

COLE₁-PLASMIDPRODUKTION IN *ESCHERICHIA COLI*:
SIMULATION UND EXPERIMENT

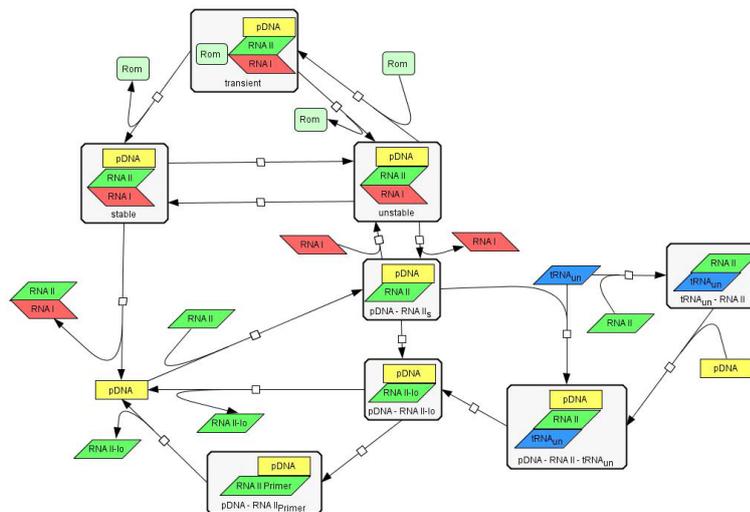
Dissertation

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

an der Fakultät für Biologie der Universität Bielefeld

(genehmigte Fassung)



vorgelegt von

Inga Freudenau

Juni 2014

„Wer hohe Türme bauen will, muß lange beim Fundament verweilen.“

— Anton Bruckner

Für Beate

INHALTSVERZEICHNIS

I	ZUSAMMENFASSUNG	1
1	ZUSAMMENFASSUNG	2
II	EINLEITUNG	4
2	DIE BEDEUTUNG VON PLASMID-DNA FÜR DIE MEDIZINISCHE ANWENDUNG	5
III	GRUNDLAGEN	9
3	BIOLOGISCHER HINTERGRUND	10
3.1	Plasmid-DNA-Produktion für die medizinische Anwendung	10
3.2	Zwei Ebenen zur Kontrolle der Plasmidreplikation	14
3.3	Plasmidreplikationskontrolle von <i>ColE1</i> -ähnlichen Plasmiden	16
3.4	Metabolomanalysen	20
3.5	¹³ C-Flussanalysen	22
4	THEORETISCHER HINTERGRUND	28
4.1	Modellierung in der Systembiologie	28
4.2	Was ist ein Modell?	29
4.3	Klassifizierung von Modellen	30
4.4	Die Erstellung eines mathematischen Modells	32
4.5	Mathematischer Hintergrund	36
5	ZIEL DER ARBEIT	40
IV	MATERIAL UND METHODEN	41
6	MATERIAL	42
6.1	Verwendete Bakterienstämme	42
6.2	Primer, Transposon und Plasmidvektoren	43
6.2.1	Primer	43
6.2.2	Transposon	44
6.2.3	Plasmidvektoren	44
6.3	Enzyme, Chemikalien, Kits und Verbrauchsmaterial	46
6.3.1	Enzyme	46

6.3.2	Chemikalien	47
6.3.3	Kits	48
6.4	Laborgeräte	49
6.4.1	Verbrauchsmaterial	50
6.5	Software	51
6.6	Nährmedien und Medienzusätze	52
6.6.1	Definierte Medien zur Kultivierung von Bakterien	52
6.7	Medienzusätze	53
6.7.1	Antibiotika	53
6.7.2	Agar	53
6.8	Puffer und Lösungen	54
6.8.1	Lösungen zur Herstellung von Dauerkulturen	54
6.8.2	Puffer zur Herstellung chemisch kompetenter Zellen	54
6.8.3	Puffer für die DNA-Agarosegelelektrophorese	55
6.8.4	Puffer für die HPLC-Analysen	55
7	METHODEN	56
7.1	Kultivierung von Bakterien	56
7.1.1	Anzucht von Bakterien	56
7.1.2	Herstellung von Dauerkulturen	56
7.1.3	Bestimmung des Bakterientiters	56
7.2	Isolierung und Darstellung von DNA	57
7.2.1	Isolierung von Plasmid-DNA mittels „Thermo Scientific - GeneJET Plasmid Miniprep Kit“	57
7.2.2	Konzentrationsbestimmung einer DNA-Lösung	57
7.2.3	Agarosegelelektrophorese	58
7.2.4	Größenbestimmung von DNA-Fragmenten	59
7.2.5	Isolierung von DNA-Fragmenten aus Agarosegelen mittels „MN - PCR clean-up and Gel extraction Kit“	59
7.3	Restriktion und Modifikation von DNA	60
7.3.1	DNA- Restriktion mit Typ II Restriktionsendonukleasen	60
7.3.2	Dephosphorylierung von 5'-DNA-Enden	61
7.3.3	Polymerasekettenreaktion (PCR)	61
7.3.4	Aufreinigung von PCR Produkten mittels „MN - PCR clean-up and Gel extraction Kit“	62

7.3.5	Ligation	63
7.4	DNA-Transfer mittels chemischer Transformation	63
7.4.1	Herstellung chemisch kompetenter <i>E. coli</i> Zellen mittels CaCl ₂ -Methode (Protokoll der Plasmid Factory)	63
7.4.2	Transformation kompetenter <i>E. coli</i> Zellen	64
7.5	Erstellung eines RNA-Standards und Durchführung von qRT-PCR	65
7.5.1	PCR zur Amplifikation des RNAI- und RNAII-Gens zum Einsatz in der T7-RNA-Polymerase-Reaktion	65
7.5.2	T7-RNA-Polymerase-Reaktion	66
7.5.3	Isolierung von RNA mittels „Qiagen - RNeasy plus Mini Kit“	67
7.5.4	Entfernung von DNA mittels DNaseI-Verdau	68
7.5.5	Konzentrationsbestimmung einer RNA-Lösung	68
7.5.6	qRT-PCR mittels „Bioline - SensiMix Probe One-Step Kit“	68
7.6	Absolute Quantifizierung von Plasmid-DNA	70
7.6.1	Absolute Quantifizierung der Plasmid-DNA durch die Firma Plasmid Factory	70
7.6.2	Absolute Quantifizierung der Plasmid-DNA durch die Firma CARPEGEN	71
7.6.3	Absolute Quantifizierung der Plasmid-DNA mittels Hochleistungsflüssigkeitschromatographie	71
7.7	Durchführung von Metabolomanalysen	72
7.7.1	Zellernte durch Schockgefrieren	72
7.7.2	Extraktion und Derivatisierung der intrazellulären Metabolite	72
7.7.3	Gaschromatographie-Massenspektrometrie	73
7.7.4	Auswertung der GC-MS-Messungen	74
7.8	Durchführung von ¹³ C-Flussanalysen	75
7.8.1	Anzucht der Bakterien und Ernte durch Schockgefrieren	75
7.8.2	Proteinhydrolyse und Derivatisierung	75
7.8.3	Gaschromatographie-Massenspektrometrie	76
7.8.4	Auswertung der GC-MS-Messungen	77
V	ERGEBNISSE	78
8	MODELLIERUNG DER PLASMIDREPLIKATIONSKONTROLLE	79
8.1	Erstellung des strukturellen Modells	81
8.2	Erfassung der experimentellen Daten	85

8.2.1	Bestimmung der Wachstumsraten	86
8.2.2	Messung der RNAI- und RNAII-Konzentrationen	92
8.2.3	Messung der intrazellulären Plasmidkonzentrationen	98
8.3	Erstellung des dynamischen Modells	104
8.4	Simulationen der Plasmidreplikationskontrolle	112
8.4.1	Simulationen zur Parameteranpassung	112
8.4.2	Prädiktive Simulationen	116
8.5	Entwicklung einer unbeladbaren Transfer-RNA	119
8.5.1	Die Transkriptionseinheit	119
8.5.2	Konstruktion des modifizierten Leucin-tRNA-Gens	122
8.5.3	Vorgehen zur Integration des modifizierten tRNA-Gens in das <i>E. coli</i> Chromosom	126
9	METABOLOMANALYSEN	131
9.1	Untersuchungen zur Metabolitzusammensetzung	131
10	¹³ C-FLUSSANALYSEN	141
10.1	Ermittlung der Isotopomer-Verteilung	141
10.2	Das metabolische Flussmodell	147
10.3	Berechnung des intrazellulären Glukoseflusses	150
VI	DISKUSSION	156
11	MODELLIERUNG DER PLASMIDREPLIKATIONSKONTROLLE	157
11.1	Das Modell	157
11.2	Untermauerung des Modells mit experimentellen Daten	159
11.3	Simulationen der Plasmidreplikationskontrolle	165
11.3.1	Simulationen zur Parameteranpassung	165
11.3.2	Prädiktive Simulationen	167
12	DAS METABOLOM UND DER INTRAZELLULÄRE GLUKOSEFLUSS	169
12.1	Das Metabolom in Abhängigkeit des Plasmidgehalts	169
12.2	Betrachtung des Glukoseflusses via ¹³ C-Flussanalysen	174
12.3	Fazit und Ausblick	179
VII	ANHANG	182
	LITERATURVERZEICHNIS	183

ABKÜRZUNGSVERZEICHNIS

AcCoA	Acetyl-Coenzym A
AKG	α -Ketoglutarat
Amp ^R	Ampizillin-Resistenz
bp	Basenpaare
bzw.	beziehungsweise
ca.	circa
cDNA	<i>copy</i> DNA
Cm ^R	Chloramphenicol-Resistenz
CMV	<i>Cytomegalievirus</i>
CO ₂	Kohlenstoffdioxid
CT	<i>cycle threshold</i>
DGL	Differentialgleichung
DNA	Desoxyribonukleinsäure
F6P	Fruktose-6-phosphat
FDA	Food and Drug Administration
E ₄ P	Erythrose-4-phosphat
E. coli	<i>Escherichia coli</i>
G6P	Glukose-6-phosphat
GC	Gaschromatographie
Glc _E x	Extrazelluläre Glukose
Gm ^R	Gentamycin-Resistenz
hGH	Human growth hormone)
HPLC	<i>High performance liquid chromatography</i>
ICT	Isocitrat
MAL	Malat

MBDSTFA	N-Methyl-N-Tert-Butyldimethylsilyl-Trifluoroacetamid
MS	Massenspektrometer
MSTFA	N-Methyl-N-Trimethylsilyl-Trifluoroacetamid
MW	Mittelwert
n. d.	Nicht detektierbar
Nm ^R	Neomycin-Resistenz
NMR	<i>Nuclear magnetic resonance</i>
OAA	Oxalacetat
OD	Optische Dichte
P ₅ P	Pentose-5-phosphat
PCA	<i>Principal component analysis</i>
PCN	<i>Plasmid copy number</i>
PCR	Polymerasekettenreaktion
PEP	Phosphoenolpyruvat
PGA	Phosphoglycerat
pDNA	Plasmid-DNA
PPP	<i>Pentosephosphate pathway</i>
PRPP	Phosphoribosylpyrophosphat
PYR	Pyruvat
qRT-PCR	quantitative Reverse Transkription-Polymerasekettenreaktion
RNA	Ribonukleinsäure
Rom	<i>RNAI modulating protein</i>
Rop	<i>Repressor of Primer</i>
RT	Reverse Transkription
S ₇ P	Sedoheptulose-7-phosphat
T ₃ P	Triose-3-phosphat
u. U.	unter Umständen
vs.	versus
z. B.	zum Beispiel
6PG	6-Phosphogluconat

ANMERKUNG ZUR ZITIERTECHNIK

In der vorliegenden Arbeit wurde folgende Zitiertechnik verwendet:

Abschnittsbezug: Bezieht sich die Quellenangabe auf den Inhalt des gesamten Abschnitts, wurde die Zitatangabe hinter dem letzten Punkt des Abschnitts gesetzt.

Satzbezug: Bezieht sich die Quellenangabe auf den Inhalt eines Satzes, wurde die Zitatangabe vor dem letzten Punkt des Abschnitts gesetzt.

Begriffsbezug: Bezieht sich die Quellenangabe nur auf einen bestimmten Begriff, wurde die Zitatangabe direkt hinter diesem gesetzt.

Teil I

ZUSAMMENFASSUNG

ZUSAMMENFASSUNG

Im Rahmen dieser Arbeit wurde ein mathematisches Modell zur Simulation der *ColE1*-Plasmidreplikationskontrolle für den *low copy* plasmidtragenden Stamm DH5 α – pSUP 201-3 und für den *high copy* plasmidtragenden Stamm DH5 α -pCMV-lacZ unter Verwendung von MATLAB, *Mathworks*[®] erstellt. Das Modell beinhaltet die Plasmidreplikationskontrolle durch die regulatorischen RNA-Moleküle, RNAI und RNAII sowie die Regulation durch unbeladene tRNA-Moleküle. Zur Untermauerung des Modells wurden experimentelle Daten, wie die RNAI- und RNAII-Konzentration, die Plasmidkopienzahl sowie die Wachstumsrate für jeweils drei verschiedene Zeitpunkte bestimmt. Die für DH5 α -pSUP 201-3 durchschnittlich gemessenen RNA-Konzentrationen betragen, abhängig vom Messzeitpunkt, 6 ± 0.7 bis 34 ± 7 RNAI-Moleküle pro Zelle und 0.44 ± 0.1 bis 3 ± 0.9 RNAII-Moleküle pro Zelle. Die durchschnittlich gemessenen Plasmidkopienzahlen von pSUP 201-3 betragen 46 ± 26 bis 48 ± 30 . Die für DH5 α -pCMV-lacZ durchschnittlich gemessenen Plasmidkopienzahlen liegen mit 1514 ± 1301 bis 5806 ± 4828 Plasmiden pro Zelle deutlich höher. Ebenso verhält es sich mit den gemessenen RNAI- und RNAII-Konzentrationen. Die durchschnittlich ermittelte Anzahl an RNAI-Molekülen pro Zelle für DH5 α -pCMV-lacZ bewegt sich zwischen 345 ± 203 und 1086 ± 298 . Die durchschnittlich gemessene RNAII-Konzentration beträgt 22 ± 2 bis 75 ± 10 Moleküle pro Zelle. Die experimentellen Daten dienen als Anfangsbedingungen für das Modell, wenn die *ColE1*-Plasmidreplikationskontrolle für eine einzelne Zelle simuliert wird und die Simulationen an einem der drei entsprechenden Messzeitpunkte beginnen.

Mit Hilfe des Modells wurde untersucht, welchen Effekt es auf die Plasmidproduktion hätte, würde man tRNA-Moleküle derart modifizieren, dass die Beladung mit einer Aminosäure durch die Aminoacyl-tRNA-Synthetase nicht mehr möglich wäre. Die Simulationen zur Untersuchung dieser Fragestellung haben gezeigt, dass solche modifizierten tRNA-Moleküle die Plasmidproduktion *in silico* erhöhen. Um diese *in silico*-Vorhersagen im Labor zu überprüfen, wurde das *leuQ*-Gen dahingehend verändert, dass das kodierte Genprodukt eine Leucin-tRNA darstellt, die von der Leucyl-tRNA-Synthetase nicht mehr erkannt wird und somit unbeladen bleibt. Das modifizierte *leuQ*-Gen wurde mittels Transposon-Mutagenese in das Chromosom von *E. coli* DH5 α inseriert. Erste

Versuche bestätigten die *in silico*-Vorhersagen und zeigten eine leichte Erhöhung der intrazellulären Plasmidkonzentration.

Neben der Regulation der Plasmidreplikation auf Genomebene lässt sich die Plasmidreplikation auf Metabolomebene durch die Verfügbarkeit von Nukleotiden kontrollieren. Inwieweit die Plasmidproduktion durch den Metabolismus der Zelle limitiert ist, wurde bisher wenig untersucht. Aus diesem Grund wurde in dieser Arbeit die Regulation auf Metabolomebene durch Metabolitprofilanalysen und ^{13}C -Flussmessungen untersucht. Im Rahmen dieser Analysen sollte der Frage nachgegangen werden, welchen Einfluss der Plasmidgehalt einer Zelle auf ihre Metabolitzusammensetzung hat. Dabei wurde der plasmidfreie Stamm DH5 α sowie die Stämme DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP untersucht. Anhand einer Hauptkomponentenanalyse konnte gezeigt werden, dass sich die drei Stämme hinsichtlich ihrer Metabolitprofile deutlich voneinander unterscheiden. Untersucht man die Konzentrationen von Ribose, Ribose-5-phosphat und Xylose sieht man Unterschiede in den Metabolitpools, die mit dem unterschiedlichen Plasmidgehalt begründet werden können, da diese Metabolite von allen identifizierten Metaboliten über den Pentosephosphatweg am direktesten mit der Nukleotidbiosynthese in Verbindung stehen.

Die Ergebnisse der ^{13}C -Flussmessungen deuten darauf hin, dass der Glukosefluss in den Pentosephosphatweg bei den untersuchten Plasmidproduktionsstämmen W3110-pHN, VH33-pHN, VH33 Δ *recA*-pHN, VH33 Δ (*recA deoR*)-pHN vom Plasmidgehalt der Zelle abhängt, so dass bei Zellen mit einer höheren Plasmidproduktion ein größerer Anteil der aufgenommenen Glukose in den Pentosephosphatweg fließt und dort zur Synthese von Vorstufen für die Nukleotidbiosynthese zur Verfügung steht.

Teil II

EINLEITUNG

DIE BEDEUTUNG VON PLASMID-DNA FÜR DIE MEDIZINISCHE ANWENDUNG

Vor gut 20 Jahren wurde zum ersten Mal entdeckt, dass man durch Injektion von DNA, die für ein körperfremdes Protein kodiert, bei Tieren eine Antwort des Immunsystems auslösen konnte. Dazu wurden Mäuse mit Plasmid-DNA überzogenen Mikroprojektilen geimpft. Die Plasmide enthielten die Kopie eines Gens, das für ein menschliches Wachstumshormon (hGH) kodiert und unter der Kontrolle des humanen β -Actin Promotors oder des Cytomegalovirus (CMV) Promotors stand. Nach der Transfektion konnte eine Antikörperproduktion direkt gegen hGH beobachtet werden. (Tang *et al.*, 1992)

Seit dieser Entdeckung nahm das wissenschaftliche Interesse am Gebiet der DNA-Vakzination zu, so dass im Jahr 2005 die ersten DNA-Vakzine zugelassen wurden. Einer der ersten DNA basierten Impfstoffe ist der *West Nile Virus Innovator*. Er wurde von der Firma *Fort Dodge Animal Health* entwickelt und schützt Pferde vor dem West-Nil-Virus. Ein weiterer im Jahr 2005 zugelassener DNA-Impfstoff mit dem Namen *Apex-IHN* wurde von *Novartis Animal Health* entwickelt und wird bei Lachsen angewendet, um sie gegen die infektiöse hämatopoetische Nekrose zu impfen (Prazeres, 2011).

Das Prinzip der DNA-Vakzination ist die Induzierung einer zellulären und einer humoralen Immunantwort durch DNA-Impfstoffe (Abbildung 2.1). Dabei werden Gene in Plasmid-Vektoren eingebaut, die für virale (oder auch bakterielle) Proteine kodieren, die zur Immunisierung notwendig sind. Zuvor werden die entsprechenden Gensequenzen mittels *polymerase chain reaction (PCR)* amplifiziert. Die Plasmide werden in den gewünschten Produktionsstamm transformiert und darin vermehrt. Anschließend wird die Plasmid-DNA (pDNA) aufgereinigt und entsprechend der Anwendung verpackt. Im nächsten Schritt werden die Plasmide unter die Haut oder in den Muskel injiziert. Hier kommt es sowohl zur Transfektion der umgebenden Muskelzellen, als auch zur Transfektion von Antigen präsentierenden Zellen (APC von engl. *antigen presenting cell*). Diese Zellen synthetisieren die plasmidkodierten, körperfremden Proteine und bilden entsprechende Antigene. Die Antigene werden mittels sogenannter Haupthistokompatibilitätskomplexen (engl. *major histocompatibility complex*) auf den Zelloberflächen präsentiert.

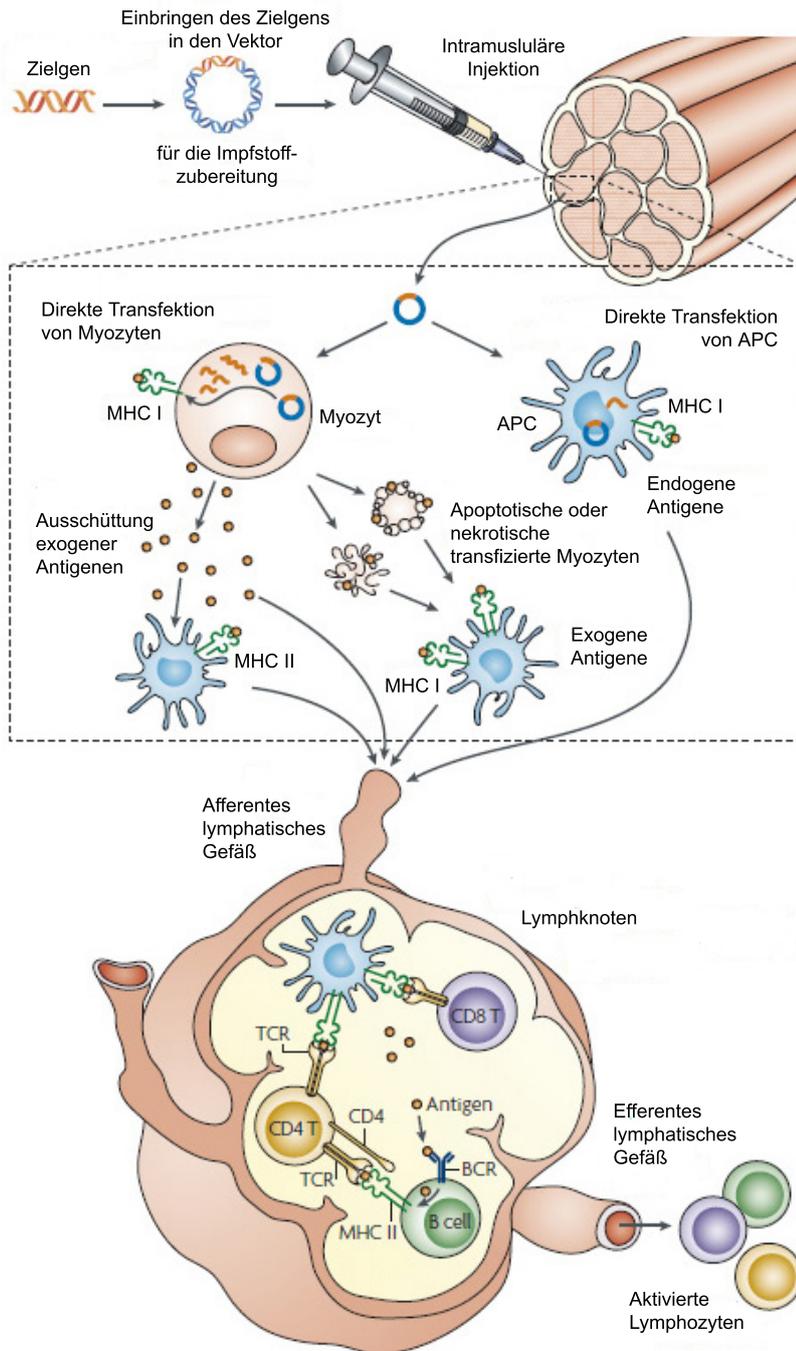


Abbildung 2.1: Prinzip der DNA-Vakzination.

Das Prinzip der DNA-Vakzination ist die Induzierung einer zellulären und einer humoralen Immunantwort durch DNA-Impfstoffe. Der Hauptbestandteil eines DNA-Impfstoffs ist Plasmid-DNA, in die Gene eingebaut sind, die für virale (oder auch bakterielle) Proteine kodieren, die zur Immunisierung notwendig sind. Die Plasmide werden unter die Haut oder in den Muskel injiziert, wo es zur Transfektion der umgebenden Muskelzellen kommt. Neben den Muskelzellen werden auch Antigen-präsentierende Zellen (APC) transfiziert. Die transfizierten Zellen synthetisieren die plasmidkodierte, körperfremde Proteine und bilden entsprechende Antigene. Die Antigene werden mittels sogenannter Haupthistokompatibilitätskomplexen (MHC) auf den Zelloberflächen präsentiert. Außerdem werden durch apoptotische und nekrotische Zellen Antigene an die Umgebung abgegeben, so dass weitere APCs die Antigene auf ihren Oberflächen exprimieren können. (Kutzler & Weiner, 2008)

Zusätzlich bilden sich apoptotische und nekrotische Zellen und es werden Antigene an die Umgebung abgegeben, wodurch weitere APCs in der Lage sind, die Antigene auf ihren Oberflächen zu exprimieren. Die APCs wandern in einen Lymphknoten, in dem über mehrere Schritte T- und B-Zellen aktiviert und Antikörper produziert werden. (Kutzler & Weiner, 2008)

Im Vergleich zu konventionellen Impfstoffen haben DNA-Vakzine den Vorteil, dass ihre Anwendung sicherer ist. Es besteht kein Risiko der Rückkehr in eine virulente Form oder einer Sekundärinfektion, da DNA-Vakzine keine abgeschwächte Pathogenität aufweisen und in Eukaryoten nicht replizieren können. Des Weiteren entstehen keine Unverträglichkeiten, weil sich die Immunantwort nur gegen solche Proteine richtet, die zur Immunisierung erforderlich sind. Ein zusätzlicher Aspekt, der für die Verwendung von DNA-Impfstoffen spricht, ist ihre kurze Herstellungszeit.

Es gibt dennoch einige Gruppen, die Bedenken bezüglich einer Integration ins Wirtsgenom und der Ausbildung einer anti-DNA Immunantwort äußern. Um diese möglichen Nachteile zu untersuchen, wurden umfangreiche Studien durchgeführt, die ergaben, dass das Risiko einer genomischen Integration sehr klein ist, da es deutlich geringer ist, als das Auftreten einer natürlich vorkommenden Mutation. Zur Ausbildung einer anti-DNA Immunantwort wurden ebenfalls eingehende Untersuchungen durchgeführt, wobei es in keiner Studie Hinweise auf eine erhöhte anti-DNA Immunantwort gab. (Kutzler & Weiner, 2008) (Ferraro *et al.*, 2011)

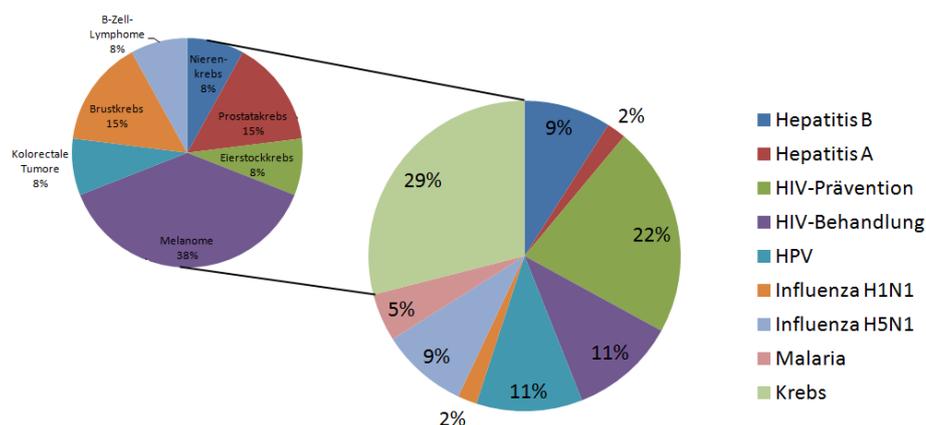


Abbildung 2.2: Klinischen Studien zu DNA-Vakzin-Angriffszielen.

Im Jahr 2011 befanden sich 43 DNA-Vakzine zur Beurteilung in klinischen Studien (großes Kreisdiagramm). Zwölf davon betrafen die Untersuchung von Krebserkrankungen, u. a. Brust- und Nierenkrebs (kleines Kreisdiagramm). (Ferraro *et al.*, 2011). Im Januar dieses Jahres ist die Anzahl der klinischen Studien, die sich mit der Erkrankung an Krebs beschäftigen, bereits auf 17 gestiegen (Senovilla *et al.*, 2013). (Ferraro *et al.*, 2011)

Neben der Anwendung zum Schutz vor viralen oder bakteriellen Infektionen könnten DNA-Vakzine auch Verwendung bei der Indikation von verschiedenen Krebsarten finden. Im Jahr 2011 wurden in 43 klinischen Studien DNA-Vakzine beurteilt, von denen 12 (28 %) die Untersuchung von Krebs betrafen (Ferraro *et al.*, 2011). Eine Übersicht über die Bereiche dieser 43 Studien ist in Abbildung 2.2 dargestellt. Im Januar 2013 beschäftigten sich bereits 17 klinische Studien mit der Erkrankung an Krebs (Senovilla *et al.*, 2013).

Neben den bereits erwähnten Plasmid-DNA-Vakzin Produkten sind im Jahr 2009 noch zwei weitere Plasmid-DNA-Vakzine für die veterinärmedizinische Anwendung zugelassen worden (Williams *et al.*, 2010).

Da in der Tiermedizin bereits einige Plasmid-Biopharmazeutika zugelassen sind und 43 DNA-Vakzine zur humanen Anwendung in klinischen Studien überprüft werden, wird die Nachfrage an pDNA für die medizinische Anwendung in den kommenden Jahren steigen (Bower & Prather, 2009). Eine notwendige Voraussetzung um diesem größer werdenden Interesse nachzukommen, ist die Produktion von pDNA in großem Maßstab und in bester Qualität. Dazu müssen die Prozesse der Plasmidproduktion mit entsprechenden Produktionsstämmen und Plasmidvektoren optimiert und weiterentwickelt werden.

Teil III

GRUNDLAGEN

BIOLOGISCHER HINTERGRUND

3.1 PLASMID-DNA-PRODUKTION FÜR DIE MEDIZINISCHE ANWENDUNG

„I propose *plasmid* as a generic term for any extrachromosomal hereditary determinant.“
(Lederberg, 1952)

Im Jahr 1952 führte Joshua Lederberg das Wort „Plasmid“ als einen allgemeinen Ausdruck für extrachromosomale Erbfaktoren ein. Zuvor waren extrachromosomale, genetische Elemente unter einer Vielzahl an Namen bekannt, wie „pangenes, bioblasts, plasmagenes, plastogenes, chondriogenes, cytogenes and proviruses“ (Lederberg, 1952). Heute definiert man Plasmide als „extrachromosomale, doppelsträngige DNA-Moleküle, die unabhängig von dem Wirtschromosom replizieren und mit diesem koexistieren.“ (Schuhmann, 1990)

In der Natur vorkommende Plasmide besitzen inhärente und phänotypische Funktionen. Zu den inhärenten Funktionen gehören Fähigkeiten wie Konjugation, Mobilisierung, Replikation, Kontrolle der Kopienzahl, Mechanismen der stabilen Vererbung von Plasmiden und Inkompatibilität. Unter phänotypischen Funktionen versteht man z. B. Antibiotika- oder Schwermetallresistenzen. Die für phänotypische Funktionen kodierenden Gene sind für das Bakterium nicht essentiell. Sie können ihm aber einen Wachstumsvorteil gegenüber anderen Bakterien verschaffen oder im Fall von sich ändernden Umweltbedingungen das Überleben sichern. (Schuhmann, 1990)

Plasmide sind seit den 70er Jahren des zwanzigsten Jahrhunderts ein wichtiges Werkzeug in der Molekulargenetik und in der Biotechnologie (Lederberg, 1998). Seit einigen Jahren finden sie zunehmend auch als Basis für diverse vielversprechende DNA-Biopharmazeutika Anwendung (Kutzler & Weiner, 2008). Dazu gehört neben der Verwendung in der DNA-Impfung auch der Einsatz in der Gentherapie. Im Rahmen einer Gentherapie wird eine Erkrankung aufgrund eines genetischen Defektes durch das Einbringen einer Kopie des intakten Gens in den Patienten behandelt. Zur erfolgreichen und gleichzeitig ökonomischen Durchführung von Gentherapien und DNA-Impfungen ist eine *large-scale* Produktion von qualitativ hochwertiger pDNA essentiell. Die Plasmid-

Produktion für DNA-Vakzine erfolgt ausschließlich in *Escherichia coli* K-12 Derivaten, die von der *Food and Drug Administration (FDA)* überprüft wurden. Dazu gehören z. B. DH5 α (Carnes *et al.*, 2006), DH5 (Listner *et al.*, 2006), DH1 (Cooke *et al.*, 2004), SCS1-L (Singer *et al.*, 2009) oder DH10B (Lahijani *et al.*, 1996). Diese Stämme besitzen den sogenannten GRAS-Status (GRAS = *Generally recognised as safe*) und sind demnach weder pathogen noch bilden sie toxische oder antibiotische Nebenprodukte.

Die Schritte eines Produktionsprozess von pDNA sind in Abbildung 3.1 als Übersicht dargestellt.

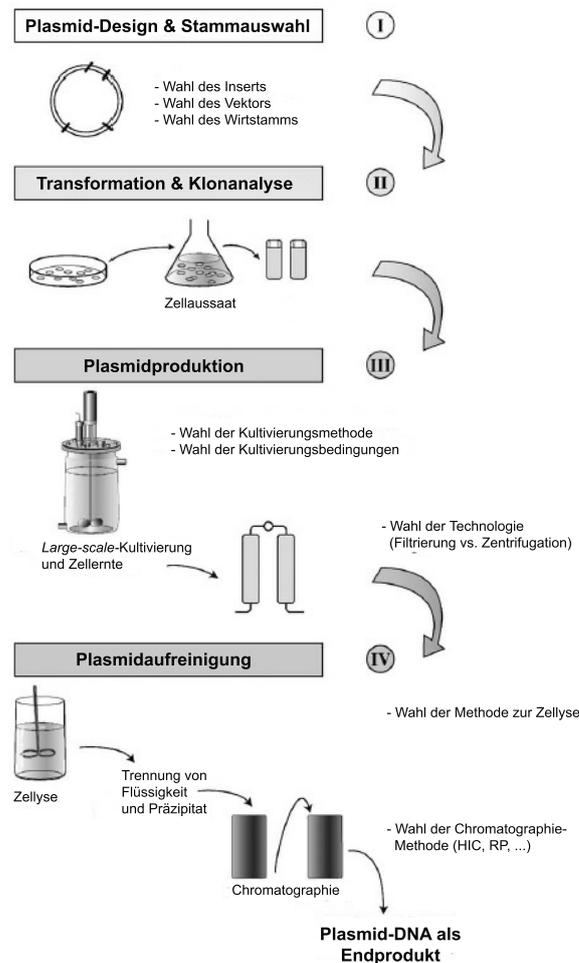


Abbildung 3.1: Schritte eines Plasmid-Produktionsprozesses.

Die Produktion von Plasmid-DNA beginnt mit dem Plasmid-Design und der Wahl des Wirtstamms (Schritt I). Anschließend wird das zu produzierende Plasmid in den gewählten Produktionsstamm transformiert und die richtigen Klone mittels einer Klonanalyse identifiziert (Schritt II). Die Zellen werden zur Vermehrung ausgesät. Im nächsten Schritt wird die Kultivierungsmethode mit passenden Kultivierungsbedingungen gewählt und eine frische Kultur der zuvor ausgesäten Zellen wird zur Plasmidproduktion eingesetzt (Schritt III). Der letzte Schritt besteht aus der Aufreinigung der Plasmid-DNA mit Hilfe einer chromatografischen Methode. (Prather *et al.*, 2003)

Zu Beginn des Produktionsprozesses müssen ein geeignetes Insert und ein passender Vektor gewählt werden. Ein mögliches Insert ist z. B. ein Antigen für ein DNA-Vakzin. Bei der Wahl eines Inserts sind viele Faktoren zu beachten, damit die Effektivität des Vakzins nicht negativ beeinflusst wird. Dazu gehört unter anderem, dass die Kodonverwendung der Aminosäuren an den Zielorganismus angepasst wird. (Williams *et al.*, 2010), (Brand *et al.*, 2012)

Bei der Auswahl des Vektors gilt generell: Je kleiner desto besser, da die Transfektionseffizienz mit zunehmender Größe abnimmt. Außerdem sollte er eine hohe Kopienzahl besitzen (Brand *et al.*, 2012). Die meisten therapeutischen Vektoren basieren auf pBR322- oder pUC-Plasmiden und besitzen *ColE1*-ähnliche Replikationsursprünge (Williams *et al.*, 2010). Die Unterschiede zwischen pBR322- und pUC-Plasmiden sind zum einen die Abwesenheit eines Gens in pUC-Plasmiden, welches für ein regulatorisches Protein namens Rom kodiert und zum anderen eine zusätzliche Punktmutation im RNAIL-Transkript (Lin-Chao *et al.*, 1992), (Yanisch-Perron *et al.*, 1985).

Neben *ColE1*-ähnlichen Replikationsursprüngen werden noch R6K- und R1-Replikons verwendet (Williams *et al.*, 2010). In Abbildung 3.2 ist eine allgemeine Darstellung eines aus Insert und Vektor bestehenden pDNA-Vakzin-Konstrukts zu sehen. Es enthält einen Replikationsursprung und einen Selektionsmechanismus zur stabilen Replikation im Wirtstamm. Des Weiteren ist das Insert unter der Kontrolle eines eukaryotischen Promotors mit einem Enhancer auf dem Vektor lokalisiert. Um die Transkription des eingefügten Gens zu beenden, ist zusätzlich ein Transkriptionsstop-Element enthalten. (Prather *et al.*, 2003)

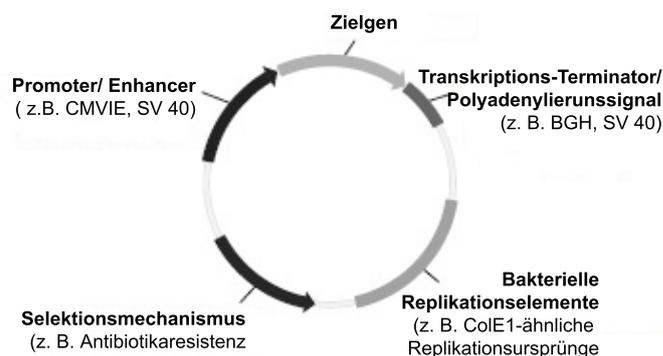


Abbildung 3.2: Ein typisches pDNA-Vakzin-Konstrukt.

Der Vektor eines typischen DNA-Vakzins weist neben einem Replikationsursprung und einem Selektionsmechanismus das Insert unter der Kontrolle eines eukaryotischen Promotors mit einem Enhancer auf. Des Weiteren ist ein Transkriptionsstop-Element enthalten, welches die Transkription des eingefügten Gens beendet. (Prather *et al.*, 2003)

Bevor die Produktion der pDNA beginnen kann, muss als letztes der Produktionsstamm gewählt werden. Ein wichtiges Kriterium bei der Wahl des Produktionsstammes ist, dass er gute Produktionseigenschaften besitzt. Dazu gehört neben der Bildung qualitativ guter pDNA, dass er nach der Kultivierung eine leichte Aufreinigung der Plasmide ermöglicht. Qualitativ gute pDNA zeichnet sich dadurch aus, dass die Plasmide in einer geschlossenen, kovalenten und zirkulären Form (*ccc* von engl. *covalently closed circular*) vorliegen. Neben der sogenannten *ccc*-Form können Plasmide noch linear und in der *oc*-Form (*oc* von engl. *open circular*) in der Zelle vorhanden sein. (Schleef, 2001)

Für den pharmazeutischen Einsatz von pDNA ist die Form der Plasmide ein wichtiger Aspekt, da die Tranfektionseffizienz von der Plasmid-Topologie beeinflusst wird. Der Grund dafür ist, dass die kompaktere *ccc*-Form besser von den Zellen aufgenommen wird, als die relaxierten Formen. Die FDA betrachtet die anderen Topologien sogar als Verunreinigung, so dass man im Rahmen der Plasmidproduktion für DNA-Vakzine an einem hohen *ccc*-Plasmidanteil interessiert ist. (Brand *et al.*, 2012) Es konnte gezeigt werden, dass der Wirtsstamm einen entscheidenden Einfluss auf die Plasmidausbeute und deren Qualität hat. Von den kommerziellen *E. coli* K-12-Stämmen eignet sich der DH5 α am besten zur Plasmidproduktion, da er mit 120,1 $\mu\text{g}/\text{ml}$ die größte Ausbeute an qualitativ hochwertiger Plasmid-DNA besitzt. Seine Leistung wird nur von dem Stamm HB101 mit einer Ausbeute von 134,9 $\mu\text{g}/\text{ml}$ übertroffen. Allerdings ist dieser Stamm nicht für eine *large-scale* Produktion geeignet, da er große Mengen an Kohlenhydraten enthält, die während der Lyse austreten und so die Plasmidaufreinigung erschweren. (Yau *et al.*, 2008)

Nach der Wahl eines geeigneten Produktionsstamms wird das entworfene Plasmid-Konstrukt in den Stamm transformiert und dieser im dritten Schritt des Produktionsprozesses kultiviert. Für die Kultivierung ist es wichtig, optimale Bedingungen zu wählen, da diese den größten Einfluss auf das Wachstum und die Plasmidausbeute haben (Carnes, 2005). Neben einer guten Sauerstoffverfügbarkeit gehört dazu ein Medium, das ein gutes Wachstum ermöglicht, so dass das Verhältnis von produzierter pDNA zu Biomasse zugunsten der Plasmide ausfällt. Hierbei ist zu beachten, dass ein zu schnelles Wachstum zu einer geringeren DNA-Ausbeute führt (Prather *et al.*, 2003).

Im Anschluss an die Kultivierung werden die Zellen geerntet. Dies kann sowohl durch Filtration oder auch Zentrifugation erfolgen.

Im letzten Schritt muss die pDNA aufgereinigt werden. Dies erfolgt in der Regel über eine Zelllyse mit einer anschließenden chromatographischen Auftrennung. Als Endpro-

dukt erhält man saubere pDNA mit einem hohen *ccc*-Anteil, die in der DNA-Vakzination eingesetzt werden kann.

Zur Zeit ist die pDNA-Ausbeute mit 2-4 g/l im Vergleich zur Produktion von proteinbasierten Pharmazeutika (10-100 g/l) sehr gering (Brand *et al.*, 2012). Diese Angaben beziehen sich auf die volumetrische Ausbeute in einem optimierten, effizienten *fed-batch*-Produktionsprozess und entsprechen einer spezifischen pDNA-Ausbeute von 51 mg/g CDW = 5 % des Zelltrockengewichts [CDW von engl. *cell dry weight*] (Williams *et al.*, 2010), (Brand *et al.*, 2012). Wird von einer pDNA-Ausbeute gesprochen, wird zwischen der volumetrischen und der spezifischen Ausbeute unterschieden. Die volumetrische Ausbeute bezieht sich auf ein Volumen und wird angegeben in [mg/l], wohingegen sich die spezifische Ausbeute auf das Zelltrockengewicht bezieht und die Einheit [mg/g·CDW] besitzt.

Ziel eines optimalen Plasmidproduktionsprozesses ist die Maximierung der volumetrischen und der spezifischen *ccc*-pDNA-Ausbeute (Williams *et al.*, 2010). Um den Einsatz von pDNA-Vakzinen erfolgreich zu machen, muss die Effizienz der Plasmidproduktionsprozesse deutlich erhöht werden. Eine wichtige Aufgabe dabei ist zu verstehen, welche Faktoren die Plasmidproduktion und somit die Ausbeute beeinflussen. Da der Prozess aus vielen aufeinander folgenden Schritten besteht, ergeben sich diverse Ansatzpunkte zur Optimierung.

Im Rahmen dieser Arbeit wird die Ebene des Wirtsstamms näher betrachtet, wobei die Faktoren untersucht werden sollen, die die Plasmidreplikation in dem *Escherichia coli* Plasmidproduktionsstamm DH5 α beeinflussen.

3.2 ZWEI EBENEN ZUR KONTROLLE DER PLASMIDREPLIKATION

Campbell & Reece (2003) beschreiben eine lebende Zelle als eine „chemische Fabrik im Miniaturformat“ in der alle Stoffwechselaktivitäten genau koordiniert und kontrolliert werden. Im Rahmen dieser Stoffwechselaktivitäten setzen die Zellen in Zuckern und anderen Betriebsstoffen gespeicherte Energie frei und benutzen sie um verschiedene Arten von Arbeit innerhalb des Zellstoffwechsels zu verrichten. Der Zellstoffwechsel kann dabei als großes Reaktionsnetzwerk betrachtet werden, das sich in viele verschiedene Stoffwechselwege aufteilt. (Campbell & Reece, 2003)

Die Plasmidreplikation von den in dieser Arbeit betrachteten *ColE1*-Plasmiden ist eine dieser Arten von Arbeit, die die Zelle unter Aufwendung von Energie verrichtet. Da die

Plasmidreplikation Energie kostet, ist es notwendig, dass sie einer strengen Kontrolle unterliegt, um nicht Plasmide im Überfluss zu synthetisieren und dadurch wertvolle Energie zu vergeuden. Betrachtet man alle bekannten Faktoren, die die Plasmidreplikation beeinflussen können, stellt man fest, dass es zwei zentrale Ebenen gibt, die den Plasmidgehalt einer Zelle regulieren (Abb. 3.3).

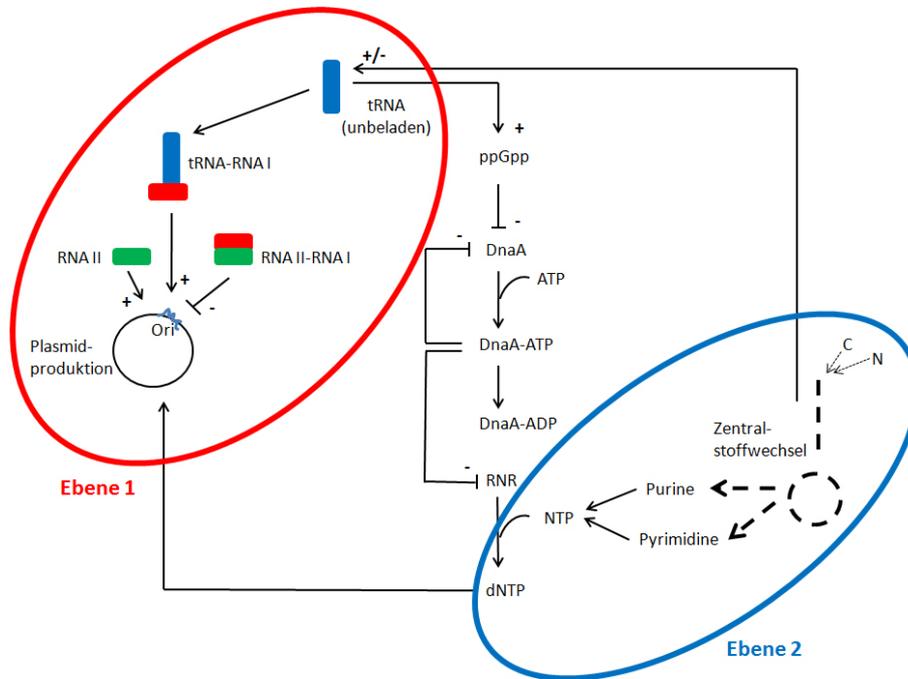


Abbildung 3.3: Zwei Ebenen zur Regulation der Plasmidkopienzahl.

Die Plasmidreplikationskontrolle kann sowohl auf Genomebene, als auch auf Metabolomebene reguliert werden. Auf Genomebene wird die Replikation des Plasmids durch die DNA-Polymerase über die Bindung von RNA-Molekülen und unbeladenen tRNA-Molekülen an Bereiche des Replikationsursprungs gesteuert (Ebene 1). Die Kontrolle über die Metabolomebene erfolgt über die Verfügbarkeit von Nucleotiden und deren Vorstufen (Ebene 2).

Die erste Ebene ist die der Plasmidreplikationskontrolle auf Genomebene. Hier wird die Replikation des Plasmids durch die DNA-Polymerase über die Bindung von RNA-Molekülen und unbeladenen tRNA-Molekülen an Bereiche des Replikationsursprungs gesteuert.

Die zweite Ebene beeinflusst die Plasmidreplikationskontrolle durch die Verfügbarkeit von Nucleotiden und deren Vorstufen und betrifft die Metabolomebene. Sind z. B. nicht ausreichend „Bausteine“ für die Synthese von Plasmid-DNA in der Zelle vorhanden, wird die Plasmidreplikation verlangsamt bzw. kommt zum Erliegen.

Aufgrund dieser beiden Möglichkeiten die Plasmidsynthese (und somit auch die Plasmidproduktion) zu regulieren, befasst sich ein Teil dieser Arbeit mit der Untersuchung

der Plasmidreplikationskontrolle auf Genomebene mit Hilfe eines mathematischen Modells. In dem anderen Teil soll die Plasmidreplikation auf Metabolomebene näher betrachtet werden. Dazu wird zum einen ein Metabolitprofil zur Untersuchung der Metabolitzusammensetzung in An- und Abwesenheit eines Plasmids erstellt und zum anderen der Metabolitfluss durch ^{13}C -Flussmessungen analysiert.

3.3 PLASMIDREPLIKATIONSKONTROLLE VON *colE1*-ÄHNLICHEN PLASMIDEN

Die Plasmidreplikationskontrolle spielt eine wichtige Rolle für die Plasmidproduktion, da sie die Plasmidanzahl pro Zelle, die sogenannte Plasmidkopienzahl (PCN von engl. *plasmid copy number*) reguliert. An dieser Stelle soll zunächst kurz die Plasmidreplikation erklärt werden. Als Plasmidreplikation bezeichnet man die Vervielfältigung der pDNA in einer Zelle (Abbildung 3.4).

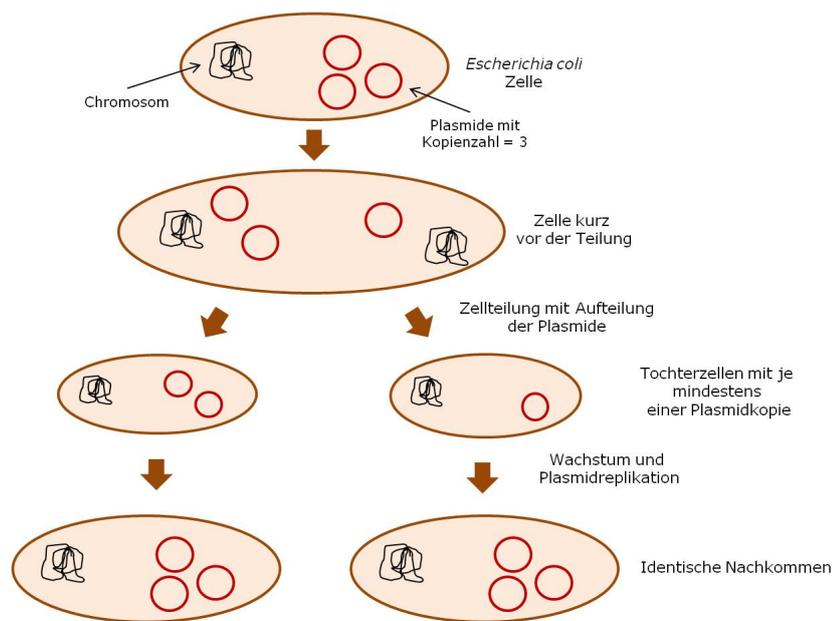


Abbildung 3.4: Überblick über die Plasmidreplikation.

Mit Plasmidreplikation bezeichnet man die Vervielfältigung der Plasmid-DNA in einer Zelle. Plasmide liegen daher in der Regel in mehreren Kopien in der Zelle vor (hier Plasmidkopienzahl = 3). Auf diese Weise wird gewährleistet, dass, bevor sich die Zelle teilt, mindestens ein Plasmid an jede Tochterzelle weitergegeben wird. Nach erfolgreicher Zellteilung beginnt jede Tochterzelle zu wachsen und repliziert ihre pDNA bis wieder eine bestimmte Kopienzahl erreicht ist.

Anhand von Abbildung 3.4 ist zu erkennen, dass Plasmide in mehreren Kopien in der Zelle vorliegen. Man unterscheidet dabei zwischen sogenannten *high copy*- und *low copy* Plasmiden. Ein *high copy* Plasmid besitzt im Gegensatz zum *low copy* Plasmid eine

höhere Kopienzahl pro Zelle. Vor der Zellteilung wird gewährleistet, dass mindestens ein Plasmid an jede Tochterzelle weitergegeben wird. Nach erfolgreicher Zellteilung beginnt jede Tochterzelle zu wachsen und repliziert ihre pDNA, damit sie, kurz bevor sie sich selbst teilt, sicherstellen kann, dass auch jede ihrer Tochterzellen mindestens ein Plasmid erhält.

Da die Mehrheit der therapeutisch angewandten Plasmide auf pBR322- oder pUC-Plasmiden basieren und diese, mit dem pMB1-Replikon, einen *ColE1*-ähnlichen Replikationsursprung besitzen (Williams *et al.*, 2010) (Bhagwat & Person, 1981), beschäftigt sich diese Arbeit mit der Untersuchung der *ColE1*-Plasmidreplikationskontrolle.

Ein Grund für die verbreitete Anwendung von *ColE1*-ähnlichen Vektoren als Gerüst für DNA-Vakzine ist, dass diese Vektoren in den vergangenen Jahrzehnten intensiv analysiert wurden. Sie werden zudem in der Mehrzahl der Produktionsprozesse für rekombinante Proteine eingesetzt, so dass die Produktionsbedingungen bereits sehr gut untersucht und eine Vielzahl an pUC-Vektor-Konstrukten erhältlich sind. Des Weiteren lassen sich, aufgrund früherer Modifikationen, mit *ColE1*-ähnlichen Plasmiden hohe Kopienzahlen erzielen. (Prather *et al.*, 2003)

Generell kontrollieren bakterielle Plasmide ihre *PNC* durch negativ wirkende Kontrollsysteme, die die Replikationsrate an sich verändernde Bedingungen anpassen (del Solar & Espinosa, 2000). Bei *ColE1*-ähnlichen Plasmiden geschieht dies durch ein inhibitorisch wirkendes *antisense* RNA-Molekül, das die Bildung eines RNA-Primermoleküls verhindert (Grabherr & Bayer, 2002).

Die Replikation von *ColE1*-Plasmiden verläuft unidirektional und benötigt kein plasmidkodiertes Initiatorprotein (Williams *et al.*, 2010). Es kann zwischen drei möglichen Szenarien unterschieden werden: Im ersten Fall (Abbildung 3.5 A) bindet ein Teil eines plasmidkodierten RNA-Vorläufer-Transkriptes (RNAII) an den Replikationsursprung, wo es eine R-Schleife bildet und einen persistenten Hybridkomplex mit der pDNA formt (Williams *et al.*, 2010) (Itoh & Tomizawa, 1980). Da das 3'-Ende der RNAII variabel ist, entstehen bei der Transkription RNAII-Moleküle verschiedener Längen. Diese unterschiedlich langen RNAII-Transkripte werden im nächsten Schritt von dem Enzym RNase H gespalten und somit auf eine Länge gebracht. Dabei wird eine 'OH-Gruppe frei und die RNAII kann als Primer dienen. RNase H ist ein Enzym, das spezifisch DNA-RNA-Hybridmoleküle erkennt und die RNA hydrolysiert. (Schuhmann, 1990)

Die Modifikation der RNAII ist essentiell für die Replikation, da ansonsten die Bildung des sogenannten *leading*-Stranges durch die DNA Polymerase I nicht starten kann (Itoh

& Tomizawa, 1980). Anschließend wird die Elongation durch die DNA PolymeraseIII fortgesetzt und das Plasmid wird repliziert.

Bei einer Hemmung der Replikation wird die Modifikation der RNAII durch die RNase H verhindert. Dies ist im zweiten Szenario, in Abbildung 3.5 B gezeigt: Ein kleines komplementäres RNA-Transkript (RNAI) bindet an die RNAII, so dass diese nicht gespalten werden und als Primer dienen kann (Tomizawa & Itoh, 1981). Die beiden RNA-Transkripte stammen von derselben DNA-Region, nahe dem Replikationsursprung, wobei RNAI konstitutiv in *antisense*-Richtung transkribiert wird. Durch die Bindung von RNAI an RNAII entsteht ein transienter, sogenannter *kissing complex*, wodurch die Primer-Reifung blockiert wird. Auf diese Weise reguliert die RNAI die Initiation der Replikation. (Tomizawa, 1984) (Tomizawa, 1985) Zusätzlich wird dieser RNAI-RNAII-pDNA-Komplex durch die Bindung des Rom-Proteins (Rom = *RNAI modulating protein*; auch als Rop-Protein bezeichnet von Rop = *Repressor of primer*) stabilisiert. (Tomizawa, 1986)(Tomizawa, 1990b)

Des Weiteren kann die Plasmidreplikation unter Aminosäuremangel von unbeladenen tRNA-Molekülen beeinflusst werden (Abbildung 3.5 C). Unbeladene tRNA-Moleküle können an RNAI- oder RNAII-Moleküle binden, mit der Konsequenz, dass kein *kissing complex* entstehen kann (Grabherr & Bayer, 2002). Da die Zelle unter diesen Mangelbedingungen weiterhin Aminosäuren benötigt, werden die restlichen an tRNAs gebundenen Aminosäuren verbraucht und die Anzahl unbeladener tRNAs in der Zelle nimmt zu. Gleichzeitig wurde eine Erhöhung der pDNA-Konzentration beobachtet, was bedeutet, dass eine hohe Konzentration unbeladener tRNAs mit einer hohen PCN korreliert. Wróbel & Węgrzyn (1998) konnten in ihren Experimenten zeigen, dass die Zunahme der pDNA-Konzentration unter Leucin-Mangel am größten war. (Wróbel & Węgrzyn, 1998)

In der Literatur existieren zwei Modelle, die eine mögliche Interaktion zwischen unbeladenen tRNAs und den regulatorischen RNAs beschreiben. Das erste Modell wurde von Yavachev & Ivanov (1988) entwickelt. In ihrer Arbeit fanden die beiden Autoren strukturelle Ähnlichkeiten zwischen den Schleifen von RNAI beziehungsweise RNAII und der Kleeblattstruktur von tRNAs. Zusätzlich entdeckten sie Sequenzhomologien über 40 % zwischen den Anticodon-Schleifen von tRNAs und den zweiten Schleifen von den RNA-Molekülen. Zu den tRNAs mit einer Ähnlichkeit zur RNAI-Schleife gehören Arginin-, Histidin-, Leucin-, Lysin-, Phenylalanin- und Threonin-tRNAs. Ähnlichkeit mit der RNAII-Schleife besitzen Tyrosin- und Glycin-tRNAs. Die tRNA-Moleküle für

Methionin und Valin zeigen eine Homologie von über 40 % mit den Schleifen beider RNA-Moleküle. Aufgrund dieser strukturellen und sequenziellen Ähnlichkeiten schlugen Yavachev & Ivanov (1988) vor, dass die Bindung zwischen der Anticodon-Schleife einer unbeladenen tRNA und der zweiten Schleife der regulatorischen RNA erfolgt. (Yavachev & Ivanov, 1988)

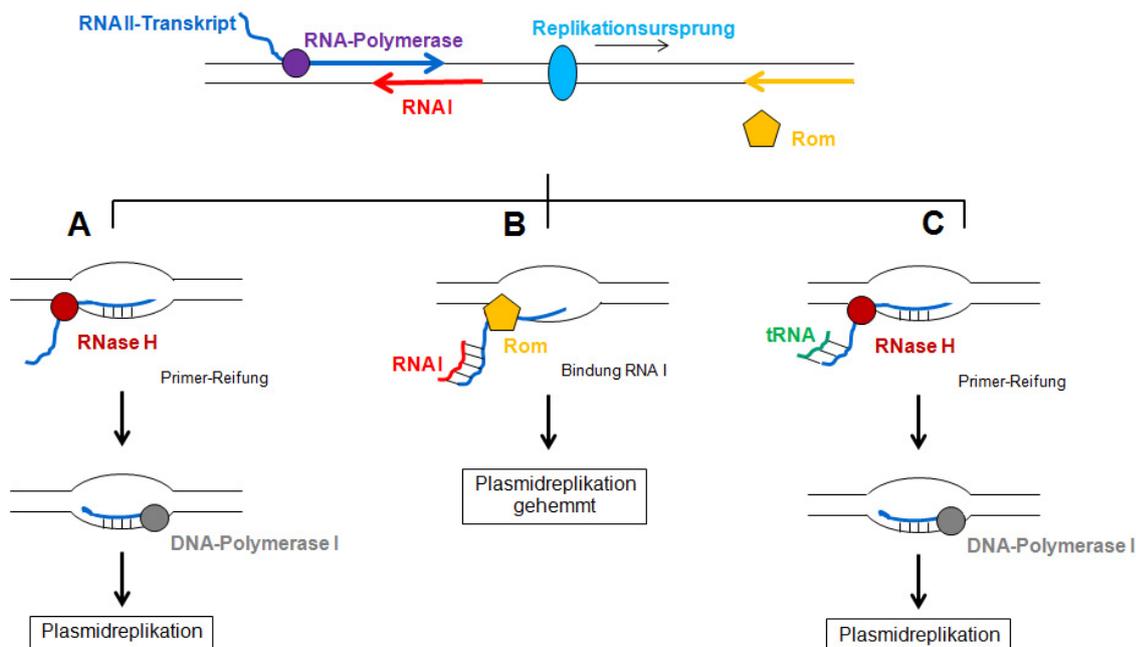


Abbildung 3.5: *ColE1*-Plasmidreplikationskontrolle.

Die Synthese eines RNA-Vorläufer-Transkripts durch die RNA-Polymerase (RNAP) ist für die Plasmidreplikation ein essentieller Schritt. In Abwesenheit einer Interaktion mit dem RNAI-Transkript (A) formt die RNAII einen persistenten Hybridkomplex mit dem Replikationsursprung der Plasmid-DNA. Dieser Komplex wird von der RNase H gespalten, so dass das RNAII-Transkript als Primer dient und die Replikation starten kann. Findet eine Interaktion zwischen der inhibitorischen RNAI und der dazu komplementären Region auf der RNAII statt (B), bildet sich ein transienter Komplex, wodurch die Primer-Reifung verhindert wird. Bei einem *low copy* Plasmid wird dieser Komplex durch das Rom-Protein stabilisiert. Neben der RNAI können unter Nährstoffmangelbedingungen auch unbeladene tRNAs mit der RNAII interagieren (C). Durch diese Interaktion wird die Bindung einer RNAI verhindert, wobei die Primer-Reifung durch die RNase H trotzdem erfolgen und die Plasmidreplikation stattfinden kann. Es gibt auch unbeladene tRNAs, die mit RNAI-Molekülen interagieren und so deren inhibitorische Wirkung aufheben. Da diese jedoch in dem Modell dieser Arbeit nicht berücksichtigt werden, ist diese Interaktion in der obigen Abbildung nicht dargestellt.

Ein zweites Modell zur tRNA-RNA-Interaktion wurde von Wang *et al.* (2002) vorgeschlagen. Hierbei wird angenommen, dass die Bindung zwischen der Aminosäurebin-

destelle der unbeladenen tRNA und der regulatorischen RNAI stattfindet. (Wang *et al.*, 2002) Dieses Modell der Interaktion konnte von Wang *et al.* (2006c) bestätigt werden.

3.4 METABOLOMANALYSEN

Die DNA als Träger der Erbinformation hat eine grundlegende Funktion in jeder lebenden Zelle. Sie besteht aus vier verschiedenen Nukleotiden, wobei jedes Nukleotid aus einem Phosphatrest, der Desoxyribose und einer von den vier organischen Basen Adenin, Thymin, Guanin oder Cytosin aufgebaut ist. Diese einzelnen Bausteine werden über verschiedene Stoffwechselwege synthetisiert, so dass die DNA als ein Produkt des Stoffwechsels betrachtet werden kann.

Seit einigen Jahren beschäftigten sich viele Forschergruppen mit den sogenannten „-omik“-Technologien, mit dem Ziel, einer nicht-zielgerichteten Identifizierung aller Genprodukte (Transkripte, Proteine und Metabolite), die in einer biologischen Probe zu finden sind (Weckwerth, 2003). Dabei teilen sich die einzelnen „-omik“-Gebiete in Genomik, Transkriptomik, Proteomik und Metabolomik auf. Die Genomik beschäftigt sich hierbei mit dem kompletten Genom, welches als ein statisches Gebilde betrachtet werden kann. Das Genom liegt physikalisch vor und ist durch die Anzahl, Art und Abfolge seiner Nukleotide genau definiert. Das Transkriptom und das Proteom sind hingegen hoch dynamisch, da sie die Gen- bzw. die Proteinexpressionsmuster innerhalb eines biologischen Systems nie allgemein, sondern immer nur unter bestimmten Bedingungen darstellen. Systembiologisch betrachtet, schließen Transkriptomik, Proteomik und Metabolomik die Lücke zwischen Genotyp und Phänotyp. (Lottspeich & Engels, 2006)

Die Metabolomik ist die jüngste Disziplin und wird als Abschluss der „-omik“-Kaskade betrachtet (Dettmer & Hammock, 2004). Sie umfasst die Analyse von niedermolekularen, nichtpolymeren organischen Stoffwechselprodukten (sogenannte Metabolite) eines biologischen Systems unter definierten Bedingungen. Als Metabolom bezeichnet man die Gesamtheit dieser Metabolite. (Lottspeich & Engels, 2006) Nach Fiehn (2002) lässt sich das Metabolom auch als Endprodukt von zellulären Prozessen und somit als direkte Antwort auf veränderte Bedingungen bezeichnen.

Das Metabolom einer Zelle besteht aus einer hohen Anzahl verschiedenster Metabolite mit unterschiedlichen chemischen und physikalischen Eigenschaften, die zugleich in unterschiedlich hohen Konzentrationen vorliegen. Es ist daher nicht möglich mit

einer einzigen Analytik alle Metabolite zu bestimmen, aber man kann verallgemeinern, dass Metabolomuntersuchungen in erster Linie auf Kernspinresonanzspektroskopie (NMR-Spektroskopie von engl. *nuclear magnetic resonance spectroscopy*) und auf einer chromatographischen Auftrennung mit anschließender Massenspektrometrie basieren. (Dettmer *et al.*, 2007) (Fiehn, 2002)

Bei der Untersuchung des Metaboloms gibt es verschiedene Vorgehensweisen: Bei der Erstellung eines Metabolitprofils (engl. *metabolite profiling*) konzentriert sich die Analyse auf eine Gruppe von Metaboliten innerhalb eines bestimmten Stoffwechselweges oder auf eine bestimmte Metabolitklasse, wie z. B. Aminosäuren.

