

**Integrative Auswertung von Multi-Omics-Daten
aus dem Zentralstoffwechsel von
*Corynebacterium glutamicum***

Dissertation
zur Erlangung des akademischen Grades
DOKTOR-INGENIEUR

vorgelegt von
Stephan Noack
aus Hoyerswerda

eingereicht dem
Fachbereich Maschinenbau
der Universität Siegen

Tag der Einreichung: 16.06.2009

Gutachter: Herr Prof. Dr. rer. nat. habil. W. Wiechert
Herr Prof. Dr.-Ing. habil. R. Takors

gedruckt auf alterungsbeständigem holz- und säurefreiem Papier

Zusammenfassung

Eine wesentliche Zielsetzung im Bereich der Systembiologie besteht in der Untersuchung der regulatorischen Prozesse einer Zelle unter Berücksichtigung der Hierarchie-Ebenen Genom, Transkriptom, Proteom, Metabolom und Fluxom. Aufgrund der rasch fortschreitenden Entwicklung experimenteller und analytischer Technologien ist es möglich, die hierfür benötigten quantitativen Multi-Omics-Daten bereitzustellen.

Den ersten Schwerpunkt der vorliegenden Arbeit bildet die Untersuchung zweier verschiedener Ansätze zur ^{13}C -Stoffflussanalyse. Hierbei wird der Organismus *C. glutamicum* erstmalig im Rahmen eines isotopisch instationären Markierungsexperiments im Fed-Batch-Ansatz untersucht. Auf diese Weise werden sowohl isotopisch transiente als auch stationäre Markierungsanreicherungen von Zentralstoffwechselmetaboliten erzeugt.

Die somit generierten Rohdaten werden auf der Basis von isotopisch stationären und instationären Modellierungsansätzen unter Verwendung verschiedener Netzwerk- und Messkonfigurationen ausgewertet. Eine anschließende Verifizierung und Interpretation der erhaltenen quantitativen Stoffflussdaten erfolgt auf Grundlage von Flussbilanzanalysen in einem weiterentwickelten genomweiten Stoffwechselmodell für *C. glutamicum*.

Die modellgestützte Auswertung von Multi-Omics-Daten bildet den zweiten Schwerpunkt innerhalb dieser Arbeit. Dabei wird ein neuartiges Modellierungskonzept entwickelt, welches auf einem vertikalen mechanistischen Netzwerkansatz basiert. Dieser ist entlang der Regulationsebenen einer Zelle (Genom, Proteom, Metabolom) formuliert.

Mit der Fokussierung auf den Zitratzyklus und der Anaplerose von *C. glutamicum* kann unter Einbezug des aktuellen Wissensstandes dieses Stoffwechselweges insbesondere auf metabolischer Ebene ein detailliertes Regulationsnetzwerk aufgebaut werden. Die Implementierung des vertikalen Netzwerkmodells erfolgt mit Hilfe der Modellierungssprache Modelica. Zu diesem Zweck wird eine geeignete Komponentenbibliothek entwickelt, mit deren Hilfe biochemische Netzwerke unter Nutzung der Objektorientiertheit dieser Sprache modelliert werden können. Die im Zuge kontinuierlicher Kultivierungen bereitgestellten umfassenden quantitativen Multi-Omics-Datensätze bilden schließlich die Grundlage zur Validierung des vertikalen Netzwerkmodells.

Vorwort

Die vorliegende Arbeit entstand am Institut für Biotechnologie 2 der Forschungszentrum Jülich GmbH im Rahmen des Projektes „Systemorientierte Analyse des Zentralstoffwechsels mikrobieller Aminosäureproduzenten“ (SysMAP). Während meiner Promotionszeit standen mir viele Menschen mit Rat und Tat zu Seite, wofür ich mich an dieser Stelle herzlich bedanken möchte.

- Mein besonderer Dank gilt meinem Doktorvater, Prof. Wolfgang Wiechert, für die Hilfestellung bei der Konzeption dieser Arbeit sowie die zahlreichen fachlichen Diskussionen, in denen er mir durch entscheidende Fragen immer wieder neuen Anstoss zum Nachdenken gab.
- Prof. Ralf Takors danke ich für seine wertvollen Hinweise und sein kritisches Hinterfragen aus der industriellen Perspektive heraus sowie für die freundliche Übernahme des Zweitgutachtens.
- Für die fachliche Betreuung vor Ort bedanke ich mich vor allem bei Dr. Marco Oldiges und Prof. Christian Wandrey für seine oft ansteckende Begeisterungsfähigkeit, die er unter anderem den Fragestellungen und Herausforderungen der Modellierung entgegenbrachte. Aber auch die vielen kleinen und großen Diskussionen mit Dr. Katharina Nöh, in denen oftmals mathematisches gegen biologisches Wissen ausgetauscht wurde, haben zur Vervollständigung dieser Arbeit beigetragen. Für die arbeits- und zeitintensive Messung und Auswertung zahlreicher Daten danke ich darüber hinaus u.a. Bianca Klein und Dipl.-Ing. Matthias Moch.
- Außerdem möchte ich mich bei allen SysMAP-Kooperationspartnern bedanken, welche mit der Bereitstellung von weiteren Omics-Daten eine wesentliche Voraussetzung zur Erstellung dieser Arbeit geschaffen haben.

Besonderer Dank gilt nicht zuletzt meiner kleinen Familie sowie meinen Eltern ohne deren Verständnis und Unterstützung ein Gelingen dieser Arbeit nicht möglich gewesen wäre.

Inhaltsverzeichnis

I	Einleitung und Methoden	1
1	Systembiologie im Omics-Zeitalter	3
1.1	Multi-Omics-Daten	3
1.2	Rolle der Systembiologie	5
1.3	Modellorganismus <i>C. glutamicum</i>	7
1.4	Fokus Zitratzyklus	7
1.5	Motivation dieser Arbeit	8
2	Zielsetzung und Gliederung	9
2.1	Schwerpunkte der Arbeit	9
2.2	Gliederung der Arbeit	13
3	Modellierung biochemischer Netzwerke	15
3.1	Mechanismen der Stoffwechselregulation	15
3.2	Vereinfachende Annahmen	18
3.3	Metabolische Netzwerkmodellierung	20
3.3.1	Massenbilanzen von Stoffwechselintermediaten	20
3.3.2	Metabolisches Beispielnetzwerk	23
3.3.3	Massenbilanzen markierter Stoffwechselintermediate	23
3.3.4	Stöchiometrische Modellierung	24
3.3.5	Kinetische Modellierung	28
3.4	Genetische Netzwerkmodellierung	33
3.4.1	Genomweite qualitative Modellierung	33
3.4.2	Fokussierte quantitative Modellierung	35
4	Methodenentwicklung mit Modelica	37
4.1	Modellierungssprache Modelica	37
4.2	Modellierung biochemischer Netzwerke mit Modelica	38
4.2.1	Gleichungsbasiertes Modellierungskonzept	38
4.2.2	Aufbau von Komponenten und Bibliotheken	39
4.2.3	Automatische Codegenerierung	43
4.2.4	Nebenbedingungen und Konsistenzprüfung	45
4.2.5	Hierarchische Modellierung und Klassenabstraktion	47
4.3	Entwicklung einer Komponentenbibliothek	48
4.3.1	Entwicklungsumgebung Dymola	48

4.3.2	Struktur der Komponentenbibliothek	49
4.4	Methoden zur Simulation und Systemanalyse	51
II	Gewinnung quantitativer Fluxom-Daten	53
5	¹³C-Stoffflussanalyse eines Lysinproduzenten	55
5.1	Kriterien isotopisch stationärer und instationärer ¹³ C-MFA	55
5.2	Experimentelle Rahmenbedingungen	57
5.3	Bioprozess- und Analysemethoden	58
5.3.1	Stamm und Kultivierungsbedingungen	58
5.3.2	Prozesskontrolle	58
5.3.3	Prozessmodell	59
5.3.4	Instationäres Markierungsexperiment	60
5.3.5	LC-MS-Analytik intrazellulärer Intermediate	60
5.4	Prozessverlauf	61
5.4.1	Intrazelluläre Markierungsdynamiken	61
5.4.2	Schätzung extrazellulärer Raten und Biomasseabflüsse	64
5.5	Modellierung isotopisch stationärer Daten	65
5.5.1	Fokussiertes Modell des Zentralstoffwechsels	65
5.5.2	Erweitertes Modell mit Aminosäurebiosynthesewegen	69
5.6	Modellierung isotopisch instationärer Daten	74
5.7	Statistische Analyse	79
5.7.1	Linearisierungsansatz	79
5.7.2	Monte-Carlo-Simulationen	80
5.7.3	Methodenvergleich	83
5.8	Quantifizierbarkeit intrazellulärer Stoffflüsse	85
6	Interpretation geschätzter Stoffflüsse	87
6.1	Genomweites Netzwerkmodell von <i>C. glutamicum</i>	87
6.2	Flussbilanzanalyse	90
6.2.1	Vergleich optimaler Stoffflusslagen	90
6.2.2	Maximierung der Lysinproduktion	94
III	Integrative mechanistische Modellierung	97
7	Gewinnung quantitativer Multi-Omics-Daten	99
7.1	Bioprozess- und Analysemethoden	99
7.1.1	Stämme und Kultivierungsbedingungen	99
7.1.2	Genomweite Transkriptom-Daten	100
7.1.3	Proteom-Daten aus Enzymaktivitätsmessungen	102
7.1.4	Metabolom-Daten intra- und extrazellulärer Metabolite	103
7.1.5	Fluxom-Daten extrazellulärer Flüsse	103
7.2	Vergleichende Analysen	103

7.2.1	Stamm- und wachstumsspezifische Unterschiede	103
8	Vertikales Netzwerkmodell des Zitratzyklus	109
8.1	Vertikale Modellierung	109
8.1.1	Massenbilanzen von mRNA, Proteinen und Metaboliten . . .	110
8.2	Modellierung genetischer Regulation	111
8.2.1	Vereinfachende Annahmen	112
8.2.2	Phänomenologische Beziehungen	112
8.3	Modellierung metabolischer Regulation	114
8.3.1	Vereinfachende Annahmen	114
8.3.2	Reaktionskinetiken	114
8.3.3	Stöchiometrische Gleichungen	116
8.4	Modellimplementierung mit Modelica	118
9	Validierung des vertikalen Netzwerkmodells	121
9.1	Integration von Multi-Omics-Daten	121
9.2	Gekoppelte Parameterschätzung	122
9.3	Modellanpassungen	124
9.4	Simulationsergebnisse	125
9.5	Sensitivitätsanalyse	130
9.5.1	Fokus Lysinbiosynthese	132
10	Zusammenfassung, Diskussion und Ausblick	139
10.1	Gewinnung quantitativer Fluxom-Daten	139
10.2	Integrative mechanistische Modellierung	143
10.3	Ausblick	146
A	Methoden zur Systemanalyse	149
A.1	Parameterschätzung	149
A.2	Sensitivitätsanalyse	151
A.2.1	Berechnung lokaler Sensitivitäten	151
A.2.2	Kontrollkoeffizienten	152
A.3	Fehlerübertragung	153
A.3.1	Linearisierungsansatz	153
A.3.2	Monte-Carlo-Simulation	154
B	¹³C-MFA von <i>C. glutamicum</i> DM1730	157
B.1	Prozessmodell	157
B.2	Biomasseabflüsse	157
B.3	Modelle zur ¹³ C-MFA	157
B.4	Modellanpassungen der ¹³ C-MFA	157

C	Vertikales Netzwerkmodell von <i>C. glutamicum</i>	163
C.1	Modellkonstanten	163
C.2	Modellreaktionen	164
C.3	Globale Bilanzen	165
C.4	Parameterschätzungen	166

Abkürzungen und Formelzeichen

Abkürzung	Bezeichnung
13CFLUX	Werkzeug zur Auswertung von ^{13}C -Stoffflussanalyse-Experimenten
AD	Automatische Differentiation
ADModelica	Werkzeug zur AD von Modelica-Modellen
ANA	Anaplerose
BM	Biomassesyntesewege
BTM	Biotrockenmasse
<i>C. glutamicum</i>	<i>Corynebacterium glutamicum</i>
cDNA	Komplementäre DNA einer RNA
DAE	Differential-Algebraische-Gleichung
DNA	Desoxyribonucleinsäure
<i>E. coli</i>	<i>Escherichia coli</i>
GC	Gaschromatographie
FBA	Flussbilanzanalyse
FQS	Fehlerquadratsumme
HPLC	Hochdruck-Flüssigchromatographie
IBT	Institut für Biotechnologie der Forschungszentrum Jülich GmbH
<i>in vitro</i>	(lat. im Reagenzglas) Experimente unter künstlichen Bedingungen
<i>in vivo</i>	(lat. im Leben) Experimente im lebenden System
LC-MS	Kopplung von Flüssigchromatographie und Massenspektrometrie
IM	Metabolisches Intermediat
LYS	Lysinbiosynthese
MCA	Metabolische Kontrollanalyse
MFA	Metabolische Flussanalyse
^{13}C -MFA	isotopisch stationäre MFA
INST- ^{13}C -MFA	isotopisch instationäre ^{13}C -MFA
MM	Michaelis-Menten-Kinetik
mRNA	Boten-RNA

Fortsetzung nächste Seite

INHALTSVERZEICHNIS

Abkürzung	Bezeichnung
rRNA	Ribosomale-RNA
tRNA	Transport-RNA
NMR	Kernspinresonanzspektroskopie
NT	Nukleotid
OD	Optische Dichte
ODE	Gewöhnliche Differentialgleichung
Omix	Visualisierungswerkzeug für biochemische Netzwerke
PPP	Pentosephosphat-Weg
RNA	Ribonucleinsäure
RNAP	RNA-Polymerase
RP	Regulatorprotein
SBML	XML-Format zur Beschreibung metabolischer Netzwerke
SRE	Stimulus-Response-Experiment
TCA	Trikarbonsäure-Zyklus
TKF	Transkriptionsfaktor
TLF	Translationsfaktor
ZSW	Zentralstoffwechsel

Symbol	Bezeichnung	Einheit
c_{IM}	Konzentration eines Intermediats	[mM]
$c_{IM,0}$	Anfangskonzentration eines Intermediats	[mM]
\mathbf{c}_{IM}	Vektor aller Intermediatpools	[mM]
$\mathbf{c}_{IM}^{\#}$	Vektor markierter Intermediatpools	[mM]
\mathbf{c}_S	Vektor aller externen Poolgrößen	[mM]
$\mathbf{c}_S^{\#}$	Vektor markierter Substratpools	[mM]
$\text{Conf}_{\hat{\theta}}(90)$	90%-Konfidenzintervall der Schätzfunktion	
E	Enzymkonzentration	[mM]
ex_{CO_2}	CO ₂ -Bildungsrate	[mmol g _{BTM} ⁻¹ h ⁻¹]
ex_{LYS}	Lysinbildungsrate	[mmol g _{BTM} ⁻¹ h ⁻¹]
feed_{GLC0}	Zufluss an unmarkierter Glucose	[mmol g _{BTM} ⁻¹ h ⁻¹]
feed_{GLC1}	Zufluss an C1-markierter Glucose	[mmol g _{BTM} ⁻¹ h ⁻¹]
feed_{GLCU}	Zufluss an vollmarkierter Glucose	[mmol g _{BTM} ⁻¹ h ⁻¹]
GP	Konzentration an Gen-Polymerase-Komplex	[mM]
k^{\rightarrow}	Geschwindigkeitskonstante der Vorwärtsreaktion	[s ⁻¹]
k^{\leftarrow}	Geschwindigkeitskonstante der Rückwärtsreaktion	[s ⁻¹]
$k_{dg,mRNA}$	Degradationskonstante einer mRNA	[h ⁻¹]

Fortsetzung nächste Seite

Symbol	Bezeichnung	Einheit
$k_{\text{dg,P}}$	Degradationskonstante eines Protein	$[\text{h}^{-1}]$
K_{m}	Affinitätskonstante	$[\text{mM}]$
K_{eq}	Gleichgewichtskonstante	
m_{RNA}	Konzentration einer mRNA	$[\text{mM}]$
m_{X}	Biomasse	$[\text{g}]$
N	Stöchiometrische Matrix	
n_{IM}	Stoffmenge eines Intermediats	$[\text{mmol}]$
$n_{\text{IM,X}}$	Biomassespezifische Stoffmenge eines Intermediats	$[\text{mmol g}_{\text{BTM}}^{-1}]$
r_{P}	Produktbildungsrate	$[\text{g l}^{-1} \text{h}^{-1}]$
r_{S}	Substratverbrauchsrate	$[\text{g l}^{-1} \text{h}^{-1}]$
r_{X}	Biomassebildungsrate	$[\text{g l}^{-1} \text{h}^{-1}]$
t	Zeit	$[\text{s}]$
upt_{GLC}	Glucoseaufnahmerate	$[\text{mmol g}_{\text{BTM}}^{-1} \text{h}^{-1}]$
v	Vektor aller Flüsse bzw. Reaktionen	$[\text{mM s}^{-1}]$
V	Volumen	$[\text{l}]$
v^{\rightarrow}	Vorwärtsfluss in Bezug auf definierte Nettorichtung	$[\text{mM s}^{-1}]$
v^{\leftarrow}	Rückwärtsfluss in Bezug auf definierte Nettorichtung	$[\text{mM s}^{-1}]$
v^{→←}	Vektor der Vorwärts- und Rückwärtsflüsse	$[\text{mM s}^{-1}]$
v_{bm}	Fluss in Richtung Biomassebildung	$[\text{mM s}^{-1}]$
V_{cell}	Intrazelluläres Volumen einer Zelle	$[\text{ml}]$
$V_{\text{cell,X}}$	Biomassespezifisches intrazelluläres Volumen	$[\text{ml g}_{\text{BTM}}^{-1}]$
v_{ct}	Enzymatische Reaktionsrate	$[\text{mM s}^{-1}]$
$v_{\text{dg,mRNA}}$	Degradationsrate einer mRNA	$[\text{mM s}^{-1}]$
$v_{\text{dg,P}}$	Degradationsrate eines Proteins	$[\text{mM s}^{-1}]$
$v_{\text{el,mRNA}}$	Elongationsrate einer mRNA	$[\text{NT RNAP}^{-1} \text{min}^{-1}]$
$v_{\text{el,P}}$	Elongationsrate eines Proteins	$[\text{NT RNAP}^{-1} \text{min}^{-1}]$
v_{exc}	Fluss aus der Zelle (Produktexkretion)	$[\text{mM s}^{-1}]$
v_i	Fluss der i -ten Reaktion	$[\text{mM s}^{-1}]$
\dot{V}_{in}	Volumenstrom in den Bioreaktor	$[\text{l h}^{-1}]$
v_{met}	Metabolische Reaktion	$[\text{mM s}^{-1}]$
v_{net}	Nettofluss einer reversiblen Reaktion	$[\text{mM s}^{-1}]$
\dot{V}_{out}	Volumenstrom aus dem Bioreaktor	$[\text{l h}^{-1}]$
V_{R}	Arbeitsvolumen des Bioreaktors	$[\text{l}]$
v_{tc}	Transkriptionsrate	$[\text{mM s}^{-1}]$
v_{tl}	Translationsrate	$[\text{mM s}^{-1}]$
v_{upt}	Fluss in die Zelle (Substrataufnahme)	$[\text{mM s}^{-1}]$

Fortsetzung nächste Seite

INHALTSVERZEICHNIS

Symbol	Bezeichnung	Einheit
$x_{IM}^{\#}$	Vektor der Markierungsanteile eines Intermediatpools	
$x_S^{\#}$	Vektor der Markierungsanteile eines Substrats	
$Y_{CO_2/S}$	Substratspezifischer CO ₂ -Ertrag	[g g ⁻¹]
$Y_{P/S}$	Substratspezifischer Lysin'ertrag	[g g ⁻¹]
$Y_{X/S}$	Substratspezifischer Biomasse'ertrag	[g g ⁻¹]
$Y_{IM/X}$	Biomassespezifischer Precursorbedarf	[mmol g _{BTM} ⁻¹]
α	Parametervektor	
μ	Biomassespezifische Wachstumsrate	[h ⁻¹]
$\hat{\mu}$	Geschätzter Mittelwert	
μ_{max}	Maximale Wachstumsrate	[h ⁻¹]
$\hat{\sigma}$	Geschätzte Standardabweichung	
ω	Skalierungsfaktor	

Abkürzung	Enzym
aceA	Isocitrat-Lyase
aceB	Malat-Synthase
acn	Aconitase
aspB	Aspartat-Aminotransferase
bsLYS	Lysin-Biosynthese
c3	Fluss im EMP nach Spaltung in C3-Körper
c5	Fluss im PPP
dapA	Dihydrodipicolinate-Synthase
dapC	Succinyl-Amino-Ketopimelate-Transaminase
dapD	Tetrahydrodipicolinate-Succinylase
dapE	Succinyl-Diaminopimelate-Desuccinylase
ddh	Diaminopimelate-Dehydrogenase
eno	Enolase
fda	Aldolase
fum	Fumarase
gapA	Glyceraldehyd-3-Phosphat-Dehydrogenase
gdh	Glutamat-Dehydrogenase
gltA	Citrat-Synthase
gnd	6-Phospho-Gluconat-Dehydrogenase
icd	Isocitrat-Dehydrogenase
lysA	Diaminopimelate-Decarboxylase
lysC	Aspartate-Kinase

Fortsetzung nächste Seite

Abkürzung	Enzym
mdh	Malat-Dehydrogenase
mez	Malat-Enzym
mqq	Malat-Quinone-Oxireductase
pck	PEP-Carboxykinase
pfk	Phosphofruktokinase
pgi	Glucose-6-Phosphat-Isomerase
pgk	Phosphoglycerat-Kinase
pdh	Pyruvat-Dehydrogenase
ppc	PEP-Carboxylase
pqq	Pyruvate-Quinone-Oxidoreductase
pta	Phosphotransacetylase
pyc	Pyruvat-Carboxylase
pyk	Pyruvat-Kinase
pts	Phosphotransferasesystem
odh	α -Ketoglutarat-Dehydrogenase
odx	Oxalacetat-Decarboxylase
sdh	Succinat-Menaquinone-Oxireductase
suc	Succinyl-CoA-Synthetase
tal	Transaldolase
tpi	Triosephosphat-Isomerase
tkt	Transketolase

Abkürzung	Metabolit
13BPG	1,3-Biphosphoglycerat
2PG	2-Phosphoglycerat
3PG	3-Phosphoglycerat
ACCOA	Acetyl-CoA
ADP	Adenosindiphosphat
AKG	α -Ketoglutarat
AMP	Adenosinmonophosphat
ASP	Aspartat
ATP	Adenosintriphosphat
CIT	Citrat
CITICIT	Gelumpter Pool aus CIT und ICIT
DHAP	Dihydroxyacetonphosphat
E4P	Erythrose-4-Phosphat

Fortsetzung nächste Seite

INHALTSVERZEICHNIS

Abkürzung	Metabolit
F6P	Fructose-6-Phosphat
FUM	Fumarat
FBP	Fructose-1,6-bisphosphat
G6P	Glucose-6-Phosphat
GA3P	Glyceraldehyd-3-Phosphat
GLC	Glucose
GLN	Glutamin
GLU	Glutamat
GLX	Glyoxylat
ICIT	Isocitrat
LYS	Lysin
MAL	Malat
MET	Methionin
MQ	Menaquinone
MQH ₂	Reduziertes MQ
NAD	Nicotinamid-adenindinucleotid
NADH	Reduziertes NAD
NADP	Nicotinamid-adenindinucleotidphosphat
NADPH	Reduziertes NADP
OAA	Oxalacetat
PEP	Phosphoenolpyruvat
PGP	Gelumpfter Pool aus 13BPG, 3PG und 2PG
P5P	Gelumpfter Pool aus RU5P, XU5P und R5P
PYR	Pyruvat
R5P	Ribose-5-Phosphat
RU5P	Ribulose-5-Phosphat
S7P	Sedoheptulose-7-Phosphat
SER	Serin
SUC	Succinat
SUCCOA	Succinyl-CoA
SUCFUM	Gelumpfter Pool aus SUC und FUM
THF	Tetrahydrofolat
THR	Threonin
TYR	Tyrosin
VAL	Valin
XU5P	Xylulose-5-Phosphat

Teil I

Einleitung und Methoden

Kapitel 1

Systembiologie im Omics-Zeitalter

Aufgrund der rasch fortschreitenden Entwicklung experimenteller und analytischer Methoden auf den Gebieten des Genoms, Transkriptoms, Proteoms, Metaboloms und Fluxoms wurden in den letzten Jahren verschiedene Werkzeuge zur Verfügung gestellt, die quantitative Einblicke in die gesamtzellulären Abläufe der einzelnen Regulationsebenen ermöglichen (Abb. 1.1).

1.1 Multi-Omics-Daten

Gegenstand der meisten Untersuchungen sind dabei nicht veränderte (native) Mikroorganismen. Zunehmend rücken aber auch genetisch veränderte Stämme, welche z.B. durch Überexpression oder Deletion spezifischer Gene gekennzeichnet sind, in den Fokus systembiologischer Fragestellungen.

Transkriptomics: Die auf diese Weise erzeugten Mutanten können mit Hilfe von genomweiten Microarrays oder im Fall von einzelnen Genen durch *Reverse Transkriptase-Polymerase-Kettenreaktion* (RT-PCR) charakterisiert werden [Bus00].

Durch die Anwendung bioinformatischer Methoden lassen sich zudem Bindestellen von Regulatoren der Transkription vorhersagen, welche anschließend mit *Electrophoretic Mobility Shift Assays* (EMSA) verifiziert werden können [HF07]. Die darauffolgende Aufreinigung solcher Regulatorproteine, z.B. über eine Kopplung von *His-tag*-Fusion und Affinitätschromatographie, ermöglicht die Suche nach potentiellen Effektoren, welche die DNA-Bindeeigenschaften der Transkriptionsfaktoren kontrollieren [JR89].

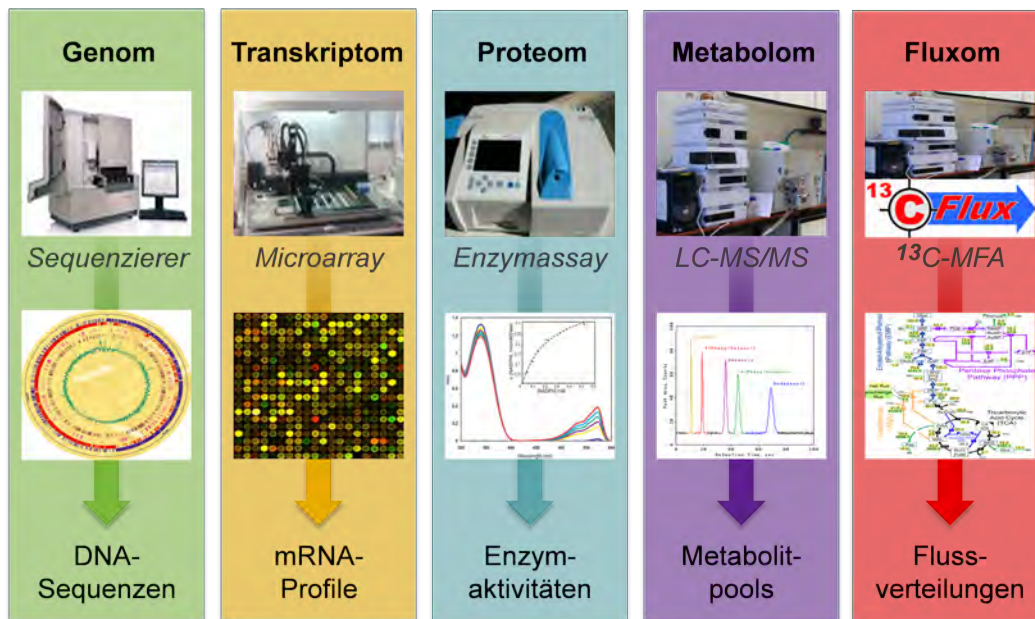


Abbildung 1.1: Beispiele verschiedener Technologien zur Untersuchung der Omics-Ebenen biologischer Systeme.

