

Modellbasierte Bestimmung von Interventionsstrategien zur Optimierung der Produktion von Biokraftstoffen in Cyanobakterien

Dissertation

zur Erlangung des akademischen Grades

**Doktoringenieur
(Dr.-Ing.)**

von Diplom-Ingenieur Philipp Erdrich
geb. am 26.08.1986 in Dresden

genehmigt durch die Fakultät für Verfahrens- und Systemtechnik
der Otto-von-Guericke-Universität Magdeburg

Promotionskommission: Prof. Dr.-Ing. Udo Reichl
Dr.-Ing. Steffen Klamt
Prof. Dr.-Ing. Achim Kienle
Jun.-Prof. Dr. Oliver Ebenhöf

eingereicht am: 10. Mai 2016

Promotionskolloquium am: 03. November 2016

Kurzzusammenfassung

Die Abhängigkeit von begrenzt verfügbaren Energieträgern zu reduzieren, ist eine der wichtigsten Herausforderungen des 21. Jahrhunderts. Die biobasierte und nachhaltige Produktion von Kraftstoffen rückt dabei immer mehr in den Fokus. Mit Hilfe von Cyanobakterien, die oxygene Photosynthese betreiben, können sogenannte Biokraftstoffe auf landwirtschaftlich unnutzbaren Flächen produziert werden. Sie können damit einen substantiellen Beitrag zur nachhaltigen Biokraftstoffproduktion leisten. Allerdings ist diese Produktion häufig durch eine geringe Produktivität charakterisiert. In der vorliegenden Arbeit sollten deshalb Interventionsstrategien zur Optimierung der Biokraftstoffproduktion in Cyanobakterien auf verschiedenen Netzwerkskalen modellgestützt identifiziert werden.

Dafür ist zunächst ein neuer Reduktionsalgorithmus, *NetworkReducer*, entwickelt worden, der es erlaubt, genomskalige stöchiometrische Stoffwechselmodelle auf Modelle kleinerer Netzwerkskalen zu reduzieren. Dabei können gewünschte Eigenschaften des Ursprungsmodells konserviert werden. *NetworkReducer* ist genutzt worden, um ein genomskaliges Stoffwechselmodell einer cyanobakteriellen Spezies auf zwei kleinere Modelle mit unterschiedlicher Komplexität zu reduzieren. Die Optimierung der Biokraftstoffproduktion unter phototrophen Wachstumsbedingungen konnte damit auf drei verschiedenen Skalen am Beispiel von Ethanol in einem Cyanobakterium untersucht werden. Die darin berechneten Interventionsstrategien zeigen, dass eine Kopplung von Ethanol- und Biomassesynthese unter phototrophen Wachstumsbedingungen erzwungen werden kann. Dabei konnte ein allgemeines Prinzip für das Metabolic Engineering von Cyanobakterien identifiziert werden, dass auch für die Produktion anderer Biokraftstoffe sehr vielversprechend erscheint.

Abstract

Reducing the dependence on limited fossil resources constitutes a key challenge of the 21st century. The bio-based and sustainable production of fuels is therefore taking center stage. Cyanobacteria are increasingly recognized as promising cell factories for the production of renewable biofuels. They perform oxygenic photosynthesis and can grow on non-arable land. Hence, they could contribute substantially to the sustainable production of biofuels. However, most biotechnological applications of these organisms are still characterized by low productivity. In this work, based on various network scales, intervention strategies for the optimization of biofuel production with cyanobacteria were to be identified.

In a first step a new reduction algorithm called *NetworkReducer* was developed, that allows reduction of genome-scale metabolic models to meaningful core models of different scales. Specific functions and functionalities of the genome-scale model can be conserved by this algorithm. *NetworkReducer* was used to reduce a genome-scale model of a cyanobacterial species to obtain two meaningful core models of different complexity for analyzing the metabolism on different network scales. Using these models optimization of cyanobacterial ethanol production under phototrophic growth conditions was analyzed across three scales. The identified strategies showed that obligate coupling of biomass synthesis with ethanol synthesis is possible under phototrophic growth conditions. Moreover, a general principle of metabolic engineering was identified which shows high potential for the production of various biofuels by different cyanobacterial organisms.

Danksagung

Die vorliegende Dissertation ist während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Max-Planck-Institut für Dynamik komplexer technischer Systeme in Magdeburg im Zeitraum von März 2012 - Mai 2016 entstanden.

Ich möchte an dieser Stelle Dr.-Ing. Steffen Klamt meinen Dank aussprechen, der mich für diese Arbeit begeistert hat. Insbesondere danke ich ihm für seine Unterstützung in verschiedenen wissenschaftlichen Bereichen während meiner gesamten Zeit am Max-Planck-Institut. Diese Unterstützung und die Möglichkeit jederzeit wissenschaftliche Inhalte zu diskutieren, waren maßgebend für den Erfolg dieser Dissertation.

Ebenso danke ich Prof. Achim Kienle (Otto-von-Guericke-Universität Magdeburg) und Jun.-Prof. Oliver Ebenhöf (Heinrich-Heine-Universität Düsseldorf) für die Begutachtung der Arbeit.

Weiterhin möchte ich mich bei Henning Knoop und Ralf Steuer für den wissenschaftlichen Austausch bedanken. Außerdem möchte ich mich bei meinen Kollegen bedanken, die sowohl wissenschaftlich als auch privat für viel Gesprächsstoff gesorgt haben. Insbesondere sind dabei zu nennen: Anke Ryll (*âtT*), Bernhard Peppi Kramer, Katha(rina) Holstein, André Franz, Robert LazyL Flassig, Oli(ver) Hädicke und Michael Mangold. Darüber hinaus geht ein besonderer Dank an meine Familie. Einerseits an meine Eltern, die mich immer unterstützt haben; andererseits an meine Frau Anne und meine Töchter Lena und Frida, die mir jeden Tag wieder zeigen, was das Leben lebenswert macht.

Inhaltsverzeichnis

1	Einleitung	1
2	Biologische Grundlagen	7
2.1	Cyanobakterien	7
2.2	Stoffwechsel von Cyanobakterien	10
2.3	Produktionswege von Biokraftstoffen in Cyanobakterien	17
2.4	Betrachtete Spezies	20
3	Mathematische Grundlagen	23
3.1	Stoffwechselnetzwerke	23
3.2	Stöchiometrische Modellierung von Stoffwechselnetzen	25
3.2.1	Stationäre Flussverteilungen	27
3.2.2	Flussbilanzanalyse	30
3.2.3	Flussvariabilitätsanalyse	32
3.2.4	Elementarmodenanalyse	33
3.3	Methoden zum modellgestützten Metabolic Engineering	37
3.3.1	Übersicht über Methoden und das Prinzip der Kopplung von Pro- duktsynthese und Wachstum	37
3.3.2	Minimal Cut Sets & constrained Minimal Cut Sets	43
3.3.3	CASOP	50
3.4	Verwendete Software	51

Inhaltsverzeichnis

3.5	Stöchiometrische Stoffwechselmodelle von Cyanobakterien	52
4	Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen	57
4.1	Motivation	57
4.2	Algorithmus <i>NetworkReducer</i>	59
4.3	Machbarkeitsstudie am Beispiel eines genomskaligen <i>E. coli</i> Modells	65
4.4	Ableitung von Modellen des Zentralstoffwechsels in Cyanobakterien	75
5	Metabolic Engineering von Cyanobakterien zur Biokraftstoffproduktion	79
5.1	Motivation	79
5.2	Die Balance der ATP- und NADPH-Produktion in der Photosynthese	80
5.3	Ethanol-Produktion in <i>Synechocystis</i> sp. PCC 6803	84
5.3.1	Analyse des Minimalmodells von <i>Synechocystis</i> (<i>SyneMin</i>)	85
5.3.2	Analyse des <i>SynePruned</i> -Modells	87
5.3.2.1	CASOP-Analyse	87
5.3.2.2	Elementarmodenanalyse	92
5.3.2.3	Analyse der constrained Minimal Cut Sets	93
5.3.2.4	ATP-Wasting als alternative/zusätzliche Interventionsstrategie	100
5.3.2.5	Einfluss der Stickstoffquelle (Ammonium vs. Nitrat)	104
5.3.3	Analyse des genomskaligen <i>SyneGS</i> -Modells	105
5.4	Produktion von weiteren Biokraftstoffen mit phototrophen Cyanobakterien	107
6	Zusammenfassung	117
A	Anhang	121
A.1	Synthesereaktionen der untersuchten Biokraftstoffe	121
A.2	Vollständige Tabelle der CASOP-Werte	121
A.3	Ergebnisse der FVA: aerobes Wachstum: $\mu = 0,85 \text{ h}^{-1}$	124

A.4 Ergebnisse der FVA: anaerobes Wachstum: $\mu = 0,2 \text{ h}^{-1}$ 127

A.5 Minimalmodell von *Synechocystis* sp. PCC 6803 (*SyneMin*) 129

Inhaltsverzeichnis

Abbildungsverzeichnis

1.1	Die Dissertation in einer Graphik.	6
2.1	Schematischer Aufbau einer Cyanobakterienzelle	10
2.2	Idealisierte Darstellung der linearen photosynthetischen ETK	12
2.3	Produktionswege der untersuchten Biokraftstoffe	20
2.4	Aufnahme von <i>Synechocystis</i> sp. PCC 6803	21
2.5	Aufnahme von <i>Synechococcus elongatus</i> PCC 7942	21
2.6	Aufnahme von <i>Cyanothece</i> sp. ATCC 51142	22
3.1	Stöchiometrische Modellierung von Stoffwechselnetzen	28
3.2	Das Konzept des „constraint-based modeling“	30
3.3	Die Elementarmoden des Beispielnetzwerkes	36
3.4	Darstellung aller Minimal Cut Sets im Beispielnetzwerk	45
3.5	Constrained Minimal Cut Sets des Beispielnetzwerkes	47
3.6	Zusammenhang zwischen den Elementarmoden und Minimal Cut Sets eines biochemischen und eines dualen Netzwerkes	48
3.7	Resultierende Ausbeutediagramme nach Intervention	49
4.1	Ablauf des Reduktionsalgorithmus <i>NetworkReducer</i>	64
5.1	Linearer Elektronenfluss durch die photosynthetische Elektronentrans- portkette	81

Abbildungsverzeichnis

5.2	Elektronenflüsse in der photosynthetischen Elektronentransportkette . . .	83
5.3	Produktionsweg für Ethanol. Graphik angelehnt an [3].	84
5.4	Die EM von <i>SyneMin</i>	87
5.5	Ausbeutediagramm der EM für Ethanol- und Biomasse.	92
5.6	Eingeteiltes Ausbeutediagramm der EM für Ethanol- und Biomasse . . .	94
5.7	Die berechneten cMCS1 - cMCS62 für die wachstumsgekoppelte Ethanol- produktion	95
5.8	Die berechneten cMCS-63 - cMCS-124 für die wachstumsgekoppelte Etha- nolproduktion	96
5.9	Histogramm der cMCSs	97
5.10	Ausbeutediagramm der EM für Ethanol- und Biomasse nach Anwendung des cMCS-1	98
5.11	Ausbeutediagramm der EM für Ethanol- und Biomasse nach Anwendung der essentiellen sechs KOs	99
5.12	Zusammenfassung von Interventionsstrategien	100
5.13	Steigerung des ATP-Konsums durch schrittweise Erhöhung des R-ATPM- Flusses	102
5.14	Einfluss des ATPM-Flusses und der Ethanolausbeute auf die cMCSs . . .	103
5.15	Ausbeutediagramme der EM für Ethanol und Biomasse: Ammonium als Stickstoffquelle	106

Tabellenverzeichnis

2.1	Morphologie der Cyanobakterien	8
3.1	Methoden des modellgestützten Metabolic Engineering	43
3.2	Publizierte <i>Synechocystis</i> Modelle	52
3.3	Strategien zur Stammkonstruktion und -optimierung in Cyanobakterien	55
4.1	Eigenschaften des originalen genomskaligen <i>E. coli</i> Modells	67
4.2	Vergleich der stöchiometrischen Koeffizienten des <i>ColiPrunedComp</i> und <i>ColiCore</i> in der BSR	71
4.3	Übereinstimmung der Flussvariabilität der betrachteten Reaktionsraten von <i>ColiCore</i> und <i>ColiPrunedComp</i> mit denen des <i>ColiGS</i> in [mmol/gTM/h].	73
4.4	Vergleich der topologischen Eigenschaften der <i>Synechocystis</i> -Modelle	76
5.1	EM des stark komprimierten Zentralstoffwechselmodells von <i>Synechocystis</i>	86
5.2	Die durch CASOP identifizierte Top-Knockout- und Top-Überexpressionskandidaten für die Ethanolproduktion	88
5.3	Liste ausgewählter Reaktionen des <i>SyneGS</i> - und <i>SynePruned</i> -Modells	91
5.4	Der Bedarf an Photonen, ATP und NADPH zur Produktion jeweils eines Moleküls Biokraftstoff im <i>SyneGS</i>	108

Tabellenverzeichnis

5.5	Die minimale Größe der identifizierten cMCSs für die gekoppelte Produktion von Biokraftstoff und Biomasse in <i>SyneGS</i>	109
5.6	Der Bedarf an Photonen, ATP und NADPH zur Produktion jeweils eines Moleküls Biokraftstoff im genomskaligen Netzwerk von <i>Synechococcus</i> . . .	111
5.7	Die minimale Größe der identifizierten cMCSs für die gekoppelte Produktion von Biokraftstoff und Biomasse im genomskaligen Modell von <i>Synechococcus</i>	112
5.8	Der Bedarf an Photonen, ATP und NADPH zur Produktion jeweils eines Moleküls Biokraftstoff im genomskaligen Netzwerk von <i>Cyanothece</i>	114
5.9	Die minimale Größe der identifizierten cMCSs für die gekoppelte Produktion von Biokraftstoff und Biomasse im genomskaligen Modell von <i>Cyanothece</i>	115

Abkürzungsverzeichnis

ADH	Alkoholdehydrogenase
AEF	Alternativer Elektronenfluss
BSR	Biomassesyntesereaktion
BT	Biotechnologie
CASOP	Computational Approach for Strain Optimization aiming at high Productivity
CBM	constraint-based modeling
cMCS(s)	constrained Minimal Cut Set(s)
Cox	Cytochrom-c-Oxidase
Cyd	Cytochrom-bd-Oxidase
Cyt b_6f	Cytochrom- b_6f -Komplex
dFBA	Dynamische FBA
EFE	Ethylen-formendes-Enzym
EM	Elementarmodus
EMA	Elementarmodenanalyse
EMiLiO	Enhancing Metabolism with iterative Linear Optimization
ETK	Elektronentransportkette
FBA	Flussbilanzanalyse
Fd	Ferredoxin
FG	Freiheitsgrad
FNR	Ferredoxin:NADPH:Oxidoreduktase

Abkürzungsverzeichnis

FQR	Ferredoxin:(Plasto)quinon:Reduktase
FSEOF	Flux Scanning based on Enforced Objective Flux
FVA	Flussvariabilitätsanalyse
G3P	Glycerat-3-Phosphat
GAP	Glycerinaldehyd-3-Phosphat
GDLS	Genetic Design through Local Search
gTM	Gramm Trocken-Biomasse
irrev	Irreversibel
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	Knockout
LEF	Linearer Elektronenfluss
MCS(s)	Minimal Cut Set(s)
ME	Metabolic Engineering
Mehler	Mehler Reaktion
MMF	Minimale Metabolische Funktionalität
NDH	NAD(P)H Dehydrogenase
ox	Oxidiert
P ₆₈₀	Reaktionszentrum des PSII
P ₇₀₀	Reaktionszentrum des PSI
PC	Plastocyanin
PDC	Pyruvatdecarboxylase
PEF	Pseudozyklischer Elektronenfluss
PHB	Polyhydroxybuttersäure
PQ	Plastoquinon
PQH ₂	Plastoquinol
PSI	Photosystem I
PSII	Photosystem II
Q	Gebundenes Quinon
R5P	Ribose-5-Phosphat

Abkürzungsverzeichnis

red	Reduziert
rev	Reversibel
rFBA	regulatorische FBA
ROS	Reaktive Sauerstoff Spezies
Ru5P	Ribulose-5-Phosphat
RubP2	Ribulose-1,5-Bisphosphat
RuBisCO	Ribulose-1,5-Bisphosphat-Carboxylase/Oxygenase
SD	Strain Design
TCA	Tricarboxylic Acid Cycle; Citratzyklus
tll	Thylakoid-Lumen
TOOF	Tilting of the Objective Function
ZEF	Zyklischer Elektronenfluss

Abkürzungsverzeichnis

Symbolverzeichnis

α_{Produkt}	Molare Masse des Produkts (CASOP)
α_i	Untere Schranke der Reaktion r_i
β_i	Obere Schranke der Reaktion r_i
γ	Relativer Masseanteil des Produkts an der Biomasse
ω	Maß für die Relevanz einer Reaktion (CASOP)
μ	Wachstumsrate
A	Artifizieller Metabolit (CASOP)
\mathbf{b}	Vektor der bekannten Raten
$\mathbf{c}(t)$	Metabolitkonzentrationsvektor
\mathbf{c}	Vektor der Koeffizienten der Zielfunktion
C	Cut Set
$\mathbf{d}_k, \mathbf{D}_k$	Vektor und dazugehörige Matrix zur Beschreibung der zu erhaltenden Funktionalität k
\mathbf{D}	Menge der Desired-Elementarmoden
\mathbf{e}	Flussvektor (Elementarmodus)
\mathbf{E}	Menge der Elementarmoden
F	Flusskegel
FG	Freiheitsgrad
k_C	Gewichtungsparameter (CASOP)
k	Laufvariable der zu erhaltenden Funktionalitäten

Symbolverzeichnis

n_{KO}	Anzahl der Knockouts (OptKnock)
K	Kernmatrix
m	Anzahl der Metabolite
$m_{i,C}(\gamma)$	Menge der EM des Szenarios γ , an welchen die Reaktion r_i beteiligt ist (CASOP)
M	Menge der Reaktionen (OptKnock)
n_{cMCS}	Minimale Anzahl der zu erhaltenden Desired-EM
$n_C(\gamma)$	Gesamtanzahl der EM für Szenario γ (CASOP)
n_{ji}	stöchiometrischer Koeffizient des Metaboliten j in Reaktion i
N	Stöchiometrische Matrix
p	Parametervektor
P^M	Menge der geschützten Metabolite
P^R	Menge der geschützten Reaktionen
q	Anzahl der Reaktionen
r	Reaktionsratenvektor
s	Anzahl der zu erhaltenden Funktionalitäten
S	Substrat
T	Menge der Target-Elementarmoden
y	binäre Variable (OptKnock)
Y	Ausbeute
z	Zielfunktion
Z	Bewertungsmaß (CASOP)

1 Einleitung

Die Biotechnologie (BT) ist eine der wichtigsten Technologien der heutigen Zeit und wird auch in der Zukunft eine Schlüsseltechnologie sein [181]. Prinzipiell nutzt die BT die enormen Möglichkeiten von Mikroorganismen und deren Enzymen z. B. zur Abwasseraufbereitung, zur Produktion von Chemikalien, Nahrung und Kosmetika oder zur Saatzucht. Ein besonderes Anwendungsfeld stellt die biobasierte und nachhaltige Produktion verschiedener Güter im Bereich von Chemikalien, Textilien und Energie dar. Im Vergleich zur konventionellen Produktion bieten biotechnologische Verfahren verschiedene Vorteile wie z. B. die Senkung der CO₂-Emissionen, der Produktionskosten, der Abfälle und des Ressourcenverbrauchs [105]. Die Biotechnologie wird klassisch in drei Teilbereiche eingeteilt: die rote, grüne und weiße Biotechnologie [68]. Die rote BT beschäftigt sich mit der Entwicklung neuer therapeutischer und diagnostischer Verfahren in der Medizin. Die grüne Biotechnologie befasst sich mit (tw. genetisch modifizierten) Pflanzen und der Landwirtschaft während die weiße Biotechnologie (auch industrielle BT genannt) biologische und biochemische Kenntnisse in technische Anwendungen übersetzt. Allerdings existiert keine strenge Abgrenzung der Teilbereiche, da diese teilweise fließend ineinander übergehen (manche Veröffentlichungen sprechen sogar von zehn verschiedenen Teilbereichen [54]).

Biotechnologische Anwendungen gibt es schon seit Jahrtausenden. Beispiele dafür sind die Herstellung von Brot, Bier und Wein. Die ersten biotechnologischen Prozesse, die als solche verstanden worden, sind erst in den letzten drei Jahrhunderten zu finden. Ein klassisches Beispiel dafür ist die Entdeckung der Gärung durch Louis Pasteur. Er schuf damit das Fundament zum Verständnis der Fermentation und begründete die moderne Mikrobiologie, die wiederum die Basis für die Entwicklung der Biotechnologie war. Heutzutage hat die Biotechnologie einen festen Platz in der Entwicklung und Produktion von z. B. Enzymen, Antibiotika, Vitaminen, Lebensmitteln, Biokunststoffen, Biokraftstoffen, etc.

1 Einleitung

Es gibt mittlerweile eine große Anzahl von biotechnologischen Unternehmen, die mit der „klassischen“ Industrie bzgl. der genannten Produkte im Wettbewerb stehen. Das bedeutet, dass diese Unternehmen ihre biotechnologischen Prozesse optimieren müssen (u. a. Senkung der Produktionskosten). Damit ergeben sich verschiedenste Herausforderungen für die genutzten Produktionsstämme, wie z. B. Wachstum auf günstigen Substraten, oder hohe Produktausscheideraten. Des Weiteren sind die Populationen von Mikroorganismen i. d. R. nicht homogen, sondern besitzen eine genetische und physiologische Variabilität. Diese muss bei der Kultivierung berücksichtigt und beherrscht werden [172]. Weitere wichtige Aufgaben sind z. B. die Kopplung von biotechnologischen und chemischen Produktionslinien [36], oder auch die Entwicklung stabiler Kultivierungsprozesse, die gerade beim Scale-up von großer Bedeutung sind. Aber auch Prozesse wie die Modifikation von nativen Bakterienstämmen hin zu Produktionsstämmen (mit hoher Produktivität und Spezifität) sind von essentieller Bedeutung.

Die „klassische“ Vorgehensweise für diese Modifikation ist eine zufällige Veränderung der Stämme, die häufig auf einer Mutagenese beruht, die durch Strahlung oder Zugabe von Chemikalien erfolgt. Neuerdings wird diese Methode abgelöst durch rationale (tw. evolutive) Verfahren, die es ermöglichen die Bakterienstämme gezielt zu modifizieren. Beispiele dafür sind das *evolutionary engineering*, bei dem ganze Zellen in Richtung eines bestimmten Phänotyps entwickelt werden; und auch das *directed engineering* das Gene, Promotoren und/oder Transkriptionsfaktoren beeinflusst, um diese gezielt zu optimieren (z. B. Veränderung des Expressionslevels bestimmter Gene) [264, 27, 14].

Die beschriebenen Methoden können dem Begriff des „Metabolic Engineering“ (ME) zugeordnet werden. ME beschäftigt sich klassischer Weise mit der Analyse und Modifikation metabolischer Pfade [238] und wurde 1991 das erste mal als eigenständige Disziplin beschrieben [24]. Im 1998 veröffentlichten Grundlagenwerk *Metabolic Engineering* von Stephanopoulos [238] wird diese Disziplin definiert als: „*direct improvement for product formation or cellular properties through the modification of specific biochemical reaction(s) or the introduction of new one(s) with the use of recombinant DNA technology*“. Eine der großen Herausforderungen dieser Disziplin ist es, Stoffwechselmodelle zu konstruieren, die den Metabolismus adäquat beschreiben können. Basierend darauf können Methoden entwickelt werden, um z. B. die Auswirkungen genetischer Modifikationen in Produktionsorganismen zur Ertragsteigerung oder Synthese neuer Produkte zu analysieren (und später experimentell umzusetzen) [148].

An erster Stelle steht das gewünschte Produkt, das synthetisiert werden soll. Dafür muss ein geeigneter Stamm identifiziert werden, der verschiedene Kriterien erfüllt (leichte genetische Modifizierbarkeit, gute Kultivierungsmöglichkeiten, etc.). Sobald ein geeigne-

ter Stamm identifiziert worden ist, wird für diesen ein Stoffwechselmodell entwickelt (oder ein vorhandenes genutzt) und analysiert. Danach wird beispielsweise versucht Interventionsstrategien und/oder ein rationales Design abzuleiten, wie der Stoffwechsel modifiziert werden kann, um den Organismus zu optimieren. Anschließend wird versucht, dies im Labormaßstab (z. B. in kleinen Bioreaktoren) umzusetzen. Zeigt sich, dass das Modell die aus dem Laborversuch gewonnenen Daten nicht wiedergeben kann, muss das mathematische Modell entsprechend angepasst werden. Damit entsteht ein Kreislauf bis das Modell in der Lage ist, den Stoffwechsel des betrachteten Stamms adäquat beschreiben zu können. Nun können neue Strategien berechnet werden, um die Synthese des gewünschten Produkts zu optimieren. Gelingt die Umsetzung im Labormaßstab, kann versucht werden, die Produktion im größerem Maßstab umzusetzen (scale-up). Es gibt viele Beispiele für die erfolgreiche Anwendung mittels Modellen identifizierter Interventionsstrategien für verschiedene Organismen. Zum Beispiel: *E. coli*: Produktion von 1,4-Butandiol [268], 1-Butanol [20] und Succinat [152]; *Saccharomyces cerevisiae*: Produktion von Ethanol [176] und n-Butanol [237]; Cyanobakterien: Produktion von Ethanol [56] und 1-Butanol [141].

Oft ist die Zielstellung, bekannte Produkte auf alternativen Wegen zu synthetisieren, um z. B. kostengünstiger, nachhaltiger oder auch unabhängiger von bestimmten Rahmenbedingungen zu sein. Zum Beispiel besteht eine enorme Abhängigkeit von fossilen Energieträgern. Allein in Deutschland besteht (bis auf Braunkohle) eine Nettoimportabhängigkeit von mindestens 85% für Steinkohle oder Mineralöle (gemessen am Primärenergieverbrauch [37]). Um diese Abhängigkeit von fossilen Brennstoffen zu reduzieren, muss der Beitrag der nachhaltigen industriellen Produktion von Kraftstoffen und Chemikalien aus biologischen Ressourcen steigen. Die EU fordert in ihrer Erneuerbare-Energien-Richtlinie, dass der Anteil der erneuerbaren Energien im Bereich Verkehr bis 2020 in jedem Mitgliedsstaat auf 10% gesteigert werden soll (gemessen am Gesamtanteil von Energie aus erneuerbaren Quellen [53]). Die Kraftstoffe aus biologischen Ressourcen (Biokraftstoffe) bilden damit ein zentrales Element der Energieproduktion der Zukunft.

Die Biokraftstoffe werden in zwei (teilweise in drei) Generationen eingeteilt [41, 173], allerdings gibt es keine strikte Abgrenzung. In der ersten Generation der Biokraftstoffe wird Öl, Zucker oder Stärke als Grundlage verwendet, um z. B. Biodiesel zu produzieren (der verbleibende Rest der Pflanze wird z. B. als Futtermittel verwendet [117, 166]). In der zweiten Generation der Biokraftstoffe wird zusätzlich die Cellulose verarbeitet. Mit anderen Worten, es wird versucht, die gesamte Biomasse in Energie umzuwandeln [270]. Allerdings ist der große Nachteil der so gewonnenen Biokraftstoffe, dass diese stark mit der Nahrungsmittelindustrie konkurrieren, da landwirtschaftlich nutzbare Flächen

1 Einleitung

benötigt werden. Das hat wiederum negative Auswirkungen auf das Preisgefüge von Lebensmitteln (Erhöhung der Kosten [103]). Auch aus diesem Grund wird versucht, Biokraftstoffe der nächsten Generation zu entwickeln. Dabei wird der Kraftstoff mit Hilfe aquatischer Mikroorganismen wie Algen oder Cyanobakterien produziert, die in der Lage sind, Photosynthese zu betreiben. Diese Mikroorganismen können Sonnenenergie, Kohlenstoffdioxid und Wasser in Kohlenwasserstoffe umwandeln [141].

Cyanobakterien sind seit geraumer Zeit im Fokus der Biotechnologie, da diese weitere Vorteile bieten. Das Genom vieler Cyanobakterienstämme ist sequenziert und es gibt viele Werkzeuge mit denen das Genom modifiziert werden kann. Des Weiteren bieten sie den Vorteil, dass diese nicht auf landwirtschaftlich nutzbaren Fläche eingesetzt werden müssen. Damit steht die Produktion von Biokraftstoffen mit Cyanobakterien nicht im Wettbewerb zur Nahrungsmittelproduktion. Weiterhin können Cyanobakterien in Meer- und Abwasser wachsen und verbrauchen kein Süßwasser [77, 239]. Es gibt bereits zahlreiche biotechnologische Anwendungen mit Cyanobakterien [9, 17, 101, 202, 244], deren Produktpalette von bioaktiven Verbindungen [8, 32, 52] über sekundäre Metabolite [188, 229] bis hin zu Bioplastik (Polyhydroalkanoate [98, 189, 234]) reicht [59]. Auch für die Produktion verschiedener Biokraftstoffe mit Cyanobakterien existieren erste Entwicklungen von Produktionsprozessen (für z. B. Ethanol [56], Isobutanol [21, 257], Isopren [153], Butanol [141] oder auch Ethylen [240]).

Mittlerweile wird z. B. von *Algenol* und *Joule Unlimited* getestet, ob sich Biokraftstoffe auch im größeren Maßstab produzieren lassen. Beide Unternehmen arbeiten mit Cyanobakterien und nutzen diese, um Sonnenlicht, Kohlenstoffdioxid und Wasser in Kraftstoffe wie Ethanol oder Diesel umzuwandeln. Die Bakterien werden dafür in speziellen Modulen mit Licht und CO₂ versorgt, um dieses anschließend kontinuierlich in Biokraftstoff umzuwandeln. Dabei wird der Kraftstoff in das Medium abgegeben, das abgetrennt und aufgereinigt wird. Der Gesamtprozess läuft über mehrere Wochen bis die Module gereinigt werden müssen. *Algenol* verarbeitet danach die Restbiomasse weiter (z. B. zu Jetfuel). Dieses Unternehmen geht davon aus, insgesamt rund 8000 Gallonen pro Morgen pro Jahr Biokraftstoff produzieren zu können, das entspricht rund 75.000 L/ha/Jahr [1]. Dies wäre beachtlich, da im Vergleich dazu aus Mais lediglich 420 Gallonen pro Morgen pro Jahr Biokraftstoff gewonnen werden (das entspricht rund 4000 L/ha/Jahr) [1].

Viele Anwendungen der nachhaltigen Produktion von Biokraftstoffen mit Cyanobakterien sind dennoch durch eine geringe Produktivität charakterisiert. Interventionsstrategien für die Modifikation von Cyanobakterien zu identifizieren, welche die Produktionsparameter erhöhen können, ist ein Anliegen dieser Arbeit. Es gibt verschiedene

Methoden und Strategien, deren Anwendung beispielsweise die Ausbeute verbessern kann. Bezüglich der Cyanobakterien sind die meisten davon bisher ungerichtet und nicht ausreichend evaluiert. Deshalb widmet sich die vorliegende Arbeit der Optimierung der Produktion von Biokraftstoffen in Cyanobakterien mit dem Ziel, konkrete Interventionsstrategien zu identifizieren. Ausgehend von einem genomskaligen Stoffwechselmodell sollen diese Strategien auf verschiedenen Netzwerkskalen berechnet und analysiert werden. Abbildung 1.1 fasst die Dissertation graphisch zusammen.

In **Kapitel 2** werden die biologischen Grundlagen von Cyanobakterien erläutert. Dies beinhaltet eine allgemeine Einordnung, den Stoffwechsel, die Produktionswege für verschiedene Biokraftstoffe und die untersuchten cyanobakteriellen Spezies.

Kapitel 3 beschreibt die grundlegenden mathematischen Methoden der stöchiometrischen Modellierung von Stoffwechselnetzwerken und gibt eine Übersicht über verschiedene stöchiometrische Modelle von Cyanobakterien.

Da zu Beginn nur genomskalige Modelle zur Beschreibung des cyanobakteriellen Stoffwechsels vorhanden waren, bestand eine weitere Aufgabenstellung darin, einen Algorithmus zu entwickeln, der es ermöglicht, genomskalige stöchiometrische Stoffwechselmodelle eigenschaftserhaltend auf Modelle kleinerer Netzwerkskalen zu reduzieren. Dieser neue Reduktionsalgorithmus wird in **Kapitel 4** vorgestellt. Es wird gezeigt, wie dieser angewendet werden kann und welche Fragestellungen bzw. Probleme mit dieser Methodik beantwortet werden können.

In **Kapitel 5** werden skalenübergreifend Interventionsstrategien in phototrophen Cyanobakterien zur Optimierung der Produktion von Biokraftstoffen beschrieben. Es werden an dieser Stelle Strategien gezeigt, die es ermöglichen, phototrophes Wachstum und Biokraftstoffproduktion aneinander zu koppeln.

Die wesentlichen Resultate der Dissertation sind bereits veröffentlicht worden. Der Reduktionsalgorithmus, *NetworkReducer*, ist in (Erdrich et al. 2015 [67]) erschienen und identifizierte Interventionsstrategien zur Optimierung der Biokraftstoffproduktion sind in (Erdrich et al. 2014, [66]) präsentiert worden.

1 Einleitung

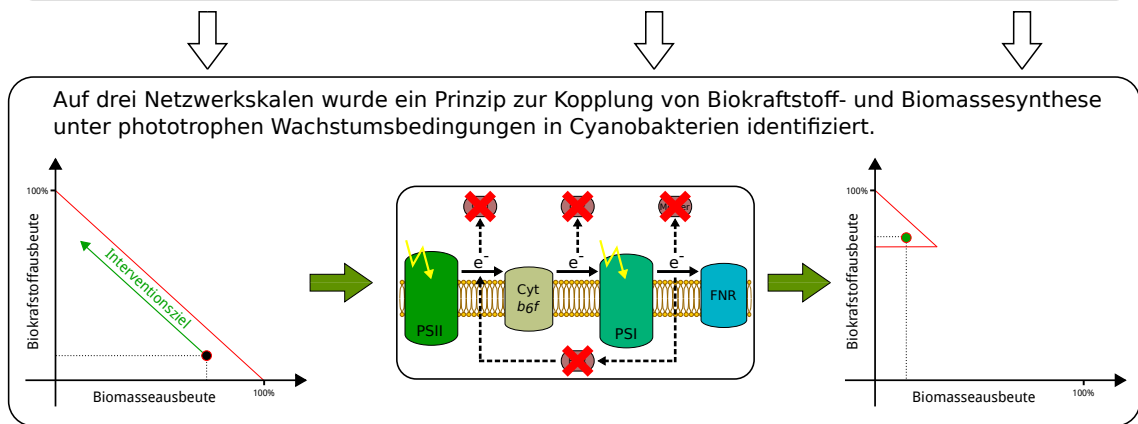
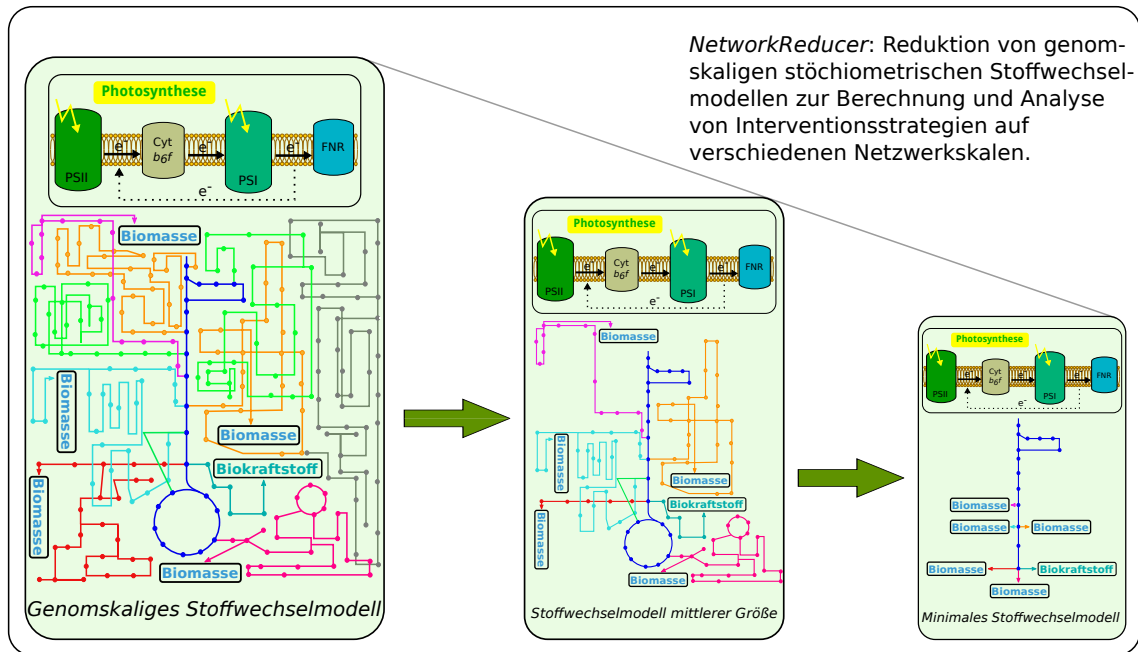
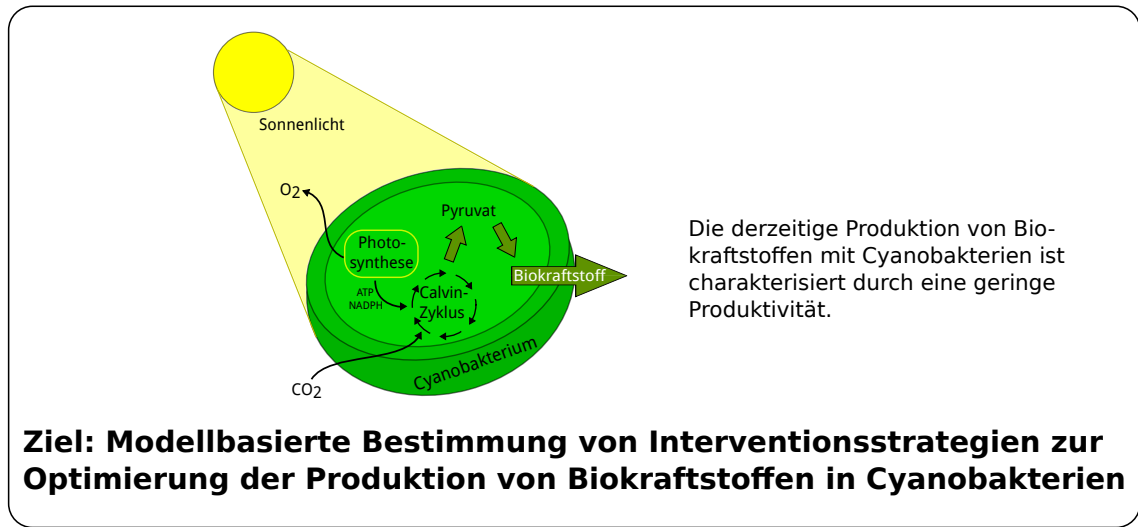


Abbildung 1.1: Die Dissertation in einer Graphik.

2 Biologische Grundlagen

Im folgenden Kapitel werden die biologischen Grundlagen von Cyanobakterien erläutert. Dies beinhaltet eine allgemeine Einführung (2.1), sowie eine Beschreibung des Stoffwechsels (2.2) und der Produktionswege verschiedener Biokraftstoffe (2.3). Anschließend werden in Kapitel 2.4 ausgewählte Vertreter der Cyanobakterien diskutiert, auf die im Verlauf der Arbeit detailliert eingegangen wird.

2.1 Cyanobakterien

Cyanobakterien, früher etwas irreführend als Blaualgen bezeichnet, stellen die größte Gruppe phototropher Bakterien dar. Sie sind in der Lage, oxygene Photosynthese zu betreiben, bei der Sauerstoff (O_2) produziert wird [28]. Cyanobakterien sind rund 2,6 - 3,5 Milliarden Jahre alt und haben durch ihre Fähigkeit zur Produktion von elementarem Sauerstoff einen wesentlichen Anteil an der Oxygenierung der Erdatmosphäre [97]. Sie sind in der Lage CO_2 als einzige Kohlenstoffquelle zu verwenden und viele Stämme können zusätzlich Stickstoff fixieren [217]. Durch diese Fähigkeiten spielen die Cyanobakterien global eine wichtige Rolle in der Assimilierung von CO_2 und der Regeneration von Sauerstoff.

Lebensraum

Sie sind als Primärproduzenten in vielen extremen und nährstoffarmen Umgebungen vertreten. Sie können in bis zu 73 Grad Celsius heißen Quellen leben, in Polarregionen, in Süß- und in Salzwasser, in tiefen Gewässern mit wenig verfügbarem Licht und in Regionen mit sehr hoher Sonneneinstrahlung [145]. Diese extreme Vielseitigkeit der Cyanobakterien hat zur Folge, dass sie eine ökologisch wichtige Rolle in verschiedenen Bereichen haben. So z. B. in den Ozeanen, in denen sie neben den Algen der wichtigste Primärproduzent sind und das Meer mit Stickstoff versorgen. In Süßwasser sind Cyano-

Tabelle 2.1: Morphologie der Cyanobakterien. Angelehnt an [145].

Subsektion	Ordnung	Merkmale
I	<i>Chroococcale</i>	Unizellulär; die Zellen liegen einzeln oder im Verbund vor; gewöhnliche Zellteilung
II	<i>Pleurocapsale</i>	Unizellulär; die Zellen liegen einzeln oder im Verbund vor; mehrfache Zellteilung
III	<i>Oscillatoriale</i>	Zellen liegen als Kette (Trichome) vor; Zellteilung in einer Ebene (uniseriate Trichome); keine Ausbildung von Heterocysten oder Akineten
IV	<i>Nostocale</i>	Zellen liegen als Kette (Trichome) vor; Zellteilung in einer Ebene (uniseriate Trichome); Ausbildung von Heterocysten möglich
V	<i>Stigonematale</i>	Zellen liegen als Kette (Trichome) vor; Zellteilung in mehreren Ebenen (multiseriate Trichome)

bakterien als Bestandteile des Phytoplanktons zu finden (z. B. in Seen oder Teichanlagen; allerdings ist deren Vorkommen dort nicht wünschenswert, da bestimmte Arten (z. B. *Mycrocystis*) Toxine bilden). Cyanobakterien wie z. B. *Chroococcidiopsis* kommen in terrestrischer Umgebung vor, in der sie eine entscheidende Rolle bei der Gesteinserosion spielen und damit von besonderer Bedeutung für den Erhalt von Gesteinsmonumenten sind [145]. Es gibt Cyanobakterien die symbiotische Verbindungen eingehen und ihren Wirt mit verschiedenen Nährstoffen versorgen (z. B. die Bereitstellung von fixiertem Stickstoff).

Morphologie

Die Vielseitigkeit der Cyanobakterien bzgl. der Umweltbedingungen spiegelt sich in ihrer hohen morphologischen Diversität wider [196]. Diese Morphologie wird laut Schlegels „Allgemeiner Mikrobiologie“ (im Folgenden Schlegel genannt) in fünf Gruppen eingeteilt [217] und reicht von einzelligen Stäbchen oder Kokken, deren Zellen oft Kolonien bilden und von einer Schleimsubstanz („Galerte“) zusammen gehalten werden, bis hin zu Gruppen, die Zellfäden (sogenannte Trichome) bilden und deren Zellteilung innerhalb dieses Fadens durch Zerbrecen dieser erfolgt [207]. In Tabelle 2.1 sind die morphologischen Gruppen detailliert dargestellt. Die Größe einer cyanobakteriellen Zelle variiert je nach Typ zwischen $0,5 \mu\text{m}$ (*Prochlorococcus*) und $100 \mu\text{m}$ (*Chroococcus*) [145].

Zelldifferenzierung

Wie bereits angedeutet, gibt es besondere Zelldifferenzierungen, so z. B. Hormogonien, Heterocysten oder Akineten. Die Hormogonien, die von einigen filamentösen Cyanobakterien gebildet werden können, sind Zellketten von parentalen Trichomen (Bruchstücke von Zellfäden) [207]. Die Hauptaufgabe, die den Hormogonien zugeschrieben wird, besteht in der Verbreitung [163]. Die Heterocysten sind in der Lage, Stickstoff zu fixieren. Sie besitzen kein Photosystem II (PSII) und erzeugen damit keinen Sauerstoff (dieser würde die Nitrogenase inaktivieren [69]). Die Heterocysten tauschen den Stickstoff (meist in Form von Glutamin) mit den Nachbarzellen aus und erhalten dafür organischen Kohlenstoff. Die Akineten repräsentieren Dauerzellen, die dickwandig sind, Glycogen und Cyanophycin beinhalten und der Überdauerung dienen [217].

Wachstumsverhalten

Im Allgemeinen sind die Cyanobakterien durch ganz unterschiedliche Wachstumsraten gekennzeichnet. Einige haben eine Verdopplungszeit die kleiner als sechs Stunden ist, während andere sich erst in einem oder sogar mehreren Tagen verdoppeln. Das langsame Wachstum der Cyanobakterien könnte der Grund sein für die starke ökologische Rolle [145].

Prinzipiell erfolgt das Wachstum durch Teilung, allerdings gibt es einige Besonderheiten. Bei der Teilung wird zwischen unizellulären und filamentösen Arten unterschieden (siehe Tabelle 2.1). Bei den unizellulären Typen findet eine gewöhnliche und/oder mehrfache Zellteilung statt. Gewöhnliche Zellteilung meint, dass innerhalb der Mutterzelle alle notwendigen „Tochterbestandteile“ heranwachsen, bis sich die Zelle teilt. Es kann bei unizellulären Arten auch zu mehrfacher Teilung kommen. Dabei entstehen mehrere Tochterzellen (sogenannte Baeocyten). Bei der filamentösen Art läuft die Teilung innerhalb der Kette ab. Es kann unterschieden werden, in welcher Ebene die Teilung erfolgt. Wenn sich die Ebene der Teilung verschiebt, können sich Abzweigungen (Ausläufer der Kette) ausprägen; verzweigte Trichome entstehen.

Aufbau und Bestandteile einer Cyanobakterienzelle

Cyanobakterien besitzen eine mehrschichtige Zellwand aus Murein, Proteinen, Lipopolysacchariden und Polysacchariden [217]. Unter Verwendung der Transmissionselektronenmikroskopie konnten viele cytoplasmatische Bestandteile der Cyanobakterien identifiziert werden (vgl. Bergeys „Handbuch der Systematischen Bakteriologie“) [44]:

2 Biologische Grundlagen

- Thylakoide, als intrazelluläre Membran, an welcher der Photosyntheseapparat (inkl. der Phycobilisome) lokalisiert ist.
- Glycogen-Granula, als zellinterner Speicherstoff für Kohlenstoff [279] (einige Cyanobakterien nutzten auch Poly- β -Hydroxybutyrat als Speicherstoff [234]).
- Cyanophycin-Granula als Reservoir für Stickstoff [278].
- Polyphosphat-Granula, als zellinterner Speicherstoff für Phosphat [13].
- Carboxysomen, als Struktur in der das Enzym Ribulose-1,5-Bisphosphat-Carboxylase/Oxygenase (RuBisCO) vorliegt.
- Gasvakuolen, die hauptsächlich in aquatischen Cyanobakterien vorhanden sind und ihnen damit einen Auftrieb ermöglichen (um in Richtung der/des Wasseroberfläche/Lichts zu treiben) [263].
- Das „Zentroplasma“ als Ort an dem die zelluläre DNA vorzufinden ist.

Eine schematischer Aufbau einer Cyanobakterienzelle ist in Abbildung 2.1 dargestellt.

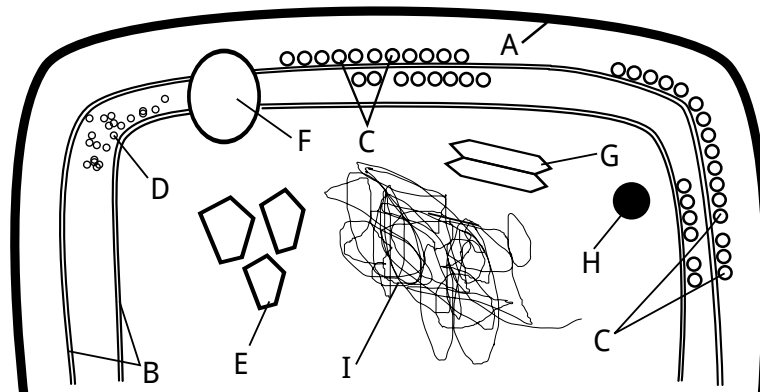


Abbildung 2.1: Schematischer Aufbau einer Cyanobakterienzelle: A) Zellwand, B) Thylakoidmembran, C) Phycobilisome, D) Glycogen-Granula, E) Carboxysomen, F) Cyanophycin-Granulum, G) Gasvakuolen, H) Polyphosphat-Granulum, I) DNA. Graphik angelehnt an [4].

2.2 Stoffwechsel von Cyanobakterien

Die Cyanobakterien sind in der Lage, Strahlungsenergie in für sie verwertbare chemische Energie umzuwandeln. Sie verfügen über ein duales Photosystem, das ihnen ermöglicht, Wasser (H_2O) als Elektronendonator zu verwenden. Sie erzeugen bei der Wasserspaltung

molekularen Sauerstoff (oxygene Photosynthese). Allerdings gibt es auch Cyanobakterien die Schwefelwasserstoff als Elektronendonator nutzen und damit anoxygene Photosynthese betreiben können [48]. Im Unterschied zu Pflanzen besitzen die Cyanobakterien nur Chlorophyll a [149] (bis auf die Gattung *Prochlorococcus*, die zusätzlich Chlorophyll b besitzt [246]).

Lichtsammel-Komplexe

Das Licht bzw. die Photonen werden über die Lichtsammel-Komplexe (Phycobilisome) absorbiert [86]. Diese Phycobilisome sind membrangebunden und bestehen aus Phycobiliproteinen. Die häufigsten Vertreter dieser Proteine in Cyanobakterien sind: Phycoerythrin, Phycocyanin und Allophycocyanin. Diese Proteine werden auch als Chromoproteine bezeichnet. Die Proteine besitzen einen Proteinanteil und eine chromophore Gruppe. Die Chromophore der Photosynthese in Cyanobakterien werden als Phycobiline bezeichnet und deren wichtigste Vertreter sind Phycocyanobilin als Bestandteil von Phycocyanin, Phycoerythrobilin als Bestandteil von Phycoerythrin und Allophycocyanobilin als Bestandteil von Allophycocyanin. Die Chromophore verfügen über anregbare Elektronen und sind damit essentiell für die Lichtabsorption. Das Maximum der Lichtabsorption ist für Allophycocyanin bei einer Wellenlänge von 650 nm, für Phycocyanin bei 620 nm und für Phycoerythrin bei 500 nm lokalisiert. Mit Hilfe dieser verschiedenen Phycobiliproteine wird das Lichtabsorptionsspektrum der Photosynthese erweitert. Der Phycobilisom-Komplex des Cyanobakteriums *Synechocystis* sp. PCC 6803 zum Beispiel besteht aus einem Allophycocyanin-Kern. Um diesen Kern herum sind sechs „Stangen“ arrangiert, die aus jeweils drei Phycocyanin-Scheiben bestehen [144]. Diese „Scheiben“ absorbieren das vorhandene Licht. Lea-Smith et al. konnten zeigen, dass für *Synechocystis* sp. PCC 6803 die Photosyntheserate abnimmt, sofern diese Phycocyanin-Scheiben-Anzahl reduziert wird [144].

Photosynthese und lineare Elektronentransportkette

Die Photosynthese der Cyanobakterien bildet das Zentrum der Energiegewinnung. Sie erzeugen durch Nutzung von Sonnenlicht Energie und Reduktionsäquivalente in Form von ATP und NADPH. Die zentralen Bestandteile des Photosyntheseapparates in Cyanobakterien sind die Lichtsammel-Komplexe, die Photosysteme I und II, der Cytochrom-*b₆f*-Komplex, die ATP Synthase, die Ferredoxin:NADPH:Oxidoreduktase (FNR) und die elektronenübertragenden Moleküle, welche die genannten Komplexe miteinander verbinden. Abbildung 2.2 zeigt eine idealisierte Darstellung der Photosynthese.

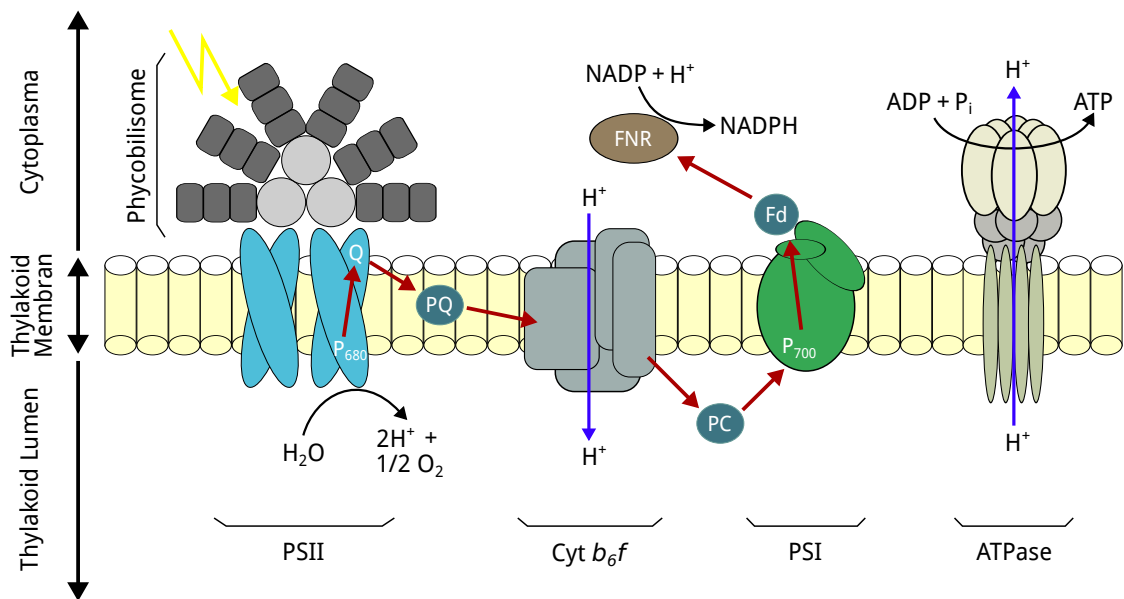


Abbildung 2.2: Idealisierte Darstellung der linearen photosynthetischen Elektronentransportkette (ETK). Die roten Pfeile stellen den Fluss der Elektronen dar. Die blauen Pfeile repräsentieren den Protonentransport über die Membran. PSII: Photosystem II, P₆₈₀: Reaktionszentrum des PSII, Q: gebundenes Quinon, PQ: Plastoquinon, Cyt *b₆f*: Cytochrom-*b₆f*-Komplex, PC: Plastocyanin, PSI: Photosystem I, P₇₀₀: Reaktionszentrum des PSI, Fd: Ferredoxin, FNR: Ferredoxin:NADPH:Oxidoreduktase. Grafik angelehnt an [145].

Die durch die Phycobiliproteine absorbierte Lichtenergie wird genutzt, um ein entsprechendes Elektron in ein höheres Orbital zu heben. Dessen Anregungsenergie wird von den äußeren Antennen zu dem innen liegenden Kern „transportiert“ (in Form eines Excitons) und von dort auf das Reaktionszentrum (P₆₈₀) des PSII übertragen (das PSII und das Phycobilisom bilden zusammen einen Superkomplex [25]). Das PSII (ein homodimerer Multiprotein-Kofaktor Komplex [156]) fungiert als Wasser:Plastoquinon:Oxidoreduktase. Es katalysiert den Transfer der aus der Wasserspaltung frei werdenden Elektronen auf ein membrangebundenes Quinon. Bei der Spaltung des Wassers entsteht molekularer Sauerstoff. Es folgt eine Reihe von Redoxreaktionen (Elektronentransportkette). Die Elektronen des Quinons werden auf ein Plastoquinon (PQ) übertragen und fließen damit in den PQ-Pool. Dieser Pool verbindet das PSII mit dem Cytochrom-*b₆f*-Komplex. Dieser Komplex fungiert einerseits als Protonenpumpe und andererseits als Komplex, an dem die Elektronen auf ein Plastocyanin (PC) übertragen werden.

Das PC verbindet den Cytochrom-*b₆f*-Komplex mit dem Photosystem I (PSI). Das PSI wird auch als Plastocyanin:Ferredoxin:Oxidoreduktase bezeichnet und bildet einen Protein-Kofaktor-Komplex [88]. Sofern das PSI (und dessen Reaktionszentrum P700) angeregt wird, katalysiert das PSI die Oxidation des reduzierten PC. Zeitgleich wird ein Ferredoxin (Fd) reduziert. Dieses reduzierte Fd wird von einer Ferredoxin:NADPH:Oxidoreduktase als Elektronendonator genutzt, um NADP^+ zu reduzieren (NADPH wird gebildet).

Durch die Anhäufung von Protonen entsteht ein elektrochemischer Gradient (protonenmotorische Kraft) über die Thylakoidmembran, der von der membrangebundenen ATP Synthase genutzt wird, um beim Einstrom von Protonen in das Stroma ATP zu generieren (ADP wird zu ATP phosphoryliert).

Zyklischer Elektronenfluss

Durch die Absorption von Lichtenergie entsteht für die Zelle nutzbare Energie in Form von ATP und Reduktionsäquivalente in Form von NADPH. Die bisher beschriebenen Prozesse der Photosynthese werden als lineare photosynthetische ETK bezeichnet. Allerdings gibt es auch einen zyklischen Elektronenfluss beispielsweise um das PSI. Es werden die Elektronen des reduzierten Fd zurück auf das PQ übertragen. Dies verhindert einerseits, dass Reduktionsäquivalente gebildet werden und ermöglicht andererseits, dass Protonen über die Thylakoidmembran transportiert werden. Daraus folgt, dass ATP unabhängig von NADPH generiert werden kann. Der zyklische Elektronenfluss gibt den Cyanobakterien die Möglichkeit, ihren Bedarf an Energie und Reduktionsäquivalenten unabhängig voneinander zu steuern. Es gibt noch weitere zyklische und pseudozyklische Elektronentransportwege, auf die in einem späteren Kapitel detailliert eingegangen wird.