Die zweite Vorgehensweise ist die Erstellung eines metabolischen Fingerabdrucks, das sogenannte *metabolite fingerprinting*. Hierbei ist nicht das Ziel jeden einzelnen Metaboliten zu identifizieren, sondern mit dieser Analysemethode möchte man Metabolitmuster detektieren. Bestimmte Metabolitmuster können als Antwort des biologischen Systems auf Krankheiten oder andere Einflüsse entstehen, so dass *metabolite fingerprinting* Untersuchungen helfen, eine Probe anhand ihrer Herkunft einzuordnen (Kruger *et al.*, 2008). Im nächsten Schritt folgt die Identifizierung der Metabolite, die für die Unterschiede in den Metabolitmustern verantwortlich sind. Die Identifizierung erfolgt dabei über die zugehörigen Retentionszeiten und Massenspektren (Kind *et al.*, 2009). *Metabolite fingerprinting* Analysen wurden bereits mit verschiedenen biologischen Proben durchgeführt, wie z. B. Urin, Plasma, Serum oder Gewebezellen, mit dem Ziel Biomarker im Metabolom zu finden, die spezifisch für eine bestimmte Erkrankung sind. Das langfristige Ziel dabei ist, das *metabolite fingerprinting* zur Diagnose von Erkrankungen eines Patienten aufgrund seines veränderten metabolischen Fingerabdrucks zu verwenden.

Neben dem *metabolite fingerprinting* gibt es noch das *metabolite footprinting* (Erstellung eines metabolischen Fußabdrucks). Hierbei handelt es sich um das gleiche Vorgehen wie beim *metabolite fingerprinting*, mit dem Unterschied, dass bei diesem Vorgehen extrazelluläre Metabolite untersucht werden. Diese Metabolite stammen aus dem Zellkulturmedium und dienen als Nachweis, welche Metabolite von den Zellen abgegeben oder aufgenommen werden. (Dettmer *et al.*, 2007) (Fiehn, 2002)

In dieser Arbeit soll das Stoffwechselprodukt DNA mittels Metabolomanalysen untersucht werden, indem ein Metabolitprofilerstellt wird. Dies erfolgt unter Anwendung von Gaschromatographie-Massenspektrometrie (GC-MS) mit anschließender Analyse der Chromatogramme durch die MeltDB-Software (Kessler *et al.*, 2013). Das Ziel ist mögliche Unterschiede im Metabolom eines plasmidfreien und zwei plasmidtragende

E. coli-DH5 α -Stämmen zu entdecken. Der Fokus liegt dabei auf Vorstufen in der Nukleotidbiosynthese, insbesondere auf den 5-fach-Zuckern.

3.5 ¹³C-FLUSSANALYSEN

Bei der Durchführung von Metabolomanalysen werden Metabolitpools unter bestimmten Bedingungen in der Zelle gemessen. Diese Messungen beschreiben die relativen intrazellulären Metabolitkonzentrationen zu einem definierten Zeitpunkt; man erhält sozusagen eine Momentaufnahme. In diesen Daten sind keine dynamischen Informationen darüber enthalten, über welche Stoffwechselwege die gemessenen Metabolitpools mit der Zeit umgewandelt werden. An diese Informationen gelangt man mittels metabolischen Stoffflussanalysen. (Winter & Krömer, 2013)

Metabolische Flussanalysen verwenden *in vivo* Messungen von Stoffwechselflüssen, die in ein stoichiometrisches Modell integriert werden. Mithilfe des Modells lassen sich anschließend die absoluten Flüsse des zentralen Kohlenstoff-Metabolismus berechnen. Dabei stellen die in dem Modell definierten Stöchiometrien der einzelnen Reaktionen Bedingungen dar, die bei der Berechnung der Metabolitflüsse berücksichtigt werden müssen. (Winter & Krömer, 2013) (Orth *et al.*, 2010)

Um den Glukosefluss innerhalb verschiedener *E. coli*-Plasmidproduktionsstämme zu analysieren, werden in dieser Arbeit zusätzlich zu den Metabolomanalysen ¹³C-Flussanalysen durchgeführt. Von besonderem Interesse ist dabei der Glukosefluss in den Pentosephosphatweg (PPP von engl. *pentose phosphate pathway*), da hier die Nukleotidvorstufe Ribose-5-Phosphat entsteht.

Die verwendeten Produktionsstämme wurden von unserem Kooperationspartner Prof. Dr. Alvaro R. Lara von der *Universidad Nacional Autónoma de México* zur Verfügung gestellt und sind Derivate des *E. coli*-Stamms W3110. Die Stämme besitzen verschiedene genetische Manipulationen und unterscheiden sich in der pDNA- und Acetatproduktion. Der Wildtyp *E. coli* W3110 produziert mit $1,14 \pm 0,10$ mg/l die geringste Menge pDNA. Zusätzlich entstehen dabei $0,32 \pm 0,03$ g/l Acetat (Borja *et al.*, 2012).

Eine Akkumulation von Acetat im Medium wirkt sich nachteilig auf den Produktionsprozess aus, da es das Zellwachstum hemmt (Bauer *et al.*, 1990). Acetat wird hauptsächlich über den Phosphotranscetylase-Acetat-Kinase-Stoffwechselweg synthetisiert, wobei Acetyl-Coenzym A (Acetyl-CoA) als Substrat dient. Die Acetatproduktion in *E. coli* er-

folgt unter Sauerstoffmangel-Bedingungen oder bei einem Wachstum unter sehr hohen Glukosekonzentrationen im Medium. Durch diese Bedingungen kommt es zu einem Ungleichgewicht zwischen dem Glykolyse- und dem Citratzyklusfluss. Glukose wird dabei sehr schnell von den Zellen aufgenommen und über die Glykolyse verstoffwechselt, so dass viel Acetyl-CoA entsteht. Da dies nicht schnell genug über den Citratzyklus metabolisiert werden kann, stellt der Phosphotranscetylase-Acetat-Kinase Stoffwechselweg ein Ventil dar, über das Acetyl-CoA als Acetat von der Zelle abgegeben wird. (Majewski & Domach, 1990) (Delgado & Liao, 1997) (Farmer & Liao, 1997)

Dieser sogenannte *overflow*-Metabolismus stellt eine Verschwendung der Zelle an energiereichen Molekülen dar und geschieht unter aeroben Bedingungen nur bei Glukoseüberfluss.

Um die Acetatproduktion zu minimieren und gleichzeitig die Plasmidproduktion zu erhöhen, haben Prof. Dr. Lara und seine Arbeitsgruppe eine Reihe von Mutanten erstellt. Zunächst wurde in dem Wildtypstamm *E. coli* W₃₁₁₀ das Phosphotransferasesystem (PTS), welches dem Glukoseimport dient, inaktiviert. Im nächsten Schritt wurde die Promoterregion des chromosomalen Galaktosepermeasegens *galP* durch einen starken Promoter ersetzt. Die modifizierte Galactosepermease stellt einen alternativen, effizienteren Glukosetransport dar, wodurch der resultierende *E. coli* Stamm VH₃₃ mit $0,14 \pm 0,01$ g/l eine stark verringerte Acetatproduktion und eine erhöhte Plasmidproduktion ($2,78 \pm 0,10$ mg/l) zeigt. Die verwendeten alternativen Glukoseaufnahmesysteme sowie die damit verbundenen metabolischen Stoffwechselwege zur Acetatbildung sind in der Abbildung 3.6 graphisch dargestellt. (De Anda *et al.*, 2006) (Borja *et al.*, 2012)

Zur Erhöhung der Plasmidproduktion wurden weitere genetische Modifikationen an dem Stamm VH₃₃ vorgenommen. Die Gene *recA*, *deoR*, *nupG* und *endA* wurden einzeln und in verschiedenen Kombinationen inaktiviert (Borja *et al.*, 2012).

In dieser Arbeit wurden die Stämme W₃₁₁₀, VH₃₃, VH₃₃Δ *recA*, VH₃₃Δ (*recA deoR*) mit einem 6,1 kb Plasmid transformiert und anschließend mittels ¹³C-Flussanalysen untersucht. Das *recA* Gen kodiert für eine Rekombinase A. Das Ausschalten von *recA* verringert das Auftreten ungewollter Rekombinationsereignisse und stabilisiert den Plasmidvektor durch eine Herabsetzung der Oligomerbildung. Der Stamm VH₃₃Δ *recA* produziert $7,09 \pm 0,10$ mg/l pDNA und akkumuliert $0,08 \pm 0,01$ g/l Acetat. Durch eine zusätzlich Mutation im *deoR* Gen kann die Plasmidproduktion weiter verbessert werden. Die Doppelmutante VH₃₃Δ (*recA deoR*) produziert mit $7,98 \pm 0,10$ mg/l pDNA 12 % mehr pDNA als VH₃₃Δ *recA*. Die Inaktivierung des *deoR* Gens bewirkt eine Zunahme

der dNTP-Synthese und somit einer Erhöhung der produzierten Plasmid-DNA. (Borja *et al.*, 2012)

Die oben erwähnten Angaben zur Plasmid- und Acetatproduktion beziehen sich auf eine Kultivierung im Schüttelkolben. Bei den beschriebenen Stämmen soll nun mittels ¹³C-Flussanalysen untersucht werden, ob eine hohe Plasmidproduktionsrate mit einem erhöhten Glukosefluss über den Pentosephosphatweg in Richtung Ribose-5-Phosphat einhergeht.

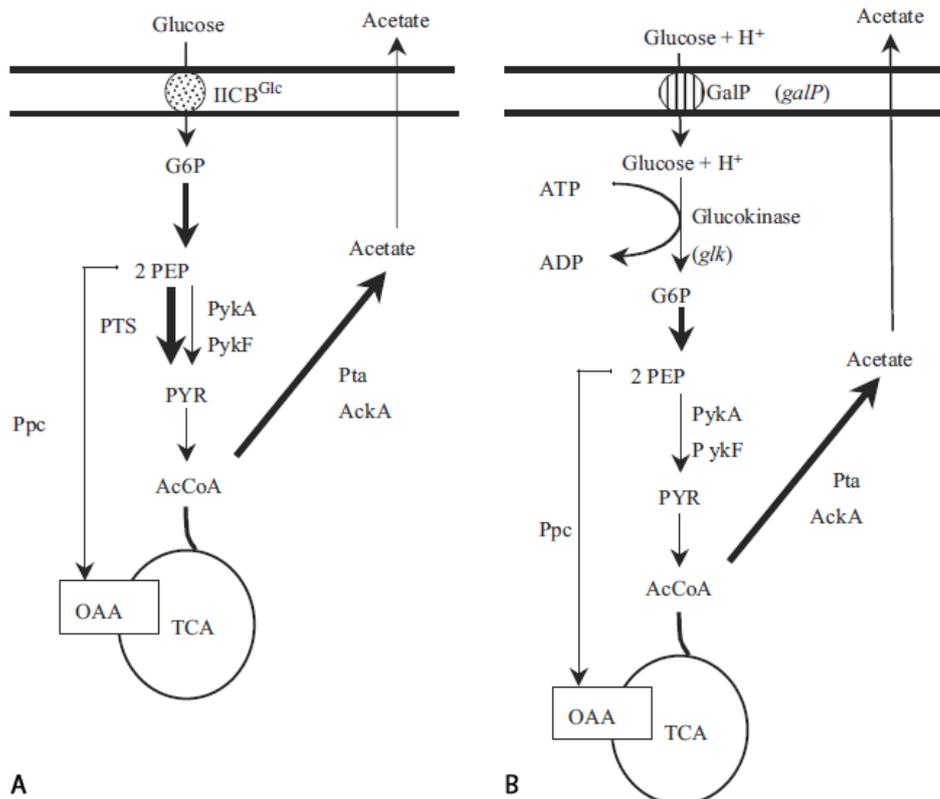


Abbildung 3.6: Glukoseimportsysteme sowie die damit verbundenen metabolischen Stoffwechselwege zur Acetatbildung.

(A) Glukoseaufnahme in dem Wildtypstamm *E. coli* W₃₁₁₀ über das Phosphotransferasesystem. Die dick dargestellten Pfeile beschreiben multiple enzymatische Reaktionen. Die Zelle verbraucht ein Mol Glukose um 2 Mol Phosphoenolpyruvat (PEP) zu synthetisieren. Ein Mol PEP wird durch das PTS zur Phosphorylierung und zum Transport von Glukose verwendet, wobei ein zusätzliches Mol Pyruvat entsteht. In dem *E. coli* Stamm VH₃₃ ist das PTS inaktiviert, so dass diese Reaktionen hier nicht mehr stattfinden (B). Die Glukoseaufnahme erfolgt hier über eine modifizierte Galactosepermease. Abkürzungen: G6P, Glukose-6-phosphat; PEP, Phosphoenolpyruvat; PYR, Pyruvat; AcCoA, Acetyl-CoA; Ppc, Phosphoenolpyruvat Carboxylase; TCA, Citratzyklus; PykA und PykF, zwei Pyruvatkinasen; Pta, Phosphotransacetylase; AckA, Acetatekinase; IICB^{Glc}, Glukose spezifischer PTS Membran Transporter. (De Anda *et al.*, 2006)

Das Prinzip von Flussanalysen mit ¹³C-markierter Glukose ist die indirekte Berechnung intrazellulärer Reaktionsgeschwindigkeiten durch eine Computer-modellbasierte Interpretation von stabilen Isotopenmustern in Stoffwechselprodukten (Sauer, 2006).

Die Durchführung besteht aus einem theoretischen und aus einem experimentellen Teil (Abbildung 3.7). Der theoretische Teil beinhaltet die Erstellung eines stöchiometrischen Modells und die Simulation der metabolischen Stoffwechselflüsse. Zunächst wird ein metabolisches Modell des Zentralstoffwechsels erstellt. Dieses Modell enthält Informationen zur Stöchiometrie der beteiligten Reaktionen und zur Weitergabe der Kohlenstoffatome des ¹³C-markierten Glukosemoleküls entlang der einzelnen Stoffwechselwege. Auf diese Weise ist die Weitergabe jedes C-Atoms festgelegt und somit das Markierungsmuster für jedes Stoffwechselzwischenprodukt bestimmt. Zusätzlich sind Biomassereaktionen mit bestimmten Raten definiert. Dadurch wird festgelegt, wieviel von dem aufgenommenen Substrat für den Aufbau von Biomasse verwendet wird. Die Definition von Biomassereaktionen ist notwendig, da das Wachstum einer Zelle garantiert sein muss. Die dafür notwendigen Moleküle werden dem Stoffwechsel entzogen und stehen somit nicht zur Produktion eines z. B. biotechnologisch relevanten Produktes zur Verfügung. (Wiechert, 2001) (Sauer, 2004)

Der experimentelle Teil besteht aus der Kultivierung der Zellen auf dem ¹³C-markiertem Substrat und der anschließenden Bestimmung der Markierungsmuster in den Stoffwechselprodukten. Er beginnt mit dem Wachstum der Zellen auf dem markierten Substrat, bis die Isotopenmarkierung im gesamten Stoffwechsel verteilt ist und ein Fließgleichgewicht angenommen werden kann. Abhängig von der metabolischen Flussverteilung in der Zelle entstehen spezifische Markierungsmuster in den Stoffwechselzwischenprodukten. (Wiechert, 2001) (Sauer, 2004)

Um die metabolische Flussverteilung berechnen zu können, müssen die spezifischen Markierungsmuster in den Stoffwechselprodukten bestimmt werden. Eine Möglichkeit ist die Messung der Isotopomer-Massenverteilung der intrazellulären Aminosäuren mittels Gaschromatographie-Massenspektrometrie. Dabei werden die Zellen nach der Kultivierung geerntet und es folgt die Metabolitextraktion mit anschließender Derivatisierung. Die Derivatisierung ist notwendig da die meisten für Flussanalysen relevanten, Metabolite geladen und somit nicht ausreichend flüchtig für die GC-MS-Analytik sind. Dies gilt insbesondere für Aminosäuren, organische Säuren oder Zucker. Bei der Derivatisierung findet eine Silylierung, Alkylierung oder Acetylierung der Metabolite statt, wodurch deren Flüchtigkeit erhöht wird. Für die Analyse von Aminosäuren bietet sich ei-

ne Silylierung mit N-Methyl-N-t-Butyldimethylsilyl-Trifluoro-Acetamid (MBDSTFA) an, da dies zu Derivaten mit besonders guter Signalintensität des $(M - 57)^+$ -Fragment-Ions führt. Im Anschluss an die Derivatisierung folgt die Vermessung via GC-MS. (Wittmann, 2007)

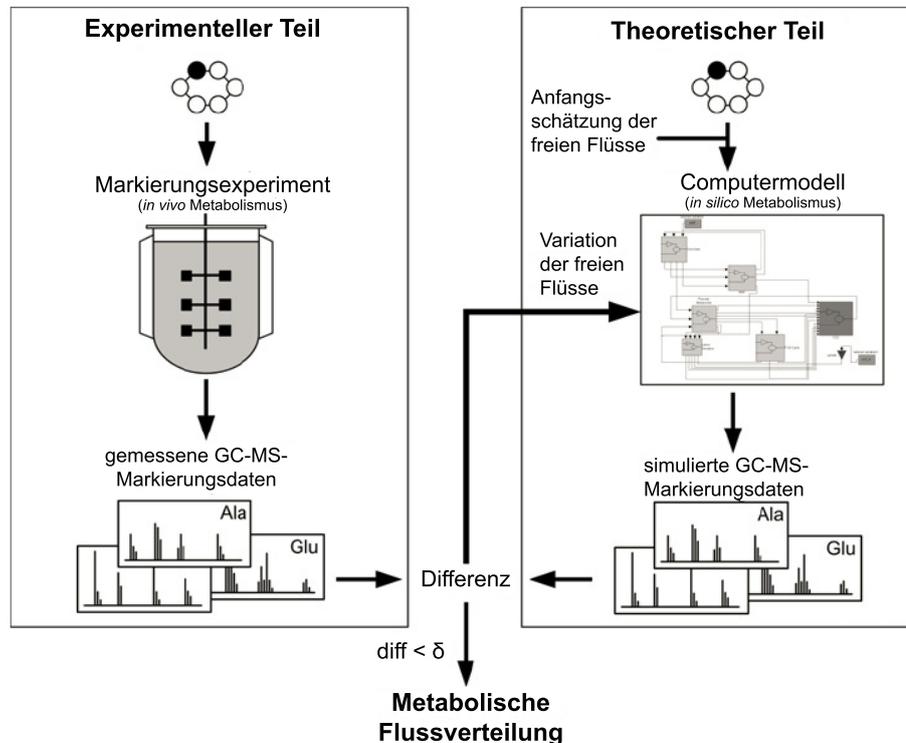


Abbildung 3.7: Durchführung von ¹³C-Flussanalysen.

Die Durchführung von ¹³C-Flussanalysen besteht aus einem experimentellen und einem theoretischen Teil. Der experimentelle Teil beinhaltet die Kultivierung der Zellen auf dem ¹³C-markiertem Substrat und die anschließende Bestimmung der Markierungsmuster in den Stoffwechselprodukten mittels GC-MS-Analysen. Der theoretische Teil beinhaltet die Erstellung eines stöchiometrischen Modells und die Berechnung der metabolischen Stoffwechselflüsse. Die Berechnung der metabolischen Flüsse basiert auf der Minimierung der Differenz (δ) zwischen den gemessenen und den simulierten Markierungsmustern. (Wittmann, 2007)

Nach der Erstellung des Modells und der Vermessung der Isotopomeren-Massenverteilung der Aminosäuren folgt die Bestimmung des metabolischen Flusses innerhalb der Zelle. Dazu werden Daten der gemessenen Massenverteilung in das metabolische Modell integriert. Der Lösungsalgorithmus beginnt mit einer beliebig gewählten, metabolischen Flussverteilung und berechnet die daraus entstehende Isotopomerverteilung. Diese wird mit den realen Daten verglichen und die Differenz δ wird berechnet. Unterscheiden sich die gemessenen Daten von den simulierten Daten, wird ein Optimierungsalgorithmus angewendet, der eine veränderte Flussverteilung wählt und die

Isotopomerverteilung erneut simuliert. Anschließend werden die Simulationen wieder mit dem Experiment verglichen und die Differenz berechnet. Dies ist ein iterativer Prozess, der so lange abläuft, bis mit der gewählten Flussverteilung eine Isotopomerverteilung simuliert werden kann, die mit den experimentell bestimmten Massenverteilungen übereinstimmt, bzw. bis die Differenz δ minimiert ist. Das bedeutet, die resultierende Flussverteilung basiert auf der Minimierung der Abweichung (δ in 3.7) zwischen den gemessenen und den simulierten Massenverteilungen. (Wiechert, 2001)

THEORETISCHER HINTERGRUND

4.1 MODELLIERUNG IN DER SYSTEMBIOLOGIE

Biologische Systeme, wie z. B. komplexe Biomoleküle, Zellen oder gesamte Organismen besitzen in ihren Strukturen und Funktionen einen hohen Organisationsgrad. Es ist daher für das Verständnis dieser Systeme notwendig, die beteiligten Komponenten nicht nur einzeln, sondern in ihrem Kontext zu untersuchen. (Klipp *et al.*, 2009) Ein wichtiges Werkzeug zum Verständnis dieser Interaktionen innerhalb biologischer Systeme ist die mathematische Modellierung.

Vor gut 10 Jahren ist mit der Systembiologie eine neue Forschungsdisziplin entstanden, mit dem Ziel, biologische Systeme ganzheitlich zu betrachten und zu verstehen (Chuang *et al.*, 2010). Um dieses Ziel zu erreichen, ist die Expertise aus vielen verschiedenen Wissensgebieten, wie Mathematik, Physik, Chemie, Biologie und Bioinformatik notwendig. Die Systembiologie ist dadurch ein sehr interdisziplinäres Forschungsgebiet und bietet einen guten Rahmen, um die Struktur biologischer Systeme mithilfe von Modellen darzustellen und ihre Dynamik zu simulieren. Die Simulationen dienen zum einen dazu, aufgrund neu gewonnener Erkenntnisse weiterführende *in vivo* und *in vitro* Experimente zu designen. Zum anderen helfen sie, die Modelldefinitionen anzupassen und gegebenenfalls zu verbessern, um das reale Systemverhalten stetig besser darstellen und verstehen zu können. (Kherlopian *et al.*, 2008)

Die Modelle in der Systembiologie basieren oft auf bekannten physikalischen Gesetzen, wie z. B. den Gesetzen der Thermodynamik oder auf dem Massenwirkungsgesetz. Neben dieser mathematischen Beschreibung bilden experimentell bestimmte Daten eine weitere wichtige Grundlage, um das biologischen System möglichst realistisch beschreiben zu können. (Klipp *et al.*, 2009) Grob gesprochen kann man sich dabei die experimentell bestimmten Daten als Bausteine und die mathematische Modellierung als Zement vorstellen, der die Bausteine zusammenfügt (Kherlopian *et al.*, 2008).

Die Modellierung besitzt den Vorteil, dass sie eine Unabhängigkeit von dem modellierten System bietet. Das bedeutet, das Systemverhalten kann mit einem einzigen Modell

unter verschiedenen Bedingungen untersucht und beliebig oft wiederholt werden. Zusätzlich kann die Zeit in Abhängigkeit der zu simulierenden Bedingungen *ad libitum* gedehnt oder gerafft werden. Dies ist beispielsweise sehr nützlich, will man die Entwicklung eines Baumbestandes untersuchen. Die Durchführung der entsprechenden Versuche würde Jahrzehnte dauern, wohingegen die Simulation des Baumwachstums mit Hilfe eines Modells schnell neue Erkenntnisse liefern kann. Ein adäquates Modell bietet somit für die experimentelle Arbeit eine gute Ergänzung, da auf diese Weise verschiedene Szenarien getestet werden können, die zu viel Zeit in Anspruch nehmen würden oder u. U. experimentell nicht zugänglich sind. So lassen sich mit Hilfe der Modellierung zum Beispiel zeitliche Verläufe verfolgen, die aufgrund ihrer hohen Geschwindigkeit nicht messbar wären. (Klipp *et al.*, 2009) (Bossel, 2004)

Zusammenfassend kann man sagen, dass die Verwendung von Modellen in der Systembiologie eine wichtige Rolle in der Überprüfung wissenschaftlicher Hypothesen spielt und dazu dient, neue Erkenntnisse zu liefern, die bei Betrachtung des Systems ohne Verhaltenssimulationen nicht direkt ableitbar wären (Klipp *et al.*, 2009) (Bossel, 2004).

4.2 WAS IST EIN MODELL?

Beschäftigt man sich mit dem Gebiet der Modellierung, stellt sich zwangsläufig die Frage: „Was ist ein Modell?“. Eine allgemeine Definition des Modellbegriffs lautet: Ein Modell ist eine abstrakte Darstellung von Objekten oder Prozessen, welche die Eigenschaften dieser erklärt (Klipp *et al.*, 2009). Da Modelle in vielen verschiedenen Fachrichtungen verwendet werden, hängt eine genaue Definition von der jeweiligen Disziplin ab, in der das Modell erstellt wird. In der Zahnmedizin z. B. versteht man unter einem Modell die Abbildung eines Gebisses, wohingegen der Künstler an eine Person denkt, die jemandem Modell steht. (Klipp *et al.*, 2009)

Wissenschaftliche Modelle kann man prinzipiell in vier Kategorien unterteilen (Haefner, 1996):

1. **Konzeptionelle oder verbale Modelle** - Sprachliche, begriffliche Beschreibungen von Objekten oder Prozessen
2. **Diagrammartige Modelle** - Graphische Darstellungen von Objekten oder Prozessen

3. **Physikalische Modelle** - Reale, physikalische Attrappen von Objekten; können sowohl vergrößert als auch verkleinert sein (z. B. Modell der DNA-Doppelhelix oder Modell eines Flugzeugs)
4. **Formale Modelle** - Mathematische Beschreibungen von Prozessen (z. B. durch algebraische Gleichungen oder Differentialgleichungen)

In der Systembiologie werden hauptsächlich diagrammartige und formale Modelle betrachtet, mit denen biochemische Systeme dargestellt werden. Beispiele für biochemische Systeme sind metabolische Stoffwechselwege, Signaltransduktionswege sowie transkriptionelle oder genregulatorische Netzwerke. Im Zusammenhang mit biochemischen Systemen versteht man daher unter einem Modell eine abstrakte Beschreibung eines biologischen Prozesses durch die Verwendung mathematischer und bioinformatischer Methoden (Reiß, 2002). Biochemische Systeme haben gemeinsam, dass sie sehr komplex sind und aus vielen Elementen bestehen, die sich gegenseitig beeinflussen. Daher besteht die größte Herausforderung darin, die hohe Komplexität des Systems mit zum Teil großen Mengen experimenteller Daten und teilweise noch lückenhaftem Wissen in einem Modell zusammen zu tragen (Ghosh *et al.*, 2011).

Man kann dabei in eine strukturelle und in eine dynamische Beschreibung unterteilen. Bei der strukturellen Darstellung werden die Komponenten des Systems beispielsweise durch Punkte und deren Interaktionen durch Pfeile gekennzeichnet. Sie dient der Visualisierung aller beteiligten Komponenten mit ihren Interaktionen und gibt einen Überblick über das gesamte Netzwerk. Die dynamische Beschreibung stellt das Netzwerk durch ein System von Differentialgleichungen dar. Die Differentialgleichungen beschreiben die Kinetik der beteiligten Komponenten, mit deren Hilfe sich das dynamische Verhalten simulieren lässt. (Klipp *et al.*, 2009), (Poolman *et al.*, 2004)

4.3 KLASSIFIZIERUNG VON MODELLEN

Zur Erstellung dynamischer Modelle gibt es, je nach Anwendung, verschiedene Modellklassifizierungen:

Systemerklärend - verhaltensbeschreibend: Das systemerklärende Modell beschreibt das biochemische System mit allen seinen Komponenten und ihren Interaktionen, mit dem Ziel, das reale Systemverhalten mit der gleichen Wirkungsstruktur darzustellen. Die zuvor gemachten Beobachtungen zum realen Verhalten des Systems dienen dabei zur Validierung der späteren Simulationsergebnisse. Bei dem verhaltensbeschreibenden

Modell hingegen bilden zuvor gemachte Beobachtungen zum realen Verhalten des Systems die Grundlage des Modells. Das Ziel dabei ist, dieses Verhalten nachzusimulieren, wobei die Struktur des Modells nicht mit der Struktur des realen Systems übereinstimmen muss. Es werden nur das reale Systemverhalten (*Output*) und die dafür verantwortlichen, äußeren Einflüsse (*Input*) berücksichtigt. Der Zusammenhang zwischen *Input* und *Output* wird in diesem Fall durch eine geeignete mathematische Funktion beschrieben, so dass das Innere des Systems in diesem Fall als *black box* behandelt wird. (Bossel, 2004)

Der Unterschied dieser beiden Begriffe wird am Beispiel der Modellierung eines Stoffwechselweges deutlich. Ein Stoffwechselweg kann sowohl systemerklärend, als auch verhaltensbeschreibend modelliert werden. In dem systembeschreibenden Modell sind alle beteiligten Metabolite und die dazugehörigen enzymatischen Reaktionen enthalten, so dass mit Hilfe der Simulationen die Veränderungen der einzelnen Metabolitkonzentrationen genau untersucht werden können. Zuvor gemachte Beobachtungen bezüglich Metabolitkonzentrationsänderungen dienen dabei der Validierung des Modells.

Für das verhaltensbeschreibende Modell müssen aufgrund gemachter Beobachtungen zum realen Systemverhalten ein *Input* und ein *Output* des Systems definiert werden, z. B. kann aufgenommene Glukose den *Input* und ein bestimmtes Stoffwechselprodukt (z. B. Acetat) den *Output* darstellen. Die Beziehung zwischen Glukoseaufnahme und der Produktion des Stoffwechselproduktes wird durch eine mathematische Funktion beschrieben, so dass mit Hilfe des Modells der Einfluss einer veränderten Glukoseaufnahme auf das Stoffwechselprodukt untersucht werden kann. Im Gegensatz zum systemerklärenden Modell erlaubt das verhaltensbeschreibende Modell keine Untersuchung aller am Stoffwechselweg beteiligten Metabolite, sondern ist *Input-Output*-orientiert.

Deterministisch - stochastisch: In einem deterministischen Modell werden zufällige Parameteränderungen ausgeschlossen, so dass das Systemverhalten für alle nachfolgenden Zustände basierend auf den vorherigen vorausgesagt werden kann. Jeder Simulationslauf liefert daher bei unverändertem Parameterset das gleiche Ergebnis. Die Modellierung eines Stoffwechselweges ist z. B. deterministisch, da keine zufälligen Änderungen in den kinetischen Konstanten erlaubt sind. Im Gegensatz dazu ist die Modellierung von DNA-Hybridisierung aufgrund zufälliger Änderungen der Bindewahrscheinlichkeiten stochastisch. In einem stochastischen Modell sind zufällige Ereignisse explizit erwünscht (z. B. zufällige Schwankungen der Umweltbedingungen aufgrund von Wettereinflüssen). Das bedeutet, dass jede Simulation ein unterschiedliches Ergebnis liefert.

Wird eine Vielzahl an Simulationsläufen durchgeführt, ist das Ergebnis eine Häufigkeitsverteilung von Systemzuständen. Ein stochastisches Modell liefert somit einen Überblick, welche Mittelwerte zu erwarten sind und wie groß die Streuung sein wird. (Klipp *et al.*, 2009) (Bossel, 2004)

Kontinuierlich - Diskret: Biochemische Systeme können kontinuierlich oder diskret sein. Kontinuierliche Modelle beschreiben ein System, das zu jedem beliebigen Zeitpunkt (zeitkontinuierlich) oder über den gesamten Raum verteilt (raumkontinuierlich) definiert ist. Das bedeutet, dass zu jedem Zeitpunkt oder zu jeder Position im Raum der Zustand des Systems berechnet werden kann. Die Modellierung eines Stoffwechselweges beispielsweise ist kontinuierlich, da die Metabolitkonzentrationen zu jedem Zeitpunkt berechnet werden können. Kontinuierliche Systeme können daher mittels gewöhnlicher Differentialgleichungen (Differentialquotienten dX/dt) beschrieben werden. Diskrete Modelle verwenden dagegen Differenzenquotienten $((X_t - X_{t-\Delta t})/\Delta t)$, da bei einem diskreten System die Zustandsänderung nur zu diskreten Zeitpunkten oder Positionen im Raum bestimmt werden kann. (Bossel, 2004)

Ein Beispiel für ein diskretes Modell ist die Modellierung der Fortpflanzung von einjährigen Pflanzen, wobei für jede Generation n (mit $n = 1, 2, 3, \dots, n$) die Anzahl der aus keimenden Samen entstehenden Pflanzen berechnet wird (Edelstein-Keshet, 2005).

Biochemische Systeme können bestimmte Eigenschaften besitzen, die bestimmte Modellklassifizierungen verlangen oder ausschließen. Der erste Schritt bei der Erstellung eines Modells ist daher die Überlegung, welche Fragestellung mit dem Modell beantwortet werden soll und welcher Modellansatz das System in geeigneter Weise repräsentiert.

4.4 DIE ERSTELLUNG EINES MATHEMATISCHEN MODELLS

Die einzelnen Schritte zur Erstellung eines mathematischen Modells sind in Abbildung 4.1 als Übersicht dargestellt. Es beginnt mit der Formulierung der Fragestellung, die das Problem enthält, das mit Hilfe des Modells gelöst werden soll. Eine gut ausgearbeitete Fragestellung sollte Antworten auf die folgenden Fragen beinhalten:

1. Welches System soll modelliert werden?
2. Welche Frage(n) soll(en) mit Hilfe des Modells beantwortet werden?
3. Wo liegt die Grenze der Modellierungsaktivität? (Wie gut muss das Modell sein und ab wann lohnt es sich nicht mehr die Modelldefinition weiter zu verfeinern?)

4. Wie wird das Modellergebnis analysiert und angewandt? (Haefner, 1996)

Ist die Fragestellung und damit auch die Zielsetzung möglichst genau formuliert, werden im nächsten Schritt die Hypothesen gebildet, die mit Hilfe des Modells untersucht werden sollen. Die Hypothesenbildung erfolgt dabei in der Regel verbal (z. B. für ein metabolisches Modell: Reduziert sich die Glukoseaufnahme, verringert sich auch die Ausbeute eines bestimmten Stoffwechselproduktes). An dieser Stelle ist der Modellierungsprozess soweit vorangeschritten, dass das zu modellierende System bereits als strukturelles Modell beschrieben werden kann. (Haefner, 1996) Die strukturelle Modellbeschreibung kann mit Hilfe verschiedener *Software-Tools* erfolgen. Ein mögliches *Tool* ist die *CellDesigner-Software*, die auch in dieser Arbeit verwendet wird. Mit ihr lassen sich die Komponenten eines zu modellierenden Systems mit ihren Interaktionen visuell als Netzwerk darstellen (Funahashi *et al.*, 2008).

Im nächsten Schritt erfolgt die Wahl einer adäquaten Disziplin als Grundlage für die Modellierung. Hierbei müssen die qualitativen Hypothesen in spezifische, quantitative Beziehungen umgewandelt werden, so dass sie sich mittels mathematischen Gleichungen beschreiben lassen. Die mathematischen Gleichungen basieren dabei in der Regel auf bekannten naturwissenschaftlichen Gesetzen aus den Disziplinen der Physik oder der Chemie. Zusätzlich beinhalten sie chemische und biologische Informationen, die den Ablauf der zu beschreibenden Reaktionen spezifizieren. Dies können z. B. Angaben darüber sein, ob die Reaktionen reversibel oder irreversibel sind, oder ob es eine hemmende Komponente gibt.

Der nächste Schritt, die Modelldefinition, umfasst die Aufstellung der mathematischen Gleichungen, bei der alle notwendigen und verfügbaren Informationen zu dem biochemischen System in dem Modell zusammengetragen werden.

Anschließend folgt der Prozess der Modellimplementierung mit dem Ziel, das mathematische Modell mit einer geeigneten Software richtig zu lösen. Aufgrund ihrer Komplexität lassen sich Modelle häufig nicht analytisch lösen, so dass man sich eines numerischen Lösungsansatzes bedient. Hierbei werden die Gleichungen in einen Programmcode übersetzt und unter Verwendung eines passenden Algorithmus mit Hilfe eines Computers gelöst (Klipp *et al.*, 2009).

Im Anschluss an die Modellimplementierung muss ein Parameterset gefunden werden, mit dem das Modell ein experimentelles Datenset in geeigneter Weise repräsentiert. Im Zuge dessen müssen die in den Gleichungen vorkommenden konstanten Parameter mit einem numerischen Wert versehen werden. Hierfür eignen sich am besten experimentell

ermittelte Werte, z. B. über gekoppelte Enzym-Assays gemessene kinetische Konstanten. Die Verwendung experimentell bestimmter Parameter ist eine wichtige Voraussetzung, um das biologische System möglichst realistisch beschreiben zu können. Neben den Parametern müssen für den numerischen Lösungsansatz die Startbedingungen festgelegt werden. Dies können z. B. experimentell gemessene Startkonzentrationen sein. (Klipp *et al.*, 2009)

Um bei dem Beispiel eines metabolischen Modells zu bleiben, würde das bedeuten, dass jede beteiligte Stoffwechselreaktion mit einem mathematischen Term beschrieben wird. Häufig wird dabei die sogenannte Michaelis-Menten-Kinetik verwendet, die als konstanter Parameter die Michaelis-Menten-Konstante besitzt. Wenn diese Konstanten im Versuch gemessen wurden, könnte man sie an dieser Stelle in das Modell integrieren. Als Startbedingungen könnten in diesem Beispiel gemessene Metabolitkonzentrationen dienen.

Oft ist man aber in der Modellierung mit dem Fall konfrontiert, dass nur sehr wenige experimentell bestimmte Daten existieren und viele Parameter im Modell unbekannt sind. In diesem Fall müssen die unbekannt Parameter durch Vergleichen der Simulationen mit realen Beobachtungen zum Systemverhalten abgeschätzt werden. Dazu wählt man eine *Output*-Variable zu der bereits quantitative Experimente vorliegen oder bestimmt deren *Output* im Labor. Anschließend werden die Simulationen unter Variation der unbekannt Parameter so lange wiederholt, bis die Parameter so weit angepasst sind, dass die Simulationen dem experimentell bestimmten *Output* entsprechen. Diesen Vorgang bezeichnet man als Parameteranpassung, bei dem das Vergleichen der Simulationen mit den experimentellen Messwerten und das Variieren der Parameter manuell durchgeführt wird. Eine Möglichkeit die Parameteranpassung zu automatisieren, ist die Parameteridentifikation mit Hilfe einer Optimierungsfunktion. Dabei werden viele Simulationsdurchläufe unter Variation verschiedener Parameter durchgeführt. Nach jedem Durchlauf werden die Ergebnisse der Simulation mit den experimentellen Messwerten verglichen und die Abweichung berechnet. Die Optimierungsfunktion dient dazu, die Abweichung zu minimieren, so dass nach einer endlichen Anzahl von Simulationsdurchläufen der optimale Parametersatz in dem Parameterraum identifiziert wird. Bei der Parameteridentifikation ist jedoch zu beachten, dass zunächst eine Optimierungsfunktion definiert werden muss und die Durchführung unter Umständen viel Rechenzeit in Anspruch nehmen kann. (Klipp *et al.*, 2009)

Nach der Implementierung folgt die Validierung des Modells. Dabei muss das Modell-

verhalten mit weiteren experimentell bestimmten Daten verglichen werden, um zu testen, ob das Modell das reale Systemverhalten für jedes beliebige experimentelle Datenset in geeigneter Weise repräsentiert. Häufig ist dies jedoch nicht sofort der Fall, so dass das Modell überarbeitet werden muss. Die Überarbeitung kann dabei an mehreren Schritten im Prozess der Modellentwicklung stattfinden (in Abb. 4.1 durch die runden Pfeile gekennzeichnet), und sie erfolgt so lange, bis das Modell das reale Systemverhalten für jedes beliebige experimentelle Datenset in geeigneter Weise repräsentiert. Ist dies letztendlich der Fall, kann das Modell zur Analyse des biochemischen Systems verwendet werden. (Haefner, 1996)

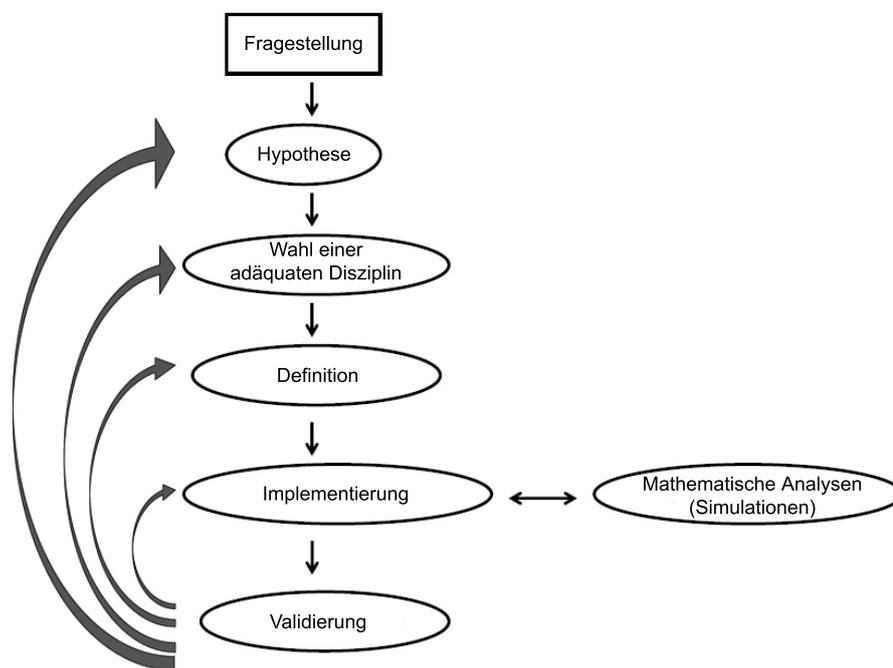


Abbildung 4.1: Erstellung eines mathematischen Modells.

Die Erstellung eines mathematischen Modells beginnt mit der Formulierung des Problems, das mit Hilfe des Modells gelöst werden soll (Fragestellung). Der nächste Schritt beinhaltet die Bildung der Hypothesen, die mit Hilfe des Modells untersucht werden sollen. Basierend auf der Fragestellung und den formulierten Hypothesen erfolgt die Wahl einer adäquaten Disziplin als Grundlage für die Modellierung und die genaue Definition des Modells. Im darauf folgenden Schritt wird das Modell zunächst implementiert und anschließend durch mathematische Analysen untersucht und validiert. Die Erstellung eines mathematischen Modells ist ein iterativer Prozess, was anhand der Pfeile verdeutlicht wird. Bei der Validierung wird das Modellverhalten mit experimentell bestimmten Daten verglichen, um zu testen, ob das Modell das reale Systemverhalten für jedes beliebige experimentelle Datenset in geeigneter Weise repräsentiert. Wenn dies nicht der Fall ist, müssen die einzelnen Schritte der Modellerstellung überprüft und korrigiert werden, bis das Modellverhalten dem realen Verhalten in der geforderten Genauigkeit entspricht. (Mit freundlicher Genehmigung von Petra Lutter)

4.5 MATHEMATISCHER HINTERGRUND ZUR QUANTITATIVEN BESCHREIBUNG EINER MODELLS MIT HILFE DES MASSENWIRKUNGSGESETZES

Vor der Erstellung eines dynamischen Modells wird das zu modellierende biologische System häufig zunächst in Form eines strukturellen Modells visualisiert, um eine Übersicht über die beteiligten Komponenten und ihre Interaktionen zu erhalten. Für die anschließende mathematische Formulierung der quantitativen Beziehungen innerhalb dieses Modells gibt es verschiedene Möglichkeiten.

Im folgenden Abschnitt werden zunächst einige grundsätzliche Punkte vorgestellt, die bei der Erstellung eines dynamischen Modells auf Grundlage des Massenwirkungsgesetzes beachten werden sollten. Zunächst ist bei der quantitativen Beschreibung des Modells zu beachten, dass jedes Element des strukturellen Modells eine Zustandsvariable ist, die eine Differentialgleichung (DGL) benötigt. Dabei repräsentiert die linke Seite der Gleichung die Änderung in Abhängigkeit der Zeit und die rechte Seite beschreibt, wie diese Änderungen erfolgen. Des Weiteren ist zu berücksichtigen, dass jeder Materialfluss in eine Zustandsvariable (Zufluss) und jeder Materialfluss aus ihr heraus (Abfluss) durch mindestens eine explizite algebraische Gleichung festgelegt sein muss. (Klipp *et al.*, 2009)

In dem strukturellen Modell sind die Zu- und Abflüsse durch Pfeile dargestellt. Das bedeutet, jeder Pfeil ist durch einen mathematischen Ausdruck beschrieben und die Summe aller Ausdrücke, die einen Zufluss und Abfluss beschreiben, bilden die rechte Seite der DGL (Haefner, 1996).

Werden alle Zu- und Abflüsse zusammengefasst, kann für eine einzelne Zustandsvariable S folgende allgemeine DGL aufgestellt werden:

$$\frac{dS}{dt} = \sum (\text{Zuflüsse}) - \sum (\text{Abflüsse}). \quad (4.1)$$

Man kann sagen, dass das Problem der quantitativen Beschreibung eines Modells „nur“ darin besteht, die passenden mathematischen Ausdrücke für die Zu- und Abflüsse zu finden. Der nächste Punkt ist, dass viele biologische Systeme gemeinsame Grundprozesse besitzen, die durch ähnliche mathematische Ausdrücke beschrieben werden können. Das erscheint auf den ersten Blick nicht offensichtlich, da biologische Systeme sehr komplex sind und es viele verschiedene Ebenen gibt, auf denen sie betrachtet werden können (z. B. auf biochemischer, zellulärer, physiologischer oder

ökologischer Ebene). Man könnte denken, dass sich kaum Gemeinsamkeiten zwischen den beteiligten Prozessen finden lassen, was aber nur bedingt richtig ist. Es gibt verschiedene mathematische Ausdrücke, die man zur Beschreibung von ähnlichen Prozessen in unterschiedlichen Systemen anwenden kann. (Haefner, 1996)

Ein Ausdruck, der zur mathematischen Beschreibung bei einer Vielzahl von Systemprozessen angewendet wird, ist das Massenwirkungsgesetz, auf das im Folgenden genauer eingegangen werden soll.

Das Ziel dieser Arbeit ist die mathematische Modellierung der *ColE1*-Plasmidreplikationskontrolle. Dabei wird bei der Erstellung des dynamischen Modells die Annahme getroffen, dass alle beteiligten Reaktionen auf dem Massenwirkungsgesetz beruhen. Das Massenwirkungsgesetz beschreibt eine Abhängigkeit der Reaktionsgeschwindigkeit von der Substratkonzentration und der kinetischen Konstante. Theoretisch ist jede Reaktion reversibel, was bedeutet, dass sie in beide Richtungen ablaufen kann. Die Geschwindigkeit der Hinreaktion wird dabei konventionsgemäß als positiv und die der Rückreaktion als negativ angenommen (Klipp *et al.*, 2009).

Für die Reaktionsgeschwindigkeiten einer einfachen reversiblen Ein-Substratreaktion



gilt

$$v_{\text{hin}} = A \cdot k_{\text{hin}} \quad (4.3)$$

$$v_{\text{rück}} = B \cdot k_{\text{rück}} \quad (4.4)$$

$$v = v_{\text{hin}} - v_{\text{rück}} = A \cdot k_{\text{hin}} - B \cdot k_{\text{rück}} \quad (4.5)$$

wobei v die Bildungsgeschwindigkeit für B , v_{hin} , $v_{\text{rück}}$ die Geschwindigkeiten der Hin- und Rückreaktionen und k_{hin} , $k_{\text{rück}}$ die jeweiligen kinetischen Konstanten sind. Der oben beschriebene Sachverhalt lässt sich auch auf Mehrsubstratreaktionen übertragen (Bisswanger, 1994).

Eine allgemeine Form des Massenwirkungsgesetzes für eine Reaktion, bei der m_i Sub-

strate mit den Substratkonzentrationen S_i in m_i Produkte mit den Produktkonzentrationen P_j umgewandelt werden, lautet

$$v = v_{\text{hin}} - v_{\text{rück}} = k_{\text{hin}} \cdot \prod_{i=1}^{m_i} S_i^{n_i} - k_{\text{rück}} \cdot \prod_{j=1}^{m_j} P_j^{n_j} \quad (4.6)$$

wobei n_i und n_j die jeweiligen Molaritäten von S_i und P_j beschreiben. (Klipp *et al.*, 2009)

Im Fall eines Reaktionsgleichgewichtes gilt für die Nettoreaktionsgeschwindigkeit $v = 0$. Für die unter (4.2) beschriebene Reaktion mit den Gleichgewichtskonzentrationen A_{eq} und B_{eq} gilt dann

$$v_{\text{hin}} = v_{\text{rück}} \quad (4.7)$$

$$A_{\text{eq}} \cdot k_{\text{hin}} = B_{\text{eq}} \cdot k_{\text{rück}} \quad (4.8)$$

$$\frac{k_{\text{hin}}}{k_{\text{rück}}} = \frac{B_{\text{eq}}}{A_{\text{eq}}} = K_{\text{eq}} \quad (4.9)$$

mit K_{eq} als Gleichgewichtskonstante. (Klipp *et al.*, 2009)

In einem biologischen System liegen in der Regel keine Gleichgewichtskonzentrationen der Substrate und Produkte vor, sondern es herrscht ein ständiger Stofffluss (z. B. entlang eines Stoffwechselweges) mit sich ändernden Substratkonzentrationen. Um diese zeitlichen Konzentrationsänderungen mathematisch zu beschreiben, bedient man sich gewöhnlichen DGLs. Für die Beispielreaktion (4.2) erhält man folgende DGLs (Bisswanger, 1994)

$$\frac{dA}{dt} = -v_{\text{hin}} \quad (4.10)$$

$$\frac{dB}{dt} = v_{\text{hin}} \quad (4.11)$$

Eine gewöhnliche DGL ist eine Gleichung, in der neben einer unabhängigen Veränderlichen t und einer Funktion dieser Veränderlichen x auch noch die Ableitungen dieser

Funktion nach der unabhängigen Veränderlichen auftreten (Bronstein *et al.*, 2000). Die explizite Form einer gewöhnlichen DGL sieht folgendermaßen aus:

$$\chi^{(n)} = f(t; \chi; \chi'; \dots; \chi^{(n-1)}). \quad (4.12)$$

Um das zeitliche Verhalten eines biologischen Systems zu beschreiben, verwendet man ein System von gewöhnlichen DGLs. Die allgemeine Form eines solchen DGL-Systems für ein dynamisches Modell schreibt sich wie folgt:

Sei G eine Teilmenge von $\mathbb{R} \times \mathbb{R}^n$ und

$$f: G \rightarrow \mathbb{R}^n \quad (4.13)$$

und $(\vec{t}, \vec{s}) \rightarrow f(\vec{t}, \vec{s})$ eine kontinuierliche Funktion. Dann ist

$$\frac{d\vec{s}}{dt} = f(\vec{s}, \vec{k}, \vec{t}) \quad (4.14)$$

mit $\vec{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$, $\vec{k} = \begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix}$ und $\vec{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$ ein System mit n Differentialgleichungen

erster Ordnung. Dabei repräsentieren s_i die Variablen, k_j die Parameter und t den Zeitvektor. In biologischen Systemen beschreiben die Funktionen f_i meistens aufbauende und abbauende Reaktionen. (Klipp *et al.*, 2009)

In der Regel ist man bei der Modellierung biologischer Systeme an den zeitlichen Veränderungen von s_i interessiert. Verwendet man dabei das Massenwirkungsgesetz zur Beschreibung der Reaktionsgeschwindigkeiten, besteht $f(\vec{s}, \vec{k}, \vec{t})$ in Gleichung (4.14) aus Termen, basierend auf dem Massenwirkungsgesetz. Um zeitliche Veränderungen von s_i untersuchen zu können, muss das DGL-System aus Gleichung (4.14) gelöst werden. Dies erfolgt nach Festlegung der Anfangsbedingungen unter Anwendung eines numerischen Lösungsansatzes mit Hilfe eines Computers.

ZIEL DER ARBEIT

Das Ziel dieser Arbeit ist, die regulatorischen Mechanismen der *ColE1*-Plasmidreplikationskontrolle anhand eines mathematischen Modells sowie durch Metabolomanalysen zu untersuchen.

Mit Hilfe des mathematischen Modells soll der Fragestellung nachgegangen werden, welchen Effekt es auf die Plasmidproduktion hat, würde man tRNA-Moleküle derart modifizieren, dass eine Beladung mit einer Aminosäure durch die Aminoacyl-tRNA-Synthetase nicht mehr möglich ist. Die Hypothese zur Beantwortung der Fragestellung ist, dass sich die Plasmidproduktion erhöht, da die Zellen keinen Stressbedingungen ausgesetzt sind, aber dennoch eine erhöhte Menge unbeladener tRNA-Moleküle in Form von modifizierten tRNAs aufweisen, die die Plasmidreplikation positiv beeinflussen. Zur Untermauerung des mathematischen Modells durch experimentell bestimmten Daten sollen im Labor die Konzentrationen der regulatorischen RNA-Moleküle, RNAI und RNAII sowie die intrazelluläre Plasmidkonzentration bestimmt werden.

Die zusätzlichen Analysen der Metabolomebene beinhalten die Untersuchungen der Metabolitzusammensetzung in Zellen mit unterschiedlichem Plasmidgehalt und die Durchführung von ^{13}C -Flussmessungen. Anhand der Untersuchungen zur Metabolitzusammensetzung soll analysiert werden, welchen Einfluss die Anwesenheit eines Plasmids auf das Metabolitprofil des Wirtsstamms hat und wo besonders große Unterschiede in den Metabolitpools zu finden sind. Ergänzend dazu soll unter Anwendung von ^{13}C -Flussmessungen untersucht werden, ob der Plasmidgehalt in der Zelle gleichzeitig auch die metabolische Flussverteilung beeinflusst.

Teil IV

MATERIAL UND METHODEN

MATERIAL

6.1 VERWENDETE BAKTERIENSTÄMME

Tabelle 6.1: In dieser Arbeit verwendete Bakterienstämme

Bakterienstamm	Relevante Eigenschaften	Referenz/ Herkunft
<i>E. coli</i> DH5 α	<i>supE44</i> Δ , <i>lacU169</i> (Φ 80 <i>lacZ</i> Δ M15), <i>hsdR17</i> , <i>recA</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>relA1</i>	(Hanahan, 1983)
<i>E. coli</i> W3110	λ ⁻ F ⁻ <i>mcrA mcrB</i> IN(<i>rrnD-rrnE</i>) ₁	(Bachmann, 1972)
<i>E. coli</i> VH33	Derivat des <i>E. coli</i> W3110, Δ <i>ptsH</i> , Δ <i>ptsI</i> , Δ <i>acI</i> , <i>lacZ::loxP</i>	(De Anda <i>et al.</i> , 2006)
<i>E. coli</i> VH33 Δ <i>recA</i>	Derivat des <i>E. coli</i> W3110, Δ <i>ptsH</i> , Δ <i>ptsI</i> , Δ <i>acI</i> , <i>lacZ::loxP</i> , Δ <i>recA</i>	(Borja <i>et al.</i> , 2012)
<i>E. coli</i> VH33 Δ <i>recA deoR</i>	Derivat des <i>E. coli</i> W3110, Δ <i>ptsH</i> , Δ <i>ptsI</i> , Δ <i>acI</i> , <i>lacZ::loxP</i> , Δ <i>recA</i> , Δ <i>deoR</i>	(Borja <i>et al.</i> , 2012)

6.2 PRIMER, TRANSPOSON UND PLASMIDVEKTOREN

6.2.1 *Primer*

Tabelle 6.2: In dieser Arbeit verwendete Primer

Bezeichnung	Sequenz (5'-3')	Herkunft
RNAIT71	GAAATTAATACGACTCACTATAGG GACAGTATTTGGTATCTGCGCTC	Metabion
RNAIT72	AACCACCGCTACCAGCGG	Metabion
RNAIIT71	GAAATTAATACGACTCAC TATAGGGTTTCCATAGGCTCCGCCC	Metabion
RNAIIT72	GCAAACAAAAAAACCACCGCTAC	Metabion
mRNAIIIc1	TAACGGCTTCAGCAGAGCGCAGAT	Metabion
mRNAII_lc2neu	TCGGGTAGGTCGTTTCGCTCCAA	Metabion
mRNAIIT71	GAAATTAATACGACTCACTATAGGGGC AAACAAAAAAACCACCGCTACCA	Metabion
mRNAIIT72	TTCCATAGGCTCCGCCCCC	Metabion
pk18mob_Cre_for	GATATCGCACCCCAGGCTTTACTACTT	Metabion
pk18mob_Cre_rev	GATATCGGCTAATCGCCATCTTCCAG	Metabion
Cre-Sma-F	CCCGGGATGTCCAATTTACTGACCG TACACC	Metabion
Cre-Sma-R	CCCGGGCTAATCGCCATCTTCCAGC AGGCGC	Metabion
LTTSac	GAGCTCTCAATTCAGGGTGGTGAATG	Metabion
LTTBam	GGATCCAATTCATGAGCGGATACATA	Metabion
HB18-Tn5-ME	CTGTCTCTTATACACATCT	Metabion
pKD20 Primer for	TTTATTATTGGATCCGATATCATGGG TATGGACAGTTTTTC	Metabion
pKD20 Primer rev	TTATATTATTGAGCTCGCAGGTGGCA CTTTTCGGGG	Metabion

6.2.2 Transposon

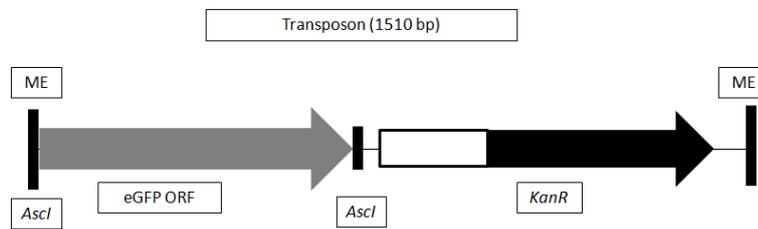


Abbildung 6.1: Transposon nach (Bednarz, 2012)

6.2.3 Plasmidvektoren

Tabelle 6.3: In dieser Arbeit verwendete Plasmide

Plasmid	Beschreibung	Referenz
pSUP 201-3	Low copy Plasmid (7896 bp), pMB1-Ori <i>E. coli</i> , Amp ^R , Cm ^R	(Simon <i>et al.</i> , 1983)
pCMV-GFP	High copy Plasmid (3487 bp), pUC-Ori <i>E. coli</i> , Kan ^R	Plasmid Factory
pDG	High copy Plasmid (21846 bp), pUC-Ori <i>E. coli</i> , Amp ^R	Plasmid Factory
pCMV-lacZ	High copy Plasmid (7164 bp), pUC-Ori <i>E. coli</i> , Amp ^R	Plasmid Factory
pHN	Phagemid (5995 bp), Nm ^R pUC-Ori <i>E. coli</i> , f1-Ori f1-Phage	(Herrera <i>et al.</i> , 2010)
pK18mob	Kan ^R , Kloniervektor, OriV <i>E. coli</i>	(Kirchner & Tauch, 2003)
pHB14	OriV <i>E. coli</i> , Kan ^R , Gm ^R Ptrc Promoter, lacIq Repressor, hyperaktive Transposase Gen egfp-nptII-Transposon	(Bednarz, 2012)

Tabelle 6.4: In dieser Arbeit verwendete Plasmide (Fortsetzung)

Plasmid	Beschreibung	Referenz
pHB15	Gm ^R , <i>cre</i> Gen	(Bednarz, 2012)
pKD20	Expressionsvektor, Amp ^R , Temperatursensitiver <i>OriR101 E. coli repA101ts</i> <i>GenBank accession no.</i> J02459	(Datsenko & Wanner, 2000)
pIF11	Beinhaltet temperatursensitiven <i>OriR101 E. coli repA101ts</i> und Amp ^R aus pKD20 mit zusätzlichen <i>Bam</i> HI/ <i>Sac</i> I/ <i>Eco</i> RV-Schnittstellen	Diese Arbeit
pIF12	<i>OriR101 E. coli repA101ts</i> , Amp ^R und <i>egfp-nptII</i> -Transposon über <i>Eco</i> RV-Schnittstelle in pIF11 kloniert	Diese Arbeit
pIF13	<i>OriR101 E. coli repA101ts</i> , Amp ^R , hyperaktive Transposase Gen und <i>egfp-nptII</i> -Transposon über <i>Eco</i> RV-Schnittstelle in pIF12 kloniert	Diese Arbeit
pIF14	<i>OriR101 E. coli repA101ts</i> , Amp ^R , hyperaktive Transposase Gen, <i>egfp-nptII</i> -Transposon und <i>Leucin-tRNA-RNAI</i> -Gen über <i>Asc</i> I-Schnittstelle in pIF13 kloniert	Diese Arbeit & Masterarbeit S. Wiebe
pIF15	<i>OriR101 E. coli repA101ts</i> , Amp ^R , hyperaktive Transposase Gen, <i>egfp-nptII</i> -Transposon und <i>Leucin-tRNA-WT</i> -Gen über <i>Asc</i> I-Schnittstelle in pIF13 kloniert	Diese Arbeit & Masterarbeit S. Wiebe
pIF16	<i>Cre</i> -Gen über <i>Sam</i> I-Schnittstelle in pK18mob kloniert	Diese Arbeit
pIF17	<i>OriR101 E. coli repA101ts</i> , Amp ^R und <i>Cre</i> -Gen über <i>Eco</i> RV-Schnittstelle in pIF11 kloniert	Diese Arbeit & Masterarbeit S. Wiebe

6.3 ENZYME, CHEMIKALIEN, KITS UND VERBRAUCHSMATERIAL

6.3.1 Enzyme

Für die Verwendung der Enzyme wurden jeweils die von den Herstellern mitgelieferten Reaktionspuffer eingesetzt.

Tabelle 6.5: In dieser Arbeit verwendete Enzyme

Restriktionsenzyme	Erkennungssequenz (5'-3')	Bezugsquelle
<i>EcoRI</i>	G-AATTC	Fermentas
<i>EcoRV</i>	GAT-ATC	Fermentas
<i>BglIII</i>	A-GATCT	Fermentas
<i>NcoI</i>	C-CATGG	Fermentas
<i>SacI</i>	GAGCT-C	Fermentas
<i>BamHI</i>	G-GATCC	Fermentas
<i>NotI</i>	GC-GGCCGC	Fermentas
<i>PstI</i>	CTGCA-G	Fermentas
<i>SmaI</i>	CCC-GGG	Fermentas
<i>AscI</i>	GG-CGCGCC	New England Biolabs
DNA-Polymerasen	Beschreibung	Bezugsquelle
GoTaq-Polymerase	A-Überhänge	Promega
Phusion-Polymerase	<i>proof-reading</i>	Fermentas
Sonstige Enzyme	Beschreibung	Bezugsquelle
T4-Ligase	Ligation von DNA-Fragmenten	Fermentas
T7-RNA-Polymerase	Synthese von RNA mit DNA als <i>Template</i>	Roche
DNAaseI	Verdau von DNA	Fermentas

6.3.2 Chemikalien

Tabelle 6.6: In dieser Arbeit verwendete Chemikalien

Chemikalie	Bezugsquelle
Aceton	VWR
Agarose	PeqLab
Agar Select	Invitrogen
Antibiotika	Serva und Sigma Aldrich
Bromphenolblau	Serva
CaCl ₂	Merck
C ₁₂ , C ₁₅ , C ₁₈ , C ₁₉ , C ₂₂ , C ₂₈ , C ₃₂ , C ₃₆	Sigma Aldrich
Alkane	
Dimethylformamid (DMF)	Merck
dNTP-Mix	Fermentas
EDTA	Roth
Ethanol	Roth
Ethidiumbromid	Serva
FeSO ₄	Merck
Glucose	VWR
Glycerin	Roth
Harnstoff	Merck
HCl	Sigma Aldrich
LB broth	Invitrogen
MBDSTFA	Macherey-Nagel
Methanol	VWR
Methoxylamin Hydrochlorid (MeOX)	Sigma Aldrich
MgSO ₄	Merck
MilliQ H ₂ O	Merck-Millipore
MSTFA	Macherey-Nagel

Tabelle 6.7: In dieser Arbeit verwendete Chemikalien (Fortsetzung)

Chemikalie	Bezugsquelle
NaCl	Merck
NaOH	VWR
NTP-Mix	Roche
Pyridin	Sigma
Ribitol	Sigma
RNase-freies Wasser	Roth
Thiamin-HCl	Sigma
Tris	Roth
¹³ C-Glukose	Eurisotop

6.3.3 Kits

Tabelle 6.8: In dieser Arbeit verwendete Kits

Kits	Verwendung	Bezugsquelle
GeneJET Plasmid Miniprep Kit	Aufreinigung von Plasmid DNA	Thermo Scientific
PCR clean-up and Gel extraction Kit	Aufreinigung von DNA aus Agarosegelen	Macherey- Nagel
SensiMix Probe One-Step Kit	Real-Time PCR Kit	Bioline
RNeasy Plus Mini Kit	Aufreinigung von RNA	Qiagen

6.4 LABORGERÄTE

Tabelle 6.9: In dieser Arbeit verwendete Laborgeräte

Produkt	Hersteller
Autoklav /VX-120)	Systemec
Brutschränke	Memmert
Eismaschine (AF-80)	Scotsman
Elektrophoresekammern für Agarosegele	Eigenbau
HPLC Spectra Physics	Thermo Fisher/ Finnigan Spectra Systems
Ribolyser (Precellys 24)	Bertin
Klimakammer/Kühlraum	Viessmann
Kühlschränke	AEG
Kühlzentrifuge	Sigma
Lyophylle	Christ
PCR-Gerät (PTC-100)	MJ Research Lab Recyclers
Photometer	Eppendorf Biophotometer
Pipettierhilfen	Gilson
pH-Meter (calimatic pH Meter 766)	Knick
Magnetrührer	IKAMAG® RCT
NanoDrop(2000c)	Peqlab
Reacti-Therm III Heating/Stirring Modul	Pierce
Reinraumwerkbank (Hera Safe)	Thermo Electron Corporation
Roller zum Kultivieren	Memmert
RT-PCR Gerät (Opticon2™)	Bio-Rad
Spannungsgeber	Pharmacia
SpeedVac (Concentrator 5301)	Eppendorf

Tabelle 6.10: In dieser Arbeit verwendete Laborgeräte (Fortsetzung)

Produkt	Hersteller
Thermoschüttler (Innova 44)	New Brunswick Scientific
Thermoschüttler (für Eppendorf-Gefäße)	Biozym
Tischzentrifuge	Eppendorf
TRACE GC Ultra ITQ 900 GC-Ion Trap MS	Thermo Scientific™
Trace GC Polaris Q Massenspektrometer	Thermo Finnigan
Transilluminator (Alphamanager HP)	Biozym
Vortex (Vortex Genie 2)	Bender & Hobein AG
Waagen	Sartorius

6.4.1 Verbrauchsmaterial

Tabelle 6.11: In dieser Arbeit verwendete Verbrauchsmaterialien

Produkt	Bezugsquelle
Reaktionsgefäße, andere Plastikwaren	Sarstedt, Greiner und Biozym
Pipettenspitzen	Starlab
Pipettenspitzen mit Filter	Biozym
Glaswaren	Schott
Sterilfilter	Roth
Spritzen	Brau
Glasgefäß mit Deckel (HPLC und GC-MS)	VWR und Wicom
Ultrafree-CL Filter	Millipore
Lysing Matrix B (Fast Protein Tube)	MP Biomedicals
RNaseZAP	Sigma-Aldrich
Glasbeads (Durchmesser 0.1 mm)	Roth

6.5 SOFTWARE

Tabelle 6.12: In dieser Arbeit verwendete Software

Programm	Verwendung	Quelle
CellDesigner 4.2	Visualisierung der Modellkomponenten	(Funahashi <i>et al.</i> , 2008)
MATLAB R2009b	Modellprogrammierung und Simulation	Mathworks©
Xcalibur	Auswertung der GC-MS- und HPLC-Daten	Thermo Scientific™
MeltDB 2.0	Auswertung der GC-MS- und HPLC-Daten	(Kessler <i>et al.</i> , 2013)
OpenFlux	Berechnung des intrazellulären Glukoseflusses	(Quek <i>et al.</i> , 2009)
Clonemanager 5.0	Planung und Kontrolle von Klonierungen und Restriktionen	Sci Ed Central
Opticon Monitor 3	Auswertung der qRT-PCR	BioRad
RNASHapes	tRNA-Strukturvorhersage	(Huang & W, 1991)
LALIGN	Erstellung von Alignments	(Steffen <i>et al.</i> , 2006)

6.6 NÄHRMEDIEN UND MEDIENZUSÄTZE

6.6.1 Definierte Medien zur Kultivierung von Bakterien

M9-Medium supplementiert mit FeSO₄(Sambrook *et al.*, 1989)

8.5 g Na₂HPO₄
 3 g KH₂PO₄
 0.5 g NaCl
 1 g NH₄Cl
 3 mg FeSO₄ · 7H₂O
 1 ml 1 M MgSO₄
 1 ml 0.1 M CaCl₂
 25 ml 20% Glucose-Lösung
 → auf 973 ml mit Millipore H₂O auffüllen,
 autoklavieren

1 M MgSO₄ und 0.1 M CaCl₂ werden als Stammlösungen getrennt autoklaviert und mit der sterilfiltrierten 20% Glucose-Lösung zugegeben.

