptidsynthese zu ungewollten Fehlverknüpfungen an allen anderen funktionellen Aminosäuregruppen, die daher chemisch geschützt werden müssen [39]. Die Festphasensynthese, die auf MERRIFIELD (1963) zurückgeht [195], hat sich als Standardmethode zur chemischen Peptidsynthese durchgesetzt [196], [197]. Die chemische Peptidsynthese kann automatisiert ablaufen, der Einbau nicht-kanonischer Aminosäuren und die Aufreinigung der synthetisierten Peptide gestaltet sich einfach [198]. Beispielsweise konnten mehrere nicht-kanonische Aminosäuren in Fragmente eines Neuropeptid-Hormons an mehreren Positionen gezielt platziert werden [199].

Allerdings akkumulieren sich bei fortwährender chemischer Peptidsynthese Fehlverknüpfungen oder Reaktionen verlaufen unvollständig, sodass sich immer mehr Nebenprodukte ansammeln, die zwar abgetrennt werden können, aber die Ausbeute des erwünschten Peptids drastisch reduzieren [39]. Daher lassen sich bei der chemischen Peptidsynthese nur kleine Peptide mit einer Länge von bis zu etwa 50 Aminosäuren synthetisieren [37], [38]. Zur Erzeugung größerer Peptide oder Proteine bedarf es aufwendiger Mechanismen, u.a. der Peptidligation [198], [200], [201].

Aus diesem Grund hat sich die ribosomale Proteinbiosynthese beim Einbau nicht-kanonischer Aminosäuren in Modellproteine mittels homo- oder heterologer Genexpression durchgesetzt [39]. Bei der homo- bzw. heterologen Genexpression werden wirtseigene bzw. dem Expressionswirt unbekannt Gene exprimiert, die entsprechenden Proteine in großer Menge synthetisiert, isoliert und purifiziert [202]. Sowohl der *in-vivo*- als auch der zellfreie *in-vitro*-Ansatz ermöglichen durch verschiedene Strategien, nicht-kanonische Aminosäuren in Proteine zu inkorporieren. Auf diese Strategien wird in Abschn. 4.1.2 eingegangen.

4.1.2. Einbau nicht-kanonischer Aminosäuren *in vivo* und *in vitro*

Beim ribosomalen Einbau nicht-kanonischer Aminosäuren in Proteine existieren zwei komplementäre Ansätze [36]: der seitenkettenspezifische und ortsspezifische Einbau.

4.1.2.1. Seitenkettenspezifischer Einbau nicht-kanonischer Aminosäuren in Proteine

Die kanonischen Aminoacyl-tRNA-Synthetasen (Singular: aaRS, Plural: aaRSs) besitzen Diskriminierungsmechanismen, um die richtige Aminosäure molekular zu erkennen, zu aktivieren und auf die zugehörige tRNA zu laden [74] (vgl. Abschn. 2.2.2.4, Abb. 4.1). Die Fähigkeit einer aaRS, zwischen strukturell ähnlichen Aminosäuren differenzieren zu können, optimierte sich im Verlauf der Evolution für die in hoher Konzentration und natürlich vorkommenden Aminosäuren [74], [203]. Allerdings weisen die kanonischen aaRSs eine gewisse Substratflexibilität bei der Erkennung von nicht-kanonischen Aminosäuren auf, die ihrer zugehörigen kanonischen Aminosäure strukturell ähneln und nicht im natürlichen Umfeld des Enzyms vorhanden sind [203], [204].

Der seitenkettenspezifische Einbau nicht-kanonischer Aminosäuren beruht auf dieser Substratflexibilität der aaRSs gegenüber solchen nicht-kanonischen, isostrukturellen Analoga [205], [206]: Die nicht-kanonische Aminosäure wird fälschlicherweise von der kanonischen aaRS erkannt, aktiviert und auf die zugehörige tRNA geladen und substituiert letztendlich die kanonische Aminosäure global im translatierten Protein. Diese Entdeckung geht auf MUNIER und COHEN zurück, die zeigten, dass Bakterien L-Selenomethionin anstelle von Met in Proteine einbauen [207].

Auf Grundlage des seitenkettenspezifischen Einbaus nicht-kanonischer Aminosäuren in Proteine haben sich einige Anwendungen entwickelt, die beispielsweise vom Übersichtsartikel [208] zusammengefasst werden und von denen einige prominente Beispiele im Folgenden kurz dargestellt werden. Die Entdeckung von MUNIER und COHEN verbesserte die Röntgenstrahlenstrukturanalyse von Proteinen, da sich durch den seitenkettenspezifischen Einbau von L-Selenomethionin schwere Atome (Selen anstelle von Schwefel) in das zu analysierende Protein einbauen lassen, ohne dessen Struktur nennenswert zu verändern [209]. Dies erlaubt es, die Beugungsintensitäten eines bestimmten Atoms innerhalb der unmodifizierten Proteinstruktur mit den Beugungsintensitäten der mit der nicht-kanonischen Aminosäure modifizierten Proteinstruktur zu vergleichen [210]. Dadurch lässt sich die Proteinstruktur

herleiten [210]. Neben L-Selenomethionin wurden weitere nicht-kanonische Aminosäuren seitenketten-spezifisch in Proteine zur Strukturanalyse inkorporiert [211]–[214]. Der seitenkettenspezifische Einbau nicht-kanonischer Aminosäuren in Proteine ermöglicht es, diese Proteine mittels „Klick-Chemie“ post-translational chemisch zu modifizieren. Beispielsweise kommen die Met-Analoga L-Azidohomoalanin (Aha) oder L-Homopropargylglycin (Hpg), die eine Azidogruppe bzw. eine Alkingruppe tragen, bei „Klick“-Reaktionen zum Einsatz [215]. Durch „Klick-Reaktionen“ lassen sich Affinitäts-Tags [216], [217] oder fluoreszierende Tags [215], [218]–[220] an den nicht-kanonischen Aminosäuren *in vivo* anbringen, mit denen sich neu synthetisierte Proteine extrahieren, quantifizieren und analysieren [216], [217] bzw. lokalisieren und visualisieren [215], [218]–[220] lassen. Darüber hinaus können Proteine über eine solche „Klick-Reaktion“ fusioniert werden [221].

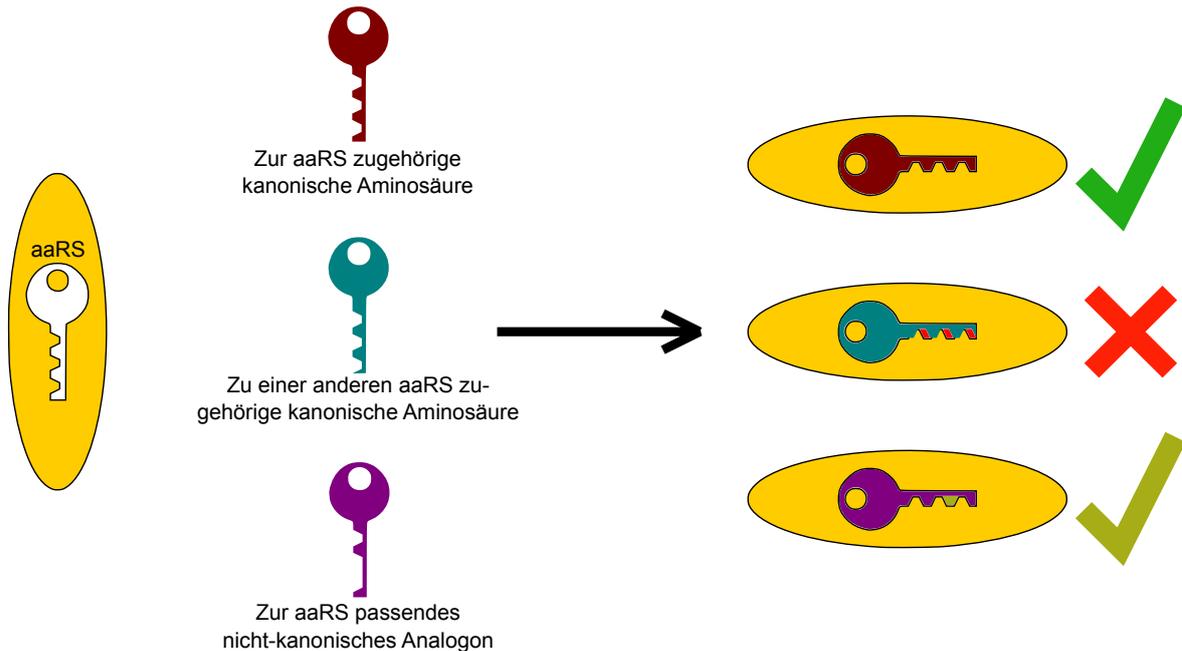


Abbildung 4.1.: Prinzip des seitenkettenspezifischen Einbaus nicht-kanonischer Aminosäuren in Proteine. Eine aaRS kann zwar zwischen den unterschiedlichen kanonischen Aminosäuren unterscheiden (oben und Mitte), besitzt aber eine gewisse Substratflexibilität gegenüber isostrukturellen nicht-kanonischen Analoga, die ihrer zugehörigen kanonischen Aminosäure ähneln (unten) [203], [204].

Konkurrieren eine kanonische Aminosäure und ein nicht-kanonisches Analogon um die Bindung an die kanonische aaRS, präferiert das Enzym im Allgemeinen seine kanonische Aminosäure [222]. Daraus ergeben sich Bedingungen für die effiziente *in-vivo*-Umsetzung des seitenkettenspezifischen Einbaus nicht-kanonischer Analoga in ein Zielprotein [223]: Es wird ein auxotropher¹ Expressionswirt benötigt, der die zu ersetzende kanonische Aminosäure nicht selbst synthetisieren kann; die Aminosäuren-Versorgung des Wirts und dessen Fermentationsbedingungen müssen kontrollierbar sein; es wird ein starkes und induzierbares Expressionssystem gebraucht; die Proteinexpression muss trotz Präsenz des nicht-kanonischen Analogons ablaufen.

Die *in-vivo*-Standardmethode, die diese Bedingungen erfüllt, ist die *selective-pressure-incorporation*-Methode (im Folgenden SPI abgekürzt) [223], [225], [226], die anhand von [223], [225] in diesem Abs.

¹Ein Bakterienstamm, der auxotroph in Bezug auf eine Substanz ist, kann diese Substanz nicht selbst synthetisieren und muss über das Nährmedium mit dieser Substanz versorgt werden [224].

erklärt wird. Das Gen, das für das Zielprotein kodiert, wird unter die Kontrolle eines induzierbaren Promotors gestellt und das resultierende DNA-Konstrukt in den auxotrophen Expressionswirt gebracht. Die Wirtszellen-Kultur wächst in einem minimalen synthetischen Nährmedium, das die zu ersetzende kanonische Aminosäure in limitierter Konzentration enthält, bis die exponentielle Wachstumsphase erreicht ist. Mit Erschöpfen des Vorrats der kanonischen Aminosäure wird dem Wirt das nicht-kanonische Analogon zur Verfügung gestellt und die Expression des Zielgens induziert. Dieser „Selektionsdruck“ zwingt den Wirt, das nicht-kanonische Analogon anstelle der kanonischen Aminosäure an jeder Position im Zielprotein einzubauen.

Dadurch, dass beim seitenkettenspezifischen Einbau meist mehrere nicht-kanonische Aminosäuren auf einmal in das Zielprotein eingebaut werden, können sich dessen Eigenschaften drastisch verändern [227], [228]. Beispielsweise wurde ein cyan-fluoreszierendes Protein durch den seitenkettenspezifischen Einbau von L-4-Aminotryptophan anstelle von Trp in ein gold-fluoreszierendes Protein umgewandelt, das stark rotverschoben emittierte und darüber hinaus stabiler wurde [227]. Durch den seitenkettenspezifischen Einbau nicht-kanonischer Analoga von Met, Pro, Phe und Tyr konnten modifizierte Enzyme erzeugt werden, die beispielsweise im Vergleich zum nativen Enzym höhere Aktivitäten aufwiesen oder andere optimale Temperatur- und pH-Wert-Bereiche besaßen [228].

Der beim SPI-Ansatz erzeugte „Selektionsdruck“ führt auch zum Einbau des nicht-kanonischen Analogons in die endogenen Proteine des Expressionswirts, insbesondere in dessen Enzyme, sodass diese von solchen im letzten Absatz beschriebenen tiefgreifenden Veränderungen ebenso betroffen sein können [40]. Insofern inhibieren die meisten nicht-kanonischen Analoga das weitere Zellwachstum [36], [40]. Folglich unterliegt der *in-vivo*-Ansatz gewissen Beschränkungen. Dadurch ist vor allem der *in-vivo*-Einbau toxischer Aminosäuren sowie solcher nicht-kanonischer Aminosäuren, die starken Einfluss auf die Proteinfaltung haben, besonders schwierig. Allerdings zählen genau derartige Aminosäuren zu den interessantesten Kandidaten, da sie in besonderer Weise durch ihren Einbau die Proteinstruktur beeinflussen und somit außergewöhnliche Proteineigenschaften erzeugen können.

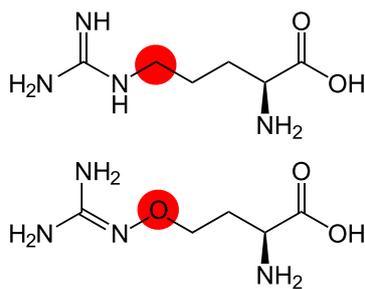


Abbildung 4.2.: Strukturformeln von Arg (oben) und Can (unten).

Ein entsprechender Kandidat ist die nicht-kanonische Aminosäure L-Canavanin (Can, Abb. 4.2), die ein natürlich vorkommendes Arg-Analogon und hauptsächlich in Leguminosen zu finden ist [48]. Can schützt diese Pflanzen vor Insekten und anderen Pflanzenfressern [229], [230] und ist aus medizinischer Sicht eine interessante und vielversprechende Substanz. Can ist aufgrund seiner anti-Krebs-Eigenschaften beispielsweise Gegenstand zahlreicher Studien, u.a. in [44]–[49] und steht im Verdacht, Stimulator für menschliche Autoimmunkrankheiten zu sein [43]. Dieses nicht-kanonische Analogon zeigt darüber hinaus anti-virale Eigenschaften [42].

Die Substitution des C-Atoms durch ein O-Atom in Can (Abb. 4.2) verringert die Elektronendichte in der Guanidin-Gruppe, sodass Can weniger basisch als Arg ist [48]. Während unter physiologischen Bedingungen mehr als 99 % aller Arg-Moleküle protoniert sind, trifft dies nur für etwa 30 % der Can-Moleküle zu [48]. Durch die reduzierte Basizität verlieren Can-modifizierte Proteine meist ihre Konformation, darüber hinaus werden viele regulatorische und katalytische Reaktionen im Arg-Metabolismus durch Can beeinträchtigt, sodass Can toxisch auf viele Organismen wirkt und zum direkten Zelltod führen kann [36], [41], [48], [230], [231].

Die Toxizität von Can macht den *in-vivo*-seitenkettenspezifischen Einbau in Proteine mit *E. coli* oder anderen gewöhnlichen Expressionswirten besonderes schwierig: Bisher wurde der vollständige *in-vivo*-Einbau von Can in ein Zielprotein erst einmal berichtet [50]. Das verwendete System ist ein spezielles Einzel-Protein-Produktions-System, in dem eine RNA-Interferase jede mRNA mit einer kurzen, häufig vorkommenden Erkennungssequenz degradiert [50]. Durch Entfernen der Erkennungssequenz im Gen des Zielproteins kommt es somit zur ausschließlichen Expression des Zielproteins [50].

Die katalytische Komposition der zellfreien Umgebung von zellfreien Proteinsynthesystemen passt sich im Gegensatz zur molekularen Umgebung der lebenden Zellen nicht mehr an äußere Reize an [61]. Somit lassen sich viele mit dem *in-vivo*-Einbau verbundene Probleme durch die zellfreie Proteinsynthese umgehen. Allerdings findet der zellfreie, Extrakt-basierte seitenkettenspezifische Einbau nicht-kanonischer Aminosäuren in Proteine weniger häufig Anwendung als der komplementäre, ortsspezifische Ansatz (vgl. Abschn. 4.1.2.2). Unter anderem gelang der *in-vitro*-seitenkettenspezifische Einbau von L-3,4-Dihydroxyphenylalanin (DOPA) [232], von chlorierten Analoga [233] und verschiedenen nicht-kanonischen Aminosäuren, die sich *in situ* entschützen ließen [234]. Häufig entstehen Proteinspezies mit Teilsubstitutionen, in denen das nicht-kanonische Analogon nicht vollständig in das Zielprotein eingebaut wurde [232]–[234]. Nach spezieller Diafiltration eines Zellextraktes zur Entfernung von kanonischen Aminosäureresten konnte der Einbau mehrerer nicht-kanonischer Analoga in eine GFP-Variante über Fluoreszenzmessungen und Gelelektrophorese gezeigt werden [235]. [236] fand hingegen trotz intensiver Dialyse des verwendeten Zellextraktes weiterhin Trp-Reste im ZFPS-System vor. Die Verwendung eines Trp-auxotrophen *E.-coli*-Stammes zur Gewinnung des Extraktes und weitere Modifikationen bei der Herstellung des Zellsats ermöglichten letztendlich den vollständigen Trp-Analogon-Einbau in ein Protein [236].

Wichtige Beiträge zum seitenkettenspezifischen Einbau nicht-kanonischer Aminosäuren lieferte die Gruppe um SZOSTAK. Es wurde ein einfaches Verfahren entwickelt, indem mittels MALDI-TOF-Massenspektrometrie systematisch über 190 nicht-kanonische Analoga als potentielle Substrate für die kanonischen aaRSs untersucht wurden [237]. Etwas mehr als 90 nicht-kanonische Analoga bewährten sich in dieser Studie [237] als Substrate für die kanonischen aaRSs. Beim seitenkettenspezifischen Einbau nicht-kanonischer Analoga in Proteine verfolgten SZOSTAK und Mitarbeiter eine alternative Strategie zum Extrakt-basierten Ansatz. Durch Verwendung des artifiziellen, aus rekombinanten Komponenten zusammengesetzten PURE-Systems (vgl. Abschn. 2.4.3.3) gelang es, bis zu 13 nicht-kanonische Analoga in ein kurzes Modellpeptid einzubauen [238], [239]. Das PURE-System wurde von der SZOSTAK-Gruppe, z.B. in [240]–[242], verwendet, um kurze Peptide gezielt durch den Einbau nicht-kanonischer Aminosäuren *in vitro* zu evolvieren, ein Ansatz, den auch andere Gruppen verfolgten [243].

4.1.2.2. Ortsspezifischer Einbau nicht-kanonischer Aminosäuren in Proteine

Der ortsspezifische Einbau nicht-kanonischer Aminosäuren in Proteine ist komplementär zum seitenkettenspezifischen Ansatz [36]. Das grundlegende Prinzip dieser Einbaustrategie wird in diesem Abs. anhand von [244], [245] erklärt. Die endogenen kanonischen aaRS/tRNA-Paare des Expressionswirts werden um ein zusätzliches, sogenanntes orthogonales exogenes aaRS/tRNA-Paar (o-Paar) ergänzt. Ebenso wird der kanonische Aminosäuren-Satz um die einzubauende nicht-kanonische Aminosäure erweitert. Für den erfolgreichen ortsspezifischen Einbau einer nicht-kanonischen Aminosäure müssen mehrere Bedingungen erfüllt sein: Die nicht-kanonische Aminosäure darf von keiner endogenen

kanonischen aaRS, sondern nur von der orthogonalen aaRS (o-aaRS) erkannt, aktiviert und auf die orthogonale tRNA (o-tRNA) geladen werden. Zusätzlich darf die o-aaRS weder eine der 20 kanonischen Aminosäuren erkennen, noch die nicht-kanonische Aminosäure auf eine endogene, kanonische tRNA laden. Darüber hinaus muss die nicht-kanonische Aminosäure im Expressionswirt oder der zellfreien Umgebung metabolisch stabil sein. Des Weiteren muss sie vom Elongationsfaktor EF-Tu und vom Ribosom akzeptiert werden. Sind diese Bedingungen erfüllt, wird darüber hinaus noch ein ungenutztes Codon benötigt, zu dem das Anticodon der o-tRNA komplementär sein muss. Meist ist dies ein Stopcodon, oft das amber-Stopcodon. Durch Insertion der Stopcodon-Sequenz an der gewünschten Stelle des Gens des zu exprimierenden Zielproteins kann bei der Translation die mit der nicht-kanonischen Aminosäure beladene o-tRNA das Stopcodon supprimieren und die nicht-kanonische Aminosäure an dieser Position ortsspezifisch in das Zielprotein inkorporieren. Neben der Stopcodon-Suppression werden auch Vier-Buchstaben-Codons von Vier-Buchstaben-Suppressor-tRNAs supprimiert, beispielsweise in [246].

Durch den ortsspezifischen Einbau einer nicht-kanonischen Aminosäure an einer bestimmten Position in einem Protein lassen sich dessen Struktur und Eigenschaften häufig eher lokal als global verändern [208], was im Folgenden beispielhaft [247], [248] demonstriert wird. Als Reaktion auf Endotoxin (Lipopolysaccharid-Bestandteil von Zellwänden gramnegativer Bakterien), Bestandteile grampositiver Bakterien oder anderer Mikroorganismen, synthetisiert das Immunsystem vermehrt den Tumornekrosefaktor- α (TNF- α), der als primärer Vermittler in der Pathogenese einer Sepsis gilt [249]. Ein mutierter TNF- α , dem die nicht-kanonische Aminosäure p-Nitrophenylalanin an der Position 11 oder 86 ortsspezifisch eingebaut wurde, führte bei Mäusen zu einer Immunreaktion, die sich nicht nur gegen den mutierten, sondern auch gegen den nativen, im Überschuss produzierten TNF- α richtete und die Mäuse somit vor dem Tod durch eine Lipopolysaccharid-induzierte Sepsis bewahrte [247], [248]. Eine reine Substitution des Tyr an der Stelle 86 durch die kanonische Aminosäure Phe schützte hingegen nicht gegen die Sepsis [247]. Die minimale, lokale Veränderung in der Proteinstruktur des TNF- α führte in diesem Beispiel zu einer Immunantwort, die sich aber auch auf den nativen TNF- α übertug, da beide Proteine ansonsten gleiche Strukturen besaßen und somit von den Antikörpern erkannt werden konnten.

Die größte Schwierigkeit beim ortsspezifischen Ansatz ist es, ein zum kanonischen Satz der aaRSs und tRNAs orthogonales aaRS/tRNA-Paar zu finden, das darüber hinaus eine hohe Spezifität gegenüber der einzubauenden nicht-kanonischen Aminosäure besitzt und von den Ribosomen inklusive seiner an der Translation beteiligten Faktoren möglichst gut akzeptiert wird [244]. Eine erfolgreiche Strategie besteht darin, auf aaRS/tRNA-Paare aus anderen Domänen des Lebens zurückzugreifen und diese dann gegebenenfalls durch molekulare Evolution zu optimieren [244]. Inzwischen gibt es eine Vielzahl an effizienten o-Paaren für den ortsspezifischen Einbau von nicht-kanonischen Aminosäuren in Proteine mit Hilfe von *E. coli* und anderen üblichen Expressionswirten, über die beispielsweise [244] einen Überblick gibt. In *E. coli* finden vor allem zwei o-Paare Anwendung [250]: das *Methanococcus jannaschii* TyrRS/tRNA_{CUA}^{Tyr}- und das *Methanosarcina barkeri*/*Methanosarcina mazei* PylRS/tRNA_{CUA}^{Pyl}-Paar. Die Tyrosyl-tRNA-Synthetase (TyrRS) aus *Methanocaldococcus jannaschii* besitzt keinen Korrekturlesemechanismus und wechselwirkt bei der tRNA-Erkennung nur minimal mit der Anticodonschleife ihrer zugehörigen tRNA^{Tyr}, weswegen die tRNA^{Tyr} zur amber-Suppressor-tRNA_{CUA}^{Tyr} mutiert werden kann, ohne dass die Beladungseffizienz signifikant abnimmt [251].

Im Folgenden wird kurz auf das in der vorliegenden Arbeit für den ortsspezifischen Einbau verwendete Pyrrolsilyl-tRNA-Synthetase/tRNA_{CUA}^{Pyl}-Paar (PylRS/tRNA_{CUA}^{Pyl}-Paar), das zum Einbau der nicht-kanonischen Aminosäure N^ε(*tert*-butyloxycarbonyl)-L-Lysin (BocLys, Abb. 4.3) in ein Zielprotein

benutzt wird, eingegangen. In *Methanosarcina barkeri* ist der sonst übliche kanonische Satz an Aminosäuren um das Lys-Derivat L-Pyrrolysin (Pyl, Abb. 4.3) erweitert, das auf natürlichem Weg durch eine amber-Stopcodon-Suppression in endogene Proteine eingebaut werden kann [33]. Pyl ist das natürliche Substrat für die PylRS, die Pyl spezifisch auf die $tRNA_{CUA}^{Pyl}$ lädt [252]. Darüber hinaus ist das PylRS/ $tRNA_{CUA}^{Pyl}$ -Paar orthogonal zur endogenen *E.-coli*-Translationsmaschinerie [253]. Außerdem können kostengünstigere, kommerziell erhältliche Pyl-Derivate effizient von der PylRS erkannt, aktiviert und auf die $tRNA_{CUA}^{Pyl}$ geladen werden [253]. Mit der mutierten Form der PylRS (Tyr384Phe) aus *Methanosarcina mazei* konnte sowohl die Effizienz der *in-vitro*-Aminoacylierung als auch die der *in-vivo*-Stopcodon-Suppression mit einer mit Pyl oder BocLys beladenen $tRNA_{CUA}^{Pyl}$ gesteigert werden [254].

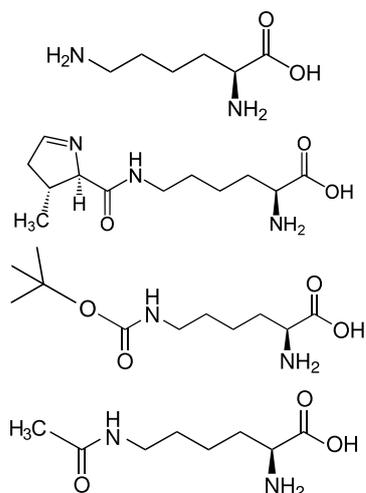


Abbildung 4.3.: Strukturformeln von oben nach unten von Lys, Pyl, BocLys und AcLys.

Im Allgemeinen konkurrieren bei der Stopcodon-Suppression die beladene Suppressor-tRNA und die Terminationsfaktoren um die jeweilige Bindung auf dem Stopcodon bzw. auf der P-Stelle des Ribosoms [245], [250]. Bindet der Terminationsfaktor 1 an der P-Stelle im Ribosom, bevor die amber-Suppressor-tRNA das Stopcodon supprimiert, wird die Translation terminiert, und es entsteht nur ein unvollständiges Protein (vgl. Abschn. 2.2.2.4) [245]. Ein Ansatz, dieses Problem zu lösen, besteht darin, das *E.-coli*-Genom gezielt so zu mutieren, dass das *prfA*-Gen, das für den Terminationsfaktor 1 kodiert, ausgeknockt werden kann, ohne dass das Bakterium stirbt [245], [255]–[257]. Es gibt weitere Ansätze, die Suppressor-Effizienz zu erhöhen, über die beispielsweise [245], [250] einen Überblick geben.

Vor allem die Gruppe um SWARTZ hat mit ihrem gekoppelten zellfreien T7-Transkriptions-Translations-System Strategien entwickelt, die den ortsspezifischen Einbau nicht-kanonischer Aminosäuren in Proteine sehr effizient machten. In anfänglichen Versuchen wurden die Gene des o-Paars in den *E.-coli*-Zellen, die zur Extrakt-Präparation verwendet wurden, exprimiert, sodass nur noch die einzubauende nicht-kanonische Aminosäure in die zellfreie Reaktion gegeben werden musste [258]. Die OmpT-Protease-Aktivität, die sowohl die T7-RNA-Polymerase als auch die orthogonale aaRS degradierte, reduzierte die Proteinausbeute jedoch drastisch (teilweise $< 35 \mu\text{g/ml}$) [258]. Die Verwendung optimierter *E.-coli*-Stämme zur Extrakt-Gewinnung, in denen nur noch das *o-trNA*-Gen exprimiert wurde, und durch die spätere Hinzugabe der aufgereinigten o-aaRS sowie der nicht-kanonischen Aminosäure, konnten deutlich höhere Proteinausbeuten erzielt werden, da die *E.-coli*-Zellen durch die Separation des o-Paares weniger Stress während des Wachstums ausgesetzt waren [258], [259]. Vielversprechend ist vor allem der Transzym-Ansatz [260]. Die Stopcodon-Suppression verläuft dort so effizient, dass mit nicht-kanonischen Aminosäuren modifizierte Modellproteine mit einer Endkonzentration von $0,9 \text{ mg/ml}$ – $1,7 \text{ mg/ml}$ zellfrei synthetisiert werden können [260]. Bei diesem Ansatz wird die o-tRNA als sogenanntes Transzym kosynthetisiert, während die o-aaRS als aufgereinigtes Enzym in die zellfreie Reaktion gegeben wird [260]. Bei der für das Transzym kodierenden DNA grenzt die Sequenz des sich autokatalytisch spleißenden Hammer-Head-Ribozyms an die Sequenz der o-tRNA [260]. Nach der Transkription durch die T7-RNA-Polymerase trennt sich das Hammer-Head-Ribozym selbst aus dem mRNA-Transkript heraus und setzt somit die aktive o-tRNA in der zellfreien Reaktion frei [260]. Wird darüber hinaus der Transzym-Ansatz in ein ZFPS integriert, das auf einem Zellextrakt basiert, der aus einem *E.-coli*-Stamm erzeugt wird, bei dem der Terminationsfaktor 1 ausgeknockt ist, lassen

sich effizient mehrere Amber-Stopcodons supprimieren [261].

Auch beim ortsspezifischen Ansatz spielt das PURE-System eine Rolle [139] (vgl. Abschn. 2.4.3.3). In [139] wurde ein Stopcodon mittels einer chemisch synthetisierten UAG-spezifischen, mit Val beladenen Suppressor-tRNA effizient überschrieben, indem der Terminationsfaktor 1 dem System entzogen wurde. Ein Vergleich [261] zwischen dem PURE-System und einem Extrakt-basierten ZFPS, in dem der Terminationsfaktor 1 ebenfalls fehlt, zeigt, dass das PURE-System bei höheren Kosten nur ca. ein Fünftel der Proteinausbeute liefert.

4.1.3. Zielsetzung und Strategie

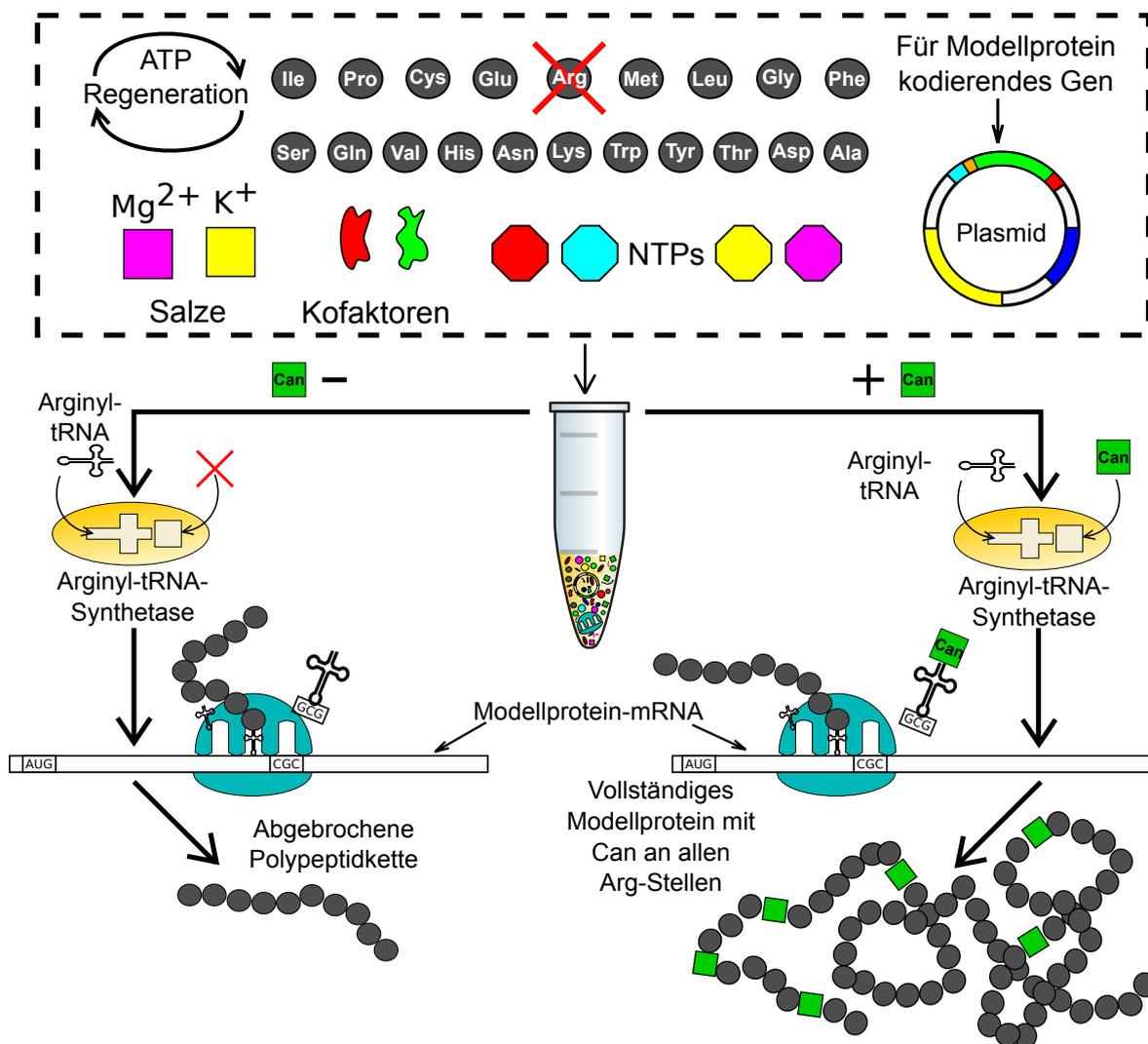


Abbildung 4.4.: Prinzip des zellfreien seitenspezifischen Einbaus nicht-kanonischer Aminosäuren in ein Modellprotein am Beispiel des nicht-kanonischen Arg-Analogons Can. Abb. 4.4 ist eine modifizierte Darstellung des graphischen Abstraktes aus [193].

Die Bakteriophagen-Transkription bildet in der zellfreien Proteinsynthese den heutigen Standard (vgl. Abschn. 2.4.3.4) [61], [106]. Solche zellfreien Systeme generieren zwar hohe Proteinausbeuten, eignen sich jedoch weniger, *E.-coli*-endogene Regulationsmechanismen zu studieren [52]. Das verwendete zellfreie Expressionssystem (vgl. Abschn. 2.4.3.4) bietet einen alternativen Ansatz, indem es auf

die endogene *E.-coli*-Transkriptions-Translations-Maschinerie zurückgreift und dennoch vergleichbar hohe Proteinausbeuten wie durch die Nutzung der Bakteriophagen-Transkription erzeugt [51], [52]. Die native Komposition des zellfreien Systems, das sich aus der molekularen *E.-coli*-Transkriptions-Translations-Maschinerie zusammensetzt, ermöglicht es, komplexe genetische Schaltkreise in dieses zellfreie System zu implementieren [262]. Dies motiviert die in Kap. 4 verfolgte Strategie, jenes zellfreie System bei einer komplexen biochemischen Prozedur, wie dem Einbau nicht-kanonischer Aminosäuren, insbesondere dem Einbau von Can, in Proteine zu nutzen.

Zusätzlich zum seitenkettenspezifischen Einbau von Can, wurde der seitenkettenspezifische Einbau des Lys-Derivats L-Hydroxylysin (Hyl) und des Met-Derivats L-Azidohomoalanin (Aha) untersucht. Als Modell- oder Zielprotein diente deGFP (vgl. Abschn. 2.5), das am C-Terminus mit einem His-Tag ausgestattet war. Abb. 4.4 stellt das in dieser Arbeit angewendete Prinzip bzw. die experimentelle Realisierung des zellfreien seitenkettenspezifischen Einbaus nicht-kanonischer Aminosäuren in Modellproteine schematisch am Beispiel des nicht-kanonischen Arg-Analogons Can dar. Der *E.-coli*-Extrakt liefert die endogene *E.-coli*-Transkriptions-Translations-Maschinerie, die mit weiteren Kofaktoren, Salzen, verschiedenen Biomolekülen, dem 3-PGA-Regenerationssystem, der für das Modellprotein kodierenden Plasmid-DNA, Ribonukleotidtriphosphate und den Aminosäuren ergänzt werden muss [51], [52]. Bei Substitution von Arg durch Can (rechter Pfad in Abb. 4.4) sollte Can aufgrund seiner strukturellen Ähnlichkeit zu Arg von der Arginyl-tRNA-Synthetase (ArgRS) erkannt, aktiviert und auf die tRNA^{Arg} geladen werden. In diesem Fall wird ein Modellprotein synthetisiert, in dem das Analogon die kanonische Aminosäure an jeder Position substituiert. Im Fall, dass die nicht-kanonische Aminosäure nicht substituiert wird, sollte die Translation beim ersten Codon, das für die kanonische Aminosäure kodiert, abbrechen (linker Pfad in Abb. 4.4). Beim seitenkettenspezifischen Einbau von Hyl und Aha wurde entsprechend diesem Schema vorgegangen. In diesen Fällen wurde Lys durch Hyl bzw. Met durch Aha substituiert. Der seitenkettenspezifische Einbau der nicht-kanonischen Aminosäure in das Modellprotein wurde per *sodium dodecyl sulfate polyacrylamide gel electrophoresis* (SDS-PAGE) [263] und HPLC-ESI-Massenspektrometrie nach vorangegangener His-Tag-Purifikation analysiert.

Neben dem seitenkettenspezifischen Ansatz wurde der ortsspezifische Einbau nicht-kanonischer Aminosäuren in Modellproteine durch amber-Stopcodon-Suppression untersucht. Als Modellprotein diente EGFP^{N149amber} (vgl. Abschn. 2.5), das am C-Terminus mit einem His-Tag ausgestattet war. Zur Stopcodon-Suppression wurde das in Abschn. 4.1.2.2 beschriebene PylRS^{Y384F}/tRNA_{CUA}^{Pyl}-Paar aus *Methanosarcina mazei* verwendet, das orthogonal zum endogenen Satz der *E.-coli*-aaRSs und -tRNAs ist [253]. Als einzubauende nicht-kanonische Aminosäure wurde BocLys (Abb. 4.3) gewählt, die sich bereits als Pyl-Derivat und Substrat für die PylRS bewiesen hat [254]. In Vorbereitungsexperimenten wurde der ortsspezifische Einbau einer mit der nicht-kanonischen Aminosäure N_ε-Acetyllysin (AcLys, Abb. 4.3) aminoacylierten amber-Stopcodon-Suppressor-tRNA (AcLys-tRNA^{amber}) analysiert. Abb. 4.6 stellt das in dieser Arbeit verwendete Prinzip anhand des ortsspezifischen Einbaus der nicht-kanonischen Aminosäure BocLys in das Modellprotein schematisch dar.

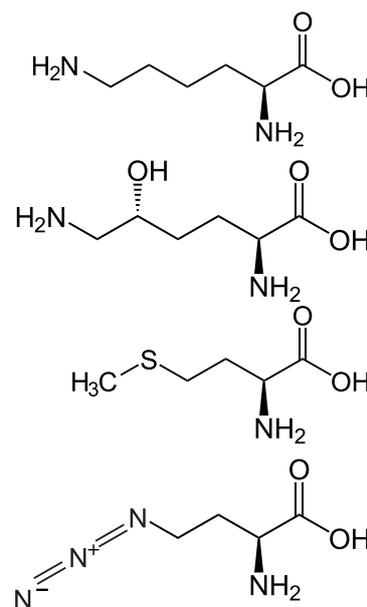


Abbildung 4.5.: Strukturformeln von Lys (oben), Hyl (2. von oben) und Met (3. von oben) und Aha (unten).

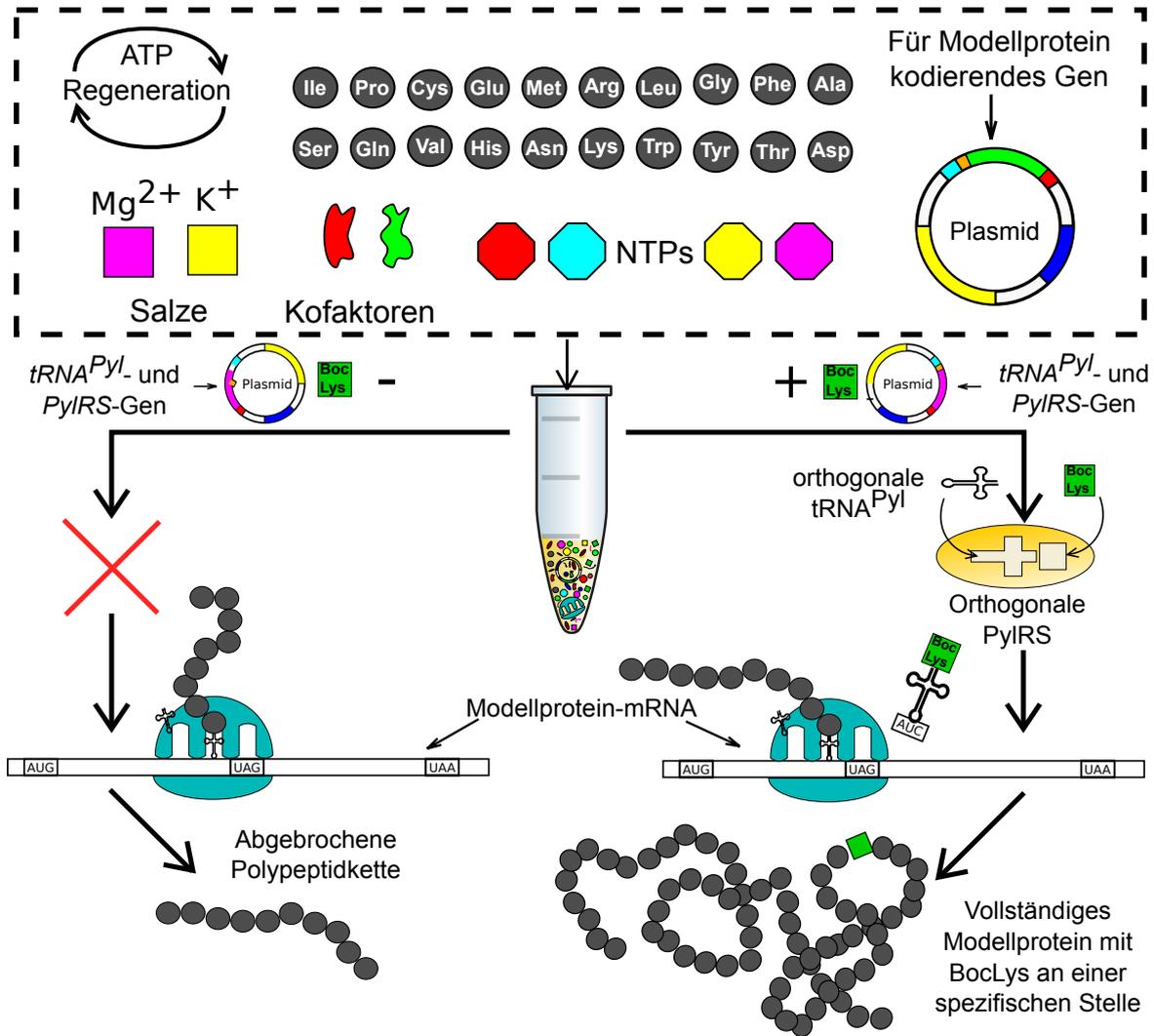


Abbildung 4.6.: Prinzip des zellfreien ortsspezifischen Einbaus der nicht-kanonischen Aminosäure BocLys in ein Modellprotein mittels Stopcodon-Suppression. Zur Erstellung von Abb. 4.6 wurden Elemente des graphischen Abstraktes aus [193] verwendet. In Anlehnung an die Idee dieses graphischen Abstraktes aus [193] erstellt.

Der experimentelle Ansatz für den ortsspezifischen Einbau einer nicht-kanonischen Aminosäure in ein Modellprotein unterscheidet sich vom seitenkettenspezifischen Ansatz dadurch, dass keine der 20 kanonischen Aminosäuren ersetzt wird, sondern der Satz der kanonischen Aminosäuren um die nicht-kanonische Aminosäure erweitert werden muss (rechter Pfad in Abb. 4.6). Um das Stopcodon zu suppressieren, müssen zusätzlich die für die $\text{PylRS}^{\text{Y384F}}$ und die $\text{tRNA}_{\text{CUA}}^{\text{Pyl}}$ kodierende DNA in das zellfreie System integriert werden. Neben dem in Abb. 4.6 dargestellten Ansatz sowohl das Gen, das für die $\text{PylRS}^{\text{Y384F}}$ kodiert, als auch jenes Gen, das für die $\text{tRNA}_{\text{CUA}}^{\text{Pyl}}$ kodiert, auf demselben Plasmid unter die Kontrolle desselben Promotors zu stellen, können beide Gene auch getrennt voneinander exprimiert oder aufgereinigte o-Paar-Moleküle dem zellfreien System zur Verfügung gestellt werden. Stehen dem zellfreien System sowohl BocLys als auch das o-Paar zur Verfügung, sollte die $\text{PylRS}^{\text{Y384F}}$ die $\text{tRNA}_{\text{CUA}}^{\text{Pyl}}$ mit BocLys beladen, das Stopcodon suppressieren und BocLys an die Position 149 in EGFP platziert werden. Die Position N149 (bzw. N150, sofern das abgespaltene Met mitgezählt wird) an der EGFP-Oberfläche im Übergang eines β -Faltblatts zu einer Schleife stellt eine geeignete Position zur Platzierung eines Stopcodons dar, da der Einbau einer nicht-kanonischen Ami-

nosäure die globale Proteinstruktur wenig beeinflusst [152]–[154]. Insofern eignet sich EGFP^{N149amber} als Reporter-Molekül in Stopcodon-Experimenten, da sich das Fluorophor nur bei erfolgter Stopcodon-Suppression bildet und in diesem Fall das Protein fluoresziert [152]–[154]. Fehlt nur eines dieser drei Moleküle (linker Pfad in Abb. 4.6), bricht die Translation beim Stopcodon an Position 149 ab. In diesem Fall fluoresziert das unvollständige Protein nicht, da sich das EGFP-Fluorophor nicht bildet [152]–[154]. Der ortsspezifische Einbau von BocLys wurde mittels Messung der Fluoreszenz-Intensität analysiert. Basierend auf dem Ansatz [51], [52] eignet sich die Messung der Fluoreszenz-Intensität, die Proteinausbeute qualitativ abzuschätzen oder zu quantifizieren, weil die Fluoreszenz-Intensität mit der Menge des zellfrei synthetisierten fluoreszierenden Moleküls ansteigt.

In Bezug auf den ortsspezifischen Einbau nicht-kanonischer Aminosäuren in Proteine, sollte herausgefunden werden, ob das verwendete zellfreie System die Möglichkeit bietet, das verwendete o-Paar neben dem zu modifizierenden Modellprotein in derselben zellfreien Reaktion zu kotsynthetisieren und gleichzeitig eine hohe Ausbeute des Zielproteins zu erzielen. Ein solcher zellfreier Ansatz würde eine flexible Plattform darstellen, um verschiedene nicht-kanonische Aminosäuren ortsspezifisch in Proteine zu inkorporieren, da nur die für das o-Paar und Modellprotein kodierende DNA in die zellfreie Reaktion gegeben werden muss².

²Zu Beginn dieser Studie war das verwendete zellfreie Expressionssystem zur Nutzung für den Einbau nicht-kanonischer Aminosäuren in Proteine unerforscht. Allerdings wurde während der Anfertigung der hier präsentierten Studie der ortsspezifische Einbau der nicht-kanonischen Aminosäure BocLys und weiterer Pyl-Derivate mit dem PylRS/tRNA_{CUA}^{Pyl}-Paar von der Gruppe um ALFONTA in einer modifizierten Variante des hier verwendeten zellfreien Systems realisiert [264]. Der dort gewählte Ansatz unterscheidet sich hinsichtlich der in dieser Arbeit gewählten Strategie, alle für den ortsspezifischen Einbau einer nicht-kanonischen Aminosäure notwendigen Komponenten in derselben zellfreien Reaktion zu synthetisieren.

4.2. Material und Methoden

Alle wässrigen Lösungen wurden mit Reinstwasser (*arium*[®] 611UV Reinstwasseranlage, *Sartorius Stedim Biotech GmbH*, Göttingen), das autoklaviert (Varioklav[®] Dampfsterilisator, Typ 75 S, *H+P Labortechnik AG*, Oberschleißheim) wurde, hergestellt. Alle Materialien, die in Kontakt mit biologischem Material kamen, waren steril oder wurden zuvor autoklaviert (Varioklav[®] Dampfsterilisator, Typ 75 S, *H+P Labortechnik AG*, Oberschleißheim). DNA-, Enzym- oder Protein-Lösungen wurden auf Eis zubereitet. Die verschiedenen Ingredienzien der Lösungen wurden durch Auf- und Abpipettieren vermischt, kurz (2s-3s) auf mittlerer Drehzahl gevortext (*Vortex 3*, *IKA*[®]-*Werke GmbH*, Staufen) und mittels einer Tisch-Minizentrifuge für etwa 2s-3s abzentrifugiert (*Sprout Minizentrifuge*, *Biozym Scientific GmbH*, Hessisch Oldendorf). Alle Enzyme, deren zugehörige Reaktionspuffer und Komponenten als auch DNA-Lösungen wurden bei -20 °C gelagert. Zur Präparation einer Reaktion wurden die benötigten Komponenten auf Eis aufgetaut. Die Präparation der Reaktionen wurde ebenfalls auf Eis durchgeführt.

4.2.1. Verwendete Chemikalien

4.2.1.1. Standard-Chemikalien

Alle Standard-Chemikalien wurden über das *Referat Zentrale Beschaffung und Dienste - Sachgebiet 2 Gefahrstoffe und Laborbedarf* (ehemals *Zentrales Chemikalienlager*) der Universität des Saarlandes beschafft.

4.2.1.2. Nicht-kanonische Aminosäuren und aminoacylierte Amber-Stopcodon-Suppressor-tRNA

Die folgenden nicht-kanonischen Aminosäuren stellte die Kooperationsarbeitsgruppe von Nediljko Budisa (Institut für Chemie, Technische Universität Berlin) zur Verfügung: L-Canavanin (Can, Abb. 4.2), L-Hydroxylysin (Hyl, Abb. 4.5), L-Azidohomoalanin (Aha, Abb. 4.5), N^ε(*tert*-butyloxycarbonyl)-L-Lysin (BocLys, Abb. 4.3). BocLys wurde zusätzlich von *Carbolution Chemicals GmbH* (St. Ingbert) bezogen. Die mit der nicht-kanonischen Aminosäure N_ε-Acetyllysin (AcLys, Abb. 4.3) aminoacylierte amber-Stopcodon-Suppressor-tRNA (AcLys-tRNA^{amber}) wurde von *tRNA Probes LLC* (College Station, USA) geliefert.

4.2.2. Nährmedien und Antibiotika-Selektion

Wenn nicht anders angegeben, wurden *E.-coli*-Zellen in flüssigem LB-Medium (*Lysogeny Broth*) oder auf LB-Agar-Platten kultiviert. Ampicillin diente der Selektion. Das LB-Medium geht auf BERTANI zurück [265]. Das verwendete LB-Flüssigmedium setzte sich aus 10 g/l Trypton, 5 g/l Hefeextrakt und 5 g/l NaCl zusammen. Zur Herstellung von LB-Agarplatten wurde LB-Flüssigmedium mit zusätzlichen 15 g/l Agar verwendet. Nach dem Autoklavieren wurde das LB-Agar-Medium auf etwa 50 °C abgekühlt, bevor Ampicillin hinzupipettiert wurde, sodass das Antibiotikum eine Konzentration von 100 µg/ml im Medium besaß. Die gegossenen LB-Agarplatten wurden mit Deckeln abgedeckt, mit Parafilm abgedichtet und bei 4 °C gelagert. Das flüssige LB-Medium wurde ebenfalls bei 4 °C gelagert,

allerdings wurde Ampicillin erst kurz vor dem Ansetzen einer Bakterienkultur hinzugegeben, sodass im Flüssigmedium eine Konzentration von 50 µg/ml herrschte.

Ausschließlich zur Kultivierung der zur zellfreien Extrakt-Herstellung verwendeten *E.-coli*-Zellen wurde 2xYT bzw. 2xYTP-Medium verwendet, zur Antibiotika-Selektion diente dabei Chloramphenicol [52]. Das flüssige und feste 2xYTP-Medium setzte sich, wie in [52] angegeben, folgendermaßen zusammen: 31 g/l 2xYT, 40 mM Dikaliumhydrogenphosphat und 22 mM Kaliumdihydrogenphosphat. Zur Herstellung von 2xYTP-Agarplatten wurden diesem Flüssigmedium zusätzlich 15 g/l Agar hinzugegeben [52]. Nach dem Autoklavieren wurde das sterile 2xYTP-Agar-Medium auf etwa 50 °C abkühlt, bevor Chloramphenicol hinzupipettiert wurde, sodass das Antibiotikum eine Konzentration von 34 µg/ml im Medium besaß [52].

4.2.3. *E.-coli*-Stämme

Zur Plasmid-Amplifizierung wurde der *E.-coli*-Stamm KL740 (Coli Genetic Stock Center (Yale), CGSG #: 4382, [266]) verwendet, der von der Kooperationsarbeitsgruppe von Vincent Noireaux (*School of Physics and Astronomy, University of Minnesota*) bereitgestellt wurde. Der Genotyp ist Hfr(PO1), *LAMcI857(ts)*, *relA1*, *spoT1*, *thiE1*. Relevant für die Plasmid-Amplifikation war die Expression des Gens des thermosensitiven λ Repressors CI857, der die Expressionsrate des unter der Kontrolle des starken, an einer Stelle mutierten λ-Repressor-Cro-Promotors (GenBank ID: J02459.1) (vgl. Abschn. 4.2.9.1) stehenden Gens während der Plasmid-Amplifikation (vgl. Abschn. 4.2.9) senkt [51].

Zur Extrakt-Präparation wurde der *E.-coli*-Stamm Rosetta™ 2 (*Novagen, Merck KGaA*, Darmstadt) verwendet [51], [52]. Der Genotyp dieses Stamms lautet F⁻, *ompT*, *hsdS_B*(r_B m_B), *gal*, *dcm*. Dieser Stamm trägt das Plasmid pRARE2, auf dem sich Gene für tRNAs mit selten in *E. coli* verwendeten Codons befinden und das Chloramphenicol-Resistenz vermittelt.

Alle *E.-coli*-Stämme wurden bei -80 °C gelagert.

4.2.4. Agarosegelelektrophorese

Mittels nativer horizontaler Agarosegelelektrophorese wurden doppelsträngige DNA-Stränge gemäß ihrer Länge unter einfachen Tris-Borat-EDTA-Pufferbedingungen (89 mM Tris, 89 mM Borsäure, 2 mM EDTA bei pH 8,3, autoklaviert) aufgetrennt. In der Regel wurden 0,8 %-ige Agarosegele verwendet, die mit UltraPure™ Agarose (*Fisher Scientific GmbH*, Schwerte) hergestellt wurden. Zur Auftrennung von DNA-Strängen geringerer Längen (< 200 bp) wurde *Biozym Sieve Genetic Pure Agarose* (*Biozym Scientific GmbH*, Hessisch Oldendorf) verwendet. In Abhängigkeit der Länge der aufzutrennenden DNA-Stränge betrug in diesen Fällen die Agarosekonzentration zwischen 4 %-5 %. Die Agarosegele enthielten bereits Ethidiumbromid (EthBr), das der flüssigen Agarose-1x-TBE-Lösung mit einer Konzentration von 0,5 µg/ml-1µg/ml hinzugegeben wurde.

Die DNA-Lösungen, die die aufzutrennenden DNA-Stränge enthielten, wurden mit 6x konzentriertem Probenpuffer (*Fisher Scientific GmbH*, Schwerte) vor der Gelelektrophorese vermischt. Der Probenpuffer setzte sich wie folgt zusammen: 10 mM Tris-HCl (pH 7.6), 0,15 % Orange G, 0,03 % Xylencyanol FF, 60 % Glycerin und 60 mM EDTA. Das hinzupipettierte Volumen des 6x Probenpuffers betrug stets ein Sechstel des Gesamtvolumens aus DNA-Lösung und Probenpuffer. Als DNA-Standards wurden die *O'RangeRuler 10 bp DNA Ladder* (verwendungsbereit, *Thermo Fisher Scientific*, Waltham, USA),

2-Log DNA Ladder (verwendungsbereit oder noch mit Reinstwasser und 6x Ladepuffer zu vermischen, *New England Biolabs GmbH*, Frankfurt am Main), der *DNA Molecular Weight Marker VII* und *XVII* (*F. Hoffmann-La Roche AG*, Basel, Schweiz) verwendet. Typischerweise dauerte die Gelelektrophorese zwischen 30 min-40 min bei einer angelegten Spannung von 140 V. Die fluoreszierende EthBr-gefärbte DNA wurde mit dem *GelVue UV Transilluminator* sichtbar gemacht (*Syngene Europe*, Cambridge, UK).

4.2.5. Polymerase chain reaction (PCR)

Die ursprüngliche Idee der *Polymerase chain reaction* (PCR) geht auf [267] zurück. Die Entwicklung der modernen Form der PCR wird MULLIS [268]–[270] zugeschrieben. Der Rest dieses Absatzes führt in das Prinzip der PCR durch Zusammenfassen der relevanten Ausführungen aus Abschn. 12.2.5 aus [62] und Kap. 4 aus [69] ein. Die PCR ist eine Standardmethode zur basengenauen Vervielfältigung von DNA-Strängen beliebiger Sequenzen. Die Komponenten einer typischen PCR-Reaktion bilden ein DNA-Template, eine thermostabile DNA-Polymerase, ein Vorwärts- und Rückwärtsprimer, die vier Desoxyribonukleosidtriphosphate dATP, dCTP, dGTP und dTTP und ein geeignetes Puffersystem. Bei einer PCR wird in zyklischer Weise die Temperatur variiert. Ein solcher PCR-Zyklus besteht aus drei Teilschritten bei unterschiedlichen Temperaturen: Denaturierung der Doppelstränge zu Beginn des Zyklus zur Minimierung der Produktinhibition, Primer-Hybridisierung oder Annealing auf dem Template, damit die DNA-Polymerase am freien 3'-OH-Ende des Primers die Strang-Polymerisierung initiieren kann und Elongation oder Polymerisation bei optimaler Enzym-Temperatur beginnend mit dem Primer entlang des DNA-Templates. Die Annealing-Temperatur ist sowohl von den Längen und Sequenzen der Primer als auch vom Puffersystem abhängig.

Die PCR-amplifizierte DNA diente als Template-DNA bei der zellfreien Genexpression oder wurde in weiteren Klonierungsschritten verwendet. Wenn nicht anders angegeben, wurden Standard-PCRs entweder mit der Taq-DNA-Polymerase oder der Q5[®]-*High-Fidelity*-DNA-Polymerase (beide *New England Biolabs GmbH*, Frankfurt am Main) durchgeführt. Als Reaktionspuffer bei einer PCR mit der Taq-DNA-Polymerase dienten entweder der 10x Standard-Taq-Reaktionspuffer oder der 10x ThermoPol[®]-Reaktionspuffer (beide *New England Biolabs GmbH*, Frankfurt am Main). Als Reaktionspuffer bei einer PCR mit der Q5[®]-*High-Fidelity*-DNA-Polymerase wurde der 5x Q5[®]-Reaktionspuffer verwendet (*New England Biolabs GmbH*, Frankfurt am Main). Während der Strang-Polymerisierung wurden die DNA-Monomere mittels des Desoxyribonukleosidtriphosphat-Lösung-Mixes (dNTP-Lösungs-Mix) (*New England Biolabs GmbH*, Frankfurt am Main) bereitgestellt. Als Template-DNA diente Plasmid-DNA oder lineare DNA. Alle DNA-Primer wurden von *Metabion international AG* (Martinsried) synthetisiert und entsalzt geliefert.

Kleine PCR-Ansätze mit Volumina bis 150 µl wurden direkt in verschließbaren DNase- und RNase-freien 0,2 ml-PCR-Reaktionsgefäßen (*Biozym Scientific GmbH*, Hessisch Oldendorf) zubereitet. Größere Ansätze wurden zuerst in DNase- und RNase-freien 1,5 ml-Reaktionsgefäßen (*Eppendorf AG*, Hamburg) präpariert und dann in gleiche Volumina auf PCR-Reaktionsgefäße verteilt (max. 150 µl pro Gefäß). Die PCR wurde mit dem *MultiGene™ 96-Well Gradient Thermal Cycler* (*Labnet International Inc.*, Edison, USA) durchgeführt.

Bei Verwendung der Taq-DNA-Polymerase ergaben sich folgende Konzentrationen bzw. Mengen für die verschiedenen Komponenten in der fertigen PCR-Reaktion: je 200 µM dATP, dCTP,

dGTP und dTTP, je 0,2 μ M Vorwärts- und Rückwärtsprimer, typischerweise < 1 ng Template-DNA/50 μ l Reaktionsvolumen und 0,025 units/ μ l Taq-DNA-Polymerase in 1x Taq- bzw. ThermoPol[®]-Reaktionspufferbedingungen.

Bei Verwendung der Q5[®]-*High-Fidelity*-DNA-Polymerase ergaben sich folgende Konzentrationen bzw. Mengen für die verschiedenen Komponenten in der fertigen PCR-Reaktion: je 200 μ M dATP, dCTP, dGTP und dTTP, je 0,5 μ M Vorwärts- und Rückwärtsprimer, typischerweise < 1 ng Template-DNA/50 μ l Reaktionsvolumen und 0,02 units/ μ l Q5[®]-*High-Fidelity*-DNA-Polymerase in 1x Q5[®]-Reaktionspufferbedingungen.

Bei Verwendung der Taq-DNA-Polymerase war ein typisches PCR-Protokoll wie folgt aufgebaut: Die DNA-Doppelstränge wurden initial und während der PCR bei 95 °C für 30 s denaturiert. Die Primer-Hybridisierung fand bei Temperaturen zwischen 45 °C–68 °C statt und dauerte 30 s. Der Elongationsschritt wurde bei einer Temperatur von 68 °C durchgeführt, und die Dauer richtete sich nach der Länge des zu verlängernden DNA-Abschnitts (Polymerisierungsrate der Taq-DNA-Polymerase: 1000 b pro Minute). Nach 25 Zyklen–30 Zyklen folgte ein fünfminütiger Extensionsschritt bei 68 °C. Danach kühlte der Thermal Cycler die Temperatur auf 4 °C ab.

Bei Verwendung der Q5[®]-*High-Fidelity*-DNA-Polymerase war ein typisches PCR-Protokoll wie folgt aufgebaut: Die DNA-Doppelstränge wurden initial bei 98 °C für 30 s denaturiert, während der PCR wurde dieser Schritt ohne Veränderung der Temperatur auf 10 s verkürzt. Die Primer-Hybridisierung fand bei Temperaturen zwischen 50 °C–72 °C statt und dauerte 20 s. Der Elongationsschritt wurde bei einer Temperatur von 72 °C durchgeführt. Die Dauer richtete sich nach der Länge des zu verlängernden DNA-Abschnitts (Q5[®]-*High-Fidelity*-DNA-Polymerase benötigt etwa 20 s–30 s für 1000 b). Nach 25 Zyklen–30 Zyklen folgte ein zweiminütiger Extensionsschritt bei 72 °C. Danach kühlte der Thermal Cycler die Temperatur auf 4 °C ab.

Die Hybridisierungs-Temperatur der Primer wurde mittels des *Tm Calculators* (*New England Biolabs Inc., Ipswich, USA*), dessen Berechnungen der Hybridisierungs-Temperatur auf [271] basieren und der Salz-Korrekturen gemäß [272] vornimmt, ermittelt. Die Primer wurden manuell mit Hilfe des Sequenz-Editors *A plasmid Editor* von Wayne Davis erstellt.