Proteomics: Darüber hinaus stehen zur Analyse des Proteoms neben den klassischen Enzymaktivitätstests auch Techniken zur Identifizierung unbekannter Proteine mittels Massenspektrometrie (z.B. MALDI-TOF-MS) [EGS⁺02] sowie zur Generierung von Expressionsprofilen durch 2D-Gelelektrophorese [HPB⁺01, BMKB07] zur Verfügung.

Metabolomics: Auf Ebene des Metaboloms wurden in der Vergangenheit verschiedene analytische Methoden zur Quantifizierung intrazellulärer Metabolite in biologischen Extrakten entwickelt. Dazu zählen beispielsweise enzymatische Assays, die Kernspinresonanzspektroskopie (NMR) sowie die Kopplung von Massenspektrometrie mit Gaschromatographie (GC-MS) oder Flüssig-Chromatographie (LC-MS) [OLP⁺07, MRDM⁺07, BCE⁺09]. Durch die gleichzeitige Entwicklung von experimentellen Verfahren zur schnellen Entnahme von Kultivierungsproben und Quenching des Zellstoffwechsels im Sub-Sekundenbereich können zudem intrazelluläre Metabolitdynamiken z.B. nach einem Substratpuls verfolgt werden [OT05].

Fluxomics: Schließlich können im Zusammenspiel von ¹³C-Markierungsexperiment, Messung der ¹³C-Markierungen mittels NMR oder GC/LC-MS und einer modellgestützten Messdatenauswertung metabolische Stoffflüsse quantifiziert werden.

Es liegt nahe, dass die daraus resultierende heterogene Datenvielfalt viel zu komplex ist, um sie mit klassischen reduktionistischen Denkansätzen der einfachen Gen-Funktion-Zuweisung zu hinterfragen. Vielmehr sind systembiologische Methoden notwendig, um die dynamischen, nicht-linearen und adaptiven Teilsysteme einer Zelle in Hinblick auf ihre Funktion und den Kontext der gesamten Zelle zu verstehen.

1.2 Rolle der Systembiologie

Mehr als ein halbes Jahrhundert Forschung im Bereich der Molekularbiologie und Biochemie lässt den klaren Schluss zu: Information fließt von einer größtenteils stabilen Gen-Bibliothek (DNA) über instabile Botenmoleküle (mRNA), welche zur Synthese molekularer Maschinen (Enzyme) genutzt werden, zu funktionellen Prozessen (metabolische Reaktionen), die für das Leben einer Zelle notwendig sind (Abb. 1.2).

Unter dem damit verbundenem wissenschaftlichen Paradigma, dass ein Fluss über eine metabolische Reaktion direkt von der Konzentration des jeweiligen Enzyms abhängig ist, wurden in der Vergangenheit eine Vielzahl biochemischer Untersuchungen mit isolierten Enzymen im Reagenzglas (*in vitro*) durchgeführt. Mittlerweile hat sich diese Sichtweise jedoch radikal geändert und sich die grundlegende Erkenntnis durchgesetzt, dass in der lebenden Zelle (*in vivo*) aufgrund der starken Vernetztheit metabolischer Reaktionen viele Enzyme gleichzeitig einen Reaktionsfluss bestimmen. Ursache dafür sind komplexe Mechanismen der Stoffwechselregulation auf Ebene des Genom, Transkriptom, Proteom und Metabolom.

Die systembiologische Betrachtung mikrobieller Systeme kommt somit einem Paradigmenwechsel gleich. Im Gegensatz zu früheren Ansätzen stehen nicht mehr einzelne, lokale Gen-Wirkungs-Analysen im Mittelpunkt des Interesses, sondern globale, die ganze Zelle umfassende Untersuchungen unter Berücksichtigung der unterschiedlichen Hierarchie-Ebenen und Regulationsprozesse in den lebenden Zellen.

Damit kommt insbesondere der mathematischen Modellbildung in der Systembiologie eine zentrale Rolle zu. Nur durch eine quantitative Beschreibung der Mechanismen zellulärer Regulation, z.B. auf der Basis kinetischer Stoffwechselmodelle, besteht die Möglichkeit zum Verständnis der Zelle als komplexes Wechselwirkungsnetzwerk. Darüber hinaus können derartige Modelle für Prädiktionen in Bezug auf

molekularbiologische Eingriffe genutzt werden, z.B. zur Verbesserung der Produktbildung, Substrat-Selektivität oder Toleranz gegenüber Umwelteinflüssen.

Um der Komplexität der systembiologischen Aufgabenstellung gerecht zu werden, konstituieren sich weltweit Forschungsverbünde aus (Bio-)Chemikern, Molekular- und Mikrobiologen, Biotechnologen, Ingenieuren und Mathematikern zur Bearbeitung der experimentellen, analytischen und modellierungstechnischen Aufgabenstellungen. Im Mittelpunkt steht dabei jeweils ein ausgewähltes biologisches System, wie z.B. im Fall der Systembiologie-Verbünde zu *Saccharomyces cerevisiae* in den Niederlanden, des deutschen Hepatozyten-Systembiologie-Projektes oder der International *Escherichia coli* Alliance (IECA).

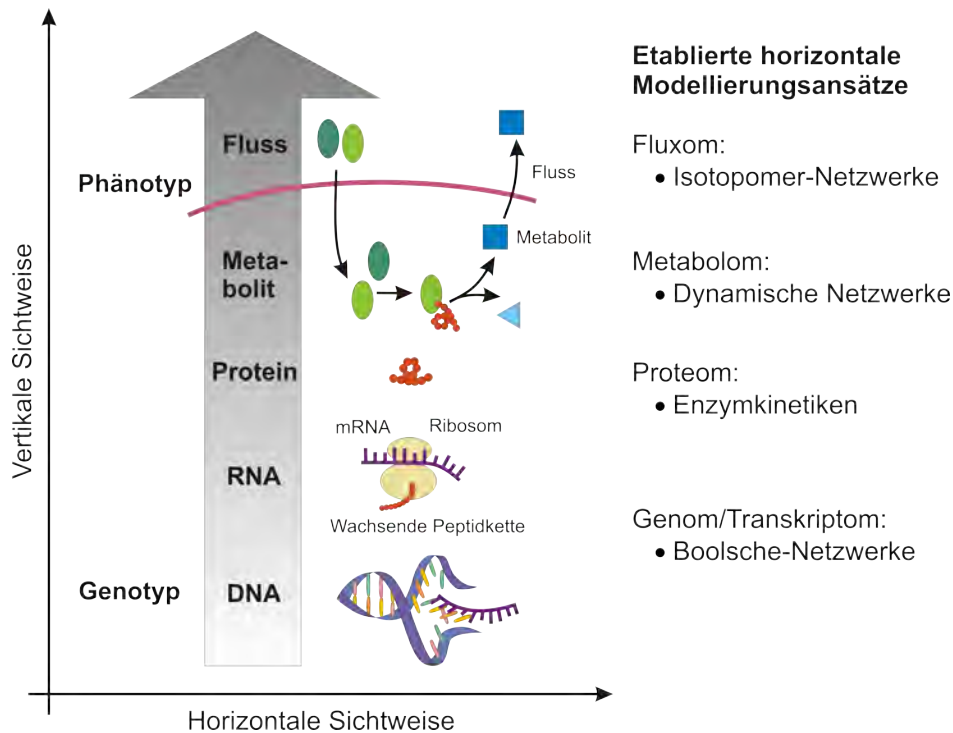


Abbildung 1.2: Schematische Darstellung des Informationsflusses innerhalb der Zelle. Auf der Grundlage der Genominformation in Form von DNA werden innerhalb der Transkription mRNA-Moleküle synthetisiert, welche anschließend zu Proteinen translatiert werden. Einige dieser Proteine fungieren als Enzyme, die verschiedene metabolische Reaktionen katalysieren. Am Ende der gesamten Kaskade stehen die metabolischen Flüsse, welche gleichzeitig als phänotypische Indikatoren dienen. „Horizontale“ Modellierungsansätze dienen der Beschreibung von Stoffwechselprozessen innerhalb der verschiedenen zellulären Regulationsebenen.

1.3 Modellorganismus *C. glutamicum*

Mit dem Projekt „Systemorientierte Analyse des Zentralstoffwechsels mikrobieller Aminosäureproduzenten“ (SysMAP) wurde Ende 2005 ein vergleichbarer Forschungsverbund geschaffen, bei welchem der Organismus *Corynebacterium glutamicum* im Mittelpunkt steht. Für die Wahl dieses gram-positiven Bakteriums als Modellorganismus im Bereich der Systembiologie gibt es mehrere Gründe:

- Seit seiner Entdeckung im Jahr 1957 ist dieser Organismus Gegenstand zahlreicher biochemischer und prozesstechnischer Untersuchungen, d.h. es existieren mittlerweile fundierte Informationen, die u.a. für die Formulierung von Stoffwechselmodellen notwendig sind.
- Das Genom von *C. glutamicum* ist vollständig sequenziert und es existiert eine Vielzahl molekulargenetischer Methoden zur gezielten Genexpression und -deletion.
- *C. glutamicum* zählt zu den weltweit wichtigsten Organismen für die biotechnologische Aminosäureproduktion.

Innerhalb des SysMAP-Projektes arbeiten neben den beiden industriellen Kooperationspartnern Evonik Degussa GmbH (Halle/Westfalen) und Genedata GmbH (Martinsried) die Institute für Biotechnologie 1 und 2 der Forschungszentrum Jülich GmbH sowie als universitäre Einrichtungen die Institute für Mikrobiologie (Erlangen), Mikrobiologie und Biotechnologie (Ulm), Genomforschung (Bielefeld) und Biochemie (Köln) eng zusammen.

1.4 Fokus Zitratzyklus

Im Unterschied zu Projekten, bei denen die Modellierung der gesamten Zelle im Vordergrund steht, erfolgt im SysMAP-Projekt eine klare Beschränkung auf einen Stoffwechselweg, den Tricarbonsäurezyklus (TCA). Dieser ist von zentraler Bedeutung, da er:

- in vollständiger oder unvollständiger Form in allen Zellen vorkommt und grundlegende neue Einsichten in die Regulation dieses Stoffwechselweges daher von großem Interesse sind,

- Aufgaben in der Energieversorgung (Bereitstellung von ATP und Reduktionsäquivalenten für die Atmungskette) und im Baustoffwechsel übernimmt,
- die Vorstufen für biotechnologisch produzierte Aminosäuren (z.B. Glutamat, Aspartat, Lysin, Threonin usw.) liefert, gleichzeitig aber für einen signifikanten Verlust von Kohlenstoff in Form von CO₂ verantwortlich ist und
- untrennbar mit der Stickstoffassimilation verbunden ist, ein weiterer wesentlicher Aspekt der mikrobiellen Aminosäure-Produktion.

Der TCA ist demnach ein ideales Objekt zur Optimierung von Produktionsstämmen und zur Etablierung von Methoden, die eine systembiologische Analyse der ganzen Zelle erlauben.

1.5 Motivation dieser Arbeit

Die vorliegende Arbeit ist in das SysMAP-Projekt eingebunden, in welchem der Modellorganismus *C. glutamicum* unter vorrangig systembiologischen Gesichtspunkten untersucht wird. Grundlegende Ziele sind hierbei, durch existierende und neu zu etablierende Methoden der Transkriptom-, Proteom-, Metabolom- und Fluxom-Analyse zuverlässige quantitative Daten auf allen Ebenen der Stoffwechselregulation zu erhalten, d.h. Genexpressionsdaten, Enzymaktivitäten, intrazelluläre Metabolitkonzentrationen sowie metabolische Stoffflüsse unter definierten Reaktionsbedingungen.

Diese Multi-Omics-Daten sollen anschließend in einem iterativen Prozess zur Modellierung des TCA eingesetzt werden, um schliesslich in einem quantitativ-mechanistischen Verständnis dieses Stoffwechselweges im Kontext der gesamten Zelle zu resultieren.

Die hier vorgestellte Arbeit ist im Bereich der Modellierung und Simulation angesiedelt, wobei insbesondere Methoden zur modellgestützten Gewinnung und Auswertung quantitativer Multi-Omics-Datensätze untersucht werden.

Kapitel 2

Zielsetzung und Gliederung

2.1 Schwerpunkte der Arbeit

Für die Schaffung einer geeigneten Multi-Omics-Datenbasis werden zunächst Experimente mit unterschiedlichen *C. glutamicum*-Stämmen (Wildtyp, Lysinproduzenten) durchgeführt und entsprechende Kultivierungsproben hinsichtlich Transkriptom, Proteom, Metabolom und Fluxom analysiert (Abb. 2.1). Die Messung und Präprozessierung von genomweiten mRNA-Daten sowie Enzymaktivitäten des Zitratzyklus von *C. glutamicum* erfolgt dabei von den SysMAP-Kooperationspartnern. Hingegen ist die Bereitstellung von Nutzdaten intra- und extrazellulärer Metabolitkonzentrationen sowie metabolischer Stoffflüsse im Zentralstoffwechsel dieses Organismus unter anderem Gegenstand der eigenen Arbeit.

Gewinnung quantitativer Fluxom-Daten

Zur experimentellen *in-vivo*-Quantifizierung unbekannter Stoffflüsse (Fluxom) hat sich die metabolische Stofffluss-Analyse auf der Grundlage von ^{13}C -Markierungsexperimenten (^{13}C -MFA) im Bereich der Systembiologie als ein zuverlässiges Standardwerkzeug etabliert. Im Gegensatz zu vielen anderen biochemischen Messmethoden erfordert die ^{13}C -MFA allerdings ein modellgestütztes Auswertungsverfahren, mit dem aus den Rohdaten (^{13}C -Markierungsmessungen) die gesuchten Stoffflüsse ermittelt werden können.

Mittlerweile wurde die klassische, isotopisch stationäre Methode [Wie95] erweitert und es können auch Experimente unter Verfolgung der transienten Isotopenanreicherung durchgeführt und mit Hilfe eines isotopisch instationären Modellansatzes (INST- ^{13}C -MFA) ausgewertet werden [Nöh06].

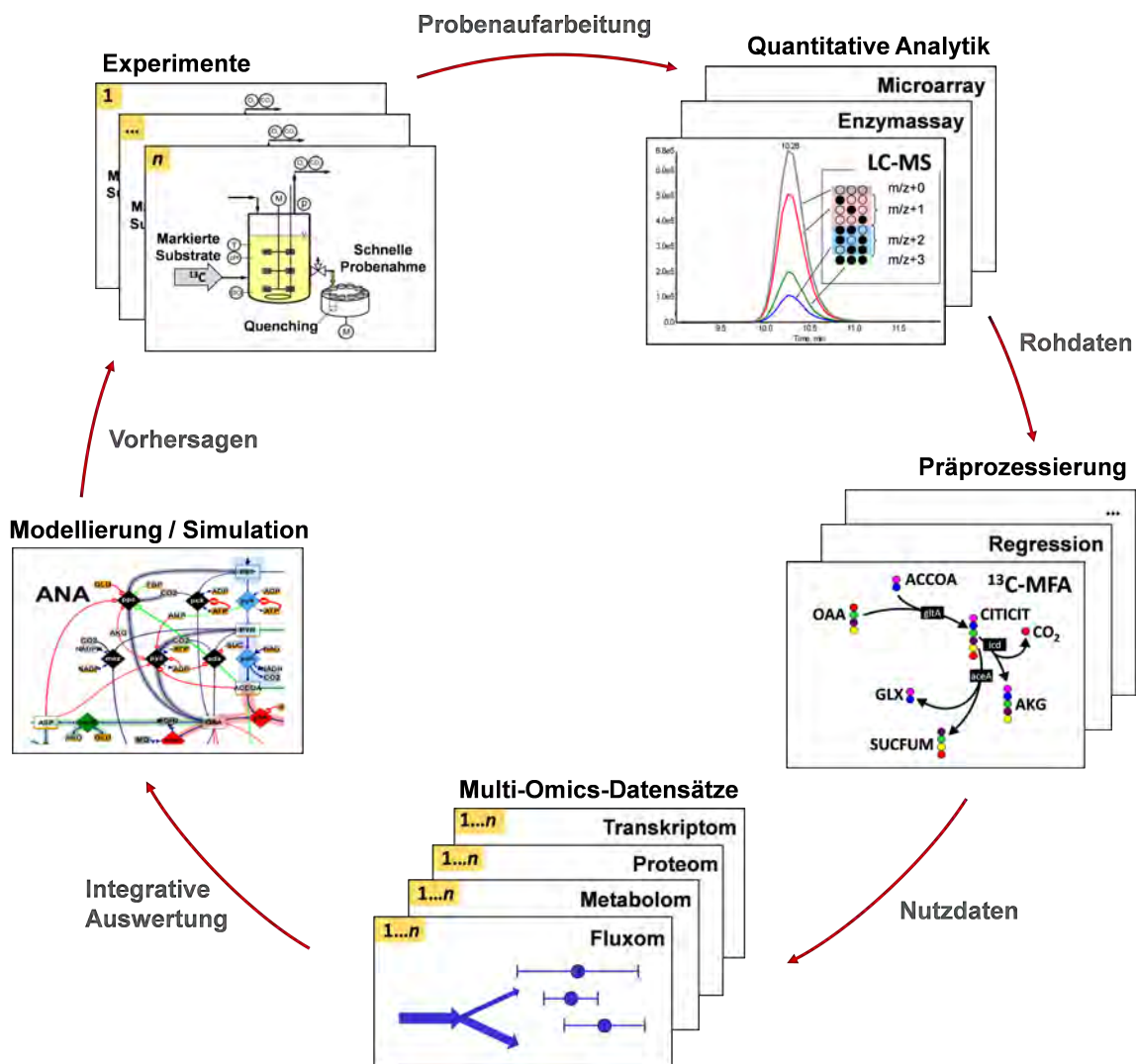


Abbildung 2.1: Zyklus der Auswertung von Multi-Omics-Daten. In einem ersten Schritt werden auf der Grundlage mehrerer Experimente mit unterschiedlichen Stämmen und Wachstumsbedingungen eine Vielzahl von Kultivierungsproben gewonnen. Nach einer entsprechenden Aufarbeitung werden die Probenextrakte durch Anwendung verschiedener Omics-Technologien analysiert. Die daraus resultierenden Rohdaten werden anschließend modellgestützt (z.B. Stoffflussanalyse zur Gewinnung von Fluxom-Daten) oder mittels statistischer Methoden (z.B. Regression zur Ermittlung von Metabolom-Daten) in Nutzdaten umgewandelt. Auf diese Weise stehen eine Vielzahl von Multi-Omics-Daten zur Verfügung, die mit Hilfe integrativer Methoden ausgewertet werden müssen. Hierbei spielt die Modellierung eine entscheidende Rolle, wobei insbesondere durch die Entwicklung vertikaler Modellierungsansätze eine vollständige Integration der kohärenten Daten möglich ist. Schließlich können auf der Basis von validierten Netzwerkmodellen Vorhersagen zur Durchführung neuer Experimente bzw. zur Stammoptimierung getroffen werden.

Während die klassische ^{13}C -MFA bereits Gegenstand zahlreicher Untersuchungen verschiedener Organismen (einschließlich *C. glutamicum*) war [MdGW⁺96, PdGE⁺00, WMS⁺04, CMB⁺07, Sch08], handelt es sich bei der INST- ^{13}C -MFA noch um ein relativ neues Verfahren, welches bisher für wenige Organismen wie beispielsweise *E. coli* angewandt wurde [NGL⁺07, SMR08]. Wenngleich mit einem höheren experimentellen, analytischen und modellierungstechnischem Aufwand verbunden, bietet die INST- ^{13}C -MFA ein Reihe wesentlicher Vorteile, die ihre Anwendung auch für *C. glutamicum* interessant machen.

Ein direkter Vergleich zwischen beiden Stoffflussanalyse-Methoden auf der Grundlage eines einheitlichen Datensatzes ist bisher nicht erfolgt.

Vor diesem Hintergrund ergeben sich für den Bereich der Präprozessierung von Fluxom-Daten folgende Zielsetzungen der Arbeit:

- Erstmalige Untersuchung des Modellorganismus *C. glutamicum* im Rahmen eines instationären ^{13}C -Markierungsexperimentes unter Fed-Batch-Kultivierungsbedingungen.
- Auswertung der gewonnenen Markierungsdaten auf der Basis von isotopisch stationären und instationären Modellierungsansätzen. Mit Hilfe der berechneten Stoffflüsse wird anschließend ein direkter Vergleich beider Methoden vorgenommen.
- Ermittlung von Konfidenzintervallen quantitativer Stoffflussdaten, wobei sowohl Verfahren der linearen als auch der nichtlinearen Statistik zum Einsatz kommen.
- Interpretation und Beurteilung der Fluxom-Daten im Rahmen von Simulationsstudien mit Hilfe eines erweiterten genomweiten metabolischen Netzwerkmodells für *C. glutamicum*.

Integrative mechanistische Modellierung

Für die modellgestützte Auswertung von Omics-Daten wurden in der Vergangenheit eine Reihe „horizontaler“ Modellierungsansätze entwickelt, mit deren Hilfe Stoffwechselprozesse innerhalb einer bestimmten Regulationsebene der Zelle (Genom, Proteom oder Metabolom) beschrieben werden können (vgl. Abb. 1.2). Somit konnte das qualitative aber auch quantitative Verständnis über die jeweilige Stoffwechselebene erweitert werden.

Wie bereits angedeutet, werden durch Nutzung der verschiedenen Omics-Technologien heutzutage eine Vielzahl quantitativer Messdaten entlang aller Regulationsebenen einer Zelle erhoben (vgl. Abb. 1.1). Allerdings gestaltet sich in vielen Fällen die Interpretation solcher Daten, ohne den zugehörigen Netzwerkkontext zu berücksichtigen, als schwierig und teilweise erfolglos.

Es werden daher dringend „vertikale“ Modellierungsansätze benötigt, welche in der Lage sind, die zugrunde liegenden zellulären Regulationsprozesse der verschiedenen Omics-Ebenen quantitativ zu beschreiben. Grundsätzlich ist eine Kopplung von Modellen, welche einerseits den Stofffluss auf Metabolomebene und andererseits die zugehörige genetische Regulation beschreiben, über die Aktivität einzelner Enzyme möglich.

Weitere Zielsetzungen der vorliegenden Arbeit sind daher:

- Entwicklung eines möglichst universellen vertikalen Modellierungsansatzes auf einem bezüglich der Daten sinnvollen Detaillierungsniveau. Mit der Fokussierung auf den Stoffwechselabschnitt des Zitratzyklus kann ein Modell mittlerer Komplexität formuliert werden, für dessen Validierung die oben erwähnte Datenqualität gewährleistet werden kann.
- Technische Implementierung des vertikalen Netzwerkmodells unter Nutzung der Modellierungssprache Modelica, die im Bereich der Biotechnologie bisher kaum eingesetzt wurde. Zu diesem Zweck ist daher die Erstellung einer geeigneten Komponentenbibliothek notwendig, mit deren Hilfe biochemische Netzwerkmodelle formuliert werden können.
- Validierung des vertikalen Modells unter Nutzung der zuvor generierten Multi-Omics-Datensätze. Diese müssen zunächst hinsichtlich ihrer Konsistenz (Massen-, C-Bilanzen) geprüft und mittels statistischer Verfahren (Fehlerabschätzung, Reproduzierbarkeit usw.) analysiert werden.
- Durchführung einer Sensitivitätsanalyse auf Grundlage des zuvor validierten Netzwerkmodells. Ziel ist hierbei die Prädiktion geeigneter Stellen für molekularbiologische Eingriffe zur verbesserten Produktbildung.

Die im Zuge der Stoffflussanalyse bzw. integrativen Modellierung erzeugten umfangreichen Simulationsergebnisse werden mit Hilfe neuartiger Visualisierungstechniken ausgewertet. Dadurch besteht die Möglichkeit die komplexen regulatorischen

Wechselwirkungen im jeweiligen biochemischen Netzwerk aufzulösen und damit einen Zugang zur Interpretation der Daten zu gewährleisten.

2.2 Gliederung der Arbeit

Die vorliegende Arbeit ist in drei Teile und einen Anhang gegliedert. Im ersten Teil wird einführend ein Überblick der Modellierung biochemischer Netzwerke gegeben. Neben den Mechanismen der Stoffwechselregulation wird hierbei insbesondere auf bereits etablierte horizontale Modellierungsansätze zur Beschreibung von Omics-Daten eingegangen. Anschließend wird eine Einführung in die Modellierungssprache Modelica gegeben. Die im Rahmen dieser Arbeit entwickelten Methoden zur Nutzung von Modelica als Modellierungswerkzeug biochemischer Netzwerke werden dabei anhand einiger Beispiele erläutert.

Der zweite Teil beschäftigt sich mit der metabolischen Stoffflussanalyse im Modellorganismus *C. glutamicum*. Grundlage hierfür bildet die Durchführung eines isotopisch instationären Markierungsexperimentes im Fed-Batch-Prozess, welches im Zusammenhang mit einer Prozessmodellierung beschrieben wird. Danach werden die genutzten Methoden zur Berechnung quantitativer intrazellulärer Stoffflüsse erläutert. Der Schwerpunkt liegt hierbei auf einem Vergleich des isotopisch stationären (^{13}C -MFA) und instationären (INST- ^{13}C -MFA) Modellierungsansatzes unter Verwendung verschiedener Netzwerk- und Messkonfigurationen. Für eine Verifizierung und Interpretation der auf diese Weise erhaltenen Stoffflussverteilungen werden Simulationsstudien unter Nutzung eines genomweiten Stoffwechsellmodells durchgeführt.

Im dritten Teil wird ein neuartiges Modellierungskonzept zur integrativen Auswertung von Multi-Omics-Daten vorgestellt. Die auf Basis von kontinuierlichen Kultivierungen gewonnenen konsistenten Messdatensätze werden zunächst unter stamm- und wachstumsspezifischen Gesichtspunkten miteinander verglichen. Im Anschluss daran wird der zur Integration dieser Messdaten entwickelte Modellierungsansatz ausführlich beschrieben. Hierbei wird insbesondere Wert auf die damit verbundenen notwendigen Modellannahmen gelegt. Der Aufbau des vertikalen Netzwerkmodells erfolgt mit Hilfe von Modelica, wobei die zuvor entwickelte Komponentenbibliothek zum Einsatz kommt. Abschließend werden Methoden zur Validierung dieses Modells unter Einbindung der verfügbaren experimentellen Daten erläutert.