PD1-Medium(Lara *et al.*, 2008)

17 g K₂HPO₄
 5.3 g KH₂PO₄
 2.5 g (NH₄)₂SO₄
 1 g (NH₄)Cl
 0.01 g Thiaminhydrochlorid
 1 g MgSO₄ · 7H₂O
 25 ml 20% Glucose-Lösung
 1 ml Spurenelementlösung
 → auf 974 ml mit Millipore H₂O auffüllen
 autoklavieren

1 M MgSO_4 und die Spurenelementlösung werden getrennt autoklaviert und mit der sterilfiltrierten 20% Glucose-Lösung zugegeben.

Spurenelementlösung:

17 g $\text{Zn}(\text{CH}_3\text{COO})_2 \cdot 2\text{H}_2\text{O}$
 7 g $\text{Na} \cdot \text{EDTA}$
 1.25 g $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$
 7.5 g $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$
 0.75 g $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$
 1.5 g H_3BO_3
 1.05 g $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$
 50 g Fe(III)Citrat
 → auf 1 l mit Millipore H_2O auffüllen

6.7 MEDIENZUSÄTZE

6.7.1 Antibiotika

Tabelle 6.13: In dieser Arbeit verwendete Antibiotika

Antibiotikum	Lösungsmittel	Konzentration für <i>E. coli</i>
Kanamycin (Kan)	H_2O	50 $\mu\text{g}/\text{ml}$
Ampicillin (Amp)	H_2O	100 $\mu\text{g}/\text{ml}$

6.7.2 Agar

Zur Herstellung von Festmedien in Form von Agarplatten werden 15 g/l Agar zum flüssigen Medium zugegeben und autoklaviert. Nach Abkühlung des Mediums auf mindestens 60°C werden, wenn benötigt, Antibiotika hinzugegeben.

6.8 PUFFER UND LÖSUNGEN

6.8.1 *Lösungen zur Herstellung von Dauerkulturen***PS-Puffer**7 g Na_2HPO_4 3 g KH_2PO_4 5 g NaCl → auf 1 l mit Millipore H_2O auffüllen

→ pH 7 einstellen und autoklavieren

62.5 % Glycerin/ PS-Puffer

PS-Puffer

62.5 % Glycerin → autoklavieren

6.8.2 *Puffer zur Herstellung chemisch kompetenter Zellen***P1-Puffer**0.1 M CaCl_2

2 mM Tris – HCl

→ pH 7.4 einstellen und autoklavieren

P2-Puffer0.1 M CaCl_2

2 mM Tris – HCl

10 % Glycerin

→ pH 7.4 einstellen und autoklavieren

6.8.3 *Puffer für die DNA-Agarosegelelektrophorese***TAE-Puffer (50x)**

2 M Tris

500 mM Natriumacetat

50 mM EDTA

→ mit Essigsäure auf pH 7.8 einstellen

DNA-Probenpuffer (6x)0.1 M CaCl₂

10 ml TAE-Puffer

10 ml Glycerin

0,2 % Bromphenolblau

Agarose-Lösung

0.8-1.5 % (w/v) Agarose in TAE (1x) aufkochen

Ethidiumbromid-Lösung10 mg/ml in Millipore H₂O6.8.4 *Puffer für die HPLC-Analysen***Puffer A**

(Mackerey-Nagel, 2013)

10 mM K₂HPO₄10 mM KH₂PO₄

5 M Harnstoff

→ mit Millipore H₂O auffüllen, pH 6.9**Puffer B**

(Mackerey-Nagel, 2013)

Puffer A

1.5 M KCl

→ mit Millipore H₂O auffüllen, pH 6.9

METHODEN

7.1 KULTIVIERUNG VON BAKTERIEN

7.1.1 *Anzucht von Bakterien*

Die Anzucht von *E. coli* erfolgte sowohl auf festem, als auch in flüssigem Medium, ggf. unter Zusatz von Antibiotika. Die Anzucht fand jeweils bei 37°C statt. Im Fall eines *E. coli* Stamms mit einem temperatursensitiven Plasmid ist die Kultivierung der Bakterien bei 30°C durchgeführt worden. Flüssigkulturen wurden durch Beimpfen von 5 ml LB-Medium mit einer Einzelkolonie und anschließender Inkubation über Nacht im Roller angezogen. Größere Volumina wurden im Erlenmeyer-Kolben inkubiert.

7.1.2 *Herstellung von Dauerkulturen*

- Bakterien von einer frischen Agarplatte in ein Eppendorf-Gefäß überführen
- Zellen in 100 µl PS-Puffer resuspendieren
- 500 µl 62.5 % Glycerin/PS-Puffer zugeben und invertieren
- Lagerung bei –20°C oder –80°C

7.1.3 *Bestimmung des Bakterientiters*

Das Wachstum einer Bakterienkultur wurde durch photometrische Messungen bei einer Wellenlänge von 600 nm (OD₆₀₀) überprüft. Als Referenz diente dabei steriles Nährmedium, in dem die Bakterien kultiviert wurden. Um den exakten Lebendzelltiter zu bestimmen, wurden Verdünnungsreihen angelegt, auf Festmedium ausplattiert und nach der Inkubation ausgezählt.

7.2 ISOLIERUNG UND DARSTELLUNG VON DNA

7.2.1 *Isolierung von Plasmid-DNA mittels „Thermo Scientific - GeneJET Plasmid Miniprep Kit“*

- 1-5 ml *E. coli* LB-Kultur abzentrifugieren und den Überstand verwerfen
- Zellpellet in 250 µl Resuspensionslösung in einem Eppendorf-Gefäß resuspendieren (alternativ können auch auf einer Agarplatte gewachsene *E. coli*-Zellen abgekratzt und resuspendiert werden)
- 250 µl Lyselösung dazugeben und sofort vorsichtig 4-6 mal invertieren (Scherkräfte vermeiden!)
- 350 µl Neutralisationslösung dazugeben und sofort vorsichtig 4-6 mal invertieren
- Zelltrümmer 5 min bei 14.000 rpm in einer Eppendorf-Tischzentrifuge abzentrifugieren
- Überstand auf eine GeneJET-Säule überführen
- 1 min bei 14.000 rpm abzentrifugieren (Eppendorf-Tischzentrifuge) und Durchlauf verwerfen
- 500 µl Waschlösung auf die Säule geben und 1 min abzentrifugieren (Eppendorf-Tischzentrifuge 14.000 rpm), Durchlauf verwerfen
- 500 µl Waschlösung auf die Säule geben und 1 min abzentrifugieren (Eppendorf-Tischzentrifuge 14.000 rpm), Durchlauf verwerfen
- 1 min bei 14.000 rpm in einer Eppendorf-Tischzentrifuge abzentrifugieren
- GeneJET-Säule auf ein neues Eppendorf-Gefäß setzen und 50 µl Elutionspuffer auf die Säule geben
- 2 min bei Raumtemperatur stehen lassen und anschließend 2 min zentrifugieren (Eppendorf-Tischzentrifuge 14.000 rpm)
- Säule verwerfen und eluierte Plasmid-DNA bei -20°C lagern

7.2.2 *Konzentrationsbestimmung einer DNA-Lösung*

Die Konzentrationsbestimmung einer DNA-Lösung erfolgte durch eine spektrophotometrische Messung bei 260 nm je eines Mikroliters der DNA-Lösung im NanoDrop-Gerät (Peqlab).

7.2.3 Agarosegelelektrophorese

Die Agarosegelelektrophorese dient der Auftrennung und Charakterisierung von DNA-Molekülen. Dabei erlauben die negativ geladenen Phosphatreste den DNA-Molekülen in dem elektrischen Feld zu wandern. Die Wanderungsgeschwindigkeit ist abhängig von der Größe und der Konformation (einzelnsträngig, doppelsträngig, linear, offenkettige Form, *supercoil*-Form) der DNA-Moleküle. Je höher die Agarosekonzentration, desto kürzer ist die Wanderungstrecke der DNA-Moleküle unter gleichbleibenden Bedingungen.

Tabelle 7.1: Agarosekonzentrationen mit den zugehörigen Trennbereichen der Fragmente. Zur Orientierung ist die Länge der Fragmente aufgeführt, die mit dem Farbstoff Bromphenol zusammen wandern (aus (Mühlhardt, 2009)).

Agarosekonzentration (w/v)	Fragmentlänge [kb]	Bromphenol
0.5 %	1 bis 30	1000 bp
0.7 %	0.8 bis 12	700 bp
1 %	0.5 bis 7	300 bp
1.2 %	0.4 bis 6	200 bp
1.5 %	0.2 bis 3	120 bp
2 %	0.1 bis 2	< 100 bp

- Agarose in TAE-Puffer in der Mikrowelle zum Kochen bringen
- auf ca. 60°C abkühlen lassen, in den Gelträger gießen und Kamm einsetzen
- Gel fest werden lassen, die Elektrophoresekammer mit TAE-Puffer auffüllen und den Kamm entfernen
- DNA-Proben mit DNA-Probenpuffer (6x) versetzen (mit ca. 1/5 Volumen des Probenvolumens)
- Geltaschen mit Proben beladen
- Auftrennung bei einer Spannung von 90 - 120 V
- anschließend das Gel in einer Ethidiumbromidlösung färben und bei Bedarf mit Wasser entfärben
- Gel unter UV-Licht ($\lambda = 302 \text{ nm}$) photographieren

7.2.4 Größenbestimmung von DNA-Fragmenten

Die Laufstrecke von DNA-Molekülen ist über einen weiten Bereich umgekehrt proportional zum natürlichen Logarithmus der Fragmentlänge. Durch Vergleich der Laufstrecke eines DNA-Fragments unbekannter Größe mit einem DNA-Längenmarker, der ebenfalls auf das Gel aufgetragen wird, lässt sich eine Aussage über die gesuchte Fragmentgröße treffen.

In dieser Arbeit wurden der 250-12.000 bp DNA-Marker und der 100 bp-Leiter DNA-Marker (hergestellt in dem Molekularbiologischen & Biochemischen Labor (MBBL) Dr. Bartling GmbH aus Bielefeld) verwendet.

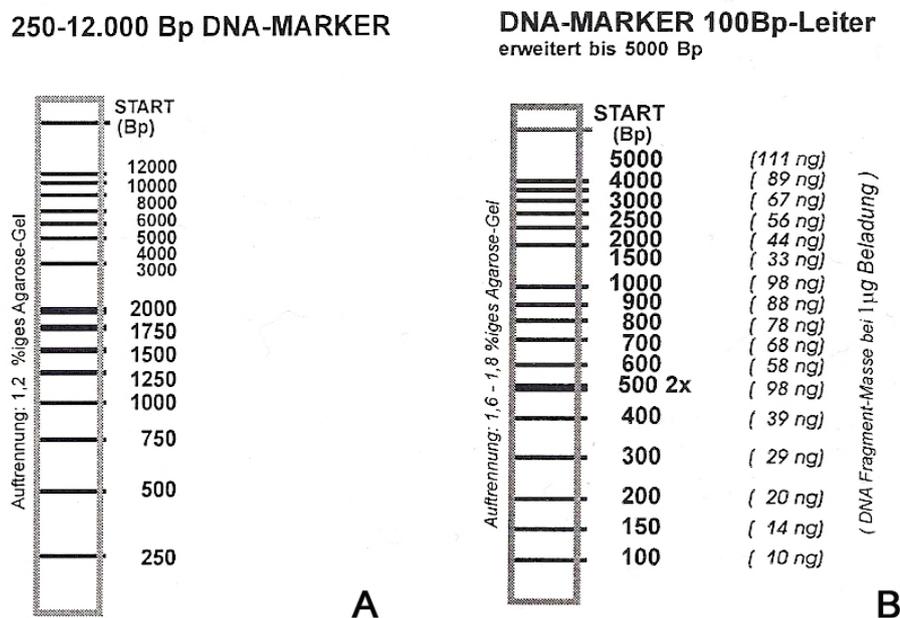


Abbildung 7.1: In dieser Arbeit verwendete Längenmarker:

7.1a. 250-12.000 bp DNA-Marker (MBBL), 7.1b. 100 bp-Leiter (MBBL)

7.2.5 Isolierung von DNA-Fragmenten aus Agarosegelen mittels „MN - PCR clean-up and Gel extraction Kit“

Restriktionsfragmente für die Klonierungen wurden nach der Auftrennung unter Verwendung des „MN - PCR clean-up and Gel extraction Kits“ aus Agarosegelen isoliert.

- das Gel nur unter langwelligem UV-Licht darstellen
- DNA-Fragment im Gel kennzeichnen und mit Skalpell heraustrennen
- ausgeschnittenes DNA-Fragment in einem 1,5 ml Eppendorf-Gefäß wiegen

- auf je 100 mg Agarosegel 200 µl NTI-Puffer zugeben
- 5-10 min bei 50°C inkubieren und den Ansatz alle 2 bis 3 min vortexen (alternativ kann auch auf dem Eppendorf-Thermomixer inkubiert werden)
- nach vollständiger Suspension des Agarosegels den gesamten Inhalt auf eine NucleoSpin-Säule® geben und 30 s bei 13.000 rpm zentrifugieren
- Durchfluss verwerfen und zum Waschen 700 µl NT3-Puffer auf die NucleoSpin-Säule® geben
- 30 s bei 13.000 rpm zentrifugieren und den Durchfluss verwerfen
- erneut Waschen mit 700 µl NT3-Puffer
- 30 s bei 13.000 rpm zentrifugieren und den Durchfluss verwerfen
- zum Trocknen der Silica-Membran 1 min bei 13.000 rpm zentrifugieren
- den oberen Teil der NucleoSpin-Säule® auf ein neues Eppendorf-Gefäß setzen
- 15 - 30 µl NE-Puffer zum Eluieren des DNA-Fragments auf die Membran geben und 1 min bei Raumtemperatur inkubieren
- 1 min bei 13.000 rpm zentrifugieren
- oberen Teil der NucleoSpin-Säule® entsorgen

7.3 RESTRIKTION UND MODIFIKATION VON DNA

7.3.1 DNA- Restriktion mit Typ II Restriktionsendonukleasen

Im Rahmen dieser Arbeit wurden für enzymatische DNA-Spaltungen Restriktionsendonukleasen von Typ II verwendet. Diese erkennen und schneiden DNA spezifisch innerhalb bestimmter palindromischer Sequenzen. Dabei entstehen entweder versetzte (*sticky-ends*) oder glatte Enden (*blunt-ends*). Die für jedes Enzym optimalen Reaktionsbedingungen wurden durch die vom Enzymhersteller empfohlenen Reaktionsbedingungen (zumeist 37°C, enzimspezifische Ionenkonzentrationen) erreicht.

Exemplarischer Spaltungsansatz mit einem Endvolumen von 20 µl:

- Vorlegen von sterilem Millipore Wasser in ein Eppendorf-Gefäß (Wasservolumen = Endvolumen - Volumen der Spaltungskomponenten)
- Gewünschte Menge DNA-Lösung dazupipettieren
- 1/10 des Gesamtvolumens TA-Restriktionspuffer (10x)
- kalkuliertes Restriktionsendonukleasen-Volumen (1U/µg zu spaltender DNA)
- Spaltungsansatz mischen und beim Temperaturoptimum des eingesetzten Restriktionsenzym je nach Bedarf 1 h bis ü/N inkubieren

- Inaktivierung des Enzyms durch Inkubation für 15 min bei -80°C oder für 15 min bei 65°C

7.3.2 Dephosphorylierung von 5'-DNA-Enden

Die Durchführung einer Dephosphorylierung von 5'-DNA-Enden mit einer alkalischen Phosphatase verhindert die Religierung eines linearisierten Vektors ohne Insertion des gewünschten DNA-Fragments. Das Enzym katalysiert die Entfernung von 5'-Phosphatgruppen am Vektor, die zum Ringschluss des mit nur einem Restriktionsenzym gespaltenen Vektors notwendig sind.

Exemplarischer Dephosphorylierungsansatz mit einem Endvolumen von 20 μl :

- 17 μl gesplattene DNA
- 2 μl *Shrimp Alkaline* (SA) Phosphatase-Puffer (10x)
- 1 μl SA-Phosphatase (1 U/ μl)
- Reaktionsansatz mischen und 10 min bis 1 h bei 37°C inkubieren
- Inaktivierung des Enzyms durch Inkubation für 10 min bei 65°C

7.3.3 Polymerasekettenreaktion (PCR)

Mit Hilfe der Polymerasekettenreaktion ist es möglich, geringste DNA-Mengen *in vitro* innerhalb kürzester Zeit zu amplifizieren, wenn jeweils kurze Bereiche der flankierenden Sequenzen des DNA-Fragments bekannt sind. Bei dieser Methode werden die vorhandenen DNA-Moleküle in jedem Schritt durch thermostabile DNA-Polymerasen verdoppelt. Als Negativkontrolle wurde in dieser Arbeit anstelle des DNA-Templates 1 μl steriles Millipore Wasser pipettiert.

Exemplarischer PCR-Ansatz mit einem Endvolumen von 25 μl :

- 5 μl 5x-Puffer
- 0.5 μl dNTP-Mix (jeweils 200 μM)
- 1 μl je Primer (jeweils 10 μM)
- 1 μl DNA
- 0.5 μl DNA-Polymerase (0.02 U/ μl)
- 8 μl steriles Millipore Wasser

Im Rahmen dieser Arbeit wurden zwei verschiedene DNA-Polymerasen verwendet: Die GoTaq- und die Phusion-Polymerase. Beim Einsatz der GoTaq-Polymerase entstehen

an den Enden des amplifizierten Fragments A-Überhänge. Diese sind wichtig, möchte man die Enden des Fragments für die weitere Verwendung mit einem Restriktionsenzym spalten, da viele Restriktionsenzyme kurze Sequenzen benötigen, die ihre spezifischen Schnittstellen flankieren. Die Phusion-Polymerase hingegen besitzt eine *Proofreading*-Aktivität, wodurch sie eine sehr geringe Amplifikationsfehlerrate aufweist. Die in dieser Arbeit verwendeten Standard-PCR-Bedingungen waren für beide Polymerasen identisch.

Tabelle 7.2: In dieser Arbeit verwendete Standard-PCR-Bedingungen

Schritt	Temperatur	Dauer	Zyklus
1. Initiale Denaturierung	98°C	30 s	1
2. Denaturierung	98°C	10 s	35
3. <i>Annealing</i>	primerabhängig	10-30 s	
4. Amplifikation	72°C	15-30 s/kb	
5. Finale Extension	72°C	10 min	1
6. Kühlung	4°C	bis Ende	

7.3.4 *Aufreinigung von PCR Produkten mittels „MN - PCR clean-up and Gel extraction Kit“*

Die Aufreinigung von PCR Produkten mit Hilfe des *PCR clean-up and Gel extraction Kit* basiert auf der Bindung von einzel- oder doppelsträngigen DNA-Fragmenten (50 bp bis ~20 kb) an eine Ionenaustauschersäule. Bestandteile, wie Primer, Enzyme, weitere PCR-Zusätze (z.B. DMSO) oder kleinere DNA-Fragmente können nicht an das Säulenmaterial binden und werden durch Waschung der Säule entfernt.

- aufzureinigendes PCR-Fragment mit zweifachem Volumen NTI-Puffer mischen
- Gemisch auf die Ionenaustauschersäule geben
- 30 s zentrifugieren, 13.000 rpm und Durchlauf verwerfen
- 700 µl NT3-Puffer auf die Säule geben
- 30 s zentrifugieren, 13.000 rpm und Durchlauf verwerfen
- zum Trocknen der Silica-Membran 1 min bei 13.000 rpm zentrifugieren
- den oberen Teil der NucleoSpin-Säule® in ein neues Eppendorf-Gefäß überführen

- 15 - 30 µl NE-Puffer zum Eluieren des DNA-Fragments auf die Membran geben und 1 min bei Raumtemperatur inkubieren
- 1 min bei 13.000 rpm zentrifugieren
- oberen Teil der NucleoSpin-Säule® entsorgen

7.3.5 Ligation

Mit Hilfe des Enzyms T₄-DNA-Ligase ist es möglich die endständigen 5'-Phosphatgruppen und die 3'-Hydroxylgruppen von DNA-Molekülen miteinander zu verknüpfen. Für die Verbindung von DNA-Molekülen ist es notwendig, dass die Enden zueinander kompatibel sind. Das bedeutet, sie müssen entweder zueinander komplementäre Einzelstränge (*sticky-ends*) oder glatte Enden (*blunt-ends*) aufweisen.

Exemplarischer Ligationsansatz mit einem Endvolumen von 20 µl:

- Vorlegen von sterilem Millipore Wasser in ein Eppendorf-Gefäß (Wasservolumen = Endvolumen - Volumen der Ligationskomponenten)
- vorgespaltene DNA-Moleküle im Mischverhältnis Vektor zum zu klonierenden DNA-Fragment von etwa 1:3 dazupipettieren
- 1/10 des Gesamtvolumens Ligationspuffer (10x)
- kalkuliertes Ligase-Volumen (3 U) zugeben und gut mischen
- Bei 16°C ü/N inkubieren
- Inaktivierung des Enzyms durch 15 min bei 65°C

7.4 DNA-TRANSFER MITTELS CHEMISCHER TRANSFORMATION IN KOMPETENTE *E. coli* -ZELLEN

7.4.1 Herstellung chemisch kompetenter *E. coli* Zellen mittels CaCl₂-Methode (Protokoll der Plasmid Factory)

Kompetente Zellen sind in der Lage freie Plasmid-DNA aus dem Medium aufzunehmen. Da *E. coli* kein eigenes Transformationssystem besitzt, muss es im Labor zunächst transformationskompetent gemacht werden. Um dies zu erreichen, gibt es verschiedene Methoden, z. B. die CaCl₂-Methode, die Rubidiumchlorid-Methode oder durch Elektroporation. Die klassische Methode ist die CaCl₂-Methode. Diese Behandlung macht die Zellmembran der kompetenten Zellen für DNA-Moleküle durchlässig. Alle Schritte zur Herstellung kompetenter Zellen sollten bei 4°C durchgeführt werden.

- Ansetzen einer *E. coli* DH5 α \bar{u} /N-Kultur
- Animpfen der Hauptkultur in 250 ml LB-Medium in einem 1 l Schüttelkolben mit einer OD₆₀₀=0.2
- Zellen bis zu einer OD₆₀₀ von 0.8 bis 0.9 wachsen lassen
- Kultivierung abbrechen und Zellen für 15 min auf Eis stellen
- 10 min bei 6000 rpm und 4°C abzentrifugieren
- Überstand verwerfen und im Zentrifugenbecher mit Zellstoff die Überstandsreste absaugen (Dabei nicht ans Pellet kommen!)
- 10 ml eiskalten Puffer 1 zugeben und die Zellen durch Schwenken resuspendieren
- 10 min bei 6000 rpm und 4°C abzentrifugieren und Überstand verwerfen
- 10 ml eiskalten Puffer 2 zugeben und die Zellen durch Schwenken resuspendieren
- Zellen in Eppendorf-Gefäße aliquotieren
- Zellen in flüssigem Stickstoff schockgefrieren und bei -80°C lagern

7.4.2 Transformation kompetenter *E. coli* Zellen

Bei der Transformation werden die kompetenten Zellen einem kurzfristigen Hitzeschock bei 42°C ausgesetzt, der sie befähigt, die Plasmide aus dem Medium ins Cytoplasma aufzunehmen.

- 200 μ l kompetente Zellen auf Eis auftauen
- die zu transformierende DNA zugeben und vorsichtig mischen
- 30 min auf Eis inkubieren
- 2 min Hitzeschock bei 42°C
- 1 ml LB-Medium zugeben und vorsichtig mischen
- 60 min bei 37°C (bei der Transformation eines temperatursensitiven Plasmids bei 30°C) inkubieren
- 100 μ l des Transformationsansatzes auf entsprechender Selektionsplatte ausplattieren
- Rest 5 min abzentrifugieren und Überstand verwerfen
- Zellpellet im Rücklauf des Überstandes resuspendieren und auf entsprechender Selektionsplatte ausplattieren
- Platten über Nacht bei 37°C bzw. 30°C inkubieren

7.5 ERSTELLUNG EINES RNA-STANDARDS UND DURCHFÜHRUNG VON QUANTITATIVER RT-PCR

7.5.1 PCR zur Amplifikation des RNAI- und RNAII-Gens zum Einsatz in der T7-RNA-Polymerase-Reaktion

Um die absoluten RNAI- und RNAII-Konzentrationen mittels quantitativer Reverse-Transkriptase-Polymerasekettenreaktion (qRT-PCR) bestimmen zu können, wird jeweils eine Eichreihe benötigt, die den Zusammenhang zwischen einem gemessenen CT-Wert und der dazugehörigen RNA-Konzentration beschreibt. Für die Erstellung der Eichreihe wurden bekannte RNAI- bzw. RNAII-Konzentrationen vermessen. Anschließend wurden die gemessenen CT-Werte gegen die Konzentrationen aufgetragen und eine Regressionsgerade erstellt. Die Erstellung von RNAI- und RNAII-Standard-Lösungen mit bekannten Konzentrationen erfolgte zunächst über eine Amplifikation des jeweiligen Gens durch PCR. Dabei sind RNAI- bzw. RNAII-Primer eingesetzt worden, die eine Extension besitzen, die aus der Promotersequenz für die T7-RNA-Polymerase besteht. Die entstandenen PCR-Produkte konnten im nächsten Schritt in der T7-RNA-Polymerase-Reaktion eingesetzt werden, um eine RNAI- bzw. eine RNAII-Standard-Lösung zu synthetisieren.

Exemplarischer PCR-Ansatz mit einem Endvolumen von 20 μl :

- 4 μl 5x-Puffer
- 0.4 μl dNTP-Mix (jeweils 200 μM)
- 1 μl RNAIT₇₁ Primer bzw. RNAIIT₇₁ Primer (jeweils 10 μM)
- 1 μl RNAIT₇₂ Primer bzw. RNAIIT₇₂ (jeweils 10 μM)
- 1 μl Plasmid-DNA
- 0.2 μl Phusion-DNA-Polymerase (0.5 U/ μl)
- 0.6 μl DMSO
- 11.8 μl steriles Millipore Wasser

Tabelle 7.3: In dieser Arbeit verwendete PCR-Bedingungen zur Amplifikation des *RNAI*- bzw. *RNAII*-Gens zum Einsatz als *Template* in der T7-RNA-Polymerase-Reaktion

Schritt	Temperatur	Dauer	Zyklus
1. Initiale Denaturierung	98°C	1 min	1
2. Denaturierung	98°C	15 s	35
3. <i>Annealing</i>	primerabhängig	25 s	
4. Amplifikation	72°C	5 s	
5. Finale Extension	72°C	30 s	1
6. Kühlung	4°C	bis Ende	

7.5.2 T7-RNA-Polymerase-Reaktion

Die T7-RNA-Polymerase-Reaktion ist eine gute Möglichkeit das Transkript eines bestimmten Gens zu amplifizieren. Die Methode wurde in dieser Arbeit angewendet, um eine Lösung mit bekannter RNAI- bzw. RNAII-Konzentration herzustellen. Die entstandenen RNAI- bzw. RNAII-Standard-Lösungen sind anschließend in verschiedenen Verdünnungsstufen mittels qRT-PCR vermessen worden, um die jeweilige Eichreihe zu erstellen. Als *Template* der T7-RNA-Polymerase-Reaktion diente das zuvor synthetisierte PCR-Produkt des *RNAI*- bzw. *RNAII*-Gens.

Exemplarischer T7-RNA-Polymerase-Reaktionsansatz mit einem Endvolumen von 20 µl:

- 1 µg DNA des RNAI- bzw. RNAII-PCR-Produkts
- 0.8 µl NTP-Mix (jeweils 1 mM)
- 2 µl T7-RNA-Polymerase (40 U)
- 2 µl Reaktionspuffer (10x)
- 0.5 µl RNase Inhibitor (20 U)
- mit RNase-freiem H₂O auf 20 µl auffüllen

→ 2 Std. bei 37°C inkubieren

7.5.3 Isolierung von RNA mittels „Qiagen - RNeasy plus Mini Kit“

Zur Bestimmung der absoluten RNAI- und RNAII-Konzentrationen in einem *low copy*- und einem *high copy* plasmidtragenden Stamm ist zunächst die Gesamt-RNA mit Hilfe des Qiagen - RNeasy plus Mini Kits isoliert worden. Dabei wurde die genomische DNA über die *gDNA Eliminator spin column* vom Zelllysate abgetrennt und die Gesamt-RNA an die *RNeasy Mini spin column* gebunden und aufgereinigt. Im nächsten Schritt wurde die isolierte RNA mittels qRT-PCR vermessen, so dass die unbekanntenen RNAI- und RNAII-Konzentrationen anschließend anhand ihrer CT-Werte mit Hilfe der Eichreihen bestimmt werden konnten.

- 1 ml *E. coli*-Kultur abzentrifugieren (5 min und 13.000 rpm) und den Überstand verwerfen
- Zellpellet in flüssigem Stickstoff schockgefrieren und sofort für die RNA-Isolierung verwenden oder bei -80°C einfrieren
- Vorbereiten: Pro Zellpellet 700 µl RLT-Puffer versetzt mit 7 µl β-Mercaptoethanol frisch ansetzen und in ein *Fast Protein Tube* geben
- Zellpellet in 200 µl Tris-HCl (10 mM) durch auf- und abpipettieren resuspendieren und in das vorbereitete *Fast Protein Tube* überführen
- Zellen im Hybrid ribolyser bei 30 s, Level 6.5 m/s aufschließen
- Zelllysate bei 3 min auf Eis abkühlen und anschließend bei 4°C für 3 min in einer Eppendorf-Tischzentrifuge die Zelltrümmer abzentrifugieren (13.000 rpm)
- Die nachfolgenden Schritte werden alle auf Eis durchgeführt!
- Überstand in ein frisches RNase-freies Eppendorf-Gefäß überführen und erneut bei 4°C für 3 min und 13.000 rpm zentrifugieren
- Überstand auf die *gDNA Eliminator spin column* geben, um genomische DNA abzutrennen
- 30 s bei 4°C und 13.000 rpm zentrifugieren (Schritt bei Bedarf wiederholen)
- Durchlauf in ein frisches RNase-freies Eppendorf-Gefäß (bei Volumen > 800 µl in ein 12 ml Greiner-Röhrchen) überführen
- 1,5-faches Volumen 100 % Ethanol dazugeben und vortexen
- *RNeasy Mini spin column* auf ein frisches RNase-freies 2 ml Eppendorf-Gefäß setzen
- 700 µl der Probe (inklusive Präzipitat) auf die Säule geben und 15 s bei 13.000 rpm und 4°C zentrifugieren
- Durchfluss verwerfen und den vorherigen Schritt wiederholen bis die komplette Probe die Säulenmembran passiert hat

- *RNeasy Mini spin column* auf ein frisches RNase-freies 2 ml Eppendorf-Gefäß setzen
- 500 µl RPE-Puffer zugeben und bei 4°C für 2 min und 13.000 rpm zentrifugieren
- Durchfluss verwerfen
- 500 µl RPE-Puffer zugeben und bei 4°C für 15 s und 13.000 rpm zentrifugieren
- *RNeasy Mini spin column* auf ein frisches RNase-freies 2 ml Eppendorf-Gefäß setzen
- bei 4°C für 1 min und 13.000 rpm zentrifugieren
- *RNeasy Mini spin column* auf ein frisches RNase-freies 1.5 ml Eppendorf-Gefäß setzen und 30 µl RNase-freies Wasser direkt auf die Säulenmembran geben
- 1 min bei 4°C und 13.000 rpm zentrifugieren um die Gesamt-RNA zu eluieren
- Säule verwerfen und eluierte RNA bei –80°C lagern

7.5.4 Entfernung von DNA mittels DNaseI-Verdau

Um eventuell verbliebene DNA-Reste (sowohl aus der T7-RNA-Polymerase-Reaktion, als auch aus der RNA-Isolierung) zu entfernen, wurde ein Verdau mit dem Enzym DNaseI durchgeführt. Dazu wurden 2 µl DNaseI (2 U) auf die eluierte RNA gegeben und 30 min bei 37°C inkubiert. Die Inaktivierung des Enzyms erfolgte durch die Zugabe von 1 µl EDTA und einer Inkubation bei 65°C für 10 min.

7.5.5 Konzentrationsbestimmung einer RNA-Lösung

Die Konzentrationsbestimmung einer RNA-Lösung erfolgte durch eine spektrophotometrische Messung bei 260 nm je eines Mikroliters der RNA-Lösung im NanoDrop- Gerät (Peqlab).

7.5.6 qRT-PCR mittels „Bioline - SensiMix Probe One-Step Kit“

Die Methode der qRT-PCR dient der Vervielfältigung von Nukleinsäuren mit einer gleichzeitigen Quantifizierung anhand von Fluoreszenzmessungen. Für die Erzeugung der Fluoreszenzsignale wird dem PCR-Ansatz ein Fluoreszenzfarbstoff zugegeben, der in die doppelsträngigen DNA-Moleküle interkaliert. Dabei nimmt die Fluoreszenz mit der Bildung der PCR-Produkte zu, so dass anhand der Fluoreszenzmessungen eine Aussage über die amplifizierte DNA-Menge getroffen werden kann. Mit Hilfe der qRT-PCR lassen sich neben DNA-Mengen auch RNA-Mengen quantifizieren. Dazu wird die RNA

zunächst mittels eines Enzyms namens Reverse Transkriptase in eine copy-DNA (cDNA) umgeschrieben, welche als *Template* für die PCR-Reaktion dient. Im Rahmen dieser Arbeit wurden die qRT-PCR-Analysen im Opticon 2TM (Bio-Rad) unter Verwendung des Fluoreszenzfarbstoffes SYBR[®] Green I durchgeführt. Das verwendete *SensiMix Probe One-Step Kit* der Firma Biorline ist so konzipiert, dass die reverse Transkription (RT) und die quantitative PCR in einem Ansatz durchgeführt werden. Als Negativkontrolle wurde in dieser Arbeit anstelle des RNA-*Templates* 1 µl RNase-freies Wasser pipettiert.

Exemplarischer qRT-PCR-Ansatz mit einem Endvolumen von 19 µl:

- 6.2 µl RNase-freies Wasser
- 10.8 µl 2x SensiMix SYBR No-Rox One Step inklusive 10 U/µl RiboSafe RNase Inhibitor
- 1 µl RNAI-Ic1 Primer bzw. mRNAI-Ic2neu Primer (jeweils 10 pM)
- 1 µl Gesamt-RNA

Bei der reversen Transkription wurde nur ein Primer zugegeben, da das RNAI-Molekül komplementär zum RNAI-Molekül ist. Hätte man beide Primer (*forward* und *reverse* Primer) dazu gegeben, wären sowohl RNAI- als auch RNAI-Moleküle in cDNA umgeschrieben worden und die später amplifizierte RNA-Menge hätte nicht mehr in RNAI und RNAI unterschieden werden können. Die reverse Transkription wurde für 10 min bei 42°C in dem PCR-Gerät *MJ Research PTC-100 Programmable Thermal Controller* durchgeführt. Nach Abschluss der RT-Reaktion wurde 1 µl des zweiten Primers hinzugegeben und die quantitative PCR im Opticon 2TM (Bio-Rad) wurde unter Verwendung des folgenden Programms gestartet:

Tabelle 7.4: In dieser Arbeit verwendete Bedingungen für die quantitative PCR

Schritt	Temperatur	Dauer	Zyklus
1. Initiale Denaturierung	95°C	10 min	1
2. Denaturierung	95°C	15 s	40
3. <i>Annealing</i>	69°C	15 s	
4. Amplifikation	72°C	15 s	
5. Finale Extension	72°C	30 s	1
6. Kühlung	4°C	bis Ende	

7.6 ABSOLUTE QUANTIFIZIERUNG VON PLASMID-DNA

7.6.1 Absolute Quantifizierung der Plasmid-DNA durch die
Firma Plasmid Factory

Bei der Messung der intrazellulären Plasmidkonzentration durch die Plasmid Factory GmbH & Co. KG (Meisenstraße 96, 33607 Bielefeld; Kooperationspartner dieses Promotionsprojekts) wurde die pDNA zunächst über eine NucleoBond PC20-Säule (Macherey-Nagel) aufgereinigt und im NanoDrop-Gerät bei 260 nm vermessen. Anschließend erfolgte die gelelektrophoretische Auftrennung und die densitometrische Bestimmung des pDNA-Anteils an der Gesamt-DNA mit Hilfe der Software LabImage 1D L 340 (Intas Science Imaging GmbH).

- 400 µl S₁-Puffer zum Zellpellet ins Eppendorf-Gefäß zugeben und durch Vortexen resuspendieren
- 400 µl S₂-Puffer für die Zellyse zugeben und mehrfach invertieren
- 400 µl S₃-Puffer für die Neutralisation zugeben und mehrfach invertieren
- 7 min bei 14.500 rpm zentrifugieren und währenddessen die NucleoBond PC20-Säule (Macherey-Nagel) mit 1 ml N₂-Puffer äquilibrieren
- Zellysat (Überstand) möglichst ohne Präzipitat auf die Säule pipettieren
- zum Waschen der pDNA 3 mal 1 ml N₃-Puffer auf die Säule geben
- währenddessen ein neues Eppendorf-Gefäß unter die Säule stellen
- 800 µl N₅-Puffer zur Elution auf die Säule geben
- Säule verwerfen
- 600 µl Isopropanol zur Fällung der pDNA zugeben und kräftig schütteln
- 30 min bei 14.500 rpm zentrifugieren, Überstand abnehmen und verwerfen
- 500 µl 70% Ethanol zum Waschen der pDNA zugeben
- kurz bei 14.500 rpm zentrifugieren und den Überstand vorsichtig abnehmen und verwerfen
- pDNA im offenen Eppendorf-Gefäß bei Raumtemperatur ca. 30 min trocknen lassen
- 20 µl WFI- oder TE-Puffer zugeben und pDNA zum Lösen mindestens 1 h bei Raumtemperatur stehen lassen

7.6.2 Absolute Quantifizierung der Plasmid-DNA durch die Firma CARPEGEN

Die Bestimmung der Plasmidkonzentration durch die Firma CARPEGEN GmbH (Mendelstr. 11, 48149 Münster) erfolgte unter Anwendung einer sondenbasierten RT-PCR. Mit Hilfe dieser Methode werden die ribosomalen Operons der gDNA quantifiziert und so der Anteil genomischer DNA bestimmt. Anschließend wird die errechnete gDNA-Menge in Relation zur photometrisch bestimmten Gesamt-DNA-Konzentration gesetzt und die pDNA-Konzentration ermittelt (GesamtDNA - gDNA = pDNA).

7.6.3 Absolute Quantifizierung der Plasmid-DNA mittels Hochleistungsflüssigkeitschromatographie

Zur absoluten Quantifizierung der Plasmid-DNA mittels Hochleistungsflüssigkeitschromatographie (HPLC von engl. *high performance liquid chromatography*) wurde zunächst eine Verdünnungsreihe mit bekannten pDNA-Konzentrationen vermessen, um eine Eichgerade zu erstellen. Die Eichgerade dient dazu, die unbekanntes Plasmidmengen zu ermitteln. Anschließend wurden die, über alkalische Lyse, hergestellten Zellysate der Proben mittels HPLC vermessen. Die Puffer zur Herstellung der Zellysate stammen aus dem Plasmid-Isolierungskit „Thermo Scientific - GeneJET“

- Zellpellet in 125 µl Resuspensionslösung in einem Eppendorf-Gefäß resuspendieren
- 125 µl Lyselösung dazugeben und sofort vorsichtig 4-6 mal invertieren (Scherkräfte vermeiden!)
- 175 µl Neutralisationslösung dazugeben und sofort vorsichtig 4-6 mal invertieren
- Zelltrümmer 30 min bei 14.000 rpm in einer Eppendorf-Tischzentrifuge abzentrifugieren
- 300 µl Überstand als Zellysate für die HPLC-Messungen einsetzen und davon 200 µl injizieren

HPLC-Bedingungen:

Säule: VA 125/6 NUCLEOGEN® 4000-7 DEAE (Macherey-Nagel)

Detektor: UV-Detektor, 260 nm

Laufmittel A: 20 mM Kaliumphosphatpuffer (10 mM KH₂PO₄ + 10 mM K₂HPO₄)

Laufmittel B: Laufmittel A + 1.5 M KCl, linearer Gradient von 300 mM KCl (20 % B, 80 % A) bis 1.5 mM KCl (100 % B) in 50 min

Flussrate: 1 ml/min

Temperatur: 22°C

7.7 DURCHFÜHRUNG VON METABOLOMANALYSEN

7.7.1 *Zellernte durch Schockgefrieren*

Für die Untersuchung von intrazellulären Metaboliten ist es sehr wichtig, dass die Ernte der Zellen so schnell wie möglich durchgeführt wird, weil Metabolite innerhalb kürzester Zeit enzymatisch umgewandelt werden können. Jede Art von Stress für die Zellen (z. B. Temperaturwechsel durch langes Stehen oder hohe Zentrifugationszeiten) führt zu Veränderungen der intrazellulären Metabolitzusammensetzung. Möchte man die intrazellulären Metabolitkonzentration unter definierten Kultivierungsbedingungen untersuchen, ist es daher notwendig diese während der Ernte möglichst konstant zu lassen bzw. unvermeidbare Schwankungen so klein wie möglich zu halten, bevor die Zellen in flüssigem Stickstoff schockgefroren werden um alle enzymatischen Aktivitäten zu stoppen.

- 100 ml M9-Medium supplementiert mit FeSO₄ mit einer OD₆₀₀=0.1 animpfen
- Bakterienkulturen bis in die spät-exponentielle Phase bei 200 rpm wachsen lassen
- 100 ml Bakterienkultur auf zwei 50 ml verschraubbare Reaktionsgefäße aufteilen
- 3 min bei 4000 rpm und 37°C zentrifugieren (Sigma-Zentrifuge)
- Überstand abgießen und Zellpellet in flüssigem Stickstoff schockgefrieren
- Bei -80°C lagern

7.7.2 *Extraktion und Derivatisierung der intrazellulären Metabolite*

Die Extraktion der intrazellulären Metabolite erfolgte durch Aufschluss der Zellen im Ribolyser (Bertin) unter Zugabe von Methanol. Für die Derivatisierung wurde das Silylierungsmittel MSTFA (N-Methyl-N-Trimethylsilyl-Trifluoroacetamid) verwendet. Bei der Behandlung mit MSTFA werden die Wasserstoffatome polarer Gruppen gegen Trimethylsilyl-Gruppen ausgetauscht.

- Gefrorenes Zellpellet durch Lyophilisierung trocken und 5 bis 10 mg in Schraubdeckel-Eppendorf-Gefäße abwiegen

- 500 mg Glasbeads in das Schraubdeckel-Eppendorf-Gefäß hinzufügen
- 1 ml Extraktionslösung bestehend aus 80% Methanol mit 10 mM Ribitol zugeben (Ribitol dient dabei als interner Standard für die spätere Quantifizierung)
- Zellen im Ribolyser aufschließen (3 x 45 s, 15 s Pause)
- 20 min bei 14.000 rpm abzentrifugieren
- Reacti-Therm III Gerät auf 37°C stellen
- 800 µl vom Überstand in ein Reacti-Glasgefäß geben
- Reacti-Glasgefäß in das Reacti-Therm III Gerät stellen und das Methanol durch „Abblasen“ mit einem Stickstoffstrom entfernen
- Nachdem das Methanol verdampft ist, einen Magnetührfisch und 100 µl Methoxylamin Hydrochlorid in Pyridin (20 mg/ ml) zugeben
- Magnetührfunktion im Reacti-Therm III Gerät anstellen und 1.5 Std. bei 37°C inkubieren
- 100 µl MSTFA und 20 µl eines Alkan-Mixes zugeben (Der Alkan-Mix besteht aus C12-, C15-, C18-, C19-, C22-, C28-, C32-Alkanen je 0.24 mg/ ml und dem C36-Alkan (0.48 mg/ ml) in Pyridin gelöst. Er dient dazu, die systembedingte Verschiebungen in den Retentionszeiten auszugleichen.)
- 30 min bei 37°C im Reacti-Therm III Gerät inkubieren
- Nach der Inkubation 5 min bei 4.000 rpm gegen Schwebstoffe abzentrifugieren
- Überstand mit Hilfe einer Pasteurpipette in ein GC-MS-Glasgefäß pipettieren
- Probe im GC-MS vermessen

7.7.3 Gaschromatographie-Massenspektrometrie

Die GC-MS-Messungen wurden auf einem „Trace GC Polaris Q Massenspektrometer“ der Firma *Thermo Scientific*TM durchgeführt. Das Gerät besteht aus einem, an ein Massenspektrometer gekoppelten Gaschromatographen und ist mit der Säule *Rtx*[®]-5MS column (30 m, iD 0.25, df 0.25 m) der Firma Restek ausgerüstet. Bei allen GC-MS-Messungen wurde jeweils 1 µl Probe splittlos injiziert.

Temperaturprogramm des GC-Ofens:

- 3 min bei 80°C
- Mit 5°C/ min auf 325°C heizen
- 2 min bei 325°C heizen

Die Temperatur der Überleitung betrug 250°C und die der Ionenquelle lag bei 220°C. Das Massenspektrum wurde in einem Scanbereich von 50 bis 750 m/z aufgenommen.

Nach vier Messungen wurde jeweils eine Leerprobe vermessen, um überprüfen zu können, ob Metabolitüberträge stattgefunden haben. Die Dateien der Rohdaten wurden mit Hilfe der Xcalibur-Software in cdf-Dateien umgewandelt und anschließend in die MeltDB-Software importiert.

7.7.4 Auswertung der GC-MS-Messungen

Für die Metabolomanalysen wurden drei biologische Replikate à zwei technische Replikaten vermessen. Die Auswertung der GC-MS-Messungen fand unter Verwendung der MeltDB-Software (Kessler *et al.*, 2013), (Neuweger *et al.*, 2008) statt. Sie begann mit der Durchführung der *MeltDB Warped Peak Detection* Methode mit einem *Signal to Noise* Verhältnis von 10. Diese Methode sucht nach Massenspuren, die zu einer bestimmten Zeit eine Peak-Form zeigen und ein lokales Maximum besitzen. In dem zugehörigen Massenspektrum sind die Intensitäten aller Massen zu dieser Zeit enthalten. Auf die Peak-Detektion folgte das sogenannte *Multiple Profiling*. Dabei werden Peaks gefunden, die sich in ihren Massenspektren und ihren Retentionszeiten ähneln und dabei in einer bestimmten Anzahl der Chromatogramme vorkommen. Diese „bestimmte Anzahl“ kann als Parameter auf z. B. 50 % eingestellt werden. In dieser Arbeit wurden die Einstellungen so gewählt, dass solche Peaks gefunden werden, die in 80 % der Chromatogramme vorkommen. Da diese Peaks noch nicht identifiziert sind, werden sie als „TAG“ markiert. Dies ist ein generischer Name, der sich aus der Retentionszeit und der Hauptmasse des Spektrums zusammensetzt. Im Anschluss daran erfolgte die Identifikation bekannter Metabolite durch den Vergleich ihrer Retentionszeiten und Massenspektren mit den Retentionszeiten und Massenspektren zuvor vermessener Referenzsubstanzen. Für die Berechnung der Retentionszeiten wurden die in dem Alkan-Mix vermessenen Alkane als Retentionindizes zu Hilfe genommen. Im Fall von unbekanntem Metaboliten, die nicht als Referenzsubstanz vorlagen, wurden die zugehörigen Massenspektren mit Spektren verglichen, welche in den Datenbanken *Golm metabolite database* (GMD, 2008) und *NIST mainlib* (*National Institute for Standards and Technology*, 2009) hinterlegt sind. Nach der Identifikation erfolgte die relative Quantifizierung der Metabolite über die Peakfläche mit einer Normalisierung auf den internen Standard Ribitol und das zu Beginn eingewogene Zellrockengewicht. Dabei war sicherzustellen, dass Ribitol nicht als intrazelluläre Komponente vorliegt. Im letzten Schritt wurden mit Hilfe der MeltDB-

Software (Kessler *et al.*, 2013), (Neuweger *et al.*, 2008) verschiedene statistische Analysen (Hauptkomponentenanalyse, *Heatmaps*) durchgeführt.

7.8 DURCHFÜHRUNG VON ^{13}C -FLUSSANALYSEN

7.8.1 Anzucht der Bakterien und Ernte durch Schockgefrieren

Das experimentelle Vorgehen bei der Durchführung der ^{13}C -Flussanalysen geschah in Anlehnung an Weckwerth (2007). Die Anzucht erfolgte über zwei Vorkulturen auf natürlich markierter Glukose und einer Hauptkultur mit ^{13}C -markierter Glukose. Zum Ernten wurden die Zellen in flüssigem Stickstoff schockgefroren.

- 10 ml LB Medium (wenn benötigt mit Antibiotika versetzt) mit einer einzelnen *E. coli* Kolonie animpfen und bei 37°C (300 rpm) 8 bis 12 Std. kultivieren (erste Vorkultur)
- Die Zellen 5 min bei 4.000 rpm zentrifugieren und mit steriler 0.9 % NaCl-Lösung waschen
- 25 ml PD1-Medium (mit 5 g/l Glukose und, wenn benötigt, mit Antibiotika) mit den gewaschenen Zellen bei einer $\text{OD}_{600}=0.1$ animpfen (zweite Vorkultur)
- Nach Erreichen der exponentiellen Phase die Zellen durch Zentrifugation ernten (5 min bei 4.000 rpm)
- Zweimal mit steriler 0.9 % NaCl-Lösung waschen
- 25 ml PD1-Medium (mit 5 g/l ^{13}C -markierter Glukose) mit den gewaschenen Zellen bei einer $\text{OD}_{600}=0.03$ animpfen (Hauptkultur)
- Nach Erreichen der exponentiellen Phase 1 ml Zellkultur in einem Eppendorf-Gefäß abzentrifugieren (10 min bei 12.000 rpm und 4°C)
- Überstand verwerfen und Zellpellet in flüssigem Stickstoff schockgefrieren
- bei -80°C lagern

7.8.2 Proteinhydrolyse und Derivatisierung

Die Proteinhydrolyse erfolgte durch Zugabe von 50 μl 6 M HCl pro 1 mg Zelltrockengewicht. Das Zelltrockengewicht wurde mit folgender Gleichung berechnet

$$\text{Zelltrockengewicht}[\text{mg}] = 0.38[\text{mg}] \cdot \text{OD}_{600}(\text{Erntezeitpunkt}), \quad (7.1)$$

wobei OD_{600} die optische Dichte bei einer Wellenlänge von 600 nm beschreibt und dimensionslos ist. Die Gleichung wurde von Juan Carlos Sigala Anlanis (*Universidad Nacional Autónoma de México*) zur Verfügung gestellt.

- Zellpellet auftauen und zweimal mit 700 μl sterilem Wasser bei Raumtemperatur waschen
- Bei 12.000 rpm 8 min abzentrifugieren und Überstand verwerfen
- 6 M HCl zugeben (50 μl pro 1 mg Zelltrockengewicht) und Pellet darin lösen
- 24 Std. bei 105°C inkubieren
- Die abgekühlten Hydrolysate mit 6 M NaOH neutralisieren (50 μl pro 1 mg Zelltrockengewicht)
- Probe durch Filtration (Ultrafree-CL Filter, Millipore) von partikulären Verunreinigungen befreien
- 50 μl des Durchlaufs in ein frisches Eppendorf-Gefäß geben und bei 37°C Vakuum trocknen (Eppendorf-SpeedVac)
- Komplette wasserfreie Probe in 50 μl frischem Dimethylformamid/ Pyridin (DMF + 0.1% Pyridin) lösen und in ein frisches Eppendorf-Gefäß überführen
- Zur Derivatisierung 50 μl MBDSTFA zugeben und für 30 min und 80°C im Thermoshaker inkubieren
- Nach der Inkubation 5 min bei 4000 rpm gegen Schwebstoffe abzentrifugieren
- Überstand mit Hilfe einer Pasteurpipette in ein GC-MS-Glasgefäß pipettieren
- Probe im GC-MS vermessen

7.8.3 Gaschromatographie-Massenspektrometrie

Die GC-MS-Messungen wurden auf dem Gerät „TRACE GC Ultra GC ITQ 900 GC-Ion Trap MS“ der Firma *Thermo Scientific*TM durchgeführt. Das Gerät besteht aus einem, an ein Massenspektrometer gekoppelten Gaschromatographen und ist mit der Säule *Rtx*[®]-5MS column (30 m, iD 0.25, df 0.25 m) der Firma Restek ausgerüstet.

Bei allen GC-MS-Messungen der ^{13}C -markierten Proben wurde jeweils 1 μl Probe splittlos injiziert.

Temperaturprogramm des GC-Ofens:

- 2 min bei 150°C
- Mit 10°C/ min auf 280°C heizen
- 3 min bei 280°C heizen

Die Temperatur der Überleitung betrug 280°C und die der Ionenquelle lag bei 200°C . Das Massenspektrum wurde in einem Scanbereich von 50 bis 570 m/z aufgenommen. Nach vier Messungen wurde jeweils eine Leerprobe vermessen, um überprüfen zu können, ob Metabolitüberträge stattgefunden haben.

7.8.4 Auswertung der GC-MS-Messungen

Für jeden Bakterienstamm wurden zwei biologische Replikate à zwei technische Replikaten vermessen. Die Auswertung der Chromatogramme und Massenspektren erfolgte mit Hilfe der Xcalibur-Software. Dabei wurde das Massenspektrum jedes detektierbaren Aminosäurepeaks hinsichtlich des Isotopomer-Verteilungsmusters des zu betrachtenden Fragments der Aminosäure untersucht. Die Massen der leichtesten Isotopomere der in dieser Arbeit betrachteten Aminosäurefragmente sind in Tabelle 10.2 dargestellt. Die auf diese Weise ermittelten spezifischen Markierungsmuster wurden für die Berechnung der metabolischen Flussverteilung mit Hilfe des stoichiometrischen Modells verwendet.

Teil V

ERGEBNISSE

MODELLIERUNG DER PLASMIDREPLIKATIONSKONTROLLE VON *COLE*₁-ÄHNLICHEN PLASMIDEN

Es ist das Ziel dieser Modellierungsarbeit, die regulatorischen Mechanismen der *ColE1*-Plasmidreplikationskontrolle mit Hilfe eines mathematischen Modells zu untersuchen und zu verstehen. An dieser Stelle soll zunächst eine Übersicht über den Modellierungsprozess gegeben werden (Abbildung 8.1). Eine detailliertere Beschreibung der einzelnen Schritte erfolgt in den weiteren Unterkapiteln.

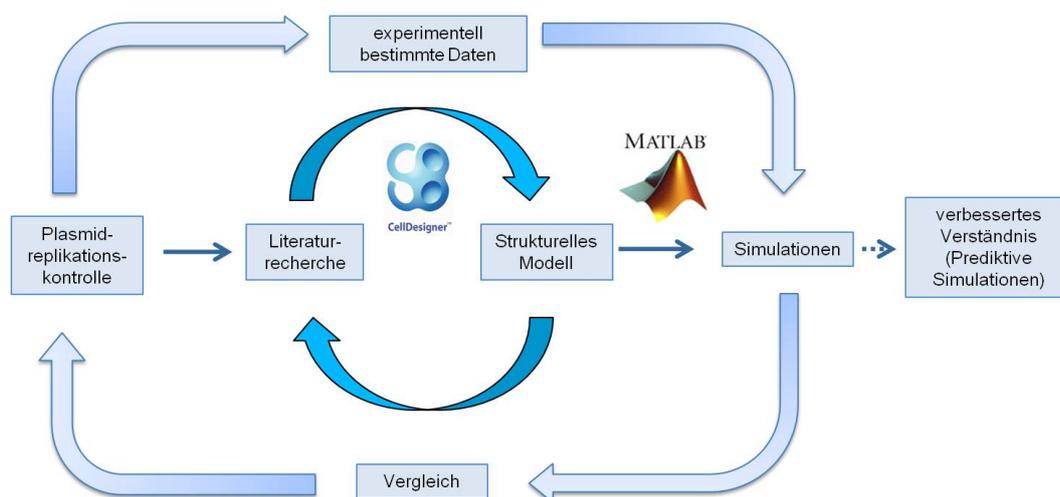


Abbildung 8.1: Übersicht zum Vorgehen bei der Modellierung der Plasmidreplikationskontrolle.

Der Modellierungsprozess begann mit der Literaturrecherche zur Plasmidreplikationskontrolle. Mit Hilfe des Literaturwissens wurde dann zunächst ein strukturelles Modell erstellt. Nachdem überprüft wurde, ob alle relevanten Reaktionen in dem Modell enthalten sind, wurde auf Grundlage des strukturellen Modells, im nächsten Schritt ein dynamisches Modell erstellt. In dem dynamischen Modell sind alle beteiligten Reaktionen durch mathematische Gleichungen beschrieben, welche mit experimentellen Daten untermauert sind. Mit Hilfe dieses Modells ist es möglich, die Plasmidreplikationskontrolle zu simulieren. Die Simulationen wurden mit experimentellen Messungen zum Verhalten der Plasmidreplikationskontrolle verglichen. Stimmt sie nicht überein wurde das Modell kontinuierlich angepasst, bis die Simulationen dem realen Verhalten mit der geforderten Genauigkeit entsprachen.

Mit Hilfe des in dieser Arbeit erstellten Modells zur *ColE1*-Plasmidreplikationskontrolle soll folgender Fragestellung nachgegangen werden: Welchen Effekt hätte es auf die Plasmidproduktion, würde man tRNA-Moleküle

derart modifizieren, dass die Beladung mit einer Aminosäure durch die Aminoacyl-tRNA-Synthetase nicht mehr möglich wäre? Die Hypothese zur Beantwortung der Fragestellung ist, dass die Zelle sich nicht unter Stressbedingungen befindet, aber dennoch eine erhöhte Menge unbeladener tRNA-Moleküle in Form von modifizierten tRNAs aufweist, was sich positiv auf die Plasmidproduktion auswirkt.

Zur Untersuchung dieser Hypothese wurde zunächst eine intensive Literaturrecherche durchgeführt, um zu erfahren, was bereits über die *ColE1*-Plasmidreplikationskontrolle bekannt ist. Der Schwerpunkt lag dabei auf der Identifizierung aller relevanten Komponenten und ihren Interaktionen. Des Weiteren galt es herauszufinden, ob zusätzlich verwendbare experimentelle Daten vorliegen, die sich in das Modell integrieren lassen. Basierend auf Literaturdaten wurde ein erstes strukturelles Modell mit der CellDesigner Software (Funahashi *et al.*, 2008) erstellt, welches sämtliche, bekannte Komponenten und Reaktionen zur *ColE1*-Plasmidreplikationskontrolle visualisiert. Anschließend folgte ein iterativer Prozess, in dem das Modell mit weiterer Literatur verglichen wurde, um zu überprüfen, ob alle eingebauten Reaktionen für die zu untersuchende Fragestellung notwendig sind oder ob überflüssige Verbindungen enthalten sind. Gleichzeitig ist dabei kontrolliert worden, ob wichtige Interaktionen oder Elemente übersehen wurden, die in dem Modell noch zu berücksichtigen sind. Dieser Prozess wurde fortgesetzt, bis das strukturelle Modell einen guten Überblick über das Netzwerk der *ColE1*-Plasmidreplikationskontrolle und alle relevanten Komponenten und Reaktionen darstellte.

Der nächste Schritt bestand aus der mathematischen Beschreibung der zuvor im strukturellen Modell visualisierten Reaktionen unter der Verwendung des Massenwirkungsgesetzes. Nun konnte das, in dieser Arbeit, mittels MATLAB (*Mathworks*®) programmierte Modell zur Durchführung von Simulationen für die Plasmidproduktion verwendet werden. Zunächst wurde dabei überprüft, ob die mathematische Beschreibung stimmig ist und die Simulationen reibungslos ablaufen. Dazu wurde, z. B. durch den Vergleich verschiedener Integrationsalgorithmen, untersucht, ob das Modell die biologischen Verhältnisse korrekt abbildet: Es sollte für den *ColE1 high copy* Plasmid eine größere pDNA-Produktion simulieren, als für den *ColE1 low copy* Plasmid. Des Weiteren wurde die Kinetik kontrolliert, indem untersucht wurde, ob die Produktion konstant abläuft und nicht oszilliert oder plötzlich einbricht.

Nach Abschluss dieser Prüfung sind die Simulationen in den weiteren Schritten mit gemessenen Ergebnissen zur Plasmidproduktion quantitativ verglichen worden.

Da die Simulationen zu diesem Zeitpunkt noch nicht mit dem realen Verhalten übereinstimmten, wurden experimentell bestimmte Daten in das Modell integriert. Die dabei eingefügten Daten stammen zum Teil aus der Literatur, wurden im Labor gemessen oder durch Parameteranpassung bestimmt. Auf die genaue Vorgehensweise zur Generierung dieser Daten wird in dem Kapitel 8.2 näher eingegangen.

Anschließend wurde die Plasmidproduktion erneut simuliert und die Simulationen wurden wieder mit dem realen Verhalten verglichen. Diese Schritte wurden wiederholt, bis die Simulationen mit den gemessenen Ergebnissen übereinstimmten und das Modell zur Untersuchung der *ColE1*-Plasmidreplikationskontrolle und der zu Beginn aufgestellten Hypothese verwendet werden konnte.

8.1 ERSTELLUNG DES STRUKTURELLEN MODELLS

Das strukturelle Modell der *ColE1*-Plasmidreplikationskontrolle (Abb. 8.2) zeigt schematisch alle beteiligten Elemente und Reaktionen, um die Replikationskontrolle eines *ColE1*-Plasmids zu beschreiben, wobei die Kontrolle eines *high copy* und eines *low copy* Plasmids simultan dargestellt sind. Der Unterschied zwischen der Beschreibung eines *high copy* und eines *low copy* Plasmids besteht aus drei zusätzlichen Reaktionen in der Replikationskontrolle des *low copy* Plasmids. Diese Reaktionen (Reaktion 1, 2 und 13 des, in Kapitel 8.3 dargestellten, Differentialgleichungssystems) sind in dem strukturellen Modell mit einem roten Kasten markiert und beschreiben die Regulation durch das Rom-Protein, welche im *high copy* Plasmid aufgrund der Deletion des *rom*-Gens nicht stattfindet (Lin-Chao *et al.*, 1992), (Yanisch-Perron *et al.*, 1985).

Bei der Durchführung der Literaturrecherche zur Erstellung des strukturellen Modells wurde festgestellt, dass bereits kinetische Modelle zur Regulation der *ColE1*-Plasmidreplikation existieren (Ataai & Shuler, 1986) (Brendel & Perelson, 1993), (Bremer & Lin-Chao, 1986), (Keasling & Palsson, 1989a), (Keasling & Palsson, 1989b), (Wang *et al.*, 2002). Da diese Modelle jedoch nur bestimmte Bereiche der Replikationskontrolle betrachten, soll in dieser Arbeit ein dynamisches Modell vorgestellt werden, welches die Regulation der *ColE1*-Plasmidreplikation als Ganzes beschreibt.