Nach Beendigung der PCR wurden anhand einer kleinen Menge der PCR-Reaktion (typischerweise 10 μ l) die PCR-Produkte mittels horizontaler Agarosegelelektrophorese (vgl. Abschn. 4.2.4) auf Reinheit überprüft. Sofern nur DNA-Fragmente erwarteter Größe detektiert werden konnten, wurde der restliche PCR-Ansatz mit Hilfe des *PureLink[®] PCR Purification Kits* (*Fisher Scientific GmbH, Schwerte*) gemäß der Hersteller-Instruktionen purifiziert. Die Lösung der DNA erfolgte allerdings in Reinstwasser. Anderenfalls wurden die PCR-Produkte des verbleibenden PCR-Ansatzes mittels Agarosegelelektrophorese getrennt, die DNA-Banden der DNA-Stränge der richtigen Länge mittels eines geeigneten Skalpells aus dem Gel herausgeschnitten und die DNA mit Hilfe des *PureLink[®] Quick Gel Extraction Kit* (*Fisher Scientific GmbH, Schwerte*) nach Hersteller-Instruktionen aus dem Gel extrahiert, purifiziert und in Reinstwasser gelöst.

4.2.6. Restriktionsverdau von DNA-Strängen

Dieser Absatz führt allgemein in den Restriktionsverdau durch Zusammenfassen der relevanten Ausführungen aus Abschn. 12.2.3 aus [62] ein. Restriktionsendonukleasen vom Typ II katalysieren die Öffnung der Phosphodiesterbindung zwischen zwei DNA-Nukleotiden innerhalb einer spezifischen, typischerweise 4 bp–8 bp langen palindromischen Erkennungssequenz. Dieser Vorgang wird als Schneiden

bezeichnet. Ein Teil der Enzyme dieses Typs spaltet die beiden Doppelstrang-bildenden Einzelstränge 1 bp–4 bp versetzt voneinander, sodass die beiden durch das Schneiden entstehenden Doppelstränge einzelsträngige Überhänge besitzen, die als klebrige Enden (*sticky ends*) bezeichnet werden.

Es wurden nur Restriktionsenzyme (alle von *New England Biolabs GmbH*, Frankfurt am Main), die klebrige Enden erzeugen, und deren mitgelieferte 10x Reaktionspuffer (alle von *New England Biolabs GmbH*, Frankfurt am Main) zum Restriktionsverdau verwendet. Die Restriktionsansätze wurden in DNase- und RNase-freien 1,5 ml Reaktionsgefäßen (*Eppendorf AG*, Hamburg) präpariert.

Ein typischer Restriktionsdoppelverdau-Ansatz zum Verdau eines PCR-Produkts hatte ein Volumen von 30 µl und setzte sich wie folgt zusammen: 25 µl PCR-Produkt-DNA-Lösung, 3 µl des 10x Reaktionspuffers und je 1 µl der beiden Restriktionsenzyme.

Bei einem typischen Doppelverdau eines Plasmids wurde das Volumen der in den Restriktionsansatz pipettierten Plasmid-Lösung so eingestellt, dass etwa 1 µg–1,5 µg Plasmid-DNA in der finalen 30 µl-Reaktion vorhanden waren. Somit setzte sich eine solche Reaktion wie folgt zusammen: y µl Plasmid-Lösung (1 µg–1,5 µg Plasmid-DNA), (25–y) µl Reinstwasser, 3 µl des 10x Reaktionspuffers und je 1 µl der beiden Restriktionsenzyme.

Bei Restriktionsenzymen, die Rinderalbumin (BSA) benötigten, und bei Fehlen von BSA beim verwendeten Reaktionspuffer, wurde durch Hinzugabe von BSA (*New England Biolabs GmbH*, Frankfurt am Main) eine BSA-Konzentration von 100 µg/ml hergestellt.

Der Restriktionsverdau wurde bei 37 °C in einem Wasserbad durchgeführt. Je nach Enzym und Hersteller-Instruktion dauerte der Restriktionsverdau zwischen 15 min–3 h. Im Anschluss wurden die verdauten DNA-Moleküle mittels Agarosegelelektrophorese getrennt. Die DNA-Banden der DNA-Stränge mit der richtigen Länge wurden mit einem geeigneten Skalpell aus dem Gel herausgeschnitten und die DNA mit Hilfe des *PureLink® Quick Gel Extraction Kit* nach Hersteller-Instruktionen aus dem Gel extrahiert, purifiziert und in Reinstwasser gelöst.

4.2.7. Ligation von DNA-Strängen mit klebrigen Enden

Dieser Absatz führt allgemein in die Ligation von DNA-Fragmenten mit klebrigen Enden durch Zusammenfassen der relevanten Ausführungen aus Abschn. 12.2.4 aus [62] ein. DNA-Moleküle mit komplementären klebrigen Enden können mittels DNA-Ligasen kovalent verbunden werden. Werden ein Plasmid und ein PCR-Produkt mit dem gleichen Restriktionsenzym-Paar verdaut, besitzen sie komplementäre klebrige Enden, die es ermöglichen, das PCR-Produkt sowie das linearisierte Plasmid zu ligieren und erneut in ein geschlossenes Plasmid zu überführen. Dazu müssen die komplementären klebrigen Enden zuerst hybridisieren. Im Anschluss werden die reaktiven 3'- und 5'-Enden der sich durch die Hybridisierung in räumliche Nähe gebrachten Enden der DNA-Stränge durch die DNA-Ligase kovalent verbunden.

Zur Ligation von zuvor verdauten und komplementäre klebrige Enden besitzenden DNA-Fragmenten wurde die T4-DNA-Ligase (*New England Biolabs GmbH*, Frankfurt am Main) verwendet. Optimale Pufferbedingungen für das Enzym wurden durch Verwendung des mitgelieferten 10x T4-Ligase-Reaktionspuffers (*New England Biolabs GmbH*, Frankfurt am Main) hergestellt. In einer typischen 10 µl- oder 20 µl-Ligationsreaktion wurden 4 fmol–60 fmol von der verdauten, linearisierten Plasmid-DNA mit der 3- bis 10-fachen Stoffmenge des verdauten PCR-Produkts gelöst. Feste Volumina besaßen in einer 20 µl-Reaktion die T4-DNA-Ligase und der T4-DNA-Ligase-Reaktionspuffer mit 1 µl bzw.

2 µl. In einer 10 µl-Reaktion besaßen die T4-DNA-Ligase und der T4-DNA-Ligase-Reaktionspuffer ein festes Volumen von 0,5 µl bzw. 1 µl. Die Volumina des Reinstwassers und der hinzupipettierten DNA-Lösungen des verdauten linearisierten Plasmids und des PCR-Produkts wurden gemäß der oben angegebenen Ziel-DNA-Stoffmenge angepasst. Die Reaktionen wurden in 1,5 ml-Reaktionsgefäßen (*Eppendorf AG*, Hamburg) präpariert und anschließend über Nacht bei 16 °C in einem Wasserbad inkubiert, um die Hybridisierung der komplementären klebrigen Enden zu gewährleisten.

4.2.8. Gibson-Assemblierung von DNA-Strängen

Eine andere Methode, DNA-Stränge kovalent zu verbinden, ist die Gibson-Assemblierung, deren Prinzip aus [273], [274] in diesem Abs. zusammengefasst wird. Sie ermöglicht die kovalente Verbindung von DNA-Strängen in einer einzigen Reaktions-Mixtur, in der drei verschiedene Enzyme involviert sind: eine T5-Exonuklease, DNA-Polymerase und DNA-Ligase. Die DNA-Doppelstränge müssen an ihren zu verbindenden Enden einen kleinen Überlapp besitzen, bei dem die Sequenzen beider Stränge übereinstimmen. Die DNA-Enzym-Mischung wird optimalerweise zwischen 15 min-60 min bei einer Temperatur von 50 °C inkubiert. Die T5-Exonuklease degradiert Sequenz-unspezifisch die 5'-Enden der DNA-Stränge, sodass DNA-Doppelstränge entstehen, die zueinander komplementäre klebrige Enden besitzen, die lang genug sein müssen, damit sie bei 50 °C hybridisieren können. Das Enzym verliert während der Inkubation seine enzymatische Aktivität. Die hybridisierten klebrigen 3'-Enden bilden die notwendigen Primer, an denen die thermostabile DNA-Polymerase die Polymerisierung initiiert. Die Strang-Polymerisierung endet am gegenüberliegenden, hybridisierten klebrigen 5'-Ende des anderen Doppelstrangs. Die beiden final resultierenden „Löcher“ werden durch die thermostabile DNA-Ligase geschlossen, sodass beide Doppelstränge zu einem einzigen Doppelstrang verbunden werden.

Zur Gibson-Assemblierung von DNA-Strängen wurde das *2x Gibson-Assembly[®]-Master-Mix* (*New England Biolabs GmbH*, Frankfurt am Main) verwendet. Die Präparation der 20 µl-Assemblierungsreaktion und deren Inkubation erfolgten gemäß der Hersteller-Anweisungen. In dieser Reaktion lag die Gesamtkonzentration aller zu assemblierenden DNA-Fragmente zwischen 1 nM-25 nM. 10 µl des Gesamtvolumens bildete die *2x Gibson-Assembly[®]-Master-Mix*-Lösung. Die restlichen 10 µl konstituierten sich aus Reinstwasser und den DNA-Lösungen der unterschiedlichen zu assemblierenden DNA-Stränge. Die Volumina des Reinstwassers und der DNA-Lösungen wurden so angepasst, dass die Gesamtkonzentration aller DNA-Stränge im oben angegebenen Bereich lag. Die Gesamtkonzentration setzte sich stets aus gleich großen Teilkonzentrationen der unterschiedlichen, zu assemblierenden DNA-Stränge zusammen. Die Gibson-Assemblierungs-Reaktionen wurden in verschließbaren DNase- und RNase-freien 0,2 ml-PCR-Reaktionsgefäßen (*Biozym Scientific GmbH*, Hessisch Oldendorf) zubereitet und zwischen 15 min-30 min bei 50 °C im *MultiGene[™] 96-Well Gradient Thermal Cycler* (*Labnet International Inc.*, Edison, USA) inkubiert.

4.2.9. Transformation kompetenter *E.-coli*-Zellen, Plasmid-Amplifikation und Plasmidpräparation

4.2.9.1. Plasmide

Dieser Abs. fasst die relevanten Ausführungen aus [70] und Abschn. 6.2.1 aus [69] zusammen. Plasmide sind doppelsträngige und ringförmige DNA-Moleküle, die auch als Klonierungsvektoren bezeichnet werden. Sie können von Zellen aufgenommen (vgl. Abschn. 2.2.2.1, 4.2.9.2) und amplifiziert werden.

Die Replikation von Plasmiden verläuft autonom, unabhängig vom Chromosom des Trägers, da Plasmide einen eigenen Replikationsursprung oder -start besitzen. Zur weiteren Minimalausstattung eines Vektors gehört neben einem Selektionsgen, das in der Regel den Träger des Plasmids mit einer gewissen Antibiotikaresistenz ausstattet, sowie eine Klonierungsstelle, an der fremde DNA gezielt in das Plasmid integriert werden kann. Je nach Art des Replikationsursprungs unterscheidet sich die resultierende Anzahl der Plasmidkopien pro Bakterium. Liegen weniger als 20 Kopien pro Zelle vor, wird von einem *low-copy*-Plasmid gesprochen. Hingegen können von *high-copy*-Plasmiden mehrere hundert Kopien pro Zelle vorzufinden sein.

Alle zur zellfreien Genexpression verwendeten Plasmide basieren auf dem *low-copy*-Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP-T500 [51], das von der Kooperationsarbeitsgruppe von Vincent Noireaux (*School of Physics and Astronomy, University of Minnesota*) zur Verfügung gestellt und bereits zuvor in [60] verwendet wurde. Dieses Plasmid ist eine Variante des Plasmids pBEST-OR2-OR1-PR-UTR1-deGFP-T500 (Addgene Plasmid # 40019 mit vollständiger Sequenz), das ebenfalls auf [51] zurückgeht, aber anstelle des ColE1- den p15a-Replikationsursprung besitzt [51]. Das Vektor-Rückgrat bildet pBEST luc^{TM} (*Promega Corporation, Fitchburg, USA*), bei dem der PTacI-Promotor [275] durch den an einer Stelle mutierten λ -Repressor-Cro-Promotor (GenBank ID: J02459.1) ersetzt wurde [51]. Der untranslatierte Bereich von pBEST luc^{TM} wurde durch die T7-*g10*-Leitsequenz [276] enthaltende UTR1 substituiert [51]. Der Transkriptionsterminator T500 [277] wurde der DNA-Konstruktion hinzugefügt [51]. Das *degfp*-Gen (Abkürzung für Protein EGFP-Del6-229, vgl. Abschn. 2.5) kodiert für die am C- und N-terminal verkürzte Variante von EGFP [51], baut auf den Arbeiten von [151] auf und ersetzt das *luc*-Gen in pBEST luc^{TM} [51].

4.2.9.2. Herstellung kompetenter Zellen und Transformation

Dieser Abs. führt allgemein in die Transformation durch Zusammenfassen der relevanten Ausführungen aus Abschn. 12.3.1 aus [62] und Abschn. 6.4 aus [69] ein. Jede Methode, durch die eine Zelle DNA aus der Umgebung aufnimmt, wird als Transformation bezeichnet. Als kompetent gilt eine Zelle, die in der Lage ist, DNA aufzunehmen. Die Kompetenz, sofern nicht natürlich vorhanden, kann durch verschiedene Verfahren hergestellt werden. Eine mögliche Methode, Bakterien kompetent zu machen, ist die Kalziumchlorid-Methode. Dieses Verfahren besteht aus der Behandlung der zu transformierenden Zellen mit einer Kalziumchloridlösung. Werden diese Zellen dann mit der einzubringenden DNA auf Eis inkubiert und anschließend schnell auf eine Temperatur von etwa 37 °C-42 °C erwärmt, nimmt ein Teil der Zellen die DNA auf.

100 ml LB-Medium (Erlenmeyerkolben mit Schikane) wurden mit 100 μ l einer *E.-coli*-KL740-Übernacht-Kultur (29 °C, 230 rpm im Schüttelinkubator) inokuliert. Die Zellen wurden bei 29 °C und 230 rpm im Schüttelinkubator (*Universalschüttler SM 30 Control* mit *Inkubationshaube TH 30*, *Edmund Bühler GmbH*, Hechingen) bis $OD_{595\text{ nm}} \approx 0,4$ (*Cary® 50 UV-Vis Spektrophotometer*, *Agilent Technologies Deutschland GmbH & Co. KG*, Waldbronn) inkubiert. Die Kultur wurde zuerst für etwa 5 min-10 min auf Eis gekühlt. Die nachfolgenden Schritte wurden auf Eis oder gekühlt (4 °C) durchgeführt. 30 ml der Kultur wurden in gleichen Volumina von jeweils 15 ml in zwei mit Drehverschluss verschließbare 15 ml-Zentrifugationsröhrchen (*Greiner Bio-One GmbH*, Frickenhausen) überführt, die Zellen durch Zentrifugation bei 3000 x g (*Allegra™ 25R*, *Beckman Coulter GmbH*, Krefeld) für 10 min geerntet. Der Überstand wurde verworfen, weitere 30 ml der Kultur erneut in gleichen Volumina von jeweils 15 ml in die beiden Zentrifugationsröhrchen gefüllt, die Zellen durch 10-minütige Zentrifugation bei 3000 x g geerntet, der Überstand verworfen. Die beiden Bakterien-Pellets wurden jeweils mit 4 ml

auf Eis gekühlter CaCl_2 -Lösung (60 mM CaCl_2 , 15 % Glycerin, 10 mM PIPES, pH 7) resuspendiert, die resuspendierten Zellen während eines 10-minütigen Zentrifugationsschrittes bei 3000 x g pelletiert, der Überstand verworfen. Die Zellen wurden auf die gleiche Weise nochmals resuspendiert, 30 min auf Eis inkubiert, erneut während eines 10-minütigen Zentrifugationsschrittes bei 3000 x g pelletiert, der Überstand verworfen. Die beiden Bakterien-Pellets wurden jeweils in 1 ml der CaCl_2 -Lösung resuspendiert, die resultierende Lösung in Volumina von 20 μl in 1,5 ml Reaktionsgefäße (*Eppendorf AG*, Hamburg) aliquotiert und mit flüssigem Stickstoff schockgefroren, die Aliquots bei -80°C gelagert.

Die kompetenten *E.-coli*-KL740-Zellen wurden ca. 5 min auf Eis aufgetaut. Etwa 1 μl -3 μl der Plasmid-DNA-Lösung wurde in ein 20 μl -Aliquot pipettiert, dieses Gemisch für 10 min auf Eis gekühlt. Es erfolgte ein Hitzeschock bei 37°C für 20 s in einem Wasserbad. Anschließend wurden die Zellen erneut 5 min auf Eis inkubiert, bevor 500 μl LB-Medium hinzugegeben wurden. Nach Hinzugabe des LB-Mediums wurden die Zellen im Schüttelinkubator bei 29°C und 230 rpm 40 min lang inkubiert, anschließend bei 10000 rpm (*MiniSpin[®] Zentrifuge*, *Eppendorf AG*, Hamburg) pelletiert, der Überstand verworfen. In den Reaktionsgefäßen verblieb stets ein minimaler Überstand (schätzungsweise 20 μl -30 μl), in dem das Bakterien-Pellet durch Auf- und Abpipettieren resuspendiert wurde. Etwa 5 μl -10 μl dieser Resuspension wurden auf eine Agarplatte (mit Ampicillin) ausgestrichen, die ausgestrichenen Zellen über Nacht bei 29°C inkubiert. Auf der Agarplatte gewachsene Zellkolonien wurden mittels einer Pipettenspitze in flüssiges LB-Medium (mit Ampicillin) überführt und bei 230 rpm sowie 29°C über Nacht für die darauffolgende Plasmidpräparation inkubiert.

Von erfolgreich transformierten Zellen, die häufig verwendete Plasmide trugen, wurden Glycerin-Stöcke angefertigt. Dazu wurde eine Teilmenge der *E.-coli*-KL740-Übernacht-Kultur (mit Ampicillin) mit 100 %-igem Glycerin (25 % Glycerin / 75 % Übernacht-Kultur) durch Auf- und Abpipettieren vermischt, in flüssigem Stickstoff schockgefroren und bei -80°C gelagert. Um eine Übernacht-Kultur mit Hilfe eines Glycerin-Stocks für eine anschließende Plasmid-Präparation zu erzeugen, wurde der Glycerin-Stock auf Eis angetaut und mit einer sterilen Pipettenspitze etwas gefrorene Masse herausgenommen. Sie wurde in flüssiges LB-Medium (mit Ampicillin) überführt und bei 230 rpm sowie 29°C über Nacht für die darauffolgende Plasmidpräparation inkubiert.

4.2.9.3. Plasmidpräparation

Die alkalische Lyse zur Isolierung von Plasmid-DNA [278] aus Zellen bildet die Grundlage fast aller kommerziellen Plasmid-Purifizierungs-Kits (vgl. Abschn. 2.2.3 aus [69]), so auch bei den in dieser Arbeit verwendeten Kits. Mini- und Midipräparationen von durch *E.-coli*-KL740-Zellen amplifizierter Plasmid-DNA wurden mit Hilfe des GenElute[™] HP Plasmid Miniprep Kits bzw. GenElute[™] HP Plasmid Midiprep Kits (beide *Sigma-Aldrich Chemie GmbH*, Taufkirchen) gemäß der Hersteller-Instruktionen durchgeführt. Bei Minipräparationen wurden typischerweise zwischen 5 ml-10 ml einer *E.-coli*-KL740-Übernacht-Kultur pro Bindungssäule verwendet³. Bei Midipräparationen wurde das Volumen V_{Kultur} der pro Säule verwendeten *E.-coli*-Kultur nach Hersteller-Instruktionen durch die Formel $V_{\text{Kultur}}=250\text{ ml}/\text{OD}_{600\text{ nm}}$ (*Cary[®] 50 UV-Vis Spektrophotometer*, *Agilent Technologies Deutschland GmbH & Co. KG*, Waldbronn) bestimmt. Die extrahierte Plasmid-DNA wurde in Reinstwasser eluiert. Der *Concentrator 5301* (*Eppendorf AG*, Hamburg) fand Anwendung, um die DNA-Lösung bis auf ein resultierendes Volumen zwischen 50 μl -100 μl zu verdampfen und somit die gelöste DNA aufzukonzentrieren. Die Aufkonzentrierung geschah unter ständiger Zentrifugation der Lösung

³Der Hersteller empfiehlt 5 ml. Aufgrund der niedrigen Inkubationstemperatur von 29°C und der damit verbundenen geringeren Zelldichte nach Übernacht-Inkubation wurde das Volumen erhöht.

bei 30 °C und Unterdruck, der den Siedepunkt des Lösungsmittels erniedrigte. Die aufkonzentrierte Plasmid-DNA wurde mit Hilfe des *PureLink® PCR Purification Kits* (Fisher Scientific GmbH, Schwerte) gemäß der Hersteller-Instruktionen purifiziert und in Reinstwasser gelöst.

4.2.10. DNA-Sequenzierung

Dieser Abs. erklärt das Prinzip der Sanger-Sequenzierung, die die gebräuchlichste Form der Sequenzierung darstellt und nach ihrem Entwickler SANGER benannt ist, anhand einer Zusammenfassung der relevanten Ausführungen aus Abschn. 12.2.6 aus [62]. Als Sequenzierung wird die Bestimmung der Basenabfolge eines DNA-Moleküls bezeichnet. Während dieser modifizierten Form der Standard-PCR, bei der das zu sequenzierende DNA-Molekül das PCR-Template bildet, werden neben den vier Standard-DNA-Monomeren dATP, dCTP, dGTP und dTTP entsprechende Analoga verwendet, denen die OH-Gruppe am 3'-C-Atom des Fünffachzuckers fehlt (2'-3'-Dideoxynukleosidtriphosphat (ddNTP)). Im Gegensatz zur Standard-PCR wird nur ein Primer pro Reaktion verwendet, der den Startpunkt für die Sequenzierung festlegt. Zur Sequenzierung eines DNA-Moleküls werden vier Reaktionen präpariert, in die jeweils eines der vier ddNTPs in geringerer Konzentration als die vier dNTPs hinzugegeben wird. Beträgt beispielsweise der Konzentrationsunterschied zwischen dATP und ddATP 1000:1, wird durchschnittlich für jedes tausendste dATP ein ddATP in den wachsenden DNA-Strang eingebaut. Durch die fehlende OH-Gruppe bricht an dieser Stelle die Polymerisierung ab. Dies geschieht in gleicher Weise in den drei anderen Reaktionen mit den jeweils anderen ddNTPs. Im Anschluss werden die so entstandenen DNA-Stränge der vier Reaktionen in jeweils einer eigenen Gelbahn mittels Gelelektrophorese gemäß ihrer Größe getrennt. Da der Sequenzierungsprimer bei allen neu synthetisierten Fragmenten das 5'-Ende bildet und diese DNA-Stränge somit alle das gleiche 5'-Ende besitzen, variieren die Längen dieser Fragmente nur aufgrund der Polymerisierungsabbrüche durch den Einbau eines ddNTPs. Durch einen Vergleich aller vier Gelbahnen und den unterschiedlich langen DNA-Strängen kann die Sequenz determiniert werden.

Die relevanten transkriptions-translations-regulatorischen und kodierenden DNA-Sequenzen wurden von *Seqlab Sequence Laboratories Göttingen GmbH* (Göttingen) sequenziert.

4.2.11. Herstellung der verwendeten DNA-Konstrukte

Wie bereits in Abschn. 4.2.9.1 erläutert, basieren alle zur zellfreien Genexpression verwendeten Plasmide auf pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP-T500, das von der Kooperationsarbeitsgruppe von Vincent Noireaux (*School of Physics and Astronomy, University of Minnesota*) zur Verfügung gestellt und bereits zuvor in [60] verwendet wurde. Details zu diesem Plasmid und zu den transkriptions-translations-regulatorischen Elementen können Abschn. 4.2.9.1 entnommen werden. Weitere lineare DNA-Konstrukte wurden auf Grundlage dieses Plasmids oder den im Folgenden angegebenen Plasmiden konstruiert. In Abschn. 4.2.11 wird auf die Herstellung der in den Experimenten verwendeten DNA-Konstrukte eingegangen, die mit den Methoden, wie in den Abschn. 4.2.5-4.2.9 beschrieben, hergestellt wurden.

4.2.11.1. pBEST-p15a-OR2-OR1-Pr-UTR1-deGFPHis₆-Tag-T500

Um das zellfrei synthetisierte deGFP mittels His-Tag Purifikation (vgl. Abschn. 4.2.16) aus der zellfreien Reaktion extrahieren zu können, wurde dem *degfp*-Gen (Details in Abschn. 2.5)

vor dem Stopcodon eine His₆-Tag-Sequenz mittels PCR eingefügt, sodass das Modellprotein in C-Terminal mit einem His-Tag ausgestattet war. Die PCR erfolgte mit dem Vorwärtsprimer 5'-CACCATCAGCCAGAAAACC-3' und dem Rückwärtsprimer 5'-GAATTCCTCGAGTTA-[GTGGTGATGGTGATGATG]-[GATCCCGGCGGCGGT]-3', der die Sequenz für den His-Tag enthielt (linke eckige Klammer). Das PCR-Template bildete das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP-T500, zu dessen Sequenz der Rückwärtsprimer eine komplementäre Sequenz zum Hybridisieren besitzt (rechte eckige Klammer). Die Sequenz für das 5'-Ende des Rückwärtsprimers bilden die palindromischen Erkennungssequenzen GAATTC und CTCGAG für die Enzyme EcoRI bzw. XhoI. Nach der PCR wurde das so entstandene PCR-Produkt 5'-Vorwärtsprimer-...-OR2-OR1-Pr-UTR1-NcoI-deGFPHis₆-Tag-XhoI-EcoRI-3' mit den Restriktionsenzymen NcoI und XhoI verdaut. Ein aus einem mit den gleichen Enzymen durchgeführter Restriktionsverdau des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP-T500 führte zu einem linearisierten Plasmid, in das das verdauten PCR-Produkt ligiert wurde, sodass pBEST-p15a-OR2-OR1-Pr-UTR1-deGFPHis₆-Tag-T500 entstand. Die Sequenzen der transkriptions-translations-regulatorischen und kodierenden Elemente sind in Anh. A.2.2.1 vorzufinden. Diese Sequenzen wurden mit den Sequenzierungsprimern 5'-CACCATCAGCCAGAAAACC-3' und 5'-GAGCTGACTGGGTTGAAGG-3' verifiziert, die stromaufwärts bzw. -abwärts des transkriptions-translations-regulatorischen und kodierenden Bereichs auf dem Plasmid binden (vgl. Plasmidsequenz in Addgene Datenbank, Plasmid # 40019).

Das zellfrei synthetisierte deGFPHis₆-Tag wird in Abschn. 4.3.2 und 4.4 vereinfacht als deGFP bezeichnet.

4.2.11.2. pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}His₆-Tag-T500

Diese Klonierung wurde von Matthias P. Exner in der Kooperationsarbeitsgruppe von Nediljko Budisa (Institut für Chemie, Technische Universität Berlin) durchgeführt. Das für EGFP^{N149amber}His₆-Tag kodierende Gen (Details in Abschn. 2.5) wurde in das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP-T500 anstelle des *degfp*-Gens vom Kooperationspartner eingefügt. Dieses Modellprotein mit dem amber-Stopcodon an Position 149 wurde bereits in seiner Dissertation [154] verwendet. Das Gen, das für EGFP^{N149amber}His₆-Tag kodiert, enthielt vor dem Stopcodon eine His₆-Tag-Sequenz, sodass das Modellprotein am C-Terminus mit einem His-Tag ausgestattet war. Die Sequenzen der transkriptions-translations-regulatorischen und kodierenden Elemente sind in Anh. A.2.2.2 vorzufinden. Diese Sequenzen wurden mit den Sequenzierungsprimern 5'-CACCATCAGCCAGAAAACC-3' und 5'-GAGCTGACTGGGTTGAAGG-3' verifiziert, die stromaufwärts bzw. -abwärts des transkriptions-translations-regulatorischen und kodierenden Bereichs auf dem Plasmid binden (vgl. Plasmidsequenz in Addgene Datenbank, Plasmid # 40019).

Das zellfrei synthetisierte EGFP^{N149amber}His₆-Tag wird in Abschn. 4.3.3 und 4.4 vereinfacht als EGFP^{N149amber} bezeichnet.

4.2.11.3. T7-tRNA^{PyI}_{CUA}

Das für die tRNA^{PyI}_{CUA} kodierende Gen (Details in Abschn. 4.1.2.2) wurde von der Firma *Eurofins Genomics GmbH* (Ebersberg) in den Vektor pEX-A2 kloniert und geliefert, der als PCR-Template zur Synthese des linearen DNA-Fragments T7-tRNA^{PyI}_{CUA} fungierte. Als Vorwärtsprimer diente 5'-GAAATTAATACGACTCACTATA[GGAACCTGATCATGTAGATC]-3'. Die eckige Klammer markiert den zum für die tRNA^{PyI}_{CUA} kodierenden komplementären Bereich, der zur Primer-Hybridisierung

notwendig ist. Die restliche Sequenz entspricht der Sequenz des T7-Promotors. Die Sequenz des Rückwärtsprimers war 5'-TGGCGGAAACCCCGGGAA-3'. Der Rückwärtsprimer konnte durch seine Sequenz passgenau auf dem 3'-Ende des für die tRNA_{CUA}^{Py1} kodierenden Gens binden.

Die Sequenzierung von T7-tRNA_{CUA}^{Py1} erfolgte nach einer Elongation dieses Konstrukts durch eine PCR. Bei dieser PCR wurden der Vorwärtsprimer 5'-GCATGTCACTACGAAATTAATACGACTCACTATAGGAAACCTG-3' und der Rückwärtsprimer 5'-GCTTGTAATGAATGGCGGAAACCCCGG-3' verwendet. Das verlängerte PCR-Produkt wurde mit den Sequenzierungsprimern 5'-GCATGTCACTACGAAATTAATACG-3' und 5'-GCTTGTAATGAATGGCGGA-3' sequenziert. Die relevanten transkriptions-translations-regulatorischen und kodierenden Elemente sind in Anh. A.2.2.3 dargestellt.

4.2.11.4. OR2-OR1-Pr-UTR1-T7-RNA-Polymerase-T500

Von Shixin Ye-Lehmann (*Département de Biologie, École Normale Supérieure de Paris*) wurde ein unbekanntes Plasmid zur Verfügung gestellt, auf dem sich das Gen, das für die T7-RNA-Polymerase mit der stillen K378R Mutation [279], [280] kodiert, befand. Es wurden drei lineare DNA-Fragmente hergestellt, die komplementäre, doppelsträngige Enden besaßen und somit mittels einer Gibson-Assemblierung kovalent verbunden werden konnten. Das resultierende DNA-Konstrukt trug die transkriptions-translations-regulatorischen und kodierenden Elemente OR2-OR1-Pr-UTR1-T7-RNA-Polymerase^{K378R}-T500. Es wurde mittels anschließender PCR amplifiziert und in den Experimenten verwendet. Um die Möglichkeit zu besitzen, die T7-RNA-Polymerase^{K378R} zellfrei zu synthetisieren, zu purifizieren und aufkonzentrieren zu können, um sie in zellfreien Expressions-Experimenten als synthetisiertes Enzym den Reaktionen hinzufügen zu können, wurde eine His₆-Tag-DNA-Sequenz hinter dem Startcodon eingefügt. Somit besaß das Enzym einen N-Terminalen His-Tag. Angemerkt sei, dass im Folgenden die Bezeichnung T7-RNA-Polymerase zur Vereinfachung verwendet wird, aber damit die T7-RNA-Polymerase mit N-Terminalen His₆-Tag und der stillen Mutation K378R gemeint ist.

Das erste lineare DNA-Fragment wurde mittels PCR mit dem Vorwärtsprimer 5'-ATG[CATCATCACCATCACCAC]AACACGATTAACATCGCTAAGAA-3', dem Rückwärtsprimer 5'-CTTTCGGCGGGCTTTGCTCGAG[TTACGCGAACGCGAAGTCC]-3' und dem unbekanntem Plasmid mit dem Gen, das für die T7-RNA-Polymerase kodiert, als PCR-Template synthetisiert. In der Sequenz des Vorwärtsprimers markiert die eckige Klammer die His₆-Tag-Sequenz. Stromabwärts dieser Sequenz liegt die zum Gen der T7-RNA-Polymerase komplementäre Teilsequenz, die der Primer-Hybridisierung dient. Die eckige Klammer in der Sequenz des Rückwärtsprimers markiert den zum Gen der T7-RNA-Polymerase komplementären Bereich, der notwendig für das Primer-Annealing ist. Stromaufwärts folgen die Schnittstellen XhoI und eine Teilsequenz von T500.

Das zweite lineare DNA-Fragment wurde mittels PCR mit dem Vorwärtsprimer 5'-CACCATCAGCCAGAAAACC-3', dem Rückwärtsprimer 5'-[GTTGTGGTGTGATGGTGTGATG]CATGGTATATCTCCTTCTTAAAGTT-3' und dem Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP-T500 als PCR-Template synthetisiert. Der Vorwärtsprimer bindet in etwa 140 bp vor OR2-OR1-Pr. Der zum Plasmid komplementäre Bereich der Sequenz des Rückwärtsprimers überlappt mit einem Teil der UTR1 und endet mit dem Startcodon. Die eckige Klammer markiert die Sequenz des His₆-Tags und ein Codon des Gens, das für die T7-RNA-Polymerase kodiert. Dieser Teil bildet den für die Gibson-Assemblierung notwendigen Überlapp mit dem ersten Fragment.

Das dritte lineare Fragment entstand durch Restriktionsverdau des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-deGFPHis₆-Tag-T500 mit den Restriktionsenzymen XhoI und SphI. Das mit SphI verdaute Ende des Plasmids besaß den für die Gibson-Assemblierung notwendigen komplementären Überlapp mit dem zweiten Fragment, das andere Ende besaß einen komplementären Überlapp mit dem ersten Fragment.

Durch eine Gibson-Assemblierung wurden die drei Fragmente vereinigt. Das vereinigte Fragment diente als PCR-Template mit dem Vorwärtsprimer 5'-CACCATCAGCCAGAAAACC-3' und dem Rückwärtsprimer 5'-CCTTCAACCCAGTCAGCTC-3', sodass das lineare DNA-Konstrukt, das die transkriptions-translations-regulatorischen und kodierenden Elemente OR2-OR1-Pr-UTR1-T7-RNA-Polymerase-T500 trägt, amplifiziert wurde. Es wird im Folgenden mit OR2-OR1-Pr-UTR1-T7-RNA-Polymerase-T500 bezeichnet. Seine relevanten transkriptions-translations-regulatorischen und kodierenden Elemente sind in Anh. A.2.2.5 dargestellt.

Die Sequenzierung dieser Elemente erfolgte mit folgenden Sequenzierungsprimern: 5'-CTTATGCCACGAAGACCACG-3', 5'-CCTTGCGGGTTGAACATTGA-3', 5'-GCGATTGAGCGTGAAGAAGACT-3' und 5'-GTCCAAAGAGTTCGGCTTCC-3'⁴.

4.2.11.5. pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA^{Pyl}_{CUA}-PylRS^{Y384F}-T500

Das Plasmid pJZ_Trp_1xMmtRNAPyl_Strep_MmPylRS(Y384F), das die Gene, die für die tRNA^{Pyl}_{CUA} und die PylRS^{Y384F} kodieren (Details in Abschn. 4.1.2.2), trägt, wurde von Matthias P. Exner aus der Kooperationsarbeitsgruppe von Nediljko Budisa (Institut für Chemie, Technische Universität Berlin) zur Verfügung gestellt. Die Mutation Y384F geht auf [254] zurück (vgl. Abschn. 4.1.2.2). Zusätzlich enthält das zur Verfügung gestellte Gen, das für die PylRS^{Y384F} kodiert, N-Terminal eine Streptavidin-Tag-Sequenz. Weitere Details zum pJZ-System und zur bereitgestellten, gekoppelten DNA-Konstruktion bestehend aus den Genen, die für die tRNA^{Pyl}_{CUA} und PylRS^{Y384F} kodieren, können der Dissertation von Matthias P. Exner [154] und [281] entnommen werden. Das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA^{Pyl}_{CUA}-PylRS^{Y384F}-T500 wurde mittels Gibson-Assemblierung zweier DNA-Fragmente hergestellt.

Um das erste Teilfragment zu synthetisieren, wurde eine PCR durchgeführt, bei der das zur Verfügung gestellte Plasmid pJZ_Trp_1xMmtRNAPyl_Strep_MmPylRS(Y384F) das PCR-Template bildete. Mittels des Vorwärtsprimers 5'-AAGGAGATATA[CCATGG]GGAAACCTGATCATGTAGATC-3' und Rückwärtsprimers 5'-CGGGCTTTG[GATCC]TTACAGGTTGGTAGAAATCC-3' wurde ein PCR-Produkt hergestellt, das die Gene, die für die tRNA^{Pyl}_{CUA} und PylRS^{Y384F} kodieren, in gekoppelter Weise trug. Das 5'-Ende des Vorwärtsprimers verlängerte das Gen, das für die tRNA^{Pyl}_{CUA} kodiert, um die Schnittstelle NcoI (eckige Klammer) und einen Teil der Sequenz der UTR1. Das 5'-Ende des Rückwärtsprimers verlängerte das Gen, das für PylRS^{Y384F} kodiert, um die Schnittstelle BamHI (eckige Klammer) und einen Teil der Sequenz des T500.

Um das zweite DNA-Fragment zu synthetisieren, wurde eine PCR durchgeführt, bei der das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP-T500 das PCR-Template war. Mittels des Vorwärtsprimers 5'-[GGATCC]CAAAGCCCGCCGAAAGGC-3' und Rückwärtsprimers 5'-[CCATGG]TATATCTCCTTCTTAAAGTTAAACAAA-3' wurde ein PCR-Produkt hergestellt, dessen Enden jeweils komplementäre Überlappungen mit dem ersten DNA-Fragment aufwiesen. Durch

⁴Die Aufbereitung der DNA zur Sequenzierung wurde von Marc Finkler während dessen Einarbeitung unter Anleitung durch den Autor dieser Dissertation durchgeführt.

den Vorwärtsprimer wurde die Plasmid-Sequenz bis zur Sequenz des T500 amplifiziert und die Schnittstelle für BamHI (eckige Klammer) angehängt. Der Rückwärtsprimer amplifizierte die Plasmid-Sequenz bis zur Sequenz der UTR1 inklusive der Schnittstelle von NcoI (eckige Klammer).

Durch eine Gibson-Assemblierung wurden die beiden linearen DNA-Fragmente zum Plasmid pBEST-OR2-OR1-Pr-UTR1-tRNA^{PyI}_{CUA}-PyIRS^{Y384F}-T500 vereinigt. Diese Sequenzen wurden mit den Sequenzierungsprimern 5'-CACCATCAGCCAGAAAACC-3' und 5'-GAGCTGACTGGGTTGAAGG-3' verifiziert, die stromaufwärts bzw. -abwärts des transkriptions-translations-regulatorischen und kodierenden Bereichs auf dem Plasmid binden (vgl. Plasmidsequenz in Addgene Datenbank, Plasmid # 40019). Die relevanten transkriptions-translations-regulatorischen und kodierenden Elemente sind in Anh. A.2.2.4 dargestellt.

4.2.11.6. pBEST-p15a-OR2-OR1-Pr-UTR1-PyIRS^{Y384F}-T500

Das Plasmid wurde mittels Ligation zweier Teilfragmente hergestellt, die komplementäre klebrige Enden besaßen. Das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA^{PyI}_{CUA}-PyIRS^{Y384F}-T500 wurde mit den Restriktionsenzymen NcoI und BamHI verdaut, um das erste Teilfragment herzustellen.

Das zweite Teilfragment wurde mittels PCR hergestellt, bei der das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA^{PyI}_{CUA}-PyIRS^{Y384F}-T500 als PCR-Template diente. Mittels des Vorwärtsprimers 5'-[AAGCTTCCATGG]ATAAAAAACCACTAAACACTC-3' und Rückwärtsprimers 5'-GCTTTG[GGATCC]TTACAGGTTGG-3' wurde ein PCR-Produkt hergestellt, das das Gen, das für die PyIRS^{Y384F} kodiert (Details in Abschn. 4.1.2.2), trug. Das 5'-Ende des Vorwärtsprimers bildet die Schnittstellen HindIII und NcoI (eckige Klammer). Das 5'-Ende des Rückwärtsprimers bildet die Schnittstelle BamHI (eckige Klammer) und eine Teilsequenz des T500. Das PCR-Produkt wurde ebenfalls mit den Restriktionsenzymen NcoI und BamHI verdaut.

Durch Ligation der beiden Teilfragmente entstand das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-PyIRS^{Y384F}-T500. Diese Sequenzen wurden mit den Sequenzierungsprimern 5'-CACCATCAGCCAGAAAACC-3' und 5'-GAGCTGACTGGGTTGAAGG-3' verifiziert, die stromaufwärts bzw. -abwärts des transkriptions-translations-regulatorischen und kodierenden Bereichs auf dem Plasmid binden (vgl. Plasmidsequenz in Addgene Datenbank, Plasmid # 40019). Die relevanten transkriptions-translations-regulatorischen und kodierenden Elemente sind in Anh. A.2.2.6 dargestellt.

4.2.12. Konzentrationsbestimmung von DNA

Es wurden zwei Verfahren zur Bestimmung der DNA-Konzentration verwendet.

4.2.12.1. Fluoreszenzspektroskopische Bestimmung der DNA-Konzentration

Bei Anwendung dieses Verfahrens wurde die Konzentration der DNA mit Hilfe des *QuantiFluor*[®] dsDNA Systems (Promega GmbH, Mannheim) bestimmt. Grundlage dieses Systems bildet der fluoreszierende *QuantiFluor*[®]-Farbstoff, der spezifisch an doppelsträngiger DNA bindet. Die hergestellten Lösungen besaßen inklusive der Farbstofflösung ein Volumen von 250 µl, es herrschten 1x TE-Pufferbedingungen (10 mM Tris-HCl (pH 7,5), 1 mM EDTA (pH 8,0); mit mitgeliefertem 20x TE-Puffer (pH 7,5) hergestellt). Die Lösungen wurden vor der Fluoreszenz-Messung zwischen 15 min-

30 min bei Raumtemperatur inkubiert. Die Messung der Fluoreszenz-Intensität erfolgte in einer 96-Well-Mikroplatte (*Greiner Bio-One GmbH*, Frickenhausen) mit dem Mikroplattenleser *POLARstar® Optima* (*BMG Labtech GmbH*, Offenburg). Bei der Messung der Fluoreszenz-Intensität betragen die Anregungs- und Emissionswellenlängen 485 nm bzw. 520 nm. Die Konzentration einer DNA-Lösung wurde mit Hilfe einer Standardkurve errechnet, die aus einer Konzentrationsreihe von DNA-Lösungen mit definierter Konzentration ermittelt wurde. Die DNA-Lösungen dieser Konzentrationsreihe wurden mittels des mitgelieferten DNA-Standards präpariert.

4.2.12.2. Absorptionsspektroskopische Bestimmung der DNA-Konzentration

Dieser Absatz führt allgemein in die absorptionsspektroskopische Bestimmung der DNA-Konzentration durch Zusammenfassen der relevanten Ausführungen aus Abschn. 2.5 aus [69] ein. Die Konzentration einer DNA-Lösung lässt sich mittels der Messung der optischen Dichte (OD) bei einer Wellenlänge von 260 nm realisieren. Proteinkontaminationen können über Messung der optischen Dichte bei 280 nm und Bildung des Quotienten $OD_{260\text{ nm}}/OD_{280\text{ nm}}$ überprüft werden. Während Quotienten zwischen 1,8-2,0 auf eine Nukleinsäure-Lösung ohne Proteinkontaminationen hindeuten, können niedrigere Quotienten durch Verunreinigungen von Proteinen entstehen.

Zur Quantifizierung der DNA-Konzentration einer Lösung auf absorptionsspektroskopischen Weg wurde das *NanoDrop 2000c UV-Vis Spectrometer* (*Thermo Fisher Scientific Inc.*, Waltham, USA) verwendet. Auf Grundlage des umgeformten Beer-Lambertschen Gesetzes

$$c = \frac{OD_{260} \cdot \epsilon}{d} \quad (4.1)$$

errechnet das Gerät mit 1 µl DNA-Lösung die Konzentration c der gelösten DNA. Dabei sind die Werte der Weglänge d des Lichts durch die DNA-Lösung und den von der Wellenlänge abhängigen Extinktionskoeffizienten ϵ im Gerät hinterlegt.

4.2.13. Konzentrationsbestimmung von Proteinen

Die Proteinkonzentration in einer Lösung wurde mittels des *Bradford-Tests* [282] quantifiziert: Der Farbstoff Coomassie-Brillant-Blau G-250 bindet an Proteine und verschiebt das Extinktionsmaximum des Farbstoffs von 465 nm auf 595 nm, sodass die Extinktion einer Proteinlösung bei einer Wellenlänge von 595 nm in Abhängigkeit der Proteinkonzentration zu- oder abnimmt.

Zur Quantifizierung der Proteinkonzentration einer Lösung wurde das *Protein Assay Kit* mit BSA Standard (*Bio-Rad Laboratories GmbH*, München) verwendet. Das Volumen der für den *Bradford-Test* hergestellten Lösungen inklusive der Farbstoff-Lösung betrug 1 ml. Vor Messung der Extinktion bei einer Wellenlänge von 595 nm wurden die Lösungen zwischen 15 min-30 min bei Raumtemperatur inkubiert. Die Messung der Extinktion erfolgte entweder mit dem *Cary® 50 UV-Vis Spectrophotometer* (*Agilent Technologies Deutschland GmbH & Co. KG*, Waldbronn) oder dem *NanoDrop 2000c UV-Vis Spectrometer* (*Thermo Fisher Scientific Inc.*, Waltham, USA). Die Konzentration einer Protein-Lösung wurde mit Hilfe einer Standardkurve errechnet, die aus einer Konzentrationsreihe von BSA-Lösungen mit definierter Konzentration ermittelt wurde. Die BSA-Lösungen dieser Konzentrationsreihe wurden mittels des mitgelieferten BSA-Standards oder eines anderen BSA-Standards (*New England Biolabs GmbH*, Frankfurt am Main) präpariert.

4.2.14. Herstellung des *E.-coli*-Zellextraktes, Präparation der zellfreien Reaktionen und Reaktionsbedingungen

4.2.14.1. Herstellung des *E.-coli*-Zellextraktes und allgemeine Reaktionsbedingungen

Das in dieser Arbeit zum Einbau nicht-kanonischer Aminosäuren verwendete zellfreie, endogene *E.-coli*-Expressionssystem geht auf [51] zurück. Ein detailliertes Protokoll (zusätzlich in Video-Format) zur Herstellung des zugrundeliegenden Zellextraktes, zur Präparation von zellfreien Reaktionen und aller dafür notwendigen Puffer, Lösungen und anderer Komponenten sowie eine Materialliste wurden in [52] veröffentlicht. Zusammengefasst enthielt eine typische zellfreie Reaktion die folgenden Komponenten in den angegebenen Konzentrationen [52]: 8,9 mg/ml-9,9 mg/ml endogene *E.-coli*-Proteine vom Zellextrakt, 50 mM HEPES, 1,5 mM ATP und GTP, 0,9 mM CTP und UTP, 0,2 mg/ml tRNA, 0,26 mM CoA, 0,33 mM NAD, 0,75 mM cAMP, 0,068 mM Folsäure, 1 mM Spermidin, 30 mM 3-PGA, 2 % PEG-8000. Die Expressionseffizienz der zellfreien Reaktion lässt sich durch die Hinzugabe von zusätzlichem Magnesium- und Kaliumglutamat optimieren [51], [52]. Die optimalen Konzentrationen dieser beiden Stoffe variieren leicht bei jeder Produktionsreihe des zugrundeliegenden Zellextraktes und sollen für jede Produktionsreihe optimiert werden [51], [52]. Für den eingesetzten Zellextrakt lagen die optimalen Konzentrationen für zusätzliches Magnesium- und Kaliumglutamat zwischen 5 mM-8 mM bzw. 8 mM-20 mM (vgl. Abschn. 4.3.1). Die verwendete Magnesiumglutamat- und Kaliumglutamat-Konzentration betrug 6 mM bzw. 10 mM (vgl. Abschn. 4.3.1). Die Konzentration jeder kanonischen Aminosäure kann zwischen 0,5 mM und 1,5 mM (Leu zwischen ca. 0,4 mM und 1,3 mM) variiert werden [51].

4.2.14.2. Präparation der zellfreien Reaktionen zum seitenkettenspezifischen Einbau nicht-kanonischer Aminosäuren in Proteine

Die in Abschn. 4.2.14.2 beschriebene Präparation der Aminosäuren-Mixes und der zellfreien Reaktionen zum seitenkettenspezifischen Einbau von nicht-kanonischen Aminosäuren in Modellproteine wurde bereits in [193], [194] veröffentlicht, baut auf [51], [52] auf und wird kurz zusammengefasst. Zur Herstellung des Aminosäuren-Mixes wurde der *RTS Amino Acid Sampler (5 PRIME GmbH, Hilden; auch erhältlich biotechrabbit GmbH, Hennigsdorf)* verwendet, der alle 20 kanonischen Aminosäuren in einzelnen Stocklösungen mit einer Konzentration von je 168 mM (Leu nur 140 mM) bereitstellt. Die einzubauenden nicht-kanonischen Aminosäuren wurden in Reinstwasser in der gleichen Konzentration (168 mM) gelöst. Nicht-kanonische Aminosäuren, die als Racemat vorlagen, wurden doppelt so hoch konzentriert (336 mM). Diese einzelnen Stocklösungen wurden in Reinstwasser gelöst, um neben einem Standardaminosäuren-Mix, der alle 20 kanonischen Aminosäuren in einer Konzentration von 6 mM (Leu nur 5 mM) enthielt, für jede einzubauende nicht-kanonische Aminosäure zwei weitere Mixes zu erhalten: Beim ersten Aminosäure-Mix wurde die kanonische Aminosäure durch das nicht-kanonische Analogon in der gleichen Konzentration (6 mM) ersetzt. Wenn die nicht-kanonische Aminosäure als Racemat vorlag, besaß sie in diesem Aminosäuren-Mix hingegen eine Konzentration von 12 mM. Beim zweiten Aminosäuren-Mix wurde weder die kanonische, noch ihr nicht-kanonisches Analogon hinzugegeben. Das fehlende Volumen wurde stattdessen durch Reinstwasser ersetzt, sodass dieser Mix die anderen 19 kanonischen Aminosäuren jeweils in der gleichen Konzentration wie in den beiden anderen Mixes bereitstellte.

Es wurden zellfreie Reaktionen hergestellt, die standardmäßig ein Volumen von 90 μ l besaßen [52] und die in Abschn. 4.2.14.1 aufgeführten Komponenten in den angegebenen Konzentrationen enthielten.

30 μl dieser 90 μl stammten vom Zellextrakt. Darüber hinaus war das DNA-Template pBEST-p15a-OR2-OR1-Pr-UTR1-deGFPHis₆-Tag-T500, das die zellfreie Synthese des Modellproteins deGFPHis₆ ermöglicht, in den zellfreien Reaktionen 10 nM konzentriert, da ab dieser Konzentration die Proteinausbeute bei Verwendung dieses Plasmids sättigt [51]. Von den 90 μl einer zellfreien Reaktion stammten 15 μl von einem der drei Aminosäuren-Mixes, sodass alle Aminosäuren jeweils eine Konzentration von 1 mM (0,83 mM Leu) in der zellfreien Reaktion besaßen. Lag eine nicht-kanonische Aminosäure als Racemat vor, so besaß sie eine doppelt so hohe Konzentration (2 mM).

Die zellfreien Reaktionen unterschieden sich in ihren Zusammensetzungen nur durch den hinzugegebenen Aminosäuren-Mix. Bei der Präparation von zellfreien Referenzreaktionen wurde der Aminosäuren-Mix, der alle 20 kanonischen Aminosäuren enthielt, hinzugegeben. Bei der Herstellung zellfreier Negativ-Kontrollreaktionen wurde der Aminosäuren-Mix, der weder die einzubauende nicht-kanonische Aminosäure noch ihr kanonisches Analogon bereitstellte, hinzupipettiert. In zellfreien Reaktionen, in denen die nicht-kanonische Aminosäure in deGFPHis₆ eingebaut werden sollte, wurde der Aminosäuren-Mix, in dem die nicht-kanonische Aminosäure ihr kanonisches Analogon substituierte, pipettiert. Um die zellfreien Reaktionen ausreichend mit Sauerstoff zu versorgen, wurden die 90 μl in Volumina von jeweils 6 μl gesplittet. Die Reaktionen wurden über Nacht bei 29 °C inkubiert. Im Anschluss an die zellfreie Expression wurden die gesplitteten Reaktionen wieder zu einer Reaktion vereinigt. Um eine größere Menge vom Modellprotein für die anschließende Analyse zu erzeugen, wurde das Volumen der zellfreien Reaktionen je nach Bedarf in Schritten von 90 μl nach oben skaliert.

4.2.14.3. Präparation der zellfreien Reaktionen zum ortsspezifischen Einbau nicht-kanonischer Aminosäuren in Proteine

Die Präparation der zellfreien Reaktionen zum ortsspezifischen Einbau nicht-kanonischer Aminosäuren baute ebenfalls auf [51], [52] auf. Zu Beginn eines Experiments wurden Mastermix-Reaktionen hergestellt. Erneut stammten 30 μl vom Zellextrakt und, wenn nicht anders angegeben, 15 μl vom Aminosäuren-Mix, der alle 20 kanonischen Aminosäuren (jeweils 6 mM, nur Leu mit 5 mM) bereitstellte (vgl. Abschn. 4.2.14.2). Weitere Komponenten wurden hinzugegeben, jedoch die Mastermix-Reaktionen nicht bis zum Gesamtvolumen von 90 μl vervollständigt, sondern ein gewisser Volumenanteil eingespart. Die unvollständig präparierten Mastermix-Reaktionen wurden in gleiche Volumina gesplittet. Die gesplitteten zellfreien Reaktionen wurden dann durch Hinzugabe weiterer Biomoleküle vervollständigt. Diese Strategie fand Anwendung, um Konzentrationen einer oder mehrerer Biomoleküle innerhalb einer Serie gesplitteter zellfreier Reaktionen zu variieren und die zellfreie Synthese des Modellproteins in Abhängigkeit dieser Variationen zu analysieren. Nach Vervollständigung der gesplitteten zellfreien Reaktionen enthielten diese die in Abschn. 4.2.14.1 angegebenen Komponenten in den dort beschriebenen Konzentrationen, jede kanonische Aminosäure besaß eine Konzentration von 1 mM (0,83 mM Leu). Bei Verwendung von linearen DNA-Fragmenten beinhalteten diese Reaktionen die verkürzte Version GamS des λ -Phagen Proteins Gam in rekombinanter und purifizierter Form (3,3 μM), das die *E.-coli*-RecBCD-Nuklease-Aktivität inhibiert [134] (vgl. Abschn. 2.4.3.1.3)⁵. Die Konzentration weiterer Biomoleküle, insbesondere der DNA-Templates und BocLys, variierte bei den Experimenten und wird im Ergebnisteil 4.3.3 für das jeweilige Experiment angegeben.

⁵Das rekombinante und His-Tag purifizierte GamS-Protein wurde von der Kooperationsarbeitsgruppe von Vincent Noireaux (*School of Physics and Astronomy, University of Minnesota*) zur Verfügung gestellt. Die transkriptions-translations-regulatorischen und kodierenden DNA-Sequenzen und weitere Details sind in Anh. A.2.2.7 zu finden.

Die Strategie der Präparation der zellfreien Reaktionen wird anhand des folgenden Beispiels illustriert, das den einfachsten Fall, die Variation einer einzigen Konzentration eines bestimmten Biomoleküls innerhalb einer Serie zellfreier Reaktionen aus einer Mastermix-Reaktion, darstellt (vgl. Abschn.4.3.3.1): Es wurde eine zellfreie Mastermix-Reaktion eines Volumens von 81 μl hergestellt, in der bereits alle anderen Komponenten des Experiments pipettiert wurden, außer dem Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500. Der Mastermix wurde in sechs gleiche Volumina von jeweils 10,8 μl gesplittet. Es wurden sechs Lösungen hergestellt, in denen pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 in jeweils einer der folgenden Konzentrationen gelöst war: 0 nM, 10 nM, 20 nM, 30 nM, 40 nM und 50 nM. In jede der sechs gesplitteten zellfreien Reaktionen wurden jeweils 1,2 μl von jeweils einer dieser sechs Plasmidlösungen pipettiert, sodass jede dieser Reaktionen nach Vervollständigung die in Abschn. 4.2.14.1 angegebenen Komponenten in den dort beschriebenen Konzentrationen und jede kanonische Aminosäure in einer Konzentration von 1 mM (0,83 mM Leu) besaß. Die Konzentration von pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 variierte jedoch innerhalb der Serie von 0 nM-5 nM. Weitere Biomoleküle, deren Konzentrationen innerhalb des Experiments konstant gehalten wurden, konnten beispielsweise in den Mastermix in definierter Konzentration vor dem Splitten gegeben werden. Eine Möglichkeit, die Konzentration mehrerer Biomoleküle innerhalb einer Serie zu variieren, bestand darin, einen größeren Volumenanteil bei der Herstellung des Mastermixes einzusparen, die Reaktion dementsprechend in kleinere Volumina zu splitten und weitere Stocklösungen den gesplitteten Reaktionen hinzuzupipettieren.

Die gesplitteten 12 μl -Reaktionen wurden 16 h bei 29 °C inkubiert. Jeweils 10 μl dieser 12 μl -Reaktionen wurden in jeweils ein *Well* einer 384-*Well*-Platte (*Nunc 384-well optical bottom plate, Thermo Fisher Scientific Inc.*, Waltham, USA) gefüllt. Die entsprechenden *Wells* wurden mit einer Klebefolie (*Nunc sealing tape, Thermo Fisher Scientific Inc.*, Waltham, USA) verschlossen. Die Messung der Fluoreszenz-Intensität erfolgte mit dem Mikroplattenleser *POLARstar® Optima (BMG Labtech GmbH, Offenburg)*. Bei der Messung der Fluoreszenz-Intensität betrug die Anregungs- und Emissionswellenlängen 485 nm bzw. 520 nm.

4.2.15. SDS-PAGE

Zur Auftrennung von Proteinen gemäß ihres Molekulargewichtes wurde die *sodium dodecyl sulfate polyacrylamide gel electrophoresis* (SDS-PAGE) durchgeführt, die auf LAEMMLI [263] zurückgeht. Diese Standard-Methode diente der Auftrennung, Visualisierung und Identifizierung der zellfrei synthetisierten Modellproteine. Die gelelektrophoretische Auftrennung der Modellproteine mittels SDS-PAGE, von der eine Realisierung vorab in [194] detailliert beschrieben wurde, erfolgte im purifizierten oder unpurifizierten Zustand (mit den endogenen aus dem Zellextrakt stammenden *E.-coli*-Proteinen vermischt). Alle verwendeten Tris-Glycin-SDS-Gele (4 %-20 %, 10 % und 12 %) wurden von *anamed Gelelektrophorese GmbH* (Groß-Bieberau) als Fertiggele geliefert. Die Zusammensetzung des Tris-Glycin-SDS-Laufpuffers betrug 25 mM Tris, 192 mM Glycerin, 0,1 % SDS, pH 8,3 mit HCl. Der Laufpuffer wurde als Zehnfach-Konzentrat von *anamed Gelelektrophorese GmbH* (Groß-Bieberau) geliefert oder selbst hergestellt und vor der Gelelektrophorese 10-fach mit Reinstwasser verdünnt. Als Probenpuffer wurde ein Zweifach-Konzentrat (*anamed Gelelektrophorese GmbH* (Groß-Bieberau)) verwendet. Als Protein-Standard diente der *Unstained Protein Marker (Broad Range (2 kDa–212 kDa), New England Biolabs GmbH, Frankfurt am Main)*.

Um das zellfreie Reaktionsgemisch, in dem die synthetisierten Modellproteine gelöst waren, gelelektrophoretisch zu trennen, wurde dieses Gemisch 5-10-fach verdünnt, sodass eine adäquate Migration

der Proteine innerhalb der Gelmatrix gewährleistet ist und verhindert wird, dass es nach späterer Färbung der Proteine zu einer Übersättigung kommt. In der Regel wurden dementsprechend 1 µl oder 2 µl einer zellfreien Reaktion mit 4 µl bzw. 8 µl Reinstwasser verdünnt. Lösungen, in denen purifizierte Modellproteine gelöst waren, wurden vor der SDS-PAGE nicht verdünnt. Proben-Lösungen, verdünnt und unverdünnt, wurden im Volumenverhältnis 1:1 mit dem 2x Probenpuffer vermischt, es herrschten die folgenden Pufferbedingungen: 62,5 mM Tris/Cl⁻, 10 % Glycerin, 2 % SDS, 0,0025 % Bromphenolblau, pH 6,8. Diese Mischungen und der Protein-Standard wurden bei 95 °C-100 °C 3 min-5 min lang (*Heizblock Thermostat (HBT) 130-2, HLC - Haep Labor Consult, Bovenden*) erhitzt, um die Proteine zu denaturieren, sodass sich SDS an das Proteinrückgrat lagern konnte. Um die Lösung nach der Erhitzung am Gefäßboden (*Eppendorf AG, Hamburg*) homogen vermischt zu konzentrieren, wurde die Lösung zuerst mittels einer Tisch-Minizentrifuge für etwa 2 s-3 s abzentrifugiert (*Sprout Minizentrifuge, Biozym Scientific GmbH, Hessisch Oldendorf*), gründlich gevortex (*Vortex 3, IKA®-Werke GmbH, Staufen*) und erneut für etwa 2 s-3 s abzentrifugiert. Nach Auftragen der präparierten Proben und des Standards auf das Gel, wurde die SDS-PAGE bei einer Spannung von 125 V bei einer Stromstärke von etwa 20 mA-40 mA mit einer ungefähren Dauer von 90 min durchgeführt. Im Anschluss an die SDS-PAGE wurden die Proteine 30 min in der Gelmatrix fixiert (50 % Methanol, 10 % Essigsäure, 40 % Reinstwasser), 60 min mit Coomassie-Brillant-Blau G-250 (*Biozym Scientific GmbH, Hessisch Oldendorf*) gefärbt (0,025 % Coomassie-Brillant-Blau G-250, 10 % Essigsäure, 90 % Reinstwasser), die Gele anschließend 60 min-120 min entfärbt (10 % Essigsäure, 90 % Reinstwasser), ein Protokoll, das sich vor allem für Proteine kleiner Molekulargewichte eignet [283], [284].

4.2.16. His-Tag-Purifikation zellfrei synthetisierter Modellproteine mittels IMAC

Dieser Absatz führt in die *Immobilized metal ion affinity chromatography* (IMAC) zur Purifikation von Proteinen, die mit einem Polyhistidin-Affinitäts-Tag (His-Tag) ausgestattet sind, anhand von [285], [286] ein. Die Purifikation von Proteinen mit His-Tags geht auf [287] zurück. In der Regel besteht ein His-Tag aus sechs aufeinanderfolgenden Histidinen, deren 18 bp lange DNA-Sequenz entweder N-Terminal oder C-Terminal in das Gen, das für das zu synthetisierende Modellprotein kodiert, kloniert wird. Die Aminosäure His wechselwirkt über ihren Imidazolring mit verschiedenen Übergangsmetallen, die innerhalb einer Matrix immobilisiert werden. Beispielsweise beinhaltet eine solche Matrix Nickel-Nitrilotriessigsäure (Ni²⁺-NTA) [288]. Die Purifikation von mit His-Tag ausgestatteten Proteinen, die unter nativen oder denaturierenden Bedingungen erfolgen kann, beginnt mit dem selektiven Binden dieser Proteine über ihre His-Tags an der Matrix. Ein Waschschriff entfernt die anderen, nicht an die Matrix gebundenen Proteine. Um im Anschluss die gebundenen Proteine von der Matrix zu lösen, kann unter anderem ein Elutionspuffer verwendet werden, der hohe Konzentrationen Imidazol (> 100 mM) beinhaltet. Imidazol konkurriert mit den Imidazolringen der His-Tags um die Bindung an die Matrix und verdrängt somit die gebundenen Proteine.

Die zellfrei synthetisierten deGFP-Moleküle, die C-Terminal einen His₆-Tag besaßen, wurden mittels des *His-Spin Protein Miniprep™*-Kits (*Zymo Research Europe GmbH, Freiburg im Breisgau*) nach Hersteller-Anweisungen purifiziert, um die Modellproteine aus der zellfreien Reaktion zu extrahieren. Die durchgeführte Prozedur wurde bereits vorab in [194] detailliert veröffentlicht und wird an dieser Stelle kurz zusammengefasst: Das verwendete Kit basiert auf IMAC und liefert Zentrifugationssäulen, in denen eine mit Nickel geladene Gelmatrix die mit dem His₆-Tag ausgestatteten deGFP-Moleküle während der Zentrifugation bis zur Elution zurückhält.