Kapitel 3

Modellierung biochemischer Netzwerke

In den nachfolgenden Abschnitten wird in kompakter Form auf die verschiedenen Regulationsebenen innerhalb des bakteriellen Stoffwechsels eingegangen. Dabei soll insbesondere vor Augen geführt werden, dass sich aufgrund der Vielzahl von regulatorischen Wechselwirkungen eine Modellierung derartiger Prozesse als problematisch erweist. Neben ihrer Komplexität sind die meisten Regulationsprinzipien bisher nur qualitativ und nicht quantitativ verstanden. Aus diesem Grund besteht von vornherein die zwingende Notwendigkeit starke Vereinfachungen bei der Prozessbeschreibung vorzunehmen, um auf der Basis von Modellen geringerer Komplexität prädiktive Aussagen zu ermöglichen.

3.1 Mechanismen der Stoffwechselregulation

Der metabolische Fluss durch einen Stoffwechselweg wird vor allem von den Aktivitäten der beteiligten Enzyme bestimmt. Eine Regulation dieser Aktivitäten kann im Rahmen der Enzymsynthese (Transkriptions-, Translationskontrolle), nach vollendeter Translation (Interkonversion) sowie während der enzymatischen Katalyse (Modulation) erfolgen (Abb. 3.1).

Transkriptionskontrolle: Im ersten Schritt der Biosynthese von Enzymen wird die genetische Information der DNA auf den Zwischenträger mRNA (messenger Ribonucleinsäure) kopiert (Transkription). Die dafür notwendigen Enzyme (RNAP, RNA-Polymerasen) besitzen spezielle Untereinheiten (σ -Faktoren), welche den je-

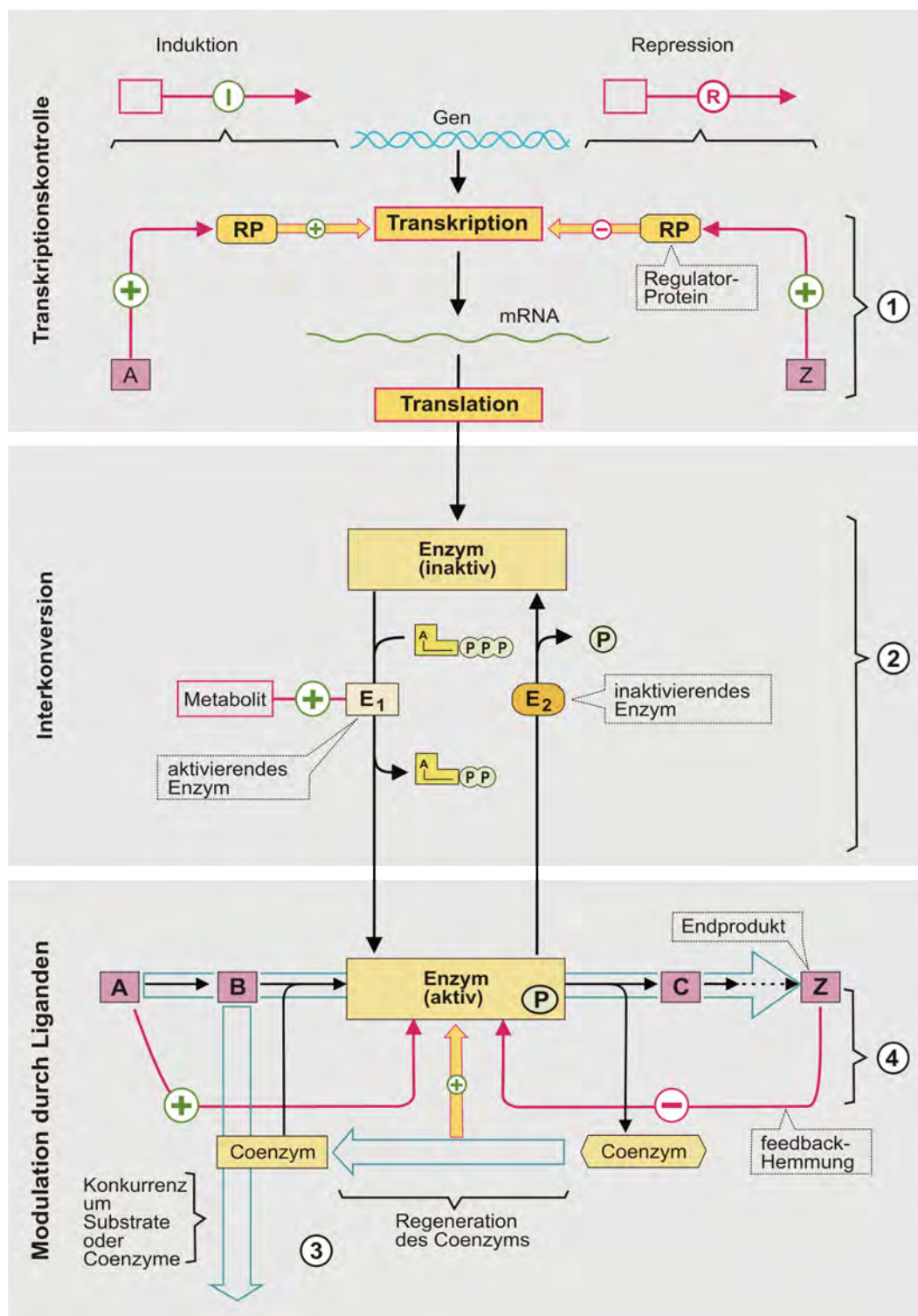


Abbildung 3.1: Schematische Darstellung von Mechanismen der Stoffwechselregulation. 1. Regulatorproteine (RP) können als Induktoren (I) oder Repressoren (R) den Beginn der Enzymsynthese aktivieren bzw. inhibieren. 2. Post-translationale Regulation durch reversible Phosphorylierung. 3. Steuerung der Verfügbarkeit von Substraten (B) und Coenzymen. 4. Beschleunigung oder Hemmung der Enzymkatalyse durch Metabolit-Effektoren (A, Z) (aus [KR03]).

weiligen Kontrollbereich (Promotor) vor einer abzulesenden DNA-Sequenz (Gen) erkennen [BN08]. Der Transkriptionsstart wird durch die Bindung von Regulatorproteinen (RP) an einen Bereich in der Nähe des Promoters (Operator) reguliert [McC85]. Darüber hinaus können weitere Transkriptionsfaktoren (TKF), zu denen auch ungeladene tRNAs (Transport-RNA) und RNA-Aptamere gehören, ein vorzeitiges Ende der mRNA-Elongation bewirken [Stu02, GH03, NM04]. In der Regel wird die Wirksamkeit von TKFs durch Liganden (z.B. Metabolite) beeinflusst, wodurch eine Feedback-Regulation zwischen der Metabolom- und Genomebene entsteht [TCN03].

Translationskontrolle: Innerhalb der Translation wird die Information der mRNA zur Synthese neuer Proteine an den Ribosomen genutzt. Ähnlich wie bei der Transkription können spezifische Faktoren (TLF) an die mRNA binden und damit den Translationsbeginn, z.B. durch Blockierung der Shine-Dalgarno-Sequenz, verhindern [Nud06]. Darüber hinaus hängt die Translationsrate von der Verfügbarkeit an Ribosomen sowie den mit Aminosäuren beladenen tRNAs ab. Die Transkription der dazu notwendigen stabilen RNA-Formen (rRNA und tRNA) unterliegt einer Reihe verschiedener Mechanismen, zu denen z.B. die stringente Kontrolle gehört, worunter eine Reduktion der RNA-Syntheserate als Antwort auf eine Aminosäurelimitation zu verstehen ist [KRA07].

Degradation: Die Zersetzung von mRNA und Proteinen spielt eine wesentliche Rolle für die Regulation der Genexpression. Damit sich ein stationärer Zustand bilden kann, muss ein Gleichgewicht zwischen Synthese und Degradation dieser Komponenten herrschen. Der Abbau von mRNAs in Nukleotide durch unterschiedliche Arten von Nukleasen ist stark reguliert, weshalb auch die Halbwertszeit mit 40 s bis 60 min je nach mRNA stark schwankt [RSSK08]. Die Hydrolyse von Proteinen in Aminosäuren wird durch Proteasen katalysiert.

Interkonversion: Die Regulation der Proteinverfügbarkeit über *de novo*-Synthese oder Degradation ist zu langsam, um den Stoffwechsel an sich ändernde Bedingungen schnell adaptieren zu können. Deshalb verfügt die Zelle über Mechanismen der reversiblen (In)aktivierung von Proteinen (auch Post-Translationale Regulation genannt), wodurch sich einzelne Enzyme direkt an- bzw. abschalten lassen. Interkonversionsprozesse bestehen in den meisten Fällen in einer ATP-abhängigen

Phosphorylierung des Enzyms durch eine Proteinkinase bzw. seiner Dephosphorylierung durch eine Phosphatase [BBS⁺03].

Modulation: Eine sehr schnelle Form der Regulation von Enzymaktivitäten stellt die Modulation durch Liganden (Substrate, Produkte, andere Metabolit-Effektoren) dar. Hierbei existiert eine Vielzahl verschiedener Wirkmechanismen, die sich u.a. in der Bindungsart zwischen Enzym und Effektor unterscheiden.

Im Fall einer kompetitiven Inhibierung bindet der Effektor am aktiven Zentrum des Enzyms, wodurch das jeweilige Substrat verdrängt und die Reaktion verlangsamt wird. Darüber hinaus können Effektoren auch an Stellen außerhalb des aktiven Zentrums binden (allosterische Regulation). Die dadurch bedingte Änderung der sterischen Konformation des Enzyms kann einen aktivierenden oder inhibierenden Effekt auf die Enzymaktivität ausüben. Häufig zu beobachten ist die Inhibierung von Schlüsselenzymen, z.B. bei Eintrittsreaktionen in Aminosäuresynthesewege, durch Endprodukte der betreffenden Reaktionskette (Feedback-Inhibierung).

Neben dem Vorhandensein aktiver Enzyme kann der Fluss durch einen Stoffwechselweg auch durch das Angebot an Cofaktoren (z.B. ATP, NAD usw.) kontrolliert werden. Die Geschwindigkeit einer Cofaktor-abhängigen Reaktion, z.B. Phosphofruktokinase (ATP-gekoppelt), kann dabei durch die Regenerationsrate des Cofaktors in einem anderen Stoffwechselweg, z.B. oxidative Phosphorylierung (Atmungskette), limitiert sein.

3.2 Vereinfachende Annahmen

Für die in den nachfolgenden Kapiteln dargestellten und verwendeten Modellierungsansätze gelten folgende allgemeine Grundannahmen:

- **Durchschnittszelle:** Mikroorganismenpopulationen wachsen in der Regel asynchron, d.h. einzelne Zellen durchlaufen den Zellzyklus unterschiedlich schnell und befinden sich demzufolge zu einem bestimmten Zeitpunkt (z.B. Probenahme) in verschiedenen physiologischen Zuständen (Größe, Gewicht, Alter, chemische Zusammensetzung, Morphologie usw.) [NKB08]. Experimentell erfasste Messgrößen im makroskopischen Bereich stellen somit im Vergleich zur Einzelzellanalyse grundsätzlich Durchschnittswerte der Gesamtpopulation dar.

- **Intrazelluläre Homogenität:** Aufgrund der hohen Dichte an Makromolekülen (z.B. Proteine) innerhalb der Zelle ist es fraglich, ob mittels ungehinderter Diffusion eine homogene Stoffverteilung aller Moleküle vorliegt. Insbesondere für die Makromoleküle selbst ist der Einfluss von stochastischen Effekten im Grunde nicht vernachlässigbar, so dass eine Modellierung auf der Basis von Konzentrationen nicht möglich ist. Für die Erstellung handhabbarer Modelle ist die Annahme einer räumlichen Gleichverteilung aller betrachteten Zellkomponenten jedoch notwendig.
- **Extrazelluläre Homogenität:** Prozessgrößen des Bioreaktors (T , pH , pO_2 usw.) können infolge unzureichender Durchmischung lokalen Schwankungen unterliegen. Ein Einfluss derartiger Konzentrationsgradienten auf das Zellwachstum wird vernachlässigt.
- **Ausschluss von Metabolic Channeling:** In der Literatur gibt es Hinweise, dass bestimmte Enzyme nicht nur die im Cytosol frei vorliegenden Metabolite binden, sondern auch auf direktem Weg die Reaktanden des Vorläufer- bzw. Nachfolgerenzym aufzunehmen können. Das Ablaufen einer solchen Kaskade wird als „Metabolic Channeling“ bezeichnet. Aufgrund der damit verbundenen Konzentrationsunterschiede zwischen Cytosol und Channel gestaltet sich eine mechanistische Beschreibung derartiger Enzymkinetiken als äußerst schwierig [Sav91].

Darüber hinaus gibt es eine Vielzahl weiterer Einflussfaktoren (z.B. Probenahme und -prozessierung) auf das zu untersuchende Zellsystem, die zwangsläufig unberücksichtigt bleiben oder als vernachlässigbar klein angenommen werden müssen. Mit Hilfe dieser Annahmen ist es möglich, biochemische Netzwerke auf der Grundlage von Zustandsgrößen, wie z.B. Metabolitpools, Flüsse, mRNA oder Proteinen, zu beschreiben. In den folgenden Abschnitten wird insbesondere auf die datengetriebene Modellierung im Bereich metabolischer und genetischer Netzwerke eingegangen.

3.3 Metabolische Netzwerkmodellierung

Die fortlaufende Entwicklung von experimentellen und analytischen Methoden zur Messung metabolischer Intermediate mit und ohne Markierungsinformationen [OLP⁺07] führte zur Formulierung verschiedener quantitativer Modellierungsansätze für die Beschreibung dieser Messdaten.

3.3.1 Massenbilanzen von Stoffwechselintermediaten

Auf der Grundlage der im vorangegangenen Kapitel eingeführten Annahmen lassen sich prinzipiell sämtliche Stoffflüsse einer Zelle in zwei Gruppen einteilen (Abb. 3.2):

1. **Extrazelluläre Flüsse**, welche Aufnahmesysteme (\mathbf{v}_{upt}) für eine Vielzahl von Substraten (S) in die Zelle sowie Exkretionssysteme (\mathbf{v}_{exc}) für Produkte (P) aus der Zelle umfassen.
2. **Intrazelluläre Flüsse**, welche Reaktionen (\mathbf{v}_{met}) zwischen metabolischen Intermediaten (IM) und Abflüsse \mathbf{v}_{bm} von Intermediat-Vorstufen in Richtung Biomassekomponenten (BM) umfassen.

Auf der Grundlage dieser Einteilung lässt sich eine Massenbilanz für die molare Stoffmenge n_{IM} jedes Intermediates einer Zelle formulieren:

$$\frac{dn_{\text{IM}}}{dt} = V_{\text{cell}} \cdot \left(\sum v_{\text{upt}} - \sum v_{\text{exc}} \pm \sum v_{\text{met}} - \sum v_{\text{bm}} \right) \quad (3.1)$$

Das intrazelluläre Volumen einer Zelle ist hierbei mit V_{cell} gekennzeichnet.

In der Regel werden metabolische Netzwerke nicht auf Basis einzelner Zellen, sondern ganzer Zellpopulationen formuliert. In diesem Fall hängt die zeitliche Änderung der Intermediatstoffmenge auch von der Biomasse m_X und dem biomassespezifischen intrazellulären Volumen $V_{\text{cell},X}$ ab:

$$\frac{d(m_X \cdot n_{\text{IM},X})}{dt} = n_{\text{IM},X} \cdot \frac{dm_X}{dt} + m_X \cdot \frac{dn_{\text{IM},X}}{dt} = m_X \cdot V_{\text{cell},X} \cdot \sum_i v_i \quad (3.2)$$

Durch Einführung der Beziehung $n_{\text{IM},X} = c_{\text{IM}} \cdot V_{\text{cell},X}$ und der Annahme eines konstanten intrazellulären Volumens $dV_{\text{cell},X}/dt = 0$, kann Gl. 3.2 in eine Massenbilanz

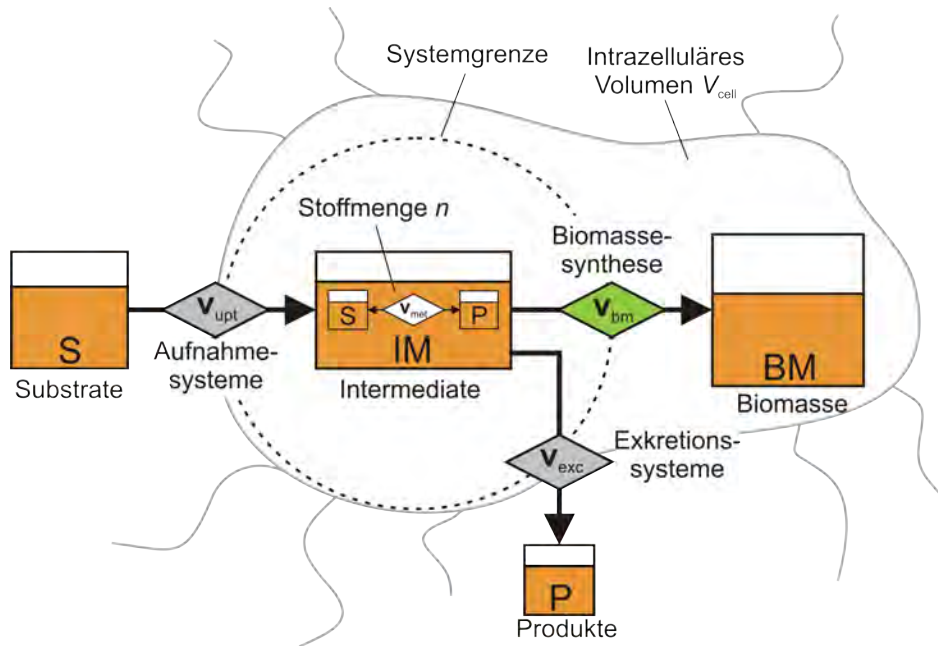


Abbildung 3.2: Schema der Modellierung metabolischer Prozesse innerhalb einer Zelle. Metabolitpools innerhalb der Systemgrenzen (gestrichelte Linie) werden durch Massenbilanzen beschrieben. Die Zelle verstoffwechselt aufgenommene Substrate über verschiedene Intermediate in Richtung Biomassebausteine und Exkretionsprodukte.

für Intermediatkonzentrationen c_{IM} überführt werden:

$$V_{cell,X} \cdot \left(c_{IM} \cdot \frac{dm_X}{dt} + m_X \cdot \frac{dc_{IM}}{dt} \right) = m_X \cdot V_{cell,X} \cdot \sum_i v_i \quad (3.3)$$

unter Berücksichtigung der spezifischen Wachstumsrate:

$$\mu \equiv \frac{1}{m_X} \frac{dm_X}{dt} \quad (3.4)$$

gilt dann:

$$\frac{dc_{IM}}{dt} = \sum_i v_i - \mu \cdot c_{IM} \quad (3.5)$$

Der Term $\mu \cdot c_{IM}$ repräsentiert die Verdünnung intrazellulärer Metabolite infolge von Biomassewachstum. Dieser Term ist vernachlässigbar, wenn das Zeitfenster der Beobachtung intrazellulärer Dynamiken sehr klein (< 1 min) gegenüber dem Biomassewachstum (z.B. $\mu_{max} = 0.8 \text{ h}^{-1}$ für *E. coli*) ist.

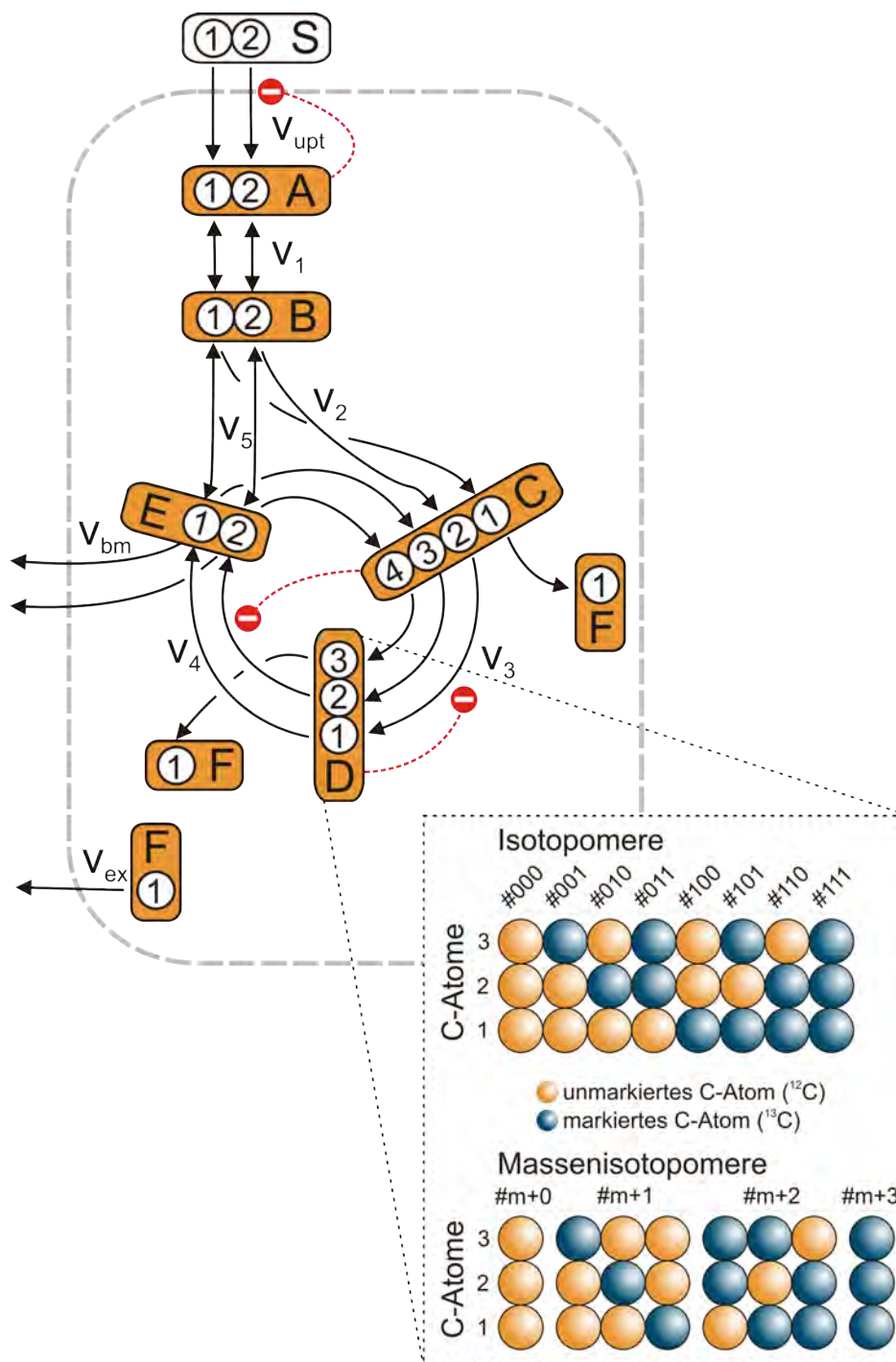


Abbildung 3.3: Beispielnetzwerk mit verschiedenen Ebenen metabolischer Netzwerkmodellierung. Ebene 1 (orange Boxen): Metabolisches Netzwerk mit Flüssen zwischen Intermediatpools. Ebene 2 (weiße Kreise): C-Atom Transitionsnetzwerk mit Umsetzungen aller Kohlenstoffatome innerhalb der Intermediate. Für den Pool D sind zwei verschiedene Beschreibungsarten des isotopischen Zustands dargestellt: Isotopomere (isotopische Isomere) und Massenisotopomere (Isotopomere gleicher Masse). Ebene 3 (rote Kreise): Regulationsnetzwerk mit Effektoren einzelner enzymatischer Reaktionen.

3.3.2 Metabolisches Beispielnetzwerk

In Abbildung 3.3 ist ein einfaches Beispielnetzwerk dargestellt. Dessen zeitliche Änderungen intrazellulärer Metabolitkonzentrationen können wie folgt formuliert werden:

$$\begin{aligned}
 \frac{dc_A}{dt} &= v_{\text{upt}} - v_1 - \mu \cdot c_A & \frac{dc_B}{dt} &= v_1 - v_2 - v_5 - \mu \cdot c_B \\
 \frac{dc_C}{dt} &= v_2 - v_3 - \mu \cdot c_C & \frac{dc_D}{dt} &= v_3 - v_4 - \mu \cdot c_D \\
 \frac{dc_E}{dt} &= v_4 + v_5 - v_2 - v_{\text{bm}} - \mu \cdot c_E & \frac{dc_F}{dt} &= v_3 + v_4 - v_{\text{exc}} - \mu \cdot c_F
 \end{aligned} \tag{3.6}$$

Das System kann vereinfacht durch Matrixnotation

$$\frac{d\mathbf{c}_{\text{IM}}}{dt} = \mathbf{N} \cdot \mathbf{v} - \mu \cdot \mathbf{c}_{\text{IM}}$$

mit $\mathbf{N} =$

$$\begin{matrix}
 & v_{\text{upt}} & v_1 & v_2 & v_3 & v_4 & v_5 & v_{\text{bm}} & v_{\text{exc}} \\
 c_A & \left(\begin{array}{cccccccc}
 +1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & +1 & -1 & 0 & 0 & -1 & 0 & 0 & 0 \\
 0 & 0 & +1 & -1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & +1 & -1 & 0 & 0 & 0 & 0 \\
 0 & 0 & -1 & 0 & +1 & +1 & -1 & 0 & 0 \\
 0 & 0 & 0 & +1 & +1 & 0 & 0 & 0 & -1
 \end{array} \right) &
 \end{matrix} \tag{3.7}$$

mit den Vektoren aller Intermediatpools $\mathbf{c}_{\text{IM}} = (c_A, c_B, \dots, c_F)^T$ und Flüssen $\mathbf{v} = (v_1, v_2, \dots, v_{\text{exc}})^T$ dargestellt werden. \mathbf{N} ist die stöchiometrische Matrix der Größe $m \times n$, wobei m der Anzahl aller bilanzierten Metabolite und n der Anzahl aller Reaktionen entspricht.

3.3.3 Massenbilanzen markierter Stoffwechselintermediate

Für die Beschreibung von ^{13}C -Markierungsexperimenten werden Modelle benötigt, die neben der Stöchiometrie des metabolischen Netzwerks auch die C-Atom-Transitionen aller Reaktionen enthalten.

In Abhängigkeit des mathematischen Formalismus für die Modellierung des Markierungszustandes eines bestimmten metabolischen Netzwerks resultiert daraus eine kombinatorische Vervielfachung der Gleichungsanzahl [WMI⁺99, vWHV02, AKS07b].

Generell gilt, dass alle Ansätze, welche auf dem Konzept der Isotopomere beruhen, eine vollständige Beschreibung des Markierungszustandes aller im Modellfokus liegenden Intermediate ermöglichen (vgl. Abb. 3.3). In diesen Fällen existieren nicht nur eine Massenbilanz für die Gesamtkonzentration eines Intermediats, sondern 2^n Gleichungen zugehöriger Isotopomer-Konzentrationen, wobei n der Anzahl an C-Atomen eines IM entspricht. Auf der Basis von Isotopomeren lassen sich Massenbilanzen für markierte Intermediatpools $\mathbf{c}_{\text{IM}}^\#$ wie folgt formulieren:

$$\frac{d\mathbf{c}_{\text{IM}}^\#}{dt} = \mathbf{f}(\mathbf{v}^\rightleftharpoons, \mathbf{c}_{\text{IM}}^\#, \mathbf{c}_S^\#) - \mu \cdot \mathbf{c}_{\text{IM}}^\#, \quad \mathbf{c}_{\text{IM}}^\#(0) = \mathbf{c}_{\text{IM},0}^\# \quad (3.8)$$

mit $c_{\text{IM}}(t) = \sum_{i=1}^{2^n} c_{\text{IM},i}^\#(t)$

Die Funktion \mathbf{f} ist dabei linear abhängig von den intrazellulären Flüssen $\mathbf{v}^\rightleftharpoons$, den Substrat-Isotopomeren $\mathbf{c}_S^\#$ und der Wachstumsrate μ und kann nichtlinear in Bezug auf die Intermediat-Isotopomere $\mathbf{c}_{\text{IM}}^\#$ sein.