Das Modell von Brendel & Perelson (1993) beschreibt die *in vivo* Kontrolle der *ColE1*-Plasmidreplikation durch RNAI und RNAII in An- und Abwesenheit des Rom-Proteins. Mit ihrem Modell konnten sie zum einen zeigen, dass die Plasmidkopienzahl in Anwesenheit eines intakten *rom*-Gens abnimmt und zum anderen, dass die Plasmidkon-

zentration bei geringer Wachstumsrate zunimmt. (Brendel & Perelson, 1993)

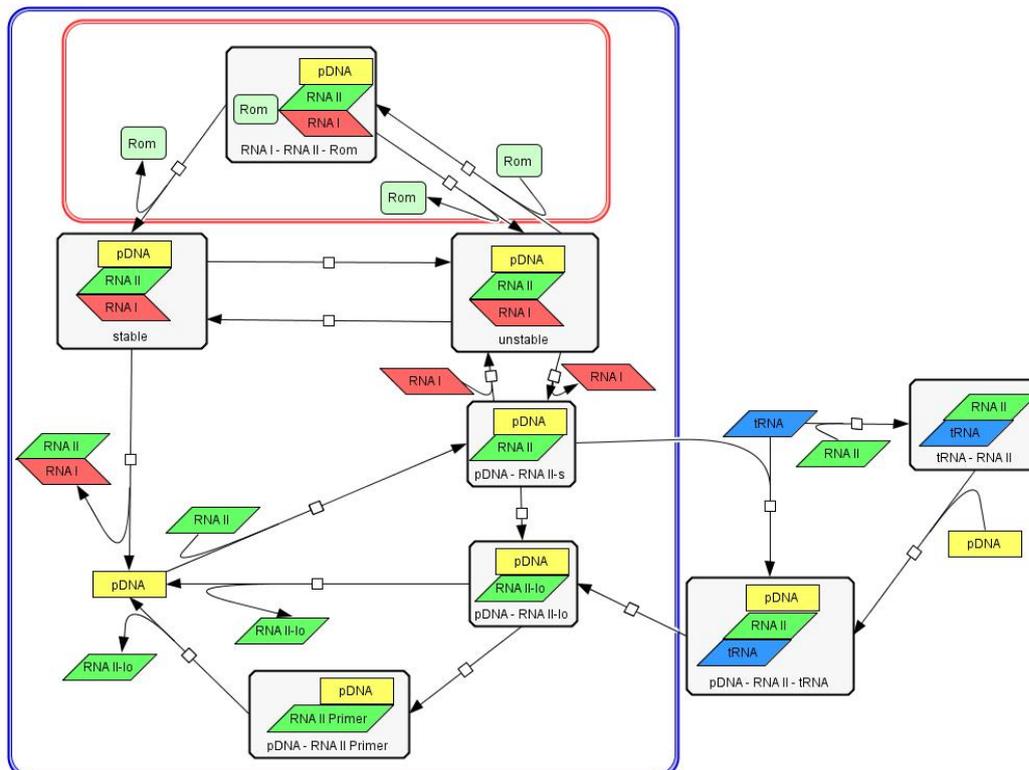


Abbildung 8.2: Strukturelles Modell der *ColE1*-Plasmidreplikationskontrolle.

Das strukturelle Modell stellt eine Erweiterung des von Brendel & Perelson (1993) vorgeschlagenen Modells (blauer Kasten) zur *ColE1*-Plasmidreplikationskontrolle dar. Der Unterschied zwischen dem Modell zur Beschreibung eines *high copy* und eines *low copy* Plasmids besteht aus drei zusätzlichen Reaktionen in der Replikationskontrolle des *low copy* Plasmids (roter Kasten). Diese Reaktionen beschreiben die Kontrolle durch das Rom-Protein und finden bei einem *high copy* Plasmid nicht statt. Die Reaktionen außerhalb des blauen Kastens wurden im Rahmen dieser Arbeit hinzugefügt und beschreiben die Kontrolle durch unbeladene tRNA-Moleküle.

Es ist mit Hilfe der in der Literatur vorhandenen Modelle nicht möglich, den Effekt einer modifizierten tRNA auf die Plasmidproduktion zu untersuchen, da diese Modelle jeweils nur bestimmte Bereiche der Replikationskontrolle betrachten. Das in dieser Arbeit erstellte Modell ist eine Erweiterung des von Brendel & Perelson (1993) vorgeschlagenen Modells. Es beinhaltet die *in vivo ColE1*-Plasmidreplikationskontrolle durch die regulatorischen RNAs (RNA I und RNA II) in An- und Abwesenheit des Rom-Proteins und beschreibt zusätzlich die Regulation durch unbeladene tRNA-Moleküle. Der Teil des von Brendel & Perelson (1993) erstellten Modells ist in der Abbildung 8.2 blau eingerahmt dargestellt. Die übrigen Reaktionen beschreiben die Kontrolle durch die unbeladenen tRNA-Moleküle. Dabei gibt es zwei Möglichkeiten: Die unbeladenen tRNA-Moleküle

können entweder an plasmidgebundene RNAII-Moleküle binden oder zunächst an freie RNAII-Molekülen komplexieren und anschließend an die Plasmid-DNA binden. In beiden Fällen wird die inhibitorische Bindung der RNAI-Moleküle verhindert, so dass die Elongation des RNAII-Primers stattfinden kann.

Die in dem strukturellen Modell (Abb. 8.2) dargestellte Plasmidreplikation beginnt mit der Synthese eines Primer-Vorläufer-Moleküls, genannt RNAII. Dazu wird von einer 555 bp aufwärts des Replikationsursprungs liegenden Region ein 100-300 Nukleotide langes Transkript synthetisiert. Dieses Transkript hybridisiert anschließend mit einer Region nahe dem Replikationsursprung zu einem RNA-DNA-Komplex (pDNA – RNAII – s). (Tomizawa, 1986) Zu diesem Zeitpunkt der Plasmidreplikation gibt es drei mögliche Szenarien für den weiteren Verlauf:

In dem ersten Szenario bindet ein RNA-Molekül mit dem Namen RNAI an das Plasmidgebundene RNAII-Molekül. Das regulatorische RNAI-Molekül wird vom Gegenstrang kodiert und seine Synthese beginnt 445 bp aufwärts des Replikationsursprungs. Sie erfolgt in entgegengesetzter Richtung zur RNAII-Transkription, so dass das RNAI-Transkript komplementär zum RNAII-Transkript ist. Das RNAI-Molekül ist daher in der Lage den Plasmid-RNAII-Hybrid zu binden und einen sogenannten *kissing complex* zu bilden (pDNA – RNAII – RNAI_{unstable}). Durch die Ausbildung dieses transienten Komplexes ist die Elongation des RNAII-Moleküls blockiert und es findet keine Primer-Reifung statt (Abb. 3.5 A). (Tomizawa, 1986)

In dem zweiten Szenario (Abb. 3.5 B) findet die Elongation des RNAII-Moleküls bis zu einer Länge von 360 bp statt (pDNA – RNAII – lo), bevor ein inhibitorisches RNAI-Molekül binden kann. Anschließend wird das RNAII-Molekül von dem Enzym RNase H gespalten, so dass eine 'OH-Gruppe frei wird und das RNAII-Molekül als Primer für die Plasmidreplikation dienen kann (pDNA – RNAII – Primer). In der darauffolgenden Reaktion findet die Elongation des RNAII-Primers durch die DNA Polymerase I statt und die pDNA wird verdoppelt. (Itoh & Tomizawa, 1980)

Das dritte Szenario beschreibt die Bindung einer unbeladenen tRNA (tRNA_{un}) an das pDNA-gebundene RNAII-Molekül (pDNA – RNAII – tRNA_{un}). In diesem Fall kann kein RNAI-Molekül mehr binden, so dass die inhibitorische Regulation durch RNAI verhindert wird. Die Elongation und anschließende Modifikation des RNAII-Moleküls durch die RNase H ist jedoch weiterhin möglich, so dass die Plasmidreplikation stattfinden kann (Grabherr & Bayer, 2002). In diesem Modell wird davon ausgegangen, dass unbeladene tRNA- und RNAII-Moleküle frei in der Zelle diffundieren, so dass auch un-

gebundene RNAII-Moleküle an unbeladene tRNAs binden können ($\text{RNAII} - \text{tRNA}_{\text{un}}$). Ist dies der Fall, kann ein RNAII-Molekül mit gebundener unbeladener tRNA an die pDNA binden und einen Komplex aus Plasmid, RNAII und unbeladener tRNA bilden ($\text{pDNA} - \text{RNAII} - \text{tRNA}_{\text{un}}$), so dass die Replikation weiterhin erfolgen kann. In dem Modell von Brendel & Perelson (1993) wurde dieses dritte Szenario (Abb. 3.5 C), welches die Plasmidreplikationskontrolle durch die unbeladenen tRNA-Moleküle beschreibt, nicht berücksichtigt.

Der Unterschied zwischen einem *low copy* und einem *high copy* Plasmid wird in dem ersten Szenario deutlich: Der gebildete Komplex aus Plasmid, RNAII und RNAI ist zu diesem Zeitpunkt noch instabil, was bedeutet, dass das RNAI-Molekül mit einer gewissen Rate wieder dissoziieren kann. Im Fall eines *low copy* Plasmids kann dieser instabile Komplex durch das Rom-Protein stabilisiert werden. Wie in dem Modell von Brendel & Perelson (1993) vorgeschlagen, wird auch in dem Modell dieser Arbeit angenommen, dass dabei zunächst ein transienter Komplex entsteht ($\text{pDNA} - \text{RNAII} - \text{RNAI} - \text{Rom}$), der anschließend in einen stabilen Komplex aus Plasmid, RNAII und RNAI ($\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{stable}}$) überführt wird (Brendel & Perelson, 1993). In Anwesenheit des Rom-Proteins befindet sich daher eine größere Anzahl stabiler pDNA-RNAII-RNAI-Komplexe in der Zelle, wodurch die Plasmidreplikation „gebremst“ wird. Im Fall eines *high copy* Plasmids findet die Stabilisierung durch das Rom-Protein nicht statt, es besteht jedoch die Möglichkeit, dass der instabile Komplex mit einer geringen Rate ohne Einfluss des Rom-Proteins in einen stabilen Komplex konvertiert, so dass auch ohne Anwesenheit des Rom-Proteins eine geringe Anzahl stabiler pDNA-RNAII-RNAI-Komplexe in der Zelle vorkommen. Diese Reaktion existiert sowohl bei einem *high copy*, als auch bei einem *low copy* Plasmid.

Zusätzlich zu den in Abbildung 8.2 dargestellten Reaktionen beinhaltet das strukturelle Modell die Synthese- und Degradierungsreaktionen von RNAI, RNAII, unbeladener tRNA und dem Rom-Protein sowie eine Reaktion zur Beschreibung der Interaktion zwischen freien RNAI- und RNAII-Molekülen zur Ausbildung eines RNAI-RNAII-Komplexes (Abbildung 8.3).

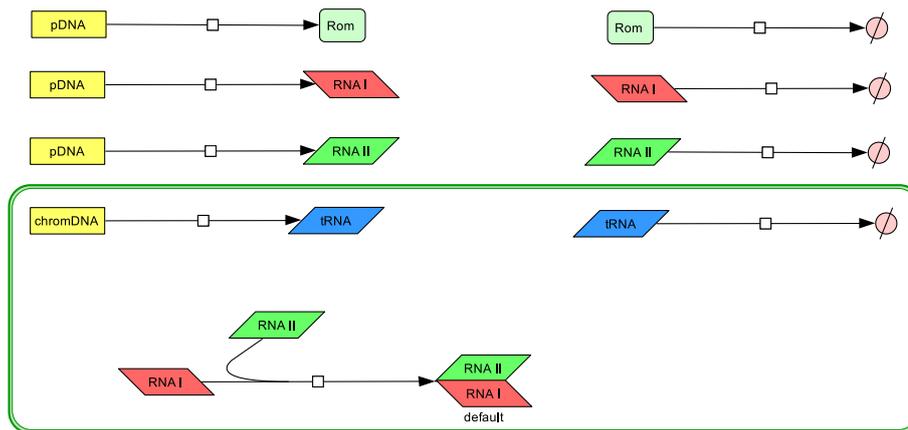


Abbildung 8.3: Zusätzliche Reaktionen des strukturellen Modells der *ColE1*-Plasmidreplikationskontrolle. Neben den in Abbildung 8.2 dargestellten Reaktionen beinhaltet das Modell noch Synthese- und Degradationsreaktionen von RNAI, RNAII, unbeladener tRNA-Moleküle und dem Rom-Protein sowie eine Reaktion zur Beschreibung der Interaktion zwischen freien RNAI- und RNAII-Molekülen zur Ausbildung eines RNAI-RNAII-Komplexes. Dabei wurden die in grün eingerahmten Reaktionen dem Modell von Brendel & Perelson (1993) im Rahmen dieser Arbeit hinzugefügt.

Zusammengefasst ergänzt das in dieser Arbeit erstellte strukturelle Modell das von Brendel & Perelson (1993) vorgeschlagene Reaktionsnetzwerk um die in Abbildung 8.3 grün eingerahmten Reaktionen und die in Abbildung 8.2 außerhalb des blauen Rahmens dargestellten Reaktionen.

8.2 ERFASSUNG DER EXPERIMENTELLEN DATEN

Die Erfassung der experimentellen Daten wurde zu verschiedenen Zeitpunkten durchgeführt und umfasste die Messung der RNAI- und RNAII-Konzentration sowie die Bestimmung des Plasmidgehalts und der Wachstumsrate. Dazu wurden die Stämme DH5 α -pSUP 201-3 (*low copy* plasmidtragend) und DH5 α -pCMV-lacZ (*high copy* plasmidtragend) über einen Zeitraum von ca. 30 Stunden kultiviert. Als Erntezeitpunkte wurden drei Zeitpunkte aus der logarithmischen Phase gewählt, weil bei einem Gespräch mit Martin Schleaf (Plasmid Factory) herauskam, dass die Plasmidproduktion in den Zellen im Verlauf der logarithmischen Phase zunimmt. Um diese Zunahme experimentell beschreiben zu können, wurden ebendiese Erntezeitpunkte gewählt. Die Berechnung der Wachstumsraten erfolgte über die Bestimmung der Steigung der Wuchskurve an den entsprechenden Erntezeitpunkten. Die RNA-Konzentrationsbestimmung wurde unter Anwendung von qRT-PCR durchgeführt und der Plasmidgehalt ist durch spektrophotometrische Messung im NanoDrop-Gerät und anschließender Agarosegelelektrophorese

quantifiziert worden. Zusätzlich zu den bereits erwähnten Messungen wurde für jeden Erntezeitpunkt die Zellzahl über das Ausplattieren einer Verdünnungsreihe ermittelt. Dies war notwendig, um die Anzahl der RNA-Moleküle bzw. der Plasmidkopienzahl pro Zelle bestimmen zu können.

8.2.1 Bestimmung der Wachstumsraten

8.2.1.1 DH5 α -pSUP 201-3

Für die Bestimmung der Wachstumsraten wurde zunächst der Wachstumsverlauf der Kultur durch die Messung der optischen Dichte bei 600 nm (OD_{600}) zu verschiedenen Zeitpunkten verfolgt. Es wurden vier biologische Replikate gemessen und der zugehörige Mittelwert (MW) berechnet. Die experimentellen Daten zur Messung der optischen Dichte sind in Tabelle 8.1 dargestellt und die Zeitpunkte der Zellernte durch ein * Punkt gekennzeichnet. Anschließend wurde mit Hilfe der MATLAB-Funktion „polyfit“ eine Ausgleichsfunktion gesucht, die möglichst nah an den OD_{600} -Messpunkten liegt. Die Funktion „polyfit“ verwendet dabei die „Methode der kleinsten Quadrate“, um die Kurvenparameter so zu bestimmen, dass die Summe der quadratischen Abweichungen zwischen den Messpunkten und der Kurve minimal ist. Das auf diese Weise bestimmte Polynoms 4-ten Grades ist eine mathematische Beschreibung für das Wachstumsverhalten des Stammes DH5 α -pSUP 201-3:

$$y(x) = 2.8108 \cdot 10^{-5} \cdot x^4 - 0.0019326 \cdot x^3 + 0.035299 \cdot x^2 + 0.050843 \cdot x - 0.0084259 \quad (8.1)$$

In der Abbildung 8.4 sind die einzelnen Messpunkte zusammen mit der polynomialen Ausgleichsfunktion und den Erntezeitpunkten zu sehen. Zusätzlich sind in der Abbildung die kleinsten und die größten gemessene OD_{600} -Werte zu jedem Zeitpunkt mit einer Linie verbunden dargestellt.

Tabelle 8.1: Verfolgung des Wachstumsverlaufs von DH5 α -pSUP 201-3 durch Messung der OD₆₀₀. Die Zeitpunkte der Zellernte sind durch ein * gekennzeichnet.

Zeit [min]	Zeit [h]	MW (OD)	$\ln(MW/MW_0)$	OD _{min}	OD _{max}
0	0	0.165 ± 0.001	0	0.164	0.165
75	1.25	0.181 ± 0.004	0.094067858	0.177	0.187
150	2.5	0.227 ± 0.017	0.32162156	0.212	0.246
250	4.17	0.322 ± 0.016	0.66856837	0.301	0.337
315*	5.25	0.433 ± 0.035	0.966885755	0.399	0.469
420	7	0.733 ± 0.071	1.492716529	0.659	0.799
495*	8.25	1.011 ± 0.058	1.814018735	0.943	1.068
585	9.75	1.868 ± 0.151	2.427926744	1.760	2.090
630*	10.5	2.065 ± 0.226	2.528456332	1.860	2.280
720	12	2.918 ± 0.281	2.874053191	2.520	3.180
780	13	3.680 ± 0.108	3.106238858	3.590	3.820
840	14	4.893 ± 0.239	3.391029526	4.670	5.190
900	15	6.250 ± 0.631	3.63670725	5.640	7.130
960	16	7.220 ± 0.526	3.780181059	6.670	7.930
1590	26.5	9.313 ± 0.726	4.034773171	8.360	9.920
1650	27.5	8.563 ± 0.328	3.950718309	8.140	8.910
1725	28.75	8.268 ± 1.025	3.915658271	7.300	9.560

Vergleicht man die Kurve der Ausgleichsfunktion mit den Messwerten, stellt man fest, dass sie das Wuchsverhalten der DH5 α -pSUP 201-3-Kultur in geeigneter Weise repräsentiert. Die Kurve verläuft sehr nah an den einzelnen Messwerten innerhalb des Minima-Maxima-Bereichs, lediglich in der stationären Phase bricht sie aus. Dies stellt jedoch für diese Arbeit kein Problem dar, da sich in diesem Bereich keine Erntezeitpunkte befinden und er somit für die Berechnung der dazugehörigen Wachstumsraten nicht relevant ist.

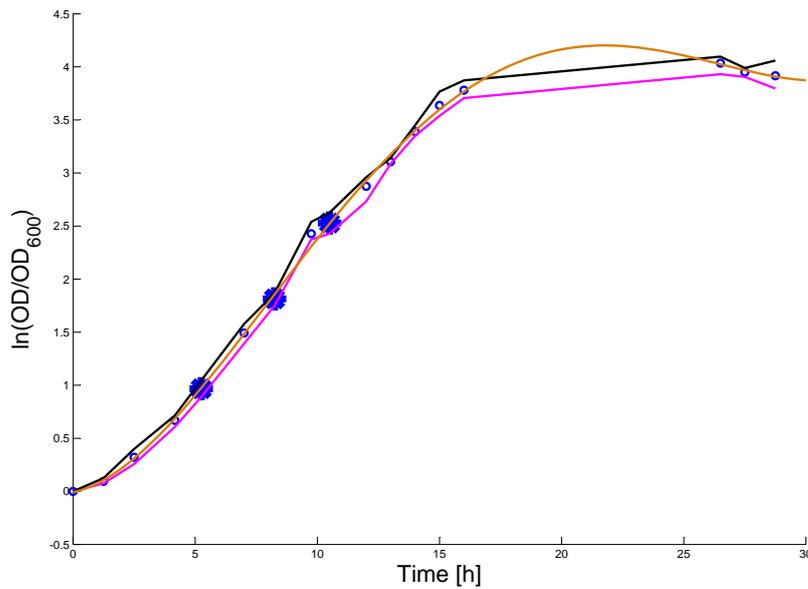


Abbildung 8.4: Wachstumskurve von DH5 α -pSUP 201-3.

Die einzelnen Datenpunkte zu den OD₆₀₀-Messungen sind durch blaue Kreise gekennzeichnet und die zugehörige polynomiale Ausgleichsfunktion ist in braun dargestellt. Die Zeitpunkte der Zellernte sind durch einen dicken Punkt markiert. Der Minima-Maxima-Bereich, innerhalb dem die Messwerte liegen müssen, ist durch die pinkfarbene (kleinste gemessene OD₆₀₀-Werte) und die schwarze Kurve (größte gemessene OD₆₀₀-Werte) veranschaulicht.

Die Wachstumsraten μ wurden mit Hilfe der ersten Ableitung der polynomialen Ausgleichsfunktion,

$$y'(x) = \mu = \frac{8296017318781249 \cdot x^3}{73786976294838206464} - \frac{13368816598819155 \cdot x^2}{2305843009213693952} + \frac{1271780505972409 \cdot x}{18014398509481984} + \frac{1831812126835185}{36028797018963968} \quad (8.2)$$

bestimmt, wobei für x der jeweilige Erntezeitpunkt eingesetzt wurde. Des Weiteren wurde die zugehörige Generationszeit g mittels

$$g = \frac{\ln(2)}{\mu} \quad (8.3)$$

berechnet. Die Formel zur Berechnung der Generationszeit kann folgendermaßen nach Fuchs & Schlegel (2007) hergeleitet werden:

Während des exponentiellen Wachstums gilt

$$\frac{d[X]}{dt} = \mu \cdot X \quad (8.4)$$

wobei X die Zellzahl zum Zeitpunkt t mit der Wachstumsrate μ beschreibt. Durch Integration erhält man

$$X = X_0 \cdot e^{\mu \cdot t} \quad (8.5)$$

X_0 beschreibt die Zellzahl zum Zeitpunkt t_0 . Eine Verdopplung von X_0 findet mit einer Generationszeit g statt. Es gilt

$$2 \cdot X_0 = X_0 \cdot e^{\mu \cdot g} \quad (8.6)$$

$$2 = e^{\mu \cdot g} \quad (8.7)$$

$$\ln(2) = \mu \cdot g \quad (8.8)$$

Nach Auflösung nach der Generationszeit g erhält man

$$g = \frac{\ln(2)}{\mu} \quad (8.9)$$

Die anhand von Gleichung 8.2 und 8.3 ermittelten Wachstumsraten mit den zugehörigen Generationszeiten sind in Tabelle 8.2 dargestellt. Sie werden in den später beschriebenen Simulationen verwendet, indem sie in das dynamische Modell eingesetzt werden, wenn die Simulation zu dem jeweiligen Erntezeitpunkt beginnt. Die Simulationszeit entspricht dabei der Generationszeit und die Wachstumsrate bleibt die gesamte Simulation hindurch konstant.

Tabelle 8.2: Anhand von Gleichung 8.2 und 8.3 ermittelte Wachstumsraten mit den zugehörigen Generationszeiten für DH5 α -pSUP 201-3. Diese Werte werden in den im folgenden Text beschriebenen Simulationen verwendet, indem sie in das Modell eingesetzt werden, wenn die Simulation zu dem jeweiligen Erntezeitpunkt beginnt. Dabei entspricht die Simulationszeit der Generationszeit und die Wachstumsrate bleibt die gesamte Simulation hindurch unverändert.

	Erntezeitpunkt [h]	Wachstumsrate [h^{-1}]	Generationszeit [min]
T ₁	5.25	0.278	150
T ₂	8.25	0.302	138
T ₃	10.5	0.283	147

8.2.1.2 DH5 α -pCMV-lacZ

Um den Wachstumsverlauf der DH5 α -pCMV-lacZ Kultur verfolgen zu können, wurde an verschiedenen Zeitpunkten eine Messung der optischen Dichte bei 600 nm durchgeführt. Es sind dabei drei biologische Replikate vermessen und der zugehörige Mittelwert

(MW) berechnet worden. Die experimentellen Daten zur Messung der OD_{600} sind in Tabelle 8.3 dargestellt, wobei die Zeitpunkte der Zellernte durch ein „*“ gekennzeichnet sind.

Tabelle 8.3: Verfolgung des Wachstumsverlaufs von DH5 α -placZ durch Messung der OD_{600} . Die Zeitpunkte der Zellernte sind durch ein * gekennzeichnet.

Zeit [min]	Zeit [h]	MW (OD)	$\ln(MW/MW_0)$	OD_{\min}	OD_{\max}
0	0	0.164 ± 0.004	0	0.159	0.167
60	1	0.180 ± 0.004	0.091236854	0.177	0.184
120	2	0.217 ± 0.005	0.278493646	0.213	0.222
200	3.33	0.293 ± 0.004	0.581443191	0.289	0.297
255	4.25	0.382 ± 0.008	0.845554181	0.377	0.391
315	5.25	0.485 ± 0.008	1.084282463	0.476	0.489
355	5.92	0.598 ± 0.010	1.294281584	0.591	0.61
425	7.08	0.832 ± 0.018	1.623565292	0.815	0.851
490*	8.17	1.125 ± 0.043	1.925968139	1.095	1.174
545	9.08	1.823 ± 0.025	2.408555179	1.8	1.85
625	10.42	2.577 ± 0.107	2.754385425	2.51	2.7
690*	11.5	3.023 ± 0.396	2.914248827	2.78	3.48
750	12.5	4.280 ± 0.560	3.261841861	3.67	4.77
810	13.5	5.353 ± 0.592	3.485608271	4.67	5.72
880	14.67	5.967 ± 0.528	3.594077275	5.59	6.57
930	15.5	7.920 ± 0.979	3.877280057	7.07	8.99
990	16.5	7.863 ± 0.170	3.870099456	7.7	8.04
1050*	17.5	8.837 ± 0.843	3.986798583	8.12	9.6
1110	18.5	8.750 ± 1.371	3.976942552	7.9	10.41
1590	26.5	8.593 ± 0.282	3.95887556	8.43	8.96
1675	27.92	8.593 ± 0.305	3.95887556	8.29	8.9
1830	30.5	8.813 ± 0.677	3.984154577	8.3	9.58
1890	31.5	8.410 ± 0.166	3.937310325	8.29	8.6

Die Berechnung der Wachstumsraten für DH5 α -pCMV-lacZ erfolgte analog zum Vorgehen für DH5 α -pSUP 201-3. Das mit Hilfe der „polyfit“-Funktion bestimmte Parameterset aus Koeffizienten wurde in die allgemeine Gleichung eines Polynoms 4-ten Grades eingesetzt, um das Wachstumsverhalten des Stammes DH5 α -pCMV-lacZ mathematisch zu beschreiben, wobei folgende Ausgleichsfunktion entstand

$$y(x) = 3.0143 \cdot 10^{-5} \cdot x^4 - 0.0019533 \cdot x^3 + 0.032587 \cdot x^2 + 0.097665 \cdot x - 0.029988 \quad (8.10)$$

Die Kurve der Ausgleichsfunktion zusammen mit den einzelnen Messpunkten sind in Abbildung 8.5 zu sehen. Die Zeitpunkte der Zellernte sind, wie in Abbildung 8.4, durch einen dicken Punkt markiert. Des Weiteren sind die kleinsten und die größten gemessene OD₆₀₀-Werte zu jedem Zeitpunkt mit einer Linie verbunden dargestellt, um den Minima-Maxima-Bereich zu kennzeichnen.

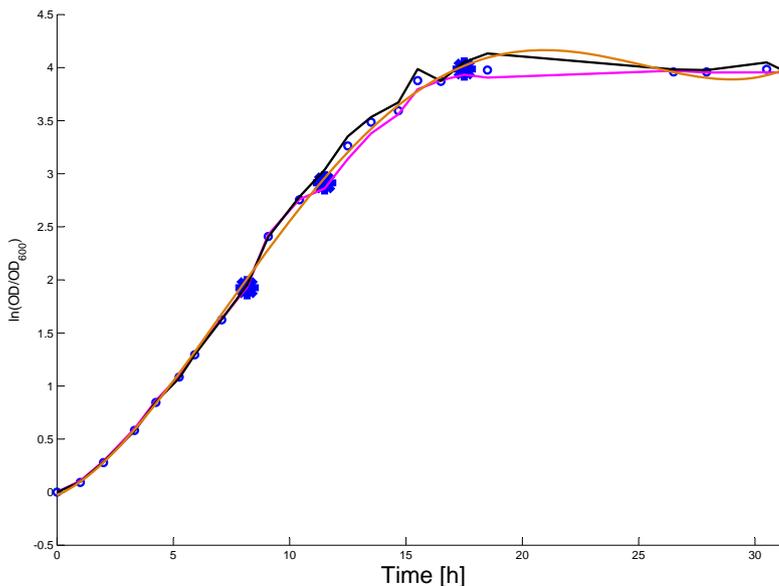


Abbildung 8.5: Wachstumskurve von DH5 α -pCMV-lacZ.

Die einzelnen Datenpunkte zu den OD₆₀₀-Messungen sind durch blaue Kreise gekennzeichnet und die zugehörige polynomiale Ausgleichsfunktion ist in braun dargestellt. Die Zeitpunkte der Zellernte sind durch einen dicken Punkt markiert. Der Minima-Maxima-Bereich, innerhalb dem die Messwerte liegen müssen, ist durch die pinkfarbene (kleinste gemessene OD₆₀₀-Werte) und die schwarze Kurve (größte gemessene OD₆₀₀-Werte) veranschaulicht.

Betrachtet man die Kurve der Ausgleichsfunktion und die einzelnen Messwerte, stellt man fest, dass sie das Wuchsverhalten der DH5 α -pCMV-lacZ-Kultur in geeigneter Weise darstellt. Die Kurve verläuft sehr nah an den einzelnen Messwerten bzw. innerhalb

des Minima-Maxima-Bereichs und bricht lediglich in der stationären Phase leicht aus. Wie für DH5 α -pSUP 201-3 bereits beschrieben, stellt dieses Verhalten für diese Arbeit kein Problem dar, da sich in diesem Bereich keine Erntezeitpunkte befinden.

Die Wachstumsraten μ wurden, unter Verwendung der ersten Ableitung der Ausgleichsfunktion (Gleichung 8.11), für jeden Erntezeitpunkt bestimmt und sind zusammen mit ihren zugehörigen Generationszeiten in der Tabelle 8.4 aufgeführt.

$$y'(x) = \mu = \frac{556040206613827 \cdot x^3}{4611686018427387904} - \frac{3378002362422831 \cdot x^2}{576460752303423488} + \frac{4696281633827915 \cdot x}{72057594037927936} + \frac{879688115214279}{9007199254740992} \quad (8.11)$$

Wie bei dem *low copy* plasmidtragenden Stamm DH5 α -pSUP 201-3 werden die Wachstumsraten mit ihren zugehörigen Generationszeiten in den später beschriebenen Simulationen verwendet, indem sie in das dynamische Modell eingesetzt werden, wenn die Simulation zu dem jeweiligen Erntezeitpunkt beginnt. Die Simulationszeit entspricht dabei der Generationszeit und die Wachstumsrate bleibt die gesamte Simulation hindurch konstant.

Tabelle 8.4: Anhand von Gleichung 8.11 und 8.3 ermittelte Wachstumsraten mit den zugehörigen Generationszeiten für DH5 α -pCMV-lacZ. Diese Werte werden in den, im folgenden Text beschriebenen Simulationen verwendet, indem sie in das Modell eingesetzt werden, wenn die Simulation zu dem jeweiligen Erntezeitpunkt beginnt. Dabei entspricht die Simulationszeit der Generationszeit und die Wachstumsrate bleibt die gesamte Simulation hindurch unverändert.

	Erntezeitpunkt [h]	Wachstumsrate [h ⁻¹]	Generationszeit [min]
T ₁	8.17	0.305	136
T ₂	11.5	0.256	162
T ₃	17.5	0.09	347

8.2.2 Messung der intrazellulären RNAI- und RNAII-Konzentrationen mittels qRT-PCR

Für die absolute Quantifizierung der RNAI- und RNAII-Konzentrationen pro Zelle war es zunächst notwendig, die genauen Zellzahlen für jeden Erntezeitpunkt zu bestimmen.

Dies erfolgte durch das Ausplattieren von Verdünnungsreihen für DH5 α -pSUP 201-3 und DH5 α -pCMV-lacZ. Die ermittelten Zellzahlen sind in Tabelle 8.5 zu finden.

Tabelle 8.5: Ermittelte Zellzahlen von DH5 α -pSUP 201-3 und DH5 α -pCMV-lacZ für die Erntezeitpunkte T₁ - T₃

	T ₁	T ₂	T ₃
DH5 α -pSUP 201-3	$1.3 \cdot 10^8$	$3.92 \cdot 10^8$	$1.19 \cdot 10^9$
DH5 α -pCMV-lacZ	$3.91 \cdot 10^7$	$2.35 \cdot 10^7$	$3.72 \cdot 10^7$

Der nächste Schritt bestand aus der Herstellung von Standardlösungen mit bekannten RNAI- und RNAII-Konzentrationen, um die notwendigen Eichgeraden erstellen zu können. Dies erfolgte durch PCR-Amplifikation des *rnaI*- und *rnaII*-Gens mit Primern, die eine T7-RNA-Polymerasepromoterregion als Extension aufweisen. Nach anschließender T7-RNA-Polymerase-Reaktion und DNaseI-Verdau konnte ein RNAI-Standard mit einer Konzentration von $c(\text{RNAI Standard}) = 1187,56 \text{ ng}/\mu\text{l}$ und ein RNAII-Standard mit $c(\text{RNAII Standard}) = 1378,86 \text{ ng}/\mu\text{l}$ hergestellt werden. Diese Standardlösungen wurden in verschiedenen Verdünnungsstufen mittels qRT-PCR vermessen. Im Anschluss daran sind die sogenannte CT-Werte (von engl. *cycle threshold*) bestimmt worden. Der CT-Wert beschreibt den Zyklus in der Real Time-PCR an dem die Fluoreszenzkurve in die exponentielle Phase eintritt. Die vermessenen Verdünnungsstufen und die dazu ermittelten CT-Werte für RNAI und RNAII sind in den Tabellen 8.6 und 8.7 aufgeführt.

Tabelle 8.6: Ermittelte CT-Werte der qRT-PCR-Messungen von verschiedenen Verdünnungsstufen der RNAI-Standardlösung

Verdünnungsstufen	Messung 1	Messung 2	Messung 3	Mittelwert
unverdünnt	n.d.	n.d.	n.d.	-
1:2	n.d.	8.45	n.d.	-
1: 2·10	10.23	11.47	8.86	10.19 ± 1.31
1:2·10 ²	13.69	14.44	9.42	12.52 ± 2.71
1:2·10 ³	17.94	17.95	11.47	15.79 ± 3.74
1:2·10 ⁴	20.54	21.5	15.52	19.19 ± 3.21
1:2·10 ⁵	24.92	25.1	18.33	22.78 ± 3.86
1:2·10 ⁶	27.97	28.39	19.71	25.36 ± 4.89

Tabelle 8.7: Ermittelte CT-Werte der qRT-PCR-Messungen von verschiedenen Verdünnungsstufen der RNAII-Standardlösung

Verdünnungsstufen	Messung 1	Messung 2	Messung 3	Mittelwert
unverdünnt	n.d.	n.d.	n.d.	-
1:2	n.d.	n.d.	n.d.	-
1:2·10	n.d.	8.7	8.6	8.65±0.07
1:2·10 ²	12.76	11.9	9.79	11.48±1.53
1:2·10 ³	16.47	14.43	13.13	14.68±1.68
1:2·10 ⁴	19.92	18.54	17.65	18.7±1.14
1:2·10 ⁵	24.44	23.05	21.12	22.87±1.67
1:2·10 ⁶	29.54	27.07	24.27	26.96±2.64
1:2·10 ⁷	31.68	28.54	28.4	29.54±1.85

Basierend auf den in den Tabellen 8.6 und 8.7 aufgeführten Mittelwerten der CT-Werte wurden die jeweiligen Eichreihen für RNAI und RNAII erstellt. Dazu wurde zunächst für jede Verdünnungsstufe die Molekülanzahl in 1000 berechnet. Anschließend wurde der Zehnerlogarithmus von den zuvor ermittelten Molekülanzahlen in 1000 gegen die gemessenen CT-Werte der Verdünnungsstufen aufgetragen und unter Verwendung der Excel-Software eine geeignete Regressionsgerade hineingelegt. Die Gleichung der Regressionsgerade für die RNAI-Eichreihe lautet

$$y(x) = -3.1443x + 38.584 \quad (8.12)$$

mit dem Bestimmtheitsmaß $R^2 = 0.9965$. Eine graphische Darstellung der Ausgleichsgerade zusammen mit den Rohdaten befindet sich in Abbildung 8.6.

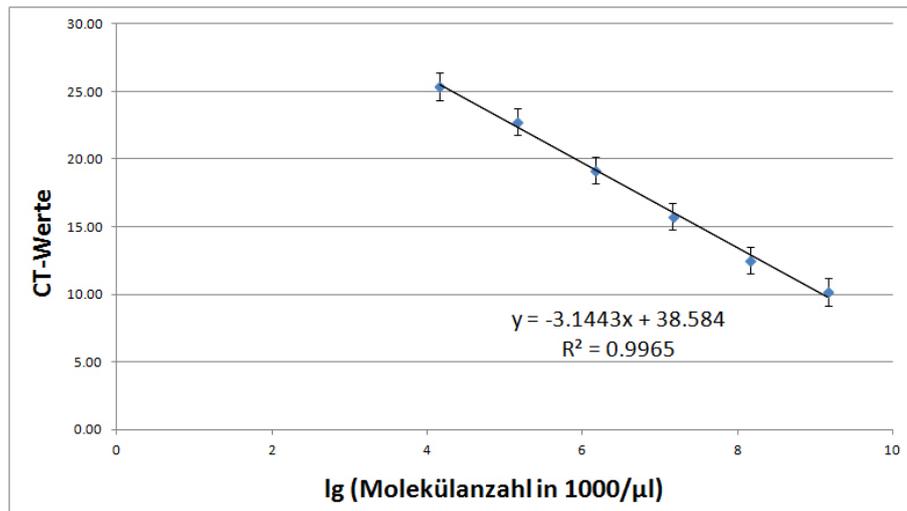


Abbildung 8.6: Eichgerade zur Bestimmung der RNAI-Konzentration

Die Rohdaten werden optisch durch die Gerade gut repräsentiert, was durch das Bestimmtheitsmaß von $R^2 \sim 1$ bestätigt wird.

Für die RNAII-Eichreihe verhält es sich ähnlich: Die Gerade bietet eine gute Annäherung an die Rohdaten, da das Bestimmtheitsmaß auch hier einen Wert von $R^2 \sim 1$ besitzt. Die Gleichung der Regressionsgerade für die RNAII-Eichreihe lautet

$$y(x) = -3.6363x + 39.328 \quad (8.13)$$

mit dem Bestimmtheitsmaß $R^2 = 0.9957$. Die graphische Darstellung der Ausgleichsgerade befindet sich zusammen mit den Rohdaten in Abbildung 8.7.

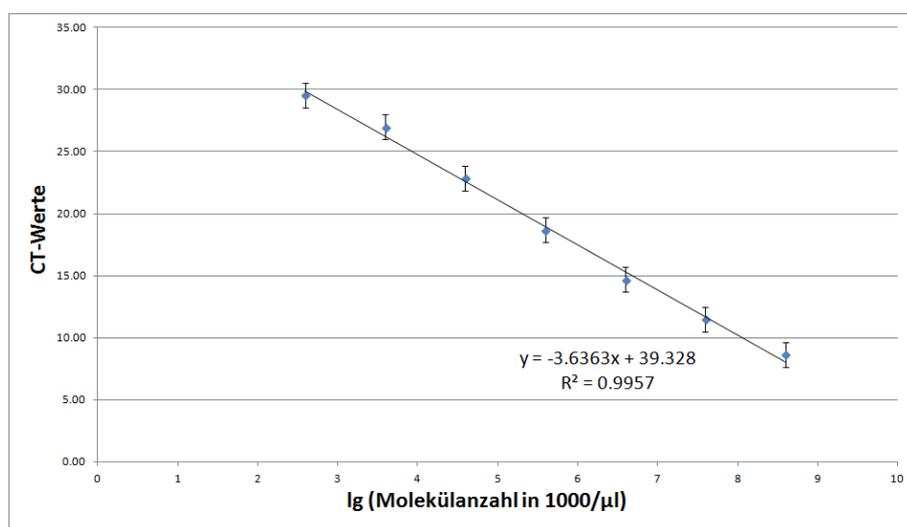


Abbildung 8.7: Eichgerade zur Bestimmung der RNAII-Konzentration

Mit Hilfe der beiden Regressionsgeraden ist es nun möglich die absoluten intrazellulären RNAI- und RNAII-Konzentrationen für DH5 α -pSUP 201-3 und DH5 α -pCMV-lacZ zu bestimmen.

8.2.2.1 Messung der RNAI- und RNAII-Konzentrationen von DH5 α -pSUP 201-3

Die Messung der RNAI- und RNAII-Konzentrationen von DH5 α -pSUP 201-3 erfolgte für jeweils drei biologische Replikate pro Erntezeitpunkt. Es wurde zunächst die Gesamt-RNA der Proben isoliert und anschließend revers transkribiert. Im nächsten Schritt erfolgte die Durchführung der qRT-PCR. Anhand der dabei gemessenen CT-Werte wurde der Mittelwert für jeden Erntezeitpunkt berechnet. Mit Hilfe dieser CT-Mittelwerte und den zuvor erstellten Eichgeraden konnten die RNAI- und RNAII-Konzentrationen bestimmt werden. Die ermittelten RNAI- und RNAII-Konzentrationen zu jeden Erntezeitpunkt sind in den Tabellen 8.8 und 8.9 aufgeführt.

Tabelle 8.8: Berechnete RNAI-Konzentrationen anhand der ermittelten CT-Werte für DH5 α -pSUP 201-3 zu den drei Erntezeitpunkten

Zeitpunkt	MW (CT-Werte)	Konzentration [mol/l]	Konzentration [Moleküle/Zelle]
T ₁	22.37 \pm 0.31	8.92 \cdot 10 ⁻⁸ \pm 1.89 \cdot 10 ⁻⁸	34 \pm 7
T ₂	21.78 \pm 0.39	4.61 \cdot 10 ⁻⁸ \pm 1.29 \cdot 10 ⁻⁸	17 \pm 5
T ₃	21.78 \pm 0.16	1.58 \cdot 10 ⁻⁸ \pm 1.84 \cdot 10 ⁻⁹	6 \pm 0.7

Tabelle 8.9: Berechnete RNAII-Konzentrationen anhand der ermittelten CT-Werte für DH5 α -pSUP 201-3 zu den drei Erntezeitpunkten

Zeitpunkt	MW (CT-Werte)	Konzentration [mol/l]	Konzentration [Moleküle/Zelle]
T ₁	20.8 \pm 0.55	7.77 \cdot 10 ⁻⁹ \pm 2.28 \cdot 10 ⁻⁹	3 \pm 0.9
T ₂	20.2 \pm 0.13	3.69 \cdot 10 ⁻⁹ \pm 1.04 \cdot 10 ⁻⁹	1 \pm 0.4
T ₃	19.89 \pm 0.39	1.16 \cdot 10 ⁻⁹ \pm 3.48 \cdot 10 ⁻¹⁰	0.44 \pm 0.1

Es fällt auf, dass die Konzentrationen der beiden RNA-Transkripte mit voranschreitendem Wachstum abnehmen und es zu jedem Zeitpunkt mehr freie RNAI- als RNAII-

Moleküle in der Zelle gibt. Die größte RNAI- und RNAII-Konzentration mit 34 ± 7 RNAI-Moleküle/Zelle und 3 ± 0.9 RNAII-Moleküle liegt zum Zeitpunkt T_1 vor. Zum letzten Erntezeitpunkt (T_3) misst man nur noch 6 ± 0.7 RNAI-Moleküle/Zelle und $0,44 \pm 0.1$ RNAII-Moleküle/Zelle.

Betrachtet man dabei das Verhältnis von RNAI zu RNAII, so stellt man fest, dass es zum Zeitpunkt T_1 am kleinsten ist und für die anderen beiden Zeitpunkte stabil bei 17 liegt. Insgesamt liegen die Werte jedoch nicht weit auseinander (Tabelle 8.10).

Tabelle 8.10: RNAI/RNAII-Verhältnisse von DH5 α -pSUP 201-3 für die Erntezeitpunkte T_1 - T_3

Stamm/Zeitpunkt	T_1	T_2	T_3
DH5 α -pSUP 201-3	11	17	17

8.2.2.2 Messung der RNAI- und RNAII-Konzentrationen von DH5 α -pCMV-lacZ

Für die Bestimmung der RNAI- und RNAII-Konzentrationen von DH5 α -pCMV-lacZ wurden ebenfalls 3 biologische Replikate vermessen und sie erfolgte analog zu dem Vorgehen für den Stamm DH5 α -pSUP 201-3. Die aus dem Mittelwert der gemessenen CT-Werte errechneten RNAI- und RNAII-Konzentrationen für DH5 α -pCMV-lacZ sind in den Tabellen 8.11 und 8.12 dargestellt.

Tabelle 8.11: Berechnete RNAI-Konzentrationen anhand der ermittelten CT-Werte für DH5 α -pCMV-lacZ zu den drei Erntezeitpunkten

Zeitpunkt	MW (CT-Werte)	Konzentration [mol/l]	Konzentration [Moleküle/Zelle]
T_1	24.02 ± 0.49	$1.44 \cdot 10^{-6} \pm 1.37 \cdot 10^{-8}$	541 ± 51
T_2	23.96 ± 0.49	$2.89 \cdot 10^{-6} \pm 7.9 \cdot 10^{-7}$	1086 ± 298
T_3	22.77 ± 0.4	$9.17 \cdot 10^{-7} \pm 5.4 \cdot 10^{-7}$	345 ± 203

Tabelle 8.12: Berechnete RNAII-Konzentrationen anhand der ermittelten CT-Werte für DH5 α -pCMV-lacZ zu den drei Erntezeitpunkten

Zeitpunkt	MW (CT-Werte)	Konzentration [mol/l]	Konzentration [Moleküle/Zelle]
T ₁	23.05 \pm 0.17	5.77 \cdot 10 ⁻⁸ \pm 6.13 \cdot 10 ⁻⁹	22 \pm 2
T ₂	22.23 \pm 0.1	1.71 \cdot 10 ⁻⁷ \pm 1.09 \cdot 10 ⁻⁸	64 \pm 4
T ₃	21.26 \pm 0.22	2.0 \cdot 10 ⁻⁷ \pm 2.66 \cdot 10 ⁻⁸	75 \pm 10

Im Gegensatz zu dem *low copy* plasmidtragenden Stamm lassen sich hier Schwankungen der RNAI- und RNAII-Konzentrationen innerhalb der drei Erntezeitpunkte beobachten. Es befinden sich jedoch bei DH5 α -pCMV-lacZ zu jedem Zeitpunkt mehr freie RNAI- und RNAII-Moleküle pro Zelle, als bei DH5 α -pSUP 201-3. Die größte Anzahl RNAI-Moleküle mit 1086 \pm 298 RNAI-Moleküle/Zelle misst man zum Zeitpunkt T₂. Die Anzahl der RNAII-Moleküle ist mit 75 \pm 10 RNAII-Moleküle/Zelle zum Zeitpunkt T₃ am größten. Die geringste Menge gemessener RNAI-Moleküle beträgt 345 \pm 203 (T₃). Am wenigsten freie RNAII-Moleküle lassen sich mit 22 \pm 2 zum Zeitpunkt T₁ messen.

Betrachtet man das Verhältnis von RNAI zu RNAII, so stellt man fest, dass es mit voranschreitendem Wachstum abnimmt. Dabei ist der Bereich, über den sich die Werte verteilen, deutlich größer als bei DH5 α -pSUP 201-3 (Tabelle 8.13).

Tabelle 8.13: RNAI/RNAII-Verhältnisse von DH5 α -pCMV-lacZ für die Erntezeitpunkte T₁ - T₃

Stamm/Zeitpunkt	T ₁	T ₂	T ₃
DH5 α -pCMV-lacZ	25	17	5

8.2.3 Messung der intrazellulären Plasmidkonzentrationen

Die Bestimmung der absoluten Plasmidkonzentrationen der geernteten Zellen wurde vom Kooperationspartner dieses Promotionsprojektes, der Plasmid Factory (Bielefeld), durchgeführt. Die Messungen erfolgten mit drei biologischen Replikaten pro Erntezeitpunkt und Bakterienstamm. Zu Beginn wurde die Plasmid-DNA mittels einer Nucleo-

Bond PC20 Säule der Firma Macherey-Nagel aufgereinigt. Es wurde stets darauf geachtet, dass die Konzentration der aufzureinigenden Plasmid-DNA die Kapazität der Säule nicht überschreitet, da ansonsten die Gefahr bestünde, dass ein Teil der in der Zelle vorkommenden Plasmid-DNA verworfen wird. Nach der Plasmidisolierung wurde zunächst die DNA-Konzentration der Proben mit Hilfe des NanoDrop-Geräts bestimmt. Anschließend erfolgte die Auftrennung der Proben mittels Agarosegelelektrophorese, wobei die genomische DNA (gDNA) von der Plasmid-DNA getrennt wird. In der Abbildung 8.8A ist exemplarisch das Gelbild der Probe DH5 α -pSUP 201-3-A für die Erntezeitpunkte T₁ - T₃ gezeigt. Man sieht eine deutliche Auftrennung in gDNA, pDNA in *oc*-Form und pDNA in *ccc*-Form. Die genomische DNA befindet sich in den oberen beiden Banden, wobei die DNA-Bande in den Geltaschen nicht bei allen Proben zu sehen war. Im Anschluss an die gelelektrophoretische Auftrennung erfolgte die Bestimmung des gDNA:pDNA-Verhältnisses mit Hilfe der Software LabImage 1D, so dass aus der im NanoDrop-Gerät gemessenen DNA-Konzentration die entsprechende pDNA-Konzentration berechnet werden konnte. Die Ermittlung des gDNA:pDNA-Verhältnisses und somit des pDNA-Anteils ist am Beispiel der Probe DH5 α -pSUP 201-3-A in der Abbildung 8.8B dargestellt.

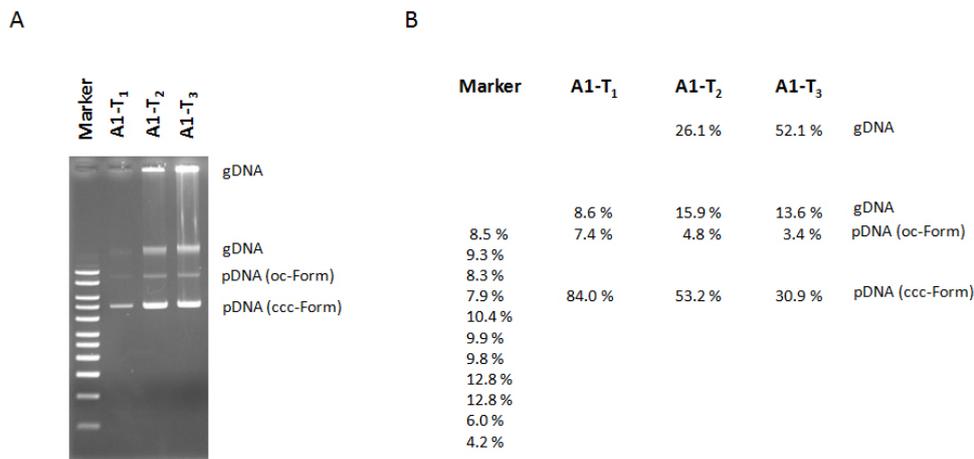


Abbildung 8.8: Ergebnisse der gelelektrophoretischen Auftrennung der isolierten DNA von DH5 α -pSUP 201-3 (Probe A) durch die Plasmid Factory für die Erntezeitpunkte T₁ - T₃. **A:** Auftrennung in die verschiedenen DNA-Formen. **B:** Mit Hilfe der Software LabImage 1D berechneter prozentualer Anteil der aufgetrennten DNA-Formen

Für die Probe DH5 α -pSUP 201-3-A zum Zeitpunkt T₁ wurde ein pDNA-Anteil von 91.4 % (7.4 % + 84 % = 91.4 %) ermittelt. Anhand der zuvor im NanoDrop-Gerät gemessenen DNA-Konzentration und den Zellzahlen zu jedem Erntezeitpunkt (Tab. 8.5) kann im nächsten Schritt die entsprechende pDNA-Konzentration pro Zelle

bestimmt werden, welche mit Hilfe des Molekulargewichts des Plasmids pSUP 201-3 in die dazu zugehörige Plasmidkopienzahl umgerechnet werden kann. Auf diese Weise wurden für alle DH5 α -pSUP 201-3- und DH5 α -pCMV-lacZ Proben die absoluten pDNA-Konzentrationen pro Zelle bestimmt.

In den Tabellen 8.14 und 8.15 sind die berechneten Plasmidkonzentrationen für DH5 α -pSUP 201-3 aufgeführt. Die Tabelle 8.14 zeigt die Plasmidkonzentrationen aller vermessenen DH5 α -pSUP 201-3-Proben, wobei die biologischen Replikate durch die Buchstaben „A“, „B“ und „C“ gekennzeichnet sind. Der zugehörige Erntezeitpunkt wird durch den Index beschrieben. Für den *low copy* plasmidtragenden Stamm wurde eine durchschnittliche Plasmidmenge von 46 bzw. 48 Plasmiden pro Zelle berechnet. Vergleicht man diese mittleren Kopienzahlen mit den Literaturangaben zu einem *low copy* Plasmid mit einem pBR322-Replikationsursprung (15-20 Plasmide/ Zelle) (Cooper, 2004), stellt man fest, dass sie ungefähr doppelt so hoch sind.

Tabelle 8.14: Messungen zur Berechnung der Plasmidkonzentrationen von der Firma Plasmid Factory für DH5 α -pSUP 201-3 an den Erntezeitpunkten T₁ - T₃

Probe	Plasmid-DNA [mol/l]	Plasmid-DNA [PCN]
A ₁	$4.12 \cdot 10^{-8}$	16
A ₂	$6.00 \cdot 10^{-8}$	23
A ₃	$1.17 \cdot 10^{-7}$	44
B ₁	$1.60 \cdot 10^{-7}$	60
B ₂	$1.02 \cdot 10^{-7}$	39
B ₃	$7.99 \cdot 10^{-8}$	30
C ₁	$1.62 \cdot 10^{-7}$	61
C ₂	$2.16 \cdot 10^{-7}$	81
C ₃	$1.71 \cdot 10^{-7}$	64

Tabelle 8.15: Berechnung des mittleren Plasmidgehalts pro Zelle für DH5 α -pSUP 201-3 (Mittelwert (MW) und Standardabweichung (SD))

Zeitpunkt	MW PCN	SD (PCN)	MW Plasmid-DNA [mol/l]	SD Plasmid-DNA [mol/l]
T ₁	46	26	$1.21 \cdot 10^{-7}$	$6.90 \cdot 10^{-8}$
T ₂	48	30	$1.26 \cdot 10^{-7}$	$8.09 \cdot 10^{-8}$
T ₃	46	17	$1.23 \cdot 10^{-7}$	$4.59 \cdot 10^{-8}$

In den Tabellen 8.16 und 8.17 sind die berechneten Plasmidkonzentrationen für DH5 α -pCMV-lacZ dargestellt. Die Tabelle 8.16 zeigt die Plasmidkonzentrationen aller vermessenen Proben, wobei die biologischen Replikate durch die Buchstaben „A“, „B“ und „C“ gekennzeichnet sind und der zugehörige Erntezeitpunkt durch den Index beschrieben wird. Betrachtet man die ermittelten Plasmidkopienzahlen für DH5 α -pCMV-lacZ fällt auf, dass die Werte sogar deutlich größer sind, als in der Literatur für einen Plasmid mit pUC-Replikationsursprung beschrieben. Nach Cooper (2004) besitzt ein pUC-Plasmid eine Anzahl von 500-700 Plasmide pro Zelle. Die in dieser Arbeit durchschnittlich gemessenen Kopienzahlen liegen in einem Bereich von 1514 bis 5806, wobei die Plasmidanzahl pro Zelle mit der Zeit zunimmt.

Tabelle 8.16: Messungen zur Berechnung der Plasmidkonzentrationen von der Firma Plasmid Factory für DH5 α -pCMV-lacZ an den Erntezeitpunkten T₁ - T₃

Probe	Plasmid-DNA [mol/l]	Plasmid-DNA [PCN]
A ₁	$1.46 \cdot 10^{-6}$	550
A ₂	$7.59 \cdot 10^{-6}$	2858
A ₃	$3.02 \cdot 10^{-5}$	11367
B ₁	$6.58 \cdot 10^{-6}$	2478
B ₂	$7.36 \cdot 10^{-6}$	2770
B ₃	$8.91 \cdot 10^{-6}$	3352
C ₁	n.d.	n.d.
C ₂	$4.20 \cdot 10^{-6}$	1580
C ₃	$7.17 \cdot 10^{-6}$	2697

Tabelle 8.17: Berechnung des mittleren Plasmidgehalts pro Zelle für DH5 α -pCMV-lacZ (Mittelwert (MW) und Standardabweichung (SD))

Zeitpunkt	MW PCN	SD (PCN)	MW Plasmid-DNA [mol/l]	SD Plasmid-DNA [mol/l]
T ₁	1514	1301	$4.02 \cdot 10^{-6}$	$3.62 \cdot 10^{-6}$
T ₂	2403	713	$6.38 \cdot 10^{-6}$	$1.20 \cdot 10^{-6}$
T ₃	5806	4828	$1.54 \cdot 10^{-5}$	$1.28 \cdot 10^{-5}$

Da die gemessenen Werte die Literaturwerte deutlich übersteigen, wurden zwei der DH5 α -pCMV-lacZ-Proben (DH5 α -pCMV-lacZ-A₂ und DH5 α -pCMV-lacZ-A₃) von der Firma CARPEGEN aus Münster unter Anwendung einer sondenbasierten RT-PCR vermessen. Mit Hilfe dieser Methode lässt sich der Anteil genomischer DNA bestimmen, wobei die ribosomalen Operons der gDNA quantifiziert werden. Basierend auf dieser Quantifizierung wird die errechnete gDNA-Menge in Relation zur photometrisch bestimmten Gesamt-DNA-Konzentration gesetzt und der pDNA-Anteil ermittelt (GesamtDNA - gDNA = pDNA). Diese Messungen dienen der Überprüfung, ob die Größenordnungen (1000er Bereich) der vermessenen Proben realistisch sind. Die Ergebnisse sind in der Tabelle 8.18 aufgeführt.

Tabelle 8.18: Messungen zur Berechnung der Plasmidkonzentrationen von der Firma CARPEGEN GmbH

Probe	PCN
DH5 α -pCMV-lacZ-A ₂	6960
DH5 α -pCMV-lacZ-A ₃	8453

Man sieht, dass die für die Probe DH5 α -pCMV-lacZ-A₂ gemessene Kopienzahl von 6960 ungefähr doppelt so hoch ist, wie die densitometrisch bestimmte Plasmidmenge (2858 Plasmide pro Zelle). Im Fall von DH5 α -pCMV-lacZ-A₃ ist sie mit 8453 Plasmiden pro Zelle etwas geringer als der zuvor gemessene Wert von 11367 Plasmiden pro Zelle. Zusammenfassend kann man jedoch festhalten, dass die mit Hilfe der sondenbasierten RT-PCR ermittelten Ergebnisse die Größenordnungen der zuvor gemessenen Plasmidkopienzahlen für DH5 α -pCMV-lacZ bestätigen.

Neben den sehr hohen Plasmidkonzentrationen pro Zelle fällt auf, dass die Plasmid-

mengen sowohl für DH5 α -pSUP 201-3, als auch für DH5 α -pCMV-lacZ innerhalb der biologischen Replikate stark schwanken. Diese Schwankungen spiegeln sich auch bei der Berechnung der Mittelwerte und den dazu gehörigen Standardabweichungen wider. Möglicherweise sind diese großen Unterschiede durch die Messmethode bedingt, da die densitometrische Bestimmung des gDNA:pDNA-Verhältnisses anhand der Gelbandenhelligkeit aufgrund von Sättigungseffekten fehlerbehaftet ist.

Um diese hohen Schwankungen zu untersuchen wurden, alle Proben erneut mittels einer Hochleistungsflüssigkeitschromatographie (HPLC von engl. *high performance liquid chromatography*) vermessen. Dazu wurde zunächst eine Verdünnungsreihe mit bekannten pDNA-Konzentrationen vermessen, um eine Eichgerade zu generieren, anhand derer die unbekanntes Plasmidmengen in den Proben bestimmt werden können. Nach Erstellung der Eichgerade wurden die Lysate der Proben mittels HPLC vermessen. Die Mittelwerte der gemessenen Plasmidkopienzahlen mit ihren dazu gehörigen Standardabweichungen sind in der Tabelle 8.19 aufgeführt.

Tabelle 8.19: Messungen zur Berechnung der Plasmidkonzentrationen mittels HPLC-Analytik für die Erntezeitpunkte T₁ - T₃

Stamm/Zeitpunkt	T ₁	T ₂	T ₃
DH5 α -pSUP 201-3	1249 \pm 60	478 \pm 11	188 \pm 3
DH5 α -pCMV-lacZ	4576 \pm 98	10423 \pm 998	10960 \pm 397

Man sieht, dass die Schwankungen innerhalb der biologischen Replikate geringer sind, als bei den vorigen Messungen. Es ist jedoch festzuhalten, dass die, unter Anwendung der HPLC ermittelten, durchschnittlichen Plasmidmengen deutlich größer sind, als die Kopienzahlen der densitometrischen Bestimmung und die der RT-PCR-Messung. Eine mögliche Erklärung ist, dass die in den Zellen enthaltenen Plasmidkonzentrationen zu niedrig sind, als dass der UV-Detektor sie detektieren kann. Aufgrund dieser, unter Anwendung der HPLC ermittelten, sowohl für DH5 α -pSUP 201-3, als auch für DH5 α -pCMV-lacZ, extrem hohen Plasmidkonzentrationen, wurden die Daten der densitometrischen Messung (trotz der hohen Schwankungen) für das dynamische Modell zur Simulation der ColE1-Plasmidreplikationskontrolle eingesetzt.

8.3 ERSTELLUNG DES DYNAMISCHEN MODELLS

Das in dieser Arbeit erstellte dynamische Modell der Plasmidreplikationskontrolle ist eine mathematische Beschreibung des zuvor vorgestellten strukturellen Modells (Abbildung 8.2). Die Kinetik jeder Reaktion wird dabei durch das Massenwirkungsgesetz beschrieben und die Änderungen der einzelnen Substratkonzentrationen werden durch gewöhnliche Differentialgleichungen (im Folgenden nur noch DGL genannt) dargestellt. Die allgemeine Form einer hier verwendeten DGL lautet

$$\frac{d[S]}{dt} = v_1 + \dots + v_n - \mu \cdot [S] \quad (8.14)$$

wobei $[S]$ die Substratkonzentration, v_i ($i = 1 \dots n$) die beteiligten Reaktionsgeschwindigkeiten, t die Zeit und μ die Wachstumsrate beschreibt.