Die zellfreien Reaktionen wurden mit dem mitgelieferten Bindungspuffer im Volumenverhältnis 1:1

vermischt. Überstieg das Gesamtvolumen 300 µl, wurde das Gemisch für die Purifikation in gleiche Volumina gesplittet. Sofern das Gemisch gesplittet wurde, verlief die Purifikation für alle gesplitteten Volumina gleich, weswegen die Purifikation anhand eines Gemisches zusammengefasst wird. Nach Präparation der Zentrifugationssäule wurde das Gemisch in die Säule transferiert. Die anschließende Zentrifugation (*MiniSpin[®] Zentrifuge, Eppendorf AG, Hamburg*) bei maximaler Drehzahl entfernte die nicht an die Matrix bindenden Moleküle, es verblieb das gebundene deGFP. Nach zweifachem Waschen mit dem mitgelieferten Wasch-Puffer, wurde das gebundene deGFP mittels des mitgelieferten Elutionspuffers von der Matrix gelöst und in diesem Puffer in einem 1,5 ml Reaktionsgefäß (*Eppendorf AG, Hamburg*) eluiert.

4.2.17. Pufferaustausch und Aufkonzentrierung

Nach der in Abschn. 4.2.16 beschriebenen His-Tag-Purifikation waren die purifizierten Modellproteine in 50 mM Natriumphosphatpuffer (pH 7,7), 300 mM NaCl und 250 mM Imidazol eluiert (Elutionspuffer vom *His-Spin Protein Miniprep[™]*-Kit, *Zymo Research Europe GmbH, Freiburg im Breisgau*). Dieser Puffer wurde für die massenspektroskopische Analyse der Modellproteine ausgetauscht, sodass nach dem Austausch folgende Pufferbedingungen herrschten: 50 mM Tris-Cl (pH 8), 100 mM NaCl, 10% Glycerin (Proteinlagerungspuffer). Der Pufferaustausch erfolgte nach Hersteller-Instruktionen mit Hilfe der *CentriPure Z25 Mini Spin Columns (Genaxxon bioscience GmbH, Ulm)* und wurde bereits vorab in [194] detailliert veröffentlicht, sodass er an dieser Stelle nur kurz zusammengefasst wird. Der Pufferaustausch zielte darauf ab, ein besseres Signal-Rausch-Verhältnis während der massenspektroskopischen Analyse der Modellproteine zu erzielen, da hohe Konzentrationen an Imidazol (> 150 mM) oder Salzen wie NaH₂PO₄ (> 300 mM) oder NaCl (> 50 mM) zu hohem Hintergrundrauschen führen können [289]. Die Gelfiltrationssäulen ermöglichen mittels Zentrifugation die Abtrennung kleinerer Moleküle oder Salze von Proteinen mit Molekulargewichten von über 5 kDa. Die Gelmatrix bindet kleinere Moleküle oder Salze länger als große Moleküle, die somit später eluiert werden.

Nachdem die Gelfiltrationszentrifugationssäulen auf Raumtemperatur aufgewärmt waren, erfolgte ein Zentrifugationsschritt bei 1000 x g für 2 min (*Microfuge[®] 22R Centrifuge, kühlbar, Beckman Coulter GmbH, Krefeld*), der einen Großteil der Salze und kleinen Moleküle des Gellagerungspuffers, mit dem das Gel hydriert war, entfernte. Um die restlichen Salze und kleinen Moleküle aus der Gelmatrix zu entfernen und die Gelmatrix mit den Salzen des Proteinlagerungspuffers zu beladen, wurden 400 µl des Proteinlagerungspuffers auf das Gelbett pipettiert, es folgte ein Zentrifugationsschritt (1000 x g, 2 min). Weitere 400 µl vom Proteinlagerungspuffer wurden auf das Gelbett aufgebracht, es folgte ein erneuter Zentrifugationsschritt (1000 x g, 2 min). Abschließend wurden die im Elutionspuffer des *His-Spin Protein Miniprep[™]*-Kits gelösten Modellproteine auf das Gelbett aufgetragen und durch einen weiteren Zentrifugationsschritt (1000 x g, 2 min) im Proteinlagerungspuffer eluiert, während die kleineren Moleküle und Salze des Elutionspuffers in der Gelmatrix zurückblieben. Für größere Volumina als 100 µl wurden mehrere Säulen zum Pufferaustausch verwendet. Die Prozedur verlief wie in diesem Absatz beschrieben, nur dass die im Elutionspuffer des *His-Spin Protein Miniprep[™]*-Kits gelösten Modellproteine auf mehrere Säulen verteilt wurden.

Gegebenenfalls wurden die Proteinlösungen mit Hilfe eines Vakuumkonzentrators *Concentrator 5301 (Eppendorf AG, Hamburg)* bei einer Temperatur von 30 °C für die anschließende HPLC-ESI Massenspektroskopie aufkonzentriert. Als für die massenspektrometrische Analyse optimal stellten sich Proteinkonzentrationen zwischen 200 µg/ml-500 µg/ml und Probenvolumina zwischen 20 µl-50 µl heraus. Im Fall einer Aufkonzentrierung wurde der Proteinlagerungspuffer vor dem Pufferaustausch um

den Faktor verdünnt, um den das Volumen durch die Aufkonzentrierung verringert wurde, sodass die oben angegebenen Pufferkonzentrationen (50 mM Tris-Cl (pH 8), 100 mM NaCl, 1 % Glycerin) galten.

4.2.18. HPLC-ESI massenspektrometrische und fluoreszenzspektroskopische Analyse der Modellproteine

Die HPLC-ESI Massenspektrometrie und die Fluoreszenzspektroskopie wurden von Matthias P. Exner in der Kooperationsarbeitsgruppe von Nediljko Budisa (Institut für Chemie, Technische Universität Berlin), wie bereits in Matthias P. Exners Dissertation [154] und den gemeinsamen Publikationen [193], [194] beschrieben, von ihm durchgeführt.

4.3. Ergebnisse und Diskussion

4.3.1. Optimierung der zellfreien Reaktion

Die Genexpressions-Effizienz des verwendeten zellfreien Systems hängt von den transkriptions-translations-regulatorischen Elementen des benutzten DNA-Konstrukts ab [51]. Beim verwendeten Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-deGFPHis₆-Tag-T500 sättigt die Proteinausbeute bei einer Konzentration von ca. 10 nM [51]. Sofern nicht anders angegeben, wurde diese Konzentration beim seitenkettenspezifischen Einbau nicht-kanonischer Aminosäuren verwendet. Im verwendeten zellfreien System besitzt jede kanonische Aminosäure allgemein eine Konzentration zwischen 0,5 mM und 1,5 mM (Leu 0,4 mM-1,3 mM) [51]. Wenn nicht anders erwähnt, besaß in den Expressionsexperimenten in Kap. 4 jede Aminosäure eine Konzentration von 1 mM (Leu ca. 0,8 mM) in den zellfreien Reaktionen. Die Expressionseffizienz lässt sich durch die Hinzugabe von zusätzlichem Magnesium- und Kaliumglutamat optimieren [51], [52]. Die optimalen Konzentrationen dieser beiden Stoffe können von Produktionsreihe zu Produktionsreihe variieren und sollen für jede Produktionsreihe optimiert werden [51], [52]. Kalibrierungsmessungen für den verwendeten Zellextrakt zeigen, dass sich die optimalen Konzentrationen für zusätzliches Magnesium- und Kaliumglutamat zwischen 5 mM-8 mM bzw. 8 mM-20 mM in der zellfreien Reaktion bewegen (Abb. 4.7). In den Expressionsexperimenten in Kap. 4 wurde eine Magnesiumglutamat- und Kaliumglutamat-Konzentration von 6 mM bzw. 10 mM eingestellt.

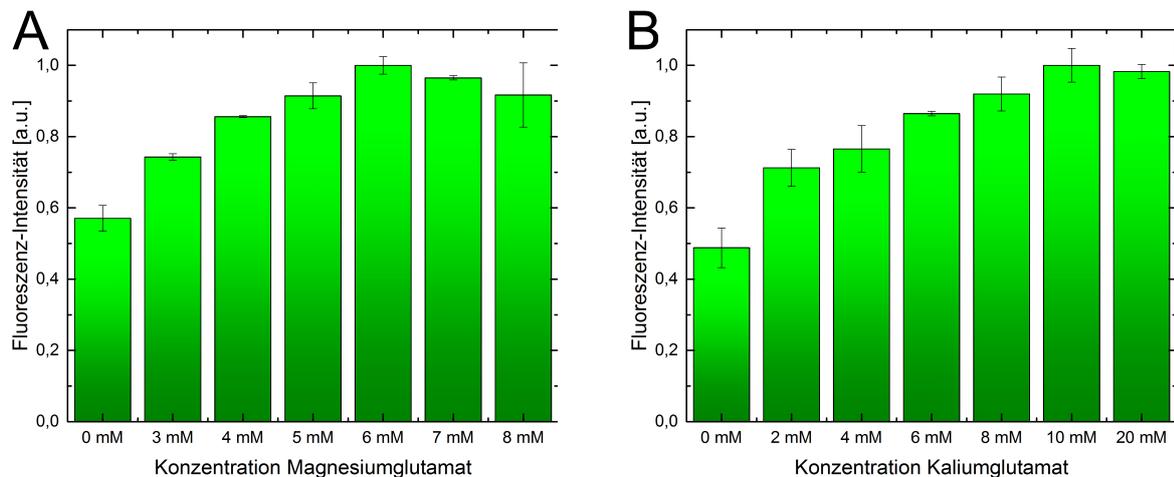


Abbildung 4.7.: Kalibrierungsmessungen nach [51], [52] zur Ermittlung eines optimalen Konzentrationsbereichs für zusätzlich zur zellfreien Reaktion hinzugegebenes Magnesium- (A) und Kaliumglutamat (B). Die Fluoreszenz-Intensitäten sind die Mittelwerte aus zwei Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung. Die Fluoreszenz-Intensitäten sind in (A) und (B) jeweils auf den höchsten Mittelwert normiert.

4.3.2. Seitenkettenspezifischer Einbau nicht-kanonischer Aminosäuren in deGFP

Das zellfrei synthetisierte deGFPHis₆-Tag wird in Abschn. 4.3.2 und 4.4 vereinfacht als deGFP bezeichnet.

4.3.2.1. Seitenkettenspezifischer Einbau von Can in deGFP

Bei einem vollständigen seitenkettenspezifischen Einbau von Can in deGFP substituiert Can Arg an jeder der insgesamt sechs Positionen in diesem Modellprotein.

4.3.2.1.1. Gelelektrophoretische Analyse des seitenkettenspezifischen Einbaus von Can in deGFP

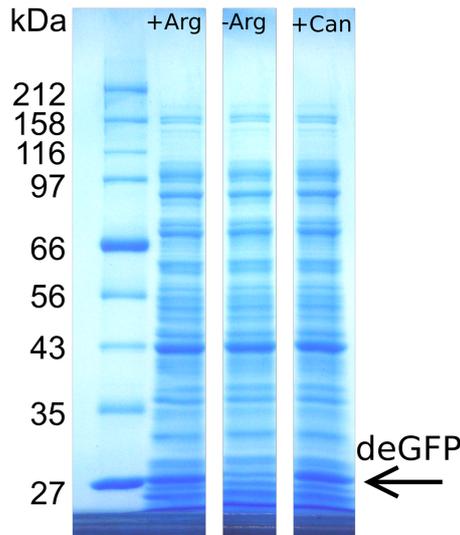


Abbildung 4.8.: SDS-PAGE zur vorläufigen Evaluierung des seitenkettenspezifischen Einbaus von Can in deGFP. Von links nach rechts: Auftrennung des Proteinstandards, der Referenz-Reaktion, Negativkontrolle und Can-Reaktion. Abb. 4.8 modifiziert aus [193], [194] übernommen.

Es wurden drei zellfreie Reaktionen angesetzt. Der ersten Reaktion (Referenz, +Arg) wurden alle 20 kanonischen Aminosäuren, die zur zellfreien Synthese eines unmodifizierten deGFP-Moleküls benötigt werden, hinzugegeben. Dem zweiten zellfreien Ansatz wurden alle kanonischen Aminosäuren außer Arg beigefügt (Negativkontrolle, -Arg). Die dritte zellfreie Reaktion setzte sich wie die erste Reaktion zusammen, nur dass dort Arg durch Can substituiert wurde (Can-Reaktion, +Can).

Abb. 4.8 zeigt die gelelektrophoretische Auftrennung der drei zellfreien Reaktionen nach Übernacht-Inkubation bei 29 °C. Sowohl in der Referenz-, als auch in der Can-Reaktion ließ sich deGFP nachweisen (+Arg- und +Can-Bahn in Abb. 4.8), während in der Negativkontrolle keine deGFP-Proteinbande auf der zu erwartenden Höhe zu sehen war (-Arg-Bahn in Abb. 4.8). Demnach waren im zellfreien System keine Arg-Reste enthalten oder Arg war so gering konzentriert, dass es zu keiner detektierbaren deGFP-Synthese kam. Insofern war davon auszugehen, dass in der Can-Reaktion der seitenkettenspezifische Einbau von Can in deGFP erfolgte. Allerdings musste

im Folgenden per Massenspektroskopie überprüft werden, ob der vollständige seitenkettenspezifische Einbau von Can möglich ist.

4.3.2.1.2. HPLC-ESI massenspektroskopische Analyse des seitenkettenspezifischen Einbaus von Can in deGFP

Eingelassen in Abb. 4.9 ist die gelelektrophoretische Auftrennung von His-Tag-purifizierten und umgepufferten deGFP-Molekülen, die aus einer zellfreien Can-Reaktion und einer entsprechenden Referenz-Reaktion nach Übernacht-Inkubation bei 29 °C hervorgingen, zu sehen. Die deGFP-Molekül-Massen wurden per HPLC-ESI Massenspektroskopie bestimmt.

Nach richtiger Faltung, Fluorophor-Bildung und der damit verbundenen reduzierten Masse (vgl. Abschn. 2.5) besitzt das native deGFP^{6 Arg} eine errechnete Masse von 26192,3 Da. Die Molekulargewichte von Arg (174,2 Da) und Can (176,18 Da) unterscheiden sich um 1,98 Da. Mit jedem durch Can substituierten Arg vergrößert sich somit die Masse des modifizierten deGFP-Moleküls um 1,98 Da, sodass die errechnete Masse des deGFP^{6 Can}-Moleküls, in der Can alle 6 Arg ersetzt, bei 26204 Da liegt. Abb. 4.9

zeigt die dekonvolutionierten Massenspektren der aufgereinigten deGFP-Moleküle. Das rote Spektrum korrespondiert zu deGFP, das aus der Referenz-Reaktion stammte. Das schwarze Spektrum gehört zu deGFP, das in der zellfreien Can-Reaktion synthetisiert wurde. Die gemessene deGFP-Masse aus der Referenz-Reaktion betrug 26192,8 Da (Haupt-Peak) und stimmte mit der errechneten Masse für das native deGFP im Rahmen des Dekonvolutions-Fehlers überein. Für deGFP aus der Can-Reaktion wurde eine Masse von 26202,5 Da gemessen, die im Rahmen des Dekonvolutions-Fehlers mit der errechneten Masse für das deGFP^{6Can} übereinstimmte. Die beiden Nebenpeaks bei 26213,1 Da und 26223,3 Da gehörten zu den jeweils um 20 Da schwereren unmodifizierten deGFP^{6Arg}- bzw. modifizierten deGFP^{6Can}-Molekülen, die kein Fluorophor ausgebildet hatten.

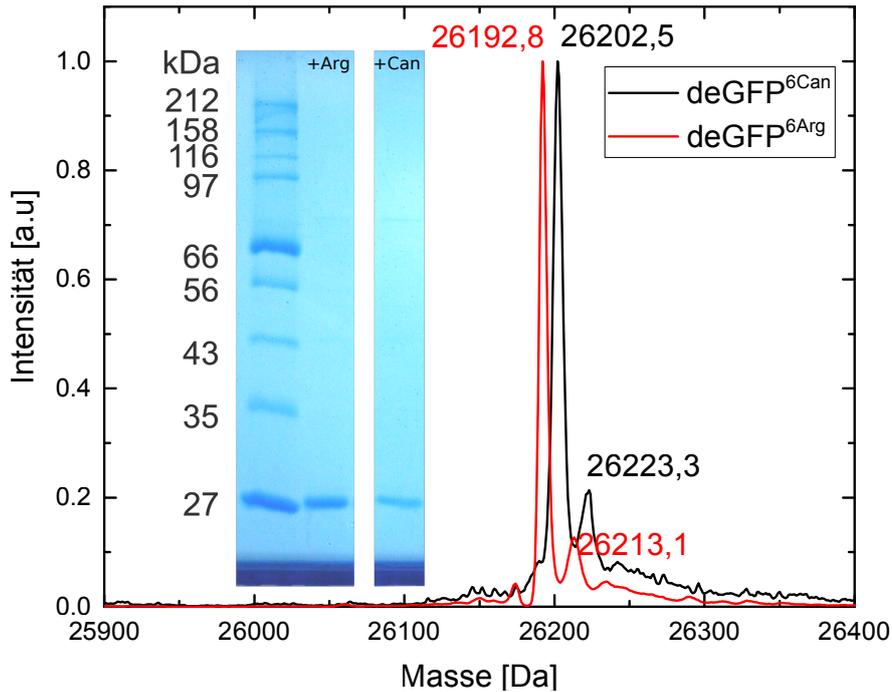


Abbildung 4.9.: Dekonvolutionierte Massenspektren zum Nachweis des vollständigen Einbaus von Can an allen 6 Arg-Stellen in deGFP und gelelektrophoretische Auftrennung per SDS-PAGE der aufgereinigten deGFP-Moleküle (eingelassen). Die errechneten Massen von deGFP^{6Arg} und deGFP^{6Can} betragen 26193 Da bzw. 26204 Da. Die entsprechenden gemessenen Massen (Haupt-Peaks) besaßen 26192,8 Da (rotes Spektrum) und 26202,5 Da (schwarzes Spektrum) und lagen somit innerhalb des Fehlers, der durch die Dekonvolution des Spektrums entstand. Die beiden Nebenpeaks bestätigten ebenso den vollständigen Can-Einbau. Sie sind deGFP^{6Arg} und deGFP^{6Can} zuzuordnen, die kein Fluorophor gebildet hatten und somit 20 Da schwerer waren. SDS-PAGE von links nach rechts: Proteinstandard, aufgereinigte, umgepufferte deGFP-Moleküle aus Referenz- und Can-Reaktion. Das Spektrum ist auf seinen höchsten Wert normiert und wurde modifiziert aus [193], [194] übernommen. Gel-Bild modifiziert aus [193], [194] übernommen. Massenspektroskopie von Matthias P. Exner durchgeführt.

Der Zell-Extrakt musste nicht wie in [235], [236] von Resten der kanonischen Aminosäure depletiert werden, um Can vollständig in deGFP einzubauen. Es mussten keine auxotrophen *E. coli* wie in [236] verwendet werden, sondern der Extrakt wurde aus Standard-Zellen gewonnen (vgl. Abschn. 4.2.3).

4.3.2.2. *In-vitro*- vs. *in-vivo*-Expression mit der toxischen Aminosäure Can

Im Folgenden wurde der zellfreie *in-vitro*-Einbau von Can in deGFP mit der *in-vivo*-Inkorporation von Can in EGFP per SPI-Methode verglichen. Die Experimente zum *in-vivo*-Einbau von Can in

EGFP, das eine Arg-Position im Vergleich zu deGFP mehr besitzt, wurden von Alessandro De Simone durchgeführt (experimentelle Details in [193]), die Ergebnisse analysiert und interpretiert.

Die massenspektroskopische Analyse von EGFP in Abb. 4.10 belegt, dass auch dort ein vollständiger Einbau von Can in das Zielprotein möglich war. Die errechnete Masse des EGFP^{7Can}-Moleküls, in dem sich Can an allen sieben Arg-Positionen befindet, beträgt 28333 Da. Die gemessene Masse stimmte mit der errechneten Masse überein (Abb. 4.10)⁶.

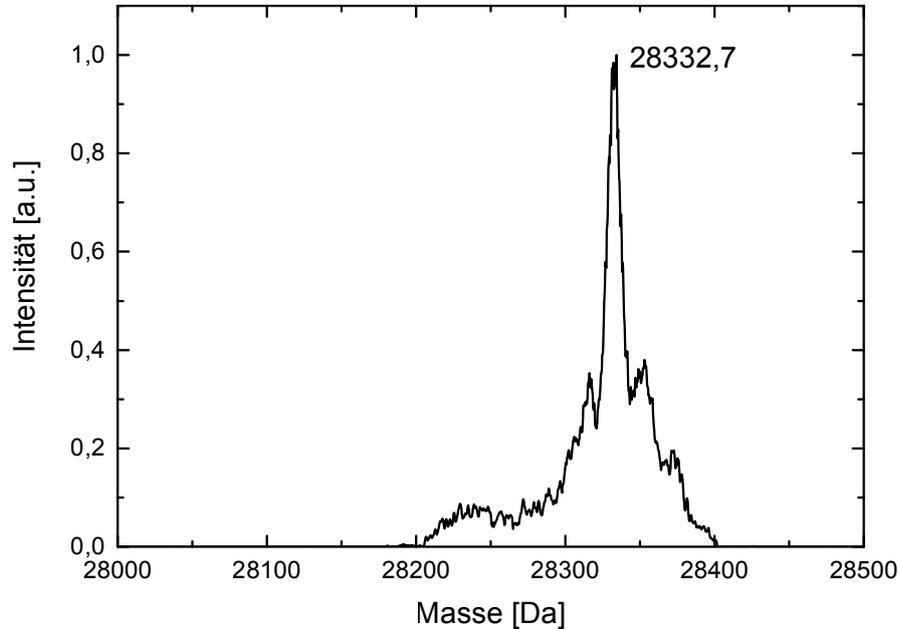


Abbildung 4.10.: Dekonvolutionsmassenspektrum zum Nachweis des vollständigen Einbaus von Can an allen sieben Arg-Positionen in EGFP. Die errechnete Masse des EGFP^{7Can}-Moleküls beträgt 28333 Da, mit der die gemessene Masse von EGFP^{7Can} übereinstimmt. Die Messung sowie die Analyse und Interpretation der Daten wurde von Alessandro De Simone durchgeführt. Das Spektrum wurde auf seinen höchsten Wert normiert. Abb. 4.9 modifiziert aus [193] übernommen.

In einigen *in-vivo*-Expressionen in Präsenz von Can konnte auch EGFP mit einer reduzierten Masse nachgewiesen werden⁷. Die reduzierte Masse wich 5 Da von der errechneten Masse des EGFP^{7Can} ab, bei dem das vollständig eingebaute Can zu Homoserin und Hydroxyguanidin degradiert wird und einem gut beschriebenen Reaktionsweg [290], [291] folgt. Die größere Abweichung könnte sich durch den größeren Dekonvolutionsfehler aufgrund des schlechten Signal-Rausch-Verhältnisses ergeben. Hingegen konnte bei keiner der massenspektroskopischen Analysen des zellfrei synthetisierten deGFP^{6Can}-Moleküls eine entsprechend reduzierte Masse detektiert werden. Insofern scheint die zellfreie, molekulare *E.-coli*-Maschinerie das eingebaute Can im Gegensatz zur lebenden Zelle zu stabilisieren.

Die zellfreie *in-vitro*-Proteinausbeute pro Liter Zellkultur wurde basierend auf einer zellfreien Reaktion eines Volumens von 90 µl, die aus etwa 50 ml Zellkultur gewonnen wurde, hochgerechnet und mit der *in-vivo*-Ausbeute verglichen: Typischerweise ließen sich mit dem zellfreien *in-vitro*-Ansatz etwa 0,25 mg deGFP^{6Can} pro Liter *E.-coli*-Zellkultur synthetisieren, während die SPI-Methode eine Ausbeute von 0,1 mg pro Liter *E.-coli*-Zellkultur lieferte (vgl. [193]). Das zellfreie Expressionssystem ermöglichte

⁶Das Referenz-Spektrum, das die Masse des nativen EGFP^{7Arg}-Moleküls zeigt, kann in [193] gefunden werden.

⁷Ein entsprechendes Spektrum ist in [193] abgebildet.

demnach, stabile Can-modifizierte Modellproteine für weitere Studien in kleinen Volumina herzustellen und lieferte typischerweise sogar höhere Ausbeuten als *in vivo*.

Der *in-vitro*-Ansatz ermöglicht des Weiteren die effiziente Produktion von stabilen Can-modifizierten Modellproteinen, während er die Grenzen von *in-vivo*-Systemen, die aus der Can-bedingten Toxizität oder der starken Abhängigkeit der mRNA-Sequenz bei Einzel-Protein-Produktions-Systemen [50] entstehen, überwindet. Darüber hinaus müssen keine speziellen Arg-auxotrophen *E.-coli*-Stämme zur Extrakt-Herstellung verwendet werden.

4.3.2.2.1. Fluoreszenzspektroskopische Analyse des nativen und modifizierten deGFP-Moleküls

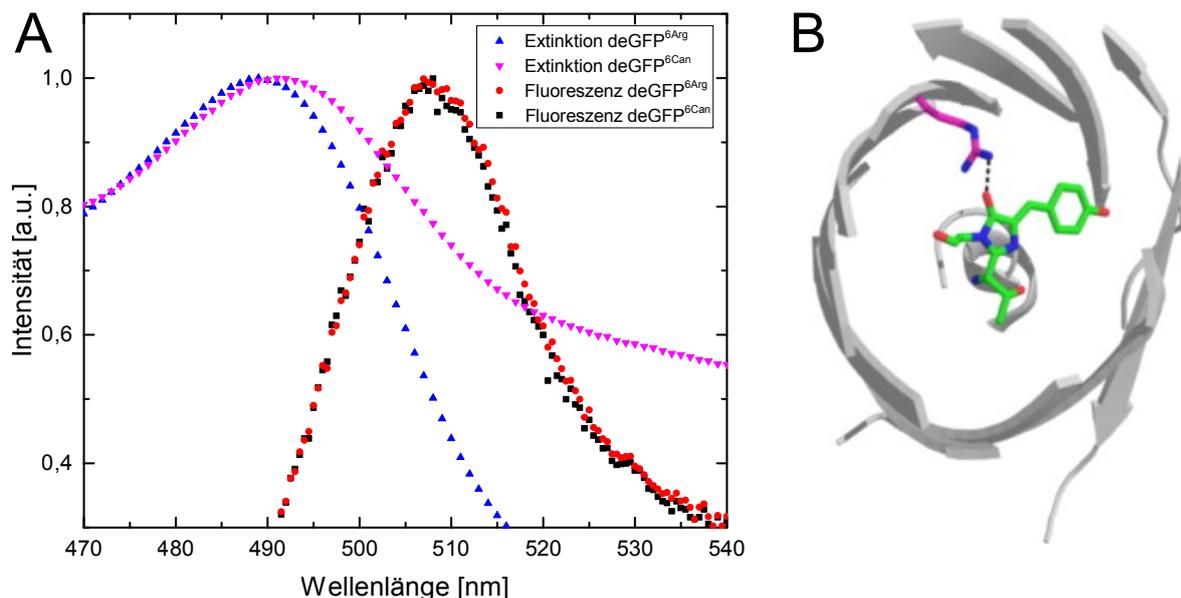


Abbildung 4.11.: Fluoreszenzspektroskopische Analyse von deGFP^{6Arg} und deGFP^{6Can} (A) sowie Draufsicht der Proteinstruktur von EGFP (B). A: Durch den Einbau von Can in deGFP war das Extinktionsmaximum leicht rot verschoben (491 nm vs. 489 nm), während die Emissionsspektren (Maximum bei 508 nm) keine Unterschiede aufwiesen. Fünf Spektren wurden von jeder Probe aufgenommen und auf den höchsten Wert normiert. B: Kristallstruktur von EGFP (Draufsicht, (PDB ID: 2Y0G) nach [292]). Das Fluorophor (grün) und die Seitenkette eines Arg bzw. Can (magentafarben) sind im polaren Kontakt miteinander. Fluoreszenzmessungen und Normierung von Matthias P. Exner durchgeführt. Abb. B von Matthias P. Exner erstellt. Abb. 4.11 modifiziert aus [193] übernommen.

Im Folgenden wurde untersucht, ob sich die fluoreszenzspektroskopischen Eigenschaften von deGFP durch den Einbau von Can verändern.

Abb. 4.11A zeigt die Extinktions- und Emissionsspektren von deGFP^{6Arg} und deGFP^{6Can}. Während sich die Emissionsspektren von deGFP^{6Arg} (rote Punkte) und deGFP^{6Can} (schwarze Quadrate) nicht unterschieden, war das Extinktionsmaximum von deGFP^{6Can} (magentafarbene Dreiecke) im Vergleich zu deGFP^{6Arg} (blaue Dreiecke) leicht rotverschoben (491 nm vs. 489 nm). Obwohl sich nur das Extinktionsmaximum von deGFP durch den Einbau von Can leicht verschob, nahm die Fluoreszenzintensität des modifizierten Proteins deutlich ab. Dies und das leicht rotverschobene Extinktionsmaximum könnten auf die Präsenz einer Can-Seitenkette, die im polaren Kontakt zum Fluorophor steht, zurückzuführen sein (Abb. 4.11B).

4.3.2.3. Seitenkettenspezifischer Einbau von Hyl in deGFP

Bei einem vollständigen seitenkettenspezifischen Einbau von Hyl in deGFP, substituiert Hyl Lys an jeder der insgesamt 18 Positionen in diesem Modellprotein.

4.3.2.3.1. Gelelektrophoretische Analyse des seitenkettenspezifischen Einbaus von Hyl in deGFP

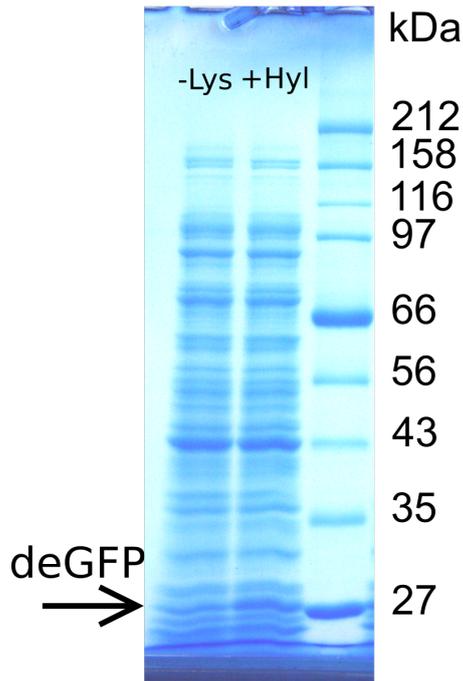


Abbildung 4.12.: SDS-PAGE zur vorläufigen Evaluierung des seitenkettenspezifischen Einbaus von Hyl in deGFP. Von links nach rechts: Auftrennung der Negativkontrolle, der Hyl-Reaktion und des Proteinstandards. Abb. 4.12 modifiziert aus [194] übernommen.

Es wurden zwei zellfreie Reaktionen präpariert. Der ersten zellfreien Reaktion wurden alle kanonischen Aminosäuren außer Lys hinzugegeben (Negativkontrolle, -Lys). In die zweite zellfreie Reaktion wurden alle kanonischen Aminosäuren außer Lys gegeben, jedoch wurde Hyl nun ergänzt (Hyl-Reaktion).

Abb. 4.12 zeigt die gelelektrophoretische Auftrennung der zwei zellfreien Reaktionen nach Übernacht-Inkubation (29 °C). In der Hyl-Reaktion zeigte sich eine Proteinbande, die deGFP zugeordnet werden kann. Allerdings konnte auch eine schwache Bande auf der gleichen Höhe bei der Negativkontrolle beobachtet werden. Dies legt den Schluss nahe, dass sich geringe Lys-Reste im zellfreien Expressions-System befanden. Sofern sich Lys-Reste im zellfreien Expressionssystem befunden haben sollten, hätten Lys und Hyl in der Hyl-Reaktion um die Bindung an die Lysyl-aaRS und somit um den Einbau in deGFP konkurriert. Anschließend wurde per Massenspektroskopie diese Hypothese überprüft.

4.3.2.3.2. HPLC-ESI massenspektroskopische Analyse des seitenkettenspezifischen Einbaus von Hyl in deGFP

Für die massenspektroskopische Analyse wurde eine zellfreie Hyl-Reaktion präpariert und über Nacht bei 29 °C inkubiert. Das synthetisierte deGFP wurde vor

der HPLC-ESI Massenspektroskopie His-Tag-purifiziert und umgepuffert. Eingelassen in Abb. 4.13 wird eine typische gelelektrophoretische Auftrennung der zu analysierenden deGFP-Moleküle vor der Massenspektroskopie gezeigt.

Die Molekulargewichte von Lys (146,19 Da) und Hyl (162,19 Da) unterscheiden sich um 16 Da. Mit jedem durch Hyl substituierten Lys vergrößert sich somit die Masse des modifizierten deGFP-Moleküls um 16 Da, sodass die errechnete Masse des deGFP^{18Hyl}-Moleküls, in der Hyl alle 18 Lys ersetzt, bei 26481 Da liegt. Abb. 4.13 zeigt das dekonvolutionierte Massenspektrum der His-Tag-purifizierten deGFP-Moleküle aus der Hyl-Reaktion. Der Haupt-Peak (26193,2 Da) und der Neben-Peak (26211,6 Da) in Abb. 4.13A können den nativen deGFP^{18Lys}-Proteinen zugeordnet werden, die ihr Fluorophor ausgebildet bzw. nicht gebildet hatten (errechnete Massen: 26193 Da bzw. 26213 Da).

Dies bestätigt die Hypothese, dass sich Lys-Reste im zellfreien System befanden, die vorzugsweise auf die Lysyl-aaRS geladen und in deGFP eingebaut wurden. Eine Vergrößerung des Spektrums im Bereich hinter dem Haupt-Peak (Abb. 4.13B) zeigt, dass jede mögliche deGFP-Spezies entstand (deGFP^{18Hyl}, deGFP^{17Hyl+1Lys}, ..., deGFP^{2Hyl+16Lys}). Allerdings überlappt der Peak, der deGFP^{1Hyl+17Lys} zugeordnet werden kann, mit dem Peak der nativen deGFP-Moleküle, die kein Fluorophor gebildet hatten. Aufgrund des schwachen Signal-Rausch-Verhältnisses wichen die gemessenen Massen teilweise mehr als 2 Da von der errechneten Masse ab.

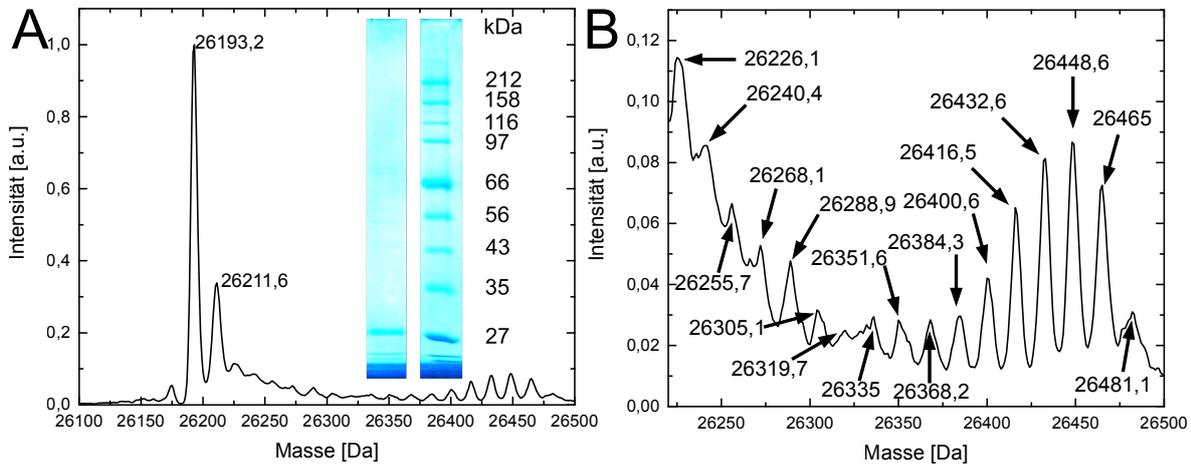


Abbildung 4.13.: Dekonvolutionsiertes Massenspektrum in unterschiedlicher Vergrößerung der His-Tag-purifizierten und umgepufferten deGFP-Moleküle, die über Nacht bei 29 °C in der Hyl-Reaktion synthetisiert wurden und typische gelelektrophoretische Auftrennung per SDS-PAGE entsprechender Proteine (eingelassen in A). A: Der Haupt-Peak (26193,2 Da) und der Neben-Peak (26211,6 Da) können dem nativen deGFP^{18Lys} zugeordnet werden, das sein Fluorophor ausgebildet oder nicht gebildet hatte (errechnete Massen: 26193 Da bzw. 26213 Da). SDS-PAGE von links nach rechts: aufgereinigte und umgepufferte deGFP-Moleküle der Hyl-Reaktion und Proteinstandard. B: Die Vergrößerung des Bereichs zwischen 26220 Da und 26500 Da zeigt, dass alle möglichen deGFP-Spezies synthetisiert wurden (deGFP^{18Hyl}, deGFP^{17Hyl+1Lys}, ..., deGFP^{2Hyl+16Lys}). Der Peak, der zu deGFP^{1Hyl+17Lys} gehört, überlappt mit dem Peak vom nativen deGFP, das kein Fluorophor gebildet hatte. Aufgrund des schwachen Signal-Rausch-Verhältnisses wichen die Massen teilweise mehr als 2 Da von ihrer errechneten Masse ab. Abb. 4.13 modifiziert aus [194] übernommen. Massenspektroskopie von Matthias P. Exner durchgeführt. Das Spektrum ist auf seinen höchsten Wert normiert.

Dass der seitenkettenspezifische Hyl-Einbau in deGFP nicht vollständig war, lag an den im zellfreien System präsenten Lys-Resten. Insofern handelte es sich nicht um ein systematisches Problem wie beispielsweise, dass die nicht-kanonische Aminosäure nicht von der zellfreien Translations-Maschinerie erkannt wurde. Für den vollständigen seitenkettenspezifischen Einbau von Hyl müssten die Lys-Reste zuvor aus dem zellfreien System entfernt werden.

4.3.2.4. Seitenkettenspezifischer Einbau von Aha in deGFP

Bei einem vollständigen seitenkettenspezifischen Einbau von Aha in deGFP substituiert Aha Met an jeder der insgesamt fünf Positionen, inklusive dem N-Terminalen Start-Met im Modellprotein.

4.3.2.4.1. Gelelektrophoretische Analyse des seitenkettenspezifischen Einbaus von Aha in deGFP

Es wurden drei zellfreie Reaktionen angesetzt. Der ersten Reaktion (Referenz, +Met) wurden alle 20

kanonischen Aminosäuren, die zur zellfreien Synthese eines unmodifizierten deGFP-Moleküls benötigt werden, hinzugegeben. Dem zweiten zellfreien Ansatz wurden alle kanonischen Aminosäuren außer Met beigefügt (Negativkontrolle, -Met). Die dritte zellfreie Reaktion setzte sich wie die erste Reaktion zusammen, nur dass dort Met durch Aha substituiert wurde (Aha-Reaktion, +Aha). Die Konzentration von pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP^{His-Tag}-T500 lag höher als sonst, bei 10,7 nM.

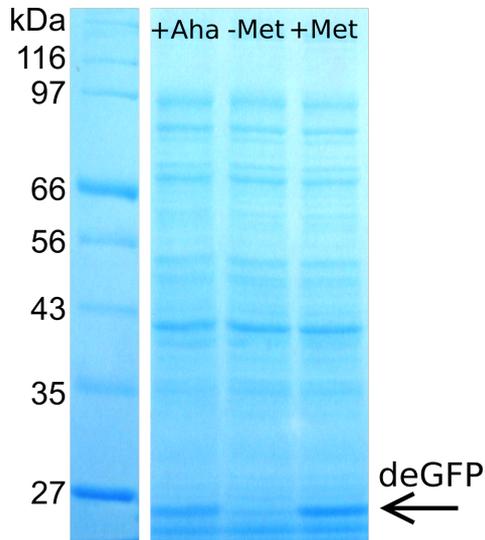


Abbildung 4.14.: SDS-PAGE zur vorläufigen Evaluierung des seitenkettenspezifischen Einbaus von Aha in deGFP. Von links nach rechts: Auftrennung des Proteinstandards, der Aha-Reaktion, Negativkontrolle und Referenz-Reaktion.

Abb. 4.14 zeigt die gelelektrophoretische Auftrennung der drei zellfreien Reaktionen nach Übernacht-Inkubation bei 29 °C. Sowohl in der Referenz- als auch in der Aha-Reaktion wurden Proteine synthetisiert, deren Proteinbanden deGFP zugeordnet werden können. Im Fall der Negativkontrolle ließ sich schwer sagen, ob deGFP synthetisiert wurde oder nicht, da die Intensität der Proteinbanden im Allgemeinen schwach war. Im Folgenden wurde per Massenspektroskopie überprüft, ob der vollständige seitenkettenspezifische Einbau von Aha in deGFP möglich ist.

4.3.2.4.2. HPLC-ESI massenspektroskopische Analyse des seitenkettenspezifischen Einbaus von Aha in deGFP

Für die massenspektroskopische Analyse wurde eine zellfreie Aha-Reaktion präpariert, über Nacht (29 °C) inkubiert. Das synthetisierte deGFP wurde vor der HPLC-ESI Massenspektroskopie His-Tag-purifiziert und umgepuffert. Eingelassen in Abb. 4.15 wird die gelelektrophoretische Trennung der deGFP-Moleküle gezeigt.

Das Molekulargewicht von Met (149,21 Da) und Aha (144,06 Da) unterscheidet sich um 5,15 Da. Mit jedem durch Aha substituierten Met verkleinert sich somit die Masse des modifizierten deGFP-Moleküls um 5,15 Da. Abb. 4.15 zeigt die massenspektroskopische Analyse der His-Tag-purifizierten und umgepufferten deGFP-Moleküle. Es konnte kein deGFP^{5 Aha} detektiert werden, bei dem Aha an allen fünf Met-Positionen eingebaut wurde (erwartete Masse 26167 Da). Der Hauptpeak kann deGFP^{1 Met+4 Aha} mit vier von fünf Aha-Substitutionen zugeordnet werden (errechnete Masse: 26172 Da, gemessene Masse: 26173,4 Da).

Das Massenspektrum zeigt Neben-Peaks, deren Zuordnung schwierig ist. Die gemessene Masse 26189,8 Da ist ca. 2 Da größer als die errechnete Masse eines deGFP^{4 Met+1 Aha}-Moleküls, bei dem eines der fünf Met-Moleküle durch Aha ersetzt wurde (errechnete Masse: 26187,9 Da). deGFP^{1 Met+4 Aha}, das kein Fluorophor ausgebildet hat, besitzt eine errechnete Masse von 26192 Da und könnte im Rahmen eines größeren Dekonvolutionsfehlers aufgrund eines schlechten Signal-Rausch-Verhältnisses auch diesem Peak zugeordnet werden. Wird ein großer Dekonvolutionsfehler in Betracht gezogen, muss zudem die Möglichkeit berücksichtigt werden, dass auch das native deGFP^{5 Met} zu diesem Peak beitragen könnte (errechnete Masse: 26193 Da). Die gemessene Masse 26201,2 Da könnte einem deGFP^{3 Met+2 Aha} zugeordnet werden, bei dem zwei von fünf Met-Molekülen durch Aha substituiert wurden und das kein Fluorophor ausgebildet hatte (errechnet: 26203 Da).

Derseitenkettenspezifische Einbau von Aha in deGFP war nicht vollständig, da das zellfreie System nicht frei von Met-Resten war. Anscheinend waren diese Reste geringer konzentriert als die Hyl-Reste, da der Haupt-Peak in diesem Fall $\text{deGFP}^{4\text{Met}+1\text{Aha}}$ und nicht dem nativen $\text{deGFP}^{5\text{Met}}$ zugeordnet werden kann. Die hohe Intensität dieses Peaks lässt hingegen die Frage nach einem Peak aufkommen, der mit $\text{deGFP}^{5\text{Aha}}$ assoziiert werden kann. Selbst im Hyl-Einbau-Experiment, in dem der dominante Peak dem nativen $\text{deGFP}^{18\text{Lys}}$ zugeordnet werden kann, wurden vollständig modifizierte $\text{deGFP}^{18\text{Hyls}}$ -Spezies detektiert, in der Lys an allen Positionen substituiert wurde. Beim seitenkettenspezifischen *in-vivo*-Einbau von Aha ist die Translations-Initiation mit Aha möglich [293]. Eine eventuelle Erklärung für die fehlende Synthese eines deGFP -Moleküls, in dem Met vollständig durch Aha substituiert wurde, könnte auf die Unfähigkeit der zellfreien molekularen Komposition zurückzuführen sein, die Translation mit Aha zu initiieren. Bei der Präparation des Extraktes könnten wichtige Faktoren verloren gehen, die der lebenden Zelle für die Initiierung der Translation mit Aha zur Verfügung stehen. Um diese Frage zu klären, sind weitere zellfreie Experimente unter ansonsten gleichen Bedingungen notwendig, in denen beispielsweise ein Modellprotein synthetisiert wird, bei dem das N-Terminale Met nach der Translation abgespalten wird. In diesem Fall sollte zumindest ein Teil der synthetisierten Modellproteine vollständig modifiziert werden.

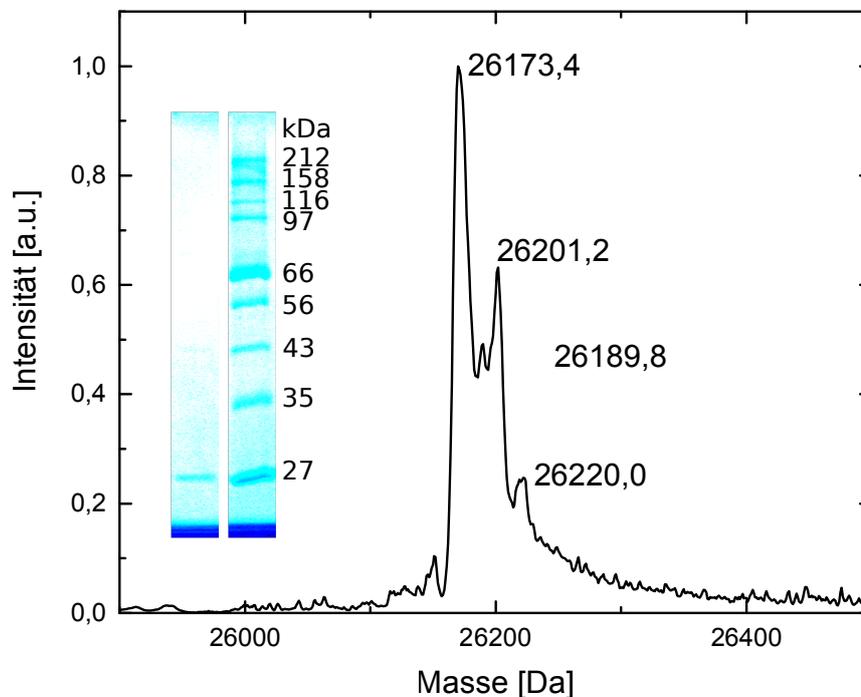


Abbildung 4.15.: Dekonvolutionsiertes Massenspektrum der His-Tag-purifizierten, umgepufferten deGFP -Moleküle der über Nacht inkubierten (29°C) Aha-Reaktion und SDS-PAGE der analysierten Modellproteine (eingelassen). Der Hauptpeak kann $\text{deGFP}^{1\text{Met}+4\text{Aha}}$ mit vier von fünf Aha-Substitutionen zugeordnet werden (errechnete Masse: 26172 Da, gemessene Masse: 26173,4 Da). Die Neben-Peaks werden im Text diskutiert. Das Spektrum ist auf seinen höchsten Wert normiert. Messung von Matthias P. Exner durchgeführt.

4.3.3. Ortsspezifischer Einbau von BocLys in $\text{EGFP}^{\text{N149amber}}$

Das zellfrei synthetisierte $\text{EGFP}^{\text{N149amberHis}_6\text{-Tag}}$ wird in Abschn. 4.3.3 und 4.4 vereinfacht als $\text{EGFP}^{\text{N149amber}}$ bezeichnet.

Erste Ergebnisse von Experimenten, die dazu dienten, die Stopcodon-Suppression bei Variation verschiedener Parameter zu untersuchen, wurden in der Dissertation von Matthias P. Exner bereits als Zusammenfassung gezeigt (Abschn. 5.1.6 in [154]). Abschn. 4.3.3 greift diese Zusammenfassung auf, geht auf die zugrundeliegenden Reaktionsbedingungen ein und zeigt fortführende Messungen.

4.3.3.1. Kontrollexperiment: Test auf Stopcodon-Suppression ohne o-Paar und BocLys

Das folgende Experiment sollte darüber Aufschluss geben, ob es tatsächlich im verwendeten zellfreien System zum Abbruch der Translation kommt, wenn weder das o-Paar (PylRS^{Y384F} und tRNA^{Pyl_{CUA}}) noch BocLys in der zellfreien Reaktion präsent sind. Dies ist eine minimale Anforderung an die zellfreie experimentelle Realisierung, um in anschließenden Experimenten bei einem Anstieg der Fluoreszenz-Intensität bei Vorhandensein dieser drei Moleküle von einer erfolgreichen Stopcodon-Suppression auszugehen.

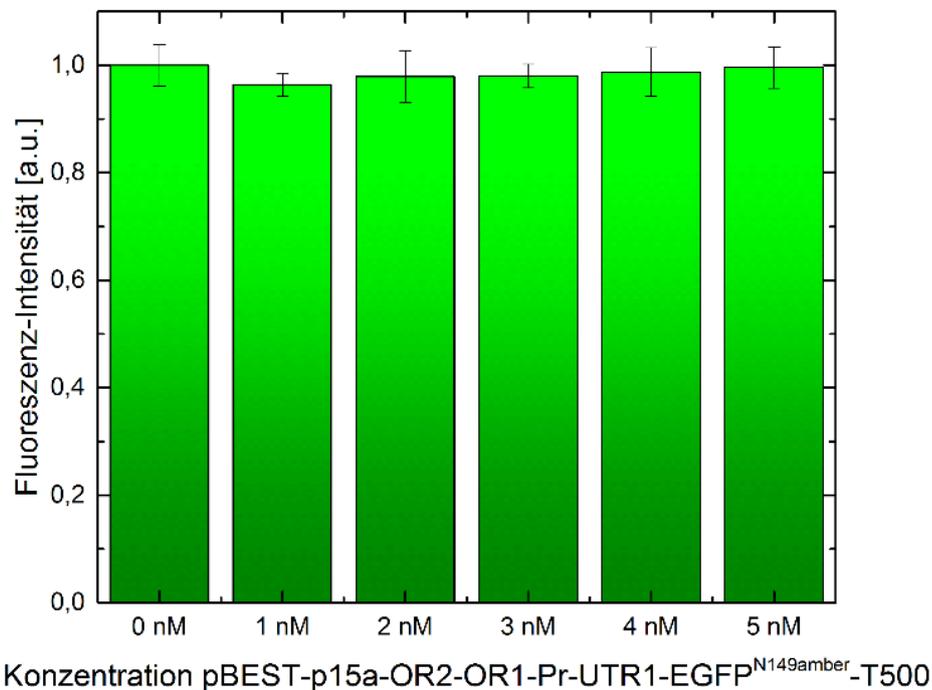


Abbildung 4.16.: Fluoreszenz-Intensität der zellfreien Reaktionen in Abhängigkeit der Konzentration des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500. Trotz zunehmender Plasmid-Konzentration wuchs die Fluoreszenz-Intensität nicht an, weil das Stopcodon nicht supprimiert und EGFP^{N149amber} nicht vollständig synthetisiert wurde. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus drei Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung. Dieses Experiment diente der Ergänzung des Experiments, das in Fig. 25A in [154] gezeigt und in gleicher Weise wie hier sowie in Abschn. 4.2.14.3 beschrieben, durchgeführt wurde.

Es wurden zellfreie Reaktionen präpariert, in denen die Konzentration des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 unter sonst gleichen Reaktionsbedingungen variiert wurde. Die Reaktionen wurden über Nacht (16 h) bei 29 °C inkubiert.

Abb. 4.16 zeigt, dass die Fluoreszenz-Intensität der zellfreien Reaktionen, wie für weitere Experimente erforderlich, trotz zunehmender Plasmid-Konzentration von pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 annähernd konstant blieb. Wird das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-

deGFP-T500 zur Expression des *degfp*-Gens, in dem sich kein Stopcodon befindet, verwendet, steigt die Fluoreszenz-Intensität hingegen mit zunehmender Plasmid-Konzentration stark an, z.B. [51] oder Fig. 25A in [154].

4.3.3.2. Stopcodon-Suppression durch eine aminoacylierte Suppressor-tRNA

Im folgenden Experiment wurde die Stopcodon-Suppression mittels einer mit der nicht-kanonischen Aminosäure N_ε-Acetyllysin (AcLys) aminoacylierten amber-Stopcodon-Suppressor-tRNA (AcLys-tRNA^{amber}) untersucht. Dies ermöglichte, den eigentlichen Prozess der Stopcodon-Suppression zu untersuchen, ohne dass dabei weitere Prozesse wie die Beladung der Suppressor-tRNA durch die o-aaRS eine Rolle spielen. Die Konzentration der AcLys-tRNA^{amber} wurde innerhalb einer Reihe zellfreier Reaktionen unter sonst gleichen Reaktionsbedingungen variiert. Das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 war in allen Reaktionen 5 nM konzentriert, die Reaktionen wurden über Nacht (16 h) bei 29 °C inkubiert.

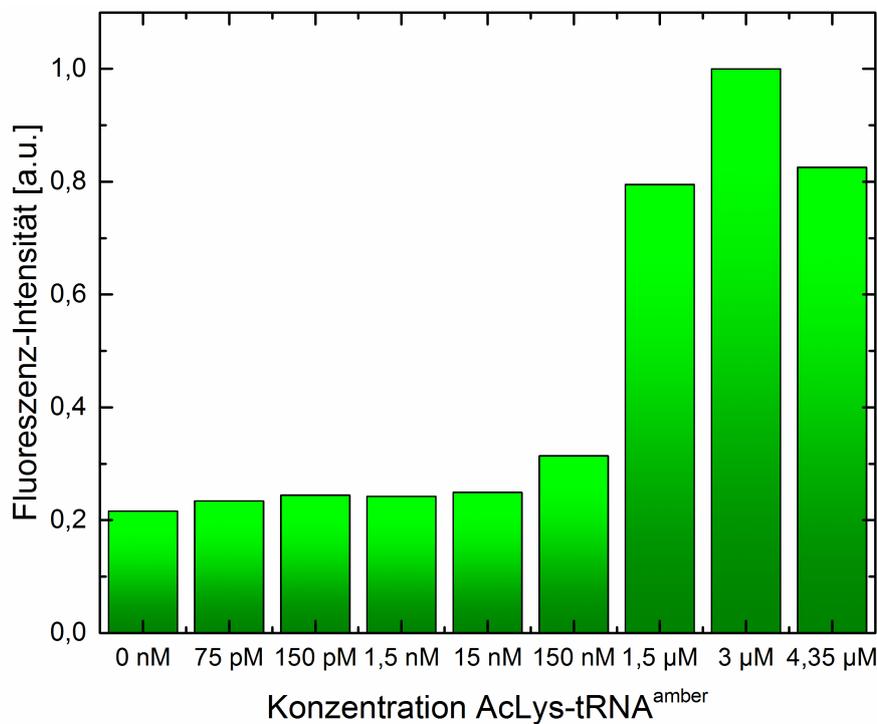


Abbildung 4.17.: Fluoreszenz-Intensität der zellfreien Reaktionen in Abhängigkeit der Konzentration der mit N_ε-Acetyllysin (AcLys) aminoacylierten amber-Stopcodon-Suppressor-tRNA (AcLys-tRNA^{amber}). Das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 war in allen Reaktionen 5 nM konzentriert, die Reaktionen wurden über Nacht (16 h) bei 29 °C inkubiert. Dieses Ergebnis wurde vorab in Fig. 25B in [154] in einer modifizierten Darstellung präsentiert. Abb. 4.17 ist eine modifizierte Darstellung der in Fig. 25B aus [154] gezeigten Daten.

Die Stopcodon-Suppression war erst ab einer AcLys-tRNA^{amber}-Konzentration von 0,15 μM detektierbar, sie ließ sich mit zunehmender AcLys-tRNA^{amber}-Konzentration bis zu einem Maximum bei einer AcLys-tRNA^{amber}-Konzentration von 3 μM weiter steigern (Abb. 4.17). Durch Einbringen dieser aminoacylierten amber-Suppressor-tRNA AcLys-tRNA^{amber} konnte demnach das Stopcodon während der Translation von EGFP^{N149amber} supprimiert werden.

Mit einem GFP-Standard wurde die Konzentration des vollständig translatierten EGFP^{N149amber}-Moleküls, in dem AcLys ortsspezifisch eingebaut wurde, abgeschätzt (Anh. A.2.1). In der zellfreien Reaktion, in der AcLys-tRNA^{amber} eine Konzentration von 3 μ M besaß, entstand EGFP^{N149amber} in einer Konzentration von etwa 0,34 μ M (ca. 9,4 μ g/ml).

Die chemische Aminoacylierung von Suppressor-tRNAs stellt keine Standardprozedur in Laboratorien dar [294], sodass für dieses Experiment die amber-Suppressor-tRNA AcLys-tRNA^{amber} käuflich erworben werden musste. Aufgrund der hohen Kosten stellt diese Möglichkeit der Stopcodon-Suppression keine Alternative dar.

4.3.3.3. Stopcodon-Suppression durch entkoppelte Kosynthese der orthogonalen Komponenten in Präsenz von BocLys

Im folgenden Experiment wurde das o-Paar in der zellfreien Reaktion neben EGFP^{N149amber} in Präsenz der nicht-kanonischen Aminosäure BocLys (2 mM) kosynthetisiert.

Das verwendete zellfreie System ermöglicht neben der Transkription mit der endogenen *E.-coli*-RNA-Polymerase auch die Integration eines parallelen Transkriptionssystems, beispielsweise der transkriptions-stärkeren T7-RNA-Polymerase und dem zugehörigen T7-Promotor [52]. Auf die parallele Transkription wurde in diesem Experiment zurückgegriffen, um die Transkription der Gene, die für die PylRS^{Y384F} und tRNA^{Pyl_{CUA}} kodieren, voneinander zu entkoppeln.

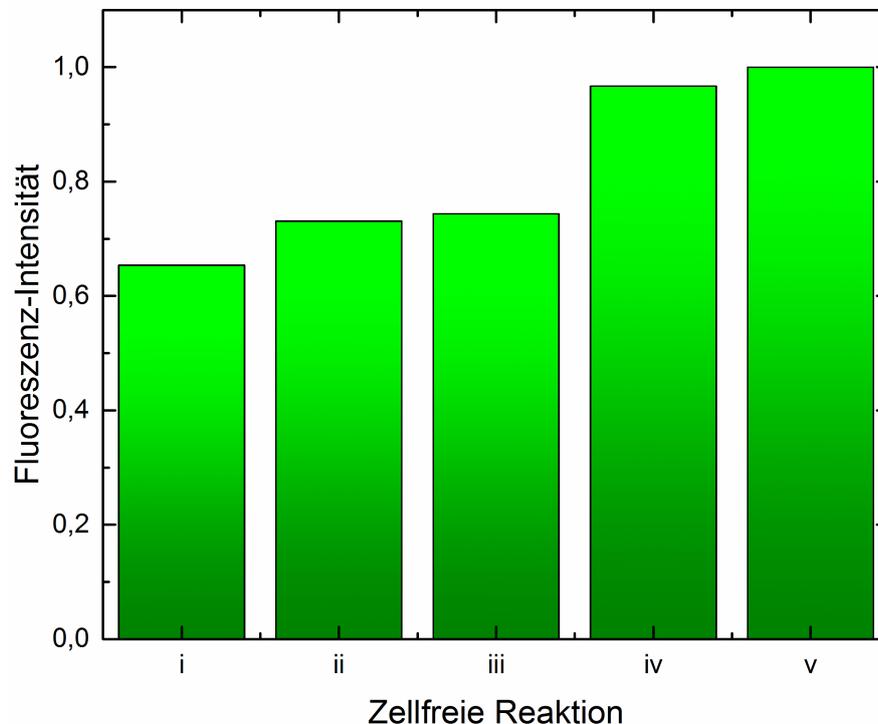


Abbildung 4.18.: Fluoreszenz-Intensität der zellfreien Reaktionen in Abhängigkeit der Konzentration verschiedener DNA-Konstruktionen, die die zellfreie Synthese der PylRS^{Y384F}, tRNA^{Pyl_{CUA}} und von EGFP^{N149amber} gewährleisten. Weitere Details können dem Text entnommen werden. Diese Ergebnisse wurden bereits in Fig. 25C in [154] vorab gezeigt.

Während die Gene, die für die PylRS^{Y384F} bzw. für EGFP^{N149amber} kodieren, jeweils unter Kontrolle des *E.-coli*-Promotors OR2-OR1-Pr standen (pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PylRS^{Y384F}-T500), wurde die Transkription des Gens, das für die tRNA_{CUA}^{Pyl} kodiert, vom T7-Promotor reguliert (lineares DNA-Template T7-tRNA_{CUA}^{Pyl}). Da die T7-RNA-Polymerase nicht im verwendeten *E.-coli*-Extrakt enthalten war, musste sie allen zellfreien Reaktionen als rekombinantes Protein zur Verfügung gestellt werden. In einer Reaktion wurde das Gen, das für die T7-RNA-Polymerase mit der stillen Mutation K378R [279], [280] kodiert, unter Kontrolle des *E.-coli*-Promotors OR2-OR1-Pr gestellt und zusätzlich koexprimiert (OR2-OR1-Pr-UTR1-T7-RNA-Polymerase-T500).

Es wurden insgesamt fünf zellfreie Reaktionen präpariert, in denen die T7-RNA-Polymerase 10,3 nM (1,01 µg/ml) konzentriert war. Ebenso wurde in alle fünf Reaktionen nach einer halben Stunde Präinkubation bei 29 °C das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 pipettiert, sodass es überall eine Konzentration von 2,9 nM hatte. Unter sonst gleichen Reaktionsbedingungen wurden in diesen fünf Reaktionen die Konzentrationen der verschiedenen DNA-Konstrukte variiert. In Reaktion (i) (Negativ-Kontrolle) wurde kein weiteres DNA-Konstrukt gegeben. In Reaktion (ii) wurde zusätzlich die Konstruktion pbest-p15a-OR2-OR1-Pr-UTR1-PylRS^{Y384F}-T500 (2,5 nM) hinzugefügt, die die PylRS-Synthese ermöglicht. Jedoch fehlte das DNA-Konstrukt, mit dem die tRNA_{CUA}^{Pyl} synthetisiert werden konnte. In Reaktion (iii) wurde das Konstrukt T7-tRNA_{CUA}^{Pyl} (307 nM) pipettiert, aber es fehlte im Gegenzug das PylRS^{Y384F}-Plasmid. In Reaktion (iv) wurden sowohl das lineare DNA-Konstrukt T7-tRNA_{CUA}^{Pyl} (307 nM) als auch das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-PylRS^{Y384F}-T500 (2,5 nM) hinzugegeben. Reaktion (v) besaß die gleiche Komposition wie Reaktion (iv), nur dass zusätzlich das lineare DNA-Konstrukt, das die zellfreie Synthese der T7-RNA-Polymerase ermöglichte, hinzugegeben wurde (0,6 nM). Nach 30 min Präinkubation und Hinzugabe des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 wurden die Reaktionen für weitere 15,5 h über Nacht bei 29 °C inkubiert. Durch die Präinkubation sollte der zellfreien Synthese des o-Paares und der T7-RNA-Polymerase eine Vorlaufzeit eingeräumt werden, damit diese Moleküle zu Beginn der zellfreien Synthese von EGFP^{N149amber} in einer gewissen Konzentration vorhanden waren, sodass es direkt zur Stopcodon-Suppression kommen konnte.

Die Stopcodon-Suppression verlief in Reaktion (v) etwa 1,5-fach effizienter als die der Referenz-Reaktion (Fluoreszenz-Intensitäts-Unterschied zwischen der Referenz-Reaktion (i) und Reaktion (v) etwa 35 %, Abb. 4.18). Die zusätzliche zellfreie Synthese der T7-RNA-Polymerase hatte, wenn überhaupt, wenig Einfluss auf die Stopcodon-Suppression (Abb. 4.18). Der gemessene Fluoreszenz-Intensitäts-Anstieg in Reaktion (iv) und (v) ist ein Indiz dafür, dass BocLys ortsspezifisch mittels des kossynthetisierten o-Paares in EGFP^{N149amber} eingebaut wurde. In vergleichbaren Experimenten, in denen keine Präinkubation stattfand, diese verlängert wurde oder die verschiedenen DNA-Konstrukte anders konzentriert waren, fiel der Fluoreszenz-Intensitäts-Anstieg jedoch geringer aus oder er fand nicht statt. Aus diesem Grund wurde nach einer besser geeigneten Strategie gesucht, um die Stopcodon-Suppression im verwendeten zellfreien Expressionssystem zu realisieren.