Bei der Modellierung isotopischer Systeme besteht die Notwendigkeit der Formulierung von Massenbilanzen, welche auf einer Trennung der Nettoraten $v^{\text{net}} = v^\rightarrow - v^\leftarrow$ in Hin- und Rückflüsse ($v^\rightarrow, v^\leftarrow$) basieren. Die Ursache dafür liegt in der Änderung der Markierung eines Intermediats, welche von den Markierungsanteilen aller zufließenden Pools abhängt. Im Fall einer reversiblen Reaktion hängt demzufolge die Markierung, welche einen Substratpool durch einen Rückfluss erreicht, von der Markierung des Produktpools ab und stimmt nicht notwendigerweise mit der Markierung des Substrats überein, außer die betreffende Reaktion ist durch einen hohen Austauschfluss gekennzeichnet.

Beispielhaft ist die Isotopomeren-Massenbilanz um den Knoten B des Netzwerks in Abb. 3.3 gegeben durch:

$$\frac{dc_B^{\#ij}}{dt} = v_1^\rightarrow \cdot \frac{c_A^{\#ij}}{c_A} + v_5^\leftarrow \cdot \frac{c_E^{\#ij}}{c_E} - (v_1^\leftarrow + v_2^\rightarrow + v_5^\rightarrow) \cdot \frac{c_B^{\#ij}}{c_B} - \mu \cdot c_B^{\#ij} \quad i, j \in \{0, 1\} \quad (3.9)$$

3.3.4 Stöchiometrische Modellierung

Metabolische Stoffflussanalyse

Metabolische Stoffflüsse stellen die letztendliche Ausprägung zellulärer Funktionen einer Zelle unter verschiedenen physiologischen Bedingungen dar (vgl. Abb. 1.2). Zur Quantifizierung der intrazellulären Flüsse einer Zelle wurde die Methode der

metabolischen Stoffflussanalyse (MFA) entwickelt [Ste99]. Um diesen Ansatz nutzen zu können, muss die Bedingung der metabolischen Stationarität $dc_{\text{IM}}/dt = 0$ erfüllt sein, d.h. die Konzentrationen aller Intermediate ändern sich im Zeitraum des Experimentes nicht. Es ist im Allgemeinen akzeptiert, dass diese Bedingung in Bioreaktoren unter kontinuierlicher Prozessführung (z.B. Chemostat) sowie in der exponentiellen Wachstumsphase von Zellpopulationen in Batch-/Fed-Batch-Kulturen annähernd erfüllt ist.

Unter der Annahme eines metabolisch stationären Systemzustandes vereinfacht sich Gl. 3.7 zu:

$$\mathbf{0} = \mathbf{N} \cdot \mathbf{v} - \mu \cdot \mathbf{c}_{\text{IM}} \quad (3.10)$$

In den meisten Fällen umfasst das betrachtete Netzwerk mehr Reaktionen als Intermediatpools ($n > m$) wodurch das resultierende Gleichungssystem unterbestimmt ist. Die Einbeziehung von Flussmessungen (z.B. Substrataufnahme- und Produktsekretionsraten sowie Biomasseabflüsse) reicht für gewöhnlich nicht aus, um alle intrazellulären Flüsse getrennt zu identifizieren. Folglich können nur Linearkombinationen aller möglichen Flüsse ermittelt werden.

Für das Beispielnetzwerk kann eine solche Linearkombination durch Festlegung des Zuflusses v_{upt} sowie der beiden Abflüsse v_{bm} und v_{exc} erhalten werden.

Flussbilanzanalyse

Eine Möglichkeit, das Problem der Unterbestimmtheit stöchiometrischer Netzwerke einzugrenzen, besteht in der Anwendung von Methoden der linearen Programmierung/Optimierung. Dabei wird der Lösungsraum mittels Maximierung oder Minimierung einer bestimmten Zielfunktion beschränkt (Abb. 3.4) [SP92, SELP00, ECP02].

In den meisten Fällen wird dabei angenommen, dass der metabolische Phänotyp einer Zelle durch die Tendenz zum maximalen Wachstum definiert ist [ERP02]. Diese Annahme sollte zumindest in substratreichen Umgebungen, wie z.B. Bioreaktoren, zutreffen. Eine entsprechende Zielfunktion ist dann gegeben durch:

$$\sum_i = Y_{\text{IM}/X,i} \cdot \text{IM}_i \xrightarrow{v_{\text{bm}}} X \quad (3.11)$$

$$\min \kappa = -v_{\text{bm}}$$

Mit $Y_{IM/X,i}$ wird dabei der Bedarf des i -ten Precursors IM_i für die Synthese von Biomasse X bezeichnet.

Daneben sind eine Reihe anderer Zielfunktionen, wie z.B. Maximierung der ATP-Produktion oder der Synthese eines bestimmten Zielproduktes, denkbar [REMP01, LPK⁺07].

Per Definition liefern lineare Optimierungsmethoden extreme Lösungen, welche an den Endpunkten des zulässigen Definitionsbereiches liegen. In Abhängigkeit des zugrundeliegenden Netzwerks und der definierten Zielfunktion können jedoch auch multiple Optima gefunden werden [LPDG00].

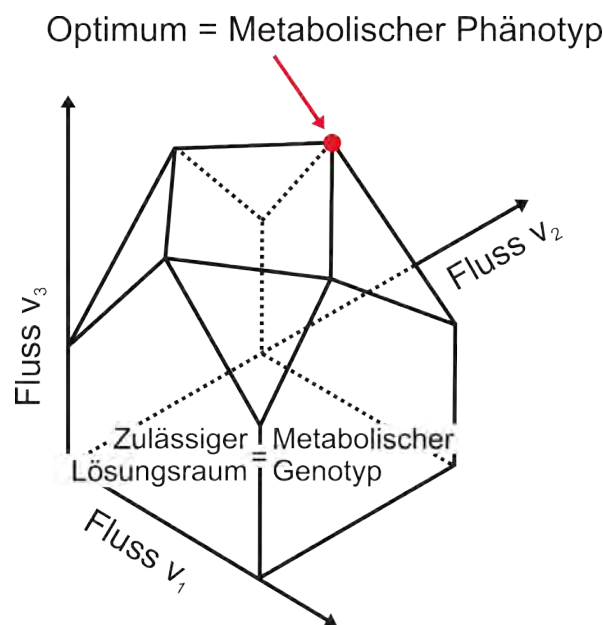


Abbildung 3.4: Prinzip der Flussbilanzanalyse (FBA). Auf Grundlage der stöchiometrischen Nebenbedingungen spannen alle Reaktionen eines Netzwerks einen n -dimensionalen Raum auf, der konzeptionell als „metabolischer Genotyp“ betrachtet wird. Eine bestimmte Kombination an Flusswerten (Stoffflussverteilung) innerhalb dieses zulässigen Definitionsbereiches wird demnach als Ausprägung eines bestimmten „metabolischen Phänotyps“ angesehen (aus [VP94]).

¹³C-Stoffflussanalyse

Eine andere Vorgehensweise zur Ermittlung quantitativer Fluxom-Daten in unterbestimmten stöchiometrischen Netzwerkmodellen beruht auf der Integration zusätzlicher Messdaten in Form intrazellulärer Markierungsanreicherungen in den Intermediatpools einer Zelle [WdG97].

Durch Einführung der Beziehung $\mathbf{c}_{\text{IM}}^{\#} = c_{\text{IM}} \cdot \mathbf{x}_{\text{IM}}^{\#}$, wobei c_{IM} und $\mathbf{x}_{\text{IM}}^{\#}$ die Gesamtkonzentration eines Intermediats sowie zugehörige Isotopomerenanteile kennzeichnet, und Anwendung der Produktregel lässt sich Gl. 3.8 umformen:

$$\frac{d\mathbf{c}_{\text{IM}}^{\#}}{dt} = \frac{d(c_{\text{IM}} \cdot \mathbf{x}_{\text{IM}}^{\#})}{dt} = c_{\text{IM}} \cdot \frac{d\mathbf{x}_{\text{IM}}^{\#}}{dt} + \mathbf{x}_{\text{IM}}^{\#} \cdot \frac{dc_{\text{IM}}}{dt} \quad (3.12)$$

Damit lassen sich aus Gl. 3.8 und 3.12 zwei Spezialfälle an Modellierungsansätzen ableiten, welche im Rahmen der ^{13}C -Stoffflussanalyse (^{13}C -MFA) Anwendung finden:

Isotopisch stationäre Zustände: Unter der Bedingung stationärer Markierungspools ($d\mathbf{c}_{\text{IM}}^{\#}/dt = 0$) folgt:

$$\mathbf{0} = \mathbf{f}(\mathbf{v}^{\rightleftharpoons}, \mathbf{x}_{\text{IM}}^{\#}, \mathbf{x}_{\text{S}}^{\#}) - \mu \cdot \mathbf{c}_{\text{IM}}^{\#} \quad (3.13)$$

Isotopisch instationäre Zustände: Unter der Bedingung stationärer Intermediatpools ($d\mathbf{c}_{\text{IM}}/dt = 0$) folgt:

$$c_{\text{IM}} \cdot \frac{d\mathbf{x}_{\text{IM}}^{\#}}{dt} = \mathbf{f}(\mathbf{v}^{\rightleftharpoons}, c_{\text{IM}}, \mathbf{x}_{\text{IM}}^{\#}, \mathbf{x}_{\text{S}}^{\#}) - \mu \cdot \mathbf{c}_{\text{IM}}^{\#}, \quad \mathbf{x}_{\text{IM}}^{\#}(0) = \mathbf{x}_{\text{IM},0}^{\#} \quad (3.14)$$

In beiden Fällen hängt die Funktion \mathbf{f} linear von den intrazellulären Flüssen $\mathbf{v}^{\rightleftharpoons}$, den Isotopomeranteilen des Substrats $\mathbf{x}_{\text{S}}^{\#}$ und der Wachstumsrate μ ab. In Bezug auf die Markierungsanteile $\mathbf{x}_{\text{IM}}^{\#}$ der Intermediate kann \mathbf{f} nichtlinear sein. Dies trifft für alle bimolekulare Reaktionen zu, bei denen zwei Edukte zu einem gemeinsamen Produkt kondensieren.

Als Beispiel sei die Massenbilanz um den Knoten C für den isotopisch instationären Zustand betrachtet (vgl. Abb. 3.3):

$$c_{\text{C}} \cdot \frac{dx_{\text{C}}^{\#ijkl}}{dt} = v_2^{\rightarrow} \cdot x_{\text{B}}^{\#ij} \cdot x_{\text{E}}^{\#kl} - v_3^{\rightarrow} \cdot x_{\text{C}}^{\#ijkl} - \mu \cdot c_{\text{C}} \quad i, j, k, l \in \{0, 1\} \quad (3.15)$$

In den letzten Jahren hat sich insbesondere die ^{13}C -MFA zu einer der Hauptwerkzeuge des Metabolic Engineering entwickelt und wurde erfolgreich für die Untersuchung verschiedener Organismen von Bakterien über Pflanzen bis hin zu tierischen Zellen angewendet [Sau06, Sch08, TMM⁺09].

3.3.5 Kinetische Modellierung

Im Gegensatz zu rein stöchiometrischen Ansätzen wird bei der kinetischen Modellierung auch versucht, die beobachteten Regulationsmechanismen unter *in vivo*-Bedingungen abzubilden. Im Kontext metabolischer Netzwerke bedeutet dies, dass alle enzymkatalysierten Reaktionsschritte durch entsprechende mathematische Modelle beschrieben werden.

Mechanistische Ansätze

Häufig werden die enzymatischen Reaktionen mechanistisch modelliert, indem Massenbilanzen um einzelne Zustandsgrößen wie Enzym (E), Substrat (S), Produkt (P), Effektoren (I, A) und zugehörige Komplexe (ES, EP, EI usw.) formuliert werden (Abb. 3.5). In der Regel wird dabei eine schnelle Einstellung des stationären Zustandes für den ES- bzw. EP-Komplex (Michaelis-Menten Quasi-Stationarität) angenommen.

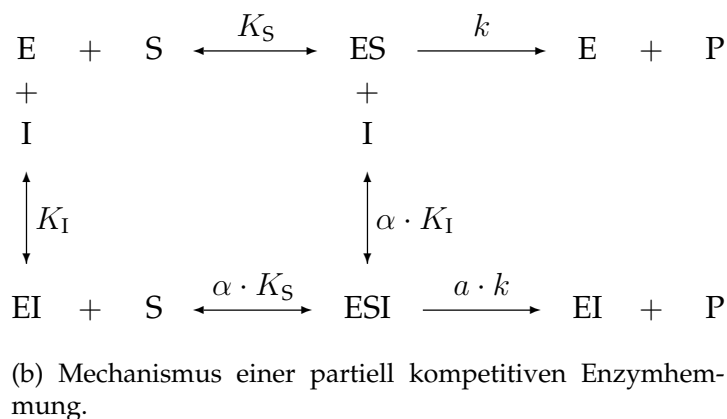
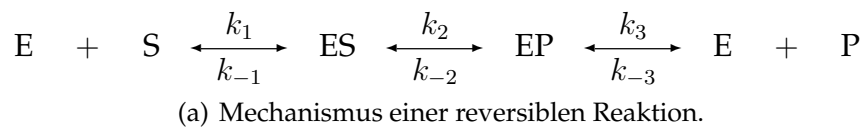


Abbildung 3.5: Modelle verschiedener Enzymmechanismen, aus denen sich kinetische Gleichungen nach Michaelis-Menten ableiten lassen. (E: Enzym, S: Substrat, I: Inhibitor, P: Produkt, ES: Enzym-Substrat-Komplex, EI: Enzym-Inhibitor-Komplex, ESI: Enzym-Substrat-Inhibitor-Komplex)

Zum Beispiel kann die Reaktionsrate v_1 des Beispielnetzwerks (vgl. Abb. 3.3) durch eine reversible Kinetik des Michaelis-Menten-Typs folgendermaßen beschrieben wer-

den [Seg75]:

$$v_1 = \frac{v_{\max,1}^{\rightarrow} \cdot \frac{c_A}{K_{A,1}} - v_{\max,1}^{\leftarrow} \cdot \frac{c_B}{K_{B,1}}}{1 + \frac{c_A}{K_{A,1}} + \frac{c_B}{K_{B,1}}}, \quad \text{mit} \quad K_{\text{eq},1} = \frac{v_{\max,1}^{\rightarrow} \cdot K_{B,1}}{v_{\max,1}^{\leftarrow} \cdot K_{A,1}} \quad (3.16)$$

Mit v_{\max}^{\rightarrow} und v_{\max}^{\leftarrow} wird dabei die maximale Reaktionsrate der Hin- bzw. Rückreaktion gekennzeichnet. Die Terme K_A und K_B stellen Affinitätskonstanten bezüglich der Reaktanden A und B dar. Je höher der Wert dieser Konstanten ist, um so niedriger ist die Affinität des Enzyms zum jeweiligen Metabolit. K_{eq} kennzeichnet die Gleichgewichtskonstante der reversiblen Reaktion.

Unter der Annahme, dass die Enzymaffinität zum Reaktand B unendlich klein ist ($K_{B,1} \rightarrow \infty$), kann die Rückreaktion vernachlässigt werden und Gl. 3.16 vereinfacht sich zur irreversiblen Form der Michaelis-Menten-Kinetik:

$$v_1 = \frac{v_{\max,1} \cdot c_A}{c_A + K_{A,1}} \quad (3.17)$$

Für die Beschreibung von Reaktionen, welche durch Effektoren beeinflusst werden, können je nach zugrunde liegendem Wirkmechanismus der Inhibitoren (I) bzw. Aktivatoren (A) verschiedene Kinetiken hergeleitet werden (Abb. 3.5 b).

Beispielhaft sei hier die irreversible Reaktion v_3 aus Abb. 3.3 betrachtet, für die eine allgemeine Inhibierungskinetik in folgender Form abgeleitet werden kann:

$$v_3 = \frac{v_{\max,3} \cdot c_C}{K_{C,3} \left(\frac{1 + \frac{c_D}{K_{I,D,3}}}{1 + \frac{a \cdot c_D}{\alpha \cdot K_{I,D,3}}} \right) + c_C \left(\frac{1 + \frac{c_D}{\alpha \cdot K_{I,D,3}}}{1 + \frac{a \cdot c_D}{\alpha \cdot K_{I,D,3}}} \right)}, \quad a \geq 0, \alpha > 0 \quad (3.18)$$

Der Term $K_{I,D}$ kennzeichnet dabei die Affinität des Enzyms zum Inhibitor D. Je nach Wahl der Parameter a und α können aus Gl. 3.18 die bekanntesten Inhibierungsmechanismen abgeleitet werden, d.h. kompetitive ($\alpha \rightarrow \infty$), nicht-kompetitive ($a = 0, \alpha = 1$) and teilweise kompetitive ($a < 1, \alpha > 1$) Inhibierung [Seg75].

Grundsätzlich gilt, dass auf der Basis von mechanistischen Ansätzen alle kinetischen Parameter einen direkten Bezug zum biochemischen Kontext besitzen.

Approximative Ansätze

An einer Vielzahl enzymatischer Reaktionen sind zwei, manchmal sogar drei Substrate sowie mehrere Effektoren beteiligt, die sich auch gegenseitig beeinflussen können. Dementsprechend kann sich die Herleitung einer entsprechenden mechanistischen Kinetik, z.B. nach der King-Altman-Methode [Bis02], als sehr aufwendig gestalten. Darüber hinaus werden für die Beschreibung eine hohe Anzahl an Modellparametern benötigt, welche in der Regel nicht durch die zur Verfügung stehenden Messdaten mit ausreichender statistischer Genauigkeit identifiziert werden können [WHOW06].

Aus diesem Grund wurden in der Vergangenheit eine Reihe alternativer mathematischer Ansätze entwickelt, mit deren Hilfe versucht wird, den wahren Enzymmechanismus zu approximieren [Hei05]. Meist wird dabei angenommen, dass die Wechselwirkungen zwischen umzusetzenden Reaktanden und Effektoren vernachlässigbar gering sind oder ausschließlich einem bestimmten Regulationstyp (z.B. nicht-kompetitive Inhibierung) unterliegen. Durch die damit verbundene Betrachtungsweise der Substanzen als mehr oder weniger unabhängige Reaktionspartner kann deren Einfluss auf die Geschwindigkeit der enzymatischen Katalyse mit wesentlich weniger Parametern beschrieben werden.

Grundsätzlich unterscheiden sich alle approximativen Ansätze in der Art der Datenbeschreibung, Anzahl kinetischer Parameter aber auch in ihrer Vorhersagefähigkeit [HNW08]. Im Folgenden werden drei Ansätze in kurzer Form erläutert, welche häufig in der Literatur zu finden sind.

PowerLaw-Approximation: Der PowerLaw-Ansatz wurde im Rahmen der „Biochemical Systems Theory“ (BST) von [SV82] entwickelt und stellt eine generalisierte Form des Massenwirkungsgesetzes dar. Eine enzymatische Reaktion wird dabei durch Produkte von Potenztermen aller Reaktanden beschrieben:

$$v_i = a_i \cdot \prod_j c_{S,j}^{b_{j,i}} \cdot \prod_k c_{P,k}^{-c_{k,i}} \cdot \prod_l c_{A,l}^{d_{l,i}} \cdot \prod_m c_{I,m}^{-e_{m,i}} \quad (3.19)$$

Der Koeffizient a sowie die Exponenten b, c, d, e stellen hierbei kinetische Parameter dar.

Ein Nachteil dieses Ansatzes ist, dass bei gegen Null laufende Konzentrationswerte eines Produktes P oder Inhibitors I die Reaktionsrate eine Singularität besitzt und unendlich groß wird.

LinLog-Approximation: Dieser weit verbreitete Approximationsansatz wurde in Analogie zu thermodynamischen Betrachtungen enzymatischer Reaktionen entwickelt [WvD87]. Dabei wird die Reaktionsgeschwindigkeit durch Summen logarithmischer Terme aller Reaktanden approximiert:

$$v_i = a_i + \sum_j b_{j,i} \cdot \ln(c_{S,j}) - \sum_k c_{k,i} \cdot \ln(c_{P,k}) + \sum_l d_{l,i} \cdot \ln(c_{A,l}) - \sum_m e_{m,i} \cdot \ln(c_{I,m}) \quad (3.20)$$

Die Anzahl kinetischer Parameter beträgt wie beim PowerLaw-Ansatz $n+1$, wobei n der Anzahl aller Reaktanden der betreffenden Reaktion entspricht. Allerdings treten hier unendlich kleine Werte für die Reaktionsrate auf, wenn die Konzentration eines Reaktanden extrem kleine Werte aufweist.

Convenience-Kinetik: Im Gegensatz zu den beiden vorangestellten Approximationsmethoden beruht der von [LK06] vorgeschlagene Convenience-Ansatz auf dem mechanistischen Konzept von Michaelis-Menten. Als grundlegender molekularer Mechanismus aller Reaktionen wird dabei eine reversible Rapid-Equilibrium Random-Order-Kinetik angenommen, welche mit entsprechenden generischen Produktermen für die Wirkung von Effektoren erweitert wird:

$$v_i = \frac{v_{\max,i}^{\rightarrow} \cdot \prod_j \frac{c_{S,j}}{K_{S,j,i}} - v_{\max,i}^{\leftarrow} \cdot \prod_k \frac{c_{P,k}}{K_{P,k,i}}}{\prod_j \left(1 + \frac{c_{S,j}}{K_{S,j,i}}\right) + \prod_k \left(1 + \frac{c_{P,k}}{K_{P,k,i}}\right) - 1} \cdot \prod_l \frac{c_{A,l} + K_{A,l,i}}{K_{A,l,i}} \cdot \prod_m \frac{K_{I,m,i}}{c_{I,m} + K_{I,m,i}} \quad (3.21)$$

mit

$$K_{\text{eq},i} = \frac{v_{\max,i}^{\rightarrow} \cdot \prod_k K_{P,k,i}}{v_{\max,i}^{\leftarrow} \cdot \prod_j K_{S,j,i}}$$

Aufgrund der Trennung in Hin- und Rückreaktion wird hier ein Parameter (v_{\max}^{\leftarrow} bzw. K_{eq}) mehr benötigt als beim PowerLaw- und LinLog-Ansatz. Demgegenüber handelt es sich hier um Sättigungsterme, d.h. die oben beschriebenen Probleme bei kleinen Konzentrationswerten treten nicht auf. Darüber hinaus haben alle kinetischen Parameter einen direkten Bezug zu den Enzymeigenschaften.

Insbesondere unter Verwendung mechanistischer Enzymkinetiken sowie des approximativen LinLog-Ansatzes wurden im Bereich des Zentralstoffwechsels sowie bestimmter Aminosäurebiosynthesewege eine Vielzahl metabolischer Netzwerkmodelle unterschiedlicher Organismen formuliert [RBTR97,VBRR99,CNRS⁺02,VSM⁺04,AEZ⁺06,MHOT06].

In vorangegangenen Arbeiten wurde auch ein Netzwerkmodell für *C. glutamicum* formuliert und mit experimentellen Daten validiert [Pet00,vL03]. Der Schwerpunkt lag in diesem Fall auf der mechanistischen Beschreibung der anaplerotischen Reaktionen dieses Organismus.

3.4 Genetische Netzwerkmodellierung

In den letzten Jahren wurden große Fortschritte im Bereich der Genomsequenzierung und -annotation gemacht, jedoch ist das Wissen über die Funktion der Gene sowie deren Regulation noch sehr gering. Dies betrifft auch das mechanistische Verständnis von Prozessen wie Transkription und Translation. Demzufolge gibt es im Bereich der Modellierung von Transkriptom- und Proteom-Daten bisher keine Standardvorgehensweise, wie dies bei der Auswertung von Metabolom-Daten auf Grundlage metabolischer Netzwerke der Fall ist.

Im Folgenden soll daher ein kurzer Überblick der am häufigsten verwendeten Ansätze zur Beschreibung genetischer Netzwerke unter dem gezielten Aspekt der Messdatenauswertung erfolgen. Für genauere Informationen in Bezug auf die mathematischen Details jeder Methode sowie das bei ihrer Anwendung resultierende komplexe Systemverhalten sei an dieser Stelle auf die Literatur verwiesen [SBB00b,dJ02,TCN03].

3.4.1 Genomweite qualitative Modellierung

Qualitative Methoden erlauben eine effiziente Analyse großer regulatorischer Netzwerke, indem stark vereinfachende Annahmen im Hinblick auf die Struktur und Dynamik genetischer Regulationssysteme getroffen werden.

Boolsche Netzwerke

Eine der am stärksten untersuchten Modellansätze zur qualitativen Modellierung genetischer Netzwerke basiert auf der sogenannten „Boolschen-Methode“ [Kau69]. Dabei gelten die Annahmen, dass:

1. ein Gen nur die zwei Expressionszustände AN (1) oder AUS (0) besitzt und
2. die regulatorische Kontrolle der Genexpression durch Regeln der Boolschen Logik approximiert werden kann.

Eine allgemeine mathematische Beschreibung ist dann gegeben durch:

$$G_i(t+1) = b_i(\mathbf{G}(t)), \quad 1 \leq i \leq n \quad (3.22)$$

Hierbei kennzeichnen die Variablen $G_i(t)$ und $\mathbf{G}(t)$ die zeitabhängigen Zustände des i -ten Gens bzw. aller n Gene. Jede Variable besitzt den Zustand 0 oder 1, so dass

insgesamt 2^n verschiedene Netzwerkzustände möglich sind. Der Zustand von $G_i(t)$ zum Zeitpunkt $t + 1$ wird mit Hilfe der Booleschen-Funktion b_i aus den vorherigen Zuständen von k abhängigen Genen berechnet.

In Abbildung 3.6 ist ein einfaches Beispiel eines Booleschen-Netzwerks dargestellt. Die zugehörige Boolesche-Funktion für Gen 1 ist in diesem Fall dadurch gegeben, dass dieses Gen zu einem bestimmten Zeitpunkt nur dann AN ist, wenn es zum Zeitpunkt davor durch das Produkt von Gen 3 aktiviert wurde. Demgegenüber ist Gen 3 nur dann angeschaltet, wenn im vorherigen Zeitschritt die Gene 1 und 2 ausgeschaltet wurden.