CO₂-Fixierung

Die aus der Photosynthese gewonnene Energie wird genutzt, um CO₂ im reduktiven Pentose Phosphatweg (Calvin Zyklus) zu binden. Um ein Molekül CO₂ zu fixieren, sind drei ATP und zwei NADPH notwendig; für die Bildung eines Glycerinaldehyd-3-Phosphats (GAP) müssen drei CO₂ fixiert werden; dafür werden neun ATP und sechs NADPH benötigt. Das dafür verantwortliche Enzym ist die Ribulose-1,5-Bisphosphat-Carboxylase/Oxygenase. Der Name verdeutlicht, dass es sich um ein bifunktionales Enzym handelt, das in allen photosynthetischen Organismen vorhanden ist [16]. Aufgrund der Bifunktionalität kann RuBisCO sowohl CO₂ als auch O₂ binden. Bei der Carboxylierung von Ribulose-1,5-Bisphosphat (RubP2) werden zwei Moleküle Glycerat-3-Phosphat

2 Biologische Grundlagen

(G3P) gebildet. Wird statt dessen RubP2 oxygeniert, entsteht ein Molekül G3P und ein Molekül 2-Phosphoglycolat. Der Prozess der Oxygenierung von RubP2 durch RuBisCO wird als Photorespiration verstanden. Nach wie vor wird diskutiert, ob die Photorespiration ein Verbrauch von Ressourcen ist, der die photosynthetische Effizienz senkt [277], oder ob es eine protektive Funktion hat (Energiedissipation bei hoher Lichtkonzentration [91]). Eine Studie von Eisenhut et al. zeigt zum Beispiel, dass der photorespiratorische Glycolat Metabolismus für Cyanobakterien essentiell ist [64].

Um die „Oxygenierungsfunktion“ von RuBisCO zu reduzieren (und die photosynthetische Effizienz zu steigern) haben Cyanobakterien Mechanismen entwickelt, die lokal zu einer Erhöhung der CO₂ Konzentration führen [23]. Wie bereits beschrieben, dient das Carboxysom als Struktur in der RuBisCO vorzufinden ist. Der zu fixierende Kohlenstoff kann in Form von CO₂ oder HCO₃⁻ (Hydrogencarbonat) in die Zelle aufgenommen werden und in Form von cytosolischem Hydrogencarbonat in das Carboxysom diffundieren, wo es in CO₂ umgewandelt wird, welches dann fixiert werden kann (lokale Konzentrationserhöhung) [197].

Tag-Nacht-Rhythmik

Cyanobakterien unterliegen einem Tag-Nacht-Rhythmus (circadiane Rhythmik). Sofern Licht vorhanden ist, betreiben sie Photosynthese und assimilieren Speicherstoffe wie z. B. Glycogen, das in Dunkelperioden als endogenes Substrat genutzt werden kann. Es ist zu beobachten, dass eine hohe Verfügbarkeit von CO₂ dazu führt, dass mehr Glycogen gespeichert werden kann [65]. Die Energiegewinnung in der Nacht erfolgt i. d. R. durch Atmung (unter Verwendung von Sauerstoff als Elektronenakzeptor [235]). Dabei wird Glycogen über den oxidativen Pentose-Phosphat-Weg abgebaut. Das gewonnene NADPH wird in der respiratorischen Elektronentransportkette, an deren Ende Oxidasen stehen, verbraucht. Die Komponenten der Atmungskette sind sowohl in der Thylakoid- als auch in der cytoplasmatischen Membran lokalisiert.

TCA-Zyklus

Lange Zeit ist angenommen worden, dass es in Cyanobakterien aufgrund des nicht vorhandenen Enzyms 2-Oxoglutarat Dehydrogenase einen unvollständigen TCA-Zyklus gibt und dass die Gewinnung von Reduktionsäquivalenten nur über den oxidativen Pentose-Phosphat Weg erfolgen kann. Zhang et al. konnten 2011 zeigen, dass die Gene, die notwendig sind, um den TCA zu schließen, in fast allen Cyanobakterien vorhanden sind (bis auf *Prochlorococcus* und marine *Synechococcus* Spezies) [273]. Es handelt sich dabei

um die Gene, welche die 2-Oxoglutarat Decarboxylase und die Succinat-Semialdehyd Dehydrogenase kodieren. Unter Nutzung des geschlossenen TCA-Zyklus können Reduktionsäquivalente auch über diesen Weg produziert werden.

Alternative Energiegewinnung

Da es Cyanobakterien gibt, die in anoxygenen Bedingungen leben, existieren auch alternative Wege der Energiegewinnung unter Lichtausschluss. Es gibt Cyanobakterien (*Oscillatoria limnetica*), die in der Nacht endogene Speicherstoffe fermentieren und elementaren Schwefel reduzieren [183]. Einige Stämme betreiben Fermentation und wachsen anaerob auf exogenen C-Quellen [107, 145]. Es gibt Cyanobakterien, die Laktat (*Oscillatoria limnetica* [183]) oder Acetat (*Oscillatoria limosa* [102]) produzieren können oder gemischte Säure-Gärung (*Cyanothece* PCC 7822 [253]) betreiben.

Trotz der Vielfalt sind die meisten cyanobakteriellen Stämme obligat photoautotroph und nutzen in der Nacht die am Tage generierten (endogenen) Speicherstoffe. Unter aeroben Bedingungen und Lichtausschluss kann die Respiration ausreichend sein, um Wachstum zu ermöglichen [235]. Sofern Fermentation betrieben wird, kann i. d. R. davon ausgegangen werden, dass aufgrund der kleinen Energieausbeute nur Erhaltungsstoffwechsel (maintenance) stattfindet, obwohl es auch Ausnahmen gibt (z. B. *Microcystis aeruginosa* [235]).

Zusätzliche Nährstoffe

Neben den bisher beschriebenen Notwendigkeiten für das cyanobakterielle Leben (im Wesentlichen CO₂, Licht und Wasser) gibt es einige Komponenten, die zusätzlich aufgenommen werden müssen, um wachsen zu können. Dabei kann zwischen Mikro- und Makronährstoffen unterschieden werden [87]. Laut Bryants „The Molecular Biology of Cyanobacteria“ sind Mikronährstoffe z. B. Eisen, Kupfer oder nicht-organisches Phosphat und Makronährstoffe z. B. Stickstoff [35]. Es gibt Mechanismen, die sich im Laufe der Zeit entwickelt haben, um Defizite bestimmter Nährstoffe auszugleichen. So können einige Cyanobakterien unter Eisenmangel das eisenhaltige Ferredoxin durch Flavodoxin ersetzen [143], ohne die Photosynthesefunktion zu reduzieren [213]. Ähnliches gilt für Kupfer. Unter Kupfermangel können einige Cyanobakterien den Elektronentransporter Plastocyanin (der Kupfer enthält) durch Cytochrom c₅₃₃ ersetzen [106].

Die Assimilation von Stickstoff hingegen ist für Cyanobakterien essentiell. Ungefähr 8 - 10% des Trockengewichts einer cyanobakteriellen Zelle bestehen aus Stickstoff [90]. Der Stickstoff wird in den meisten Fällen in Form von Ammonium (NH₄⁺) oder Ni-

2 Biologische Grundlagen

trat (NO_3^-) aufgenommen. Allerdings gibt es auch Stämme, die Stickstoff in Form von Harnstoff oder als gasförmigen Stickstoff (N_2) aufnehmen können. Für die Aufnahme von Stickstoff (NO_3^- und NH_4^+) gibt es membrangebundene Transportsysteme, während N_2 per Diffusion in die Zelle gelangt [145]. Zellintern wird Ammoniak (NH_3) benötigt. Das aufgenommene Nitrat muss schrittweise reduziert und in NH_3 umgewandelt werden. Dieser Prozess benötigt sowohl Energie (ATP) als auch Reduktionsäquivalente (in Form von reduziertem Ferredoxin).

Einige Cyanobakterien sind in der Lage, elementaren Stickstoff (N_2) zu fixieren. Das dafür notwendige Enzym, die Nitrogenase, verbraucht dabei ATP, reduziert Ferredoxin und reduziert zeitgleich Protonen zu H_2 . Die Nitrogenase ist sehr sensitiv gegenüber Sauerstoff und wird durch diesen irreversibel inaktiviert [208]. Die Cyanobakterien haben Strategien entwickelt, dieses Problem zu umgehen. Wie zu Beginn erwähnt, gibt es Zelldifferenzierungen (Heterocysten), die kein Photosystem II haben und demnach keinen Sauerstoff produzieren, der die Nitrogenase beeinflussen könnte.

Sekundärmetabolismus

Viele Cyanobakterien sind in der Lage, verschiedene sekundäre Metabolite zu synthetisieren [80]. Sie produzieren z. B. Toxine, Siderophore, Phytohormone und photoprotektive Metabolite [266]. Beispiele für Cyanotoxine sind Microcystin, welches von *Microcystis* und *Nostoc* produziert wird. Dieses Toxin wird in der Leber von Wirbeltieren akkumuliert und zeigt einen hepatotoxischen Effekt [80]. Weitere Toxine sind Alkaloide wie Anatoxin und Saxitoxin. Dabei handelt es sich um Neurotoxine, die in hohen Dosen zum Tod führen können [161, 205]. Cyanobakterienstämme wie *Anabena* sp. PCC 7120 oder *Synechococcus* sp. PCC 7002 sind in der Lage, Siderophore zu produzieren (kleine organische Liganden, die Eisen binden können) [84, 112]. Phytohormone sind sekundäre Metabolite, die das Wachstum und die Entwicklung von Pflanzen stimulieren [205]. Auxin, ein sehr bedeutendes Phytohormon, wird z. B. von symbiotisch stickstofffixierenden Cyanobakterien produziert [224]. Einige cyanobakterielle Stämme sind in der Lage, photoprotektive Komponenten, wie z. B. Mycosporin-artige Aminosäuren oder Scytonemin, zu synthetisieren, die sie vor zu hoher/intensiver Strahlung schützen [75].

Die biotechnologische Anwendbarkeit vieler cyanobakterieller sekundärer Metabolite ist in vielen Fällen gezeigt worden [39, 232]. Deren biologische Aktivität reicht von „antimalarial“ [154] über „anticancer“ [231] und „antibiotic“ [104] bis zu „anti-HIV“ [58].

2.3 Produktionswege von Biokraftstoffen in Cyanobakterien

Cyanobakterien stellen einen vielversprechenden Produzenten für die Synthese verschiedener Produkte dar. Durch gezielte Eingriffe in den Stoffwechsel der Cyanobakterien wird versucht, die Produktpalette dieser Organismen zu vergrößern. Insbesondere die Produktion von Biokraftstoffen mit Cyanobakterien ist aussichtsreich, da damit die erdölbasierte Produktion von Kraftstoffen hin zur einer nachhaltigen und biobasierten Produktion mit schnell wachsenden und robusten Bakterien verschoben werden kann.

Deng und Coleman haben 1999 die Grundlage für Produktion von Ethanol mit Hilfe eines Cyanobakteriums gelegt [56]. Sie haben den cyanobakteriellen Stamm *Synechococcus* sp. PCC 7942 genetisch modifiziert, sodass dieser in der Lage war, Ethanol zu produzieren. Dafür haben sie eine Pyruvatdecarboxylase (PDC) und eine Alkoholdehydrogenase II (ADH) des Bakteriums *Zymomonas mobilis* in das Genom integriert. Der cyanobakteriellen Stamm ist damit derart modifiziert worden, dass dieser Sonnenenergie und CO₂ in Ethanol umwandeln konnte.

Mittlerweile gibt es eine Reihe an Biokraftstoffen die mit Hilfe von Cyanobakterien produziert werden können, so z. B. 1-Butanol [141, 142], Ethylen [114, 212, 240, 252], Isopren [153, 226], Isobutanol [21, 255] und weitere wertvolle Produkte [59]. Die Produktion von Biokraftstoff mit Cyanobakterien bietet im Vergleich zu Biokraftstoffproduzenten vorheriger Generationen einige Vorteile [214].

1. Cyanobakterien beanspruchen keine landwirtschaftlich nutzbare Fläche und konkurrieren damit nicht mit der Agrarindustrie.
2. Cyanobakterielles Wachstum ist relativ substratminimal. Sie benötigen lediglich Sonnenlicht, CO₂, (Salz)Wasser und eine kleine Menge an anorganischen Spurenelementen.
3. Viele cyanobakterielle Stämme sind einfach (und stabil) genetisch modifizierbar. Das Genom vieler Cyanobakterien ist bekannt und in Datenbanken verfügbar (z. B. CyanoBase [174] und CyanoGEBA [228]).

Die in der vorliegenden Arbeit untersuchten Produkte und deren Synthesewege sind in der Abbildung 2.3 dargestellt (siehe auch Tabelle A.1). Es handelt sich dabei um 1-Butanol, Ethanol, Ethylen, Isobutanol und Isopren. Diese Biokraftstoffe können bisher bereits mit Hilfe von Cyanobakterien produziert werden. Häufig ist die Produktion jedoch durch eine kleine Produktausbeute gekennzeichnet [210].

2 Biologische Grundlagen

Ethanol

Wie bereits beschrieben haben Deng und Coleman die Grundlage für die Produktion von Ethanol mit *Synechococcus elongatus* PCC 7942 [56] gelegt. Eine Weiterentwicklung ihrer Arbeit haben Dexter und Fu 2009 vorgestellt [57]. Sie haben (wie Deng und Coleman) eine PDC und eine ADH genutzt und in das Genom von *Synechocystis* sp. PCC 6803 integriert. Im Unterschied zu Deng und Coleman haben sie aber einen anderen Vektor zur Integration der Gene verwendet. Die Ausbeute der Ethanolproduktion war allerdings nicht wesentlich höher. Gao et al. haben 2012 postuliert, dass nach Integration der genannten Reaktionen und Deletion der Poly- β -Hydroxybuttersäureproduktion die Ausbeute erheblich gesteigert werden kann [82].

Isobutanol

Ein Ansatz Isobutanol mit Hilfe von Cyanobakterien zu produzieren, ist von 2009 Atsumi et al. vorgestellt worden [21]. Um Isobutyrylaldehyd, den direkten Präkursor für Isobutanol, produzieren zu können, ist eine Ketosäure Decarboxylase aus *Lactococcus Lactis* in das cyanobakterielle Genom integriert worden. Zusätzlich sind Gene integriert worden, die den Kohlenstofffluss in Richtung 2-Oxoisovalerat erhöhen und Gene die Isobutyrylaldehyd zu Isobutanol umwandeln können (Alkoholdehydrogenasen). Mit Hilfe dieser Veränderungen konnte Isobutanol in dem cyanobakteriellen Stamm *Synechococcus elongatus* PCC 7942 erfolgreich produziert werden.

1-Butanol

Der gleiche Stamm ist auch genutzt worden, um 1-Butanol zu produzieren. Lan et al. haben dafür den CoA-abhängigen Pfad von Acetyl-CoA zu 1-Butanol in das Genom von *Synechococcus elongatus* PCC 7942 integriert [141]. Dafür mussten insgesamt fünf verschiedene Gene in das Genom integriert werden (eine Acetyl-CoA Acetyltransferase aus *E. coli*, eine 3-Hydroxybutyryl-CoA Dehydrogenase, eine Crotonase und eine bifunktionale Aldehyd/Alkohol Dehydrogenase aus *Clostridium acetobutylicum*, die Trans-2-enoyl-CoA Reduktase aus *Treponema denticola*). Die Etablierung dieses Pfades innerhalb *Synechococcus elongatus* PCC 7942 gewährleistet die Produktion von 1-Butanol.

Isopren

Die Produktion von Isopren mit Hilfe von *Synechocystis* sp. PCC 6803 ist 2010 von Lindberg et al. gezeigt worden [153] (die Produktion von Isopren mit anderen Cyano-

2.3 Produktionswege von Biokraftstoffen in Cyanobakterien

bakterien, wie z. B. *Prochlorococcus* konnte schon früher demonstriert werden [226]). Das dafür notwendige Gen (die Isopren Synthase) ist aus der Pflanze *Pueraria montana* gewonnen worden. Bentley et al. konnten zeigen, dass die Produktion von Isopren gesteigert werden kann, wenn zusätzlich der Mevalonatweg in *Synechocystis* sp. PCC 6803 integriert und exprimiert wird, da dieser die Präkursoren für Isopren liefert [31].

Ethylen

Ethylen wird mit Hilfe des Ethylen-formenden-Enzyms (EFE) produziert. Das Enzym kodierende Gen ist aus *Pseudomonas syringae* gewonnen und in *Synechococcus elongatus* PCC 7942 integriert worden [212, 240]. Die Ethylen-produzierende Reaktion wandelt 2-Oxoglutarat in CO₂ und Ethylen um (dabei wird Sauerstoff verbraucht). Guerro et al. haben das EFE-Gen in *Synechocystis* sp. PCC 6803 integriert und konnten zeigen, dass auch dieser cyanobakterielle Stamm in der Lage ist, Ethylen zu produzieren [89].

Es gibt viele weitere Produkte, die mit Cyanobakterien synthetisiert werden können, wie z. B. Wasserstoff [10], Hexosen [175], Alkane [216], Fettsäuren [155] und extrazelluläre Polysaccharide [177] (für detaillierte Informationen siehe [191, 210, 214]).

Inzwischen gibt es einige Unternehmen, die intensiv an der nachhaltigen und biobasierten Produktion von Biokraftstoffen und weiteren wertvollen Produkten mit Hilfe von Cyanobakterien forschen und arbeiten. Beispiele solcher Firmen sind:

- Algenol
- Joule Unlimited
- AlgaEnergy s. A.
- Cyano Biotech GmbH
- Proterro

Viele der bisher vorgestellten Produkte werden mit Hilfe des cyanobakteriellen Stamms *Synechocystis* sp. PCC 6803 oder *Synechococcus elongatus* PCC 7942 synthetisiert. Aufgrund dessen werden in der vorliegenden Arbeit diese Cyanobakterien als Modellorganismen genutzt. Als zusätzlicher Vertreter wird die Spezies *Cyanothece* sp. ATCC 51142 analysiert, da diese in der Lage ist, einerseits Stickstoff zu fixieren und andererseits nativ bereits kleine Mengen Kohlenwasserstoffe produzieren kann [211]. Der Schwerpunkt der Analysen liegt auf *Synechocystis* sp. PCC 6803. Im Folgenden werden die untersuchten Cyanobakterien kurz dargestellt.

2 Biologische Grundlagen

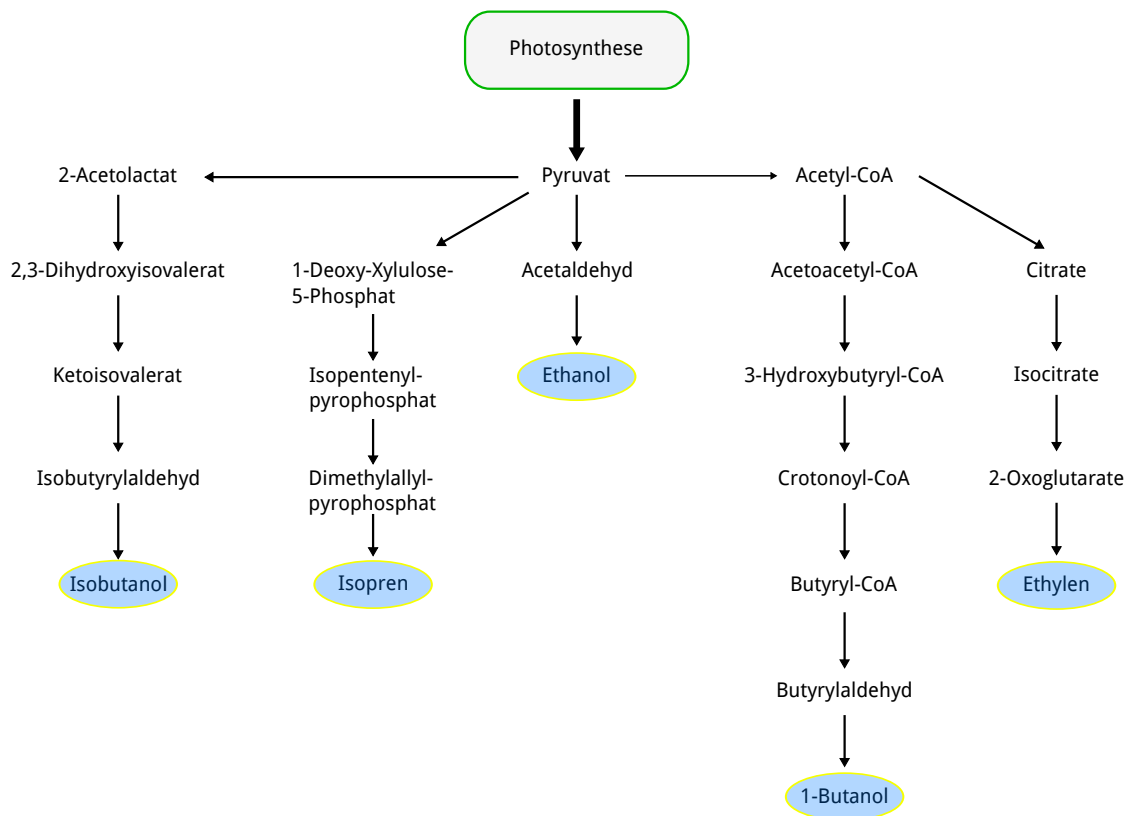


Abbildung 2.3: Produktionswege der untersuchten Biokraftstoffe (siehe auch Tabelle A.1).

2.4 Betrachtete Spezies

Synechocystis sp. PCC 6803

Synechocystis sp. PCC 6803 (im Folgenden *Synechocystis* genannt; siehe Abbildung 2.4) ist ein unizelluläres Cyanobakterium. I. d. R. beträgt der Durchmesser der Zellen nicht mehr als 4 μm . Sie betreiben gewöhnliche Zellteilung und besitzen keine Gasvakuolen. Des Weiteren fixieren diese Cyanobakterien keinen molekularen Stickstoff. Dieser Stamm ist das erste mal 1968 von R. Kunisawa aus Süßwasser isoliert worden (in Kalifornien, USA [236]). Dessen Genom ist 1996 durch Kaneko et al. komplett sequenziert worden [121]. *Synechocystis* ist auch in der Lage, photoheterotroph zu wachsen [15] und besitzt die Fähigkeit, extrazelluläre DNA aufzunehmen [60, 85]. Mittlerweile sind viele Informationen über *Synechocystis* vorhanden (auch aufgrund der frühen Sequenzierung), sodass

dieses Cyanobakterium als ein sehr guter Modellorganismus betrachtet werden kann [171].

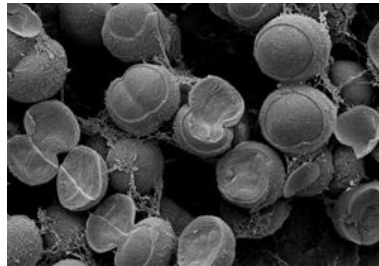


Abbildung 2.4: Aufnahme von *Synechocystis* sp. PCC 6803 [6].

Synechococcus elongatus PCC 7942

Synechococcus elongatus PCC 7942 (im Folgenden *Synechococcus* genannt; siehe Abbildung 2.5; früher *Anacystis nidulans* genannt) ist ein unizelluläres Süßwassercyanobakterium. Der Durchmesser dieser Cyanobakterien ist i. d. R. nicht größer als $1,2 \mu\text{m}$. Sie sind obligat photoautotroph [165] und besitzen eine geringe Salz-Toleranz [145]. Dass das Genom von *Synechococcus* genetisch transformierbar ist, wurde bereits 1970 gezeigt [227]. Dieser Organismus ist in der Forschung intensiv genutzt worden, um verschiedene Prozesse zu untersuchen und aufzuklären. Beispiele dafür sind: der Nitrat Metabolismus [146], der sogenannte Kohlenstoff-Konzentrations-Mechanismus [34] und die circadiane Genexpression [168]. Wie bereits gezeigt werden konnte, ist *Synechococcus* auch für die Produktion von Biokraftstoffen ein geeigneter Modellorganismus [21, 56].

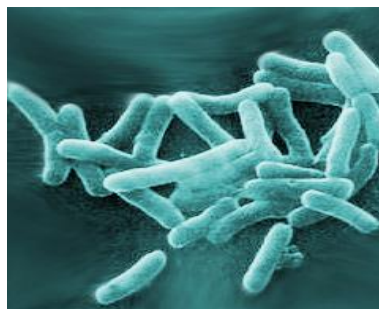


Abbildung 2.5: Aufnahme von *Synechococcus elongatus* PCC 7942 [5].

2 Biologische Grundlagen

Cyanothece sp. ATCC 51142

Cyanothece sp. ATCC 51142 (im Folgenden *Cyanothece* genannt; siehe Abbildung 2.6) ist ein unizelluläres Cyanobakterium [145], das Stickstoff fixieren kann. Das Genom ist komplett entschlüsselt. Leider stehen für dieses Cyanobakterium bisher wenige genetische Werkzeuge zur Verfügung, obwohl es interessante native Produktionswege für z. B. Ethanol besitzt [211, 265]. Deshalb wird auch an diesem Cyanobakterium weiter geforscht [74, 169] und genomskalige Modelle des Stoffwechsels stehen zur Verfügung [211, 262]. Aus den genannten Gründen wird *Cyanothece* als zusätzlicher Modellorganismus analysiert.

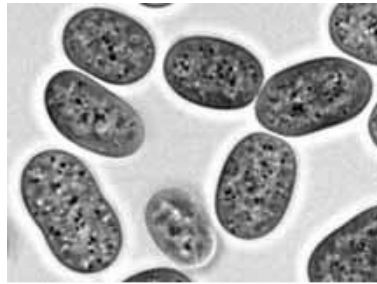


Abbildung 2.6: Aufnahme von *Cyanothece* sp. ATCC 51142 [2].

3 Mathematische Grundlagen

Im folgenden Kapitel werden alle für diese Arbeit relevanten mathematischen Grundlagen erläutert. In Kapitel 3.1 wird erklärt, was Stoffwechselnetzwerke sind. (3.2) zeigt, wie diese modelliert und analysiert werden können. Des Weiteren werden Methoden zum modellgestützten Metabolic Engineering vorgestellt (3.3). Kapitel 3.5 gibt eine Übersicht über verschiedene stöchiometrische Stoffwechselmodelle von Cyanobakterien.

3.1 Stoffwechselnetzwerke

Der Begriff Stoffwechsel, oder auch Metabolismus, beschreibt die Gesamtheit aller biochemischen Reaktionen in einem Organismus. Dabei gliedert sich dieser in den Katabolismus, also den Abbau von Nährstoffen und den Anabolismus, der den Aufbau von zelleigenen Stoffen beschreibt.

Der Begriff Netzwerk findet in vielen Bereichen des täglichen Lebens Anwendung. So z. B. in der Elektrotechnik, als eine gemeinsame Schaltung von elektronischen Bauelementen; in der elektronischen Datenverarbeitung als Rechnernetzwerk, also einem Zusammenschluss von verschiedenen elektronischen Systemen, um eine Kommunikation der einzelnen Elemente untereinander zu erlauben; aber auch im sozialen Bereich (sogenannte soziale Netzwerke wie beispielsweise Facebook). Der Begriff Netzwerk beschreibt also eine Art Gemeinschaft oder Zusammenschluss von verschiedenen Elementen, die auf unterschiedliche Art und Weise miteinander interagieren.

Wenn von Stoffwechselnetzwerken die Rede ist, sind die Elemente verschiedene Stoffwechselspezies (Metabolite), die durch biochemische Reaktionen miteinander in Wechselwirkung stehen. Diese Stoffwechselreaktionen werden durch verschiedene Charakteristika beschrieben.

- Die **Stöchiometrie** beschreibt das Verhältnis der beteiligten Edukte und Produkte in einer Reaktionsgleichung durch die jeweiligen stöchiometrischen Koeffizienten

3 Mathematische Grundlagen

(auch Verhältniszahlen genannt). Diese Koeffizienten geben an, wie viel Stoffmenge eines Edukts verbraucht und wie viel Stoffmenge eines Produkts dabei gebildet wird.

- **Reversibilität:** Im Allgemeinen gilt, dass alle chemischen Reaktionen reversibel sind. Das heißt, dass eine Reaktion vor- und rückwärts ablaufen kann. Unter biologisch relevanten Bedingungen können allerdings viele Reaktionen als praktisch irreversibel betrachtet werden, da viele fast ausschließlich in eine Richtung ablaufen.
- **Enzyme**, die eine Reaktion katalysieren und deren Aktivierungsenergie heruntersetzen. Fast alle Stoffwechselreaktionen werden durch Enzyme katalysiert. Damit wird ein Zusammenhang zum Genom erstellt, da die Enzyme in der DNA kodiert sind. Allerdings ist diese Relation (Gen-Enzym-Reaktion) nicht immer eindeutig, da bestimmte Enzyme auch mehrere Reaktionen katalysieren können, oder eine Reaktion durch mehrere Enzyme katalysiert werden kann. Des Weiteren können Enzymkomplexe in verschiedenen Genen kodiert sein. Trotzdem ermöglicht diese Relation einen Rückschluss von Genen auf Reaktionen und bildet damit die Grundlage für die Rekonstruktion von Stoffwechselnetzwerken.
- Die **Kinetik** einer Reaktion beschreibt ihre Dynamik auf Grundlage eines bestimmten kinetischen Mechanismus (z. B. Massenwirkungskinetik).

Um den Stoffwechsel eines bestimmten Organismus adäquat beschreiben zu können, bedarf es der Identifikation aller relevanten Reaktionen und Metabolite. Mit anderen Worten: am Anfang steht die Rekonstruktion des Stoffwechsels. Diese Rekonstruktion basiert im Allgemeinen auf verschiedenen Informationsquellen, wie Primärliteratur, verschiedenen Übersichtsartikeln, sowie Datenbanken, wie z. B. „Kyoto Encyclopedia of Genes and Genomes“ (KEGG [119, 120]) oder BioCyc [45]. Der Ablauf der Rekonstruktion ist bereits beschrieben worden [70, 245] und kann in vier Hauptarbeitsschritten zusammengefasst werden.

1. Die Entwicklung eines „Vorläufernetzwerkes“ basierend auf Informationen aus der Literatur und verschiedenen Datenbanken (KEGG, BioCyc).
2. Eine Überprüfung und Validierung des entwickelten Modells, um Lücken zu schließen und/oder Fehler zu beheben.
3. Das entwickelte Modell muss anschließend in ein mathematisches Modell übersetzt werden, das nach Möglichkeit in einer einheitlichen Sprache (wie z. B. „Systems

3.2 Stöchiometrische Modellierung von Stoffwechselnetzen

Biology Markup Language“ [110]) formuliert werden sollte, um einen einfachen Austausch des Stoffwechselmodells zu ermöglichen.

4. Dieses Modell kann dann entsprechend analysiert und genutzt werden, um tiefere Einblicke in den Stoffwechsel des betrachteten Organismus zu gewinnen.

Die Schritte 2 - 4 können bzw. sollten iterativ durchlaufen werden, da i. d. R. nicht davon auszugehen ist, dass das erzeugte Modell fehlerfrei sein wird. Das modellierte Stoffwechselnetz kann dann beispielsweise mit Methoden des „constraint-based modeling“ analysiert werden.

3.2 Stöchiometrische Modellierung von Stoffwechselnetzen

(„constraint-based modeling“)

Wie im vorherigen Kapitel beschrieben, ist eine biochemische Reaktion durch ihre Stöchiometrie, Reversibilität, Kinetik und ihre Assoziation mit Enzymen und Genen charakterisiert. Für die stöchiometrische Modellierung von Stoffwechselnetzen wird die Kinetik einer Reaktion nicht berücksichtigt und nur die Topologie des Stoffwechsel betrachtet. Um die verschiedenen Reaktionen des Stoffwechsels besser beschreiben zu können, ist es hilfreich, zwischen den verschiedenen Komponenten des Stoffwechsels zu unterscheiden [238].

- Das **Substrat** oder Nährstoffe, die metabolisiert werden können (z. B. Zucker, Stickstoff, etc.). Die Aufnahme externer Substanzen wird dabei als eine Transportreaktion betrachtet und in das Modell integriert.
- **Produkte**, welche die Zelle synthetisiert und ausscheiden kann (z. B. Ethanol, CO₂ oder sekundäre Metabolite). Der Export dieses Produkts kann wieder als eine Transportreaktion betrachtet und in das Modell integriert werden.
- Die **Bestandteile der Biomasse** sind verschiedene Makromoleküle (wie z. B. Proteine, DNA, RNA, etc.). Die sogenannte Biomassesynthesereaktion (BSR) beschreibt dann die molaren Verhältnisse von Energie, Redoxkraft und weiteren Bestandteilen, die notwendig sind, um ein Gramm Trocken-Biomasse (gTM) zu synthetisieren. Die BSR wird dabei als eine Pseudoreaktion in das Modell integriert. Ihre Rate ist die Wachstumsrate μ [h⁻¹] und die Stöchiometrie der Metabolite in der BSR tragen die Einheit mmol/gTM.

3 Mathematische Grundlagen

- Alle anderen beteiligten Stoffwechselspezies stellen die **intrazellulären Metabolite** dar. Diese internen Spezies werden bei der stöchiometrischen Modellierung explizit bilanziert.

Substrate und Produkte, die außerhalb der Zelle vorliegen, werden als externe Metabolite bezeichnet und bilden Quellen und Senken des Systems. Formal lässt sich das stöchiometrische Stoffwechselnetz beschreiben durch:

- m : die Anzahl der internen Metabolite,
- q : die Anzahl der Reaktionen,
- \mathbf{N} : die stöchiometrische Matrix bestehend aus m Zeilen und q Spalten mit den stöchiometrischen Koeffizienten n_{ji} des Metaboliten j in Reaktion i .

Des Weiteren wird berücksichtigt, welche Reaktionen reversibel und welche irreversibel sind:

- rev : Indexmenge aller reversiblen Reaktionen und
- $irrev$: Indexmenge aller irreversiblen Reaktionen.

Es gilt:

$$rev \cap irrev = \emptyset \quad \text{und} \quad |rev \cup irrev| = q. \quad (3.1)$$

Die Konzentrationsänderungen der Metabolite können durch die folgende Differentialgleichung beschrieben werden:

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{N}\mathbf{r}(t). \quad (3.2)$$

Der Reaktionsratenvektor $\mathbf{r}(t)$ ist wiederum eine Funktion der Metabolitkonzentrationen $\mathbf{c}(t)$ und verschiedenen (kinetischen) Parametern \mathbf{p} :

$$\mathbf{r}(t) = \mathbf{f}(\mathbf{c}(t), \mathbf{p}). \quad (3.3)$$

Der Reaktionsratenvektor $\mathbf{r}(t)$ beinhaltet die Informationen über die Kinetik und die Regulation. Die stöchiometrische Matrix \mathbf{N} hingegen beschreibt die Interaktionen der Metaboliten untereinander, die unabhängig sind von der Kinetik, der Zeit und der Konzentration. Dieser Teil der Gleichung (3.2) kann folglich als invarianter Teil der Differentialgleichung angesehen werden und bildet die Grundlage der stöchiometrischen und „constraint-based“ Modellierung von Stoffwechselnetzen.

3.2 Stöchiometrische Modellierung von Stoffwechselnetzen

3.2.1 Stationäre Flussverteilungen

Im Vergleich mit regulatorischen Prozessen innerhalb eines Organismus ist der Metabolismus geprägt von sehr schnell ablaufenden Reaktionen, wodurch Metabolite vor allem im Zentralstoffwechsel sehr schnell umgesetzt werden (Verhältnis von Reaktionsgeschwindigkeiten zu Metabolitkonzentration ist sehr hoch). Es kann daher vereinfacht angenommen, dass sich die Metabolitkonzentrationen nicht verändern und in einem Fließgleichgewicht befinden [238]:

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{N}\mathbf{r}(t) = \mathbf{0}. \quad (3.4)$$

Des Weiteren werden die Reaktionsraten als unabhängige Variablen betrachtet, womit sich Gleichung (3.2) vereinfacht zu:

$$\mathbf{N}\mathbf{r} = \mathbf{0}. \quad (3.5)$$

Bei Gleichung (3.5) handelt es sich nun um ein homogenes System linearer Gleichungen. Bzgl. der internen Metabolite muss sich die Summe aus Verbrauch und Produktion aufheben. Es sei an dieser Stelle noch einmal darauf hingewiesen, dass extrazelluläre Spezies (externe Metabolite) bei der expliziten Modellierung (Massenbilanz) nicht berücksichtigt werden.

Alle Lösungen \mathbf{r} des Gleichungssystems (3.5) bilden sogenannte stationäre Flussverteilungen. Der Kern oder auch der rechte Nullraum der stöchiometrischen Matrix \mathbf{N} enthält die Menge aller Lösungen. Die Dimension des Nullraums spiegelt dabei die Anzahl der linear unabhängigen Lösungen (die Anzahl der Freiheitsgrade) für Gleichung (3.5) wider und kann durch $q - \text{Rang}(\mathbf{N})$ berechnet werden. Sofern die Anzahl der linear unabhängigen Reaktionen größer ist als die Anzahl der Metabolite m , existieren unendlich viele Lösungen für Gleichung (3.5) (das gilt für die meisten Stoffwechselmodelle). Gilt $q = \text{Rang}(\mathbf{N})$ existiert nur die Lösung $\mathbf{r} = \mathbf{0}$ für Gleichung (3.5), das ein System mit null Freiheitsgraden abbildet. Der Kern des Nullraums kann als Matrix dargestellt werden und enthält eine Basis des Nullraums:

$$\mathbf{N}\mathbf{K} = \mathbf{0}. \quad (3.6)$$

Die Kernmatrix \mathbf{K} erlaubt Rückschlüsse auf bestimmte strukturelle Zusammenhänge des Stoffwechselmodells. Es können blockierte Reaktionen (Reaktionen deren Raten immer null sind) und gekoppelte Reaktionen (Reaktionen die immer gemeinsam in einem bestimmte Verhältnis zueinander ablaufen) identifiziert werden.

3 Mathematische Grundlagen

Anhand des Beispielnetzwerks in Abbildung 3.1 (A und B) können die bisher gezeigten Grundlagen verdeutlicht werden. Das Modell besteht aus neun Reaktionen, fünf intrazellulären und vier extrazellulären Metaboliten. Die stöchiometrische Matrix \mathbf{N} besteht demnach aus fünf Zeilen und neun Spalten. Es beschreibt die Aufnahme von zwei Substraten ($S1_{\text{extern}}$ und $S2_{\text{extern}}$). Allerdings kann $S2_{\text{extern}}$ auch als Produkt exportiert werden. Des Weiteren stellt es den Export der Produkte $P1_{\text{extern}}$ und $P2_{\text{extern}}$ dar. Von den insgesamt neun Reaktionen sind zwei reversibel und die verbleibenden sieben Reaktionen irreversibel. Der Rang der stöchiometrischen Matrix ist fünf. Demnach verfügt das System über vier Freiheitsgrade.

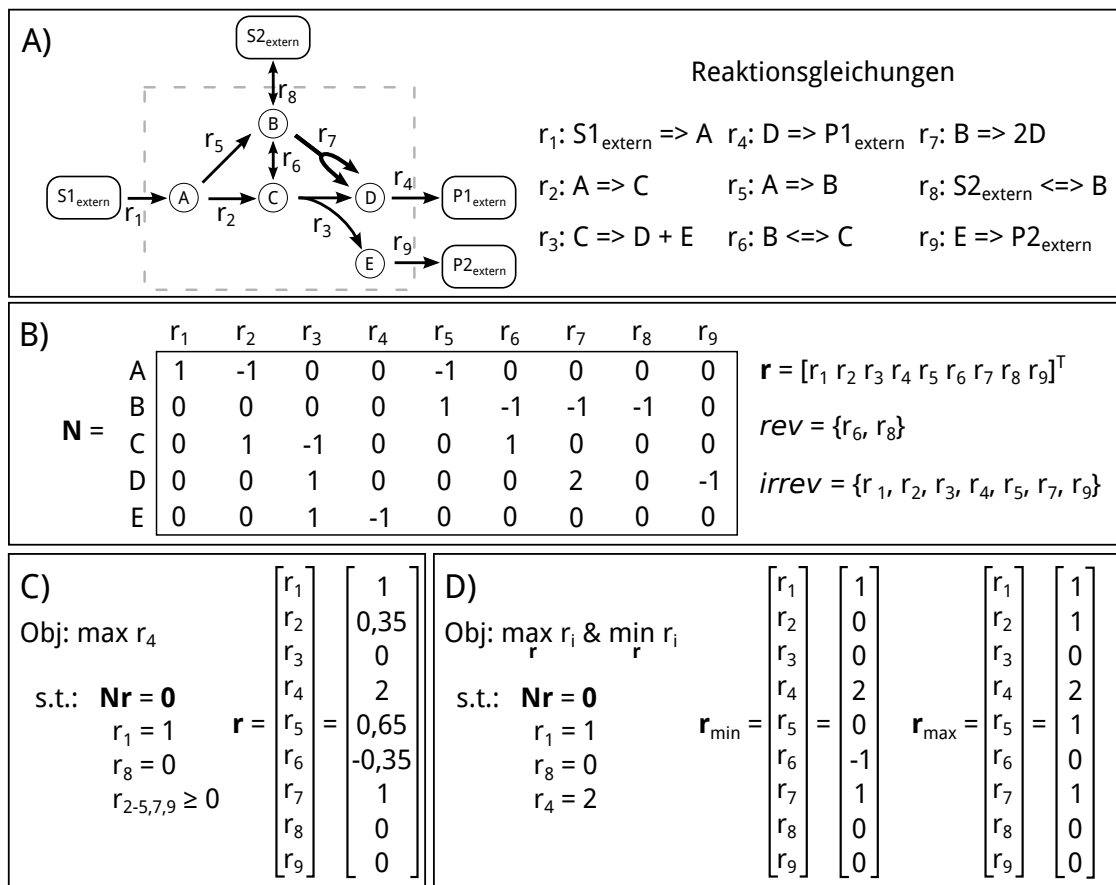


Abbildung 3.1: Stöchiometrische Modellierung von Stoffwechselnetzen. A) Das verwendete Beispielnetzwerk und die dazugehörigen Reaktionsgleichungen. B) Die stöchiometrische Matrix \mathbf{N} , der Reaktionsratenvektor \mathbf{r} und die Menge der reversiblen (rev) und irreversiblen ($irrev$) Reaktionen. C) Beispiel für eine Flussbilanzanalyse. D) Beispiel für eine Flussvariabilitätsanalyse, welche die Variabilität der Flüsse bzgl. des Optimierungsproblems in (C) zeigt. Grafik angelehnt an [251].

3.2 Stöchiometrische Modellierung von Stoffwechselnetzen

Es existieren i. d. R. unendlich viele Lösungen (Flussverteilungen) für Gleichung (3.5) und der Lösungsraum ist unbeschränkt. Die Verwendung zusätzlicher Nebenbedingungen (constraints) ermöglicht es, den Lösungsraum einzuschränken [198, 199]:

$$(NB1) \quad r_i \geq 0 \quad \forall i \in irrev \quad (\text{Reversibilität}), \quad (3.7)$$

$$(NB2) \quad \alpha_i \leq r_i \leq \beta_i \quad (\text{Kapazität/Reversibilität}), \quad (3.8)$$

$$(NB3) \quad \mathbf{A}\mathbf{r} \leq \mathbf{b} \quad (\text{inhomogene Nebenbedingungen, z. B. gemessene Flussraten}). \quad (3.9)$$

Die Verwendung der Gleichung (3.5) und der genannten Nebenbedingung(en) repräsentiert das Fundament des „constraint-based modeling“ (CBM). *NB2* (3.8) beschreibt bekannte obere (β_i) und untere (α_i) Schranken von Reaktionsraten. Damit können Grenzen für Austauschreaktionen (Aufnahme/Ausscheidung) gesetzt werden. Falls keine Grenzen bekannt sind, kann die Kapazität von Reaktion r_i als unbeschränkt definiert werden ($\pm\infty$); für alle irreversiblen Reaktionen kann zudem $\alpha_i = 0$ gesetzt werden (*NB1* (3.7)). Dies kann auch direkt in (3.8) integriert werden. Die Verwendung von Gleichung (3.5) zusammen mit den Reversibilitätsnebenbedingungen *NB1* (3.7) schränkt den Lösungsraum derart ein, dass ein konvexer polyedrischer Kegel geformt wird [209] (auch Flusskegel genannt und hier mit F bezeichnet; siehe Abbildung 3.2B). Es gilt:

$$F = \{\mathbf{r} \in \Re^q \mid \mathbf{N}\mathbf{r} = \mathbf{0}, r_i \geq 0 \quad \forall i \in irrev\}. \quad (3.10)$$

Durch Hinzunahme der *NB2* (3.8) und *NB3* (3.9) wird der Lösungsraum auf einen beschränkten und/oder unbeschränkten Polyeder begrenzt. Im Folgenden ist, wenn nicht anders angegeben, $r_i \geq 0 \quad \forall i \in irrev$ in $\alpha_i \leq r_i \leq \beta_i$ enthalten.

3 Mathematische Grundlagen

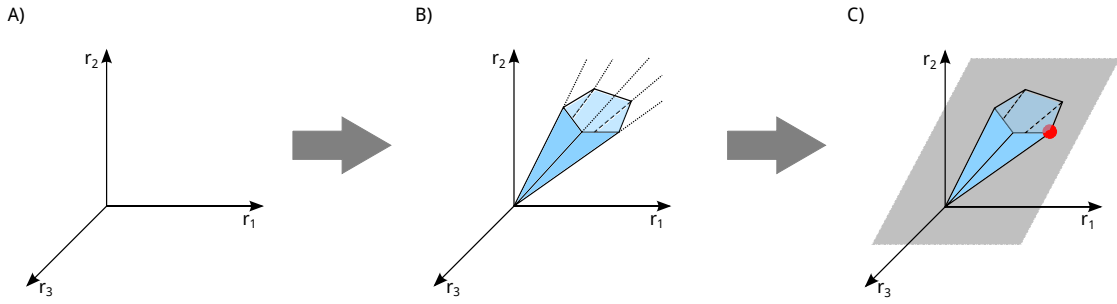


Abbildung 3.2: Das Konzept des constraint-based modeling. A) Zeigt den Raum aller Flussverteilungen an. B) Durch Nutzung der Gleichung (3.5) und der Reversibilitätsnebenbedingungen *NB1* (3.7) wird der zulässige Lösungsraum auf einen konvexen polyedrischen Kegel beschränkt. C) Unter Hinzunahme von *NB2* (3.8) und/oder *NB3* (3.9) entsteht ein unbeschränkter oder beschränkter Polyeder. Hier wird der Polyeder z. B. durch eine max. Substrataufnahmerate nach oben beschränkt (graue Fläche). Der rote Punkt zeigt die optimale Lösung für eine Flussbilanzanalyse an. Grafik angelehnt an [186].

3.2.2 Flussbilanzanalyse

Die Flussbilanzanalyse (FBA) ist ein mächtiges Werkzeug des CBM [186]. Im Prinzip dient diese Analyse dazu, unter Nutzung der genannten Nebenbedingungen und der Annahme einer bestimmten Zielfunktion, eine optimale Flussverteilung in einem stöchiometrischen Stoffwechselmodell zu berechnen. Dies kann unter Einbeziehung der Gleichung (3.5) und der Nebenbedingungen (3.7) - (3.9) formal geschrieben werden:

$$\begin{aligned}
 & \mathbf{c}^T \mathbf{r} = c_1 r_1 + c_2 r_2 + \dots + c_q r_q = \max ! \\
 \text{s.t.} \quad & \mathbf{N} \mathbf{r} = \mathbf{0} \\
 & \mathbf{A} \mathbf{r} \leq \mathbf{b} \\
 & \alpha_i \leq r_i \leq \beta_i.
 \end{aligned} \tag{3.11}$$

Dieses lineare Optimierungsproblem liefert ein eindeutiges Optimum bzgl. der Zielfunktion. Allerdings ist die Flussverteilung nicht immer eindeutig. Das lineare Optimierungsproblem kann mit Methoden der linearen Programmierung gelöst werden [186].

Die Definition der Zielfunktion ist dabei von entscheidender Bedeutung. Häufig wird die Maximierung der Wachstumsrate als Zielfunktion definiert [72]. Es können auch andere Kriterien als Zielfunktion definiert werden [218], wie z. B. die Minimierung der Substrataufnahmerate (um die effizienteste Nutzung des Substrates abzubilden [182]),

3.2 Stöchiometrische Modellierung von Stoffwechselnetzen

die Maximierung der ATP Produktion [61] oder die Minimierung der Summe der zellinternen Flüsse (um die Effizienz der Enzyme bzgl. des Zellwachstums abzubilden [33]).

Abbildung 3.1C beschreibt die FBA anhand des Beispielnetzwerks. In diesem Fall soll die Rate r_4 maximiert werden unter der Nebenbedingung, dass $r_1 = 1$ und $r_8 = 0$ ist (Aufnahme von $S1_{\text{extern}}$ als einziges Substrat). Des Weiteren soll die FBA die Nebenbedingungen erfüllen, dass $r_{2-5,7,9} \geq 0$ ist. Mit den gegebenen Bedingungen wird r_4 maximiert. Der gefundene optimale Flussvektor zeigt an, dass unter diesen Umständen r_4 nicht größer als zwei werden kann. Diese Rate ist damit eindeutig bestimmt. Dies gilt allerdings nicht für alle Raten. So wird z. B. deutlich, dass der Fluss durch r_1 beliebig zwischen der Reaktion r_5 und r_2 aufgeteilt werden kann, ohne dass dies einen Einfluss auf r_4 hat.

Obwohl die FBA keine kinetischen Parameter verwendet und auch keine Aussage über Metabolitkonzentrationen zulässt, sind die Anwendungsmöglichkeiten enorm [122, 147, 203]. Mit Hilfe der FBA können z. B. Fragen nach optimaler Ausbeute (Produkt, oder Biomasse), bestimmter Funktionen des Modells (ist die Synthese eines Metaboliten realisierbar) oder die Frage nach bestimmten (optimalen) physiologischen Leistungen des Modells beantwortet werden. Weitere Anwendungsbeispiele sind physiologische Studien in denen die Wachstumsraten untersucht werden [180]; sogenannte „Gap-filling“ Analysen in denen identifiziert wird, in welchen Bereichen des Stoffwechselmodells noch Unvollständigkeiten bzw. Lücken sind [200]; oder auch sogenannte „Gene-Essentiality“ Studien [62, 63], in denen analysiert wird, welche Reaktionen essentiell sind, um Wachstum zu garantieren.

Es gibt mittlerweile Methoden, die über die klassische FBA hinaus gehen und weitere Informationen berücksichtigen können. Beispiele dafür sind regulatorische FBA (rFBA [49, 51, 99]), die Integration von thermodynamischen Nebenbedingungen [29, 109] oder dynamische FBA (dFBA [254]). Die rFBA integriert regulatorische Prozesse mit Hilfe Boolescher Logik (ein bestimmtes Gen ist aktiv in Abhängigkeit von bestimmten internen und/oder externen Metaboliten und/oder Transkriptionsfaktoren). Dieses regulatorische Modell wird mit dem FBA Modell kombiniert und nacheinander gelöst. Die aus dem regulatorischen Modell abgeleiteten Erkenntnisse dienen dann als Grundlage der FBA (z. B. unter bestimmten Umweltbedingungen sind nur gewisse Reaktionen aktiv; Realisierung über die Kapazitätsnebenbedingungen). Shlomi et al. [230] und Covert et al. [50] konnten zeigen, dass rFBA genauere Ergebnisse als die klassische FBA liefern kann. Die Berücksichtigung von thermodynamischen Nebenbedingungen ist ein weiteres Beispiel für den erweiterten Einsatz der FBA. Hoppe et al. [109] nutzen dafür die Änderung der freien Gibbs'schen Energie, die einen Rückschluss auf die Flussrichtung

3 Mathematische Grundlagen

zulässt. Die Änderung der freien Gibbs'schen Energie ist wiederum abhängig von den Metabolitkonzentrationen und somit können experimentelle Daten bzgl. der Metabolitkonzentrationen über diesen Ansatz berücksichtigt werden. Ein weiteres Beispiel ist die dynamische FBA (dFBA), bei welcher der Konzentrationsverlauf und die Aufnahme von externen Spezies dynamisch modelliert wird, während für die internen Metabolite Quasistationarität (Gleichung (3.5)) angenommen wird [254].

3.2.3 Flussvariabilitätsanalyse

Die Flussvariabilitätsanalyse (FVA), eine weitere Technik des CBM, bestimmt den zulässigen Wertebereich von Reaktionsraten in einem Flusssszenario unter Berücksichtigung der Gleichung (3.5) und gegebener Nebenbedingungen (3.7) - (3.9) [160]. Um die zulässigen oberen und unteren Schranken einer Reaktionsrate zu berechnen, werden zwei FBA-Probleme formuliert. Prinzipiell kann diese Berechnung ohne die Berücksichtigung einer biologischen Zielfunktion erfolgen. Um z. B. multiple Lösungen bzgl. eines bestimmten zellulären Verhaltens zu berechnen, können bekannte Reaktionsraten als zusätzliche Nebenbedingung in das FBA-Problem integriert werden ($\mathbf{Ar} \leq \mathbf{b}$). Unabhängig davon ergeben sich insgesamt $2q$ Optimierungsprobleme, da jede Reaktionsrate minimiert und maximiert wird. Dies kann formal dargestellt werden als:

$$\begin{array}{ll}
 \max_{\mathbf{r}} r_i & \min_{\mathbf{r}} r_i \\
 s.t. \quad \mathbf{Nr} = \mathbf{0} & s.t. \quad \mathbf{Nr} = \mathbf{0} \\
 \alpha_i \leq r_i \leq \beta_i & \alpha_i \leq r_i \leq \beta_i \\
 \mathbf{Ar} \leq \mathbf{b}, & \mathbf{Ar} \leq \mathbf{b}.
 \end{array} \tag{3.12}$$

Die Lösung der $2q$ Optimierungsprobleme liefert dann die zulässigen Schranken einer jeden Reaktionsrate für ein gegebenes Szenario. Die Variabilität einer Reaktionsrate lässt dann verschiedene Rückschlüsse zu. Beispielsweise kann abgeleitet werden, ob eine Reaktion essentiell ist. Dafür muss entweder das Minimum größer null oder das Maximum einer Reaktionsrate kleiner null sein. Entspricht der Wert des Minimums dem des Maximums, ist diese Reaktionsrate eindeutig berechenbar. Ein spezieller Fall ist gegeben, sobald beide Werte null sind. Dann ist diese Reaktion blockiert.

Die FVA kann anhand des Beispielnetzwerks verdeutlicht werden (siehe Abbildung 3.1D). Dafür wird das Minimum und das Maximum aller neun Reaktionsraten berechnet, wenn $r_1 = 1$ und $r_8 = 0$ wie im vorherigen Szenario gesetzt werden (Aufnahme von $S1_{\text{extern}}$). Zusätzlich wird der im vorherigen Anwendungsbeispiel berechnete optimale

3.2 Stöchiometrische Modellierung von Stoffwechselnetzen

Wert für die Reaktion r_4 ($r_4 = 2$) in das Optimierungsproblem integriert, um multiple Lösungen für dieses Szenario beschreiben zu können. Die Ergebnisse zeigen nun, im Gegensatz zur FBA, die Variabilitäten aller Reaktionsraten unter Berücksichtigung der genannten Nebenbedingungen. Aus den Ergebnissen der FVA kann u. a. abgeleitet werden, dass $r_3 = 0$ und $r_9 = 0$ ist. Das bedeutet, dass diese Reaktionen in diesem Szenario blockiert sind. Die Reaktion r_7 hingegen ist essentiell. Es wird deutlich, dass in diesem Szenario nur die Reaktionen r_2, r_5 und r_6 in einem bestimmten Wertebereich variabel sind. Alle anderen Reaktionsraten sind eindeutig berechenbar. Es gibt zahlreiche Beispiele in denen die FVA als Analysemethode erfolgreich angewendet wurde [40, 67, 96, 186].

3.2.4 Elementarmodenanalyse

Die Elementarmodenanalyse hat das Ziel, minimale funktionale Einheiten (oder Pathways) eines stöchiometrischen Netzwerks im Fließgleichgewicht zu identifizieren. Minimal bezieht sich darauf, dass, sobald ein Teil dieser Einheit entfernt wird, diese keine Funktion mehr im steady-state wahrnehmen kann.

Prinzipiell gibt es zwei etablierte ähnliche Ansätze zur mathematischen Beschreibung solcher Einheiten: die „Extreme Pathways“ [215] und die Elementarmoden [219, 220, 221]. Allerdings sind die extremen Pathways eine Untermenge der Elementarmoden und damit selbst Elementarmoden [272]. Es wird deshalb an dieser Stelle nicht weiter auf die extremen Pathways eingegangen (eine vergleichende Analyse der genannten Ansätze ist zu finden in [132, 190]).

Für die Beschreibung der Elementarmoden wird die Notation des *supports* eingeführt. Der *support* eines Flussvektors \mathbf{r} ist definiert als $\text{supp}(\mathbf{r}) = \{i | r_i \neq 0\}$ und beschreibt eine Menge von Indizes der Nicht-Null Elemente in \mathbf{r} . Ein Elementarmodus bildet damit eine eindeutige und support-minimale Menge von Reaktionen, die als stationäre Flussvektoren operieren. Das bedeutet, dass sobald nur eine Reaktion dieser support-minimalen Menge an Reaktionen entfernt wird, dieser Elementarmodus inaktiv ist [219, 220, 221]. Formal ist ein Flussvektor \mathbf{e} ein Elementarmodus (EM), sofern dieser die folgenden Kriterien erfüllt:

$$(K1) \quad \mathbf{N}\mathbf{e} = \mathbf{0},$$

$$(K2) \quad e_i \geq 0 \quad \forall i \in \text{irrev},$$

$$(K3) \quad \text{es existiert keine nicht-triviale Flussverteilung } \mathbf{r} \text{ die (K1) und (K2) erfüllt und für die gilt: } \text{supp}(\mathbf{r}) \subsetneq \text{supp}(\mathbf{e}).$$

3 Mathematische Grundlagen

Die Kriterien (K1) und (K2) entsprechen genau der steady-state-Bedingung (Gleichung (3.5)) bzw. der Nebenbedingung *NB1* (3.7) und bilden den schon besprochenen Flusskegel (Gleichung (3.10)). Ein Flussvektor \mathbf{e} kann per Definition nur ein EM sein, sofern dieser das Fließgleichgewicht (K1) erfüllt, alle irreversiblen Reaktionen in die thermodynamisch mögliche Richtung ablaufen (K2) und es (K3) keinen nicht-trivialen Flussvektor \mathbf{r} gibt, der (K1) und (K2) erfüllt und zerlegbar ist. Daraus folgt, dass, sofern eine Reaktion $i \in \text{supp}(\mathbf{e})$ des EM \mathbf{e} entfernt wird, dieser EM nicht mehr funktional ist. Mit anderen Worten, es kann keine nicht-triviale Flussverteilung des verbleibenden Teilnetzes realisiert werden. Ein Elementarmodus ist durch seinen support eindeutig bestimmt (wird ein EM \mathbf{e} mit einem positiven Skalar s multipliziert, betrachten wir den Vektor $s\mathbf{e}$ als denselben EM) und lässt sich mit Hilfe verschiedener Algorithmen berechnen [81, 126, 243]. Die meisten Programme zur Berechnung von EM (wie z. B. Metatool [260] oder efntool [243]) basieren auf der „double description“ Methode [79]. Basierend auf einem Tableau generiert diese Methode neue Elementarmodus-Kandidaten durch die Anwendung von Gaußschen Kombinationen. Allerdings müssen diese Kandidaten auf Elementarität überprüft werden. Die Berechnung der EM für große Netze ist sehr aufwendig, da sich mit ansteigender Netzwerkgröße die Anzahl der EM aufgrund der kombinatorischen Komplexität dramatisch erhöht [131].