4.3.3.4. Stopcodon-Suppression durch gekoppelte Kosynthese der orthogonalen Komponenten in Präsenz von BocLys

Im folgenden Experiment wurde das o-Paar in gekoppelter Weise unter Kontrolle des Promotors OR2-OR1-Pr gestellt (pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA_{CUA}^{Pyl}-PylRS^{Y384F}-T500). Die tRNA_{CUA}^{Pyl}-

und PylRS^{Y384F}-Synthese verlief parallel zur zellfreien Synthese von EGFP^{N149amber} (pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500) in Präsenz von BocLys (2 mM). Aus den in Abschn. 4.3.3.3 beschriebenen Gründen wurde das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 erst nach einer halben Stunde Präinkubation bei 29 °C den zellfreien Reaktionen hinzugegeben, sodass es eine Konzentration von 5 nM hatte. In der Negativ-Kontrolle war die für das o-Paar kodierende DNA nicht vorhanden. In der anderen Reaktion besaß pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA^{Py1}_{CUA}-PylRS^{Y384F}-T500 eine Konzentration von 4,9 nM. Nach 30 min Präinkubation und Hinzugabe des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 wurden die Reaktionen für weitere 14,5 h über Nacht bei 29 °C inkubiert.

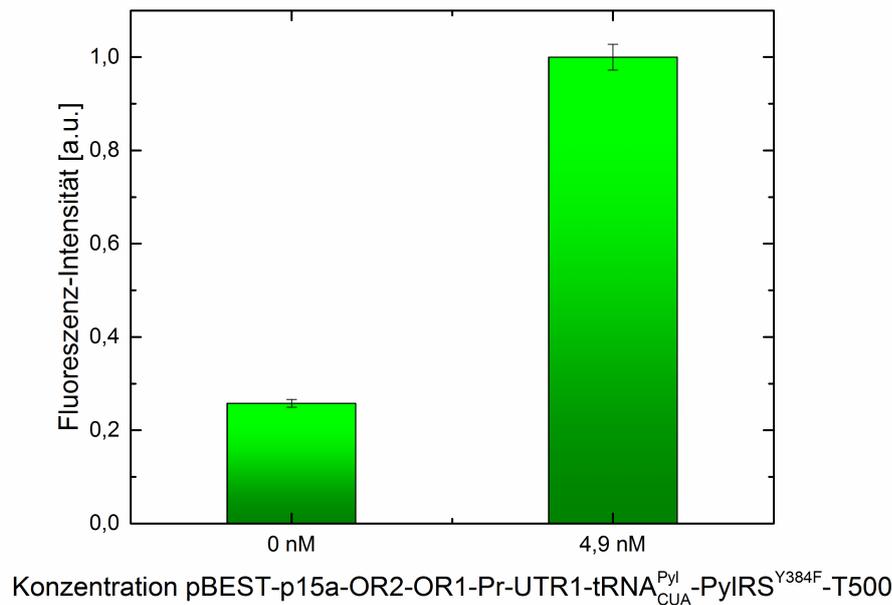


Abbildung 4.19.: Fluoreszenz-Intensität der zellfreien Reaktionen in Abhängigkeit der Konzentration des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA_{CUA}^{Py1}-PylRS^{Y384F}-T500. Die Fluoreszenz-Intensität stieg in Präsenz von pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA_{CUA}^{Py1}-PylRS^{Y384F}-T500 (4,9 nM) etwa vierfach im Vergleich zur Negativ-Kontrolle an. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus drei Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung.

Die Stopcodon-Suppression fand in der zellfreien Reaktion, in der das o-Paar synthetisiert wurde, statt (etwa Vervielfachung der Fluoreszenz-Intensität im Vergleich zur Negativ-Kontrolle, Abb. 4.19). Eine Quantifizierung mittels der bereits in Abschn. 4.3.3.2 verwendeten GFP-Eichgerade ergab, dass EGFP^{N149amber} in einer Konzentration von 0,19 µM (5,25 µg/ml) zellfrei synthetisiert wurde.

Die Kosynthese vom o-Paar und dem Reporter-Molekül EGFP^{N149amber} in der gleichen zellfreien Reaktion zum Zweck der Stopcodon-Suppression ist demnach möglich. Die gekoppelte Transkription der Gene, die für das o-Paar kodieren und gemeinsam unter Kontrolle des *E.-coli*-Promotors stehen, erwies sich als geeignet, um eine ausreichende tRNA_{CUA}^{Py1}- und PylRS^{Y384F}-Menge für die Stopcodon-Suppression bereitzustellen. *E.-coli*-Enzyme, die die tRNA_{CUA}^{Py1} prozessieren, scheinen im Zellextrakt vorhanden zu sein. Allerdings fiel auch bei diesem Experiment die Proteinausbeute niedrig aus. Das Volumen der zellfreien Reaktion müsste deutlich nach oben skaliert werden, um eine adäquate EGFP^{N149amber}-Menge zu synthetisieren, die genügt, um nach His-Tag-Purifikation und

Umpufferung genug Substanz für eine massenspektroskopische Analyse zur Verfügung zu stellen. Eine massenspektroskopische Analyse des zellfrei synthetisierten EGFP^{N149amber}-Moleküls müsste hingegen noch durchgeführt werden, um den ortsspezifischen Einbau von BocLys in EGFP^{N149amber} tatsächlich nachzuweisen.

Das verwendete zellfreie System besitzt das Potential, mittels der endogenen Transkriptions-Translations-Maschinerie bei gleichzeitiger Kosynthese des o-Paares und des zu modifizierenden Zielproteins, nicht-kanonische Aminosäuren ortsspezifisch in das Zielprotein zu inkorporieren. Damit das zellfreie System in der Lage ist, bei gleichzeitiger Synthese des o-Paares und des zu modifizierenden Zielproteins, eine adäquate Ausbeute des Zielproteins synthetisieren zu können, müssen weitere Verbesserungen an den experimentellen Bedingungen erfolgen.

Mittels einer alternativen Strategie und einer Modifizierung des hier verwendeten zellfreien Systems gelang der Gruppe von ALFONTA der ortsspezifische Einbau von BocLys und weiterer Pyl-Derivate in deGFP [264]. Im Gegensatz zu dem in dieser Arbeit gewählten Ansatz, das o-Paar gleichzeitig mit dem zu modifizierenden Zielprotein zellfrei zu synthetisieren, wurde von [264] bereits der Extrakt mit dem o-Paar ausgestattet: Die o-Paar-Synthese wurde in den *E.-coli*-Zellen, aus denen der Extrakt im weiteren Verlauf gewonnen wurde, induziert. Außerdem wurde in dieser Studie [264] zur Extrakt-Präparation der *E.-coli*-Stamm C321.Δ*prfA* [257] verwendet, bei dem der Terminationsfaktor 1 ausgeknockt ist. Die Strategie von [264] erzeugte eine modifizierte Zielprotein-Ausbeute von 0,2 mg/ml, das Stopcodon wurde mit nahezu 100%-iger Effizienz supprimiert. Diese Ausbeute war in etwa 40-fach höher als die in der vorliegenden Studie erzeugte Ausbeute. [264] zeigte ebenfalls, dass die Suppressions-Effizienz sehr stark abnimmt, sobald *E.-coli*-Stämme zur Extrakt-Gewinnung verwendet werden, in denen der Terminationsfaktor 1 vorhanden ist. Dies bietet einen möglichen Erklärungsansatz dafür, warum die Zielprotein-Ausbeute in der vorliegenden Studie niedriger als in [264] ausfiel: Zur Extrakt-Herstellung wurden *E.-coli*-Zellen, die mit dem Terminationsfaktor 1 ausgestattet sind, benutzt (vgl. Abschn. 4.2.3).

4.4. Zusammenfassung, Konklusion und Ausblick

In der vorliegenden Studie wurde der Einbau von nicht-kanonischen Aminosäuren in Modellproteine mittels des zellfreien Expression-Systems von NOIREAUX [51], [52] untersucht. Es ermöglicht die Genexpression mittels der endogenen *E.-coli*-Transkriptions-Translationsmaschinerie [51], [52]. Diese native endogene molekulare Zusammensetzung erlaubt es, komplexe genetische Schaltkreise zu implementieren [262]. Daraus leitete sich die Motivation ab, jenes zellfreie System bei einer komplexen biochemischen Prozedur, wie dem Einbau nicht-kanonischer Aminosäuren in Proteine, zu nutzen.

Der erste Teil von Kap. 4 beschäftigte sich mit dem seitenkettenspezifischen Einbau nicht-kanonischer Aminosäuren-Analoga in das Modellprotein deGFP, das einen His-Tag zur späteren Purifikation besaß. Das zugrundeliegende Protokoll zur Präparation der zellfreien Reaktionen [51], [52] wurde erweitert, indem drei unterschiedlich zusammengesetzte Aminosäuren-Mixes eingesetzt wurden. Wurde der zellfreien Reaktion ein Aminosäuren-Mix zur Verfügung gestellt, in dem eine einzubauende nicht-kanonische Aminosäure ihr kanonisches Analogon substituierte, konnten verschiedene nicht-kanonische Aminosäuren vollständig oder teilweise in das Zielprotein eingebaut werden. Mittels eines Standard-Aminosäuren-Mixes, der aus den 20 kanonischen Aminosäuren bestand, wurde das unmodifizierte Zielprotein als Referenz synthetisiert. Zellfreie Reaktionen, die mit Aminosäuren-Mixes, die weder die zu ersetzende kanonische Aminosäure noch ihr nicht-kanonisches Analogon enthielten, ausgestattet wurden, dienten als Negativ-Kontrollen: Sofern keine Reste der zur nicht-kanonischen Aminosäure analogen kanonischen Aminosäure im zellfreien System vorhanden waren, wurde kein Zielprotein synthetisiert.

Die gelelektrophoretische Auftrennung der zellfreien Reaktion per SDS-PAGE erwies sich als nützliche vorläufige Analyse-Methode. Ließen sich bei der Negativkontrolle ebenfalls Proteinbanden finden, die dem synthetisierten Zielprotein zugeordnet werden konnten, war davon auszugehen, dass sich Reste der analogen kanonischen Aminosäure im zellfreien System befanden, die zum unvollständigen Einbau der nicht-kanonischen Aminosäure in das Zielprotein geführt hatten. Der Einsatz zweier Spin-Säulen-Systeme erwies sich als geeignet, das Modellprotein aus den zellfreien Reaktionen zu extrahieren, purifizieren und umzupuffern. Der finale Nachweis des vollständigen oder unvollständigen Einbaus einer nicht-kanonischen Aminosäure in deGFP wurde per HPLC-ESI-Massenspektrometrie erbracht, die von Matthias P. Exner (Kooperationsarbeitsgruppe von Nediljko Budisa, Institut für Chemie, Technische Universität Berlin) durchgeführt wurde.

Der seitenkettenspezifische Einbau der toxischen nicht-kanonischen Aminosäure Can in deGFP war vollständig und verlief effizient, ohne dass der zellfreie Extrakt in spezieller Weise von Arg depletiert werden musste. Der vollständige Can-Einbau in deGFP wurde von der zellfreien Umgebung des Systems stabilisiert. Hingegen wies in einem Vergleichsexperiment *in vivo* synthetisiertes und mit Can modifizierte EGFP häufig eine reduzierte Masse auf, die auf eine Degradierung von Can in Homoserin und Hydroxyguanidin, ein gut beschriebener Reaktionsweg [290], [291], hindeutet. Der hier gewählte zellfreie Ansatz umging weitere Limitationen des *in-vivo*-Einbaus von Can in Zielproteine, wie auch dessen Toxizität auf den Metabolismus des lebenden Expressionswirts [36], [41], [48], [230], [231] oder die starke mRNA-Abhängigkeit eines Einzel-Protein-Produktions-Systems [50].

Im Gegensatz zur Can-Inkorporation verlief der seitenkettenspezifische Einbau von Hyl und Aha unvollständig. Dies ist auf die Präsenz von Resten der entsprechenden analogen kanonischen Aminosäuren zurückzuführen. Dass Reste verschiedener kanonischer Aminosäuren in Extrakt-basierten zellfreien Systemen den seitenkettenspezifischen Einbau ihrer nicht-kanonischen Analoga erschweren, ist bekannt [232]–[236]. Ein Erklärungsansatz, warum Arg viel geringer oder gar nicht im verwendeten

zellfreien System vorhanden war, könnte auf die Aktivität der Arginin-Decarboxylase zurückzuführen sein, die häufig in Extrakt-basierten zellfreien Systemen die Proteinausbeute reduziert, da das Enzym Arg degradiert [114], [129], [131]. Möglicherweise wurden Arg-Reste während des *Run-Offs* bei der Extrakt-Herstellung durch die im Zellextrakt vorhandene Arginin-Decarboxylase entfernt. Das Arg-Analogon Can ist zwar ebenfalls ein Substrat für dieses Enzym, jedoch wird es nur mit 40%-iger Effizienz vom Enzym attackiert [295]. Außerdem bildet Spermidin einen Bestandteil des zellfreien Systems [51], [52], das als Arginin-Decarboxylase-Inhibitor bekannt ist [61], [295]. Somit kann erklärt werden, warum Arg zwar während des *Run-Offs* degradiert, jedoch Can in der zellfreien Reaktion während der Expression stabilisiert wurde.

Um auch solche nicht-kanonischen Aminosäuren seitenkettenspezifisch in Proteine einbauen zu können, deren kanonische Analoga in Resten im zellfreien System vorhanden sind, müssen weitere Strategien zur Entfernung dieser Reste entwickelt werden.

Die Proteinausbeute des Zielproteins hängt stark von der Natur der einzubauenden nicht-kanonischen Aminosäure ab. In dieser Studie besaßen die nicht-kanonischen Aminosäuren die gleiche Konzentration wie die kanonischen Aminosäuren im zellfreien System. Dies führte stets zu adäquaten Ausbeuten, die ausreichten, um die modifizierten Modellproteine weiter zu analysieren. Höhere Modellprotein-Ausbeuten können durch Variation der Konzentration der nicht-kanonischen Aminosäuren möglich sein [235]. Diese Abhängigkeit sollte in weiteren Experimenten untersucht werden.

Resümierend festzuhalten bleibt, dass sich das verwendete zellfreie System für den seitenkettenspezifischen Einbau nicht-kanonischer Aminosäuren in Modellproteine eignet, solange die einzubauende nicht-kanonische Aminosäure von der endogenen *E.-coli*-Translations-Maschinerie erkannt wird.

Der zweite Teil von Kap. 4 beschäftigte sich mit dem ortsspezifischen Einbau von nicht-kanonischen Aminosäuren in Proteine mittels Stopcodon-Suppression. Zur Stopcodon-Suppression wurde das PylRS^{Y384F}/tRNA_{CUA}^{Pyl}-Paar aus *Methanosarcina mazei* verwendet, das orthogonal zum endogenen Satz der *E.-coli*-aaRSs und -tRNAs ist [253]. Die einzubauende nicht-kanonische Aminosäure war BocLys, die sich bereits als geeignetes Pyl-Derivat und Substrat für die PylRS erwiesen hat [254]. Als Modellprotein und Reportermolekül diente EGFP^{N149amber} (mit His-Tag), dessen Gen an der Stelle 149 anstatt eines Asn-Codons ein Amber-Stopcodon enthielt und deswegen nur bei erfolgreicher Stopcodon-Suppression fluoreszierte [152]–[154].

Mittels der endogenen *E.-coli*-Transkriptions-Translations-Maschinerie des verwendeten zellfreien Systems ließ sich, bei gleichzeitiger Kosynthese des PylRS^{Y384F}/tRNA_{CUA}^{Pyl}-Paares und des zu modifizierenden Zielproteins, die nicht-kanonische Aminosäure BocLys ortsspezifisch in das Zielprotein inkorporieren. Obwohl sich mit dem hier gewählten Ansatz nicht-kanonische Aminosäuren flexibel ortsspezifisch in Zielproteine inkorporieren ließen, da nur die für das o-Paar und das Modellprotein kodierende DNA zusätzlich in die zellfreie Reaktion gegeben werden musste, fiel die Proteinausbeute zu niedrig aus, um den vollständigen Einbau von BocLys in das Zielprotein per Massenspektrometrie ohne drastische Vergrößerung der Reaktionsvolumina nachweisen zu können. Jedoch kann die erfolgte Zunahme der Fluoreszenz-Intensität bei zellfreier Synthese des PylRS^{Y384F}/tRNA_{CUA}^{Pyl}-Paares und Präsenz von BocLys als starkes Indiz gesehen werden, dass die Stopcodon-Suppression erfolgreich verlief. Weitere experimentelle Verbesserungen müssen erfolgen, damit adäquate Proteinausbeuten des Zielproteins mit dem gewählten Ansatz erzeugt werden können.

Mit Hilfe einer alternativen Strategie und einer Modifizierung des hier verwendeten zellfreien Systems gelang der Gruppe von ALFONTA der ortsspezifische Einbau von BocLys und weiterer Pyl-Derivate in deGFP [264]. Anders als beim hier gewählten Ansatz, das PylRS^{Y384F}/tRNA_{CUA}^{Pyl}-Paar gleichzeitig

mit dem zu modifizierenden Zielprotein zu synthetisieren, stattete [264] bereits den Extrakt mit dem o-Paar durch induzierte PylRS^{Y384F}/tRNA_{CUA}^{Pyl}-Paar-Synthese in den *E.-coli*-Zellen aus, die im weiteren Verlauf den Extrakt lieferten. Darüber hinaus wurde dieser aus dem *E.-coli*-Stamm C321.Δ*prfA* [257] gewonnen, dem der Terminationsfaktor 1 fehlt. Mit dem von [264] verfolgten Ansatz ließen sich etwa 40-fach höhere Ausbeuten des modifizierten Modellproteins generieren. Diese experimentelle Strategie von [264] motiviert für zukünftige Studien die Verwendung eines C321.Δ*prfA*-Zellextraktes, der das PylRS^{Y384F}/tRNA_{CUA}^{Pyl}-Paar nicht bereits enthält, aber vom Terminationsfaktor 1 befreit ist. Ein daraus resultierendes zellfreies System könnte in der Lage sein, bei gleichzeitiger Synthese des PylRS^{Y384F}/tRNA_{CUA}^{Pyl}-Paares und des zu modifizierenden Zielproteins, eine adäquate Ausbeute des Zielproteins zu generieren. In solch einem System könnten flexibel verschiedene o-Paare synthetisiert und verschiedene nicht-kanonische Aminosäuren getestet und ausgetauscht werden.

5. Zellfreie Analyse des epigenetischen *Escherichia-coli*-Pili-Phasenvariationsmechanismus

5.1. Einführung, Stand der Wissenschaft, Strategie und Zielsetzung

5.1.1. Urinaltraktinfektionen durch *E. coli* und ihre Folgen

Einer Schätzung aus dem Jahr 1994 [296] zufolge gehören Urinaltraktinfektionen (UTI) zu den häufigsten menschlichen bakteriellen Infektionskrankheiten, an denen jährlich mehr als 150 Millionen Menschen weltweit erkranken. Beispielsweise wurden 2015 die annualen Kosten, die in den USA durch derartige Infektionskrankheiten entstanden, mit rund 3,5 Milliarden US \$ beziffert [58]. Ausgehend vom unteren Urinaltrakt können die infizierenden Erreger über die Harnblase zur Niere aufsteigen und Entzündungen verursachen [56], [58]. Klinisch werden diese Entzündungen nach Ort der Infektion in Zystitis (untere UTI) oder in Pyelonephritis (obere UTI) unterschieden [297], [298]. Bei einer Zystitis wird nur die Harnblase von Erregern kolonisiert [297]. Ist (zusätzlich) die Niere, vor allem das Nierenbecken, infiziert und entzündet, so wird von einer Pyelonephritis gesprochen [298]. UTI-assoziierte Komplikationen, beispielsweise eine Urosepsis, können zum Tod des Wirts führen [56], [298]. Mindestens 65 % aller UTI gehen auf uropathogene *E. coli* (UPEC) zurück [58].

5.1.2. Pyelonephritis-assoziierte Pili von uropathogenen *E. coli*

Gramnegative Bakterien, denen auch UPEC angehören, können Pili, auch Fimbrien genannt, ausbilden [56], [57]. Pili sind haarartige, lineare, unverzweigte und polymerartige Strukturen, die in der äußeren Zellmembran verankert sind [56], [299]. Durch Pili können die Bakterien an der Epithelzellschicht der ableitenden Harnwege, dem Urothel des Wirts, adhären [56], [57].

Die für diesen Teil der Arbeit relevanten Pili vom Typ P (P-Pili) gehören der Klasse der Chaperon-Usher-Pili an [56]. Da Pili dieses Typs vor allem mit Pyelonephritis assoziiert sind, werden sie auch als Pyelonephritis-assoziierte Pili (Pap) bezeichnet. Der P-Pilus untergliedert sich in eine flexible, adhäsive Spitze ($\varnothing \approx 2$ nm und 15 nm lang) und einen rechtsgängigen helikalen Stab ($\varnothing \approx 7$ nm und 2 μ m lang) [57] (Abb. 5.1). Pap setzen sich aus verschiedenen Proteinen zusammen, die in unterschiedlich hoher Anzahl im Pilus eingebaut und nicht-kovalent miteinander verbunden sind [300], [301] (Abb. 5.1).

Das äußere Ende der Spitze bildet eine Einheit des Adhäsins PapG [56], [57], [299], [300]. PapG bindet an Gal- α (1-4)-Gal-Verbindungen der Kohlenhydratanteile von Glykosphingolipiden der Globo-Serie, die sich an der Zellmembranoberfläche des Wirts befinden [300], [302], [303]. In Richtung Verankerung folgen nacheinander die Adapter-Moleküle PapF (eine Einheit) und PapE (5–10 Einheiten), die zusammen mit PapG die elastische Spitze bilden [57], [299].

Es folgen mehr als 1000 Einheiten von PapA, die den unelastischen Stab aufbauen [57], [299], [300]. Der Pilus-Stab wird durch PapH in der äußeren Membran terminiert und verankert [299], [300]. Die Pilus-Spitze und der Pilus-Stab sind durch eine Einheit des Adapters PapK verbunden [304]. Die Pilus-Polymerisierung wird durch das Chaperon-*Usher*-System reguliert [304], das aus dem sich

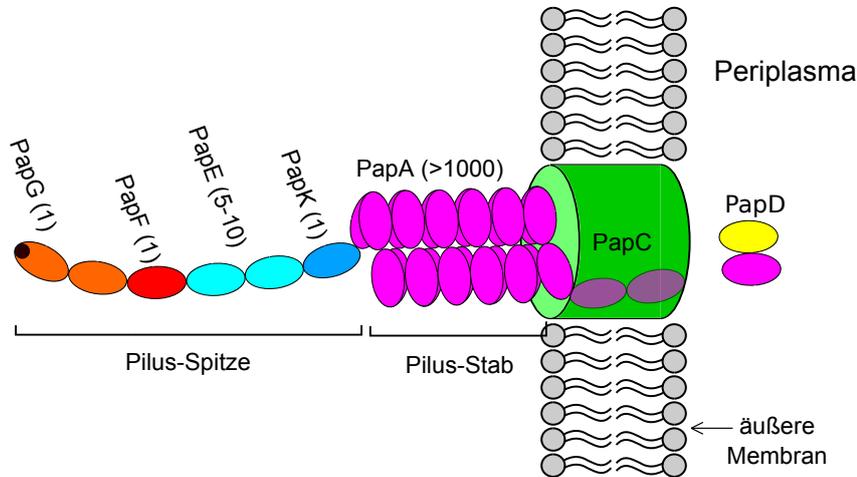


Abbildung 5.1.: Schematischer Pap-Aufbau. Solch ein P-Pilus besteht aus sechs verschiedenartigen Proteinen, die in unterschiedlich hoher Anzahl mit Hilfe des CU-Systems zum Pilus assembliert werden [300], [304]. Abb. 5.1 ist eine erweiterte und somit modifizierte Darstellung der Idee von Abb. 1 aus [301].

im Periplasma befindlichen Chaperon PapD und dem in der äußeren Zellmembran sitzenden *Usher*-Protein PapC [305] besteht. Während der Pap-Biogenese werden die strukturbildenden Untereinheiten nach ihrer Synthese vom Cytosol ins Periplasma gebracht [304]. Das Chaperon ist für die Stabilisierung und Faltung der Untereinheiten entscheidend, hindert diese an vorzeitiger Polymerisierung innerhalb des Periplasmas [304]. Schlussendlich übermittelt es die Untereinheiten an das *Usher*-Protein, das die sukzessive Polymerisierung des Pilus an der Außenseite der äußeren Zellmembran reguliert [304].

5.1.3. Epigenetische Regulation des UPEC-Pili-Phasenvariationsmechanismus

Durch einen Phasenvariationsmechanismus, der auf Transkriptions-Ebene reguliert wird, können UPEC zwischen der An-Phase, bei der die *pap*-Gene exprimiert werden, und der Aus-Phase, bei der die Transkription der *pap*-Gene reprimiert wird, wechseln [59]. Der *pap*-regulatorische Lokus der UPEC-DNA, auf den im Folgenden eingegangen wird, ist in Abb. 5.2 schematisch dargestellt. Er beginnt stromaufwärts mit dem *papI*-Gen und endet stromabwärts mit dem *papB*-Gen [306]. Die beiden in die Regulation involvierten *papI*- und *papB*-Gene werden divergent transkribiert [307] und umschließen einen ca. 400 bp langen DNA-Bereich, auf dem sich sechs DNA-Bindestellen für das *leucine-responsive regulatory protein* (Lrp) befinden [306]. Das 18,8 kDa schwere Lrp ist ein globaler Transkriptions-Regulator in *E. coli* [308], [309] und kann sich zu einem Oktamer aus vier Dimeren zusammensetzen [310], [311]. Die sechs Lrp-Bindestellen gruppieren sich zu zwei Dreier-Gruppen [312]: Während die Bindestellen 4–6 distal vom *papBA*-Promotor lokalisiert sind, befinden sich die Bindestellen 1–3 proximal zum *papBA*-Promotor. Die Lrp-Bindestelle 3 liegt zwischen dem -35- und -10-Element des *E.-coli*-RNA-Polymerase-Promotors [312]. Die Bindestellen 2 und 5 tragen jeweils die Erkennungssequenz GATC für das Enzym DNA-Adenin-Methylase (Dam), das Adenin innerhalb dieses Sequenzmotivs methyliert [83], [306], [312], [313] (vgl. Abschn. 2.3.2). Das *papI*-Gen besitzt seinen eigenen Promotor [314], [315]. Mittels *in-vitro*-Bindungsexperimenten konnten drei Bindestellen

im *pap*-regulatorischen Bereich auffindig gemacht werden, auf denen das erste Expressions-Produkt PapB des *pap*-Gen-Clusters spezifisch bindet [316]: Bindestelle 1 befindet sich zwischen dem *papI*-Promotor und Lrp-Bindestelle 4, die zweite PapB-Bindestelle überlappt mit der Erkennungssequenz der *E. coli*-RNA-Polymerase des *papBA*-Promotors und die dritte PapB-Bindestelle ist innerhalb des *papB*-Gens selbst lokalisiert (PapB-Bindestellen 2 und 3 sind nicht in Abb. 5.2 eingezeichnet).

Dass die Methylierung der Lrp-Bindestellen bei der transkriptionellen Regulation [59] der Phasenvariation eine entscheidende Rolle spielt, konnte früh gezeigt werden [317]: Bei der DNA aus Phase-Aus-Zellen ist die distale GATC-Stelle ($\text{GATC}^{\text{dist}}$) methyliert, bei Phasen-An-Zellen ist hingegen die proximale GATC-Stelle ($\text{GATC}^{\text{prox}}$) methyliert. Ein essentieller Regulator ist Lrp, ohne den es in *E. coli*-Stämmen, in denen das *lrp*-Gen ausgeknockt ist, zu fast keiner Transkription der *pap*-Gene kam [317]. In dieser frühen Studie [317] wurde aufgrund der *in-vivo*-Methylierungsmuster in Kombination mit *in-vitro*-DNase-Fußabdruck-Untersuchungen geschlussfolgert, dass Lrp und PapI in Phasen-An-Zellen auf der Lrp-Bindestelle 5 binden und $\text{GATC}^{\text{dist}}$ vor Dam-Methylierung schützen, während in Phase-Aus-Zellen Lrp auf der Seite 2 bindet und somit die Dam-Methylierung von $\text{GATC}^{\text{prox}}$ verhindert. Lrp bindet jeweils als Dimer kooperativ auf mehreren seiner Bindestellen [312], [318]. Des Weiteren, basierend auf Kristallstrukturanalysen von Lrp und Lrp-ähnlichen Regulatoren, wurde vermutet, dass Lrp als aus vier Dimeren bestehendes Oktamer kooperativ insgesamt drei seiner Bindestellen besetzen kann, sodass ein Dimer ungebunden bleibt [310], [311], [319]. Diese kooperative Lrp-Oktamer-Bindung auf dem *pap*-regulatorischen Bereich wird teilweise aufgegriffen, z.B. in [320], [321] oder nicht, beispielsweise in [56], [322].

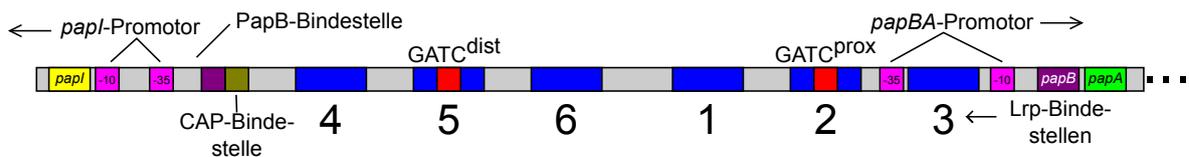


Abbildung 5.2.: Schematischer Aufbau des *pap*-regulatorischen Bereichs inklusive des *papA*-Gens. Diese Abbildung lehnt sich an die Ideen der Abb. 1 aus [306], Abb. 1a aus [320] und Abb. 1 aus [323] an.

Die kooperative Lrp-Bindung aufgreifend, schlug [306] vor, dass in Phase-Aus-Zellen Lrp kooperativ auf den Bindestellen 1–3 bindet, die Methylierung von $\text{GATC}^{\text{prox}}$ und letztlich das Binden der RNA-Polymerase hindert, und dass Lrp in Phase-An-Zellen mit Hilfe von PapI und weiterer Regulatoren (s. unten) die Bindestellen 4–6 kooperativ besetzt, sodass durch Methylierung von $\text{GATC}^{\text{prox}}$ das dortige kooperative Binden von Lrp geblockt wird und die Transkription der *pap*-Gene ablaufen kann. Dass die kooperative Bindung von Lrp auf den Seiten 1–3 die Transkription der *pap*-Gene blockiert, wurde zuvor in *in-vitro*-Transkriptions-Experimenten gezeigt, und es wurde ebenfalls geschlussfolgert, dass dies darauf zurückzuführen sei, dass die RNA-Polymerase aus sterischen Gründen nicht auf dem *papBA*-Promotor binden kann [324]. Des Weiteren konnte gezeigt werden, dass auf den Seiten 1–3 gebundenes Lrp die Affinität von Lrp, auf den gegenüberliegenden Seiten 4–6 zu binden, etwa zehnfach verringert [306]. Dieser Mechanismus wird als *mutual exclusion effect* (Ausschlussphänomen) bezeichnet, ist wenig verstanden, tritt nur bei überspiralierter *pap*-regulatorischer DNA auf und stabilisiert den Aus-Zustand [306]. Außerdem hindert die methylierte $\text{GATC}^{\text{dist}}$ -Stelle das Binden von Lrp und stabilisiert den Aus-Zustand zusätzlich [313]. Für die Phasenvariation werden die Koregulatoren PapI und PapB benötigt [306], [313]. Das Binden von PapB auf der Bindestelle 1, zu der es die größte Bindungsaffinität aufweist, aktiviert die *papI*-Transkription [316]. Durch die Transkription der *pap*-Gene entsteht demnach eine positive Rückkopplung, bei der PapB die *papI*-Transkription verstärkt

[313], [316]. Mittels *Gel-Shift*-Analyse konnte gezeigt werden, dass PapI die Bindungsaffinität von Lrp gegenüber den Lrp-Bindestellen 2 und 5 aufgrund der in diesen Stellen enthaltenen ACGATC-Sequenz erhöht [313]. Für die hemimethylierte GATC^{dist} weist der Lrp-PapI-Komplex eine höhere Bindungsaffinität als für die vollständig methylierten GATC^{dist} auf [313]. Des Weiteren konnte mittels *Gel-Shift*-Analysen beobachtet werden, dass PapI für sich alleine nicht spezifisch an DNA bindet und eine geringe Bindungsaffinität gegenüber freiem Lrp aufweist [314]. *In vitro*-Bindungs-Experimente offenbarten, dass geringe PapI-Konzentrationen die Affinität von Lrp, kooperativ auf den Seiten 4–6 zu binden, erhöhen [313].

Eine Möglichkeit, um vom Phase-Aus- zum Phase-An-Zustand zu gelangen, bietet die DNA-Replikation, da durch diesen Vorgang das an die Lrp-Bindestellen 1–3 gebundene Lrp dissoziiert und die GATC^{dist}-Stelle hemimethyliert vorliegt [313], [320], [325]. Durch die Dissoziation von Lrp von den Bindestellen 1–3 wird das Ausschließungsphänomen aufgehoben, was die Wahrscheinlichkeit, dass Lrp kooperativ auf den Stellen 4–6 bindet, erhöht [306]. Direkt nach der DNA-Replikation konkurrieren Lrp und Dam in einem stochastischen Prozess um die Bindung an die Lrp-Bindestellen 2 bzw. 5, eine Konkurrenzsituation, die in der Methylierung dieser GATC-Stellen im Fall der Dam-Bindung oder des Blockens der Methylierung dieser Stellen aufgrund des gebundenen Lrp endet [313]. Für die Phasenvariation vom Aus- zum An-Zustand ist es notwendig, dass Lrp kooperativ auf den Seiten 4–6 bindet, damit GATC^{dist} vor erneuter vollständiger Methylierung durch Dam geschützt und dass GATC^{prox} von Dam vollständig methyliert wird, um das kooperative Binden von Lrp auf den Seiten 1–3 zu verhindern [313], [320]. Ein Erklärungsansatz, den Übergang zum Phase-An-Zustand zu initiieren, schlägt vor, dass der geringe PapI-Level direkt nach der DNA-Replikation das kooperative Binden von Lrp auf den Lrp-Bindestellen 4–6 hervorruft und dass eine vollständige Methylierung der GATC^{prox} das kooperative Binden von Lrp auf den Stellen 1–3 hindert [313]. Bindet Lrp jedoch kooperativ auf den Bindestellen 1–3, blockt es die Transkription des *papBA*-Promotors, und der Phase-Aus-Phänotyp bleibt erhalten [313]. Eine *in-vitro*-Analyse legte den Schluss nahe, dass das kooperative Binden von Lrp auf den Seiten 4–6 nur bei geringer Lrp-Konzentration möglich ist [320]. Die Dam-Methylierung der GATC-Sequenzen innerhalb der *pap*-regulatorischen DNA ist aufgrund der flankierenden Sequenz wenig prozessiv, sodass trotz Methylierung der einen GATC-Seite selten die andere auch methyliert wird [325], [326]. Dies könnte ein weiterer wichtiger Faktor für das kooperative Binden vom Lrp-PapI-Komplex auf den Seiten 4–6 sein, weil somit die für die Phasenvariation notwendige Methylierung der GATC^{prox}-Stelle nicht automatisch zur erneuten vollständigen Methylierung der GATC^{dist}-Stelle führt, die ein Binden von PapI-Lrp auf den Stellen 4–6 blocken würde [325]. In jedem Fall wird der Phase-An-Zustand nach einer zweiten DNA-Replikations-Runde komplettiert, wenn der Lrp-PapI-Komplex auf der vollständig unmethylierten GATC^{dist}-Stelle bzw. kooperativ auf den Lrp-Bindestellen 4–6 bindet und GATC^{prox} letztendlich vollständig methyliert wird, sodass die Bindestellen 1–3 vor der Lrp-Bindung geschützt sind [313]. Auch wenn der Mechanismus der Phasenvariation vom An-Zustand in den Aus-Zustand noch nicht im Detail untersucht wurde [321], wurde in *in-vivo*-Experimenten gefunden, dass dieser Wechsel etwa hundertfach häufiger als die Variation vom Phase-Aus- zum Phase-An-Zustand stattfindet [327].

Ebenso essentiell für die Expression der *pap*-Gene ist offenbar das *Catabolite Activator Protein* (CAP), indem dieses Protein an seine Bindestelle, die sich an die PapB-Bindestelle 1 anschließt (Abb. 5.2), bindet [328], [329]. Es wurde vermutet, dass die Bindung von Lrp-PapI auf den Bindestellen 4–6 zu einer Krümmung der DNA führt, die es der zum *papBA*-Promotor zugewandten Untereinheit von CAP ermöglicht, mit der C-terminalen Domäne der α -Untereinheit der RNA-Polymerase wechselzuwirken [329]. Jedenfalls resultiert die CAP-abhängige Transkription der *pap*-Gene nicht aus der Wechselwirkung von CAP mit dem *papI*-Promotor oder der kodierenden Region, und CAP und Lrp binden

unabhängig voneinander auf ihren korrespondierenden Bindestellen innerhalb des *pap*-regulatorischen Bereichs [329]. Die Expression der *pap*-Gene wird zusätzlich durch Umwelteinflüsse, die weitere Regulatoren wie das CpxA-CpxR-Zwei-Komponenten-System [330] oder das *Histone-like Nucleoid Structuring Protein* (H-NS) [331], [332] involvieren, beeinflusst.

Mit dem Phasenvariationsmechanismus beschäftigen sich einige Übersichtsartikel, beispielsweise [321], [322], [333], an denen sich bei der Literaturrecherche zur Erstellung von Abschn. 5.1.3 orientiert wurde.

5.1.4. Vorarbeiten in zellfreier *E.-coli*-ähnlicher Umgebung

Durch eine Verstärkung der Dam-Expression *in vivo* kommt die Transkription der *pap*-Gene zum Erliegen, sodass daraus abgeleitet werden kann, dass die Konzentrationen der involvierten Regulatorproteine in einem aufeinander angepassten Verhältnis zueinander stehen müssen [325]. Ein ähnliches Ergebnis kommt bei hohen Konzentrationen von Lrp zustande [325]. Auch eine Veränderung der PapI-Konzentration hat Einfluss auf die Bindungsaffinität von Lrp gegenüber der *pap*-regulatorischen DNA [313]. In *in-vivo*-Experimenten kann sich die molekulare Zusammensetzung stetig ändern, da die Zelle in der Lage ist, auf verschiedene Reize zu reagieren [61]. Insofern ist es herausfordernd, ein solch komplexes und sensibel aufeinander abgestimmtes Zusammenspiel von verschiedenen Proteinen und der *pap*-regulatorischen DNA *in vivo* zu untersuchen. *In-vitro*-Bindungsexperimente, in denen die Bindung oder Wechselwirkung der in die Phasenvariation involvierten Proteine untereinander und mit der *pap*-regulatorischen DNA untersucht wurden, haben sich als geeignet erwiesen, verschiedene Aspekte der Phasenvariation zu untersuchen, so geschehen u.a. in [306], [313], [320], [325].

Das bereits in Kap. 3 verwendete zellfreie Expressionssystem von NOIREAUX [51], [52], das die Genexpression mittels der nativen *E.-coli*-RNA-Polymerase und ihrem σ^{70} -Faktor ermöglicht, bietet eine im Vergleich zur lebenden Zelle kontrollierbarere *E.-coli*-ähnliche Umgebung, auf die bereits eine vorangegangene Studie [60] zurückgriff, um Aspekte des Phasenvariationsmechanismus zu untersuchen. Die experimentelle Strategie und Ergebnisse aus [60] werden in diesem Abs. zusammengefasst. Sechs DNA-Fragmente, die jeweils eine der sechs Lrp-Bindestellen enthielten, wurden in dieser Studie in unterschiedlichen Kombinationen miteinander gekoppelt, mit dem Gen, das für das Reporterprotein deGFP kodiert, fusioniert und bildeten DNA-Templates für zellfreie Expressions-Experimente. Dabei wurden maximal drei Lrp-Bindestellen tragende DNA-Fragmente in verschiedenen Kombinationen mit dem *degfp*-Gen in einem DNA-Template verbunden. Damit das Reportergen dieser DNA-Templates zellfrei exprimiert werden konnte, wurde stets die letzte, sich vor dem *degfp*-Gen befindliche Lrp-Bindestelle vom -35- und -10-Element des σ^{70} -Faktors der *E.-coli*-RNA-Polymerase flankiert. Die verschiedenen DNA-Templates waren entweder vollständig methyliert oder unmethyliert. Die Gene, die für die Regulatorproteine Lrp und PapI kodieren, wurden getrennt voneinander von zwei Plasmiden, die sich in verschiedenen Konzentrationen ebenfalls in den zellfreien Reaktionen befanden, koexprimiert. Da die deGFP-Fluoreszenz-Intensität mit zunehmender Konzentration der zellfrei synthetisierten deGFP-Moleküle steigt, diente sie somit als Maß für die stattfindende zellfreie Expression des *degfp*-Gens. Durch kontinuierliche (Kinetik) oder Endpunkt-Messung der deGFP-Fluoreszenz-Intensität wurde somit auf die Effizienz der durch Lrp und PapI vermittelten Transkriptions-Repression des *degfp*-Gens bei den verschiedenen DNA-Templates geschlossen. Dies ermöglichte, Rückschlüsse auf die Wechselwirkung zwischen Lrp, PapI, der RNA-Polymerase und den verschiedenen kombinierten Lrp-Bindestellen zu ziehen. Es zeigte sich, dass die Repressions-Effizienzen bei den verschiedenen DNA-Templates nicht durch einfaches, von der Methylierung vermitteltes Binden oder nicht-Binden von Lrp und PapI und

dem damit verbundenen Blocken oder nicht-Blocken der RNA-Polymerase erklärt werden kann. Nicht notwendigerweise ging eine stärkere Bindung der Koregulatoren Lrp und PapI auf den DNA-Templates mit einer stärkeren Repression einher. Ein Beispiel dafür ist das untersuchte DNA-Template, das die Lrp-Bindestelle 5 einfach oder dreifach gekoppelt enthielt. Die *degfp*-Transkription ließ sich beim unmethylierten DNA-Template mit einfach vorhandener Lrp-Bindestelle 5 bei zunehmender Konzentration des *lrp*-Gen tragenden Plasmids und mit konstanter Konzentration des *papI*-Gen tragenden Plasmids stärker reprimieren, während die *degfp*-Expression beim gleichen DNA-Template in methylierter Form die gegenteilige Korrelation zeigte. Würde die Repression nur dem auf der DNA gebundenen Lrp-PapI, das die RNA-Polymerase blockt, basieren, wäre zu erwarten gewesen, dass sich bei dreifacher Kopplung der Lrp-Bindestelle 5 die für die einzelne Bindestelle 5 gesehenen Effekte verstärkten. Allerdings wurde die stärker werdende Repression der *degfp*-Expression bei steigender Konzentration des *lrp*-Gen tragenden Plasmids und bei konstanter Konzentration des *papI*-Gen tragenden Plasmids bei der dreifach gekoppelten Lrp-Bindestelle 5 nicht nur im unmethylierten Fall, sondern nun auch im methylierten Fall beobachtet.

Im Einklang mit diesen Ergebnissen wurde ebenfalls in [306] vermutet, dass die Topologie der DNA Einfluss auf den Phasenvariationsmechanismus hat.

5.1.5. Strategie und Zielsetzung der Studie

Die vorliegende Studie zielte darauf ab, an die zellfreien Expressions-Experimente aus den Vorarbeiten [60] anzuknüpfen. Wie bereits die vorangegangene Dissertation [60] beabsichtigte die vorliegende Studie, in *E.-coli*-ähnlicher, zellfreier Umgebung den UPEC-Phasenvariationsmechanismus in Abhängigkeit des Methylierungszustandes des *pap*-regulatorischen Bereichs und der Konzentration der Koregulatoren Lrp und PapI auf transkriptionellem Level zu analysieren. Jedoch bildeten in dieser Arbeit alle sechs Lrp-Bindestellen, im nativen Verbund miteinander gekoppelt, den zu analysierenden *pap*-regulatorischen Bereich.

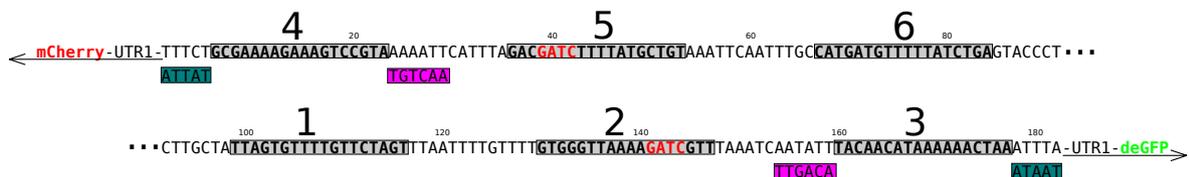


Abbildung 5.3.: Schematischer Aufbau des synthetischen *pap*-regulatorischen DNA-Konstrukts und Sequenzvergleich zur nativen Sequenz. Die Sequenz des nativen *pap*-regulatorischen Bereichs (GenBank ID: X14471.1, zurückgehend auf [59]) wird in den oberen Zeilen präsentiert (durchnummeriert). Abweichungen von dieser Sequenz beim synthetischen Konstrukt werden unterhalb angegeben. Die DNA-Sequenzen der sechs Lrp-Bindestellen (nach [313]) sind grau markiert, fett gedruckt und nummeriert. Die äußeren Lrp-Bindestellen 3 und 4 werden jeweils von den -35- (pinke Markierung) und -10-Elementen (türkisfarbene Markierung) des σ^{70} -Faktors der *E.-coli*-RNA-Polymerase flankiert. Stromaufwärts und -abwärts schließen sich die Gene, die für die fluoresszierenden Reporterproteine mCherry bzw. deGFP kodieren, an. Die GATC^{prox}- (141–144) und GATC^{dist}-Stelle (39–42) sind jeweils rot markiert. Werden beide oder eine dieser GATC-Stellen zu GCTC mutiert, wird diese Stelle nicht mehr Dam-methyliert, aber die Bindungsaffinität von Lrp und PapI gegenüber der mutierten Stelle bleibt unverändert [317]. Weitere Informationen sind dem Text zu entnehmen.

Das Kernstück der zellfreien Expressions-Experimente bildete ein synthetisches DNA-Konstrukt (Abb. 5.3), das dem nativen *pap*-regulatorischen Bereich (GenBank ID: X14471.1, zurückgehend auf [59]) in der Umgebung der sechs Lrp-Bindestellen bis auf die für die zellfreie Expression notwendigen Modifikationen entspricht. In Analogie zum Aufbau der DNA-Templates aus [60] wurde die native

DNA-Sequenz jeweils um die äußeren Lrp-Bindestellen 3 und 4 herum, wie in Abb. 5.3 angegeben, mutiert, sodass diese beiden Bindestellen von dem -35- und -10-Element des σ^{70} -Faktors der *E.-coli*-RNA-Polymerase flankiert werden. Die beiden Elemente stimmen mit denen des an einer Stelle mutierten λ -Repressor-Cro-Promotors OR2-OR1-Pr aus [51] überein. Die nicht-kodierenden Sequenzen und die *pap*-Gene, die sich stromaufwärts und -abwärts des korrespondierenden nativen DNA-Abschnitts befinden, wurden beim synthetischen DNA-Konstrukt beidseitig durch den untranslatierten Bereich UTR1, der die T7 *g10* Leitsequenz enthält [51], [276], ersetzt (Abb. 5.3). Es schließen sich stromaufwärts und -abwärts bei der synthetischen Sequenz die Gene, die für die fluoreszierenden Reporterproteine mCherry bzw. deGFP kodieren, an (Abb. 5.3)¹. Mittels der Protokolle von [60] wurde der Methylierungszustand von GATC^{prox} und GATC^{dist} vor der zellfreien Expression durch Dam-Methylierung des synthetischen DNA-Konstrukts festgelegt. Um die distale oder proximale GATC-Stelle selektiv unmethyliert lassen zu können, wurde auf eine Idee von [317] zurückgegriffen: Die Mutation der Sequenz der GATC-Stellen zu GCTC verändert die Bindungsaffinität der Koregulatoren Lrp und PapI gegenüber den involvierten Lrp-Bindestellen nicht, verhindert aber, dass diese Stellen Dam-methyliert werden können.

Durch diese Mutation von GATC^{dist} und Methylierung des DNA-Konstrukts entsteht das *pap*^(0,1)-DNA-Konstrukt mit dem Methylierungszustand (0,1). Der Methylierungszustand (1,0) des *pap*^(1,0)-Konstrukts entsteht dementsprechend durch die besagte Mutation von GATC^{prox} und anschließender Methylierung des Konstrukts. Wird die unmutierte, synthetische *pap*-DNA Dam-methyliert, entsteht das vollständig methylierte Konstrukt *pap*^(1,1) mit dem Methylierungszustand (1,1). Unter der Voraussetzung, dass die GCTC-Mutation das Bindungsverhalten von Lrp und PapI gegenüber dem *pap*-regulatorischen Bereich nicht verändert [317], und der Annahme, dass diese Punktmutation keinen Einfluss auf die Wechselwirkung der DNA mit sich selbst hat, sind die Sequenzunterschiede zwischen den drei *pap*-DNA-Konstrukten zu vernachlässigen; sie legen nur den Methylierungszustand fest². Insofern wird im Folgenden nur noch von einem synthetischen *pap*-regulatorischen DNA-Konstrukt oder vereinfacht dem *pap*-DNA-Konstrukt gesprochen, das die Methylierungszustände (0,1), (1,0) oder (1,1) haben kann.

Das Konzept von [60] übernehmend und in diesem Abs. zusammenfassend dienen die im deGFP- und mCherry-Kanal gemessenen Fluoreszenz-Intensitäten als Maß für das zellfrei synthetisierte deGFP bzw. mCherry. Die Transkription des *degfp*- und des *mcherry*-Gens wird durch das Binden des σ^{70} -Faktors der *E.-coli*-RNA-Polymerase auf den -35- und -10-Elementen, die sowohl die Lrp-Bindestelle 4 als auch die Lrp-Bindestelle 3 flankieren, initiiert. Somit kann die Transkriptionsrate der beiden Gene mit der entsprechenden Fluoreszenz-Intensität in Korrelation gesetzt werden [51], [52]: Je höher die gemessene Fluoreszenz-Intensität ist, desto höher ist auch die Transkriptionsrate des korrespondierenden Gens. Dieser Zusammenhang ermöglichte es, Rückschlüsse auf die Wechselwirkung zwischen dem *pap*-regulatorischen Bereich und den beiden Koregulatoren Lrp und PapI in Abhängigkeit des Methylierungszustandes zu ziehen. Die *lrp*- bzw. *papI*-Transkriptionsrate und somit die Konzentration des zellfrei synthetisierten Lrp-Moleküls bzw. PapI-Proteins steigt mit zunehmender Konzentration des Plasmids, das die Expression vom *lrp*- bzw. *papI*-Gen ermöglicht.

Dadurch, dass beim synthetischen *pap*-DNA-Konstrukt das *degfp*-Gen Platzhalter für die aneinandergereihten *pap*-Gene im nativen DNA-Konstrukt ist, konnte über die Messung der deGFP-Fluoreszenz-

¹Der untranslatierte Bereich und die beiden Gene wurden wie in [60] mit dem *pap*-regulatorischen Bereich verknüpft. Weitere Details zu den transkriptions-translations-regulatorischen Sequenzen UTR1 und OR2-OR1-Pr sowie zu den Reportergenen können den Abschn. 2.5 und 4.2.9.1 entnommen werden.

²Die Sequenzen der drei *pap*-DNA-Konstrukte und einen Überblick über alle weiteren GATC-Stellen sind in Anh. A.3.1.2 zu finden.

Intensität der Phasenvariationsmechanismus unter Involvierung aller sechs Lrp-Bindestellen untersucht werden. Durch die Integration des *mcherry*-Gens in die synthetische *pap*-regulatorische DNA und die Messung der mCherry-Fluoreszenz-Intensität, war es darüber hinaus möglich, die Wechselwirkung zwischen den proximalen und distalen Lrp-Bindestellen inklusive der u.a. in [306], [313], [320] postulierten kooperativen Bindung von Lrp auf den Lrp-Seiten 4–6 und 1–3, in An- oder Abwesenheit von PapI, zu analysieren. Wie bereits in [60] verlief die *papI*-Transkription dabei unabhängig von PapB und wurde gezielt durch Hinzugabe eines *papI*-Gen-tragenden Plasmids in die zellfreie Umgebung gesteuert. In den Vorarbeiten [60] wurde die zellfreie Expression des an die verschiedenen miteinander gekoppelten Lrp-Bindestellen fusionierten *degfp*-Gens, ohne die Integration der CAP-Bindestelle in die jeweiligen DNA-Templates, analysiert. Auch in dieser Studie wurde bei dem verwendeten synthetischen DNA-Konstrukt auf den Einbau dieser Stelle verzichtet.

Der minimalistische Aufbau der synthetischen *pap*-DNA und die experimentellen Bedingungen ermöglichten es, wesentliche Aspekte der Phasenvariation zu studieren und entkoppelten dabei involvierte Prozesse, wie beispielsweise die PapB-abhängige *papI*-Transkription. Durch Anwendung des Ansatzes [60], die Interaktion zwischen der *pap*-regulatorischen DNA und den beiden Koregulatoren in zellfreier *E.-coli*-ähnlicher Umgebung über *E.-coli*-RNA-Polymerase-Transkriptionsraten zu analysieren, wurden experimentelle Bedingungen geschaffen, die dem nativen *pap*-regulatorischen System möglichst nahe kommen. Insbesondere im Gegensatz zum *in-vivo*-Ansatz, bei dem verschiedene Umwelteinflüsse die molekulare Zusammensetzung der Zelle ständig verändern können [61], variierten die Konzentrationen von Lrp und PapI nur dann, wenn sie gezielt zellfrei synthetisiert wurden. Erweitert auf einen *pap*-regulatorischen Bereich, der alle sechs Lrp-Bindestellen in nativer Anordnung umfasst, sollte der Phasenvariationsmechanismus vor allem auf der Basis der Ergebnisse aus [60], [306] analysiert werden, dass die reine Topologie der DNA bereits Einfluss auf den Phasenvariationsmechanismus hat.

5.2. Material und Methoden

Alle wässrigen Lösungen wurden mit Reinstwasser (*arium*[®] 611UV Reinstwasseranlage, *Sartorius Stedim Biotech GmbH*, Göttingen), das autoklaviert (Varioklav[®] Dampfsterilisator, Typ 75 S, *H+P Labortechnik AG*, Oberschleißheim) wurde, hergestellt. Alle Materialien, die in Kontakt mit biologischem Material kamen, waren steril oder wurden zuvor autoklaviert (Varioklav[®] Dampfsterilisator, Typ 75 S, *H+P Labortechnik AG*, Oberschleißheim). DNA-, Enzym- oder Protein-Lösungen wurden auf Eis zubereitet. Die verschiedenen Ingredienzien der Lösungen wurden durch Auf- und Abpipettieren vermischt, kurz (2 s-3 s) auf mittlerer Drehzahl gevortext (*Vortex 3, IKA*[®]-*Werke GmbH*, Staufen) und mittels einer Tisch-Minizentrifuge für etwa 2 s-3 s abzentrifugiert (*Sprout Minizentrifuge, Biozym Scientific GmbH*, Hessisch Oldendorf). Alle Enzyme, deren zugehörige Reaktionspuffer und Komponenten als auch DNA-Lösungen wurden bei -20 °C gelagert. Zur Präparation einer Reaktion wurden die benötigten Komponenten auf Eis aufgetaut. Die Präparation der Reaktionen wurde ebenfalls auf Eis durchgeführt.

Wie bereits in Abschn. 5.1.5 beschrieben, haben die unterschiedlichen Sequenzen der *pap*-DNA-Konstrukte *pap*^(1,0), *pap*^(0,1) und *pap*^(1,1) keinen Einfluss auf die Interpretation der experimentellen Ergebnisse, weswegen von einem *pap*-DNA-Konstrukt gesprochen wird, das drei Methylierungszustände aufweisen kann. Für die Beschreibung der Herstellung und Präparation der in den Experimenten verwendeten DNA-Konstrukte muss die im Material- und Methoden-Teil (Abschn. 5.2) vollzogene Sequenz-Unterscheidung erfolgen.

Die verwendeten Standard-Methoden wurden bereits in Abschn. 4.2 eingeführt und beschrieben. Abschn. 5.2 baut auf den dortigen Ausführungen auf, jedoch ergaben sich Besonderheiten, sodass Standard-Protokolle, insbesondere für die Herstellung der *Pap*-DNA-Konstrukte, modifiziert wurden. Die Abweichungen von den in Abschn. 4.2 beschriebenen Standard-Protokollen werden angegeben. Die Synthese und Amplifizierung der verschiedenen *pap*-DNA-Konstrukte in geeigneten Mengen gestaltete sich schwierig und brachte nur mit den im Folgenden angegebenen Konfigurationen adäquate Ausbeuten.

Im Rahmen seiner Einarbeitung in die zellfreie Genexpression und die molekularbiologischen Methoden unterstützte Marc Finkler diese Studie unter Anleitung des Autors dieser Dissertation, indem er DNA synthetisierte und präparierte. Mit Hilfe der Protokolle (vgl. Abschn. 5.2.1.1), die vom Autor dieser Dissertation entwickelt wurden, synthetisierte, vervielfältigte und präparierte Marc Finkler die *pap*-DNA-Konstrukte, mit denen die in den Abschn. 5.3.1 und 5.3.2 vorgestellten Messungen durchgeführt wurden. Zusätzlich amplifizierte und präparierte er die Plasmide pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500, bereitete die DNA für die Sequenzierung vor und half bei der Analyse der Sequenzierungs-Ergebnisse. Des Weiteren waren Julia Flesch und Lucina Kainka an Vorarbeiten beteiligt, deren Beiträge an geeigneter Stelle angegeben werden.

5.2.1. Herstellung der verwendeten DNA-Konstrukte

5.2.1.1. Herstellung der *pap*-DNA-Konstrukte

Die linearen *pap*-DNA-Konstrukte wurden durch Gibson-Assemblierung (vgl. Abschn. 4.2.8) von PCR-Produkten (vgl. Abschn. 4.2.5) und durch anschließende PCR-Amplifikation hergestellt.

5.2.1.1.1. *pap*^(1,1)-Konstrukt

Zuerst wurde ein lineares, unmethyliertes *pap*^(1,1)-Konstrukt durch Gibson-Assemblierung zweier PCR-Produkte und anschließender PCR-Amplifikation hergestellt. Erst durch anschließende Dam-Methylierung dieses Produkts (vgl. Abschn. 5.2.2) wurde der Methylierungszustand (1,1) erreicht.

Als PCR-Template-DNA zur Synthese von Teilfragment I diente das Plasmid pBEST-p15a-OR2-OR1-PR-UTR1-mCherry-T500, das von der Kooperationsarbeitsgruppe von Vincent Noireaux (*School of Physics and Astronomy, University of Minnesota*) bereitgestellt wurde. Es basiert auf dem Plasmid pBEST-OR2-OR1-PR-UTR1-deGFP-T500 (Addgene Plasmid # 40019), das auf [51] zurückgeht³. In dieser DNA-Konstruktion substituiert das *mcherry*-Gen (GenBank ID: AY678264.1, auf [157] zurückgehend, vgl. Abschn. 2.5), das bereits in [60] als Reportergen verwendet wurde, das *degfp*-Gen. Die Sequenzen des Vorwärtsprimers und des Rückwärtsprimers waren 5'-GAGCTGACTGGGTTGAAGG-3' bzw. 5'-GCTAGCAATAATTTTGTTTAACTTTAAGAAGGAGATATAC-3'.

Ein DNA-Konstrukt, das sich aus den Sequenzen der transkriptions-translations-regulatorischen und kodierenden Elemente 5'-UTR1-*pap*-regulatorischer-Bereich-UTR1-deGFP-T500-...-3' zusammensetzt, wurde von *Eurofins Genomics GmbH* (Ebersberg) synthetisiert, in den Vektor pEX-K4 kloniert und geliefert. Die vollständige Sequenz des synthetisierten DNA-Konstrukts, das in den Vektor kloniert wurde, ist in Anh. A.3.1.1 angegeben. Die Sequenz des *pap*-regulatorischen Bereichs entspricht der Sequenz, die in Abb. 5.3 dargestellt wird. Das gelieferte Plasmid diente als PCR-Template zur Herstellung des Teilfragments II. Die Sequenz des Vorwärtsprimers war 5'-GTAAACAAAATTATTGCTAGCTGCAACCATTATGCG-3', die des Rückwärtsprimers lautete 5'-CACCATCAGCCAGAAAACC-3'. Das 5'-Ende des Vorwärtsprimers bildete den für die Gibson-Assemblierung notwendigen Überlapp mit Teilfragment I.

Für die Gibson-Assemblierung der beiden Teilfragmente wurde im Vergleich zum Standard-Protokoll (vgl. Abschn. 4.2.8) das folgende modifizierte Temperatur-Zeit-Programm verwendet: 1 min bei 50 °C, 1 min bei 45 °C, 1 min bei 50 °C, 1 min bei 45 °C, 26 min bei 50 °C. Im Anschluss an die Gibson-Assemblierung erfolgte eine PCR unter Verwendung des Vorwärts- und Rückwärtsprimers mit der Sequenz 5'-GAGCTGACTGGGTTGAAGG-3' bzw. 5'-CACCATCAGCCAGAAAACC-3'. Es wurde herausgefunden, dass bei der PCR das DNA-Konstrukt der richtigen Länge unter der Bedingung synthetisiert werden konnte, dass die PCR-Template-DNA direkt aus der Gibson-Assemblierungsreaktion stammte: 10 µl der Gibson-Assemblierungs-Reaktion wurden in eine insgesamt 300 µl große PCR-Reaktion transferiert. Die PCR verlief mit der Taq-DNA-Polymerase unter ThermoPol®-Reaktionspuffer-Bedingungen, die Annealing-Temperatur lag bei 50 °C, die Elongationszeit betrug 1 min und 55 s, die PCR durchlief 18 Zyklen. Aufgrund der Entstehung von DNA-Nebenprodukten anderer Längen während der PCR, musste das DNA-Konstrukt mit der richtigen Länge mittels Agarosegelelektrophorese purifiziert werden (vgl. Abschn. 4.2.4, Abschn. 4.2.5). Es konnten keine Bedingungen gefunden werden, unter denen die purifizierte DNA, als PCR-Template dienend, mittels PCR effizient amplifiziert werden konnte⁴. Hingegen lieferten PCRs, bei denen die PCR-Template-DNA, wie oben beschrieben, über Hinzugabe eines Teils der Gibson-Assemblierungsreaktion bereitgestellt wurde, auf reproduzierbare Weise das PCR-Produkt richtiger Länge, auch wenn stets

³Details zum Plasmid, zu den transkriptions-translations-regulatorischen Elementen und dem *degfp*-Gen können Abschn. 4.2.9.1 und 2.5 entnommen werden.

⁴Experimente, die darauf abzielten, die purifizierte DNA mittels PCR zu amplifizieren, wurden an Julia Fleisch als Teilprojekt ihrer Masterarbeit und im Rahmen ihrer Anstellung als studentische Hilfskraft am Lehrstuhl für Biologische Experimentalphysik an der Universität des Saarlandes übertragen und angeleitet. Dabei unterstützte Lucina Kainka ebenfalls im Rahmen ihrer Anstellung als studentische Hilfskraft.

Nebenprodukte entstanden. Es konnte keine Erklärung gefunden werden, weswegen die PCR unter diesen Bedingungen adäquate Ergebnisse erzielte und in den anderen Fällen nicht.