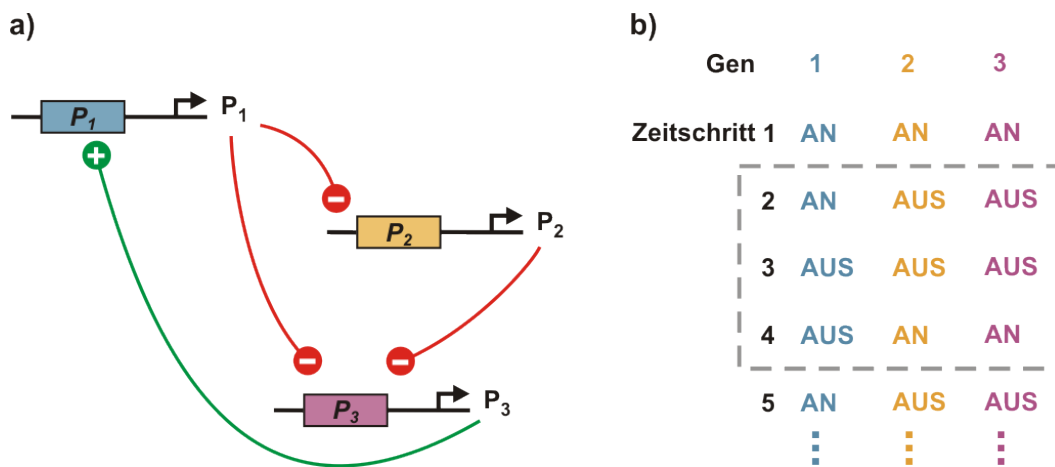


Abbildung 3.6: Beispiel eines Booleschen-Netzwerks, welches Gene als AN-AUS-Schalter betrachtet. a) Netzwerk aus drei Genen. Das Produkt von Gen 1 inhibiert die Transkription der Gene 2 und 3. Zusätzlich wird Gen 3 durch das Genprodukt P_2 inhibiert. Hingegen wird Gen 1 durch das Genprodukt P_3 aktiviert. b) Simulation des Netzwerks unter Anwendung der Booleschen-Logik. Aufgrund der regulatorischen Beziehungen stellt sich unabhängig vom Anfangszustand des Netzwerkes schnell ein Wiederholungszyklus von drei möglichen Zuständen ein (modifiziert aus [SBB00a]).

Auf der Basis solcher einfachen Regeln können somit große Zeitschritte simuliert und die Rechenzeit erheblich minimiert werden. Darüber hinaus sind Boolesche-Netzwerke interessant, da sie ähnliche Eigenschaften wie biologische Systeme besitzen, z.B. Selbstorganisation, Stabilität, Redundanz und Periodizität [SS96].

Für die Konstruktion genomweiter genetischer Netzwerke mit Hilfe der Booleschen-Methode werden in der Regel Zeitserien von Transkriptom-Daten benötigt, welche z.B. durch Beobachtung des Wachstumsverlaufs oder durch gezielte Störung des biologischen Systems durch Pulsexperimente gewonnen werden [WFM⁺98].

Eine andere Form der qualitativen Auswertung von Transkriptom-Daten basiert auf der Integration in genomweite Modelle zur Flussbilanzanalyse. Dabei werden meta-

bolische Reaktionen als inaktiv betrachtet, wenn keine Expression der betreffenden Gene stattgefunden hat, die für das jeweilige Enzym kodieren [AFN04, CXCK08].

Statistische Methoden

In der Literatur existiert mittlerweile auch ein breite Palette an statistischen Methoden, die zur Minimierung von Fehlern bei der Präprozessierung von Microarray-Daten oder der anschließenden Extraktion von Informationen über genetische Regulationsnetzwerke angewendet werden können. Dazu zählen beispielsweise die Bayes-Statistik [FLNP00, PREF01], Clusteranalyse-Verfahren [Azu02, BA03] sowie die Haupt- und Netzwerk-Komponentenanalyse [HMM⁺00, HMC⁺01, KYB⁺04].

3.4.2 Fokussierte quantitative Modellierung

Mit Hilfe quantitativer Methoden besteht die Möglichkeit, den gesamten Bereich des nichtlinearen Verhaltens von Genregulationsprozessen gründlich zu untersuchen. Aufgrund mangelnder Informationen über die Gesamtheit der dynamischen Wechselwirkungen innerhalb eines spezifischen genetischen Netzwerks werden dabei in der Regel einzelne Teilsysteme und -prozesse fokussiert modelliert.

Dynamische Systemansätze

Vergleichbar zu metabolischen Netzwerken werden auf der Basis von Differentialgleichungen die zeitlichen Änderungen von Zustandsgrößen mit Hilfe von Massenbilanzen modelliert:

$$\frac{dx_i}{dt} = g_i(\mathbf{x}), \quad 1 \leq i \leq n \quad (3.23)$$

Der Vektor $\mathbf{x} = [x_1, \dots, x_n]^T$ kennzeichnet dabei z.B. die Konzentrationen an mRNA, Proteinen oder Metaboliteffektoren. Die Funktion g_i ist für gewöhnlich nichtlinear und enthält kinetische Terme, welche die Regulation der Genexpression durch das Zusammenspiel von Regulatorproteinen und Effektormolekülen bzw. Degradationsprozessen beschreiben.

Auf Grundlage fokussierter dynamischer Modelle wurden in der Vergangenheit eine Reihe spezifischer Genregulationsysteme unterschiedlicher Organismen untersucht [XZD97, KTY98, SM01, YM03, WO03, SMR⁺04, BFK⁺06, TAKP07].

Neben gewöhnlichen Differentialgleichungen kommen vermehrt auch stochastische Ansätze zum Einsatz, welche die Prozesse auf Ebene einzelner Zellen beschreiben

bzw. den teilweise hohen Rauschanteil in den Expressiondaten berücksichtigen [Dre01, RZ06, LPC⁺06, WYH⁺08, ZFL09, SM09].

Die Simulation dynamischer Netzwerkmodelle erfordert in der Regel eine höhere Rechenkapazität, jedoch ist durch die Verwendung kontinuierlicher Variablen die physikalische Systembeschreibung in vielen Fällen akkurater.

Ein wesentlicher Nachteil bei der Verwendung detaillierter mechanistischer Kinetiken ergibt sich aus der hohen Anzahl unbekannter Modellparameter und der messtechnischen Unzugänglichkeit relevanter Zustandsgrößen.

Kapitel 4

Methodenentwicklung mit Modelica

Ein entscheidender Punkt bei der Formulierung von biochemischen Netzwerkmodellen ist die Auswahl einer geeigneten Beschreibungssprache zur Implementierung einzelner Modellkomponenten, einer Komponentenbibliothek des verknüpfenden Regelwerkes sowie der erzeugten Teilmodelle.

Innerhalb dieser Arbeit wurde die Modellierungssprache Modelica zur integrativen mechanistischen Modellierung von Multi-Omics-Daten auf der Grundlage eines vertikalen Modellierungskonzeptes genutzt (vgl. Kap. 8).

Da Modelica im Bereich der Biotechnologie bisher kaum eingesetzt wurde, war die Entwicklung von grundlegenden Methoden für die Modellierung biochemischer Netzwerke zunächst Voraussetzung und damit ein wichtiger Gegenstand dieser Arbeit.

4.1 Modellierungssprache Modelica

Die Anwendung einer objektorientierten Programmiersprache (z.B. C++, Java) zur Implementierung von Netzwerkmodellen lässt viele Freiheiten und bietet nahezu alle Möglichkeiten der Umsetzung spezifischer Zielstellungen. Voraussetzung dafür ist allerdings, dass praktisch alle Modellierungsbestandteile in Form separater Eigenlösungen erstellt werden müssen. Dies reicht von der numerischen Behandlung mathematischer Probleme, der Implementierung von Algorithmen, der graphischen Visualisierung von Objektstrukturen, der Repräsentation von Ergebnissen bis hin zur Realisierung einer Bibliotheksstruktur.

Demgegenüber steht die Verwendung einer domänenspezifischen Beschreibungssprache wie SBML („Systems Biology Markup Language“) zur mathematischen Mo-

dellbildung [HFS⁺03]. SBML basiert auf dem XML-Standard („Extensible Markup Language“) und es stehen mittlerweile eine ganze Reihe von Softwarewerkzeugen für die Formulierung, Prozessierung, Simulation und Analyse von SBML-Modellen zur Verfügung [FTM03, HSG⁺06]. Aufgrund der festen Struktur von SBML ergeben sich jedoch auch deutliche Einschränkungen bei der Umsetzung von neuen Modellierungskonzepten. Dies betrifft insbesondere die Formulierung von komplexen Differential-Algebraischen-Gleichungssystemen (DAE), welche im Fall von vertikalen Netzwerkmodellen eine Rolle spielen können.

Eine Alternative stellt die Modellierungssprache Modelica dar, welche die jeweiligen Vorzüge der obigen Ansätze kombiniert. Modelica ist in den Ingenieurwissenschaften des Maschinenbaus und der Elektrotechnik seit Jahren etabliert und erfreut sich zunehmender Verbreitung in der Industrie. Hingegen wurde Modelica im Bereich der Biotechnologie und insbesondere zur Modellierung biochemischer Netzwerke bisher kaum eingesetzt. Dies liegt vor allem an der fehlenden Verfügbarkeit entsprechender Komponentenbibliotheken.

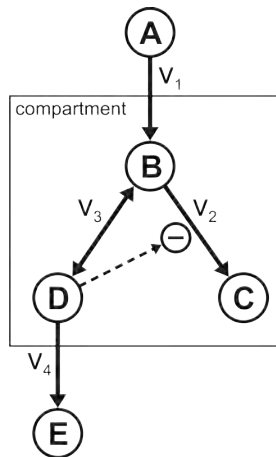
Die Entstehung von Modelica in den 90er Jahren war das Ergebnis einer wachsenden Nachfrage nach einer multidisziplinären Modellierungssprache. Da Mathematik die einzige Sprache ist, welche allen wissenschaftlichen Disziplinen zugrunde liegt, basiert Modelica auf den elementaren Bausteinen der differentiellen und algebraischen Gleichungen. Aus diesem Grund wird Modelica auch als „gleichungsbasierte“ Sprache bezeichnet und steht damit im Gegensatz zu SBML, welche auf einem „reaktionsbasierten“ Konzept beruht.

4.2 Modellierung biochemischer Netzwerke mit Modelica

Im Folgenden werden einige wesentliche Prinzipien von Modelica sowie entwickelter Methoden zur Modellierung biochemischer Netzwerke am Beispiel eines einfachen metabolischen Netzwerks erläutert (Abb. 4.1).

4.2.1 Gleichungsbasiertes Modellierungskonzept

In Modelica sind einzelne Gleichungen zu physikalisch sinnvollen Blöcken zusammengefasst, die dann wiederum zu größeren Modellen verknüpft werden können. Damit besteht die Möglichkeit der Formulierung anwendungsspezifischer Biblio-



$$v_1 = v_{\max,1} \cdot \frac{c_A}{K_{A,1} + c_A}$$

$$v_2 = v_{\max,2} \cdot \frac{c_B}{K_{B,2} + c_B} \cdot \frac{K_{I,D,2}}{K_{I,D,2} + c_D}$$

$$v_3 = \frac{v_{\max,3}^{\rightarrow} \cdot \frac{c_B}{K_{B,3}} - v_{\max,3}^{\leftarrow} \cdot \frac{c_D}{K_{D,3}}}{1 + \frac{c_B}{K_{B,3}} + \frac{c_D}{K_{D,3}}}$$

$$v_4 = k_4 \cdot c_D$$

Abbildung 4.1: Einfaches metabolisches Netzwerk mit mechanistischen Ausdrücken für die Beschreibung der Enzymkinetiken.

theiken, z.B. für mechanische Mehrkörpersysteme, elektrische Schaltkreise oder Anwendungen der Prozesskontrolle.

Prinzipiell ist Modelica mit einer objektorientierten Programmiersprache vergleichbar. Ein wesentlicher Unterschied besteht jedoch im eigentlichen Ziel der Kompilierung, welches kein ausführbares Computerprogramm sondern ausführbarer Simulationscode ist. Ebenso wie Programm-Quellcode ist ein gut strukturierter und dokumentierter Modelica-Code menschenlesbar, insbesondere wenn spezielle Editoren mit Syntaxhervorhebung zum Einsatz kommen (siehe nachfolgende Quellcodes).

Aufgrund der Tatsache, dass es in Modelica keine vordefinierten Strukturen gibt, in denen ein biochemisches Netzwerk formuliert sein muss, bleiben dem Modellierer viele Freiheitsgrade bei der Festlegung der Modellspezifikationen. Die grundlegendste Form einer solchen Spezifikation stellt hierbei der sogenannte „flat code“ dar, in welchem alle benötigten differentiellen und algebraischen Gleichungen sequenziell definiert sind.

Der in Abbildung 4.2 dargestellte Quellcode zeigt eine derartige Implementierung für das Beispielnetzwerk. Dieser Code enthält bereits alle notwendigen Informationen um ein ausführbares Simulationsprogramm zu erzeugen.

4.2.2 Aufbau von Komponenten und Bibliotheken

Obwohl sich die Spezifikation von Gleichungssystemen in obiger Form als sehr einfach erweist, ist dieser Ansatz für die Implementierung großer Modelle oder die

Erstellung wiederverwendbarer Komponenten nicht empfehlenswert. Insbesondere für die Anwendung zur modularen und hierarchischen Modellierung verfügt Modelica über leistungsfähige Konzepte der objektorientierten Programmierung. Dazu gehört beispielsweise die Zerlegung von Gleichungssystemen in unabhängige Teilsysteme, welche ihrerseits einzelne Modellkomponenten enthalten. Die jeweiligen Komponenten sind dabei über ein spezielles Schnittstellenobjekt namens **connector** verbunden.

```

model SimpleNetwork
  //Deklarationsteil
  //Metabolitkonzentrationen;
  Real A(start=10.0), B(start=0.1), ...;
  //Reaktionsraten
  Real v1, v2, ...;
  //Kinetische Parameter
  parameter Real vmax1=10.0, kmA1=0.05, ...;
equation
  //Massenbilanzen
  der(A)=-v1;
  der(B)=v1-v2-v3;
  der(C)=v2;
  der(D)=v3-v4;
  der(E)=v4;
  //Kinetiken
  v1=vmax1*A/(A+kmA1);
  v2=vmax2*B/(kmB2+B)*kiD2/(D+kiD2);
  v3=(vmax3f*B/kmB3-vmax3b*D/kmD3)/(1+B/kmB3+D/kmD3);
  v4=k4*D;
end SimpleNetwork;

```

Abbildung 4.2: „Flat code“-Spezifikation des Beispielnetzwerks aus Abbildung 4.1 in Modelica. Aus dem Modellkontext ersichtliche Aufzählungen sind durch Punkte abgekürzt.

Konnektoren stellen bidirektionale Informationskanäle dar, welche die Identifikation der Variablen aus verschiedenen Komponenten ermöglichen. Im laufenden Beispiel umfasst ein solcher Konnektor (hier mit *cp* bezeichnet) zwei Variablen, eine für Metabolitkonzentrationen und eine für Reaktionsraten (Abb. 4.3). Die Bedeutung des Schlüsselwortes **flow** wird im nachfolgenden Kapitel näher erläutert.

Auf Basis der Anwendung von Konnektoren besteht nun die Möglichkeit der Formulierung verschiedener Teilmodelle, z.B. für eine Michalis-Menten-Kinetik zur Beschreibung von Reaktionsraten (Abb. 4.4). In diesem Fall ist die Reaktion mit ihrem Substrat *S* und Produkt *P* über zwei Konnektoren verbunden.

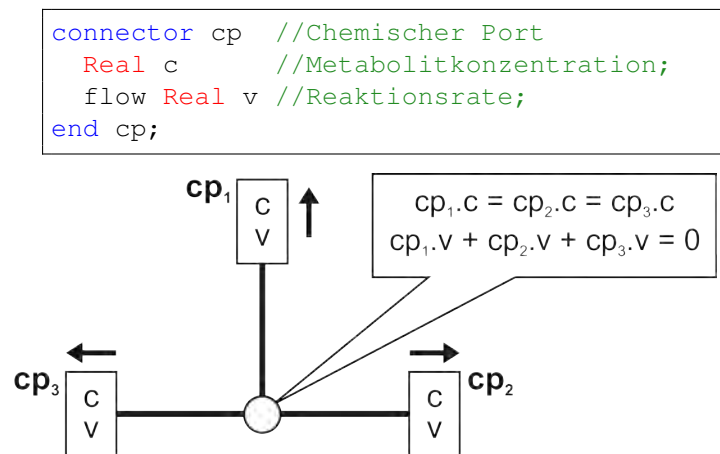


Abbildung 4.3: Definition von Konnektoren in Modelica. Die Verbindung von 3 Konnektoren bedingt gleichzeitig die Generierung der gezeigten Gleichungen. Die Konnektoren können dabei zu unabhängigen Komponenten (z.B. Metabolit, Reaktion, Effektor) gehören.

Die Bedeutung der beiden assoziierten Flussvariablen liegt in der Menge an Material, welches von S in Richtung P fließt. Natürlich stehen diese Flüsse über die Stöchiometrie der Reaktion miteinander in Beziehung, in diesem einfachen Fall wird 1 mol Substrat zu 1 mol Produkt umgesetzt. Das Minuszeichen ist notwendig, da durch die Standardkonvention von Modelica positive Flüsse immer von innen nach außen relativ zu einer Komponente gerichtet sind. Durch Einführung zusätzlicher stöchiometrischer Parameter können auch ungeradzahlige Stöchiometrien berücksichtigt werden.

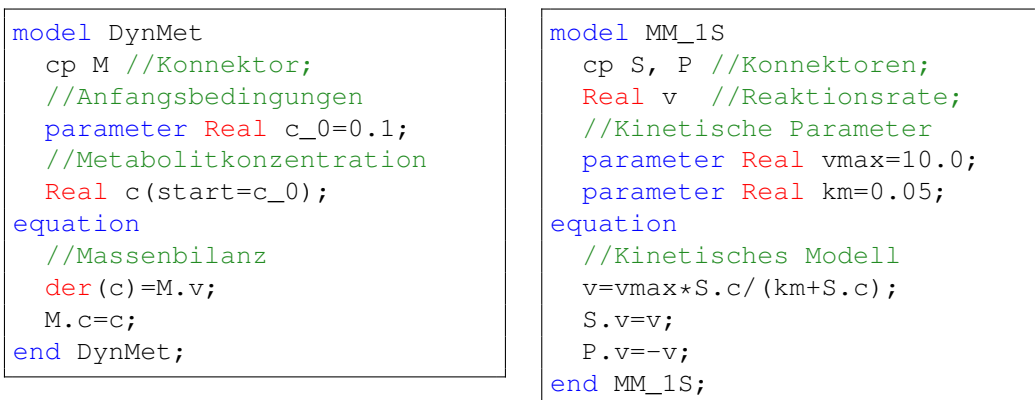


Abbildung 4.4: Implementierung von Modelica-Komponenten am Beispiel von Metaboliten und Reaktionen als grundlegende Teile eines metabolischen Netzwerks.

Die Formulierung anderer Reaktionstypen, z.B. mit mehreren Substraten und Pro-

dukten, kann auf ähnliche Weise erfolgen. Auch die Einbindung von Effektorwirkungen von Inhibitoren und Aktivatoren ist möglich, indem einfach der zugehörige Flusswert eines Effektor-Konnektors (`cp_I`) auf Null gesetzt wird. Dies bedeutet gleichzeitig, dass ein Effektor eine Reaktionsrate beeinflussen kann, jedoch nicht an der eigentlichen enzymatischen Umsetzung teilnimmt (Abb. 4.5).

```

model MM_1S1I
  //Konnektoren
  cp S, P;
  cp_I I;
  Real v //Reaktionsrate;
  //Kinetische Parameter
  parameter Real vmax=0.6;
  parameter Real km=1.0;
  parameter Real ki=0.4;
equation
  //Kinetisches Modell
  v=vmax*S.c/(km+S.c)*ki/(I.c+ki);
  S.v=v;
  P.v=-v;
  I.v=0; //keine Umsetzung des Inhibitors
end MM_1S1I;

```

Abbildung 4.5: Modelica-Komponente zur Modellierung einer inhibierten Reaktion. Eine separate Definition des Konnektors für den Inhibitor (`cp_I`) ist insbesondere für die grafische Modellierung mit Modelica-Komponenten von Vorteil (vgl. Abb. 4.7).

Im Gegensatz zu Reaktionsmodellen enthält die Spezifikation für einen Metabolitpool nur einen Konnektor (vgl. Abb. 4.4). Dieser Konnektor ist gleichsam bedeutend mit einer Massenbilanz aller Flüsse um diesen Metabolitknoten. Im nachfolgenden Kapitel wird gezeigt, wie diese Flussbilanz implizit der Variablen zugewiesen wird. Damit hängt die Änderung einer Poolgröße von der Flussvariable ab, welche ein positives oder negatives Vorzeichen besitzen kann.

Diese einfache Beschreibung funktioniert aufgrund des Modelica-Mechanismus zur Verbindung von Komponenten. Zwei Komponenten werden nicht direkt durch einen Konnektor verbunden, sondern zwischen diesen gibt es einen Kopplungsknoten. Mit Hilfe der **connect**-Anweisung ist es möglich, beliebig viele Konnektoren mit einem Kopplungsknoten zu verbinden (vgl. Abb. 4.3).

Nach der einmaligen Spezifikation einer Komponente kann diese im Systemmodell wiederholt verwendet werden. Der Modelica-Sprachformalismus gewährleistet dabei automatisch die eindeutige Identifikation der Größen jedes Teilmodells innerhalb eines größeren Modells. Zu diesem Zweck wird die in der objektorientierten

Programmierung bekannte „Punktnotation“ zur Verbindung von Objekten genutzt. Eine entsprechende Umsetzung für das Beispielnetzwerk ist in Abbildung 4.6 dargestellt.

```

model SimpleNetwork
  //Metabolite
  DynMet A(c_0=10.0);
  DynMet B, C, D, E;
  //Kinetische Modelle
  MM_1S v1(vmax=10.0, km=0.05);
  MM_1S1I v2(vmax=0.6, km=1.0, ki=0.4);
  MM_1S1P v3(vmaxf=0.3, vmaxb=0.4, kmS=0.4, kmP=0.5);
  Trp_1S v4(k=0.5);
equation
  connect(A.M, v1.S);
  connect(v1.P, B.M);
  connect(B.M, v2.S);
  connect(v2.P, C.M);
  connect(D.M, v2.I);
  connect(B.M, v3.S);
  connect(v3.P, D.M);
  connect(D.M, v4.S);
  connect(v4.P, E.M);
end SimpleNetwork;

```

Abbildung 4.6: Spezifikation des Beispielnetzwerks aus Abbildung 4.1 in Modelica auf Basis vordefinierter Modellkomponenten. DynMet: Dynamischer Metabolitknoten, MM: Michaelis-Menten-Kinetik, Trp: Transport-Kinetik (vgl. Abb. 4.4 und 4.5).

Für die Organisation der verschiedenen Komponenten können in Modelica Bibliotheken angelegt werden. Dies ist besonders bei der Nutzung grafischer Modelleditoren von Vorteil, da hier ein leichter Zugang zu einzelnen grafischen Modellblöcken aus verschiedenen domainspezifischen Bibliotheken ermöglicht wird (Abb. 4.7).

4.2.3 Automatische Codegenerierung

Was bedeutet es nun, drei verschiedene Konnektoren (*cp*) mit assoziierten Konzentrations- (*c*) und Flussvariablen (*v*) an einem einzigen Knoten zu verbinden? Die Antwort hängt vom Variablentyp in der Konnektorspezifikation ab (vgl. Abb. 4.3). Fehlt hier das Schlüsselwort **flow**, so wird eine Potentialvariable definiert. Dies bedeutet, dass alle zugehörigen Variablen von verbundenen Konnektoren gleich groß sind. Im betrachteten Fall erfolgt deshalb am Kopplungsknoten die Generierung folgender

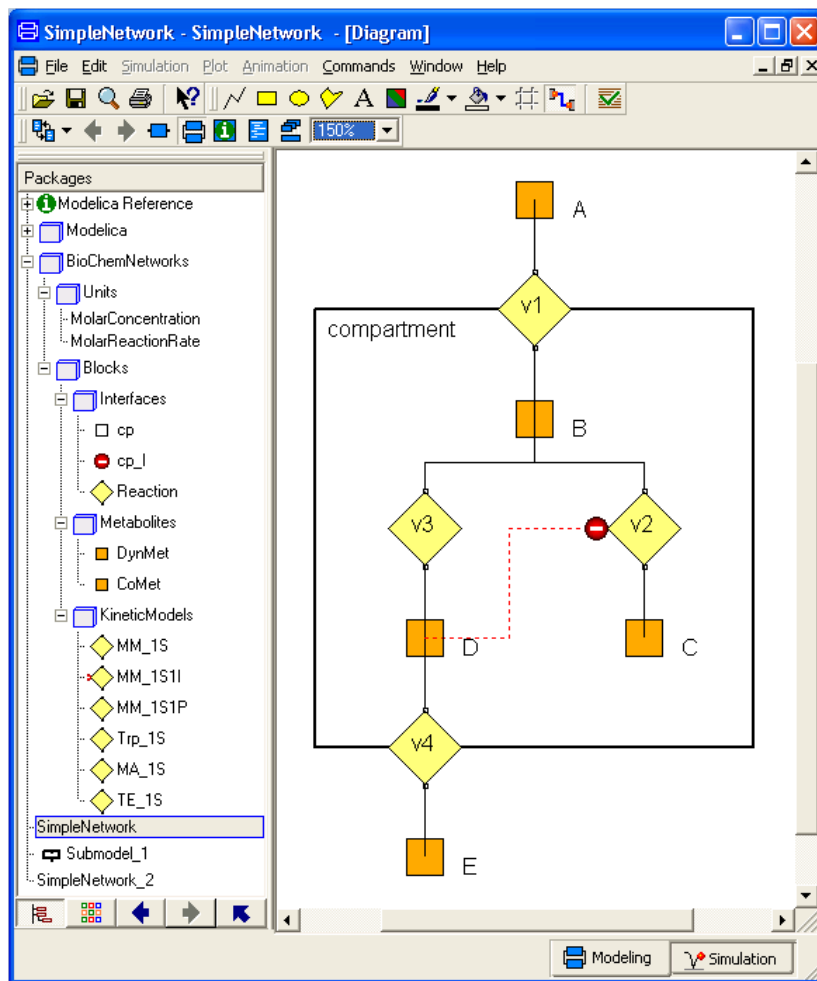


Abbildung 4.7: Grafische Benutzeroberfläche von Dymola. Modelle können durch bereits vorhandene oder selbstdefinierte Komponentenbibliotheken aufgebaut und simuliert werden.