Es gibt mittlerweile einige algorithmische Entwicklungen, die verschiedene Aspekte der Berechnung von Elementarmoden verbessert haben [81, 126, 242]. Trotz dessen kann in genomskaligen Netzwerken häufig nicht die komplette Menge der EM bestimmt werden. Es gibt einige Anwendungen, die versuchen, zusätzliche Informationen zu integrieren, um den Lösungsraum einzuschränken (so z. B. der Gebrauch von transkriptionellen regulatorischen Regeln, um die biologisch relevanten EM zu finden [115]). Andere Ansätze versuchen (1) das genomskalige Netzwerk in Subsysteme zu zerlegen, um darin die EM zu berechnen [222], (2) nur die kürzesten EM zu berechnen [55] oder (3) Metabolomics Daten zu integrieren, um nur thermodynamisch mögliche EM zu berechnen [83]. Kaleta et al. haben eine Methode vorgestellt, um EM zu sampeln [116]. Sie argumentieren, dass es nicht unbedingt notwendig ist, alle EM zu kennen, um das Netzwerk beschreiben zu können.

Durch nicht-negative Linearkombination der Elementarmoden kann der schon genannte Flusskegel (siehe Gleichung (3.10)) generiert werden:

$$F = \{\mathbf{r} \in \mathfrak{R}^q \mid \mathbf{r} = \sum_{h=1}^{|\mathbf{E}|} \varphi_h \mathbf{e}^h\}, \varphi_h \geq 0. \quad (3.13)$$

3.2 Stöchiometrische Modellierung von Stoffwechselnetzen

\mathbf{E} entspricht der Menge der EM ($\mathbf{E} = \{\mathbf{e}^1, \mathbf{e}^2, \dots\}$). Gleichung (3.13) zeigt, dass alle zulässigen stationären Flussverteilungen \mathbf{r} durch die EM erzeugt werden können. Allerdings lassen sich konkrete Flussverteilungen \mathbf{r} i. d. R. nicht eindeutig in EM zerlegen. Die EM enthalten alle Kanten („extreme rays“ = Extremalstrahlen) des Flusskegels (i. A. sind aber nicht alle Elementarmoden Kanten des Flusskegels).

Die Elementarmodenanalyse (EMA) ist ein wichtiges Werkzeug des CBM und ermöglicht vielfältige und detaillierte Studien von Stoffwechselmodellen [251]. Sie lassen z. B. Aussagen über optimale und suboptimale Routen durch das Netzwerk zu (z. B. bzgl. der Ausbeute). Zusätzlich können die EM genutzt werden, um blockierte, gekoppelte und essentielle Reaktionen zu identifizieren. Des Weiteren lässt sich die Auswirkung von Knockouts leicht überprüfen, indem die gewünschten Reaktionen aus dem Modell entfernt werden und die verbleibenden aktiven EM analysiert werden.

Abbildung 3.3 beschreibt die EM des Beispielnetzwerks graphisch. $EM_{2,4,8}$ beschreiben die gekoppelte Produktion von $P1_{\text{extern}}$ und $P2_{\text{extern}}$. EM_1 und EM_3 zeigen, dass es auch Pfade durch das Netzwerk gibt, in denen weder $P1_{\text{extern}}$ noch $P2_{\text{extern}}$ produziert wird. $EM_{6,7}$ stellen die EM dar, die eine optimale Ausbeute von $P1_{\text{extern}}$ bezogen auf Substrat $S1_{\text{extern}}$ ermöglichen. EM_5 repräsentiert den EM, der eine optimale Ausbeute von $P1_{\text{extern}}$ bezogen auf Substrat $S2_{\text{extern}}$ ermöglicht.

3 Mathematische Grundlagen

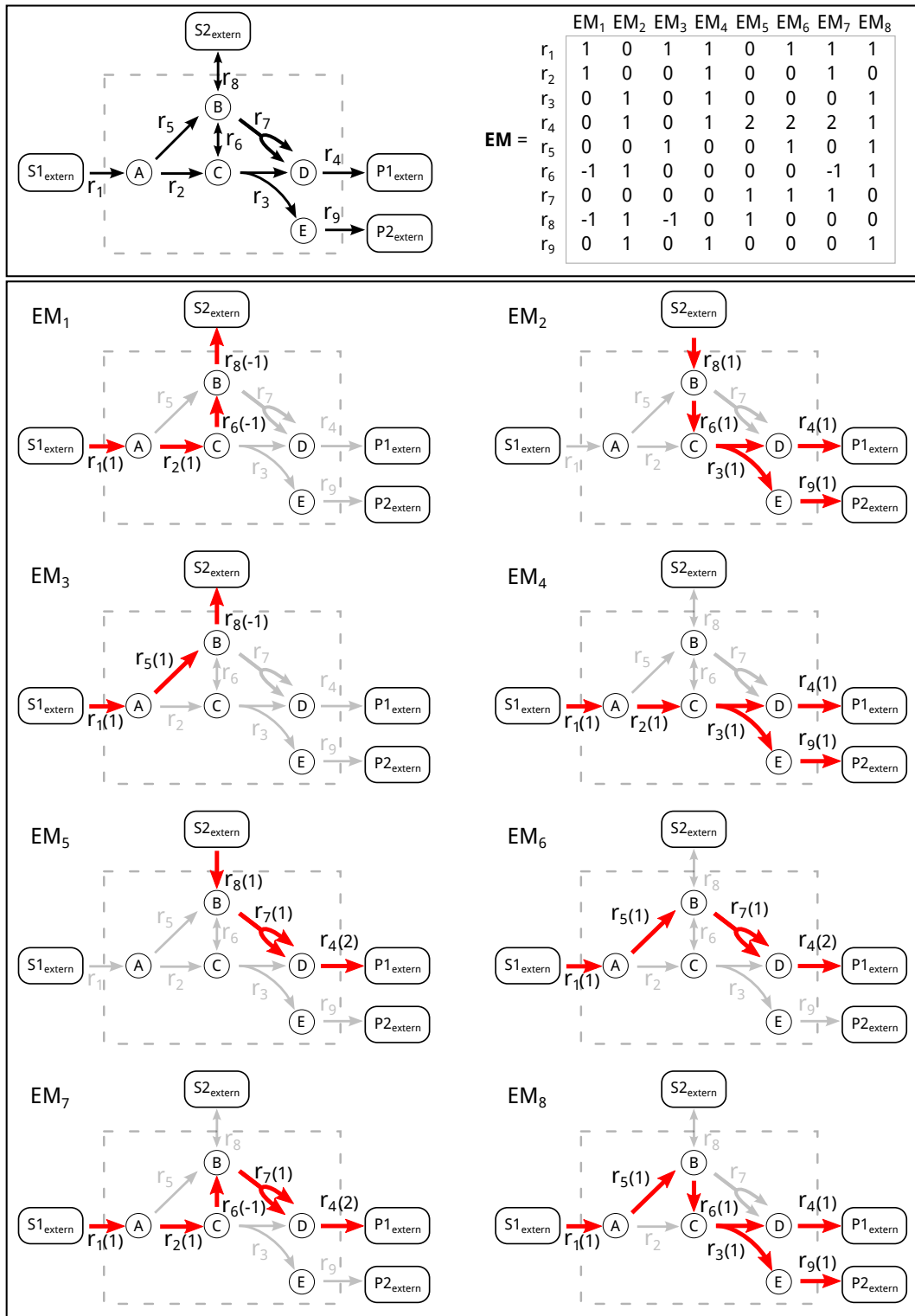


Abbildung 3.3: Die Elementarmoden des dargestellten Beispielnetzwerks. In Klammern hinter den Reaktionsbezeichnungen steht der jeweilige stöchiometrische Koeffizient der Reaktion.

3.3 Methoden zum modellgestützten Metabolic Engineering

Die grundlegende Idee des Metabolic Engineering (ME) ist es, eine Veränderung zellulärer Eigenschaften zu erreichen, um verfügbare, zellinterne Ressourcen beispielsweise zur Synthese gewünschter (neuer) Produkte zu nutzen [238]. Um dieses Ziel zu verwirklichen, werden mikrobielle Stämme genetisch verändert. Die Voraussetzung für das sogenannte „Strain Design“ (SD) ist die genaue Kenntnis des metabolischen Netzwerks, um einen Stamm gezielt zu verändern. Es sind bereits verschiedene Methoden des CBM, wie FBA, FVA und EMA, vorgestellt worden, die es ermöglichen, ein stöchiometrisches Modell auf seine Funktionalität zu untersuchen. Im Folgenden werden Methoden vorgestellt, die sich mit der Identifikation konkreter Interventionsstrategien auseinandersetzen.

3.3.1 Übersicht über Methoden und das Prinzip der Kopplung von Produktsynthese und Wachstum

Prinzipiell können zwei Klassen zur modellbasierten Identifikation von Interventionsstrategien unterschieden werden. Die eine basiert auf der Elementarmodenanalyse und die andere auf mathematischer Optimierung [92, 157, 159, 162, 247].

Elementarmodenbasierte Methoden zum modellgestützten Metabolic Engineering

Ein Beispiel für einen solchen Ansatz ist die Entwicklung eines Organismus mit minimaler Funktionalität, der erstmals von Carlson und Sircienc beschrieben worden ist [42, 43]. Die Idee ist, einen Stamm iterativ durch Deletion von Reaktionen in einen Zustand zu bringen, in dem dieser nur die gewollte Funktionalität besitzt. Dafür sind drei Schritte entscheidend.

1. Die Identifikation aller Pfade bzw. Elementarmoden, die im Netzwerk enthalten sind.
2. Die Selektion der gewünschten Funktionalität.
3. Durch Knockout von suboptimalen Pfaden werden iterativ die Reaktionen gelöscht, deren Auswirkung auf die maximale Ausbeute nicht negativ ist, aber deren Entfernung die Anzahl der EM am meisten senkt. Dieser Schritt wird wiederholt, bis keine weitere Deletion mehr möglich ist (oder die maximale Ausbeute nicht mehr gegeben ist).

Diese Methode dient der Identifikation von Knockout-Strategien, um Zellen mit minimaler metabolischer Funktionalität (MMF) zu entwickeln. Dieser Ansatz ist bereits erfolg-

3 Mathematische Grundlagen

reich eingesetzt worden, z. B. für die Entwicklung des am effizientesten Biomasse produzierenden *E. coli* Bakteriums [249] oder der Entwicklung einer Minimal-*E. coli*-Zelle zur Produktion von Ethanol [250]. Weitere Beispiele für elementarmodenbasierte Methoden zum modellgestützten Metabolic Engineering sind: FluxDesign [167], (constrained) Minimal Cut Sets [94, 127, 125] und CASOP [93]. FluxDesign ist ein heuristischer Ansatz, der basierend auf den EM und einer Korrelationsanalyse Knockout-, Knockdown und Überexpressionskandidaten identifiziert. Diese Methode ist bereits erfolgreich eingesetzt worden [30]. Der (constrained) Minimal Cut Set und der CASOP Ansatz werden später ausführlich diskutiert.

Das Prinzip der Kopplung von Produktsynthese und Wachstum

Viele Methoden des SD verwenden ein gemeinsames Designprinzip, nämlich die Produktsynthese an das Zellwachstum zu koppeln [38, 94, 268]. Das grundlegende Ziel der Kopplung ist es, den Metabolismus eines geeigneten Stamms durch den gezielten KO verschiedener Reaktionen derart zu modifizieren, dass dieser Organismus nur dann wachsen kann, wenn das gewünschte Produkt synthetisiert und ausgeschieden wird [129]. Damit kann garantiert werden, dass, sobald der Organismus wächst, das gewünschte Produkt mit einer Mindestausbeute synthetisiert wird. Für diese Art der Kopplung hat sich der Begriff der wachstumsgekoppelten Produktsynthese (*growth-coupled product synthesis*) etabliert, weshalb dieser in der vorliegenden Arbeit verwendet wird. Es sei darauf hingewiesen, dass es dennoch möglich sein kann, dass das Produkt auch unabhängig vom Wachstum synthetisiert werden kann. Es ist aber auf jeden Fall ausgeschlossen, dass der Organismus wachsen kann, ohne dabei das Produkt zu synthetisieren.

Ein weiterer Vorteil dieser Strategie ist, dass nachfolgend die Produktivität durch die Anwendung von *evolutionary engineering* verbessert werden kann. So kann z. B. eine Änderung der Umweltbedingungen das Wachstumsverhalten beeinflussen. Führt dies nach optimaler, durch Selektion getriebener genetischer Anpassung des Stoffwechsels an die neuen Umweltbedingungen zu einer erhöhten Wachstumsrate, erhöht sich auch die Produktivität, da Produktsynthese und Wachstum gekoppelt sind. Die im Folgenden vorgestellten optimierungsbasierten Methoden versuchen, solche Kopplungsstrategien in einem stöchiometrischen Stoffwechselmodell zu identifizieren.

3.3 Methoden zum modellgestützten Metabolic Engineering

Methoden zum modellgestützten Metabolic Engineering basierend auf mathematischer Optimierung

Es gibt mittlerweile viele verschiedene Methoden, die auf mathematischer Optimierung basieren und Strategien bzgl. Knockouts, Knockdowns, Überexpressionen oder der Integration neuer Pfade identifizieren. Diesen Methoden liegen verschiedene mathematische Formulierungen des jeweiligen Optimierungsproblems zugrunde, so z. B. linear [46], bilevel [38, 47], evolutionär [192] oder kybernetisch [256]. Nachfolgend werden einige Methoden detailliert dargestellt.

Der OptKnock Ansatz von Burgard et al. soll Reaktions-Knockouts identifizieren, so dass das gewünschte Produkt synthetisiert wird, während der Organismus wächst (siehe oben erläuterte Kopplungsstrategie) [38]. Die grundlegende Annahme ist, dass Organismen ihre zellinternen Stoffflüsse bei z. B. umgebungsbedingten oder genetischen Änderungen derart umlenken, dass sie dennoch optimal wachsen können [111]. Dann muss ein Reaktions-Knockout derart identifiziert werden, dass dieser im Einklang steht mit der Zellantwort bzgl. der Verschiebung der zellinternen Flüsse in Richtung des optimalen Wachstums. Um dieses Problem zu lösen, haben Burgard et al. eine bilevel Optimierung implementiert, die zwei verschiedene Zielfunktionen nutzt. Die innere Zielfunktion maximiert i. d. R. die Wachstumsrate (bzw. die Biomasseausbeute). Dabei werden zusätzliche Nebenbedingungen, wie z. B. Substrataufnahmeraten oder ATP Maintenance (Erhaltungsstoffwechsel) berücksichtigt. Die äußere Zielfunktion maximiert die Synthese des gewünschten Produkts durch Deletion von Reaktionen. Die Entfernung von Reaktionen wird unter Nutzung einer binären Variable y in den Algorithmus implementiert. Wie in [38] gezeigt, entsteht folgendes zweistufiges gemischt-ganzzahliges Optimierungsproblem.

$$\begin{array}{ll}
 \underset{y_i}{\text{maximiere}} & r_{\text{Produkt}} \\
 \text{s.t.} & \underset{r_i}{\text{maximiere}} \quad \mu \\
 & \text{s.t.} \quad \mathbf{Nr} = \mathbf{0} \\
 & r_{\text{Aufnahme}} = r_{\text{Aufnahme}}^{\text{Ziel}} \\
 & r_{\text{ATP}} \geq r_{\text{ATP Maintenance}} \\
 & \mu \geq \mu^{\text{Ziel}} \\
 & r_i^{\min} \cdot y_i \leq r_i \leq r_i^{\max} \cdot y_i, \forall i \in M \\
 & y_i = \{0, 1\} \quad \forall i \in M \\
 & \sum_{i \in M} (1 - y_i) \leq n_{KO}
 \end{array}$$

3 Mathematische Grundlagen

M beschreibt die Indexmenge der Reaktionen und n_{KO} die erlaubte maximale Anzahl der Knockouts. Dabei maximiert die interne Zielfunktion die Wachstumsrate μ unter Berücksichtigung von verschiedenen Nebenbedingungen. Es wird dabei berücksichtigt, dass eine bestimmte Menge an Substrat aufgenommen wird, dass es einen basalen Verbrauch von ATP gibt (Maintenance) und dass die Wachstumsrate nicht kleiner ist als die vorgegebene Zielwachstumsrate. Die Nebenbedingung $(r_i^{\min} \cdot y_i \leq r_i \leq r_i^{\max} \cdot y_i, \forall i \in M)$ bewirkt, dass eine Reaktion r_i dann ausgeschaltet ist, wenn deren korrespondierende binäre Variable y_i null ist. Diese Methode ist z. B. für die Laktatproduktion [76], für die Überproduktion von Aminosäuren [193] sowie für die Produktion von 1,4-Butandiol [268] in *E. coli* erfolgreich eingesetzt worden.

Eine Weiterentwicklung von OptKnock ist RobustKnock [241]. Ein Problem des OptKnock Ansatz ist es, dass nur die maximale Produktausbeute bei optimalem Wachstumsverhalten analysiert wird. Dabei wird nicht berücksichtigt, dass evtl. andere optimale Lösungen der inneren Optimierung dafür sorgen, dass Produktsynthese und Wachstum entkoppelt werden könnten. RobustKnock umgeht dieses Problem, indem die minimale Produktausbeute bei optimalem Wachstumsverhalten maximiert wird (dadurch entsteht ein dreistufiges Optimierungsproblem). Diese Methode ist erfolgreich in einem genomskaligen Stoffwechselmodell von *E. coli* (*i*JR904 [206]) eingesetzt worden, um Knockoutstrategien zur Produktion von verschiedenen Chemikalien zu identifizieren. Der Nachteil beider Methoden ist, dass die gefundenen Interventionsstrategien auf der Annahme beruhen, dass der Organismus optimal wächst. Damit kann unter nicht-optimalem Wachstumsverhalten nicht garantiert werden, dass die gefundene Kopplung von Produktsynthese an Wachstum weiterhin gegeben ist.

Weitere Methoden, die Knockouts identifizieren können, sind z. B. OptGene (ein evolutionärer Ansatz, der auf einem genetischen Algorithmus beruht und Mutationskombinationen nutzt, um Interventionen zu finden [192]) und „tilting of the objective function“ (TOOF) (simuliert den „worst-case“ für Produktsynthese mit OptGene und OptKnock [73]).

Andere Verfahren versuchen den Organismus zu verbessern, indem (zusätzlich) nicht-native Reaktionen integriert werden, so z. B. OptStrain [194] und SimOptStrain [124]. OptStrain nutzt eine universelle Datenbank um (bezogen auf ein bestimmtes Substrat) Pfade zu identifizieren, die optimal sind bzgl. der Ausbeute eines gewünschten Produkts oder die Synthese bestimmter Produkte erst zulassen. Anschließend wird das Netzwerk des gegebenen Organismus um eine minimale Anzahl an nicht-nativen Reaktionen erweitert, die eine optimale Ausbeute ermöglichen. In einem zweiten Schritt werden KO-Kandidaten identifiziert, die verhindern, dass Produktsynthese und Wachstum gekoppelt

3.3 Methoden zum modellgestützten Metabolic Engineering

werden können. Im Gegensatz zu OptStrain findet SimOptStrain die beschriebenen Interventionen (Addition nicht-nativer Reaktionen und KO-Strategien) simultan [124].

Zusätzlich zu den bisher beschriebenen Ansätzen gibt es Methoden, die in der Lage sind, weitere Arten von Interventionen, wie Hoch- und Runterregulierung, zu kombinieren, so z. B. OptReg [195], OptORF [123] und OptForce [204]. OptReg identifiziert Reaktionen die hoch-/runterreguliert oder eliminiert werden müssen, damit das gewünschte Produkt überproduziert werden kann. Zur Bewertung, ob eine Reaktion hoch- oder runterreguliert werden muss, wird der Fluss durch eine Reaktion in einer Mutante mit höherer Ausbeute verglichen mit der Wildtyp-Flussverteilung. Ist der Fluss durch eine Reaktion signifikant größer, ergibt sich eine Hochregulierung. Ist dieser Fluss signifikant kleiner als im Wildtyp, muss diese Reaktion runterreguliert werden, um die Ausbeute zu erhöhen. Prinzipiell basiert OptReg auf einer bilevel Optimierung, deren äußere Zielfunktion das gewünschte Produkt maximiert. Die innere Zielfunktion maximiert z. B. Wachstum und minimiert zusätzlich den absoluten Gesamtfluss. Da es drei Möglichkeiten der Intervention gibt, ergeben sich drei Mengen binärer Variablen für jede Reaktion.

OptORF ist in der Lage, Überexpressions- und KO-Kandidaten zu identifizieren, die es ermöglichen, die Synthese des gewünschten Produkts zu maximieren [123]. Dabei werden Gen-Protein-Reaktions-Assoziationen explizit berücksichtigt, sodass die identifizierten Interventionen direkt den betroffenen Genen zugeordnet werden können. Zusätzlich werden transkriptionelle regulatorische Nebenbedingungen berücksichtigt (ein Gen bzw. dessen Expression wird in Abhängigkeit der Aktivität seines Transkriptionsfaktors modelliert). Sowohl die Gen-Protein-Reaktions-Assoziationen, als auch die regulatorischen Nebenbedingungen werden mit Hilfe Boolescher Logik modelliert.

OptForce nutzt eine Referenzflussverteilung, die den maximalen zulässigen Wertebereich (Flussvariabilität) aller metabolischen Reaktionen beschreibt und die, sofern vorhanden, auf experimentellen Daten beruht [204]. Diese Verteilung wird mit einer „Flussvariabilitätsverteilung“ verglichen, die es ermöglicht, das gewünschte Produkt mit einer bestimmten Ausbeute zu synthetisieren. Dieser Vergleich erlaubt einen Rückschluss auf die notwendigen Veränderungen (Hoch-/Runterregulierung, KO), die durchgeführt werden müssen, um das gewünschte Ziel zu erreichen. Nachfolgend werden die minimalen, nicht redundanten Mengen an Interventionen (Hoch-/Runterregulierung, KO) identifiziert, welche die notwendigen Veränderungen erzwingen, um die gewünschte Flussverteilung zu generieren. Laut Ranganathan et al. ist OptForce im Gegensatz zu OptReg nicht auf die Maximierung einer Fitnessfunktion angewiesen, um die metabolischen Flüsse vorherzusagen [204].

3 Mathematische Grundlagen

Weitere Beispiele für Methoden die auf mathematischer Optimierung basieren, sind „flux scanning based on enforced objective flux“ (FSEOF) [46], „genetic design through local search“ (GDLS) [158] und „enhancing metabolism with iterative linear optimization“ (EMiLio) [267].

Die bisher besprochenen Methoden haben i. d. R. das Ziel, die Produktausbeute durch Anwendung verschiedener Interventionsstrategien zu maximieren und die Produktsynthese obligatorisch an das Wachstum koppeln. Das bedeutet, dass sobald der Organismus (optimal) wächst, dieser das gewünschte Produkt synthetisiert. An dieser Stelle wird die Limitierung einiger Methoden wie z. B. OptKnock oder RobustKnock deutlich. Denn sobald der Organismus nicht optimal wächst, kann es passieren, dass dieser gar kein Produkt synthetisiert. Das gilt auch für die (Sim)OptStrain Methoden. Ein weiterer Nachteil dieser Methoden ist es, dass mehrfache Lösungen nicht explizit betrachtet werden.

Eine Methode, die darauf abzielt, alle Interventionsstrategien zu identifizieren, welche die Kopplung von Produktsynthese und Wachstum garantieren kann, ist der Ansatz der constrained Minimal Cut Sets (cMCSs), der im folgenden Kapitel erklärt wird. Tabelle 3.1 gibt einen abschließenden Überblick über die genannten Methoden.

Tabelle 3.1: Methoden des modellgestützten Metabolic Engineering.

Art der Intervention	Methode	Referenz
Knockout	MMF	[42, 43, 249]
	(c)MCS	[94, 125, 127]
	OptKnock	[38]
	RobustKnock	[241]
	OptGene	[192]
	TOOF	[73]
Knockout und Integration nicht-nativer Reaktionen	OptStrain	[194]
	SimOptStrain	[124]
Knockout, Hoch- und Runterregulierung von Flüssen	FluxDesign	[167]
	OptReg	[195]
	OptForce	[204]
	GDLS	[158]
	EMiLio	[267]
	OptORF	[123]
	CASOP	[93]
	FSEOF	[46]

3.3.2 Minimal Cut Sets & constrained Minimal Cut Sets

2004 sind die Minimal Cut Sets (MCSs) das erste mal von Klamt und Gilles zur Analyse von Stoffwechselnetzen beschrieben worden [127]. Diese Methodik identifiziert, basierend auf EM, die kleinste (minimale) Menge (set) an Knockouts (cuts), um eine bestimmte metabolische Funktionalität zu unterdrücken. Der zu unterdrückenden Funktionalität werden alle EM zugeordnet, die dieses Verhalten ermöglichen. Die so selektierten EM werden Target-EM genannt. Im Folgenden soll dies an dem Beispielnetzwerk erklärt werden (Abbildung 3.1 auf Seite 28). Es soll die Produktion von $P2_{\text{extern}}$ unterdrückt werden. Die Target-EM sind demnach $EM_{2,4,8}$ (siehe Abbildung 3.3, Seite 36). Es ergeben sich sieben MCSs, die das unerwünschte Verhalten unterdrücken (Abbildung 3.4). Das

3 Mathematische Grundlagen

bedeutet, dass es sieben (support-minimale) Mengen an Reaktions-KOs gibt, die das Verhalten der Target-EM ausschalten.

Alle gefundenen sieben MCSs sind minimal. Das bedeutet, dass die Entfernung nur eines Cuts (diese Reaktion wäre dann nicht inaktiv) innerhalb eines Cut Sets dazu führt, dass die unerwünschte Funktion wieder hergestellt ist. Durch geschickte Wahl der Target-EM können auch komplexere Ziele definiert werden [125]. Dennoch kann es passieren, dass neben der unerwünschten Funktionalität auch gewünschte Eigenschaften unterdrückt werden. Angenommen die Synthese von $P2_{\text{extern}}$ soll blockiert werden, die Produktion von $P1_{\text{extern}}$ aber dennoch möglich sein. In diesem Fall können nicht alle gefundenen MCSs in Abbildung 3.4 angewendet werden. So unterdrückt MCS_1 beispielsweise die Produktion von $P1_{\text{extern}}$ und $P2_{\text{extern}}$; MCS_6 blockiert die Aufnahme von Substraten. Um Interventionen berechnen zu können, die in der Lage sind, bestimmte erwünschte Netzwerkeigenschaften zu erhalten, ist das Konzept der constrained Minimal Cut Sets (cMCSs) eingeführt worden [94].

Für das gezeigte Beispiel (keine Produktion von $P2_{\text{extern}}$, aber Synthese von $P1_{\text{extern}}$ ist möglich) wird die Menge der EM, welche die Produktion von $P2_{\text{extern}}$ beschreiben, als Target-EM definiert. Die EM, welche die Synthese von $P1_{\text{extern}}$ ermöglichen, werden der Menge der Desired-EM zugeordnet. Dies sind die EM, die das gewünschte (zu erhaltende) Netzwerkverhalten repräsentieren.

3.3 Methoden zum modellgestützten Metabolic Engineering

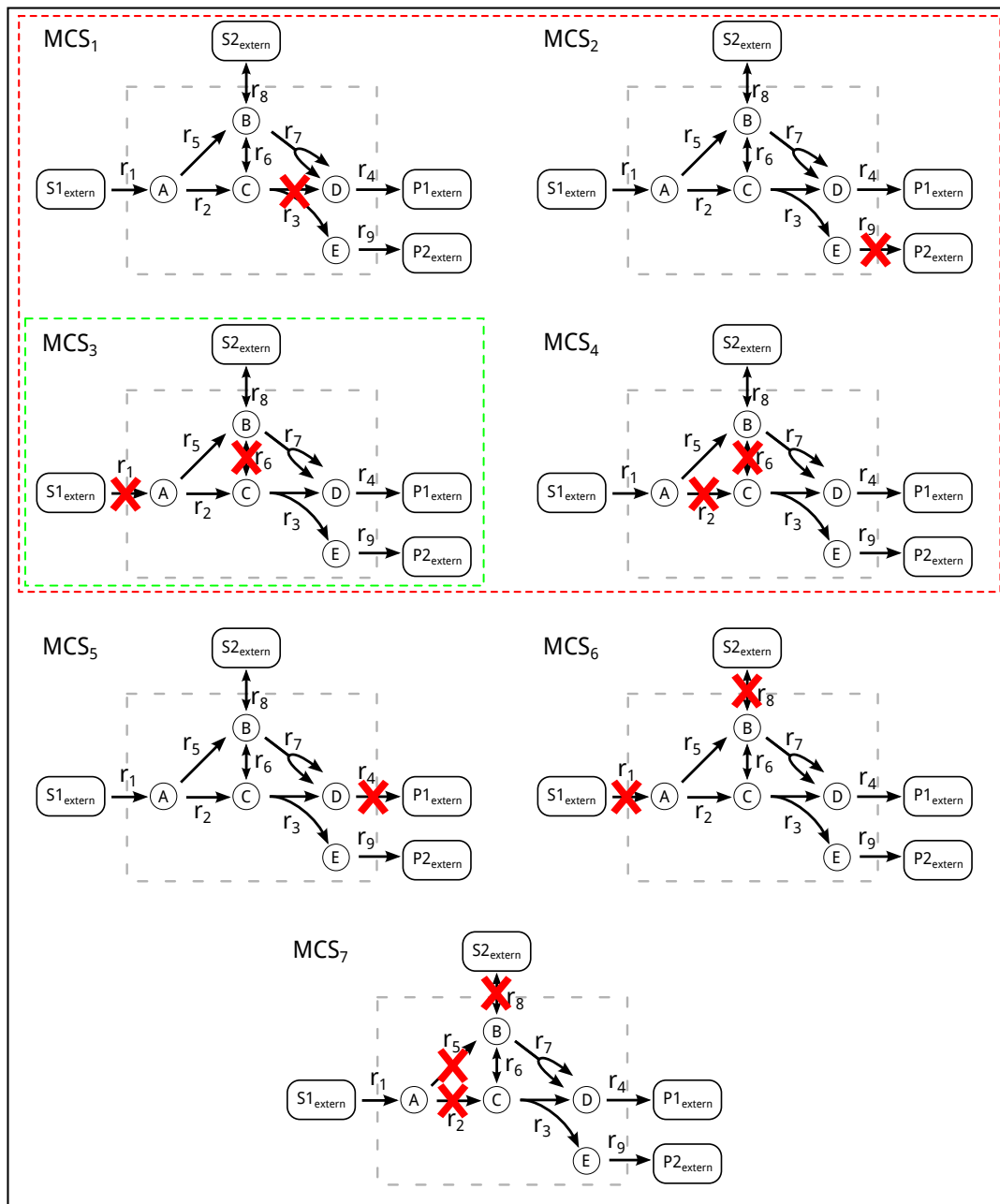


Abbildung 3.4: Darstellung aller Minimal Cut Sets im Beispielnetzwerk (Abbildung 3.1, Seite 28), die eine Produktion von $P2_{\text{extern}}$ verhindern. Die MCSs innerhalb des roten Rahmens ermöglichen weiterhin die Produktion von $P1_{\text{extern}}$ und repräsentieren damit constrained Minimal Cut Sets. Bei Anwendung von MCS₃ (grüner Rahmen) gibt es nur einen verbleibenden EM, bei dem $P1_{\text{extern}}$ zwingend synthetisiert wird.

3 Mathematische Grundlagen

Formal lässt sich der Ansatz der MCSs bzw. cMCSs wie folgt beschreiben. Die Menge aller Elementarmoden entspricht \mathbf{E} (hier: jeder EM \mathbf{e} wird durch seinen Support $\text{supp}(\mathbf{e})$ beschrieben), die Menge der Target-EM wird durch \mathbf{T} und die Menge der Desired-EM durch \mathbf{D} beschrieben. Es gilt: $\mathbf{T} \subset \mathbf{E}$ und $\mathbf{D} \subset \mathbf{E}$. Laut Hädicke und Klamt [94] ist ein Cut Set C als Menge von Reaktionen definiert, sodass in allen Target-EM T jeweils mindestens eine Reaktion aus C enthalten ist. Es gilt:

$$\forall T \in \mathbf{T} : C \cap T \neq \emptyset. \quad (3.14)$$

Die Minimal Cut Sets fordern zusätzlich, dass es keine echte Teilmenge C' von C geben darf, die Gleichung (3.14) erfüllt. Damit wird gewährleistet, dass jeder Cut Set minimal ist. Die Erweiterung der MCSs, die constrained MCSs, fordert zusätzlich, dass mindestens n_{cMCS} EM aus der Menge der Desired-EM \mathbf{D} erhalten bleiben, um die gewünschte Funktionalität gewährleisten zu können. Es gilt:

$$\mathbf{D}^C = \{D \in \mathbf{D} \mid C \cap D = \emptyset\}. \quad (3.15)$$

\mathbf{D}^C beschreibt die Menge der Desired-EM, die nicht von einem Cut Set C betroffen sind. Wie bereits beschrieben, muss ein constrained MCS zusätzlich die folgende Bedingung erfüllen:

$$|\mathbf{D}^C| \geq n_{\text{cMCS}}. \quad (3.16)$$

Wenn nicht anders angegeben, wird im Folgenden immer angenommen, dass $n_{\text{cMCS}} = 1$. Bezogen auf das beschriebene Beispiel bleibt die Menge der Target-EM unverändert ($\mathbf{T} = \{\text{EM}_{2,4,8}\}$). Zusätzlich wird eine Menge Desired-EM definiert: $\mathbf{D} = \{\text{EM}_{5,6,7}\}$. Damit ergeben sich vier cMCSs ($\text{MCS}_1 - \text{MCS}_4$), die in Abbildung 3.4 gezeigt sind und die Synthese von $\text{P2}_{\text{extern}}$ blockieren, aber die Produktion von $\text{P1}_{\text{extern}}$ ermöglichen. Diese Abbildung verdeutlicht, dass die cMCSs grundsätzlich eine Teilmenge der MCSs sind. Nach Anwendung des MCS_3 (grüner Rahmen) bleibt sogar nur eine Möglichkeit (EM) bestehen: der Organismus muss $\text{P1}_{\text{extern}}$ synthetisieren.

Wie im vorherigen Abschnitt gezeigt, versuchen viele Ansätze Zellwachstum und Produktsynthese obligat aneinander zu koppeln. Häufig ist es aber der Fall, dass die gefundene Kopplungsstrategie nur unter wachstumsoptimalen Bedingungen funktioniert. Der Ansatz der cMCSs ermöglicht es, Interventionsstrategien zu identifizieren, die diese Kopplung für alle Biomasseausbeuten erzwingen.

Im folgenden Beispiel wird daher angenommen, dass $\text{P1}_{\text{extern}}$ das Zellwachstum (oder die Synthese eines Biomassebestandteils) repräsentiert und $\text{P2}_{\text{extern}}$ die Synthese eines

3.3 Methoden zum modellgestützten Metabolic Engineering

gewünschten Produkts. Um die Kopplung zu erreichen, wird die Menge der Desired-EM definiert durch: $\mathbf{D} = \{\text{EM}_{2,4,8}\}$ und die Menge der Targetmoden durch $\mathbf{T} = \{\text{EM}_{1,3,5,6,7}\}$. Es ergeben sich fünf mögliche Szenarien (cMCSs) mit denen *garantiert* werden kann, dass während des Wachstums das gewünschte Produkt synthetisiert wird. Die dafür notwendigen Interventionen sind in der Abbildung 3.5 dargestellt.

Wie gezeigt, eröffnet der Ansatz der cMCSs die Möglichkeit, Strategien zu identifizieren, die obligat Produktsynthese und Wachstum koppeln können (unabhängig von der Annahme der Wachstumsoptimalität). Das stellt einen enormen Vorteil zu den Methoden dar, die auf der Annahme beruhen, dass sich eine Zelle wachstumsoptimal verhalten muss, kann aber auch mehr Reaktionsknockout erfordern.

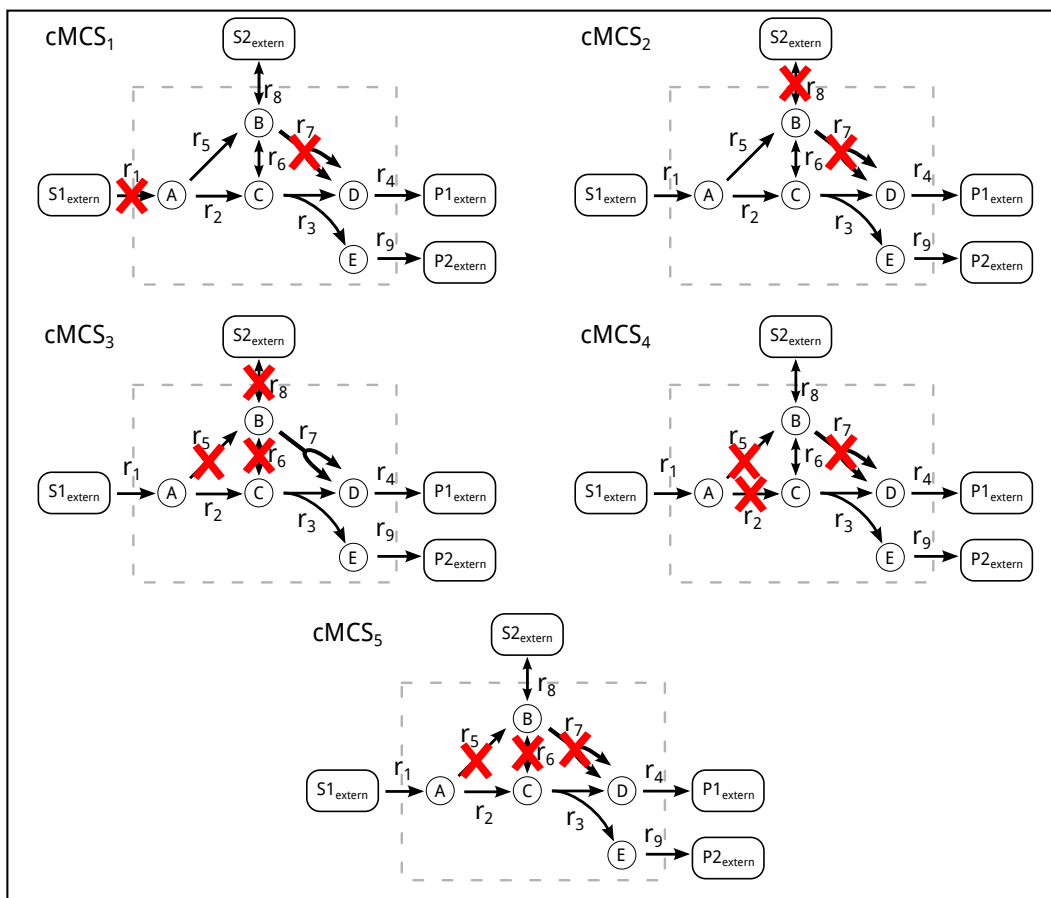


Abbildung 3.5: Constrained Minimal Cut Sets des Beispielnetzwerkes (Abbildung 3.1, Seite 28) für das Szenario, dass ein Biomassebestandteil (hier: $P1_{\text{extern}}$) und Produkt ($P2_{\text{extern}}$) gekoppelt synthetisiert werden.

3 Mathematische Grundlagen

Im Allgemeinen basiert die Berechnung der (c)MCSs auf der Menge der Elementarmoden. Allerdings kann diese gerade in großen stöchiometrischen Netzwerken selten komplett berechnet werden. Zur Lösung dieses Problems haben von Kamp und Klamt 2014 einen Ansatz präsentiert, der in der Lage ist, (c)MCSs bis zu einer bestimmten Größe in genomskaligen Netzwerken zu berechnen [259]. Dieser Ansatz basiert auf dem Dualitätsprinzip von EM und MCSs. Dieses besagt, dass die MCSs des metabolischen (oder primalen) Netzwerks, die EM eines dualen Netzwerks sind (und umgekehrt) [26].

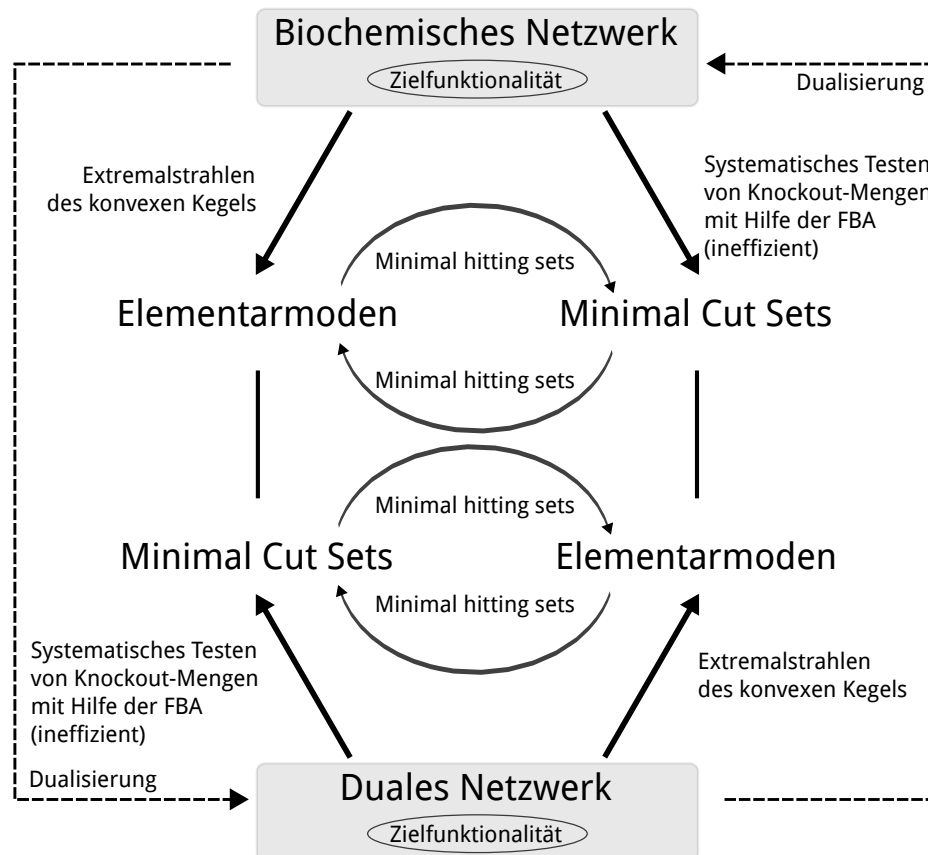


Abbildung 3.6: Zusammenhang zwischen den Elementarmoden und Minimal Cut Sets eines biochemischen und eines dualen Netzwerks. Der neben den Pfeilen stehende Text beschreibt den Ansatz, wie die EM (MCSs) aus dem jeweiligen Netzwerk berechnet werden können. Die MCSs können prinzipiell mit dem Ansatz der „minimal hitting sets“ aus den EM berechnet werden (und umgekehrt; allerdings müssen die jeweiligen Mengen separat berechnet werden [125]). Die EM des biochemischen Netzwerks entsprechen den MCSs des dualen Netzwerks (und umgekehrt). Graphik angelehnt an [26].

3.3 Methoden zum modellgestützten Metabolic Engineering

Dieses Prinzip ist von von Kamp und Klamt derart genutzt worden, dass die kürzesten EM in dem dualen Netzwerk berechnet worden sind, die den kleinsten MCSs im primalen (metabolischen) Netzwerk entsprechen. Es gelang damit z. B. alle MCSs bis zur Kardinalität fünf zu berechnen, die in dem genutzten (genomskaligen) Modell das Wachstum verhindern. Des Weiteren konnten alle cMCSs bis zur Kardinalität sieben berechnet werden, welche die wachstumsgekoppelte Ethanolsynthese erzwingen. Die Abbildung 3.6 stellt den Zusammenhang graphisch dar. Abschließend zeigt Abbildung 3.7 die typischen Ausbeutediagramme nach Anwendung einer OptKnock-, RobustKnock-, OptStrain- und cMCS-Strategie. Es wird deutlich, dass nur die cMCS-Strategie eine wachstumsgekoppelte Produktsynthese garantieren kann.

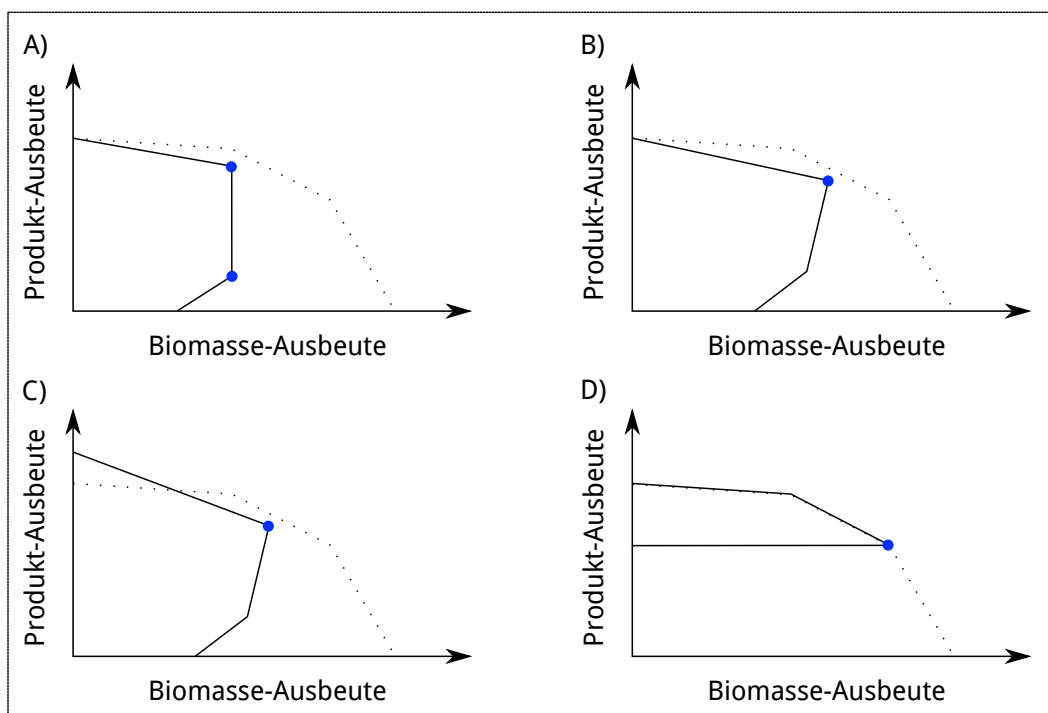


Abbildung 3.7: Mögliche resultierende Ausbeutediagramme nach Anwendung von Interventionsstrategien, die durch die Nutzung verschiedener Methoden (A - D) identifiziert worden sind. Die blauen Punkte deuten die biomasseausbeuteoptimalen Lösungen der Mutanten an. Die gepunktete Linie zeigt die Ausbeutediagramme vor der Intervention. A) OptKnock-Ansatz. B) RobustKnock-Ansatz. C) OptStrain-Ansatz. Die maximale Produktausbeute ist höher als im Wildtyp. Dies ist auf eine Reaktionsaddition zurückzuführen. D) cMCS-Ansatz. In diesem Fall kann garantiert werden, dass Produkt immer synthetisiert wird, egal mit welcher Biomasseausbeute der Organismus wächst.

3.3.3 CASOP

Die bisher beschriebenen Methoden sind daran interessiert, die Ausbeute eines bestimmten Produkts zu optimieren. Der Ansatz von CASOP („Computational Approach for Strain Optimization aiming at high Productivity“) versucht hingegen, die Produktivität als kombiniertes Maß für Ausbeute und Netzwerkkapazität zu optimieren [93].

Der Ansatz beruht ebenfalls auf den EM. Im Prinzip werden die EM für den Wildtyp berechnet (nur Wachstum) und die EM für einen Phänotyp, der den Großteil des aufgenommenen Substrates während des Wachstums in ein gewünschtes Produkt umwandelt. Diese beiden Mengen an EM werden verglichen und bewertet. Diese Bewertung erlaubt einen Rückschluss darauf, ob eine Reaktion überexprimiert, herunterreguliert oder ausgeschaltet werden soll, um den gewünschten Phänotyp zu erhalten. Notwendig für diese Bewertung (bzw. Berechnung der EM des gewünschten Phänotyps), ist die Einführung einer artifiziellen Reaktion r_A , die einen artifiziellen Metaboliten A generiert:

$$r_A : (1 - \gamma) \cdot \text{Biomasse} + \gamma \cdot \alpha_{\text{Produkt}} \cdot \text{Produkt} \implies 1 \text{ [Gramm] A}, \quad \gamma \in [0, 1]. \quad (3.17)$$

Sowohl die Biomasse als auch das gewünschte Produkt werden dafür als interne Metabolite definiert. Sofern $\gamma = 0$, wird im Modell ausschließlich Biomasse produziert (Wildtyp); für $\gamma = 1$ nur Produkt. γ steuert demnach, in welchem Verhältnis Produkt und Biomasse im Modell synthetisiert werden. Um als Einheit für den Metaboliten A Gramm zu erhalten, muss das Produkt mit seiner reziproken molaren Masse (α_{Produkt}) multipliziert werden.

Die EM der beiden oben genannten EM-Mengen (EM für den Wildtyp (nur Wachstum); EM für einen Phänotyp, der den Großteil des aufgenommenen Substrates während des Wachstums in ein gewünschtes Produkt umwandelt) werden, in Abhängigkeit der Ausbeute des artifiziellen Metaboliten A bezogen auf ein Substrat S, gewichtet. Dieses Gewicht $\nu_{p,\gamma}$ für den EM_p^γ ($p = 1, \dots, n_C(\gamma)$) wird laut Hädicke und Klamt definiert als [93]:

$$\nu_{p,\gamma} = \frac{\left(Y_p^{A/S}\right)^{k_C}}{\sum_{p=1}^{n_C(\gamma)} \left(Y_p^{A/S}\right)^{k_C}}. \quad (3.18)$$

$n_C(\gamma)$ beschreibt die Anzahl der EM mit einer positiven Ausbeute $Y^{A/S}$ für ein bestimmtes γ . Die Gewichtung in 3.18 ermöglicht, dass ein EM mit einer höheren Ausbeute ein höheres Gewicht bekommt. Der Parameter (Exponent) k_C ermöglicht eine zusätzliche Gewichtung. Ein kleines k_C erlaubt eine Bewertung mit höherem Einfluss der Netzwerk-

flexibilität ($k_C = 0$: alle EM sind gleich gewichtet sind). Große k_C hingegen verschieben die Bewertung hin zur Ausbeute der EM. Für die Bewertung der Relevanz einer Reaktion r_i in einem definierten Szenario γ wird das Maß $\omega_\gamma(r_i)$ eingeführt. ω_γ der Reaktion r_i ist definiert als die Summe aller Gewichte ($\nu_{p,\gamma}$) der EM, in denen die Reaktion $r_i \neq 0$:

$$\omega_\gamma(r_i) = \sum_{h:EM_h^\gamma(r_i) \neq 0} \nu_{h,\gamma}. \quad (3.19)$$

Der Index h beschreibt die EM, in denen die Reaktionsrate $r_i \neq 0$. Das Maß $\omega_\gamma(r_i)$ erlaubt nun eine Bewertung, wie wichtig eine Reaktion r_i in einem definierten Szenario ist. Der Wertebereich von $\omega_\gamma(r_i)$ liegt in $[0, 1]$. Hädicke und Klamt nutzen nun eine Referenzlösung (Wildtyp) und vergleichen diese mit dem „Produktionsszenario“ ($\gamma = 0, 9$: 90% des Metaboliten A besteht aus Produkt):

$$Z(r_i) = \omega_{0,9}(r_i) - \omega_0(r_i). \quad (3.20)$$

Das Bewertungsmaß Z kann für jede Reaktion r_i berechnet werden, die an dem Szenario γ beteiligt ist. Der Wertebereich von $Z(r_i)$ liegt in $[-1, 1]$. Werte kleiner null zeigen einen potentiellen KO-Kandidaten an und Werte größer als null einen potentiellen Überexpressionskandidaten (je (absolut) größer der Wert ist, desto eher ist dieser Kandidat für den entsprechenden Eingriff geeignet).

3.4 Verwendete Software

Als Werkzeug für die Berechnungen und Analysen wurde die MATLAB-Toolbox *CellNetAnalyzer* genutzt [130] (in MATLAB 2010). Diese Toolbox ermöglicht die strukturelle und funktionelle Analyse stöchiometrischer Modelle mit Hilfe vieler verschiedener CBM-basierter Methoden, wie z. B. FBA, FVA, EMA, (c)MCS-Analyse und CASOP-Analyse. Der in Kapitel 4 beschriebene Algorithmus wurde in diese Toolbox integriert.

3.5 Stöchiometrische Stoffwechselmodelle von Cyanobakterien

Im folgenden Abschnitt werden Arbeiten vorgestellt, die sich mit der stöchiometrischen Modellierung von Stoffwechselnetzen und dem rationalen Stamm-Design von Cyanobakterien beschäftigen. Der Fokus liegt dabei auf den hier betrachteten Spezies *Synechocystis* sp. PCC 6803, *Synechococcus elongatus* PCC 7942 und *Cyanothece* sp. ATCC 51142. Die folgende Tabelle stellt eine Zusammenfassung bisheriger Netzwerkrekonstruktionen für den Stoffwechsel von *Synechocystis* sp. PCC 6803 dar, der wohl wichtigsten cyanobakteriellen Modellspezies.

Tabelle 3.2: Publierte *Synechocystis* sp. PCC 6803 Modelle.

Entwickler	Erscheinungsjahr	Anzahl Reaktionen	Anzahl Metabolite	Referenz
Shastri und Morgan	2005	70	46	[225]
Hong und Lee	2007	56	63	[108]
Fu	2009	831	704	[78]
Knoop et al.	2010	380	291	[136]
Montagud et al.	2010	882	790	[170]
Yoshikawa et al.	2010	493	465	[269]
Montagud et al.	2011	956	911	[171]
Nogales et al.	2012	863	795	[178]
Saha et al.	2012	1156	996	[211]
Knoop et al.	2013	759	601	[134]

Das in der vorliegenden Arbeit verwendete (und zu diesem Zeitpunkt aktuellste) Modell von *Synechocystis* sp. PCC 6803 (im Folgenden *Synechocystis* genannt), ist von Knoop et al. entwickelt worden [134]. Dieses Modell basiert auf der Rekonstruktion aus dem Jahr 2010 [136] und ist um den Fettsäurestoffwechsel, die Biosynthese von Peptidoglycan, Chlorophyll, Terpenoiden, Carotenoiden, Quinonen, Tocopherolen und diversen Vitaminen und einigen durch Stress induzierten Reaktionswegen erweitert worden [134]. Das Netzwerk besteht aus 759 Stoffwechselreaktionen und 601 Metaboliten, die in sechs Kompartimenten verteilt sind (Cytosol, Thylakoid- und Plasmamembran, Thylakoidlumen, periplasmatischer Raum, Carboxysomen). Die Biomassesynthesereaktion stammt

3.5 Stöchiometrische Stoffwechselmodelle von Cyanobakterien

aus der Rekonstruktion von Nogales et al. [178] und ist modifiziert und angepasst worden. Das Modell ist in der Lage, sowohl phototrophes Wachstum als auch Wachstum unter Lichtausschluss zu beschreiben (Nachtstoffwechsel). Des Weiteren erfasst es u. a. die Bildung reaktiver Sauerstoffspezies. In der dazugehörigen Veröffentlichung werden die möglichen Szenarien der Schließung des TCA-Zyklus diskutiert. Dabei werden der Nebenweg von Zhang und Bryant [273], sowie der GABA-Shunt als wahrscheinlichste Reaktionswege beschrieben, da diese die niedrigsten Investitionskosten haben und damit evolutionär von Vorteil sind [134]. Experimentelle Ergebnisse legen dagegen nahe, dass der Glyoxylatzyklus nicht vorhanden ist (dieser ist daher nicht im Modell enthalten).

In der vorliegenden Arbeit wird das Modell ohne Berücksichtigung von Isoenzymen genutzt, da der Fokus der Analysen auf den Reaktionen liegt. Zur detaillierteren Beschreibung der zyklischen Elektronenflüsse um die Photosysteme ist eine Ferredoxin-Plastoquinon-Reduktase und die Mehler-Reaktion in das Modell integriert worden (in Anlehnung an Nogales et al. [178]). Um die Biokraftstoffproduktion zu ermöglichen, sind die entsprechenden Bildungs- und Exportreaktionen (siehe Tabelle A.1) in das Modell integriert worden. Der Nachtmetabolismus wird durch eine Glycogenaufnahmereaktion abgebildet. Basierend auf der Annahme, dass Wasser und cytoplasmatische Protonen im Überfluss vorhanden sind, werden diese als externe Metabolite betrachtet. Die für die ATP-Bildung (durch die ATPase) notwendigen periplasmatischen und im Thylakoidlumen befindliche Protonen werden aber explizit bilanziert. Das genomskalige Modell hat mit diesen Anpassungen eine Größe von 612 Reaktionen und 526 internen Metaboliten.

Neben den oben genannten Netzwerkrekonstruktionen für *Synechocystis* gibt es auch einige weitere für z. B. *Synechococcus* sp. PCC 7002 [95, 233, 261], *Cyanothece* sp. ATCC 51142 [211, 262] und *Synechococcus elongatus* PCC 7942 [133, 248]. In der vorliegenden Arbeit werden zwei weitere Vertreter des cyanobakteriellen Metabolismus genauer analysiert. Es handelt sich dabei um *Synechococcus elongatus* PCC 7942 (Modell von Knoop [133]) und *Cyanothece* sp. ATCC 51142 (Modell von Saha et al. [211]).