In jeder DGL befindet sich ein Ausdruck, in dem die negative Wachstumsrate mit der jeweiligen Substratkonzentration multipliziert wird ($-\mu \cdot [S]$). Die Verwendung dieses Terms wurde von Brendel & Perelson (1993) vorgeschlagen, um Verdünnungseffekte in Folge von Zellwachstum zu berücksichtigen. Mit Verdünnungseffekt ist hier gemeint, dass die Verteilung der zellulären Komponenten nach der Zellteilung auf die Tochterzellen von der Wachstumsrate abhängig ist. Im Fall einer schnell wachsenden Zelle nimmt das Volumen des Cytoplasmas schneller zu, als die Konzentration der intrazellulären Komponenten. Das hat zur Folge, dass die Tochterzellen, im Vergleich zur Mutterzelle, nach der Zellteilung eine sehr viel geringere Konzentration an intrazellulären Komponenten aufweisen. Dieser Verdünnungseffekt zeigt sich auch bei einer langsam wachsenden Zelle, jedoch ist er hier nicht so ausgeprägt, da die Zelle mehr Zeit hat ihre intrazellulären Komponenten neu zu synthetisieren bevor sie auf zwei Tochterzellen aufgeteilt werden. Das bedeutet, die Stärke des Verdünnungseffektes ist abhängig von der Wachstumsrate. (Brendel & Perelson, 1993)

Basierend auf dem strukturellen Modell wurde unter Berücksichtigung des Massenwirkungsgesetzes und des Verdünnungsterms folgendes DGL-System aufgestellt:

$$\frac{d[\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{unstable}}]}{dt} = -v_1 + v_2 + v_3 - v_4 - v_{10} + v_{11} - \mu \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{unstable}}] \quad (8.15)$$

$$\frac{d[\text{Rom}]}{dt} = -v_1 + v_2 + v_{13} + v_{20} - v_{23} - \mu \cdot [\text{Rom}] \quad (8.16)$$

$$\frac{d[\text{pDNA} - \text{RNAII} - \text{RNAI} - \text{Rom}]}{dt} = v_1 - v_2 - v_{13} - \mu \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI} - \text{Rom}] \quad (8.17)$$

$$\frac{d[\text{pDNA} - \text{RNAII} - s]}{dt} = -v_3 + v_4 - v_5 + v_9 - v_{15} - \mu \cdot [\text{pDNA} - \text{RNAII} - s] \quad (8.18)$$

$$\frac{d[\text{RNAI}]}{dt} = -v_3 + v_4 + v_{18} - v_{21} - v_{25} - \mu \cdot [\text{RNAI}] \quad (8.19)$$

$$\frac{d[\text{pDNA} - \text{RNAII} - \text{lo}]}{dt} = v_5 - v_6 - v_8 + v_{16} - \mu \cdot [\text{pDNA} - \text{RNAII} - \text{lo}] \quad (8.20)$$

$$\frac{d[\text{pDNA} - \text{RNAII} - \text{Primer}]}{dt} = v_6 - v_7 - \mu \cdot [\text{pDNA} - \text{RNAII} - \text{Primer}] \quad (8.21)$$

$$\frac{d[\text{RNAII} - \text{lo}]}{dt} = v_7 + v_8 - \mu \cdot [\text{RNAII} - \text{lo}] \quad (8.22)$$

$$\frac{d[\text{pDNA}]}{dt} = 2 \cdot v_7 + v_8 - v_9 + v_{12} - v_{17} - \mu \cdot [\text{pDNA}] \quad (8.23)$$

$$\frac{d[\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{stable}}]}{dt} = v_{10} - v_{11} - v_{12} + v_{13} - \mu \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{stable}}] \quad (8.24)$$

$$\frac{d[\text{tRNA}_{\text{un}}]}{dt} = -v_{14} - v_{15} + v_{24} - v_{26} - \mu \cdot [\text{tRNA}_{\text{un}}] \quad (8.25)$$

$$\frac{d[\text{RNAII}]}{dt} = -v_9 - v_{14} + v_{19} - v_{22} - v_{25} - \mu \cdot [\text{RNAII}] \quad (8.26)$$

$$\frac{d[\text{tRNA}_{\text{un}} - \text{RNAII}]}{dt} = v_{14} - v_{17} - \mu \cdot [\text{tRNA}_{\text{un}} - \text{RNAII}] \quad (8.27)$$

$$\frac{d[\text{pDNA} - \text{RNAII} - \text{tRNA}_{\text{un}}]}{dt} = v_{15} - v_{16} + v_{17} - \mu \cdot [\text{pDNA} - \text{RNAII} - \text{tRNA}_{\text{un}}] \quad (8.28)$$

$$\frac{d[\text{RNAI} - \text{RNAII}]}{dt} = v_{25} - \mu \cdot [\text{RNAI} - \text{RNAII}] \quad (8.29)$$

Das DGL-System beinhaltet mit $v_i (i = 1 \dots 26)$ folgende Gleichungen, die die Geschwindigkeiten der einzelnen Reaktionen beschreiben:

Reaktion 1:

Bildung des transienten Komplexes aus Plasmid-gebundener RNAII, RNAI und Rom-Protein

$$v_1 = k_1 \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{unstable}}] \cdot [\text{Rom}] \quad (8.30)$$

Reaktion 2:

Aufbrechen des transienten Komplexes aus Plasmid-gebundener RNAII, RNAI und Rom-Protein

$$v_2 = k_2 \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI} - \text{Rom}] \quad (8.31)$$

Reaktion 3:

Bindung von RNAI an Plasmid-gebundene RNAII

$$v_3 = k_3 \cdot [\text{pDNA} - \text{RNAII} - s] \cdot [\text{RNAI}] \quad (8.32)$$

Reaktion 4:

Aufbrechen des instabilen Komplexes aus Plasmid-gebundener RNAII und RNAI

$$v_4 = k_4 \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{unstable}}] \quad (8.33)$$

Reaktion 5:

Elongation der Plasmid-gebundenen RNAII

$$v_5 = k_5 \cdot [\text{pDNA} - \text{RNAII} - s] \quad (8.34)$$

Reaktion 6:

Modifikation der Plasmid-gebundenen RNAII

$$v_6 = k_6 \cdot [\text{pDNA} - \text{RNAII} - lo] \quad (8.35)$$

Reaktion 7:

Plasmidreplikation

$$v_7 = k_7 \cdot [\text{pDNA} - \text{RNAII} - \text{Primer}] \quad (8.36)$$

Reaktion 8:

Loslösen der verlängerten RNAII

$$v_8 = k_8 \cdot [\text{pDNA} - \text{RNAII} - \text{lo}] \quad (8.37)$$

Reaktion 9:

Hybridisierung des RNAII-Transkriptes mit der pDNA

$$v_9 = k_9 \cdot [\text{pDNA}] \quad (8.38)$$

Reaktion 10:

Umwandlung des instabilen Komplexes aus Plasmid-gebundener RNAII und RNAI in einen stabilen Komplex

$$v_{10} = k_{10} \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{unstable}}] \quad (8.39)$$

Reaktion 11:

Umwandlung des stabilen Komplexes aus Plasmid-gebundener RNAII und RNAI in einen instabilen Komplex

$$v_{11} = k_{11} \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{stable}}] \quad (8.40)$$

Reaktion 12:

Loslösen des RNAI-RNAII Komplexes von der pDNA

$$v_{12} = k_{12} \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI}_{\text{stable}}] \quad (8.41)$$

Reaktion 13:

Loslösen des Rom-Proteins von der pDNA

$$v_{13} = k_{13} \cdot [\text{pDNA} - \text{RNAII} - \text{RNAI} - \text{Rom}] \quad (8.42)$$

Reaktion 14:

Bindung von RNAII an eine unbeladene tRNA

$$v_{14} = k_{14} \cdot [\text{tRNA}_{\text{un}}] \cdot [\text{RNAII}] \quad (8.43)$$

Reaktion 15:

Bindung einer unbeladenen tRNA an eine Plasmid-gebundene RNAII

$$v_{15} = k_{15} \cdot [\text{tRNA}_{\text{un}}] \cdot [\text{pDNA} - \text{RNAII} - \text{s}] \quad (8.44)$$

Reaktion 16:

Elongation der Plasmid-gebundenen RNAII und Loslösen der gebundenen unbeladenen tRNA

$$v_{16} = k_{16} \cdot [\text{pDNA} - \text{RNAII} - \text{tRNA}_{\text{un}}] \quad (8.45)$$

Reaktion 17:

Bindung der pDNA an RNAII, welche eine unbeladene tRNA gebunden hat

$$v_{17} = k_{17} \cdot [\text{tRNA}_{\text{un}} - \text{RNAII}] \cdot [\text{pDNA}] \quad (8.46)$$

Reaktion 18:

RNAI Synthese

$$v_{18} = k_{18} \cdot [\text{pDNA}] \quad (8.47)$$

Reaktion 19:

RNAII Synthese

$$v_{19} = k_{19} \cdot [\text{pDNA}] \quad (8.48)$$

Reaktion 20:

Synthese des Rom-Proteins

$$v_{20} = k_{20} \cdot [\text{pDNA}] \quad (8.49)$$

Reaktion 21:

RNAI Degradierung

$$v_{21} = k_{21} \cdot [\text{RNAI}] \quad (8.50)$$

Reaktion 22:

RNAII Degradierung

$$v_{22} = k_{22} \cdot [\text{RNAII}] \quad (8.51)$$

Reaktion 23:

Degradierung des Rom-Proteins

$$v_{23} = k_{23} \cdot [\text{Rom}] \quad (8.52)$$

Reaktion 24:

Synthese einer unbeladenen tRNA

$$v_{24} = k_{24} \quad (8.53)$$

Reaktion 25:

Bildung eines RNAI-RNAII Komplexes

$$v_{25} = k_{25} \cdot [\text{RNAI}] \cdot [\text{RNAII}] \quad (8.54)$$

Reaktion 26:

Degradierung einer unbeladenen tRNA

$$v_{26} = k_{26} \cdot [\text{tRNA}_{\text{un}}] \quad (8.55)$$

Die in den Gleichungen verwendeten kinetischen Konstanten k_i ($i = 1 \dots 26$) sind in Tabelle 8.20 aufgelistet. Die Spalte „Quelle“ gibt die Herkunft der jeweiligen Parameter an. Alle kinetischen Konstanten, die mit einem „c“ gekennzeichnet sind, wurden in dieser Arbeit durch Parameteranpassung bestimmt. Dazu wurden die Simulationen mit der zum Zeitpunkt T_2 gemessenen pDNA-Konzentration verglichen und die jeweiligen Parameter wiederholt angepasst, bis die simulierte pDNA-Konzentration der experimentell bestimmten Plasmidkopienzahl entsprach.

Die Werte der mit „g“ gekennzeichneten Konstanten wurden im Rahmen dieser Arbeit durch bestimmte Annahmen abgeschätzt: Der Parameter k_{16} wurde gewählt, aufgrund der Annahme, dass ein Plasmid-gebundenes RNAII-Molekül mit einer gleichzeitig gebundenen unbeladenen tRNA mit der gleichen Rate prozessiert wird, wie in Reaktion 5, in der das RNAII-Molekül nur die Plasmid-DNA gebunden hat. Für die kinetische Konstante der Reaktion 17 (k_{17}) wurde vorausgesetzt, dass ein Komplex aus unbeladener tRNA und RNAII die Plasmid-DNA in gleicher Weise bindet, wie ein freies RNAII-Molekül. Für k_{25} gilt: $k_{25} = k_3$, unter der Annahme, dass ein RNAI-Molekül ein freies RNAII-Molekül mit der gleichen Rate bindet, wie ein Plasmid-gebundenes RNAII-Molekül.

Für alle Parameter gilt, dass sie von der Wachstumsphase unabhängig sind. Das bedeutet sie bleiben für jede Wachstumsrate konstant.

Tabelle 8.20: Verwendete Parameter zur Simulation der *ColE1*-Plasmidreplikationskontrolle

^a (Brendel & Perelson, 1993), ^b (Wang *et al.*, 2002) mit der Annahme, dass tRNA_{un} die RNAII-Moleküle mit derselben Geschwindigkeitskonstante binden, wie RNAI-Moleküle in Reaktion 3; ^c diese Arbeit (angepasst durch *in vivo* Messungen); ^d k_{18} und k_{19} wurden unter der Annahme einer RNA-Polymerase-Transkriptionsrate von 50 Nucleotiden/s berechnet (von Hippel *et al.*, 1984); ^e k_{24} wurde unter der Annahme einer Transkriptionsrate von 42 Nucleotiden/s berechnet (Gotta *et al.*, 1991) und (Klumpp, 2011); ^f k_7 wurde unter der Annahme einer DNA-Polymerase-Elongationsrate von 500 bp/s (Alberts *et al.*, 2007) mit Berücksichtigung der Plasmidgröße berechnet; ^g diese Arbeit (geschätzt).

Reaktion	Geschwindigkeitskonstante	Low copy Plasmid	High copy Plasmid	Quelle
1	k_1 [$M^{-1} \cdot \text{min}^{-1}$]	$1.7 \cdot 10^8$	non-existent	a,-
2	k_2 [min^{-1}]	0.17	non-existent	a,-
3	k_3 [$M^{-1} \cdot \text{min}^{-1}$]	$1.02 \cdot 10^8$	$3.05 \cdot 10^6$	c,c
4	k_4 [min^{-1}]	48	20	a,c
5	k_5 [min^{-1}]	12	12	a,a
6	k_6 [min^{-1}]	4.3	4.3	a,a
7	k_7 [min^{-1}]	3.8	4.19	f,f
8	k_8 [min^{-1}]	4.3	4.3	a,a
9	k_9 [$M^{-1} \cdot \text{min}^{-1}$]	0.25	0.25	a,a
10	k_{10} [min^{-1}]	44	44	a,a
11	k_{11} [min^{-1}]	0.085	0.085	a,a
12	k_{12} [min^{-1}]	17	17	a,a
13	k_{13} [min^{-1}]	34	non-existent	a,-
14	k_{14} [min^{-1}]	$1.8 \cdot 10^5$	$1.8 \cdot 10^5$	b,b
15	k_{15} [$M^{-1} \cdot \text{min}^{-1}$]	$1.8 \cdot 10^5$	$1.8 \cdot 10^5$	b,b
16	k_{16} [min^{-1}]	12	12	g,g
17	k_{17} [min^{-1}]	0.25	0.25	g,g
18	k_{18} [min^{-1}]	6	6	a,a
19	k_{19} [min^{-1}]	0.25	0.25	a,a
20	k_{20} [min^{-1}]	4	non-existent	a,-
21	k_{21} [min^{-1}]	0.35	0.35	a,a
22	k_{22} [min^{-1}]	0.35	0.35	a,a
23	k_{23} [min^{-1}]	0.14	non-existent	a,-
24	k_{24} [$M \cdot \text{min}^{-1}$]	$7.99 \cdot 10^{-6}$	$7.99 \cdot 10^{-6}$	e,e
25	k_{25} [$M^{-1} \cdot \text{min}^{-1}$]	$1.02 \cdot 10^8$	$3.05 \cdot 10^6$	g,g
26	k_{26} [min^{-1}]	0.35	0.35	c,c

Nach Festlegung der kinetischen Konstanten konnte das zuvor beschriebene DGL-System unter Verwendung eines numerischen Lösungsansatzes mit Hilfe des Computers gelöst werden. Als Anfangsbedingungen wurden dabei die experimentell bestimmten RNAI- und RNAII-Konzentrationen zusammen mit dem halbierten Wert der gemessenen pDNA-Konzentrationen und der entsprechenden Wachstumsrate verwendet. Die Motivation für das Einsetzen der halbierten Werte der gemessenen pDNA-Konzentrationen ist die Annahme, dass die experimentell bestimmten Plasmidkonzentrationen den Plasmidkopienzahlen direkt vor der Zellteilung entsprechen und sich die Plasmidmenge innerhalb einer Generationszeit verdoppelt. Alle unbekanntes Substratkonzentrationen bekamen den Wert null zugewiesen, da davon ausgegangen wird, dass diese Stoffe erst synthetisiert werden müssen.

Nachdem die zuvor beschriebenen Parameter und Anfangsbedingungen in das dynamische Modell integriert worden sind, kann das Modell zur Simulation der *ColE1*-Plasmidreplikationskontrolle verwendet werden.

Das dynamische Modell besteht aus vier verschiedenen Dateien: „Main_low_copy_plasmid.m“, „DGL_low_copy_plasmid.m“ und „Main_high_copy_plasmid.m“, „DGL_high_copy_plasmid.m“. Mit den Dateien „Main_low_copy_plasmid.m“ und „DGL_low_copy_plasmid.m“ lässt sich die *ColE1*-Plasmidreplikationskontrolle für den *low copy* plasmidtragenden Stamm DH5 α -pSUP 201-3 simulieren. Die Datei „Main_low_copy_plasmid.m“ ist eine sogenannte *Script-File* und die Datei „DGL_low_copy_plasmid.m“ ist ein sogenanntes *Function-File*. In dem *Script-File* kann der zu simulierende Zeitpunkt (übergeben als Variable „tend“) mit der zugehörigen Wachstumsrate (übergeben als Vektor „gr“) und den jeweiligen Anfangsbedingungen (übergeben als Vektor „s“) ausgewählt werden. Diese werden zusammen mit den zeitpunktunabhängigen Parametern über den Aufruf der Funktion „DGL_low_copy_plasmid“ an das *Function-File* übergeben. Im *Function-File* ist das DGL-System definiert, welches mit Hilfe des *MATLAB-solvers* „ode15s“ gelöst wird. Das Ergebnis (Simulation) sind zwei Vektoren: Ein Zeitvektor und ein Konzentrationsvektor, der die Konzentrationen aller beteiligten Substrate zu jedem Zeitpunkt enthält. Plottet man den Konzentrationsvektor gegen den Zeitvektor, lassen sich die Substratkonzentrationen im Verlauf der Zeit bildlich darstellen.

Mit den Dateien „Main_high_copy_plasmid.m“ und „DGL_high_copy_plasmid.m“ lässt sich die *ColE1*-Plasmidreplikationskontrolle für den *high copy* plasmidtragenden Stamm DH5 α -pCMV-lacZ simulieren. Der Aufbau der Dateien gleicht dem von

„Main_low_copy_plasmid.m“ und „DGL_low_copy_plasmid.m“ und die Ausführung erfolgt analog zu denen des *low copy* plasmidtragenden Stamms. Der einzige Unterschied zwischen den Dateien für DH5 α -pSUP 201-3 und für DH5 α -pCMV-lacZ sind die drei Reaktionen (Reaktion 1, 2 und 13), an denen das Rom-Protein beteiligt ist. Diese Reaktionen sind in „Main_high_copy_plasmid.m“ und „DGL_high_copy_plasmid.m“ nicht definiert.

8.4 SIMULATIONEN DER *colE1*-ÄHNLICHEN PLASMIDREPLIKATIONSKONTROLLE FÜR EINE EINZELNE ZELLE

8.4.1 Simulationen zur Parameteranpassung

Wie bereits erwähnt wurden die in Tabelle 8.20 mit einem „c“ gekennzeichneten Parameter durch Parameteranpassung anhand der experimentell bestimmten Daten ermittelt. Sowohl für den *low copy*, als auch für den *high copy* plasmidtragenden Stamm wurden dafür die zum Zeitpunkt T_2 gemessenen Daten verwendet. Die anderen beiden Zeitpunkte (T_1 und T_3) wurden für die Validierung des Modells genutzt.

Die Simulationen wurden für eine einzelne Zelle durchgeführt, wobei die für die Zeitpunkte T_1 bis T_3 , experimentell bestimmten Daten als Anfangsbedingungen gewählt wurden. Die Simulationszeit entspricht der Generationszeit, die mit Hilfe der, für jeden Erntezeitpunkt ermittelten, Wachstumsraten berechnet wurde. Des Weiteren wurde für alle Simulationen in dieser Arbeit angenommen, dass die RNAI-, RNAII- und die unbeladenen tRNA-Moleküle sowie das Rom-Protein frei in der Zelle diffundieren können. In Anlehnung an das von Brendel & Perelson (1993) vorgeschlagene Modell wird ebenfalls angenommen, dass die Aktivität der RNase H konstant ist.

Die für DH5 α -pSUP 201-3 simulierten Plasmidkopienzahlen sind in der Tabelle 8.21 dargestellt. Die Simulationen wurden dabei zum jeweiligen Erntezeitpunkt begonnen und für die Dauer einer Zellverdopplung (d. h. Simulationszeit = Generationszeit) durchgeführt, wobei eine einzelne Zelle betrachtet wurde. Am Beispiel der Simulation des dritten Erntezeitpunkts bedeutet das, dass die experimentellen Daten zum Zeitpunkt T_3 als Anfangsbedingungen gewählt wurden und die Simulationszeit 147 min betrug. Das heißt, eine Zelle beginnt zum Zeitpunkt T_3 zu wachsen, teilt sich nach 147 min und besitzt direkt vor ihrer Teilung 49 Plasmide.

Tabelle 8.21: Ergebnisse der *in vitro* und *in silico* bestimmten Plasmidkopienzahlen für DH5 α -pSUP 201-3. Bei den *in vitro* bestimmten Plasmidkopienzahlen handelt es sich um die von der Plasmid Factory durchschnittlich bestimmten Plasmidkopienzahlen und die *in silico* bestimmten Plasmidkopienzahlen basieren auf den Simulationen des dynamischen Modells.

Zu den verwendeten Anfangsbedingungen: Bei den *in silico* Berechnungen wurden die zu den drei Erntezeitpunkten (Abb. 8.4) bestimmten experimentellen Daten als Anfangsbedingungen der Simulationen gewählt. Dazu gehören die gemessenen RNAI- und RNAII-Konzentrationen (Tab. 8.8 und 8.9) sowie die von der Plasmid Factory durchschnittlich bestimmten Plasmidkopienzahlen (Tab. 8.15). Des Weiteren wurde die Wachstumsraten mit der zugehörigen Generationszeiten (Tab. 8.2) von dem Erntezeitpunkt, an dem die Simulation beginnt, in das Modell eingesetzt.

Zeitpunkt	Gemessene PCN	Simulierte PCN
T ₁	46 \pm 26	49
T ₂	47 \pm 30	47
T ₃	46 \pm 17	49

Vergleicht man die simulierten Plasmidkopienzahlen mit den von der Plasmid Factory durchschnittlich gemessenen Plasmidkopienzahlen (Tab. 8.15), lässt sich festhalten, dass die Simulationen die real stattfindende Plasmidreplikationskontrolle gut widerspiegeln. In den experimentellen Daten lässt sich zu allen Zeitpunkten ein stabiler durchschnittlicher Plasmidgehalt pro Zelle beobachten. Dieses Verhalten lässt sich mit Hilfe des dynamischen Modells reproduzieren.

Betrachtet man die simulierten Plasmidkonzentrationen des *high copy* plasmidtragenden Stamms DH5 α -pCMV-lacZ sieht man, dass die zu den Zeitpunkten T₁ und T₃ von der Plasmid Factory durchschnittlich gemessenen Plasmidkopienzahlen ebenfalls gut durch die Simulationen wiedergegeben werden (Tabelle 8.22). Des Weiteren ist hier anzumerken, dass simulierten Plasmidkopienzahlen für alle Zeitpunkte innerhalb der Spanne liegen, die seitens der durchschnittlich gemessenen PCN durch die Fehlerbalken entsteht.

Tabelle 8.22: Ergebnisse der *in vitro* und *in silico* bestimmten Plasmidkopienzahlen für DH5 α -pCMV-lacZ. Bei den *in vitro* bestimmten Plasmidkopienzahlen handelt es sich um die von der Plasmid Factory durchschnittlich bestimmten Plasmidkopienzahlen und die *in silico* bestimmten Plasmidkopienzahlen basieren auf den Simulationen des dynamischen Modells.

Zu den verwendeten Anfangsbedingungen: Bei den *in silico* Berechnungen wurden die zu den drei Erntezeitpunkten (Abb. 8.5) bestimmten experimentellen Daten als Anfangsbedingungen der Simulationen gewählt. Dazu gehören die gemessenen RNAI- und RNAII-Konzentrationen (Tab. 8.11 und 8.12) sowie die von der Plasmid Factory durchschnittlich bestimmten Plasmidkopienzahlen (Tab. 8.17). Des Weiteren wurde die Wachstumsraten mit der zugehörigen Generationszeiten (Tab. 8.4) von dem Erntezeitpunkt, an dem die Simulation beginnt, in das Modell eingesetzt.

Zeitpunkt	Gemessene PCN	Simulierte PCN
T ₁	1514 \pm 1301	1718
T ₂	2403 \pm 713	2405
T ₃	5805 \pm 4828	5421

Um das Modellverhalten weiter zu überprüfen, wurde der Zusammenhang zwischen Wachstumsrate und Plasmidproduktion simuliert und mit Experimenten von Atlung *et al.* (1999) verglichen.

Atlung und seine Arbeitsgruppe haben den Einfluss des Rom-Proteins auf die Plasmidkopienzahl eines pBR322-Plasmids bei verschiedenen Wachstumsraten untersucht. Dabei haben sie festgestellt, dass die Plasmidkopienzahl eines pBR322-Plasmids bei einer Zunahme der Generationszeit von 29 min auf 150 min, um das drei- bis vierfache zunimmt (Atlung *et al.*, 1999). Nach Cooper (2004) besitzen pBR322-Plasmide eine durchschnittliche Kopienzahl von 15 - 20 Plasmiden pro Zelle. Geht man davon aus, dass sich, bei einer Generationszeit von 29 min, 15 - 20 Plasmide in einer Zelle befinden, so liegen nach Atlung *et al.* (1999) bei einer Generationszeit von 150 min zwischen 45 und 60 pBR322-Plasmide in einer Zelle vor. Um das Modellverhalten zu überprüfen wurden die simulierten Plasmidkopienzahlen für DH5 α -pSUP 201-3 (Tab. 8.21) mit den nach Atlung *et al.* (1999) und Cooper (2004) berechneten Kopienzahlen verglichen (Abb. 8.9). Man sieht, dass die simulierten Plasmidkopienzahlen mit 49, 47 und 49 bei Generationszeiten um 150 min innerhalb der berechneten Spannweite von 45 bis 60 Plasmiden liegen. Das Modell ist somit in der Lage, die Plasmidkopienzahl einer Zelle in der geforderten Genauigkeit wiederzugeben.

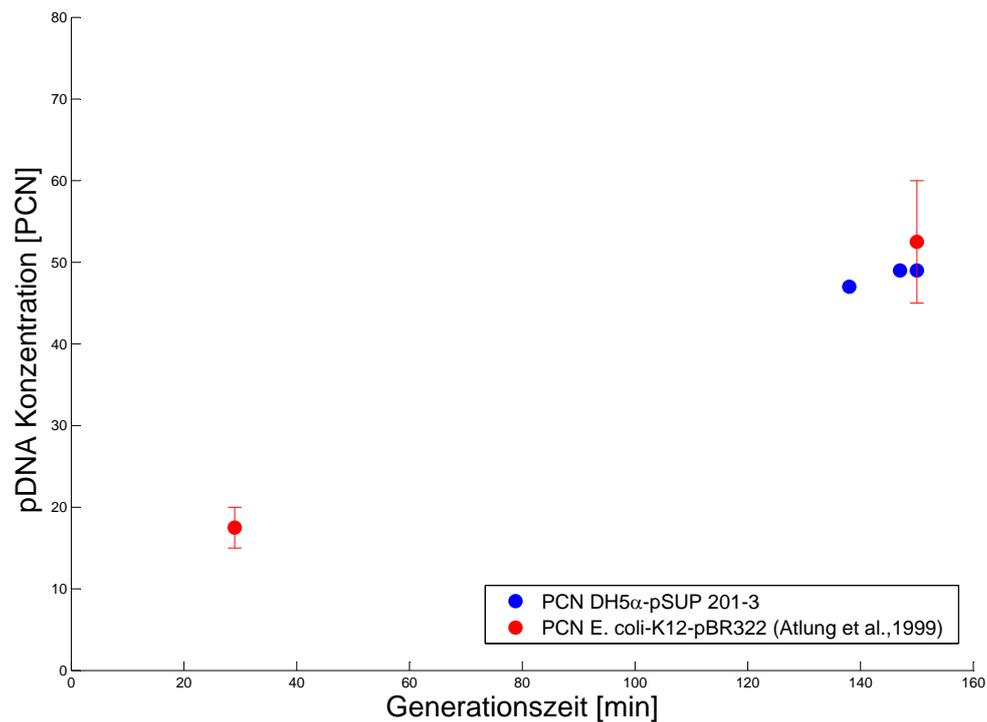


Abbildung 8.9: Überprüfung des dynamischen Modells durch Vergleichen der *in silico* berechneten Plasmidkopienzahlen für DH5α-pSUP 201-3 (Tab. 8.21) (blau •) mit den nach Atlung *et al.* (1999) und Cooper (2004) berechneten Kopienzahlen (rot •). Die *in silico* berechneten Plasmidkopienzahlen für DH5α-pSUP 201-3 wurden durch Simulationen, beginnend an den drei Erntezeitpunkten T_1 , T_2 und T_3 , ermittelt. Zu den verwendeten Anfangsbedingungen: Bei den *in silico* Berechnungen wurden die zu den drei Erntezeitpunkten (Abb. 8.4) bestimmten experimentellen Daten als Anfangsbedingungen der Simulationen gewählt. Dazu gehören die gemessenen RNAI- und RNAII-Konzentrationen (Tab. 8.8 und 8.9) sowie die von der Plasmid Factory durchschnittlich bestimmten Plasmidkopienzahlen (Tab. 8.15). Des Weiteren wurde die Wachstumsraten mit der zugehörigen Generationszeiten (Tab. 8.2) von dem Erntezeitpunkt, an dem die Simulation beginnt, in das Modell eingesetzt. Die Generationszeiten lagen dabei um 150 min herum, was der Generationszeit aus den Versuchen von Atlung *et al.* (1999) entspricht. Atlung *et al.* (1999) haben festgestellt, dass die Plasmidkopienzahl eines pBR322-Plasmids bei einer Zunahme der Generationszeit von 29 min auf 150 min um das Drei- bis Vierfache zunimmt (Atlung *et al.*, 1999). Nach Cooper (2004) besitzen pBR322-Plasmide eine durchschnittliche Kopienzahl von 15 - 20 Plasmiden pro Zelle (in der Abbildung bei $x=29$ min eingezeichnet). Geht man davon aus, dass sich, bei einer Generationszeit von 29 min, 15 - 20 Plasmide in einer Zelle befinden, so liegen nach Atlung *et al.* (1999) durch eine drei- bis vierfache Zunahme der PCN, bei einer Generationszeit von 150 min, zwischen 45 und 60 pBR322-Plasmide in einer Zelle vor (in der Abbildung bei $x=150$ min eingezeichnet).

Zusammen mit den schon zuvor beschriebenen Simulationen zur Validierung des Modells bestätigen diese Ergebnisse, dass das Modell das angestrebte Verhalten in der ge-

forderten Genauigkeit zeigt, so dass es zur Durchführung von prädiktiven Simulationen und somit zur Untersuchung der zu Beginn dieses Kapitels vorgestellten Fragestellung eingesetzt werden kann.

8.4.2 Prädiktive Simulationen

Mit Hilfe des erstellten Modells soll der Fragestellung nachgegangen werden, ob es einen Effekt auf die *ColE1*-Plasmidproduktion hätte, würde man tRNA-Moleküle derart modifizieren, dass die Beladung mit einer Aminosäure durch die Aminoacyl-tRNA-Synthetase nicht mehr möglich ist.

Eine hohe Menge unbeladener tRNA-Moleküle kommt in der Regel nur in Situationen vor, in denen sich die Zelle unter Stressbedingungen befindet. Dies ist z. B. unter Nährstoffknappheit und dem damit verbundenen Aminosäuremangel der Fall. Da die Zelle unter diesen Umständen weiterhin Aminosäuren verbraucht, nimmt die Zahl der unbeladenen tRNAs in der Zelle zu. Wróbel & Węgrzyn (1998) konnten bereits zeigen, dass diese Zunahme an unbeladenen tRNAs mit einer hohen PCN in Zusammenhang steht. Das bedeutet, für die Plasmidproduktion wäre es von Vorteil, wenn sich die Zellen unter Nährstoffmangel befinden, da dies die PCN erhöhen würde. Diese Bedingungen sind jedoch in einem Langzeit-Produktionsprozess nicht anwendbar, da die Zellen nicht bis zum Ende der Kultivierungen überleben würde.

Eine Möglichkeit dieses Problem zu umgehen, ist die Überexpression eines *tRNA*-Gens, wodurch sich die Konzentration an unbeladenen tRNAs in der Zelle erhöhen würde. Es ist jedoch zu befürchten, dass diese Erhöhung nicht ausreicht, um einen Einfluss auf die Plasmidproduktion zu haben, da die zusätzlichen unbeladenen tRNAs mit einer hohen Wahrscheinlichkeit von den Aminoacyl-tRNA-Synthetasen beladen werden, bevor sie die Plasmidreplikation beeinflussen können. Des Weiteren ist es wichtig, keine zu hohe Anzahl unbeladener tRNAs zu provozieren, um der Zelle keinen Aminosäuremangel vorzutäuschen. Denn wird eine bestimmte Konzentration unbeladener tRNAs überschritten, bekommt die Zelle den Eindruck, sie stünde unter Aminosäuremangel und löst die sogenannte *stringent response* aus. Dies hat eine Aktivierung des Stresssignals ppGpp (Guanosin-5'-Diphosphate-3'-Diphosphate) zur Folge (Węgrzyn, 1999). Durch das ppGpp-Signal werden in der Zelle weitere Stressantworten ausgelöst, wie z.B. die Inhibition der Chromosom- und Plasmidreplikation, was nachteilig für die Plasmidproduktion wäre (Guzman *et al.*, 1988), (Hecker *et al.*, 1988).

Eine andere Möglichkeit, das Problem der mit einer hohen Konzentration unbeladener tRNAs einhergehenden Mangelbedingungen zu umgehen, ist die Integration eines Gens, das für eine modifizierte tRNA kodiert. Diese veränderte tRNA besitzt die Besonderheit, dass sie nicht mehr beladen werden kann, gleichzeitig jedoch noch in der Lage ist, ihre Rolle bei der Plasmidreplikationskontrolle auszuüben. Der Vorteil dieser modifizierten tRNA ist, dass die Zelle keinen Stressbedingungen mehr ausgesetzt wäre, aber dennoch eine erhöhte Menge unbeladener tRNA-Moleküle aufweisen würde, was sich positiv auf die Plasmidproduktion auswirken könnte. Diese Hypothese gilt es mit dem in dieser Arbeit erstellten Modell zu überprüfen.

Zur Überprüfung der Hypothese wurde die *ColE1*-Plasmidreplikationskontrolle von DH5 α -pCMV-lacZ, beginnend zum Zeitpunkt T_3 , für folgende drei Bedingungen simuliert: Wachstum unter normalen Nährstoffbedingungen, Wachstum unter Mangelbedingungen und Wachstum mit modifiziertem tRNA-Gen ohne Mangelbedingungen. Die Simulationen wurden für alle drei Bedingungen zum Erntezeitpunkt T_3 begonnen und für die Dauer einer Zellverdopplung (d. h. Simulationszeit = Generationszeit(T_3) = 347 min) durchgeführt, wobei eine einzelne Zelle betrachtet wurde. Als Anfangsbedingungen wurden die experimentell bestimmten Daten zum Zeitpunkt T_3 verwendet. Dazu gehören die gemessenen RNAI- und RNAII-Konzentrationen sowie der halbierte Wert der pDNA-Konzentration und die Wachstumsrate. Die übrigen Substrate bekamen die Konzentration null zugewiesen. Auf biologischer Ebene bedeutet das, es wird eine Zelle betrachtet, die zum Zeitpunkt T_3 zu wachsen beginnt, ihre Plasmide repliziert und sich nach 347 min in zwei Tochterzellen teilt. An dieser Stelle endet die Simulation und die Tochterzellen werden nicht mehr betrachtet. Der Zeitpunkt T_3 wurde exemplarisch gewählt, da zu diesem Zeitpunkt die höchste Plasmidkopienzahl gemessen wurde. Es können jedoch mit Hilfe des Modells ebensogut die Zeitpunkte T_1 und T_2 betrachtet werden. Des Weiteren wurden die Simulationen für den *high copy* plasmidtragenden Stamm durchgeführt, da bei einer Plasmidproduktion im großen Maßstab in der Regel nur *high copy* Plasmide verwendet werden. Die Simulationen lassen sich jedoch auch für den *low copy* plasmidtragenden Stamm durchführen, wobei qualitativ das gleiche Ergebnis zu beobachten ist.

Die Simulationen zu den drei verschiedenen Bedingungen sind in Abbildung 8.10 graphisch dargestellt. In der ersten Situation wurde die Plasmidproduktion für normale Nährstoffbedingungen simuliert. Da unter diesen Umständen ausreichend Aminosäuren vorhanden sind, wurde für die Simulation die kinetische Konstante zur Synthese

unbelasteter tRNA-Moleküle (k_{24}) mit dem Faktor 0.01 multipliziert. Dies beruht auf der Annahme, dass bei ausreichender Nährstoffversorgung lediglich $\approx 1\%$ der tRNA-Moleküle in der Zelle unbeladen sind. Stehen der Zelle genügend Nährstoffe zur Verfügung, wird nach Ablauf einer Generationszeit (347 min) eine Zunahme der PCN um 2523 (5421 Plasmide zum Zeitpunkt T_3 – Start-PCN = 2523) Plasmide durch das Modell vorhergesagt. Die zweite Simulation wurde für eine Zelle durchgeführt, die unter Nährstoffmangel leidet. Dabei wurde die Annahme getroffen, dass aufgrund von schlecht ablaufender Proteinbiosynthese und langsamerem Wachstum bei unzureichender Nährstoffversorgung sich die Konzentration aller Proteine verringert, so dass alle Reaktionen nur zu 10 % ablaufen. Der Wert 10 % wurde aufgrund mangelnder Angaben in der Literatur für diese Arbeit geschätzt. Es wäre auch jeder andere Wert zwischen 1 % und 99 % denkbar. Ein anderer Wert hätte lediglich zur Folge, dass die Verringerung der pDNA-Produktion für die Bedingung „Nährstoffmangel“ unterschiedlich deutlich ausfallen würde. Für die mathematische Beschreibung bedeutet das, dass alle kinetischen Konstanten und die Wachstumsrate mit dem Faktor 0.1 multipliziert werden. Des Weiteren wurde der Faktor 0.1 vor k_{24} durch eine 1 ersetzt, da unter Nährstoff- bzw. Aminosäuremangel eine hohe Anzahl unbelasteter tRNAs in der Zelle vorkommen. Man sieht, dass für diesen Fall, nach Ablauf einer Generationszeit, nur noch eine geringe Plasmidproduktion vorhergesagt wird und die Zelle lediglich 379 Plasmide produziert. In der dritten Situation wurde die zuvor beschriebene Hypothese untersucht. Dafür wurde eine Situation simuliert, in der ausreichend Nährstoffe zur Verfügung stehen, aber gleichzeitig durch das modifizierte *tRNA*-Gen viele unbeladene tRNA-Moleküle in der Zelle vorkommen, die die Plasmidreplikation fördern. Für die mathematische Beschreibung bedeutet das, dass alle Reaktionen zu 100 % ablaufen, da kein Nährstoffmangel herrscht. Des Weiteren wurde der Faktor 0.1 vor k_{24} durch eine 1 ersetzt (entspricht einem Ablauf der Reaktion zu 100 %), da in dieser Situation aufgrund des modifizierten *tRNA*-Gens eine hohe Anzahl unbelasteter tRNA-Moleküle in der Zelle vorkommen. Unter diesen Umständen wird durch das Modell, nach Ablauf einer Generationszeit, eine Plasmidproduktion von 3822 Plasmiden pro Zelle vorhergesagt.

Vergleicht man alle drei Simulationen (Abbildung 8.10) miteinander, stellt man fest, dass die vorhergesagte Plasmidproduktion unter Nährstoffknappheit am geringsten ist. Die meisten Plasmide werden laut Modell produziert, wenn man ein Gen einfügen würde, das für eine modifizierte tRNA kodiert, die die Plasmidproduktion erhöhen würde. Auf

den genauen Aufbau eines solchen modifizierten *tRNA*-Gens wird im nächsten Unterkapitel näher eingegangen.

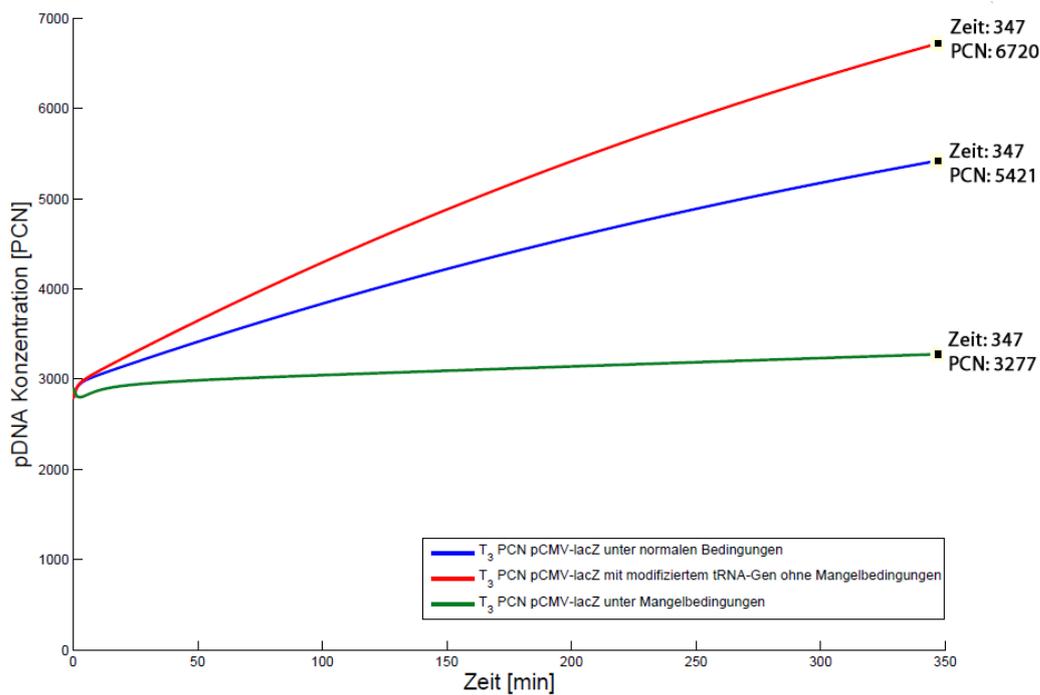


Abbildung 8.10: Simulation der Plasmidkopienzahl für DH5 α -lacZ (beginnend zum Zeitpunkt T_3) für drei verschiedene Bedingungen:

blau: Plasmidproduktion für normale Nährstoffbedingungen

grün: Plasmidproduktion unter Nährstoffmangel

rot: Normale Nährstoffbedingungen in Anwesenheit von unbeladenen *tRNA*-Molekülen (kodiert durch das modifizierte *Leucin-tRNA*-Gen)

8.5 ENTWICKLUNG EINER UNBELADBAREN TRANSFER-RNA

8.5.1 Die Transkriptionseinheit

Unbeladene Transfer-RNA-Moleküle (*tRNA*-Moleküle) spielen in der *ColE1*-Plasmidreplikationskontrolle eine besondere Rolle, da sie die Plasmidkopienzahl positiv beeinflussen können (Abbildung 3.5). Aus diesem Grund wird in dieser Arbeit ein *tRNA*-Gen konstruiert, das für ein *tRNA*-Molekül kodiert, welches nicht mehr mit einer Aminosäure beladen werden kann.

Das Genom von *E. coli* besitzt 79 *tRNA*-Gene, welche über das gesamte Chromosom verteilt vorliegen. In der Regel sind sie dabei in Clustern angeordnet. Lediglich 20 dieser

Gene befinden sich in einer monocistronischen Transkriptionseinheit. Die übrigen 59 tRNA-Gene sind in polycistronischen Operons sowohl mit weiteren tRNA-Genen, als auch mit rRNA- und proteinkodierenden Genen organisiert. Dabei sind die Sequenzen zwischen den einzelnen, aufeinanderfolgenden Genen sehr variabel bezüglich ihrer Basenabfolge und Länge. Alle bisher bekannten tRNA-Gene sind 75 bp bis 90 bp lang und haben gemeinsam, dass sie keine Intronelemente besitzen und am 3'-Terminus eine CCA-Sequenz tragen. (Neidhardt, 1987), (Inokuchi & Yamao, 1995)

Wróbel & Węgrzyn (1998) fanden heraus, dass für die *ColE1*-Plasmidreplikationskontrolle insbesondere die unbeladenen tRNAs der Aminosäuren Isoleucin, Leucin, Histidin, Arginin und Threonin eine Rolle spielen. Sie kultivierten *E. coli*-Bakterien mit verschiedenen *ColE1*-ähnlichen Plasmiden jeweils unter Mangel der zu untersuchenden Aminosäure und dabei zeigte sich, dass die Erhöhung der Plasmidkopienzahl unter Leucin-Mangel am Größten ist, also eine große Menge unbeladener Leucin-tRNAs in der Zelle vorliegt. (Wróbel & Węgrzyn, 1998)

Der genaue Kontrollmechanismus ist noch nicht gänzlich geklärt, aber aufgrund von Sequenzhomologien der Leucin-tRNA-Anticodon-Schleife mit der zweiten Schleife des RNAI-Moleküls wird vermutet, dass die unbeladenen Leucin-tRNA's mit ihrer Anticodon-Schleife an RNAII-Moleküle binden können und sie so vor den inhibitorisch wirkenden RNAI-Molekülen schützen Yavachev & Ivanov (1988). Des Weiteren wird angenommen, dass auch die Aminosäurebindestelle an der tRNA-RNA-Interaktion beteiligt ist (Wang *et al.*, 2002).

Auf der Grundlage der Erkenntnisse von Yavachev & Ivanov (1988) und Wang *et al.* (2002) wurde in dieser Arbeit das Gen für eine modifizierte Leucin-tRNA konstruiert. Das Genprodukt soll eine veränderte Leucin-tRNA sein, welche nicht mehr von der Leucyl-tRNA-Synthetase erkannt und beladen werden kann.

Die modifizierte Leucin-tRNA ist so aufgebaut, dass die für die Plasmidreplikationskontrolle relevanten Elemente noch vorhanden sind, die Erkennungsstellen zur Beladung durch die Leucyl-tRNA-Synthetase jedoch fehlen. Die Grundlage für die Konstruktion des modifizierten Leucin-tRNA-Gens bildet die Sequenz des *E. coli* Leucin-tRNA Gens *leuQ*. Auf dem Genom von *E. coli* liegen acht Leucin-tRNA-Gene (*leuP*, *leuQ*, *leuT*, *leuU*, *leuV*, *leuW*, *leuX* und *leuZ*), welche sich in ihren Sequenzen sehr ähneln (Soerensen *et al.*, 2005). Sechs von diesen tRNA-Genen befinden sich in einer Operon-Struktur und zwei liegen in einer monocistronischen Transkriptionseinheit (Komine *et al.*, 1990). Das Gen *leuQ* liegt zusammen mit den Genen *leuP* und *leuV* in einer polycistronischen

Transkriptionseinheit und besteht aus 76 bp (Duester *et al.*, 1981). Die Ergebnisse von Asahara *et al.* (1993) zeigen, dass die Leucyl-tRNA-Synthetase bis zu sechs verschiedene Leucin-tRNA-Moleküle der Klasse II belädt. Daher wird in dieser Arbeit angenommen, dass alle dieser tRNAs die gleichen Strukturen besitzen, an denen sie von dem Enzym erkannt werden, so dass das *leuQ*-Gen exemplarisch als Grundlage für die Sequenz der modifizierten tRNA gewählt wurde.

Das Gen für die modifizierte Leucin-tRNA soll nur als einfache Kopie in das Chromosom inseriert werden, so dass es notwendig war, dass es als monocistronische Transkriptionseinheit designt wird. Die Grundlage dafür sollte, wenn möglich, eine der beiden monocistronischen Transkriptionseinheiten der Leucin-tRNA-Gene *leuU* oder *leuX* bilden. Diese kamen jedoch beide nicht in Frage, da für das monocistronische *leuX*-Gen in der Literatur wenig über die Promotorregion und die Terminationselemente bekannt ist und das *leuU*-Gen einen präsumtive Promotorbereich innerhalb eines anderen Gens (*secG*) besitzt. Das Problem im Fall von *leuU* war, dass nicht bekannt ist, wieviele Basenpaare vor der Promotorregion möglicherweise noch für die Transkription benötigt werden. Die konstruierte Transkriptionseinheit sollte daher mit einem großzügig gewählten Sequenzbereich vor dem Promotor beginnen. Dabei bestand die Gefahr, dass dieser Bereich zu großzügig gewählt wird und so die Funktionalität z. B. durch die überlappende *secG*-Sequenz gestört wird. Um dies auszuschließen, wurde die monocistronische Transkriptionseinheit des Asparagin-tRNA-Gen *asnT* als Grundlage gewählt, da sie keine Überlappungen mit bekannten kodierenden Sequenzen auf dem Chromosom besitzt und die Promotorregion und die Terminationselemente bekannt sind (Komine *et al.*, 1990).

Die für einen prokaryotischen Promotor typische Konsensussequenz, bestehend aus dem -35- und dem -10-Element, findet sich auch in einigen *upstream* Regionen von tRNA-Transkriptionseinheiten in *E. coli* wieder. Der Bereich zwischen diesen beiden Elementen besteht aus 15 bis 18 Nukleotiden, was ebenfalls charakteristisch für einen prokaryotischen Promotor ist. Zusätzlich befindet sich in der Promotorregion eine 7 bp lange GC-reiche Sequenz. Diese Sequenz entspricht einer hoch konservierten Region, die alle unter stringenter Kontrolle liegenden Gene in *E. coli* gemeinsam haben. (Inokuchi & Yamao, 1995)

Ikemura & Dahlberg (1973) konnten eine Antwort der gesamten tRNA-Synthese auf stringente Kontrolle zeigen, allerdings ist der Zusammenhang zwischen stringenter Kontrolle und individuellen tRNA-Genen insgesamt noch nicht gut untersucht (Travers,

1980) (Travers, 1984). (Inokuchi & Yamao, 1995)

Der präsumptive Promotor der Transkriptionseinheit des Asparagin-tRNA-Gens *asnT* besitzt das -35-Element mit der Sequenz 5'-TTGACA-3' und das -10-Element mit 5'-TAATAT-3'. Zwischen diesen Teilen des Promotors befinden sich 17 bp und nach dem -10-Element folgt eine GC-reiche 7 bp lange Sequenz (5'-CGCCCCG-3'), die darauf hindeutet, dass das *asnT*-Gen einer stringenten Kontrolle unterliegt. (Komine *et al.*, 1990)

Über die Terminationselemente von tRNA-Transkriptionseinheiten ist zur Zeit noch wenig bekannt. DNA-Sequenzierungen haben ergeben, dass fast alle tRNA-Gene im *downstream*-Bereich rho-unabhängige Terminationsstrukturen besitzen, die aus palindromischen Sequenzen mit einer Reihe von 'T'-Resten bestehen. Das präsumptive Terminationselement der Asparagin-tRNA-Transkriptionseinheit ist rho-unabhängig und besteht aus einem Palindrom mit der folgenden Sequenz 5'-CAATGT'...'ACATTG-3'. Diese beiden Bereiche liegen 16 bp voneinander getrennt. (Komine *et al.*, 1990)

8.5.2 Konstruktion des modifizierten Leucin-tRNA-Gens

Im Allgemeinen besitzen tRNAs eine Kleeblattstruktur, die in DHU-Schleife, Anticodon-Schleife, variablen Arm, T ψ C-Schleife und in Akzeptorstamm inklusive Anheftungsstelle für die Aminosäure unterteilt werden kann (Berg *et al.*, 2007). Bei allen tRNAs existieren spezifische Erkennungselemente für die Beladung durch Aminoacyl-tRNA-Synthetasen (Asahara *et al.*, 1993). Diese Elemente können im Akzeptorstamm, im variablen Arm, in der T ψ C-Schleife oder in der DHU-Schleife liegen (Asahara *et al.*, 1993). Des Weiteren können die Anticodonbasen an der Erkennung beteiligt sein (Asahara *et al.*, 1993).

Das *leuQ*-Gen kodiert für eine Leucin-tRNA mit dem Anticodon 'CAG', die von einem Enzym namens Leucyl-tRNA-Synthetase mit ihrer Aminosäure beladen wird (Jühling *et al.*, 2009). Dafür besitzt die Leucin-tRNA bestimmte strukturelle Merkmale, an denen sie von der Leucyl-tRNA-Synthetase erkannt wird (Sohm *et al.*, 2004).

Eine wichtige Rolle bei der Erkennung spielt die sogenannte *discriminator*-Base. Sie befindet sich im Akzeptorstamm, direkt vor der Aminosäurebindestelle, und wird als A₇₃ bezeichnet (Asahara *et al.*, 1993). Ersetzt man die Base A₇₃ durch C₇₃ wird die Leucin-tRNA nicht mehr mit der Aminosäure Leucin beladen. Des Weiteren sind die Basen U₅₅ und C₅₆ der T ψ C-Schleife an der Erkennung beteiligt (Larkin *et al.*, 2002).

Larkin *et al.* (2002) konnten zeigen, dass bei einem Austausch der Basen U₅₅ in C₅₅ und C₅₆ in U₅₆ keine Aminoacylierung stattfindet. Dies ist ebenfalls der Fall bei Deletion der DHU-Schleife (Larkin *et al.*, 2002).

Basierend auf diesen Erkenntnissen wurde das *leuQ*-Gens modifiziert. Im Rahmen der Modifikation wurden die Erkennungselemente der Leucyl-tRNA-Synthetase verändert bzw. entfernt. Gleichzeitig wurde darauf geachtet, dass die, für die *ColE1*-Plasmidreplikationskontrolle relevanten Strukturen, erhalten bleiben. Zu diesen Strukturen gehören die Aminosäurebindestelle am Akzeptorstamm und die Anticodon-Schleife (Yavachev & Ivanov, 1988), (Wang *et al.*, 2002), (Wang *et al.*, 2006c), so dass die einfachste Möglichkeit zur Unterbindung der tRNA-Beladung, die daraus besteht die Aminosäurebindestelle zu entfernen, nicht durchgeführt werden konnte.

In der Abbildung 8.11 ist die Sequenz des *leuQ*-Gens mit den Erkennungselementen für die Leucyl-tRNA-Synthetase und den für die *ColE1*- Plasmidreplikationskontrolle relevanten Strukturen dargestellt.

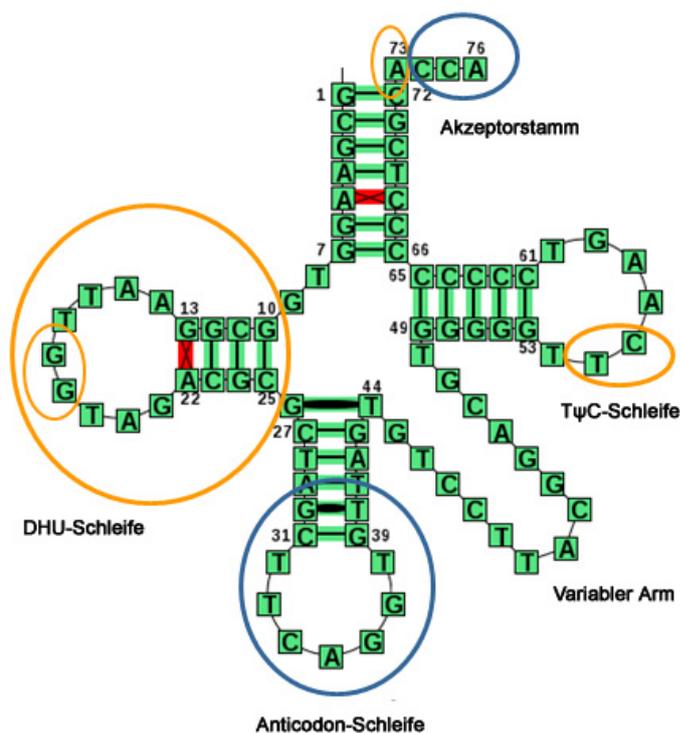


Abbildung 8.11: Sequenz des *leuQ*-Gens mit den Erkennungselementen für die Leucyl-tRNA-Synthetase (gelb) und den für die *ColE1*-Plasmidreplikationskontrolle relevanten Strukturen (blau). (Abbildung der Sequenz aus tRNAdb (Jühling *et al.*, 2009))

Das *leuQ*-Gen wurde derart modifiziert, dass die DHU-Schleife komplett entfernt und durch vier Thyminbasen ersetzt wurde. Gleichzeitig wurden die Basen A₇₃ in

C₇₃, U₅₅ in C₅₅ und C₅₆ in U₅₆ ausgetauscht. Eine weitere Veränderung umfasste den Anticodon-Stamm. Hier wurde die Sequenz der Anticodon-Schleife von 5'-TTCAGGT-3' in 5'-TTGGTAG-3' geändert. Dies geschah aufgrund der Sequenzhomologien zwischen der Anticodon-Schleife der Leucin-tRNA und der zweiten Schleife des RNAI-Moleküls. Die Sequenz 5'-TTGGTAG-3' stellt dabei eine Konsensussequenz dar, die durch den Vergleich von acht tRNA-Sequenzen mit der RNAI-Sequenz ermittelt wurde (Yavachev & Ivanov, 1988). Zusätzlich wurde noch eine weitere Variante des modifizierten *leuQ*-Gens erstellt, mit dem Unterschied, dass die Sequenz der Anticodon-Schleife in dieser Variante nicht durch die Konsensussequenz ausgetauscht wurde.

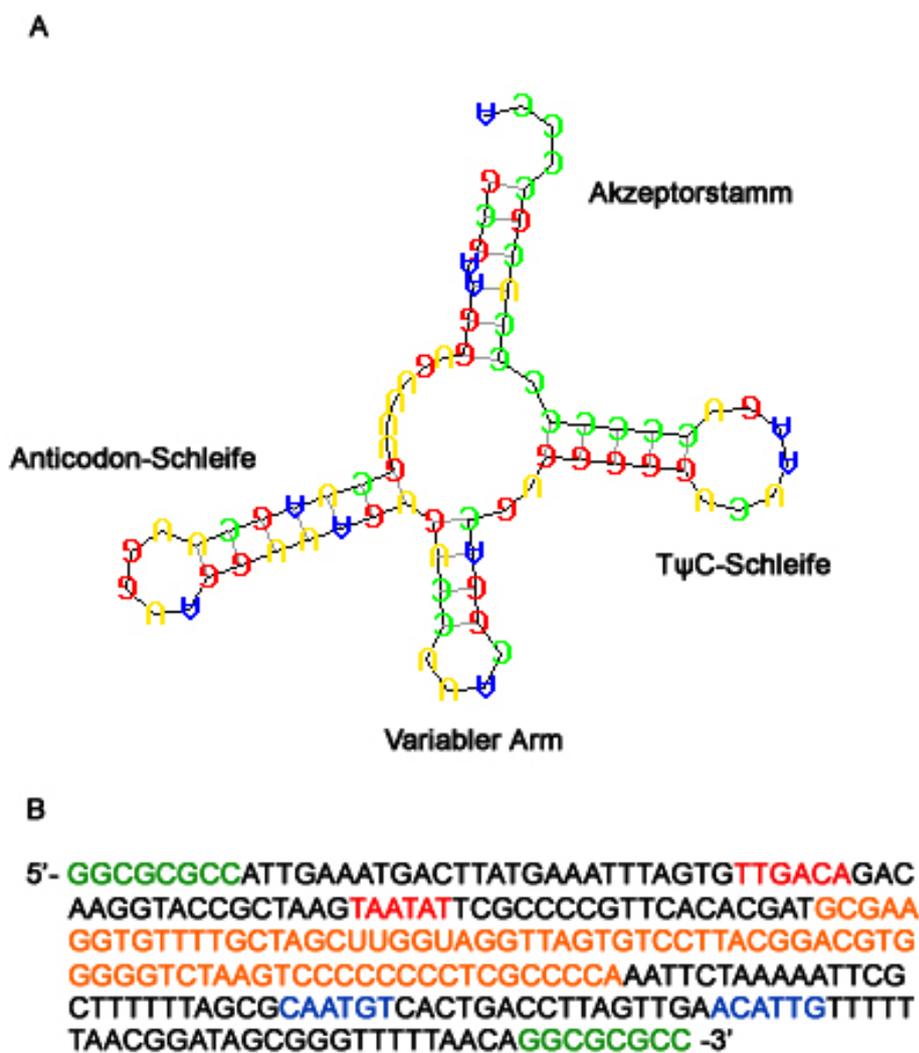
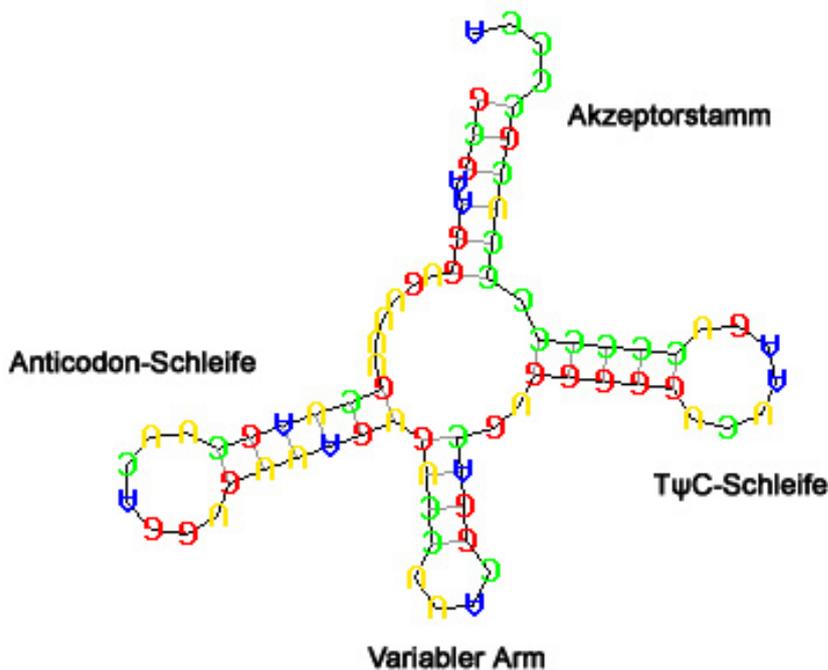


Abbildung 8.12: A: Vorhergesagte tRNA-Struktur für die Leucin-tRNA-RNAI mit dem Tool RNAsnapes (Steffen *et al.*, 2006) mit einer Wahrscheinlichkeit von 98 %. B: Sequenz der Transkriptionseinheit des Leucin-tRNA-RNAI Gens; grün -> Ascl-Schnittstellen, rot -> Promotorelemente, orange -> Leucin-tRNA-RNAI-Gen, blau -> Terminationselemente

Es wurden in dieser Arbeit somit zwei Varianten eines modifizierten *leuQ*-Gens erstellt, die sich in der Sequenz ihrer Anticodon-Schleife unterscheiden: Die erste Variante, genannt Leucin-tRNA-RNAI, enthält die Konsensussequenz (5'-TTGGTAG-3') als Anticodon-Schleife und die zweite Variante, genannt Leucin-tRNA-WT, besitzt die Wildtyp-Sequenz der Anticodon-Schleife (5'-TTCAGGT-3'). In den Abbildungen 8.12 und 8.13 sind die Transkriptionseinheiten der modifizierten Leucin-tRNAs mit den jeweiligen vorhergesagten Strukturen dargestellt. Die Strukturvorhersage wurde mit der Software „RNAshapes“ (Steffen *et al.*, 2006) durchgeführt.

A



B

```

5'- GGCGCGCCATTGAAATGACTTATGAAATTTAGTGTTGACAGAC
  AAGGTACCGCTAAGTAATATTCGCCCGTTACACGATGCGAA
  GGTGTTTTGCTAGCTTCAGGTGTTAGTGTCTTACGGACGTG
  GGGGTCTAAGTCCCCCCCCTCGCCCCAAATTCTAAAAATTCG
  CTTTTTAGCGCAATGTCACCTGACCTTAGTTGAACATTGTTTTT
  TAACGGATAGCGGGTTTTTAACAGGCGCGCC -3'

```

Abbildung 8.13: A: Vorhergesagte tRNA-Struktur für die Leucin-tRNA-WT mit dem Tool RNAshapes (Steffen *et al.*, 2006) mit einer Wahrscheinlichkeit von 94 %. B: Sequenz der Transkriptionseinheit des Leucin-tRNA-WT Gens; grün -> AscI-Schnittstellen, rot -> Promotorelemente, orange -> Leucin-tRNA-RNAI-Gen, blau -> Terminationselemente

Die Software „RNAshapes“ berechnet die Faltungsmöglichkeiten von RNA-Molekülen und den dazugehörigen Wahrscheinlichkeiten, basierend auf der Minimierung der zur Faltung notwendigen Energie (Steffen *et al.*, 2006). Die vorhergesagte Faltung, der in Abbildung 8.12 dargestellten Leucin-tRNA-RNAI entspricht der typischen tRNA-Kleeblattstruktur mit dem Unterschied, dass die DHU-Schleife fehlt. Die Wahrscheinlichkeit für diese Struktur beträgt 98 %.

Für die in Abbildung 8.13 gezeigte Leucin-tRNA-WT verhält es sich ähnlich: Hier wird, mit einer Wahrscheinlichkeit von 94 %, ebenfalls eine Kleeblattstruktur mit fehlender DHU-Schleife vorhergesagt.

8.5.3 Vorgehen zur Integration der Leucin-tRNA-Transkriptionseinheiten in das *E. coli* Chromosom

Für die Konstruktion der Transkriptionseinheit der modifizierten Leucin-tRNA wurde die Sequenz des *asnT*-Gen gegen die veränderten *leuQ*-Sequenzen ausgetauscht, so dass zwei Varianten einer modifizierten Leucin-tRNA-Transkriptionseinheit entstanden sind: Die erste Variante besteht aus der *asnT*-Transkriptionseinheit mit dem *Leucin-tRNA-RNAI*-Gen und die zweite Variante besteht aus der *asnT*-Transkriptionseinheit mit dem *Leucin-tRNA-WT*-Gen.

Im weiteren Vorgehen sollen die Leucin-tRNA-Transkriptionseinheiten getrennt in das DH5 α Chromosom integriert werden, so dass am Ende zwei DH5 α -Derivate entstehen: DH5 α -*Leucin-tRNA-RNAI*⁺ und DH5 α -*Leucin-tRNA-WT*⁺.

Die Integration der Leucin-tRNA-Transkriptionseinheiten in das *E. coli* Chromosom soll mit Hilfe von Tn5 Transposon-Mutagenese erfolgen. Dazu wurden zunächst die Vektoren pIF11 und pIF13 konstruiert (Abbildung 8.14). Der Vektor pIF11 beinhaltet den temperatursensitiven Replikationsursprung *OriR101 E. coli repA101ts* und das *bla*-Gen für eine Ampicillinresistenz. Er ist durch PCR-Amplifikation und anschließender Ligation aus dem Vektor pKD20 entstanden. Im nächsten Schritt wurde der Vektor pIF12 konstruiert. Dieser basiert auf pIF11 und enthält zusätzlich ein Transposon-Element, welches ein, mit *loxP*-Sequenzen flankiertes, *nptII*-Gen kodiert und ein von *AscI*-Schnittstellen eingerahmtes Sequenzstück (*egfp*) enthält. Das Transposon-Element wurde ebenfalls mittels PCR amplifiziert und anschließend *blunt-end* in die *EcoRV*-Schnittstellen des Plasmids pIF11 integriert. Das auf dem Transposon-Element enthaltende *nptII*-Gen vermittelt Resistenzen gegenüber den Antibiotika Neomycin und Kanamycin. Der Vektor pIF13 be-

sitzt als Grundgerüst das Plasmid pIF12 und zusätzlich das Gen einer hyperaktiven Tn5 Transposase unter Kontrolle des *P_{trc}/lacI* Promotersystems. Das Gen der hyperaktiven Tn5 Transposase wurde unter Verwendung von Primern mit *SacI*- und *BamHI*-Erkennungssequenzen aus dem Vektor pHB14 (Bednarz, 2012) mittels PCR amplifiziert und über *SacI*- und *BamHI*-Schnittstellen nach pIF12 kloniert.

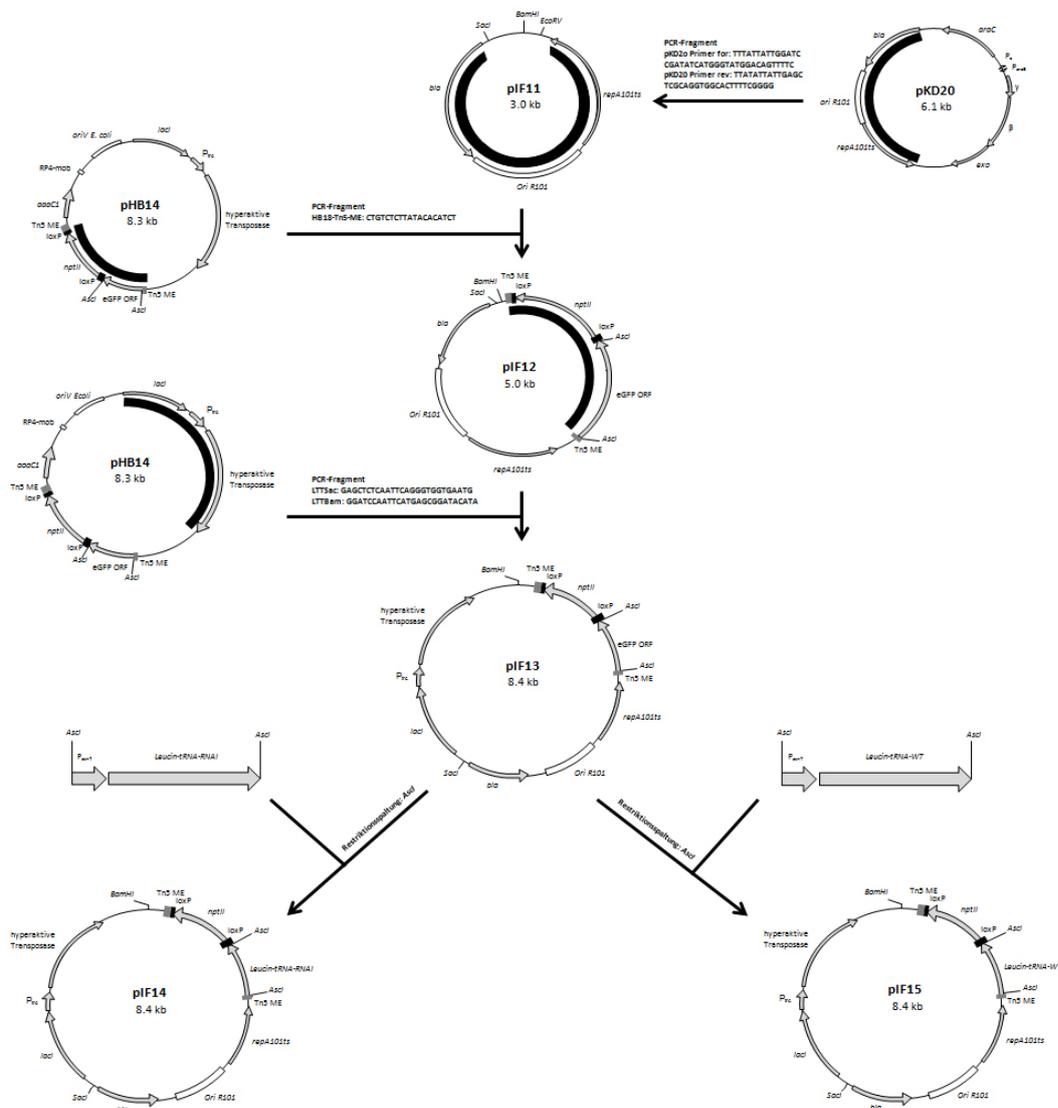


Abbildung 8.14: Konstruktionsschema für die Plasmidvektoren pIF14 und pIF15 zur Durchführung einer transposonbasierten Insertion der tRNA-Transkriptionseinheiten. Details siehe Text.