Gleichheitsbedingungen:

$$cp_1.c = cp_2.c = cp_2.c \quad (4.1)$$

Hingegen wird bei Verwendung des Schlüsselwortes **flow** eine Variable als Flussgröße definiert und eine Kirchhoffsche Massenbilanz generiert:

$$cp_1.v + cp_2.v + cp_3.v = 0 \quad (4.2)$$

Das resultierende Gleichungssystem für das Beispielnetzwerk, welches aus dem Vernetzungsschema der Komponenten generiert wird, besitzt folgende Form:

$$\begin{aligned}
 \dot{A}_{M_c} &= A_{M_v} & v_{3v} &= \frac{v_{3f} \cdot \frac{v_{3S_c}}{v_{3k_{mS}}} - v_{3b} \cdot \frac{v_{3P_c}}{v_{3k_{mP}}}}{1 + \frac{v_{3S_c}}{v_{3k_{mS}}} + \frac{v_{3P_c}}{v_{3k_{mP}}}} \\
 A_{M_c} &= v_{1S_c} & v_{3S_v} &= v_{3v} \\
 0 &= A_{M_v} + v_{1S_v} & v_{3P_v} &= -v_{3v} \\
 v_{1v} &= v_{1v_{max}} \cdot \frac{v_{1S_c}}{v_{1S_c} + v_{1k_m}} & \dot{D}_{M_c} &= D_{M_v} \\
 v_{1S_v} &= v_{1v} & D_{M_c} &= v_{3P_c} + v_{4S_c} + v_{2I_c} \\
 v_{1P_v} &= -v_{1v} & 0 &= D_{M_v} + v_{3P_v} + v_{4S_v} + v_{2I_v} \\
 \dot{B}_{M_c} &= B_{M_v} & v_{4v} &= v_{4k} \cdot v_{4S_c} \\
 B_{M_c} &= v_{1P_c} + v_{2S_c} + v_{3S_c} & v_{4S_v} &= v_{4v} \\
 0 &= B_{M_v} + v_{1P_v} + v_{2S_v} + v_{3S_v} & v_{4P_v} &= -v_{4v} \\
 v_{2v} &= v_{2v_{max}} \cdot \frac{v_{2S_c}}{v_{2S_c} + v_{2k_m}} \cdot \frac{v_{2k_i}}{v_{2I_c} + v_{2k_i}} & \dot{E}_{M_c} &= \dot{E}_{M_v} \\
 v_{2S_v} &= v_{2v} & E_{M_c} &= v_{4P_c} \\
 v_{2P_v} &= -v_{2v} & 0 &= \dot{E}_{M_v} + v_{4P_v} \\
 0 &= v_{2I_v} \\
 \dot{C}_{M_c} &= C_{M_v} \\
 C_{M_c} &= v_{2P_c} \\
 0 &= C_{M_v} + v_{2P_v}
 \end{aligned}$$

Es ist zu erkennen, dass die Identifizierung der Potentialvariablen und die Anwendung der Kirchhoffschen Gesetze zu einer Vielzahl überflüssiger Variablen führt. Dies ist jedoch kein grundsätzliches Problem, da fortgeschrittene Modelica-Compiler über Methoden der Modellvereinfachung durch Variablen-Eliminierung und automatische Formelmanipulation verfügen [Elm93, MGL96]. Darüber hinaus werden auch nichtlineare Beziehungen durch Methoden der Computeralgebra vereinfacht und komplexe Modelle mittels Graphenalgorithmen in einfacher lösbarer Teilsysteme zerlegt [Mur87, Cel06].

4.2.4 Nebenbedingungen und Konsistenzprüfung

Wie bereits erläutert ist Modelica eine gleichungsbasierte Modellierungssprache und demzufolge ist das Hinzufügen jeglicher Art von Gleichung problemlos möglich. Insbesondere müssen keine neuen Konzepte eingeführt werden, um differentielle und algebraische Gleichungen miteinander zu kombinieren. Nach der Formulierung eines Modells auf Basis beider Gleichungstypen entsteht ein DAE-System und genau für diesen Typ von mathematischer Modellstruktur wurde Modelica entwi-

ckelt.

Wird beispielsweise für den dritten Reaktionsschritt $B \rightleftharpoons D$ des Beispielnetzwerks ein thermodynamisches Gleichgewicht angenommen, muss in diesem Fall keine Reaktionsrate definiert werden, sondern es kann folgende algebraische Gleichung zwischen den Pools B und D abgeleitet werden:

$$k^{\rightarrow} \cdot B = k^{\leftarrow} \cdot D \quad (4.3)$$

Eine entsprechende Modelica-Komponente für die Modellierung thermodynamischen Gleichgewichts zeigt Abbildung 4.8. Diese Komponente kann jedes andere reaktionskinetische Modell mit der gleichen Anzahl an Substraten und Produkten ersetzen.

```

model TE_1S
  cp S, P //Konnektoren;
  Real v //Reaktionsrate;
  //Kinetische Parameter
  parameter Real kf=1.0;
  parameter Real kb=0.5;
equation
  //Thermodynamisches Gleichgewicht
  kf*S.c=kb*P.c;
  S.v=v;
  P.v=-v;
end TE_1S;

```

Abbildung 4.8: Formulierung von Modellen mit Nebenbedingungen in Modelica am Beispiel einer thermodynamischen Gleichgewichtsbeziehung.

Die Formulierung zusätzlicher Nebenbedingungen in einem Modell kann im resultierenden Gleichungssystem zu Inkonsistenzen führen. Insbesondere Über- oder Unterbestimmtheiten müssen durch den Algorithmus der Modellgenerierung erkannt werden [LH01]. Weitaus problematischer ist allerdings das sogenannte Indexproblem von DAEs zu handhaben [MS93, Pan98]. Zur Veranschaulichung dieses Problems sei erneut das Beispielnetzwerk (vgl. Abb. 4.1) betrachtet, nun wird jedoch die Michaelis-Menten-Kinetik der Reaktion v_1 durch eine Gleichgewichtsbeziehung zwischen A und B ersetzt (vgl. Gl. 4.3):

$$\dot{A} = -v_x, \quad \dot{B} = v_x - v_2 - v_3, \quad v_x = f(A, B) \quad (4.4)$$

Offensichtlich liefert die Gleichgewichtsbeziehung in diesem Modell nicht direkt die

Information, wie nach der unbekanntem Reaktionsrate v_x aufgelöst werden kann. Die fehlende Information kann jedoch vom ursprünglichen Modell durch zusätzliche Gleichungsdifferentiation gewonnen werden. Eine Ableitung der Gleichgewichtsbeziehung in Gl. 4.3 nach der Zeit liefert folgende neue Differentialgleichung:

$$k^{\rightarrow} \cdot \dot{A} = k^{\leftarrow} \cdot \dot{B} \Leftrightarrow v_x = -\frac{k^{\leftarrow}}{k^{\rightarrow}} \cdot \dot{B} \quad (4.5)$$

Damit kann das Problem gelöst werden, da Gl. 4.5 eine direkte Beziehung zwischen der unbekanntem Variable v_x und den Zustandsvariablen A und B herstellt.

Seit den 80er Jahren wurde die Theorie der DAEs ständig weiterentwickelt und sowohl die Gemeinsamkeiten als auch die Unterschiede zu gewöhnlichen Differentialgleichungen (ODE) ohne algebraische Anteile sind mittlerweile gut verstanden [Cel06]. Auf der Grundlage computeralgebraischer Methoden ist Modelica in der Lage, die oben erwähnten Inkonsistenzen und Indexprobleme zu erkennen und, falls möglich, zu lösen. Desweiteren sind heutzutage verschiedene numerische DAE-Löser verfügbar, womit selbst komplexe Systeme in Echtzeit simuliert werden können [AP98].

4.2.5 Hierarchische Modellierung und Klassenabstraktion

Ein wesentliches Anliegen der Systembiologie ist die Entwicklung von strukturierten Beschreibungen komplexer biologischer Systeme unter Nutzung hierarchisch organisierter Komponenten. Bekannte Hierarchien sind zum Beispiel auf genetischer Ebene: *Gen* \rightarrow *Operon* \rightarrow *Regulon* bzw. auf Ebene des Metabolom: *Reaktion* \rightarrow *Reaktionspfad* \rightarrow *Stoffwechselabschnitt*.

Bereits aus den vorangegangenen Abschnitten wird deutlich, dass Modelica insbesondere für den Aufbau modularer Systeme geeignet ist. Ein einfaches Beispiel ist in Abbildung 4.9 dargestellt, wo ein Teil der Reaktionen des Beispielnetzwerks durch Integration in ein Teilmodell zu einem Reaktionspfad zusammengefasst wurde. Auf diese Weise kann jedes Teilmodell zu größeren Modellen zusammengesetzt werden, welche wiederum Grundbausteine für die nächste hierarchische Ebene bilden [KG01].

Darüber hinaus wurde in Modelica der hierarchische Ansatz durch Konzepte der objektorientierten Modellierung erweitert. Vor allem durch Nutzung des Prinzips der Klassenabstraktion können zunächst allgemeingültige Modellklassen erstellt und anschließend als ausformulierte Komponenten im Modell integriert werden.

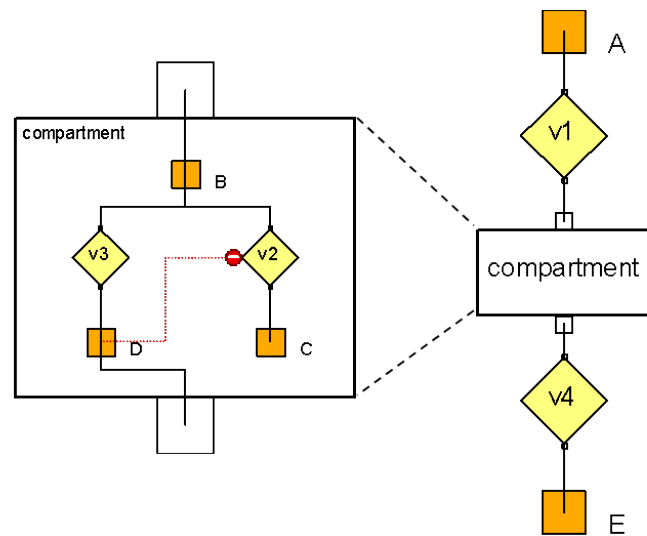


Abbildung 4.9: Hierarchische Modellierung in Modelica. Alle Reaktionen innerhalb des Kompartiments wurden in ein Teilmodell integriert, welches mit den übrigen Netzwerkeilen verbunden ist.

Beispielhaft sei hier ein Reaktionsschritt betrachtet, der noch keine Gesetzmäßigkeit für die Reaktionskinetik enthält (Abb. 4.10). Innerhalb dieser abstrakten Klasse sind lediglich die Schnittstellen in Form von Konnektoren spezifiziert. Mit Hilfe des Schlüsselwortes **extends** können dann die Eigenschaften dieser Klasse zur Formulierung konkreter Reaktionsmodelle übernommen werden. Ein wesentlicher Vorteil der Klassenabstraktion ist somit der schnelle Austausch von Reaktionsgesetzen in einem Netzwerkmodell, ohne dessen Verbindungen auflösen zu müssen.

4.3 Entwicklung einer Komponentenbibliothek

4.3.1 Entwicklungsumgebung Dymola

Die Spezifikationen von Modelica werden von dem öffentlichen Konsortium „Modelica Association“ (www.modelica.org) in Form von Open Source herausgegeben und sind Gegenstand fortlaufender Entwicklungen. Aufgrund der Komplexität der Sprache und des großen Umfangs verschiedener Methoden und Konzepte, auf denen Modelica basiert, sind relativ wenig Compiler und Entwicklungsumgebungen verfügbar. Die gegenwärtig frei nutzbaren Compiler wie OpenModelica sind deutlich ineffizienter im Vergleich zu kommerziellen Entwicklungsumgebungen wie Dymola (Dynasim AB) oder MathModelica (MathCore Engineering AB), die auch für


```

partial model Reaction
  cp S, P //Konnektoren;
  Real v //Reaktionsrate;
equation
  S.v=v;
  P.v=-v;
end Reaction;

```

```

model MM_1S
  extends Reaction;
  parameter Real kf=1.0;
  parameter Real kmS=0.5;
equation
  //Michaelis-Menten
  v=kf*S.c/(kmS+S.c);
end MM_1S;

```

```

model MA_1S
  extends Reaction;
  parameter Real kf=1.0;
  parameter Real kb=0.5;
equation
  //Massenwirkung
  v=kf*(S.c-P.c/kb);
end MA_1S;

```

```

model TE_1S
  extends Reaction;
  parameter Real kf=1.0;
  parameter Real kb=0.5;
equation
  //Thermodynamisches GGW
  kf*S.c=kb*P.c;
end TE_1S;

```

Abbildung 4.10: Konzept der Klassenabstraktion in Modelica. Durch die Definition einer abstrakten Reaktionsklasse können verschiedene kinetische Modelle durch Ergänzung des spezifischen Reaktionsmechanismus abgeleitet werden.

den Aufbau, Simulation und Analyse großer Modelle geeignet sind.

Innerhalb dieser Arbeit wurde Dymola genutzt, da hier eine direkte Anbindung zu MATLAB (Mathworks) möglich ist. Auf diesen Fakt wird später näher eingegangen. Dymola verfügt über eine Benutzeroberfläche mit einem grafischen Modelleditor, mit deren Hilfe einzelne Komponenten aus Bibliotheken in Form grafischer Blöcke zu einem Modell verknüpft werden können (vgl. Abb. 4.7).

Im Gegensatz zu bestimmten Fachbereichen innerhalb des Maschinenbaus und der Elektrotechnik, in denen Modelica seit Jahren ein breites Anwendungsfeld hat, sind derzeit noch keine Bibliotheken für den Bereich der Biotechnologie verfügbar. Allerdings wurden bereits erste Konzepte für den Aufbau einer solchen Bibliothek für die Modellierung biologischer und biochemischer Systeme entwickelt [NF05].

Im Rahmen dieser Arbeit wurden einige dieser grundlegenden Ansätze aufgegriffen und eine entsprechende Komponentenbibliothek für Dymola entwickelt.

4.3.2 Struktur der Komponentenbibliothek

In Abbildung 4.11 ist ein Teil der erstellten Komponenten dargestellt, die für den grafischen Aufbau biochemischer Netzwerke in Dymola genutzt werden können. Eine grundlegende Einteilung aller Komponenten erfolgt hierbei im Hinblick auf die verschiedenen Omics-Ebenen (Genom, Proteom, Metabolom).

- **Units:** Für die Deklaration von Variablen und Parametern sind eine Reihe physikalischer Typendefinitionen notwendig. Die dabei verwendeten Typen basieren auf den SI-Einheiten mol (Stoffmenge) und s (Zeit) sowie der Nicht-SI-Einheit l (Volumen).
- **Interfaces:** Hier sind die zur Verbindung von Komponenten notwendigen Schnittstellen definiert. Diese umfassen Konnektoren für die Reaktanden und Effektoren metabolischer Reaktionen sowie für die Enzyme als Produkte der Genexpression.
- **Nodes:** Als Knoten dienen verschiedene Spezifikationen für Metabolite (stationäre Massenbilanz, stöchiometrische Beziehung, Cometabolit) sowie für mRNA- und Enzym pools (Genexpression).
- **Fluxes:** Damit können Reaktionen ins Netzwerk eingebunden werden, denen kein enzymkinetisches Modell zugrunde liegt (z.B. Biomasseabflüsse).
- **Kinetics:** Mit Hilfe abstrakter Klassen werden spezifische enzymkinetische Modellkomponenten formuliert, die je nach Mechanismus eine unterschiedliche Anzahl an Schnittstellen für Reaktanden, Effektoren und Enzymen beinhalten können.

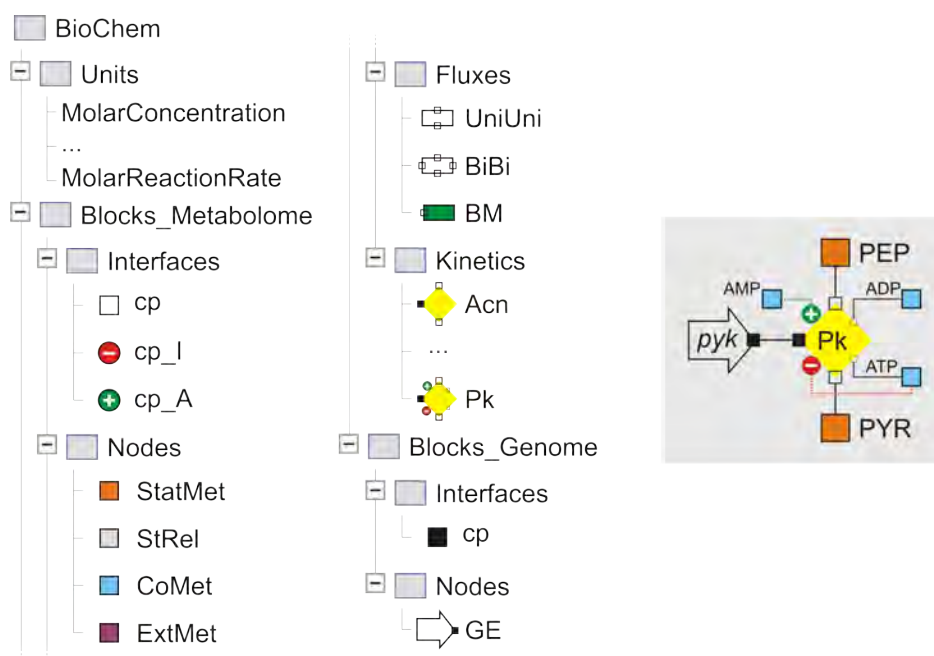


Abbildung 4.11: Modelica-Bibliothek zur Modellierung biochemischer Netzwerke.

4.4 Methoden zur Simulation und Systemanalyse

Nach der Kompilierung von Modellen in Dymola steht ausführbarer Code zur Verfügung, mit dem das zugrunde liegende Netzwerk frei parametrisierbar simuliert werden kann. Auf dieser Grundlage können mit Hilfe externer Programme Simulationsstudien durchgeführt sowie Methoden der Systemanalyse, wie z.B. Parameterschätzungen oder Sensitivitätsanalysen, angewendet werden.

Innerhalb dieser Arbeit wurde die Entwicklungsumgebung MATLAB (Mathworks, R2008b) genutzt, um die erstellten Modelica-Modelle im Zuge von Anpassungen an experimentelle Daten zu validieren und anschließend auf Basis weiterführender Simulationsstudien näher zu charakterisieren. Eine Entwicklung der dafür notwendigen m-Files war ebenfalls Gegenstand dieser Arbeit.

Darüber hinaus wurde für die Durchführung von Sensitivitätsanalysen auf der Basis von Modelica-Modellen ein Werkzeug namens ADModelica verwendet, welches von Attyah Elsheikh am Institut für Biotechnologie 2 im Rahmen seiner Dissertation entwickelt wurde [ENW08]. Mit Hilfe von ADModelica werden unter Verwendung einer effizienten Methode zur automatischen Differentiation von Modelica-Code maschinengenaue partielle Ableitungen der Zustandsvariablen nach allen Modellparametern erzeugt. Eine nähere Erläuterung zu dieser Methode ist im Anhang A.2.1 zu finden.

Insgesamt steht damit ein geschlossener Workflow zur Verfügung, mit dessen Hilfe biochemische Netzwerke in Modelica modelliert und analysiert werden können.

Teil II

Gewinnung quantitativer Fluxom-Daten

Kapitel 5

¹³C-Stoffflussanalyse eines Lysinproduzenten

Dieses Kapitel beschäftigt sich mit der Untersuchung zweier verschiedener Methoden zur Gewinnung quantitativer Fluxom-Daten auf der Basis von ¹³C-Markierungsexperimenten. Am Beispiel dieser Omics-Technologie wird hierbei die Qualität der erhaltenen Nutzdaten in Form intra- und extazellulärer Stoffflüsse unter dem Aspekt der weiteren Verwendung zur integrativen Modellierung kritisch analysiert. Dabei wird erstmals eine direkte Gegenüberstellung der isotopisch stationären und instationären ¹³C-MFA durchgeführt.

5.1 Kriterien isotopisch stationärer und instationärer ¹³C-MFA

Prinzipiell gibt es eine Vielzahl von Kriterien, die zum Vergleich beider methodischer Ansätze herangezogen werden können. Einige davon sind:

- **Anwendbarkeit:** Aufgrund der allgemeinen Notwendigkeit metabolischer Stationarität für die MFA sind beide Methoden auf experimentelle Zustände beschränkt, in denen alle intrazellulären Stoffflüsse näherungsweise konstant sind, d.h. kontinuierliche Prozessbedingungen oder exponentielle Wachstumsphase von Zellen, welche im Batch-/Fed-Batch-Modus unter nicht-substratlimitierten Bedingungen kultiviert werden [WN05]. Wird darüber hinaus ein ausgeglichener Markierungszustand angenommen (vgl. Gl. 3.13), muss das Markierungsexperiment solange durchgeführt werden bis mindestens nähe-

rungsweise isotopische Stationarität in den Intermediatpools erreicht ist. Für den Fall, dass Markierungsinformationen von hydrolysierten Proteinen zur Flussschätzung genutzt werden, betrifft dies auch die Biomassekomponenten [Szy95].

- **Bioreaktordesign:** Mit dem Ziel der Kostenminimierung für die isotopische Markierung sind Bioreaktoren mit kleinem Arbeitsvolumen generell zu bevorzugen [EMSD⁺03]. Für die Durchführung eines isotopisch instationären Experimentes (vgl. Gl. 3.14) wird darüber hinaus eine schnelle Probenahme benötigt, die eine Probengenerierung im Sub-Sekundenbereich ermöglicht und gleichzeitig eine sofortige Inaktivierung des Zellstoffwechsels garantiert [SFL09].
- **Messtechnik:** Prinzipiell können Markierungsinformationen durch Messung proteinogener Aminosäuren mittels NMR [MdGW⁺96, Szy98] und MS [FZS04] oder intrazellulärer Metabolite mittels MS [vWvDR⁺05, HSG⁺07, AKS07a] gewonnen werden. Bei der Anwendung des isotopisch instationären Ansatzes werden insbesondere Messdaten von gering konzentrierten Intermediaten (einschließlich Markierungsanteile und Poolgrößen, vgl. Gl. 3.14) benötigt, weshalb die Nutzung der deutlich sensitiveren LC-MS-Methoden eine wesentliche Voraussetzung ist [WN05].
- **Rechenaufwand:** Da die numerische Behandlung beider Ansätze verschieden ist, d.h. Lösung eines algebraischen (Gl. 3.13) bzw. differentiellen Gleichungssystems (Gl. 3.14), hängt der Rechenaufwand für die Flussschätzung und statistische Analyse stark vom untersuchten System ab, ist jedoch im letzteren Fall bedeutend höher [NWW06].
- **Ergebnis:** Bisher konnte gezeigt werden, dass in Abhängigkeit des Verhältnisses zwischen Messdaten und Netzwerkkomplexität jede Methode mehr oder weniger in der Lage ist, statistisch signifikante intrazelluläre Flussdaten zu produzieren. Darüber hinaus ermöglicht die isotopisch instationäre Methode auch Zugang zu den nicht-messbaren Metabolitkonzentrationen, welche durch das jeweilige metabolische Netzwerkmodell abgebildet werden [NGL⁺07].

Mit der Zielsetzung des Methodenvergleichs sind u.a. für den ersten und dritten Punkt verschiedene Kombinationen von Kultivierungsbedingungen (einschließlich Stammvariabilität) und Messkonfigurationen denkbar.

5.2 Experimentelle Rahmenbedingungen

Da bisher noch kein direkter Vergleich zwischen beiden Ansätzen zur ^{13}C -MFA durchgeführt worden ist, lag die Konzentration im Rahmen dieser Arbeit zunächst auf dem Ergebnisvergleich beider Methoden unter Maßgabe eines definierten experimentellen Setups und einheitlicher Messkonfigurationen. Aus diesem Grund wurden folgende allgemeine Rahmenbedingungen gewählt:

1. **Organismus:** *C. glutamicum* ist ein weit verbreiteter Organismus für die industrielle L-Lysin-Produktion [PMBM03]. Der Zentralstoffwechsel von *C. glutamicum* einschließlich des Lysin-Biosyntheseweges ist gut erforscht und war bereits das Ziel einer Reihe von Untersuchungen mit Anwendung der klassischen ^{13}C -MFA [WdGSE00, PMdG⁺01, WH02, DEMM⁺03, KSK⁺04, SMK⁺06]. Im Gegensatz zu *E. coli* wurde der instationäre Modellierungsansatz auf Basis von isotopisch transienten Daten im Sub-Sekundenbereich bisher nicht für diesen Organismus angewendet [NGL⁺07].
2. **Kultivierungsbedingungen:** Die vorherrschenden industriellen Prozessbedingungen sind Batch und Fed-Batch. Mittlerweile wurde in verschiedenen Untersuchungen gezeigt, dass eine ^{13}C -MFA auch unter diesen Bedingungen durchgeführt werden kann, wenn bestimmte Randbedingungen erfüllt sind (vgl. Abschn. 5.1). Um beide Methoden direkt miteinander vergleichen zu können, wurde ein isotopisch instationäres Fed-Batch-Experiment durchgeführt, bei dem gleichzeitig die Markierungsanreicherungen in den metabolischen Intermediaten bis in den näherungsweise stationären Markierungszustand verfolgt wurden.
3. **Messdaten:** Für die Messung der Massenisotopomere von Intermediaten des Zentralstoffwechsels von *C. glutamicum* kamen hochempfindliche LC-MS-Analytikmethoden zum Einsatz. Unter der Voraussetzung, dass im Verlauf des Experiments alle gemessenen Intermediate zumindest einen isotopisch quasi-stationären Zustand erreichen, können diese Daten auch für den klassischen ^{13}C -MFA-Ansatz verwendet werden. Somit ist keine fehleranfällige Datenextrapolation durch die sogenannte „Wash-Out-Korrektur“ notwendig, wie dies zum Beispiel bei der Verwendung von Messinformationen aus hydrolysierten Biomasseproben der Fall wäre [WN05].

Auf die konkrete Umsetzung der oben genannten Punkte wird nun im nachfolgenden Abschnitt eingegangen.

5.3 Bioprozess- und Analysemethoden

5.3.1 Stamm und Kultivierungsbedingungen

Untersucht wurde der L-Lysin produzierende Stamm DM1730 $\Delta pck pyc(P458S) hom(V59A) lysC(T311I)$, welcher durch rationale Stammentwicklung auf der Basis von *C. glutamicum* ATCC13032 (Biotin-auxotropher Wildtyp) beim Industriepartner Evonik Degussa GmbH erzeugt wurde. Der Genotyp dieses Stammes ist durch die Deletion des Gens für das Enzym Phosphoenolpyruvat-Carboxykinase (*pck*) sowie durch Punktmutationen in den Genen für die Enzyme Pyruvat-Carboxylase (*pyc*), Aspartat-Kinase (*lysC*) und Homoserin-Dehydrogenase (*hom*) gekennzeichnet.

Kryokulturen wurden bei -80°C in LB (lysogeny broth) Medium mit 10% ($v v^{-1}$) Glycerol gelagert. Die Medien der Vor- und Hauptkulturen basierten auf einem modifiziertem CGXII-Medium [WWR⁺08]. Vorkulturen zur 10% ($v v^{-1}$) Animpfung der Hauptkultur wurden über Nacht bei 30°C kultiviert. Nach dem Animpfen erfolgte die Kultivierung in einem 2 l-Bioreaktor [EMSD⁺03] mit einem Arbeitsvolumen von 1 l beginnend mit einer Batch-Phase auf 10.5 g l^{-1} Glucose. Die Temperatur wurde auf 30°C geregelt und der pH-Wert durch Zugabe von 25% Ammonium auf 7.0 eingestellt. Die Belüftungsrate wurde konstant auf 1 vvm gehalten und die Gelöstsauerstoff-Konzentration durch Anpassung der Rührerdrehzahl auf 20% Sättigung geregelt. Die Abgasanalyse für O_2 und CO_2 erfolgte durch einen Gasanalysator (Fischer-Rosemount, Germany). Bevor die Glucoselimitation einsetzte, wurde der Fed-Batch-Modus mit Hilfe einer Feedstrecke gestartet, wobei durch manuelle Anpassungen eine nicht-limitierende Glucosekonzentration während des Prozessverlaufes sichergestellt wurde.