Das genomskalige Modell von *Synechococcus elongatus* besteht aus 666 Reaktionen und 574 Metaboliten. Es ist im Wesentlichen aus dem *Synechocystis* Modell von Knoop et al. abgeleitet worden und beschreibt phototrophes Wachstum und den Abbau des zellinternen Speicherstoffes Glycogen unter Lichtausschluss. Die wesentlichen Unterschiede innerhalb des Stoffwechsels sind laut Knoop gegeben durch: keine Synthese von Cyanophycin und PHB (zellinterne Speicherstoffe); verschiedene Tocopheroltypen können nicht synthetisiert werden; der Harnstoffmetabolismus fehlt (veränderte Synthese von Putrescin und Spermidin); Glycerat kann nicht über Glycerol zu Glycerol-1-Phosphat

3 Mathematische Grundlagen

abgebaut werden; ein zyklischer Fluss durch den TCA ist nicht möglich, da weder die Gene des GABA-Shunts noch der 2-Oxoglutarat-Decarboxylase vorhanden sind (unvollständiger TCA) [133]. Insgesamt weist das Modell eine geringere Variabilität aus. Unter phototrophen Bedingungen wirkt sich dies allerdings kaum auf die Wachstumsrate aus (diese ist geringfügig kleiner als bei *Synechocystis* [133]). Unter Lichtausschluss wächst der Organismus langsamer [133].

Als zusätzlicher Modellorganismus ist *Cyanothece* sp. ATCC 51142 (im Folgenden *Cyanothece* genannt) ausgewählt worden. Das Genom von *Cyanothece* ist ca. 35% größer als das von *Synechocystis* [211]. Das Modell *iCyt773* von Saha et al. [211] beruht auf den Informationen einer vorherigen Rekonstruktion von Vu et al. [262] und ist im wesentlichen durch die Integration der Lipid-, Pigment- und Pentadecansynthese erweitert worden [211]. Das genomskalige Modell von *Cyanothece* besteht aus 946 Reaktionen und 811 Metaboliten, die in vier Kompartimenten verteilt sind (Periplasma, Cytosol, Carboxysomen, Thylakoidlumen). Die Biomassekomposition von *Cyanothece* unterscheidet sich im Bezug auf das Vorhandensein von Stickstoff und Licht [211] (z. B. ist der Bedarf an Pigmenten wie Chlorophyll unter Lichtbedingungen größer). Laut Saha et al. [211] ist der Stoffwechsel verglichen mit *Synechocystis* durch folgende Unterschiede gekennzeichnet: das Fehlen der Threonin-Ammonium-Lyase (veränderte Synthese von Isoleucin); einen unvollständigen Urea-Zyklus; das Vorhandensein einer Laktat Dehydrogenase (fermentative Produktion von Laktat ist möglich); das Fehlen einer Protoporphyrin Cyclase (notwendig für anaerobe Synthese von Chlorophyll); Unterschiede im Fettsäurestoffwechsel [211].

Bisherige Studien zum Metabolic Engineering von Cyanobakterien

Die Produktpalette der Cyanobakterien ist mittlerweile sehr groß und deren potentielle Anwendungsbreite enorm (siehe Kapitel 2.2). Aufgrund des gesteigerten Interesses, Cyanobakterien als Produktionsplattform zu nutzen, sind viele cyanobakterielle Wildtyp-Stämme zu Produktionsstämmen modifiziert worden.

Angermayr et al. beschreiben am Beispiel von Ethanol diese Entwicklung wie folgt [18]. Initial konnte Ethanol in *Synechocystis* mit einem Titer von 0,021 mg/L synthetisiert werden und nachdem die Stammkonstruktion und die Kultivierungsbedingungen optimiert worden sind, war dieser in der Lage 596 mg/L zu produzieren [57]. Durch weitere genetische Modifikationen und Nutzung eines optimalen Kofaktors (NADPH) konnte der Titer weiter gesteigert werden auf 5500 mg/L [82]. Tabelle 3.3 gibt eine Übersicht über verschiedene erfolgreiche Strategien der Stammoptimierung.

Tabelle 3.3: Erfolgreiche Strategien zur Stammkonstruktion und -optimierung für Produktsynthese in Cyanobakterien. Für mehr Beispiele siehe [18, 140, 276].

Produktionsstamm	Produkt	Strategien
<i>Synechocystis</i> sp. PCC 6803	Ethanol	Integration von Genen aus <i>Z. mobilis</i> und KO der PHB-Synthese [82]
	Aceton	KO der PHB- und Acetat-Synthese (PHB-Synthase und Phosphotransacetylase) [275]
	Ethylen	Integration eines Gens aus <i>Pseudomonas syringae</i> pv. <i>Phaseolicola</i> [252]
	D-Lactat	Integration der D-Lactat Dehydrogenase und KO der PHB- und Acetat-Synthese [274]
<i>Synechococcus</i> <i>elongatus</i> PCC 7942	Isopropanol	Integration von Genen aus <i>Clostridium acetobutylicum</i> und <i>E. coli</i> [139]
	Isobutanol	KO der Glycogen Synthese [151]
	1-Butanol	Integration von Genen aus <i>C. acetobutylicum</i> , <i>Treponema denticola</i> und <i>E. coli</i> [141]
<i>Synechococcus</i> sp. PCC 7002	D-Mannitol	Integration von Genen aus <i>E. coli</i> und <i>Eimeria tenella</i> , KO der Glycogen Synthese [113]

Ebenso gibt es Strategien, die stärker auf der Analyse eines zu Grunde liegenden metabolischen Netzwerks beruhen. Knoop et al. haben 2012 das stöchiometrische Potential von Cyanobakterien (speziell *Synechocystis*) für kohlenwasserstoffbasierte Treibstoffe analysiert, um Aussagen darüber zu treffen, welche Eigenschaften ein Produktionsstamm haben muss, um wirtschaftlich zu sein [118]. In einer späteren Veröffentlichung haben Knoop et al. beschrieben, wie sich die Flussverteilung innerhalb des Modells von *Synechocystis* verschiebt, wenn ein Wildtypstamm (nur Wachstum) zu einem Produktionsstamm verändert wird [135].

Nogales et al. [178] haben 2012 den Photosyntheseapparat von *Synechocystis* modelltheoretisch untersucht. Bei der Modellierung des Stoffwechselnetzes ist die Photosynthese sehr detailliert beschrieben worden, um deren optimale Konfiguration für verschiedene Stoffwechselszenarien zu identifizieren. In einer weiterführenden Veröffentlichung aus dem Jahr 2013 [179] wird demonstriert, wie sich diese metabolischen Eigenschaften auf die biotechnologischen Kapazitäten auswirken. Aus den Analysen wird überraschend geschlussfolgert, dass der lichtgetriebene Metabolismus die biotechnologischen Möglich-

3 Mathematische Grundlagen

keiten von *Synechocystis* limitiert. Nogales et al. [179] haben versucht, Strategien zu finden, die es erlauben, Wachstum und Produktion zu koppeln. Sie konnten in vier von sechs Fällen keine Kopplungsstrategien unter phototrophen Bedingungen finden (für die Produktion von Ethanol, 1-Butanol, Sucrose, und L-Lactat) und schreiben: „[...] our search did not reveal any growth-coupled mutants overproducing these metabolites.“ Und weiter: „[...] the results [...] strongly suggest that a re-routing of the carbon flux is more difficult to achieve in photosynthetic organisms than in heterotrophs, such as *E. coli*.“

Vu et al. [261] haben 2013 in einem Modell von *Synechococcus* sp. PCC 7002 verschiedene Szenarien angedeutet, mit welchen die Kopplung unter phototrophen Bedingungen erreicht werden könnte. Sie sagen, dass wenigstens neun Reaktionsknockouts notwendig sind, um das gewünschte Produkt (u. a. Ethylen und Ethanol) an Wachstum zu koppeln. Allerdings wird in dieser Veröffentlichung weder die für die Kopplung notwendige Strategie erklärt, noch gesagt, welche Reaktionsknockouts dafür notwendig sind.

Trotz der gezeigten Ansätze konnten bisher kaum konkrete Prinzipien für phototrophe Organismen identifiziert werden, mit denen wachstumsgekoppelte Produktionsstämme entwickelt werden können. Ein wesentlicher Aspekt der vorliegenden Arbeit ist es, Strategien zu finden, die es ermöglichen, die Biokraftstoffproduktion an das Wachstum verschiedener Cyanobakterien zu koppeln. Im Verlauf der Arbeit soll daher untersucht werden, ob es, ähnlich zu heterotrophen Organismen wie *E. coli* [259], Interventionsstrategien gibt, die während des Zellwachstums die Synthese eines gewünschten Produkts erzwingen (Kopplung). Basierend auf unterschiedlichen Netzwerkskalen sollen die dafür notwendigen KO-Szenarien analysiert und diskutiert werden.

Neben dem in Kapitel 3.5 vorgestellten genomskaligen Stoffwechselmodell von *Synechocystis*, sollen zwei weitere *Synechocystis*-Modelle verschiedener Skalen betrachtet werden: ein Modell mittlerer Größe und ein stark reduziertes Zentralstoffwechselmodell. Aufbauend auf diesem Minimalmodell sollen skalenübergreifend KO-Strategien gefunden, ausgewertet und verglichen werden, die eine wachstumsgekoppelte Produktion von Biokraftstoffen unter phototrophen Bedingungen ermöglichen. Da zu Beginn dieser Arbeit nur die genomskalige Version des genannten *Synechocystis*-Modells zur Verfügung stand, wurde ein Reduktionsalgorithmus entwickelt, der stöchiometrische Stoffwechselmodelle eigenschaftserhaltend auf Modelle kleinerer Netzwerkskalen reduzieren kann. Dieser Algorithmus wird im folgenden Kapitel vorgestellt.

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen metabolischer Netzwerke

Dieses Kapitel beruht auf der folgenden Publikation:

[67] Erdrich, P., Steuer, R. & Klamt, S. (2015) An Algorithm for the Reduction of Genome-Scale Metabolic Network Models to Meaningful Core Models. *BMC Systems Biology* **9**:48

Im folgenden Kapitel wird der Algorithmus zur eigenschaftserhaltenden Reduktion stöchiometrischer Stoffwechselmodelle eingeführt. Nach einer einleitenden Motivation (4.1), wird der Reduktionsalgorithmus formal vorgestellt (4.2). Kapitel 4.3 zeigt eine Machbarkeitsstudie (Proof-of-Concept) am Beispiel eines genomskaligen *E. coli* Modells. Anschließend wird in Kapitel 4.4 die Reduktion des in Kapitel 3.5 dargestellten genomskaligen *Synechocystis*-Modells, *SyneGS*, auf ein Modell mittlerer Größe (*SynePruned*) und auf ein sehr kleines Zentralstoffwechselmodell (*SyneMin*) beschrieben. Diese drei Modelle repräsentieren damit drei verschiedene Netzwerkskalen.

4.1 Motivation

Die bisher vorgestellten Methoden der stöchiometrischen Netzwerkanalyse liefern tiefe Einblicke in Eigenschaften und Funktionalitäten des zellulären Metabolismus [128, 150, 164]. Der Erfolg dieser Methoden ist nicht zuletzt an der Entwicklung von großen stöchiometrischen Modellen zu erkennen. Mitte bis Ende der 90er Jahre bestanden die damals größten Modelle aus bis zu 100 Reaktionen, die meist im Zentralstoffwechsel lokalisiert sind. Mittlerweile haben genomskalige Modelle eine Größe von tw. über 2000 Reaktionen und es existieren weit über 100 genomskalige Modelle für verschiedene Organismen [7].

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

Am Beispiel von *E. coli* kann die Evolution von stöchiometrischen Modellen des Stoffwechsels gut beschrieben werden. Das erste genomskalige Modell von *E. coli* (*iJE660*) ist im Jahr 2000 veröffentlicht worden [63]. Dieses Modell besteht aus 627 Reaktionen und 438 Metaboliten. Das darauffolgende Netzwerk, *iJR904*, das von Reed et. al 2003 vorgestellt worden ist, hat eine Netzwerkgröße von 931 Reaktionen und 625 Metaboliten [206]. 2007 folgte mit *iAF1260* von Feist et al. [71] die nächste Version eines genomskaligen *E. coli* Netzwerks. Das *iAF1260* ist eines der meist zitierten genomskaligen Netzwerkmodelle und besteht aus 2077 Reaktionen und 1039 Metaboliten. Die aktuellste genomskalige Rekonstruktion des Metabolismus von *E. coli*, das *iJO1366*, ist im Jahr 2011 von Orth et. al veröffentlicht worden und enthält 2251 Reaktionen und 1136 Metabolite [184]. Alle dargestellten Modelle haben die Grundlage für die jeweils nächste Version geschaffen und es ist davon auszugehen, dass diese Entwicklung weiter fortschreiten wird. Durch die Berücksichtigung zusätzlicher Informationen, wie z. B. der Proteinexpression (ME-Modelle [187]), nimmt die Komplexität solcher Modelle weiter zu.

Allerdings steigt mit der Zunahme der Komplexität und Größe solcher Modelle auch der Aufwand, diese Systeme zu analysieren. Einige Methoden können zum Teil nicht mehr angewendet werden. So ist z. B. die Berechnung der EM in Netzwerken mit mehreren tausend Reaktionen und die Nutzung der darauf aufbauenden Methoden nicht möglich [251]. Die metabolische Flussanalyse kann i. d. R. nicht genutzt werden, da intrazelluläre Flüsse im Gesamtsystem oft nicht bestimmbar sind [201], selbst wenn Daten aus Isotop-Tracer Experimenten vorliegen [271]. Ähnliches gilt für kinetische Modellierungsansätze, bei denen kinetische Parameter und Mechanismen oft nur für kleine Subnetzwerke bekannt sind.

Im Gegensatz zu genomskaligen Modellen können in kleineren Modellen (beispielsweise des Zentralstoffwechsels) häufig die grundlegenden Prinzipien und Funktionen eines Systems besser analysiert und ausgewertet werden. Die Informationen eines komplexen genomskaligen Modells sollten dabei, sofern es möglich ist, in derartigen Zentralstoffwechselmodellen abgebildet werden. Bisher gibt es keinen formalen Algorithmus, mit dem genomskalige Netzwerkmodelle reduziert werden können. Es gibt allerdings einige Beispiele für reduzierte bzw. kleinere Modelle, die manuell aus genomskaligen Modellen abgeleitet worden sind [66, 185, 201]. Ein Ansatz für eine formale Reduktion stöchiometrischer Stoffwechselmodelle wird im folgenden Abschnitt vorgestellt.

4.2 Algorithmus *NetworkReducer*

Wie bereits im Abschnitt 3.2 beschrieben, wird ein stöchiometrisches Modell durch m interne Metabolite und q Reaktionen beschrieben. Damit ergibt sich eine stöchiometrische Matrix \mathbf{N} der Dimension $m \times q$. Die Annahme des Fließgleichgewichts führt zur zentralen Gleichung des CBM (siehe Gleichung (3.5)):

$$\mathbf{N}\mathbf{r} = \mathbf{0}. \quad (4.1)$$

Der Freiheitsgrad (FG) eines solchen Systems entspricht der Dimension des Nullraums und ist definiert als:

$$FG = q - \text{Rang}(\mathbf{N}). \quad (4.2)$$

Die oberen und unteren Grenzen der Reaktionen schränken den Lösungsraum weiter ein (analog zu *NB2* (3.8)):

$$\alpha_i \leq r_i \leq \beta_i. \quad (4.3)$$

Alle Flussvektoren \mathbf{r} , die Gleichung (4.1) und (4.3) erfüllen, formen im Allgemeinen einen beschränkten, oder unbeschränkten Polyeder. Um den zulässigen Wertebereiche (Range) von Reaktionsraten im Lösungsraum zu bestimmen, kann eine Flussvariabilitätsanalyse durchgeführt werden (siehe Kapitel 3.2.3). Der Range von Reaktionsraten erlaubt Rückschlüsse auf die Relevanz einer jeden Reaktion. Ist das Minimum einer Reaktionsrate größer null oder das Maximum einer Reaktionsrate kleiner null, ist diese Reaktion im gesamten Lösungsraum aktiv. Daraus folgt, dass, sobald diese Reaktion entfernt wird, das darauf basierende Netzwerkverhalten nicht mehr realisierbar ist. Dieser Range bildet das Fundament des *NetworkReducer*-Algorithmus, der im Folgenden erklärt wird.

Ein Vorteil des *NetworkReducer*-Ansatzes ist es, dass bestimmte Eigenschaften oder Bestandteile des genomskaligen Modells erhalten bleiben, wenn dies gewünscht ist. Der Nutzer hat verschiedene Möglichkeiten zu definieren, welche Eigenschaften bzw. zelluläre Funktionen erhalten bleiben sollen:

- (a) Eine Menge P^M an zu erhaltenden Metaboliten, die nicht gelöscht werden dürfen und in dem reduzierten Modell erhalten bleiben sollen.
- (b) Eine Menge P^R an zu erhaltenden Reaktionen, die nicht gelöscht werden dürfen und in dem reduzierten Modell erhalten bleiben sollen. Zusätzlich kann gefordert werden, dass diese Reaktionen nach der Reduktion nicht geblockt sein dürfen (das bedeutet, dass es zumindest eine zulässige Flussverteilung \mathbf{r} für jede Reaktion $i \in P^R$ geben muss, die $r_i \neq 0$ erfüllt).

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

- (c) Spezifische Funktionen und Funktionalitäten des genomskaligen Modells können während der Reduktion geschützt werden, sodass diese in dem reduzierten Modell weiterhin vorhanden sind. Diese werden über Ungleichungen definiert (s. u.).
- (d) Der minimale zu erhaltende Freiheitsgrad (FG_{\min}), sodass im reduzierten Modell gilt: $FG \geq FG_{\min}$.
- (e) Die minimale Anzahl an zu erhaltenden Reaktionen, sodass im reduzierten Modell gilt: $q \geq q_{\min}$.

Mit den Eigenschaften (a) und (b) kann der Anwender festlegen, welche Metabolite und Reaktionen in dem reduzierten Modell verbleiben sollen. Es ist möglich, dass geschützte Reaktionen nach der Reduktion blockiert sind. Um dies zu umgehen, kann in (b) gefordert werden, dass die zu erhaltenden Reaktionen in mindestens einer gültigen Flussverteilung nicht null sein dürfen. Die zu erhaltenden Metabolite werden derart geschützt, dass es mindestens eine nicht-geblockte Reaktion in diesem Modell geben muss, die den zu erhaltenden Metaboliten beinhaltet. Eigenschaften (d) und (e) werden als Abbruchkriterium genutzt. Die spezifischen Funktionen und Funktionalitäten, die unter (c) definiert werden können, gewährleisten, dass die gewünschten Charakteristika des genomskaligen Modells in der reduzierten Version erhalten bleiben. Jede zu erhaltende Eigenschaft $k, k \in (1, \dots, s)$, wird durch ein korrespondierendes Set an Gleichungen und Ungleichungen beschrieben:

$$\mathbf{D}_k \mathbf{r} \leq \mathbf{d}_k \quad k = 1 \dots s. \quad (4.4)$$

Jedes Ungleichungssystem (4.4) wird im *NetworkReducer*-Algorithmus berücksichtigt, sodass im reduzierten Modell die gewünschten Funktionalitäten gewährleistet werden können. Beispiele solcher Funktionalitäten sind: max. Wachstumsraten, Produktsyntheseraten, etc. Ein Pseudocode des *NetworkReducer*-Algorithmus wird im Folgenden dargestellt.

4.2 Algorithmus NetworkReducer

```

/* Eingaben */
network : stöchiometrisches Netzwerk
PR : geschützte Reaktionen
PM : geschützte Metabolite
FGmin : minimal zu erhaltender Freiheitsgrad
qmin : minimale Anzahl an zu erhaltenden Reaktionen
D1...Ds, d1...ds : Spezifikationen der (s) geschützten Funktionen und Funktionalitäten

/* Entferne geblockte Reaktionen! Checke Zulässigkeit der (s) geschützten Funktionen und Funktionalitäten! */
network = Präprozessierung(network, PR, PM, D1...Ds, d1...ds);

/* removables beinhalten alle entfernbaren Reaktionen. */
removables = {{1...#Reaktionen(network)} \ PR};

/* Hauptschleife: Beschneidung des Netzwerks. */
WHILE FG(network) > FGmin AND NOTEMPTY(removables) AND #Reaktionen(network) > qmin

    /* Bestimme für jede Reaktion die Vereinigung aller Ranges über alle geschützte Funktionen! Eine entfernbar
    Reaktion wird nicht-entfernbar, wenn diese für eine der geschützten Funktionalitäten essentiell ist. */
    [Ranges, removables] = FlussVariabilitätsAnalyse(network, removables, PR, D1...Ds, d1...ds);

    success = false;

    WHILE success = false AND NOTEMPTY(removables)

        /* Selektiere eine entfernbar Reaktion mit dem kleinsten Range für die Deletion! */
        reac = ReaktionsKandidatFürDeletion(network, Ranges, removables)

        network = EntferneReaktion(network, reac);
        removables = removables \ reac;
        success = CheckGeschützteFunktionalitäten(network, PR, PM, D1...Ds, d1...ds);

        IF success = false
            network = Reintegration(network, reac);
        ENDIF

    ENDWHILE

ENDWHILE

ENDWHILE

/* Deletion der durch Reaktionsentfernung entstandenen nichtverbundenen Metabolite */
network = EntferneNichtVerbundeneMetabolite(network);

/* Optional: (verlustfreie) Kompression des Netzwerks durch Zusammenschieben von linearen Ketten */
network = NetzwerkKompression(network, PR, PM);

```

In der Hauptschleife (Beschneidung des Netzwerks) wird das Ursprungsmodell durch Deletion von nicht geschützten Reaktionen reduziert. Unter Erhaltung der Eigenschaften (a) - (e) wird dieser Schritt iterativ ausgeführt und geht einher mit einer Reduktion des

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

Freiheitsgrades, also einer Beschneidung des Lösungsraumes. Der Algorithmus benutzt in jeder Iteration eine FVA und berechnet für jede nicht geschützte, entfernbare Reaktion i den Range (für jede zu erhaltende Eigenschaft k separat). Die Vereinigung F_i der Ranges F_i^k einer jeden Reaktion i für die Funktionalitäten $k, k = 1 \dots s$, wird definiert als: $F_i = \cup_{k=1}^s F_i^k$. Falls keine geschützte Funktionalität vorgegeben wurde, wird der gesamte Range einer jeden Reaktion betrachtet. Sofern eine Reaktion einen Range F_i^k hat, der die Null nicht enthält, handelt es sich um eine essentielle Reaktion für die zu schützende Funktionalität k . Damit darf diese Reaktion nicht gelöscht werden und wird von der Liste der potentiell deletierbaren Reaktionen gestrichen. Alle verbleibenden deletierbaren Reaktionen werden entsprechend des jeweiligen Ranges geordnet und die Reaktion mit dem kleinsten Gesamtrange F_i wird entfernt. Die zu Grunde liegende Annahme für die Entfernung der Reaktion mit dem kleinsten Range ist, dass die verbleibende Variabilität innerhalb des Modells höher ist.

Nachdem eine Reaktion deletiert worden ist, wird überprüft, ob alle geschützten Funktionen gewährleistet werden können. Sofern eine Eigenschaft nicht mehr erfüllt werden kann, wird die zuletzt gelöschte Reaktion wieder in das Modell integriert und als „nicht-entfernbar“ definiert. Statt dieser Reaktion wird die Reaktion des zweit kleinsten Ranges gelöscht usw. Diese gesamte Schleife wird wiederholt durchlaufen, bis eine gewünschte Funktionalität nicht mehr aufrecht erhalten werden kann, oder der zu erhaltenden Freiheitsgrad (die zu erhaltende Anzahl an Reaktionen) erreicht ist. Anschließend werden alle nicht verbundenen Metabolite gelöscht, die durch das Entfernen von Reaktionen entstanden sind.

Nach der Beschneidung des Modells kann optional ein Kompressionsschritt ausgeführt werden. Durch die Beschneidung können einige lineare Ketten von (gekoppelten) Reaktionen entstehen. Solche linearen Ketten können unter Beachtung der geschützten Reaktionen und Metabolite in dem genannten Kompressionsschritt zu einer Reaktion mit kumulativer Stöchiometrie zusammengefasst werden. Der Vorteil dieser verlustfreien Kompression ist, dass der Freiheitsgrad und z. B. auch alle EM des beschnittenen Modells erhalten bleiben, während die Anzahl der Reaktionen und Metabolite weiter reduziert wird.

Angenommen ein genomskaliges Reaktionsnetzwerk soll auf ein Zentralstoffwechselmodell reduziert und alle Reaktionen des Zentralstoffwechsels sowie die Funktionalität des optimalen Wachstums sollen geschützt werden. Dann wird die Beschneidung des genomskaligen Modells für jede Komponente, die in der BSR enthalten ist, einen optimalen Pfad erhalten und die alternativen Routen entfernen. Sofern die an diesem (optimalen) Pfad beteiligten Reaktionen und Metabolite nicht geschützt sind, werden

diese in dem Kompressionsschritt zusammengefasst. Die Stöchiometrie jedes ursprünglich in der BSR enthaltenen Metaboliten wird nun durch eine kumulierte Stöchiometrie der für die Bildung des Metaboliten notwendigen Präkursoren und Kometabolite (ATP, NAD(P)H, etc.) ersetzt. Ein Beispiel: der am Zentralstoffwechsel beteiligte Präkursor P sei geschützt. Die Reaktionen: $2 P \implies D$, $D + \text{NADPH} \implies F$ und $F + \text{ATP} \implies A$ bilden einen Pathway zur Synthese von A, der wiederum ein Bestandteil der BSR des Modells ist (z. B. könnte A eine Aminosäure sein). In der BSR geht A mit einer Stöchiometrie von 2 ein: $2 A + \dots \implies \text{Biomasse}$. Der Kompressionsschritt des *NetworkReducer*s entfernt in diesem Beispiel den Metaboliten A in der BSR, da dieser nicht geschützt werden soll. Statt dessen werden P, ATP und NADPH mit den Stöchiometrien 4, 2, 2 in die BSR eingetragen. Die Metabolite D und F werden in der angepassten BSR nicht explizit integriert, da sie sich jeweils durch ihren Präkursor beschreiben lassen. Die komprimierte BSR sieht demnach wie folgt aus: $4 P + 2 \text{ATP} + 2 \text{NADPH} + \dots \implies \text{Biomasse}$. Im Anschluss werden die drei oben genannten Reaktionen sowie die Metabolite D, F und A entfernt. Das Resultat ist eine kompaktere, aber stöchiometrisch konsistente Darstellung des Reaktionsnetzwerks.

An dieser Stelle wird noch einmal darauf hingewiesen, dass dieser Kompressionsschritt ohne eine Reduktion des Freiheitsgrades oder einer Veränderung der Funktionalität des Stoffwechselmodells einhergeht. Die Abbildung 4.1 fasst den Ablauf des *NetworkReducer*-Algorithmus noch einmal zusammen.

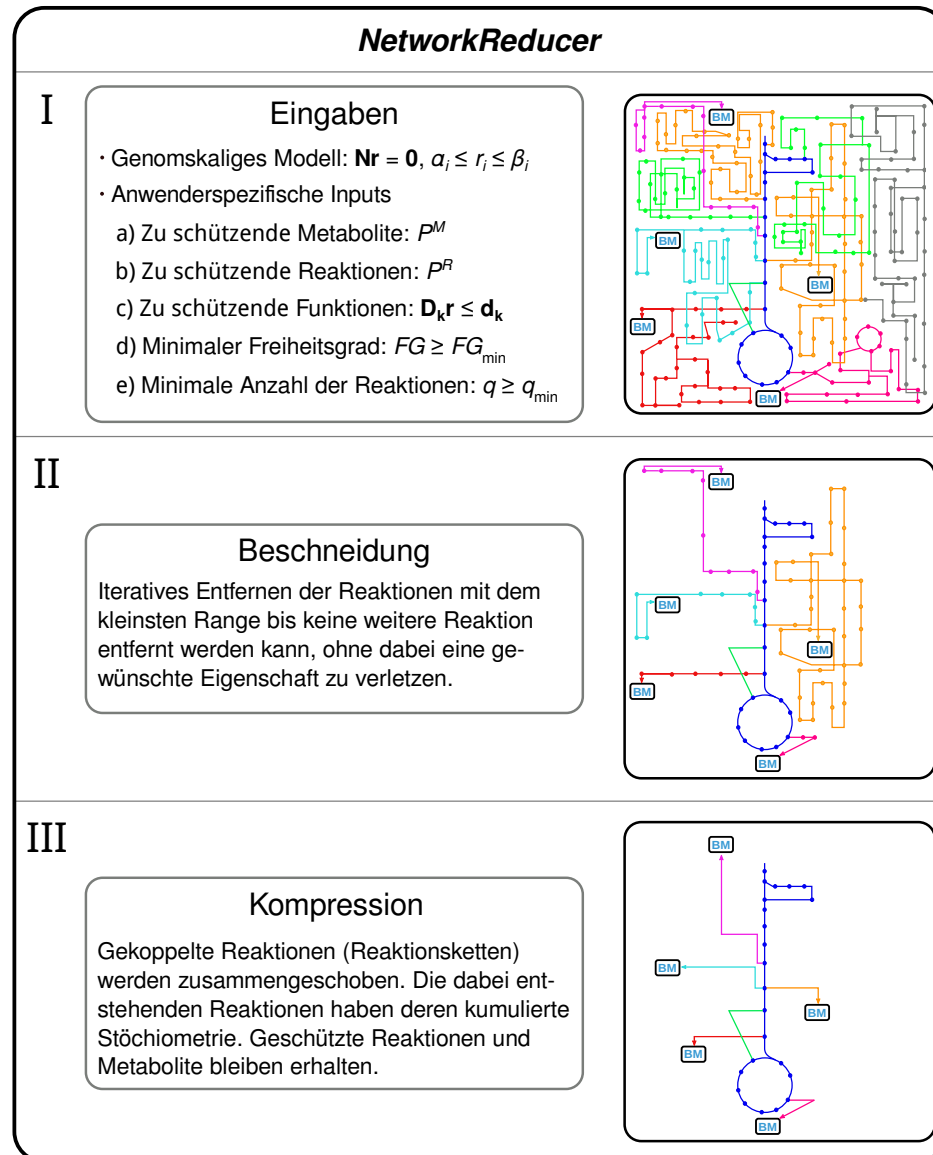


Abbildung 4.1: Ablauf des Reduktionsalgorithmus *NetworkReducer*. I) Spezifikationen der notwendigen Informationen, die der Algorithmus braucht, um ein stöchiometrisches Stoffwechselmodell zu reduzieren. Rechts: Beispiel eines genomskaligen Stoffwechselmodells („BM“ deutet Abflüsse von Monomeren, wie z. B. Aminosäuren, in die Biomassesynthese an). II) Der Algorithmus beschneidet das Modell unter Berücksichtigung der zu erhaltenden Funktionalitäten. Rechts: Zentralstoffwechsel und optimales Wachstum bleiben erhalten. III) Optionale Kompression des beschneideten Modells. Rechts: Nicht-geschützte Reaktionen in Reaktionsketten werden zusammengefasst.

4.3 Machbarkeitsstudie am Beispiel eines genomskaligen *E. coli* Modells

In diesem Abschnitt wird ein realistisches Beispiel zur Anwendung des *NetworkReducer*s dargestellt. I. d. R. wird versucht, ein genomskaliges Stoffwechselmodell mit mehreren tausend Reaktionen und Metaboliten auf ein Zentralstoffwechselmodell mit einer Größe von ca. 80 - 150 Reaktionen zu reduzieren. Neben der Protektion der Reaktionen sollen bestimmte Eigenschaften, wie zum Beispiel optimales Wachstum und/oder Produktsynthese, in dem reduzierten Modell erhalten bleiben.

Als Machbarkeitsnachweis wird das genomskalige Stoffwechselmodell *iAF1260* von *E. coli* (Feist et al., 2007 [71]) reduziert. Dieses Modell, bestehend aus 2382 Reaktionen und 1668 Metaboliten (im Folgenden als *ColiGS* bezeichnet), dient als Basis für die Reduktion. Das daraus abgeleitete Zentralstoffwechselmodell wird anschließend mit einem manuell erstellten Modell von Orth et al. aus dem Jahre 2010 verglichen [185], welches aus 95 Reaktionen und 72 Metaboliten besteht und im Folgenden *ColiCore* genannt wird. *ColiCore* wurde von Orth et al. [185] aus dem genomskaligen Modell abgeleitet. Daher sind die Reaktions- und Metabolitbezeichnungen identisch. Diese Tatsache erlaubt einen direkten Vergleich des *ColiCore* mit dem durch den *NetworkReducer* reduzierten Modell. *ColiCore* wird im Folgenden als Zielnetzwerk betrachtet. Das bedeutet, dass dieses Zentralstoffwechselmodell ausgehend von dem genomskaligen Modell (*ColiGS*) generiert werden soll.

Im Vergleich zu *ColiGS* haben Orth et al. in ihrem per Hand erstellten Modell einige Reaktionen minimal verändert [185], was wie folgt berücksichtigt wurde. *ColiCore* verwendet für die Fumarat-Reduktase-Reaktion Ubiquinol-8 anstatt Menaquinol-8 als Redox-Carrier. Diese Reaktion wird als solche in das *ColiGS* integriert, um diese Reaktion explizit schützen zu können. Des Weiteren wird die Biomasse als intrazellulärer Metabolit eingeführt und zusätzlich eine Exportreaktion für die Biomasse in das *ColiGS* eingefügt. Durch diese minimale Erweiterung verändert sich die Größe des *ColiGS* auf 2384 Reaktionen und 1669 Metabolite. Bei beiden Modellen (*ColiGS* und *ColiCore*) soll Glukose die einzige Kohlenstoffquelle sein. Da das *ColiCore* auch andere Substrate aufnehmen kann (z. B. Aminosäuren), wird dieses Netz ausschließlich für die Aufnahme von Glukose konfiguriert und andere Substrataufnahmen eliminiert. Wie in *ColiGS* wird in *ColiCore* die Biomasse als zusätzlicher Metabolit und eine dazugehörige Exportreaktion eingefügt. Mit diesen Veränderungen besteht das *ColiCore* aus 88 Reaktionen und 69 internen Metaboliten. Es sollen nun verschiedene Eigenschaften und Bestandteile des *ColiGS* erhalten werden, auf die im Folgenden eingegangen wird.

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

Die 88 Reaktionen des *ColiCore* werden als geschützte Reaktionen definiert (bzw. streng genommen werden 87 Reaktionen explizit durch die Indexmenge P^R geschützt und die BSR über die Nebenbedingungen in (c)). Zusätzlich wird gefordert, dass alle geschützten Reaktionen nicht geblockt sein dürfen (also deren Fluss in mindestens einer Flussverteilung nicht null ist). Da alle Reaktionen des Zielmodells geschützt werden, ist es nicht notwendig die Metabolite explizit zu schützen, da deren Erhalt implizit (durch die Reaktionen) gesichert wird ($P^M = \emptyset$). Das reduzierte Modell soll 99,9% der maximalen Wachstumsrate des *ColiGS* mit Wachstum auf Glukose unter aeroben sowie unter anaeroben Bedingungen ermöglichen. Die maximale Glukoseaufnahmerate wird auf 10 mmol/gTM/h beschränkt (gTM: Gramm Trocken-Biomasse). Unter diesen Bedingungen ergeben sich im *ColiGS* folgende maximale Wachstumsraten: $0,9290 \text{ h}^{-1}$ unter aeroben Bedingungen und $0,2309 \text{ h}^{-1}$ unter anaeroben Bedingungen. Diese Nebenbedingungen werden als zwei getrennte Ungleichungssysteme formuliert:

$$\begin{array}{ll}
 \mathbf{D}_1/\mathbf{d}_1 : & \mathbf{D}_2/\mathbf{d}_2 : \\
 r_{\text{GlukoseAufnahme}} \leq 10 & r_{\text{GlukoseAufnahme}} \leq 10 \\
 -\mu \leq -0,999\mu_{\max\text{ColiGS,aerob}} & r_{\text{O}_2\text{Aufnahme}} \leq 0 \\
 & -\mu \leq -0,999\mu_{\max\text{ColiGS,anaerob}}.
 \end{array} \tag{4.5}$$

Für den Erhaltungsstoffwechsel (ATP maintenance) wird angenommen, dass dieser mindestens $8,39 \text{ mmol ATP/gTM/h}$ benötigt. Diese Bedingung ist in den allgemeinen Rategrenzen als untere Schranke integriert (siehe Gleichung (4.3)). Da das *ColiGS* soweit wie möglich reduziert werden soll, wird der zu erhaltende minimale Freiheitsgrad auf eins gesetzt ($FG_{\min} = 1$). Die minimale Anzahl der zu erhaltenden Reaktionen ist in diesem Szenario nicht entscheidend, da die relevanten Reaktionen durch die Indexmenge P^R geschützt werden.

Mit den dargestellten Bedingungen kann der *NetworkReducer*-Algorithmus gestartet werden. Wird nur der erste Teil des Algorithmus angewendet (Beschneidung des Modells), entsteht aus dem *ColiGS* mit 2384 Reaktionen und 1669 Metaboliten ein reduziertes Modell mit 455 Reaktionen und 438 Metaboliten. Dieses reduzierte Modell (im Folgenden *ColiPruned* genannt) ist ein echtes Subnetzwerk des *ColiGS*. Es ist in der Lage, die geforderten Mindest-Wachstumsraten (aerob und anaerob) wiederzugeben. Genauer gesagt: unter aeroben Bedingungen ist die maximale Wachstumsrate fast ($0,9288 \text{ h}^{-1}$ statt $0,9290 \text{ h}^{-1}$) und unter anaeroben Bedingungen exakt identisch (siehe Tabelle 4.1).

4.3 Machbarkeitsstudie am Beispiel eines genomskaligen *E. coli* Modells

Tabelle 4.1: Eigenschaften des originalen genomskaligen *E. coli* Modells, des Zentralstoffwechselmodells von Orth et al. und zweier mit *NetworkReducer* abgeleiteter Reduktionen.

	Genomskaliges <i>E. coli</i> Modell (<i>ColiGS</i>)	Beschnittenes <i>E. coli</i> Modell (<i>ColiPruned</i>)	Reduziertes <i>E. coli</i> Modell (<i>ColiPruned-Comp</i>)	Zentralstoff- wechselmodell Orth et al. [185](<i>Coli- Core</i>)
Anzahl Reaktionen	2384	455	105	88
Anzahl interner Metabolite	1669	438	85	69
Anzahl externer Metabolite	305	33	33	17
Freiheitsgrad	753	26	26	24
Anzahl linearer Reaktionsketten (und darin enthaltener Reaktionen)	321 (890)	23 (413)	22 (62)	23 (55)
μ_{max} (aerob)	0,9290 h ⁻¹	0,9288 h ⁻¹	0,9288 h ⁻¹	0,8739 h ⁻¹
μ_{max} (anaerob)	0,2309 h ⁻¹	0,2309 h ⁻¹	0,2309 h ⁻¹	0,2117 h ⁻¹

Nicht nur die Anzahl der Reaktionen ist bei der Beschneidung dramatisch gesunken, sondern auch die Anzahl der Freiheitsgrade (von 753 auf 26). Dennoch verwundert es auf den ersten Blick, dass bei 26 Freiheitsgraden das *ColiPruned* durch 455 Reaktionen und 438 internen Metaboliten beschrieben wird. Der Grund ist, dass viele (gekoppelte) Reaktionen in den Biomassessynthesewegen lineare Ketten bilden. Das *ColiPruned* besitzt 23 solcher Ketten, die bereits 413 Reaktionen enthalten.

Der kleinere Freiheitsgrad sorgt dafür, dass trotz der vielen Reaktionen die Enumeration der Elementarmoden möglich ist. Damit ist ein wichtiges Anliegen erreicht; nämlich die Berechnung der EM. *ColiPruned* besitzt 1.410.322 EM. Der Freiheitsgrad von *ColiPruned* entspricht fast dem des *ColiCore* Modells ($FG_{ColiPruned} = 26$; $FG_{ColiCore} = 24$). Bei der Beschneidung sind viele suboptimale und redundante Pfade deletiert wor-

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

den. Das verbleibende Modell *ColiPruned* kann unter Nutzung des zweiten Teils des *NetworkReducers*, der Kompression, weiter reduziert werden.

Wie eingangs erklärt, ist diese Kompression nicht mit einem Verlust an Freiheitsgraden assoziiert (siehe Tabelle 4.1). Unter Erhaltung der geschützten Reaktionen (in P^R) kann das *ColiPruned* mit einer Netzwerkgröße von 455 Reaktionen und 438 internen Metaboliten auf 105 Reaktionen und 85 interne Metabolite komprimiert werden ($FG_{ColiPrunedComp} = 26$). Ebenso unverändert wie der Freiheitsgrad bleibt die Menge der EM im *ColiPrunedComp*. Da das *ColiPruned* lediglich zusammengeschoben wird, bleiben die Eigenschaften (max. Wachstum unter aeroben und anaeroben Bedingungen) des *ColiPruned* im *ColiPrunedComp* erhalten.

Trotz der Ähnlichkeit des *ColiPrunedComp* und des *ColiCore* Modells von Orth et al. [185] gibt es nachfolgende Unterschiede. Insgesamt hat das *ColiPrunedComp* im Vergleich zu *ColiCore* 17 Reaktionen und 16 interne Metabolite mehr. Der Großteil dieser Reaktionen und Metabolite kann direkt auf die Darstellung der Transportreaktionen durch die verschiedenen Kompartimente zurückgeführt werden. Genauer gesagt, unterscheidet sich die Anzahl der Kompartimente in den genannten Modellen. Das *ColiCore* von Orth et al. bilanziert explizit zwei Kompartimente (nämlich das Cytoplasma und den extrazellulären Raum) und implizit die Umgebung [185]. Daraus ergeben sich drei Transportreaktionen, um z. B. verschiedene Spurenelemente aufzunehmen. Dementsprechend ergeben sich drei Metabolite M der jeweiligen Spezies (für jedes Kompartiment ein anderer Metabolit): $M_{Umgebung} \implies M_{extrazellulärer\ Raum} \implies M_{Cytoplasma}$ (analog für Exportreaktionen). Das *ColiGS* wiederum wird durch drei Kompartimente dargestellt: das Cytoplasma, den extrazellulären Raum, das Periplasma und implizit die Umgebung. Damit ergeben sich vier Austauschreaktionen mit vier dazugehörigen Spezies

($M_{Umgebung} \implies M_{extrazellulärer\ Raum} \implies M_{Periplasma} \implies M_{Cytoplasma}$). Da nur jeweils drei der vier Transportreaktionen explizit geschützt werden konnten, da nur die Reaktionen des *ColiCore* geschützt wurden, verbleibt für jede Transportkette eine weitere Reaktion, die nicht komprimiert werden konnte. 15 der oben genannten 17 zusätzlichen Reaktionen sind daher Transportreaktionen, die durch den geforderten Erhalt der Transportprozesse im *ColiPrunedComp* verbleiben. Dazugehörig gibt es für das zusätzliche Kompartiment (Periplasma) jeweils einen weiteren Metaboliten (in der Summe 15 zusätzliche). Da diese 15 Reaktionen und Metabolite umgeben sind von geschützten Reaktionen, können diese nicht zusammengeschoben werden. Der Großteil der Differenz des *ColiPrunedComp* und *ColiCore* Modells ist damit geklärt.

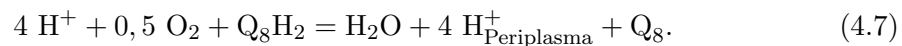
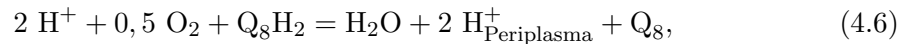
Der verbleibende Metabolit ist eine Phosphatspezies (M_{Pi_p}), die im Periplasma vorkommt. Die Transportprozesse für Phosphat sind in der gleichen Art und Weise geschützt

4.3 Machbarkeitsstudie am Beispiel eines genomskaligen *E. coli* Modells

wie die anderen Transportreaktionen. Allerdings bildet, wie in *ColiCore*, die zwischen dem Kompartiment extrazellulärer Raum und Periplasma liegende Phosphattransportreaktion eine lineare Kette mit der BSR, die nicht geschützt worden ist (statt der BSR wird der Biomasseexport geschützt). Diese lineare Kette bedingt, dass diese Transportreaktion in der neuen BSR integriert wird und dementsprechend der dazugehörige Metabolit M.Pi_p. Daraus folgt, dass trotz 15 zusätzlicher Transportreaktionen 16 zusätzliche Metabolite in dem *ColiPrunedComp* vorhanden sind.

Die verbleibenden Unterschiede basieren auf zwei zusätzlichen Reaktionen, die nicht im *ColiCore* enthalten sind. Diese sorgen dafür, dass die Wachstumsraten des *ColiPrunedComp* denen des *ColiGS* entsprechen (beide Raten sind größer als die des *ColiCore*). Des Weiteren sind diese beiden Reaktionen für den größeren Freiheitsgrad des *ColiPrunedComp* im Vergleich zu *ColiCore* verantwortlich.

Die erste der beiden Reaktionen produziert molekularen Wasserstoff. Nur mit dieser Reaktion kann die maximale Wachstumsrate unter anaeroben Bedingungen gewährleistet werden (das selbe trifft für das *ColiGS* zu; sofern der Export von molekularem Wasserstoff blockiert wird, sinkt die anaerobe Wachstumsrate drastisch). Bei der zweiten zusätzlichen Reaktion handelt es sich um eine Cytochrom-Oxidase. Das *ColiGS* verfügt über zwei Cytochrom-Oxidasen: eine Cytochrom-*bd*- und eine Cytochrom-*bo₃*-Oxidase. Beide oxidieren Ubiquinol und pumpen Protonen über die Membran, um einen Protonengradienten zu erzeugen. Der Unterschied liegt in der Anzahl (Stöchiometrie) der gepumpten Protonen:



Um unter aeroben Bedingungen die maximale Wachstumsrate zu erreichen, ist es notwendig, dass das Modell die Reaktion entsprechend Gleichung (4.7) nutzen kann. Die größere Anzahl der gepumpten Protonen führt zu einer besseren ATP-Ausbeute und damit dazu, dass der Organismus schneller wachsen kann. Um die durch Gleichung (4.5) geforderten Nebenbedingungen einhalten zu können, muss der *NetworkReducer*-Algorithmus diese Reaktion behalten. Im Gegensatz dazu transportiert Reaktion (4.6) nur zwei Protonen. Dies geht mit einer geringeren ATP- und Biomasseausbeute einher.

Um zu überprüfen, ob diese Reaktion allein für die unterschiedlichen Wachstumsraten des *ColiCore* und *ColiPrunedComp* im Vergleich zu *ColiGS* verantwortlich ist, wird die Reaktion (4.7) in das *ColiCore* Modell integriert. Unter der gegebenen Bedingungen (Aufnahme von 10 mmol Glukose/gTM/h und aerobes Wachstum) steigt die maximale

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

Wachstumsrate des *ColiCore* signifikant von $0,8739 \text{ h}^{-1}$ auf $0,9647 \text{ h}^{-1}$ und ist damit größer als die maximale Wachstumsrate des *ColiPrunedComp* und des *ColiGS* Modells. Das wiederum bedeutet, dass diese Reaktion allein nicht verantwortlich sein kann für die unterschiedlichen Wachstumsraten der genutzten Modelle (*ColiGS* und *ColiCore*).

Da die restlichen Reaktionen der beiden Modelle allerdings identisch sind, kann angenommen werden, dass zusätzlich die Stöchiometrie der BSR der Grund für die Unterschiede sein muss. Die stöchiometrischen Koeffizienten der einzelnen Präkursoren (Pyruvat, Acetyl-CoA, etc.) für die Monomere der Biomasse sind in *ColiCore* manuell berechnet worden. Laut Orth et al. ist dies nicht trivial [185]: „*Since most of the subunits of the cellular macromolecules, such as nucleic acids and amino acids, are not present in the core model, they could not be directly accounted for in the biomass reaction. The metabolites in the core model that those macromolecular subunits are synthesized from are included instead. These are precursor metabolites. For example, the amino acid L-alanine is synthesized from pyruvate and L-glutamate, so both of these metabolites are consumed in the biomass reaction.*“

Anhand dieser Aussage wird deutlich, dass alle Bestandteile bzw. Komponenten der alten BSR stöchiometrisch an die Metabolite angepasst werden müssen, die in dem Zentralstoffwechselmodell erhalten bleiben sollen. Diese Transformation ist nicht trivial und umso anspruchsvoller, sofern dies manuell durchgeführt wird. Der hier vorgestellte *NetworkReducer* hilft, diesen Prozess zu automatisieren und eine stöchiometrisch korrekte Abbildung der BSR zu generieren, die konsistent ist mit der BSR des genomskaligen Modells. Um dies genauer zu analysieren, wird im Folgenden ein Vergleich der BSR des *ColiPrunedComp* und des *ColiCore* gezeigt.

Tabelle 4.2 stellt den Vergleich der beiden Biomassessynthesereaktionen dar. Auf den ersten Blick ähneln sich die beiden BSR. Insgesamt setzt sich die BSR des *ColiCore* aus 23 Komponenten zusammen (Edukte und Produkte). 22 dieser Komponenten der BSR_{*ColiCore*} sind auch in der BSR des *ColiPrunedComp* wiederzufinden.

Die BSR des *ColiPrunedComp* ist ein wenig komplexer. Ein Vorteil ist, dass die BSR_{*ColiPrunedComp*} alle Spurenelemente und Mineralstoffe und ihren Bedarf stöchiometrisch abbildet. All diese Spurenelemente sind für die Biomassessynthese notwendig und im *ColiGS* vorhanden. Während der Reduktion werden diese Metabolite und deren Transportreaktionen berücksichtigt und verbleiben im Modell. Durch die anschließende Kompression werden diese (linearen) Transportreaktionen zusammengeschoben, sodass die Spurenelemente als externe Metabolite in die BSR_{*ColiPrunedComp*} eingehen (Transport und Konsum ist dadurch als ein Schritt modelliert). Da diese Elemente als externe Metabolite integriert werden, verändern diese die Eigenschaften des Modells nicht. Dennoch ist

Tabelle 4.2: Vergleich der stöchiometrischen Koeffizienten des *ColiPrunedComp* und *ColiCore* in der BSR. Positive Werte beschreiben die Produktion und negative Werte den Konsum des jeweiligen Metaboliten für die Biomassensynthese. Die BSR des *ColiPrunedComp* beinhaltet auch zusätzlich alle Spurenelemente des *ColiGS* (nicht dargestellt). Alle Werte sind in [mmol] angegeben.

Metabolit	BSR des <i>Coli-Core</i>	BSR des <i>Coli-PrunedComp</i>	Metabolit	BSR des <i>Coli-Core</i>	BSR des <i>Coli-PrunedComp</i>
2-Oxoglutarat	4,1182	7,4661	Glyoxylat		0,0007
3-Phospho-D-Glycerat	- 1,496	- 1,7175	H _c	59,8100	64,5223
Acetat		0,5810	H _p		- 0,1686
Acetyl-CoA	- 3,7478	- 3,8560	H ₂ O	- 59,8100	- 57,4196
ADP	59,8100	67,7163	NAD	- 3,5470	- 0,3971
AMP		2,2653	NADH	3,5470	0,3971
ATP	- 59,8100	- 69,9816	NADP	13,0279	7,2399
CO ₂		1,7341	NADPH	- 13,0279	- 7,2399
Coenzym-A	3,7487	4,3809	NH ₄		- 0,4352
Dihydroxyaceton-Phosphat		- 0,1413	Oxaloacetat	- 1,7867	- 2,9257
Erythrose-4-Phosphat	- 0,3610	- 0,3720	Phosphoenolpyruvat	- 0,5191	- 0,8102
Format		- 0,1080	Pi _c	59,8100	75,3394
Fruktose-6-Phosphat	- 0,0709	- 0,0945	Pyruvat	- 2,8328	- 2,7842
Fumarat		0,7063	Ribose-5-Phosphat	- 0,8977	- 0,9325
Glukose-6-Phosphat	- 0,2050		Ribulose-5-Phosphat		- 0,0398
Glutamat	- 4,9414	- 6,7339	Succinat		0,8586
Glutamin	- 0,2557	- 1,8075	Succinyl-CoA	- 0,5249	
Glycerinaldehyd-3-Phosphat	- 0,1290	0,0540			

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

es nützlich zu wissen, in welchen molaren Verhältnissen diese Elemente benötigt werden, um Biomasse zu synthetisieren. Es ist demnach von Vorteil, dass der *NetworkReducer*-Algorithmus diese Metabolite erhält. In direktem Zusammenhang damit steht der Metabolit H_p , der beim aktiven Transport via Antiport für die Aufnahme verschiedener Spurenelemente dient und damit in der BSR von *ColiPrunedComp* auftaucht.

Des Weiteren ist AMP im *ColiPrunedComp* im Gegensatz zu *ColiCore* explizit modelliert ($ATP + ADP + (AMP) = \text{const.}$). Weitere Unterschiede entstehen dadurch, dass verschiedene Referenzpunkte für Präkursormetabolite genutzt werden, so z. B. für Glyceraldehyd-3-Phosphat. *ColiCore* nutzt diesen Metaboliten als Präkursor für die BSR während *ColiPrunedComp* statt dessen (in einer ähnlichen Menge) Dihydroxyacetonphosphat benutzt. Da beide Metabolite über eine Triose-Isomerase Reaktion verknüpft sind, also direkt ineinander umwandelbar sind, können auch beide als Präkursor genutzt werden. In ähnlicher Weise verhält es sich mit Ribose-5-Phosphat (R5P) und Ribulose-5-Phosphat (Ru5P). Die BSR_{*ColiCore*} berücksichtigt nur R5P und die BSR_{*ColiPrunedComp*} sowohl R5P als auch Ru5P. Durch eine Isomerase Reaktion (Ribose-5-Phosphat Isomerase) können beide Metabolite ineinander umgewandelt werden und damit auch beide als Präkursor für die BSR dienen.

Des Weiteren fällt auf, dass der Bedarf an Energie im *ColiCore* deutlich geringer ist. Da die Stöchiometrie der BSR_{*ColiPrunedComp*} aber exakt ist, kann angenommen werden, dass der ATP-Bedarf des *ColiCore* unterschätzt wird. Bzgl. der Reduktionsäquivalente zeigt sich ein gegensätzliches Bild. Die BSR des *ColiCore* verbraucht deutlich mehr Reduktionsäquivalente in Form von NADPH und NADH als die BSR_{*ColiPrunedComp*}.

Aufgrund der beschriebenen Differenzen der Modelle und der unterschiedlichen BSR des *ColiCore* und des *ColiPrunedComp* ist das durch den *NetworkReducer* erzeugte Modell in der Lage, die Wachstumsraten des *ColiGS* zu beschreiben. Das gelingt dem manuell erzeugten *ColiCore* nicht. *ColiPruned* ist ein echtes Subnetzwerk des *ColiGS* und beinhaltet die identische BSR im Vergleich zu *ColiGS*. Die Kompression hat lediglich lineare Pfade zusammengeschoben, um die Größe des Modells weiter zu reduzieren (ohne den Verlust eines Freiheitsgrades und damit potentieller Eigenschaften). Damit kann die erzeugte BSR des *ColiPrunedComp* als reproduzierbare und unverfälschte Beschreibung der Biomassesynthese des genomskaligen Modells betrachtet werden.

Um die Qualität des Algorithmus zu untermauern, wird eine zusätzliche Analyse präsentiert, in der für zwei verschiedene Szenarien (I: aerobes Wachstum mit fester Wachstumsrate; II: anaerobes Wachstum mit fester Wachstumsrate) die Flussvariabilität aller Reaktionen der drei Modelle *ColiGS*, *ColiCore* und *ColiPrunedComp* vergli-

4.3 Machbarkeitsstudie am Beispiel eines genomskaligen *E. coli* Modells

chen wird. Die vollständigen dazugehörigen Tabellen sind im Anhang A.3 und A.4 zu finden. Es wird an dieser Stelle lediglich ein Ausschnitt vorgestellt.

Im Folgenden wird jeweils die Übereinstimmung des zulässigen Wertebereichs der Reaktionsraten von *ColiPrunedComp* und *ColiCore* mit denen des *ColiGS* in den zwei Szenarien verglichen. Es werden drei verschiedene Klassen betrachtet: exakte (Abweichung im Bereich $[0, 0, 1]$ mmol/gTM/h), gute (Abweichung im Bereich $(0, 1, 1]$ mmol/gTM/h) und geringe (Abweichung im Bereich $(1, 10]$ mmol/gTM/h) Übereinstimmung. Die folgende Tabelle zeigt die Ergebnisse.

Tabelle 4.3: Übereinstimmung der Flussvariabilität der betrachteten Reaktionsraten von *ColiCore* und *ColiPrunedComp* mit denen des *ColiGS* in [mmol/gTM/h].

I: aerobes Wachstum ($\mu = 0,85h^{-1}$)	<i>ColiPrunedComp</i>	<i>ColiCore</i>
exakte Übereinstimmung	33/87	1/87
gute Übereinstimmung	4/87	6/87
geringe Übereinstimmung	27/87	31/87
II: anaerobes Wachstum ($\mu = 0,20h^{-1}$)	<i>ColiPrunedComp</i>	<i>ColiCore</i>
exakte Übereinstimmung	49/87	6/87
gute Übereinstimmung	10/87	17/87
geringe Übereinstimmung	20/87	55/87

Tabelle 4.3 zeigt, dass das *ColiPrunedComp* in beiden Szenarien gerade im Bereich der exakten Übereinstimmung wesentlich besser als das *ColiCore* mit den zulässigen Wertebereichen der betrachteten Reaktionsraten des *ColiGS* übereinstimmt. Trotz der deutlich höheren Übereinstimmung der Flussvariabilitäten des *ColiPrunedComp* im Vergleich zu *ColiGS*, gibt es einige Abweichungen. Diese lassen sich im wesentlichen darauf zurückführen, dass (1) einige Reaktionen des *ColiGS* interne Zyklen bilden und damit der gesamte, in Gleichung (4.3) definierte Wertebereich zulässig ist. Dies ist in *ColiPrunedComp* nicht mehr möglich, da durch den *NetworkReducer* einige Reaktionen entfernt wurden, die Bestandteil interner Zyklen sind. Dies stellt einen weiteren Vorteil des *NetworkReducers* dar, da diese Zyklen biologisch nicht sinnvoll sind. Weiterhin gibt es (2) verschiedene Aufnahmesysteme für bestimmte Metabolite, sodass im *ColiGS* der Wertebereich solcher Reaktionsraten sehr variabel ist, während das *ColiPrunedComp* für die selben Metabolite nur ein Aufnahmesystem besitzt, welches unter diesen Bedingungen

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

genutzt werden muss. Das gleiche gilt (3) für Reaktionen, deren Ko-Metabolit austauschbar ist (z. B. NADPH oder NADH). Außerdem besitzt das *ColiGS* einen deutlich höheren Freiheitsgrad, sodass bei suboptimalen Wachstumsraten ein größerer Wertebereich für verschiedene Reaktionsraten zulässig ist.