Des Weiteren zeigte sich, dass die Verwendung anderer PCR-Primer zu höheren Ausbeuten führte: Basierend auf dem oben beschriebenen Protokoll wurde das unmethylierte *pap*^(1,1)-Konstrukt durch PCR unter Verwendung des Vorwärts- und Rückwärtsprimers mit den Sequenzen 5'-CGCTATGAATTCCTCGAGTTACTTGTACAGCTC-3' bzw. 5'-CCATGGAAGCTTTTAGATCCCGGCGGCGGTCA-3' synthetisiert. Dabei lieferten erneut 10 µl der Gibson-Assemblierungsreaktion die PCR-Templates für die PCR-Reaktion, die wiederum ein Volumen von insgesamt 300 µl hatte. Diese Gibson-Assemblierungsreaktion wurde, wie im letzten Abs. beschrieben, durchgeführt. Die PCR verlief mit der Taq-DNA-Polymerase unter ThermoPol®-Reaktionspuffer-Bedingungen, die Annealing-Temperatur lag bei 56 °C, die Elongationszeit betrug 1 min und 55 s und die PCR durchlief 15 Zyklen. Da erneut neben dem PCR-Produkt der richtigen Länge DNA-Stränge anderer Längen entstanden, wurde das PCR-Produkt der richtigen Länge mittels Agarosegel-Extraktion von den anderen Strängen getrennt und purifiziert⁵. Die Sequenz des unmethylierten bzw. methylierten *pap*^(1,1)-Konstrukts ist in Anh. A.3.1.2 angegeben. Es wurde einmalig vollständig mit den Sequenzierungsprimern 5'-CGCTATGAATTCCTCGAGTTACTTGTACAGCTC-3', 5'-CCATGGAAGCTTTTAGATCCCGGCGGCGGTCA-3', 5'-AAGACCATGGGCTGGGAG-3', 5'-CTCCATGTGCACCTTGAAG-3' und 5'-AACATCCTGGGGCACAAG-3' sequenziert. Weitere Sequenzierungen erfolgten mit den Sequenzierungs- bzw. PCR-Primern 5'-CGCTATGAATTCCTCGAGTTACTTGTACAGCTC-3' und 5'-CCATGGAAGCTTTTAGATCCCGGCGGCGGTCA-3', mit denen die Sequenzen des *pap*-regulatorischen Bereichs und der umschließenden UTR1-Regionen vollständig sowie das *degfp*- bzw. *mcherry*-Gen bis auf wenige Nukleotide am Gen-Ende verifiziert werden konnten.

Wie zu Beginn des letzten Abs. beschrieben, ließen sich typischerweise mit den PCR-Primern mit den Sequenzen 5'-CGCTATGAATTCCTCGAGTTACTTGTACAGCTC-3' bzw. 5'-CCATGGAAGCTTTTAGATCCCGGCGGCGGTCA-3' höhere DNA-Ausbeuten des *pap*-PCR-Produkts der richtigen Länge erzielen. Durch diese Primer wurde nur der DNA-Bereich amplifiziert, der zwischen den beiden Terminatoren T500 liegt, die sich nach der Gibson-Assemblierung jeweils an den Enden der DNA-Konstrukte befanden. Möglicherweise wurde durch das Fehlen der beiden endständigen Terminatoren, deren Sequenzen zueinander komplementär sind, die Bildung von Sekundärstrukturen während der PCR verhindert.

5.2.1.1.2. *pap*^(1,0)-Konstrukt

Das unmethylierte, lineare *pap*^(1,0)-Konstrukt wurde durch Gibson-Assemblierung zweier PCR-Produkte (Teilfragmente) und anschließender PCR-Amplifikation hergestellt. Erst durch anschließende Dam-Methylierung dieses Produkts (vgl. Abschn. 5.2.2) wurde der Methylierungszustand (1,0) erreicht.

⁵Marc Finkler synthetisierte zwei weitere Teilfragmente I und II, die auf die gerade beschriebene Weise mittels Gibson-Assemblierung verbunden wurden und somit ebenfalls auf alternativem Weg zum unmethylierten *pap*^(1,1)-Konstrukt führten. Bei der Standard-PCR zur Herstellung dieses alternativen Paares an Teilfragmenten dienten der Vorwärtsprimer 5'-CGCTATGAATTCCTCGAGTT-3' und der Rückwärtsprimer 5'-AGATAAAAACATCATGGCAAATTG-3' zur Synthese von Teilfragment I, zur Synthese von Teilfragment II wurden der Vorwärtsprimer 5'-TTTGCCATGATGTTTTTATCTGA-3' und der Rückwärtsprimer 5'-CCATGGAAGCTTTTAGATCC-3' verwendet. Das PCR-Template für die Herstellung dieser beiden Teilfragmente bildete das unmethylierte DNA-Konstrukt *pap*^(1,1).

Als PCR-Template für beide Teilfragmente diente das *pap*-DNA-PCR-Produkt, das mit Hilfe des Vorwärts- und Rückwärtsprimers mit den Sequenzen 5'-GAGCTGACTGGGTTGAAGG-3' bzw. 5'-CACCATCAGCCAGAAAACC-3', wie in Abschn. 5.2.1.1.1 beschrieben, synthetisiert und mittels Agarosegel-Extraktion purifiziert wurde.

Teilfragment I entstand durch eine PCR unter Verwendung des Vorwärts- und Rückwärtsprimers mit der Sequenz 5'-CATCAGCCAGAAAACCGAGCTGACTGGGTTGAAG-3' bzw. 5'-AGATTTAAACGA[G]CTTTTAACCCACAAAAC-3'.

Teilfragment II wurde durch eine PCR unter Verwendung des Vorwärts- und Rückwärtsprimers mit der Sequenz 5'-TAAAAG[C]TCGTTTAAATCTTGACATACAACATAAAAAAC-3' bzw. 5'-CTGACTGGGTTGAAGGCACCATCAGCCAGAAAAC-3' synthetisiert.

Der Rückwärtsprimer zur Synthese von Teilfragment I als auch der Vorwärtsprimer zur Herstellung von Teilfragment II erzeugten den zur Gibson-Assemblierung notwendigen Überlapp der Teilfragmente I und II. Diese Primer, die jeweils im Bereich der Lrp-Bindestelle 2 während der PCR auf dem PCR-Template hybridisierten, waren an der proximalen GATC-Stelle am Adenin-Basen-Nukleotid nicht komplementär zum Template (eckige Klammer). Die entstehenden Teilfragmente I und II besaßen daher in ihrer überlappenden Sequenz die mutierte GCTC-Stelle nach [317]. Die Gibson-Assemblierung erfolgte wie in Abschn. 5.2.1.1.1 beschrieben.

Das unmethylierte *pap*^(1,0)-Konstrukt wurde, wie das unmethylierte *pap*^(1,1)-Konstrukt, durch PCR unter Verwendung des Vorwärts- und Rückwärtsprimers mit den Sequenzen 5'-CGCTATGAATTCCTCGAGTTACTTGTTACAGCTC-3' bzw. 5'-CCATGGAAGCTTTTATAGATCCCGGCGGCGGTCA-3' synthetisiert. Dabei lieferten 10 µl der Gibson-Assemblierungsreaktion die PCR-Templates für die PCR-Reaktion, die ein Volumen von insgesamt 300 µl hatte. Die PCR verlief mit der Taq-DNA-Polymerase unter ThermoPol[®]-Reaktionspuffer-Bedingungen, die Annealing-Temperatur lag bei 56 °C, die Elongationszeit betrug 1 min und 55 s und die PCR durchlief 15 Zyklen. Auch nach dieser PCR mussten die PCR-Produkte der richtigen Länge mittels Agarosegel-Extraktion von den entstandenen Strängen anderer Längen getrennt und purifiziert werden. Die Sequenz des unmethylierten bzw. methylierten *pap*^(1,0)-Konstrukts ist in Abschn. A.3.1.2 angegeben. Die Sequenzierung verlief wie für das unmethylierte *pap*^(1,1)-DNA-Konstrukt in Abschn. 5.2.1.1.1 beschrieben.

5.2.1.1.3. *pap*^(0,1)-Konstrukt

Das unmethylierte, lineare *pap*^(0,1)-Konstrukt wurde durch Gibson-Assemblierung zweier PCR-Produkte (Teilfragmente) und anschließender PCR-Amplifikation hergestellt. Erst durch anschließende Dam-Methylierung dieses Produkts (vgl. Abschn. 5.2.2) wurde der Methylierungszustand (0,1) erreicht.

Als PCR-Template für beide Teilfragmente diente das *pap*-DNA-PCR-Produkt, das mittels des Vorwärts- und Rückwärtsprimers mit den Sequenzen 5'-GAGCTGACTGGGTTGAAGG-3' bzw. 5'-CACCATCAGCCAGAAAACC-3', wie in Abschn. 5.2.1.1.1 beschrieben, synthetisiert und mit Hilfe einer Agarosegel-Extraktion purifiziert wurde.

Teilfragment I entstand durch eine PCR unter Verwendung des Vorwärts- und Rückwärtsprimers mit der Sequenz 5'-CATCAGCCAGAAAACCGAGCTGACTGGGTTGAAG-3' bzw. 5'-CAGCATAAAAAGA[G]CGTCTAAATGTTGACATAC-3'.

Teilfragment II wurde durch eine PCR unter Verwendung des Vorwärts- und Rückwärtsprimers mit der Sequenz 5'-AGACG[C]TCTTTTATGCTGTAAATTCAATTTGC-3' bzw. 5'-CTGACTGGGTTGAAGGCACCATCAGCCAGAAAAC-3' synthetisiert.

Der Rückwärtsprimer zur Synthese von Teilfragment I als auch der Vorwärtsprimer zur Herstellung von Teilfragment II erzeugten den zur Gibson-Assemblierung notwendigen Überlapp der Teilfragmente I und II. Diese Primer, die jeweils im Bereich der Lrp-Bindestelle 5 während der PCR auf dem PCR-Template hybridisierten, waren an der distalen GATC-Stelle am Adenin-Basen-Nukleotid nicht komplementär zum Template (eckige Klammer). Die entstehenden Teilfragmente I und II besaßen daher in ihrer überlappenden Sequenz die mutierte GCTC-Stelle nach [317]. Die Gibson-Assemblierung erfolgte wie in Abschn. 5.2.1.1.1 beschrieben.

Das unmethylierte *pap*^(0,1)-Konstrukt wurde, wie die unmethylierten *pap*^(1,1)- und *pap*^(1,0)-Konstrukte, durch PCR unter Verwendung des Vorwärts- und Rückwärtsprimers mit den Sequenzen 5'-CGCTATGAATTCCTCGAGTTACTTGTACAGCTC-3' bzw. 5'-CCATGGAAGCTTTTAGATCCCGGCGGCGGTCA-3' synthetisiert. Das PCR-Protokoll und die darauffolgende Purifizierung verliefen wie in Abschn. 5.2.1.1.1 und 5.2.1.1.2 beschrieben. Die Sequenz des unmethylierten bzw. methylierten *pap*^(0,1)-Konstrukts ist in Anh. A.3.1.2 angegeben. Die Sequenzierung verlief wie für das unmethylierte *pap*^(1,1)-Konstrukt in Abschn. 5.2.1.1.1 beschrieben.

5.2.1.2. pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500

Die Konstruktion der Plasmide pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500 wurde von Marc Schenkelberger in den Vorarbeiten [60] durchgeführt⁶. Diese Konstruktionen dienten bereits in seiner Dissertation [60], auf der diese Studie aufbaut, zur zellfreien Biosynthese des Proteins Lrp (GenBank ID: 949051) und PapI (GenBank ID: 1039535). Die Sequenzen der transkriptions-translations-regulatorischen und kodierenden Elemente sind in Anh. A.3.1.3 und A.3.1.4 angegeben. Die relevanten transkriptions-translations-regulatorischen Bereiche des Plasmids wurden mit dem Sequenzierungsprimer 5'-CACCATCAGCCAGAAAACC-3' sequenziert. Wenn nötig, wurde zusätzlich mit dem Sequenzierungsprimer 5'-GAGCTGACTGGGTTGAAGG-3' sequenziert.

5.2.2. Dam-Methylierung der *pap*-DNA-Konstrukte

Die *in-vitro*-Methylierung der purifizierten, unmethylierten *pap*^(1,1)-, *pap*^(1,0)- und *pap*^(0,1)-Konstrukte erfolgte mit Hilfe einer DNA-Adenin-Methylase (Dam), einem mitgelieferten 10x Reaktionspuffer und S-Adenosylmethionin (SAM) (*New England Biolabs GmbH*, Frankfurt am Main) (vgl. Abschn. 2.3.2). Das verwendete, von den Herstellerangaben abweichende Protokoll, wurde in [60] etabliert, für diese Studie übernommen und wird im Folgenden zusammengefasst. Das Reaktionsvolumen wurde auf 50 µl fixiert. Es herrschten 1x Dam-Reaktionspufferbedingungen. SAM besaß eine Konzentration von 80 µM. Innerhalb dieser 50 µl-Reaktion wurde die Menge des hinzugegebenen Enzyms Dam in Abhängigkeit der zu methylierenden DNA-Masse skaliert: Für 10 µg DNA wurden 25 µl Enzym-Stocklösung

⁶Details zu den Plasmiden bzw. den transkriptions-translations-regulatorischen Elementen können Abschn. 4.2.9.1 entnommen werden.

(8 units/ μ l) verwendet. Die 50 μ l-Reaktionen wurden für 3 h in einem Wasserbad oder mit dem *MultiGene™ 96-Well Gradient Thermal Cycler* (Labnet International Inc., Edison, USA) bei 37 °C inkubiert. Im Anschluss wurde die DNA mit Hilfe des *PureLink® PCR Purification Kits* (Fisher Scientific GmbH, Schwerte) gemäß der Hersteller-Instruktionen purifiziert und in Reinstwasser gelöst. Das Funktionieren dieses Methylierungs-Protokolls wurde durch einen Restriktionsverdau mit dem Enzym DpnI (*New England Biolabs GmbH*, Frankfurt am Main) verifiziert. DpnI schneidet DNA nur an methylierten GATC-Stellen zwischen A und T und hinterlässt glatte Enden. Dieser Standard-Verdau verlief in der Regel nach Hersteller-Instruktionen: In einer 50 μ l-Reaktion wurden in etwa 1 μ g DNA verdaut, von den 50 μ l stammten 5 μ l vom mitgelieferten 10x Reaktionspuffer (*New England Biolabs GmbH*, Frankfurt am Main) und 1 μ l von der DpnI-Stocklösung (20 units/ μ l). Je nach DNA-Konzentration der zu verdauenden DNA wurde die DNA-Masse bis auf 0,5 μ g, die eingesetzte DpnI-Menge bis auf 0,5 μ l und das Reaktionsvolumen bis auf 25 μ l herunterskaliert. Die durch den Verdau bei erfolgreicher Methylierung entstehende Fragmentierung der verdauten DNA wurde mittels Agarosegelelektrophorese analysiert (vgl. Abschn. 4.2.4).

5.2.3. Präparation der zellfreien Reaktionen, Reaktionsbedingungen und Fluoreszenz-Intensitäts-Messungen

Die Analyse der Regulation des epigenetischen *E.-coli*-Pili-Phasenvariationsmechanismus wurde ebenfalls im zellfreien Expressionssystem [51], [52] realisiert, das bereits zum zellfreien Einbau nicht-kanonischer Aminosäuren (vgl. Kap. 4) verwendet wurde. Zur Präparation der zellfreien Reaktionen wurde Zellextrakt aus der Produktionsreihe verwendet, die Marc Schenkelberger herstellte und in seiner Studie [60] benutzte, um eine bessere Vergleichbarkeit mit seinen Ergebnissen zu gewährleisten. Eine typische zellfreie Reaktion enthielt die in Abschn. 4.2.14.1 angegebenen Komponenten in den dort angegebenen Konzentrationen mit Ausnahme der folgenden Abweichungen: Die Konzentrationen für zusätzliches Magnesium- und Kaliumglutamat beliefen sich auf 3 mM bzw. 50 mM, da diese Konzentrationen zu einer optimalen Expressionseffizienz führten [60]; wie in [60] besaß jede kanonische Aminosäure eine Konzentration von 0,5 mM, bis auf Leu, das 0,42 mM konzentriert war (7,5 μ l vom Aminosäuren-Mix, der alle 20 kanonischen Aminosäuren (jeweils 6 mM, nur Leu mit 5 mM) bereitstellte (vgl. Abschn. 4.2.14.2)).

Die Präparation der zellfreien Reaktionen baute auf den Protokollen von [51], [52] auf. Sie entsprach dem Prinzip der Strategie, die bereits bei der Präparation der zellfreien Reaktion zum ortsspezifischen Einbau nicht-kanonischer Aminosäuren in Modellproteine Anwendung fand (vgl. Abschn. 4.2.14.3). Die Mastermix-Reaktionen wurden nicht bis zum Gesamtvolumen von 90 μ l vervollständig, sondern ein gewisser Volumenanteil wurde eingespart. Die unvollständig präparierten Mastermix-Reaktionen wurden in gleiche Volumina gesplittet. Diese gesplitteten Volumina wurden dann durch Hinzugabe weiterer Biomoleküle vervollständig. Somit ließen sich Konzentrationen einer oder mehrerer Biomoleküle innerhalb einer Serie gesplitteter zellfreier Reaktionen variieren und die zellfreie Synthese des zu synthetisierenden Modellproteins in Abhängigkeit dieser Variationen analysieren. Um die linearen *pap*-DNA-Templates vor enzymatischer Degradation zu schützen, beinhalteten die Reaktionen 3,3 μ M der verkürzten Version GamS des λ -Phagen Proteins Gam in rekombinanter und purifizierter Form, das die *E.-coli*-RecBCD-Nuklease-Aktivität inhibiert [134] (vgl. Abschn. 2.4.3.1.3)⁷. In der Regel wurde

⁷Das rekombinante und His-Tag purifizierte GamS-Protein wurde von der Kooperationsarbeitsgruppe von Vincent Noireaux (*School of Physics and Astronomy, University of Minnesota*) zur Verfügung gestellt. Die transkriptions-translations-regulatorischen und kodierenden DNA-Sequenzen und weitere Details sind in Anh. A.2.2.7 zu finden.

das Volumen der Mastermix-Reaktionen und aller zugehörigen Komponenten nach oben skaliert, um größere Versuchsreihen unter möglichst gleichen Reaktionsbedingungen durchführen zu können.

Bei den in Abschn. 5.3.1 präsentierten Experimenten wurden 20% Volumen bei der Herstellung der Mastermixes eingespart. Die Mastermixes wurden in Volumina von 9,6 μl gesplittet. Von jedem der drei unterschiedlich methylierten *pap*-DNA-Konstrukte wurde eine 25 nM DNA-Lösung hergestellt. Die gesplitteten 9,6 μl -Reaktionen wurden mit je 2,4 μl von einer der drei 25 nM DNA-Lösungen oder mit 2,4 μl Reinstwasser auf das Gesamtvolumen von 12 μl ergänzt. Demnach betragen die Konzentrationen der *pap*-DNA-Konstrukte 5 nM bzw. 0 nM (mit Reinstwasser ergänzte Reaktionen). Die zellfreien Reaktionen, die mit Reinstwasser ergänzt wurden, dienten als Referenz, in der keine Template-DNA vorhanden war.

Bei den in Abschn. 5.3.2 präsentierten Experimenten wurden ebenfalls 20% Volumen bei der Herstellung der Mastermixes eingespart. Die Mastermixes wurden in Volumina von 9,6 μl gesplittet. Von den beiden Plasmiden pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500 wurden jeweils 0 nM, 1 nM und 10 nM wässrige Lösungen hergestellt. Die gesplitteten 9,6 μl -Reaktionen wurden je mit 1,2 μl einer der Lösungen von pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und 1,2 μl einer der Lösungen von pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500 ergänzt. In Referenz-Reaktionen, in denen keines der beiden oder nur eines der beiden Plasmide vorkommen sollte, wurden die 9,6 μl -Reaktionen mit 2,4 μl bzw. 1,2 μl Reinstwasser vervollständigt. Die in den zellfreien Reaktionen gelösten *pap*-DNA-Konstrukte besaßen nach der Vervollständigung eine Konzentration von 5 nM, die Konzentrationen der beiden Plasmide betragen entweder je 0 nM, 0,1 nM oder 1 nM.

Bei den in Abschn. 5.3.1 präsentierten Experimenten wurden die zellfreien Reaktionen auf zwei unterschiedliche Weisen inkubiert. Die Hälfte aller 12 μl -Reaktionen wurde auf zufällige Weise im Inkubator verteilt und inkubiert. Von der anderen Hälfte der 12 μl -Reaktionen wurden jeweils 10 μl in je ein *Well* einer 384-*Well*-Platte (*Nunc 384-well optical bottom plate*, *Thermo Fisher Scientific Inc.*, Waltham, USA) gefüllt. Es wurde darauf geachtet, dass sich die Reaktionen in unterschiedlichen Arealen der 384-*Well*-Platte befanden. Die entsprechenden *Wells* wurden mit einer Klebefolie (*Nunc sealing tape*, *Thermo Fisher Scientific Inc.*, Waltham, USA) verschlossen. Die Inkubation dieser Reaktionen fand im Mikroplattenleser *POLARstar® Optima* (*BMG Labtech GmbH*, Offenburg) statt. Die Inkubationsdauer von 16 h und die Temperatur von 29 °C waren bei beiden Methoden gleich. Bei den im Mikroplattenleser inkubierten Reaktionen wurde eine Kinetik-Messung vorgenommen, bei der die Messung der Fluoreszenz-Intensität alle 231 s erfolgte. Jeweils 10 μl der im Inkubator inkubierten Reaktionen wurden am Ende der Inkubation in je ein *Well* einer 384-*Well*-Platte pipettiert und das entsprechende Areal mit Klebefolie verschlossen. Bei diesen Reaktionen erfolgte eine Endpunkt-Messung der Fluoreszenz-Intensität.

Bei den in Abschn. 5.3.2 präsentierten Experimenten wurden hingegen nur Endpunkt-Messungen auf die gleiche Weise, wie im vorangegangenen Abs. beschrieben, durchgeführt.

Wie auch schon in [60] betrug bei den Messungen der deGFP-Fluoreszenz-Intensität die Anregungs- bzw. Emissionswellenlänge 485 nm bzw. 520 nm; bei den Messungen der mCherry-Fluoreszenz-Intensität betrug die Anregungs- bzw. Emissionswellenlänge 595 nm bzw. 620 nm. Um alle Messungen innerhalb dieser Studie untereinander vergleichen zu können, wurden für alle deGFP-Fluoreszenz-Messungen der gleiche *gain* verwendet. Dies wurde ebenfalls für die mCherry-Fluoreszenz-Intensitäts-Messungen beachtet.

5.2.4. Vorhersage der *pap*-DNA-Sekundärstruktur

Die Sekundärstruktur wurde mit Hilfe des *mfold web servers* [334] vorhergesagt. Die zugrundeliegende DNA-Sequenz entsprach der Sequenz des synthetischen *pap*-regulatorischen Bereichs inklusive der stromaufwärts und -abwärts liegenden UTR1-Sequenzen und den jeweiligen Anfängen des *degfp*- und *mcherry*-Gens (Basenpaare 685–1053 in Sequenz aus Abschn. A.3.1.2). Die Struktur wurde bei einer Temperatur von 29 °C (Inkubationstemperatur der zellfreien Reaktionen) und mittels der Standardeinstellungen vorhergesagt (DNA-Typ: lineare DNA, ionische Bedingungen: 1 M Na⁺, Korrektur-Typ: Oligomer, *percent suboptimality number*: 5, *upper bound on the number of computed foldings*: 50, maximale Distanz zwischen gepaarten Basen: kein Limit). Die vorhergesagten Strukturen wurden in das Programm *Rna Viz* [335] importiert, mit dem die Abb. 5.11, A.16 und A.17 erstellt wurden.

5.3. Ergebnisse und Diskussion

Die Regulation des epigenetischen *E.-coli*-Pili-Phasenvariationsmechanismus vollzieht sich auf Ebene der Transkription [59] durch Modulation des DNA-Methylierungsmusters des *pap*-regulatorischen Bereichs und involviert neben weiteren Koregulatoren vor allem Lrp und PapI [317]. Die folgenden Experimente untersuchten in zellfreier *E.-coli*-ähnlicher Umgebung den Phasenvariationsmechanismus in Abhängigkeit des Methylierungsmusters der synthetischen *pap*-regulatorischen DNA bei gleichzeitiger zellfreier Synthese oder ohne gleichzeitige zellfreie Synthese von Lrp und PapI. Zusätzlich wurde die Wechselwirkung des *pap*-regulatorischen Bereichs mit sich selbst theoretisch analysiert.

5.3.1. Epigenetische Regulation des *pap*-regulatorischen Bereichs ohne zellfreie Lrp- und PapI-Synthese

Das folgende Experiment diente der Analyse des *pap*-regulatorischen Bereichs in Abhängigkeit des Methylierungsmusters der synthetischen *pap*-regulatorischen DNA, ohne dabei die Koregulatoren Lrp oder PapI zusätzlich zellfrei zu synthetisieren. Zu diesem Zweck wurden drei zellfreie Reaktionen präpariert, die sich in ihrer Zusammensetzung beim Methylierungsmuster der hinzugegebenen *pap*-regulatorischen DNA unterschieden. Den Reaktionen (0,1), (1,0) und (1,1) wurde das *pap*-DNA-Konstrukt mit den Methylierungszuständen (0,1), (1,0) bzw. (1,1) hinzugefügt (5 nM). Das Methylierungsmuster (0,1) entspricht dem *pap*-regulatorischen Bereich, dessen GATC^{dist}-Stelle unmethyliert und GATC^{prox}-Stelle methyliert ist. Dementsprechend besitzt das *pap*-DNA-Konstrukt mit dem Methylierungszustand (1,0) eine methylierte GATC^{dist}- und unmethylierte GATC^{prox}-Stelle, beim *pap*-DNA-Konstrukt mit dem Methylierungsmuster (1,1) sind sowohl GATC^{dist} als auch GATC^{prox} methyliert. Zusätzlich wurde eine Referenz-Reaktion (x) hergestellt, die die gleiche Zusammensetzung wie die anderen drei Reaktionen hatte, aber keine *pap*-regulatorische DNA beinhaltet. Alle zellfreien Reaktionen wurden 16 h bei einer Temperatur von 29 °C inkubiert, im Anschluss folgten Endpunkt-Messungen der deGFP- und mCherry-Fluoreszenz-Intensität. Abb. 5.4 zeigt die gemessenen Endpunkt-Fluoreszenz-Intensitäten der zellfreien Reaktionen im deGFP- und mCherry-Kanal.

In allen zellfreien Reaktionen, in denen die *pap*-Template-DNA vorhanden war, kam es unabhängig vom Methylierungszustand des präsenten *pap*-DNA-Konstrukts zur zellfreien Biosynthese von deGFP und mCherry. Folglich band die *E.-coli*-RNA-Polymerase jeweils auf den -35- und -10-Elementen, die die Lrp-Bindestellen 3 bzw. 4 flankieren, sodass die beiden entsprechenden Gene divergent transkribiert wurden. Die Transkription beider Gene bildete die notwendige Voraussetzung dafür, dass eine Analyse der Wechselwirkung zwischen der *E.-coli*-RNA-Polymerase und den beiden Promotoren auf der linken bzw. rechten „Hälfte“ des synthetischen *pap*-DNA-Konstrukts in Abhängigkeit verschiedener Parameter möglich war.

Ein zentrales Ergebnis aus [317] ist, dass eine unmethylierte GATC^{dist}- und eine methylierte GATC^{prox}-Stelle notwendig für die Transkription der *pap*-Gene bzw. den Phase-An-Zustand sind. Dieses Resultat wurde mit dem hier durchgeführten Experiment bestätigt: Das Methylierungsmuster (0,1) führte zu einer deutlich erhöhten Transkriptionsrate des *degfp*-Gens, das beim synthetischen DNA-Konstrukt als Ersatz für die nativen *pap*-Gene fungiert, im Vergleich zu den durch die anderen Methylierungsmuster generierten *degfp*-Transkriptionsraten (Abb. 5.4A). Hingegen war es der Methylierungszustand (1,0), der, im Vergleich zu den anderen Methylierungsmustern, die höchste *mcherry*-Transkriptionsrate erzeugte (Abb. 5.4B). Offenbar kann die *E.-coli*-RNA-Polymerase besser auf einem der beiden Promotoren binden, wenn die direkt benachbarte GATC-Stelle methyliert

ist, aber die gegenüberliegende GATC-Stelle unmethyliert bleibt. Der Unterschied zwischen den *degfp*-Transkriptionsraten, die aufgrund des Methylierungszustands (1,1) bzw. (0,1) entstanden (Abb. 5.4A), legt eine methylierungsabhängige Wechselwirkung zwischen der linken und rechten „Hälfte“ des *pap*-regulatorischen Bereichs nahe: In beiden Fällen ist die dem Promotor benachbarte GATC-Stelle methyliert, sodass angenommen werden kann, dass die unterschiedlich hohen Transkriptionsraten aus dem unterschiedlichen Methylierungszustand der gegenüberliegenden GATC-Stelle resultieren. Diese Hypothese wird durch die unterschiedlich hohen *mcherry*-Transkriptionsraten, die die Methylierungszustände (1,1) und (1,0) erzeugen (Abb. 5.4B), unterstützt.

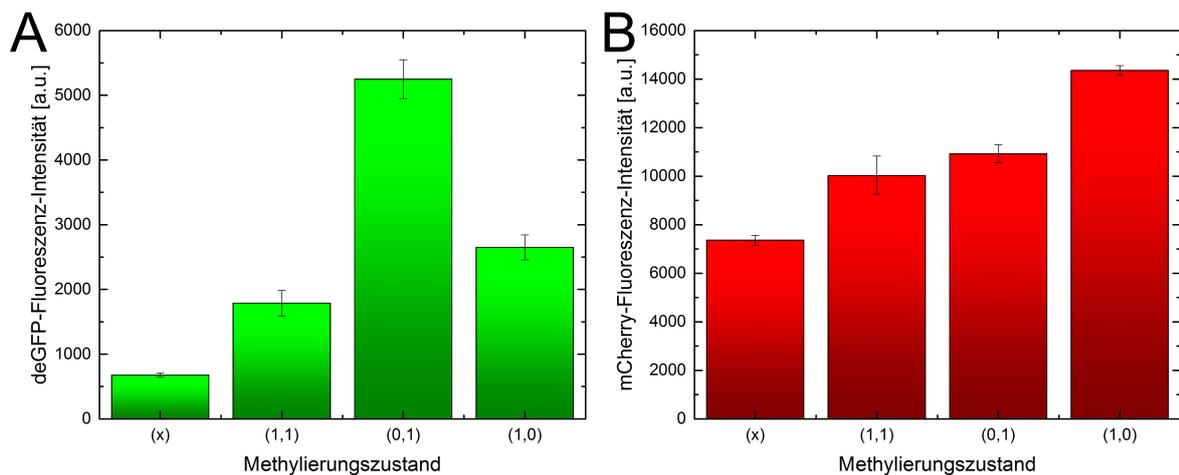


Abbildung 5.4.: deGFP- (A) und mCherry-Endpunkt-Fluoreszenz-Intensitäten (B) nach 16 h Inkubation bei 29 °C in Abhängigkeit des Methylierungszustands des hinzupipettierten *pap*-DNA-Konstrukts. Die Referenz-Reaktion x enthielt keine *pap*-DNA, den Reaktionen vom Typ (0,1), (1,0) und (1,1) wurde das *pap*-DNA-Konstrukt mit den Methylierungszuständen (0,1), (1,0) bzw. (1,1) hinzugegeben. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus fünf Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung.

Die deGFP-Fluoreszenz-Intensität der Reaktion, in der sich das *pap*-DNA-Konstrukt mit dem Methylierungszustand (1,0) befand, war durchschnittlich ungefähr 1,48-fach höher als die Intensität der Reaktion, in der das *pap*-DNA-Konstrukt mit dem Methylierungszustand (1,1) gelöst war (Abb. 5.4A). Offensichtlich führt das asymmetrische Methylierungsmuster (1,0) zu einer erhöhten *degfp*-Transkriptionsrate im Vergleich zum symmetrischen Methylierungsmuster (1,1). Es ist unklar, ob dieses Ergebnis auch auf die *mcherry*-Transkriptionsrate übertragen werden kann: Zwar liegt das mCherry-Fluoreszenz-Intensitäts-Signal der zellfreien Reaktion mit dem *pap*-DNA-Konstrukt mit dem Methylierungszustand (0,1) durchschnittlich etwa 1,09-fach über dem Signal der Reaktion, in der das *pap*-DNA-Konstrukt mit dem Methylierungszustand (1,1) präsent war, allerdings überlappen die Fehlerbalken der beiden Signale (Abb. 5.4B).

Es ist nicht bekannt, ob oder wie hoch der globale *E.-coli*-Regulator Lrp in den zellfreien Reaktionen konzentriert war. Die Dialyse während der Präparation des zellfreien Extraktes hält Proteine, deren Molekulargewicht größer als 10 kDa ist, im Zellysate zurück (Dialysenkassetten mit *molecular weight cutoffs* (MWCOs) von 10 kDa, vgl. Materialliste aus [52]). Daher sollten sich Reste des rund 19 kDa schweren Lrp im Zellextrakt befunden haben. Hingegen kann davon ausgegangen werden, dass sich kein PapI im Zellextrakt befand, da sich das für UPEC charakteristische *papI*-Gen nicht im bakteriellen Genom der zur Extrakterstellung verwendeten *E.-coli*-Zellen befindet (vgl. Abschn. 4.2.3). Zusätzlich würde vorhandenes PapI (ca. 8,7 kDa) das für die Dialyse maßgebliche Molekulargewicht von 10 kDa unterschreiten.

Angenommen, dass sich Lrp-Reste in den zellfreien Reaktionen befänden, deuten die Ergebnisse darauf hin, dass dieser Koregulator, wie u.a in [60], [306], [310]–[312], [318], [319] vorgeschlagen, in Abhängigkeit der Methylierungszustände von $\text{GATC}^{\text{prox}}$ und $\text{GATC}^{\text{dist}}$ kooperativ auf mehreren Lrp-Bindestellen gebunden hat. Diese Schlussfolgerung lässt sich daraus ableiten, dass die Methylierung der GATC-Stellen innerhalb der Lrp-Bindestellen 2 und 5 Einfluss auf die Transkription hatte (Abb. 5.4), obwohl die -35- und -10-Elemente der beiden Promotoren der jeweils benachbarten äußeren Lrp-Bindestellen flankieren. Unkooperativ bindendes Lrp ließe sich nicht durch die Methylierung der Lrp-Bindestellen 2 oder 5 davon abhalten, auf den benachbarten Lrp-Bindestellen 3 bzw. 4 zu binden. Insofern bietet das beispielsweise in [306], [313], [320] diskutierte kooperative Binden von Lrp auf den Lrp-Bindestellen 1–3 bzw. 4–6 ebenfalls in dieser Studie eine Möglichkeit, den Einfluss der Methylierung von $\text{GATC}^{\text{prox}}$ und $\text{GATC}^{\text{dist}}$ auf die Lrp-Bindung zu erklären, die die Genexpressionsrate durch wie das in [324] vorgeschlagene sterische Blockieren der RNA-Polymerase mindert.

[313] zeigte, dass die Präsenz von PapI die Bindungsaffinität von Lrp gegenüber den unmethylierten Lrp-Bindestellen 2 und 5 erhöht. In dam^- -Zellen führte die Mutation der nativen Sequenz $\text{CGATC}^{\text{prox}}$ zu $\text{TGATC}^{\text{prox}}$, die ein PapI-abhängiges Binden von Lrp auf Lrp-Bindestelle 2 verhinderte, zum Phase-An-Zustand [313]. Im Gegensatz dazu führte die native Sequenz $\text{CGATC}^{\text{prox}}$, die das PapI-abhängige Binden von Lrp auf Lrp-Bindestelle 2 ermöglichte, in der dam^- -*in-vivo*-Umgebung zum Phase-Aus-Zustand [313]. [313] schlussfolgerte, dass die Methylierung von $\text{GATC}^{\text{prox}}$ dem Zweck dient, das PapI-abhängige Binden von Lrp auf Lrp-Bindestelle 2 zu verhindern. Des Weiteren zeigte sich in Abwesenheit von PapI, dass die Methylierung von $\text{GATC}^{\text{prox}}$ keinen Einfluss auf das kooperative Binden von Lrp auf den Lrp-Bindestellen 1–3 hat, während sich im Gegensatz dazu die Bindungsaffinität von Lrp gegenüber den Lrp-Bindestellen 4–6 durch die Methylierung von $\text{GATC}^{\text{dist}}$ deutlich reduzieren ließ [313].

Unter der vorher gemachten Annahme, dass Lrp, aber kein PapI in den zellfreien Reaktionen vorhanden war, lässt sich das kooperative Binden von Lrp auf den Lrp-Bindestellen 1–3 durch die Methylierung von $\text{GATC}^{\text{prox}}$ auch ohne den Einfluss von PapI verhindern, sofern $\text{GATC}^{\text{dist}}$ unmethyliert ist. Dies zeigt sich durch die erhöhte *degfp*-Transkriptionsrate, die der Methylierungszustand (0,1) verursacht, im Vergleich zur durch das Methylierungsmuster (1,0) erzeugten Transkriptionsrate des *degfp*-Gens (Abb. 5.4A). Dieses von [306] abweichende Ergebnis resultiert möglicherweise aus den unterschiedlichen experimentellen Bedingungen: Beispielsweise wurde bei den *in-vitro*-Bindungsexperimenten von [306] ein *pap*-DNA-Fragment verwendet, das nur die Lrp-Bindestellen 1–3 umfasst. Die erhöhte Transkriptionsrate des *mcherry*-Gens beim Methylierungszustand (1,0) im Vergleich zum Methylierungszustand (0,1) zeigt, dass sich das kooperative Binden von Lrp ohne den Einfluss von PapI auch auf den gegenüberliegenden Lrp-Bindestellen 4–6 reduzieren lässt, sofern auch hier die gegenüberliegende $\text{GATC}^{\text{prox}}$ -Stelle unmethyliert ist (Abb. 5.4B). Dieses Resultat ist wiederum im Einklang mit [313]. Im Rahmen der Betrachtungsweise, dass Lrp im zellfreien Expressionssystem vorhanden war, stellt sich die Frage, warum sich die *degfp*- bzw. *mcherry*-Transkriptionsraten, die durch die Methylierungsmuster (1,1) und (0,1) bzw. (1,1) und (1,0) erzeugt wurden, unterscheiden, obwohl bei den verglichenen beiden Muster-Paaren jeweils eine methylierte $\text{GATC}^{\text{prox}}$ bzw. $\text{GATC}^{\text{dist}}$ vorlag. Auch unter dieser Perspektive scheint eine methylierungsabhängige Wechselwirkung zwischen der linken und rechten „Hälfte“ des *pap*-regulatorischen Bereichs zentral zu sein. Diese Hypothese stellt auch einen Ansatz dar, die Diskrepanz zu den gerade diskutierten Ergebnissen aus [313] zu erklären, da dort nur jeweils eine „Hälfte“ des *pap*-regulatorischen Bereichs ohne den Einfluss der anderen untersucht wurde.

Das Methylierungsmuster und die Wechselwirkung der beiden „Hälften“ des *pap*-regulatorischen Bereichs scheinen demnach zentral für den Phasenvariationsmechanismus zu sein. Dass die reine To-

pologie der DNA eine Rolle beim Phasenvariationsmechanismus spielt, schlug bereits [306] vor und dass der Phasenvariationsmechanismus nicht ausschließlich durch methylierungsabhängiges Blocken oder Binden von Lrp auf der *pap*-DNA und dem damit verbundenen Binden oder Blocken der RNA Polymerase beruhen kann, diskutierte bereits [60]. Ob sich die methylierungsabhängige Wechselwirkung zwischen der linken und rechten „Hälfte“ der *pap*-regulatorischen DNA auch bei zusätzlicher und gezielter zellfreier Synthese von Lrp und PapI weiterhin beobachten lässt, wird in Abschn. 5.3.2 untersucht.

Für dieses Experiment wurden neben den Endpunkt-Fluoreszenz-Messungen ergänzende Kinetik-Messungen durchgeführt, deren Ergebnisse in Anh. A.3.2 gezeigt werden. Sie untermauern die Ergebnisse der Endpunkt-Messungen, wenngleich die Fehlerbalken größer sind. Welche Faktoren bei den Kinetik-Messungen zu einer Vergrößerung der Standardabweichung unter diesen Versuchsbedingungen geführt haben, konnte nicht geklärt werden.

5.3.2. Epigenetische Regulation des *pap*-regulatorischen Bereichs bei zusätzlicher Lrp- und PapI-Synthese

Im Folgenden wurde die epigenetische Regulation des *pap*-regulatorischen Bereichs bei gleichzeitiger zellfreier Synthese der Koregulatoren Lrp und PapI analysiert.

Für jedes der drei Experimente, deren Ergebnisse in den Abschn. 5.3.2.1–5.3.2.3 präsentiert und diskutiert werden, wurden jeweils neun zellfreie Reaktionen präpariert. In jedem dieser aus neun Reaktionen bestehenden drei Sätze wurden die Konzentrationen der beiden Plasmide pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500, die als Templates zur zellfreien Expression der Koregulatoren-Gene *lrp* und *papI* dienten, zwischen 0 nM, 0,1 nM und 1 nM auf die gleiche Weise variiert, sodass die folgenden Plasmidkonzentrationen als Paare resultierten: (0 nM; 0 nM), (0 nM; 0,1 nM), (0 nM; 1 nM), (0,1 nM; 1 nM) ..., (1 nM; 1 nM). In jedem der drei Sätze wurde das *pap*-DNA-Konstrukt mit jeweils einem anderen Methylierungszustand in einer Konzentration von 5 nM gelöst. Alle zellfreien Reaktionen wurden 16 h bei einer Temperatur von 29 °C inkubiert, im Anschluss folgten Endpunkt-Messungen der deGFP- und mCherry-Fluoreszenz-Intensität. In den Abb. 5.5A–5.10A werden die entsprechenden deGFP- und mCherry Endpunkt-Fluoreszenz-Intensitäten präsentiert. Um die aus den drei Experimenten resultierenden Fluoreszenz-Intensitäten besser miteinander vergleichen zu können, wurden diese Intensitäten jeweils auf den höchsten deGFP- bzw. mCherry-Fluoreszenz-Intensitäts-Messwert normiert (Abb. 5.5B–5.10B).

5.3.2.1. Methylierungszustand (1,1)

Die Abb. 5.5 und Abb. 5.6 zeigen deGFP-Endpunkt- bzw. mCherry-Endpunkt-Fluoreszenz-Intensitäten der zellfreien Reaktionen, in denen das *pap*-DNA-Konstrukt mit dem Methylierungszustand (1,1) vorhanden war.

[314] beobachtete in *Gel Shift*-Experimenten, dass PapI nicht allein spezifisch an die *pap*-DNA binden kann. Unter dieser Voraussetzung deutet der Abfall der *mcherry*-Transkriptionsrate bei steigender Plasmidkonzentration von pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500, folglich wachsender PapI-Konzentration, und bei fehlendem Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500, darauf hin (Abb. 5.6), dass sich Lrp-Reste in den zellfreien Reaktionen befanden, deren Bindungsaffinität

gegenüber der Bindestelle 5 durch den Einfluss von PapI trotz Methylierung von $GATC^{dist}$ erhöht und weswegen die *mcherry*-Transkription stärker reprimiert wurde.

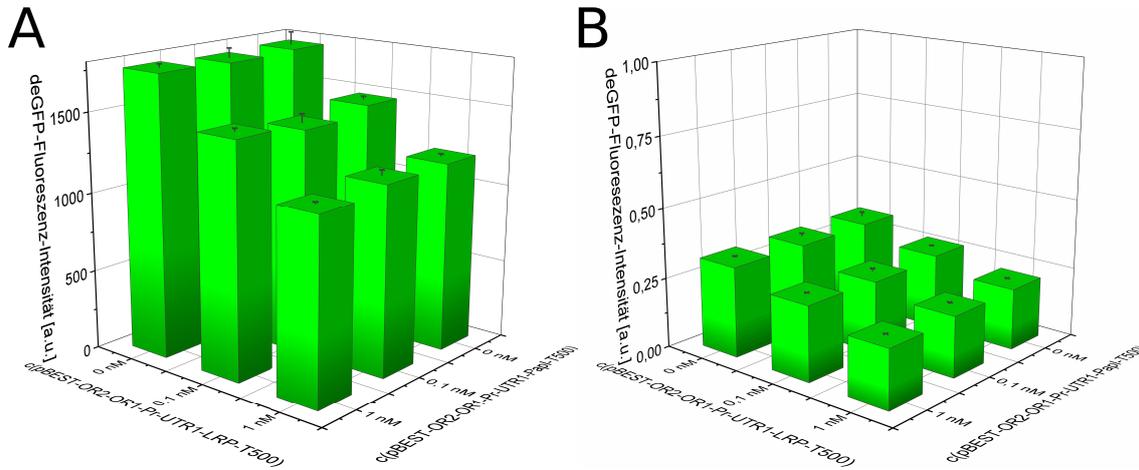


Abbildung 5.5.: deGFP- (A) und normierte deGFP-Endpunkt-Fluoreszenz-Intensitäten (B) der zellfreien Reaktionen nach 16 h Inkubation bei 29 °C in Abhängigkeit der Plasmidkonzentrationen pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500. Die Konzentration des *pap*-DNA-Konstrukts mit dem Methylierungszustand (1,1) betrug in allen zellfreien Reaktionen 5 nM. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus zwei Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung. Die Normierung erfolgte auf den höchsten Messwert der deGFP-Fluoreszenz-Intensitäts-Messungen, die jeweils in (A) in den Abb. 5.5, 5.7 und 5.9 gezeigt werden.

Die angenommene Restkonzentration von Lrp scheint zu gering gewesen zu sein, um in sättigender Weise auf der *pap*-DNA zu binden, da sich durch steigende *lrp*-Plasmidkonzentrationen und demnach wachsenden Lrp-Konzentrationen die *degfp*- und *mcherry*-Transkriptionsraten weiter senken ließen (Abb. 5.5 und 5.6). Eine Möglichkeit, diese Beobachtung zu erklären, bietet das von [324] vorgeschlagene Binden von Lrp auf der *pap*-DNA, was die Bindung der RNA-Polymerase sterisch blockiert.

Es zeigte sich, wie auch in [60], [306] zu beobachten, dass die Methylierung der $GATC^{prox}$ - und $GATC^{dist}$ -Stellen Lrp nicht vollständig blockte, sondern höhere Lrp-Konzentrationen dazu führten, dass Lrp dennoch auf kooperative Weise auf den Bindestellen 1–3 bzw. 4–6 band (Abb. 5.5 und 5.6).

[313] beobachtete in *in-vitro*-Bindungsexperimenten, dass die Methylierung von $GATC^{prox}$ ohne die Präsenz von PapI keinen Einfluss auf das kooperative Binden von Lrp auf den Lrp-Bindestellen 1–3 hat, und schlussfolgerte, basierend auf weiteren Ergebnissen, dass die Methylierung von $GATC^{prox}$ dem Verhindern des PapI-abhängigen Bindens von Lrp auf Lrp-Bindestelle 2 dient. Unter der Annahme der kooperativen Lrp-Bindung konnte ein solches PapI-abhängiges Blocken der Lrp-Bindung durch die methylierte $GATC^{prox}$ -Stelle nicht beobachtet werden, da sich die Bindungsaffinität von Lrp gegenüber der Lrp-Bindestelle 2 mit zunehmender Konzentration des zellfrei synthetisierten PapI-Proteins nicht veränderte, sich jedoch mit zunehmender Konzentration des zellfrei synthetisierten Lrp-Moleküls auch in Abwesenheit von PapI erhöhte (Abb. 5.5).

Hingegen ließ sich die Bindungsaffinität von Lrp gegenüber der Bindestelle 5 mit zunehmender Konzentration des zellfrei synthetisierten PapI-Proteins schwach erhöhen (Abb. 5.6). Im Einklang mit diesem Ergebnis, zeigten die *in-vitro*-Bindungsexperimente von [313] ebenfalls, dass eine methylierte $GATC^{dist}$ -Stelle den PapI-Lrp-Komplex blocken kann.

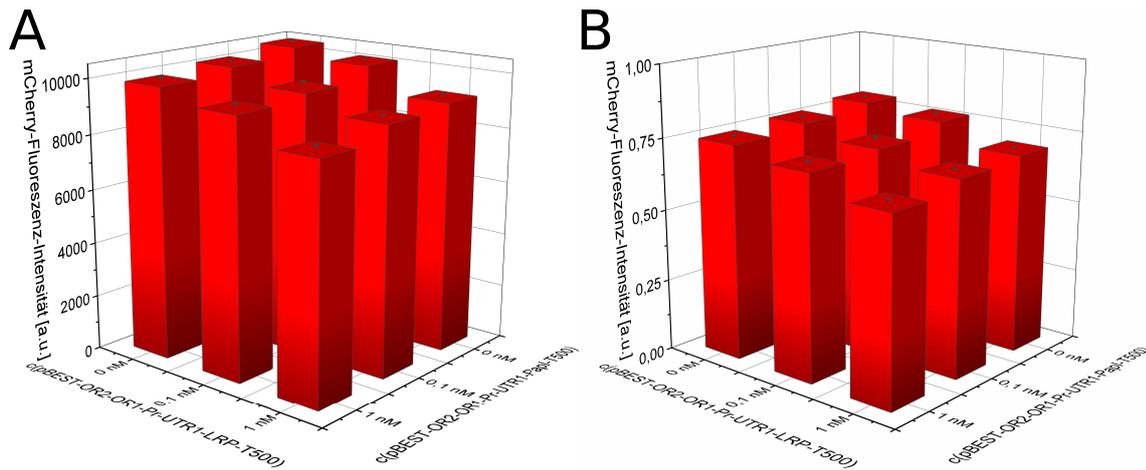


Abbildung 5.6.: mCherry- (A) und normierte mCherry-Endpunkt-Fluoreszenz-Intensitäten (B) der zellfreien Reaktionen nach 16 h Inkubation bei 29 °C in Abhängigkeit der Plasmidkonzentrationen pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500. Die Konzentration des *pap*-DNA-Konstrukts mit dem Methylierungszustand (1,1) war in allen neun zellfreien Reaktionen 5 nM. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus zwei Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung. Die Normierung erfolgte auf den höchsten Messwert der deGFP-Fluoreszenz-Intensitäts-Messungen, die jeweils in (A) in den Abb. 5.6, 5.8 und 5.10 gezeigt werden.

5.3.2.2. Methylierungszustand (0,1)

Die Abb. 5.7 und Abb. 5.8 zeigen deGFP-Endpunkt- bzw. mCherry-Endpunkt-Fluoreszenz-Intensitäten der zellfreien Reaktionen, in denen das *pap*-DNA-Konstrukt mit dem Methylierungszustand (0,1) vorhanden war.

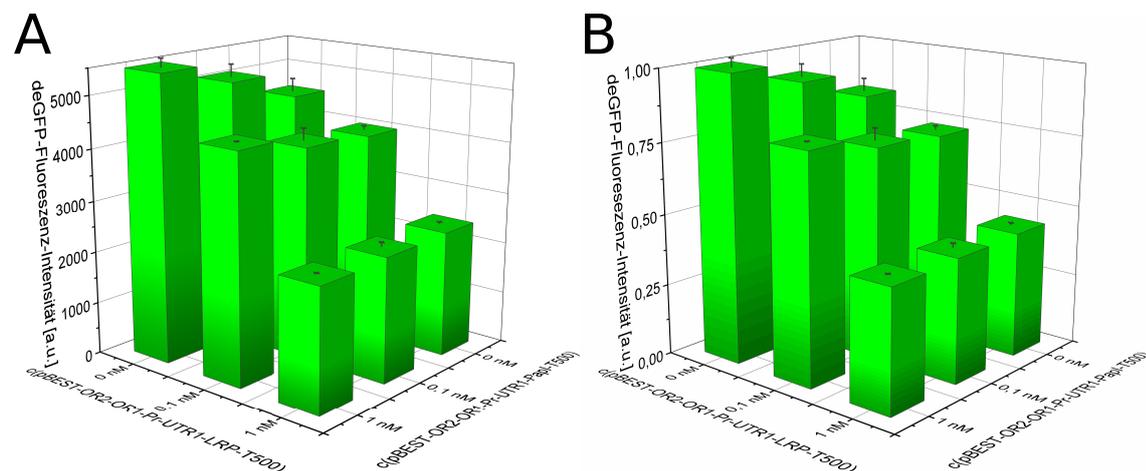


Abbildung 5.7.: deGFP- (A) und normierte deGFP-Endpunkt-Fluoreszenz-Intensitäten (B) der zellfreien Reaktionen nach 16 h Inkubation bei 29 °C in Abhängigkeit der Plasmidkonzentrationen pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500. Die Konzentration des *pap*-DNA-Konstrukts mit dem Methylierungszustand (0,1) betrug in allen zellfreien Reaktionen 5 nM. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus zwei Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung. Die Normierung erfolgte auf den höchsten Messwert der deGFP-Fluoreszenz-Intensitäts-Messungen, die jeweils in (A) in den Abb. 5.5, 5.7 und 5.9 gezeigt werden.

Wie in Abschn. 5.3.2.1 bereits erläutert, fand [313] heraus, dass die Methylierung von $GATC^{prox}$ in Abwesenheit von PapI keinen Einfluss auf das kooperative Binden von Lrp auf den Lrp-Bindestellen 1–3 hat, und schlussfolgerte, dass die Methylierung von $GATC^{prox}$ das PapI-abhängige Binden von Lrp auf Lrp-Bindestelle 2 blockt. Im Gegensatz dazu und im Einklang mit den Ergebnissen aus Abschn. 5.3.1 reichte erneut nur die Methylierung von $GATC^{prox}$ aus, um das kooperative Binden von Lrp auf den Lrp-Bindestellen 1–3 zu blocken, sofern $GATC^{dist}$ unmethyliert war (Abb. 5.7). Dies erschließt sich aus der Beobachtung, dass auch in Abwesenheit von PapI das Methylierungsmuster (0,1) die höchste *degfp*-Transkriptionsrate im Vergleich zu den *degfp*-Transkriptionsraten, die durch die anderen Methylierungszustände generiert wurden, erzeugte (vgl. Abb. 5.5B, 5.7B und 5.9B). Jedoch scheint der PapI-Lrp-Komplex bei geringeren Lrp-Konzentrationen stärker von der Methylierung der $GATC^{prox}$ -Stelle geblockt zu werden als das reine Lrp-Molekül, da sich die *degfp*-Transkriptionsrate mit zunehmender Plasmidkonzentration von pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500 und somit größer werdender Konzentration von PapI weiter steigern ließ (Abb. 5.7). In diesem Rahmen konnte die in [313] beobachtete PapI-abhängige Sensitivität von Lrp gegenüber der methylierten $GATC^{prox}$ -Stelle ebenfalls gesehen werden.

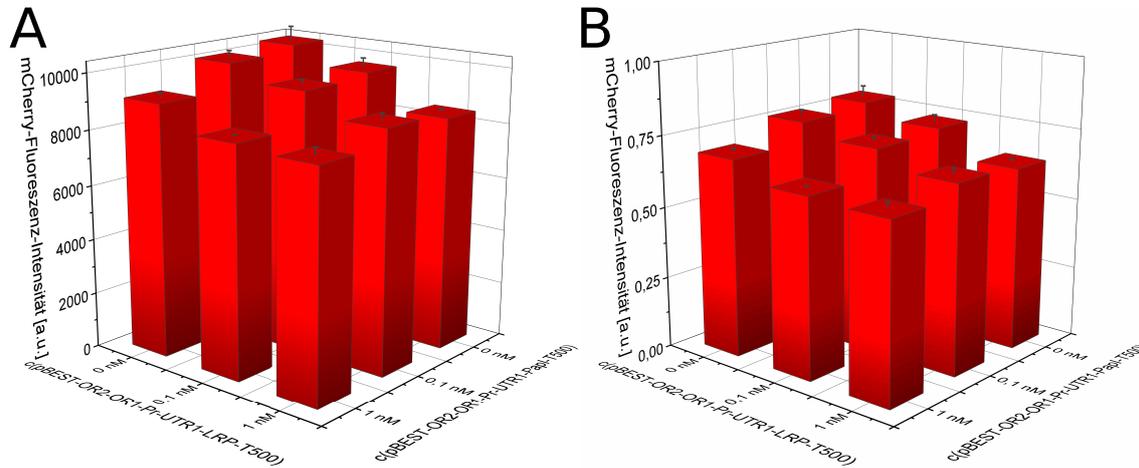


Abbildung 5.8.: mCherry- (A) und normierte mCherry-Endpunkt-Fluoreszenz-Intensitäten (B) der zellfreien Reaktionen nach 16 h Inkubation bei 29 °C in Abhängigkeit der Plasmidkonzentrationen pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500. Die Konzentration des *pap*-DNA-Konstrukts mit dem Methylierungszustand (0,1) war in allen neun zellfreien Reaktionen 5 nM. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus zwei Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung. Die Normierung erfolgte auf den höchsten Messwert der deGFP-Fluoreszenz-Intensitäts-Messungen, die jeweils in (A) in den Abb. 5.6, 5.8 und 5.10 gezeigt werden.

Jedoch, ebenfalls in Übereinstimmung mit den Ergebnissen aus Abschn. 5.3.1, legt ein Vergleich der *degfp*-Transkriptionsraten, die durch die Methylierungsmuster (1,1) und (0,1) entstanden (vgl. Abb. 5.5 und 5.7), die Vermutung einer methylierungsabhängigen Wechselwirkung zwischen der linken und rechten „Hälfte“ der *pap*-regulatorischen DNA nahe, die ein zentraler Mechanismus der Phasenvariation sein könnte. Erneut sind bei beiden Methylierungsmustern der betrachteten *pap*-DNA-Konstrukte die dem *degfp*-Promotor benachbarten $GATC^{prox}$ -Stellen methyliert, sodass angenommen werden kann, dass die unterschiedlich hohen Transkriptionsraten aus dem unterschiedlichen Methylierungszustand der gegenüberliegenden $GATC^{dist}$ -Stelle resultierten. Das Methylierungsmuster (0,1) lieferte selbst bei maximaler Lrp- und PapI-bedingter Repression (tiefste deGFP-Fluoreszenz-Intensität) eine *degfp*-Transkriptionsrate, die im gleichen Größenbereich wie die durch das Methylierungsmuster (1,0) verursachte maximale *degfp*-Transkriptionsrate (höchste deGFP-Fluoreszenz-Intensität) lag

(vgl. Abb. 5.7B und 5.9B). Sie war ca. 1,35-fach höher als die durch das Methylierungsmuster (1,1) generierte, maximale *degfp*-Transkriptionsrate (vgl. Abb. 5.5B und 5.7B). Durchschnittlich brachte das Methylierungsmuster (0,1) eine *degfp*-Expressionsrate hervor, die ca. 2,3-fach und 2,6-fach über der *degfp*-Transkriptionsrate, die vom Methylierungsmuster (1,0) bzw. (1,1) verursacht wurde, lag (vgl. Abb. 5.5B, 5.7B und 5.9B). Insofern scheint das Methylierungsmuster und die damit verbundene Wechselwirkung zwischen der linken und rechten „Hälfte“ des *pap*-regulatorischen Bereichs entscheidend für die Transkription des *degfp*-Gens bzw. der *pap*-Gene beim nativen DNA-Konstrukt zu sein. Die Ergebnisse legen eine zusätzliche Variation der Transkriptionsrate durch die Koregulatoren nahe.

Die Vermutung, dass eine methylierungsabhängige Wechselwirkung zwischen der linken und rechten „Hälfte“ der *pap*-regulatorischen DNA ein zentraler Mechanismus der Phasenvariation sein könnte, wird zusätzlich dadurch untermauert, dass beim Methylierungszustand (0,1) die von [313] beobachtete PapI-abhängige Sensitivität von Lrp gegenüber der methylierten GATC^{prox}-Stelle gesehen werden konnte, während dies beim Methylierungszustand (1,1) nicht möglich war (vgl. Abb. 5.5 und 5.7). In beiden Fällen ist die GATC^{prox}-Stelle methyliert, der Unterschied liegt demnach nur im Methylierungszustand der gegenüberliegenden GATC^{dist}-Stelle.

Ein Aspekt des von [313] vorgeschlagenen Ansatzes zur Erklärung des Phasenvariationsmechanismus ist, dass die methylierte GATC^{prox}-Stelle das PapI-abhängige Binden des Lrp-PapI-Komplexes blockt, und dass der Komplex bei unmethylierter GATC^{dist}-Stelle kooperativ auf den gegenüberliegenden Lrp-Bindestellen 4–6 binden kann. Eine methylierte GATC^{dist}-Stelle blockt jedoch ebenfalls den PapI-Lrp-Komplex [313]. Somit wäre zu erwarten gewesen, dass die durch das Methylierungsmuster (0,1) verursachte *mcherry*-Transkriptionsrate generell niedriger ist, als die, die durch den Methylierungszustand (1,1) erzeugt wird, da beim Methylierungszustand (1,1) die methylierte GATC^{dist}-Stelle den Koregulator-Komplex blocken sollte. Tatsächlich war die *mcherry*-Transkriptionsrate beim Methylierungsmuster (1,1) durchschnittlich 1,05-fach höher als die durch den Methylierungszustand (0,1) erzeugte *mcherry*-Transkriptionsrate (vgl. Abb. 5.6B und 5.8B).

Wie bereits in Abschn. 5.3.2.1 gezeigt und diskutiert, blockte auch hier die Methylierung der GATC^{prox}- und GATC^{dist}-Stelle Lrp nicht vollständig, sondern höhere Lrp-Konzentrationen führten dazu, dass Lrp dennoch kooperativ auf den Bindestellen 1–3 bzw. 4–6 band. Unter der Bedingung, dass PapI nicht allein auf der *pap*-DNA binden kann [314], gemeinsam mit der Beobachtung, dass sich auch hier ohne die zusätzliche zellfreie Synthese von Lrp die *degfp*- und *mcherry*-Transkriptionsrate bei Variation der pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500-Konzentration änderte (Abb. 5.7 und 5.8), deutet darauf hin, dass Lrp-Reste im zellfreien System vorhanden waren.

5.3.2.3. Methylierungszustand (1,0)

Die Abb. 5.9 und Abb. 5.10 zeigen deGFP-Endpunkt- bzw. mCherry-Endpunkt-Fluoreszenz-Intensitäten der zellfreien Reaktionen, in denen das *pap*-DNA-Konstrukt mit dem Methylierungszustand (1,0) präsent war.

Im Einklang mit den Ergebnissen aus Abschn. 5.3.1 reichte bereits die Methylierung von GATC^{dist} aus, um das kooperative Binden von Lrp auf den Lrp-Bindestellen 4–6 zu blocken, sofern GATC^{prox} unmethyliert war, da das Methylierungsmuster (1,0) in Abwesenheit von PapI die höchste *mcherry*-Transkriptionsrate im Vergleich zu den anderen Methylierungsmustern hervorbrachte (vgl. Abb. 5.6B,

5.8B und 5.10B): In den zellfreien Reaktionen, in denen das Plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500 fehlte, lag die vom Methylierungszustand (1,0) erzeugte *mcherry*-Transkriptionsrate durchschnittlich 1,26-fach über der durchschnittlichen *mcherry*-Transkriptionsrate, die durch die Methylierungsmuster (1,1) oder (0,1) generiert wurden. Im Gegensatz zu der durch das Methylierungsmuster (0,1) erzeugten *degfp*-Transkriptionsrate ließ sich die *mcherry*-Transkriptionsrate hier nicht mehr weiter mit zunehmender PapI-Konzentration steigern (vgl. Abb. 5.10 und 5.7).

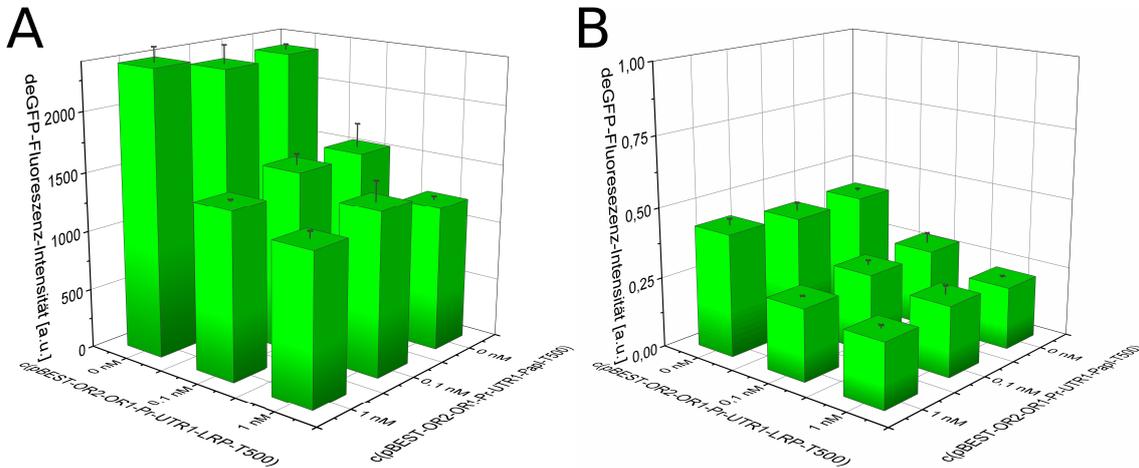


Abbildung 5.9.: deGFP- (A) und normierte deGFP-Endpunkt-Fluoreszenz-Intensitäten (B) der zellfreien Reaktionen nach 16 h Inkubation bei 29 °C in Abhängigkeit der Plasmidkonzentrationen pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500. Die Konzentration des *pap*-DNA-Konstrukts mit dem Methylierungszustand (1,0) betrug in allen zellfreien Reaktionen 5 nM. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus zwei Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung. Die Normierung erfolgte auf den höchsten Messwert der deGFP-Fluoreszenz-Intensitäts-Messungen, die jeweils in (A) in den Abb. 5.5, 5.7 und 5.9 gezeigt werden.