Die Transkriptionseinheiten der Leucin-tRNA-RNAi und der Leucin-tRNA-WT wurden von der Firma Entelechon (Bad Abbach) vollständig synthetisiert. Anschließend sind sie über die *AscI*-Schnittstellen in jeweils eine Kopie des Vektors pIF13 integriert worden, so dass am Ende die Vektoren pIF14 und pIF15 vorlagen (Abbildung 8.14). Der Vektor pIF14 besitzt das Plasmid pIF13 als Rückgrat und die Transkriptionseinheit der Leucin-tRNA-RNAi. Der Vektor pIF15 enthält ebenfalls das Plasmid pIF13 als Rückgrat

und die Transkriptionseinheit der Leucin-tRNA-WT.

Um die Leucin-tRNA-Transkriptionseinheiten in das Chromosom integrieren zu können, wird neben den Vektoren pIF14 und pIF15 noch ein weiteres Plasmid benötigt, welches das *Cre*-Gen besitzt, das für ein Enzym mit dem Namen Cre-Rekombinase kodiert (Abbildung 8.15). Dazu wurde das *Cre*-Gen mittels PCR aus dem Vektor pHB15 (Bednarz, 2012) amplifiziert und über *SmaI*-Schnittstellen in den Vektor pK18mob (Kirchner & Tauch, 2003) hinter den *lacZ*-Promotor kloniert. Das *Cre*-Gen steht somit nun unter der Kontrolle des *lacZ*-Promotors (pIF16). Im darauffolgenden Schritt wurde das *Cre*-Gen inklusive *lacZ*-Promotor mit Primern, die *EcoRV*-Schnittstellen als Extension besitzen, unter Anwendung von PCR amplifiziert. Gleichzeitig wurde der Vektor pIF11 mit dem Restriktionsenzym *EcoRV* linearisiert. Nach der Ligation entstand das Plasmid pIF17, welches den Vektor pIF11 als Rückgrat und das *Cre*-Gen unter Kontrolle des *lacZ*-Promotors enthält.

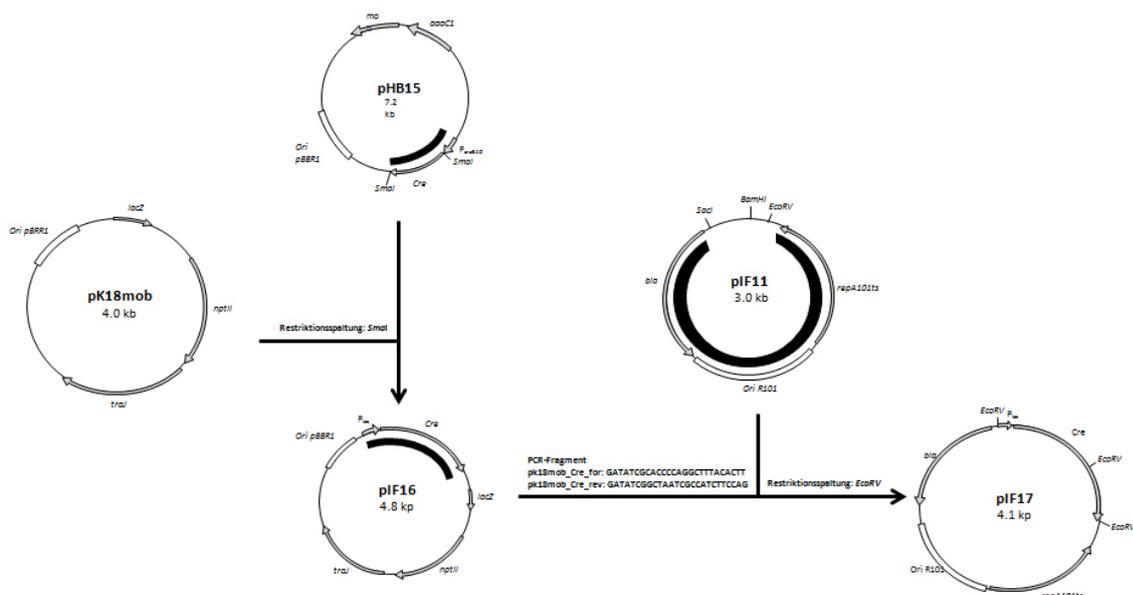


Abbildung 8.15: Konstruktionsschema für die Exzision des Selektionsmarkers durch das Enzym Cre-Rekombinase. Details siehe Text.

Der Vektor pIF17 wird im Anschluss an die Integration der Transkriptionseinheit in das Chromosom durch die Transposon-Mutagenese benötigt, da die darauf kodierte Cre-Rekombinase über die *loxP*-Schnittstellen das *nptII*-Gen herausschneiden kann. Das Entfernen der Antibiotikaresistenz ist wichtig im Hinblick auf eine spätere industrielle Nutzung der Plasmid-Produktionsstämme für eine medizinischen Anwendung.

Die im folgenden beschriebenen Analysen wurden von Sebastian Wiebe im Rahmen seiner Masterarbeit durchgeführt. Zunächst wurden dabei die Vektoren pIF14, pIF15

und pIF17 mittels Sequenzierung auf ihre Integrität überprüft. Nachdem auf diese Weise sichergestellt worden war, dass die Vektorkonstruktionen erfolgreich sind, wurden sie jeweils in eine DH5 α -Zelle transformiert, wobei drei verschiedene Stämme (DH5 α -pIF14, DH5 α -pIF15 und DH5 α -pIF17) entstanden. Anschließend wurde in den Stämmen DH5 α -pIF14 und DH5 α -pIF15 jeweils die Transposon-Mutagenese durch Induzierung der hyperaktiven Transposase ausgelöst. Dabei ist das auf dem Transposonelement enthaltene *Leucin-tRNA-RNAI*-Gen bzw. das *Leucin-tRNA-WT*-Gen in das DH5 α -Chromosom inseriert. Mit einer Inkubation bei 30°C ist das entsprechende Plasmid aus der Zelle entfernt worden und die Stämme DH5 α -*Leucin-tRNA-RNAI*⁺ bzw. DH5 α -*Leucin-tRNA-WT*⁺ sind entstanden.

Im nächsten Schritt ist überprüft worden, ob die inserierten modifizierten Leucin-tRNA-Gene und ihre Produkte einen Einfluss auf die Plasmidproduktion haben. Dazu wurden die Stämme DH5 α -*Leucin-tRNA-WT*⁺ und DH5 α -*Leucin-tRNA-RNAI*⁺ bezüglich ihres Plasmidgehaltes untersucht und mit dem Stamm DH5 α verglichen.

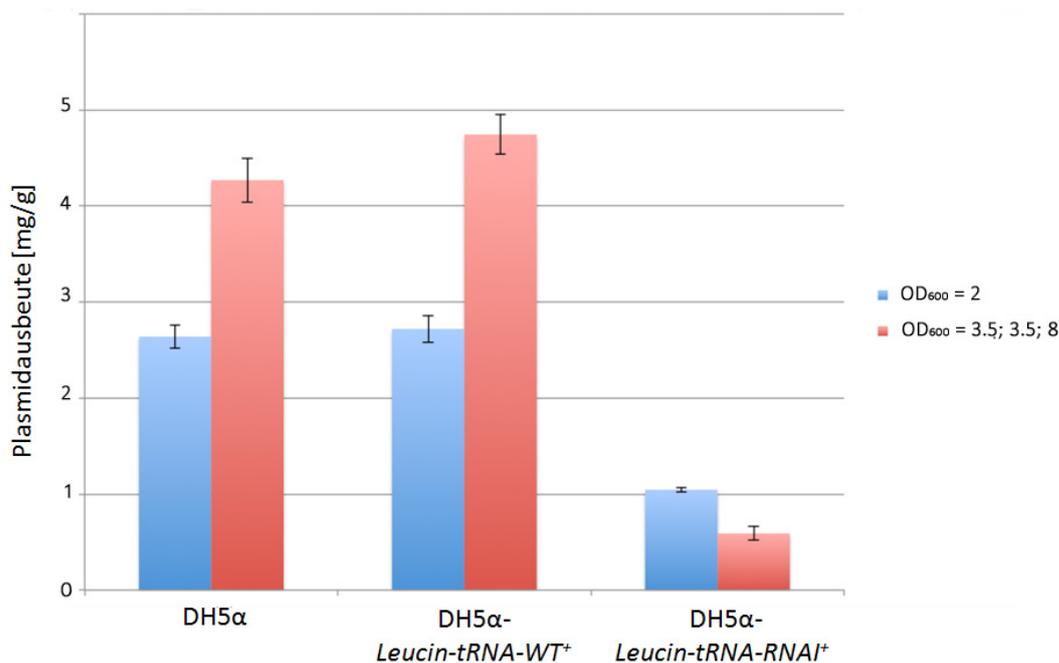


Abbildung 8.16: Ergebnisse zur Untersuchung des Einflusses der modifizierten Leucin-tRNA-Gene in den Stämmen DH5 α -*Leucin-tRNA-WT*⁺ und DH5 α -*Leucin-tRNA-RNAI*⁺ auf die Plasmidproduktion (Daten erhoben von Sebastian Wiebe). Die Plasmidausbeute wurde für zwei verschiedene Zeitpunkte in mg pro g Biomasse bestimmt. Die OD₆₀₀ zum ersten Zeitpunkt der Zellernte (blau) betrug bei allen drei Stämmen den Wert 2. Zum zweiten Erntezeitpunkt (rot) besaßen die Stämme DH5 α und DH5 α -*Leucin-tRNA-WT*⁺ eine optische Dichte von 3,5 und der Stamm DH5 α -*Leucin-tRNA-RNAI*⁺ eine optische Dichte von 8.

Für diese Untersuchungen wurden die Zellen an zwei verschiedenen Zeitpunkten geerntet und die Plasmidausbeute gemessen. Zum Erntezeitpunkt der ersten Messung besaßen alle drei Stämme eine optische Dichte von 2. Zum zweiten Erntezeitpunkt betrug die optische Dichte der Stämme DH5 α und DH5 α -*Leucin-tRNA-WT*⁺ einen Wert von 3.5, während der Stamm DH5 α -*Leucin-tRNA-RNAI*⁺ eine OD₆₀₀ von 8 aufwies. Die dazu von Sebastian Wiebe ermittelten Daten zur Plasmidausbeute sind in Abbildung 8.16 dargestellt.

Man sieht, dass die Plasmidproduktion in dem Stamm DH5 α -*Leucin-tRNA-RNAI*⁺ im Vergleich zu DH5 α sowohl bei einer Ernte-OD₆₀₀ von 2, als auch bei Ernte-OD₆₀₀ von 8 deutlich gesunken ist. Das bedeutet, die von dem *Leucin-tRNA-RNAI*-Gen kodierten tRNA-Moleküle haben einen negativen Effekt bezüglich der Plasmidproduktion. In dem Stamm DH5 α -*Leucin-tRNA-WT*⁺ konnte man im Vergleich zu DH5 α zu beiden Zeitpunkten eine Zunahme der Plasmidproduktion beobachten, wobei diese zum ersten Erntezeitpunkt (Ernte-OD₆₀₀ = 2) sehr gering ausfällt. Beim zweiten Erntezeitpunkt (Ernte-OD₆₀₀ = 3.5) ist die Zunahme der Plasmidproduktion deutlicher. Die von dem *Leucin-tRNA-WT*-Gen kodierten tRNA-Moleküle fördern somit die Plasmidproduktion in DH5 α . Es soll an dieser Stelle jedoch darauf hingewiesen werden, dass es sich hierbei um erste Versuche zur Untersuchung des Einflusses der modifizierten tRNAs auf die Plasmidproduktion in den Stämmen DH5 α -*Leucin-tRNA-WT*⁺ und DH5 α -*Leucin-tRNA-RNAI*⁺ handelt, welche in weiteren Experimenten wiederholt und bestätigt werden sollen.

METABOLOMANALYSEN

9.1 UNTERSUCHUNGEN ZUR METABOLITZUSAMMENSETZUNG

Betrachtet man alle bekannten Faktoren, die die Plasmidreplikation beeinflussen können, stellt man fest, dass es zwei zentrale Ebenen gibt, die den Plasmidgehalt einer Zelle bestimmen. Zum einen ist das die Ebene der Plasmidreplikationskontrolle und zum anderen die Verfügbarkeit von Nukleotiden und deren Vorstufen. Aus diesem Grund soll in dieser Arbeit neben der Plasmidreplikationskontrolle ein Metabolitprofil erstellt werden, um die Metabolitzusammensetzung in einem plasmidfreien und zwei plasmidtragenden Stämmen zu untersuchen.

Hiller *et al.* (2007) stellen in Ihrem Artikel die Bedeutung von *metabolic profiling*-Analysen heraus, in dem sie sagen, die Messung der intrazellulären Metabolitkonzentrationen (*metabolic profiling*) ist ein notwendiges Hilfsmittel, um biologische Prozesse verstehen und optimieren zu können (Hiller *et al.*, 2007). Die in dieser Arbeit durchgeführten Analysen sollen daher helfen, den Plasmidproduktionsprozess innerhalb der Zelle besser nachvollziehen und optimieren zu können.

Die dabei verwendeten Stämme sind der plasmidfreie *E. coli* Stamm DH5 α sowie der *low copy* plasmidtragende DH5 α -pSUP 201-3 und der *high copy* plasmidtragende DH5 α -pCMV-GFP. Als Kohlenstoffquelle wurde für diese Analysen Glycerin verwendet, da hierbei weniger Acetat produziert wird, als bei der Verwendung von Glukose (El-Mansi & Holms, 1989). Die Analysen wurden für vier biologische Replikate mit jeweils einem technischen Replikat durchgeführt. Nach Zellaufschluss wurden die Metabolite mittels eiskaltem Methanol extrahiert und anschließend mit MSTFA (N-Methyl-N-(trimethylsilyl) trifluoroacetamid) derivatisiert. Bei der Derivatisierung wird die Flüchtigkeit von Substanzen erhöht, so dass auf diese Weise Verbindungen, die unter normalen Umständen nicht flüchtig sind, für die GC-Analytik zugänglich gemacht werden. Dies geschieht durch den Austausch von polaren Gruppen (-OH, -NH, -SH) durch weniger polare Gruppen (Silyl-, Acyl- oder Alkylgruppen), was typischerweise für Aminosäuren, organische Säuren oder Zucker notwendig ist. Ein weiterer Vorteil derivatisierter

Metabolite ist ihre dadurch verbesserte thermische Stabilität. Man muss allerdings dazu sagen, dass längst nicht alle Metabolite verdampfbar gemacht werden können. Manche Substanzen lassen sich aufgrund ihrer Größe auch durch Derivatisierung nicht verflüchtigen. Die anschließende Messung der derivatisierten Proben erfolgte durch GC-MS. In der Abbildung 9.1 ist jeweils ein exemplarisches Gaschromatogramm für den plasmidfreien Stamm DH5 α sowie für die plasmidtragenden Stämme DH5 α -pSUP 201-3 (*low copy Plasmid*) und DH5 α -pCMV-GFP (*high copy Plasmid*) dargestellt.

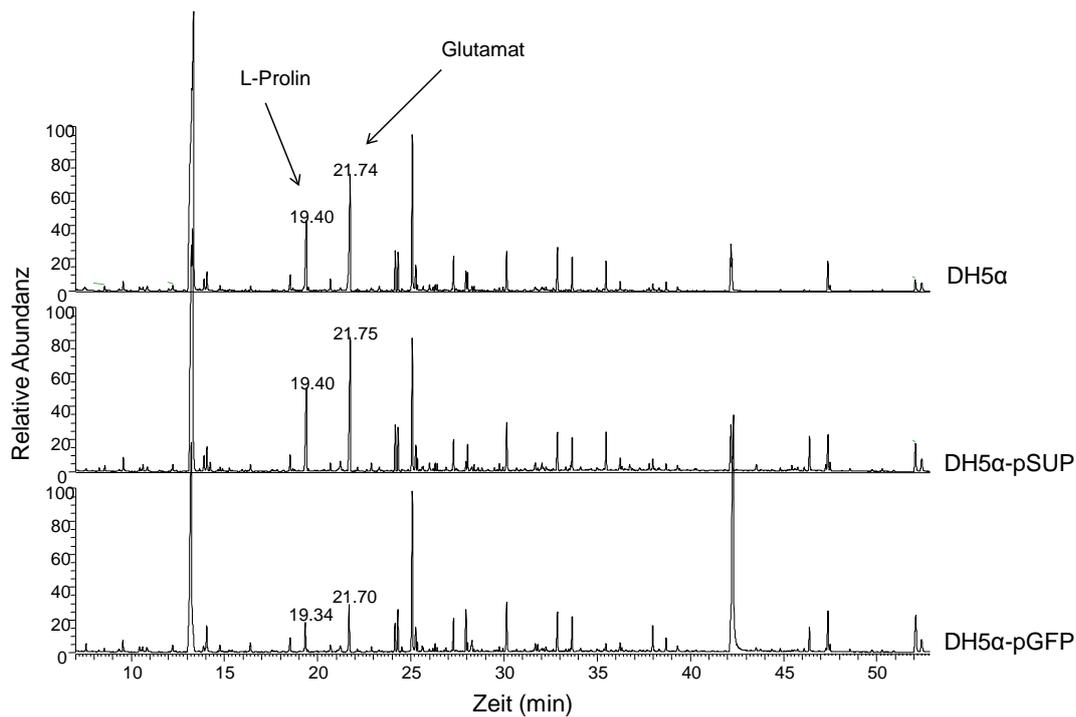


Abbildung 9.1: Untersuchung der Metabolitzusammensetzung für DH5 α sowie für DH5 α -pSUP 201-3 (*low copy Plasmid*) und DH5 α -pCMV-GFP (*high copy Plasmid*) mittels GC-MS. (Abbildung wurde aus der Xcalibur-Software exportiert.)

Auf der x-Achse ist die Retentionszeit aufgetragen und auf der y-Achse ist die relative Abundanz dargestellt. Vergleicht man die Chromatogramme untereinander, fällt auf, dass das Muster der Peaks ähnlich ist, sie sich jedoch in ihren relativen Höhen unterscheiden. Die Skalierung der y-Achse ist bei der Darstellung in Abbildung 9.1 so gewählt, dass man die Peaks der einzelnen Chromatogramme direkt untereinander vergleichen kann. Betrachtet man z. B. in jedem Chromatogramm den Peak mit einer Retentionszeit von ~ 21.7 min, so sieht man, dass die Stämme DH5 α und DH5 α -pSUP 201-3 eine ungefähr gleich große intrazelluläre Glutamatkonzentration besitzen und der Stamm DH5 α -pCMV-GFP ungefähr halb soviel. Im Fall von Prolin (Retentionszeit von

~ 19,4 min) lassen sich ebenfalls Unterschiede erkennen. Am wenigsten Prolin befindet sich in den DH5 α -pCMV-GFP-Zellen, der plasmidfreie Stamm enthält etwas mehr und die größte Menge an Prolin findet man in DH5 α -pSUP 201-3. Die Identifizierung der zuvor erwähnten Peaks erfolgte durch Datenbankvergleiche, durchgeführt mit der MeltDB-Software. Dabei werden die für jeden Peak innerhalb eines Chromatogramms durchgeführten Massenspektren mit Spektren bereits identifizierter Metabolite verglichen, welche in speziellen Datenbanken hinterlegt sind. Auf diese Weise lassen sich einige Metabolite zuordnen. Für eine Vielzahl an Substanzen gibt es allerdings noch keinen Datenbankeintrag, so dass sich derzeit noch nicht alle Verbindungen identifizieren lassen. Es konnten daher in den durchgeführten Messungen Metabolite detektiert werden, deren Konzentrationen sich zwar innerhalb der drei Stämme unterschieden, welche sich jedoch nicht zuordnen ließen.

Da sich diese Arbeit mit der Plasmidreplikation in DH5 α beschäftigt, liegt das Hauptaugenmerk auf der Untersuchung der annotierten und somit bekannten Metabolite, insbesondere auf Unterschieden in den Metaboliten, die für die Nukleotidbiosynthese relevant sind. Um mögliche Unterschiede detektieren zu können, wurden die Chromatogramme der GC-MS-Messungen unter Verwendung der MeltDB-Software genauer analysiert. Dabei wurden zunächst alle annotierten Metabolite über ihre Peakfläche quantifiziert und anschließend auf die den Proben zugegebene Referenzsubstanz Ribitol und das jeweilige zelluläre Trockengewicht normalisiert. Die Normalisierung und die darauffolgenden auswertenden Analysen wurden ebenfalls mit Hilfe der MeltDB-Software durchgeführt.

Eine Übersicht aller bekannten Metabolite, die in den analysierten Chromatogrammen detektiert wurden, befindet sich als Datei auf der beigefügten CD in Form einer *Heatmap*. Innerhalb der annotierten Metabolite sollen die Metabolite Glukose, Glukose-6-phosphat, Fruktose-6-phosphat, Glycerinaldehyd-3-phosphat, 2-Phosphoglycerat, 3-Phosphoglycerat, Phosphoenolpyruvat, Pyruvat, Citrat, cis-Aconitat, 2-Oxoglutarat, Succinat, Fumarat, Malat, Ribose, Ribose-5-phosphat, Xylose, Glutamat, Glutamin, Glycin und Aspartat im Folgenden näher analysiert werden.

Citratzyklus

Das Vorkommen der Metabolite des Citratzyklus ist in Abbildung 9.2 durch eine *Heatmap* dargestellt. Die einzelnen Metabolitkonzentrationen werden dabei relativ zu der maximal gemessenen Konzentration dieses Metaboliten dargestellt, wobei Blau für eine geringe Konzentration steht, Rot für eine hohe Konzentration und Schwarz für nicht-

detektiert. Es fällt auf, dass das Vorkommen der einzelnen Metabolite in den Stämmen DH5 α und DH5 α -pSUP 201-3 sehr ähnlich ist, sich jedoch deutlich von dem in DH5 α -pCMV-GFP unterscheidet. Die Konzentrationen von Citrat, α -Ketoglutarat (= 2-Oxoglutarat) und Succinat sind in dem Stamm DH5 α -pCMV-GFP wesentlich höher, als in den beiden anderen Stämmen. Für Fumarat, Malat und cis-Aconitat verhält es genau sich umgekehrt, hier weist der *high copy* Stamm eine geringere Menge auf bzw. cis-Aconitat ist gar nicht detektierbar.

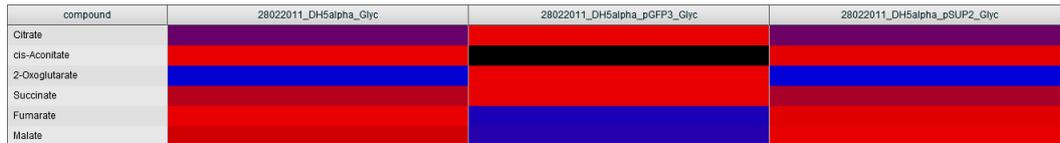


Abbildung 9.2: Darstellung der relativen Abundanzen einiger Metabolite des Citratzyklus in Form einer Heatmap für DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP. Eine ansteigende Konzentration ist durch die Farbabstufungen von blau (niedrig) bis rot (hoch) gekennzeichnet.

Um diese Unterschiede genauer zu betrachten ist in Abbildung 9.3 die relative Abundanz in einem Säulendiagramm dargestellt.

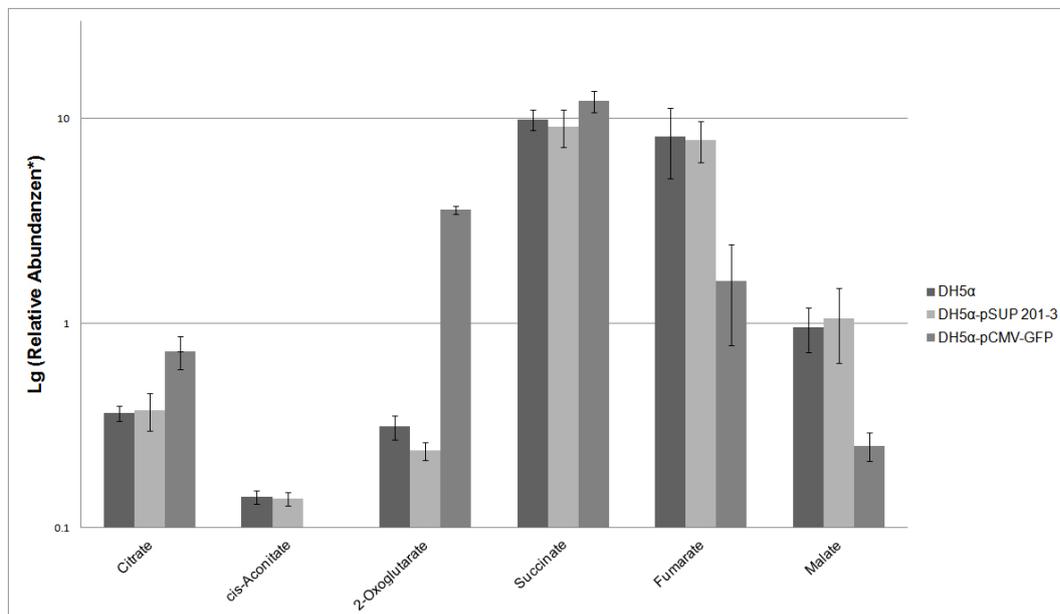


Abbildung 9.3: Darstellung der logarithmierten relativen Abundanzen einiger Metabolite des Citratzyklus in Form eines Säulendiagramms für die Stämme DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP.

* Die Berechnung der relativen Abundanz erfolgte mit Hilfe der MeltDB-Software durch die Quantifizierung der Metabolite über ihre Peakfläche mit einer anschließenden Normalisierung auf den internen Standard Ribitol und das zu Beginn eingewogene Zellrockengewicht.

Man sieht, dass die α -Ketoglutarat-Konzentration innerhalb von DH5 α -pCMV-GFP sehr viel größer als in den beiden anderen Stämmen ist. Die Citrat- und

Succinat-Konzentrationen sind ebenfalls erhöht, allerdings weniger stark. Die Fumaratkonzentration in dem *high copy* Stamm ist im Vergleich mit DH5 α und DH5 α -pSUP 201-3 deutlich geringer, wobei die Unterschiede in der Malatmenge geringer ausfallen. Bei der Betrachtung von cis-Aconitat fällt auf, dass dieser Metabolit in DH5 α -pCMV-GFP nicht detektiert wurde und in den beiden anderen Stämmen nur in sehr geringer Konzentration vorliegt.

Glykolyse

Als nächstes sollen die Metabolite der Glykolyse näher analysiert werden. Anhand der *Heatmap* in Abbildung 9.4 wird deutlich, dass die Stämme DH5 α und DH5 α -pSUP 201-3, im Gegensatz zum Citratzyklus, hier kein konstant ähnliches Metabolitkonzentrationslevel mehr zeigen.

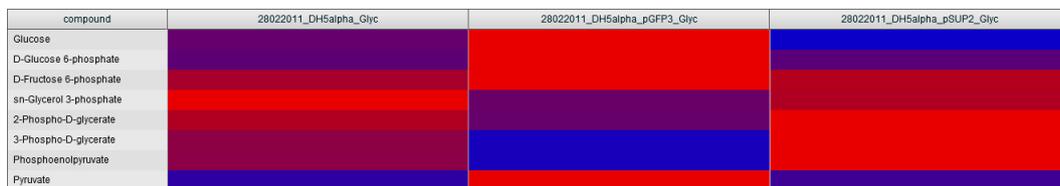


Abbildung 9.4: Darstellung der relativen Abundanzen einiger Metabolite der Glykolyse in Form einer *Heatmap* für die Stämme DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP. Eine ansteigende Konzentration ist durch die Farbabstufungen von blau (niedrig) bis rot (hoch) gekennzeichnet.

Um die einzelnen relativen Unterschiede genauer betrachten zu können, wurde die relative Abundanz der einzelnen Metabolite erneut in einem Säulendiagramm dargestellt. Man kann beobachten, dass die relative Konzentrationen von Glukose-, Glukose-6-phoshat, Fruktose-6-phoshat und Pyruvat in DH5 α -pCMV-GFP am höchsten sind, wobei die Unterschiede in Fruktose-6-phoshat am geringsten ausfallen. Im Fall der Metabolite Glycerinaldehyd-3-phosphat, 2-Phosphoglycerat, 3-Phosphoglycerat, Phosphoenolpyruvat besitzt der *high copy* plasmidtragende Stamm im Vergleich stets die geringste Konzentration. Dabei fällt auf, dass die Glycerinaldehyd-3-phosphat-Konzentration im DH5 α am größten ist; sie nimmt somit mit der Plasmidkopienzahl ab. Des Weiteren ist anzumerken, dass DH5 α -pSUP 201-3 die geringste intrazelluläre Glukosemenge besitzt. Insgesamt betrachtet fallen die Unterschiede bezüglich der Glykolyse-Metabolite zwischen DH5 α und DH5 α -pSUP 201-3 jedoch nicht so deutlich aus, wie im Vergleich zu DH5 α -pCMV-GFP.

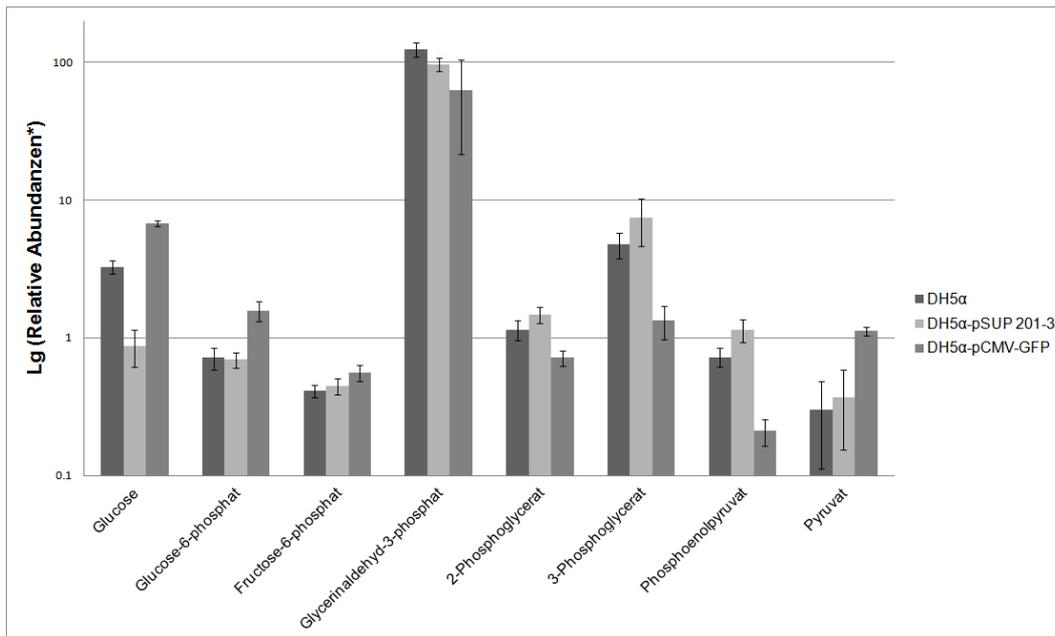


Abbildung 9.5: Darstellung der logarithmierten relativen Abundanzen einiger Metabolite der Glykolyse in Form eines Säulendiagramms für die Stämme DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP.

* Die Berechnung der relativen Abundanz erfolgte mit Hilfe der MeltDB-Software durch die Quantifizierung der Metabolite über ihre Peakfläche mit einer anschließenden Normalisierung auf den internen Standard Ribitol und das zu Beginn eingewogene Zelltrockengewicht.

Nukleotidbiosynthese

Da das Hauptaugenmerk dieser Arbeit auf der Plasmidreplikation liegt, sollen im Folgenden Abschnitt die für die Nukleotidbiosynthese relevanten Metabolite genauer betrachtet werden. Dazu gehören die Fünffachzucker Ribose, Ribose-5-Phosphat und Xylose sowie die an der Synthese der Basen beteiligten Verbindungen Glutamat, Glutamin, Aspartat und Glycin.

Zunächst sollen die Fünffachzucker untersucht werden. Die Abbildung 9.6 zeigt die relativen Metabolitkonzentrationen der Fünffachzucker in Form einer *Heatmap*.

compound	28022011_DH5alpha_Glyc	28022011_DH5alpha_pGFP3_Glyc	28022011_DH5alpha_pSUP2_Glyc
D-Ribose	Red	Blue	Red
D-Ribose 5-phosphate	Red	Blue	Red
D-Xylose	Red	Blue	Red

Abbildung 9.6: Darstellung der relativen Abundanzen von Ribose, Ribose-5-Phosphat und Xylose in Form einer Heatmap für die Stämme DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP. Eine ansteigende Konzentration ist durch die Farbabstufungen von blau (niedrig) bis rot (hoch) gekennzeichnet.

Es wird deutlich, dass die Ribose-5-phosphat-Konzentration in dem Stamm DH5 α -pCMV-GFP am höchsten ist, gefolgt von DH5 α -pSUP 201-3. Der plasmidfreie Stamm besitzt am wenigsten freie Ribose-5-phosphat-Moleküle, weist jedoch die höchsten Ribose- und Xylose-Konzentrationen auf. Für Ribose und Xylose verhält es sich umgekehrt: Für

DH5 α -pCMV-GFP wurde dabei die geringste Konzentration gemessen und für DH5 α die höchste Konzentration. Wie ausgeprägt diese Unterschiede lässt sich anhand von Abbildung 9.7 erkennen.

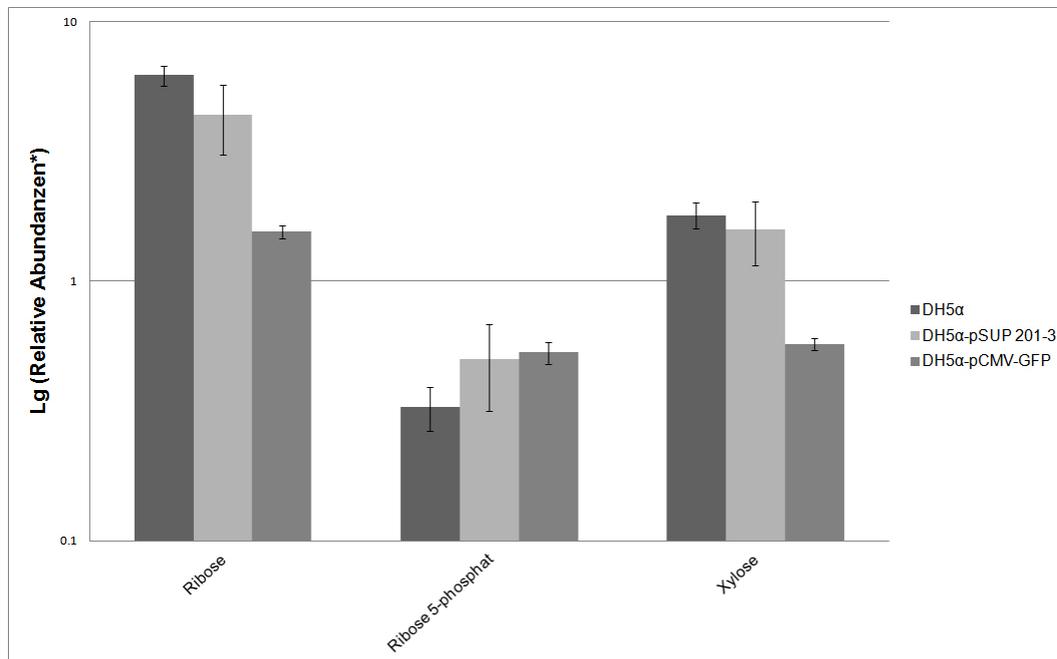


Abbildung 9.7: Darstellung der logarithmierten relativen Abundanzen von Ribose, Ribose-5-Phosphat und Xylose in Form eines Säulendiagramms für die Stämme DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP. * Die Berechnung der relativen Abundanz erfolgte mit Hilfe der MeltDB-Software durch die Quantifizierung der Metabolite über ihre Peakfläche mit einer anschließenden Normalisierung auf den internen Standard Ribitol und das zu Beginn eingewogene Zelltrockengewicht.

Man sieht, dass der DH5 α -pCMV-GFP im Vergleich deutlich weniger Ribose und Xylose besitzt. Die Unterschiede in der Ribose-5-phosphat-Konzentration sind dagegen nicht so stark ausgeprägt. Man kann festhalten, dass die Konzentration in DH5 α wesentlich geringer ist. Die plasmidtragenden Stämme weisen ein ähnliches Konzentrationslevel auf, wobei DH5 α -pCMV-GFP man eine minimal höhere Konzentration besitzt. Neben den Fünffachzuckern sind die Metabolite Glutamat, Glutamin, Aspartat und Glycin von besonderem Interesse, da sie an der Synthese der Purin- und Pyrimidinbasen beteiligt sind. Bei der Synthese von Purinbasen wird Glutamin in Glutamat umgewandelt, wobei die NH₂-Gruppe über eine Substitution auf die phosphorylierte Ribose übertragen wird. Aspartat und Glycin wird neben Formiat und CO₂ für den Aufbau des Purinringes benötigt. Zusätzlich ist Aspartat an dem Aufbau des Pyrimidinringes beteiligt. Die Unterschiede zwischen den drei Stämmen bezüglich dieser Substanzen sind in der Abbildung 9.8 durch eine *Heatmap* und in Abbildung 9.9 in Form eines Säulendiagramms dargestellt.



Abbildung 9.8: Darstellung der relativen Abundanzen von Glutamat, Glutamin, Aspartat und Glycin in Form einer Heatmap für die Stämme DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP. Eine ansteigende Konzentration ist durch die Farbabstufungen von blau (niedrig) bis rot (hoch) gekennzeichnet.

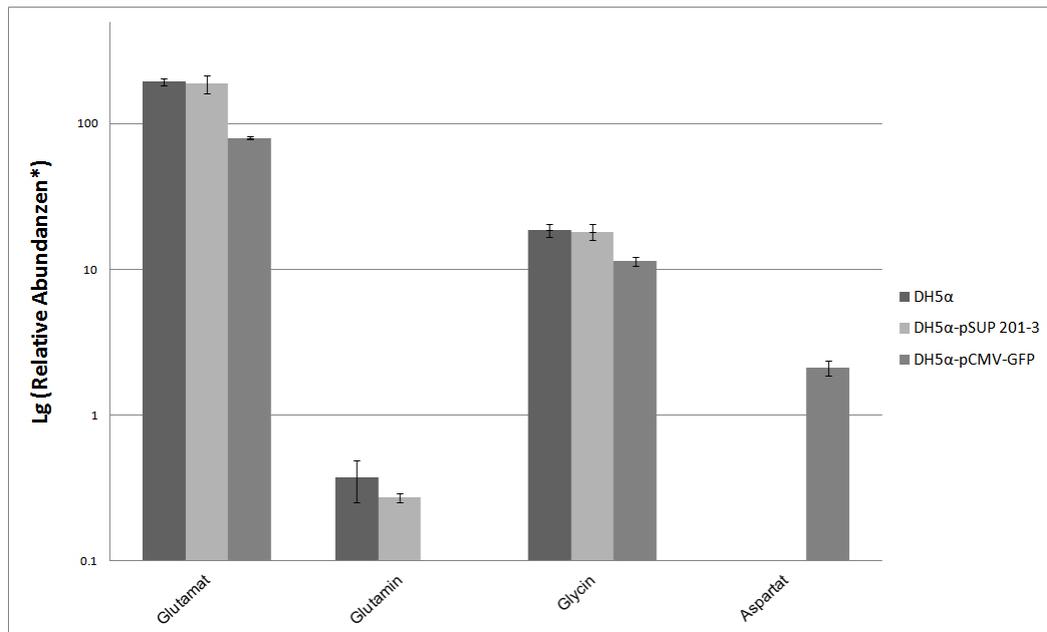


Abbildung 9.9: Darstellung der logarithmierten relativen Abundanzen von Glutamat, Glutamin, Aspartat und Glycin in Form eines Säulendiagramms für die Stämme DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP.

* Die Berechnung der relativen Abundanz erfolgte mit Hilfe der MeltDB-Software durch die Quantifizierung der Metabolite über ihre Peakfläche mit einer anschließenden Normalisierung auf den internen Standard Ribitol und das zu Beginn eingewogene Zelltrochgewicht.

Die Abbildungen zeigen, dass die gemessene Glutamatkonzentration in allen Stämmen sehr hoch ist. Der plasmidfreie Stamm besitzt mit einem minimalen Unterschied zu DH5 α -pSUP 201-3 den höchsten Glutamatgehalt. Im *high copy* Plasmid Stamm ist die Konzentration am geringsten. Die Glycinkonzentration ist insgesamt niedriger, aber auch hier wurde die geringste Konzentration in DH5 α -pCMV-GFP gemessen. In den anderen beiden Stämmen ist der Unterschied sehr gering; im DH5 α konnten, im Vergleich zu DH5 α -pSUP-3, minimal größere Mengen Glycin detektiert werden. Betrachtet man die Glutaminkonzentration von DH5 α und DH5 α -pSUP stellt man fest, dass sie sehr gering ist. Für DH5 α -pCMV-GFP konnte dieser Metabolit nicht detektiert werden. Im Gegensatz dazu konnte der Aspartat lediglich in DH5 α -pCMV-GFP gemessen werden. In der Abbildung 9.10 sind die Ergebnisse einer Hauptkomponentenanalyse für alle

annotierten Metabolite zu sehen. Eine Hauptkomponentenanalyse ist eine Methode der Statistik mit der multidimensionale Daten auf eine geringere Anzahl Variablen (die sogenannten Hauptkomponenten) reduziert und so übersichtlicher dargestellt werden kann. Bei der Reduktion werden die Hauptmuster des Datensatzes beibehalten, so dass die Hauptkomponentenanalyse eine Möglichkeit bietet, in einem sehr umfangreichen Datensatz versteckte Muster und Zusammenhänge aufzudecken. In dieser Arbeit wurde die Hauptkomponentenanalyse durchgeführt, um zu untersuchen, wie deutlich die Unterschiede in der Metabolitzusammensetzung unter Berücksichtigung aller annotierten Metabolite zwischen den Stämmen DH5 α , DH5 α -pSUP und DH5 α -pCMV-GFP sind. Anhand von Abbildung 9.10 wird deutlich, dass es Unterschiede in den einzelnen Metabolitpools gibt.

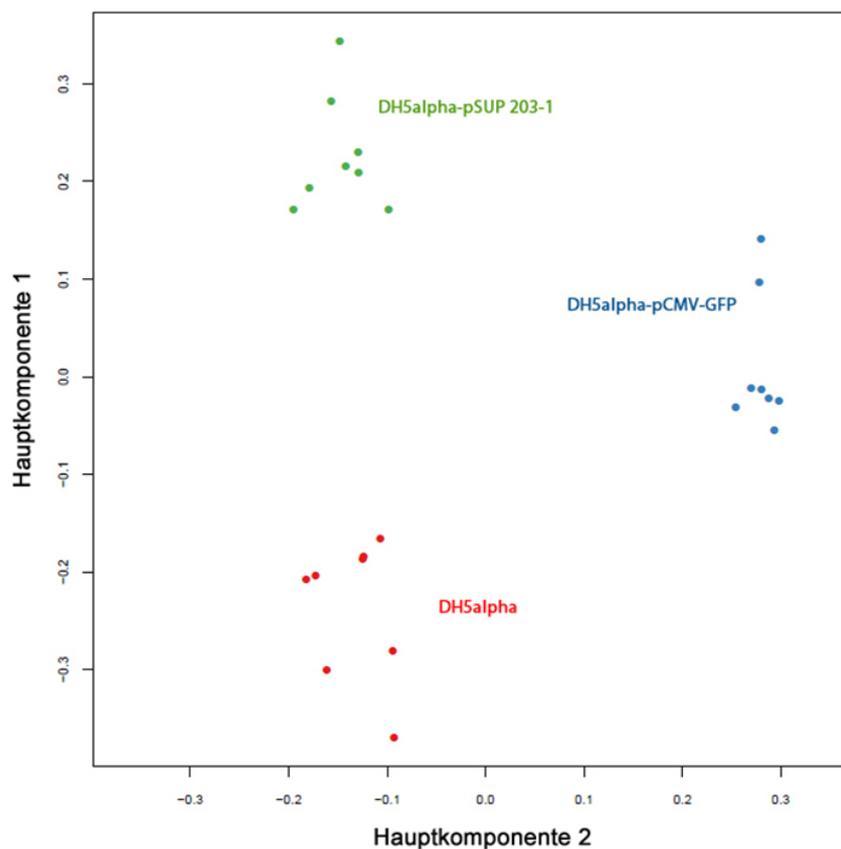


Abbildung 9.10: Hauptkomponentenanalyse aller annotierten Metabolite für DH5 α (rot), DH5 α -pSUP 201-3 (grün) und DH5 α -pCMV-GFP (blau). Die Bakterienstämme wurden auf M9-Medium (supplementiert mit FeSO₄) mit Glycerin als Kohlenstoffquelle kultiviert und in der späten exponentiellen Phase geerntet.

Für die PCA wurden alle annotierten Metabolite durchgeführt, welche über ihre Peakfläche quantifiziert und anschließend auf die, den Proben zugegebene Referenzsubstanz Ribitol und das jeweilige zelluläre Trockengewicht normalisiert worden sind. Man sieht

eine eindeutige Auftrennung der drei untersuchten Stämme bezüglich ihres Plasmidgehaltes. In rot sieht man den DH5 α , in grün den DH5 α -pSUP 201-3 und in blau den DH5 α -pCMV-GFP.

Insgesamt betrachtet zeigen die Metabolitprofil-Messungen, dass es aufgrund eines unterschiedlichen Plasmidgehaltes Unterschiede in den einzelnen Metabolitpools gibt.

^{13}C -FLUSSANALYSEN

10.1 ERMITTLUNG DER ISOTOPOMER-VERTEILUNG

Zusätzlich zu den bereits beschriebenen Metabolomanalysen wurden in dieser Arbeit ^{13}C -Flussanalysen für *E. coli* Stämme mit unterschiedlichen Plasmidproduktionsraten durchgeführt. Das Ziel dabei war, zu untersuchen, ob sich der Glukosefluss innerhalb der Zelle aufgrund verschiedener Plasmidproduktionsraten zwischen den Stämmen unterscheidet. Die Analysen erfolgten im Rahmen eines Kooperationsprojektes mit Prof. Dr. Alvaro R. Lara von der Universität *Autónoma Metropolitana-Cuajimalpa* in Mexiko. Seine Arbeitsgruppe beschäftigt sich mit der Entwicklung von Strategien, um die biotechnologische Produktion von Biopharmazeutika (z. B. Plasmid-DNA) zu optimieren. Dabei untersuchen Dr. Lara und seine Kollegen die Auswirkungen bestimmter genetischer Manipulationen in dem *E. coli* Stamm W₃₁₁₀ auf die Plasmidproduktion. Das Gesamtziel dieses Projektes sind die Identifizierung möglicher limitierender Faktoren und Schritte in der pDNA-Synthese. Außerdem sollen Ansatzpunkte für weitere genetische Manipulationen aufgefunden werden, um die Produktionsrate positiv zu beeinflussen und auf diese Weise verbesserte Produktionsstämme zu erhalten.

Für die durchgeführten Analysen wurden die Stämme W₃₁₁₀, VH₃₃, VH₃₃ Δ *recA*, VH₃₃ Δ (*recA deoR*) zunächst mit dem 6,1 kb großen *high copy* Plasmid pHN transformiert. Anschließend ist der Glukosefluss innerhalb dieser Stämme über Markierungsexperimente mit $1\text{-}^{13}\text{C}$ -markierter Glukose bestimmt worden.

Die Durchführung von ^{13}C -Flussanalysen besteht insgesamt aus drei Teilen:

1. Kultivierung auf dem markierten Substrat
2. Analyse der Markierungsmuster mittels Massenspektrometrie
3. Verarbeitung der unter Punkt 1 und 2 gewonnenen Daten mit Hilfe eines mathematischen Modells. (Quek *et al.*, 2009)

Die Stämme W₃₁₁₀-pHN, VH₃₃-pHN, VH₃₃ Δ *recA*-pHN, VH₃₃ Δ (*recA deoR*)-pHN wurden in dem Minimalmedium PD₁ mit einer $1\text{-}^{13}\text{C}$ -Glukose Konzentration von 5 g/l

kultiviert. In Tabelle 10.1 sind die jeweils die Mittelwerte der gemessenen OD_{600} zum Zeitpunkt der Zellernte für die durchgeführten Messungen dargestellt. Dabei wurde darauf geachtet, dass die Stämme innerhalb einer Messung ungefähr zur selben Wachstumsphase geerntet wurden.

Tabelle 10.1: Mittelwerte der gemessenen OD_{600} zum Zeitpunkt der Zellernte für die durchgeführten Flussmessungen

	W₃₁₁₀- pHN	VH₃₃- pHN	VH₃₃Δ<i>recA</i>- pHN	VH₃₃Δ <i>recA</i> <i>deoR</i>-pHN
Ernte- OD_{600}				
Messung 1	2.41	2.16	1.78	1.43
Ernte- OD_{600}				
Messung 2	2.98	2.98	2.83	2.15

Es wurden zwei Messungen mit je zwei biologischen und zwei technischen Replikaten durchgeführt. Die Erntezeitpunkte der ersten Messung liegen in der frühen exponentiellen Phase. Die Erntezeitpunkte der zweiten Messung wurden so gewählt, dass sie sich in der mittleren exponentiellen Wachstumsphase befinden.

Nach der Zellernte wurden die Proben mit MBDSTFA derivatisiert und anschließend mittels GC-MS vermessen. Für die Auswertung wurden anhand der Chromatogramme und den jeweiligen Massenspektren die Isotopomer-Verteilungsmuster aller detektierbaren Aminosäuren untersucht. Dabei wurde für jede gemessene Aminosäure das zugehörige Massenspektrum hinsichtlich ihrer charakteristischen Massen untersucht. Die charakteristischen Massen beschreiben für jede Aminosäure typische Fragmente, in welche sie im derivatisierten Zustand bei der Ionisierung zerfallen kann. Die prominentesten Fragmente mit ihren „Sollbruchstellen“, die bei MBDSTFA-derivatisierten Aminosäuren entstehen, sind in Abbildung 10.1 verallgemeinert dargestellt.

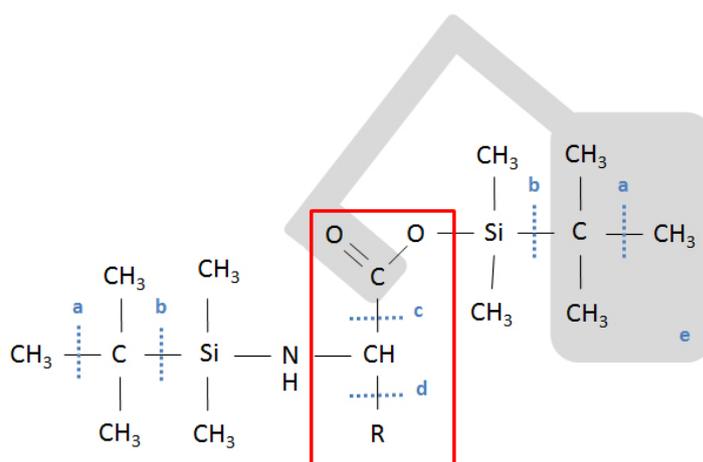


Abbildung 10.1: Die prominentesten Fragmente mit ihren „Sollbruchstellen“, die bei der Ionisierung von MBDSTFA-derivatisierten Aminosäuren entstehen. Rot umrahmt eine allgemeine Aminosäure mit dem Rest „R“.

Mit den kleinen Buchstaben „a“ bis „e“ in blau sind die Bruchstellen gekennzeichnet, wodurch folgende Fragmente entstehen: a - $(M - 15)^+$ -Fragment und eine Methylgruppe; b - $(M - 57)^+$ -Fragment und eine Tert-Butyl-Gruppe; c - $(M - 159)^+$ -Fragment und das C(O)O-TBDMS-Ion; d - $(f302)^+$ -Fragment und die Seitenkette R mit evtl. noch weiteren TBDMS-Gruppen (je nachdem um welchen Aminosäurerest es sich handelt); e - $(M - 85)^+$ -Fragment, das CO der Aminosäure und eine Tert-Butyl-Gruppe (grau hinterlegt).
Abbildung erstellt nach (Weckwerth, 2007).

Man sieht in der Mitte, rot umrahmt eine allgemeine Aminosäure mit dem Rest R. Mit den kleinen Buchstaben „a“ bis „e“ in blau sind die Bruchstellen gekennzeichnet, wodurch folgende Fragmente entstehen:

a - $(M - 15)^+$ -Fragment und eine Methylgruppe

b - $(M - 57)^+$ -Fragment und eine Tert-Butyl-Gruppe

c - $(M - 159)^+$ -Fragment und das C(O)O-TBDMS-Ion

d - $(f302)^+$ -Fragment und die Seitenkette R mit evtl. noch weiteren TBDMS-Gruppen (je nachdem um welchen Aminosäurerest es sich handelt)

e - $(M - 85)^+$ -Fragment, das CO der Aminosäure und eine Tert-Butyl-Gruppe (grau hinterlegt)

Diese einzelnen Fragmente haben für jede Aminosäure charakteristische Massen, anhand derer sie identifiziert werden können. Je nach Isotopomer unterscheiden sie dabei. Die Masse des leichtesten Isotopomers wird als m_0 bezeichnet und beinhaltet kein markiertes C-Atom. Ist eins der C-Atome markiert, addiert sich die Massenladungszahl des Fragmentes mit eins und wird m_1 genannt. m_1 beschreibt somit die Masse des zweit leichtesten Isotopomers. Dies setzt sich für alle möglichen Isotopomere eines Fragments fort, so dass man für jede Aminosäure m_0 bis m_n Massen für das entsprechende Frag-

ment erhält. Das n beschreibt dabei die maximale Anzahl der C-Atome, die von der Aminosäure stammen. In der Tabelle 10.2 sind die Massen der leichtesten Isotopomere für die in dieser Arbeit detektierten Aminosäuren aufgeführt. Die Zahl in Klammern hinter der Massenladungszahl beschreibt die Anzahl der C-Atome, die von der Aminosäure stammen.

Tabelle 10.2: Massen der leichtesten Isotopomere der in dieser Arbeit betrachteten Aminosäure-Fragmente. Die Zahl in Klammern beschreibt die Anzahl der C-Atome, die von der Aminosäure stammen. Tabelle erstellt nach (Weckwerth, 2007).

Aminosäure	$(M - 15)^+$	$(M - 57)^+$	$(M - 85)^+$	$(M - 159)^+$	$(f302)^+$
Alanin		260 (3)	232 (2)		
Glycin	288 (2)	246 (2)	218 (1)	144 (1)	
Valin		288 (5)	269 (4)		302 (2)
Leucin	344 (6)		274 (5)	200 (5)	
Isoleucin	344 (6)		274 (5)	200 (5)	
Prolin	328 (5)	286 (5)	258 (4)	184 (4)	
Serin	432 (3)	390 (3)	362 (2)	288 (2)	302 (2)
Threonin	446 (4)	404 (4)	376 (3)		
Phenylalanin		336 (9)	308 (8)	234 (8)	302 (2)
Aspartat	460 (4)	418 (4)	390 (3)	316 (3)	302 (2)
Glutamat	474 (5)	432 (5)	404 (4)	330 (4)	302 (2)
Lysin		431 (6)		329 (5)	
Tyrosin	508 (9)	466 (9)	438 (8)	364 (8)	302 (2)

Anhand der Abbildungen 10.2 und 10.3 soll das Vorgehen zur Bestimmung der Isotopomer-Verteilungsmuster der Aminosäuren am Beispiel von Alanin in dem Stamm VH33-pHN verdeutlicht werden. Die Abbildung 10.2 zeigt das dazugehörige Gaschromatogramm der hydrolysierten, mit MBDSTFA derivatisierten Biomasse des Stammes VH33-pHN. Zu den detektierten Aminosäuren gehören Alanin, Glycin, Valin, Leucin, Isoleucin, Prolin, Serin, Threonin, Phenylalanin, Aspartat, Glutamat, Lysin und Tyrosin. Sie sind in der Abbildung mit ihrem Buchstabensymbol gekennzeichnet. Um die Isotopomer-Verteilungsmuster untersuchen zu können wurde für jede gemessene Aminosäure das Massenspektrum hinsichtlich der in Tabelle 10.2 charakteristischen Massen analysiert.

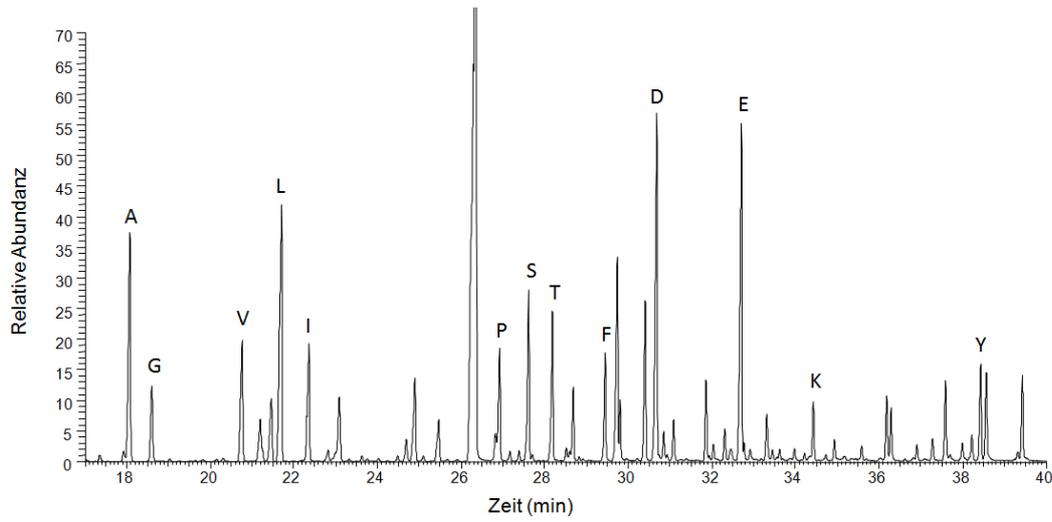


Abbildung 10.2: Gaschromatogramm der hydrolysierten, mit MBDSTFA derivatisierten Biomasse von VH33-pHN. Die detektierbaren Aminosäuren sind mit ihrem Buchstabensymbol gekennzeichnet.

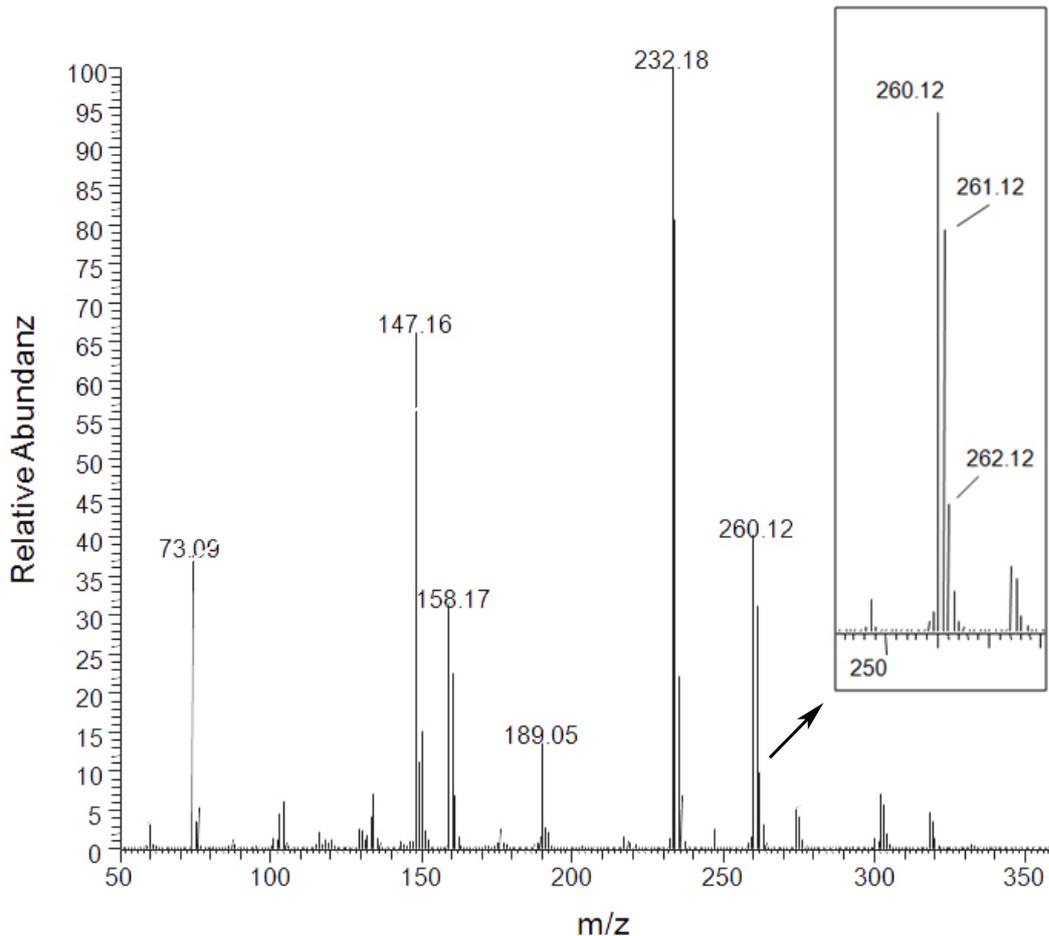


Abbildung 10.3: Massenspektrum von Alanin.

Derivatisiert man Alanin mit MBDSTFA, zerfällt es hauptsächlich in die Fragmente $(M - 57)^+$ und $(M - 85)^+$. Das leichteste Isotopomer des $(M - 57)^+$ -Fragments besitzt dabei die Masse 260 m/z und beinhaltet drei von Alanin stammende C-Atome. Das $(M - 85)^+$ -Fragment besitzt zwei von Alanin stammende C-Atome und das schwerste Isotopomer hat die Masse 232 m/z.

In Abbildung 10.3 ist das Massenspektrum für Alanin zu sehen. Derivatisiert man Alanin mit MBDSTFA, zerfällt es hauptsächlich in die Fragmente $(M - 57)^+$ und $(M - 85)^+$. Das leichteste Isotopomer des $(M - 57)^+$ -Fragments besitzt dabei die Masse 260 m/z und beinhaltet drei von Alanin stammende C-Atome. Das $(M - 85)^+$ -Fragment besitzt zwei von Alanin stammende C-Atome und das leichteste Isotopomer hat die Masse 232 m/z. Diese beiden Massen sind in dem Spektrum gut zu sehen. In dem kleinen Ausschnitt ist das $(M - 57)^+$ -Fragment vergrößert dargestellt. Hier sind neben der Masse 260 m/z (m_0) deutlich die Massen m_1 (261 m/z), m_2 (262 m/z) und m_3 (263 m/z) zu sehen. Die jeweiligen Nachkommastellen sind dabei zu vernachlässigen, da sie auf die Messgenauigkeit des Massenspektrometers zurückzuführen sind. Bei beiden Fragmenten besitzt das leichteste Isotopomer die höchste relative Häufigkeit, welche

mit zunehmender Markierung abnimmt. Im letzten Schritt berechnet man mit Hilfe der Xcalibur-Software das Isotopomer-Verteilungsmuster der Fragmente $(M - 57)^+$ und $(M - 85)^+$ und erhält die in Tabelle 10.3 aufgeführten relativen Häufigkeiten für Alanin.

Tabelle 10.3: Isotopomer-Verteilungsmuster (der relativen Häufigkeiten) der Fragmente $(M - 57)^+$ und $(M - 85)^+$ für Alanin in dem Stamm VH33-pHN

	m_0	m_1	m_2	m_3
$(M - 57)^+$	0.473522452 ± 0.007	0.37646928 ± 0.009	0.11510733 ± 0.003	0.034900938 ± 0.002
$(M - 85)^+$	0.479574488 ± 0.003	0.383236748 ± 0.001	0.106058456 ± 0.002	0.031130308 ± 0.002

Auf diese Weise wurden die Isotopomer-Verteilungsmuster für alle detektierbaren Aminosäuren aus Messung 1 und Messung 2 bestimmt. Dabei wurde für jede Aminosäure das Fragment gewählt, das im Massenspektrum am deutlichsten zu erkennen war. Die Bestimmung der Isotopomer-Verteilungsmuster wurde nicht für alle Fragmente einer Aminosäure durchgeführt, da für die spätere Modellierung nur ein Fragment berücksichtigt werden konnte.

10.2 DAS METABOLISCHE FLUSSMODELL

Derzeit sind bereits einige stöchiometrische Modelle für ^{13}C -Flussanalysen in *E. coli* publiziert (Kayser *et al.*, 2005), (Fuhrer *et al.*, 2005), (Wang *et al.*, 2006a), (Leighty & Antoniewicz, 2012). In keinem der Fachartikel wurde das verwendete Modell jedoch so publiziert, dass man es mit Hilfe der entsprechenden Software direkt anwenden kann. Es sind lediglich die darin betrachteten Reaktionen veröffentlicht worden. Daher wurde in dieser Arbeit ein metabolisches Flussmodell erstellt, welches unter Verwendung der OpenFlux-Software in Kombination mit MATLAB, *Mathworks*® zur Berechnung des intrazellulären Glukoseflusses benutzt werden kann. Das darin beschriebene Netzwerk von Stoffwechselreaktionen wurde in Anlehnung an das von Fuhrer *et al.* (2005) veröffentlichte Reaktionsnetzwerk erstellt. Es besteht aus 56 Reaktionen und beinhaltet dabei die Glykolyse, den Pentosephosphatweg, den Citratzyklus und die Acetatbildungsreaktionen. Eine Übersicht über das Reaktionsnetzwerk ist in Abbildung 10.4 dargestellt. Die Metabolitvorstufen der Aminosäuren sind in der Abbildung blau markiert und wurden

nach Weckwerth (2007) festgelegt.