5.3.2 Prozesskontrolle

Im Verlauf der gesamten Kultivierung wurden Proben zur Analyse der Biomasse-, Glucose- und Nebenproduktkonzentrationen genommen. Messungen der optischen Dichte bei 600 nm (OD_{600} , Shimadzu, PharmaSpec UV-1700) zeigten einen linearen Zusammenhang mit gravimetrischen Trockengewichtsbestimmungen. Die Bestimmung der Glucosekonzentration erfolgte mittels Enzymassay (Roche, Accu-Check AvivaTM), organische Säuren wurden mittels HPLC auf einer Aminex HPX-87H Säule (Bio-Rad, Munich, Germany) gemessen.

Zur Ermittlung quantitativer intrazellulärer Metabolitkonzentrationen wurden Proben bei $\text{OD}_{600} 15$ ($5.5 \text{ g}_{\text{BTM}} \text{ l}^{-1}$) und $\text{OD}_{600} 30$ ($9.6 \text{ g}_{\text{BTM}} \text{ l}^{-1}$) genommen, jede in

5 Replikaten. Probevolumina von 5 ml wurden direkt in kalter Methanollösung (-50°C , $60\% \text{ v v}^{-1}$) gequenchet, um ein sofortiges Abstoppen des Zellmetabolismus zu gewährleisten. Anschließend wurden die Proben bei -25°C abzentrifugiert und die überstandsfreien Zellpellets bei -80°C bis zur Messung gelagert.

5.3.3 Prozessmodell

Um zusätzliche Messungen aller Flüsse zu erhalten, welche die Systemgrenzen des metabolischen Netzwerkmodells bilden, wurde ein Prozessmodell der Fed-Batch-Kultivierung formuliert. Dieses Modell sollte vordergründig die ermittelten Prozessdaten beschreiben und gleichzeitig statistisch signifikante Schätzungen extrazellulärer Flüsse sowie der spezifischen Wachstumsrate liefern. Deshalb wurde ein einfacher Ansatz auf Basis eines unstrukturierten Modells in folgender allgemeiner Form gewählt:

$$\begin{aligned}
 \frac{dV_R}{dt} &= \dot{V}_{\text{in}}(t) & -\dot{V}_{\text{out}}(t) & & V_R(0) &= V_{R,0} \\
 \frac{dV_R \cdot c_S}{dt} &= \dot{V}_{\text{feed}}(t) \cdot c_S & -\dot{V}_{\text{out}}(t) \cdot c_S & - r_S \cdot V_R & c_S(0) &= c_{S,0} \\
 \frac{dV_R \cdot c_X}{dt} &= & -\dot{V}_{\text{out}}(t) \cdot c_X & + r_X \cdot V_R & c_X(0) &= c_{X,0} \\
 \frac{dV_R \cdot c_P}{dt} &= & -\dot{V}_{\text{out}}(t) \cdot c_P & + r_P \cdot V_R & c_P(0) &= c_{P,0} \quad (5.1)
 \end{aligned}$$

Der Volumenstrom in den Bioreaktor hängt vom Zufluss des Substrats Glucose sowie einer pH-Wert-regulierenden Menge an Ammonium ab, d.h. $\dot{V}_{\text{in}}(t) = \dot{V}_{\text{feed}}(t) + \dot{V}_{\text{NH}_3}(t)$. Reduzierungen des Volumens durch Probenahmen sind im zeitabhängigen Abfluss $\dot{V}_{\text{out}}(t)$ berücksichtigt. Für die Beschreibung des Glucoseverbrauchs r_S sowie der Bildung an Biomasse r_X und des einzigen Nebenprodukts Lysin r_P wurden Monod-Kinetiken verwendet [BO84]:

$$\begin{aligned}
 r_S &= \frac{r_X}{Y_{X,S}} + \frac{r_P}{Y_{P,S}} + m_S \cdot c_X \\
 r_X &= \mu_{\text{max}} \cdot \frac{c_S}{c_S + K_{X,S}} \cdot c_X \\
 r_P &= \pi_{\text{max}} \cdot \frac{c_S}{c_S + K_{P,S}} \cdot c_X \quad (5.2)
 \end{aligned}$$

Hierbei sind Modellparameter in Bezug auf Biomasse- und Lysinbildung die Ertragskoeffizienten $Y_{X,S}$ und $Y_{P,S}$, die maximalen Raten μ_{max} und π_{max} sowie die Halbsättigungskonstanten $K_{X,S}$ und $K_{P,S}$. Der Term $m_S \cdot c_X$ berücksichtigt einen möglichen

Erhaltungsstoffwechsel.

Die Bildungsrate von Kohlendioxid wurde über eine Massenbilanzierung der Gasphase bestimmt:

$$r_{\text{CO}_2} = \frac{\dot{V}_G}{V_R \cdot 24.47} \left(Y_{\text{CO}_2,\text{out}} - Y_{\text{CO}_2,\text{in}} \frac{1 - Y_{\text{O}_2,\text{in}} - Y_{\text{CO}_2,\text{in}}}{1 - Y_{\text{O}_2,\text{out}} - Y_{\text{CO}_2,\text{out}}} \right) \quad (5.3)$$

Dabei bezeichnet \dot{V}_G die während des Experiments eingestellte Begasungsrate. Die Terme $Y_{\text{O}_2,\text{in}}$, $Y_{\text{CO}_2,\text{in}}$ und $Y_{\text{O}_2,\text{out}}$, $Y_{\text{CO}_2,\text{out}}$ bezeichnen jeweils die Anteile an Sauerstoff und Kohlendioxid in der Zu- und Abluft.

Auf dieser Basis wurden schließlich spezifische Raten der Glucoseaufnahme ($\text{upt}_{\text{GLC}} = r_S/c_X$), Lysinbildung ($\text{ex}_{\text{LYS}} = r_P/c_X$) und Kohlendioxidbildung ($\text{ex}_{\text{CO}_2} = r_{\text{CO}_2}/c_X$) ermittelt.

5.3.4 Instationäres Markierungsexperiment

Direkt nach der zweiten Probenahme zur Metabolomanalyse wurde das Markierungsexperiment durch schnelle Zugabe einer optimalen ¹³C-markierten Glucosemischung [NGL⁺07] gestartet. Für die anschließende Messung der isotopisch transienten Markierungsanreicherungen im Sub-Sekundenbereich wurde eine schnelle Probenahme-Einheit mit dem Bioreaktor verbunden, welche automatisch 20 Proben innerhalb eines optimierten Probenahmezeitfensters nimmt [EMSD⁺03]. In diesem Fall erfolgte eine schnelle Probenahme innerhalb der ersten 40 s, gefolgt von einem exponentiellen Probenahmeprofil, bis ein näherungsweise isotopisch stationärer Systemzustand in den metabolischen Intermediaten erreicht wurde (vgl. Abb. 5.1, unten).

5.3.5 LC-MS-Analytik intrazellulärer Intermediate

Die Extraktion intrazellulärer Metabolite erfolgte auf Basis einer Methanol/Chloroform-Prozedur [VBMA⁺05]. Die dabei gewonnenen Zellextrakte wurden anschließend auf einem Agilent 1100 HPLC System (Agilent Technologies, Waldbronn, Germany) gekoppelt an ein API 4000 (Applied Biosystems, Concord, Canada) Massenspektrometer mit TurboIon-Spray-Quelle gemessen. Für detaillierte Beschreibungen aller verwendeten Trennungsmethoden sei an dieser Stelle auf [LGT⁺07, TFS⁺08] verwiesen.

Die Berechnung von Intermediatkonzentrationen erfolgte unter Annahme eines biomassespezifischen Zellvolumens für *C. glutamicum* von $V_{\text{cell},X} = 1.93 \text{ ml g}_{\text{BTM}}^{-1}$ [Sch91].

5.4 Prozessverlauf

In Abbildung 5.1 ist der gesamte Kultivierungsverlauf dargestellt. Nach ca. 9.5 h war die Startglucose aufgebraucht und die Kultivierung wurde im Fed-Batch-Betrieb (Aufgabe einzelner Glucosepulse) fortgesetzt. Die Mikroorganismen wuchsen exponentiell mit einer spezifischen Wachstumsrate von $\mu = 0.2 \text{ h}^{-1}$. Die Lysin-konzentration nahm ebenfalls exponentiell zu, wobei bis zum Ende der wachstumsgekoppelten Produktionsphase 3.15 g l^{-1} Lysin gebildet wurden. Nach einer kurzen Anlaufphase stellte sich in der Batch-Phase eine konstante spezifische CO_2 -Bildungsrate ein, deren Wert sich mit dem Übergang in den Fed-Batch-Betrieb deutlich erhöhte. Nach 14 h wurde das isotopisch instationäre Markierungsexperiment gestartet. Infolgedessen sank der CO_2 -Wert aufgrund des Einflusses der Probenahme sowie des hohen Anteils des ^{13}C -Isotops, welches im verwendeten Gasanalysator nur in sehr geringem Maße erfasst wird.

5.4.1 Intrazelluläre Markierungsdynamiken

Die Verläufe der Markierungsanreicherungen von Intermediaten der Glykolyse (EMP), des Pentosephosphatweges (PPP), des Zitratzyklus (TCA) und der Aminosäurebiosynthesewege zeigen deutlich unterschiedliche Zeitkonstanten (Abb. 5.2). Die EMP-Intermediate Glucose-6-phosphat (G6P) und Fructose-6-phosphat (F6P) weisen hierbei die schnellste Dynamik auf und eine Sättigung wird bereits nach $t > 10 \text{ s}$ erreicht. Eine leichte Verzögerung der Anreicherung ist in den Metaboliten des PPP sowie der unteren Glykolyse erkennbar, wodurch isotopische Stationarität erst im Zeitfenster $50 \leq t \leq 1000 \text{ s}$ erreicht wird.

Ein wesentlich anderes Bild ergibt sich hingegen für die TCA-Intermediate. Erst nach $t = 5 \text{ s}$ ist ein Eintrag an Markierung in den Mischungspool Citrat/Isocitrat (CITICIT) beobachtbar. Anschließend findet entlang des TCA eine weitere Verlangsamung der Dynamik statt, so dass in den ersten 40 s kaum Markierung in den Succinat-Pool (SUC) gelangt.

Ein vergleichbarer Effekt wurde auch im instationären Experiment mit dem *E. coli*-Wildtyp beobachtet [NGL⁺07]. Hier erfolgte innerhalb der ersten 16 s kein Markierungseintrag in α -Ketoglutarat (AKG), Glutamat (GLU) und SUC. Dieser Sach-

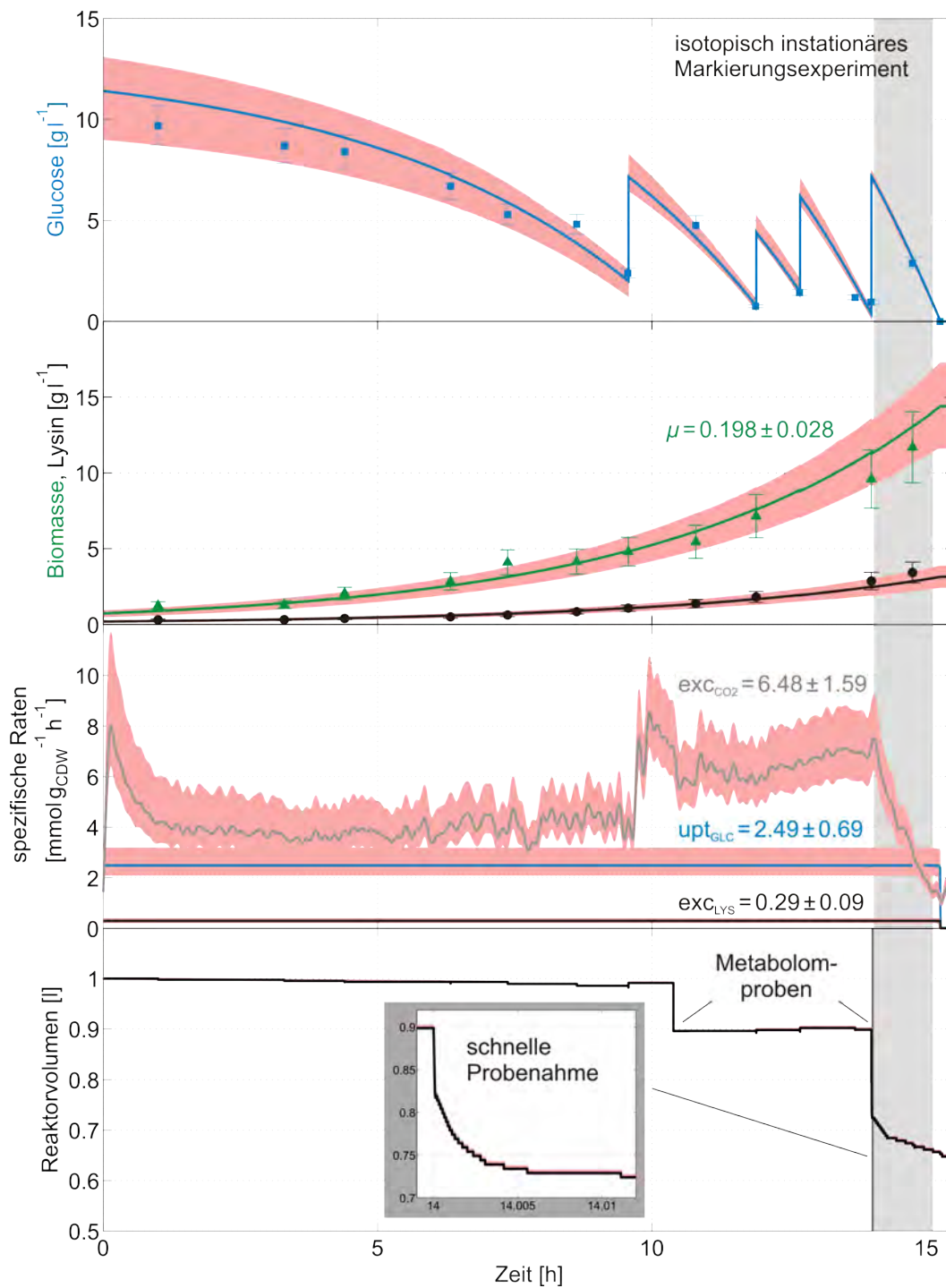


Abbildung 5.1: Verlauf der Fed-Batch-Kultivierung für das ^{13}C -Markierungsexperiment. Durchgezogene Linien stellen simulierte Werte des Prozessmodells dar. Rote Linien kennzeichnen die durch die statistische Analyse ermittelten 95%-Vertrauensbereiche der simulierten Messungen.

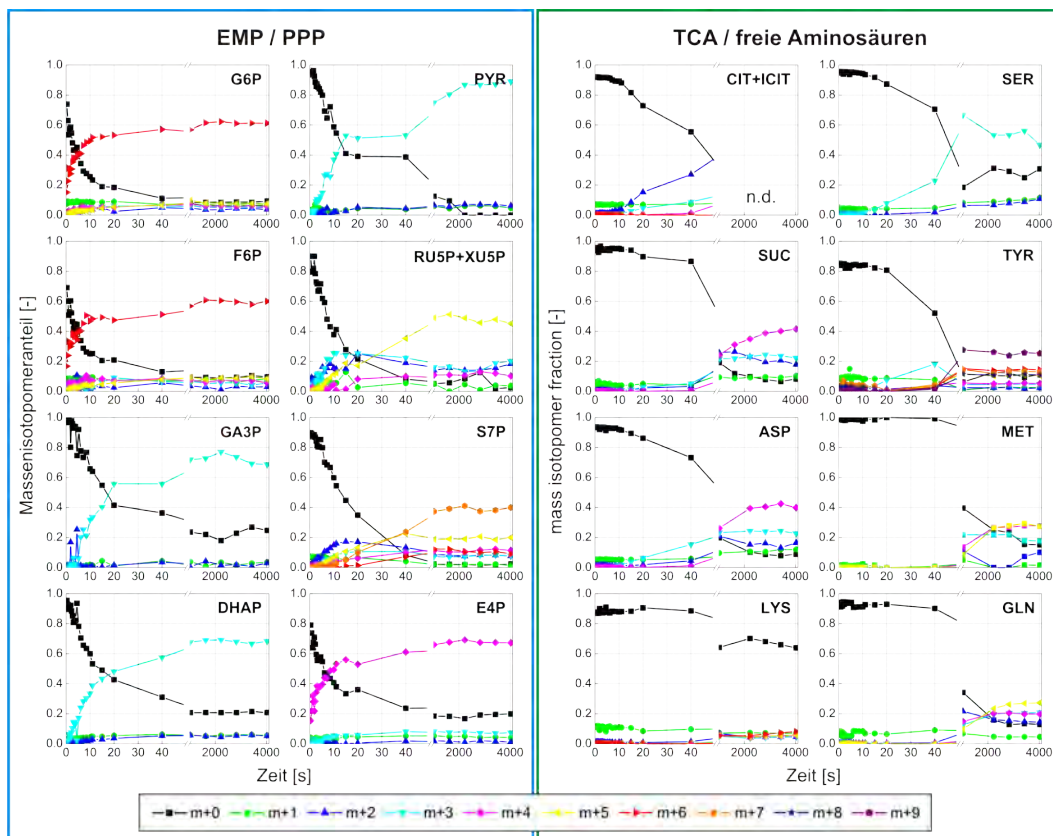


Abbildung 5.2: Experimentell ermittelte instationäre Markierungsanreicherungen in Form von Massenisotopomeren ausgewählter Metabolite im Zeitbereich kurz nach dem Markierungspuls ($0 \leq t \leq 50$ s) sowie im stationären Verlauf nach ca. 15 min Markierung.

verhält wurde mit einem möglichen Eintrag von unmarkiertem Material aus GLU in AKG (hohe Transaminase-Aktivität) erklärt. Interessanterweise stimmt diese Beobachtung gut mit den hier gemessenen Daten überein, zumal *C. glutamicum* nachweislich einen sehr hohen intrazellulären Glutamatpool besitzt. Leider waren im hier untersuchten Fall die Markierungsdaten für AKG und GLU zu stark verrauscht, jedoch unterstützen die Messdaten des Glutaminpools (GLN) diese These.

Die Markierungsanreicherungen in den freien Aminosäuren sind generell stark verzögert und ein näherungsweise isotopisch stationärer Zustand wird erst gegen Ende des Experimentes erreicht. Auffällig ist dabei der unterschiedliche stationäre Markierungszustand in Aspartat (ASP) und Methionin (MET) im Vergleich zu Lysin (LYS). Während alle Massenisotopomeren Spuren in ASP und MET deutlich erhöht sind, weist der Lysinpool lediglich einen Markierungseintrag von ca. 20% auf. Eine naheliegende Erklärung dafür ist die hohe Menge an extrazellulärem Lysin, welches bereits im Zeitraum vor dem Markierungsexperiment gebildet wurde. Im Zuge

der anschließenden Gewinnung von überstandsfreien Markierungsproben könnte ein Teil dieser Stoffmenge auch nach Zentrifugation im Zwischenzellvolumen verblieben sein und somit eine Verdünnung der Markierung des intrazellulären Lysins nach Zellaufschluss bewirkt haben.

5.4.2 Schätzung extrazellulärer Raten und Biomasseabflüsse

Mit Hilfe des Prozessmodells wurden die in Abb. 5.1 aufgeführten spezifischen Raten für Wachstum (μ), Glucoseaufnahme (upt_{GLC}) und Bildung von Lysin (exc_{LYS}) sowie CO_2 (exc_{CO_2}) ermittelt. Zur Schätzung von Standardabweichungen für alle Raten wurde das Prozessmodell genutzt, um auf Basis von Monte-Carlo-Simulationen eine Fehlerübertragung der Prozessdaten auf die gesuchten Raten zu berechnen (nähere Erläuterung dazu im Anhang A.3.2). Eine Liste der Schätzwerte aller Modellparameter befindet sich im Anhang B.1.

Für die Bestimmung der Abflüsse von Intermediatvorstufen in Richtung Biomassebildung wurde ein lineares Verhältnis zwischen Metabolitbedarf und Biomassewachstum angenommen. Entsprechende Werte wurden auf der Grundlage experimenteller Daten der Biomassezusammensetzung eines exponentiell wachsenden *C. glutamicum*-Lysinproduzenten geschätzt [MdGW⁺96] und sind im Anhang in Tabelle B.2 zu finden.

5.5 Modellierung isotopisch stationärer Daten

Die Rekonstruktion der unbekanntem intrazellulären Stoffflüsse aus den berechneten spezifischen Raten sowie den gemessenen Markierungsdaten bedingt eine komplexe Prozedur aus mathematischer Modellierung und Computersimulation. Aufgrund der teilweise zyklischen Netzwerkstruktur verwendeter Modelle können die zugrunde liegenden Bilanzgleichungen nicht explizit gelöst werden. Deshalb ist eine iterative Vorgehensweise für die Flusschätzung nötig, d.h. Minimierung der Unterschiede zwischen gemessenen und simulierten ^{13}C -Markierungsverteilungen.

An dieser Stelle wird nicht weiter auf die technischen und mathematischen Details der Durchführung einer ^{13}C -MFA eingegangen, da diese hinreichend in der Literatur beschrieben sind [Wie95, Nöh06, Wah07]. Wesentliche Aspekte, welche den mathematischen Hintergrund der verwendeten Methoden zur Simulation und Systemanalyse betreffen sind in kompakter Darstellung im Anhang A zu finden.

Im Rahmen dieser Arbeit wurde für den Großteil des gesamten Workflows bestehend aus Modellerstellung, Optimierung und Sensitivitätsanalyse das Softwarewerkzeug 13CFLUX [WMPdG01, ZFRS09] genutzt.

Im Folgenden wird zunächst die Modellierung der isotopisch stationären Daten auf der Grundlage von zwei Modellansätzen unterschiedlicher Komplexität erläutert. Der Messdatensatz umfasste neben den extrazellulären Flussmessungen die stationären Markierungsanreicherungen intrazellulärer Metabolite, wobei aus den letzten 4-5 Proben entsprechende Mittelwerte und zugehörige Standardabweichungen bestimmt wurden.

5.5.1 Fokussiertes Modell des Zentralstoffwechsels

In den Abbildungen 5.3 und 5.4 sind alle bekannten Reaktionen der Glykolyse (EMP), des Pentosephosphatweges (PPP), des Zitratzyklus (TCA), der Anaplerose (ANA) sowie der Lysinbiosynthese (LYS) von *C. glutamicum* dargestellt.

Für die Auswertung der gewonnenen Markierungsdaten von Intermediaten des Zentralstoffwechsels wurde jedoch zunächst ein fokussierter Modellansatz gewählt, welcher die Reaktionen des EMP, PPP, TCA, ANA sowie einer stark vereinfachten Lysinbiosynthese enthält (Abb. 5.5).

Für diesen nachfolgend mit ^{13}C -MFA ZSW MM1 (Messmodell 1) bezeichneten Modellansatz wurden eine Reihe von Annahmen getroffen:

1. Für *C. glutamicum* sind fünf anaplerotische Reaktionsschritte bekannt. Die-

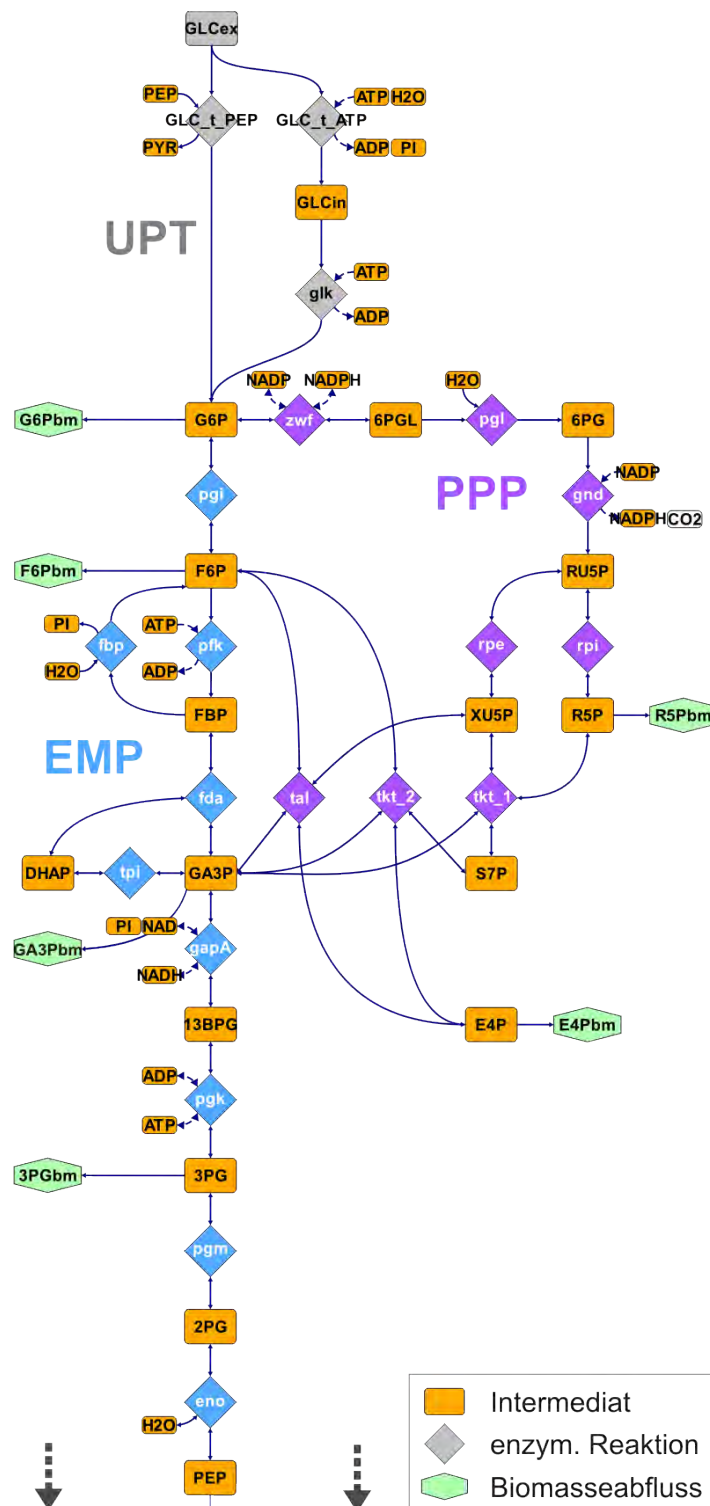


Abbildung 5.3: Metabolisches Reaktionsnetzwerk der Glykolyse und des Pentosephosphatweges von *C. glutamicum* [EB05]. Die Visualisierung des Netzwerks erfolgte unter Nutzung des Werkzeugs Omix, welches von Peter Droste am Institut für Biotechnologie 2 im Rahmen seiner Dissertation entwickelt wurde.