Jedoch, ebenfalls in Übereinstimmung mit den Ergebnissen aus Abschn. 5.3.1 und 5.3.2.2, legt ein Vergleich der *mcherry*-Transkriptionsraten, die durch das Methylierungsmuster (1,1) bzw. (1,0) erzeugt wurden, die Vermutung einer methylierungsabhängigen Wechselwirkung zwischen der linken und rechten „Hälfte“ der *pap*-regulatorischen DNA nahe und untermauert, dass dies ein zentraler Mechanismus der Phasenvariation sein könnte. Auch hier war bei beiden Methylierungszuständen des *pap*-DNA-Konstrukts die dem mCherry-Promotor benachbarte GATC^{dist}-Stelle methyliert, sodass angenommen werden kann, dass die verschiedenen hohen *mcherry*-Transkriptionsraten aus dem unterschiedlichen Methylierungszustand der gegenüberliegenden GATC^{prox}-Stelle resultierten. Selbst unter dem maximal reprimierenden Einfluss von Lrp und PapI erzeugte das Methylierungsmuster (1,0) eine *mcherry*-Transkriptionsrate (tiefste mCherry-Fluoreszenz-Intensität), die nur etwas kleiner (95 % bzw. 94 %) als die durch die Methylierungsmuster (0,1) und (1,1) verursachten maximalen *mcherry*-Transkriptionsraten (höchste mCherry-Fluoreszenz-Intensitäten) war (vgl. Abb. 5.6B, 5.8B und 5.10B). Durchschnittlich lag die *mcherry*-Transkriptionsrate beim Methylierungsmuster (1,0) ca. 1,3-fach über der des Methylierungsmusters (0,1) und 1,25-fach über der des Methylierungszustandes (1,1) (vgl. Abb. 5.6B, 5.8B und 5.10B).

Die durch das Methylierungsmuster (1,0) hervorgebrachte *degfp*-Transkriptionsrate lag durchschnittlich 1,15-fach höher als die *degfp*-Transkriptionsrate, die durch den Methylierungszustand (1,1) erzeugt wurde (vgl. Abb. 5.5B und 5.9B). In schwächerer Weise als bei den in Abschn. 5.3.1 präsentierten Ergebnissen, führte demnach das asymmetrische Methylierungsmuster (1,0) auch in Anwesenheit der

zellfrei synthetisierten Koregulatoren Lrp und PapI zu einer schwach erhöhten *degfp*-Transkriptionsrate im Vergleich zum symmetrischen Methylierungsmuster (1,1). Jedoch verwundert dieses Ergebnis: Aufgrund des PapI-abhängigen Blockens des Lrp-PapI-Komplexes durch die methylierte GATC^{prox}-Stelle [313] war zu erwarten, dass das Methylierungsmuster (1,1) zu höheren deGFP-Fluoreszenz-Intensitäten als das Methylierungsmuster (1,0) führen würde. Auch diese Beobachtung kann auf eine methylierungsabhängige Wechselwirkung zurückzuführen sein.

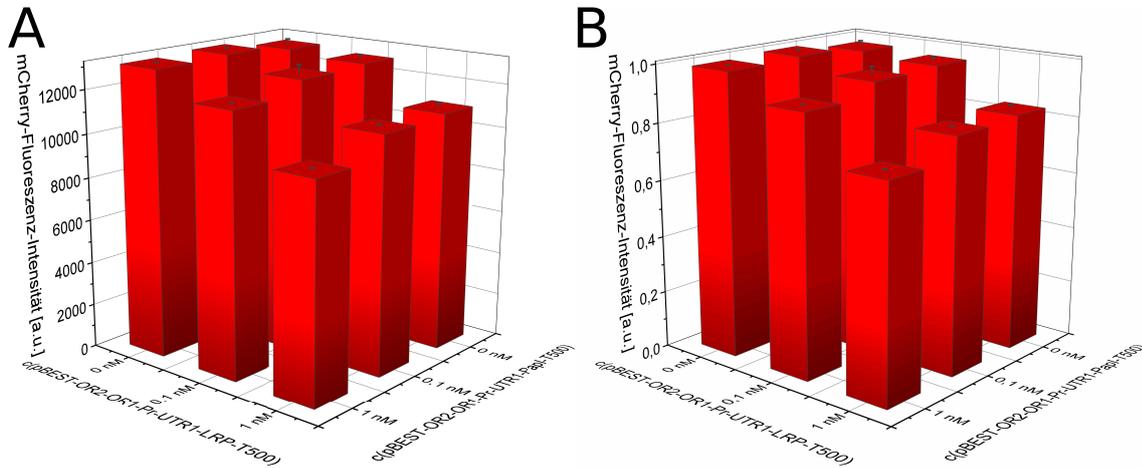


Abbildung 5.10.: mCherry- (A) und normierte mCherry-Endpunkt-Fluoreszenz-Intensitäten (B) der zellfreien Reaktionen nach 16 h Inkubation bei 29 °C in Abhängigkeit der Plasmidkonzentrationen pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 und pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500. Die Konzentration des *pap*-DNA-Konstrukts mit dem Methylierungszustand (1,0) war in allen neun zellfreien Reaktionen 5 nM. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus zwei Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung. Die Normierung erfolgte auf den höchsten Messwert der deGFP-Fluoreszenz-Intensitäts-Messungen, die jeweils in (A) in den Abb. 5.6, 5.8 und 5.10 gezeigt werden.

Wie bereits mehrfach zuvor gezeigt und diskutiert (vgl. Abschn. 5.3.2.1 und 5.3.2.2), blockte die Methylierung der GATC^{prox}- und GATC^{dist}-Stellen Lrp nicht vollständig, sondern höhere Lrp-Konzentrationen führten dazu, dass Lrp dennoch kooperativ auf den Bindestellen 1–3 bzw. 4–6 band.

5.3.3. Untersuchung der DNA-Wechselwirkung des *pap*-regulatorischen Bereichs

Die in Abschn. 5.3.1 und 5.3.2 präsentierten und diskutierten Ergebnisse legen eine methylierungsabhängige Wechselwirkung zwischen den beiden „Hälften“ des *pap*-regulatorischen Bereichs nahe. Mit Hilfe des *mfold web servers* [334] (Standardeinstellungen, 29 °C, vgl. Abschn. 5.2.4) wurde die Sekundärstruktur der *pap*-regulatorischen DNA vorhergesagt und somit deren Wechselwirkung mit sich selbst untersucht. Die zugrundeliegende DNA-Sequenz entsprach der Sequenz des synthetischen *pap*-regulatorischen Bereichs inklusive der stromaufwärts und -abwärts liegenden UTR1-Sequenzen und den jeweiligen Anfängen des *degfp*- und *mcherry*-Gens (Abschn. 5.2.4).

Unter den genannten Bedingungen wurden drei ähnliche Sekundärstrukturen vorhergesagt, von denen eine exemplarisch in Abb. 5.11 gezeigt wird. Die beiden anderen Sekundärstrukturen sind in Anh. A.3.3 zu finden. In jeder dieser drei Strukturen kommt es zumindest zur teilweisen Basenpaarung zwischen

den Nukleotiden der Lrp-Bindestelle 2 und 5 sowie 3 und 4; bei allen drei Strukturen ist eine vollständige Paarung zwischen $\text{GATC}^{\text{prox}}$ und $\text{GATC}^{\text{dist}}$ zu beobachten (Abb. 5.11, A.16 und A.17). Bei der Sekundärstruktur aus Abb. 5.11 sind die Nukleotide der Lrp-Bindestellen 2 und 5 sogar vollständig gepaart. Der betrachtete *pap*-DNA-Bereich faltet sich über weite Bereiche und wechselwirkt somit stark mit sich selbst (Abb. 5.11, A.16 und A.17). Dies bietet einen möglichen Erklärungsansatz, warum die PCR-Amplifikation der linearen *pap*-DNA-Templates nicht effizient verlief und es zur Entstehung vieler Nebenprodukte während der PCR kam (vgl. Abschn. 5.2.1.1). Im Anschluss an den Denaturierungsschritt liegt die zu amplifizierende DNA als Einzelstrang-DNA vor, die beim Abkühlen auf die Primer-Hybridisierungs-Temperatur mit sich selbst wechselwirken und Sekundärstrukturen ausbilden kann. Treten diese Sekundärstrukturen im Bereich der DNA auf, in denen es zur Primer-Hybridisierung kommen soll, kann die DNA-Faltung verhindern, dass der Primer auf der *pap*-DNA bindet. Kürzere Nebenprodukte können entstehen, wenn es zu vorzeitigen Abbrüchen der Strang-Elongation kommt. Zu solchen Abbrüchen könnte es dann kommen, wenn die DNA-Polymerase auf eine stark ausgeprägte Sekundärstruktur trifft und sich deshalb von der DNA löst.

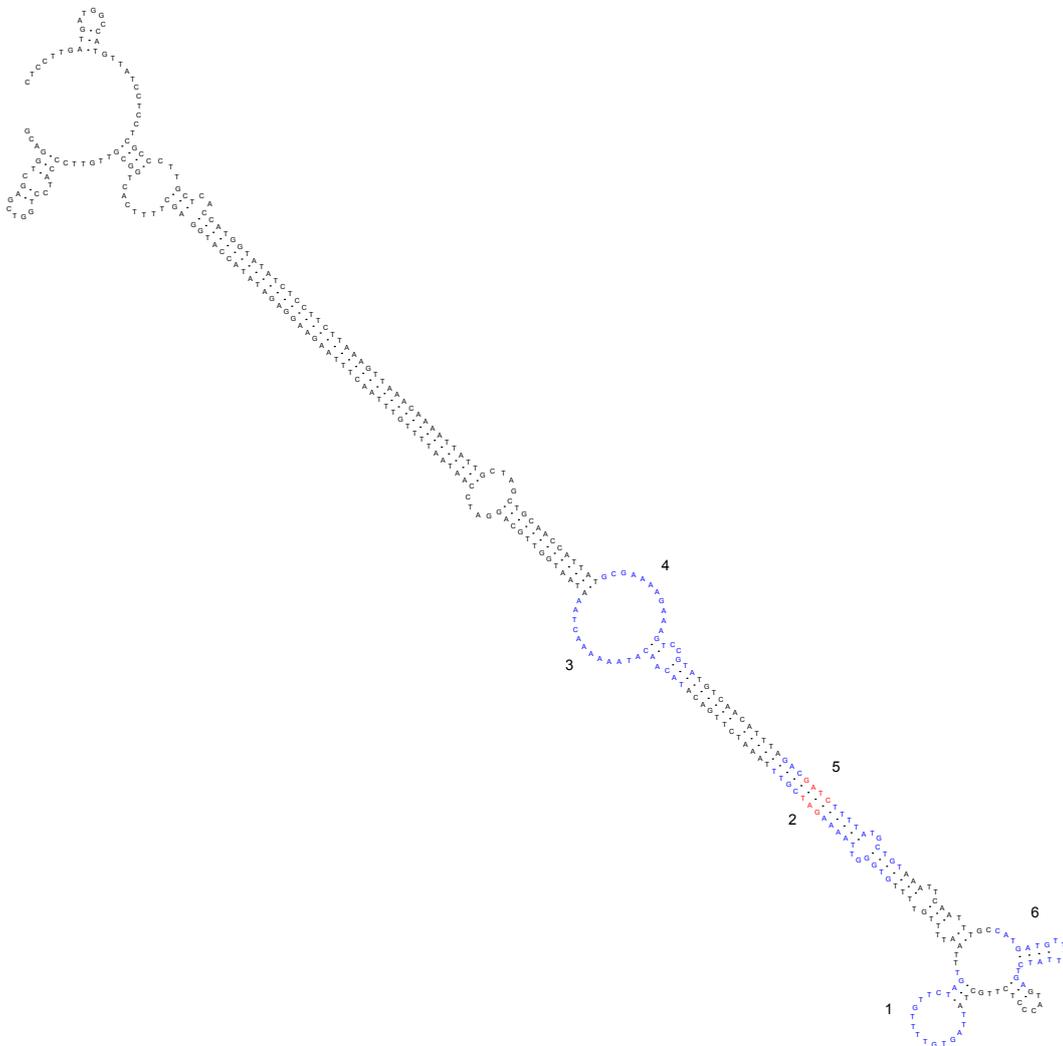


Abbildung 5.11.: Eine der drei vom *mfold web server* [334] (Standardeinstellungen, 29 °C, vgl. Abschn. 5.2.4) vorhergesagten Sekundärstrukturen der synthetischen *pap*-DNA. Die sechs Lrp-Bindestellen sind blau, $\text{GATC}^{\text{prox}}$ und $\text{GATC}^{\text{dist}}$ rot markiert. Die Abbildung wurde mit dem Programm *Rna Viz* [335] hergestellt.

Hemimethylierte GATC-Stellen besitzen verengte große Furchen und sind strukturell weniger stabil, woraus sich spezifische Wechselwirkungen zwischen Proteinen und diesen Stellen ergeben können [336]. In Abschn. 5.3.1 und 5.3.2 führten die asymmetrischen Methylierungsmuster (0,1) und (1,0) zu erhöhten Transkriptionsraten des *degfp*- bzw. *mcherry*-Gens. Falten sich die *pap*-regulatorischen Bereiche mit den Methylierungszuständen (0,1) und (1,0) wie in den Abb. 5.11, A.16 oder A.17 angegeben, sind $\text{GATC}^{\text{prox}}$ bzw. $\text{GATC}^{\text{dist}}$ unterschiedlich hemimethyliert. Es kann spekuliert werden, dass genau solche DNA-Strukturen zu einer spezifischen Wechselwirkung mit Lrp, PapI oder der *E.-coli*-RNA-Polymerase führen. Sind sowohl $\text{GATC}^{\text{prox}}$ bzw. $\text{GATC}^{\text{dist}}$ vollständig methyliert, sind diese beiden Stellen auch innerhalb der Sekundärstruktur vollständig methyliert und die angenommene Hemimethylierungs-abhängige spezifische DNA-Protein-Wechselwirkung würde beim Methylierungsmuster (1,1) nicht zum Tragen kommen. Insofern bietet diese Hypothese einen in sich stimmigen Ansatz, die unterschiedlich hohen Transkriptionsraten, die durch die unterschiedlichen Methylierungsmuster der *pap*-DNA entstehen, zu erklären. Eine mögliche DNA-Konformation, die sich aus der vorhergesagten Wechselwirkung der *pap*-DNA ergeben könnte, wird als DNA-Kreuzform bezeichnet (Abb. 5.12). Kreuzstrukturen spielen bei vielen biologischen Prozessen, wie beispielsweise der Gen-Regulation, eine entscheidende Rolle und interagieren mit vielen Proteinen, die u.a. derartige Strukturen wieder lösen können [337]. Solche Proteine können von den bei der Präparation des zellfreien Extraktes verwendeten Zellen auf das zellfreie Zellsat übertragen worden sein.

Tatsächlich wurde bereits in *in-vitro*-Bindungsexperimenten der Einfluss einer hemimethylierten $\text{GATC}^{\text{dist}}$ -Stelle auf die Bindungsaffinitäten von Lrp und PapI gegenüber den Lrp-Bindestellen 4–6 untersucht [313]. Ist die doppelsträngige $\text{GATC}^{\text{dist}}$ -Stelle nur am unteren Einzelstrang methyliert, besitzt Lrp eine schwächere Bindungsaffinität gegenüber den Lrp-Bindestellen 4–6, als wenn die vollständig methylierte $\text{GATC}^{\text{dist}}$ -Stelle als Bindungspartner zur Verfügung stände [313]. Ist hingegen der obere Einzelstrang der doppelsträngigen $\text{GATC}^{\text{dist}}$ -Stelle methyliert, nimmt die Bindungsaffinität von Lrp gegenüber den Bindestellen 4–6 im Vergleich zur vollständig methylierten $\text{GATC}^{\text{dist}}$ -Stelle zu [313]. Ist PapI zusätzlich involviert, nimmt die Affinität von Lrp, auf den Bindestellen 4–6 zu binden, bei beiden Formen der hemimethylierten $\text{GATC}^{\text{dist}}$ -Stelle im Vergleich zur vollständig methylierten $\text{GATC}^{\text{dist}}$ -Stelle zu [313]. Diese Ergebnisse von [313] belegen, dass die Hemimethylierung der $\text{GATC}^{\text{dist}}$ -Stelle Einfluss auf das Bindungsverhalten der Koregulatoren Lrp und PapI hat und dass die Art der Hemimethylierung ebenso das Bindungsverhalten der beiden Proteine beeinflusst. Auch diese Beobachtungen [313] rechtfertigen es, eine DNA-Kreuzstruktur anzunehmen, deren hemimethylierte GATC-Stellen aus gepaarten Nukleotiden der Lrp-Bindestellen 2 und 5 entstehen, sich je nach Methylierungsmuster (0,1) oder (1,0) in der Art ihrer Hemimethylierung unterscheiden und somit charakteristisch die Bindung von Lrp, PapI oder vor allem der *E.-coli*-RNA-Polymerase beeinflussen.

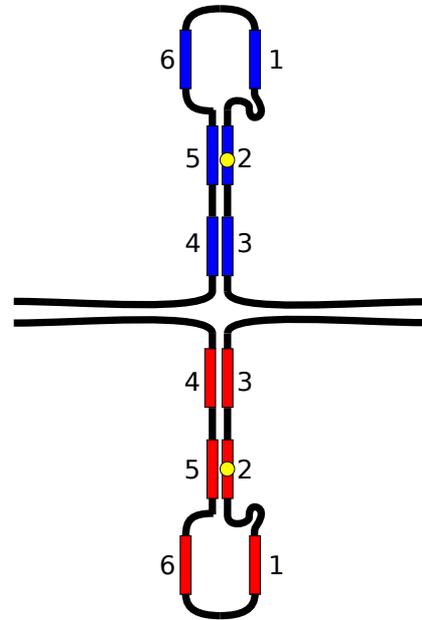


Abbildung 5.12.: Stark vereinfachte Darstellung einer möglichen Kreuzstruktur, die die *pap*-DNA mit dem Methylierungsmuster (0,1) auf Grundlage der vom *mfold web server* [334] vorhergesagten Sekundärstruktur bilden könnte. Die sechs Lrp-Bindestellen sind rot bzw. blau markiert. Beim Methylierungsmuster (1,0) liegen die Methylgruppen (gelber Kreis) jeweils auf dem gegenüberliegenden Einzelstrang.

5.4. Zusammenfassung, Konklusion und Ausblick

Die vorliegende Studie beschäftigte sich mit der Analyse des Phasenvariationsmechanismus, mit dem die Pap-Biogenese auf Transkriptions-Ebene in uropathogenen *E.-coli*-Zellen reguliert wird [59]. Sie knüpfte an Vorarbeiten [60] an, in denen bereits auf das auch in dieser Arbeit verwendete zellfreie Transkriptions-Translations-System von NOIREAUX [51], [52] zurückgegriffen wurde. Es ermöglicht eine effiziente zellfreie Genexpression mittels der nativen *E.-coli*-RNA-Polymerase und ihrem σ^{70} -Faktor [51], [52]. Insofern bietet diese *E.-coli*-ähnliche zellfreie Umgebung eine ideale Plattform, um den Phasenvariationsmechanismus in Abhängigkeit des Methylierungszustandes des *pap*-regulatorischen Bereichs und der Konzentration der Koregulatoren Lrp und PapI zu untersuchen [60]. Der Phasenvariationsmechanismus wurde vor allem in Bezug auf die Studien [60], [306] analysiert, aus denen hervorgeht, dass die reine Topologie der DNA bereits Einfluss auf den Phasenvariationsmechanismus haben könnte.

In Analogie zum Aufbau der DNA-Templates aus [60], in denen maximal drei Lrp-Bindestellen in verschiedenen Kombinationen miteinander gekoppelt wurden, erfolgte die Synthese eines *pap*-regulatorischen Bereichs, der dem nativen *pap*-regulatorischen DNA-Lokus (GenBank ID: X14471.1, zurückgehend auf [59]) in der Umgebung aller sechs Lrp-Bindestellen bis auf die für die zellfreie Expression notwendigen Modifikationen entspricht und an den sich stromaufwärts das *mcherry*- und -abwärts das *degfp*-Gen anschließen (Abb. 5.3). Die Mutation von GATC^{prox} oder GATC^{dist} zu GCTC verhindert, dass GATC^{prox} bzw. GATC^{dist} Dam-methyliert werden können, ohne dass sich die Bindungsaffinitäten von Lrp und PapI gegenüber den Lrp-Bindestellen 2 bzw. 5 verändern [317]. Auf Grundlage dieser Mutation [317] und den *in-vitro*-Dam-Methylierungs-Protokollen von [60] konnten drei Methylierungszustände des synthetischen *pap*-DNA-Konstrukts hergestellt werden. Entweder waren GATC^{prox} und GATC^{dist} methyliert (Methylierungszustand (1,1)) oder GATC^{prox} war nicht und GATC^{dist} war methyliert (Methylierungszustand (1,0)) oder GATC^{dist} war nicht und GATC^{prox} war methyliert (Methylierungszustand (0,1)). Wie in [60] wurde zusätzlich Lrp oder PapI zellfrei durch Hinzugabe eines Plasmids synthetisiert, das die Expression des *lrp*- oder *papI*-Gens ermöglichte und wodurch die PapB-abhängige *papI*-Transkription umgangen werden konnte; die *lrp*- bzw. *papI*-Transkriptionsrate und somit die Konzentration der zellfrei synthetisierten Lrp- bzw. PapI-Moleküle stieg mit zunehmender Konzentration des jeweiligen Plasmids. Gemäß der Idee, die Transkriptionsrate eines Reportergens mit der entsprechenden Fluoreszenz-Intensität in Korrelation zu setzen [51], [52], gab wie in den Vorarbeiten [60] die gemessene Fluoreszenz-Intensitäts-Höhe Aufschluss über die Menge des zellfrei synthetisierten Reporterproteins deGFP und somit über die Transkriptionsrate des *degfp*-Gens, das im *pap*-DNA-Konstrukt der vorliegenden Studie unter Kontrolle der rechten „Hälfte“ des *pap*-regulatorischen Bereichs stand. Dieses Konzept, mittels der Reportergen-Fluoreszenz-Intensität auf die Transkriptionsrate des Reportergens zu schließen, wurde durch die zusätzliche Fusion des *mcherry*-Gens an den *pap*-regulatorischen Bereich und Messung der mCherry-Fluoreszenz-Intensität um einen weiteren Kanal erweitert. Dadurch, dass alle sechs Lrp-Bindestellen in einem DNA-Konstrukt erfolgreich gekoppelt werden konnten und das *mcherry*-Gen unter Kontrolle der linken „Hälfte“ des *pap*-regulatorischen Bereichs stand, konnte die Wechselwirkung zwischen den proximalen und distalen Lrp-Bindestellen inklusive der u.a. in [306], [313], [320] postulierten kooperativen Bindung von Lrp auf den Lrp-Bindestellen 4–6 und 1–3, in An- oder Abwesenheit von PapI, analysiert werden. Da das *degfp*-Gen im synthetischen *pap*-DNA-Konstrukt Platzhalter für die *pap*-Gene ist, konnte mittels der Transkriptionsrate des *degfp*-Gens ein Rückschluss auf die Transkriptionsrate der *pap*-Gene gezogen werden. Auf die Integration der Bindestelle für das *Catabolite Activator Protein* (CAP) wurde wie zuvor [60] verzichtet. Diese Stelle benachbart die PapB-Bindestelle und ermöglicht die Bindung

von CAP, von dem vermutet wird, dass es aufgrund einer Lrp-PapI-bedingten Krümmung im *pap*-regulatorischen Bereich der DNA mit der RNA-Polymerase wechselwirkt und ebenfalls in die Regulation der Expression der *pap*-Gene involviert ist [328], [329].

Der Phasenvariationsmechanismus wurde ohne (Abschn. 5.3.1) und mit (Abschn. 5.3.2) gleichzeitiger zellfreier Synthese der Koregulatoren Lrp und PapI untersucht, indem die Transkriptionsraten der beiden Reportergene *degfp* und *mcherry* in Abhängigkeit des Methylierungsmusters der synthetischen *pap*-DNA analysiert wurden. Da die zur Herstellung des Extrakts verwendeten Zellen nicht Träger des UPEC-spezifischen *papI*-Gens sind, kann davon ausgegangen werden, dass PapI nicht im zellfreien System vorhanden war. Da PapI nicht alleine spezifisch an die *pap*-DNA binden kann [314], sich jedoch Transkriptionsraten der Reportergene bei fehlender *lrp*-Expression und mit zunehmender Konzentration des synthetisierten PapI-Proteins änderten, kann geschlossen werden, dass Reste des globalen *E.-coli*-Regulators Lrp im zellfreien System vorhanden waren, die von den Zellen, aus denen der Zellextrakt gewonnen wurde, stammten. Dass die Methylierung der Lrp-Bindestelle 2 oder 5 Einfluss auf die Transkriptionsraten der Reportergene hatte, obwohl die -35- und -10-Elemente der beiden Promotoren, die jeweils die benachbarten, äußeren Lrp-Bindestellen flankieren, untermauerte die u.a. in [60], [306], [310]–[312], [318], [319] vorgeschlagene, vom Methylierungsmuster abhängige kooperative Bindung von Lrp auf mehreren Lrp-Bindestellen. Auf unkooperativ bindendes Lrp hätte die Methylierung einer benachbarten Lrp-Bindestelle keinen direkten Einfluss gehabt, und die Methylierung hätte Lrp somit nicht daran hindern können, auf der äußeren Lrp-Bindestelle zu binden.

Insofern bietet das beispielsweise in [306], [313], [320] diskutierte kooperative Binden von Lrp auf den Lrp-Bindestellen 1–3 bzw. 4–6 ebenfalls in dieser Studie eine Möglichkeit, den Einfluss der Methylierung von $\text{GATC}^{\text{prox}}$ und $\text{GATC}^{\text{dist}}$ auf die Lrp-Bindung zu erklären: Die Genexpressionsrate wird durch, wie in [324] vorgeschlagen, sterisches Blockieren der RNA-Polymerase gemindert. Dass die Transkriptionsraten der beiden Reportergene mit zunehmender Lrp-Konzentration stets abnahmen, zeigte, dass, wie bereits zuvor berichtet [60], [306], die Methylierung der $\text{GATC}^{\text{prox}}$ - und $\text{GATC}^{\text{dist}}$ -Stellen Lrp nicht vollständig blockieren kann, sondern höhere Lrp-Konzentrationen dazu führen, dass Lrp dennoch auf kooperative Weise auf den Bindestellen 1–3 bzw. 4–6 binden kann.

Die Experimente, in denen PapI und Lrp nicht zellfrei synthetisiert wurden (Abschn. 5.3.1), zeigten, dass jedes Methylierungsmuster die Expression beider Reportergene ermöglicht, jedoch die Höhe der Expressionsrate, wie zuvor berichtet [317], stark vom Methylierungszustand des *pap*-regulatorischen Bereichs abhängt. Im Einklang mit den Beobachtungen aus [60] erfolgte die Genexpression auch ohne eine in die *pap*-DNA integrierte CAP-Bindestelle, deren Besetzung durch CAP als essentiell für die Expression der *pap*-Gene angesehen wird [328], [329]. Ob sich dies auf die Sequenzunterschiede zwischen der nativen und synthetischen Sequenz zurückführen lässt oder das Binden von CAP auf der *pap*-DNA weniger Einfluss auf die Transkription der *pap*-Gene wie gedacht [328], [329] hat, müssen weitere Experimente klären. Ebenso essentiell für die Transkription der *pap*-Gene ist, dass der *pap*-regulatorische Bereich das Methylierungsmuster (0,1) aufweist [317]. Eine zentrale Beobachtung dieser Studie deckt sich mit diesem Ergebnis: Das asymmetrische Methylierungsmuster (0,1) führte im Vergleich zu den anderen Methylierungszuständen zu einer deutlich höheren *degfp*-Transkription, die von der rechten *pap*-regulatorischen „Hälfte“ reguliert wurde. Darüber hinaus konnte durch die Kopplung des *mcherry*-Gens mit der *pap*-regulatorischen DNA herausgefunden werden, dass sich dieses Ergebnis auch auf die linke „Hälfte“ der *pap*-regulatorischen DNA übertragen lässt: Der Methylierungszustand (1,0) generierte die stärkste *mcherry*-Transkription im Vergleich zu den anderen Methylierungsmustern. Bemerkenswerterweise können diese erhöhten *degfp*- und *mcherry*-Transkriptionsraten nicht alleine auf die Methylierung der beiden Promotoren benachbarten $\text{GATC}^{\text{prox}}$ - bzw. $\text{GATC}^{\text{dist}}$ -Stellen

zurückgeführt werden, da der Methylierungszustand (1,1) in beiden Fällen zu keiner erhöhten Transkription führte, obwohl auch hier die jeweils dem Promotor benachbarten GATC-Stellen methyliert waren. Es lässt sich demnach schlussfolgern, dass die *E.-coli*-RNA-Polymerase nur dann stärker auf einem der beiden Promotoren, deren -35- und -10-Elemente entweder Lrp-Bindestelle 3 oder 4 flankieren, bindet, wenn die direkt benachbarte GATC-Stelle 2 oder 5 methyliert ist, aber die gegenüberliegende GATC-Stelle 5 bzw. 2 unmethyliert bleibt. Diese Beobachtungen legen eine methylierungsabhängige Wechselwirkung zwischen der linken und rechten „Hälfte“ des *pap*-regulatorischen Bereichs nahe, die Einfluss auf die Bindung der *E.-coli*-RNA-Polymerase auf die involvierten Promotoren hat.

Ob sich das zentrale Ergebnis einer methylierungsabhängigen Wechselwirkung zwischen der linken und rechten „Hälfte“ der *pap*-regulatorischen DNA auch bei zusätzlicher und gezielter zellfreier Synthese von Lrp und PapI weiterhin bewahrheitete, wurde im weiteren Verlauf der Studie überprüft (Abschn. 5.3.2). Zu diesem Zweck wurde der Phasenvariationsmechanismus bei zellfreier Synthese der Koregulatoren untersucht, indem die Transkriptionsraten der beiden Reportergene in Abhängigkeit des Methylierungsmusters der synthetischen *pap*-DNA und der Konzentration der Plasmide, die die *lrp*- oder *papI*-Expression ermöglichten, analysiert wurden. In allen drei Reaktions-Sätzen, in denen das *pap*-DNA-Konstrukt mit jeweils einem anderen Methylierungszustand vorhanden war, veränderten sich die Transkriptionsraten der Reportergene in Abhängigkeit der Lrp- und PapI-Konzentration.

Im Folgenden wird nur auf die wesentlichen Veränderungen der Reportergen-Transkriptionsraten durch den Einfluss der beiden Koregulatoren eingegangen. Generell ließen sich mit zunehmender Lrp-Konzentration die Expressionsraten beider Reportergene stärker reprimieren. Die PapI-Sensitivität bei der kooperativen Lrp-Bindung fiel je nach Methylierungszustand und betrachteter „Hälfte“ der *pap*-regulatorischen DNA unterschiedlich aus. Beim Methylierungsmuster (1,1) konnte die Lrp-Bindung auf den Lrp-Bindestellen 4–6 mit zunehmender PapI-Konzentration verstärkt werden, während sich die kooperative Lrp-Bindung auf den Bindestellen 1–3 dadurch nicht veränderte. Hingegen führte das Methylierungsmuster (0,1) bei geringeren Lrp-Konzentrationen und mit zunehmender PapI-Konzentration zu höheren *degfp*-Transkriptionsraten. Somit zeigte sich eine gewisse PapI-Abhängigkeit der kooperativen Lrp-Bindung, von der methylierten GATC^{prox}-Stelle geblockt zu werden, von der [313] zuvor berichtete. Dass die Methylierung von GATC^{prox} in Abwesenheit von PapI keinen Einfluss auf das kooperative Binden von Lrp auf den Lrp-Bindestellen 1–3 hat [313], konnte nicht beobachtet werden.

Jedoch richtete sich trotz des Einflusses der beiden Koregulatoren die Höhe der Reportergen-Expressionsraten hauptsächlich nach dem Methylierungsmuster des *pap*-DNA-Konstrukts. Das Methylierungsmuster (0,1) lieferte selbst bei maximaler Lrp- und PapI-bedingter Repression (tiefste deGFP-Fluoreszenz-Intensität) eine *degfp*-Transkriptionsrate, die etwa gleich groß wie die durch das Methylierungsmuster (1,0) verursachte maximale *degfp*-Transkriptionsrate (höchste deGFP-Fluoreszenz-Intensität) war. Sie war ca. 1,35-fach höher als die durch das Methylierungsmuster (1,1) generierte, maximale *degfp*-Transkriptionsrate. Die durchschnittliche durch das Methylierungsmuster (0,1) generierte *degfp*-Expressionsrate lag ca. 2,3-fach und 2,6-fach über der *degfp*-Transkriptionsrate, die vom Methylierungsmuster (1,0) bzw. (1,1) verursacht wurde. Das Methylierungsmuster (1,0) generierte selbst bei maximaler Lrp- und PapI-bedingter Repression (tiefste mCherry-Fluoreszenz-Intensität) eine *mcherry*-Transkriptionsrate, die nur etwas kleiner (95 % bzw. 94 %) war als die durch die Methylierungsmuster (0,1) und (1,1) verursachten maximalen *mcherry*-Transkriptionsraten (höchste mCherry-Fluoreszenz-Intensitäten). Durchschnittlich lag die *mcherry*-Expressionsrate beim Methylierungsmuster (1,0) ca. 1,3-fach über der des Methylierungsmusters (0,1) und 1,25-fach über der des Methylierungszustandes (1,1). Insofern stimmten auch beim Einfluss von Lrp und PapI die durch die verschiedenen Methylierungsmuster bedingten Unterschiede der Reportergen-Expressionsraten tendenziell mit den Differenzen überein, die sich ohne den Einfluss der beiden Koregulatoren durch die un-

terschiedlichen Methylierungszustände ergaben. Im Einklang mit [317] lässt sich schlussfolgern, dass das Methylierungsmuster die entscheidende Rolle beim Phasenvariationsmechanismus spielt. PapI und Lrp regulieren die Transkription jedoch zusätzlich im durch den Methylierungszustand vorgegebenen Rahmen. Somit legen auch die Resultate der Experimente, die die beiden Koregulatoren miteinbeziehen, eine methylierungsabhängige Wechselwirkung zwischen der linken und rechten „Hälfte“ des *pap*-regulatorischen Bereichs nahe, die die Bindung der *E.-coli*-RNA-Polymerase auf die involvierten Promotoren beeinflusst.

Aus diesem Grund wurde diese Wechselwirkung näher untersucht, indem die Sekundärstruktur der *pap*-regulatorischen DNA mittels des *mfold web servers* [334] vorhergesagt wurde. Die Vorhersage lieferte drei ähnliche Sekundärstrukturen. In jeder dieser drei Sekundärstrukturen kommt es zumindest zu einer teilweisen Basenpaarung zwischen den Nukleotiden der Lrp-Bindestellen 3 und 4 sowie 2 und 5. In jeder dieser Vorhersagen paaren sich GATC^{prox} und GATC^{dist} vollständig, und die betrachtete *pap*-DNA faltet sich über weite Bereiche und wechselwirkt somit erheblich mit sich selbst. Aus den vorhergesagten Sekundärstrukturen können sich DNA-Kreuzstrukturen bilden. Diese DNA-Konformation kann u.a. zur Genregulation beitragen und wechselwirkt mit Proteinen, beispielsweise mit solchen, die eine Auflösung der Kreuzstruktur verursachen [337]. Es ist bekannt, dass hemimethylierte GATC-Stellen verengte große Furchen ausbilden, weniger strukturell stabil sind und daher spezifisch mit Proteinen wechselwirken können [336]. Die Ausbildung einer Kreuzstruktur, bei der sich GATC^{prox} und GATC^{dist} paaren (Abb. 5.12), bietet einen möglichen Erklärungsansatz, die unterschiedlich hohen Reporter-Transkriptionsraten zu verstehen. Die Methylierungsmuster (0,1) und (1,0) liefern innerhalb der Kreuzstruktur unterschiedlich hemimethylierte Doppelstrang-GATC-Stellen, während das Methylierungsmuster (1,1) vollständig methylierte GATC-Stellen in der Kreuzkonformation besitzt. Es kann spekuliert werden, dass genau diese unterschiedlich hemimethylierten GATC-Stellen innerhalb der Kreuzstruktur zu einer spezifischen Wechselwirkung mit Lrp, PapI oder der *E.-coli*-RNA-Polymerase führen, die die unterschiedlich hohen Reporter-Transkriptionsraten bei den asymmetrischen Methylierungsmustern (0,1) und (1,0) hervorrufen. Dass die Art der Hemimethylierung der GATC^{dist}-Stelle das Bindungsverhalten der Koregulatoren Lrp und PapI beeinflusst, wurde bereits von [313] gezeigt.

Ob es tatsächlich zur spekulierten Ausbildung einer DNA-Kreuzstruktur kommt und ob diese den Phasenvariationsmechanismus steuert, muss in weiteren Experimenten geklärt werden. Mit Hilfe von magnetischen Pinzetten und mittels Förster-Resonanzenergietransfer (FRET) kann die DNA-Kreuzstruktur visualisiert werden, indem Donator und Akzeptor so innerhalb der zu analysierenden DNA platziert werden, dass es zu einem geringen FRET-Signal bei Ausbildung der DNA-B-Konformation kommt, aber ein hohes FRET-Signal detektierbar ist, wenn sich die beiden fluoreszierenden Moleküle durch Ausbildung einer Kreuzform annähern [338]. Durch Integration eines solchen experimentellen Aufbaus in das zellfreie System könnten die Reporter-Expressionsraten in Abhängigkeit der spekulierten DNA-Kreuzform in Ab- oder Anwesenheit der Koregulatoren PapI und Lrp untersucht werden. Zur Präparation des *E.-coli*-Extrakts können auch andere als in dieser Studie eingesetzte *E.-coli*-Zellen verwendet werden [52]. [317] benutzte in seiner Studie *lrp*⁻-*E.-coli*-Zellen. Ein zellfreier Extrakt, der aus solchen Zellen generiert würde, ermöglichte eine zellfreie Umgebung ohne Lrp-Reste. Durch Hinzugabe von PapI und Lrp in rekombinanter Form könnte der Phasenvariationsmechanismus unter kontrollierteren experimentellen Bedingungen untersucht werden.

A. Anhang

A.1. Ergänzendes Material zu Kap. 3

Anh. A.1 wurde bereits vorab in [158] veröffentlicht.

A.1.1. DNA-Sequenzen der verwendeten DNA-Moleküle

Im Folgenden werden die Sequenzen der verwendeten DNA-Moleküle in 5' → 3'-Richtung aufgelistet.

Σ = GGC ATC TTA GTA CAT ATC AG & Σ^c = CTG ATA TGT ACT AAG ATG CC.

σ = CGT TAT TAG C & σ^c = GCT AAT AAC G.

S_1 = TGT GCC TTT G & S_1^c = CAA AGG CAC A. S_2 = ACA CCT GAT C & S_2^c = GAT CAG GTG T. S_3 = GGA GTT GGT T & S_3^c = AAC CAA CTC C. S_4 = ATT TGG ACC G & S_4^c = CGG TCC AAA T. S_5 = CTT ACA CGA G & S_5^c = CTC GTG TAA G. S_6 = CTA CAT GAC G & S_6^c = CGT CAT GTA G. S_7 = GGG AAG TTG T & S_7^c = ACA ACT TCC C. S_8 = GAA GAC TAG C & S_8^c = GCT AGT CTT C.

S_1S_2 = TGT GCC TTT GAC ACC TGA TC & $S_2^cS_1^c$ = GAT CAG GTG TCA AAG GCA CA. S_3S_4 = GGA GTT GGT TAT TTG GAC CG & $S_4^cS_3^c$ = CGG TCC AAA TAA CCA ACT CC. S_5S_6 = CTT ACA CGA GCT ACA TGA CG & $S_6^cS_5^c$ = CGT CAT GTA GCT CGT GTA AG. S_7S_8 = GGG AAG TTG TGA AGA CTA GC & $S_8^cS_7^c$ = GCT AGT CTT CAC AAC TTC CC.

$S_1S_2S_3S_4$ = TGT GCC TTT GAC ACC TGA TCG GAG TTG GTT ATT TGG ACC G & $S_4^cS_3^cS_2^cS_1^c$ = CGG TCC AAA TAA CCA ACT CCG ATC AGG TGT CAA AGG CAC A. $S_5S_6S_7S_8$ = CTT ACA CGA GCT ACA TGA CGG GGA AGT TGT GAA GAC TAG C & $S_8^cS_7^cS_6^cS_5^c$ = GCT AGT CTT CAC AAC TTC CCC GTC ATG TAG CTC GTG TAA G.

$S_1S_2S_3S_4S_5S_6S_7S_8$ = TGT GCC TTT GAC ACC TGA TCG GAG TTG GTT ATT TGG ACC GCT TAC ACG AGC TAC ATG ACG GGG AAG TTG TGA AGA CTA GC & $S_8^cS_7^cS_6^cS_5^cS_4^cS_3^cS_2^cS_1^c$ = GCT AGT CTT CAC AAC TTC CCC GTC ATG TAG CTC GTG TAA GCG GTC CAA ATA ACC AAC TCC GAT CAG GTG TCA AAG GCA CA.

A.1.2. Reduzierung der Mungbohnen-Enzym-Aktivität

Im Referenz-Experiment, dessen Ergebnisse in Abb. 3.4B gezeigt werden, entfernte die Mungbohnen-Nuklease zwar die einzelsträngigen Überhänge, aber degradierte auch Sequenz-unspezifisch die spontan geformten längeren DNA-Templates und inhibierte somit deren Template-basierte Reproduktion. Daher blieb die Menge der entstehenden 40 bp-Stränge gering. Es konnten keine noch längeren Doppelstränge detektiert werden (Abb. 3.4B).

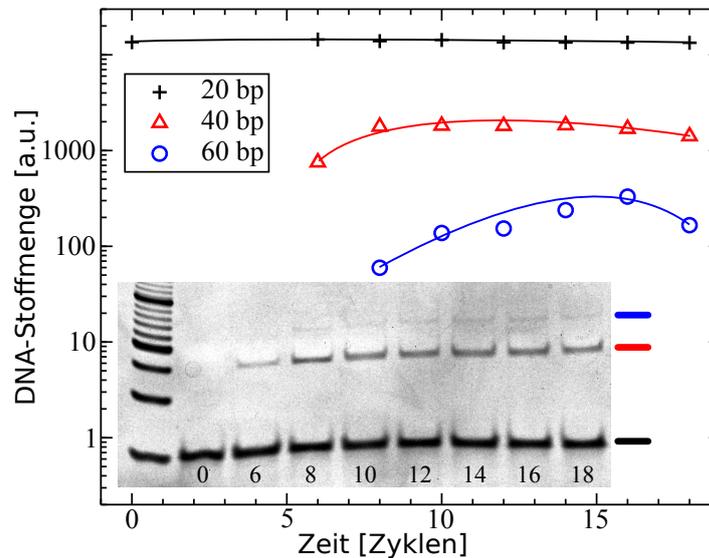


Abbildung A.1.: Zeitlicher Verlauf der relativen DNA-Stoffmenge der 20 bp-Stränge mit der Sequenz Σ bei reduzierter Enzym-Aktivität der Mungbohnen-Nuklease als zum Referenz-Experiment. Die Exonuklease wurde zu Beginn des siebten Zyklus hinzugegeben. Die Linien dienen der optischen Führung. Eingebettet zu sehen ist das zugehörige Polacrylamidgel. Die Zyklenzahl bei Probenentnahme ist angegeben. Links abgebildet ist die Auftrennung der 10 bp, 20 bp, ..., 150 bp-DNA-Stränge der DNA-Leiter (Abschn. 3.2.3). Es entstanden neben 40 bp-Strängen auch 60 bp-Stränge, aber unbegrenztes Längenwachstum der DNA-Stränge konnte nicht beobachtet werden. Abb. A.1 ist modifiziert aus [158] übernommen.

Insofern sollte im folgenden Experiment getestet werden, ob eine Reduzierung der Enzym-Aktivität der Mungbohnen-Nuklease die detektierbare Entstehung von DNA-Strängen einer Länge oberhalb von 40 bp ermöglicht. Zu diesem Zweck wurde die Mungbohnen-Nuklease einen Zyklus später als im Referenz-Experiment in die Proben gegeben. Vor dieser initialen Hinzugabe der Exonuklease wurden die Proben wie im Experiment, dessen Ergebnisse Abb. 3.4A zeigt, inkubiert. Es schloss sich ein einmaliger fünfminütiger Inkubations-Schritt bei einer Temperatur von 30 °C an. Wiederum wurden ab dem siebten Zyklus das Temperatur-Zeit-Profil der Zyklen verändert: 15 min bei 30 °C, 10 min bei 60 °C, 1 min bei 66 °C, 10 min bei 60 °C, 1 min bei 84 °C, gefolgt von 5 min bei 30 °C. Zwar wurde in gleicher Weise wie im Referenz-Experiment die Mungbohnen-Nuklease zu Beginn jedes weiteren Zyklus in die Proben gegeben, allerdings wurden nur noch 0,5 units anstelle von 1 unit in die Reaktionen pipettiert. Die beiden 20 b langen Einzelstränge mit den Sequenzen Σ und Σ^c waren zu Beginn jeweils 500 nM konzentriert, die Taq-DNA-Ligase besaß eine Konzentration von 8 units/ μ l.

Abb. A.1 zeigt, dass die reduzierte Mungbohnen-Aktivität tatsächlich zu intensiveren DNA-Banden der 40 bp-Stränge im Vergleich zum Referenz-Experiment (Abb. 3.4B) führte. Allerdings erzeugte der Überhang-Ligations-Mechanismus auch 60 b lange Stränge. Die resultierenden 60 bp-Doppelstränge

konnten schwach auf dem Polyacrylamidgel detektiert werden. DNA-Banden, die noch größeren DNA-Strängen zugeordnet werden können, waren nicht beobachtbar. Dieses Ergebnis untermauert zusätzlich die Hypothese, dass das unbegrenzte Längenwachstum der DNA-Stränge im Experiment, dessen Ergebnisse in Abb. 3.4A präsentiert werden, durch die Überhang-Ligation erzeugt wurde.

A.1.3. Bestimmung der spontanen Ligations- und autokatalytischen Vermehrungsrate

In Anh. A.1.3 wird beschrieben, wie α und β anhand der experimentellen Daten abgeschätzt werden. Um β zu bestimmen, wurden Phasen des Experiments betrachtet, in denen sich neue Moleküle hauptsächlich durch die freie Ligation bildeten. In diesem Fall kann die Vermehrung der Moleküle durch folgende Gleichung beschrieben werden:

$$\frac{d}{dt}[A_{2n}] = \beta[A_n]^2. \quad (\text{A.1})$$

In Gleichung A.1 beschreibt $[A_n]$ die Zahl der doppelsträngigen DNA-Moleküle A_n , n gibt die Zahl der Basenpaare der initial im System vorhandenen DNA-Moleküle an. Die spontane Ligation dominiert das experimentelle System stets in seinen frühen Phasen, wenn sich das System hauptsächlich aus den initial vorhandenen DNA-Strängen zusammensetzt. Dann ergibt sich der Wert für β durch folgende Gleichung:

$$\frac{[A_{2n}](t_1)}{t_1} = \beta[A_n](0)^2. \quad (\text{A.2})$$

In Gleichung A.2 ist t_1 der Zeitpunkt, an dem die ligierten DNA-Doppelstränge A_{2n} das erste Mal detektiert werden konnten.

Mit 20 bp langen DNA-Doppelsträngen der Sequenz Σ bei einer Temperatur von 60 °C startend (Abb. 3.4A), ergibt sich mit $[A_{20}](0) = 4961 \text{ fmol}$ und $[A_{40}](t_1 = 8 \text{ Zyklen}) = 266 \text{ fmol}$ $\beta = 13,5 \cdot 10^{-7} \text{ Zyklus}^{-1} \text{ fmol}^{-1}$. Mit 10 bp langen DNA-Doppelsträngen der Sequenzen S_1 – S_8 (Anh. A.1.1) bei einer Temperatur von 37 °C startend, resultiert mit $[A_{10}](0) = 3021 \text{ fmol}$ und $[A_{20}](t_1 = 18 \text{ Zyklen}) = 126 \text{ fmol}$ $\beta = 8 \cdot 10^{-7} \text{ Zyklus}^{-1} \text{ fmol}^{-1}$ (Abb. A.2). Um zu testen, ob der Wert von β von der initialen DNA-Stoffmenge der DNA abhängt, wurde mit einer höheren Stoffmenge an 10 bp-Doppelsträngen begonnen. Mit $[A_{10}](0) = 4123 \text{ fmol}$ und $[A_{20}](t_1 = 20 \text{ Zyklen}) = 270 \text{ fmol}$ ergibt sich erneut $\beta = 8 \cdot 10^{-7} \text{ Zyklus}^{-1} \text{ fmol}^{-1}$ (Abb. A.2). Anscheinend ist die β unempfindlich gegenüber der Konzentration der DNA-Stränge, der Temperatur und der Sequenz. Es wird zur weiteren Berechnung der durchschnittliche Wert $\beta = 9,8 \cdot 10^{-7} \text{ Zyklus}^{-1} \text{ fmol}^{-1}$ verwendet.

Die autokatalytische Vermehrungsrate wird durch Gleichung A.3 bestimmt:

$$\alpha[A_{2n}] = \frac{\frac{d}{dt}[A_{2n}] - \beta \sum_{k=1}^n [A_{2n-k}][A_k]}{\sum_{k=1}^n [A_{2n-k}][A_k]}. \quad (\text{A.3})$$

$\frac{d}{dt}[A_{2n}]$ wurde durch die Stoffmengendifferenz von $[A_{2n}]$ zweier aufeinanderfolgender Zeitpunkte abgeschätzt. Im Allgemeinen ist der auf diese Weise erhaltene Wert für α eine obere Abschätzung, da auch andere Ligationsprozesse zur Bildung von A_{2n} beitragen.

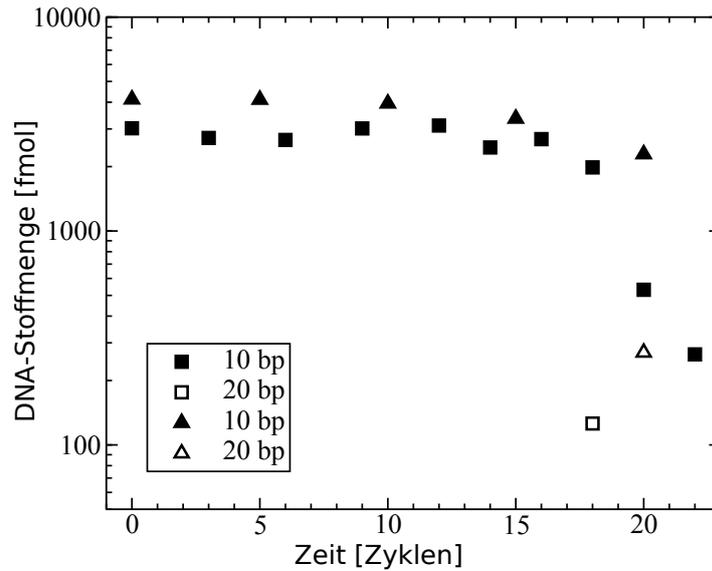


Abbildung A.2.: Zeitlicher Stoffmengen-Verlauf der 10 bp (gefüllte Symbole) und 20 bp (offene Symbole) langen DNA-Doppelstränge von zwei Experimenten, in denen freie Ligation dominierte. Die Experimente dienten der Bestimmung der Rate der freien Ligation β mit der die 10 b langen Stränge zu 20 b langen Strängen ligiert wurden. Die initiale Konzentration der 10 bp-Stränge mit den Sequenzen S_1 – S_8 betrug 3021 fmol (Quadrate) bzw. 4123 fmol (Dreiecke). Bei der niedrigeren initialen Stoffmenge konnten nach 18 Zyklen das erste Mal 20 bp-Stränge mit einer Stoffmenge von 126 fmol detektiert werden. Bei der höheren initialen 10 bp-Strang-Konzentration konnten das erste Mal nach 20 Zyklen 20 bp-Doppelstränge mit einer Stoffmenge von 270 fmol nachgewiesen werden. In beiden Experimenten resultiert, mit Hilfe der im Text angegebenen Abschätzung, $\beta = 8 \cdot 10^{-7} \text{ Zyklus}^{-1} \text{ fmol}^{-1}$. Abb. A.2 ist modifiziert aus [158] übernommen.

Die autokatalytische Vermehrungsrate $\alpha[A_n]$ und der dazugehörige Quotient ω_n zwischen autokatalytischer Vermehrungsrate $\alpha[A_n]$ und β der Experimente, deren Ergebnisse Abb. 3.4A und Abb. 3.8 präsentieren, werden in Tab. A.1 bzw. Tab. A.2 gezeigt.

Tabelle A.1.: Autokatalytische Vermehrungsrate $\alpha [A_n]$ und der dazugehörige Quotient ω_n zwischen autokatalytischer Vermehrungsrate $\alpha [A_n]$ und β des Experiments, dessen Ergebnisse in Abb. 3.4A präsentiert werden. $[A_n]$ bezeichnet die Stoffmenge der doppelsträngigen DNA-Stränge A_n , n gibt die Zahl ihrer Basenpaare an. t_1 und t_2 definieren das Zeitintervall in Zyklen, in dem die autokatalytische Vermehrung von $[A_n]$ analysiert wurde. Durchschnittlich ergab sich $\omega \approx 40$. Tab. A.1 ist modifiziert aus [158] übernommen.

$t_1 - t_2$	$\alpha \cdot [A_{40}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{40}	$t_1 - t_2$	$\alpha \cdot [A_{60}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{60}
6-8	$4,97 \cdot 10^{-6}$	5,08	14	$20,01 \cdot 10^{-6}$	20,42
8-10	$2,35 \cdot 10^{-6}$	2,4	15	$10,16 \cdot 10^{-6}$	10,37
10-12	$0,5 \cdot 10^{-6}$	0,51	16	$41,21 \cdot 10^{-6}$	42,05
13	$10,47 \cdot 10^{-6}$	10,68	17	$37,18 \cdot 10^{-6}$	37,93
14	$3,74 \cdot 10^{-6}$	3,82	18	$58,42 \cdot 10^{-6}$	59,62
15	$-0,35 \cdot 10^{-6}$	-0,35	19	$35,36 \cdot 10^{-6}$	36,08
16	$7,31 \cdot 10^{-6}$	7,46			
17	$5,53 \cdot 10^{-6}$	5,65			
18	$26,98 \cdot 10^{-6}$	27,53			
19	$1,04 \cdot 10^{-6}$	1,06			
$t_1 - t_2$	$\alpha \cdot [A_{80}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{80}	$t_1 - t_2$	$\alpha \cdot [A_{100}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{100}
16	$31,12 \cdot 10^{-6}$	31,76	18	$279,57 \cdot 10^{-6}$	285,27
17	$17,73 \cdot 10^{-6}$	18,09	19	$36,82 \cdot 10^{-6}$	37,57
18	$57,06 \cdot 10^{-6}$	58,22	20	$46,74 \cdot 10^{-6}$	47,69
19	$35,27 \cdot 10^{-6}$	35,99	21	$-74,68 \cdot 10^{-6}$	-76,20
20	$0,58 \cdot 10^{-6}$	0,60	22	$17,57 \cdot 10^{-6}$	17,92
$t_1 - t_2$	$\alpha \cdot [A_{120}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{120}	$t_1 - t_2$	$\alpha \cdot [A_{140}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{140}
18	$363,22 \cdot 10^{-6}$	370,63	20	$73,65 \cdot 10^{-6}$	75,15
19	$0,06 \cdot 10^{-6}$	0,06	21	$18,66 \cdot 10^{-6}$	19,07
20	$54,63 \cdot 10^{-6}$	55,74	22	$44,06 \cdot 10^{-6}$	44,96
$t_1 - t_2$	$\alpha \cdot [A_{160}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{160}	$t_1 - t_2$	$\alpha \cdot [A_{180}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{180}
21	$53,37 \cdot 10^{-6}$	54,45	21	$42,01 \cdot 10^{-6}$	42,87
22	$58,11 \cdot 10^{-6}$	59,23	22	$80,77 \cdot 10^{-6}$	82,42
23	$4,60 \cdot 10^{-6}$	4,70	23	$30,51 \cdot 10^{-6}$	31,13
$t_1 - t_2$	$\alpha \cdot [A_{200}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{200}			
22	$153,83 \cdot 10^{-6}$	156,97			
23	$29,73 \cdot 10^{-6}$	30,34			

Tabelle A.2.: Autokatalytische Vermehrungsrate $\alpha[A_n]$ und der dazugehörige Quotient ω_n zwischen autokatalytischer Vermehrungsrate $\alpha[A_n]$ und β des Experiments, dessen Ergebnisse in Abb. 3.8 präsentiert werden. $[A_n]$ bezeichnet die Stoffmenge der doppelsträngigen DNA-Stränge A_n , n gibt die Zahl ihrer Basenpaare an. t_1 und t_2 definieren das Zeitintervall in Zyklen, in dem die autokatalytische Vermehrung von $[A_n]$ analysiert wurde. Durchschnittlich ergab sich $\omega \approx 510$. Tab. A.2 ist modifiziert aus [158] übernommen.

$t_1 - t_2$	$\alpha \cdot [A_{20}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{20}
0-6	$10,14 \cdot 10^{-6}$	10,35
6-12	$8,75 \cdot 10^{-6}$	8,93
$t_1 - t_2$	$\alpha \cdot [A_{40}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{40}
25	$201,00 \cdot 10^{-6}$	205,11
26	$96,05 \cdot 10^{-6}$	98,01
$t_1 - t_2$	$\alpha \cdot [A_{80}]$ [fmol ⁻¹ · Zyklus ⁻¹]	ω_{80}
28	$596,00 \cdot 10^{-6}$	608,16
29	$437,15 \cdot 10^{-6}$	446,07
30	$1578,65 \cdot 10^{-6}$	1610,87
31	$1080,69 \cdot 10^{-6}$	1102,74

A.2. Ergänzendes Material zu Kap. 4

A.2.1. Ergänzungen zur Stopcodon-Suppression mittels der aminoacylierten Suppressor-tRNA

Zur Quantifizierung des vollständig translatierten EGFP^{N149amber}-Moleküls wurden zellfreie Reaktionen hergestellt, die sich von den Reaktionen aus dem in Abschn. 4.3.3.2 präsentierten Experiment (Referenz-Experiment) dadurch unterschieden, dass anstelle des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500 und der AcLys-tRNA^{amber} die Konzentration eines GFP-Standards (*Vector Laboratories Inc.*, Burlingame, CA, USA) variiert wurde. Die zellfreien Reaktionen wurden in gleicher Weise wie die Reaktionen des Referenz-Experiments über Nacht (16 h) bei 29 °C inkubiert. Ebenso wurde die Fluoreszenz-Intensität dieser Reaktionen mit den gleichen Messparametern wie bei der Fluoreszenz-Messung des Referenz-Experiments gemessen.

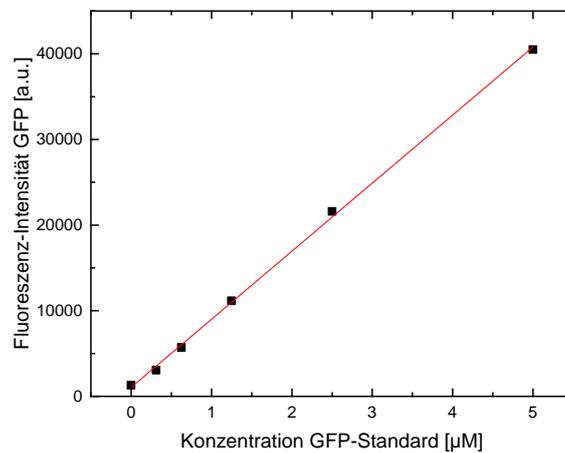


Abbildung A.3.: Fluoreszenz-Intensität der zellfreien Reaktionen in Abhängigkeit der Konzentration eines GFP-Standards. Die Steigung der Fitgeraden beträgt gerundet $7950,24\mu\text{M}^{-1}$, der Schnittpunkt mit der Y-Achse liegt bei 1050,34. Die Steigung der Fitgeraden wird in Tab. A.3 zur Berechnung der Konzentration des vollständig translatierten EGFP^{N149amber}-Moleküls verwendet.

In der ersten Spalte von Tab. A.3 ist die gemessene Fluoreszenz-Intensität der zellfreien Reaktionen aus dem Referenz-Experiment zu finden. Diese Fluoreszenz-Intensitäten gehören von oben nach unten jeweils zu den zellfreien Reaktionen, in denen die AcLys-tRNA^{amber} mit einer Konzentration von 0 nM, 75 pM, 150 pM, ..., 4,35 µM gelöst war. In der zweiten Spalte von links sind die Fluoreszenz-Intensitäten abzüglich der Hintergrund-Fluoreszenz-Intensitäten der Referenzreaktionen ($c(\text{AcLys-tRNA}^{\text{amber}})=0$) des Referenz-Experiments aufgelistet. Die Werte der dritten Spalte von links errechnen die Konzentration von GFP durch Division der Werte aus der zweiten Spalte durch die Steigung der Fitgeraden aus Abb. A.3. In der vierten Spalte werden die in der dritten Spalte gelisteten Werte korrigiert: Die Konzentration von vollständig translatiertem EGFP^{N149amber} ergibt sich durch Division der Werte aus der dritten Spalte durch sechs, da die Fluoreszenz-Intensität von EGFP bei Anregung bei einer Wellenlänge von 488 nm etwa sechs Mal höher als die von GFP ist [150]. Die Konzentrationen wurden auf die zweite Nachkommastelle gerundet.

FI [a.u.]	abzgl. HG-FI [a.u.]	c(GFP) [μ M]	c(EGFP) [μ M]
4478	0	0	0
4852	374	0,05	0
5070	592	0,07	0,01
5023	545	0,07	0,01
5169	691	0,09	0,01
6511	2033	0,26	0,04
16499	12021	1,51	0,25
20750	16272	2,05	0,34
17135	12657	1,59	0,26

Tabelle A.3.: Berechnung der Konzentration des vollständig translatierten EGFP^{N149amber}-Moleküls, in dem das Stopcodon durch den Einbau von AcLys supprimiert wurde. Weitere Informationen sind dem Text zu entnehmen.

A.2.2. DNA-Sequenzen

Im Folgenden werden die Sequenzen der transkriptions-translations-regulatorischen und kodierenden Elemente der in Kap.4 verwendeten DNA-Konstrukte angegeben.

Alle zur zellfreien Genexpression verwendeten Plasmide basieren auf pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP-T500, das von der Kooperationsarbeitsgruppe von Vincent Noireaux (*School of Physics and Astronomy, University of Minnesota*) zur Verfügung gestellt und bereits zuvor in [60] verwendet wurde. Details zu diesem Plasmid und zu den transkriptions-translations-regulatorischen Elementen können Abschn. 4.2.9.1 entnommen werden. Weitere Details sind im Methoden-Teil Abschn. 4.2.11 zu finden.

A.2.2.1. pBEST-p15a-OR2-OR1-Pr-UTR1-deGFPHis₆-Tag-T500

```

1  TGAGC TAACA CCGTG CGTGT TGACA ATTTT ACCTC TGGCG GTGAT AATGG TTGCA GCTAG
61  CAATA ATTTT GTTTA ACTTT AAGAA GGAGA TATAC CATGG AGCTT TTCAC TGGCG TTGTT
121 CCCAT CCTGG TCGAG CTGGA CGGCG ACGTA AACGG CCACA AGTTC AGCGT GTCCG GCGAG
181 GGGCA GGGCG ATGCC ACCTA CGGCA AGCTG ACCCT GAAGT TCATC TGCAC CACCG GCAAG
241 CTGCC CGTGC CCGTG CCCAC CCTCG TGACC ACCCT GACCT ACGGC GTGCA GTGCT TCAGC
301 CGCTA CCCCC ACCAC ATGAA GCAGC ACGAC TTCTT CAAGT CCGCC ATGCC CGAAG GCTAC
361 GTCCA GGAGC GCACC ATCTT CTTCA AGGAC GACGG CAACT ACAAG ACCCG CGCCG AGGTG
421 AAGTT CGAGG GCGAC ACCCT GGTGA ACCGC ATCGA GCTGA AGGGC ATCGA CTTCA AGGAG
481 GACGG CAACA TCCTG GGGCA CAAGC TGGAG TACAA CTACA ACAGC CACAA CGTCT ATATC
541 ATGGC CGACA AGCAG AAGAA CGGCA TCAAG GTGAA CTTCA AGATC CGCCA CAACA TCGAG
601 GACGG CAGCG TGCAG CTCGC CGACC ACTAC CAGCA GAACA CCCCC ATCGG CGACC GCCCC
661 GTGCT GCTGC CCGAC AACCA CTACC TGAGC ACCCA GTCCG CCCTG AGCAA AGACC CCAAC
721 GAGAA GCGCG ATCAC ATGGT CCTGC TGGAG TTCGT GACCG CCGCC GGGAT CCATC ATCAC
781 CATCA CCACT AACTC GAGCA AAGCC CGCCG AAAGG CGGGC TTTTC TGT

```

Abbildung A.4.: Transkriptions-translations-regulatorische und kodierende DNA-Sequenzen des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-deGFPHis₆-Tag-T500. Cyan: OR2-OR1-Pr. Orange: UTR1. Grün: für deGFPHis₆ kodierendes Gen. Braun: His₆-Tag. Grau: T500.