Das Reaktionsnetzwerk beginnt mit der Aufnahme der extrazellulären Glucose in die Zelle. Anschließend erfolgt die Phosphorylierung der Glucose zu Glucose-6-Phosphat, welches in den darauf folgenden, katabolen Reaktionen der Glykolyse oder des Pentosephosphatwegs verstoffwechselt werden kann. Im Gegensatz zu dem von Fuhrer *et al.* (2005) veröffentlichten Reaktionsnetzwerk wird in dem hier betrachteten Modell der Entner-Doudoroff-Weg (ED) zur Vereinfachung, in Anlehnung an Wang *et al.* (2006b), nicht berücksichtigt. Der Entner-Doudoroff-Weg trägt mit nur 1 % einen sehr geringen Beitrag zur Verstoffwechslung von Glucose bei (Dauner, 2000), da in Gegenwart von Glucose als C-Quelle das Transkriptionslevel seiner Gene vernachlässigbar gering ist (Eisenberg & Dobrogosz, 1967). Eine Erhöhung der Expression wird erst durch die Anwesenheit von Gluconat induziert (Eisenberg & Dobrogosz, 1967), (Peekhaus & Conway, 1998).

Mit Hilfe des Pentosephosphatwegs wird in seiner oxidativen Phase Glucose-6-Phosphat zu Ribose-6-Phosphat oxidiert, um das Reduktionsäquivalent NADPH zu regenerieren. Die dabei entstehende Ribose-5-Phosphat und seine Derivate, dienen unter anderem als Vorläufermoleküle für die Synthese von Nukleinsäuren und einigen Aminosäuren. In dem erstellten metabolischen Modell werden Ribose-5-Phosphat und seine Derivate als Pentose-5-phosphat zusammengefasst. Die nichtoxidative Phase des Pentosephosphatwegs besteht aus der Umwandlung von 3-, 4-, 5-, 6-, und 7-fach Zuckern ineinander durch nichtoxidative Reaktionen. Dabei können überschüssige C₅-Kohlenhydrate in Form von Zwischenstufen der Glykolyse an diese abgegeben werden. (Berg *et al.*, 2007) Insgesamt betrachtet wird nach Fuhrer *et al.* (2005) der größte Anteil der aufgenommenen Glucose in *E. coli* über die Glykolyse zu Pyruvat verstoffwechselt. Über die Reaktion der Pyruvat-Dehydrogenase tritt ein Teil des Pyruvats in Form von Acetyl-CoA in den Citratzyklus ein, um dort, unter Energiegewinnung, vollständig zu CO₂ oxidiert zu werden (Berg *et al.*, 2007). Der andere Teil des Pyruvats wird mit einer stammspezifischen Acetatproduktionsrate zu Acetat umgewandelt und an das Medium abgegeben. Der Citratzyklus hat mit der Oxidation von Acetyl-CoA nicht nur eine katabole, sondern auch eine anabole Funktion, da seine Zwischenprodukte als Vorläufermoleküle für Biosynthesen werden. Zu diesen Biosynthesen gehört u. a. die Synthese von verschiedenen Aminosäuren und Fettsäuren. Um den Citratzyklus nicht zum Erliegen zu bringen, müssen die dafür entzogenen Intermediate durch eine anaplerotische Reaktion wieder

aufgefüllt werden. Dies geschieht durch die Reaktion der PEP-Carboxylase, die Phosphoenolpyruvat in Oxalacetat umwandelt, welches dem Citratzyklus zugeführt wird.

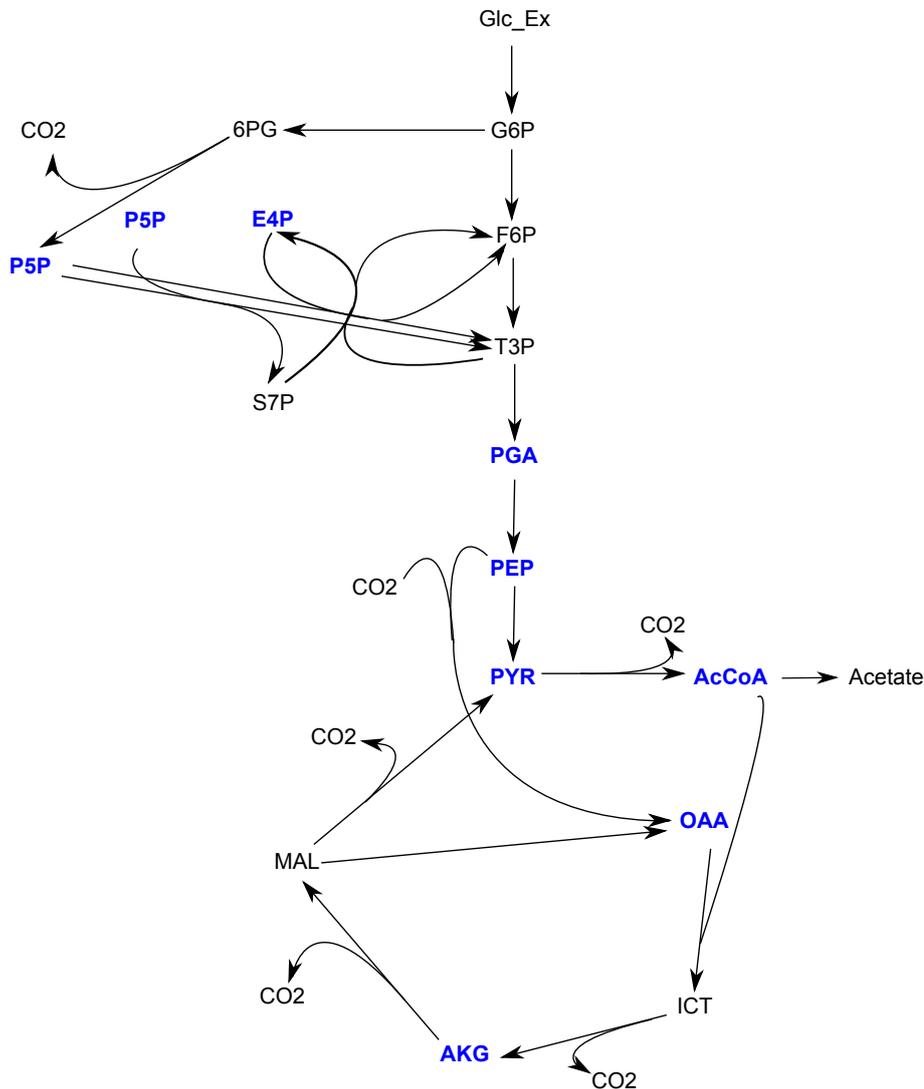


Abbildung 10.4: Reaktionsnetzwerk für *E. coli*, das als Basis zur Berechnung der metabolischen Nettoflüsse verwendet wird. Die betrachteten Metabolite sind in schwarz dargestellt. Die Vorstufen-Moleküle für die Aminosäuresynthese sind in blau markiert. Das Reaktionsnetzwerk beinhaltet neben den in Abbildung 10.4 gezeigten Reaktionen noch weitere Reaktionen, welche den Aufbau von Biomasse definieren.

(Abkürzungen: AcCoA, Acetyl-Coenzym A; AKG, α -Ketoglutarat; CO₂, Kohlenstoffdioxid; F6P, Fruktose-6-phosphat; E4P, Erythrose-4-phosphat; G6P, Glukose-6-phosphat; Glc_{Ex}, Extrazelluläre Glukose; ICT, Isocitrat; MAL, Malat; OAA, Oxalacetat; P5P, Pentose-5-phosphat; PEP, Phosphoenolpyruvat; PGA, Phosphoglycerat; PYR, Pyruvat; S7P, Sedoheptulose-7-phosphat; T3P, Triose-3-phosphat; 6PG, 6-Phosphogluconat)

Zusätzlich zu den in Abbildung 10.4 gezeigten Reaktionen wurden in dem Flussmodell noch Reaktionen zum Aufbau von Biomasse und die, für den jeweiligen Stamm typische, Plasmidproduktion und Acetatbildung berücksichtigt. Die Daten zum Aufbau der Biomasse stammen aus Neidhardt (1987). Darin wird die genaue Biomassezusam-

mensetzung einer durchschnittlichen *E. coli* Zelle beschrieben. Die verwendeten Raten zur Acetatbildung und zur Plasmidproduktion stammen von unserem Kooperationspartner, Prof. Dr. Alvaro R. Lara, und sind in der Tabelle 10.4 zusammengefasst.

Tabelle 10.4: Experimentell bestimmte Plasmid- und Acetatproduktion in den untersuchten Stämmen, angegeben in % Zelltrockengewicht. Mit freundlicher Genehmigung von Prof. Dr. Alvaro Lara.

	W ₃₁₁₀ - pHN	VH ₃₃ - pHN	VH ₃₃ Δ <i>recA</i> - pHN	VH ₃₃ Δ <i>recA</i> <i>deoR</i> -pHN
Acetatbildung [% ZTW]	0.067	0.024	0.016	0.015
Plasmidgehalt [% ZTW]	0.000169	0.00061	0.0013	0.0011

Nach Erstellung des metabolischen Flussmodells anhand der gemessenen Daten zu den Isotopomer-Verteilungsmustern, erfolgte die Berechnung des intrazellulären Glukoseflusses.

10.3 BERECHNUNG DES INTRAZELLULÄREN GLUKOSEFLUSSES

Um den Glukosefluss für die Stämme W₃₁₁₀-pHN, VH₃₃-pHN, VH₃₃ Δ *recA*-pHN, VH₃₃ Δ (*recA deoR*)-pHN berechnen zu können, wurden die Daten der GC-MS-Messungen aller detektierten Aminosäuren in das metabolische Flussmodell inseriert. Zu den detektierten Aminosäuren gehören Alanin (M – 57⁺), Valin (M – 57⁺), Threonin (M – 57⁺), Aspartat (M – 15⁺), Glutamat (M – 57⁺), Serin (M – 57⁺), Phenylalanin (M – 57⁺), Glycin (M – 57⁺), Tyrosin (M – 57⁺), Lysin (M – 57⁺), Prolin (M – 57⁺), Leucin (M – 85⁺) und Isoleucin (M – 85⁺). In Klammern ist jeweils das Fragment der Aminosäure angegeben, dessen Isotopomer-Verteilungsmuster zur Berechnung des intrazellulären Glukoseflusses eingesetzt wurde.

Der Unterschied zwischen den beiden Messungen ist die optische Dichte der Bakterienkulturen zum Zeitpunkt der Zellernte (Tab. 10.1). Die Kulturen der ersten Messung wurden in der mittleren exponentiellen Wachstumsphase geerntet und die der zweiten Messung in der späten exponentiellen Wachstumsphase.

Bei der Berechnung des Glukoseflusses mit Hilfe des metabolischen Modells entspricht die Glukoseaufnahme einem metabolischen Fluss von 100 %, so dass es sich bei den Angaben des Modells um relative Netto-Flüsse handelt. Die Berechnungen erfolgten unter Anwendung der OpenFlux-Funktion „Parameter Estimation“. Dabei wurde die

Anzahl der Iterationsschritte zur Verbesserung der Genauigkeit auf 50 erhöht und eine Lösungsclusteranzahl von 10 gewählt. Für die übrigen Optionen wurden die, von Open-Flux vorgegebenen Standardeinstellungen verwendet. Die Ergebnisse der metabolischen Flussberechnung sind in Abbildung 10.5 (Messung 1) und Abbildung 10.6 (Messung 2) dargestellt.

Da sich diese Arbeit mit der Synthese von Plasmid-DNA beschäftigt, wurde für die Stämme W₃₁₁₀-pHN, VH₃₃-pHN, VH₃₃Δ *recA*-pHN, VH₃₃Δ (*recA deoR*)-pHN insbesondere die Auftrennung des Glukoseflusses auf die Glykolyse und den Pentosephosphatweg analysiert. Betrachtet man die metabolischen Flussberechnungen der ersten Messung fällt auf, dass der Anteil der aufgenommenen Glukose, der in den Pentosephosphatweg fließt in dem Stamm W₃₁₁₀-pHN mit 21 % am geringsten ist. In den Stämmen VH₃₃-pHN und VH₃₃Δ (*recA deoR*)-pHN ist der metabolische Fluss in den Pentosephosphatweg mit 71.8 % bzw. 67.3 % mehr als dreimal so hoch. Eine Ausnahme bildet VH₃₃Δ *recA*-pHN. Dieser Stamm zeichnet sich, ebenso wie VH₃₃-pHN und VH₃₃Δ (*recA deoR*), durch eine erhöhte Plasmidproduktion aus, so dass bezüglich des metabolischen Flusses ein ähnliches Verhalten erwartet wurde. Betrachtet man jedoch die Ergebnisse der Flussberechnungen, stimmen diese nicht mit den Erwartungen überein, da der metabolischen Fluss in den Pentosephosphatweg mit 32.1 % deutlich kleiner ist, als bei VH₃₃-pHN, und VH₃₃Δ (*recA deoR*) und somit näher im Bereich dem des Wildtyps (21 %) liegt.

Die Berechnung des Anteils der aufgenommenen Glukose der über die Phosphoglucose-Isomerase-Reaktion in die Glykolyse fließt, ergab für den Wildtyp einen Nettofluss von 79 %. Für die anderen drei Stämme wurde ein Fluss von 0 % bzw. 100 % berechnet. Betrachtet man diese berechneten Flüsse vor dem Hintergrund, dass 100 % Glukose in die Zelle aufgenommen wurden, erscheint eine Aufteilung des Glukoseflusses zu 21 % auf den Pentosephosphatweg und zu 79 % auf die Glykolyse plausibel. Bei der Berechnung der Glukoseflüsse für die Stämme VH₃₃-pHN, VH₃₃Δ *recA*-pHN und VH₃₃Δ (*recA deoR*)-pHN würde man demzufolge für die Phosphoglucose-Isomerase-Reaktion metabolische Flüsse im Bereich von 28.2 % bis 67.9 % erwarten. Da für diese Stämme Flüsse von 0 % bzw. 100 % berechnet wurden ist anzunehmen, dass diese Flüsse vermutlich von der Software nicht eindeutig berechnet werden konnten. Diese Vermutung wird dadurch gestützt, dass die berechneten Flüsse einen Wert von 0 % oder 100 % betragen, was genau den Grenzen entspricht, innerhalb derer sich die Flüsse bewegen. Dabei entspricht 0 % der unteren Grenze und 100 % der oberen Grenze.

W3110-pHN
 VH33-pHN
 VH33recA-pHN
 VH33recAdeoR-pHN

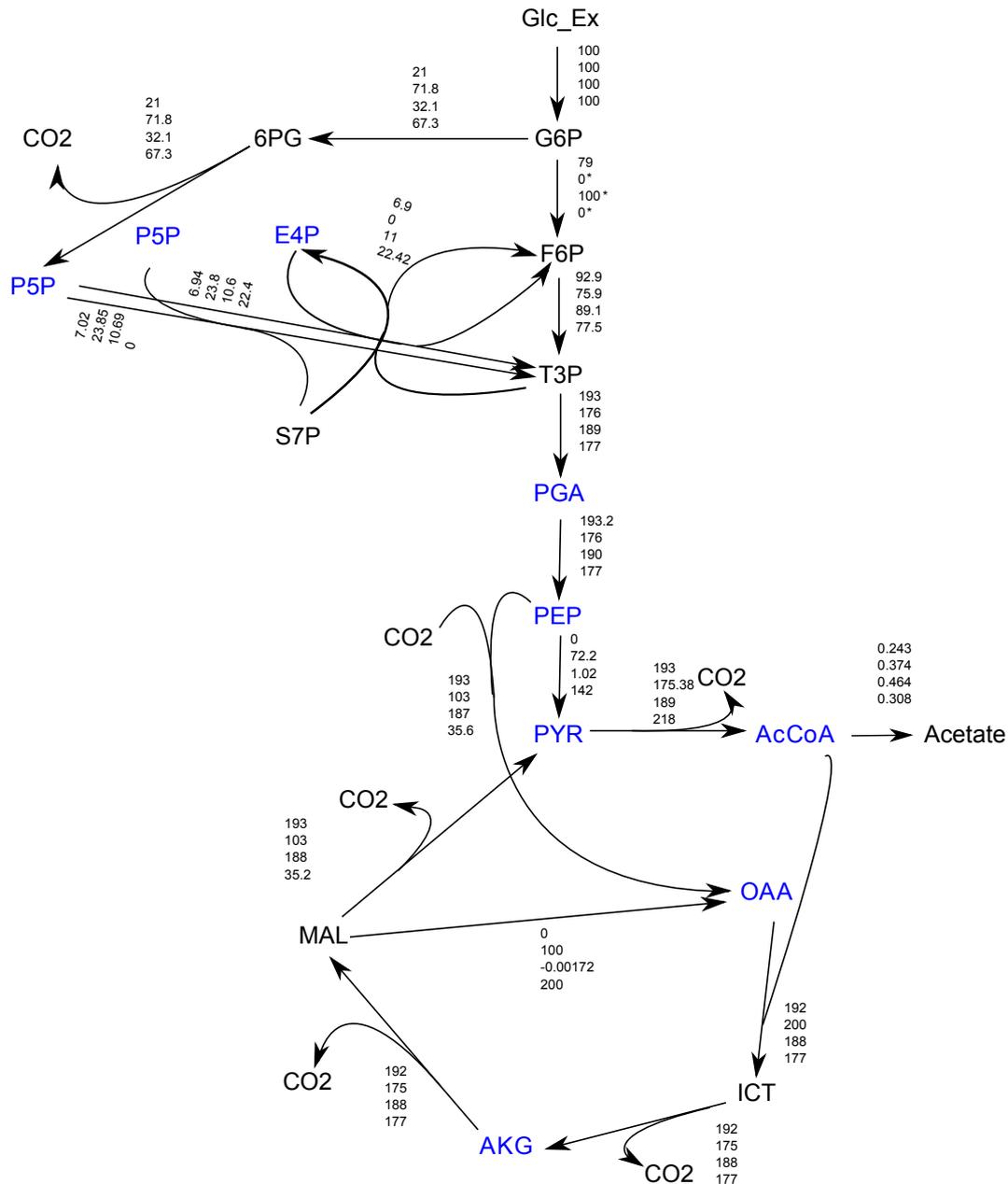


Abbildung 10.5: Ergebnis zur Berechnung der metabolischen Flüsse innerhalb der Stämme W3110-pHN, VH33-pHN, VH33Δ *recA*-pHN, VH33Δ (*recA deoR*)-pHN für Messung 1 (Tab. 10.1). Die berechneten Flüsse sind auf die jeweilige Glukoseaufnahme normalisiert und als Prozentwerte dargestellt. Anmerkung: Der metabolische Fluss der Phosphoglucose-Isomerase-Reaktion konnte für die Stämme VH33-pHN, VH33Δ *recA*-pHN und VH33Δ (*recA deoR*)-pHN von der Software nicht eindeutig berechnet werden. Die entsprechenden Flüsse sind in der Abbildung mit einem „*“ gekennzeichnet.

(Abkürzungen: AcCoA, Acetyl-Coenzym A; AKG, α -Ketoglutarat; CO₂, Kohlenstoffdioxid; F6P, Fruktose-6-phosphat; E4P, Erythrose-4-phosphat; G6P, Glukose-6-phosphat; Glc_{Ex}, Extrazelluläre Glukose; ICT, Isocitrat; MAL, Malat; OAA, Oxalacetat; P5P, Pentose-5-phosphat; PEP, Phosphoenolpyruvat; PGA, Phosphoglycerat; PYR, Pyruvat; S7P, Sedoheptulose-7-phosphat; T3P, Triose-3-phosphat; 6PG, 6-Phosphogluconat)

Es ist daher davon auszugehen, dass der metabolische Fluss der Phosphoglucose-Isomerase-Reaktion für die Stämme VH33-pHN, VH33Δ *recA*-pHN und VH33Δ (*recA deoR*)-pHN von der Software nicht eindeutig berechnet werden konnte. Der Grund dafür liegt vermutlich darin, dass die in den gemessenen Daten enthaltenen Markierungsinformationen nicht ausreichen, um den metabolischen Fluss an dieser Stelle eindeutig zu bestimmen. Die entsprechenden Flüsse sind in Abbildung 10.5 mit einem „*“ gekennzeichnet.

Vergleicht man die metabolischen Flüsse der Stämme innerhalb des Pentosephosphatwegs, fällt auf, dass der metabolische Fluss von P5P in Richtung Glykolyse größer wird, je größer die Plasmidproduktivität des Stammes ist. Der metabolische Fluss entlang der Glykolyse in Richtung PEP verhält sich dagegen umgekehrt. Hier gilt, je größer die Plasmidproduktivität des Stammes, desto geringer ist der metabolische Fluss. Beim Wildtyp W3110-pHN fließen 193.2 % in Richtung PEP; bei den Stämmen VH33-pHN und VH33Δ (*recA deoR*)-pHN sind es 176 % bzw. 177 %. Der Stamm VH33Δ *recA*-pHN bildet auch in diesem Fall eine Ausnahme, da der metabolische Fluss in Richtung PEP mit 190 % nicht, wie erwartet, innerhalb der Größenordnung der Flüsse von VH33-pHN und VH33Δ (*recA deoR*)-pHN liegt.

Betrachtet man den metabolischen Fluss innerhalb des Citratzyklus, fällt auf, dass dieser mit steigender Plasmidproduktivität des Stammes abnimmt. Bei dieser Betrachtung ist jedoch zu berücksichtigen, dass das Interesse dieser Arbeit der Synthese von Plasmid-DNA-Vorstufen gilt und die Markierungsexperimente zur Berechnung der metabolischen Flüsse daher mit ^{13}C -markierter Glukose durchgeführt wurden. Die Verwendung von ^{13}C -Glukose ermöglicht eine gute Auflösung der oberen Stoffwechselwege der Glykolyse und des Pentosephosphatwegs (Wang *et al.*, 2006b). Die unteren Stoffwechselwege der Glykolyse und der Fluss des Citratzyklus lassen sich mit diesem markierten Substrat weniger gut untersuchen. Für die Analyse dieser Stoffwechselwege sind Markierungsexperimente mit einer Mischung aus uniform markierter Glukose (U^{13}C -Glukose) und natürlich markierter Glukose besser geeignet (Wang *et al.*, 2006b). Aus diesem Grund ist die Auflösung der unteren Stoffwechselwege der Glykolyse und der Citratzyklus unter Umständen nicht optimal.

Betrachtet man die metabolischen Flussberechnungen der zweiten Messung, zeigt sich im Vergleich zur ersten Messung ein deutlicher Unterschied bei der Auftrennung des Glukoseflusses auf den Pentosephosphatweg und die Glykolyse. Im Gegensatz zur ersten Messung ist der Glukoseanteil, der über den Pentosephosphatweg verstoffwechselt

wird, bei den Stämmen VH33-pHN und VH33Δ (*recA deoR*)-pHN mit 24.4 % bzw. 38.6 % deutlich geringer. Beim Wildtyp war mit 24.8 % ein etwas höherer Fluss festzustellen, wobei bei VH33Δ *recA*-pHN mit 20.9 % überraschenderweise eine leichte Abnahme zu verzeichnen war. Für VH33Δ *recA*-pHN wurde ein Wert erwartet, der eine Größenordnung zwischen 24.4 % und 38.6 % besitzt, da VH33Δ *recA*-pHN mit seiner Plasmidproduktivität zwischen den Stämmen VH33-pHN und VH33Δ (*recA deoR*)-pHN liegt. Der berechnete metabolische Nettofluss in die Glykolyse beträgt in der Messung 2 bei allen Stämmen 100 % und entspricht nicht den Erwartungen. Aufgrund der berechneten Glukoseflüsse in den Pentosephosphatweg würde man für die Phosphoglukose-Isomerase-Reaktion metabolische Flüsse im Bereich von 79.1 % bis 61.4 % erwarten, da sich die aufgenommene Glukose auf den Pentosephosphatweg und auf die Glykolyse verteilt. Wie jedoch bereits bei den Ergebnissen der Messung 1 erläutert, deuten die für die Phosphoglukose-Isomerase-Reaktion berechneten Flüsse mit Werten von 100 % darauf hin, dass der metabolische Fluss in die Glykolyse für alle Stämme von der Software nicht eindeutig berechnet werden konnte. Die entsprechenden Flüsse sind in Abbildung 10.6 mit einem „*“ gekennzeichnet.

Innerhalb des Pentosephosphatwegs ist der metabolische Fluss von P5P in Richtung Glykolyse beim Wildtyp mit ≈ 8 % ähnlich zu dem der ersten Messung. Die Stämme VH33-pHN, VH33Δ *recA*-pHN und VH33Δ (*recA deoR*)-pHN zeigen im Vergleich zur ersten Messung mit ≈ 8 %, ≈ 8 %, ≈ 7 % und ≈ 13 % einen geringeren Fluss in Richtung Glykolyse.

Betrachtet man den metabolischen Fluss entlang der Glykolyse in Richtung PEP fällt auf, dass der Fluss bei W3110-pHN und VH33Δ *recA*-pHN mit 191 % bzw. 193 % dem der ersten Messung ähnelt. Bei VH33-pHN und VH33Δ (*recA deoR*) ist der Fluss jedoch im Vergleich zur ersten Messung gestiegen.

Zusammenfassend lässt sich festhalten, dass sich die Stämme W3110-pHN, VH33-pHN und VH33Δ (*recA deoR*)-pHN sowohl in der mittleren exponentiellen Wachstumsphase (Messung 1) als auch in der späten exponentiellen Wachstumsphase (Messung 2), hinsichtlich ihrer metabolischen Flüsse in Abhängigkeit der Plasmidproduktivität unterscheiden. Eine Ausnahme bildet der Stamm VH33Δ *recA*-pHN, bei dem die berechneten Flüsse nicht den Erwartungen entsprechen. Es ist an dieser Stelle jedoch anzumerken, dass es sich möglicherweise nicht um den richtigen Stamm handelt, da es Hinweise gibt, dass der Stamm bereits in Mexiko vor der Kultivierung von einem Studenten vertauscht wurde.

W3110-pHN
 VH33-pHN
 VH33recA-pHN
 VH33recAdeoR-pHN

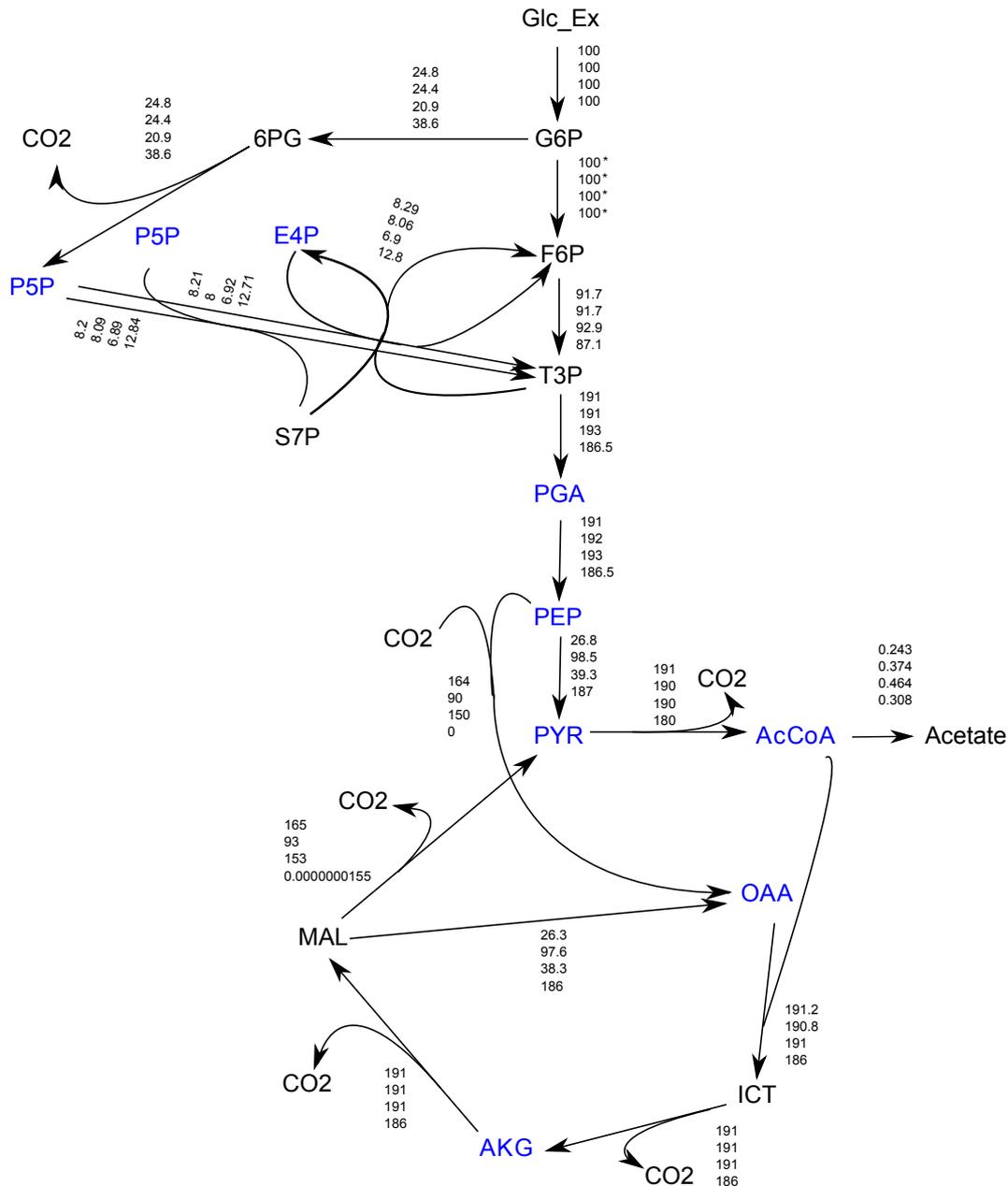


Abbildung 10.6: Ergebnis zur Berechnung der metabolischen Flüsse innerhalb der Stämme W3110-pHN, VH33-pHN, VH33 *recA*-pHN, VH33Δ (*recA deoR*)-pHN für Messung 2 (Tab. 10.1). Die berechneten Flüsse sind auf die jeweilige Glukoseaufnahme rate normalisiert und als Prozentwerte dargestellt. Anmerkung: Der metabolische Fluss der Phosphoglucose-Isomerase-Reaktion konnte für alle Stämme von der Software nicht eindeutig berechnet werden. Die entsprechenden Flüsse sind in der Abbildung mit einem „*“ gekennzeichnet.

(Abkürzungen: AcCoA, Acetyl-Coenzym A; AKG, α -Ketoglutarat; CO₂, Kohlenstoffdioxid; F6P, Fructose-6-phosphat; E4P, Erythrose-4-phosphat; G6P, Glukose-6-phosphat; Glc_{Ex}, Extrazelluläre Glukose; ICT, Isocitrat; MAL, Malat; OAA, Oxalacetat; P5P, Pentose-5-phosphat; PEP, Phosphoenolpyruvat; PGA, Phosphoglycerat; PYR, Pyruvat; S7P, Sedoheptulose-7-phosphat; T3P, Triose-3-phosphat; 6PG, 6-Phosphogluconat)

Teil VI

DISKUSSION

MODELLIERUNG DER PLASMIDREPLIKATIONSKONTROLLE VON *COLE*₁-ÄHNLICHEN PLASMIDEN

11.1 DAS MODELL

Das in dieser Arbeit vorgestellte Modell zur Simulation der *ColE*₁-Plasmidreplikationskontrolle ist eine wichtige Fortsetzung der Modellierungsarbeiten von Bremer & Lin-Chao (1986), Ataai & Shuler (1986), Brendel & Perelson (1993) und Wang *et al.* (2002), da es die bislang vorhandenen Modelle erweitert. Es unterscheidet sich dadurch, dass nicht nur die Regulation durch inhibitorische RNAI-Moleküle betrachtet wird, sondern auch die Kontrolle in Form von unbeladenen tRNA-Molekülen, die an plasmidgebundene RNAII binden und sie so vor der Hemmung durch RNAI-Moleküle schützen. Dadurch ist das dem Modell zu Grunde liegende Reaktionsnetzwerk detaillierter und in Folge dessen die Anzahl der verwendeten Differentialgleichungen größer.

Das von Bremer & Lin-Chao (1986) vorgeschlagene Modell besteht aus drei Differentialgleichungen, in denen die Zunahme der Zellmasse, die zeitliche Änderung der RNAI-Konzentration sowie die Veränderung in der Plasmidkopienzahl berücksichtigt wird. Der Einfluss des Rom-Proteins wird dabei durch einen Faktor modelliert, der die Effizienz der Hemmung durch RNAI beschreibt (Bremer & Lin-Chao, 1986). Das Reaktionsnetzwerk des von Ataai & Shuler (1986) vorgestellten Modells betrachtet neben der RNAI-Konzentration und der Plasmidkopienzahl noch die RNAII-Konzentration und das Rom-Protein. Es ist somit, im Vergleich zu dem Modell von (Bremer & Lin-Chao, 1986), eine genauere Beschreibung der *ColE*₁-Plasmidreplikationskontrolle. Der Unterschied zwischen dem Modell von Brendel & Perelson (1993) zu dem von Ataai & Shuler (1986) liegt in der ausführlicheren Beschreibung kinetischer Details, wie z. B. der Bildung eines instabilen Komplexes zwischen plasmid-gebundener RNAII und freier RNAI (Brendel & Perelson, 1993). Des Weiteren betrachten Brendel & Perelson (1993) in ihrem Modell nicht nur die Plasmidreplikationskontrolle eines *low copy* Plasmids, sondern auch die eines *high copy* Plasmids. Das von Wang *et al.* (2002)

vorgeschlagene Modell beinhaltet dagegen nur die Plasmidreplikationskontrolle eines *low copy* Plasmids und basiert auf einem kleineren Reaktionsnetzwerk. Im Gegensatz zu Brendel & Perelson (1993) berücksichtigen Wang *et al.* (2002) jedoch zusätzlich die Plasmidreplikationskontrolle durch unbeladene tRNA-Moleküle unter Nährstoffmangelbedingungen, die an RNAI-Moleküle binden und somit die Konzentration inhibitorisch aktiver RNAI-Moleküle aufheben.

Das in dieser Arbeit vorgestellte Modell stellt eine Erweiterung des von Brendel & Perelson (1993) vorgeschlagenen Modells dar. Es beschreibt die *in vivo ColE1*-Plasmidreplikationskontrolle durch RNAI und RNAII sowohl für einen *low copy*, als auch für einen *high copy* Plasmid und berücksichtigt zusätzlich die Kontrolle durch unbeladene tRNA-Moleküle. Generell kann die Regulation durch unbeladene tRNAs nach Yavachev & Ivanov (1988) über zwei Arten erfolgen: Einige der tRNAs binden im unbeladenen Zustand bevorzugt RNAI-Moleküle und die anderen hybridisieren bevorzugt mit RNAII-Molekülen (Yavachev & Ivanov, 1988). In beiden Fällen kommt es zu einer Erhöhung der Plasmidproduktion, mit dem Unterschied, dass verschiedene tRNAs an den Reaktionen beteiligt sind. In dieser Arbeit soll, mit Hilfe des Modells, der Effekt einer modifizierten unbeladenen Leucin-tRNA untersucht werden. Nach Yavachev & Ivanov (1988) binden unbeladene Leucin-tRNAs aufgrund von Sequenzhomologien zur RNAI mit großer Wahrscheinlichkeit bevorzugt RNAII-Moleküle (Yavachev & Ivanov, 1988), so dass in dem hier vorgeschlagenem Modell lediglich die Bindung von RNAII-Molekülen berücksichtigt wird.

Um den Einfluss von modifizierten, unbeladbaren tRNA-Molekülen auf die Plasmidproduktion analysieren zu können, wird die *in vivo ColE1*-Plasmidreplikationskontrolle mit Hilfe des Modells unter drei verschiedenen Nährstoffbedingungen untersucht: Normale Nährstoffbedingungen, Nährstoffmangel und die damit verbundene Anwesenheit von unbeladenen tRNA-Molekülen sowie normale Nährstoffbedingungen mit Anwesenheit von unbeladenen tRNA-Molekülen. Dabei werden die, unter normalen Nährstoffbedingungen anwesenden, unbeladenen tRNAs durch das modifizierte Leucin-tRNA-Gen kodiert.

Die in dieser Arbeit, mit Hilfe des Modells, durchgeführten Vorhersagen beziehen sich nur auf den Plasmidgehalt in einer einzelnen Zelle, da die Plasmidreplikationskontrolle nur für die Dauer einer Generationszeit simuliert wird. Es ist mit dem vorgestellten Modell auch möglich, die Plasmidproduktion, beginnend mit einer einzelnen Zelle, für mehrere Generationszeiten vorherzusagen. Dies kann sowohl für gleichbleibende, als

auch für variierende Wachstumsbedingungen, wie z.B. sich ändernde Wachstumstaten, durchgeführt werden. Dazu müssen die sich ändernden Bedingungen dem Modell durch das Einfügen sogenannter „Events“ übergeben werden. Das Hinzufügen von Events sind kleine, leicht durchführbare Anpassungen, die jedoch im Rahmen dieser Arbeit nicht gezeigt werden, da die Simulationen der Plasmidreplikationskontrolle über mehrere Generationszeiten zur Untersuchung der Fragestellung nicht relevant sind.

11.2 UNTERMAUERUNG DES MODELLS MIT EXPERIMENTELLEN DATEN

Ein großer Vorteil des hier vorgestellten Modells ist seine Untermauerung mit experimentellen Daten, die an drei verschiedenen Messzeitpunkten für die zu modellierenden Stämme erhoben wurden. Da die Datensets für jeden der drei Zeitpunkte in demselben Stamm gemessen wurden, kann man sie auch als homogene Datensets bezeichnen. Viele publizierte Modelle basieren auf heterogenen Datensets, in denen die verwendeten Parametermessungen aus verschiedenen Stämmen und zum Teil sogar aus verschiedenen Organismen stammen (Klumpp, 2011). Dies ist für die Modellierung nicht optimal, da mit Hilfe des Modells in der Regel ein spezifischer Organismus untersucht werden soll, für den ein in einem fremden Organismus gemessener Parameter, unter Umständen, ungeeignet ist. Aufgrund von mangelnden Messdaten ist es oft jedoch unumgänglich die experimentellen Daten zur Untermauerung eines Modells unter bestimmten Annahmen aus verschiedenen Stämmen und Organismen zusammenzutragen, so dass dem Modell ein heterogenes Datenset zugrunde liegt. Die in dieser Arbeit erhobenen experimentellen Messdaten stellen somit den Idealfall eines Parametersets dar, da die Messungen an drei verschiedenen Zeitpunkten für denselben Bakterienstamm (DH5 α -pSUP 201-3 bzw. DH5 α -pCMV-lacZ) im selben Labor durchgeführt wurden. Ein solches experimentelles Datenset beinhaltet, pro Messzeitpunkt, die gemessenen RNAI-, RNAII- und pDNA-Konzentrationen sowie die ermittelte Wachstumsrate.

Für die Bestimmung der Wachstumsraten wurde zunächst der Wachstumsverlauf durch die Messung der optischen Dichte bei 600 nm verfolgt. Mit Hilfe der MATLAB-Funktion „polyfit“ wurde anschließend eine Ausgleichsfunktion berechnet, die eine geeignete Repräsentation der gemessenen Wachstumskurve darstellt. Die Ausgleichsfunktion diente dabei als Hilfsmittel um die Wachstumsraten μ für die relevanten Erntezeitpunkte bestimmen zu können. Anhand der Wachstumsrate konnte anschließend die entsprechende Generationszeit berechnet werden, die zusammen mit der Wachstumsrate in das

dynamische Modell für den Erntezeitpunkt eingesetzt wird, an dem die Simulation beginnt.

Die graphischen Darstellungen der Ausgleichsfunktionen für DH5 α -pSUP 201-3 und DH5 α -pCMV-lacZ (Abb. 8.4 und Abb. 8.5) zeigen, dass die Daten des Wachstumsverlauf durch die Kurve gut wiedergegeben werden. Für beide Stämme verläuft die Kurve sehr nah an den einzelnen Messwerten innerhalb des Minima-Maxima-Bereichs. Sie bricht lediglich in der stationären Phase aus, was jedoch für diese Arbeit kein Problem darstellt, da sich in diesem Bereich keine Erntezeitpunkte befinden und der Bereich somit für die Berechnung der Wachstumsraten anhand der ersten Ableitung nicht relevant ist.

Die Wachstumsraten wurden mittels der ersten Ableitung der Ausgleichsfunktion berechnet. Bei Betrachtung der berechneten Wachstumsraten (Tab. 8.2 und Tab. 8.4) fällt auf, dass sowohl der *low copy* plasmidtragende Stamm, DH5 α -pSUP 201-3, als auch der *high copy* plasmidtragende Stamm, DH5 α -pCMV-lacZ, sehr langsam wächst. Dies liegt an der Verwendung eines Minimalmediums zur Kultivierung der Bakterien. Das Medium enthält nur für *E. coli* essentielle Bestandteile, so dass die Zellen gezwungen sind alles weitere selbst zu synthetisieren. Der Grund für die Wahl dieses Mediums liegt in der Durchführung der Metabolomanalysen. Da die Ergebnisse der Modellierung und die der Metabolomanalysen miteinander in Zusammenhang gesetzt werden sollen, sollte so wenig Variation wie möglich in der Kultivierung vorhanden sein. Für die Metabolommessungen sollte das Nährmedium keine Substanzen beinhalten, die von dem Bakterium selbst synthetisiert werden können (z. B. *Casamino acids*), um die Ergebnisse nicht zu verfälschen. Aus diesem Grund wurde für die Bestimmung der experimentellen Daten zur Modellierung ebenfalls ein Minimalmedium verwendet. Das langsame Wachstum zeigte sich jedoch nur bei der Verwendung eines Minimalmediums; bei einer Kultivierung auf einem LB-Vollmedium zeigten die beiden Stämme ein typisches Wachstumsverhalten.

Die Bestimmung der intrazellulären RNAI- und RNAII-Konzentrationen erfolgte mittels qRT-PCR. Dazu wurden zunächst, parallel zur Zellernte, die Bakterientiter der beiden Stämme für jeden Erntezeitpunkt durch das Ausplattieren von Verdünnungsreihen ermittelt. Dieses Vorgehen war notwendig, um die absoluten RNAI- und RNAII-Konzentrationen sowie die Plasmidkopienzahl berechnen zu können. Für die Ermittlung der absoluten RNAI- und RNAII-Konzentrationen wurden Standardlösungen mit bekannten RNAI- und RNAII-Konzentrationen vermessen, um mit Hilfe der Excel-Software jeweils eine Eichgerade zu erstellen. Da das Bestimmtheitsmaß R^2 der errech-

neten Eichgeraden bei beiden Stämmen ungefähr 1 beträgt, werden die Rohdaten durch die Geraden gut repräsentiert. Mit Hilfe dieser Eichgeraden wurden anschließend die RNAI- und RNAII-Konzentrationen anhand der gemessenen CT-Werte berechnet.

Betrachtet man die gemessenen RNAI- und RNAII-Konzentrationen für DH5 α -pSUP 201-3 (Tab. 8.8 und Tab. 8.9), fällt auf, dass die Anzahl der beiden RNA-Transkripte pro Zelle mit voranschreitendem Wachstum abnimmt. Des Weiteren lässt sich erkennen, dass zu jedem Zeitpunkt mehr freie RNAI-, als RNAII-Moleküle in der Zelle vorkommen. Beim Vergleich der Messungen mit den von Brenner & Tomizawa (1991) vor 22 Jahren bestimmten RNAI- und RNAII-Konzentrationen fällt auf, dass die RNAII-Konzentrationen in derselben Größenordnung liegen. Brenner & Tomizawa (1991) bestimmten eine durchschnittliche Anzahl von 1.9 ± 0.9 bis 3.7 ± 1.0 RNAII-Molekülen pro Zelle (Brenner & Tomizawa, 1991). Die Größenordnung dieser Werte stimmt sehr gut mit den Messungen dieser Arbeit überein. Bei der Anzahl der RNAI-Moleküle verhält es sich etwas anders. Die von Brenner & Tomizawa (1991) ermittelten RNAI-Konzentrationen liegen zwischen 333 ± 49 und 499 ± 36 . Sie sind somit knapp um einen Faktor von zehn größer, als die in dieser Arbeit gemessenen Molekülanzahlen. Dieser Größenunterschied könnte durch die verschiedenen Messmethoden bedingt sein. Brenner & Tomizawa (1991) führten ihre Messungen densitometrisch, unter Anwendung von quantitativen *Probe Protection*-Experimenten durch. Dazu synthetisierten sie via *in vitro*-Transkription für beide RNAs ^{32}P -markierte *Riboprobes*, die mit den *in vivo* vorhandenen RNAI- und RNAII-Molekülen hybridisieren. Anschließend erfolgte eine Behandlung mit einer RNase, die spezifisch für die Degradation einzelsträngiger Moleküle ist (RNase A und/ oder RNase T₁). Die RNAI- und RNAII-Transkripte waren aufgrund der Hybridisierung mit der markierten *Riboprobe* vor der Degradation geschützt und konnten über ihre Radioaktivität und densitometrisch quantifiziert werden (Brenner & Tomizawa, 1991). Im Vergleich zu der hier angewendeten qRT-PCR ist die Quantifizierung mittels *Probe Protection*-Experimenten möglicherweise weniger sensitiv, wodurch diese Größenunterschiede zustande kommen können. Des Weiteren verwendeten Brenner & Tomizawa (1991) zur Kultivierung ein, mit verschiedenen nicht-essentiellen Zusätzen, supplementiertes Minimalmedium. Diese etwas andere Zusammensetzung könnte das Wachstum der Zellen etwas beschleunigen und den Plasmidgehalt leicht verringern.

Das Verhältnis von RNAI zu RNAII (Tab. 8.10) ist mit 11 zum Zeitpunkt T₁ am kleinsten und besitzt zu den Zeitpunkten T₂ und T₃ den Wert 17. Trotz des leichten Unterschieds kann man insgesamt von einem stabilen RNAI/RNAII-Verhältnis innerhalb der Mess-

zeitpunkte sprechen, was auf eine stabile Plasmidkopienzahl hindeutet.

Mit den für DH5 α -pCMV-lacZ gemessenen RNAI- und RNAII-Konzentrationen (Tab. 8.11 und Tab. 8.12) verhält es sich anders. Die RNAI-Konzentration ist zum Zeitpunkt T₂ am höchsten und zum Zeitpunkt T₃ am geringsten. Die gemessene RNAII-Konzentration ist zum Erntezeitpunkt T₁ am geringsten und erreicht zum dritten Zeitpunkt ihr Maximum. Dies ist genau umgekehrt, als es bei dem *low copy* plasmidtragenden Stamm, DH5 α -pSUP 201-3, der Fall ist. Ein weiterer Unterschied zwischen DH5 α -pSUP 201-3 und DH5 α -pCMV-lacZ ist, dass für DH5 α -pCMV-lacZ kein stabiles RNAI/RNAII-Verhältnis über alle Messzeitpunkte hinaus festgestellt werden konnte (Tab. 8.13). Das Verhältnis ist zum ersten Erntezeitpunkt am größten und wird dann stetig kleiner. Das bedeutet, es verändert sich mit voranschreitendem Wachstum zugunsten der RNAII-Konzentration. Ein Vergleich der gemessenen Konzentrationen mit Literaturwerten ist für DH5 α -pCMV-lacZ nicht möglich, da bislang keine bekannten Vergleichsdaten vorliegen

Die Bestimmung der absoluten Plasmidkonzentrationen erwies sich als problematisch, da der gesamte Plasmidgehalt einer geernteten Probe aufgereinigt werden musste. Dies sollte mit Hilfe einer Anionentauschersäule durchgeführt werden, da anhand von Testisolierungen festgestellt wurde, dass sich eine Plasmidisolierung mittels alkalischer Lyse und anschließender Ethanol-fällung als weniger erfolgreich erwies. Ein wichtiger Punkt bei der Aufreinigung war, genauestens darauf zu achten, dass keine pDNA-Verluste, z. B. durch die Überladung der Anionentauschersäule entstehen.

Als erstes wurde die intrazellulären Plasmidkonzentration von der Plasmid Factory bestimmt. Dazu wurde die pDNA zunächst über eine NucleoBond PC20- Säule (Macherey-Nagel) aufgereinigt und im NanoDrop-Gerät bei 260 nm vermessen. Im Anschluss erfolgte die gelelektrophoretische Auftrennung und die densitometrische Bestimmung des Plasmid-Anteils an der Gesamt-DNA.

Die für DH5 α -pSUP 201-3 gemessenen Plasmidkopienzahlen liegen bei durchschnittlich 46 bzw. 48 Plasmiden pro Zelle (Tab. 8.15). Vergleicht man diese mittleren Kopienzahlen mit den in der Literatur beschriebenen Angaben für einen *low copy* Plasmid mit pBR322-Replikationsursprung (15-20 Plasmide/ Zelle) (Cooper, 2004), stellt man fest, dass die hier gemessenen Angaben ungefähr doppelt so hoch sind. Berücksichtigt man jedoch die hohen Standardabweichungen, lässt sich erkennen, dass die Literaturwerte noch innerhalb der seitens durch die Fehlerbalken entstandenen Spanne liegen. Der dennoch vorhandene, große Unterschied kann möglicherweise durch die Kultivierung in dem

Minimalmedium zustande kommen, da die Zellen hier sehr langsam wachsen und langsames Wachstum die Produktion von pDNA begünstigt (Bremer & Lin-Chao, 1986). Insgesamt betrachtet kann man anhand der Messungen sehen, dass sich der durchschnittliche Plasmidgehalt pro Zelle in dem Wachstumsbereich von T_1 bis T_3 stabil verhält. Diese stabilen, durchschnittlichen Plasmidkopienzahlen sind vermutlich durch das, ebenfalls relativ stabile, Verhältnis von RNAI zu RNAII (Tab. 8.10) bedingt.

Im Gegensatz zu DH5 α -pSUP 201-3 nimmt der Plasmidgehalt bei DH5 α -pCMV-lacZ im Verlauf des Wachstums deutlich zu (Tab. 8.17). Mit durchschnittlich 5806 ± 4828 Plasmiden pro Zelle zum Zeitpunkt T_3 , ist die pDNA-Konzentration damit innerhalb der betrachteten Messpunkte am größten. Zusammen mit der pDNA-Konzentration nimmt die RNAII-Konzentration bei DH5 α -pCMV-lacZ ebenfalls zu. Man kann daher davon ausgehen, dass die steigende RNAII-Konzentration für die Zunahme der PCN verantwortlich ist. Wie bereits bei den Messungen für DH5 α -pSUP 201-3 waren auch hier sehr große Standardabweichungen zu beobachten. Beim Vergleich der gemessenen Plasmidkopienzahlen für DH5 α -pCMV-lacZ mit den Literaturangaben für pUC-Plasmide (500-700 Plasmide pro Zelle (Cooper, 2004)), fällt zusätzlich auf, dass die gemessenen Werte nicht mit denen aus der Literatur übereinstimmen. Weiterführende Überlegungen dazu ergaben, dass für die hohen Plasmidkopienzahlen und die großen Standardabweichungen möglicherweise ein Fehler bei Durchführung der Plasmid-Quantifizierung verantwortlich ist. Es könnte z. B. sein, dass bei der densitometrischen Berechnung der Anteil der chromosomalen DNA an dem Gesamt-DNA-Anteil unterschätzt wurde, was die PCN fälschlicherweise erhöhen würde. Um dies zu überprüfen, wurden zwei der DH5 α -pCMV-lacZ-Proben (DH5 α -pCMV-lacZ-A2 und DH5 α -pCMV-lacZ-A3) von der Firma CARPEGEN unter Anwendung einer sondenbasierten RT-PCR erneut vermessen. Diese exemplarischen Messungen dienten der Überprüfung, ob die Größenordnungen (1000er Bereich) der vorherigen Quantifizierung realistisch sind oder ob bei der Durchführung ein Fehler passiert ist. Die Ergebnisse der Messungen ergaben für die Probe DH5 α -pCMV-lacZ-A2 eine Kopienzahl von 6960 Plasmiden pro Zelle (Tab. 8.18), was ungefähr einer Verdopplung der densitometrisch bestimmten Plasmidmenge (2858 Plasmide pro Zelle (Tab. 8.16)) entspricht. Für die Probe DH5 α -pCMV-lacZ-A3 wurde eine Kopienzahl von 8453 Plasmiden (Tab. 8.18) ermittelt. Dieser Wert ist etwas geringer, als der zuvor, für diese Probe, gemessene PCN (11367 Plasmide pro Zelle (Tab. 8.16)). Zusammenfassend lässt sich festhalten, dass die mit Hilfe der sondenbasierten RT-PCR ermittelten Ergebnisse die Größenordnungen der zuvor gemessenen Plasmidkopienzah-

len für DH5 α -pCMV-lacZ bestätigen. Aus Kostengründen musste auf die externe Vermessung aller 30 Proben durch die Firma CARPEGEN verzichtet werden.

Es konnte, obwohl die gemessenen Plasmidkopienzahlen nicht mit den von Cooper (2004) angegebenen Werten übereinstimmen, aufgrund der Messungen der Firma CARPEGEN bestätigt werden, dass eine derart große Menge an Plasmiden in der Zelle vorkommt. Dies wird zusätzlich anhand der von Atlung *et al.* (1999) gemachten Beobachtungen deutlich. Atlung *et al.* (1999) verwendeten in ihren Versuchen ein *high copy* Plasmid, wobei es sich um ein pBR322-Plasmid handelt, in dem sie das *rom*-Gen deletiert haben (pBR322 – *rom*⁻). Dabei haben sie festgestellt, dass sich die Plasmidkopienzahl, bei einer Zunahme der Generationszeit von 22 min auf 210 min, ungefähr um das 12-fache erhöht (Atlung *et al.*, 1999). Der in dieser Arbeit verwendete *high copy* Plasmid ist ein pUC-Plasmid und unterscheidet sich nur durch eine zusätzliche Punktmutation im RNAII-Transkript (Lin-Chao *et al.*, 1992), (Yanisch-Perron *et al.*, 1985) von dem von Atlung *et al.* (1999) verwendeten pBR322 – *rom*⁻. Nach Cooper (2004) hat ein pUC-Plasmid eine Kopienzahl von 500 - 700 Plasmiden pro Zelle. Unter der Annahme, dass sich bei einer Generationszeit von 22 min, 500 - 700 Plasmide in einer Zelle befinden, liegen, den Erkenntnissen Atlung *et al.* (1999) zur Folge, bei einer Generationszeit von 210 min zwischen 6000 und 8400 Plasmide in einer Zelle vor. Das würde für die in dieser Arbeit gemessenen Plasmidkopienzahlen bedeuten, dass sie durchaus realistisch sind, da sich das Zellwachstum ebenfalls durch sehr lange Generationszeiten auszeichnet. Ebenso wird an dieser Rechnung deutlich, dass große Schwankungen in den Plasmidkopienzahlen nicht ungewöhnlich sind, da selbst die rein rechnerisch ermittelten Kopienzahlen sich in einer Spannweite von 200 Plasmiden bei einer Generationszeit von 22 min bzw. in einer Spannweite von 2400 Plasmiden bei einer Generationszeit von 210 min bewegen. Dieses Erkenntnis lässt die hohen Standardabweichungen der Messungen durch die Plasmid Factory nicht mehr so überraschend erscheinen. Zusätzlich spricht das sehr langsame Wachstum der Kultur für die hohen Plasmidkopienzahlen, da nach Bremer & Lin-Chao (1986) mit langsamem Wachstum ein hoher Plasmidgehalt einhergeht.

Bei näherer Betrachtung der gemessenen Plasmidkopienzahlen für jede einzelne Probe (Tab. 8.14 und Tab. 8.16) fällt neben dem sehr hohen Plasmidgehalt pro Zelle auf, dass die Plasmidmengen sowohl für DH5 α -pSUP 201-3, als auch für DH5 α -pCMV-lacZ innerhalb der biologischen Replikate stark schwanken. Die Ursache für diese großen Unterschiede und der damit verbundenen hohen Standardabweichungen liegt möglicherweise an der Messmethode selbst, da die densitometrische Bestimmung des gDNA:pDNA-

Verhältnisses anhand der Gelbandenhelligkeit aufgrund von Sättigungseffekten fehlerbehaftet ist. Um zum einen diese hohen Schwankungen zu untersuchen, und zum anderen zur zusätzlichen Überprüfung der hohen Plasmidkopienzahlen in DH5 α -pCMV-lacZ, wurden alle Proben erneut mittels HPLC vermessen. Die Ergebnisse der HPLC-Messungen zeigen eine deutliche geringere Schwankungen innerhalb der biologischen Replikate, als die vorigen densitometrischen Messungen (Tab. 8.19). Des Weiteren fällt auf, dass die, unter Anwendung der HPLC ermittelten, durchschnittlichen Plasmidkopienzahlen noch größer sind, als die der densitometrischen Bestimmung und die der RT-PCR-Messung. Die Ursache könnte in der Sensitivität des UV-Detektors der HPLC liegen. Zur Berechnung der Plasmidkonzentration wurde zunächst in dem gemessenen Chromatogramm die Fläche des pDNA-Peaks mit Hilfe der Xcalibur-Software bestimmt. Im nächsten Schritt wurde dann anhand der Eichgerade die entsprechende Plasmidkonzentration ermittelt. Bei einigen Proben besaßen die pDNA-Peaks keine Peakform, sondern hatten eine eckige, treppenförmige Kontur. Diese Form deutet darauf hin, dass die in den Zellen enthaltenen Plasmidkonzentrationen zu niedrig sind, als dass der UV-Detektor sie hinreichend detektieren kann. Für Peaks, die keine Peakform, sondern diese eckige, treppenförmige Kontur aufwiesen, konnten daher die Peakflächen nicht immer eindeutig bestimmt werden, wodurch diese sehr hohen Plasmidkopienzahlen entstanden sein könnten. Aufgrund dieser Unsicherheiten wurden die Daten der densitometrischen Messung, trotz der hohen Standardabweichungen, für das dynamische Modell eingesetzt.

11.3 SIMULATIONEN DER PLASMIDREPLIKATIONSKONTROLLE

11.3.1 Simulationen zur Parameteranpassung

Bevor das dynamische Modell für prädiktive Simulationen der Plasmidreplikationskontrolle verwendet werden konnte, mussten zunächst Simulationen zur Parameteranpassung durchgeführt werden. Diese dienten dazu, unbekannte Parameter, zu denen weder experimentelle Daten, noch Literaturangaben vorhanden waren, abzuschätzen. Für die Simulationen zur Parameteranpassung wurde sowohl für DH5 α -pSUP 201-3, als auch für DH5 α -pCMV-lacZ das experimentelle Datenset von Messzeitpunkt T_2 verwendet. Das Modell wurde dabei unter variierenden Parametern an die für das Datenset von T_2 gemessene Plasmidproduktion angepasst, so dass auf diese Weise die Parameter k_3

(Bindekonstante von RNAI an Plasmid-gebundene RNAII) und k_{26} (Degradationskonstante einer unbeladenen tRNA) für beide Stämme und k_4 (Dissoziationskonstante des instabilen Komplexes aus Plasmid-gebundener RNAII und RNAI) nur für den *high copy* plasmidtragenden Stamm abgeschätzt werden konnten. Die betreffenden Parameter sind in Tabelle 8.20 mit einem „c“ gekennzeichnet.

Im Anschluss an die Parameteranpassung erfolgte die Validierung des Modells. Dazu wurde die Plasmidreplikationskontrolle für eine einzelne Zelle simuliert, wobei die für die Zeitpunkte T_1 bis T_3 experimentell bestimmten Daten als Anfangsbedingungen eingesetzt wurden. Vergleicht man die simulierten Plasmidkonzentrationen von DH5 α -pSUP 201-3 mit den gemessenen, lässt sich erkennen, dass die in *in vivo* stattfindende Plasmidreplikationskontrolle durch die Simulationen gut wiedergegeben wird (Tab. 8.21). Die Differenz zwischen den simulierten und den durchschnittlichen experimentell bestimmten Plasmidkopienzahlen beträgt maximal 3 Plasmide (49 Plasmide - 46 Plasmide), was einem prozentualen Fehler von 6.1 % entspricht. Damit liegt sie deutlich innerhalb der Spanne, die seitens der durchschnittlich gemessenen PCN durch die Fehlerbalken entsteht. Daher ist sie vernachlässigbar gering. Bei DH5 α -pCMV-lacZ wird der gemessene Plasmidgehalt ebenfalls in geeigneter Weise durch die Simulationen repräsentiert. Im Vergleich zu DH5 α -pSUP 201-3 sind die Unterschiede zwischen den simulierten und den durchschnittlichen experimentell bestimmten Plasmidkopienzahlen mit 384 Plasmide (5805 Plasmide - 5421 Plasmide) wesentlich größer (Tab. 8.22). Die Differenz von 384 Plasmiden entspricht einem prozentualen Fehler von 6.6 %. Betrachtet man nun die prozentualen Fehler, stellt man fest, dass die Unterschiede zwischen den simulierten und den durchschnittlichen experimentell bestimmten Plasmidkopienzahlen, prozentual gesehen, bei DH5 α -pSUP 201-3 und DH5 α -pCMV-lacZ gleich groß sind. Man kann daher festhalten, dass das Modell die Plasmidreplikationskontrolle auch für DH5 α -pCMV-lacZ in geeigneter Weise repräsentiert. Dafür spricht ebenfalls, dass die simulierten Plasmidkopienzahlen, wie auch bei DH5 α -pSUP 201-3, innerhalb der Spanne liegen, die seitens der durchschnittlich gemessenen PCN durch die Fehlerbalken entsteht.

Eine weitere Überprüfung des Modellverhaltens erfolgte durch die Simulation des Zusammenhangs zwischen Wachstumsrate und Plasmidproduktion für DH5 α -pSUP 201-3 und dem anschließenden Vergleich mit den experimentellen Ergebnissen von Atlung *et al.* (1999). Der Vergleich zeigte, dass die simulierten Plasmidkopienzahlen mit 49, 47 und 49 (Tab. 8.21) bei Generationszeiten um 150 min innerhalb der von Atlung *et al.* (1999) berechneten Spannweite von 45 bis 60 Plasmiden liegen. Das von Atlung *et al.*

(1999) beobachtete Verhalten wird somit durch das Modell wiedergegeben. Für DH5 α -pCMV-lacZ konnte dieser Vergleich nicht durchgeführt werden, da zu diesen Zweck keine experimentellen Daten zur Verfügung stehen.

11.3.2 Prädiktive Simulationen

Die prädiktiven Simulationen wurden durchgeführt, um der Fragestellung nachzugehen: Welchen Effekt hätte es auf die Plasmidproduktion, würde man tRNA-Moleküle derart modifizieren, dass die Beladung mit einer Aminosäure durch die Aminoacyl-tRNA-Synthetase nicht mehr möglich wäre?

Zur Beantwortung der Frage wurden folgende drei Bedingungen simuliert: Wachstum unter normalen Nährstoffbedingungen, Wachstum unter Mangelbedingungen und Wachstum mit modifiziertem tRNA-Gen ohne Mangelbedingungen (Abb. 8.10). Für die Simulationen wurde der *high copy* plasmidtragende Stamm gewählt, da für eine industrielle Plasmidproduktion mit dem Ziel der maximalen Plasmidausbeute in der Regel nur *high copy* Plasmide in Frage kommen. Es ist allerdings auch möglich, die Simulationen für DH5 α -pSUP 201-3 durchzuführen, wobei qualitativ das gleiche Ergebnis zu beobachten ist. Als Anfangsbedingungen wurden die experimentellen Daten des dritten Erntezeitpunkts verwendet, da die Plasmidreplikation im Verlauf der logarithmischen Phase zunehmen soll (Persönliche Kommunikation mit Martin Schleaf).

Die in Abbildung 8.10 dargestellten Simulationen zeigen, dass die vorhergesagte Plasmidproduktion unter Nährstoffknappheit am geringsten ist. Im Gegensatz dazu würde das Vorhandensein von tRNA-Molekülen, die nicht mehr durch Aminoacyl-tRNA-Synthetasen mit einer Aminosäure beladen werden können, die Plasmidproduktion gegenüber normalen Nährstoffbedingungen erhöhen. Das bedeutet, mit Hilfe des Modells wird die größte Plasmidproduktion vorausgesagt, wenn man ein Gen einfügt, das für eine modifizierte tRNA kodiert.

Um diese Vorhersage experimentell zu überprüfen, wurden zwei modifizierte Leucin-tRNA-Gene konstruiert, die beide ein unbeladbares tRNA-Molekül als Genprodukt bilden. Die Basis dieser modifizierten Gene ist das *leuQ*-Gen, das für eine Leucin-tRNA mit dem Anticodon 'CAG' kodiert. Bei der Konstruktion der modifizierten Leucin-tRNA-Gene wurden die Bereiche des *leuQ*-Gens entfernt, die für die Erkennungsstellen der Leucyl-tRNA-Synthetase kodieren und zugleich für die Plasmidreplikationskontrolle unwichtig sind. Die daraus resultierenden Gene, namens *Leucin-tRNA-RNAI*

und *Leucin-tRNA-WT*, kodieren jeweils eine unbeladbare Leucin-tRNA, die sich in ihrer Anticodon-Schleife unterscheiden (Abb. 8.13 und Abb. 8.12). Die beiden Gene wurden jeweils, anstelle des *asnT*-Gens, in eine *asnT*-Transkriptionseinheit eingebaut. In den darauffolgenden Schritten wurden diese beiden Transkriptionseinheiten von Sebastian Wiebe, im Rahmen seiner Masterarbeit, über Tn5 Transposon-Mutagenese in das *E. coli* Chromosom inseriert. Anschließend führte Sebastian Wiebe mit den dabei entstandenen Stämmen, *DH5 α -Leucin-tRNA-RNAI⁺* und *DH5 α -Leucin-tRNA-WT⁺*, erste Experimente durch, um sie bezüglich ihrer Plasmidproduktion zu untersuchen. In diesen Versuchen sollte überprüft werden, ob die inserierten modifizierten *Leucin-tRNA*-Gene einen Einfluss auf die Plasmidproduktion haben. Dabei stellte sich heraus, dass die Plasmidproduktion in *DH5 α -Leucin-tRNA-RNAI⁺* im Vergleich zum Wildtyp gesunken ist. Das bedeutet, die von dem *Leucin-tRNA-RNAI*-Gen kodierten tRNA-Moleküle üben einen negativen Effekt auf die Plasmidproduktion aus. Im Gegensatz dazu konnte man bei *DH5 α -Leucin-tRNA-WT⁺* eine leichte Zunahme der Plasmidproduktion erkennen. Aufgrund dieser ersten Versuche lässt sich daher schlussfolgern, dass die von dem *Leucin-tRNA-WT*-Gen kodierten tRNAs die pDNA-Synthese fördern. Diese Ergebnisse sollen aber mit weiteren Versuchen bestätigt und statistisch abgesichert werden. Es wird daher an dieser Stelle für eine genauere Betrachtung auf die Masterarbeit von Sebastian Wiebe verwiesen. Sollten sich die Ergebnisse bezüglich der erhöhten Plasmidsynthese bestätigen, wäre dieser Stamm ein potentieller Kandidat für die industrielle Plasmidproduktion.