Dennoch ist klar zu erkennen, dass das *ColiPrunedComp* das Verhalten des *ColiGS* unter den gezeigten Bedingungen deutlich besser wiedergeben kann als das *ColiCore*. Dies war nicht unbedingt zu erwarten, da die Flussvariabilitäten während des Reduktionsprozesses nicht explizit geschützt worden sind. Das hohe Maß an Übereinstimmung der Funktionalität und Variabilität des *ColiPrunedComp* und des *ColiGS* unterstreicht die Vorteile des *NetworkReducer*-Algorithmus.

4.4 Ableitung von Modellen des Zentralstoffwechsels in Cyanobakterien

Ein Anliegen dieser Arbeit ist es, skalenübergreifend Interventionsstrategien zur Optimierung der Biokraftstoffproduktion in *Synechocystis* sp. PCC 6803 zu identifizieren. Da zu Beginn dieser Arbeit nur das in Kapitel 3.5 vorgestellte genomskalige Modell von *Synechocystis* sp. PCC 6803, *SyneGS*, vorhanden war, wird in diesem Abschnitt der *NetworkReducer* angewendet, um dieses Modell zu reduzieren. Zwei Modelle unterschiedlicher Netzwerkskalen sollen dabei aus dem *SyneGS* abgeleitet werden: ein Modell mittlerer Größe, *SynePruned*, und ein Minimalmodell des Zentralstoffwechsels, *SyneMin*. Zusätzlich zu *SyneGS* sollen diese reduzierten Modelle in Kapitel 5 genutzt werden, um skalenübergreifend potentielle Interventionsmöglichkeiten zur wachstumsgekoppelten Produktion von Biokraftstoffen zu identifizieren.

Ein Netzwerk mittlerer Größe für *Synechocystis* (*SyneRN*) wurde von Knoop et al. manuell aus *SyneGS* abgeleitet und uns zur Verfügung gestellt. Hier sollte nun zunächst überprüft werden, ob dieses Modell mit dem *NetworkReducer* reproduzierbar ableitbar ist und somit ein konsistentes Subnetzwerk des genomskaligen Modells darstellt. Die dafür notwendigen Eingaben werden wie folgt definiert:

- das genomskalige Modell (*SyneGS*) dient als Ausgangsnetzwerk,
- alle Reaktionen des *SyneRN* werden als zu schützende Reaktionen definiert,
- Metabolite werden nicht explizit geschützt ($P^M = \emptyset$),
- und der minimal zu erhaltende Freiheitsgrad wird auf 1 gesetzt.

Das von Knoop et al. abgeleitete Netzwerk, *SyneRN*, hat eine Größe von 505 Reaktionen und 487 internen Metaboliten. Demnach hat der Indexvektor P^R eine Größe von 505 Einträgen. Die zu erhaltende Funktionalität ist das maximale phototrophe Wachstum mit einer maximalen Photonenaufnahme von 100 mmol/gTM/h:

$$\begin{aligned}
 \mathbf{D}_1/\mathbf{d}_1 : \\
 r_{\text{PhotonenAufnahme}} &\leq 100 \\
 -\mu &\leq -0,999\mu_{\max_{\text{SyneGS}}}.
 \end{aligned} \tag{4.8}$$

Mit diesen Eingaben wird der *NetworkReducer* gestartet. Der Fokus der Reduktion liegt ausschließlich auf der Beschneidung des Modells. Die Ergebnisse in der Tabelle 4.4 verdeutlichen, dass das erhaltene beschnittene Modell von *Synechocystis* (*SynePruned*)

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

tatsächlich deckungsgleich mit *SyneRN* ist. Daraus kann geschlussfolgert werden, dass

Tabelle 4.4: Vergleich der topologischen Eigenschaften der *Synechocystis*-Modelle.

	Genomskaliges Modell (<i>SyneGS</i>) [134]	Reduziertes Modell von Knoop et al. (<i>SyneRN</i>)	Beschnittenes Modell (<i>SynePruned</i>)	Minimalmodell (<i>SyneMin</i>)
Anzahl Reaktionen	612	505	505	37
Anzahl interner Metabolite	526	487	487	38
Anzahl externer Metabolite	56	34	34	33
Freiheitsgrad	102	30	30	4
Anzahl linearer Ketten (und darin enthaltener Reaktionen)	51 (431)	23 (442)	23 (442)	7 (19)
μ_{\max}	0,185 h ⁻¹	0,185 h ⁻¹	0,185 h ⁻¹	0,185 h ⁻¹

beide Modelle das gleiche (echte) Subnetzwerk mit den gewünschten Eigenschaften des genomskaligen Modells repräsentieren. Das *SynePruned* stellt nunmehr ein Zentralstoffwechselmodell mittlerer Größe dar, das die geforderte Bedingung (maximales phototrophes Wachstum) erfüllt. Der kleinere Freiheitsgrad des *SynePruned* erlaubt es außerdem, die vollständige Menge der Elementarmoden zu enumerieren und die auf den EM basierenden Methoden anzuwenden.

Das zweite reduzierte Modell, das im Folgenden von dem genomskaligen Modell von *Synechocystis*, *SyneGS*, abgeleitet wird, soll ein minimales Zentralstoffwechselmodell (*SyneMin*) darstellen. Es soll die CO₂-Fixierung im Calvin-Benson Zyklus, unspezifischen ATP-Verbrauch, optimales (phototrophes) Wachstum sowie maximale Ethanolproduktion beschreiben können. Der heterotrophe (Nacht-) Metabolismus und die Aufnahme von Ammonium werden an dieser Stelle vernachlässigt und aus dem Modell entfernt. Die Biomasse wird hier als interner Metabolit berücksichtigt und deren Export als zusätzliche Reaktion in das Modell *SyneGS* integriert.

4.4 Ableitung von Modellen des Zentralstoffwechsels in Cyanobakterien

Damit ergeben sich die folgenden Eingaben:

- das genomskalige Modell (*SyneGS*) dient als Ausgangsnetzwerk,
- alle relevanten Reaktionen des Calvin Zyklus, sowie die Ethanolproduktion, der unspezifische ATP-Verbrauch und der Biomasseexport werden explizit geschützt (insgesamt: 26 Reaktionen),
- Metabolite werden nicht explizit geschützt ($P^M = \emptyset$),
- der minimal zu erhaltende Freiheitsgrad wird auf 1 gesetzt,
- und die Anzahl der minimal zu erhaltenden Reaktionen wird auf 1 gesetzt.

Die zu erhaltenden Funktionalitäten sind einerseits die maximale phototrophe Wachstumsrate und andererseits die maximale Ethanolproduktionsrate. Die Photonenaufnahme wird beschränkt auf maximal 100 mmol/gTM/h.

$$\begin{array}{ll}
 \mathbf{D}_1/\mathbf{d}_1 : & \mathbf{D}_2/\mathbf{d}_2 : \\
 r_{\text{PhotonenAufnahme}} \leq 100 & r_{\text{PhotonenAufnahme}} \leq 100 \\
 -\mu \leq -0,999\mu_{\max_{\text{SyneGS}}} & -r_{\text{Ethanol}} \leq -0,999r_{\max_{\text{EthanolSyneGS}}}
 \end{array} \tag{4.9}$$

Mit diesen Eingaben liefert der *NetworkReducer* ein sehr kleines, beschnittenes und komprimiertes Netzwerk (*SyneMin*, siehe Anhang A.5). Dieses reduzierte Modell beinhaltet 37 Reaktionen und 38 interne Metabolite. Der Rang der stöchiometrischen Matrix ist 33. Damit ergibt sich ein Freiheitsgrad von vier. Die optimale phototrophe Wachstumsrate und die maximale Ethanolproduktionsrate bleiben erhalten. Die nicht-geschützten, aber beibehaltenen Reaktionen setzen sich zusammen aus Transportreaktionen (für Hydrogencarbonat, Sauerstoff und CO₂), Photosystem-assoziierten Reaktionen (zusammengefasste Reaktion für PSI und PSII, FNR, ATPase und eine CO₂ unabhängige NADPH Dehydrogenase), der Hydrogencarbonatequilibriumreaktion im Carboxysomen und der Biomassesynthesereaktion (BSR). Die topologischen Eigenschaften sind in Tabelle 4.4 im Vergleich zu *SyneGS* und *SynePruned* dargestellt. Auf dieses Modell wird in Kapitel 5.3.1 näher eingegangen.

Wie gezeigt werden konnte, besteht der große Vorteil des *NetworkReducers* darin, dass bestimmte (gewünschte) Eigenschaften eines genomskaligen Modells während der Reduktion erhalten werden können. Damit können wesentliche Prinzipien und grundlegende Zusammenhänge in reduzierten Modellen erhalten werden. Die Reduktion

4 Algorithmus zur Ableitung von eigenschaftserhaltenden Teilmodellen

ermöglicht es, dass bestimmte (z. B. elementarmodenbasierte) Methoden auf diese Minimalmodelle angewendet werden können und damit dessen kompletter Lösungsraum beschrieben werden kann. Der entwickelte *NetworkReducer*-Algorithmus liefert somit einen wichtigen Baustein für die automatisierte Modellreduktion stöchiometrischer Netzwerke.

Mit den drei beschriebenen Modellen *SyneMin*, *SynePruned* und *SyneGS*, können nun die Analysen zur Identifikation von Interventionsmöglichkeiten zur Optimierung der Produktion von Biokraftstoffen auf drei verschiedenen Netzwerkskalen gestartet werden. Ausgehend von dem Minimalmodell, *SyneMin*, sollen grundlegende Prinzipien identifiziert werden, welche zur wachstumsgekoppelten Biokraftstoffproduktion unter phototrophen Wachstumsbedingungen führen. Aufbauend auf den Erkenntnissen aus *SyneMin*, sollen diese Analysen skalenübergreifend ausgeweitet, ausgewertet und verglichen werden.

5 Neue Prinzipien für das Metabolic Engineering von Cyanobakterien zur Biokraftstoffproduktion

Dieses Kapitel beruht auf der folgenden Publikation:

[66] Erdrich, P., Knoop, H., Steuer, R. & Klamt, S. (2014) Cyanobacterial Biofuels: New Insights and Strain Design Strategies Revealed by Computational Modeling. *Microbial Cell Factories* **13**:128

Im folgenden Kapitel wird nach einer einleitenden Motivation (5.1), die Balance der ATP- und NADPH-Produktion in der Photosynthese erläutert (5.2), da diese Balance für verschiedene zelluläre Prozesse von entscheidender Bedeutung ist. Nachfolgend wird die Biokraftstoffproduktion in *Synechocystis* sp. PCC 6803 am Beispiel von Ethanol skalenübergreifend analysiert (5.3). Beginnend mit dem Minimalmodell, *SyneMin*, werden dabei insbesondere Interventionsstrategien gesucht, welche die Produktion von Ethanol während des Zellwachstums erzwingen (Kopplung mit Wachstums, siehe Abschnitt 3.3.1). In Kapitel 5.4 wird diese Analyse auf weitere Biokraftstoffe in genomskaligen Stoffwechselmodellen verschiedener cyanobakterieller Spezies ausgedehnt.

5.1 Motivation

Die Notwendigkeit zur nachhaltigen Produktion von Kraftstoffen ist aufgrund der begrenzten Ressourcen fossiler Energieträger bereits seit langer Zeit gegeben. Wie bereits im einleitenden Kapitel 1 dargestellt, erzeugen Cyanobakterien seit geraumer Zeit Aufmerksamkeit, da diese in der Lage sind, oxygene Photosynthese zu betreiben und atmosphärisches CO₂ in Kohlenwasserstoffe umzuwandeln. Sie besitzen die Eigenschaft in extremen Umgebungsbedingungen zu wachsen und sind bzgl. ihres Stoffwechsel sehr vielseitig. Ein großer Vorteil ist dabei, dass sie keine landwirtschaftlich nutzbare Fläche

benötigen. Aufgrund dessen werden sie vermehrt in biotechnologischen Anwendungen eingesetzt (u. a. zur Biokraftstoffproduktion).

Der Nachteil ist allerdings, dass die meisten Anwendungen für die nachhaltige Produktion mit Cyanobakterien durch eine zu geringe Produktivität charakterisiert sind [210]. Im folgenden Kapitel werden verschiedene modellbasierte Strategien präsentiert, um die Biokraftstoffproduktion in Cyanobakterien zu optimieren. Insbesondere soll die wachstumsgekoppelte Biokraftstoffsynthese unter phototrophen Bedingungen analysiert werden. Denn trotz der Tatsache, dass für phototrophe Organismen wie *Synechocystis* oder andere Cyanobakterien eine Reihe genomskaliger Netzwerkmodelle vorhanden sind [118, 178, 211], gibt es bisher keine systematische Untersuchung, ob und wie das Kopplungsprinzip für den phototrophen Metabolismus funktioniert [178, 261]. Es werden neue KO-Strategien präsentiert, welche die wachstumsgekoppelte Biokraftstoffsynthese unter phototrophen Bedingungen erzwingen.

Für die Beschreibung des phototrophen cyanobakteriellen Stoffwechsels wurde als Ausgangspunkt ein genomskaliges Modell von *Synechocystis*, *SyneGS*, genutzt (Knoop et al. [134], Kapitel 3.5). Wie in Kapitel 4.4 beschrieben, wurde *SyneGS* mit Hilfe des *NetworkReducers* auf zwei Modelle kleinerer Netzwerkskalen reduziert (ein Modell mittlerer Größe, *SynePruned*; ein Minimalmodell, *SyneMin*). Ausgehend von dem Minimalmodell, *SyneMin*, sollen im Folgenden auf allen drei Netzwerkskalen Strategien zur Optimierung der Produktion von Biokraftstoffen berechnet, analysiert und diskutiert werden.

5.2 Die Balance der ATP- und NADPH-Produktion in der Photosynthese

Wie bereits in Kapitel 2.2 beschrieben, bildet die Photosynthese das zentrale Element des phototrophen Stoffwechsels der Cyanobakterien. Sie dient einerseits zur Energiegewinnung in Form von ATP und andererseits der Erzeugung von Reduktionsäquivalenten in Form von NADPH. Die Abbildung 2.2 auf Seite 12 zeigt eine idealisierte und allgemeine Darstellung der Photosynthese.

Es ist bereits diskutiert worden, dass die Elektronen, die aus der Wasserspaltung frei werden, verschiedene Wege durch die photosynthetische Elektronentransportkette nehmen können. Der lineare Elektronenfluss (siehe Abbildung 5.1) liefert dabei ein festes Verhältnis von ATP zu NADPH. Ausgehend davon, dass die ATPase 14 Protonen in das Cytoplasma strömen lassen muss, um drei ATP zu gewinnen, ergibt sich ein ATP/NADPH Verhältnis von 9/7 bzw. 1,28 [11, 223, 258].

5.2 Die Balance der ATP- und NADPH-Produktion in der Photosynthese

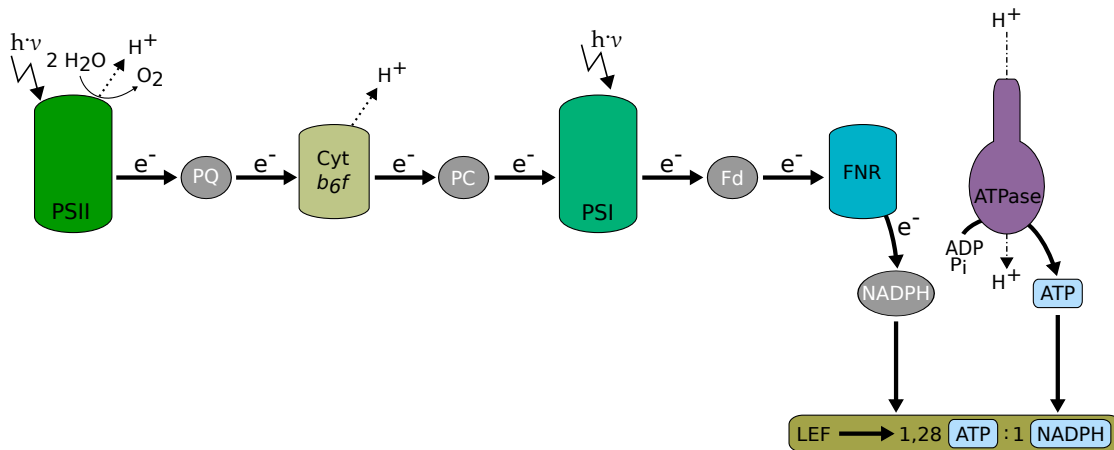


Abbildung 5.1: Linearer Elektronenfluss (LEF) durch die photosynthetische Elektronentransportkette (ETK). Gepunktete Linien stellen den Protonenübergang in das Thylakoidlumen dar. Die gepunktet-gestrichelte Linie zeigt den Protoneneinstrom aus dem Thylakoidlumen in das Cytoplasma an der ATPase. PSII: Photosystem II, PQ: Plastoquinon, Cyt *b6f*: Cytochrom-*b6f*-Komplex, PC: Plastocyanin, PSI: Photosystem I, Fd: Ferredoxin, FNR: Ferredoxin-NADPH-Oxidoreduktase.

Allerdings ist für die Fixierung von CO_2 im Calvin Zyklus ein Verhältnis von ATP zu NADPH von 3/2 notwendig. Damit dieses Verhältnis erreicht werden kann, gibt es verschiedene Wege für Elektronen in der photosynthetischen ETK, die netto ATP produzieren [12, 138]. Hier werden, in Anlehnung an Allen [11], diese Wege wie folgt klassifiziert:

- zyklischer Elektronenfluss (Ferredoxin-(Plasto)quinon-Reduktase, NADPH Dehydrogenase) und
- pseudozyklischer Elektronenfluss (Mehler-Reaktion und respiratorische Reaktionen).

Während des zyklischen Elektronentransports über die Ferredoxin-(Plasto)quinon-Reduktase (FQR), werden die Elektronen des reduzierten Ferredoxins auf das Plastoquinon übertragen [100]. Dabei werden keine Protonen gepumpt. Ein weiterer Vertreter des zyklischen Elektronentransports ist die NADPH Dehydrogenase (NDH). Diese nutzt das gebildete NADPH, um die Elektronen auf das PQ zu übertragen. In diesem Fall werden zusätzlich Protonen transportiert. Dieser Transport verschiebt das ATP/NADPH Verhältnis zu Gunsten der Energiegewinnung. Es gibt viele verschiedene NDH, die sich bzgl. ihres Elektronendonators und der Menge an Protonen, die sie pumpen können, un-

terscheiden. Die Gemeinsamkeit ist, dass die NDH Elektronen wiederverwenden können und damit einen zyklischen Elektronenfluss erzeugen.

Ein pseudozyklischer Elektronenflussmechanismus wird durch die Mehler-Reaktion repräsentiert [22]. Es werden die Elektronen des Ferredoxins genutzt, um diese auf Sauerstoff zu übertragen, wobei ein Superoxidradikal entsteht. Dieses Radikal ist eine reaktive Sauerstoffspezies (ROS) und kann zu Zellschäden führen (hohe Mengen von ROS führen zu oxidativem Stress). Dieses Radikal kann unter Nutzung der Superoxiddismutase zu Wasserstoffperoxid (H_2O_2) umgewandelt werden, das wiederum zur Wasserbildung genutzt werden kann. Damit entsteht um die Mehler-Reaktion ein Wasser-Wasser Zyklus [19]. Ein weiteres Beispiel ist die Cytochrom-c-Oxidase. Diese nutzt die Elektronen des Plastocyanins, um Sauerstoff zu Wasser zu reduzieren. Die Elektronen können folglich nicht weiter durch die photosynthetische ETK fließen. Zudem werden Protonen über die Membran gepumpt. Diese Oxidase hat demnach zwei Funktionen, die das ATP/NADPH Verhältnis verschieben (Abzug der Elektronen und Pumpen von Protonen über die Membran). Da die Elektronen genutzt werden, um ein Molekül Sauerstoff zu zwei Molekülen Wasser zu reduzieren, kann in diesem Fall wiederum von einer Art Wasser-Wasser Zyklus gesprochen werden. Denn die genutzten Elektronen sind bei der Wasserspaltung frei geworden und werden wieder auf die Wassermoleküle übertragen. Die Cytochrom-bd-Oxidase hat einen ähnlichen Effekt. Allerdings werden dabei die Elektronen (und Protonen) des Plastoquinols genutzt. Die Abbildung 5.2 fasst die genannten Möglichkeiten des alternativen Elektronenflusses (zyklisch und pseudozyklisch) in der photosynthetischen ETK graphisch zusammen.

Die dargestellten „Wiederverwendungsmöglichkeiten“ der Elektronen dienen dazu, dass ATP/NADPH Verhältnis zugunsten des ATPs zu verschieben, um den Bedarf an Energie für zellinterne Prozesse wie z. B. die Fixierung von CO_2 zu decken. Das Verhältnis von ATP zu NADPH, das notwendig ist, um Biomasse zu synthetisieren, ist in diesem Kontext sehr interessant. Dieses Verhältnis ist nicht einfach zu quantifizieren. Mit dem stöchiometrischen Modell *SyneGS* ist es möglich, zumindest eine untere Grenze zu identifizieren. Dafür werden ATP und NADPH als externe Metabolite dem Netzwerk zur Verfügung gestellt und anschließend die Biomasse unter Ausschluss der Photosynthese und der Respiration maximiert. Damit kann eine untere Grenze des notwendigen ATP/NADPH Verhältnisses von 1,51 berechnet werden. Es ist anzunehmen, dass das tatsächliche Verhältnis größer sein wird, da es verschiedene ATP-verbrauchende Prozesse gibt, die in dem Modell nicht berücksichtigt sind. Das notwendige Verhältnis für die Produktion von Biomasse ist demnach größer als das für die Fixierung von CO_2 im Calvin Zyklus.

5.2 Die Balance der ATP- und NADPH-Produktion in der Photosynthese

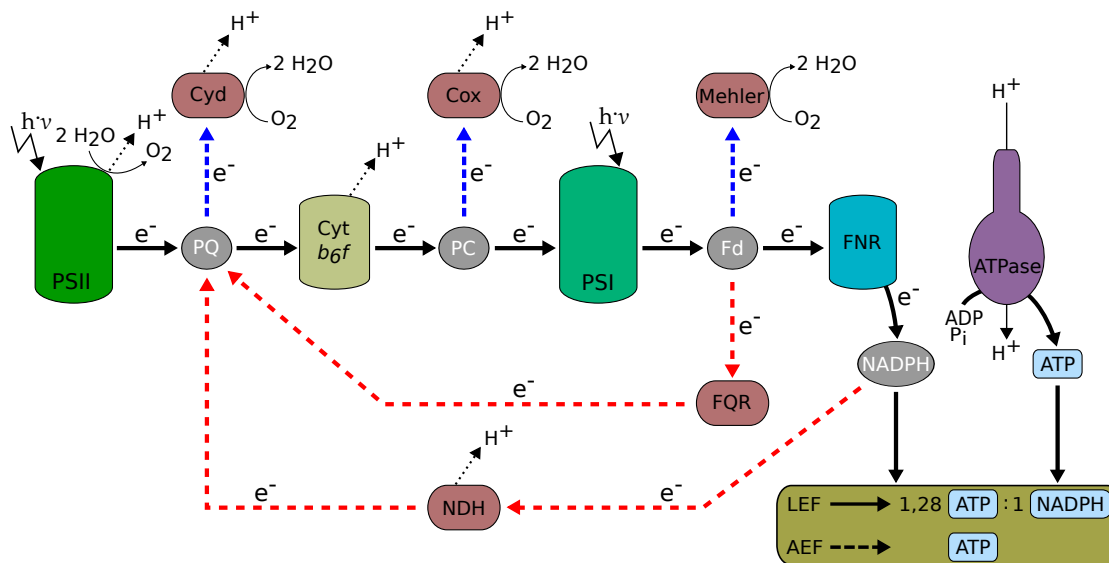


Abbildung 5.2: Elektronenflüsse in der photosynthetischen Elektronentransportkette. Die gestrichelten Linien stellen verschiedene alternative (zyklische und pseudozyklische) Elektronenflüsse dar: (---) zyklischer (ZEF), (- - -) pseudozyklischer (PEF) Elektronenfluss. Der alternative Elektronenfluss ermöglicht es, dass ATP unabhängig von NADPH generiert werden kann. Gepunktete Linien stellen den Protonenübergang in das Thylakoidlumen dar. PSII: Photosystem II, PQ: Plastoquinon, Cyd: Cytochrom-bd-Oxidase, Cyt *b6f*: Cytochrom-*b6f*-Komplex, PC: Plastocyanin, Cox: Cytochrom-c-Oxidase, PSI: Photosystem I, Fd: Ferredoxin, Mehler: Mehler Reaktion, FQR: Ferredoxin-Plastoquinon-Reduktase, FNR: Ferredoxin-NADPH-Oxidoreduktase, NDH: NAD(P)H Dehydrogenase, LEF: linearer Elektronenfluss.

Das aus der Photosynthese generierte Verhältnis von ATP zu NADPH hat einen entscheidenden Einfluss auf das Verhalten des Organismus. Dieses Verhältnis ist für photosynthetische Organismen hinsichtlich der metabolischen Selbstregulation entscheidend, aber auch für die Synthese verschiedener Produkte [137]. Am Beispiel der Produktion von Ethanol soll dies im Folgenden dargestellt werden. Nach Deng et al. ist für eine erfolgreiche Produktion von Ethanol die Integration einer Pyruvatdecarboxylase und einer Alkoholdehydrogenase notwendig [56]. Diese Reaktionen gewährleisten einen kurzen Weg von der CO_2 Fixierung durch RuBisCO bis hin zum Ethanol (siehe Abbildung 5.3).

Der Bedarf an Energie und Reduktionsäquivalenten zur Produktion von einem Molekül Ethanol beträgt dann sieben ATP und sechs NADPH [118]. Dies entspricht einem ATP/NADPH-Verhältnis von 1,17. Die Unterschiede zwischen den ATP/NADPH-Verhältnissen (1,17 Ethanolproduktion/1,28 LEF/1,5 CO_2 -Fixierung/1,51 Biomassesyntese) stellen nun ein mögliches Ziel für Interventionen dar, um die Ethanolproduktion

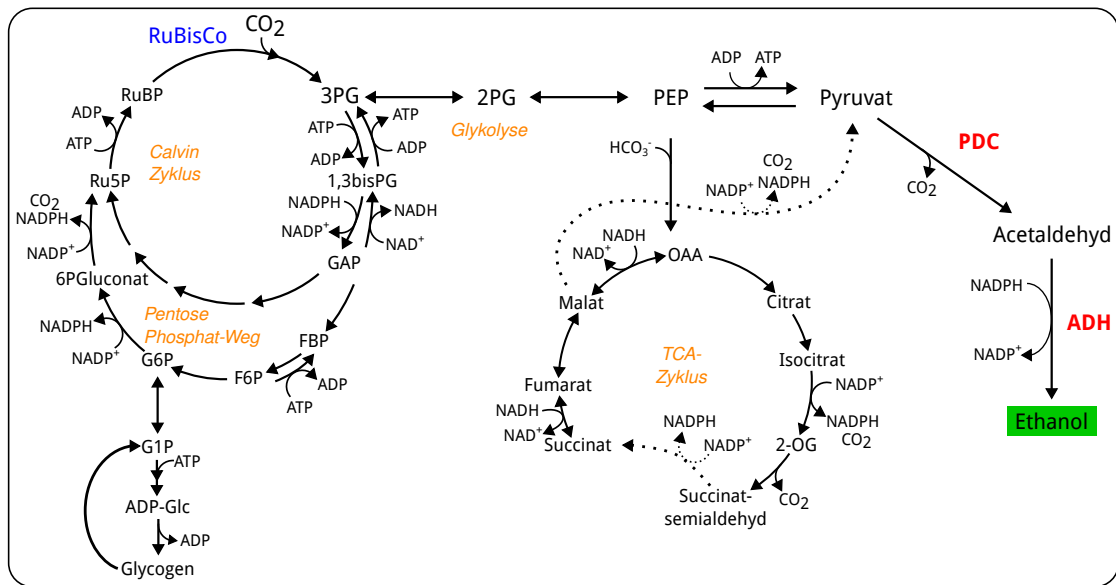


Abbildung 5.3: Produktionsweg für Ethanol. Graphik angelehnt an [3].

zu steigern. Theoretisch gibt es vier Möglichkeiten das Verhältnis von ATP zu NADPH zu verkleinern und damit die Ausbeute von Ethanol im Verhältnis zur Biomasse zu erhöhen. Dies kann durch:

- i) Senkung der ATP Produktion,
- ii) Steigerung es ATP Konsums,
- iii) Steigerung der NADPH Produktion,
- iv) oder Senkung des NADPH Konsums

erreicht werden. Da die alternativen Elektronentransportwege maßgeblich das ATP/NADPH-Verhältnis zugunsten von ATP verändern, ist zu erwarten, dass diese Elektronentransportwege ein zentrales Ziel von Interventionsstrategien sein werden.

5.3 Ethanol-Produktion in *Synechocystis* sp. PCC 6803

Beginnend mit dem in Kapitel 4.4 vorgestellten Minimalmodell von *Synechocystis*, *Syn-eMin*, sollen nachfolgend Interventionsstrategien für die Ethanolproduktion in dem Modell mittlerer Größe, *Syn-ePruned*, und dem genomskaligen Modell, *Syn-eGS*, identifiziert, ausgewertet und verglichen werden.

5.3.1 Analyse des Minimalmodells von *Synechocystis* (*SyneMin*)

Um zunächst ein grundlegendes Verständnis der Biokraftstoffproduktion in *Synechocystis* zu erhalten, wird ein stark komprimiertes Zentralstoffwechselmodell von *Synechocystis* analysiert. Dieses Modell, *SyneMin* ist mit Hilfe des in Kapitel 4 präsentierten *NetworkReducer*-Algorithmus reduziert und in Kapitel 4.4 vorgestellt worden. *SyneMin* ist in der Lage, die CO₂-Fixierung durch RuBisCO, unspezifischen ATP-Verbrauch, Wachstum und Ethanolproduktion abzubilden. Andere Produkte, sowie der Nachtmetabolismus werden in diesem Minimalmodell nicht berücksichtigt. Das Modell besteht aus 37 Reaktionen und 38 internen Metaboliten (siehe Anhang A.5). Der Rang der stöchiometrischen Matrix ist 33. Damit ergibt sich ein Freiheitsgrad von vier. Wie in Kapitel 4.4 gefordert wurde, sind alle geschützten Reaktionen Bestandteil des reduzierten Modells (insgesamt 26).

Die BSR ist stark komprimiert und zieht die Metabolite des Calvin Zyklus ab (Präkursoren, ATP und NADPH), um daraus Biomasse zu synthetisieren. Die stöchiometrischen Koeffizienten dieser Metabolite sind recht groß, da diese sowohl als direkte Präkursoren für die Biomasse dienen, als auch den indirekten Bedarf für andere Präkursoren beschreiben. Dennoch kann der *NetworkReducer* garantieren, dass die BSR des reduzierten Modells konsistent ist mit der des genomskaligen Modells. Das heißt, dass die Gesamtmenge an benötigtem Kohlenstoff, ATP und Reduktionsäquivalenten identisch ist und daher die gleiche maximale Wachstumsrate möglich ist.

In diesem Minimalmodell können insgesamt zehn Elementarmoden gefunden werden, die sowohl ausschließlich Wachstum (EM_{1,2,3}) oder Ethanolproduktion (EM_{5,6,7}) als auch die gekoppelte Produktion von Ethanol und Biomasse (EM_{8,9,10}) beschreiben können. (EM₄) repräsentiert die netto ATP-Synthese, die durch den unspezifischen ATP-Verbrauch (für z. B. Maintenance-Prozesse) ausbalanciert wird. Tabelle 5.1 und Abbildung 5.4 zeigen die EM.

Tabelle 5.1: EM des stark komprimierten Zentralstoffwechselmodells von *Synechocystis*, *SyneMin*.

Reaktion	EM1	EM2	EM3	EM4	EM5	EM6	EM7	EM8	EM9	EM10
Photonenaufnahme	100	100	100	100	100	100	100	100	100	100
Ethanolproduktion	0	0	0	0	4,17	4,17	4,17	2,72	2,72	2,72
Biomasse-synthesereaktion	0,19	0,19	0,19	0	0	0	0	0,08	0,08	0,08
NDH (AEF)	3,06	3,06	3,06	50,00	0	0	0	0	0	0
ATP-Verbrauch	0	0	0	85,72	2,98	2,98	2,98	0	0	0

Tabelle 5.1 zeigt, dass die Reaktion des AEF (NDH) funktional sein muss, sobald ausschließlich Biomasse synthetisiert wird (EM_{1,2,3}). Denn in diesem Fall muss ein ATP/NADPH Verhältnis erzeugt werden, das größer ist als das des LEF. Für den Fall, dass nur Ethanol produziert wird (EM_{5,6,7}), ist der alternative Elektronenfluss (NDH) nicht aktiv, da der LEF ausreichend Energie und Reduktionsäquivalente zur Verfügung stellt ($\text{ATP/NADPH}_{\text{LEF}} = 1,28$; $\text{ATP/NADPH}_{\text{Ethanolproduktion}} = 1,17$). In diesem Fall entsteht ein ATP-Überschuss, der im Modell durch den unspezifischen ATP-Verbrauch ausgeglichen wird.

Interessant ist, dass im Fall der gekoppelten Produktion von Ethanol und Biomasse diese Reaktion des alternativen Elektronenflusses (NDH) ebenso nicht aktiv ist. Wird die Funktionalität der NDH unterdrückt, sind (EM_{1,2,3,4}) nicht mehr realisierbar (siehe Tabelle 5.1 und Abbildung 5.4). Nur die (EM_{5,6,7,8,9,10}) sind dann noch funktional. Diese EM produzieren alle Ethanol. Insbesondere bedeutet das, dass der KO des alternativen Elektronenflusses auch die wachstumsgekoppelte Ethanolproduktion erzwingt. Das Minimalmodell *SyneMin* zeigt damit eine erste KO-Strategie, wie die wachstumsgekoppelte Ethanolproduktion erreicht werden kann.

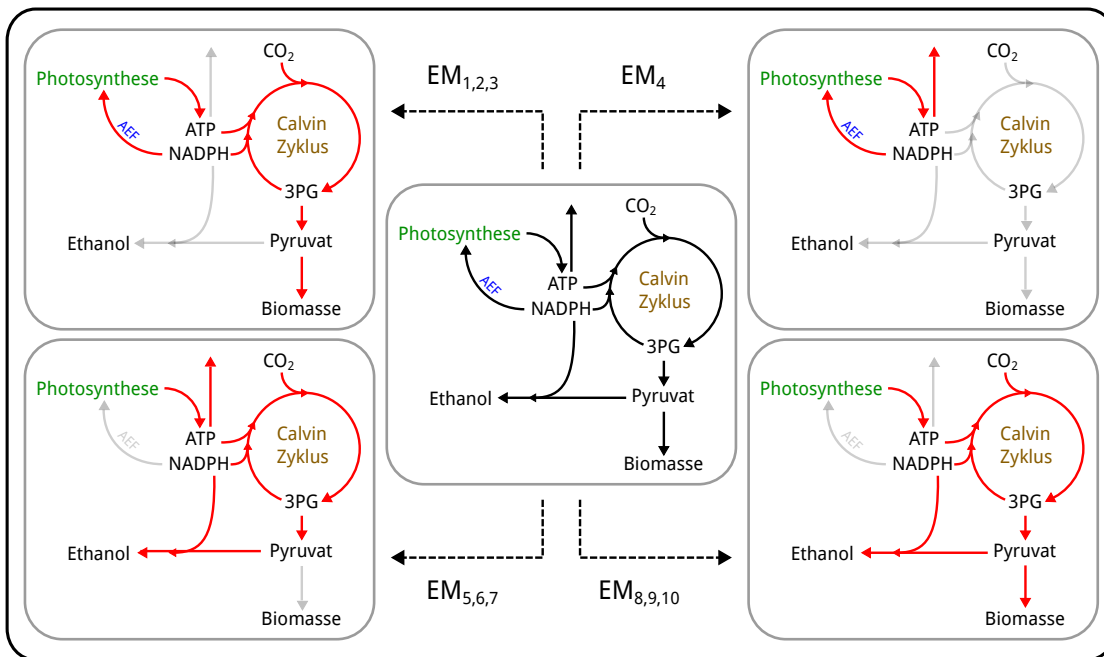


Abbildung 5.4: Die EM von *SyneMin* (siehe auch Tabelle 5.1).

5.3.2 Analyse des *SynePruned*-Modells

5.3.2.1 CASOP-Analyse

Wie in Kapitel 3.3.3 erklärt, dient die CASOP Methodik dazu, Überexpressions- und Knockoutkandidaten zu identifizieren. Der Ansatz beruht auf Elementarmodi und berücksichtigt dabei eine Verschiebung vom Wildtypwachstum (ausschließlich Biomassensynthese) hin zu gekoppelter Synthese von Produkt und Biomasse. Dabei wird ein spezielles Maß (Z-Score) ermittelt, das den Wertebereich $[-1, 1]$ annehmen kann. Sobald dieser Score kleiner null ist, weist dieser einen potentiellen Knockoutkandidaten aus. Positive Werte beschreiben potentielle Überexpressionskandidaten.

Die folgenden Analysen und Ergebnisse beruhen auf der Annahme phototrophen Wachstums, ATP-unabhängiger Hydrogencarbonataufnahme und Nitrat als Stickstoffquelle [133]. Als einziges Produkt wird Ethanol betrachtet. Um ein eher Produktausbeute-bezogenes Szenario zu simulieren, wird der Parameter k_C auf fünf gesetzt (der Exponent in Gleichung (3.18)). Die folgenden Tabellen zeigen die jeweils prominentesten Vertreter der durch CASOP identifizierten Überexpressions- bzw. Knockoutkandidaten (die vollständige Tabelle kann dem Anhang (A.2) entnommen werden).

Die Reaktionsgleichungen der Reaktionen, die im Folgenden diskutiert werden, sind in der Tabelle 5.3 dargestellt.

Tabelle 5.2: Die durch CASOP identifizierten Top-Knockout- und Top-Überexpressionskandidaten für die Ethanolproduktion in *SynePruned*. Die Reaktionsgleichungen sind in Tabelle 5.3 zu finden.

Knockoutkandidaten		Überexpressionskandidaten	
ID (kurz)	CASOP-Wert	ID (kurz)	CASOP-Wert
R1	-0,189	R29	1,000
R3	-0,189	R30	1,000
R7	-0,142	R31	1,000
R4	-0,142	R32	0,298
R2	-0,121	R33	0,173
R5	-0,118	R34	0,128
R6	-0,118	R35	0,096
		R36	0,096
		R37	0,096
		R38	0,096

Tabelle 5.2 zeigt die Top-Knockoutkandidaten und die dazugehörigen CASOP-Werte für die phototrophe Ethanolproduktion mit *Synechocystis*. Alle gefundenen KO-Ziele stehen im Zusammenhang mit der Photosynthese bzw. dem Elektronenfluss um die photosynthetische ETK. Die durch CASOP identifizierten Reaktionen lassen sich wie folgt gruppieren.

- (I) Zyklische Reaktionen (R7 (FQR), R1 (NDH1), R2 (CO₂ assoziierte NDH1)), die reduziertes Ferredoxin bzw. NADPH nutzen, um Plastoquinon zu reduzieren.
- (II) Pseudozyklische Reaktionen, wie die Mehler-Reaktion (R5, R6) und respiratorische Reaktionen (R3 (C-terminale Oxidase) und R4 (Cyd-bd-Oxidase), die einen Wasser-Wasser Zyklus bilden.

CASOP suggeriert somit, die zyklischen und pseudozyklischen Elektronenflüsse um die photosynthetische ETK zu blockieren. Dieser Eingriff bewirkt, dass das ATP/NADPH

Verhältnis reduziert wird. CASOP hat damit das gleiche Ziel identifiziert, wie die EM-Analyse des *SyneMin*. Im Vergleich zu *SyneMin* enthält das Modell *SynePruned* mehr Reaktionen, die alternative (zyklische und pseudozyklische) Elektronenflüsse ermöglichen. Deshalb werden mehr potentielle Knockoutkandidaten von CASOP identifiziert.

Bezüglich der Überexpressionskandidaten zeigt sich, dass drei Reaktionen den Wert eins annehmen (siehe Tabelle 5.2). Diese drei Reaktionen sind direkt an der Produktion von Ethanol beteiligt und sind kein Bestandteil der EM-Menge für den Fall, dass ausschließlich Biomasse produziert wird. Allerdings sind diese Reaktionen in allen EM enthalten, die Ethanol produzieren. Daraus folgt, dass diese den Wert eins annehmen müssen. Es handelt sich dabei um die Pyruvatdecarboxylase (R31), die Alkoholdehydrogenase (R30) und die Exportreaktion von Ethanol (R29). Die Reaktion R32 (Phosphoenolpyruvat zu Pyruvat) bildet einen Präkursor für Acetaldehyd (und damit für Ethanol) und CASOP suggeriert deshalb, diese Reaktion zu überexprimieren.

Die Reaktion R33 beschreibt den Export von CO_2 und es erscheint auf den ersten Blick eigenartig, dass diese Reaktion überexprimiert werden soll. Allerdings ist diese Reaktion für die maximale Ausbeute von Ethanol bezogen auf die Photonenaufnahme essentiell (und damit in den EM enthalten, die eine hohe Ausbeute von Ethanol haben). Das bedeutet, dass ohne den Export von CO_2 die Ausbeute von Ethanol geringer ist. Dies beruht auf der Art der Modellierung der CO_2 -Fixierung im stöchiometrischen Modell *SynePruned*. CO_2 wird in Form von Hydrogencarbonat in die Zelle aufgenommen. Dieses Hydrogencarbonat wird als solches in das Carboxysom transportiert und dort über H_2CO_3^- in CO_2 umgewandelt, das dann durch RuBisCO fixiert werden kann. Das CO_2 -Molekül kann als solches nicht in das Carboxysom gelangen. CO_2 , das z. B. bei der Decarboxylierung von Pyruvat entsteht, kann über zwei Wege verbraucht werden. Es kann (1) aus der Zelle exportiert werden (R33). Oder das CO_2 Molekül wird (2) unter Nutzung der CO_2 -abhängigen NDH (R2) in Hydrogencarbonat umgewandelt (Verbrauch von NADPH). Dieser zusätzliche Verbrauch von NADPH führt zu einer geringeren Ausbeute von Ethanol. Um dies zu überprüfen, wird die PDC im Modell (R31) derart modifiziert, dass diese statt CO_2 HCO_3^- freisetzt. Durch diese Veränderung kann die maximale Ausbeute von Ethanol bezogen auf die Photonenaufnahme erreicht werden, ohne Beteiligung der Reaktion R33 (der CASOP-Wert verringert sich deutlich).

Die verbleibenden Überexpressionskandidaten eint ein Mechanismus, der zunächst ebenfalls wenig intuitiv erscheint. So schlägt CASOP die Erhöhung des ATP-Verbrauchs vor (R35). Dies suggeriert, dass sich dieser Verbrauch vorteilhaft für eine erhöhte Produktion von Ethanol auswirkt. Werden die verbleibenden Reaktionen der Tabelle 5.2 betrachtet, fällt auf, dass diese Reaktionen ebenso zu einer Verschwendung von ATP

5 Metabolic Engineering von Cyanobakterien zur Biokraftstoffproduktion

führen. Reaktion R36 und R37 bilden einen Zyklus um Cyanophycin, der ATP verbraucht. Die Phosphofructo-1-Kinase (R34) und die Sedoheptulose-7-Phosphat-Kinase (R38) dienen ebenfalls dem erhöhten Verbrauch von ATP. Auch der Export von CO_2 kann neben dem genannten Effekt zusätzlich einen ATP verbrauchenden Zyklus durch eine indirekte Umwandlung von Hydrogencarbonat in CO_2 erzeugen.

Die durch CASOP vorgeschlagenen Überexpressionskandidaten können als komplementäre Ansätze zu den KO-Strategien verstanden werden. In diesem Fall wird statt des KOs der zusätzlichen Produktion von ATP, der erhöhte Verbrauch von ATP favorisiert. Beide Ansätze vereinen erneut das Ziel der Reduktion des ATP/NADPH Verhältnisses.

5.3 Ethanol-Produktion in *Synechocystis* sp. PCC 6803

Tabelle 5.3: Liste ausgewählter Reaktionen des *SyneGS*- und *SynePruned*-Modells.
tll: Thylakoidlumen, cax: Carboxysomen

ID	Reaktion
R1	$\text{NADPH} + 5 \text{H}^+ + \text{PQ} \Rightarrow \text{NADP}^+ + \text{PQH}_2 + 4 \text{H}_{\text{tll}}^+$
R2	$\text{NADPH} + 4 \text{H}^+ + \text{PQ} + \text{H}_2\text{O} + \text{CO}_2 \Rightarrow \text{NADP}^+ + \text{PQH}_2 + 4 \text{H}_{\text{tll}}^+ + \text{HCO}_3^-$
R3	$4 \text{red. Plastocyanin} + \text{O}_2 + 8 \text{H}^+ \Rightarrow 4 \text{ox. Plastocyanin} + 4 \text{H}_{\text{tll}}^+ + 2 \text{H}_2\text{O}$
R4	$2 \text{PQH}_2 + \text{O}_2 + 4 \text{H}^+ \Rightarrow 2 \text{PQ} + 4 \text{H}_{\text{tll}}^+ + 2 \text{H}_2\text{O}$
R5	$\text{O}_2 + \text{red. Ferredoxin} \Rightarrow \text{ox. Ferredoxin} + \text{O}_2^-$
R6	$2 \text{O}_2^- + 2 \text{H}^+ \Rightarrow \text{H}_2\text{O}_2 + \text{O}_2$
R7	$\text{PQ} + 2 \text{red. Ferredoxin} + 2 \text{H}^+ \Rightarrow 2 \text{ox. Ferredoxin} + \text{PQH}_2$
R8	$\text{D-Ribulose 1,5-Bisphosphat} + \text{O}_2 \Rightarrow 3\text{-Phospho-D-Glycerat} + 2\text{-Phosphoglycolat}$
R9	$\text{O}_2 \Rightarrow \text{O}_{2\text{cax}}$
R10	$2\text{-Phosphoglycolat} + \text{H}_2\text{O} \Rightarrow \text{Glycolat} + \text{Orthophosphat}$
R11	$2\text{-Phosphoglycolat}_{\text{cax}} \Rightarrow 2\text{-Phosphoglycolat}$
R12	$\text{NADPH} + \text{NAD}^+ \Leftrightarrow \text{NADP}^+ + \text{NADH}$
R13	$\text{ATP} + \text{Acetat} \Leftrightarrow \text{ADP} + \text{Acetyl-Phosphat}$
R14	$\text{Glyceronphosphat} + \text{NADH} + \text{H}^+ \Leftrightarrow \text{sn-Glycerol-3-Phosphat} + \text{NAD}^+$
R15	$\text{ATP} + \text{Glycerol} \Rightarrow \text{ADP} + \text{sn-Glycerol-3-Phosphat}$
R16	$\text{Glycerol} + \text{NADP}^+ \Leftrightarrow \text{D-Glyceraldehyd} + \text{NADPH} + \text{H}^+$
R17	$\text{D-Glyceraldehyd} + \text{NAD}^+ + \text{H}_2\text{O} \Leftrightarrow \text{D-Glycerat} + \text{NADH} + \text{H}^+$
R18	$5,10\text{-Methylentetrahydrofolat} + \text{Glycin} + \text{H}_2\text{O} \Leftrightarrow \text{Tetrahydrofolat} + \text{L-Serin}$
R19	$\text{Glycin} + \text{Tetrahydrofolat} + \text{NAD}^+ \Rightarrow 5,10\text{-Methylentetrahydrofolat} + \text{NH}_3 + \text{CO}_2 + \text{NADH} + \text{H}^+$
R20	$\text{D-Glycerat} + \text{NADP}^+ \Leftrightarrow \text{Hydroxypyruvat} + \text{NADPH} + \text{H}^+$
R21	$\text{L-Serin} + \text{Glyoxylat} \Leftrightarrow \text{Hydroxypyruvat} + \text{Glycin}$
R22	$\text{Glycin} + 2\text{-Oxoglutarat} \Leftrightarrow \text{Glyoxylat} + \text{L-Glutamat}$
R23	$3\text{-Phospho-D-Glycerat} + \text{NAD}^+ \Leftrightarrow 3\text{-Phosphonooxypyruvat} + \text{NADH} + \text{H}^+$
R24	$\text{O-Phospho-L-serin} + 2\text{-Oxoglutarat} \Leftrightarrow 3\text{-Phosphonooxypyruvat} + \text{L-Glutamat}$
R25	$\text{O-Phospho-L-serin} + \text{H}_2\text{O} \Rightarrow \text{L-Serin} + \text{Orthophosphat}$
R26	$(\text{S})\text{-Malat} + \text{NAD}^+ \Leftrightarrow \text{Oxaloacetat} + \text{NADH} + \text{H}^+$
R27	$\text{D-Fructose-6-Phosphat} + \text{Orthophosphat} \Rightarrow \text{Acetyl-Phosphat} + \text{D-Erythrose-4-Phosphat} + \text{H}_2\text{O}$
R28	$\text{D-Xylulose-5-Phosphat} + \text{Orthophosphat} \Rightarrow \text{Acetyl-Phosphat} + \text{D-Glyceraldehyd-3-Phosphat} + \text{H}_2\text{O}$
R29	$\text{Ethanol} \Rightarrow \text{Ethanol}_{\text{ext}}$
R30	$\text{Acetaldehyd} + \text{NADPH} + \text{H}^+ \Rightarrow \text{Ethanol} + \text{NADP}^+$
R31	$\text{Pyruvat} \Rightarrow \text{CO}_2 + \text{Acetaldehyd}$
R32	$\text{ADP} + \text{Phosphoenolpyruvat} \Rightarrow \text{ATP} + \text{Pyruvat}$
R33	$\text{CO}_2 \Rightarrow \text{CO}_{2,\text{ext}}$
R34	$\text{ATP} + \text{D-Fructose-6-Phosphat} \Rightarrow \text{ADP} + \text{D-Fructose-1,6-Bisphosphat}$
R35	$1 \text{ATP} + \text{H}_2\text{O} \Rightarrow 1 \text{ADP} + \text{Orthophosphat}$
R36	$\text{Cyanophycin} + 2 \text{H}_2\text{O} \Rightarrow 1 \text{L-Aspartat} + 1 \text{L-Arginin} + \text{Cyanophycin-Polymer}$
R37	$1 \text{L-Aspartat} + 1 \text{L-Arginin} + 2 \text{ATP} + \text{Cyanophycin-Polymer} \Rightarrow \text{Cyanophycin} + 2 \text{Orthophosphat} + 2 \text{ADP}$
R38	$\text{ATP} + \text{Sedoheptulose-7-Phosphat} \Rightarrow \text{ADP} + \text{Sedoheptulose-1,7-Bisphosphat}$
R39	$\text{sn-Glycerol-3-Phosphat} + \text{PQ} \Rightarrow \text{Glyceronphosphat} + \text{PQH}_2$

5.3.2.2 Elementarmodenanalyse

Im Folgenden soll die Menge der EM im mittelskaligen Stoffwechselmodell von *Synechocystis* (*SynePruned*) weiter detailliert analysiert werden. Die vollständige Menge der EM für phototrophes Wachstum und Ethanolproduktion beläuft sich auf 100803 EM. Abbildung 5.5 projiziert jeden EM (blaue Kreise) auf seine Ausbeute bzgl. der Biomasse (in gTM/mmol Photonen) und bzgl. des Ethanols (in mmol Ethanol/mmol Photonen). Die

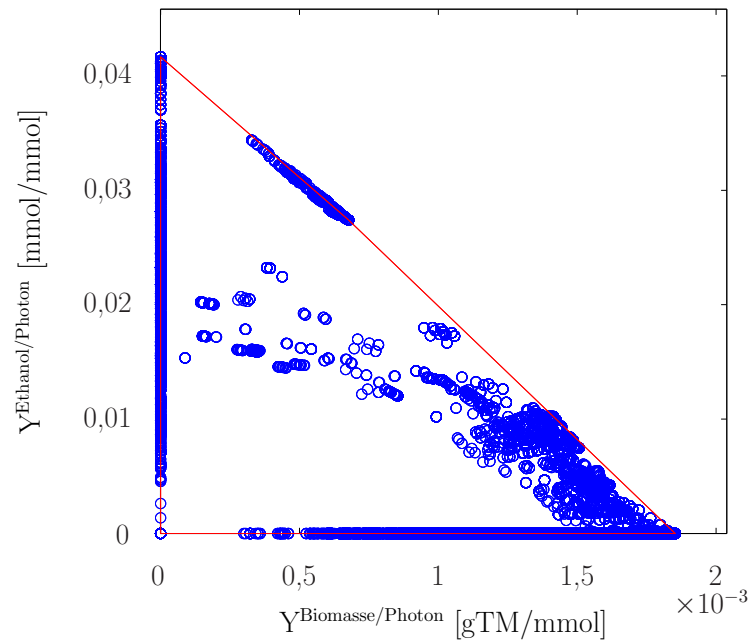


Abbildung 5.5: Ausbeutediagramm der EM für Ethanol- und Biomasse in *Syne-Pruned*. Die blauen Kreise entsprechen den EM. Die rote Linie beschreibt die konvexe Hülle der EM.

konvexe Hülle der EM wird durch die rote Linie beschrieben und diese gibt den kompletten Lösungsraum der Verhältnisse von Biomasse- und Ethanolausbeute an. EM die nur auf der Achse für die Biomasseausbeute liegen (X-Achse), beschreiben Wachstum ohne Ethanolproduktion und analog dazu, beschreiben die EM auf der Achse der Ethanolausbeute (Y-Achse) die ausschließliche Produktion von Ethanol. Aus dieser Abbildung kann die jeweilige optimale phototrophe Ausbeute bestimmt werden. Die maximale Biomasseausbeute beträgt $1,85 \cdot 10^{-3}$ gTM/mmol Photonen und die maximale Ethanolausbeute $4,167 \cdot 10^{-2}$ mmol/mmol Photonen. Letztere gibt somit auch die bekannte theoretisch maximale Ethanolausbeute wieder (24 Photonen pro Ethanolmolekül [118]).

5.3 Ethanol-Produktion in *Synechocystis* sp. PCC 6803

Das Ausbeutediagramm erlaubt weitere Rückschlüsse. Zum einen wird in Abbildung 5.5 deutlich, dass je höher die Ethanolausbeute ist, desto kleiner ist die optimale Biomasseausbeute. Dieser Zusammenhang führt zu der typischen dreieckigen Form der konvexen Hülle. Zum anderen zeigt diese Abbildung, dass es theoretisch möglich ist, die Ethanolproduktion an das Wachstum zu koppeln. Alle EM, die nicht auf den Achsen liegen, erfüllen diese Bedingung, da die Ausbeute der Biomasse und des Ethanols dann jeweils größer als null ist. Interessant sind die EM, die dabei eine hohe Ausbeute von Ethanol ermöglichen (dies geht mit einer kleineren Ausbeute von Biomasse einher). Die maximale garantierte Ausbeute von Ethanol unter phototrophen wachstumsgekoppelten Bedingungen beträgt 0,0344 mmol/mmol Photonen. Dieser Wert kann dem EM entnommen werden, der die größte Ethanolausbeute hat und zeitgleich eine Biomasseausbeute größer als null gewährleisten kann.

5.3.2.3 Analyse der constrained Minimal Cut Sets

Wie in Kapitel 3.3.2 beschrieben, können mit Hilfe der cMCSs Interventionen berechnet werden, die in der Lage sind, eine bestimmte metabolische Funktionalität zu unterdrücken und dabei eine erwünschte Netzwerkeigenschaften zu erhalten [94]. Die dargestellten Ergebnisse des Kapitels 5.3.1 und 5.3.2.1 liefern einen ersten Eindruck, welche Ziele interessant sein könnten, um eine wachstumsgekoppelte Ethanolproduktion zu erreichen. Zum Erzwingen einer solchen Kopplung werden im Folgenden cMCSs berechnet, die alle EM mit niedriger Ethanolausbeute löschen und mindestens einen Elementarmodus erhalten, der eine hohe Produktausbeute und eine minimale Biomasseausbeute garantieren kann.

Wie in Kapitel 3.3.2 dargestellt, beschreibt die Menge \mathbf{E} die Gesamtheit der Elementarmoden, die Menge \mathbf{T} die Target-EM (also jene, EM die gelöscht werden sollen) und die Menge \mathbf{D} die Desired-EM (die EM, von denen nach Anwendung der Intervention min. ein EM erhalten bleiben soll). Konkret wurde folgendes cMCS-Problem formuliert. Die Anforderungen an die Desired-EM sind, dass die minimale Produktausbeute $Y^{\text{Ethanol/Photon}} = 0,03$ mmol/mmol Photonen und die minimale Biomasseausbeute $Y^{\text{Biomasse/Photon}} = 0,0001$ gTM/mmol Photonen beträgt. Daraus ergibt sich, dass alle EM für die gilt, dass $Y^{\text{Ethanol/Photon}} < 0,03$ mmol/mmol Photonen, Target-EM sind. Dies kann wie folgt zusammengefasst werden:

$$\begin{aligned} \mathbf{T} &= \{\text{alle EM } E \in \mathbf{E} \mid Y_E^{\text{Ethanol/Photon}} < 0,03\}, \\ \mathbf{D} &= \{\text{alle EM } E \in \mathbf{E} \mid Y_E^{\text{Ethanol/Photon}} \geq 0,03 \wedge Y_E^{\text{Biomasse/Photon}} \geq 0,01 \cdot 10^{-2}\}. \end{aligned} \quad (5.1)$$

Unter diesen Bedingungen kann der in Abbildung 5.5 dargestellte Lösungsraum in verschiedene Bereiche unterteilt werden (Abbildung 5.6).