A.2.2.2. pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500

```

1  TGAGC TAACA CCGTG CGTGT TGACA ATTTT ACCTC TGGCG GTGAT AATGG TTGCA GCTAG
61  CAATA ATTTT GTTTA ACTTT AAGAA GGAGA TATAC CATGG TGAGC AAGGG CGAGG AGCTG
121 TTCAC CGGGG TGGTG CCCAT CCTGG TCGAG CTGGA CGGCG ACGTA AACGG CCACA AGTTC
181 AGCGT GTCCG GCGAG GGGCA GGGCG ATGCC ACCTA CGGCA AGCTG ACCCT GAAGT TCATC
241 TGCAC CACCG GCAAG CTGCC CGTGC CCGTG CCCAC CCTCG TGACC ACCCT GACCT ACGGC
301 GTGCA GTGCT TCAGC CGCTA CCCCC ACCAC ATGAA GCAGC ACGAC TTCTT CAAGT CCGCC
361 ATGCC CGAAG GCTAC GTCCA GGAGC GCACC ATCTT CTTCA AGGAC GACGG CAACT ACAAG
421 ACCCG CGCCG AGGTG AAGTT CGAGG GCGAC ACCCT GGTGA ACCGC ATCGA GCTGA AGGGC
481 ATCGA CTTCA AGGAG GACGG CAACA TCCTG GGGCA CAAGC TGGAG TACAA CTACA ACAGC
541 CACTA GGTCT ATATC ATGGC CGACA AGCAG AAGAA CGGCA TCAAG GTGAA CTTCA AGATC
601 CGCCA CAACA TCGAG GACGG CAGCG TGCAG CTCGC CGACC ACTAC CAGCA GAACA CCCCC
661 ATCGG CGACC GCCCC GTGCT GCTGC CCGAC AACCA CTACC TGAGC ACCCA GTCCG CCCTG
721 AGCAA AGACC CCAAC GAGAA GCGCG ATCAC ATGGT CCTGC TGGAG TTCGT GACCG CCGCC
781 GGGAT CACTC TCGGC ATGGA CGAGC TGTAC AAGCA TCACC ATCAC CATCA CTAAC TCGAG
841 CAAAG CCCGC CGAAA GGCGG GCTTT TCTGT

```

Abbildung A.5.: Transkriptions-translations-regulatorische und kodierende DNA-Sequenzen des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-EGFP^{N149amber}-T500. Cyan: OR2-OR1-Pr. Orange: UTR1. Grün: für EGFP^{N149amber} kodierendes Gen. Rot: amber-Stopcodon. Grau: T500.

A.2.2.3. T7-tRNA^{Pyl}_{CUA}

1 GCATG TCACT ACGAA ATTAA TACGA CTCAC TATAG GAAAC CTGAT CATGT AGATC GAATG
 61 GACTC TAAAT CCGTT CAGCC GGGTT AGATT CCCGG GGTTC CCGCC ATTCA TTACA AGC

Abbildung A.6.: Transkriptions-translations-regulatorische und kodierende DNA-Sequenzen des (elongierten) linearen DNA-Konstrukts T7-tRNA^{Pyl}_{CUA}. Cyan: T7-Promotor. Orange: für tRNA^{Pyl}_{CUA} kodierendes Gen. Farblos: Elongierter Bereich zur Sequenzierung.

A.2.2.4. pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA^{Pyl}_{CUA}-PylRS^{Y384F}-T500

1 TGAGC TAACA CCGTG CGTGT TGACA ATTTT ACCTC TGGCG GTGAT AATGG TTGCA GCTAG
 61 CAATA ATTTT GTTTA ACTTT AAGAA GGAGA TATAC CATGG GGAAA CCTGA TCATG TAGAT
 121 CGAAT GGACT CTAAA TCCGT TCAGC CGGGT TAGAT TCCCG GGGTT TCCGC CACCT AGGTT
 181 TAAAC TAAGG AGGTT AATTA AATGA GCGCT TGGAG CCACC CGCAG TTCGA AAAAC TCGAG
 241 ATGGA TAAAA AACCA CTAAA CACTC TGATA TCTGC AACCG GGCTC TGGAT GTCCA GGACC
 301 GGAAC AATTC ATAAA ATAAA ACACC ACGAA GTCTC TCGAA GCAAA ATCTA TATTG AAATG
 361 GCATG CGGAG ACCAC CTTGT TGTAA ACAAC TCCAG GAGCA GCAGG ACTGC AAGAG CGCTC
 421 AGGCA CCACA AATAC AGGAA GACCT GCAAA CGCTG CAGGG TTTCC GATGA GGATC TCAAT
 481 AAGTT CCTCA CAAAG GCAAA CGAAG ACCAG ACAAG CGTAA AAGTC AAGGT CGTTT CTGCC
 541 CCTAC CAGAA CGAAA AAGGC AATGC CAAAA TCCGT TCGCA GAGCC CCGAA ACCTC TTGAG
 601 AATAC AGAAG CGGCA CAGGC TCAAC CTTCT GGATC TAAAT TTTCA CCTGC GATAC CGGTT
 661 TCCAC CCAAG AGTCA GTTTC TGTCC CGGCA TCTGT TTCAA CATCA ATATC AAGCA TTTCT
 721 ACAGG AGCAA CTGCA TCCGC ACTGG TAAAA GGGAA TACGA ACCCC ATTAC ATCCA TGTCT
 781 GCCCC TGTTT AGGCA AGTGC CCCCC CACTT ACGAA GAGCC AGACT GACAG GCTTG AAGTC
 841 CTGTT AAACC CAAAA GATGA GATTT CCCTG AATTC CGGCA AGCCT TTCAG GGAGC TTGAG
 901 TCCGA ATTGC TCTCT CGCAG AAAAA AAGAC CTGCA GCAGA TCTAC GCGGA AGAAA GGGAG
 961 AATTA TCTGG GGAAA CTCGA GCGTG AAATT ACCAG GTTCT TTGTG GACAG GGGTT TTCTG
 1021 GAAAT AAAAT CCCCC ATCCT GATCC CTCTT GAGTA TATCG AAAGG ATGGG CATTG ATAAT
 1081 GATAC CGAAC TTTCA AAACA GATCT TCAGG GTTGA CAAGA ACTTC TGCCT GAGAC CCATG
 1141 CTTGC TCCAA ACCTT TACAA CTACC TGCGC AAGCT TGACA GGGCC CTGCC TGATC CAATA
 1201 AAAAT TTTTG AAATA GGCCC ATGCT ACAGA AAAGA GTCCG ACGGC AAAGA ACACC TCGAA
 1261 GAGTT TACCA TGCTG AACTT CTGCC AGATG GGATC GGGAT GCACA CGGGA AAATC TTGAA
 1321 AGCAT AATTA CGGAC TTCCT GAACC ACCTG GGAAT TGATT TCAAG ATCGT AGGCG ATTCC
 1381 TGCAT GGTCT TTGGG GATAC CCTTG ATGTA ATGCA CGGAG ACCTG GAACT TTCCT CTGCA
 1441 GTAGT CGGAC CCATA CCGCT TGACC GGGAA TGGGG TATTG ATAAA CCCTG GATAG GGGCA
 1501 GGTTC CGGGC TCGAA CGCCT TCTAA AGGTT AAACA CGACT TTAAA AATAT CAAGA GAGCT
 1561 GCAAG GTCCG AGTCT TACTA TAACG GGATT TCTAC CAACC TGTAAG GGATC CCAAA GCCCC
 1621 CCGAA AGGCG GGCTT TTCTG I

Abbildung A.7.: Transkriptions-translations-regulatorische und kodierende DNA-Sequenzen des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-tRNA^{Pyl}_{CUA}-PylRS^{Y384F}-T500. Cyan: OR2-OR1-Pr. Orange: UTR1. Braun/Gelb: für tRNA^{Pyl}_{CUA}/PylRS kodierende Gene. Grau: T500.

A.2.2.5. OR2-OR1-Pr-UTR1-T7-RNA-Polymerase-T500

1 **TGAGC TAACA CCGTG CGTGT TGACA ATTTT ACCTC TGGCG GTGAT AATGG TTGCA** GCTAG
 61 **CAATA ATTTT GTTTA ACTTT AAGAA GGAGA TATAC** CATGC ATCAT CACCA TCACC ACAAC
 121 **ACGAT TAACA TCGCT AAGAA CGACT TCTCT GACAT CGAAC TGGCT GCTAT CCCGT TCAAC**
 181 **ACTCT GGCTG ACCAT TACGG TGAGC GTTTA GCTCG CGAAC AGTTG GCCCT TGAGC ATGAG**
 241 **TCTTA CGAGA TGGGT GAAGC ACGCT TCCGC AAGAT GTTTG AGCGT CAACT TAAAG CTGGT**
 301 **GAGGT TGCGG ATAAC GCTGC CGCCA AGCCT CTCAT CACTA CCCTA CTCCC TAAGA TGATT**
 361 **GCACG CATCA ACGAC TGGTT TGAGG AAGTG AAAGC TAAGC GCGGC AAGCG CCCGA CAGCC**
 421 **TTCCA GTTCC TGCAA GAAAT CAAGC CGGAA GCCGT AGCGT ACATC ACCAT TAAGA CCACT**
 481 **CTGGC TTGCC TAACC AGTGC TGACA ATACA ACCGT TCAGG CTGTA GCAAG CGCAA TCGGT**
 541 **CGGGC CATTG AGGAC GAGGC TCGCT TCGGT CGTAT CCGTG ACCTT GAAGC TAAGC ACTTC**
 601 **AAGAA AAACG TTGAG GAACA ACTCA ACAAG CGCGT AGGGC ACGTC TACAA GAAAG CATT**
 661 **ATGCA AGTTG TCGAG GCTGA CATGC TCTCT AAGGG TCTAC TCGGT GGCGA GCGGT GGTCT**
 721 **TCGTG GCATA AGGAA GACTC TATTC ATGTA GGAGT ACGCT GCATC GAGAT GCTCA TTGAG**
 781 **TCAAC CGGAA TGGTT AGCTT ACACC GCCAA AATGC TGGCG TAGTA GGTCA AGACT CTGAG**
 841 **ACTAT CGAAC TCGCA CCTGA ATACG CTGAG GCTAT CGCAA CCCGT GCAGG TGCGC TGGCT**
 901 **GGCAT CTCTC CGATG TTCCA ACCTT GCGTA GTTCC TCCTA AGCCG TGGAC TGGCA TTA**
 961 **GGTGG TGGCT ATTGG GCTAA CGGTC GTCGT CCTCT GGCGC TGGTG CGTAC TCACA GTAAG**
 1021 **AAAGC ACTGA TGCGC TACGA AGACG TTTAC ATGCC TGAGG TGTAC AAAGC GATTA ACATT**
 1081 **GCGCA AAACA CCGCA TGGAA AATCA ACAAG AAAGT CCTAG CCGTC GCCAA CGTAA TCACC**
 1141 **AAGTG GAAGC AATTGT CCGGT CGAGG ACATC CCTGC GATTG AGCGT GAAGA ACTCC CGATG**
 1201 **AAACC GGAAG ACATC GACAT GAATC CTGAG GCTCT CACCG CGTGG **AGACG** TGCTG CCGCT**
 1261 **GCTGT GTACC GCAAG GACAA GGCTC GCAAG TCTCG CCGTA TCAGC CTTGA GTTCA TGCTT**
 1321 **GAGCA AGCCA ATAAG TTTGC TAACC ATAAG GCCAT CTGGT TCCCT TACAA CATGG ACTGG**
 1381 **CGCGG TCGTG TTTAC GCTGT GTCAA TGTTC AACCC GCAAG GTAAC GATAT GACCA AAGGA**
 1441 **CTGCT TACGC TGGCG AAAGG TAAAC CAATC GGTAA GGAAG GTTAC TACTG GCTGA AAATC**
 1501 **CACGG TGCAA ACTGT GCGGG TGTCC ATAAG GTTCC GTTCC CTGAG CGCAT CAAGT TCATT**
 1561 **GAGGA AAACC ACGAG AACAT CATGG CTTGC GCTAA GTCTC CACTG GAGAA CACTT GGTGG**
 1621 **GCTGA GCAAG AATTCT CCGTT CTGCT TCCTT GCGTT CTGCT TTGAG TACGC TGGGG TACAG**
 1681 **CACCA CGGCC TGAGC TATAA CTGCT CCCTT CCGCT GGCGT TTGAC GGGTC TTGCT CTGGC**
 1741 **ATCCA GCACT TCTCC GCGAT GCTCC GAGAT GAGGT AGGTG GTCCG CCGGT TAACT TGCTT**
 1801 **CCTAG TGAAA CCGTT CAGGA CATCT ACGGG AATTGT TGCTA AGAAA GTCAA CGAGA TTCTA**
 1861 **CAAGC AGACG CAATC AATGG GACCG ATAAC GAAGT AGTTA CCGTG ACCGA TGAGA ACACT**
 1921 **GGTGA AATCT CTGAG AAAGT CAAGC TGGGC ACTAA GGCAC TGGCT GGTCA ATGGC TGGCT**
 1981 **TACGG TGTTA CTCGC AGTGT GACTA AGCGT TCAGT CATGA CGCTG GCTTA CCGGT CAAAA**
 2041 **GAGTT CCGCT TCCGT CAACA AGTGC TGGAA GATAC CATTG AGCCA GCTAT TGATT CCGGC**
 2101 **AAGGG TCTGA TGTTC ACTCA GCCGA ATCAG GCTGC TGGAT ACATG GCTAA GCTGA TTTGG**
 2161 **GAATC TGTGA GCGTG ACGGT GGTAG CTGCC GTTGA AGCAA TGAAC TGGCT TAAGT CTGCT**
 2221 **GCTAA GCTGC TGGCT GCTGA GGTCA AAGAT AAGAA GACTG GAGAG ATTCT TCGCA AGCGT**
 2281 **TGCGC TGTGC AATTGG GTAAC TCCTG ATGGT TTCCC TGTGT GGCAG GAATA CAAGA AGCCT**
 2341 **ATTCA GACGC GCTTG AACCT GATGT TCCTC GGTCA GTTCC GCTTA CAGCC TACCA TTAAC**
 2401 **ACCAA CAAAG ATAGC GAGAT TGATG CACAC AAACA GGAGT CTGGT ATCGC TCCTA ACTTT**
 2461 **GTACA CAGCC AAGAC GGTAG CCACC TTCGT AAGAC TGTAG TGTGG GCACA CGAGA AGTAC**
 2521 **GGAAT CGAAT CTTTT GCACT GATTC ACGAC TCCTT CCGTA CCATT CCGGC TGACG CTGCG**
 2581 **AACCT GTTCA AAGCA GTGCG CGAAA CTATG GTTGA CACAT ATGAG TCTTG TGATG TACTG**
 2641 **GCTGA TTTCT ACGAC CAGTT CGCTG ACCAG TTGCA CGAGT CTCAA TTGGA CAAAA TGCCA**
 2701 **GCACT TCCGG CTA AAA GGTAA CTTGA ACCTC CGTGA CATCT TAGAG TCGGA CTTCC CGTTC**
 2761 **GCGTA ACTCG AGCAA **AGCCC** GCCGA AAGGC GGGCT TTTCT **GT****

Abbildung A.8.: Transkriptions-translations-regulatorische und kodierende DNA-Sequenzen des linearen DNA-Konstrukts OR2-OR1-Pr-UTR1-T7-RNA-Polymerase-T500. Cyan: OR2-OR1-Pr. Orange: UTR1. Gelb: für T7-RNA-Polymerase kodierendes Gen. Braun: His₆-Tag. Rot: Mutation K378R. Grau: T500.

A.2.2.6. pBEST-p15a-OR2-OR1-Pr-UTR1-PyIRS^{Y384F}-T500

1 **TGAGC TAACA CCGTG CGTGT TGACA ATTTT ACCTC TGGCG GTGAT AATGG TTGCA** GCTAG
 61 **CAATA ATTTT GTTTA ACTTT AAGAA GGAGA TATAC** CATGG ATAAA AAACC ACTAA ACACT
 121 **CTGAT ATCTG CAACC GGGCT CTGGA TGTCC AGGAC CGGAA CAATT CATAA AATAA AACAC**
 181 **CACGA AGTCT CTCGA AGCAA AATCT ATATT GAAAT GGCAT GCGGA GACCA CCTTG TTGTA**
 241 **AACAA CTCCA GGAGC AGCAG GACTG CAAGA GCGCT CAGGC ACCAC AAATA CAGGA AGACC**
 301 **TGCAA ACGCT GCAGG GTTTC GGATG AGGAT CTCAA TAAGT TCCTC ACAAA GGCAA ACGAA**
 361 **GACCA GACAA GCGTA AAAGT CAAGG TCGTT TCTGC CCCTA CCAGA ACGAA AAAGG CAATG**
 421 **CCAAA ATCCG TTGCG AGAGC CCCGA AACCT CTTGA GAATA CAGAA GCGGC ACAGG CTCAA**
 481 **CCTTC TGGAT CTAAA TTTTC ACCTG CGATA CCGGT TTCCA CCCAA GAGTC AGTTT CTGTC**
 541 **CCGGC ATCTG TTTCA ACATC AATAT CAAGC ATTTC TACAG GAGCA ACTGC ATCCG CACTG**
 601 **GTAAA AGGGA ATACG AACCC CATTa CATCC ATGTC TGCCC CTGTT CAGGC AAGTG CCCCC**
 661 **GCACT TACGA AGAGC CAGAC TGACA GGCTT GAAGT CCTGT TAAAC CAAA AGATG AGATT**
 721 **TCCCT GAATT CCGGC AAGCC TTTCA GGGAG CTTGA GTCCG AATTG CTCTC TCGCA GAAAA**
 781 **AAAGA CCTGC AGCAG ATCTA CGCGG AAGAA AGGGA GAATT ATCTG GGGAA ACTCG AGCGT**
 841 **GAAAT TACCA GGTTC TTTGT GGACA GGGGT TTTCT GGAAA TAAAA TCCCC GATCC TGATC**
 901 **CCTCT TGAGT ATATC GAAAG GATGG GCATT GATAA TGATA CCGAA CTTTC AAAAC AGATC**
 961 **TTCAG GGTG ACAAG AACTT CTGCC TGAGA CCCAT GCTTG CTCCA AACCT TTACA ACTAC**
 1021 **CTGCG CAAGC TTGAC AGGGC CCTGC CTGAT CCAAT AAAAA TTTTT GAAAT AGGCC CATGC**
 1081 **TACAG AAAAG AGTCC GACGG CAAAG AACAC CTCGA AGAGT TTACC ATGCT GAACT TCTGC**
 1141 **CAGAT GGGAT CCGGA TGCAC ACGGG AAAAT CTTGA AAGCA TAATT ACGGA CTTC TGAAC**
 1201 **CACCT GGGAA TTGAT TTCAA GATCG TAGGC GATTC CTGCA TGGTC TTTGG GGATA CCCTT**
 1261 **GATGT AATGC ACGGA GACCT GGAAC TTTCC TCTGC AGTAG TCGGA CCCAT ACCGC TTGAC**
 1321 **CGGGA ATGGG GTATT GATAA ACCCT GGATA GGGGC AGGTT TCGGG CTCGA ACGCC TTCTA**
 1381 **AAGGT TAAAC ACGAC TTTAA AAATA TCAAG AGAGC TGCAA GGTCC GAGTC TTA CT**
 1441 **GGGAT TTCTA CCAAC CTGTA AGGAT CCAA AGCCC GCCGA AAGGC GGGCT TTCT GT**

Abbildung A.9.: Transkriptions-translations-regulatorische und kodierende DNA-Sequenzen des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-PyIRS^{Y384F}-T500. Cyan: OR2-OR1-Pr. Orange: UTR1. Gelb: für PyIRS kodierendes Gen. Grau: T500.

A.2.2.7. pBADmod1-linker2-gamS

1 **GACGC TTTTT ATCGC AACTC TCTAC TGTTT CTCCA TACCC GTTTT TTAAC TTTAA GAAGG**
 61 **AGATA TACCA TGGGA AACGC TTATT ACATT CAGGA TCGTC TTGAG GCTCA GAGCT GGGCG**
 121 **CGTCA CTACC AGCAG CTCGC CCGTG AAGAG AAAGA GGCAG AACTG GCAGA CGACA TGGAA**
 181 **AAAGG CCTGC CCCAG CACCT GTTTG AATCG CTATG CATCG ATCAT TTGCA ACGCC ACGGG**
 241 **GCCAG CAAAA AATCC ATTAC CCGTG CGTTT GATGA CGATG TTGAG TTTCA GGAGC GCATG**
 301 **GCAGA ACACA TCCGG TACAT GTTG AAACC ATTGC TCACC ACCAG GTTGA TATTG ATTCA**
 361 **GAGGT ACTCG AGCAC CACCA CCACC ACCAC TGA**

Abbildung A.10.: Transkriptions-translations-regulatorische und kodierende DNA-Sequenzen des Plasmids pBADmod1-linker2-gamS (Addgene Datenbank, Plasmid # 45833), das von Vincent Noireaux (*School of Physics and Astronomy, University of Minnesota*) zur Verfügung gestellt und mit Richard Murray kreiert wurde. Cyan: pBAD-Promotor. Orange: für GamS kodierendes Gen.

A.3. Ergänzendes Material zu Kap. 5

A.3.1. DNA-Sequenzen

A.3.1.1. Relevante transkriptions-translations-regulatorische und kodierende Elemente des PCR-Templates zur Synthese von Teilfragment II für die Herstellung des unmethylierten *pap*^(1,1)-DNA-Konstrukts

```

1 GGTAT ATCTC CTTCT TAAAG TTAAA CAAAA TTATT GCTAG CTGCA ACCAT TATGC GAAAA
61 GAAAG TCCGT ATGTC AACAT TTAGA CGATC TTTTA TGCTG TAAAT TCAAT TTGCC ATGAT
121 GTTTT TATCT GAGTA CCCTC TTGCT ATTAG TGTTT TGTC TAGTT TAATT TTGTT TTGTG
181 GGTTA AAAGA TCGTT TAAAT CTTGA CATAc AACAT AAAAA ACTAA ATAAT GGTTG CAGGA
241 TCCAA TAATT TTGTT TAACT TTAAG AAGGA GATAT ACCAT GGAGC TTTTC ACTGG CGTTC
301 TTCCC ATCCT GGTCG AGCTG GACCG CGACC TAAAC GGCCA CAAGT TCAGC GTGTC CGGCC
361 AGGGC GAGGG CGATG CCACC TACGG CAAGC TGACC CTGAA GTTCA TCTGC ACCAC CGGCA
421 AGCTG CCCGT GCCCT GCGCC ACCCT CGTGA CCACC CTGAC CTACC GCGTC CAGTC CTTCA
481 GCGGC TACCC CGACC ACATG AAGCA GCACC ACTTC TTCAA GTCCG CCATG CCCGA AGGCT
541 ACGTC CAGGA GCGCA CCATC TTCTT CAAGC ACGAC GGCAA CTACA AGACC CGCGC CGAGG
601 TGAAG TTCGA GGGCG ACACC CTGGT GAACC GCATC GAGCT GAAGG GCATC GACTT CAAGG
661 AGGAC GGCAA CATCC TGGGG CACAA GCTGG AGTAC AACTA CAACA GCCAC AACGT CTATA
721 TCATG GCCGA CAAGC AGAAG AACGG CATCA AGGTG AACTT CAAGA TCCGC CACAA CATCG
781 AGGAC GGCAG CGTGC AGCTC GCCGA CCACT ACCAG CAGAA CACCC CCATC GGCGA CGGCC
841 CCGTG CTGCT GCCCG ACAAC CACTA CCTGA GCACC CAGTC CGCCC TGAGC AAAGA CCCCA
901 ACGAG AAGCG CGATC ACATG GTCCT GCTGG AGTTC GTGAC CGCCG CCGGG ATCTA ACTCG
961 AGCAA AGCCC GCCGA AAGGC GGGCT TTTCT GTGTC GACCG ATGCC CTTGA GAGGG TTTTC
1021 TGGCT GATGG TG

```

Abbildung A.11.: Relevante transkriptions-translations-regulatorische und kodierende Elemente des PCR-Templates zur Synthese von Teilfragment II für die Herstellung des unmethylierten *pap*^(1,1)-DNA-Konstrukts, das durch Dam-Methylierung (vgl. Abschn. 5.2.2) zum *pap*^(1,1)-DNA-Konstrukt wird. Braun: UTR1. Orange: -10-Element. Blau: -35-Element. Grau: Lrp-Bindestellen 4-6 und 1-3 (nach [313]). Grün: für deGFP kodierendes Gen. Die Sequenz des *pap*-regulatorischen Bereichs (49-230) entspricht der nativen Sequenz (GenBank ID: X14471.1, zurückgehend auf [59]) bis auf die in Abb. 5.3 angegebenen Mutationen. Details zu den transkriptions-translations-regulatorischen Elementen und zum *degfp*-Gen können Abschn. 4.2.9.1 und 2.5 entnommen werden.

A.3.1.2. DNA-Konstrukt $pap^{(1,1)}$, $pap^{(0,1)}$ und $pap^{(1,0)}$

```

1 CGCTA TGAAT TCCTC GAGTT ACTTC TACAG CTCGT CCATG CCGCC GGTGG AGTGG CGGCC
61 CTCGG CCGGT TCGTA CTGTT CCACC ATGGT GTAGT CCTCG TTGTG GGAGG TGATG TCCAA
121 CTTGA TGTTC ACGTT GTAGG CGCCG GGCAG CTGCA CGGGC TTCTT GGCCT TGTAG GTGGT
181 CTTGA CCTCA GCGTC GTAGT GGCCG CCGTC CTTC ACTTC AGCCT CTGCT TGATC TCGCC
241 CTTC A GGGCG CCGTC CTCGG GGTAC ATCCG CTCGG AGGAG GCCTC CCAGC CCATG GTCTT
301 CTTC T GCATT ACGGG GCCGT CGGAG GGGAA GTTGG TGCCG CGCAG CTTC A CCTTG TAGAT
361 GAACT CGCCG TCCTG CAGGG AGGAG TCCTG GGTC A CCGTC ACCAC GCCGC CGTCC TCGAA
421 GTTCA TCACG CGCTC CCACT TGAAG CCTC GGGGA AGGAC AGCTT CAAGT AGTCG GGGAT
481 GTCGG CGGGG TGCTT CACGT AGGCC TTGGA GCCGT ACATG AACTG AGGGG ACAGG ATGTC
541 CCAGG CGAAG GGCAG GGGGC CACCC TTGGT CACCT TCAGC TTGGC GGTCT GGGTG CCCTC
601 GTAGG GGGCG CCTC GCCCT CGCC TCGAT CTCGA ACTCG TGGCC GTTC A CCGAG CCCTC
661 CATGT GCACC TTGAA GCGCA TGAAC TCCTT GATGA TGGCC ATGTT ATCCT CTCG C CCTT
721 GTC A CCATG GTATA TCTCC TTCTT AAAGT TAAAC AAAAT TATTG CTAGC TGCAA CCATT
781 ATGCG AAAAG AAAGT CCGTA TGTCA ACATT TAGAC GATCT TTTAT GCTGT AAATT CAATT
841 TGCCA TGATG TTTTT ATCTG AGTAC CCTCT TGCTA TTAGT GTTTT GTTCT AGTTT AATTT
901 TGTTT TGTGG GTTAA AAGAT CGTTT AAATC TTGAC ATACA ACATA AAAAA CTAAA TAATG
961 GTTGC AGGAT CCAAT AATTT TGTTT AACTT TAAGA AGGAG ATATA CCATG GAGCT TTTCA
1021 CTGGC GTTGT TCCCA TCCTG GTCGA GCTGG ACGGC GACGT AAACG GCCAC AAGTT CAGCG
1081 TGTCC GGGGA GGGCG AGGGC GATGC CACCT ACGGC AAGCT GACCC TGAAG TTCAT CTGCA
1141 CCACC GGCAA GCTGC CCGTG CCTG GCCA CCTC GTGAC CACCC TGACC TACGG CGTGC
1201 AGTGC TTCAG CCGCT ACCCC GACCA CATGA AGCAG CACGA CTTC TCAAG TCCGC CATGC
1261 CCGAA GGCTA CGTCC AGGAG CGCAG CATCT TCTTC AAGGA CGACC GCAAC TACAA GACCC
1321 GCGCC GAGGT GAAGT TCGAG GCGCA CACCC TGGTG AACCG CATCC AGCTG AAGGG CATCC
1381 ACTTC AAGGA GGACG GCAAC ATCCT GGGGC ACAAG CTGGA GTACA ACTAC AACAG CCACA
1441 ACGTC TATAT CATGG CCGAC AAGCA GAAGA ACGGC ATCAA GGTGA ACTTC AAGAT CCGCC
1501 ACAAC ATCGA GGACG GCAGC GTGCA GCTCG CCGAC CACTA CCAGC AGAAC ACCCC CATCC
1561 GCGAC GGCC CGTGC TGCTG CCCGA CAACC ACTAC CTGAG CACCC AGTCC GCCCT GAGCA
1621 AAGAC CCCAA CGAGA AGCGC GATCA CATGG TCCTG CTGGA GTTCC TGACC GCCCG CCGGA
1681 TCTAA AAGCT TCCAT GG

```

Abbildung A.12.: DNA-Sequenz des $pap^{(1,1)}$ -Konstrukts. Die Sequenzen des $pap^{(0,1)}$ - bzw. des $pap^{(1,0)}$ -Konstrukts unterscheiden sich von dieser Sequenz nur durch die C-Mutation nach [317] an der Position 817 bzw. 919. Rot (19–729): für mCherry kodierendes Gen. Braun (732–764): UTR1. Orange (778–782): -10-Element. Grau (783–800): Lrp-Bindestelle 4. Türkis (801–806): -35-Element. Grau (813–830): Lrp-Bindestelle 5. Grau (844–861): Lrp-Bindestelle 6. Grau (876–893): Lrp-Bindestelle 6. Grau (907–924): Lrp-Bindestelle 2. Türkis (931–936): -35-Element. Grau (937–954): Lrp-Bindestelle 3. Orange (955–959): -10-Element. Braun (973–1005): UTR1. Grün (1008–1685): für deGFP kodierendes Gen. GATC^{dist.}: 816–819. GATC^{prox.}: 918–921. Weitere GATC-Stellen: 232–235, 628–631, 968–971, 1493–1496, 1641–1644, 1679–1682. Die Sequenz des pap -regulatorischen Bereichs (778–959) entspricht der nativen Sequenz (GenBank ID: X14471.1, zurückgehend auf [59]) bis auf die in Abb. 5.3 angegebenen Mutationen. Details zu den transkriptions-translations-regulatorischen Elementen sowie dem *degfp*- und *mcherry*-Gen können den Abschn. 4.2.9.1 und 2.5 entnommen werden. Die Markierung der Lrp-Bindestellen erfolgt gemäß [313].

A.3.1.3. pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500

```

1  TGAGC TAACA CCGTG CGTGT TGACA ATTTT ACCTC TGGCG GTGAT AATGG TTGCA GCTAG
61 CAATA ATTTT GTTTA ACTTT AAGAA GGAGA TATAC CATGG TAGAT AGCAA GAAGC GCCCT
121 GGCAA AGATC TCGAC CGTAT CGATC GTAAC ATTCT TAATG AGTTG CAAAA GGATG GCGCT
181 ATTTC TAACG TCGAG CTTTC TAAAC GTGTG GGACT TTCCC CAACG CCGTG CCTTG AGCGT
241 GTGCG TCGGC TGGAA AGACA AGGGT TTATT CAGGG CTATA CGGCG CTGCT TAACC CCCAT
301 TATCT GGATG CATCA CTTCT GGTAT TCGTT GAGAT TACTC TGAAT CGTGG CGCAC CGGAT
361 GTGTT TGAAC AATTC AATAC CGCTG TACAA AAACT TGAAG AAATT CAGGA GTGTC ATTTA
421 GTATC CGGTG ATTTC GACTA CCTGT TGAAA ACACG CGTGC CGGAT ATGTC AGCCT ACCGT
481 AAGTT GCTGG GGGAA ACCCT GCTGC GTCTG CCTGG CGTCA ATGAC ACACG GACAT ACGTT
541 GTTAT GGAAG AAGTC AAGCA GAGTA ATCGT CTGGT TATTA AGACG CGCTA ACTCG AGCAA
601 AGCCC GCCGA AAGGC GGGCT TTTCT GT

```

Abbildung A.13.: Transkriptions-translations-regulatorische und kodierende DNA-Sequenzen des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500. Cyan: OR2-OR1-Pr. Orange: UTR1. Gelb: für Lrp kodierendes Gen (GenBank ID: 949051). Grau: T500. Details zum Plasmid und zu den transkriptions-translations-regulatorischen Elementen können Abschn. 4.2.9.1 entnommen werden.

A.3.1.4. pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500

```

1  TGAGC TAACA CCGTG CGTGT TGACA ATTTT ACCTC TGGCG GTGAT AATGG TTGCA GCTAG
61 CAATA ATTTT GTTTA ACTTT AAGAA GGAGG ATCCA AATGA GTGAA TATAT GAAGA ATGAA
121 ATACT GGAAT TTCTG AACAG GCATG ATGGA GGGAA AACGG CAGAA ATTGC AGAGG CGCTG
181 GCGGT AACGG ATTAC CAGGC CCGCT ATTAC CTGTT ATTAC TGGAA AAAGC AGGTA TGGTT
241 CAGCG CTCAC CATT AAGAC GGGCA TGGCA ACATA CTGGT TTCTG AAAGG AGAGA AGCAG
301 GCGGG GCAGA GCTGT TCTTC CACAA CTTAA CTCGA GAGTC GACCA AAGCC CGCCG AAAGG
361 CGGGC TTTTC TGT

```

Abbildung A.14.: Transkriptions-translations-regulatorische und kodierende DNA-Sequenzen des Plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500. Cyan: OR2-OR1-Pr. Orange: UTR1. Gelb: für PapI kodierendes Gen (GenBank ID: 1039535). Grau: T500. Details zum Plasmid und zu den transkriptions-translations-regulatorischen Elementen können Abschn. 4.2.9.1 entnommen werden.

A.3.2. Kinetik-Messungen zur Analyse der epigenetischen Regulation des *pap*-regulatorischen Bereichs ohne zellfreie Lrp- und PapI-Synthese

Der zu dem in Abschn. 5.3.1 präsentierte Experiment ergänzende Versuch diente ebenfalls der Analyse der epigenetischen Regulation des *pap*-regulatorischen Bereichs, ohne dabei zusätzlich die Koregulatoren Lrp oder PapI zellfrei zu synthetisieren. In gleicher Weise wie im in Abschn. 5.3.1 präsentierte Experiment wurden drei zellfreie Reaktionen präpariert, die sich in ihrer Zusammensetzung beim Methylierungsmuster der hinzugegebenen *pap*-regulatorischen DNA unterschieden. Den Reaktionen (0,1), (1,0) und (1,1) wurde das *pap*-DNA-Konstrukt mit den Methylierungszuständen (0,1), (1,0) bzw. (1,1) hinzugefügt (5 nM). Zusätzlich wurde eine Referenzreaktion (x) hergestellt, die die gleiche Zusammensetzung wie die anderen drei Reaktionen besaß, aber keine *pap*-regulatorische DNA beinhaltete. Jedoch wurden bei diesen Experimenten die zellfreien Reaktionen im Mikroplattenleser für 16 h bei einer Temperatur von 29 °C inkubiert. Die Messung der Fluoreszenz-Intensität im deGFP- und mCherry-Kanal erfolgte kontinuierlich während der Inkubation (vgl. Abschn. 5.2.3). Abb. A.15 zeigt die gemessenen Kinetik-Fluoreszenz-Intensitäten dieser zellfreien Reaktionen im deGFP- (A) und mCherry-Kanal (B).

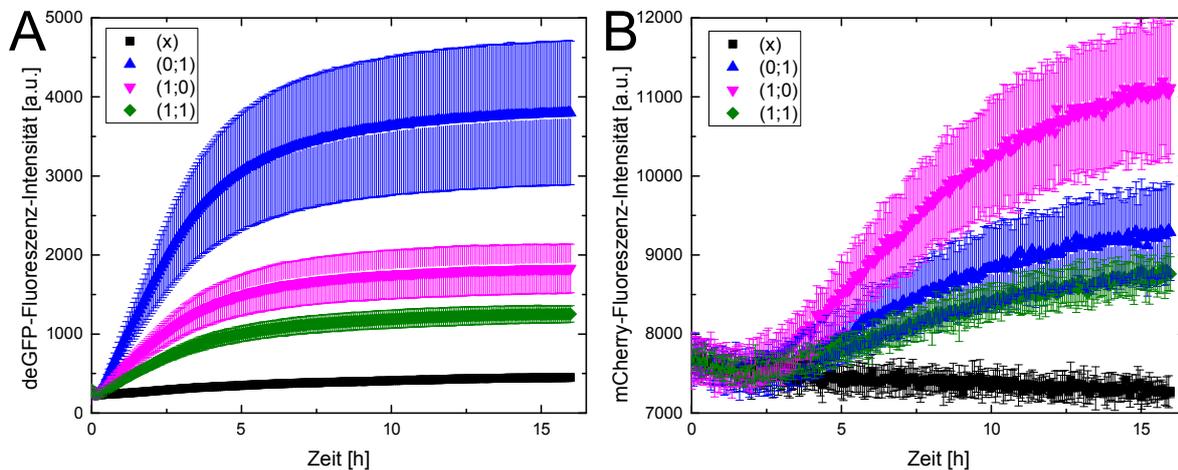


Abbildung A.15.: deGFP- (A) und mCherry-Kinetik-Fluoreszenz-Intensitäten (B) während 16 h Inkubation bei 29 °C in Abhängigkeit des Methylierungszustandes des hinzupipettierten *pap*-DNA-Konstrukts. Der Referenz-Reaktion (x) wurde kein *pap*-DNA-Konstrukt zur Verfügung gestellt. Den Reaktionen vom Typ (1,1), (0,1) und (1,0) wurde das *pap*-DNA-Konstrukt mit den Methylierungszuständen (1,1), (0,1), (1,0) hinzugegeben. Die Fluoreszenz-Intensitäten sind die Mittelwerte aus fünf Messungen. Die Fehlerbalken entsprechen der einfachen Standardabweichung.

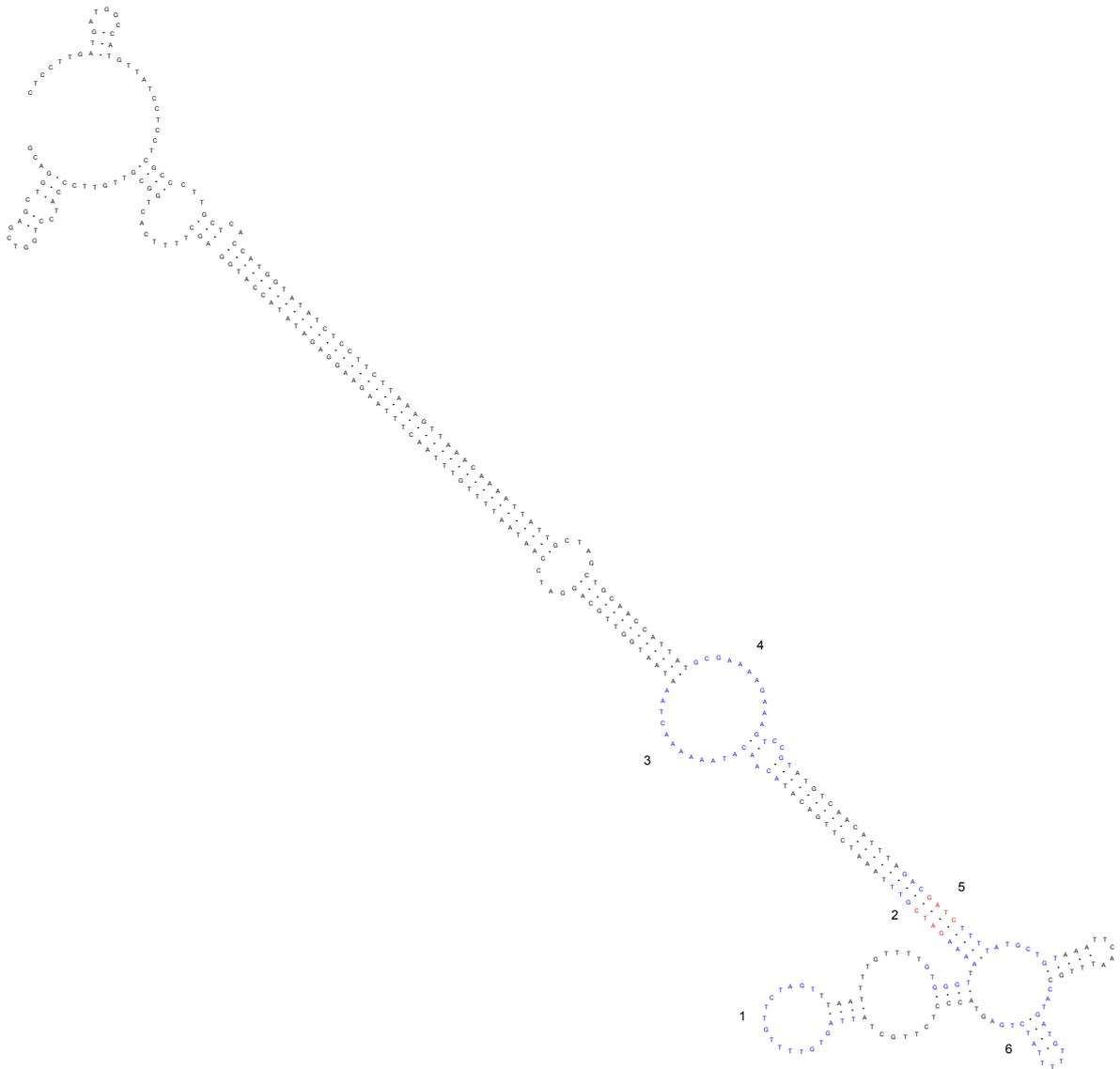
A.3.3. Weitere Sekundärstrukturen der synthetischen *pap*-DNA

Abbildung A.16.: Eine weitere von insgesamt drei (vgl. Abb. 5.11 und A.17) vom *mfold web server* [334] (Standardeinstellungen, 29 °C, s. Abschn. 5.2.4) vorhergesagte Sekundärstruktur der synthetischen *pap*-DNA. Die sechs Lrp-Bindestellen sind blau, GATC^{prox} und GATC^{dist} rot markiert. Die Abbildung wurde mit dem Programm *Rna Viz* [335] hergestellt.

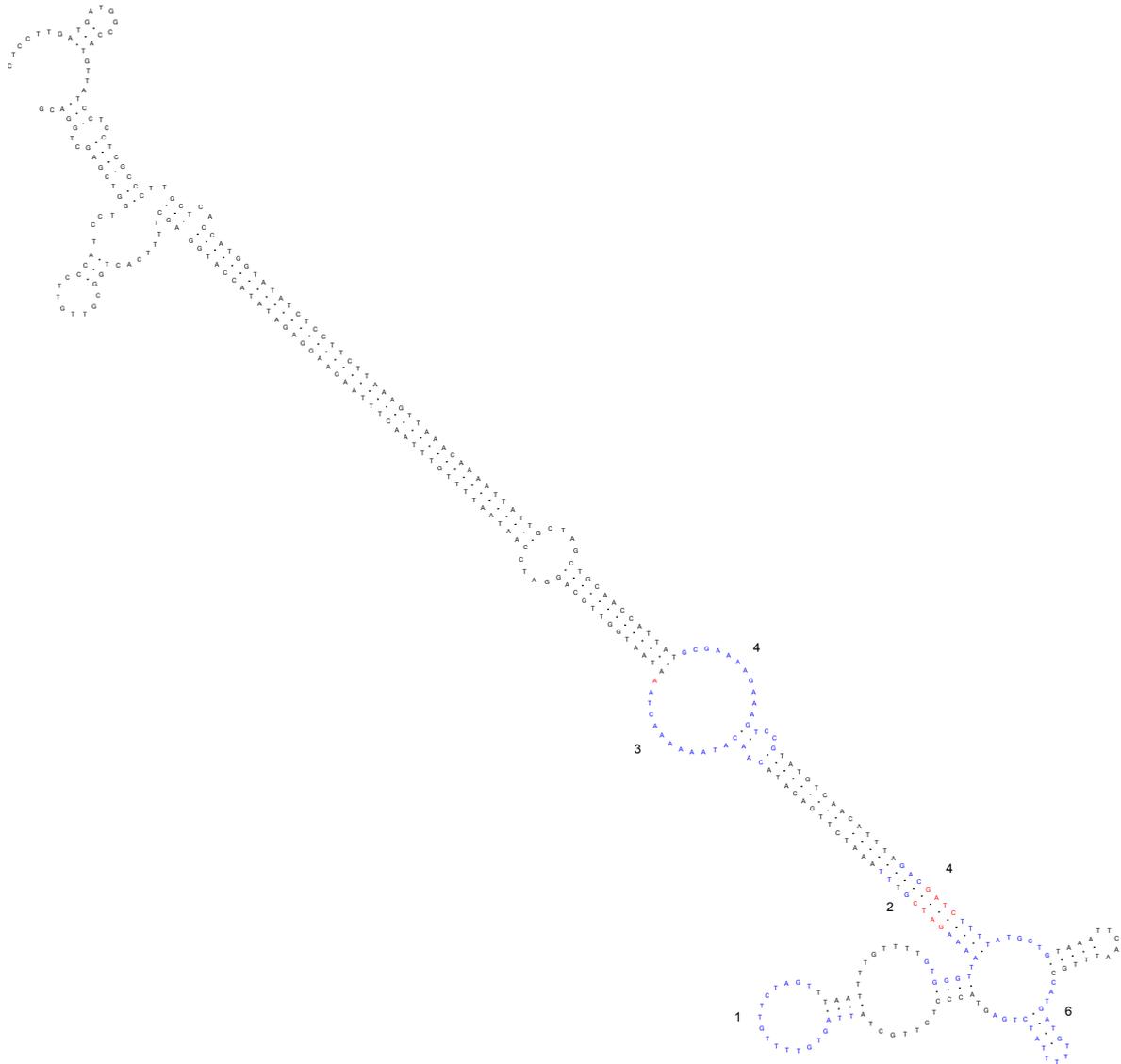


Abbildung A.17.: Eine weitere von insgesamt drei (vgl. Abb. 5.11 und A.16) vom *mfold web server* [334] (Standardeinstellungen, 29 °C, s. Abschn. 5.2.4) vorhergesagte Sekundärstruktur der synthetischen *pap*-DNA. Die sechs Lrp-Bindestellen sind blau, GATC^{prox} und GATC^{dist} rot markiert. Die Abbildung wurde mit dem Programm *Rna Viz* [335] hergestellt.

Literatur

- [1] C. Darwin, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, 1. Aufl. London: J. Murray, 1859.
- [2] S. L. Miller, A Production of Amino Acids Under Possible Primitive Earth Conditions, *Science*, Bd. **117**, Nr. 3046, S. 528–529, 1953.
- [3] E. Wollrab et al., Chemical Analysis of a “Miller-Type” Complex Prebiotic Broth. Part I: Chemical Diversity, Oxygen and Nitrogen Based Polymers, *Origins of Life and Evolution of Biospheres*, Bd. **46**, Nr. 2-3, S. 149–169, 2016.
- [4] S. Scherer et al., Chemical Analysis of a “Miller-Type” Complex Prebiotic Broth. Part II: Gas, Oil, Water and the Oil/Water-Interface, *Origins of Life and Evolution of Biospheres*, Bd. **47**, Nr. 4, S. 381–403, 2017.
- [5] C. R. Woese und G. E. Fox, The Concept of Cellular Evolution, *Journal of Molecular Evolution*, Bd. **10**, Nr. 1, S. 1–6, 1977.
- [6] C. R. Woese, The universal ancestor, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **95**, Nr. 12, S. 6854–6859, 1998.
- [7] C. R. Woese, On the evolution of cells, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **99**, Nr. 13, S. 8742–8747, 2002.
- [8] C. R. Woese, „Evolving biological organization“, in *Microbial Evolution: Concepts and Controversies*, J. Sapp, Hrsg., Oxford, New York: Oxford University Press, 2005, Kap. 4, S. 99–118, ISBN: 0195168771.
- [9] J.-M. Lehn, From supramolecular chemistry towards constitutional dynamic chemistry and adaptive chemistry, *Chemical Society Reviews*, Bd. **36**, Nr. 2, S. 151–160, 2007.
- [10] H. Arnoldt, S. H. Strogatz und M. Timme, Toward the Darwinian transition: Switching between distributed and speciated states in a simple model of early life, *Physical Review E*, Bd. **92**, Nr. 5, S. 052 909, 2015.
- [11] F. J. Odling-Smee, K. N. Laland und M. W. Feldman, Niche Construction, *The American Naturalist*, Bd. **147**, Nr. 4, S. 641–648, 1996.
- [12] K. N. Laland, F. J. Odling-Smee und M. W. Feldman, Evolutionary consequences of niche construction and their implications for ecology. *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **96**, Nr. 18, S. 10 242–10 247, 1999.

- [13] C. Salzberg und H. Sayama, „Heredity, Complexity, and Surprise: Embedded Self-Replication and Evolution in CA“, in *Cellular Automata*, P. M. A. Sloot, B. Chopard und A. G. Hoekstra, Hrsg., Berlin, Heidelberg: Springer Verlag, 2004, S. 161–171, ISBN: 9783540304791.
- [14] A. J. Bissette und S. P. Fletcher, Mechanisms of Autocatalysis, *Angewandte Chemie International Edition*, Bd. **52**, Nr. 49, S. 12 800–12 826, 2013.
- [15] L. E. Orgel, Unnatural Selection in Chemical Systems, *Accounts of Chemical Research*, Bd. **28**, Nr. 3, S. 109–118, 1995.
- [16] M. Eigen, *From Strange Simplicity to Complex Familiarity: A Treatise on Matter, Information, Life and Thought*, 1. Aufl. Oxford: Oxford University Press, 2013, ISBN: 9780198570219.
- [17] D. R. Mills, R. L. Peterson und S. Spiegelman, An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **58**, Nr. 1, S. 217–224, 1967.
- [18] R. Saffhill, H. Schneider-Bernloehr, L. Orgel und S. Spiegelman, In vitro selection of bacteriophage Q β ribonucleic acid variants resistant to ethidium bromide, *Journal of Molecular Biology*, Bd. **51**, Nr. 3, S. 531–539, 1970.
- [19] M. Eigen, Selforganization of Matter and the Evolution of Biological Macromolecules, *Die Naturwissenschaften*, Bd. **58**, Nr. 10, S. 465–523, 1971.
- [20] M. Eigen und P. Schuster, The Hypercycle. A Principle of Natural Self-Organization. Part A: Emergence of the Hypercycle. *Die Naturwissenschaften*, Bd. **64**, Nr. 11, S. 541–565, 1977.
- [21] M. Eigen und P. Schuster, The Hypercycle. A Principle of Natural Self-Organisation. Part B: The Abstract Hypercycle, *Naturwissenschaften*, Bd. **65**, Nr. 1, S. 7–41, 1978.
- [22] M. Eigen und P. Schuster, The Hypercycle. A Principle of Natural Self-Organisation. Part C: The Realistic Hypercycle, *Die Naturwissenschaften*, Bd. **65**, Nr. 7, S. 341–369, 1978.
- [23] N. G. Walter und G. Strunk, Strand displacement amplification as an in vitro model for rolling-circle replication: deletion formation and evolution during serial transfer, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **91**, Nr. 17, S. 7937–7941, 1994.
- [24] R. R. Breaker und G. F. Joyce, Emergence of a replicating species from an in vitro RNA evolution reaction, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **91**, Nr. 13, S. 6093–6097, 1994.
- [25] T. A. Lincoln und G. F. Joyce, Self-Sustained Replication of an RNA Enzyme, *Science*, Bd. **323**, Nr. 5918, S. 1229–1232, 2009.
- [26] S. F. Elena und R. E. Lenski, Microbial genetics: Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation, *Nature Reviews Genetics*, Bd. **4**, Nr. 6, S. 457–469, 2003.

- [27] Z. D. Blount, C. Z. Borland und R. E. Lenski, Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **105**, Nr. 23, S. 7899–7906, 2008.
- [28] M. Le Gac, J. Plucain, T. Hindré, R. E. Lenski und D. Schneider, Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with *Escherichia coli*, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **109**, Nr. 24, S. 9487–9492, 2012.
- [29] C. Adami, C. Ofria und T. C. Collier, Evolution of biological complexity. *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **97**, Nr. 9, S. 4463–4468, 2000.
- [30] N. Guttenberg und N. Goldenfeld, Cascade of Complexity in Evolving Predator-Prey Dynamics, *Physical Review Letters*, Bd. **100**, Nr. 5, S. 058 102, 2008.
- [31] C. Abröhl et al., *Biochemie - Zellbiologie*, K. Munk, Hrsg. Stuttgart: Georg Thieme Verlag, 2008, ISBN: 9783131448316.
- [32] A. Böck et al., Selenocysteine: the 21st amino acid, *Molecular Microbiology*, Bd. **5**, Nr. 3, S. 515–520, 1991.
- [33] G. Srinivasan, C. M. James und J. A. Krzycki, Pyrrolysine Encoded by UAG in Archaea: Charging of a UAG-Decoding Specialized tRNA, *Science*, Bd. **296**, Nr. 5572, S. 1459–1462, 2002.
- [34] B. Alberts et al., „Wie Proteine arbeiten“, in *Lehrbuch der Molekularen Zellbiologie*, J. Graw, Hrsg., 4. Aufl., Weinheim: Wiley-VCH Verlag, 2012, Kap. 4.2, S. 151–161, ISBN: 3527328246.
- [35] L. Wang, J. Xie und P. G. Schultz, Expanding The Genetic Code, *Annual Review of Biophysics and Biomolecular Structure*, Bd. **35**, Nr. 1, S. 225–249, 2006.
- [36] N. Budisa, Prolegomena to Future Experimental Efforts on Genetic Code Engineering by Expanding Its Amino Acid Repertoire, *Angewandte Chemie International Edition*, Bd. **43**, Nr. 47, S. 6426–6463, 2004.
- [37] C. T. T. Wong, C. L. Tung und X. Li, Synthetic cysteine surrogates used in native chemical ligation, *Molecular BioSystems*, Bd. **9**, Nr. 5, S. 826–833, 2013.
- [38] R. Behrendt, P. White und J. Offer, Advances in Fmoc solid-phase peptide synthesis. *Journal of Peptide Science*, Bd. **22**, Nr. 1, S. 4–27, 2016.
- [39] H. R. Kalbitzer, „Chemische Peptidsynthese“, in *Löffler/Petrides Biochemie und Pathobiochemie*, P. C. Heinrich, M. Müller und L. Graeve, Hrsg., 9. Aufl., Berlin, Heidelberg: Springer-Verlag, 2014, Kap. 6.6.1, S. 99–100, ISBN: 3642179711.
- [40] M. H. Richmond, The effect of amino acid analogues on growth and protein synthesis in microorganisms, *Bacteriological Reviews*, Bd. **26**, Nr. 4, S. 398–420, 1962.
- [41] C. F. Schachtele und P. Rogers, Canavanine death in *Escherichia coli*, *Journal of Molecular Biology*, Bd. **14**, Nr. 2, S. 474–489, 1965.

- [42] D. Bell, The Effect of Canavanine on Herpes Simplex Virus Replication, *Journal of General Virology*, Bd. **22**, Nr. 3, S. 319–330, 1974.
- [43] J. Akaogi et al., Role of non-protein amino acid L-canavanine in autoimmunity, *Autoimmunity reviews*, Bd. **5**, Nr. 6, S. 429–435, 2006.
- [44] P. F. Kruse, P. B. White, H. A. Carter und T. A. McCoy, Incorporation of Canavanine into Protein of Walker Carcinosarcoma 256 Cells Cultured in Vitro, *Cancer Research*, Bd. **19**, Nr. 1, S. 122–125, 1959.
- [45] M. H. Green und J. F. Ward, Enhancement of Human Tumor Cell Killing by L-Canavanine in Combination with γ -Radiation, *Cancer Research*, Bd. **43**, Nr. 9, S. 4180–4182, 1983.
- [46] D. S. Swaffar, C. Y. Ang, P. B. Desai und G. A. Rosenthal, Inhibition of the Growth of Human Pancreatic Cancer Cells by the Arginine Antimetabolite L-Canavanine, *Cancer Research*, Bd. **54**, Nr. 23, S. 6045–6048, 1994.
- [47] A. K. Bence, D. R. Worthen, V. R. Adams und P. A. Crooks, The antiproliferative and immunotoxic effects of L-canavanine and L-canaline, *Anti-Cancer Drugs*, Bd. **13**, Nr. 3, S. 313–320, 2002.
- [48] A. K. Bence und P. A. Crooks, The Mechanism of L-Canavanine Cytotoxicity: Arginyl tRNA Synthetase as a Novel Target for Anticancer Drug Discovery, *Journal of Enzyme Inhibition and Medicinal Chemistry*, Bd. **18**, Nr. 5, S. 383–394, 2003.
- [49] B. Vynnytska-Myronovska, Y. Bobak, Y. Garbe, C. Dittfeld, O. Stasyk und L. A. Kunz-Schughart, Single amino acid arginine starvation efficiently sensitizes cancer cells to canavanine treatment and irradiation, *International Journal of Cancer*, Bd. **130**, Nr. 9, S. 2164–2175, 2012.
- [50] Y. Ishida, J.-H. Park, L. Mao, Y. Yamaguchi und M. Inouye, Replacement of All Arginine Residues with Canavanine in MazF-bs mRNA Interferase Changes Its Specificity, *The Journal of Biological Chemistry*, Bd. **288**, Nr. 11, S. 7564–7571, 2013.
- [51] J. Shin und V. Noireaux, Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70, *Journal of Biological Engineering*, Bd. **4**, Artikel 8, 2010.
- [52] Z. Z. Sun, C. A. Hayes, J. Shin, F. Caschera, R. M. Murray und V. Noireaux, Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. *Journal of Visualized Experiments*, Nr. 79, e50762, 2013.
- [53] G. Felsenfeld, A Brief History of Epigenetics, *Cold Spring Harbor Perspectives in Biology*, Bd. **6**, Nr. 1, a018200, 2014.
- [54] J. Graw, „Epigenetik“, in *Genetik*, 6. Aufl., Berlin, Heidelberg: Springer-Verlag, 2015, Kap. 8, S. 291–338, ISBN: 9783662448175.
- [55] D. E. Gottschling, „Epigenetics: From Phenomenon to Field“, in *Epigenetics*, C. D. Allis, T. Jenuwein, D. Reinberg und M.-L. Caparros, Hrsg., 1. Aufl., New York: Cold Spring Harbor Laboratory Press, 2007, Kap. 1, S. 2, ISBN: 0879697245.

- [56] C. Russell und M. Mulvey, „Type 1 and P Pili of Uropathogenic *Escherichia coli*“, in *Bacterial Pili: Structure, Synthesis and Role in Disease*, M. A. Barocchi und J. L. Telford, Hrsg., Wallingford, Oxfordshire, Boston: CAB International, 2014, Kap. 4, S. 49–70, ISBN: 9781780642550.
- [57] G. Phan und G. Waksman, „Pilus Biogenesis by the Chaperone-Usher Pathway“, in *Bacterial Pili: Structure, Synthesis and Role in Disease*, M. A. Barocchi und J. L. Telford, Hrsg., Wallingford, Oxfordshire, Boston: CAB International, 2014, Kap. 3, S. 36–48, ISBN: 9781780642550.
- [58] A. L. Flores-Mireles, J. N. Walker, M. Caparon und S. J. Hultgren, Urinary tract infections: epidemiology, mechanisms of infection and treatment options, *Nature Reviews Microbiology*, Bd. **13**, Nr. 5, S. 269–284, 2015.
- [59] L. B. Blyn, B. A. Braaten, C. A. White-Ziegler, D. H. Rolfson und D. A. Low, Phase-variation of pyelonephritis-associated pili in *Escherichia coli*: evidence for transcriptional regulation, *The EMBO Journal*, Bd. **8**, Nr. 2, S. 613–620, 1989.
- [60] M. Schenkelberger, „Cooperative Biomolecular Binding: High specificity of competitive single stranded DNA hybridization, influence of DNA methylation on the duplex stability, and epigenetic regulation of in vitro gene switches“, Dissertation, Universität des Saarlandes, 2014.
- [61] J. Swartz, Developing cell-free biology for industrial applications, *Journal of Industrial Microbiology & Biotechnology*, Bd. **33**, Nr. 7, S. 476–485, 2006.
- [62] D. Jahn et al., *Taschenlehrbuch Biologie: Genetik*, 1. Aufl., K. Munk, Hrsg. Stuttgart: Georg Thieme Verlag, 2010, ISBN: 9783131537515.
- [63] M. T. Medigan und J. M. Martinko, *Brock Mikrobiologie*, 11. Aufl. München: Pearson Studium, 2006, ISBN: 3827371872.
- [64] J. D. Watson und F. H. C. Crick, Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid, *Nature*, Bd. **171**, Nr. 4356, S. 737–738, 1953.
- [65] T. Hinze und M. Sturm, Hrsg., *Rechnen mit DNA: Eine Einführung in Theorie und Praxis*. München: Oldenbourg Wissenschaftsverlag, 2004, ISBN: 3486275305.
- [66] W. Baltes und R. Matissek, „Essenzielle Aminosäuren, Proteinwertigkeit“, in *Lebensmittelchemie*, 7. Aufl., Berlin, Heidelberg: Springer-Verlag, 2011, Kap. 8.2, S. 170, ISBN: 9783642165399.
- [67] F. H. C. Crick, On protein synthesis, *Symposia of the Society for Experimental Biology*, Bd. **12**, S. 138–163, 1958.
- [68] F. H. C. Crick, Central dogma of molecular biology, *Nature*, Bd. **227**, Nr. 5258, S. 561–563, 1970.
- [69] C. Müllhardt, *Der Experimentator Molekularbiologie / Genomics*, 6. Aufl. Heidelberg: Spektrum Akademischer Verlag, 2009, ISBN: 9783827420367.

- [70] H. Lodish, S. L. Zipursky, P. Matsudaira, D. Baltimore und J. Darnell, „DNA Cloning with Plasmid Vectors“, in *Molecular Cell Biology*, 4. Aufl., New York: W. H. Freeman, 2000, Kap. 7.1, ISBN: 0716731363.
- [71] H. Lodish, S. L. Zipursky, P. Matsudaira, D. Baltimore und J. Darnell, „Bacterial Transcription Initiation“, in *Molecular Cell Biology*, 4. Aufl., New York: W. H. Freeman, 2000, Kap. 10.2, ISBN: 0716731363.
- [72] R. Merkl und S. Waack, „Konsensus-Sequenzen“, in *Bioinformatik Interaktiv: Algorithmen und Praxis*, 2. Aufl., Weinheim: Wiley-VCH Verlag, 2009, Kap. 10.7, S. 189–190, ISBN: 3527325948.
- [73] K. Watanabe, N. Hayashi, A. Oyama, K. Nishikawa, T. Ueda und K. Miura, Unusual anticodon loop structure found in E.coli lysine tRNA. *Nucleic Acids Research*, Bd. **22**, Nr. 1, S. 79–87, 1994.
- [74] J. M. Berg, J. L. Tymoczko und L. Stryer, „Aminoacyl-Transfer RNA Synthetases Read the Genetic Code“, in *Biochemistry*, 5. Aufl., W. H. Freeman, 2002, Kap. 29.2, ISBN: 0716746840.
- [75] G. Löffler, „Proteinbiosynthese, Proteinmodifizierung und Proteinabbau“, in *Biochemie und Pathobiochemie*, 5. Aufl., Berlin, Heidelberg: Springer-Verlag, 1997, Kap. 11, S. 273, ISBN: 9783662060636.
- [76] R. Bingel-Erlenmeyer et al., A peptide deformylase-ribosome complex reveals mechanism of nascent chain processing, *Nature*, Bd. **452**, Nr. 7183, S. 108–111, 2008.
- [77] F. Sherman, J. W. Stewart und S. Tsunasawa, Methionine or Not Methionine at the Beginning of a Protein, *Bioessays*, Bd. **3**, Nr. 1, S. 27–31, 1985.
- [78] P.-H. Hirel, J.-M. Schmitter, P. Dessen, G. Fayat und S. Blanquet, Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the side-chain length of the penultimate amino acid, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **86**, Nr. 21, S. 8247–8251, 1989.
- [79] F. Frottin et al., The Proteomics of N-terminal Methionine Cleavage, *Molecular & Cellular Proteomics*, Bd. **5**, Nr. 12, S. 2336–2349, 2006.
- [80] C. H. Waddington, The genetic control of wing development in Drosophila, *Journal of Genetics*, Bd. **41**, S. 75–139, 1940.
- [81] C. H. Waddington, Epigenetics and evolution, *Symposia of the Society for Experimental Biology*, Bd. **7**, S. 186–199, 1953.
- [82] C. H. Waddington, *The Strategy of the Genes. A Discussion of Some Aspects of Theoretical Biology*. London: Allen & Unwin, 1957.
- [83] M. G. Marinus und A. Løbner-Olesen, DNA Methylation, *Ecosal Plus*, Bd. **6**, Nr. 1, 2014.
- [84] C. End, C. Walczuch und M. Buntru, Zellfreie Proteinexpression für Forschung und Produktion, *BIOspektrum*, Bd. **20**, Nr. 1, S. 70–72, 2014.