Der Stamm *DH5 α -Leucin-tRNA-RNAI⁺* spielt im Hinblick auf die Produktion von Plasmid-DNA keine Rolle. Er könnte aber dennoch bei der Produktion rekombinanter Proteine von industriellem Nutzen sein. Bei der Produktion rekombinanter Proteine ist das Primärziel die Produktion der größtmöglichen Proteinmenge, wobei man stets auf dem schmalen Grat zwischen einer hohen Expression des rekombinanten Proteins und der metabolischen Kapazität des Wirts wandelt (Grabherr & Bayer, 2002). Das bedeutet, dass für die Produktion rekombinanter Proteine eine zu hohe intrazelluläre Plasmidkonzentration aufgrund der damit verbundenen metabolischen Belastung für den Wirtsstamm von Nachteil (Grabherr & Bayer, 2002) ist, was wiederum den Stamm *DH5 α -Leucin-tRNA-RNAI⁺* für die Industrie interessant machen könnte.

DAS METABOLOM UND DER INTRAZELLULÄRE GLUKOSEFLUSS IN ABHÄNGIGKEIT DES PLASMIDGEHALTS

12.1 DAS METABOLOM IN ABHÄNGIGKEIT DES PLASMIDGEHALTS

Neben der Modellierung der Plasmidreplikationskontrolle sollte in dieser Arbeit im Rahmen von Metabolomanalysen der Frage nachgegangen werden, welchen Einfluss der Plasmidgehalt einer Zelle auf ihre Metabolitzusammensetzung hat. Der Fokus dieser Analysen lag auf Unterschieden innerhalb der annotierten und somit bekannten Metabolite. Zu den dabei betrachteten Metaboliten gehören die detektierbaren Metabolite der Glykolyse (Glukose, Glukose-6-phosphat, Fruktose-6-phosphat, Glycerinaldehyd-3-phosphat, 2-Phosphoglycerat, 3-Phosphoglycerat, Phosphoenolpyruvat, Pyruvat), des Citratzyklus (Citrat, cis-Aconitat, 2-Oxoglutarat, Succinat, Fumarat, Malat), des Pentosephosphatwegs (Ribose, Ribose-5-phosphat, Xylose) sowie die Aminosäuren Glutamat, Glutamin, Glycin und Aspartat. Diese Metabolite wurden innerhalb der 40 annotierten Metabolite zur näheren Betrachtung ausgewählt, da sie direkt oder zum Teil entfernt für die Nukleotidbiosynthese relevant sind.

Beim Vergleich der Metabolite des Citratzyklus innerhalb der betrachteten drei Stämme DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP fällt auf, dass sich die Stämme am deutlichsten in der α -Ketoglutarat-Konzentration unterscheiden (α -Ketoglutarat = 2-Oxoglutarat). Der Stamm DH5 α -pCMV-GFP weist dabei von den drei Stämmen die höchste α -Ketoglutarat-Konzentration auf (Abb. 9.3). Eine mögliche Erklärung für die erhöhte Konzentration ist, dass in DH5 α -pCMV-GFP mehr α -Ketoglutarat verbraucht wird, um daraus Glutamin zu synthetisieren. Glutamin wird in der Reaktion der Glutamin-Phosphoribosyl-Amidotransferase zur Herstellung von Purinen benötigt (Berg *et al.*, 2007). Dabei wird vom Phosphoribosylpyrophosphat (PRPP), das die Grundstruktur zum Aufbau von Purinbasen darstellt, unter Glutaminverbrauch die Pyrophosphatgruppe durch eine Aminogruppe ausgetauscht, so dass Phosphoribosylamin und Glutamat entsteht (Berg *et al.*, 2007). Diese Erklärung wird durch die Ergebnisse der Messung von Glutamin (Abb. 9.9) unterstützt, da in DH5 α -pCMV-GFP kein Glutamin

detektiert wurde, was auf einen hohen Glutaminverbrauch hindeutet. Gegen einen Zusammenhang zwischen den hohen α -Ketoglutarat- und Glutaminkonzentrationen in DH5 α -pCMV-GFP mit der Nukleotidbiosynthese spricht jedoch die geringe Glutamatkonzentration in DH5 α -pCMV-GFP im Vergleich zu den anderen Stämmen. Bei einem bestehenden Zusammenhang würde man für DH5 α -pCMV-GFP eine höhere Konzentration erwarten, als bei DH5 α und DH5 α -pSUP 201-3, da bei einem erhöhten Glutaminverbrauch im Rahmen der Purinbiosynthese auch eine erhöhte Glutamatbildung zu sehen sein müsste. Dennoch sollte der Zusammenhang nicht komplett ausgeschlossen werden, da man bei der Bewertung berücksichtigen sollte, dass sowohl α -Ketoglutarat, als auch Glutamat und Glutamin nicht nur für die Synthesen von Nukleotiden benötigt werden, sondern auch Vorstufen für viele weitere Stoffwechselprodukte in der Zelle sind, wie z.B. für andere Aminosäuren (Prolin und Arginin) und Proteine. Dieser Aspekt ist auch bei der Betrachtung der relativen Aspartat- und Glycinkonzentrationen zu beachten.

Anhand der Ergebnisse in den Abbildungen 9.8 und 9.9 lässt sich bei DH5 α -pCMV-GFP eine etwas geringere Glycinkonzentration erkennen, als bei DH5 α und DH5 α -pSUP 201-3. Aspartat hingegen ist nur in DH5 α -pCMV-GFP detektierbar. Wie im Fall von α -Ketoglutarat, Glutamat und Glutamin lässt sich auch hier ein Zusammenhang mit der Nukleotidbiosynthese herstellen: Glycin wird bei der Synthese von Purinbasen verbraucht (Berg *et al.*, 2007), was die geringere Glycinkonzentration in DH5 α -pCMV-GFP im Vergleich zu den Stämmen DH5 α und DH5 α -pSUP 201-3 erklären könnte. Aspartat wird sowohl in der Purin-, als auch in der Pyrimidinsynthese verbraucht (Berg *et al.*, 2007), so dass man auch hier für DH5 α -pCMV-GFP die geringste Konzentration erwarten würde. Die Messungen zeigen jedoch, dass genau das Gegenteil der Fall ist: DH5 α -pCMV-GFP besitzt im Vergleich die höchste Aspartatkonzentration. Aber wie bereits zuvor erwähnt, sollte man bei der Interpretation der Ergebnisse berücksichtigen, dass Aspartat und Glycin nicht nur für die Synthesen von Nukleotiden benötigt werden, sondern auch Vorstufen für viele weitere Stoffwechselprodukte in der Zelle sind, so dass man nicht eindeutig von den relativen Aspartat- und Glycinkonzentrationen auf die Nukleotidbiosynthese schlussfolgern kann.

Führt man nun mit der Betrachtung der Messungen zu den Metaboliten des Citratzyklus fort, lassen sich neben dem Unterschied in der α -Ketoglutarat-Konzentration weitere Unterschiede bei Citrat, Fumerat und Malat beobachten. Der Fumerat- und Malatgehalt ist bei DH5 α und DH5 α -pSUP 201-3 ähnlich und bei DH5 α -pCMV-GFP am niedrigsten. Mit der relativen Citratkonzentration verhält es sich genau umgekehrt. Bei DH5 α und

DH5 α -pSUP 201-3 ist sie ebenfalls ähnlich, bei DH5 α -pCMV-GFP jedoch am höchsten. Diese Unterschiede könnten aufgrund des unterschiedlichen Plasmidgehalt entstehen, lassen sich jedoch nach derzeitigem Wissensstand nicht direkt mit der Nukleotidbiosynthese in Zusammenhang bringen. Innerhalb der Metabolite Succinat und cis-Aconitat sind keine nennenswerten Unterschiede zu sehen. Hier lässt sich allgemein nur festhalten, dass die Menge an Succinat in allen Stämmen sehr hoch und die cis-Aconitatkonzentration sehr gering ist.

Beim Vergleich der Metabolite der Glykolyse ist besonders auffällig, dass die Menge an Glycerinaldehyd-3-phosphat in allen Stämmen mit Abstand am größten ist. Bei näherer Betrachtung ist dies jedoch nicht überraschend, da bei den Metabolomversuchen Glycerin als Kohlenstoffquelle eingesetzt wurde. Nach Martínez-Gómez *et al.* (2012) wird Glycerin nach Aufnahme in die Zelle in Glycerinaldehyd-3-phosphat umgewandelt und tritt so in die Glykolyse ein.

Des Weiteren fällt auf, dass die Metabolitkonzentrationen von Glycerinaldehyd-3-phosphat aufwärts der Glykolyse bei DH5 α -pCMV-GFP im Vergleich zu DH5 α und DH5 α -pSUP 201-3 am größten sind. Das gilt insbesondere für die Konzentration von Glukose-6-phosphat. Die Metabolitkonzentrationen von Glycerinaldehyd-3-phosphat abwärts der Glykolyse sind dagegen bei DH5 α -pCMV-GFP, bis auf die von Pyruvat, am geringsten. Diese Unterschiede sind durch die erhöhte pDNA-Synthese in DH5 α -pCMV-GFP zu erklären. Es ist anzunehmen, dass in diesem Stamm ein größerer Teil des aufgenommenen Glycerins über die Reaktionen der Gluconeogenese in den Pentosephosphatweg fließt, um als Ribose-5-Phosphat der Nukleotidbiosynthese zur Verfügung zu stehen. Diese Erklärung wird durch die Ergebnisse der Messungen zu den Fünffach-Zuckern unterstützt (Abb. 9.7), wobei die Fünffach-Zucker zu den relevantesten Metaboliten für die Nukleotidbiosynthese zählen.

Betrachtet man die relativen Konzentrationen der Pentosen, Ribose, Ribose-5-phosphat, Xylose sieht man zum Teil deutliche Unterschiede zwischen allen drei Stämmen. Der Ribosegehalt ist in dem plasmidfreien Stamm DH5 α am höchsten, gefolgt von DH5 α -pSUP 201-3, und am wenigsten Ribose wurde in DH5 α -pCMV-GFP gemessen. Dasselbe gilt für Xylose, allerdings ist hier der Unterschied zwischen DH5 α und DH5 α -pSUP 201-3 wesentlich geringer. Bei Ribose-5-phosphat sieht man einen deutlichen Unterschied zwischen den plasmidtragenden Stämmen und DH5 α . DH5 α besitzt die geringste Ribose-5-phosphat-Konzentration, bei DH5 α -pSUP 201-3 ist sie deutlich höher und bei DH5 α -pCMV-GFP am höchsten, wobei der Unterschied zwischen den beiden

plasmidtragenden Stämmen gering ist. Die höhere Ribose-5-phosphat-Konzentration in DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP sowie die mit dem Plasmidgehalt abnehmenden Ribose- und Xylosekonzentrationen, sind mit dem erhöhten Nukleotidbedarf in DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP zu begründen. Je größer der Plasmidgehalt der Zelle, desto geringer sind die Ribose- und Xylosekonzentrationen, da beide Metabolite bei erhöhtem Nukleotidbedarf zu Ribose-5-phosphat umgewandelt werden. Die Ribose-5-phosphat kann nach Umwandlung in PRPP in die Nukleotidbiosynthese einfließen.

Zurückkehrend zu den Metaboliten der Glykolyse soll an dieser Stelle der Fokus auf die Metabolite 2-Phosphoglycerat, 3-Phosphoglycerat, Phosphoenolpyruvat und Pyruvat gerichtet werden. Wie bereits erwähnt, sind die Metabolitkonzentrationen von Glycerinaldehyd-3-phosphat abwärts der Glykolyse, mit Ausnahme von Pyruvat, bei DH5 α -pCMV-GFP am geringsten. Man kann davon ausgehen, dass der hohe Pyruvatgehalt nicht aus der Glykolyse stammt, sondern über die Reaktion des NADP⁺-abhängigen Malatenzyms aus Malat gebildet wird. Die hohe Malatkonzentration in DH5 α -pCMV-GFP unterstützt diese Vermutung, da dies zeigt, dass für die Umwandlung von Malat zu Pyruvat ausreichend Substrat vorhanden ist.

Insgesamt betrachtet fällt auf, dass die Konzentrationsunterschiede der detektierbaren Metabolite des Citratzyklus, der Glykolyse und der Nukleotidbiosynthese innerhalb DH5 α und DH5 α -pSUP 201-3 nicht so deutlich ausfallen, wie die Unterschiede zu DH5 α -pCMV-GFP. Man hätte erwartet, dass sich alle Stämme aufgrund ihres unterschiedlichen DNA-Gehalts deutlich voneinander unterscheiden und DH5 α -pSUP 201-3 als *low copy*-plasmidtragender Stamm mit seinem Metabolitprofil zwischen den anderen beiden liegt. Dass dies nicht der Fall ist bedeutet, dass sich DH5 α und DH5 α -pSUP 201-3 in ihrer Metabolitzusammensetzung ähnlicher sein müssen, als DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP, obwohl DH5 α -pSUP 201-3 ebenfalls ein Plasmid trägt. Dies lässt sich dadurch erklären, dass DH5 α -pSUP 201-3 mit seinem Plasmidgehalt nicht genau zwischen DH5 α und DH5 α -pCMV-GFP liegt. DH5 α -pSUP 201-3 besitzt eine sehr viel geringere Plasmidkonzentration, was anhand der experimentell bestimmten Plasmidkopienzahlen für das Replikationsmodell deutlich wird (Tab. 8.15 und 8.17). Die Plasmidkonzentration wurde dabei zwar für das Plasmid pCMV-lacZ bestimmt, aber da sich die Plasmide pCMV-lacZ und pCMV-GFP nur durch den Austausch des *lacZ*-Gens durch das *gfp*-Gen voneinander unterscheiden, kann man davon ausgehen, dass der Stamm DH5 α -pCMV-GFP eine ähnlich hohe Plasmidkopienzahl besitzt, wie DH5 α -

pCMV-lacZ.

Betrachtet man den unterschiedlichen Plasmidgehalt der Stämme DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP im Verhältnis zur gesamten Biomasse der Zelle lässt sich schlussfolgern, dass bei einer Zelle mit dem *high copy* Plasmid der prozentuale DNA-Anteil an der gesamten Biomasse deutlich größer sein muss. Nach Neidhardt (1987) beträgt der Anteil genomischer DNA in einer *E. coli*-Zelle 3% der Gesamtbiomasse, was einer Anzahl von 4600 kb entspricht. Nimmt man nun für das *low copy* Plasmid pSUP 201-3 mit einer Größe von 7896 bp eine mittlere Kopienzahl von 47 an (Tab. 8.15), muss der Stamm DH5 α -pSUP 201-3 durchschnittlich 371 kb DNA pro Zelle ($7896 \text{ bp} \times 47 = 371 \text{ kb}$) zusätzlich produzieren. In Prozent ausgedrückt sind das 8 % mehr Desoxyribonukleosidtriphosphate (dNTPs), die die Zelle zusätzlich zu ihrem Chromosom synthetisieren muss. Im Fall des *high copy* Plasmids pCMV-GFP (3487 bp) bei einer angenommenen Kopienzahl von 2403 Plasmiden (Tab. 8.17, Zeitpunkt T₂) macht die Anwesenheit des Plasmids eine zusätzliche DNA-Menge von 8379 kb ($3487 \text{ bp} \times 2403 = 8379 \text{ kb}$) in der Zelle aus. Dies entspricht einem erhöhten dNTP-Bedarf pro Zelle von 182 %, was bedeutet, dass die Zelle 182 % mehr dNTPs, zusätzlich zu ihrem Chromosom, synthetisieren muss.

Vergleicht man nun den zusätzlichen DNA-Gehalt von 8 % in Anwesenheit von pSUP 201-3 mit dem zusätzlichen DNA-Gehalt von 182 % in Anwesenheit von pCMV-GFP, lässt sich feststellen, dass dieser deutlich näher an dem DNA-Gehalt von DH5 α liegt als an dem von DH5 α -pCMV-GFP. Dies erklärt, warum die Konzentrationsunterschiede der detektierbaren Metabolite des Citratzyklus, der Glykolyse und der Nukleotidbiosynthese innerhalb DH5 α und DH5 α -pSUP 201-3 geringer ausfallen als zwischen DH5 α und DH5 α -pCMV-GFP.

Trotz der Ähnlichkeit bezüglich einiger Metabolitpools existieren aber dennoch Unterschiede in der Metabolitzusammensetzung von DH5 α und DH5 α -pSUP 201-3. Dies wird anhand einer Hauptkomponentenanalyse deutlich. In den zuvor beschriebenen Analysen wurden nur die detektierbaren Metabolite des Citratzyklus, der Glykolyse und der Nukleotidbiosynthese betrachtet, da diese aufgrund der Literatur mit der Nukleotidbiosynthese in Zusammenhang gebracht werden können. Im Rahmen der GC-MS-Messungen lassen sich noch 40 weitere Metabolite identifizieren, die jedoch nach derzeitigem Wissensstand nicht mit der Nukleotidbiosynthese in Verbindung gebracht werden. Um dennoch den Einfluss aller annotierten Metabolite auf die Metabolitzusammensetzung zu untersuchen und so ein Gesamtbild zu bekommen, wurde eine Haupt-

komponentenanalyse durchgeführt (Abb. 9.10). Diese zeigt eine eindeutige Auftrennung von DH5 α , DH5 α -pSUP 201-3 und DH5 α -pCMV-GFP bezüglich ihres Plasmidgehaltes. Das Bild der Hauptkomponentenanalyse bestätigt somit die erwartete Unterscheidung der Stämme anhand ihres Metabolitprofils. Für die Metabolomanalysen bedeutet das, dass es, neben den zuvor betrachteten Metaboliten des Citratzyklus, der Glykolyse und der Nukleotidbiosynthese noch weitere Metabolite gibt, die bisher nicht mit der Synthese von DNA (insbesondere pDNA), in Verbindung gebracht wurden, aber dennoch zu der Auftrennung der Stämme in Abhängigkeit ihres Plasmidgehalts in der Hauptkomponentenanalyse beitragen.

Zusammenfassend zeigen die Metabolomanalysen, dass sich die drei Stämme, unter Berücksichtigung aller detektierbaren Metabolite, hinsichtlich ihrer Metabolitprofile deutlich voneinander unterscheiden. Betrachtet man nur die detektierbaren Metabolite des Citratzyklus, der Glykolyse und der Nukleotidbiosynthese zeigen die Stämme DH5 α und DH5 α -pSUP 201-3 nur geringe Unterschiede in ihrem Metabolitprofil, grenzen sich dabei aber deutlich von DH5 α -pCMV-GFP ab. Dabei unterscheiden sich die Stämme am deutlichsten in den Konzentrationen von Ribose, Ribose-5-phosphat, Xylose, was dadurch begründet werden kann, dass Ribose, Ribose-5-phosphat, Xylose die Metabolite sind, die über den Pentosephosphatweg am direktesten mit der Nukleotidbiosynthese in Verbindung stehen. Da das Metabolom aber nur eine Momentaufnahme der Metabolitkonzentrationen ist und nur Überlegungen, keine Aussagen zu den intrazellulären Metabolitflüssen erlaubt, wurde in dieser Arbeit zusätzlich der intrazelluläre Glukosefluss bei Plasmidproduktionsstämmen mittels ^{13}C -Flussanalysen genauer zu untersucht.

12.2 BETRACHTUNG DES INTRAZELLULÄREN GLUKOSEFLUSSES VIA ^{13}C -FLUSSANALYSEN IN ABHÄNGIGKEIT DES PLASMIDGEHALTS

Die Untersuchung des intrazellulären Glukoseflusses mittels ^{13}C -Flussanalysen wurde für die Plasmidproduktionsstämme W₃₁₁₀-pHN, VH₃₃-pHN, VH₃₃ Δ *recA*-pHN, VH₃₃ Δ (*recA deoR*)-pHN durchgeführt. Dafür wurde zunächst ein metabolisches Flussmodell erstellt, welches aus 56 Reaktionen besteht und die Glykolyse, den Pentosephosphatweg, den Citratzyklus und die Acetatbildungsreaktionen beinhaltet (Abb. 10.4). Mit Hilfe des Modells und den experimentellen Daten der ^{13}C -Markierungsexperimente kann unter Verwendung der OpenFlux-Software in Kombination mit MATLAB, Mathworks[®] der intrazelluläre Glukosefluss berechnet werden.

Die Berechnung des intrazellulären Glukoseflusses wurde für zwei Messungen durchgeführt, die sich in der optischen Dichte der Bakterienkulturen zum Zeitpunkt der Zellern- te unterscheiden. Die Zellen der ersten Messung wurden in der mittleren exponentiellen Wachstumsphase geerntet und die der zweiten Messung in der späten exponentiellen Wachstumsphase. Das Ziel dieser Flussmessungen war zu untersuchen, ob der Plas- midgehalt in der Zelle einen Einfluss auf die metabolische Flussverteilung hat. Dabei lag das Hauptaugenmerk auf der Verteilung des Glukoseflusses auf die Glykolyse und den Pentosephosphatweg, da über den Pentosephosphatweg mit Ribose-5-phosphat ein wichtiger Baustein zur Synthese von Nukleotiden gebildet wird.

In der folgenden Diskussion der metabolischen Flussmessungen werden nur die Stäm- me W₃₁₁₀-pHN, VH₃₃-pHN und VH₃₃ Δ (*recA deoR*)-pHN betrachtet. Die Ergebnisse des Stamms VH₃₃ Δ *recA*-pHN entsprechen in beiden Messungen nicht den erwarteten Ergebnissen und sind nicht mit den Daten der anderen Stämme in Einklang zu brin- gen. Für VH₃₃ Δ *recA*-pHN wurde ein metabolischer Fluss erwartet, der zwischen den berechneten Flüssen von VH₃₃-pHN und VH₃₃ Δ (*recA deoR*)-pHN liegt. Bereits vor der Kultivierung war nicht eindeutig klar, ob es sich bei dem Stamm um den richtigen VH₃₃ Δ *recA*-pHN handelt, da es Hinweise gibt, dass der Stamm vor dem Versand von Mexiko nach Bielefeld von einem Studenten vertauscht wurde. Aufgrund der Ergebnis- se der Flussmessungen und dem Verdacht, dass es der falsche Stamm sein könnte, wird davon ausgegangen, dass es sich nicht um den richtigen Stamm namens VH₃₃ Δ *recA*- pHN handelt. Die Ergebnisse von VH₃₃ Δ *recA*-pHN wurden daher bei der Auswertung nicht berücksichtigt.

Des Weiteren ist anzumerken, dass der Anteil der aufgenommenen Glukose, der über die Phosphoglukose-Isomerase-Reaktion in die Glykolyse fließt, größtenteils von der Software nicht eindeutig berechnet werden konnte. Dies zeigt sich daran, dass die be- rechneten Flüsse einen Wert von 0 % oder 100 % besitzen, was genau den Grenzen entspricht, innerhalb derer sich die Flüsse bewegen. Dabei entspricht 0 % der unte- ren Grenze und 100 % der oberen Grenze. Die entsprechenden Flüsse sind in den Abbildungen 10.5 und 10.6 mit einem „*“ markiert. Eine Überprüfung des metabo- lischen Modells auf etwaige Fehler und die Tatsache, dass die Flussberechnung der Phosphoglukose-Isomerase-Reaktion bei dem Wildtypstamm W₃₁₁₀ in der ersten Mes- sung erfolgreich war zeigt, dass das Problem mit hoher Wahrscheinlichkeit nicht bei dem metabolischen Modell liegt. Die Ursache für die nicht eindeutige Berechnung des Flusses der Phosphoglukose-Isomerase-Reaktion liegt vermutlich darin, dass die in den

Messdaten enthaltenen Informationen über die Isotopomer-Verteilungsmuster der gemessenen Aminosäuren nicht ausreichen, um eindeutig identifizieren zu können, über welche metabolischen Flüsse die Aminosäuren synthetisiert wurden. Derartige Schwierigkeiten bei der eindeutigen Berechnung von metabolischen Flüssen sind nicht ungewöhnlich (Wiechert, 2001). Dies gilt insbesondere für reversible Reaktionen, zu denen auch die Reaktion der Phosphoglukose-Isomerase gehört (Wiechert, 2001). Für eine genaue Berechnung der intrazellulären Flussverteilung sind somit weitere Kultivierungen auf ^{13}C -markierter Glukose notwendig, um zusätzliche Informationen über die Isotopomer-Verteilungsmuster der Aminosäuren zu erhalten. Daher wird in dieser Arbeit aufgrund der Schwierigkeiten bei der Berechnung des metabolischen Flusses der Phosphoglukose-Isomerase-Reaktion in der folgenden Diskussion nicht näher auf die einzelnen prozentualen Werte eingegangen, sondern es wird lediglich untersucht, ob sich die Stämme W₃₁₁₀-pHN, VH₃₃-pHN und VH₃₃ Δ (*recA deoR*)-pHN bezüglich ihres Glukoseflusses in den Pentosephosphatweg unterscheiden.

Betrachtet man die metabolischen Flüsse der ersten Messung (Abb. 10.5) fällt auf, dass der Stamm mit der geringsten Plasmidproduktion (W₃₁₁₀-pHN), wie erwartet, auch den geringsten metabolischen Fluss in den Pentosephosphatweg in Richtung P_{5P} besitzt. Bei den Stämmen VH₃₃-pHN und VH₃₃ Δ (*recA deoR*)-pHN ist der Fluss mit 71.8 % und 67.3 % deutlich höher. Allerdings wurde für die Doppelmutante, im Vergleich zu VH₃₃-pHN, ein höherer metabolischer Fluss erwartet, da die Doppelmutante am meisten Plasmid-DNA von allen Stämmen produziert. Der unerwartet hohe Fluss in VH₃₃-pHN (bzw. der unerwartet niedrige Fluss in VH₃₃ Δ (*recA deoR*)-pHN) lässt sich aber möglicherweise durch die unterschiedlichen Wachstumsraten der Stämme erklären. Die untersuchten Stämme wachsen alle unterschiedlich schnell und besitzen verschiedene Maximum-OD₆₀₀-Werte, an denen die Kulturen stationär werden. Der Stamm VH₃₃ Δ (*recA deoR*)-pHN wächst z. B. deutlich langsamer als VH₃₃-pHN und erreicht eine geringere Maximum-OD₆₀₀. Daher ist davon auszugehen, dass die Doppelmutante mit einer geringeren Wachstumsrate wächst. Nach Kayser *et al.* (2005) ist der Glukosefluss in den Pentosephosphatweg bei kleinen Wachstumsraten geringer, als bei hohen Wachstumsraten (Kayser *et al.*, 2005). Es ist daher möglich, dass VH₃₃ Δ (*recA deoR*)-pHN bei gleicher Wachstumsrate wie VH₃₃-pHN einen höheren Glukosefluss in Richtung Pentosephosphatweg zeigen würde. Da es sich bei der Kultivierung für diese Versuche um eine Anzucht im Schüttelkolben handelt, war es nicht möglich, alle Stämme auf eine einheitliche Wachstumsrate einzustellen, so dass die Ernte-OD₆₀₀ derart gewählt wurde, dass

sich alle Stämme ungefähr in derselben Wachstumsphase befinden. Das bedeutet, zum Zeitpunkt der Zellernte wuchsen nicht alle Stämme mit derselben Wachstumsrate, befanden sich aber ungefähr in derselben Wachstumsphase. Kleine Unterschiede zwischen den gemessenen und den erwarteten Ergebnissen sind daher mit großer Wahrscheinlichkeit durch die uneinheitlichen Wachstumsraten bedingt. Es wäre daher von Vorteil die Ergebnisse mit einer *fed-batch*-Kultivierung abzusichern, da man so die Möglichkeit hat die Zellen bei einer einheitlichen Wachstumsrate zu ernten und keine Unterschiede in der metabolischen Flussverteilung aufgrund verschiedener Wachstumsraten entstehen können.

Betrachtet man den Glukosefluss in den Pentosephosphatweg der zweiten Messung, fällt auf, dass der Fluss bei VH33-pHN und VH33Δ (*recA deoR*)-pHN im Vergleich zur ersten Messung deutlich geringer ist. Dies liegt daran, dass der Erntezeitpunkt in der späten exponentiellen Wachstumsphase lag, und die Zellen sich schon am Übergang in die stationäre Phase befanden. Das bedeutet, dass sie zu diesem Zeitpunkt mit einer kleineren Wachstumsrate gewachsen sind als bei der ersten Messung. Bei geringen Wachstumsraten ist nach Kayser *et al.* (2005) auch ein geringerer Glukosefluss in den Pentosephosphatweg zu erwarten. Betrachtet man den metabolischen Fluss von W3110-pHN fällt auf, dass dieser geringfügig höher ist, als der metabolische Fluss von VH33-pHN. Dies lässt sich ebenfalls durch die Wachstumsrate erklären. Der Stamm W3110-pHN wächst von den untersuchten Stämmen am schnellsten und besaß im Vergleich zu den anderen beiden Stämmen zum Zeitpunkt der Zellernte noch eine höhere Wachstumsrate. Man kann davon ausgehen, dass der Fluss bei W3110-pHN bei gleicher Wachstumsrate kleiner ist, als bei VH33-pHN.

Neben dem Fluss in den Pentosephosphatweg in Richtung P5P sind die Flüsse aus ihm heraus in Richtung Glykolyse von Interesse. Für diese Flüsse lässt sich sowohl für Messung 1, als auch für Messung 2 festhalten, dass ungefähr ein Drittel der Glukose, die in Richtung P5P fließt ist, von dort über F6P bzw. T3P wieder in die Glykolyse eintritt.

Betrachtet man die metabolischen Flüsse von F6P und T3P entlang der Glykolyse in Richtung PEP lässt sich erkennen, dass sie mit steigendem Plasmidgehalt abnehmen. Dies gilt ebenfalls für die Flüsse des Citratzyklus. Ein möglicher Grund dafür ist, dass bei einem höheren Plasmidgehalt ein größerer Anteil der aufgenommenen Glukose im Pentosephosphatweg für die Synthese von Ribose-5-phosphat verwendet wird und somit weniger von der aufgenommenen Glukose über F6P und T3P in die Glykolyse und anschließend in den Citratzyklus fließt. Allerdings ist bei der Bewertung dieser Flüsse zu

berücksichtigen, dass Markierungsexperimente mit $1\text{-}^{13}\text{C}$ -markierter Glukose eine gute Auflösung der metabolischen Flüsse im oberen Teil der Glykolyse und dem Pentosephosphatweg bieten, für die Untersuchung des unteren Teils der Glykolyse und des Citratzyklus jedoch weniger geeignet sind (Wang *et al.*, 2006b). Für die Untersuchung des unteren Teils der Glykolyse und des Citratzyklus versprechen Versuche mit U- ^{13}C -markierter Glukose eine bessere Auflösung (Wang *et al.*, 2006b).

Trotz kleiner Unterschiede, bedingt durch die ungleichen Wachstumsraten zum Zeitpunkt der Zellernte, lässt sich zusammenfassend festhalten, dass die Plasmidproduktionsstämme mit einem höheren pDNA-Gehalt einen größeren Glukosefluss in den Pentosephosphatweg besitzen.

Vor einigen Jahren stellten Wang *et al.* (2006b) mittels ^{13}C -Flussanalysen, DNA Microarrays und Enzymaktivitätstests fest, dass die Anwesenheit eines *ColE1*-Plasmids den Stoffwechsel des *E. coli* Stamms BL21(DE3) beeinflusst. Dieser Stamm spielt zwar im Rahmen der Plasmidproduktion keine Rolle, aber dennoch ist es interessant, die Glukoseflüsse von den hier untersuchten Plasmidproduktionsstämmen mit denen von einem Nicht-Produktionsstamm zu vergleichen.

Bei ihren Markierungsexperimenten verwendeten Wang *et al.* (2006b) ein *ColE1*-Plasmid mit zwei Replikationsursprüngen als *high copy*-Plasmid mit einer Kopienzahl von 410 (Wang *et al.*, 2006b). Die Plasmidkopienzahl ist dabei jedoch mit Vorsicht zu betrachten, da sie anhand der gemessenen DNA-Konzentration des klaren Zelllysats bestimmt wurde. Es ist davon auszugehen, dass in dem klaren Zelllysats noch chromosomale DNA-Verunreinigungen vorhanden waren, die mitgemessen und fälschlicherweise als Plasmid-DNA angesehen wurde. Aus diesem Grund wird die tatsächliche Kopienzahl des von Wang *et al.* (2006b) verwendeten *high copy*-Plasmids mit großer Wahrscheinlichkeit geringer sein, als angegeben, so dass auch der Bedarf an Ribose-5-phosphat für die pDNA-Synthese geringer sein wird. Dies sollte man bei Vergleich der metabolischen Flüsse im Hinterkopf behalten. Aber trotz der möglicherweise nicht exakt bestimmten Plasmidkopienzahl lohnt sich ein Vergleich der Ergebnisse von Wang *et al.* (2006b) mit den für die Stämme W3110-pHN, VH33-pHN, und VH33 Δ (*recA deoR*)-pHN berechneten metabolischen Flüssen. Bei diesem Vergleich wird nur der Glukosefluss in den Pentosephosphatweg in Richtung P5P betrachtet, da dieser für die Plasmidproduktion von besonderer Bedeutung ist. Wang *et al.* (2006b) bestimmten den metabolischen Fluss für einen *high copy* plasmidtragenden BL21(DE3)-Stamm für zwei Wachstumsraten, die denen einer frühen stationären Phase entsprechen. Bei $\mu=0.29\text{ h}^{-1}$ bestimmten sie einen

Glukosefluss von 12 % und bei $\mu=0.2 \text{ h}^{-1}$ einen Fluss von 11 % (Wang *et al.*, 2006b). Im Vergleich zu den in dieser Arbeit bestimmten Glukoseflüssen in den Pentosephosphatweg für W_{3110} -pHN, VH_{33} -pHN, und $VH_{33}\Delta$ (*recA deoR*)-pHN (Abb. 10.5 und 10.6), ist der Fluss in dem Nicht-Plasmidproduktionsstamm BL21(DE) deutlich geringer. Obwohl das von Wang *et al.* (2006b) verwendete *high copy*-Plasmid mit großer Wahrscheinlichkeit eine geringere Kopienzahl besitzt, als das in dieser Arbeit verwendete pHN-Plasmid und obwohl die Wachstumsraten nicht exakt übereinstimmen, ist der Unterschied im Glukosefluss so deutlich, dass man dennoch davon ausgehen kann, dass die Plasmidproduktionsstämme aufgrund ihres höheren Plasmidgehalts einen größeren Glukosefluss in den Pentosephosphatweg besitzen, als der Nicht-Plasmidproduktionsstamm BL21(DE). Das bedeutet, der Plasmidgehalt einer Zelle hat einen Einfluss auf den Anteil der aufgenommenen Glukose, der in den Pentosephosphatweg einfließt und zur pDNA-Synthese verwendet wird.

Bei abschließender Betrachtung der ^{13}C -Flussmessungen zusammen mit den Metabolomanalysen lässt sich festhalten, dass der Plasmidgehalt einer Zelle einen Einfluss auf ihre Stoffwechselprozesse hat. Dies wird sowohl anhand der Metabolomanalysen durch Unterschiede in den Metabolitpools, als auch anhand der ^{13}C -Flussmessungen durch Unterschiede im Glukosefluss deutlich. Bei der Untersuchung des Metaboloms waren insbesondere Unterschiede in den Konzentrationen der Pentosen Ribose, Ribose-5-phosphat und Xylose zu messen. Die ^{13}C -Flussmessungen deuten daraufhin, dass die Plasmidproduktionsstämme mit einem höheren pDNA-Gehalt einen größeren Glukosefluss in den Pentosephosphatweg besitzen. Diese Ergebnisse werden durch die theoretische Berechnung unterstützt, dass man in Anwesenheit eines durchschnittlich großen *high copy* Plasmids (4600 bp) mit einer Kopienzahl von 1000 Plasmiden eine Erhöhung des metabolischen Flusses in die dNTP-Synthese um 100 % erwarten würde, da der dadurch entstehende zusätzliche dNTP-Bedarf ungefähr die Größe eines *E. coli*-Genoms ausmachen würde. ($1000 \times 4600 \text{ bp} = 4600 \text{ kb} \hat{=} 1 \text{ E. coli-Genom}$).

12.3 FAZIT UND AUSBLICK

In dieser Arbeit ist ein mathematisches Modell zur Simulation der *ColE1*-Plasmidreplikationskontrolle für einen *low copy* und einen *high copy* Plasmid erstellt worden. Mit Hilfe des Modell wurde gezeigt, dass das Einfügen des *Leucin-tRNA-WT*-Gens in das DH5 α -Chromosom die Plasmidproduktion *in silico* erhöht. Diese Vorhersage

wurde durch erste Versuche im Labor im Rahmen der Masterarbeit von Sebastian Wiebe bestätigt, wobei diese Ergebnisse noch durch weitere Experimente abzusichern sind. Zeigen diese Experimente ebenfalls eine Erhöhung der Plasmidkonzentration aufgrund des *Leucin-tRNA-WT*-Gens, hat der Kooperationspartner dieses Promotionsprojektes (Plasmid Factory, Bielefeld) großes Interesse daran, den Stamm *DH5 α -Leucin-tRNA-WT⁺* in einem industriellen Produktionsprozess zu testen.

Des Weiteren wurde in dieser Arbeit die Plasmidreplikation auf Metabolomebene durch Metabolitprofilanalysen und ¹³C-Flussmessungen untersucht. Dabei wurde anhand einer Hauptkomponentenanalyse gezeigt, dass sich die untersuchten Stämme aufgrund ihres Plasmidgehalts hinsichtlich ihrer Metabolitprofile deutlich voneinander unterscheiden. Besonders deutlich wurde dies bei Betrachtung der Konzentrationen von Ribose, Ribose-5-phosphat und Xylose. Mit Hilfe der ¹³C-Flussmessungen wurde untersucht, ob der Plasmidgehalt in der Zelle einen Einfluss auf die metabolische Flussverteilung hat. Dabei wurde der Glukosefluss in drei Plasmidproduktionsstämmen näher betrachtet, wobei der Fokus auf der Verteilung des Glukoseflusses auf die Glykolyse und den Pentosephosphatweg lag, da über den Pentosephosphatweg mit Ribose-5-phosphat ein wichtiger Baustein zur Synthese von Nukleotiden gebildet wird. Die Ergebnisse der ¹³C-Flussmessungen deuten daraufhin, dass die Plasmidproduktionsstämme mit einem höheren pDNA-Gehalt einen größeren Glukosefluss in den Pentosephosphatweg aufweisen. Wie in der Diskussion ausführlich beschrieben, gibt es derzeit aufgrund fehlender Informationen bezüglich der Isotopomer-Verteilungsmuster der Aminosäuren noch Schwierigkeiten den Glukosefluss über die Phosphoglukose-Isomerase-Reaktion in die Glykolyse eindeutig zu berechnen. Um diese Schwierigkeiten zu beheben, sollte die Kultivierung der Stämme mit ¹³C-markierter Glukose optimiert werden, indem man von der Schüttelkolbenkultivierung auf eine *Fed-Batch*-Kultivierung umstellt. Diese Art der Kultivierung erfolgt unter besser kontrollierbaren Bedingungen. Ein weiterer Vorteil der *Fed-Batch*-Kultivierung ist die Möglichkeit der Zellernte bei einer einheitlichen Wachstumsrate, so dass keine Unterschiede in der metabolischen Flussverteilung aufgrund verschiedener Wachstumsraten entstehen. Im Anschluss an die Kultivierung wird das Isotopomer-Verteilungsmuster der Aminosäuren gemessen, wobei auch hier über eine Optimierung des Derivatisierungsprotokolls nachzudenken ist, so dass neben den bisher detektierten Aminosäuren noch weitere Aminosäuren gemessen werden können. Zusammenfassend lässt sich festhalten, dass im Rahmen von Metabolitprofilanalysen und ¹³C-Flussmessungen gezeigt werden konnte, dass der Plasmidgehalt einer Zelle

einen Einfluss auf die Metabolitzusammensetzung und den intrazellulären Glukosefluss hat.

Aufgrund des in dieser Arbeit gezeigten Einflusses des Plasmidgehalts auf den Stoffwechsel der Zelle, wäre für die Fortsetzung dieser Arbeit die Erweiterung des Modells um die Metabolomebene, z. B. durch die Verfügbarkeit von Nukleotiden, ein interessanter Aspekt. Eine Möglichkeit dies umzusetzen wäre, in die Reaktion der Plasmidreplikation (hier Reaktion 7) Nukleotide als zusätzliches Substrat hinzuzufügen sowie Reaktionen zur Synthese und Degradierung der Nukleotide zu definieren. Auf diese Weise ließe sich über die Nukleotidkonzentration die Verfügbarkeit von Nukleotiden simulieren und deren Zusammenhang mit der Plasmidreplikationskontrolle durch die regulatorischen RNA-Moleküle namens RNAI und RNAII sowie die Regulation durch unbeladene tRNA-Moleküle *in silico* eingehender untersuchen.

Teil VII

ANHANG

LITERATURVERZEICHNIS

- Alberts, B., Johnson, A., Walter, P. & J, L. 2007. *Molecular Biology of the Cell*. 5th Revised edition (REV), Taylor & Francis.
- Asahara, H., Himeno, H., Tamura, K., Hasegawa, T., Watanabe, K. & Shimizu, M. 1993. Recognition nucleotides of escherichia coli trna^{Leu} and its elements facilitating discrimination from trna^{Ser} and trna^{Tyr}. *J. Mol. Biol.* 231, 219–229.
- Ataai, M. M. & Shuler, M. L. 1986. Mathematical model for the control of cole1 type plasmid replication. *Plasmid*, 16, 204–212.
- Atlung, T., Christensen, B. B. & Hansen, F. G. 1999. Role of the rom protein in copy number control of plasmid pbr322 at different growth rates in escherichia coli k-12. *Plasmid*, 41, 110–119.
- Bachmann, B. 1972. Pedigrees of some mutant strains of escherichia coli k-12. *Bacteriol. Rev.* 36, 525–557.
- Bauer, K., Bassat, A., Dawson, M., Puente, V. & Neway, J. 1990. Improved expression of human interleukin-2 in high-cell-density fermentor cultures of escherichia coli k-12 by a phosphotransacetylase mutant. *Appl. Environ. Microbiol.* 56, 1296–1302.
- Bednarz, H. 2012. Etablierung einer transposon-basierten methode zur fluoreszenten markierung und mikroskopischen aufklärung subzellulärer proteinlokalisierungen in sinorhizobium meliloti rm1021. *Dissertation*, .
- Berg, J. M., Tymoczko, J. L. & Stryer, L. 2007. *Biochemie, Kapitel 30*. Spektrum Akademischer Verlag.
- Bhagwat, A. S. & Person, S. 1981. Structure and properties of the region of homology between plasmids pmb1 and cole1. *Molecular and general genetics : MGG*, 182, 505–507.
- Bisswanger, H. 1994. *Enzymkinetik-Theorie und Methoden*. 2. Auflage, VCH Verlag.

- Borja, G. M., Meza Mora, E., Barrón, B., Gosset, G., Ramírez, O. T. & Lara, A. R. 2012. Engineering *Escherichia coli* to increase plasmid dna production in high cell-density cultivations in batch mode. *Microbial Cell Factories*, 11.
- Bossel, H. 2004. *Systeme, Dynamik, Simulation*. 3-8334-0984-3, Books on Demand GmbH, Norderstedt.
- Bower, D. M. & Prather, K. L. J. 2009. Engineering of bacterial strains and vectors for the production of plasmid DNA. *Applied microbiology and biotechnology*, 82 (5), 805–813.
- Brand, E., Ralla, K. & Neubauer, P. 2012. *Part One Upstream Technologies-Strategies for Plasmid DNA Production in Escherichia coli in Biopharmaceutical Production Technology* (ed. Ganapathy Subramanian. 1th edition, Wiley-VCH Verlag GmbH & Co. KGaA.
- Bremer, H. & Lin-Chao, S. 1986. Analysis of the physiological control of replication of *colE1*-type plasmids. *Journal of theoretical biology*, 123, 453–470.
- Brendel, V. & Perelson, A. S. 1993. Quantitative model of *colE1* plasmid copy number control. *J. Mol. Biol.* 229, 860–872.
- Brenner, M. & Tomizawa, J. 1991. Quantitation of *colE1*-encoded replication elements. *Proc. Natl. Acad. Sci. USA*, 88, 405–409.
- Bronstein, I. N., Semendjajew, K. A., Musiol, G. & Mühlig, H. 2000. *Taschenbuch der Mathematik*. 4., überarb. u. erw. Auflage, Harri Deutsch.
- Campbell, N. A. & Reece, J. B. 2003. *Biologie*. 6. Auflage, Spektrum Akademischer Verlag.
- Carnes, A. E. 2005. Fermentation design for the manufacture of therapeutic plasmid dna. *BioProcess Int.* Oktober, 2–7.
- Carnes, A. E., Hodgson, C. P. & Williams, J. A. 2006. Inducible *Escherichia coli* fermentation for increased plasmid dna production. *Biotechnol Appl Biochem*, 45 (7), 155–66.
- Chuang, H.-Y., Hofree, M. & Ideker, T. 2010. A decade of systems biology. *Annual Review of Cell and Developmental Biology*, 26, 721–744.
- Cooke, J. R., McKie, E. A., Ward, J. M. & Keshavarz-Moore, E. 2004. Impact of intrinsic dna structure on processing of plasmids for gene therapy and dna vaccines. *Biotechnol*, 114, 239–54.

- Cooper, J. 2004. *Biosensors - A Practical Approach*. 2nd Edition, Oxford University Press, USA.
- Datsenko, K. & Wanner, B. L. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 6640–6645.
- Dauner, M. 2000. Intracellular carbon flux analysis by ^{13}C -tracer experiments. Dissertation, ETH Zürich, .
- De Anda, R., Lara, A. R., Hernández, V., Hernández-Montalvo, V., Gosset, G., Bolívar, F. & Ramírez, O. T. 2006. Replacement of the glucose phosphotransferase transport system by galactose permease reduces acetate accumulation and improves process performance of *Escherichia coli* for recombinant protein production without impairment of growth rate. *Metabolic Engineering*, 8, 281–290.
- del Solar, G. & Espinosa, M. 2000. Plasmid copy number control: an ever-growing story. *Molecular Microbiology*, 37, 492–500.
- Delgado, J. & Liao, J. 1997. Inverse flux analysis for reduction of acetate excretion in *Escherichia coli*. *Biotechnol. Prog.* 13, 361–367.
- Dettmer, K., Aronov, P. A. & Hammock, B. D. 2007. Mass spectrometry-based metabolomics. *Mass Spectrom Rev.* 26, 51–78.
- Dettmer, K. & Hammock, B. D. 2004. Guest editorial metabolomics - a new exciting field within the 'omics' sciences. *Environmental Health Perspectives*, 112, 396–397.
- Duester, G., Campen, R. K. & Holmes, W. M. 1981. Nucleotide sequence of an *Escherichia coli* trnA (leu 1) operon and identification of the transcription promoter signal. *Nucleic Acids Research*, 9, 2121–2139.
- Edelstein-Keshet, L. 2005. *Mathematical Models in Biology*. 1. Auflage, SIAM: Society for Industrial and Applied Mathematics.
- Eisenberg, R. C. & Dobrogosz, W. J. 1967. Gluconate metabolism in *Escherichia coli*. *Journal of Bacteriology*, 93, 941–949.
- El-Mansi, E. M. T. & Holms, W. H. 1989. Control of carbon flux to acetate excretion during growth of *Escherichia coli* in batch and continuous cultures. *Journal of General Microbiology*, 135, 2875–2883.

- Farmer, W. & Liao, J. 1997. Reduction of aerobic acetate production by *escherichia coli*. *Appl. Environ. Microbiol.* 63, 3205–3210.
- Ferraro, B., Morrow, M. P., Hutnick, N. A., Shin, T. H., Lucke, C. E. & Weiner, D. B. 2011. Clinical applications of DNA vaccines: current progress. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 53 (3), 296–302.
- Fiehn, O. 2002. Metabolomics - the link between genotypes and phenotypes. *Plant Molecular Biology*, 48, 155–171.
- Fuchs, G. & Schlegel, H. G. 2007. *Allgemeine Mikrobiologie*. 8. Auflage, Georg Thieme Verlag.
- Fuhrer, T., Fischer, E. & Sauer, U. 2005. Experimental identification and quantification of glucose metabolism in seven bacterial species experimental identification and quantification of glucose metabolism in seven bacterial species. *J. Bacteriol.* 187, 1581–1590.
- Funahashi, A., Matsuoka, Y., Jouraku, A., Morohashi, M., Kikuchi, N. & Kitano, H. 2008. Celldesigner 3.5: a versatile modeling tool for biochemical networks. *Proceedings of the IEEE*, 96, 1254–1265.
- Ghosh, S., Matsuoka, Y., Asai, Y., Hsin, K.-Y. & Kitano, H. 2011. Software for systems biology: from tools to integrated platforms. *Nat. Rev. Genet.* 12, 821 –832.
- Gotta, S. L., Miller, O. L. & French, S. L. 1991. rna transcription rate in *escherichia coli*. *J. Bacteriol.* 173, 6647–6649.
- Grabherr, R. & Bayer, K. 2002. Impact of targeted vector design on *cole1* plasmid replication. *TRENDS in Biotechnology*, 20, 257–260.
- Guzman, E. C., Carrillo, F. J. & Jimenez-Sanchez, A. 1988. Differential inhibition of the initiation of dna replication in stringent and relaxed strains of *escherichia coli*. *Genet. Res.* 51, 173–177.
- Haefner, J. W. 1996. *Modeling Biological Systems - Principles and Applications*. Chapman & Hall.
- Hanahan, D. 1983. Studies on transformation of *escherichia coli* with plasmids. *J. Mol. Biol.* 166, 557–580.

- Hecker, M., Riethdorf, S., Bauer, C., Schroeter, A. & Borriss, R. 1988. Expression of a cloned beta-glucanase gene from bacillus amyloliquefaciens in an escherichia coli rela strain after plasmid amplification. *Mol. Gen. Genet.* 215, 181–183.
- Herrera, E., Barcenas, P., Hernández, R., Méndez, A., Pérez-Ishiwara, G. & Barrón, B. 2010. A 176 amino acid polypeptide derived from the mumps virus hn ectodomain shows immunological and biological properties similar to the hn protein. *Virology Journal*, 195.
- Hiller, J., Franco-Lara, E. & Weuster-Botz, D. 2007. Metabolic profiling of escherichia coli cultivations: evaluation of extraction and metabolite analysis procedures. *Biotechnology letters*, 29, 1169–1178.
- Huang, X. & W, M. 1991. A time-efficient linear-space local similarity algorithm. *Adv. Appl. Math.* 12, 337–357.
- Ikemura, T. & Dahlberg, J. E. 1973. Small ribonucleic acids of escherichia coli - ii. non-coordinated accumulation during stringent control. *J. Biol. Chem.* 248, 5033–5041.
- Inokuchi, H. & Yamao, F. 1995. *tRNA: Structure, Biosynthesis and Function, Chapter 3*. American Society for Microbiology.
- Itoh, T. & Tomizawa, J. I. 1980. Formation of an rna primer for initiation of replication of cole1 dna by ribonuclease. *H. Proc. Natl. Acad. Sci. USA*, 77, 2450–2454.
- Jühling, F., Mörl, M., Hartmann, R. K., Sprinzl, M., Stadler, P. F. & Pütz, J. 2009. trnadb 2009: compilation of trna sequences and trna genes. *Nucleic Acids Research*, 37.
- Kayser, A., Weber, J., Hecht, V. & Rinas, U. 2005. Metabolic flux analysis of escherichia coli in glucose-limited continuous culture . i . growth-rate- dependent metabolic efficiency at steady state. *Microbiology*, 151, 693–706.
- Keasling, J. D. & Palsson, B. O. 1989a. On the kinetics of plasmid replication. *Journal of theoretical biology*, 136, 487–492.
- Keasling, J. D. & Palsson, B. O. 1989b. Colei plasmid replication : a simple kinetic description from a structured model. *Journal of theoretical biology*, 141, 447–461.
- Kessler, N., Neuweger, H., Bonte, A., Langenkämper, G., Niehaus, K., Nattkemper, T. W., & Goesmann, A. 2013. Meltdb 2.0 - advances of the metabolomics software system. *Bioinformatics*, Online veröffentlicht am 5. August 2013.

- Kherlopian, A. R., Song, T., Duan, Q., Neimark, M. a., Po, M. J., Gohagan, J. K. & Laine, A. F. 2008. A review of imaging techniques for systems biology. *BMC systems biology*, 2.
- Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S. & Fiehn, O. 2009. Fiehnlib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal. Chem.* 81, 10038–10048.
- Kirchner, O. & Tauch, A. 2003. Tools for genetic engineering in the amino acid-producing bacterium *corynebacterium glutamicum*. *J. Biotechnol.* 107, 287–299.
- Klipp, E., Liebermeister, W., Wierling, C., Kowald, A., Lehrach, H. & Herwig, R. 2009. *Systems Biology - A Textbook*. Wiley-Blackwell Verlag.
- Klumpp, S. 2011. Growth-rate dependence reveals design principles of plasmid copy number control. *PloS one*, 6, e20403.
- Komine, Y., Adachi, T., Inokuchi, H. & Ozeki, H. 1990. Genomic organization and physical mapping of the transfer rna genes in *escherichia coli* k12. *Journal of molecular biology*, 212, 579–598.
- Kruger, N. J., Troncoso-Ponce, M. A. & Ratcliffe, R. G. 2008. ¹h nmr metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. *Nat. Protoc.* 3, 1001–1012.
- Kutzler, M. A. & Weiner, D. B. 2008. DNA vaccines: ready for prime time? *Nature reviews. Genetics*, 9 (10), 776–788.
- Lahijani, R., Hulley, G., Soriano, G., Horn, N. A. & Marquet, M. 1996. High-yield production of pbr322-derived plasmids intended for human gene therapy by employing a temperature-controllable point mutation. *Hum Gene Ther*, 7, 1971–1980.
- Lara, A. R., Caspeta, L., Gosset, G., Bolívar, F. & Ramírez, O. T. 2008. Utility of an *escherichia coli* strain engineered in the substrate uptake system for improved culture performance at high glucose and cell concentrations: an alternative to fed-batch cultures. *Biotechnology and Bioengineering*, 99, 893–901.
- Larkin, D. C., Williams, A. M., Martinis, S. A. & Fox, G. E. 2002. Identification of essential domains for *escherichia coli* trna^{leu} aminoacylation and amino acid editing using minimalist rna molecules. *Nucleic acids research*, 30, 2103–2113.

- Lederberg, J. 1952. Cell genetics and hereditary symbiosis. *Physiol. Rev.* 32, 403–430.
- Lederberg, J. 1998. Personal perspective plasmid (1952-1997). *Plasmid*, 39, 1–9.
- Leighty, R. W. & Antoniewicz, M. R. 2012. Parallel labeling experiments with u-¹³c glucose validate e. coli metabolic network model for ¹³c metabolic flux analysis. *Metabolic Engineering*, 14, 533–541.
- Lin-Chao, S., Chen, W. T. & Wong, T. T. 1992. High copy number of the puc plasmid results from a rom/rop-suppressible point mutation in rna ii. *Molecular microbiology*, 6, 3385–3393.
- Listner, K., Bentley, L., Okonkowski, J., Kistler, C., Wnek, R., Caparoni, A., Junker, B., Robinson, D., Salmon, P. & Chartrain, M. 2006. Development of a highly productive and scalable plasmid dna production platform. *Biotechnol Prog*, 22 (7), 1335–1345.
- Lottspeich, F. & Engels, J. W. 2006. *Bioanalytik*. 2. Auflage, Spektrum Akademischer Verlag.
- Mackerey-Nagel 2013. *www.mn-net.com*. Large scale preparation of plasmid DNA.
- Majewski, R. A. & Domach, M. M. 1990. Simple constrained-optimization view of acetate overflow in e. coli. *Biotechnol Bioeng.* 35, 732–738.
- Martínez-Gómez, K., Flores, N., Castañeda, H. M., Martínez-Batallar, G., Hernández-Chávez, G., Ramírez, O. T., Gosset, G., Encarnación, S. & Bolivar, F. 2012. New insights into escherichia coli metabolism: carbon scavenging, acetate metabolism and carbon recycling responses during growth on glycerol. *Microbial Cell Factories*, 11.
- Mühlhardt, C. 2009. *Der Experimentator-Molekularbiologie/ Genomics*. 6. Auflage, Spektrum Akademischer Verlag.
- Neidhardt, H. F. C. 1987. *Escherichia Coli and Salmonella Typhimurium*. American Society for Microbiology.
- Neuweger, H., Albaum, S. P., Dondrup, M., Persicke, M., Watt, T., Niehaus, K., Stoye, J. & Goesmann, A. 2008. Meltdb: a software platform for the analysis and integration of metabolomics experiment data. *Bioinformatics*, 24, 2726–2732.
- Orth, J. D., Thiele, I. & Palsson, B. O. 2010. What is flux balance analysis? *Nature Publishing Group*, 28, 245–248.

- Peekhaus, N. & Conway, T. 1998. What's for dinner?: entner-doudoroff metabolism in *escherichia coli*. *Journal of Bacteriology*, 180, 3495–3502.
- Poolman, M. G., Assmus, H. E. & Fell, D. A. 2004. Applications of metabolic modelling to plant metabolism. *Journal of Experimental Botany*, 55, 1177–1186.
- Prather, K. J., Sagar, S., Murphy, J. & Chartrain, M. 2003. Industrial scale production of plasmid dna for vaccine and gene therapy: plasmid design, production, and purification. *Enzyme and Microbial Technology*, 33 (7), 865–883.
- Prazeres, D. M. F. 2011. *Plasmid Biopharmaceuticals - Basics, Applications and Manufacturing*. 1. Auflage, Wiley Verlag.
- Quek, L.-E., Wittmann, C., Nielsen, L. K. & Krömer, J. O. 2009. Openflux: efficient modelling software for ^{13}C -based metabolic flux analysis. *Microbial cell factories*, 8.
- Reiß, T. 2002. *Systeme des Lebens - Systembiologie*. Bundesministerium für Bildung und Forschung.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. 1989. *Molecular Cloning*. 2. Auflage, NY Cold Spring Harbor Laboratory Press.
- Sauer, U. 2004. High-throughput phenomics: experimental methods for mapping fluxomes. *Current opinion in biotechnology*, 15, 58–63.
- Sauer, U. 2006. Metabolic networks in motion: ^{13}C -based flux analysis. *Molecular systems biology*, 2.
- Schleef, M., ed. 2001. *Plasmids for Therapy and Vaccination*. 1. Auflage, WILEY-VHC-Verlag.
- Schuhmann, W. 1990. *Biologie bakterieller Plasmide*. 1. Auflage, Vieweg-Verlag.
- Senovilla, L., Vacchelli, E., Garcia, P., Eggermont, A., Fridman, W. H., Galon, J., Zitvogel, L., Kroemer, G. & Galluzzi, L. 2013. Trial watch: DNA vaccines for cancer therapy. *Oncoimmunology*, 2 (4), e23803.
- Simon, R., Priefer, U. & Pühler, A. 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Nature Biotechnology*, 1, 784–791.

- Singer, A., Eiteman, M. A. & Altman, E. 2009. Dna plasmid production in different host strains of *Escherichia coli*. *Journal of industrial microbiology and biotechnology*, 36 (4), 521–530.
- Soerensen, M. A., Elf, J., Bouakaz, E., Tenson, T., Sanyal, S., Björk, G. R. & Ehrenberg, M. 2005. Over expression of a tRNA^{Leu} isoacceptor changes charging pattern of leucine tRNAs and reveals new codon reading. *J. Mol. Biol.* 354, 16–24.
- Sohm, B., Sissler, M., Park, H., King, M. P. & Florentz, C. 2004. Conserved features of coordinately regulated *E. coli* promoters nucleic. *Journal of molecular biology*, 339, 17–29.
- Steffen, P., Voß, B., Rehmsmeier, M., Reeder, J. & Giegerich, R. 2006. RNashapes: an integrated RNA analysis package based on abstract shapes. *Bioinformatics*, 22.
- Tang, D.-C., DeVit, M. & Johnston, S. A. 1992. Genetic immunization is a simple method for eliciting an immune response. *Nature*, pp. 152–154.
- Tomizawa, J. I. 1984. Control of *colE1* plasmid replication: the process of binding of RNA *i* to the primer transcript. *Cell*, 38, 861–870.
- Tomizawa, J. I. 1985. Control of *colE1* plasmid replication: initial interaction primer transcript of RNA *i* and the *i* is reversible. *Cell*, 40, 527–535.
- Tomizawa, J. I. 1986. Control of *colE1* plasmid replication: binding of RNA *i* to RNA *ii* and inhibition of primer formation. *Cell*, 47, 89–97.
- Tomizawa, J. I. 1990b. Control of *colE1* plasmid replication interaction of Rom protein with an unstable complex formed by RNA *i* and RNA *ii*. *J. Mol. Biol.* 212, 695–708.
- Tomizawa, J. I. & Itoh, T. 1981. Plasmid *colE1* incompatibility determined by interaction of RNA *i* with primer transcript. *Proc. Natl. Acad. Sci. USA*, 78, 6096–6100.
- Travers, A. A. 1980. Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. *Journal of bacteriology*, 141, 973–976.
- Travers, A. A. 1984. Conserved features of coordinately regulated *E. coli* promoters nucleic. *Nucleic Acids Research*, 12, 2605–2618.
- von Hippel, P. H., Bear, D. G., Morgan, W. D. & McSwiggen, J. A. 1984. Protein-nucleic acid interactions in transcription: a molecular analysis. *Annual Review of Biochemistry*, 53, 389–446.

- Wang, Z., Le, G., Shi, Y., Węgrzyn, G. & Wróbel, B. 2002. A model for regulation of *cole1*-like plasmid replication by uncharged trnas in amino acid-starved *escherichia coli* cells. *Plasmid*, 47, 69–78.
- Wang, Z., Xiang, L., Shao, J., Węgrzyn, A. & Węgrzyn, G. 2006a. Effects of the presence of *cole1* plasmid dna in *escherichia coli* on the host cell metabolism. *Microbial cell factories*, 5.
- Wang, Z., Xiang, L., Shao, J., Węgrzyn, A. & Węgrzyn, G. 2006b. Effects of the presence of *cole1* plasmid dna in *escherichia coli* on the host cell metabolism. *Microbial cell factories*, 5.
- Wang, Z., Yuan, Z., Xiang, L., Shao, J. & Węgrzyn, G. 2006c. trna-dependent cleavage of the *cole1* plasmid-encoded rna i. *Microbiology (Reading, England)*, 152, 3467–3476.
- Weckwerth, W. 2003. Metabolomics in systems biology. *Annual review of plant biology*, 54, 669–689.
- Weckwerth, W. 2007. *Metabolomics-Methods and Protocols*. Humana Press Inc. Kapitel: Determination of Metabolic Flux Ratios from ^{13}C -Experiments and Gas Chromatography - Mass Spectrometry Data.
- Węgrzyn, G. 1999. Replication of plasmids during bacterial response to amino acid starvation. *Plasmid*, 41, 1–16.
- Wiechert, W. 2001. ^{13}C metabolic flux analysis. *Metabolic Engineering*, 3, 195–206.
- Williams, J. A., Carnes, A. E. & Hodgson, C. P. 2010. Plasmid DNA Vaccine vector design: impact on efficacy, safety and upstream production. *Biotechnol Adv.* 27 (4), 353–370.
- Winter, G. & Krömer, J. O. 2013. Fluxomics - connecting 'omics' analysis and phenotypes. *Environmental Microbiology*, 15, 1901–1916.
- Wittmann, C. 2007. Fluxome analysis using gc-ms. *Microbial cell factories*, 6.
- Wróbel, B. & Węgrzyn, G. 1998. Replication regulation of *cole1*-like plasmids in amino acid-starved *escherichia coli*. *Plasmid*, 39, 48–62.
- Yanisch-Perron, C., Vieira, J. & Messing, J. 1985. Improved m13 phage cloning vectors and host strains; nucleotide sequences of the m13mp18 and puc19 vectors. *Gene*, 33, 103–119.

- Yau, S. Y., Keshavarz-Moore, E. & Ward, J. 2008. Host strain influences on supercoiled plasmid dna production in escherichia coli: implications for efficient design of large-scale processes. *Biotechnology and bioengineering*, 101, 529–544.
- Yavachev, L. & Ivanov, I. 1988. What does the homology between e . coil trnas and rnas controlling cole1 plasmid replication mean ? *J. theor. Biol.* 131, 235–241.

DANKSAGUNG

Während dieser Arbeit habe ich viele Menschen und Institutionen kennengelernt, die mich in irgendeiner Form unterstützt haben und denen ich jetzt danken möchte:

VIELEN DANK...

... meinem Doktorvater, **Prof. Dr. Karsten Niehaus**, für das mir entgegengebrachte Vertrauen, die vielen spannenden Diskussionen und die sehr gute wissenschaftliche Betreuung.

... meinem Zweitgutachter, **Herrn Prof. Dr. Karl Friehs**, für die Bereitschaft zur Begutachtung dieser Arbeit.

... **Prof. Dr. Alfred Pühler** für die Überlassung des interessanten Themas.

... **Dr. Petra Lutter**, die jederzeit bereit war, mit mir über auftretende Probleme zu diskutieren, mich stets motiviert und im Bereich der mathematischen Modellierung exzellent unterstützt hat.

... meiner Arbeitsgruppe, den „**Phytos**“, für ihre große Hilfsbereitschaft und die schöne Zeit.

... **Sebastian Wiebe** für die Unterstützung bei der Arbeit im Labor.

... dem **Team von der Plasmid Factory** für die sehr gute Zusammenarbeit und die kompetente Hilfe bei Fragen rund um Plasmide.

... **Christina, Julia** und **Susanne** für die stets fröhliche und unterhaltsame Gesellschaft beim Mittagessen.

... **Jenni**, für Deine Freundschaft, Deine moralische Unterstützung und die schöne Zeit im Büro mit vielen wissenschaftlichen und nicht-wissenschaftlichen Gesprächen.

... **Matthias**, dass Du immer an mich geglaubt und mich nach allen Kräften unterstützt hast. Danke, dass Du da bist!

... **meinen Eltern**, für Eure Unterstützung in allen Lebenslagen und Euren stetigen Zuspruch. Ohne Eure Hilfe wäre mein Studium in dieser Form nicht möglich gewesen.

KOLOPHON

Dieses Dokument wurde erstellt unter Verwendung der \LaTeX Vorlage `classicthesis` entwickelt von André Miede. Die Vorlage ist frei erhältlich unter:

<http://code.google.com/p/classicthesis/>

ERKLÄRUNG

Mit dieser Unterschrift bestätige ich, dass ich die vorliegende Dissertation selbstständig und ohne Nutzung anderer als angegebener Quellen und Hilfsmittel angefertigt habe. Die vorliegende Dissertation wurde weder vollständig noch teilweise einer anderen Fakultät oder Universität mit dem Ziel vorgelegt, einen akademischen Titel zu erwerben. Ich bewerbe mich erstmals um den Doktorgrad der Naturwissenschaften.

Bielefeld, Juni 2014

Inga Freudenau