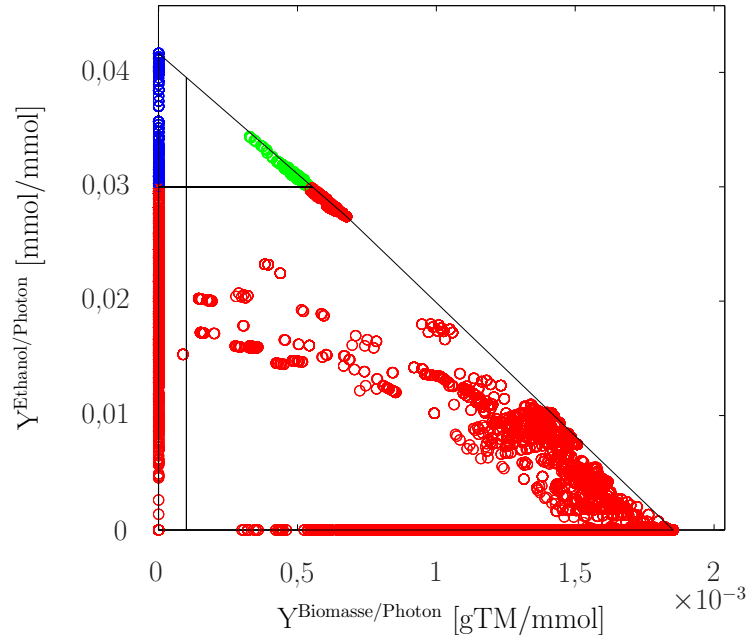


Abbildung 5.6: Eingeteiltes Ausbeutediagramm der EM für Ethanol- und Biomasse im *SynePruned*. Die roten Kreise entsprechen den Target-EM und die grünen Kreise den Desired-EM (die verbleibenden EM sind als blaue Kreise dargestellt und weder Target- noch Desired-EM). Die beschriebenen Grenzen für die Einteilung der EM sind als schwarze Linien innerhalb der konvexen Hülle dargestellt.

Die roten Kreise entsprechen den Target-EM und die grünen Kreise den Desired-EM, von denen mindestens einer erhalten bleiben soll. Mit diesen Forderungen können die cMCSs berechnet werden. Insgesamt können für das definierte Problem 124 cMCSs mit einer Größe von jeweils 8 - 13 Reaktionsknockouts identifiziert werden (Abbildung 5.7 und 5.8). Alle gefundenen Interventionsstrategien ermöglichen in der Nachtphase (unter Verbrauch von Glycogen) eine ATP-Synthese. Abbildung 5.9 zeigt das Histogramm der Größenverteilung der cMCSs. Die gefundenen 124 cMCSs erzwingen also jeweils das gewünschte Verhalten, nämlich die wachstumsgekoppelte Ethanolproduktion (mit einer hohen Produktausbeute). Exemplarisch wird der erste cMCS genauer analysiert, der die Reaktionen (R1, R2, R3, R4, R5, R7, R8, R23) ausschaltet (Abbildung 5.10).

5.3 Ethanol-Produktion in *Synechocystis* sp. PCC 6803

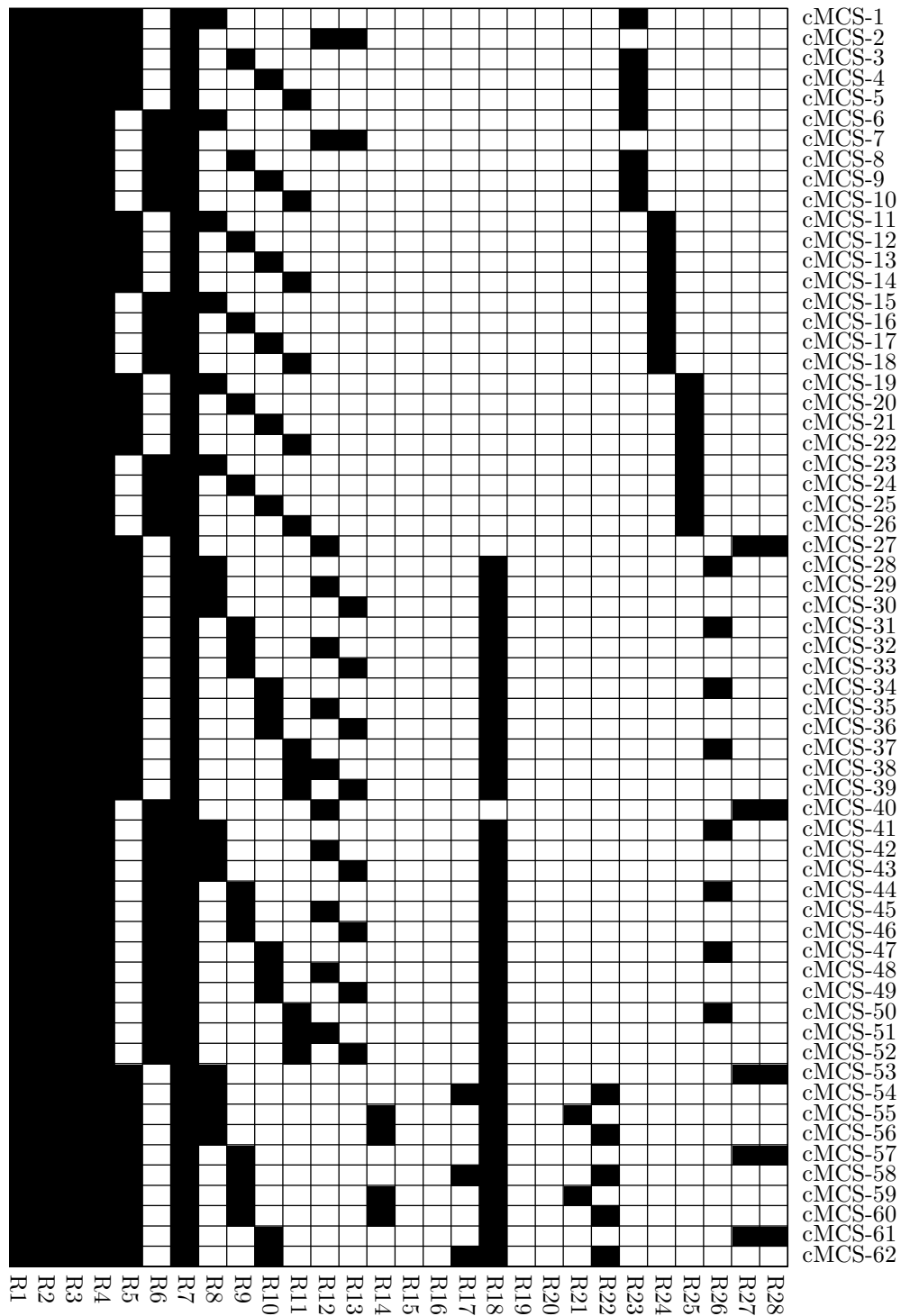


Abbildung 5.7: Die berechneten cMCS-1 - cMCS-62 für die wachstumsgekoppelte Ethanolproduktion. Die schwarzen Kacheln zeigen die notwendigen Reaktions-KOs in den jeweiligen cMCSs an.

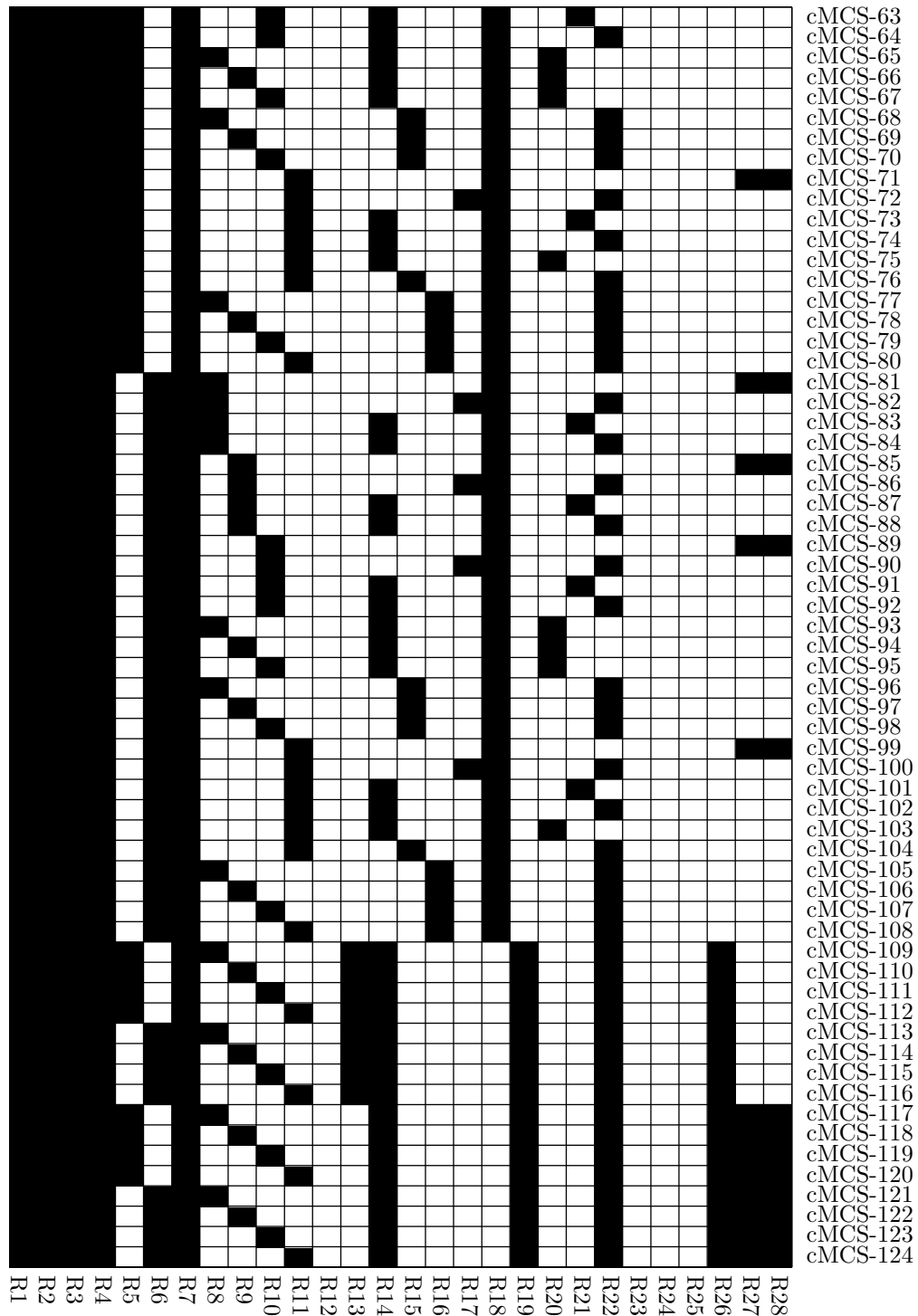


Abbildung 5.8: Die berechneten cMCS-63 - cMCS-124 für die wachstumsgekoppelte Ethanolproduktion. Die schwarzen Kacheln zeigen die notwendigen Reaktions-KOs in den jeweiligen cMCSs an.

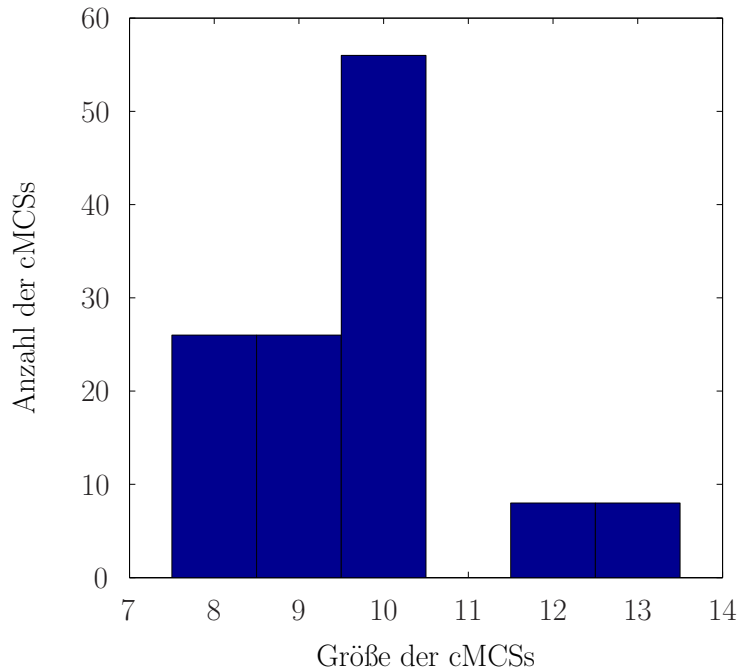


Abbildung 5.9: Histogramm der cMCSs.

Es wird deutlich, dass durch den Großteil dieser Reaktionen ATP unabhängig von NADPH generiert werden kann. Es handelt sich dabei um zyklische (R1, R2, R7) und pseudozyklische (R3, R4, R5) Elektronenflusspfade. Zusätzlich müssen zwei weitere Reaktionen unterdrückt werden. Zum einen die Photorespiration (R8) und zum anderen eine Phosphoglyceratdehydrogenase (R23). Der Knockout beider Reaktionen (R8 und R23) dient dazu, die Ausbeute von Ethanol weiter zu erhöhen (siehe Erläuterung unten).

Durch Anwendung dieser KO-Strategie (cMCS-1) kann garantiert werden, dass das gewünschte Verhalten erreicht wird. Um ATP und NADPH für die Biomassesynthese zur Verfügung zu stellen, muss die Zelle den linearen Elektronenfluss der Photosynthese nutzen. Alle anderen zyklischen und pseudozyklischen Wege, die ohne (Netto-)NADPH-Synthese ATP produzieren können, sind ausgeschaltet. Daraus folgt, dass das maximal mögliche ATP/NADPH Verhältnis, das die photosynthetische Elektronentransportkette generieren kann, 1,28 beträgt. Dieses Verhältnis liegt unter dem das notwendig ist, um Biomasse zu synthetisieren (nämlich 1,51). Das bedeutet, dass die Zelle, um wachsen zu können, einen Überschuss an NADPH generiert, der nur ausgeglichen werden kann, indem Ethanol produziert wird. Das Produkt, das synthetisiert werden soll, fungiert in

diesem Fall als Senke für die überschüssigen Reduktionsäquivalente. Dieser Effekt führt zur wachstumsgekoppelten Ethanolproduktion.

Nach Anwendung von cMCS-1 ergibt sich die in Abbildung 5.10 gezeigte Ausbeutediagramm. Diese Abbildung zeigt die verbleibenden EM (für phototrophes Wachstum)

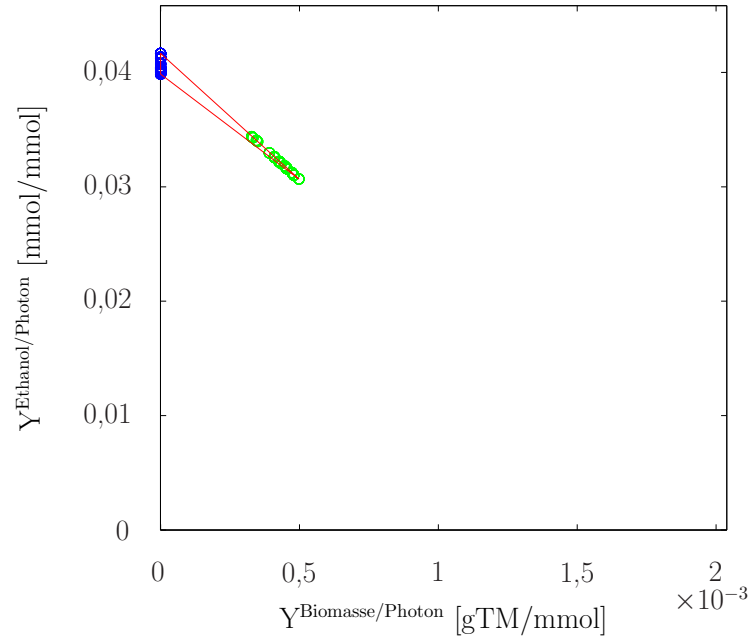


Abbildung 5.10: Ausbeutediagramm der EM für Ethanol- und Biomasse nach Anwendung des cMCS-1 im *SynePruned*. Die grünen Kreise entsprechen den verbliebenen Desired-EM. Die rote Linie beschreibt die konvexe Hülle der verbliebenen EM. Die blauen Kreise stellen andere verbleibende EM dar.

nach Anwendung der genannten KO-Strategie. Zwei Dinge sind deutlich zu erkennen:

1. Die wachstumsgekoppelte Ethanolproduktion kann garantiert werden. Um wachsen zu können, muss die Zelle Ethanol produzieren.
2. Die Ethanolausbeute beträgt mindestens 0,03 mmol Ethanol/mmol Photon. Diese Ausbeute kann unabhängig von der Wachstumsrate garantiert werden.

Die Analyse aller 124 cMCSs zeigt, dass insgesamt 28 verschiedene Reaktionen an der Gesamtmenge der KOs beteiligt sind (eine Übersicht all dieser Reaktionen ist in Tabelle 5.3 zu finden). Dabei zeigt sich, dass die Reaktionen R1, R2, R3, R4 und R7 essentielle Reaktions-KOs repräsentieren. Das bedeutet, dass diese Reaktionen nicht aktiv sein dürfen, um das geforderte Verhalten (gekoppelte Synthese von Ethanol und Biomasse) zu

erreichen. Eine der Reaktionen R5 und R6 muss zusätzlich ausgeschaltet werden. Zusammen stellen diese zwei Reaktionen einen Wasser-Wasser Zyklus dar (Mehler-Reaktion). Da sie eine lineare Kette bilden, reicht es, eine der beiden Reaktionen zu inaktivieren, da die andere dann nicht mehr funktional ist.

Nach Anwendung dieser sechs (essentiellen) Reaktions-KOs sind die Hauptwege der zyklischen und pseudozyklischen Elektronenflusspfade ausgeschaltet. Der KO dieser Reaktionen führt bereits zu der gewünschten wachstumsgekoppelten Ethanolproduktion. Allerdings ist die minimale Ethanolausbeute etwas kleiner als gefordert, nämlich 0,027 mmol Ethanol/mmol Photon (Abbildung 5.11). Dies wiederum bedeutet, dass alle weiteren KOs „lediglich“ dafür sorgen, die Ausbeute von Ethanol weiter zu erhöhen.

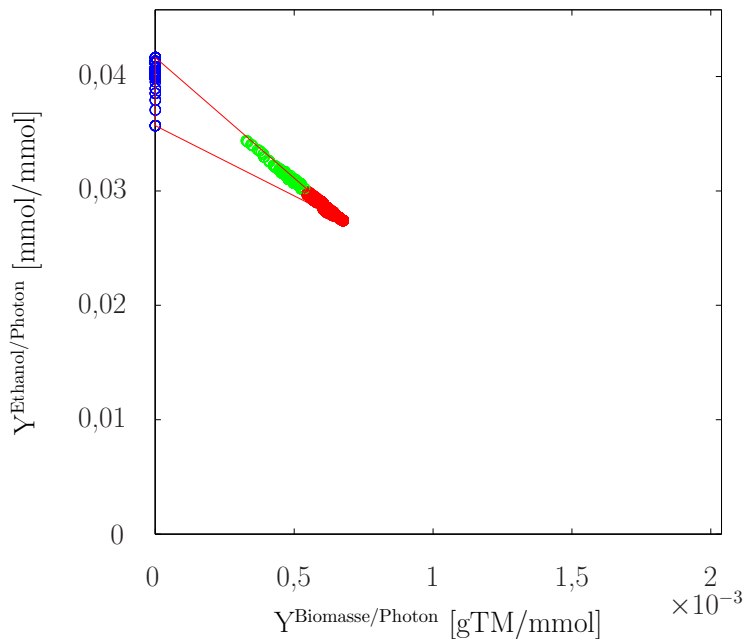


Abbildung 5.11: Ausbeutediagramm der EM für Ethanol- und Biomasse nach Anwendung der (fünf + eins) essentiellen KOs im *SynePruned*. Die grünen Kreise entsprechen den Desired-EM. Die roten Kreisen entsprechen Target-EM. Die rote Linie beschreibt die konvexe Hülle der verbliebenen EM. Die blauen Kreise stellen andere verbleibende EM dar. Der KO dieser Reaktionen bedingt eine wachstumsgekoppelte Ethanolproduktion mit einer minimalen Ausbeute von 0,027 mmol Ethanol/mmol Photon.

Insgesamt bestätigen die identifizierten Strategien die KO-Ziele, die CASOP vorgeschlagen hat. Wie oben bereits erwähnt, sind alle gefundenen Interventionsstrategien bzgl. der Nachtbedingung in der Lage, die ATP-Synthese (ausgehend von Glycogen-

Aufnahme) zu ermöglichen. Da allerdings die respiratorischen Wege durch die Reaktionsknockouts ausgeschaltet sind, erfolgt die ATP-Synthese über fermentative Pfade. Der positive Nebeneffekt ist, dass das Ethanol unter diesen Bedingungen als Fermentationsprodukt ausgeschieden wird [235]. Das bedeutet, dass die Gesamtertrag sogar noch größer ist, da die Zelle nun auch unter Lichtausschluss das gewünschte Produkt synthetisiert. Abbildung 5.12 fasst die potentiellen, durch CASOP und die cMCS-Analyse identifizierten Interventionsstrategien zusammen.

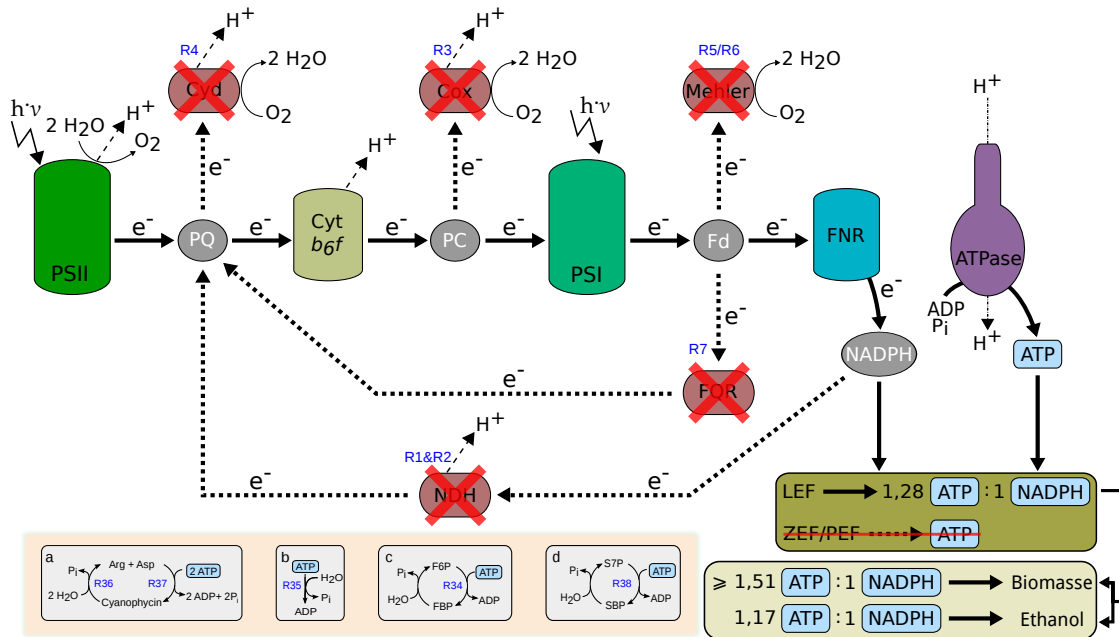


Abbildung 5.12: Zusammenfassung mittels CASOP- und cMCS-Analyse gefundener Interventionsstrategien im *SynePruned*. Die roten Kreuze (x) spiegeln die KO-Ziele wider. Die cMCS-Analyse macht deutlich, dass der KO der gezeigten Reaktionen alle zyklischen und pseudozyklischen Elektronenflüsse blockiert und damit zu einem festen ATP/NADPH Verhältnis von 1,28 führt, welches durch den verbleibenden LEF während der Photosynthese erzeugt wird. Ethanol muss dann zwingend produziert werden, um das ATP/NADPH-Verhältnis entsprechend an die Biomassensynthese anzupassen. Die Reaktionen (bzw. Reaktionszyklen) in (a - d) zeigen die durch CASOP identifizierten Überexpressionskandidaten. Diese führen alle zu einem erhöhten ATP-Verbrauch (ATP-Wasting).

5.3.2.4 ATP-Wasting als alternative/zusätzliche Interventionsstrategie

Da CASOP ATP-verbrauchende Reaktionen als Überexpressionskandidaten vorgeschlagen hat, wird überprüft, ob durch zusätzlichen Verbrauch von ATP eine wachstums-

gekoppelte Ethanolproduktion möglich ist. Wie bereits beschrieben, kann der zyklische und pseudozyklische Elektronenfluss NADPH-unabhängig zusätzliches ATP generieren. Das bedeutet, dass durch alleinige Modifikation des ATP-Verbrauchs die Kopplung nicht garantiert werden kann. Allerdings kann angenommen werden, dass Cyanobakterien die Flüsse des linearen, zyklischen und pseudozyklischen Elektronenflusses um die Photosysteme in Abhängigkeit der Umgebungsbedingungen optimiert haben. Daraus folgt, dass die Zellen einen erhöhten ATP-Verbrauch vermutlich nicht vollständig durch hochgradige Anpassung im zyklischen oder pseudozyklischen Elektronenflusspfad ausgleichen können.

Um ein solches Szenario beschreiben zu können, wird optimales Wachstum in dem Modell *SynePruned* simuliert. Basierend auf einer fixierten Aufnahmerate stellt sich ein optimales Verhältnis der verschiedenen Elektronenflüsse um die Photosysteme ein, um die maximale Wachstumsrate zu erreichen. Da es verschiedene Möglichkeiten gibt, wie der Organismus die ZEF und PEF optimal nutzen kann, können sich Variationen der Flussverteilungen ergeben, die allerdings die maximale Wachstumsrate nicht beeinflussen. Jede dieser optimalen Lösungen kann für die folgende Analyse verwendet werden.

Um den Effekt eines gesteigerten ATP-Verbrauchs und dessen Einfluss zu beobachten, wird (1) ein wachstumsoptimales Szenario simuliert (Photonenaufnahme: 100 mmol/gTM/h). Anschließend werden (2) die zyklischen und pseudozyklischen Elektronenflusspfade um die Photosysteme fixiert und (3) der ATP-Maintenance-Verbrauch (R32, im Folgenden R-ATPM) von 0 mmol/gTM/h schrittweise auf 8 mmol/gTM/h erhöht. Nach Fixierung des ATP-Maintenance-Verbrauchs wird (4) erneut das Wachstum maximiert. Abbildung 5.13 zeigt, wie sich die Biomassesynthese und die Ethanolproduktion in Abhängigkeit des ATP-Verbrauchs verändern. Es ist klar zu erkennen, dass die Steigerung des ATP-Verbrauchs (ATP-Wasting) unter diesen Bedingungen zu einer Verschiebung von Biomassesynthese in Richtung Ethanolproduktion führt. Dadurch entsteht ein NADPH Überschuss, der nur durch die Produktion von Ethanol ausgeglichen werden kann. Ein ähnliches Verhalten zeigt sich, wenn die ZEF/PEF eine obere Kapazitätsgrenze haben. Bis zu dem Erreichen dieser Grenze wird lediglich die Ausbeute der Biomasse sinken (da zusätzlich ATP verbraucht wird). Sobald die Grenzen der zyklischen und pseudozyklischen Elektronenflusspfade erreicht sind, wird es eine ähnliche Verschiebung von Biomassesynthese in Richtung Ethanolproduktion geben.

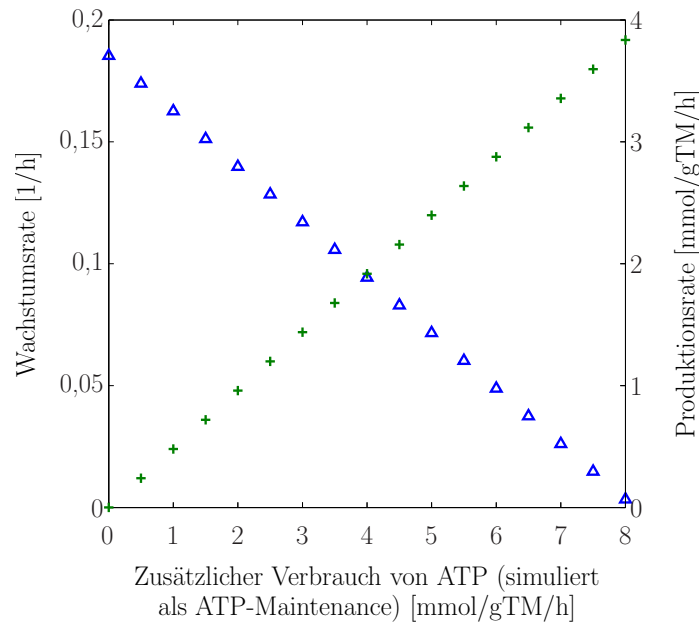


Abbildung 5.13: Steigerung des ATP-Konsums durch schrittweise Erhöhung des R-ATPM-Flusses. Δ beschreibt die Wachstumsrate und $+$ die Ethanol-Produktionsrate. Die Photonenaufnahme ist beschränkt auf 100 mmol/gTM/h. Alle ZEF/PEF sind auf einen optimalen Wert bzgl. maximalem Wachstums bei R-ATPM = 0 fixiert. Anschließend wird der Fluss durch die R-ATPM schrittweise erhöht und die Wachstumsrate maximiert. Eine Verschiebung des Flusses in Richtung Ethanol zu beobachten.

Um zu überprüfen, ob ein zusätzlicher ATP-Verbrauch einen Einfluss auf die Anzahl der für die Kopplung notwendigen Knockouts hat, wird das folgende Szenario simuliert. Die Photonenaufnahme wird beschränkt auf 100 mmol/gTM/h [178]. Die untere Grenze des Flusses der ATPM-Reaktion wird schrittweise von 0 auf 14 mmol/gTM/h erhöht (20 Schritte). Die untere Grenze der Ausbeute von Ethanol (für die Definition der Targetmoden) wird ebenso schrittweise angepasst von 0,002 auf 0,040 mmol/mmol Photonen (20 Schritte). Die untere Grenze der Biomasseausbeute wird festgesetzt auf 0,0001 gTM/mmol Photonen. Damit ergeben sich $20 \times 20 = 400$ verschiedene Szenarien, für die jeweils die kleinsten cMCSs berechnet werden, die eine wachstumsgekoppelte Ethanolproduktion ermöglichen. Dabei zeigt sich, dass bei einem fixierten ATP-Verbrauch die Anzahl der notwendigen Reaktionsknockouts sinkt, sofern die geforderte minimale Ethanolausbeute gesenkt wird (Abbildung 5.14). Das wird sehr deutlich, wenn der Bereich einer sehr hohen Ausbeute (0,03 - 0,04 mmol Ethanol/mmol Photonen) mit einem Bereich hoher Ausbeute (0,02 - 0,03 mmol Ethanol/mmol Photonen) vergli-

5.3 Ethanol-Produktion in *Synechocystis* sp. PCC 6803

chen wird. Ebenso deutlich wird dies, wenn kein zusätzlicher ATP-Verbrauch betrachtet wird. Um eine minimale Ausbeute von 0,03 mmol/mmol Photonen zu erreichen, müssen acht Reaktionen ausgeschaltet werden. Wird die minimale Ethanolausbeute auf 0,026 mmol/mmol Photonen gesenkt, sind sechs Reaktionsknockouts notwendig. Wird statt dessen die minimale Ethanolausbeute auf 0,034 mmol/mmol Photonen erhöht, sind elf Reaktionsknockouts erforderlich, um die Kopplung zu erreichen.

In dem Bereich von 0 - 2,8 mmol/gTM/h des ATP-Verbrauchs ist zu sehen, dass der zusätzliche Verbrauch von ATP einerseits zu einer Reduktion der Cut Set Größe führt und andererseits eine Kopplung mit höherer Ausbeute erst ermöglicht. Am Beispiel von 0,034 mmol Ethanol/mmol Photon ist gut zu erkennen, dass die Erhöhung des ATP-Verbrauchs die Anzahl der für die Kopplung notwendigen Reaktionsknockouts von elf auf sechs reduziert.

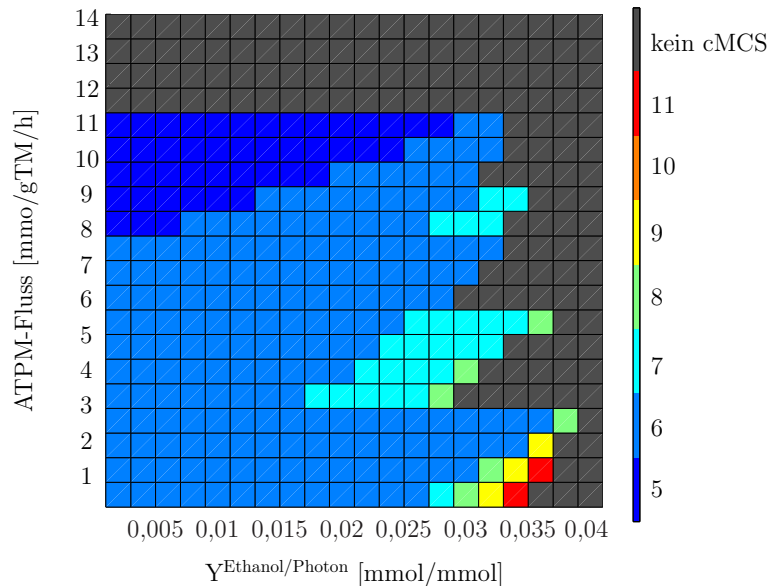


Abbildung 5.14: Einfluss des ATPM-Flusses und der Ethanolausbeute auf die minimale Anzahl von KOs innerhalb der cMCSs im *SynePruned*. Die Ordinate beschreibt die untere Grenze des zusätzlichen Verbrauchs von ATP und die Abszisse die minimale Ausbeute von Ethanol. Abgebildet ist die minimale Größe der jeweiligen cMCSs (Anzahl an notwendigen Reaktionsknockouts).

Der Einfluss des zunehmenden ATP-Verbrauchs bei einer fixierten Ethanolausbeute ist nicht so strikt und in einigen Fällen nicht-monoton. Beispielsweise sinkt, bei einer

Ethanolausbeute von 0,028 mmol Ethanol/mmol Photonen, die notwendige Anzahl der Reaktionsknockouts bei zunehmendem ATP-Verbrauch erst von sieben auf sechs, um ab einem ATPM-Fluss von mehr als 2,8 mmol/gTM/h auf acht zu springen. Wird der ATPM-Fluss allerdings weiter erhöht, sinkt die minimale Größe der cMCSs wieder auf sechs, während ab einem ATP-Verbrauch von 8 mmol/gTM/h die minimale cMCS-Größe auf sieben steigt. Diese Diskontinuitäten entstehen einerseits dadurch, dass bei höherem ATPM-Fluss einige cMCSs nicht mehr zulässig sind und die minimale Biomasseausbeute nicht mehr erfüllen können. Nur größere cMCSs können dann zulässig bleiben. Des Weiteren, kann das Modell den hohen ATP-Verbrauch teilweise nicht durch den linearen Elektronenfluss decken. Eine Reaktion, welche in den alternativen Elektronenfluss involviert ist, wird dann notwendig. In diesem Fall kann nur die Reaktion R2 (siehe Tabelle 5.3), eine CO₂-abh. NDH, dafür genutzt werden, da es Fälle gibt, in denen diese Reaktion nicht unendlich CO₂ in HCO₃⁻ umwandeln kann. Diese Reaktion ermöglicht eine höhere ATP-Ausbeute. Das zusätzliche ATP wird für den höheren ATPM-Fluss gebraucht und steht damit nicht der Biomassesynthese zur Verfügung.

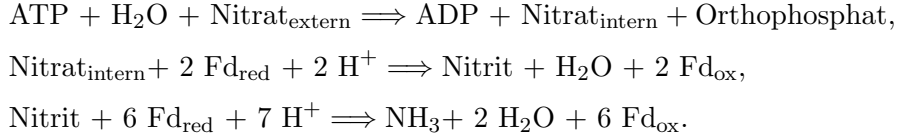
Die minimale Anzahl an Reaktionsknockouts für wachstumsgekoppelte Ethanolproduktion in diesem Modell beträgt fünf und benötigt einen R-ATPM-Fluss von 8 - 11 mmol/gTM/h. Die Abbildung 5.14 zeigt außerdem, dass der ATP-Verbrauch nicht unbegrenzt hoch sein kann, um die Ethanolproduktion und Wachstum zu koppeln. Ab einen zusätzlichen Verbrauch von mehr als 11 mmol ATP/gTM/h können keine cMCSs identifiziert werden, die es ermöglichen, die beschriebene Kopplung zu erzwingen. Unter diesen Bedingungen ist das generierte ATP/NADPH Verhältnis nicht mehr in der Lage, die Bedingungen an (i) den unspezifischen ATP-Verbrauch, (ii) die geforderte minimale Biomasseausbeute und (iii) die festgelegte Ethanolausbeute zu erfüllen. Um noch mehr ATP zu verschwenden, muss die Zelle einen unbeschränkten Elektronenflussmechanismus um die Photosysteme nutzen, der das ATP/NADPH Verhältnis verschiebt (vergrößert), wodurch die garantierte wachstumsgekoppelte Ethanolproduktion nicht mehr möglich ist.

5.3.2.5 Einfluss der Stickstoffquelle (Ammonium vs. Nitrat)

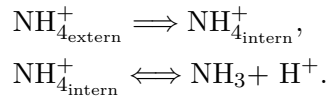
Bei allen bisher gezeigten Ergebnissen wurde angenommen, dass als Stickstoffquelle ausschließlich Nitrat zur Verfügung steht. Innerhalb der Zelle wird allerdings Stickstoff in Form von NH₃ metabolisiert. Die Aufnahme eines Nitrat-Moleküls erfolgt unter Ver-

5.3 Ethanol-Produktion in *Synechocystis* sp. PCC 6803

brauch eines ATPs. Nitrat wird nach Aufnahme intern erst zu Nitrit und danach in NH_3 umgewandelt (unter Verbrauch von Ferredoxin):



Im Modell kann als weitere Stickstoffquelle Ammonium ohne ATP-Verbrauch aufgenommen und ohne den Konsum von reduziertem Ferredoxin in NH_3 umgewandelt werden:



Unter der Annahme, dass 100 mmol Photonen/gTM/h aufgenommen werden, erhöht sich damit die Wachstumsrate im Modell von $0,185 \text{ h}^{-1}$ auf $0,228 \text{ h}^{-1}$. Abbildung 5.15 zeigt das Ausbeutediagramm für phototrophes Wachstum mit Ammonium als Stickstoffquelle. Es wird deutlich, dass die Zelle durch die Aufnahme von Ammonium anstatt von Nitrat mehr Energie (1 ATP) und Reduktionsäquivalente ($8 \text{ Fd}_{\text{red}} \hat{=} 4 \text{ NADPH}$) zur Verfügung hat.

Die maximal mögliche Kopplungsausbeute ist größer, nämlich $0,0377 \text{ mmol Ethanol/mmole Photon}$ statt $0,0344 \text{ mmol Ethanol/mmole Photon}$, da der Organismus im Modell nun weniger NADPH verbraucht, um den Stickstoff zu metabolisieren. Es ergibt sich ein höherer NADPH-Überschuss und damit eine höhere Produktausbeute. Wird die kleinste KO-Strategie (KO der Hauptwege der zyklischen und pseudozyklischen Elektronenflusspfade) für die wachstumsgekoppelte Ethanolproduktion in diesen Fall angewendet, ist die garantierte minimale Ausbeute höher, nämlich $0,036 \text{ mmol Ethanol/mmole Photon}$ statt $0,027 \text{ mmol Ethanol/mmole Photon}$ (siehe Abbildung 5.15 rechts). Dabei wären dann auch nur sechs Reaktions-KOs (anstatt acht) notwendig, um die in Abschnitt 5.3.2.3 geforderte Ausbeute von $0,03 \text{ mmol Ethanol/mmole Photon}$ zu erreichen.

5.3.3 Analyse des genomskaligen *SyneGS*-Modells

Um zu überprüfen, ob die bisher gezeigten Strategien auch im genomskaligen Modell grundlegende Prinzipien repräsentieren, wird eine cMCS-Analyse im genomskaligen Netzwerk (*SyneGS*) durchgeführt. Allerdings erlaubt die Komplexität dieses Modells nicht, die komplette Menge der cMCSs zu berechnen (im Gegensatz zu *SynePruned* und *SyneMin*). Deshalb werden nur die kleinsten cMCSs mit Hilfe des in Abschnitt 3.3.2 be-

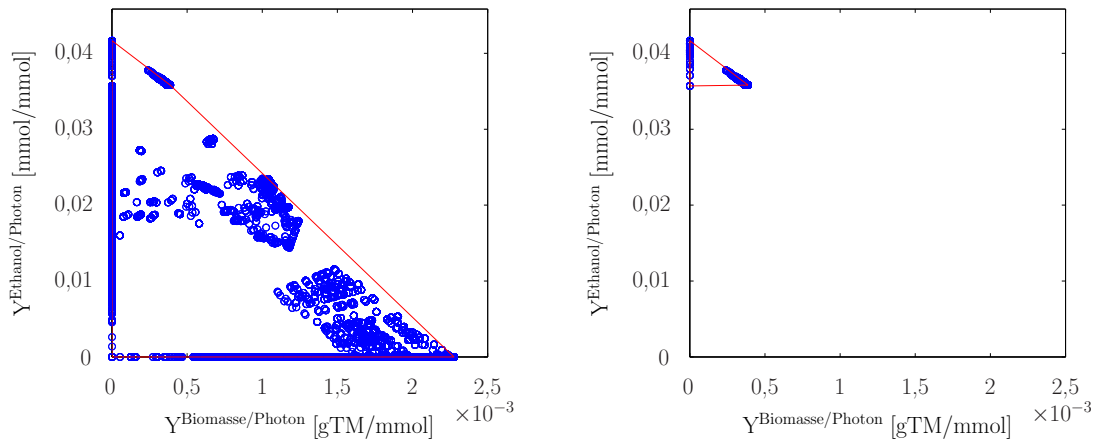


Abbildung 5.15: Ausbeutediagramme der EM für Ethanol und Biomasse im *Syne-Pruned*: Ammonium als Stickstoffquelle. Die blauen Kreise entsprechen den EM. Die rote Linie beschreibt die konvexe Hülle der EM. Links: Ausbeutediagramm der EM für Biomasse und Ethanol. Rechts: verbleibendes Ausbeutediagramm nach Anwendung des cMCS-1.

beschriebenen dualen Ansatzes berechnet. Die geforderte minimale Ausbeute für Ethanol und Biomasse bleibt bestehen (siehe Gleichung (5.1)). Des Weiteren sollen die cMCSs die ATP-Synthese ausgehend von Glycogen unter Nachtbedingungen ermöglichen.

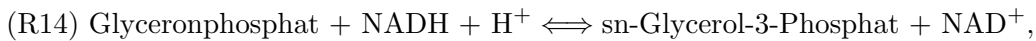
Die kleinste Interventionsstrategie, die dieses Ziel erreicht, erfordert den Knockout von 14 Reaktionen. Bis zu einer Größe von 16 Reaktionsknockouts konnten alle cMCSs identifiziert werden; Größe (Anzahl): 14 (66), 15 (244) und 16 (1254). Insgesamt existieren demnach 1564 cMCSs, die ermöglichen, dass Ethanolproduktion und Wachstum unter den genannten Bedingungen gekoppelt werden kann.

Einige Reaktionen sind in allen cMCSs vertreten. Diese sind (wie in *SynePruned*) R1, R3, R4 und R7. Zusätzlich muss die Reaktion NDH2 (NADH-abhängige Dehydrogenase), die Mehler-ähnliche Reaktion und die Reaktion: Prolin + PQ \implies 1-Pyrrolin-5-Carboxylat + PQH₂ ausgeschaltet werden. Diese Reaktionen stellen im genomskaligen Modell zusätzliche alternative Elektronenflusswege dar, deren Funktion unterdrückt werden muss, um die Kopplung zu erzwingen.

Interessant ist, dass für diese Kopplungsbedingungen die Reaktion R2 nicht zwangsläufig ausgeschaltet werden muss (in *SynePruned* muss diese Reaktion deletiert werden). Das bedeutet, dass es in dem genomskaligen Modell bestimmte (gewünschte) Szenarien gibt, für die diese Reaktion notwendig ist und deren KO damit nicht mehr essentiell

5.4 Produktion von weiteren Biokraftstoffen mit phototrophen Cyanobakterien

für alle identifizierten Interventionsstrategien ist. Um die geforderte Ethanol-Ausbeute zu erreichen, gibt es weitere Reaktionen, die als Interventionsziel auftauchen, da diese potentiell das Verhältnis von ATP zu NADPH vergrößern können. Allerdings sind viele dieser Reaktionen gegeneinander austauschbar. Ein Beispiel dafür ist ein Zyklus, an dem mehrere direkt aufeinanderfolgende Reaktionen beteiligt sind. Es reicht demnach eine dieser Reaktionen zu deletieren, um die Funktion eines solchen Zyklus zu unterdrücken. Ein Beispiel repräsentiert die Transhydrogenase (R12, Tabelle 5.3) im Zusammenspiel mit den Reaktionen:



Dieser Reaktionsweg stellt einen weiteren alternativen Elektronenflusspfad dar, der das ATP/NADPH Verhältnis zu Gunsten des ATPs verschieben kann, da die Elektronen des NADPH wieder in die Elektronentransportkette zurückgeführt werden können. Dieser Weg muss demnach unterdrückt werden. Wird die geforderte Ausbeute von Ethanol auf 0,027 mmol/mmol Photonen gesenkt, kann die Größe des kleinsten cMCSs auf zwölf Reaktionsknockouts reduziert werden.

Die gefundenen cMCSs des genomskaligen, *SyneGS*, des Modells mittlerer Größe, *SynePruned*, und das Minimalmodell, *SyneMin*, zeigen alle das gleiche fundamentale Prinzip, um die Ethanolproduktion mit einer hohen Ausbeute an das Wachstum zu koppeln. Es werden alle zyklischen und pseudozyklischen Elektronenflusspfade ausgeschaltet, die ein ATP/NADPH Verhältnis ermöglichen, das größer ist als das des LEF. Dies bedingt, dass der Organismus einen Überschuss an NADPH generiert, der nur durch die Synthese des Ethanols ausbalanciert werden kann.

Das genomskalige Modell zeigt, dass es neben den „klassischen“ Elektronenflusspfaden auch viele andere Wege durch das Netzwerk gibt, die potentiell eine Rückführung von Elektronen ermöglichen. Das hat zur Folge, dass die notwendige Anzahl an Reaktionsknockouts höher als in dem *SynePruned*-Modell ist.

5.4 Produktion von weiteren Biokraftstoffen mit verschiedenen phototrophen Cyanobakterien

Wie in Kapitel 2.3 angemerkt, soll hier die Produktion weiterer Biokraftstoffe analysiert werden. Dazu gehören Isobutanol, Isopren, Ethylen und 1-Butanol. Alle genannten Kraftstoffe sind bereits mit Hilfe von Cyanobakterien im Labor produziert worden.

5 Metabolic Engineering von Cyanobakterien zur Biokraftstoffproduktion

Im Folgenden wird untersucht, ob es im Vergleich zu Ethanol ähnliche Strategien zur wachstumsgekoppelten Biokraftstoffproduktion gibt. Da die genannten Strategien darauf basieren, das Verhältnis von ATP zu NADPH zu modifizieren, sind deren benötigten Mengen für die Produktion jeweils eines Moleküls Biokraftstoff von besonderem Interesse. Tabelle 5.4 zeigt neben dem Bedarf an ATP und NADPH auch den Bedarf an Photonen zur Produktion jeweils eines Moleküls Biokraftstoff und eines Gramms Trockenbiomasse im *SyneGS*.

Tabelle 5.4: Der Bedarf an Photonen, ATP und NADPH zur Produktion jeweils eines Moleküls Biokraftstoff bzw. eines Gramm Trockenbiomasse im *SyneGS*. Nitrat entspricht der Standard Stickstoffquelle.

Biokraftstoff	Photonenbedarf	ATP	NADPH	ATP/NADPH
1-Butanol	48,00	12,00	12,00	1,00
Ethanol	24,00	7,00	6,00	1,17
Ethylen	61,00	19,50	15,50	1,26
Ethylen _{Ammonium}	44,00	18,00	9,50	1,89
Isobutanol	48,00	14,00	12,00	1,17
Isopren	56,00	17,00	14,00	1,21
Biomasse	539,51	191,60	126,64	1,51
Biomasse _{Ammonium}	439,29	182,61	90,74	2,01

Für die Produktion von Ethylen und Biomasse hat die Nitratquelle einen Einfluss auf den Bedarf von ATP und NADPH. Im Fall des teureren Aufnahmesystems (Nitrat) ist bei beiden zu erkennen, dass der Bedarf an Photonen und ATP größer ist. Für Ethylen beträgt der Unterschied 1,5 ATP. Die Produktion von Ethylen erfordert die Aufnahme von Stickstoff für die Synthese von Arginin. Die Aufnahme von Stickstoff über den Nitratweg kostet pro Nitrat ein ATP und um ein Molekül Ethylen zu produzieren, ist die Aufnahme von 1,5 Molekülen Nitrat notwendig. Die Aufnahme von Ammonium hingegen benötigt kein ATP. Deshalb ergibt sich ein Unterschied in der ATP Bilanz von 1,5 (Ethylen_{Nitrat} vs Ethylen_{Ammonium}). Die Unterschiede in der NADPH Bilanz basieren ebenso auf dem Aufnahmesystem von Stickstoff (siehe Kapitel 5.3.2.5). Pro Nitratmolekül werden acht reduzierte Ferredoxin benötigt. Dies entspricht vier NADPH (Ferredoxin-NADPH-Reduktase: $2 \text{ reduzierte Ferredoxin} + \text{NADP}^+ + \text{H}^+ \rightleftharpoons 2 \text{ oxidierte Ferredoxin} + \text{NADPH}$).

5.4 Produktion von weiteren Biokraftstoffen mit phototrophen Cyanobakterien

dierte Ferredoxin + NADPH). Da 1,5 Moleküle Nitrat pro Ethylen aufgenommen werden müssen, beträgt die Differenz sechs NADPH.

Gleiches gilt für die Biomasse, nur dass der Bedarf an Stickstoff größer ist (für die Produktion von einem Gramm Trockenbiomasse ist die Aufnahme von rund neun mmol Nitrat bzw. Ammonium notwendig). Analog ergibt sich für die Biomasse ein Unterschied von rund 36 NADPH und rund 9 ATP (siehe Tabelle 5.4).

Im Folgenden wurden die Kopplungsstrategien für die genannten Produkte in Abhängigkeit der Stickstoffquelle im *SyneGS* berechnet. Tabelle 5.5 zeigt jeweils die geringste Anzahl an Reaktionsknockouts in den gefundenen cMCSs, welche zur wachstumsgekoppelten Biokraftstoffproduktion führen. Die kleinste Interventionsstrategie für jeden zu

Tabelle 5.5: Die minimale Größe der identifizierten cMCSs für die gekoppelte Produktion von Biokraftstoffen und Biomasse in *SyneGS*. Die minimale Produktionsrate (in [mmol/gTM/h]) wurde berechnet unter der Annahme der Aufnahme von 100 mmol Photonen/gTM/h und $Y^{\text{Biomasse/Photon}} \geq 0,0001$ gTM/mmol Photonen. Die minimale Ausbeute (in [mmol Biokraftstoff/mmol Photonen]) ergibt sich aus: min. Produktionsrate geteilt durch Photonenaufnahme (100 mmol Photonen/gTM/h).

Biokraftstoff	Stickstoffquelle: Nitrat		Stickstoffquelle: Ammonium	
	kleinster cMCS	minimale Produktionsrate	kleinster cMCS	minimale Produktionsrate
1-Butanol	12	0,93	12	1,23
Ethanol	12	2,74	12	3,28
Ethylen	12	1,54	13	1,56
Isobutanol	12	1,37	12	1,64
Isopren	12	1,36	12	1,55

produzieren Biokraftstoff beträgt zwölf Reaktionsknockouts. Für alle Produkte (bis auf Ethylen_{Ammonium}) gibt es ein gemeinsames Cut Set. Dieses entspricht dem in Kapitel 5.3.3 vorgestellten cMCS.

Das in *SyneMin* und *SynePruned* identifizierte Prinzip wird im genomskaligen Modell, *SyneGS*, also nicht nur bestätigt, sondern ebenso für andere Produkte mit einem ATP/NADPH Verhältnis, welches kleiner ist als das des LEF, wiedergefunden. Somit kann geschlossen werden, dass es sich dabei um ein generelles Interventionsprinzip für die wachstumsgekoppelte Produktion von Biokraftstoffen handelt. Des Weiteren wird

deutlich, dass die Aufnahme von Ammonium anstatt von Nitrat eine höhere minimale Ausbeute der Produkte zu Folge hat.

Für Ethylen_{Ammonium} kann die Kopplung ebenso erzwungen werden. Allerdings ist das Verhältnis von ATP zu NADPH größer als das des linearen Elektronenflusses. Damit muss eine Reaktion aktiv sein, die in den „zyklischen“ Elektronenflussmechanismus um die photosynthetische ETK involviert ist. Analog zu Kapitel 5.3.2.4 handelt es sich um eine CO₂-abhängige NDH, deren Fluss durch diese Abhängigkeit begrenzt werden kann. Dennoch reduziert sich die Anzahl der notwendigen Reaktionsknockouts nicht, da diese aktive Reaktion des alternativen Elektronenflusses ermöglicht, dass ausschließlich Biomasse synthetisiert werden kann. Deshalb kann die Kopplung mit dieser Ausbeute nur erreicht werden, wenn der Organismus gezwungen wird, einerseits diese Reaktion des alternativen Elektronenflusses zu beschränken (durch z. B. den KO der Reaktion von Malat zu Pyruvat bei der CO₂ frei wird) und andererseits die Reaktion von 3-Phosphoglycerat (3PG) zu 2-Phosphoglycerat (2PG) auszuschalten. Damit muss der Organismus den energie-/kostenintensiven Phosphoserinpfad oder die Photorespiration nutzen, um 2PG zu bilden (höherer ATP-Verbrauch). Dadurch kann die Kopplung garantiert werden. Allerdings sei noch einmal darauf hingewiesen, dass eine cytosolische Hydrogencarbonatequilibriumreaktion dieses Verhalten zerstören würde. Mit dieser Reaktion könnte die Reaktion des alternativen Elektronenflusses (R2) nicht beschränkt werden und der genannte cMCS könnte keine Kopplung garantieren.

Biokraftstoffproduktion mit *Synechococcus elongatus* PCC 7942

Um zu überprüfen, ob die gefundenen Strategien auch in anderen phototrophen Cyanobakterien wiederzufinden sind, wird mit dem Cyanobakterium *Synechococcus elongatus* PCC 7942 eine weitere Spezies untersucht.

Für die folgenden Analysen wurde das genomskalige Modell von Knoop [133] verwendet (für eine detailliertere Beschreibung siehe Kapitel 3.5). Die beschriebenen Biokraftstoffe wurden analog zu *Synechocystis* in das Modell integriert (siehe Tabelle A.1). Der Bedarf an ATP und NADPH für die Produktion eines Moleküls Biokraftstoff und eines Gramms Trockenbiomasse im Modell von in *Synechococcus* ist in Tabelle 5.6 zu finden. Die Tabelle 5.6 zeigt, dass *Synechococcus* einen identischen bzw. ähnlichen Bedarf an ATP und NADPH für die Synthese der analysierten Biokraftstoffen hat. Der Bedarf für die Synthese von Biomasse ist vergleichbar mit dem von *Synechocystis* und die Unterschiede bzgl. der Stickstoffquelle sind auch in diesem Fall auf das Aufnahmesystem zurückzuführen.

5.4 Produktion von weiteren Biokraftstoffen mit phototrophen Cyanobakterien

Tabelle 5.6: Der Bedarf an Photonen, ATP und NADPH zur Produktion jeweils eines Moleküls Biokraftstoff bzw. eines Gramm Trockenbiomasse im genomskaligen Netzwerk von *Synechococcus*. Nitrat entspricht der Standard Stickstoffquelle.

Biokraftstoff	Photonenbedarf	ATP	NADPH	ATP/NADPH
1-Butanol	48,00	12,00	12,00	1,00
Ethanol	24,00	7,00	6,00	1,17
Ethylen	64,25	22,50	15,50	1,45
Ethylen _{Ammonium}	47,50	21,00	9,50	2,29
Isobutanol	48,00	14,00	12,00	1,17
Isopren	56,00	17,00	14,00	1,21
Biomasse	542,69	194,75	126,95	1,53
Biomasse _{Ammonium}	442,47	185,77	91,05	2,04

Für die Produktion von Ethylen zeigt sich, dass in beiden Fällen mehr ATP benötigt wird. Dies ist dadurch zu erklären, dass sich der Metabolismus ein wenig unterscheidet. Konkret wird für die Produktion von Arginin (welches ein Präkursor für Ethylen ist) mehr ATP benötigt (pro Arginin vier ATP mehr) und der Abbau eines Nebenprodukts der Ethylensynthese (Succinat) ist ebenso kostenintensiver (pro Succinat zwei ATP mehr). Für die Produktion eines Moleküls Ethylen werden 0,5 Moleküle Arginin verbraucht und 0,5 Moleküle Succinat gebildet. Damit ergibt sich ein höherer Energiebedarf von zwei ATP pro Ethylen (Vergleich: Tabelle 5.4 und 5.6).

Prinzipiell werden wieder die Strategien gesucht, die eine wachstumsgekoppelte Biokraftstoffproduktion mit möglichst wenigen Reaktionsknockouts ermöglichen. Die cMCSs werden mit Hilfe des in Kapitel 3.3.2 vorgestellten dualen Ansatzes berechnet. Im Modell von *Synechococcus* finden sich ähnliche Interventionsstrategien wie in *Synechocystis*. Das kleinste cMCS, das wachstumsgekoppelten Biokraftstoffproduktion erzwingt, benötigt elf Reaktionsknockouts. Bis auf Ethylen gibt es für alle Produkte eine gemeinsame Kopplungsstrategie. Es müssen dafür verschiedene NDHs, eine FQR-Reaktion und eine Mehler-ähnliche Reaktion ausgeschaltet werden. Zusätzlich müssen respiratorische Reaktionen inhibiert werden. Fast alle Reaktionen sind damit wieder direkt auf zyklische und pseudozyklische Elektronenflusspfade zurückzuführen. Auch für dieses Modell kann somit geschlussfolgert werden, dass es sich um ein generelles Interventionsprinzip für

Tabelle 5.7: Die minimale Größe der identifizierten cMCSs für die gekoppelte Produktion von Biokraftstoffen und Biomasse im genomskaligen Modell von *Synechococcus*. Die minimale Produktionsrate (in [mmol/gTM/h]) wurde berechnet unter der Annahme der Aufnahme von 100 mmol Photonen/gTM/h und $Y^{\text{Biomasse/Photon}} \geq 0,0001$ gTM/mmol Photonen. Die minimale Ausbeute (in [mmol Biokraftstoff/mmol Photonen]) ergibt sich aus: min. Produktionsrate geteilt durch Photonenaufnahme (100 mmol Photonen/gTM/h).