- [85] A. S. Spirin, „Cell-Free Protein Synthesis“, in *Cell-Free Translation Systems*, A. S. Spirin, Hrsg., 1. Aufl., Berlin, Heidelberg, New York: Springer-Verlag, 2002, Kap. 1, S. 3–20, ISBN: 9783642639562.
- [86] S. Chong, Overview of Cell-Free Protein Synthesis: Historic Landmarks, Commercial Systems, and Expanding Applications, *Current Protocols in Molecular Biology*, Bd. **108**, Nr. 1, S. 16.30.1–16.30.11, 2014.
- [87] T. Winnick, F. Friedberg und D. M. Greenberg, The utilization of labeled glycine in the process of amino acid incorporation by the protein of liver homogenate, *The Journal of Biological Chemistry*, Bd. **175**, Nr. 1, S. 117–126, 1948.
- [88] H. Borsook, C. L. Deasy, A. J. Haagen-Smit, G. Keighley und P. H. Lowy, The incorporation of labeled lysine into the proteins of guinea pig liver homogenate, *The Journal of Biological Chemistry*, Bd. **179**, Nr. 2, S. 689–704, 1949.
- [89] P. Siekevitz, Uptake of radioactive alanine in vitro into the proteins of rat liver fractions, *The Journal of Biological Chemistry*, Bd. **195**, Nr. 2, S. 549–565, 1952.
- [90] E. F. Gale und J. P. Folkes, Effect of Nucleic Acids on Protein Synthesis and Amino-Acid Incorporation in Disrupted Staphylococcal Cells, *Nature*, Bd. **173**, Nr. 4417, S. 1223–1227, 1954.
- [91] M. W. Nirenberg und J. H. Matthaei, The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **47**, Nr. 10, S. 1588–1602, 1961.
- [92] M. W. Nirenberg, „Cell-free protein synthesis directed by messenger RNA“, in *Methods in Enzymology*, N. O. Kaplan und S. P. Colowick, Hrsg., 1. Aufl., Academic Press, 1963, Kap. 3, S. 17–23, ISBN: 9780121818067.
- [93] D. V. Liu, J. F. Zawada und J. R. Swartz, Streamlining *Escherichia Coli* S30 Extract Preparation for Economical Cell-Free Protein Synthesis, *Biotechnology Progress*, Bd. **21**, Nr. 2, S. 460–465, 2005.
- [94] J. W. Davies und P. Kaesberg, Translation of Virus mRNA: Synthesis of Bacteriophage Q β Proteins in a Cell-Free Extract from Wheat Embryo, *Journal of Virology*, Bd. **12**, Nr. 6, S. 1434–1441, 1973.
- [95] D. Efron und A. Marcus, Efficient synthesis of rabbit globin in a cell-free system from wheat embryo, *FEBS Letters*, Bd. **33**, Nr. 1, S. 23–27, 1973.
- [96] A. Marcus, B. Luginbill und J. Feeley, Polysome formation with tobacco mosaic virus RNA, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **59**, Nr. 4, S. 1243–1250, 1968.
- [97] B. E. Roberts und B. M. Paterson, Efficient Translation of Tobacco Mosaic Virus RNA and Rabbit Globin 9S RNA in a Cell-Free System from Commercial Wheat Germ, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **70**, Nr. 8, S. 2330–2334, 1973.

- [98] H. R. B. Pelham und R. J. Jackson, An Efficient mRNA-Dependent Translation System from Reticulocyte Lysates, *European Journal of Biochemistry*, Bd. **67**, Nr. 1, S. 247–256, 1976.
- [99] J. H. Matthaei und M. W. Nirenberg, Characteristics and stabilization of DNAase-sensitive protein synthesis in *E. coli* extracts, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **47**, Nr. 10, S. 1580–1588, 1961.
- [100] W. B. Wood und P. Berg, The effect of enzymatically synthesized ribonucleic acid on amino acid incorporation by a soluble protein-ribosome system from *Escherichia coli*, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **48**, Nr. 1, S. 94–104, 1962.
- [101] R. Byrne, J. G. Levin, H. A. Bladen und M. W. Nirenberg, The in Vitro Formation of a DNA-Ribosome Complex, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **52**, Nr. 1, S. 140–148, 1964.
- [102] M. Lederman und G. Zubay, DNA-directed peptide synthesis. I. A comparison of T2 and *Escherichia coli* DNA-directed peptide synthesis in two cell-free systems, *Biochimica et Biophysica Acta*, Bd. **149**, Nr. 1, S. 253–258, 1967.
- [103] G. Zubay, In Vitro Synthesis of Protein in Microbial Systems, *Annual Review of Genetics*, Bd. **7**, Nr. 1, S. 267–287, 1973.
- [104] R. Zaniewski, E. Petkaitis und M. P. Deutscher, A multiple mutant of *Escherichia coli* lacking the exoribonucleases RNase II, RNase D, and RNase BN, *Journal of Biological Chemistry*, Bd. **259**, Nr. 19, S. 11 651–11 653, 1984.
- [105] D. E. Nevin und J. M. Pratt, A coupled in vitro transcription-translation system for the exclusive synthesis of polypeptides expressed from the T7 promoter, *FEBS Letters*, Bd. **291**, Nr. 2, S. 259–263, 1991.
- [106] J. Garamella, R. Marshall, M. Rustad und V. Noireaux, The All *E. coli* TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology, *ACS Synthetic Biology*, Bd. **5**, Nr. 4, S. 344–355, 2016.
- [107] D. Craig, M. T. Howell, C. L. Gibbs, T. Hunt und R. J. Jackson, Plasmid cDNA-directed protein synthesis in a coupled eukaryotic in vitro transcription-translation system, *Nucleic Acids Research*, Bd. **20**, Nr. 19, S. 4987–4995, 1992.
- [108] A. Zemella, L. Thoring, C. Hoffmeister und S. Kubick, Cell-Free Protein Synthesis: Pros and Cons of Prokaryotic and Eukaryotic Systems, *ChemBioChem*, Bd. **16**, Nr. 17, S. 2420–2431, 2015.
- [109] E. D. Carlson, R. Gan, C. E. Hodgman und M. C. Jewett, Cell-free protein synthesis: Applications come of age, *Biotechnology Advances*, Bd. **30**, Nr. 5, S. 1185–1194, 2012.
- [110] J. F. Zawada et al., Microscale to manufacturing scale-up of cell-free cytokine production—a new approach for shortening protein production development timelines, *Biotechnology and Bioengineering*, Bd. **108**, Nr. 7, S. 1570–1578, 2011.

- [111] S. H. Hong, Y.-C. Kwon und M. C. Jewett, Non-standard amino acid incorporation into proteins using *Escherichia coli* cell-free protein synthesis, *Frontiers in Chemistry*, Bd. **2**, Artikel 34, 2014.
- [112] P. Shrestha, T. M. Holland und B. C. Bundy, Streamlined extract preparation for *Escherichia coli*-based cell-free protein synthesis by sonication or bead vortex mixing, *BioTechniques*, Bd. **53**, Nr. 3, S. 163–174, 2012.
- [113] Y.-C. Kwon und M. C. Jewett, High-throughput preparation methods of crude extract for robust cell-free protein synthesis, *Scientific Reports*, Bd. **5**, Artikel 8663, 2015.
- [114] D.-M. Kim und J. Swartz, Prolonging Cell-Free Protein Synthesis by Selective Reagent Additions, *Biotechnology Progress*, Bd. **16**, Nr. 3, S. 385–390, 2000.
- [115] D.-M. Kim und J. R. Swartz, Regeneration of Adenosine Triphosphate from Glycolytic Intermediates for Cell-Free Protein Synthesis, *Biotechnology and Bioengineering*, Bd. **74**, Nr. 4, S. 309–316, 2001.
- [116] K. A. Calhoun und J. R. Swartz, Energizing Cell-Free Protein Synthesis With Glucose Metabolism, *Biotechnology and Bioengineering*, Bd. **90**, Nr. 5, S. 606–613, 2005.
- [117] T.-W. Kim, J.-W. Keum, I.-S. Oh, C.-Y. Choi, H.-C. Kim und D.-M. Kim, An economical and highly productive cell-free protein synthesis system utilizing fructose-1,6-bisphosphate as an energy source, *Journal of Biotechnology*, Bd. **130**, Nr. 4, S. 389–393, 2007.
- [118] K. Sitaraman, D. Esposito, G. Klarmann, S. F. Le Grice, J. L. Hartley und D. K. Chatterjee, A novel cell-free protein synthesis system, *Journal of Biotechnology*, Bd. **110**, Nr. 3, S. 257–263, 2004.
- [119] L. Ryabova, L. Vinokurov, E. Shekhovtsova, Y. Alakhov und A. Spirin, Acetyl Phosphate as an Energy Source for Bacterial Cell-Free Translation Systems, *Analytical Biochemistry*, Bd. **226**, Nr. 1, S. 184–186, 1995.
- [120] T. Kigawa et al., Cell-free production and stable-isotope labeling of milligram quantities of proteins, *FEBS Letters*, Bd. **442**, Nr. 1, S. 15–19, 1999.
- [121] D.-M. Kim und J. R. Swartz, Prolonging cell-free protein synthesis with a novel ATP regeneration system, *Biotechnology and Bioengineering*, Bd. **66**, Nr. 3, S. 180–188, 1999.
- [122] T.-W. Kim, D.-M. Kim und C.-Y. Choi, Rapid production of milligram quantities of proteins in a batch cell-free protein synthesis system, *Journal of Biotechnology*, Bd. **124**, Nr. 2, S. 373–380, 2006.
- [123] T. Kigawa et al., Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression, *Journal of Structural and Functional Genomics*, Bd. **5**, Nr. 1-2, S. 63–68, 2004.
- [124] R.-G. Kim und C.-Y. Choi, Expression-independent consumption of substrates in cell-free expression system from *Escherichia coli*, *Journal of Biotechnology*, Bd. **84**, Nr. 1, S. 27–32, 2000.

- [125] D.-M. Kim und J. R. Swartz, Oxalate improves protein synthesis by enhancing ATP supply in a cell-free system derived from *Escherichia coli*, *Biotechnology Letters*, Bd. **22**, Nr. 19, S. 1537–1542, 2000.
- [126] T.-W. Kim et al., Prolonged Cell-Free Protein Synthesis Using Dual Energy Sources: Combined Use of Creatine Phosphate and Glucose for the Efficient Supply of ATP and Retarded Accumulation of Phosphate, *Biotechnology and Bioengineering*, Bd. **97**, Nr. 6, S. 1510–1515, 2007.
- [127] F. Caschera und V. Noireaux, Synthesis of 2.3 mg/ml of protein with an all *Escherichia coli* cell-free transcription-translation system, *Biochimie*, Bd. **99**, S. 162–168, 2014.
- [128] M. C. Jewett und J. R. Swartz, Mimicking the *Escherichia coli* Cytoplasmic Environment Activates Long-Lived and Efficient Cell-Free Protein Synthesis, *Biotechnology and Bioengineering*, Bd. **86**, Nr. 1, S. 19–26, 2004.
- [129] M. C. Jewett und J. R. Swartz, Substrate Replenishment Extends Protein Synthesis With an In Vitro Translation System Designed to Mimic the Cytoplasm, *Biotechnology and Bioengineering*, Bd. **87**, Nr. 4, S. 465–472, 2004.
- [130] M. C. Jewett, K. A. Calhoun, A. Voloshin, J. J. Wu und J. R. Swartz, An integrated cell-free metabolic platform for protein production and synthetic biology, *Molecular Systems Biology*, Bd. **4**, Artikel 220, 2008.
- [131] N. Michel-Reydellet, K. Calhoun und J. Swartz, Amino acid stabilization for cell-free protein synthesis by modification of the *Escherichia coli* genome, *Metabolic Engineering*, Bd. **6**, Nr. 3, S. 197–203, 2004.
- [132] K. A. Calhoun und J. R. Swartz, Total amino acid stabilization during cell-free protein synthesis reactions, *Journal of Biotechnology*, Bd. **123**, Nr. 2, S. 193–203, 2006.
- [133] M. S. Dillingham und S. C. Kowalczykowski, RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks, *Microbiology and Molecular Biology Reviews*, Bd. **72**, Nr. 4, S. 642–671, 2008.
- [134] K. C. Murphy, The λ Gam Protein Inhibits RecBCD Binding to dsDNA Ends, *Journal of Molecular Biology*, Bd. **371**, Nr. 1, S. 19–24, 2007.
- [135] N. Michel-Reydellet, K. Woodrow und J. Swartz, Increasing PCR Fragment Stability and Protein Yields in a Cell-Free System with Genetically Modified *Escherichia coli* Extracts, *Journal of Molecular Microbiology and Biotechnology*, Bd. **9**, Nr. 1, S. 26–34, 2005.
- [136] A. Spirin, V. Baranov, L. Ryabova, S. Ovodov und Y. Alakhov, A Continuous Cell-Free Translation System Capable of Producing Polypeptides in High Yield, *Science*, Bd. **242**, Nr. 4882, S. 1162–1164, 1988.
- [137] V. I. Baranov, I. Y. Morozov, S. A. Ortlepp und A. S. Spirin, Gene expression in a cell-free system on the preparative scale, *Gene*, Bd. **84**, Nr. 2, S. 463–466, 1989.

- [138] D.-M. Kim und C.-Y. Choi, A Semicontinuous Prokaryotic Coupled Transcription/Translation System Using a Dialysis Membrane, *Biotechnology Progress*, Bd. **12**, Nr. 5, S. 645–649, 1996.
- [139] Y. Shimizu et al., Cell-free translation reconstituted with purified components, *Nature Biotechnology*, Bd. **19**, Nr. 8, S. 751–755, 2001.
- [140] Y. Shimizu, T. Kanamori und T. Ueda, Protein synthesis by pure translation systems, *Methods*, Bd. **36**, Nr. 3, S. 299–304, 2005.
- [141] H. F. Kung, B. Redfield, B. V. Treadwell, B. Eskin, C. Spears und H. Weissbach, DNA-directed in Vitro Synthesis of beta-Galactosidase: Studies With Purified Factors, *The Journal of Biological Chemistry*, Bd. **252**, Nr. 19, S. 6889–6894, 1977.
- [142] M. C. Jewett und J. R. Swartz, Mimicking the Escherichia coli Cytoplasmic Environment Activates Long-Lived and Efficient Cell-Free Protein Synthesis, *Biotechnology and Bioengineering*, Bd. **86**, Nr. 1, S. 19–26, 2004.
- [143] K. A. Calhoun und J. R. Swartz, An Economical Method for Cell-Free Protein Synthesis using Glucose and Nucleoside Monophosphates, *Biotechnology Progress*, Bd. **21**, Nr. 4, S. 1146–1153, 2005.
- [144] T.-W. Kim, J.-W. Keum, I.-S. Oh, C.-Y. Choi, C.-G. Park und D.-M. Kim, Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system, *Journal of Biotechnology*, Bd. **126**, Nr. 4, S. 554–561, 2006.
- [145] O. Shimomura, F. H. Johnson und Y. Saiga, Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea, *Journal of Cellular and Comparative Physiology*, Bd. **59**, Nr. 3, S. 223–239, 1962.
- [146] D. M. Chudakov, M. V. Matz, S. Lukyanov und K. A. Lukyanov, Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues, *Physiological Reviews*, Bd. **90**, Nr. 3, S. 1103–1163, 2010.
- [147] M. Zimmer, Green Fluorescent Protein (GFP): Applications, Structure, and Related Photophysical Behavior, *Chemical Reviews*, Bd. **102**, Nr. 3, S. 759–782, 2002.
- [148] B. P. Cormack, R. H. Valdivia und S. Falkow, FACS-optimized mutants of the green fluorescent protein (GFP), *Gene*, Bd. **173**, Nr. 1, S. 33–38, 1996.
- [149] R. Heim, A. B. Cubitt und R. Y. Tsien, Improved green fluorescence, *Nature*, Bd. **373**, Nr. 6516, S. 663–664, 1995.
- [150] G. H. Patterson, S. M. Knobel, W. D. Sharif, S. R. Kain und D. W. Piston, Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy, *Biophysical Journal*, Bd. **73**, Nr. 5, S. 2782–2790, 1997.
- [151] X. Li, G. Zhang, N. Ngo, X. Zhao, S. R. Kain und C.-C. Huang, Deletions of the Aequorea victoria Green Fluorescent Protein Define the Minimal Domain Required for Fluorescence, *Journal of Biological Chemistry*, Bd. **272**, Nr. 45, S. 28 545–28 549, 1997.

- [152] J. Guo, C. E. Melançon, H. S. Lee, D. Groff und P. G. Schultz, Evolution of Amber Suppressor tRNAs for Efficient Bacterial Production of Proteins Containing Nonnatural Amino Acids, *Angewandte Chemie International Edition*, Bd. **48**, Nr. 48, S. 9148–9151, 2009.
- [153] M. G. Hoesl, „Expanding the Toolkit of Protein Engineering: Towards Multiple Simultaneous in vivo Incorporation of Noncanonical Amino Acids“, Dissertation, Technische Universität München, 2011.
- [154] M. P. Exner, „Incorporation of novel noncanonical amino acids in model proteins using rational and evolved variants of Methanosarcina mazei pyrrolysyl-tRNA synthetase“, Dissertation, Technische Universität Berlin, 2016.
- [155] M. V. Matz et al., Fluorescent proteins from nonbioluminescent Anthozoa species, *Nature Biotechnology*, Bd. **17**, Nr. 10, S. 969–973, 1999.
- [156] B. J. Bevis und B. S. Glick, Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed), *Nature Biotechnology*, Bd. **20**, Nr. 1, S. 83–87, 2002.
- [157] N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. G. Giepmans, A. E. Palmer und R. Y. Tsien, Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein, *Nature Biotechnology*, Bd. **22**, Nr. 12, S. 1567–1572, 2004.
- [158] E. G. Worst, P. Zimmer, E. Wollrab, K. Kruse und A. Ott, Unbounded growth patterns of reproducing, competing polymers-similarities to biological evolution, *New Journal of Physics*, Bd. **18**, Nr. 10, Artikel 103003, 2016.
- [159] P. Zimmer, „Ueber die Evolution der Struktur sich autokatalytisch vermehrender Molekuele“, Dissertation, Universität des Saarlandes, 2017.
- [160] G. von Kiedrowski, A Self-Replicating Hexadeoxynucleotide, *Angewandte Chemie International Edition*, Bd. **25**, Nr. 10, S. 932–935, 1986.
- [161] W. S. Zielinski und L. E. Orgel, Autocatalytic synthesis of a tetranucleotide analogue, *Nature*, Bd. **327**, Nr. 6120, S. 346–347, 1987.
- [162] G. von Kiedrowski und E. Szathmáry, Selection versus Coexistence of Parabolic Replicators Spreading on Surfaces, *Selection*, Bd. **1**, Nr. 1-3, S. 173–180, 2001.
- [163] C. Fernando, G. von Kiedrowski und E. Szathmáry, A Stochastic Model of Nonenzymatic Nucleic Acid Replication: “Elongators” Sequester Replicators, *Journal of Molecular Evolution*, Bd. **64**, Nr. 5, S. 572–585, 2007.
- [164] E. J. Hayden, G. von Kiedrowski und N. Lehman, Systems Chemistry on Ribozyme Self-Construction: Evidence for Anabolic Autocatalysis in a Recombination Network, *Angewandte Chemie International Edition*, Bd. **47**, Nr. 44, S. 8424–8428, 2008.
- [165] J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, J. J.-P. Peyralans und S. Otto, Mechanosensitive Self-Replication Driven by Self-Organization, *Science*, Bd. **327**, Nr. 5972, S. 1502–1506, 2010.

- [166] C. B. Mast und D. Braun, Thermal Trap for DNA Replication, *Physical Review Letters*, Bd. **104**, Nr. 18, S. 188–192, 2010.
- [167] H. Krammer, F. M. Möller und D. Braun, Thermal, Autonomous Replicator Made from Transfer RNA, *Physical Review Letters*, Bd. **108**, Nr. 23, S. 238–241, 2012.
- [168] J. W. Sadownik, E. Mattia, P. Nowak und S. Otto, Diversification of self-replicating molecules, *Nature Chemistry*, Bd. **8**, Nr. 3, S. 264–269, 2016.
- [169] J. W. Szostak, The eightfold path to non-enzymatic RNA replication, *Journal of Systems Chemistry*, Bd. **3**, Nr. 1, Artikel 2, 2012.
- [170] F. Olasagasti, H. J. Kim, N. Pourmand und D. W. Deamer, Non-enzymatic transfer of sequence information under plausible prebiotic conditions, *Biochimie*, Bd. **93**, Nr. 3, S. 556–561, 2011.
- [171] C. Deck, M. Jauker und C. Richert, Efficient enzyme-free copying of all four nucleobases templated by immobilized RNA, *Nature Chemistry*, Bd. **3**, Nr. 8, S. 603–608, 2011.
- [172] B. J. Cafferty und N. V. Hud, Abiotic synthesis of RNA in water: a common goal of prebiotic chemistry and bottom-up synthetic biology, *Current Opinion in Chemical Biology*, Bd. **22**, S. 146–157, 2014.
- [173] G. von Kiedrowski, „Minimal Replicator Theory I: Parabolic Versus Exponential Growth“, in *Bioorganic Chemistry Frontiers*, H. Dugas und F. P. Schmidtchen, Hrsg., Berlin, Heidelberg: Springer Verlag, 1993, S. 113–146, ISBN: 9783642781124.
- [174] F. Barany, The ligase chain reaction in a PCR world, *PCR Methods and Applications*, Bd. **1**, Nr. 1, S. 5–16, 1991.
- [175] F. Barany, Genetic disease detection and DNA amplification using cloned thermostable ligase, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **88**, Nr. 1, S. 189–193, 1991.
- [176] K. Leu et al., Cascade of Reduced Speed and Accuracy after Errors in Enzyme-Free Copying of Nucleic Acid Sequences, *Journal of the American Chemical Society*, Bd. **135**, Nr. 1, S. 354–366, 2013.
- [177] T. Li und K. C. Nicolaou, Chemical self-replication of palindromic duplex DNA, *Nature*, Bd. **369**, Nr. 6477, S. 218–221, 1994.
- [178] G. F. Joyce, Amplification, mutation and selection of catalytic RNA, *Gene*, Bd. **82**, Nr. 1, S. 83–87, 1989.
- [179] C. Tuerk und L. Gold, Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science*, Bd. **249**, Nr. 4968, S. 505–510, 1990.
- [180] A. D. Ellington und J. W. Szostak, In vitro selection of RNA molecules that bind specific ligands, *Nature*, Bd. **346**, Nr. 6287, S. 818–822, 1990.

- [181] J. Derr et al., Prebiotically plausible mechanisms increase compositional diversity of nucleic acid sequences, *Nucleic Acids Research*, Bd. **40**, Nr. 10, S. 4711–4722, 2012.
- [182] T. J. Collins, ImageJ for microscopy, *BioTechniques*, Bd. **43**, Nr. 1 Supplement, S25–S30, 2007.
- [183] C. A. Schneider, W. S. Rasband und K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, *Nature Methods*, Bd. **9**, Nr. 7, S. 671–675, 2012.
- [184] R. A. Dimitrov und M. Zuker, Prediction of Hybridization and Melting for Double-Stranded Nucleic Acids, *Biophysical Journal*, Bd. **87**, Nr. 1, S. 215–226, 2004.
- [185] N. R. Markham und M. Zuker, DINAMelt web server for nucleic acid melting prediction, *Nucleic Acids Research*, Bd. **33**, Nr. Web Server Issue, W577–W581, 2005.
- [186] N. R. Markham und M. Zuker, „UNAFold: Software for Nucleic Acid Folding and Hybridization“, in *Bioinformatics: Structure, Function and Applications*, J. M. Keith, Hrsg., Bd. 453, Totowa: Humana Press, 2008, Kap. 1, S. 3–31, ISBN: 9781603274296.
- [187] M. Takahashi, E. Yamaguchi und T. Uchida, Thermophilic DNA Ligase. Purification and Properties of the Enzyme from *Thermus Thermophilus* HB8, *The Journal of Biological Chemistry*, Bd. **259**, Nr. 16, S. 10041–10047, 1984.
- [188] D. Kowalski, W. D. Kroecker und M. Laskowski, Mung bean nuclease I. 6. Physical, chemical, and catalytic properties, *Biochemistry*, Bd. **15**, Nr. 20, S. 4457–4463, 1976.
- [189] T. F. McCutchan, J. L. Hansen, J. B. Dame und J. A. Mullins, Mung bean nuclease cleaves *Plasmodium* genomic DNA at sites before and after genes. *Science*, Bd. **225**, Nr. 4662, S. 625–628, 1984.
- [190] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *The Journal of Physical Chemistry*, Bd. **81**, Nr. 25, S. 2340–2361, 1977.
- [191] W. Krauth, *Statistical Mechanics: Algorithms and Computations*, 1. Aufl. Oxford, New York: Oxford University Press, 2006, ISBN: 0198515367.
- [192] O. Tange, GNU Parallel: The Command-Line Power Tool, *login: The USENIX Magazine*, Bd. **36**, Nr. 1, S. 42–47, 2011.
- [193] E. G. Worst et al., Cell-free expression with the toxic amino acid canavanine, *Bioorganic & Medicinal Chemistry Letters*, Bd. **25**, Nr. 17, S. 3658–3660, 2015.
- [194] E. G. Worst et al., Residue-specific Incorporation of Noncanonical Amino Acids into Model Proteins Using an *Escherichia coli* Cell-free Transcription-translation System, *Journal of Visualized Experiments*, Nr. 114, e54327, 2016.
- [195] R. B. Merrifield, Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide, *Journal of the American Chemical Society*, Bd. **85**, Nr. 14, S. 2149–2154, 1963.
- [196] S. L. Pedersen et al., Microwave heating in solid-phase peptide synthesis, *Chemical Society Reviews*, Bd. **41**, Nr. 5, S. 1826–1844, 2012.
- [197] J. M. Palomo, Solid-phase peptide synthesis: an overview focused on the preparation of biologically relevant peptides, *RSC Advances*, Bd. **4**, Nr. 62, S. 32658–32672, 2014.

- [198] B. L. Nilsson, M. B. Soellner und R. T. Raines, Chemical Synthesis of Proteins, *Annual Review of Biophysics and Biomolecular Structure*, Bd. **34**, S. 91–118, 2005.
- [199] D. Zwanziger, I. Böhme, D. Lindner und A. G. Beck-Sickinger, First selective agonist of the neuropeptide Y1-receptor with reduced size, *Journal of Peptide Science*, Bd. **15**, Nr. 12, S. 856–866, 2009.
- [200] L. Liu, Hrsg., *Protein Ligation and Total Synthesis I*, 1. Aufl., Ser. Topics in Current Chemistry. Cham, Heidelberg, New York, Dordrecht, London: Springer International Publishing, 2015, Bd. 362, ISBN: 9783319191850.
- [201] L. Liu, Hrsg., *Protein Ligation and Total Synthesis II*, 1. Aufl., Ser. Topics in Current Chemistry. Cham, Heidelberg, New York, Dordrecht, London: Springer International Publishing, 2015, Bd. 363, ISBN: 9783319191881.
- [202] S. Pöggeler, „Heterologe Genexpression“, in *Praktikum der Molekulargenetik*, U. Kück, Hrsg., Berlin, Heidelberg, New York: Springer-Verlag, 2005, Kap. 7, S. 291–322, ISBN: 3540211667.
- [203] C. G. Acevedo-Rocha und N. Budisa, Xenomicrobiology: a roadmap for genetic code engineering. *Microbial Biotechnology*, Bd. **9**, Nr. 5, S. 666–676, 2016.
- [204] N. Budisa et al., Toward the experimental codon reassignment in vivo: protein building with an expanded amino acid repertoire, *FASEB Journal*, Bd. **13**, Nr. 1, S. 41–51, 1999.
- [205] K. E. Beatty und D. A. Tirrell, „Noncanonical Amino Acids in Protein Science and Engineering“, in *Protein Engineering*, C. Koehler und U. L. RajBhandary, Hrsg., 1. Aufl., Berlin, Heidelberg: Springer-Verlag, 2009, Kap. 7, S. 127–153, ISBN: 9783540709374.
- [206] A. Singh-Blom, R. A. Hughes und A. D. Ellington, „Residue-Specific Incorporation of Unnatural Amino Acids into Proteins In Vitro and In Vivo“, in *Enzyme Engineering: Methods and Protocols*, J. C. Samuelson, Hrsg., Totowa: Humana Press, 2013, Kap. 7, S. 93–114, ISBN: 9781627032926.
- [207] R. Munier und G. N. Cohen, Incorporation d’analogues structuraux d’aminoacides dans les protéines bactériennes, *Biochimica et Biophysica Acta*, Bd. **21**, Nr. 3, S. 592–593, 1956.
- [208] J. A. Johnson, Y. Y. Lu, J. A. Van Deventer und D. A. Tirrell, Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications, *Current Opinion in Chemical Biology*, Bd. **14**, Nr. 6, S. 774–780, 2010.
- [209] W. A. Hendrickson, J. R. Horton und D. M. LeMaster, Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure, *The EMBO Journal*, Bd. **9**, Nr. 5, S. 1665–1672, 1990.
- [210] R. Cotterill, „Röntgenbeugungsanalyse und Molekülstruktur“, in *Biophysik: eine Einführung*, 1. Aufl., Weinheim: Wiley-VCH-Verlag, 2008, Kap. 6.1, S. 89–99, ISBN: 9783527406869.

- [211] J. O. Boles et al., Bio-incorporation of telluromethionine into buried residues of dihydrofolate reductase, *Nature Structural Biology*, Bd. **1**, Nr. 5, S. 283–284, 1994.
- [212] N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann und R. Huber, High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*, *European Journal of Biochemistry*, Bd. **230**, Nr. 2, S. 788–796, 1995.
- [213] J. H. Bae et al., Incorporation of β -selenolo[3,2-b]pyrrolyl-alanine into proteins for phase determination in protein X-ray crystallography, *Journal of Molecular Biology*, Bd. **309**, Nr. 4, S. 925–936, 2001.
- [214] M.-P. Strub et al., Selenomethionine and Selenocysteine Double Labeling Strategy for Crystallographic Phasing, *Structure*, Bd. **11**, Nr. 11, S. 1359–1367, 2003.
- [215] F. I. Hinz, D. C. Dieterich, D. A. Tirrell und E. M. Schuman, Noncanonical Amino Acid Labeling in Vivo to Visualize and Affinity Purify Newly Synthesized Proteins in Larval Zebrafish, *ACS Chemical neuroscience*, Bd. **3**, Nr. 1, S. 40–49, 2012.
- [216] D. C. Dieterich, A. J. Link, J. Graumann, D. A. Tirrell und E. M. Schuman, Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **103**, Nr. 25, S. 9482–9487, 2006.
- [217] D. C. Dieterich, J. J. Lee, A. J. Link, J. Graumann, D. A. Tirrell und E. M. Schuman, Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging, *Nature Protocols*, Bd. **2**, Nr. 3, S. 532–540, 2007.
- [218] K. E. Beatty et al., Fluorescence Visualization of Newly Synthesized Proteins in Mammalian Cells, *Angewandte Chemie International Edition*, Bd. **45**, Nr. 44, S. 7364–7367, 2006.
- [219] D. C. Dieterich et al., In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons, *Nature Neuroscience*, Bd. **13**, Nr. 7, S. 897–905, 2010.
- [220] I. Erdmann et al., Cell-selective labelling of proteomes in *Drosophila melanogaster*, *Nature Communications*, Bd. **6**, Artikel 7521, 2015.
- [221] S. I. Presolski, V. P. Hong und M. G. Finn, Copper-Catalyzed Azide-Alkyne Click Chemistry for Bioconjugation, *Current Protocols in Chemical Biology*, Bd. **3**, Nr. 4, S. 153–162, 2011.
- [222] C. G. Acevedo-Rocha, A.-S. Geiermann, N. Budisa und L. Merkel, Design of protein congeners containing β -cyclopropylalanine, *Molecular BioSystems*, Bd. **8**, Nr. 10, S. 2719–2723, 2012.
- [223] C. Minks, S. Alefelder, L. Moroder, R. Huber und N. Budisa, Towards New Protein Engineering: In Vivo Building and Folding of Protein Shuttles for Drug Delivery and Targeting by the Selective Pressure Incorporation (SPI) Method, *Tetrahedron*, Bd. **56**, Nr. 48, S. 9431–9442, 2000.

- [224] „Genetik“, in *Biologie: Ein Lehrbuch*, G. Czihak, H. Langer und H. Ziegler, Hrsg., 4. Aufl., Berlin, Heidelberg: Springer-Verlag, 1990, Kap. 2, S. 197, ISBN: 9783540183488.
- [225] N. Budisa, *Engineering the Genetic Code: Expanding the Amino Acid Repertoire for the Design of Novel Proteins*. Weinheim: Wiley-VCH Verlag, 2006, S. 100–102, ISBN: 3527607099.
- [226] D. Rösner, T. Schneider, D. Schneider, M. Scheffner und A. Marx, Click chemistry for targeted protein ubiquitylation and ubiquitin chain formation, *Nature Protocols*, Bd. **10**, Nr. 10, S. 1594–1611, 2015.
- [227] J. Hyun Bae et al., Expansion of the Genetic Code Enables Design of a Novel “Gold” Class of Green Fluorescent Proteins, *Journal of Molecular Biology*, Bd. **328**, Nr. 5, S. 1071–1081, 2003.
- [228] M. G. Hoesl et al., Lipase Congeners Designed by Genetic Code Engineering, *Chem-CatChem*, Bd. **3**, Nr. 1, S. 213–221, 2011.
- [229] G. A. Rosenthal, The Protective Action of a Higher Plant Toxic Product: Biochemical studies reveal how the allelochemical L-canavanine disrupts insect development, *BioScience*, Bd. **38**, Nr. 2, S. 104–109, 1988.
- [230] G. A. Rosenthal und D. L. Dahlman, Incorporation of L-Canavanine into Proteins and the Expression of Its Antimetabolic Effects, *Journal of Agricultural and Food Chemistry*, Bd. **39**, Nr. 5, S. 987–990, 1991.
- [231] G. A. Rosenthal, The Biological Effects and Mode of Action of L-Canavanine, a Structural Analogue of L-Arginine, *Quarterly Review of Biology*, Bd. **52**, Nr. 2, S. 155–178, 1977.
- [232] K. Ozawa et al., Translational incorporation of L-3,4-dihydroxyphenylalanine into proteins, *FEBS Journal*, Bd. **272**, Nr. 12, S. 3162–3171, 2005.
- [233] D. J. Stigers et al., Incorporation of chlorinated analogues of aliphatic amino acids during cell-free protein synthesis, *Chemical Communications*, Bd. **47**, Nr. 6, S. 1839–1841, 2011.
- [234] I. N. Arthur et al., In Situ Deprotection and Incorporation of Unnatural Amino Acids during Cell-Free Protein Synthesis, *Chemistry - A European Journal*, Bd. **19**, Nr. 21, S. 6824–6830, 2013.
- [235] S.-J. Oh, K.-H. Lee, H.-C. Kim, C. Catherine, H. Yun und D.-M. Kim, Translational Incorporation of Multiple Unnatural Amino Acids in a Cell-free Protein Synthesis System, *Biotechnology and Bioprocess Engineering*, Bd. **19**, Nr. 3, S. 426–432, 2014.
- [236] A. Singh-Blom, R. A. Hughes und A. D. Ellington, An amino acid depleted cell-free protein synthesis system for the incorporation of non-canonical amino acid analogs into proteins, *Journal of Biotechnology*, Bd. **178**, Nr. 1, S. 12–22, 2014.

- [237] M. C. T. Hartman, K. Josephson und J. W. Szostak, Enzymatic aminoacylation of tRNA with unnatural amino acids, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **103**, Nr. 12, S. 4356–4361, 2006.
- [238] K. Josephson, M. C. T. Hartman und J. W. Szostak, Ribosomal Synthesis of Unnatural Peptides, *Journal of the American Chemical Society*, Bd. **127**, Nr. 33, S. 11 727–11 735, 2005.
- [239] M. C. T. Hartman, K. Josephson, C.-W. Lin und J. W. Szostak, An Expanded Set of Amino Acid Analogs for the Ribosomal Translation of Unnatural Peptides, *PLoS ONE*, Bd. **2**, Nr. 10, e972, 2007.
- [240] A. O. Subtelny, M. C. T. Hartman und J. W. Szostak, Ribosomal Synthesis of N-Methyl Peptides, *Journal of the American Chemical Society*, Bd. **130**, Nr. 19, S. 6131–6136, 2008.
- [241] Y. V. G. Schlippe, M. C. T. Hartman, K. Josephson und J. W. Szostak, In Vitro Selection of Highly Modified Cyclic Peptides That Act as Tight Binding Inhibitors, *Journal of the American Chemical Society*, Bd. **134**, Nr. 25, S. 10 469–10 477, 2012.
- [242] K. Josephson, A. Ricardo und J. W. Szostak, mRNA display: from basic principles to macrocycle drug discovery, *Drug Discovery Today*, Bd. **19**, Nr. 4, S. 388–399, 2014.
- [243] T. Kawakami, H. Murakami und H. Suga, Messenger RNA-Programmed Incorporation of Multiple N-Methyl-Amino Acids into Linear and Cyclic Peptides, *Chemistry & Biology*, Bd. **15**, Nr. 1, S. 32–42, 2008.
- [244] C. C. Liu und P. G. Schultz, Adding New Chemistries to the Genetic Code, *Annual Review of Biochemistry*, Bd. **79**, Nr. 1, S. 413–444, 2010.
- [245] J. W. Chin, Expanding and Reprogramming the Genetic Code of Cells and Animals, *Annual Review of Biochemistry*, Bd. **83**, Nr. 1, S. 379–408, 2014.
- [246] T. Hohsaka, Y. Ashizuka, H. Taira, H. Murakami und M. Sisido, Incorporation of Nonnatural Amino Acids into Proteins by Using Various Four-Base Codons in an Escherichia coli in Vitro Translation System, *Biochemistry*, Bd. **40**, Nr. 37, S. 11 060–11 064, 2001.
- [247] J. Grünewald et al., Immunochemical termination of self-tolerance, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **105**, Nr. 32, S. 11 276–11 280, 2008.
- [248] J. Grünewald et al., Mechanistic studies of the immunochemical termination of self-tolerance with unnatural amino acids, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **106**, Nr. 11, S. 4337–4342, 2009.
- [249] C. Dückelmann, „Arzneimittel in der Intensivmedizin“, in *Angewandte Arzneimitteltherapie: Klinisch-Pharmazeutische Betreuung in Fallbeispielen*, H. Schneemann, L. Young und M. A. Koda-Kimble, Hrsg., 1. Aufl., Berlin, Heidelberg: Springer Verlag, 2001, Kap. 27, S. 581–613, ISBN: 9783642565052.

- [250] H. Neumann, Rewiring translation - Genetic code expansion and its applications, *FEBS Letters*, Bd. **586**, Nr. 15, S. 2057–2064, 2012.
- [251] L. Wang, A. Brock, B. Herberich und P. G. Schultz, Expanding the Genetic Code of *Escherichia coli*, *Science*, Bd. **292**, Nr. 5516, S. 498–500, 2001.
- [252] C. Polycarpo et al., An aminoacyl-tRNA synthetase that specifically activates pyrrolysine, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **101**, Nr. 34, S. 12 450–12 454, 2004.
- [253] C. R. Polycarpo, S. Herring, A. Bérubé, J. L. Wood, D. Söll und A. Ambrogelly, Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase, *FEBS Letters*, Bd. **580**, Nr. 28-29, S. 6695–6700, 2006.
- [254] T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto und S. Yokoyama, Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode N(epsilon)-(o-Azidobenzoyloxycarbonyl) lysine for Site-Specific Protein Modification, *Chemistry & Biology*, Bd. **15**, Nr. 11, S. 1187–1197, 2008.
- [255] T. Mukai et al., Codon reassignment in the *Escherichia coli* genetic code, *Nucleic Acids Research*, Bd. **38**, Nr. 22, S. 8188–8195, 2010.
- [256] D. B. F. Johnson et al., RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites, *Nature Chemical Biology*, Bd. **7**, Nr. 11, S. 779–786, 2011.
- [257] M. J. Lajoie et al., Genomically Recoded Organisms Expand Biological Functions, *Science*, Bd. **342**, Nr. 6156, S. 357–360, 2013.
- [258] A. R. Goerke und J. R. Swartz, High-Level Cell-Free Synthesis Yields of Proteins Containing Site-Specific Non-Natural Amino Acids, *Biotechnology and Bioengineering*, Bd. **102**, Nr. 2, S. 400–416, 2009.
- [259] B. C. Bundy und J. R. Swartz, Site-Specific Incorporation of p-Propargyloxyphenylalanine in a Cell-Free Environment for Direct Protein-Protein Click Conjugation, *Bioconjugate Chemistry*, Bd. **21**, Nr. 2, S. 255–263, 2010.
- [260] C. Albayrak und J. R. Swartz, Cell-free co-production of an orthogonal transfer RNA activates efficient site-specific non-natural amino acid incorporation, *Nucleic Acids Research*, Bd. **41**, Nr. 11, S. 5949–5963, 2013.
- [261] S. H. Hong, I. Ntai, A. D. Haimovich, N. L. Kelleher, F. J. Isaacs und M. C. Jewett, Cell-free Protein Synthesis from a Release Factor 1 Deficient *Escherichia coli* Activates Efficient and Multiple Site-specific Nonstandard Amino Acid Incorporation, *ACS Synthetic Biology*, Bd. **3**, Nr. 6, S. 398–409, 2014.
- [262] J. Shin und V. Noireaux, An *E. coli* Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells, *ACS Synthetic Biology*, Bd. **1**, Nr. 1, S. 29–41, 2012.
- [263] U. K. Laemmli, Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4, *Nature*, Bd. **227**, Nr. 5259, S. 680–685, 1970.

- [264] Y. Chemla, E. Ozer, O. Schlesinger, V. Noireaux und L. Alfonta, Genetically Expanded Cell-Free Protein Synthesis Using Endogenous Pyrrolysyl Orthogonal Translation System, *Biotechnology and Bioengineering*, Bd. **112**, Nr. 8, S. 1663–1672, 2015.
- [265] G. Bertani, Studies on Lysogenesis. I. The Mode of Phage Liberation by Lysogenic *Escherichia coli*, *Journal of Bacteriology*, Bd. **62**, Nr. 3, S. 293–300, 1951.
- [266] R. Sussman und F. Jacob, Sur un système de répression thermosensible chez le bactériophage λ d'*Escherichia coli*, *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, Bd. **254**, Nr. 1, S. 1517–1519, 1962.
- [267] K. Kleppe, E. Ohtsuka, R. Kleppe, I. Molineux und H. Khorana, Studies on Polynucleotides: XCVI. Repair Replication of Short Synthetic DNA's as Catalyzed by DNA Polymerases, *Journal of Molecular Biology*, Bd. **56**, Nr. 2, S. 341–361, 1971.
- [268] K. B. Mullis und F. A. Faloona, Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction, *Methods in Enzymology*, Bd. **155**, S. 335–350, 1987.
- [269] R. Saiki et al., Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase, *Science*, Bd. **239**, Nr. 4839, S. 487–491, 1988.
- [270] K. B. Mullis, The Unusual Origin of the Polymerase Chain Reaction, *Scientific American*, Bd. **262**, Nr. 4, S. 56–65, 1990.
- [271] J. SantaLucia, A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **95**, Nr. 4, S. 1460–1465, 1998.
- [272] R. Owczarzy et al., Effects of Sodium Ions on DNA Duplex Oligomers: Improved Predictions of Melting Temperatures, *Biochemistry*, Bd. **43**, Nr. 12, S. 3537–3554, 2004.
- [273] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison und H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nature Methods*, Bd. **6**, Nr. 5, S. 343–345, 2009.
- [274] D. G. Gibson, H. O. Smith, C. A. Hutchison, J. C. Venter und C. Merryman, Chemical synthesis of the mouse mitochondrial genome, *Nature Methods*, Bd. **7**, Nr. 11, S. 901–903, 2010.
- [275] H. A. de Boer, L. J. Comstock und M. Vasser, The tac promoter: A functional hybrid derived from the trp and lac promoters. *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **80**, Nr. 1, S. 21–25, 1983.
- [276] P. O. Olins, C. S. Devine, S. H. Rangwala und K. S. Kavka, The T7 phage gene 10 leader RNA, a ribosome-binding site that dramatically enhances the expression of foreign genes in *Escherichia coli*, *Gene*, Bd. **73**, Nr. 1, S. 227–235, 1988.
- [277] M. H. Larson, W. J. Greenleaf, R. Landick und S. M. Block, Applied Force Reveals Mechanistic and Energetic Details of Transcription Termination, *Cell*, Bd. **132**, Nr. 6, S. 971–982, 2008.
- [278] H. C. Birnboim und J. Doly, A rapid alkaline extraction procedure for screening recombinant plasmid DNA, *Nucleic Acids Research*, Bd. **7**, Nr. 6, S. 1513–1523, 1979.

- [279] P. E. Burmeister et al., 2'-Deoxy Purine, 2'-O-Methyl Pyrimidine (dRmY) Aptamers as Candidate Therapeutics, *Oligonucleotides*, Bd. **16**, Nr. 4, S. 337–351, 2006.
- [280] A. D. Friedman, D. Kim und R. Liu, Highly stable aptamers selected from a 2'-fully modified fGmH RNA library for targeting biomaterials, *Biomaterials*, Bd. **36**, S. 110–123, 2015.
- [281] R. S. Al Toma et al., Site-Directed and Global Incorporation of Orthogonal and Isostructural Noncanonical Amino Acids into the Ribosomal Lasso Peptide Capistrin, *ChemBioChem*, Bd. **16**, Nr. 3, S. 503–509, 2015.
- [282] M. M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical Biochemistry*, Bd. **72**, Nr. 1-2, S. 248–254, 1976.
- [283] H. Schägger und G. von Jagow, Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, *Analytical Biochemistry*, Bd. **166**, Nr. 2, S. 368–79, 1987.
- [284] H. Schägger, H. Aquila und G. von Jagow, Coomassie blue-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for direct visualization of polypeptides during electrophoresis, *Analytical Biochemistry*, Bd. **173**, Nr. 1, S. 201–205, 1988.
- [285] J. A. Bornhorst und J. J. Falke, Purification of Proteins Using Polyhistidine Affinity Tags, *Methods in Enzymology*, Bd. **326**, S. 245–254, 2000.
- [286] M. E. Kimple, A. L. Brill und R. L. Pasker, Overview of Affinity Tags for Protein Purification, *Current Protocols in Protein Science*, Bd. **73**, Nr. 1, S. 9.9.1–9.9.23, 2013.
- [287] E. Hochuli, W. Bannwarth, H. Döbeli, R. Gentz und D. Stüber, Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent, *Nature Biotechnology*, Bd. **6**, Nr. 11, S. 1321–1325, 1988.
- [288] E. Hochuli, H. Döbeli und A. Schacher, New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues, *Journal of Chromatography A*, Bd. **411**, S. 177–184, 1987.
- [289] R. Hurst, G. Kobs und T. Johnson, Mass Spectrometric Analysis of MagneHis Purified Proteins, *Promega Corporation Web Site*, aufgerufen am 02.05.2017 unter <http://www.promega.de/resources/pubhub/enotes/mass-spectrometric-analysis-of-magnehis-purified-proteins/>, 2003.
- [290] H. Kihara, J. M. Prescott und E. E. Snell, The bacterial cleavage of canavanine to homoserine and guanidine, *The Journal of Biological Chemistry*, Bd. **217**, Nr. 1, S. 497–503, 1955.
- [291] G. D. Kalyankar, M. Ikawa und E. E. Snell, The Enzymatic Cleavage of Canavanine to Homoserine and Hydroxyguanidine, *The Journal of Biological Chemistry*, Bd. **233**, Nr. 5, S. 1175–1178, 1958.

- [292] A. Royant und M. Noirclerc-Savoie, Stabilizing role of glutamic acid 222 in the structure of Enhanced Green Fluorescent Protein, *Journal of Structural Biology*, Bd. **174**, Nr. 2, S. 385–390, 2011.
- [293] A. De Simone, C. G. Acevedo-Rocha, M. G. Hoesl und N. Budisa, Towards Reassignment of the Methionine Codon AUG to Two Different Noncanonical Amino Acids in Bacterial Translation, *Croatica Chemica Acta*, Bd. **89**, Nr. 2, S. 243–253, 2016.
- [294] T. A. Cropp, J. C. Anderson und J. W. Chin, Reprogramming the amino-acid substrate specificity of orthogonal aminoacyl-tRNA synthetases to expand the genetic code of eukaryotic cells, *Nature Protocols*, Bd. **2**, Nr. 10, S. 2590–2600, 2007.
- [295] S. L. Blethes, E. A. Boeker und E. E. Snell, Arginine Decarboxylase from *Escherichia coli* I. Purification and Specificity for Substrates and Coenzyme, *The Journal of Biological Chemistry*, Bd. **243**, Nr. 8, S. 1671–1677, 1968.
- [296] G. Harding und A. Ronald, The management of urinary infections; what have we learned in the past decade?, *International Journal of Antimicrobial Agents*, Bd. **4**, Nr. 2, S. 83–88, 1994.
- [297] T. J. Hannan, M. Totsika, K. J. Mansfield, K. H. Moore, M. A. Schembri und S. J. Hultgren, Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic *Escherichia coli* bladder infection, *FEMS Microbiology Reviews*, Bd. **36**, Nr. 3, S. 616–648, 2012.
- [298] C. Spaulding und S. Hultgren, Adhesive Pili in UTI Pathogenesis and Drug Development, *Pathogens*, Bd. **5**, Nr. 1, S. 30, 2016.
- [299] R. Fronzes, H. Remaut und G. Waksman, Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria, *The EMBO Journal*, Bd. **27**, Nr. 17, S. 2271–2280, 2008.
- [300] D. G. Thanassi, J. B. Bliska und P. J. Christie, Surface organelles assembled by secretion systems of Gram-negative bacteria: diversity in structure and function, *FEMS Microbiology Reviews*, Bd. **36**, Nr. 6, S. 1046–1082, 2012.
- [301] D. C. F. Monteiro, W. V. P. Kamdoum und E. Paci, Growth Kinetics of Bacterial Pili from Pairwise Pilin Association Rates, *PLoS ONE*, Bd. **8**, Nr. 5, e63065, 2013.
- [302] S. B. Svenson, G. Källenius, R. Möllby, H. Hultberg, T. K. Korhonen und J. Winberg, P-fimbriae of pyelonephritogenic *Escherichia coli*: Identification and chemical characterization of receptors, *Infection*, Bd. **11**, Nr. 1, S. 61–67, 1983.
- [303] K. Bock et al., Specificity of Binding of a Strain of Uropathogenic *Escherichia coli* to Gal α 1- \rightarrow 4Gal-containing Glycosphingolipids, *The Journal of Biological Chemistry*, Bd. **260**, Nr. 14, S. 8545–8551, 1985.
- [304] A. Busch und G. Waksman, Chaperone-usher pathways: diversity and pilus assembly mechanism, *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, Bd. **367**, Nr. 1592, S. 1112–1122, 2012.

- [305] D. G. Thanassi, E. T. Saulino, M.-J. Lombardo, R. Roth, J. Heuser und S. J. Hultgren, The PapC usher forms an oligomeric channel: Implications for pilus biogenesis across the outer membrane, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **95**, Nr. 6, S. 3146–3151, 1998.
- [306] A. Hernday, M. Krabbe, B. Braaten und D. Low, Self-perpetuating epigenetic pili switches in bacteria, *Proceedings of the National Academy of Sciences of the United States of America*, Bd. **99**, Nr. Suppl. 4, S. 16 470–16 476, 2002.
- [307] M. Båga, M. Göransson, S. Normark und B. E. Uhlin, Transcriptional activation of a pap pilus virulence operon from uropathogenic *Escherichia coli*, *The EMBO Journal*, Bd. **4**, Nr. 13B, S. 3887–3893, 1985.
- [308] E. B. Newman, R. D’Ari und R. T. Lin, The leucine-Lrp Regulon in *E. coli*: A Global Response in Search of a Raison d’Être, *Cell*, Bd. **68**, Nr. 4, S. 617–619, 1992.
- [309] J. M. Calvo und R. G. Matthews, The Leucine-Responsive Regulatory Protein, a Global Regulator of Metabolism in *Escherichia coli*, *Microbiological Reviews*, Bd. **58**, Nr. 3, S. 466–490, 1994.
- [310] P. M. Leonard et al., Crystal structure of the Lrp-like transcriptional regulator from the archaeon *Pyrococcus furiosus*, *The EMBO Journal*, Bd. **20**, Nr. 5, S. 990–997, 2001.
- [311] S. de los Rios und J. J. Perona, Structure of the *Escherichia coli* Leucine-responsive Regulatory Protein Lrp Reveals a Novel Octameric Assembly, *Journal of Molecular Biology*, Bd. **366**, Nr. 5, S. 1589–1602, 2007.
- [312] X. Nou, B. Braaten, L. Kaltenbach und D. A. Low, Differential binding of Lrp to two sets of pap DNA binding sites mediated by Pap I regulates Pap phase variation in *Escherichia coli*, *The EMBO Journal*, Bd. **14**, Nr. 23, S. 5785–5797, 1995.
- [313] A. D. Hernday, B. A. Braaten und D. A. Low, The Mechanism by which DNA Adenine Methylase and PapI Activate the Pap Epigenetic Switch, *Molecular Cell*, Bd. **12**, Nr. 4, S. 947–957, 2003.
- [314] L. S. Kaltenbach, B. A. Braaten und D. A. Low, Specific Binding of PapI to Lrp-pap DNA Complexes, *Journal of Bacteriology*, Bd. **177**, Nr. 22, S. 6449–6455, 1995.
- [315] M. van der Woude, B. Braaten und D. Low, Epigenetic phase variation of the pap operon in *Escherichia coli*, *Trends in Microbiology*, Bd. **4**, Nr. 1, S. 5–9, 1996.
- [316] K. Forsman, M. Göransson und B. E. Uhlin, Autoregulation and multiple DNA interactions by a transcriptional regulatory protein in *E. coli* pili biogenesis, *The EMBO Journal*, Bd. **8**, Nr. 4, S. 1271–1277, 1989.
- [317] B. A. Braaten, X. Nou, L. S. Kaltenbach und D. A. Low, Methylation Patterns in Pap Regulatory DNA Control Pyelonephritis-Associated Pili Phase Variation in *E. coli*, *Cell*, Bd. **76**, Nr. 3, S. 577–588, 1994.

- [318] Y. Cui, M. A. Midkiff, Q. Wang und J. M. Calvo, The Leucine-responsive Regulatory Protein (Lrp) from *Escherichia coli*: Stoichiometry and Minimal Requirements for Binding to DNA, *The Journal of Biological Chemistry*, Bd. **271**, Nr. 12, S. 6611–6617, 1996.
- [319] T. Kawamura, A. S. Vartanian, H. Zhou und F. W. Dahlquist, The Design Involved in PapI and Lrp Regulation of the pap Operon, *Journal of Molecular Biology*, Bd. **409**, Nr. 3, S. 311–332, 2011.
- [320] S. N. Peterson und N. O. Reich, Competitive Lrp and Dam Assembly at the pap Regulatory Region: Implications for Mechanisms of Epigenetic Regulation, *Journal of Molecular Biology*, Bd. **383**, Nr. 1, S. 92–105, 2008.
- [321] J. Casadesús und D. A. Low, Programmed Heterogeneity: Epigenetic Mechanisms in Bacteria, *The Journal of Biological Chemistry*, Bd. **288**, Nr. 20, S. 13 929–13 935, 2013.
- [322] S. Adhikari und P. D. Curtis, DNA methyltransferases and epigenetic regulation in bacteria, *FEMS Microbiology Reviews*, Bd. **40**, Nr. 5, S. 575–591, 2016.
- [323] M. Totsika, S. A. Beatson, N. Holden und D. L. Gally, Regulatory interplay between pap operons in uropathogenic *Escherichia coli*, *Molecular Microbiology*, Bd. **67**, Nr. 5, S. 996–1011, 2008.
- [324] N. J. Weyand und D. A. Low, Regulation of Pap phase variation. Lrp is sufficient for the establishment of the phase off pap DNA methylation pattern and repression of pap transcription in vitro, *The Journal of Biological Chemistry*, Bd. **275**, Nr. 5, S. 3192–3200, 2000.
- [325] S. N. Peterson und N. O. Reich, GATC Flanking Sequences Regulate Dam Activity: Evidence for how Dam Specificity may Influence pap Expression, *Journal of Molecular Biology*, Bd. **355**, Nr. 3, S. 459–472, 2006.
- [326] N. Mashhoon et al., Functional Characterization of *Escherichia coli* DNA Adenine Methyltransferase, a Novel Target for Antibiotics, *The Journal of Biological Chemistry*, Bd. **279**, Nr. 50, S. 52 075–52 081, 2004.
- [327] L. B. Blyn, B. A. Braaten und D. A. Low, Regulation of pap pilin phase variation by a mechanism involving differential dam methylation states, *The EMBO Journal*, Bd. **9**, Nr. 12, S. 4045–4054, 1990.
- [328] M. Göransson, K. Forsman, P. Nilsson und B. E. Uhlin, Upstream activating sequences that are shared by two divergently transcribed operons mediate cAMP-CRP regulation of pilus-adhesin in *Escherichia coli*, *Molecular Microbiology*, Bd. **3**, Nr. 11, S. 1557–1565, 1989.
- [329] N. J. Weyand, B. A. Braaten, M. van der Woude, J. Tucker und D. A. Low, The essential role of the promoter-proximal subunit of CAP in pap phase variation: Lrp and helical phase-dependent activation of papBA transcription by CAP from -215, *Molecular Microbiology*, Bd. **39**, Nr. 6, S. 1504–1522, 2001.

-
- [330] A. D. Hernday, B. A. Braaten, G. Broitman-Maduro, P. Engelberts und D. A. Low, Regulation of the Pap Epigenetic Switch by CpxAR: Phosphorylated CpxR Inhibits Transition to the Phase ON State by Competition with Lrp, *Molecular Cell*, Bd. **16**, Nr. 4, S. 537–547, 2004.
- [331] M. W. van der Woude, L. S. Kaltenbach und D. A. Low, Leucine-responsive regulatory protein plays dual roles as both an activator and a repressor of the Escherichia coli pap fimbrial operon, *Molecular Microbiology*, Bd. **17**, Nr. 2, S. 303–312, 1995.
- [332] C. A. White-Ziegler, M. L. Angus Hill, B. A. Braaten, M. W. van der Woude und D. A. Low, Thermoregulation of Escherichia coli pap transcription: H-NS is a temperature-dependent DNA methylation blocking factor, *Molecular Microbiology*, Bd. **28**, Nr. 6, S. 1121–1137, 1998.
- [333] J. Casadesús und D. Low, Epigenetic Gene Regulation in the Bacterial World, *Microbiology and Molecular Biology Reviews*, Bd. **70**, Nr. 3, S. 830–856, 2006.
- [334] M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Research*, Bd. **31**, Nr. 13, S. 3406–3415, 2003.
- [335] P. De Rijk, J. Wuyts und R. De Wachter, RnaViz 2: an improved representation of RNA secondary structure, *Bioinformatics*, Bd. **19**, Nr. 2, S. 299–300, 2003.
- [336] S.-H. Bae, H.-K. Cheong, C. Cheong, S. Kang, D. S. Hwang und B.-S. Choi, Structure and Dynamics of Hemimethylated GATC Sites: implications for DNA-SeqA recognition, *The Journal of Biological Chemistry*, Bd. **278**, Nr. 46, S. 45 987–45 993, 2003.
- [337] V. Brázda, R. C. Laister, E. B. Jagelská und C. Arrowsmith, Cruciform structures are a common DNA feature important for regulating biological processes, *BMC Molecular Biology*, Bd. **12**, Artikel 33, 2011.
- [338] M. Swoboda, M. S. Grieb, S. Hahn und M. Schlierf, „Measuring Two at the Same Time: Combining Magnetic Tweezers with Single-Molecule FRET“, in *Fluorescent Methods for Molecular Motors*, C. P. Toseland und N. Fili, Hrsg., Basel: Springer Verlag, 2014, Kap. 12.6.2, S. 271–273, ISBN: 9783034808569.

Dank und Anerkennung

In erster Linie gilt mein Dank meinem Doktorvater Herrn Prof. Dr. Albrecht Ott für die Möglichkeit, in seiner Arbeitsgruppe geforscht zu haben und diese Dissertation zu erstellen. Seine offene Art im Einklang mit seiner Neugier und dem Spaß an interessanten Fragestellungen und Problemen haben die Promotion zu einer lehrreichen und prägenden Zeit werden lassen.

Ich möchte mich ebenso bei Herrn Prof. Dr. Dr. Karsten Kruse bedanken, der mir nicht nur in seiner Funktion als wissenschaftliche Begleitung mit Rat und Tat zur Seite stand und immer ein offenes Ohr für mich hatte, sondern stets auch reges Interesse am Gelingen dieser Arbeit und gemeinsamen Projekten zeigte.

Ich bedanke mich auch bei Frau Prof. Dr. Karin Jacobs für die Übernahme der wissenschaftlichen Wegbegleitung nach dem Weggang von Prof. Dr. Dr. Karsten Kruse sowie für ihre engagierte Unterstützung und ihr reges Interesse am Erfolg dieser Dissertation.

Selbstverständlich gilt mein Dank ebenfalls der gesamten Arbeitsgruppe: Eva Wollrab, die mich während der vorangegangenen Diplomarbeit betreute und großen Anteil daran hatte, meinen wissenschaftlichen Weg fortzusetzen; Marc Schenkelberger für die Unterstützung bei der Einarbeitung in die zellfreie Genexpression; Julia Flesch und Lucina Kainka als zu betreuende studentische Hilfskräfte für ihre experimentelle Unterstützung; Marc Finkler, meinem wissenschaftlichen Nachfolger, für seine experimentelle Unterstützung im Rahmen seiner Einarbeitung in die zellfreie Genexpression; meinen Büro-Nachbarn Sabrina Scherer und Christian Trapp für das angenehme Miteinander.

Herzlich danken möchte ich Jennifer Maria Maul, der guten Seele des Sekretariats, der nie etwas zu viel war und die bei jedem Formular, von denen es im öffentlichen Dienst bekanntlich viele gibt, half.

Letztlich gilt mein Dank meinen Kooperationspartnern Philipp Zimmer und Matthias P. Exner wie auch den Herren Prof. Dr. Vincent Noireaux und Prof. Dr. Nediljko Budisa für die gute Zusammenarbeit und ihre Beiträge bzw. Betreuung bei gemeinsamen Projekten.

Einen Freund gilt es besonders hervorzuheben, Nicolas Thewes, dem ich für die vielen Jahre danke, die wir fast täglich im Studium und während der Promotion gemeinsam verbracht haben.

Dank gegenüber meiner Familie auszudrücken, ist kaum möglich. Meine Eltern Hildegard und Peter sowie meine Lebensgefährtin Anke Marx und Max waren die emotionalen Pfeiler während der Promotion. Sie haben mich in bedingungsloser Weise liebevoll unterstützt. Ungeboren, aber sichtbar auf dem Weg, hat mich meine Tochter Leni beim letzten Feinschliff der Dissertation in höchstem Maße motiviert.