Biokraftstoff	Stickstoffquelle: Nitrat		Stickstoffquelle: Ammonium	
	kleinster cMCS	minimale Produktionsrate	kleinster cMCS	minimale Produktionsrate
1-Butanol	11	0,90	11	0,98
Ethanol	11	2,70	11	2,92
Ethylen	11	0,20	11	0,58
Isobutanol	11	1,35	11	1,46
Isopren	11	1,35	11	1,46

die wachstumsgekoppelte Produktion von Biokraftstoffen handelt, deren ATP/NADPH Verhältnis kleiner ist, als das des LEF.

Es gibt, ähnlich wie im Modell von *Synechocystis*, versteckte und alternative Wege, die das ATP/NADPH-Verhältnis so ungünstig modulieren können, dass die Kopplung aufgehoben wird. Auch diese Reaktionen sind denen des *SyneGS* ähnlich. Des Weiteren wird erneut deutlich, dass die Aufnahme von Ammonium anstatt von Nitrat eine höhere minimale Ausbeute der Produkte zu Folge hat.

Bezüglich der Ethylenproduktion gibt es eine Besonderheit. Für den Fall der Nitrataufnahme ist das ATP/NADPH Verhältnis größer als das des LEF, aber kleiner als das der Biomassesynthese. Um dies zu erreichen, können nicht alle in den „zyklischen“ Elektronenflussmechanismus involvierten Reaktionen ausgeschaltet werden. In diesem Fall bleibt eine CO₂-abhängige NDH aktiv. Dennoch sinkt die Größe des cMCSs nicht, da diese aktive Reaktion das ATP/NADPH Verhältnis so modulieren kann, dass die Kopplung aufgehoben wird. Durch die Art der Modellierung des CO₂ Haushalts kann diese NDH jedoch beschränkt werden. Denn auch in diesem Netzwerk kann statt CO₂ nur Hydrogencarbonat aufgenommen werden, das in das Carboxysom transportiert wird, um dort in Form von CO₂ fixiert zu werden. Die Hydrogencarbonatequilibriumreaktion liegt ausschließlich in diesem Kompartiment vor, sodass die cytosolische CO₂-abhängige

5.4 Produktion von weiteren Biokraftstoffen mit phototrophen Cyanobakterien

NDH beschränkt werden kann. Der statt dessen notwendige KO verhindert die interne Bildung von CO₂ aus Formiat. Dadurch wird die aktive CO₂-abh. NDH limitiert und das ATP/NADPH Verhältnis kann nicht so verändert werden, dass die Kopplung aufgehoben wird.

Interessant ist, dass derselbe cMCS angewendet werden kann, um die Produktsynthese und Wachstum zu koppeln, wenn Ammonium statt Nitrat aufgenommen wird. Dass die Kopplung auch in diesem Fall derart funktioniert ist bemerkenswert, da das ATP/NADPH Verhältnis größer ist als das der Biomassesynthese. Der Grund dafür ist, dass bei der Ethylenproduktion CO₂ frei wird (pro Molekül Ethylen: 3,5 Moleküle CO₂). Um zu wachsen muss der Organismus das dafür notwendige ATP/NADPH Verhältnis durch Nutzung der CO₂-abh. NDH generieren. Damit diese NDH in dafür ausreichendem Maße betrieben werden kann, muss der Organismus Ethylen bilden, um das daraus gewonnene CO₂ zur Verfügung zu stellen. Dadurch kann die Kopplung garantiert werden. Wird angenommen, dass kein CO₂ bei der Ethylenproduktion entsteht, kann die Kopplung nicht mehr aufrecht erhalten werden.

Biokraftstoffproduktion mit *Cyanothece* sp. ATCC 51142

Als abschließendes Beispiel einer cyanobakteriellen Spezies, wird das genomskalige Modell für *Cyanothece* sp. ATCC 51142 von Saha et al. untersucht [211]. Das Modell verfügt über zwei Photosystem-Komplexe (an der Thylakoidmembran und an der periplasmatischen Membran). Aufgrund der im Modell enthaltenen Unvollständigkeit des Photosystem-Komplexes an der periplasmatischen Membran, wird ausschließlich der Photosystem-Komplex an der Thylakoidmembran in den folgenden Analysen betrachtet.

Die beschriebenen Biokraftstoffe werden analog zu *Synechocystis* und *Synechococcus* in das Modell integriert (siehe Tabelle A.1). Der Bedarf an ATP und NADPH für die Produktion jeweils eines Moleküls Biokraftstoff und eines Gramms Trockenbiomasse in dem Modell von *Cyanothece*, ist in Tabelle 5.8 zu finden. In diesem Netzwerk wird statt Ammonium direkt NH₃ aufgenommen (ATP-unabhängig). Für Isopren ergibt sich ein höherer ATP Bedarf, da eine Reaktion fehlt, die in dem *Synechocystis* und *Synechococcus* Modell die Synthese von ATP aus ADP und Diphosphat erlaubt (Diphosphat entsteht bei der Produktion von Isopren). Die cMCSs werden mit Hilfe des in Kapitel 3.3.2 vorgestellten genomskaligen Ansatzes berechnet.

Tabelle 5.9 zeigt, dass auch in diesem Modell Interventionsstrategien zu finden sind, die es ermöglichen, die Produktion von Biokraftstoffen an die Biomassesynthese zu koppeln. Für den Fall der Aufnahme von Nitrat gibt es bis auf Isopren einen gemeinsamen cM-

Tabelle 5.8: Der Bedarf an Photonen, ATP und NADPH zur Produktion jeweils eines Moleküls Biokraftstoff bzw. eines Gramm Trockenbiomasse im genomskaligen Modell von *Cyanothece*. Nitrat entspricht der Standard Stickstoffquelle.

Biokraftstoff	Photonenbedarf	ATP	NADPH	ATP/NADPH
1-Butanol	48,00	12,00	12,00	1,00
Ethanol	24,00	7,00	6,00	1,17
Ethylen	60,25	19,50	15,00	1,30
Ethylen _{NH₃}	45,25	19,50	9,00	2,17
Isobutanol	48,00	14,00	12,00	1,17
Isopren	57,16	19,00	14,00	1,36
Biomasse	374,51	137,98	85,74	1,62
Biomasse _{NH₃}	318,59	137,98	63,37	2,18

CS, der insgesamt 14 Reaktionsknockouts erfordert. Dieses cMCS fordert analog zu den bisherigen Ergebnissen überwiegend den KO von zyklischen und pseudozyklischen Elektronenflusspfaden. Die restlichen Reaktionen stellen versteckte und alternative Wege dar, die das ATP/NADPH-Verhältnis so ungünstig modulieren können, dass die Kopplung aufgehoben wird. Beispiele solcher Reaktionen in diesem Modell sind: die Succinat Dehydrogenase, eine Glutathion:NADP:Oxidoreduktase, eine Glycerol-3-Phosphat Dehydrogenase, eine Wasserstoffperoxid Oxidoreduktase und eine Protochlorophyllid Reduktase. Der KO der Wasserstoffperoxid Oxidoreduktase verhindert die Bildung von Superoxid-Anion (die sonst Elektronen aus der Photosynthese abziehen könnte). Die Succinat Dehydrogenase würde einen Zyklus ermöglichen, in dem NADPH verbraucht wird. Für die verbleibenden Reaktionen (Glutathion:NADP Oxidoreduktase, Glycerol-3-Phosphat Dehydrogenase und Protochlorophyllid Reduktase) gilt dies ebenso. Die Protochlorophyllid Reduktase würde zusätzlich die Bildung von ATP erlauben. Das identifizierte Interventionsprinzip wird auch in diesem Modell bestätigt. Die Reaktionsknockouts beeinflussen das ATP/NADPH Verhältnis so, dass die Zelle, um wachsen zu können, einen NADPH Überschuss generiert, den diese im Modell nur in Richtung des Produkts ausgleichen kann.

Die Kopplungsstrategie für Isopren ist im Vergleich zu denen der anderen analysierten Biokraftstoffe etwas aufwendiger. Das spiegelt sich in der notwendigen Anzahl der Reaktionsknockouts wider (16 im Vergleich zu 14). Analog zu den bisher diskutierten Er-

5.4 Produktion von weiteren Biokraftstoffen mit phototrophen Cyanobakterien

Tabelle 5.9: Die minimale Größe der identifizierten cMCSs für die gekoppelte Produktion von Biokraftstoff und Biomasse im genomskaligen Modell von *Cyanothece*. Die minimale Produktionsrate (in [mmol/gTM/h]) wurde berechnet unter der Annahme der Aufnahme von 100 mmol Photonen/gTM/h und $Y^{\text{Biomasse/Photon}} \geq 0,0001$ gTM/mmol Photonen. Die minimale Ausbeute (in [mmol Biokraftstoff/mmol Photonen]) ergibt sich aus: min. Produktionsrate geteilt durch Photonenaufnahme (100 mmol Photonen/gTM/h).

Biokraftstoff	Stickstoffquelle: Nitrat		Stickstoffquelle: NH ₃	
	kleinster cMCS	minimale Produktionsrate	kleinster cMCS	minimale Produktionsrate
1-Butanol	14	1,05	14	1,32
Ethanol	14	3,05	14	3,67
Ethylen	14	1,57	17	1,71
Isobutanol	14	1,52	14	1,83
Isopren	16	1,20	16	1,25

gebnissen zeigt sich aber, dass für die Produkte, deren ATP/NADPH-Verhältnis größer ist als das des LEF, ein Mechanismus benötigt wird, der einen „zyklischen“ Elektronenfluss begrenzt zulässt.

Es können auch in diesem Modell cMCSs gefunden werden, die unter diesen Bedingungen die wachstumsgekoppelte Biokraftstoffproduktion erzwingen. Wie in den vorherigen Modellen gezeigt, beeinflusst die Nitratquelle die Produktionsrate (die Aufnahme von NH₃ anstatt von Nitrat führt zu einer höheren minimalen Produktionsrate).

Zusammenfassend ist festzustellen, dass analoge Interventionsstrategien auf allen drei Skalen (*SyneMin*, *SynePruned* und *SyneGS*) und für drei verschiedene Modellorganismen (*Synechocystis* sp. PCC 6803, *Synechococcus* PCC 7942 und *Cyanothece* sp. ATCC 51142) gefunden werden konnten. Unabhängig des genutzten Organismus bzw. dessen Stoffwechselmodells müssen überwiegend die zyklischen und pseudozyklischen Elektronenflusswege inaktiviert werden, um den Bedarf an Energie, der für die Biomassesynthese notwendig ist, ausschließlich über den linearen Elektronenfluss zu decken.

Durch die Verwendung von Ammonium anstatt von Nitrat kann die Produktionsrate erhöht werden, da für die Umwandlung von Nitrat in NH₃ viele Reduktionsäquivalente verbraucht werden. Die gezeigten Interventionsstrategien erfordern eine hohe Anzahl von

5 Metabolic Engineering von Cyanobakterien zur Biokraftstoffproduktion

Reaktionsknockouts, da die Möglichkeiten Elektronen wieder zu verwenden, vielfältig sind. Es zeigt ebenso, dass die Kopplungsprinzipien in phototrophen Bakterien ungleich komplexer sind im Vergleich zu heterotrophen Bakterien wie z. B. *E. coli*.

Für eine praktische Umsetzung kann daher versucht werden, zunächst einige ausgewählte Reaktionen des zyklischen und/oder pseudozyklischen Elektronenflusses zu inaktivieren, um eine Verschiebung des Verhältnisses von Biomassesynthese hin zur Produktion von Biokraftstoffen zu erreichen.

6 Zusammenfassung

Die biobasierte und nachhaltige Produktion von Kraftstoffen rückt immer mehr in den Fokus, um die Abhängigkeit von fossilen Energieträgern zu reduzieren. Sogenannte Biokraftstoffe können z. B. aus Stärke oder Cellulose gewonnen werden. Im Gegensatz dazu können mit Hilfe von Cyanobakterien, die oxygene Photosynthese betreiben, Biokraftstoffe auf landwirtschaftlich unnutzbaren Flächen und ohne den Verbrauch von Nahrungsmitteln produziert werden. Allerdings ist diese Produktion häufig durch eine geringe Produktivität charakterisiert [210].

Ziel dieser Arbeit war es deshalb, modellbasiert auf unterschiedlichen Netzwerkskalen Interventionsstrategien zur Optimierung der Biokraftstoffproduktion in verschiedenen Cyanobakterien zu identifizieren. Da zu Beginn dieser Arbeit nur genomskalige Modelle zur Beschreibung des cyanobakteriellen Stoffwechsels zur Verfügung standen, ist der Reduktionsalgorithmus *NetworkReducer* entwickelt worden. *NetworkReducer* erlaubt es, komplexe genomskalige Stoffwechselmodelle auf Modelle kleinerer Netzwerkskalen zu reduzieren und dabei gewünschte Eigenschaften des Ursprungsmodells zu erhalten. Obwohl nur genomskalige Modelle die komplette Funktionalität des Stoffwechsels beschreiben können, gibt es gute Gründe diese zu reduzieren. So ist z. B. die volle Enumeration von EM oft nur in kleineren Modellen möglich. Außerdem können Einblicke in grundlegende Eigenschaften des Metabolismus in Zentralstoffwechselmodellen leichter gewonnen werden.

Der *NetworkReducer*-Algorithmus gliedert sich in zwei Teilabschnitte. (1) Basierend auf einer FVA werden Reaktionen aus dem Modell entfernt (Netzwerkbescheidung). Dieser Prozess geht mit der Reduktion des Freiheitsgrades einher. Optional werden (2) lineare Reaktionsketten zusammengeschoben (Netzwerkkompression). Dieser Schritt erfolgt ohne Verlust eines Freiheitsgrades, sodass z. B. die Anzahl der EM des reduzierten Modells (Bescheidung und Kompression) der des beschnittenen Modells entspricht.

6 Zusammenfassung

Als Machbarkeitsnachweis ist ein genomskaliges *E. coli* Modell unter Erhaltung der maximalen aeroben und anaeroben Wachstumsraten reduziert worden. Das durch den *NetworkReducer* reduzierte Modell (*ColiPrunedComp*) ist mit einem früheren Zentralstoffwechselmodell (*ColiCore*) verglichen worden, welches manuell aus dem *ColiGS* reduziert worden ist. Im Gegensatz zu dem manuell generierten *ColiCore* ist das *ColiPrunedComp* in der Lage, die Wachstumsraten (aerob und anaerob) des genomskaligen Netzwerks (*ColiGS*) zu beschreiben.

Der *NetworkReducer* ist weiterhin genutzt worden, um ein genomskaliges Stoffwechselmodell von *Synechocystis* sp. PCC 6803 (*SyneGS*) auf ein stöchiometrisches Modell mittlerer Größe (*SynePruned*) und auf ein Minimalmodell (*SyneMin*) zu reduzieren. Mit Hilfe dieser drei Stoffwechselmodelle unterschiedlicher Größe (*SyneMin*, *SynePruned* und *SyneGS*), konnten skalenübergreifend Interventionsstrategien zur Optimierung der Ethanolproduktion für die cyanobakterielle Modellspezies *Synechocystis* sp. PCC 6803 identifiziert werden. Ausgehend von dem Minimalmodell, *SyneMin*, wurde insbesondere nach Strategien gesucht, welche die wachstumsgekoppelte Produktion von Ethanol erzwingen.

Die KO-Strategien, welche die wachstumsgekoppelte Ethanolproduktion mit einer hohen Ethanolausbeute (mehr als 70% der maximalen Ausbeute) ermöglichen, eint dabei ein generelles Prinzip. Sie zeigen skalenübergreifend, dass zyklische und pseudozyklische Elektronenflusswege ausgeschaltet werden müssen, um während des phototrophen Wachstums die Ethanolproduktion zu erzwingen. Dadurch muss der Organismus seinen Bedarf an ATP und NADPH über den linearen Elektronentransport der Photosynthese decken. Um dennoch wachsen zu können, muss der dabei entstehende Überschuss an Reduktionsäquivalenten ausgeglichen werden. Ethanol fungiert dann als Senke für diese Reduktionsäquivalente.

Des Weiteren wurde festgestellt, dass diese Kopplungsstrategien mit zunehmender Modellgröße immer komplexer werden. Immer mehr Reaktionsknockouts wurden notwendig, um den Organismus (im Modell) zu zwingen, während des phototrophen Wachstums Ethanol zu produzieren. Dies hebt hervor, dass es viele verschiedene und versteckte Wege im Metabolismus phototropher Cyanobakterien gibt, die Elektronen aus der Photosynthese direkt oder indirekt nutzen können, um das ATP/NADPH Verhältnis zu vergrößern. Außerdem unterstreicht es die Tatsache, dass die wachstumsgekoppelte Produktsynthese im Vergleich zu heterotrophen Organismen wie z. B. *E. coli* deutlich komplexer und komplizierter ist.

Das Kopplungsprinzip konnte zusätzlich für verschiedene Biokraftstoffe (Ethanol, Isobutanol, 1-Butanol, Isopren und Ethylen) in genomskaligen Modellen verschiedener pho-

totropher Cyanobakterien (*Synechocystis* sp. PCC 6803, *Synechococcus elongatus* PCC 7942 und *Cyanothece* sp. ATCC 51142) bestätigt werden.

Damit konnte im Gegensatz zu modellbasierten Arbeiten von Nogales et al. [179] gezeigt werden, dass die wachstumsgekoppelte Biokraftstoffproduktion unter phototrophen Bedingung erzwungen werden kann. Und weiter, dass im Kontrast zu modellgestützten Arbeiten von Vu et al. [261] konkrete Interventionsstrategien in genomskaligen cyanobakteriellen Stoffwechselmodellen identifiziert werden konnten.

Komplementär zu den Knockoutszenarien sind Überexpressionskandidaten identifiziert worden, die das ATP/NADPH Verhältnis durch den unspezifischen ATP-Verbrauch reduzieren. Dabei handelt es sich um sogenannte „futile cycles“, die netto ATP verschwenden (ATP-Wasting). Allerdings lässt sich mit dieser Strategie die Kopplung nicht erzwingen. Es konnte aber gezeigt werden, dass ein erhöhter ATP-Verbrauch die Anzahl der notwendigen Reaktionsknockouts zur wachstumsgekoppelten Ethanolproduktion senken kann.

Die wesentlichen Resultate dieser Arbeit können wie folgt zusammengefasst werden:

- Entwicklung einer neuen Methode, *NetworkReducer*. Diese reduziert unter Erhaltung benutzerdefinierter Eigenschaften genomskalige Stoffwechselmodelle auf Modelle kleinerer Netzwerkskalen.
- Skalenübergreifende Identifikation des Kopplungsprinzips von Ethanol- und Biomassensynthese in *Synechocystis* sp. PCC 6803 unter phototrophen Wachstumsbedingungen.
- Bestätigung dieses Prinzips für weitere Biokraftstoffe und andere phototrophe Cyanobakterien.
- Vorhersage konkreter Interventionsstrategien in den untersuchten cyanobakteriellen Modellspezies.

Die theoretischen Vorhersagen dieser Arbeit werden derzeit von experimentellen Partnern getestet. Dabei werden zwei konkrete Ansätze verfolgt. Der erste Ansatz versucht die ATP-Wasting Strategie durch Zugabe von Salz umzusetzen. Salz induziert Stress in der Zelle, der einen (gewünschten) erhöhten ATP-Verbrauch mit sich bringt. Der zweite Ansatz versucht eine Deletion einer der wichtigsten zyklischen Elektronenflüsse umzusetzen, nämlich der NAD(P)H-Dehydrogenase (NDH1). Derzeit wird versucht, beide Ansätze zur Optimierung der Biokraftstoffproduktion in Cyanobakterien im Labormaßstab umzusetzen.

6 Zusammenfassung

Eine zukünftige Herausforderung besteht darin, die Abhängigkeit von beschränkt verfügbaren Ressourcen, wie fossilen Energieträgern, zu reduzieren. Cyanobakterien können dabei einen substantiellen Beitrag zur nachhaltigen Biokraftstoffproduktion leisten. Die identifizierten Strategien liefern einen Beitrag zur Optimierung der Biokraftstoffproduktion in Cyanobakterien, wobei die in dieser Arbeit reduzierten Modelle, welche mit Hilfe der neuen Methode *NetworkReducer* entwickelt wurden, skalenübergreifende Einblicke in grundlegende Stoffwechselmechanismen ermöglichen.

A Anhang

A.1 Synthesereaktionen der untersuchten Biokraftstoffe

Biokraftstoff	Reaktion
1-Butanol	$2 \text{ Acetyl-CoA} \Rightarrow \text{Acetoacetyl-CoA} + \text{CoA}$
	$\text{Acetoacetyl-CoA} + \text{NADPH} + \text{H}^+ \Rightarrow \text{3-Hydroxybutyryl-CoA} + \text{NADP}^+$
	$\text{3-Hydroxybutyryl-CoA} \Rightarrow \text{Crotonoyl-CoA} + \text{H}_2\text{O}$
	$\text{Crotonoyl-CoA} + \text{NADH} + \text{H}^+ \Rightarrow \text{Butyryl-CoA} + \text{NAD}^+$
	$\text{Butyryl-CoA} + \text{NADH} + \text{H}^+ \Rightarrow \text{Butyrylaldehyd} + \text{CoA} + \text{NAD}^+$
	$\text{Butyrylaldehyd} + \text{NADPH} + \text{H}^+ \Rightarrow \text{1-Butanol} + \text{NADP}^+$
Ethanol	$\text{Pyruvat} \Rightarrow \text{Acetaldehyd} + \text{CO}_2$
	$\text{Acetaldehyd} + \text{NADPH} + \text{H}^+ \Rightarrow \text{Ethanol} + \text{NADP}^+$
Ethylen	$3 \text{ 2-Oxoglutarat} + 3 \text{ O}_2 + \text{L-Arginin} \Rightarrow 2 \text{ Ethylen} + 7 \text{ CO}_2 + \text{Succinat} + \text{Guanidin} + \text{1-Pyrrolin-5-Carboxylat}$
Isobutanol	$\text{3-Methyl-2-Oxobutanoat} \Rightarrow \text{Isobutyrylaldehyd} + \text{CO}_2$
	$\text{Isobutyrylaldehyd} + \text{NADH} + \text{H}^+ \Rightarrow \text{Isobutanol} + \text{NAD}^+$
Isopren	$\text{Dimethylallylpyrophosphat} \Rightarrow \text{Isopren} + \text{Diphosphat}$

A.2 Vollständige Tabelle der CASOP-Werte für alle Reaktionen des beschnittenen *Synechocystis* Modells (*SynePruned*)

Z-Score	Reak-ID	Z-Score	Reak-ID	Z-Score	Reak-ID	Z-Score	Reak-ID	Z-Score	Reak-ID
0	CA0008	0,0265	GS0040	0	PG0014	0	PR0034	0	VS0002
0	GL0020	0	PG0006	0	AL0016	0,0265	GS0042	0	HE0002
0	LI0054	0	PR0042	0	RI0008	0	TE0010	0	BM0003
0	RI0009	0	GM0026	0	LI0062	0	CP0015	0	QT0010
0	AG0008	0	PG0016	0	TE0011	-0,09	GM0010	0	LI0002
0	IL0017	0	LI0063	0	AL0013	0	SS0005	0	PU0022
0	GL0016	0	LI0021	0	GL0001	0	LI0042	0	TE0003
0	ST0007	0	AA0009	0	LI0001	0	LI0003	0	GL0013
0	TR0038	0	TE0004	-0,019	PT0036	0,0106	PP0012	0	RI0006
0	AG0044	0	GE0008	0,0718	PP0006	0	TR0036	0	PT0038
0	HI0002	-0,0643	GM0018	0	PU0013	0	VT0023	0	PU0020
0	VT0008	0	PC0010	0	LI0055	0,0974	PT0011	0	LI0064

A Anhang

Z-Score	Reak-ID	Z-Score	Reak-ID	Z-Score	Reak-ID	Z-Score	Reak-ID	Z-Score	Reak-ID
0	PU0015	0	AL0012	0	TR0050	0	VS0001	0	PT0016
0	AG0038	0	LI0066	0	FO0011	0	VT0015	0,0243	AG0039
0,0957	ST0014	0	TR0002	0	PY0012	0	VT0003	-0,0144	TR0053
0	PY0011	0	LI0011	0	VT0016	0	AA0003	0	QT0019
0	PG0010	0	AA0017	0	GM0027	0	IL0014	0	GE0007
0	NI0006	0	TR0044	0	LI0027	0	PU0004	0	GS0019
0	TR0018	0	AL0005	0	RI0003	0	TE0002	0	IL0002
0	LI0050	0	AL0004	0	PT0014	0	PC0003	0	VS0005
0	AL0009	0	QT0004	0	LI0051	0	CA0004	0	AL0011
0	TR0023	0	HE0003	0	AG0009	0	PU0010	0	PU0021
0	GL0003	-0,018	PT0013	0	TR0035	0,0649	PT0003	0	GM0012
0,0243	AG0043	0	PG0008	0	RI0001	0	AA0014	0	GM0033
0	TR0006	0	RI0002	0	LI0067	0	LI0015	0	AG0033
0	PP0007	0	TR0039	0	GL0018	0	PU0018	0	QT0020
0	CA0010	0	PG0018	-0,0144	TR0049	0	TP0003	0	IL0016
0	RI0004	0	HI0008	0	HI0009	0	PU0028	0	CP0018
0	AA0001	0	CP0006	0	PP0015	0	AA0010	0	FO0005
0	AG0004	0	GM0032	0	AA0007	0	PT0015	0	QT0015
0,0032	PP0010	0	NI0001	0	LI0059	0	LI0068	0	HI0004
-0,0095	PT0034	0	CA0007	0	LI0030	0	PC0006	0,0976	GS0014
0	PY0014	0	GM0001	-0,0144	PT0027	0	PR0050	0	PT0012
0	AA0006	0	AG0030	0	AL0001	0	CP0013	0	CP0008
0	LI0058	0	FO0007	0	NI0007	0	CP0004	0	RI0005
-0,0144	PP0016	0	PC0001	0	PY0016	0	PU0012	0	NI0009
0	TP0005	0	QT0017	0	BM0005	0	QT0001	0	PU0017
0	GE0011	0	GL0007	0	VT0022	0	PG0011	0	QT0005
0	TP0001	0	PY0005	0	FO0001	0,0718	PP0005	0	LI0069
0	CP0021	0	TR0055	0	AL0006	0,0062	GS0013	0	CA0003
0	FO0010	0	QT0014	0	PP0011	0	TR0037	0	NI0002
0,0454	GM0011	0	PR0003	0	TE0012	0	QT0016	0	QT0002
0	AA0013	0	PU0024	0	GS0038	0	AL0008	0	AA0004
0	VT0017	0	GE0015	0	LI0004	0	PY0015	0	VT0001
0	PC0011	0	GL0011	0	AL0007	0	PY0022	0	CA0005
0	PP0013	0	LI0019	0	NI0003	0	RI0007	0,0265	GS0041
0	BM0004	0,0243	AG0042	0	VT0010	0	PT0010	0	IL0015
0	VT0020	0	LI0036	0	HI0005	0	TE0008	0	GM0006
0,1764	TR0003	0	AG0020	0	FO0012	0	CA0012	0	BM0007
0	PU0029	0	LI0071	0	LI0061	-0,0382	GM0020	0	VT0004
-0,0185	PT0024	0	GM0004	0	PU0027	0	FO0017	0,0425	PP0003
0	PC0004	0	CP0007	0	GM0031	0	PR0036	0,0062	GS0011
0	LI0013	0	GL0017	-0,0095	PT0037	0	PT0022	0	IL0011
0	FO0018	0	VT0002	0	LI0049	0	TR0015	0	TR0047
0	PU0002	0	SS0008	0	PU0030	0	CA0009	0,0718	PP0004
0	BM0008	0,2994	PT0001	0	HI0011	0	AA0008	0	PG0007
0	TR0026	-0,1236	PR0006	0	NI0004	-0,0422	GS0043	0	CP0005
0,0957	PP0024	0,0106	PP0001	0,0243	AG0041	0,0106	GS0004	0	IL0012
0	AG0015	0	PU0014	0	PT0019	0	VT0006	0	TR0052
0	VT0013	0	PP0009	0	AG0029	0	SS0004	0	GM0002
0	LI0039	0	LI0024	0	AL0002	0	CA0001	0	QT0021
0	GL0006	0	HI0012	0	FO0004	0	FO0013	-0,189	PR0004

A.2 Vollständige Tabelle der CASOP-Werte

Z-Score	Reak-ID	Z-Score	Reak-ID	Z-Score	Reak-ID	Z-Score	Reak-ID	Z-Score	Reak-ID
0	QT0018	0	GM0008	0	AG0024	0	IL0005	0,0243	AG0040
0	PU0001	0	PG0004	0	PY0004	0	BM0006	0	IL0008
0	ST0004	0	PY0020	0	FO0006	-0,0406	PP0021	0	VT0007
-0,0438	PP0020	0	PY0026	0	CA0006	0	TR0051	0	AG0007
0	LI0017	0	GL0002	0	LI0009	0	TE0013	0	PY0021
0	PC0008	0	TE0001	0	GM0007	0	IL0009	0	QT0013
0	PU0011	0	IL0007	0	CP0014	0	AG0028	0	PY0025
0	PU0005	0	VS0004	0	AL0010	0	PU0008	0	AA0015
0	PR0002	0	VT0005	0	TE0007	0	TP0002	0	PR0001
0	TR0046	0	HI0001	0	PT0006	0	GE0004	0	PG0015
0	LI0070	0	PU0006	0	AA0019	0	AA0021	1	PDC
0	QT0003	0	CP0024	0	IL0010	0	IL0001	1	EtOH-ex
0	SS0007	0	CP0011	0	GL0012	-0,0382	GM0019	0	PR0043
0	CA0002	0	VT0021	0	GL0004	-0,189	PR0014	0	PR0044
0	GL0014	0	PY0009	0	PU0023	0	PY0010	0	PR0045
0	HI0003	0	PC0005	0	AG0023	0	LI0060	0	PR0046
0	PU0007	0	LI0046	0	IL0006	0	ST0001	0	PR0047
0	AA0012	1	PT0009	0	PG0003	0	PY0002	0	PR0048
0	PG0013	0	TR0041	0	LI0056	0	RI0011	0	PR0049
0	FO0015	0	GL0015	0,0957	ST0016	0	LI0033	0	PR0035
0	IL0003	0	TR0045	0	LI0065	0	AA0005	0	PR0022
0	VS0003	0	LI0053	0	TR0032	0	PU0026	0	PR0023
-0,0095	PT0035	0	TE0005	0	RI0010	0	PG0012	0	PR0024
0	LI0007	0	PG0001	0	AL0017	0	TE0006	0	PR0025
0	TP0004	0	CP0003	0,0187	GM0016	0	GS0018	0	PR0026
0	GM0005	0	GS0003	0,022	GM0025	0	GE0002	0	PR0027
0,0106	GS0008	0	IL0004	0	HE0001	0	FO0003	-0,1412	PR0018
0	PT0020	0	VS0007	0	LI0052	0	PC0007	0	PT0025
0	PG0002	0	CP0025	0	LI0047	0	CP0027	-0,1176	Mehler
0	PP0014	0	TR0028	0,0054	GS0010	0	GM0015	-0,1176	GE0010
0	VT0018	0	PT0018	0	PU0009	0	PY0007	0,0957	GE0001
0	PY0024	0	VT0019	0	ST0002	0	BM0001	0	TR0025
0,0638	GS0012	0	BM0002	0	LI0045	0	LI0057	-0,0721	R20131
0	GL0019	0	HI0007	0	TE0014	0	LI0048	-0,0721	R20132
0	CP0001	0	QT0011	0	CP0002	0	GL0005	-0,0721	R20133
0	VS0006	0	AA0020	0	TR0020	0	AG0025	-0,1412	FQR
0,1276	GS0006	0	PG0005	0	GM0013	0	GE0003	-0,0854	GE0009
0	PT0017	0,0187	GM0017	0	BM0009	0	PG0009	-0,0844	PT0026

A.3 Vollständige Ergebnisse der FVA für das Szenario: aerobes
Wachstum: $\mu = 0,85 \text{ h}^{-1}$ im *ColiPrunedComp*, *ColiCore* und
ColiGS Modell

Reaktions-ID	<i>ColiPrunedComp</i>		<i>ColiCore</i>		<i>ColiGS</i>	
	r-min	r-max	r-min	r-max	r-min	r-max
R-AKGDH	0	7,237	0	7,081	0	11,151
R-ATPM	8,390	28,046	8,390	13,087	8,390	28,092
R-CS	0,914	8,151	2,742	7,998	0,914	26,767
R-CYTBDpp	0	39,312	41,515	45,690	0	39,404
R-CYTBO3-4pp	7,453	46,765			7,366	46,770
R-EX-ac-e-	0	3,024	0	1,044	0	3,031
R-EX-acald-e-	0	2,125	0	0,696	0	2,130
R-EX-akg-e-	0	1,229	0	0,391	0	1,231
R-EX-etoh-e-	0	1,829	0	0,606	0	1,833
R-EX-for-e-	0	9,823	0	3,132	0	9,850
R-EX-glc-e-	9,181	10	9,739	10	9,179	10
R-EX-glu-L-e-	0	1,092	0	0,348	0	1,095
R-EX-lac-D-e-	0	1,787	0	0,587	0	1,791
R-EX-pyr-e-	0	2,125	0	0,696	0	2,130
R-EX-succ-e-	0	1,430	0	0,458	0	1,433
R-FBP	0	19,656	0	4,697	0	19,702
R-FORt2pp	0	78,625	0	18,790	0	78,808
R-FORtppi	0	78,625	0	18,790	0	78,808
R-GLCptspp	9,181	10	9,739	10	0	10
R-GLNS	1,536	21,193	0,217	4,915	1,536	21,238
R-GLUN	0	19,656	0	4,697	0	19,702
R-GLUSy	0	19,656	0	4,697	0	19,702
R-GND	0	21,711	0	15,770	0	27,759
R-ICL	0	7,237	0	5,731	0	15,165
R-MALS	0,001	7,238	0	5,731	0	9,771
R-ME1	0	12,631	0	4,077	0	16,012
R-ME2	0	15,756	0	5,352	0	18,904
R-NADH16pp	28,764	46,481	36,726	43,866	0	46,486
R-NADTRHD	0	47,887	0	17,125	0	52,831
R-PDH	0	18,081	6,205	15,564	0	19,500
R-PFK	2,026	27,867	3,898	12,353	0	27,913
R-PFL	0,092	11,106	0	3,132	0	19,920
R-PGL	0	21,711	0	15,770	0	28,493
R-PPC	0	22,875	0	7,787	0	23,032
R-PPCK	0	19,656	0	4,697	0	19,702
R-PPS	0	19,656	0	4,697	0	19,702
R-PYK	0	23,289	0	8,130	0	30,801
R-SUCct2-2pp	0	26,208	0	6,263	0	26,269
R-SUCct3pp	0	26,208	0	6,263	0	26,269
R-SUCDi	0	1000	1,825	1000	0	1000
R-THD2pp	0	41,944	0	13,141	0	42,035
R-RFD-new	0	1000	0	998,180	0	1000
Bio-T	0,850	0,850	0,850	0,850	0,850	0,850

A.3 Ergebnisse der FVA: aerobes Wachstum: $\mu = 0,85 \text{ h}^{-1}$

Reaktions-ID	<i>ColiPrunedComp</i>		<i>ColiCore</i>		<i>ColiGS</i>	
	r-min	r-max	r-min	r-max	r-min	r-max
R-ACALD	-2,125	0	-0,696	0	-19,702	52,831
R-ACALDtex	-2,125	0			-2,130	0
R-ACALDtpp	-2,125	0	-0,696	0	-2,130	0
R-ACKr	-2,530	0,494	-1,044	0	-1000	20,196
R-ACONTa	0,914	8,151	2,742	7,998	0,914	16,079
R-ACONTb	0,914	8,151	2,742	7,998	0,914	16,079
R-ACt2rpp	-3,024	0	-1,044	0	-1000	0
R-ACtex	-3,024	0			-3,031	0
R-ADK1	1,925	21,581	0	4,697	-1000	1000
R-AKGt2rpp	-1,229	0	-0,391	0	-1,231	0
R-AKGtex	-1,229	0	-0,606	0	-1,231	0
R-ALCD2x	-1,829	0			-1,833	0
R-ATPS4rpp	53,034	79,434	40,187	50,700	53,031	82,795
R-CO2tex	-25,190	-15,367			-25,193	-15,343
R-CO2tpp	-25,190	-15,367	-23,828	-20,696	-25,193	-15,343
R-D-LACt2rpp	-1,787	0	-0,587	0	-1,791	0
R-D-LACtex	-1,787	0			-1,791	0
R-ENO	9,269	16,506	11,212	16,469	-0,491	24,434
R-ETOHt2rpp	-1,829	0	-0,606	0	-1,833	0
R-ETOHtex	-1,829	0			-1,833	0
R-EX-co2-e-	15,367	25,190	20,696	23,828	15,343	25,193
R-EX-h2o-e-	36,567	47,580	26,888	30,020	36,539	552,360
R-EX-h-e-	7,807	17,630	17,051	20,183	-995,610	17,657
R-EX-nh4-e-	-10,259	-9,167	-4,983	-4,635	-16,019	-9,166
R-EX-o2-e-	-23,383	-17,335	-22,845	-20,757	-273,380	-17,323
R-EX-pi-e-	-0,817	-0,817	-3,127	-3,127	-1,790	-0,817
R-FBA	2,026	9,263	3,898	9,155	-11,547	21,521
R-FORtex	-9,823	0			-9,850	0
R-FUM	0,232	8,122	1,825	7,081	-0,097	53,715
R-G6PDH2r	0	21,711	0	15,770	0	28,493
R-GAPD	10,729	17,966	12,484	17,740	3,867	17,967
R-GLCtex	9,181	10			9,179	10
R-GLUDy	-8,352	12,396	-4,766	0,280	-60,091	12,407
R-GLUt2rpp	-1,092	0	-0,348	0	-1000	4,636
R-GLUtex	-1,092	0			-1,095	0
R-H2Otex	-47,580	-36,567			-552,360	-36,539
R-H2Otp	-47,580	-36,567	-30,020	-26,888	-67,284	-36,539
R-Htex	-17,630	-7,807			-17,657	995,610
R-ICDHyr	0,914	8,151	0,917	7,998	0,914	12,065
R-LDH-D	-1,787	0	-0,587	0	-26,269	9,001
R-MDH	-4,015	15,359	2,078	10,588	-19,246	53,716
R-NH4tex	9,167	10,259			9,166	16,019
R-NH4tpp	9,167	10,259	4,635	4,983	9,166	28,868
R-O2tex	17,335	23,383			17,323	273,380
R-O2tpp	17,335	23,383	20,757	22,845	17,323	23,385
R-PGI	-11,711	10	-5,944	9,826	-18,493	10
R-PGK	-17,966	-10,729	-17,740	-12,484	-17,967	-3,867
R-PGM	-16,506	-9,269	-16,469	-11,212	-24,434	0,491
R-PIt2rpp	0,817	0,817	3,127	3,127	-25,452	9,574

A Anhang

Reaktions-ID	<i>ColiPrunedComp</i>		<i>ColiCore</i>		<i>ColiGS</i>	
	r-min	r-max	r-min	r-max	r-min	r-max
R-PTAr	-0,494	2,530	0	1,044	-20,196	1000
R-PYRt2rpp	-2,125	0	-0,696	0	-2,130	0
R-PYRtex	-2,125	0			-2,130	0
R-RPE	-0,656	13,818	-0,611	9,902	-6,800	17,850
R-RPI	-7,860	-0,623	-5,868	-0,611	-11,805	-0,595
R-SUCCtex	-1,430	0			-1,433	0
R-SUCOAS	-6,791	0,446	-7,081	0	-1000	9,334
R-TALA	-0,170	7,067	-0,152	5,104	-3,242	9,083
R-TKT1	-0,170	7,067	-0,152	5,104	-3,242	9,083
R-TKT2	-0,486	6,751	-0,459	4,798	-3,558	8,767
R-TPI	1,906	9,143	3,898	9,155	-4,956	9,144

A.4 Ergebnisse der FVA: anaerobes Wachstum: $\mu = 0,2 \text{ h}^{-1}$

A.4 Vollständige Ergebnisse der FVA für das Szenario: anaerobes Wachstum: $\mu = 0,2 \text{ h}^{-1}$ im *ColiPrunedComp*, *ColiCore* und *ColiGS* Modell

Reaktions-ID	<i>ColiPrunedComp</i>		<i>ColiCore</i>		<i>ColiGS</i>	
	r-min	r-max	r-min	r-max	r-min	r-max
R-AKGDH	0	1,074	0	0,383	0	1,075
R-ATPM	8,390	11,613	8,390	9,443	8,390	11,614
R-CS	0,215	1,289	0,216	0,599	0,215	3,439
R-CYTBDpp	0	0	0	0	0	0
R-CYTBO3-4pp	0	0			0	0
R-EX-ac-e-	3,088	8,530	3,803	8,586	3,074	10,942
R-EX-acald-e-	0	5,157	0	1,685	0	5,158
R-EX-akg-e-	0	1,031	0	0,366	0	1,032
R-EX-etoh-e-	5,822	13,352	6,970	12,586	5,822	13,357
R-EX-for-e-	0	12,892	9,502	17,926	0	12,894
R-EX-glc-e-	9,008	10	9,617	10	9,008	10
R-EX-glu-L-e-	0	0,992	0	0,351	0	0,992
R-EX-lac-D-e-	0	5,157	0	2,808	0	5,158
R-EX-pyr-e-	0	3,223	0	1,404	0	3,224
R-EX-succ-e-	0,067	2,645	0	1,053	0,067	2,646
R-FBP	0	3,223	0	1,053	0	3,224
R-FORt2pp	0	12,892	0	4,212	0	12,894
R-FORtppi	0	12,892	9,502	22,138	0	12,894
R-GLCptspp	9,008	10	9,617	10	0	10
R-GLNS	0,362	3,585	0,051	1,104	0,362	3,585
R-GLUN	0	3,223	0	1,053	0	3,224
R-GLUSy	0	3,223	0	1,053	0	3,224
R-GND	0	3,634	0	2,290	0	3,634
R-ICL	0	1,031	0	0,337	0	1,053
R-MALS	0,000	1,032	0	0,337	0	1,032
R-ME1	0	3,223	0	1,053	0	3,224
R-ME2	0	4,694	0	2,106	0	4,694
R-NADH16pp	0	2,578	0	1,053	0	2,579
R-NADTRHD	0	6,446	0	2,106	0	6,447
R-PDH	0	8,595	0	8,424	0	8,596
R-PFK	8,477	13,050	8,977	10,854	0	13,050
R-PFL	9,206	17,800	9,502	17,926	9,204	17,800
R-PGL	0	3,634	0	2,290	0	4,694
R-PPC	0,659	5,352	0,573	2,679	0,659	5,396
R-PPCK	0	3,223	0	1,053	0	3,224
R-PPS	0	3,223	0	1,053	0	3,224
R-PYK	3,664	11,580	6,386	9,545	0	19,081
R-SUCCt2-2pp	0	4,297	0	1,404	0	4,298
R-SUCCt3pp	0,067	4,364	0	1,404	0,067	4,365
R-SUCDi	0	1000,000	0	1000	0	1000
R-THD2pp	0	9,387	0	5,535	0	9,388
R-RFD-new	0	1000	0	1000	0	1000
Bio-T	0,200	0,200	0,200	0,200	0,200	0,200

A Anhang

Reaktions-ID	<i>ColiPrunedComp</i>		<i>ColiCore</i>		<i>ColiGS</i>	
	r-min	r-max	r-min	r-max	r-min	r-max
R-ACALD	-13,352	-5,822	-12,586	-6,970	-13,357	-1,954
R-ACALDt _{tex}	-5,157	0			-5,158	0
R-ACALDt _{pp}	-5,157	0	-1,685	0	-5,158	0
R-ACKr	-8,413	-2,972	-8,586	-3,803	-1000	-2,958
R-ACONTa	0,215	1,289	0,216	0,599	0,215	1,290
R-ACONTb	0,215	1,289	0,216	0,599	0,215	1,290
R-ACt _{2rpp}	-8,530	-3,088	-8,586	-3,803	-1000,000	-3,074
R-ACt _{tex}	-8,530	-3,088			-10,942	-3,074
R-ADK1	0,453	3,676	0	1,053	-1000	1000
R-AKGt _{2rpp}	-1,031	0	-0,366	0	-1,032	0
R-AKGt _{tex}	-1,031	0			-1,032	0
R-ALCD2x	-13,352	-5,822	-12,586	-6,970	-13,357	-5,822
R-ATPS4 _{rpp}	-3,159	2,772	-6,342	-0,987	-3,160	2,774
R-CO2 _{tex}	-20,104	-4,790			-20,556	-4,788
R-CO2t _{pp}	-20,104	-4,790	-8,067	1,410	-20,556	-4,788
R-D-LACt _{2pp}	-5,157	0	-2,808	0	-5,158	0
R-D-LACt _{tex}	-5,157	0			-5,158	0
R-ENO	16,968	19,178	18,092	19,169	14,484	19,915
R-ETOHt _{2rpp}	-13,352	-5,822	-12,586	-6,970	-13,357	-5,822
R-ETOHt _{tex}	-13,352	-5,822			-13,357	-5,822
R-EX-co2-e-	4,790	20,104	-1,410	8,067	4,788	20,556
R-EX-h2o-e-	-3,162	1,135	-7,275	-3,063	-3,781	1,136
R-EX-h-e-	5,059	23,392	17,888	31,577	5,044	23,394
R-EX-nh4-e-	-3,149	-2,157	-1,442	-1,091	-4,068	-2,157
R-EX-o2-e-	0	0	0	0	0	0
R-EX-pi-e-	-0,192	-0,192	-0,736	-0,736	-0,595	-0,192
R-FBA	8,477	9,827	8,977	9,801	-2,285	11,976
R-FORT _{tex}	-12,892	0			-12,894	0
R-FUM	-2,437	0,141	-1,053	0	-2,437	6,588
R-G6PDH2r	0	3,634	0	2,290	0	4,694
R-GAPD	17,312	19,521	18,391	19,468	14,827	19,521
R-GLCt _{tex}	9,008	10			9,008	10
R-GLUDy	-2,700	1,515	-1,390	0,014	-8,155	1,515
R-GLUt _{2rpp}	-0,992	0	-0,351	0	-1000	0,758
R-GLUt _{tex}	-0,992	0			-0,992	0
R-H2O _{tex}	-1,135	3,162			-1,136	3,781
R-H2Ot _{pp}	-1,135	3,162	3,063	7,275	-1,136	3,781
R-Ht _{tex}	-23,392	-5,059			-23,394	-5,044
R-ICDHyr	0,215	1,289	0,216	0,599	0,215	1,290
R-LDH-D	-5,157	0	-2,808	0	-5,158	1,612
R-MDH	-4,552	1,173	-2,106	0,337	-4,553	6,588
R-NH4t _{tex}	2,157	3,149			2,157	4,068
R-NH4t _{pp}	2,157	3,149	1,091	1,442	2,157	5,380
R-O2t _{tex}	0	0			0	0
R-O2t _{pp}	0	0	0	0	0	0
R-PGI	6,366	10	7,669	9,959	5,306	10
R-PGK	-19,521	-17,312	-19,468	-18,391	-19,521	-14,827
R-PGM	-19,178	-16,968	-19,169	-18,092	-19,915	-14,484
R-Plt _{2rpp}	0,192	0,192	0,736	0,736	-4,106	1,625

A.5 Minimalmodell von *Synechocystis* sp. PCC 6803 (*SyneMin*)

Reaktions-ID	<i>ColiPrunedComp</i>		<i>ColiCore</i>		<i>ColiGS</i>	
	r-min	r-max	r-min	r-max	r-min	r-max
R-PTAr	2,972	8,413	3,803	8,586	2,958	1000
R-PYRt2rpp	-3,223	0	-1,404	0	-3,224	0
R-PYRtex	-3,223	0			-3,224	0
R-RPE	-0,154	2,268	-0,144	1,383	-1,512	2,268
R-RPI	-1,358	-0,146	-0,907	-0,144	-1,984	-0,119
R-SUCCtex	-2,645	-0,067			-2,646	-0,067
R-SUCOAS	-0,969	0,105	-0,383	0	-1000	1,717
R-TALA	-0,040	1,171	-0,036	0,728	-0,719	1,171
R-TKT1	-0,040	1,171	-0,036	0,728	-0,719	1,171
R-TKT2	-0,114	1,097	-0,108	0,656	-0,793	1,097
R-TPI	8,449	9,798	8,977	9,801	5,104	9,798

A.5 Minimalmodell von *Synechocystis* sp. PCC 6803 (*SyneMin*)

Reaktions-ID	Reaktion
PP0007	1 D-Ribose-5-Phosphat-cyt = 1 D-Ribulose-5-Phosphate-cyt
GS0008	1 D-Fructose-1,6-Bisphosphat-cyt = 1 Glyceronphosphat-cyt + 1 D-Glyceraldehyd-3-Phosphat-cyt
PP0014	1 D-Ribulose-5-Phosphate-cyt + 1 ATP-cyt = 1 ADP-cyt + 1 D-Ribulose-1,5-Bisphosphat-cyt
PP0009	1 Sedoheptulose-7-Phosphat-cyt + 1 D-Glyceraldehyd-3-Phosphat-cyt = 1 D-Ribose-5-Phosphat-cyt + 1 D-Xylulose-5-Phosphat-cyt
PT0009	1 NADPH-cyt + 1 Acetaldehyd-cyt + 1 H ⁺ -cyt = 1 Ethanol-cyt + 1 NADP ⁺ -cyt
GS0010	1 D-Glyceraldehyd-3-Phosphat-cyt = 1 Glyceronephosphat-cyt
GS0004	1 D-Fructose-1,6-Bisphosphat-cyt + 1 H ₂ O-cyt = 1 D-Fructose-6-Phosphat-cyt + 1 Orthophosphat-cyt
GS0018	1 2-Phospho-D-Glycerat-cyt = 1 Phosphoenolpyruvat-cyt + 1 H ₂ O-cyt
PP0010	1 D-Fructose-6-Phosphat-cyt + 1 D-Glyceraldehyd-3-Phosphat-cyt = 1 D-Xylulose-5-Phosphat-cyt + 1 D-Erythrose-4-Phosphat-cyt
PP0013	1 CO-2-cax + 1 D-Ribulose-1,5-Bisphosphat-cax + 1 H ₂ O-cax = 2 3-Phospho-D-Glycerat-cax
PT0001	1 Phosphoenolpyruvat-cyt + 1 ADP-cyt = 1 ATP-cyt + 1 Pyruvat-cyt
PP0001	1 D-Erythrose-4-Phosphat-cyt + 1 Glyceronephosphat-cyt = 1 Sedoheptulose-1,7-Bisphosphat-cyt
TR0050	1 HCO ₃ ⁻ -cyt = 1 HCO ₃ ⁻ -cax
PP0011	1 D-Ribulose-5-Phosphate-cyt = 1 D-Xylulose-5-Phosphat-cyt
GE0001	1 ATP-cyt + 1 H ₂ O-cyt = 1 ADP-cyt + 1 Orthophosphat-cyt
PP0012	1 Sedoheptulose-7-Phosphat-cyt + 1 D-Glyceraldehyd-3-Phosphat-cyt = 1 D-Fructose-6-Phosphat-cyt + 1 D-Erythrose-4-Phosphat-cyt
GS0013	1 ATP-cyt + 1 3-Phospho-D-Glycerat-cyt = 1 ADP-cyt + 1 3-Phospho-D-glyceroylphosphat-cyt
TR0051	1 D-Ribulose-1,5-Bisphosphat-cyt = 1 D-Ribulose-1,5-Bisphosphat-cax
GS0014	1 2-Phospho-D-Glycerat-cyt = 1 3-Phospho-D-Glycerat-cyt
PP0003	1 Sedoheptulose-1,7-Bisphosphat-cyt + 1 H ₂ O-cyt = 1 Sedoheptulose-7-Phosphat-cyt
GS0011	1 3-Phospho-D-glyceroylphosphat-cyt + 1 NADPH-cyt + 1 H ⁺ -cyt = 1 D-Glyceraldehyd-3-Phosphat-cyt + 1 Orthophosphat-cyt + 1 NADP ⁺ -cyt
TR0052	1 3-Phospho-D-Glycerat-cax = 1 3-Phospho-D-Glycerat-cyt

A Anhang

Reaktions-ID	Reaktion
PR0001	1 Photon-ext = 1 Photon-tlm
PDC	1 Pyruvat-cyt = 1 CO ₂ -cyt + 1 Acetaldehyd-cyt
EtOH-ex	1 Ethanol-cyt = 1 Ethanol-ext
Bio-T	1 Biomass = 1 B0-ext
PR0004	1 NADPH-cyt + 5 H ⁺ -cyt + PQ = 4 H ⁺ -tll + 1 PQH ₂ + 1 NADP ⁺ -cyt
PR0036	2 red. Ferredoxin-cyt + 1 NADP ⁺ -cyt + 1 H ⁺ -cyt = 2 ox. Ferredoxin-cyt + 1 NADPH-cyt
PR0050	14 H ⁺ -tll + 3 ADP-cyt + 3 Orthophosphat-cyt = 3 ATP-cyt + 3 H ₂ O-cyt + 14 H ⁺ -cyt
TR0036	1 Natrium-ext + 1 H ⁺ -cyt = 1 Natrium-cyt + 1 C00080-ext
TR0003	1 CO ₂ -cyt = 1 CO ₂ -ext
TR0006	1 HCO ₃ ⁻ -ext + 1 Natrium-ext = 1 HCO ₃ ⁻ -cyt + 1 Natrium-cyt
TR0002	1 O ₂ -cyt = 1 O ₂ -ext
GE0004–GE0002–lumped	
1 HCO ₃ ⁻ -cax + 1 H ⁺ -cax = 1 CO–2-cax + 1 H ₂ O-cax	
PR0022–PR0023–PR0003–PR0026–PR0027–PR0024–PR0034–PR0025–PR0035–lumped	
0,5 PQH ₂ + 1 Photon-tlm + 1 H ⁺ -cyt + 1 ox. Ferredoxin-cyt = 0,5 PQ + 2 H ⁺ -tll + 1 red. Ferredoxin-cyt	
PR0002–PR0042–TR0055–PR0045–PR0044–PR0047–PR0046–PR0048–PR0043–PR0049–lumped	
2 Photon-tlm + 1 H ₂ O-tll + 2 H ⁺ -cyt + 1 PQ = 0,5 O ₂ -cyt + 2 H ⁺ -tll + 1 PQH ₂	
BSR	
0,6238 O ₂ -cyt + 5,6262 D-Fructose-6-Phosphat-cyt + 6,0988 HCO ₃ ⁻ -cyt + 0,8034 D-Ribose-5-Phosphat-cyt + 0,0014 D-Ribulose-5-Phosphat-cyt + 75,3408 ATP-cyt + 1,1483 3-Phospho-D-Glycerat-cyt + 0,1742 Natrium-cyt + 13,3317 NADPH-cyt + 73,5368 red. Ferredoxin-cyt + 0,1717 Glyceronphosphat-cyt + 2,7159 Pyruvat-cyt + 0,1 D-Glyceraldehyd-3-Phosphat-cyt + 3,6497 Phosphoenolpyruvat-cyt + 17,2401 H ₂ O-cyt + 101,9233 H ⁺ -cyt + 0,1704 Kalium-ext + 0,0068 Fe ³⁺ -ext + 0,7245 Orthophosphat-ext + 0,2103 Glycogengranulat-cyt + 0,1786 Sulfat-ext + 0,0045 Calcium-ext + 0,0065 Ferricytochromb1-cyt + 0,0075 Fe ²⁺ -ext + 8,9748 Nitrat-ext + 0,003 Molybdat-ext + 0,003 Zink-ext + 0,0277 Magnesium-ext + 0,003 Mangan-ext + 0,0033 Cobaltion-ext + 0,003 Kupfer-ext = 82,06 Orthophosphat-cyt + 6,6466 CO ₂ -cyt + 75,3408 ADP-cyt + 4,7815 D-Erythrose-4-Phosphat-cyt + 13,3317 NADP ⁺ -cyt + 73,5368 ox. Ferredoxin-cyt 1 Biomasse-ext + 0,0002 4-Cresol-ext + 0,0065 Ferrocycyochromb1-cyt + 0,0045 H ⁺ -ext + 0,0005 5-Deoxy-D-Ribose-ext + 0,0068 S-Methyl-5-thio-D-Ribulose-1-Phosphat-ext + 0,1704 Natrium-ext + 0,0002 CO-ext	

Literaturverzeichnis

- [1] Algenol.
<http://www.algenol.com/direct-to-ethanol/the-technology>.
Zugriff: 10.08.2015.
- [2] *Cyanothece* sp. ATCC 51142.
<http://bacmap.wishartlab.com/organisms/672>.
Zugriff: 03.12.2015.
- [3] Dechema (Bundesalgenstammtisch) Vortrag Paul Woods.
http://dechema.de/events_media/07_Woods+Roundtable+2012-p-3976.pdf.
Zugriff: 24.03.2016.
- [4] Schematischer Aufbau einer vegetativen Cyanobakterien-Zelle.
<http://www.spektrum.de/lexikon/biologie/cyanobakterien/16137>.
Zugriff: 24.03.2016.
- [5] *Synechococcus elongatus* PCC 7942.
<http://bacmap.wishartlab.com/organisms/289>.
Zugriff: 29.09.2015.
- [6] *Synechocystis* sp. PCC 6803.
<http://bioenergy.asu.edu/faculty/vermaas/biofactories.html>.
Zugriff: 29.09.2015.
- [7] Systems Biology Research Group - Other Organisms.
<http://systemsbiology.ucsd.edu/InSilicoOrganisms/OtherOrganisms>.
Zugriff: 7.10.2015.
- [8] S. Abarzua, S. Jakubowski, S. Eckert, and P. Fuchs. Biotechnological investigation for the prevention of marine biofouling ii. blue-green algae as potential producers of biogenic agents for the growth inhibition of microfouling organisms. *Botanica Marina*, 42(5):459 – 465, 1999.

- [9] R. Abed, S. Dobretsov, and K. Sudesh. Applications of cyanobacteria in biotechnology. *Journal of Applied Microbiology*, 106(1):1 – 12, 2009.
- [10] S. I. Allakhverdiev, V. Thavasi, V. D. Kreslavski, S. K. Zharmukhamedov, V. V. Klimov, S. Ramakrishna, D. A. Los, M. Mimuro, H. Nishihara, and R. Carpentier. Photosynthetic hydrogen production. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, 11(2–3):101 – 113, 2010.
- [11] J. Allen. Cyclic, pseudocyclic and noncyclic photophosphorylation: New links in the chain. *Trends in Plant Science*, 8(1):15 – 19, 2003.
- [12] J. F. Allen. Photosynthesis of atp: Electrons, proton pumps, rotors, and poise. *Cell*, 110(3):273 – 276, 2002.
- [13] M. M. Allen. Cyanobacterial cell inclusions. *Annual Review of Microbiology*, 38: 1 – 25, 1984.
- [14] H. Alper, J. Moxley, E. Nevoigt, G. R. Fink, and G. Stephanopoulos. Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science*, 314(5805):1565 – 1568, 2006.
- [15] S. L. Anderson and L. McIntosh. Light-activated heterotrophic growth of the cyanobacterium *Synechocystis* sp. strain PCC 6803: A blue-light-requiring process. *Journal of Bacteriology*, 173(9):2761 – 2767, 1991.
- [16] I. Andersson. Catalysis and regulation in rubisco. *Journal of Experimental Botany*, 59(7):1555 – 1568, 2008.
- [17] S. A. Angermayr, K. J. Hellingwerf, P. Lindblad, and M. J. T. de Mattos. Energy biotechnology with cyanobacteria. *Current Opinion in Biotechnology*, 20(3): 257 – 263, 2009.
- [18] S. A. Angermayr, A. G. Rovira, and K. J. Hellingwerf. Metabolic engineering of cyanobacteria for the synthesis of commodity products. *Trends in Biotechnology*, 33(6):352 – 361, 2015.
- [19] K. Asada, J. Allen, C. Foyer, and H. Matthijs. The water-water cycle as alternative photon and electron sinks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 355(1402):1419 – 1431, 2000.

- [20] S. Atsumi, A. F. Cann, M. R. Connor, C. R. Shen, K. M. Smith, M. P. Brynildsen, K. J. Chou, T. Hanai, and J. C. Liao. Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metabolic Engineering*, 10(6):305 – 311, 2008.
- [21] S. Atsumi, W. Higashide, and J. Liao. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nature Biotechnology*, 27:1177 – 1180, 2009.
- [22] M. Badger, S. Von Caemmerer, S. Ruuska, H. Nakano, A. Laisk, J. Allen, K. Asada, H. Matthijs, and H. Griffiths. Electron flow to oxygen in higher plants and algae: Rates and control of direct photoreduction (mehler reaction) and rubisco oxygenase. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 355(1402):1433 – 1446, 2000.
- [23] M. R. Badger and G. D. Price. CO₂ concentrating mechanisms in cyanobacteria: Molecular components, their diversity and evolution. *Journal of Experimental Botany*, 54(383):609 – 622, 2003.
- [24] J. E. Bailey. Toward a science of metabolic engineering. *Science*, 252(5013):1668 – 1675, 1991.
- [25] D. Bald, J. Kruip, and M. Rögner. Supramolecular architecture of cyanobacterial thylakoid membranes: How is the phycobilisome connected with the photosystems? *Photosynthesis Research*, 49(2):103 – 118, 1996.
- [26] K. Ballerstein, A. von Kamp, S. Klamt, and U.-U. Haus. Minimal cut sets in a metabolic network are elementary modes in a dual network. *Bioinformatics*, 28(3):381 – 387, 2012.
- [27] J. E. Barrick, D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh, D. Schneider, R. E. Lenski, and J. F. Kim. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature*, 461:1243 – 1247, 2009.
- [28] F. U. Battistuzzi, A. Feijao, and S. B. Hedge. A genomic timescale of prokaryote evolution: Insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evolutionary Biology*, 4:44, 2004.
- [29] D. A. Beard, S. dan Liang, and H. Qian. Energy balance for analysis of complex metabolic networks. *Biophysical Journal*, 83(1):79 – 86, 2002.

- [30] J. Becker, O. Zelder, S. Häfner, H. Schröder, and C. Wittmann. From zero to hero - design-based systems metabolic engineering of *Corynebacterium glutamicum* for l-lysine production. *Metabolic Engineering*, 13(2):159 – 168, 2011.
- [31] F. K. Bentley, A. Zurbriggen, and A. Melis. Heterologous expression of the mevalonic acid pathway in cyanobacteria enhances endogenous carbon partitioning to isoprene. *Molecular Plant*, 7(1):71 – 86, 2014.
- [32] P. Bhadury and P. Wright. Exploitation of marine algae: Biogenic compounds for potential antifouling applications. *Planta*, 219(4):561 – 578, 2004.
- [33] L. M. Blank, L. Kuepfer, and U. Sauer. Large-scale ^{13}C -flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biology*, 6(6):R49, 2005.
- [34] D. J. Bonfil, M. Ronen-Tarazi, D. Sültemeyer, J. Lieman-Hurwitz, D. Schatz, and A. Kaplan. A putative HCO_3^- transporter in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *FEBS Letters*, 430(3):236 – 240, 1998.
- [35] D. A. Bryant. *The Molecular Biology of Cyanobacteria*. Kluwer Academic Publishers, Dordrecht, 2004.
- [36] Bundesministerium für Bildung und Forschung. Nationale Forschungsstrategie Bioökonomie 2030. Technical report, Bundesministerium für Bildung und Forschung, 2010.
- [37] Bundesministerium für Wirtschaft und Energie. Energiedaten: Gesamtausgabe. Technical report, Bundesministerium für Wirtschaft und Energie, 2015.
- [38] A. Burgard, P. Pharkya, and C. Maranas. Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and Bioengineering*, 84(6):647 – 657, 2003.
- [39] A. M. Burjaa, B. Banaigs, E. Abou-Mansour, J. G. Burgess, and P. C. Wright. Marine cyanobacteria - a prolific source of natural products. *Tetrahedron*, 57(46):9347 – 9377, 2001.
- [40] M. E. Bushell, S. I. Sequeira, C. Khannapho, H. Zhao, K. F. Chater, M. J. Butler, A. M. Kierzek, and C. A. Avignone-Rossa. The use of genome scale metabolic flux variability analysis for process feed formulation based on an investigation of

- the effects of the *zwf* mutation on antibiotic production in *Streptomyces coelicolor*. *Enzyme and Microbial Technology*, 39(6):1347 – 1353, 2006.
- [41] C. R. Carere, R. Sparling, N. Cicek, and D. B. Levin. Third generation biofuels via direct cellulose fermentation. *International Journal of Molecular Sciences*, 9(7):1342 – 1360, 2008.
- [42] R. Carlson and F. Sreenc. Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: Creation of overall flux states. *Biotechnology and Bioengineering*, 86(2):149 – 162, 2004.
- [43] R. Carlson and F. Sreenc. Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: Identification of reactions. *Biotechnology and Bioengineering*, 85(1):1 – 19, 2004.
- [44] R. W. Carstenholz. *Bergey's Manual of Systematic Bacteriology*, chapter Cyanobacteria, pages 473 – 599. Springer New York, 2001.
- [45] R. Caspi, T. Altman, R. Billington, K. Dreher, H. Foerster, C. A. Fulcher, T. A. Holland, I. M. Keseler, A. Kothari, A. Kubo, et al. The metacyc database of metabolic pathways and enzymes and the biocyc collection of pathway/genome databases. *Nucleic Acids Research*, 42(D1):D459 – D471, 2014.
- [46] H. S. Choi, S. Y. Lee, T. Y. Kim, and H. M. Woo. *In Silico* identification of gene amplification targets for improvement of lycopene production. *Applied and Environmental Microbiology*, 76(10):3097 – 3105, 2010.
- [47] A. Chowdhury, A. R. Zomorodi, and C. D. Maranas. Bilevel optimization techniques in computational strain design. *Computers & Chemical Engineering*, 72: 363 – 372, 2015.
- [48] Y. Cohen, B. B. Jorgensen, N. P. Revsbech, and R. Poplwaski. Adaptation to hydrogen sulfide of oxygenic and anoxygenic photosynthesis among cyanobacteria. *Applied and Environmental Microbiology*, 51(2):398 – 407, 1986.
- [49] M. W. Covert, E. M. Knight, J. L. Reed, M. J. Herrgard, and B. O. Palsson. Integrating high-throughput and computational data elucidates bacterial networks. *Nature*, 429(6987):92 – 96, 2004.

- [50] M. W. Covert and B. O. Palsson. Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *Journal of Biological Chemistry*, 277: 28058 – 28064, 2002.
- [51] M. W. Covert, C. H. Schilling, and B. O. Palsson. Regulation of gene expression in flux balance models of metabolism. *Journal of Theoretical Biology*, 213(1): 73 – 88, 2001.
- [52] H.-U. Dahms, X. Ying, and C. Pfeiffer. Antifouling potential of cyanobacteria: A mini-review. *Biofouling*, 22(5 - 6):317 – 327, 2006.
- [53] Das Europäische Parlament und der Rat der Europäischen Union. Richtlinie 2009/28/EG des Europäischen Parlaments und des Rates. Technical report, Europäische Union, 2009.
- [54] E. Dasilva. The colours of biotechnology: Science, development and humankind. *Electronic Journal of Biotechnology*, 7(3), 2004.
- [55] L. de Figueiredo, A. Podhorski, A. Rubio, C. Kaleta, J. Beasley, S. Schuster, and F. Planes. Computing the shortest elementary flux modes in genome-scale metabolic networks. *Bioinformatics*, 25(23):3158 – 3165, 2009.
- [56] M.-D. Deng and J. Coleman. Ethanol synthesis by genetic engineering in cyanobacteria. *Applied and Environmental Microbiology*, 65(2):523 – 528, 1999.
- [57] J. Dexter and P. Fu. Metabolic engineering of cyanobacteria for ethanol production. *Energy & Environmental Science*, 2:857 – 864, 2009.
- [58] B. Dey, D. L. Lerner, P. Lusso, M. R. Boyd, J. H. Elder, and E. A. Berger. Multiple antiviral activities of cyanovirin-n: Blocking of human immunodeficiency virus type 1 gp120 interaction with CD4 and coreceptor and inhibition of diverse enveloped viruses. *Journal of Virology*, 74(10):4562 – 4569, 2000.
- [59] D. C. Ducat, J. C. Way, and P. A. Silver. Engineering cyanobacteria to generate high-value products. *Trends in Biotechnology*, 29(2):95 – 103, 2011.
- [60] V. A. Dzelnkalns and L. Bogorad. Stable transformation of the cyanobacterium *Synechocystis* sp. PCC 6803 induced by uv irradiation. *Journal of Bacteriology*, 165(3):964 – 971, 1986.

- [61] O. Ebenhöh and R. Heinrich. Evolutionary optimization of metabolic pathways. theoretical reconstruction of the stoichiometry of atp and nadh producing systems. *Bulletin of Mathematical Biology*, 63(1):21 – 55, 2001.
- [62] J. S. Edwards, R. U. Ibarra, and B. O. Palsson. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature Biotechnology*, 19(2):125 – 130, 2001.
- [63] J. S. Edwards and B. O. Palsson. The *Escherichia coli* MG1655 *in silico* metabolic genotype: Its definition, characteristics, and capabilities. *Proceedings of the National Academy of Sciences of the United States of America*, 97:5528 – 5533, 2000.
- [64] M. Eisenhut, W. Ruth, M. Haimovich, H. Bauwe, A. Kaplan, and M. Hagemann. The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. *Proceedings of the National Academy of Sciences*, 105(44):17199 – 17204, 2008.
- [65] M. Eisenhut, E. A. von Wobeser, L. Jonas, H. Schubert, B. W. Ibelings, H. Bauwe, H. C. Matthijs, and M. Hagemann. Long-term response toward inorganic carbon limitation in wild type and glycolate turnover mutants of the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Plant Physiology*, 144(4):1946 – 1959, 2007.
- [66] P. Erdrich, H. Knoop, R. Steuer, and S. Klamt. Cyanobacterial biofuels: New insights and strain design strategies revealed by computational modeling. *Microbial Cell Factories*, 13:128, 2014.
- [67] P. Erdrich, R. Steuer, and S. Klamt. An algorithm for the reduction of genome-scale metabolic network models to meaningful core models. *BMC Systems Biology*, 9:48, 2015.
- [68] European Commission. Towards a europe knowledge-based bioeconomy. Technical report, European Commission, 2004.
- [69] P. Fay. Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiological Reviews*, 56(2):340 – 373, 1992.
- [70] A. Feist, M. Herrgard, I. Thiele, J. Reed, and B. Palsson. Reconstruction of biochemical networks in microbial organisms. *Nature Reviews Microbiology*, 7(2): 129 – 143, 2009.

- [71] A. M. Feist, C. S. Henry, J. L. Reed, M. Krummenacker, A. R. Joyce, P. D. Karp, L. J. Broadbelt, V. Hatzimanikatis, and B. O. Palsson. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 orfs and thermodynamic information. *Molecular Systems Biology*, 3:121, 2007.
- [72] A. M. Feist and B. O. Palsson. The biomass objective function. *Current Opinion in Microbiology*, 13(3):344 – 349, 2010.
- [73] A. M. Feist, D. C. Zielinski, J. D. Orth, J. Schellenberger, M. J. Herrgard, and B. O. Palsson. Model-driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. *Metabolic Engineering*, 12(3):173 – 186, 2010.
- [74] X. Feng, A. Bandyopadhyay, B. Berla, L. Page, B. Wu, H. B. Pakrasi, and Y. J. Tang. Mixotrophic and photoheterotrophic metabolism in *Cyanothece* sp. ATCC 51142 under continuous light. *Microbiology*, 156(8):2566 – 2574, 2010.
- [75] E. D. Fleming and R. W. Castenholz. Effects of periodic desiccation on the synthesis of the uv-screening compound, scytonemin, in cyanobacteria. *Environmental Microbiology*, 9(6):1448 – 1455, 2007.
- [76] S. Fong, A. Burgard, C. Herring, E. Knight, F. Blattner, C. Maranas, and B. Palsson. In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnology and Bioengineering*, 91(5):643 – 648, 2005.
- [77] C. Formighieri. *Solar-to-Fuel Conversion in Algae and Cyanobacteria*. Springer, 2015.
- [78] P. Fu. Genome-scale modeling of *Synechocystis* sp. PCC 6803 and prediction of pathway insertion. *Journal of Chemical Technology and Biotechnology*, 84(4): 473 – 483, 2009.
- [79] K. Fukuda and A. Prodon. Double Description Method Revisited. In *Combinatorics and Computer Science*. Springer, 1996.
- [80] K. Gademann and C. Portmann. Secondary metabolites from cyanobacteria: Complex structures and powerful bioactivities. *Current Organic Chemistry*, 12(4): 326 – 341, 2008.
- [81] J. Gagneur and S. Klamt. Computation of elementary modes: A unifying framework and the new binary approach. *BMC Bioinformatics*, 5:175, 2004.

- [82] Z. Gao, H. Zhao, Z. Li, X. Tan, and X. Lu. Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy & Environmental Science*, 5(12):9857 – 9865, 2012.
- [83] M. P. Gerstl, D. E. Ruckerbauer, D. Mattanovich, C. Jungreuthmayer, and J. Zanghellini. Metabolomics integrated elementary flux mode analysis in large metabolic networks. *Scientific Reports*, 5:8930, 2015.
- [84] S. J. Goldman, P. J. Lammers, M. S. Berman, and J. Sanders-Loehr. Siderophore-mediated iron uptake in different strains of *Anabaena* sp. *Journal of Bacteriology*, 156(3):1144 – 1150, 1983.
- [85] G. Grigorieva and S. Shestakov. Transformation in the cyanobacterium *Synechocystis* sp. 6803. *FEMS Microbiology Letters*, 13(4):367 – 360, 1982.
- [86] A. R. Grossman, M. R. Schaefer, G. G. Chiang, and J. L. Collier. The phycobilisome, a light-harvesting complex responsive to environmental conditions. *Microbiological Reviews*, 57(3):725 – 749, 1993.
- [87] A. R. Grossman, M. R. Schaefer, G. G. Chiang, and J. L. Collier. *The Molecular Biology of Cyanobacteria*, chapter The Responses of Cyanobacteria to Environmental Conditions: Light and Nutrients, pages 641 – 675. Kluwer Academic Publishers, 1994.
- [88] I. Grotjohann and P. Fromme. Structure of cyanobacterial photosystem i. *Photosynthesis Research*, 85(1):51 – 72, 2005.
- [89] F. Guerrero, V. Carbonell, M. Cossu, D. Correddu, and P. R. Jones. Ethylene synthesis and regulated expression of recombinant protein in *Synechocystis* sp. PCC 6803. *PLoS ONE*, 7(11):e50470, 2012.
- [90] M. G. Guerrero and C. Lara. *The Cyanobacteria*, chapter Assimilation of Inorganic Nitrogen, pages 163 – 186. Elsevier Science Publishers, 1987.
- [91] C. Hackenberg, A. Engelhardt, H. C. P. Matthijs, F. Wittink, H. Bauwe, A. Kaplan, and M. Hagemann. Photorespiratory 2-phosphoglycolate metabolism and photoreduction of O₂ cooperate in high-light acclimation of *Synechocystis* sp. strain PCC 6803. *Planta*, 230(4):625 – 637, 2009.
- [92] O. Hädicke. *Modellbasierte Identifizierung von Interventionsstrategien für Stoffwechselnetzwerke*. PhD thesis, Otto-von-Guericke-Universität Magdeburg, 2014.

- [93] O. Hädicke and S. Klamt. CASOP: A computational approach for strain optimization aiming at high productivity. *Journal of Biotechnology*, 147(2): 88 – 101, 2010.
- [94] O. Hädicke and S. Klamt. Computing complex metabolic intervention strategies using constrained minimal cut sets. *Metabolic Engineering*, 13(2):204 – 213, 2011.
- [95] J. J. Hamilton and J. L. Reed. Identification of functional differences in metabolic networks using comparative genomics and constraint-based models. *PLoS ONE*, 7(4):e34670, 2012.
- [96] J. Hay and J. Schwender. Metabolic network reconstruction and flux variability analysis of storage synthesis in developing oilseed rape *Brassica napus L.* embryos. *The Plant Journal*, 67(3):526 – 541, 2011.
- [97] S. B. Hedges, H. Chen, S. Kumar, D. Y. Wang, A. S. Thompson, and H. Watanabe. A genomic timescale for the origin of eukaryotes. *BMC Evolutionary Biology*, 1:4, 2001.
- [98] S. Hein, H. Tran, and A. Steinbüchel. *Synechocystis* sp. PCC 6803 possesses a two-component polyhydroxyalkanoic acid synthase similar to that of anoxygenic purple sulfur bacteria. *Archives of Microbiology*, 170(3):162 – 170, 1998.
- [99] M. J. Herrgard, B.-S. Lee, V. Portnoy, and B. O. Palsson. Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*. *Genome Research*, 16(5):627 – 635, 2006.
- [100] A. P. Hertle, T. Blunder, T. Wunder, P. Pesaresi, M. Pribil, U. Armbruster, and D. Leister. PGRL1 is the elusive ferredoxin-plastoquinone reductase in photosynthetic cyclic electron flow. *Molecular Cell*, 49(3):511 – 523, 2013.
- [101] W. R. Hess. Cyanobacterial genomics for ecology and biotechnology. *Current Opinion in Microbiology*, 14(5):608 – 614, 2011.
- [102] H. Heyer and W. E. Krumbein. Excretion of fermentation products in dark and anaerobically incubated cyanobacteria. *Archives of Microbiology*, 155(3):284 – 287, 1991.
- [103] J. Hill, E. Nelson, D. Tilman, S. Polasky, and D. Tiffany. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences*, 103(30):11206 – 11210, 2006.

- [104] K. Hirata, S. Yoshitomi, S. Dwi, O. Iwabe, A. Mahakhant, J. Polchai, and K. Miyamoto. Bioactivities of nostocine produced by a freshwater cyanobacterium *Nostoc spongiaeforme* TISTR 8169. *Journal of Bioscience and Bioengineering*, 95(5): 512 – 517, 2003.
- [105] P. D. T. Hirth. Schriftliche Stellungnahme zum Fragenkatalog „Biotechnologische Verfahren“. Technical report, Fraunhofer IGB, 2014.
- [106] K. K. Ho and D. W. Krogmann. Electron donors to P700 in cyanobacteria and algae: An instance of unusual genetic variability. *Biochemica et Biophysica Acta*, 766(2):310 – 316, 1984.
- [107] D. S. Hoare, L. O. Ingram, E. L. Thurston, and R. Walkup. Dark heterotrophic growth of an endophytic blue-green alga. *Archives of Microbiology*, 78(4): 310 – 321, 1971.
- [108] S.-J. Hong and C.-G. Lee. Evaluation of central metabolism based on a genomic database of *Synechocystis* PCC 6803. *Biotechnology and Bioprocess Engineering*, 12(2):165 – 173, 2007.
- [109] A. Hoppe, S. Hoffmann, and H.-G. Holzhütter. Including metabolite concentrations into flux balance analysis: Thermodynamic realizability as a constraint on flux distributions in metabolic networks. *BMC Systems Biology*, 1:23, 2007.
- [110] M. Hucka, A. Finney, H.M.Sauro, H. Bolouri, J. C. Doyle, H. Kitano, and the rest of the SBML Forum. The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524 – 531, 2003.
- [111] R. Ibarra, J. Edwards, and B. Palsson. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*, 420(6912): 186 – 189, 2002.
- [112] Y. Ito and A. Butler. Structure of synechobactins, new siderophores of the marine cyanobacterium *Synechococcus* sp. PCC 7002. *Limnology and Oceanography*, 50(6):1918 – 1923, 2005.
- [113] J. H. Jacobsen and N.-U. Frigaard. Engineering of photosynthetic mannitol biosynthesis from CO₂ in a cyanobacterium. *Metabolic Engineering*, 21:60 – 70, 2014.

- [114] S. Jindou, Y. Ito, N. Mito, K. Uematsu, A. Hosoda, and H. Tamura. Engineered platform for bioethylene production by a cyanobacterium expressing a chimeric complex of plant enzymes. *ACS Synthetic Biology*, 3(7):487 – 496, 2014.
- [115] C. Jungreuthmayer, D. Ruckerbauer, and J. Zanghellini. regEfmtool: Speeding up elementary flux mode calculation using transcriptional regulatory rules in the form of three-state logic. *Biosystems*, 113(1):37 – 39, 2013.
- [116] C. Kaleta, L. F. de Figueiredo, J. Behre, and S. Schuster. EFMEvolver: Computing elementary flux modes in genome-scale metabolic networks. *Proceedings of the 14th German Conference on Bioinformatics (GCB)*, 2:180 – 190, 2009.
- [117] R. Kalscheuer, T. S. T., and A. Steinbüchel. Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology*, 152(9):2529 – 2536, 2006.
- [118] J. Kämäräinen, H. Knoop, N. Stanford, F. Guerrero, M. Akhtar, E.-M. Aro, R. Steuer, and P. Jones. Physiological tolerance and stoichiometric potential of cyanobacteria for hydrocarbon fuel production. *Journal of Biotechnology*, 162(1): 67 – 74, 2012.
- [119] M. Kanehisa and S. Goto. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1):27 – 30, 2000.
- [120] M. Kanehisa, S. Goto, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe. Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Research*, 42:D199 – D205, 2014.
- [121] T. Kaneko, S. Sato, H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima, M. Hirosawa, M. Sugiura, S. Sasamoto, T. Kimura, T. Hosouchi, A. Matsuno, A. Muraki, N. Nakazaki, K. Naruo, S. Okumura, S. Shimpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, and S. Tabata. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. ii. sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Research*, 3(3):109 – 136, 1996.
- [122] K. J. Kauffman, P. Prakash, and J. S. Edwards. Advances in flux balance analysis. *Current Opinion in Biotechnology*, 14(5):491 – 496, 2003.
- [123] J. Kim and J. Reed. OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Systems Biology*, 4:53, 2010.

- [124] J. Kim, J. L. Reed, and C. T. Maravelias. Large-scale bi-level strain design approaches and mixed-integer programming solution techniques. *PLoS ONE*, 6(9):e24162, 2011.
- [125] S. Klamt. Generalized concept of minimal cut sets in biochemical networks. *BioSystems*, 83(2-3):233 – 247, 2006.
- [126] S. Klamt, J. Gagneur, and A. von Kamp. Algorithmic approaches for computing elementary modes in large biochemical reaction networks. *IEE Proceedings - Systems Biology*, 152(4):249 – 255, 2005.
- [127] S. Klamt and E. D. Gilles. Minimal cut sets in biochemical reaction networks. *Bioinformatics*, 20(2):226 – 234, 2004.
- [128] S. Klamt, O. Hädicke, and A. von Kamp. *Large-Scale Networks in Engineering and Life Sciences*, chapter Stoichiometric and Constraint-Based Analysis of Biochemical Reaction Networks., pages 263 – 316. Springer International Publishing, 2014.
- [129] S. Klamt and R. Mahadevan. On the feasibility of growth-coupled product synthesis in microbial strains. *Metabolic Engineering*, 30:166 – 178, 2015.
- [130] S. Klamt, J. Saez-Rodriguez, and E. Gilles. Structural and functional analysis of cellular networks with *CellNetAnalyzer*. *BMC Systems Biology*, 1:2, 2007.
- [131] S. Klamt and J. Stelling. Combinatorial complexity of pathway analysis in metabolic networks. *Molecular Biology Reports*, 29(1-2):233 – 236, 2002.
- [132] S. Klamt and J. Stelling. Two approaches for metabolic pathway analysis? *Trends in Biotechnology*, 21(2):64 – 69, 2003.
- [133] H. Knoop. *Integrative Analyse des cyanobakteriellen Stoffwechsels*. PhD thesis, Humboldt-Universität zu Berlin, 2014.
- [134] H. Knoop, M. Gründel, Y. Zilliges, R. Lehmann, S. Hoffmann, W. Lockau, and R. Steuer. Flux balance analysis of cyanobacterial metabolism: The metabolic network of *Synechocystis* sp. PCC 6803. *PLoS Computational Biology*, 9(6):e1003081, 2013.
- [135] H. Knoop and R. Steuer. A computational analysis of stoichiometric constraints and trade-offs in cyanobacterial biofuel production. *Frontiers in Bioengineering and Biotechnology*, 3:47, 2015.

- [136] H. Knoop, Y. Zilliges, W. Lockau, and R. Steuer. The metabolic network of *Synechocystis* sp. PCC 6803: Systemic properties of autotrophic growth. *Plant Physiology*, 154(1):410 – 422, 2010.
- [137] D. Kramer and J. Evans. The importance of energy balance in improving photosynthetic productivity. *Plant Physiology*, 155(1):70 – 78, 2011.
- [138] D. M. Kramer, T. J. Avenson, and G. E. Edwards. Dynamic flexibility in the light reactions of photosynthesis governed by both electron and proton transfer reactions. *Trends in Plant Science*, 9(7):349 – 357, 2004.
- [139] T. Kusakabe, T. Tatsuke, K. Tsuruno, Y. Hirokawa, S. Atsumi, J. C. Liao, and T. Hanai. Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. *Metabolic Engineering*, 20: 101 – 108, 2013.
- [140] M. C. Lai and E. I. Lan. Advances in metabolic engineering of cyanobacteria for photosynthetic biochemical production. *Metabolites*, 5(4):636 – 658, 2015.
- [141] E. I. Lan and J. C. Liao. Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metabolic Engineering*, 13(4):353 – 363, 2011.
- [142] E. I. Lan and J. C. Liao. Atp drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 109(16):6018 – 6023, 2012.
- [143] D. E. Laudenbach, M. E. Reith, and N. A. Straus. Isolation, sequence analysis, and transcriptional studies of the flavodoxin gene from *Anacystis nidulans* R2. *Journal of Bacteriology*, 170(1):258 – 265, 1988.
- [144] D. J. Lea-Smith, P. Bombelli, J. S. Dennis, S. A. Scott, A. G. Smith, and C. J. Howe. Phycobilisome-deficient strains of *Synechocystis* sp. PCC 6803 have reduced size and require carbon-limiting conditions to exhibit enhanced productivity. *Plant Physiology*, 165(2):705 – 714, 2014.
- [145] J. Lederberg. *Encyclopedia of Microbiology*. Academic Press, 2000.
- [146] H. M. Lee, M. F. Vazquez-Bermudez, and N. de Marsac. The global nitrogen regulator NtcA regulates transcription of the signal transducer PII (GlnB) and influences its phosphorylation level in response to nitrogen and carbon supplies in

- the cyanobacterium *Synechococcus* sp. strain PCC 7942. *Journal of Bacteriology*, 181(9):2697 – 2702, 1999.
- [147] J. M. Lee, E. P. Gianchandani, and J. A. Papin. Flux balance analysis in the era of metabolomics. *Briefings in Bioinformatics*, 7(2):140 – 150, 2006.
- [148] S. Y. Lee, H. U. Kim, J. H. Park, J. M. Park, and T. Y. Kim. Metabolic engineering of microorganisms: General strategies and drug production. *Drug Discovery Today*, 14(1-2):78 – 88, 2009.
- [149] J. W. Lengeler, G. Drews, and H. G. Schlegel. *Biology of the Prokaryotes*. Thieme Stuttgart, 1999.
- [150] N. E. Lewis, H. Nagarajan, and B. O. Palsson. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology*, 10(4):291 – 305, 2012.
- [151] X. Li, C. R. Shen, and J. C. Liao. Isobutanol production as an alternative metabolic sink to rescue the growth deficiency of the glycogen mutant of *Synechococcus elongatus* pcc 7942. *Photosynthesis Research*, 120(3):301 – 310, 2014.
- [152] H. Lin, G. N. Bennett, and K.-Y. San. Metabolic engineering of aerobic succinate production systems in *Escherichia coli* to improve process productivity and achieve the maximum theoretical succinate yield. *Metabolic Engineering*, 7(2):116 – 127, 2005.
- [153] P. Lindberg, S. Park, and A. Melis. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metabolic Engineering*, 12(1):70 – 79, 2010.
- [154] R. Linington, J. Gonzalez, L. Urena, L. Romero, E. Ortega-Barria, and W. Gerwick. Venturamides A and B: Antimalarial constituents of the panamanian marine cyanobacterium *Oscillatoria* sp. *Journal of Natural Products*, 70(3):397 – 401, 2007.
- [155] X. Liu, J. Sheng, and R. Curtiss. Fatty acid production in genetically modified cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 108(17):6899 – 6904, 2011.

- [156] B. Loll, J. Kern, W. Saenger, A. Zouni, and J. Biesiadka. Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem ii. *Nature*, 438(7070):1040 – 1044, 2005.
- [157] M. R. Long, W. K. Ong, and J. L. Reed. Computational methods in metabolic engineering for strain design. *Current Opinion in Biotechnology*, 34:135 – 141, 2015.
- [158] D. S. Lun, G. Rockwell, N. J. Guido, M. Baym, J. A. Kelner, B. Berger, J. E. Galagan, and G. M. Church. Large-scale identification of genetic design strategies using local search. *Molecular Systems Biology*, 5:296, 2009.
- [159] D. Machado and M. Herrgard. Co-evolution of strain design methods based on flux balance and elementary mode analysis. *Metabolic Engineering Communications*, 2:85 – 92, 2015.
- [160] R. Mahadevan and C. Schilling. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic Engineering*, 5(4): 264 – 276, 2003.
- [161] N. Mahmood, W. Carmichael, and D. Pfahler. Anticholinesterase poisonings in dogs from a cyanobacterial (blue-green algae) bloom dominated by *Anabaena flos-aquae*. *American Journal of Veterinary Research*, 49(4):500 – 503, 1988.
- [162] P. Maia, M. Rocha, and I. Rocha. In silico constraint-based strain optimization methods: the quest for optimal cell factories. *Microbiology and Molecular Biology Reviews*, 80(1):45 – 67, 2016.
- [163] I. Maldener, M. L. Summers, and A. Sukenik. *The Cell Biology of Cyanobacteria*, chapter Cellular Differentiation in Filamentous Cyanobacteria, pages 263 – 291. Caister Academic Press, 2014.
- [164] D. McCloskey, B. O. Palsson, and A. M. Feist. Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Molecular Systems Biology*, 9:661, 2013.
- [165] J. T. McEwen, I. M. P. Machado, M. R. Connor, and S. Atsumi. Engineering *Synechococcus elongatus* PCC 7942 for continuous growth under diurnal conditions. *Applied and Environmental Microbiology*, 79(5):1668 – 1675, 2013.

- [166] L. Meher, D. V. Sagar, and S. Naik. Technical aspects of biodiesel production by transesterification - a review. *Renewable and Sustainable Energy Reviews*, 10(3):248 – 268, 2006.
- [167] G. Melzer, M. E. Esfandabadi, E. Franco-Lara, and C. Wittmann. Flux design: In silico design of cell factories based on correlation of pathway fluxes to desired properties. *BMC Systems Biology*, 3:120, 2009.
- [168] H. Min, Y. Liu, C. H. Johnson, and S. S. Golden. Phase determination of circadian gene expression in *Synechococcus elongatus* PCC 7942. *Journal of Biological Rhythms*, 19(2):103 – 112, 2004.
- [169] H. Min and L. A. Sherman. Hydrogen production by the unicellular, diazotrophic cyanobacterium *Cyanothece* sp. strain ATCC 51142 under conditions of continuous light. *Applied and Environmental Microbiology*, 76(13):4293 – 4301, 2010.
- [170] A. Montagud, E. Navarro, P. Fernández de Córdoba, J. Urchueguía, and K. Patil. Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. *BMC Systems Biology*, 4:156, 2010.
- [171] A. Montagud, A. Zelezniak, E. Navarro, P. F. de Cordoba, J. F. Urchueguia, and D. K. R. Patil. Flux coupling and transcriptional regulation within the metabolic network of the photosynthetic bacterium *Synechocystis* sp. PCC 6803. *Biotechnology Journal*, 6(3):330 – 342, 2011.
- [172] S. Müller, H. Harms, and T. Bley. Origin and analysis of microbial population heterogeneity in bioprocesses. *Current Opinion in Biotechnology*, 21(1):100 – 113, 2010.
- [173] S. Naik, V. V. Goud, P. K. Rout, and A. K. Dalai. Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 14(2):578 – 597, 2010.
- [174] Y. Nakamura, T. Kaneko, M. Hirosawa, N. Miyajima, and S. Tabata. Cyanobase, a www database containing the complete nucleotide sequence of the genome of *Synechocystis* sp. strain PCC 6803. *Nucleic Acids Research*, 26(1):63 – 67, 1998.
- [175] H. Niederholtmeyer, B. T. Wolfstädter, D. F. Savage, P. A. Silver, and J. C. Way. Engineering cyanobacteria to synthesize and export hydrophilic products. *Applied and Environmental Microbiology*, 76(11):3462 – 3466, 2010.

- [176] T. L. Nissen, M. C. Kielland-Brandt, J. Nielsen, and J. Villadsen. Optimization of ethanol production in *Saccharomyces cerevisiae* by metabolic engineering of the ammonium assimilation. *Metabolic Engineering*, 2(1):69 – 77, 2000.
- [177] D. R. Nobles and R. M. Brown. Transgenic expression of *Gluconacetobacter Xylinus* strain ATCC 53582 cellulose synthase genes in the cyanobacterium *Synechococcus Leopoliensis* strain UTCC 100. *Cellulose*, 15(5):691 – 701, 2008.
- [178] J. Nogales, S. Gudmundsson, E. Knight, B. Palsson, and I. Thiele. Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 109(7):2678 – 2683, 2012.
- [179] J. Nogales, S. Gudmundsson, and I. Thiele. Toward systems metabolic engineering in cyanobacteria: Opportunities and bottlenecks. *Bioengineered*, 4(3):158 – 163, 2013.
- [180] M. A. Oberhardt, J. Puchalka, K. E. Fryer, V. A. P., M. dos Santos, and J. A. Papin. Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *Journal of Bacteriology*, 190(8):2790 – 2803, 2008.
- [181] OECD. *Future Prospects for Industrial Biotechnology*. OECD Publishing, 2011.
- [182] A. P. Oliveira, J. Nielsen, and J. Förster. Modeling *Lactococcus Lactis* using a genome-scale flux model. *BMC Microbiology*, 5:39, 2005.
- [183] A. Oren and M. Shilo. Anaerobic heterotrophic dark metabolism in the cyanobacterium *Oscillatoria limnetica*: Sulfur respiration and lactate fermentation. *Archives of Microbiology*, 122(1):77 – 84, 1979.
- [184] J. D. Orth, T. M. Conrad, J. Na, J. A. Lerman, H. Nam, A. M. Feist, and B. O. Palsson. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism - 2011. *Molecular Systems Biology*, 7:535, 2011.
- [185] J. D. Orth, R. M. T. Fleming, and B. O. Palsson. Reconstruction and use of microbial metabolic networks: the core *Escherichia coli* metabolic model as an educational guide. *EcoSal Plus*, page doi: 10.1128/ecosalplus.10.2.1., 2010.
- [186] J. D. Orth, I. Thiele, and B. O. Palsson. What is flux balance analysis? *Nature Biotechnology*, 28(3):245 – 248, 2010.

- [187] E. J. O'Brien and B. O. Palsson. Computing the functional proteome: Recent progress and future prospects for genome-scale models. *Current Opinion in Biotechnology*, 34:125 – 134, 2015.
- [188] H. Paerl, R. Fulton 3rd., P. Moisander, and J. Dyble. Harmful freshwater algal blooms, with an emphasis on cyanobacteria. *The Scientific World Journal*, 1: 76 – 113, 2001.
- [189] B. Panda and N. Mallick. Enhanced poly- β -hydroxybutyrate accumulation in a unicellular cyanobacterium, *Synechocystis* sp. PCC 6803. *Letters in Applied Microbiology*, 44(2):194 – 198, 2007.
- [190] J. A. Papin, J. Stelling, N. D. Price, S. Klamt, S. Schuster, and B. O. Palsson. Comparison of network-based pathway analysis methods. *Trends in Biotechnology*, 22(8):400 – 405, 2004.
- [191] A. Parmar, N. K. Singh, A. Pandey, E. Gnansounou, and D. Madamwar. Cyanobacteria and microalgae: A positive prospect for biofuels. *Bioresource Technology*, 102(22):10163 – 10172, 2011.
- [192] K. R. Patil, I. Rocha, J. Förster, and J. Nielsen. Evolutionary programming as a platform for *in silico* metabolic engineering. *BMC Bioinformatics*, 6:308, 2005.
- [193] P. Pharkya, A. Burgard, and C. Maranas. Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. *Biotechnology and Bioengineering*, 84(7):887 – 899, 2003.
- [194] P. Pharkya, A. Burgard, and C. Maranas. OptStrain: A computational framework for redesign of microbial production systems. *Genome Research*, 14(11): 2367 – 2376, 2004.
- [195] P. Pharkya and C. Maranas. An optimization framework for identifying reaction activation/inhibition or elimination candidates for overproduction in microbial systems. *Metabolic Engineering*, 8(1):1 – 13, 2006.
- [196] M. Potts and B. A. Whitton. *The Ecology of Cyanobacteria. Their Diversity in Time and Space*. Kluwer Academic Publishers, 1999.
- [197] G. D. Price, M. R. Badger, F. J. Woodger, and B. M. Long. Advances in understanding the cyanobacterial CO₂-concentrating-mechanism (CCM): Functional

- components, ci transporters, diversity, genetic regulation and prospects for engineering into plants. *Journal of Experimental Botany*, 59(7):1441 – 1461, 2008.
- [198] N. D. Price, J. A. Papin, C. H. Schilling, and B. O. Palsson. Genome-scale microbial in silico models: The constraints-based approach. *Trends in Biotechnology*, 21(4):162 – 169, 2003.
- [199] N. D. Price, J. L. Reed, and B. O. Palsson. Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nature Reviews Microbiology*, 2(11):886 – 897, 2004.
- [200] J. Puchalka, M. A. Oberhardt, M. Godinho, A. Bielecka, D. Regenhardt, K. N. Timmis, J. A. Papin, and V. A. P. M. dos Santos. Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology. *PLoS Computational Biology*, 4(10):e1000210, 2008.
- [201] L. E. Quek, S. Dietmair, M. Hanscho, V. S. Martinez, N. Borth, and L. K. Nielsen. Reducing recon 2 for steady-state flux analysis of HEK cell culture. *Journal of Biotechnology*, 184:172 – 178, 2014.
- [202] N. Quintana, F. Van Der Kooy, M. Van De Rhee, G. Voshol, and R. Verpoorte. Renewable energy from cyanobacteria: Energy production optimization by metabolic pathway engineering. *Applied Microbiology and Biotechnology*, 91(3):471 – 490, 2011.
- [203] K. Raman and N. Chandra. Flux balance analysis of biological systems: Applications and challenges. *Briefings in Bioinformatics*, 10(4):435 – 449, 2009.
- [204] S. Ranganathan, P. F. Suthers, and C. D. Maranas. OptForce: An optimization procedure for identifying all genetic manipulations leading to targeted overproductions. *PLoS Computational Biology*, 6(4):e1000744, 2010.
- [205] R. P. Rastogi and R. P. Sinha. Biotechnological and industrial significance of cyanobacterial secondary metabolites. *Biotechnology Advances*, 27(4):521 – 539, 2009.
- [206] J. L. Reed, T. D. Vo, C. H. Schilling, and B. O. Palsson. An expanded genome-scale model of *Escherichia coli* K-12 (*iJR904 gsm/gpr*). *Genome Biology*, 4(9):R54, 2003.

- [207] R. Rippka, J. Deruelles, J. B. Waterbury, M. Herdman, and R. Y. Stanier. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Journal of General Microbiology*, 111(1):1 – 61, 1979.
- [208] R. L. Robson and J. R. Postgate. Oxygen and hydrogen in biological nitrogen fixation. *Annual Review of Microbiology*, 34:183 – 207, 1980.
- [209] R. Rockafellar. *Convex Analysis*. University Press, 1970.
- [210] L. Rosgaard, A. J. de Porcellinisa, J. H. Jacobsenc, N.-U. Frigaardc, and Y. Sakuragi. Bioengineering of carbon fixation, biofuels, and biochemicals in cyanobacteria and plants. *Journal of Biotechnology*, 162(1):134 – 147, 2012.
- [211] R. Saha, A. Verseput, B. Berla, T. Mueller, H. Pakrasi, and C. Maranas. Reconstruction and comparison of the metabolic potential of cyanobacteria *Cyanothece* sp. ATCC 51142 and *Synechocystis* sp. PCC 6803. *PLoS ONE*, 7(10):e48285, 2012.
- [212] M. Sakai, T. Ogawa, M. Matsuoka, and H. Fukuda. Photosynthetic conversion of carbon dioxide to ethylene by the recombinant cyanobacterium, *Synechococcus* sp. PCC 7942, which harbors a gene for the ethylene-forming enzyme of *Pseudomonas syringae*. *Journal of Fermentation and Bioengineering*, 84(5):434 – 443, 1997.
- [213] G. Sandmann, M. L. Peleato, M. F. Fillat, M. C. Lazaro, and C. Gomez-Moreno. Consequences of the iron-dependent formation of ferredoxin and flavodoxin on photosynthesis and nitrogen fixation on *Anabaena* strains. *Photosynthesis Research*, 26(2):119 – 125, 1990.
- [214] F. Sarsekeyeva, B. Zayadan, A. Usserbaeva, V. Bedbenov, M. Sinetova, and D. Los. Cyanofuels: Biofuels from cyanobacteria. reality and perspectives. *Photosynthesis Research*, 125(1-2):329 – 340, 2015.
- [215] C. H. Schilling, D. Letscher, and B. O. Palsson. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *Journal of Theoretical Biology*, 203(3):229 – 248, 2000.
- [216] A. Schirmer, M. A. Rude, X. Li, E. Popova, and S. B. del Cardayre. Microbial biosynthesis of alkanes. *Science*, 329(5991):559 – 562, 2010.
- [217] H. G. Schlegel. *Allgemeine Mikrobiologie*. Georg Thieme Verlag, 2007.

- [218] R. Schuetz, L. Kuepfer, and U. U. Sauer. Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Molecular Systems Biology*, 3:119, 2007.
- [219] S. Schuster, T. Dandekar, and D. A. Fell. Detection of elementary flux modes in biochemical networks: A promising tool for pathway analysis and metabolic engineering. *Trends in Biotechnology*, 17(2):53 – 60, 1999.
- [220] S. Schuster, D. Fell, and T. Dandekar. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnology*, 18(3):326 – 332, 2000.
- [221] S. Schuster and C. Hilgetag. On elementary flux modes in biochemical reaction systems at steady state. *Journal of Biological Systems*, 2(2):165 – 182, 1994.
- [222] S. Schuster, T. Pfeiffer, F. Moldenhauer, I. Koch, and T. Dandekar. Exploring the pathway structure of metabolism: Decomposition into subnetworks and application to *Mycoplasma Pneumoniae*. *Bioinformatics*, 18(2):351 – 361, 2002.
- [223] H. Seelert, A. Poetsch, N. A. Dencher, A. Engel, H. Stahlberg, and D. J. Müller. Structural biology: Proton-powered turbine of a plant motor. *Nature*, 405(6785):418 – 419, 2000.
- [224] E. Sergeeva, A. Liaimer, and B. Bergman. Evidence for production of the phytohormone indole-3-acetic acid by cyanobacteria. *Planta*, 215(2):229 – 238, 2002.
- [225] A. A. Shastri and J. A. Morgan. Flux balance analysis of photoautotrophic metabolism. *Biotechnology Progress*, 21(6):1617 – 1626, 2005.
- [226] S. L. Shaw, S. W. Chisholmb, and R. G. Prinn. Isoprene production by *Prochlorococcus*, a marine cyanobacterium, and other phytoplankton. *Marine Chemistry*, 80(4):227 – 245, 2003.
- [227] S. V. Shestakov and N. T. Khyen. Evidence for genetic transformation in blue-green alga *Anacystis nidulans*. *Molecular and General Genetics*, 107(4):372 – 375, 1970.
- [228] P. M. Shih, D. Wu, A. Latifi, S. D. Axen, D. P. Fewer, E. Talla, A. Calteau, F. Cai, N. T. de Marsac, R. Rippka, M. Herdman, K. Sivonen, T. Coursin, T. Laurent, L. Goodwin, M. Nolan, K. W. Davenport, C. S. Han, E. M. Rubin, J. A. Eisen,

- T. Woyke, M. Gugger, and C. A. Kerfeld. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, 110(3):1053 – 1058, 2013.
- [229] Y. Shimizu. Microalgal metabolites. *Current Opinion in Microbiology*, 6(3): 236 – 243, 2003.
- [230] T. Shlomi, Y. Eisenberg, R. Sharan, and E. Ruppin. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Molecular Systems Biology*, 3:101, 2007.
- [231] T. Simmons, E. Andrianasolo, K. McPhail, P. Flatt, and W. Gerwick. Marine natural products as anticancer drugs. *Molecular Cancer Therapeutics*, 4(2): 333 – 342, 2005.
- [232] S. Singh, B. N. Kate, and U. C. Banerjee. Bioactive compounds from cyanobacteria and microalgae: An overview. *Critical Reviews in Biotechnology*, 25(3):73 – 95, 2005.
- [233] H. S. Song, R. S. McClure, H. C. Bernstein, C. C. Overall, E. A. Hill, and A. S. Beliaev. Integrated in silico analyses of regulatory and metabolic networks of *Synechococcus* sp. PCC 7002 reveal relationships between gene centrality and essentiality. *Life*, 5(2):1127 – 1140, 2015.
- [234] L. Stal. Poly(hydroxyalkanoate) in cyanobacteria: An overview. *FEMS Microbiology Reviews*, 103(2-4):169 – 180, 1992.
- [235] L. J. Stal and R. Moezelaar. Fermentation in cyanobacteria. *FEMS Microbiology Reviews*, 21(2):179 – 211, 1997.
- [236] R. Y. Stanier, R. Kunisawa, M. Mandel, and G. Cohen-Bazire. Purification and properties of unicellular blue-green algae (order *Chroococcales*). *Bacteriological Reviews*, 35(2):171 – 205, 1971.
- [237] E. J. Steen, R. Chan, N. Prasad, S. Myers, C. J. Petzold, A. Redding, M. Ouellet, and J. D. Keasling. Metabolic engineering of *Saccharomyces Cerevisiae* for the production of n-butanol. *Microbial Cell Factories*, 7:36, 2008.
- [238] G. N. Stephanopoulos, A. A. Aristidou, and J. Nielsen. *Metabolic Engineering; Principles and Methodologies*. Academic Press, 1998.

- [239] E. Stephens, I. L. Ross, and B. Hankamer. Expanding the microalgal industry - continuing controversy or compelling case. *Current Opinion in Chemical Biology*, 17(3):444 – 452, 2013.
- [240] K. Takahama, M. Matsuoka, K. Nagahama, and T. Ogawa. Construction and analysis of a recombinant cyanobacterium expressing a chromosomally inserted gene for an ethylene-forming enzyme at the *psbAI* locus. *Journal of Bioscience and Bioengineering*, 95(3):302 – 305, 2003.
- [241] N. Tepper and T. Shlomi. Predicting metabolic engineering knockout strategies for chemical production: Accounting for competing pathways. *Bioinformatics*, 26(4):536 – 543, 2010.
- [242] M. Terzer and J. Stelling. Accelerating the computation of elementary modes using pattern trees. In P. Bächer and B. M. Moret, editors, *Algorithms in Bioinformatics*, volume 4175 of *Lecture Notes in Computer Science*, pages 333 – 343. Springer Berlin Heidelberg, 2006.
- [243] M. Terzer and J. Stelling. Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics*, 24(19):2229 – 2235, 2008.
- [244] N. Thajuddin and G. Subramanian. Cyanobacterial biodiversity and potential applications in biotechnology. *Current Science*, 89(1):47 – 57, 2005.
- [245] I. Thiele and B. Palsson. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, 5(1):93 – 121, 2010.
- [246] C. S. Ting, G. Rocap, J. King, and S. W. Chisholm. Cyanobacterial photosynthesis in the oceans: The origins and significance of divergent light-harvesting strategies. *Trends in Microbiology*, 10(3):134 – 142, 2002.
- [247] N. Tomar and R. K. De. Comparing methods for metabolic network analysis and an application to metabolic engineering. *Gene*, 521(1):1 – 14, 2013.
- [248] J. Triana, A. Montagud, M. Siurana, D. Fuente, A. Urchueguia, D. Gamermann, J. Torres, J. Tena, P. F. de Cordoba, and J. F. Urchueguia. Generation and evaluation of a genome-scale metabolic network model of *Synechococcus elongatus* PCC 7942. *Metabolites*, 4(3):680 – 698, 2014.

- [249] C. Trinh, R. Carlson, A. Wlaschin, and F. Sreenc. Design, construction and performance of the most efficient biomass producing *E. coli* bacterium. *Metabolic Engineering*, 8(6):628 – 638, 2006.
- [250] C. Trinh, P. Unrean, and F. Sreenc. Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. *Applied and Environmental Microbiology*, 74(12):3634 – 3643, 2008.
- [251] C. Trinh, A. Wlaschin, and F. Sreenc. Elementary mode analysis: A useful metabolic pathway analysis tool for characterizing cellular metabolism. *Applied Microbiology and Biotechnology*, 81(5):813 – 826, 2009.
- [252] J. Ungerer, L. Tao, M. Davis, M. Ghirardi, P.-C. Manessa, and J. Yu. Sustained photosynthetic conversion of CO₂ to ethylene in recombinant cyanobacterium *Synechocystis* 6803. *Energy & Environmental Science*, 5(10):8998 – 9006, 2012.
- [253] J. van der Oost, B. A. Bulthuis, S. Feitz, K. Krab, and R. Kraayenhof. Fermentation metabolism of the unicellular cyanobacterium *Cyanothece* PCC 7822. *Archives of Microbiology*, 152(5):415 – 419, 1989.
- [254] A. Varma and B. O. Palsson. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Applied and Environmental Microbiology*, 60(10):3724 – 3731, 1994.
- [255] A. M. Varman, Y. Xiao, H. B. Pakrasi, and T. Y. J. Metabolic engineering of *Synechocystis* sp. strain PCC 6803 for isobutanol production. *Applied and Environmental Microbiology*, 79(3):908 – 914, 2013.
- [256] J. Varner and D. Ramkrishna. Metabolic engineering from a cybernetic perspective. 1. theoretical preliminaries. *Biotechnology Progress*, 15(3):407 – 425, 1999.
- [257] R. Vidal, L. López-Maury, M. Guerrero, and F. Florencio. Characterization of an alcohol dehydrogenase from the cyanobacterium *Synechocystis* sp. strain PCC 6803 that responds to environmental stress conditions via the Hik34-Rre1 two-component system. *Journal of Bacteriology*, 191(13):4383 – 4391, 2009.
- [258] M. Vollmar, D. Schlieper, M. Winn, C. Büchner, and G. Groth. Structure of the c14 rotor ring of the proton translocating chloroplast atp synthase. *Journal of Biological Chemistry*, 284(2727):18228 – 18235, 2009.

- [259] A. von Kamp and S. Klamt. Enumeration of smallest intervention strategies in genome-scale metabolic networks. *PLoS Computational Biology*, 10(1):e1003378, 2014.
- [260] A. von Kamp and S. Schuster. Metatool 5.0: Fast and flexible elementary modes analysis. *Bioinformatics*, 22(15):1930 – 1931, 2006.
- [261] T. T. Vu, E. A. Hill, L. A. Kucek, A. E. Konopka, A. S. Beliaev, and J. L. Reed. Computational evaluation of *Synechococcus* sp. PCC 7002 metabolism for chemical production. *Biotechnology Journal*, 8(5):619 – 630, 2013.
- [262] T. T. Vu, S. M. Stolyar, G. E. Pinchuk, E. A. Hill, L. A. Kucek, R. N. Brown, M. S. Lipton, A. Osterman, J. K. Fredrickson, A. E. Konopka, A. S. Beliaev, and J. L. Reed. Genome-scale modeling of light-driven reductant partitioning and carbon fluxes in diazotrophic unicellular cyanobacterium *Cyanothece* sp. ATCC 51142. *PLoS Computational Biology*, 8(4):e1002460, 2012.
- [263] A. E. Walsby. Gas vesicles. *Microbiological Review*, 58(1):94 – 144, 1994.
- [264] H. H. Wang, F. J. Isaacs, P. A. Carr, Z. Z. S. G. Xu, C. R. Forest, and G. M. Church. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, 460(7257):894 – 989, 2009.
- [265] E. A. Welsh, M. Liberton, J. Stöckel, T. Loh, T. Elvitigala, C. Wang, A. Wollam, R. S. Fulton, S. W. Clifton, J. M. Jacobs, R. Aurora, B. K. Ghosh, L. A. Sherman, R. D. Smith, R. K. Wilson, and H. B. Pakrasi. The genome of *Cyanothece* 51142, a unicellular diazotrophic cyanobacterium important in the marine nitrogen cycle. *Proceedings of the National Academy of Sciences of the United States of America*, 105(39):15094 – 15099, 2008.
- [266] C. Wiegand and S. Pflugmacher. Ecotoxicological effects of selected cyanobacterial secondary metabolites a short review. *Toxicology and Applied Pharmacology*, 203(3):201 – 218, 2005.
- [267] L. Yang, W. Cluett, and R. Mahadevan. EMILiO: A fast algorithm for genome-scale strain design. *Metabolic Engineering*, 13(3):272 – 281, 2011.
- [268] H. Yim, R. Haselbeck, W. Niu, C. Pujol-Baxley, A. Burgard, J. Boldt, J. Khandurina, J. Trawick, R. Osterhout, R. Stephen, J. Estadilla, S. Teisan, H. Schreyer, S. Andrae, T. Yang, S. Lee, M. Burk, and S. Van Dien. Metabolic engineering of

- Escherichia coli* for direct production of 1,4-butanediol. *Nature Chemical Biology*, 7(7):445 – 452, 2011.
- [269] K. Yoshikawa, Y. Kojima, T. Nakajima, C. Furusawa, T. Hirasawa, and H. Shimizu. Reconstruction and verification of a genome-scale metabolic model for *Synechocystis* sp. PCC 6803. *Applied Microbiology and Biotechnology*, 92(2):347 – 358, 2011.
- [270] A. Zabaniotou, O. Ioannidou, and V. Skoulou. Rapeseed residues utilization for energy and 2nd generation biofuels. *Fuel*, 87(8-9):1492 – 1502, 2008.
- [271] N. Zamboni and U. Sauer. Novel biological insights through metabolomics and 13C-flux analysis. *Current Opinion in Microbiology*, 12(5):553 – 558, 2009.
- [272] J. Zanghellini, D. Ruckerbauer, M. Hanscho, and C. Jungreuthmayer. Elementary flux modes in a nutshell: Properties, calculation and applications. *Biotechnology Journal*, 8(9):1009 – 1016, 2013.
- [273] S. Zhang and D. A. Bryant. The tricarboxylic acid cycle in cyanobacteria. *Science*, 334(6062):1551 – 1553, 2011.
- [274] J. Zhou, H. Zhang, H. Meng, Y. Zhang, and Y. Li. Production of optically pure d-lactate from CO₂ by blocking the phb and acetate pathways and expressing d-lactate dehydrogenase in cyanobacterium *Synechocystis* sp. pcc 6803. *Process Biochemistry*, 49(12):2071 – 2077, 2014.
- [275] J. Zhou, H. Zhang, Y. Zhang, Y. Li, and Y. Ma. Designing and creating a modularized synthetic pathway in cyanobacterium *Synechocystis* enables production of acetone from carbon dioxide. *Metabolic Engineering*, 14(4):394 – 400, 2012.
- [276] J. Zhou, T. Zhu, Z. Cai, and Y. Li. From cyanochemicals to cyanofactories: a review and perspective. *Microbial Cell Factories*, 15:2, 2016.
- [277] X.-G. Zhu, S. P. Long, and D. R. Ort. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Current Opinion in Biotechnology*, 19(2):153 – 159, 2008.
- [278] K. Ziegler, A. Diener, C. Herpin, R. Richter, R. Deutzmann, and W. Lockau. Molecular characterization of cyanophycin synthetase, the enzyme catalyzing the biosynthesis of the cyanobacterial reserve material multi-l-arginyl-poly-l-aspartate (cyanophycin). *European Journal of Biochemistry*, 254(1):154 – 159, 1998.

Literaturverzeichnis

- [279] Y. Zilliges. *The Cell Biology of Cyanobacteria*, chapter Glycogen, a Dynamic Cellular Sink and Reservoir for Carbon., pages 189 – 210. Caister Academic Press, 2014.