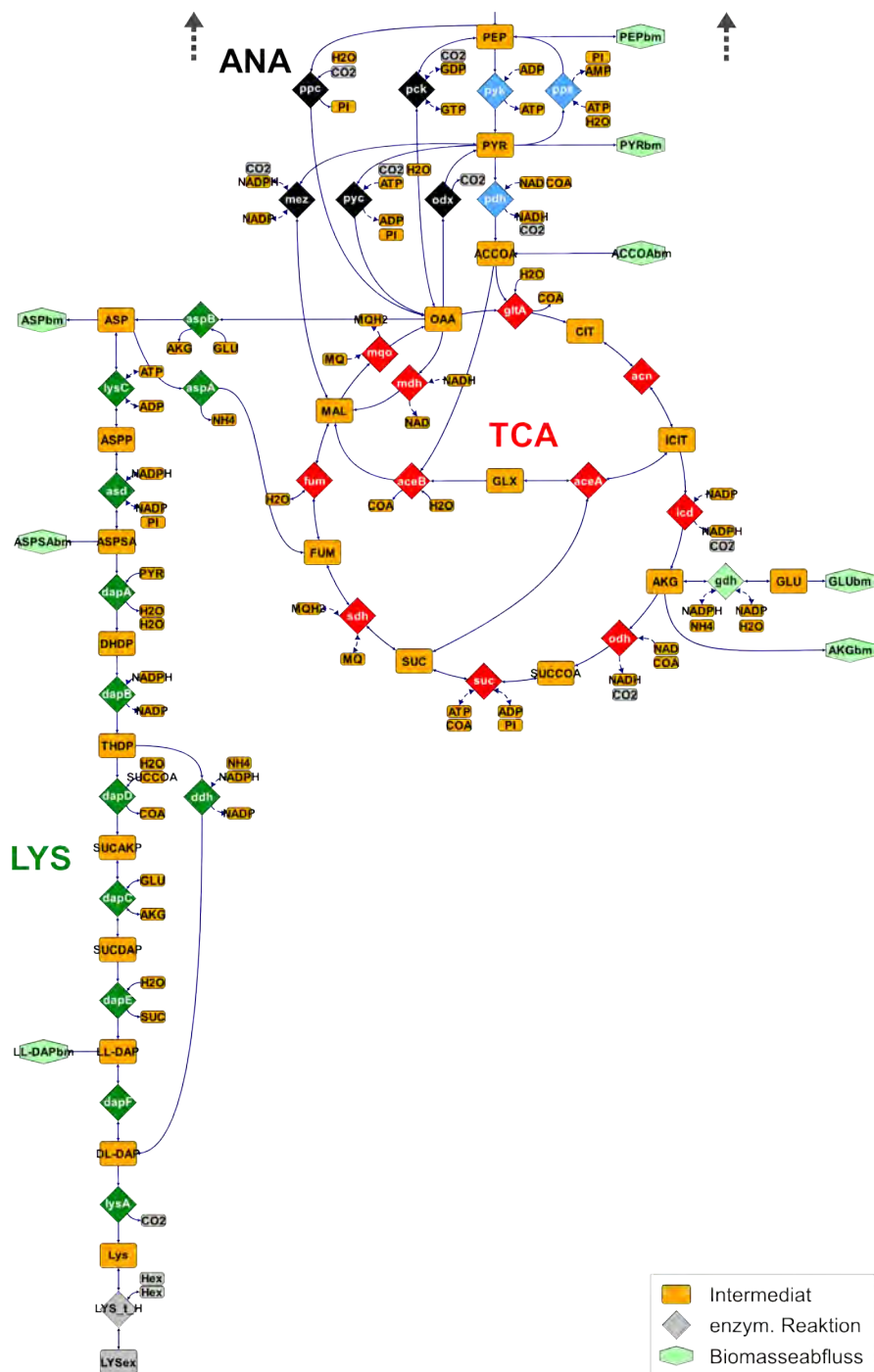


Abbildung 5.4: Metabolisches Reaktionsnetzwerk des Zitratzyklus, der Anaplerose und des Lysinbiosyntheseweges von *C. glutamicum* [EB05]. Aufgrund der Verzweigung am Metabolit Tetrahydrodipicolinat (THDP) umfasst die Lysinbiosynthese 10 bzw. 7 einzelne Reaktionsschritte ausgehend vom ersten Precursor Oxalacetat. Innerhalb der Reaktion Dihydrodipicolinat-Synthese (dapA) wird der zweite Precursor Pyruvat (PYR) in das C-Gerüst eingebunden. Darüber hinaus sind 4 Reduktionsäquivalente in Form von NADPH (inklusive Glutamat-Rezyklierung) notwendig, welche über den Zentralstoffwechsel generiert werden müssen.

se werden durch die PEP-Carboxylase (ppc), die PEP-Carboxykinase (pck), die Pyruvat-Carboxylase (pyc), die Oxalacetat-Decarboxylase (odx) und das Malat-Enzym (mez) katalysiert [EB05]. Aufgrund der Deletion im *pck*-Gen des untersuchten Stammes kann kein Rückfluss von OAA nach PEP erfolgen. Die beiden unidirektionalen Reaktionen pyc und odx bewirken jeweils den Vorwärts- und Rückwärtsschritt zwischen PYR und OAA, weshalb diese durch eine reversible Reaktion modelliert sind.

2. *C. glutamicum* verfügt über eine cytosolische Malat-Dehydrogenase (mdh) und eine membrangebundene Malat-Quinone-Oxireduktase (mqo). Es wird vermutet, dass beide Reaktionen unter *in vivo*-Bedingungen gegenläufig operieren [MvdRDY00]. Unabhängig davon bedingen beide Reaktionen die gleiche Änderung im C-Gerüst von MAL bzw. OAA, wodurch eine getrennte Identifizierung nicht möglich ist. In diesem Fall muss eine einzige reversible Reaktion angenommen werden.
3. Die Metabolitpools der Phosphoglycerate 13PG, 2PG und 3PG sind mittels LC nicht getrennt messbar und wurden deshalb in einen gemeinsamen Pool PGP gelumpt.
4. Die lineare Sequenz des oxidativen PPP (zwf, pgl, gnd) wurde zu einem Reaktionsschritt zusammengefasst. Die Pentose-5-phosphate RU5P, XU5P und R5P wurden unter der Annahme hoher Austauschraten zu einem Pool P5P gelumpt. Dies betrifft auch die TCA-Intermediate CIT, ACN, ICIT sowie SUC-COA, SUC und FUM, bei denen ebenfalls Pool-Lumping in CITICIT und SUC-FUM erfolgte.
5. Für die verbleibende Scrambling-Reaktion zwischen SUCFUM und MAL wurde ein symmetrisches Scrambling angenommen.
6. Die Lysinbiosynthese in *C. glutamicum* erfolgt auf zwei parallelen Reaktionswegen (dapDCEF und ddh) und umfasst eine Reihe linearer Reaktionsschritte. Unter Vernachlässigung zusätzlicher stöchiometrischer Kopplungen (z.B. GLU-AKG) und Berücksichtigung der Scrambling-Reaktion dapE wurde die Lysinbildung vereinfacht durch zwei Reaktionen abgebildet, bei denen das Scrambling-Verhältnis nicht festgelegt ist.
7. Die im Modell getroffenen Irreversibilitätsannahmen wurden anhand thermodynamischer Überlegungen abgeleitet.

Alle im Modell enthaltenen Reaktionen und zugehörige C-Atomtransitionen sind in Tabelle B.3 des Anhangs aufgeführt.

Simulierte Markierungsdaten

Das im Zuge der Optimierung mit Hilfe von 13CFLUX parametrisierte Modell beschreibt die gemessenen Massenisotopomerverteilungen in guter Genauigkeit (vgl. Tab. B.5, ZSW MM1). Signifikante Abweichungen zwischen gemessenen und simulierten Werten sind vor allem für die unmarkierte (m+0) und vollmarkierte (m+3) Spur von PYR zu finden. Allerdings ist hier aufgrund der verhältnismäßig hohen Mess-Standardabweichungen der Anteil am Gesamtresiduum $\kappa(\alpha) = 465.4$ gering.

Intrazelluläre Stoffflussverteilung

In Abbildung 5.5 sind die Ergebnisse der Flussschätzungen dargestellt. Ein Großteil der aufgenommenen Glucose wird demnach direkt über die Glykolyse verstoffwechselt. Hingegen scheint die Aktivität des PPP und die damit verbundene NADPH-Generierung und CO₂-Produktion verhältnismäßig gering zu sein. Das für die Lysinbiosynthese notwendige NADPH müsste somit über den TCA durch Isocitrat-Dehydrogenase (icd) oder über die Anaplerose durch das Malat-Enzym (mez) gebildet werden. In diesem Fall wurde ein moderater Fluss für icd ermittelt, die mez-Reaktion läuft hingegen in die NADPH verbrauchende Richtung (vgl. Abb. 5.4).

Darüber hinaus fällt der sehr geringe Fluss der Glutamat-Dehydrogenase (gdh) auf, welcher direkt an die Aspartat-Aminotransferase (aspB) gekoppelt ist. Für das Auffüllen des TCA mit C4-Körpern scheint neben mez auch die PEP-Carboxylase (ppc) aktiv zu sein. Letztendlich sind bei dieser Modellvariante ca. 11% der aufgenommenen Glucose im Lysin wiederzufinden und lediglich 139% wird in Form von CO₂ ausgeschleust.

5.5.2 Erweitertes Modell mit Aminosäurebiosynthesewegen

Für die Integration von Markierungsdaten aus freien Aminosäuren wurde der zuvor definierte Modellansatz um die Biosynthesewege aller Aminosäuren erweitert (vgl. Tab. B.4). Dieser Ansatz wird nachfolgend als ¹³C-MFA ZSW_AS bezeichnet. Hierbei sind folgende Punkte wesentlich:

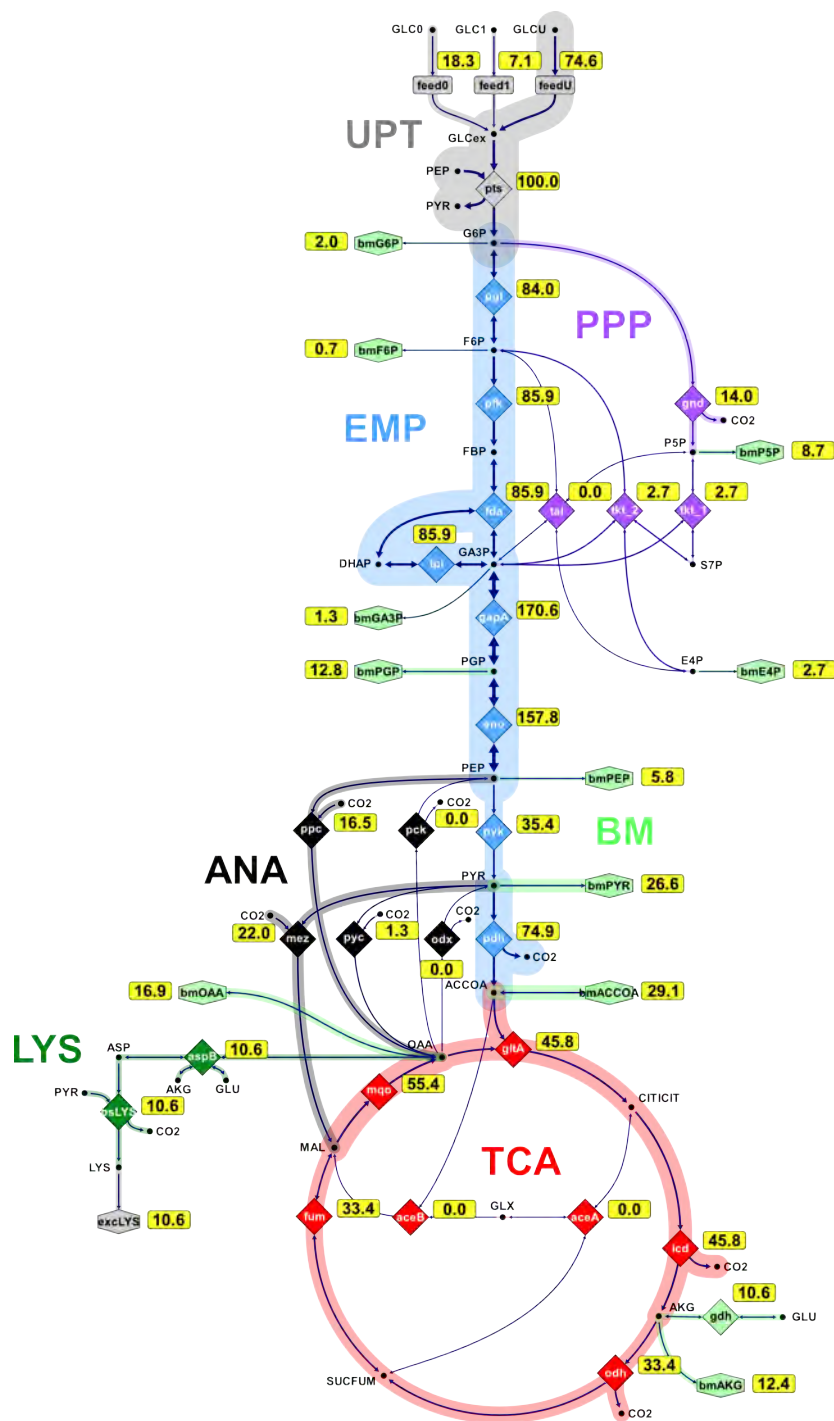


Abbildung 5.5: Verteilung geschätzter intrazellulärer Nettoflüsse für das fokussierte Netzwerk im isotopisch stationären Modellierungsansatz. Mithilfe von Omix können Simulationsdaten auch im Netzwerkkontext visualisiert werden [DNNW09]. In dieser Abbildung repräsentiert die Pfeilbreite die geschätzte Stoffflussverteilung. Die angegebenen Werte stellen molare Anteile der Glucoseaufnahme dar.

1. Beim Hinzufügen der Aminosäurereaktionen wurde insbesondere auf eine korrekte Stöchiometrie geachtet, d.h. alle Reaktionspartner, deren C-Gerüst an einer Stelle des Netzwerkes direkt umgesetzt wird (z.B. $\text{AKG} \xrightarrow{\text{gdh}} \text{GLU}$), werden auch in Massenbilanzen erfasst, in denen ihr C-Gerüst nicht beeinflusst wird (z.B. $\text{OAA} + \text{AKG} \xrightarrow{\text{aspB}} \text{ASP} + \text{GLU}$). Auf diese Weise enthält das Modell zusätzliche stöchiometrische Kopplungen, die den Lösungsraum biologisch sinnvoll einschränken sollen. Für eine entsprechende Umsetzung wurden im Gegensatz zum Ausgangsmodell getrennte Massenbilanzen für SUC-COA, SUC und FUM formuliert.
2. Sofern es die Stöchiometrie nicht beeinflusste, wurden lineare Reaktionswege zusammengefasst.
3. Im Hinblick auf die Systemgrenzen werden die Abflüsse aus Precursoren (z.B. ACCOA) in Richtung Aminosäuren (z.B. Leucin) und sonstiger Biomassekomponenten (z.B. Lipide) nun separat modelliert (vgl. Tab. B.2).

Simulierte Markierungsdaten

Um den Einfluss verschiedener Messkonfigurationen auf die zwei Modellvarianten untersuchen zu können, wurde das erweiterte Modell an die Markierungsdaten der ZSW-Intermediate (ZSW_AS MM1) sowie durch Einbindung zusätzlicher Aminosäuremessungen (ZSW_AS MM2) angepasst (vgl. Tab. B.5 und B.6). Mit Ausnahme von P5P werden in beiden Fällen die Massenisotopomere des ZSW deutlich besser beschrieben als mit Hilfe des fokussierten Modellansatzes. Darüber hinaus ist auf Grundlage des zweiten Messmodells eine gute Übereinstimmung zwischen gemessenen und simulierten Markierungen freier Aminosäuren zu finden. Größere Abweichungen treten hier im unmarkierten und vollmarkierten GLN-Pool auf. Die Gesamtresiduen betragen jeweils $\kappa(\alpha) = 304.1$ bzw. $\kappa(\alpha) = 1177.9$. Letzterer Wert ist bedeutend höher, da hier ein Drittel mehr Messungen in die Fluss-schätzung eingeflossen sind. Zudem waren die geschätzten Messfehler der Aminosäuren verhältnismäßig gering und damit deren gewichteter Einfluss auf das Residuum hoch.

Intrazelluläre Stoffflussverteilung

Im Vergleich zum fokussierten Netzwerk wurde für beide Messmodellvarianten im erweiterten Netzwerk ein deutlich höherer Fluss über den PPP berechnet (Abb. 5.6).

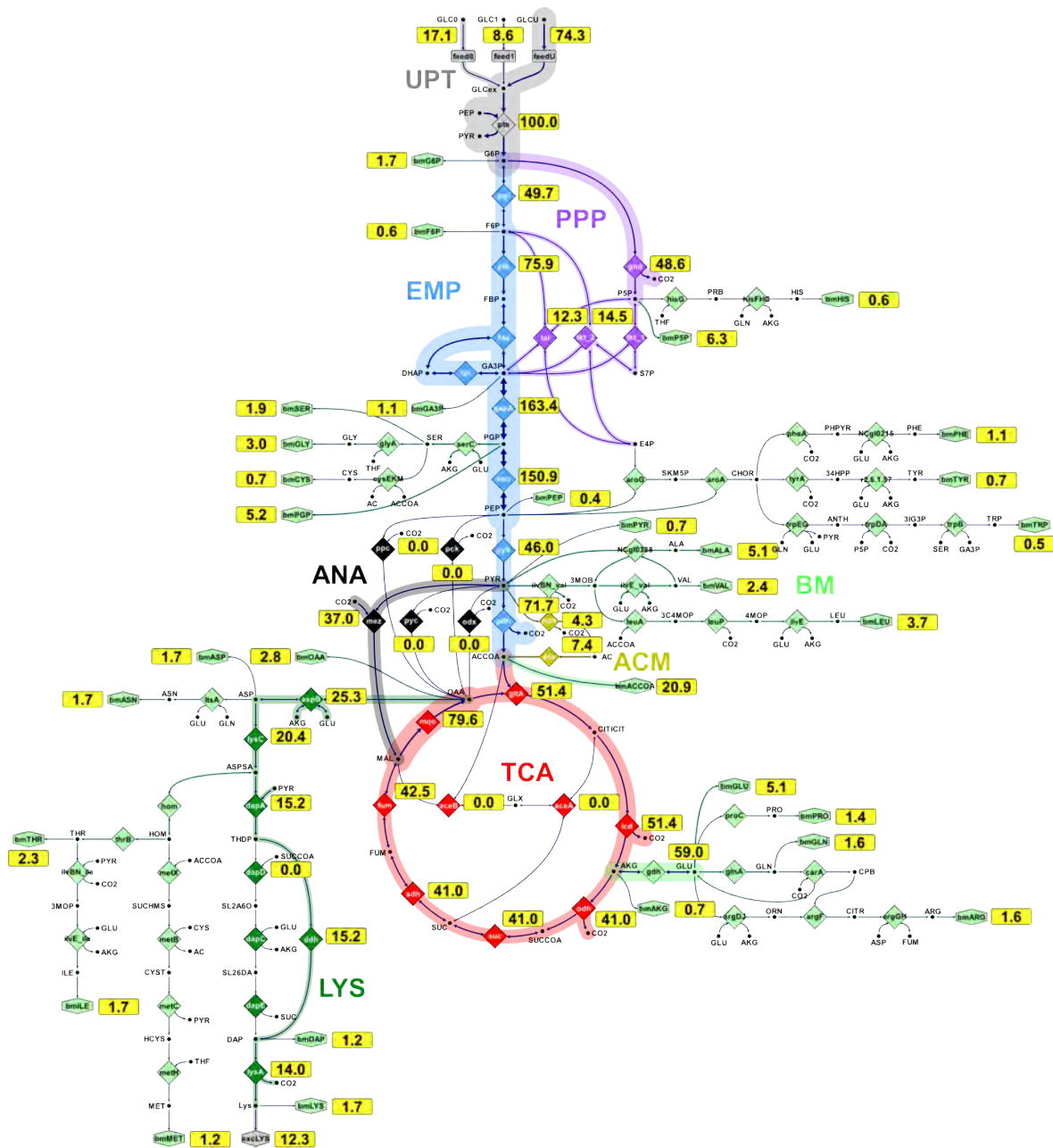


Abbildung 5.6: Verteilung geschätzter intrazellulärer Nettoflüsse für das erweiterte Netzwerk mit Aminosäurebiosynthesewegen im isotopisch stationären Modellierungsansatz. Ausgewählte Flüsse der Messmodellvariante 2 (MM2) sind als molare Anteile der Glucoseaufnahme angegeben.

Dementsprechend ist hier der Beitrag zur Generierung von NADPH und CO₂ wesentlich höher. In der in Abb. 5.6 dargestellten Lösung von Messmodellvariante 2 wurde wie im fokussierten Modellansatz ein sehr hoher Fluss über das Malat-Enzym (mez) unter NADPH-Verbrauch geschätzt.

Desweiteren wurde generell ein stärkerer Fluss über die Reaktion gdh ermittelt, der sich vor allem durch die zusätzlichen stöchiometrischen Kopplungen in den Aminotransferase-Reaktionen ergibt.

5.6 Modellierung isotopisch instationärer Daten

Im durchgeführten Fed-Batch-Experiment wurden neben den stationären Markierungsdaten auch instationäre Markierungsverläufe sowie Konzentrationen intrazellulärer Metabolite ermittelt. Somit standen erstmalig Messdaten für den Organismus *C. glutamicum* zur Verfügung, welche im Rahmen einer isotopisch instationären Stoffflussanalyse ausgewertet werden konnten.

Grundlage für diese Auswertung bildet der mit Gl. 3.14 eingeführte metabolisch stationäre und isotopisch instationäre Modellierungsansatz. Zur Berechnung intrazellulärer Stoffflüsse muss somit nicht wie bei der klassischen ¹³C-MFA ein algebraisches Gleichungssystem, sondern ein Differentialgleichungssystem mit Metabolitpools als zusätzliche Zustandsvariablen gelöst werden. Entsprechend höher gestalten sich die Anforderungen an die mathematischen Prozeduren zur effizienten Simulation, Optimierung und Sensitivitätsanalyse solcher Modelle. Für ausführliche Erläuterungen zu dieser Thematik wird wiederum auf die Literatur verwiesen [Nöh06,NWW06].

Mit der Zielstellung des Vergleichs beider Modellierungsansätze wurde im konkreten Fall das fokussierte Modell des Zentralstoffwechsels (vgl. Abschnitt 5.5.1) genutzt, um gemessene Massenisotopomeren-Anteile und Konzentrationen von ZSW-Intermediaten zu integrieren. Entsprechend der bereits eingeführten Notation wird dieser Ansatz im Folgenden als INST-¹³C-MFA ZSW MM1 bezeichnet. Die damit verbundene Modellanpassung und anschließende statistische Analyse wurde von Frau Dr. Katharina Nöh am Institut für Biotechnologie 2 durchgeführt, bei der eine von ihr entwickelte Software namens INST-13CFLUX zum Einsatz kam [Nöh06].

In diesem Fall umfasste der Messdatensatz die extrazellulären Flussmessungen sowie die instationären Markierungsanreicherungen und Poolgrößen intrazellulärer Metabolite. Für die Messungen der Massenisotopomere wurden zugehörige Standardabweichungen mit Hilfe einer Fehlerübertragungsrechnung aus den zuvor ermittelten Fehlern der stationären Markierungsdaten geschätzt. Die Messfehler der Poolgrößen wurden hingegen direkt aus den Replikatproben ermittelt.

Simulierte Markierungsanreicherungen und Poolgrößen

In den Abbildungen 5.7 und 5.8 sind die Ergebnisse der Modellanpassung für ausgewählte Massenisotopomeren-Verläufe sowie alle Poolgrößen des Modells dargestellt.

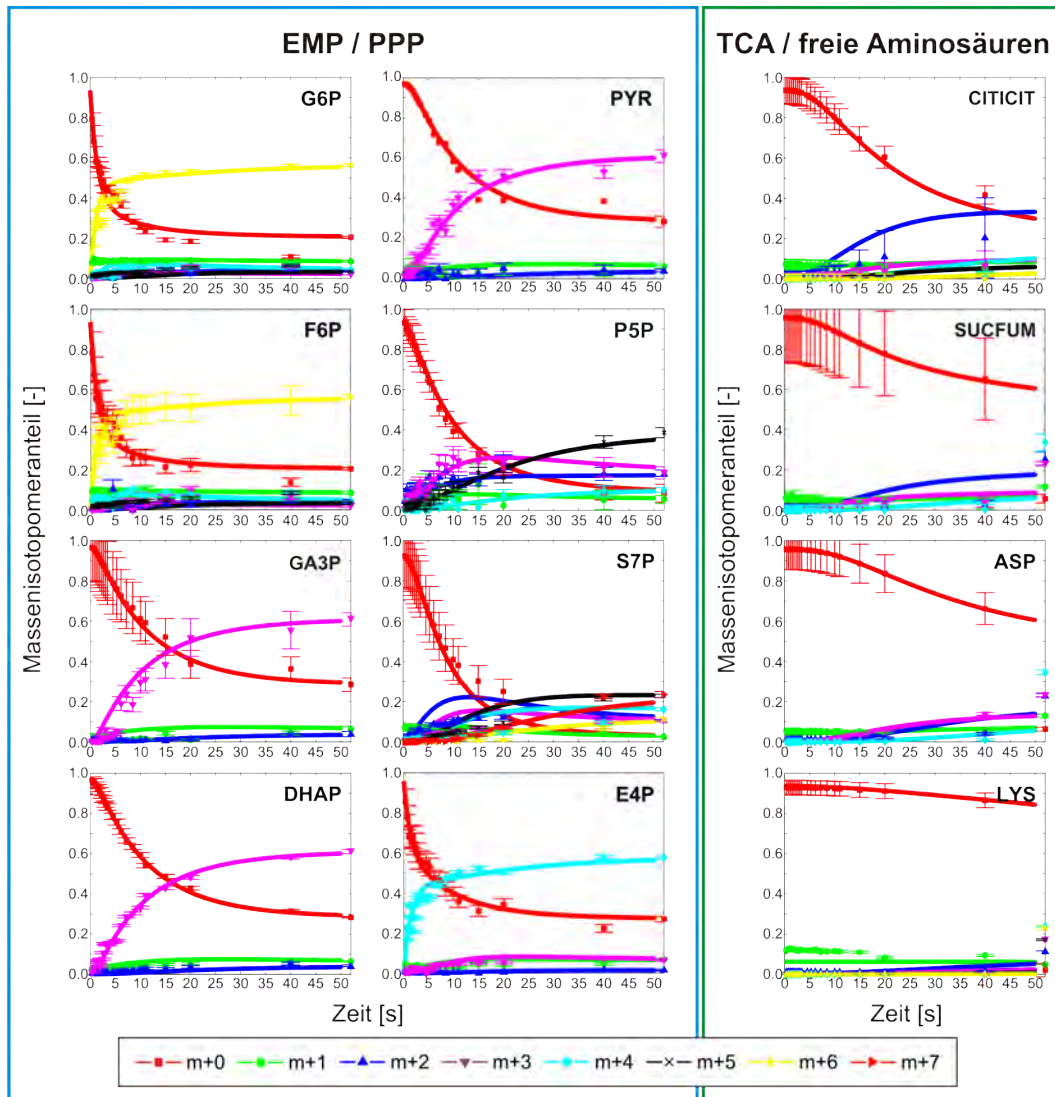


Abbildung 5.7: Zeitliche Verläufe simulierter Massen isotopomer-Anteile ausgewählter intrazellulärer Metabolite auf Basis des isotopisch instationären Modellierungsansatz. Der Modellanpassung lagen die Daten im Zeitfenster $0 \leq t \leq 40$ s nach dem Markierungspuls zugrunde. Die Messpunkte an der rechten Diagrammbegrenzung stellen den isotopisch stationären Zustand dar.

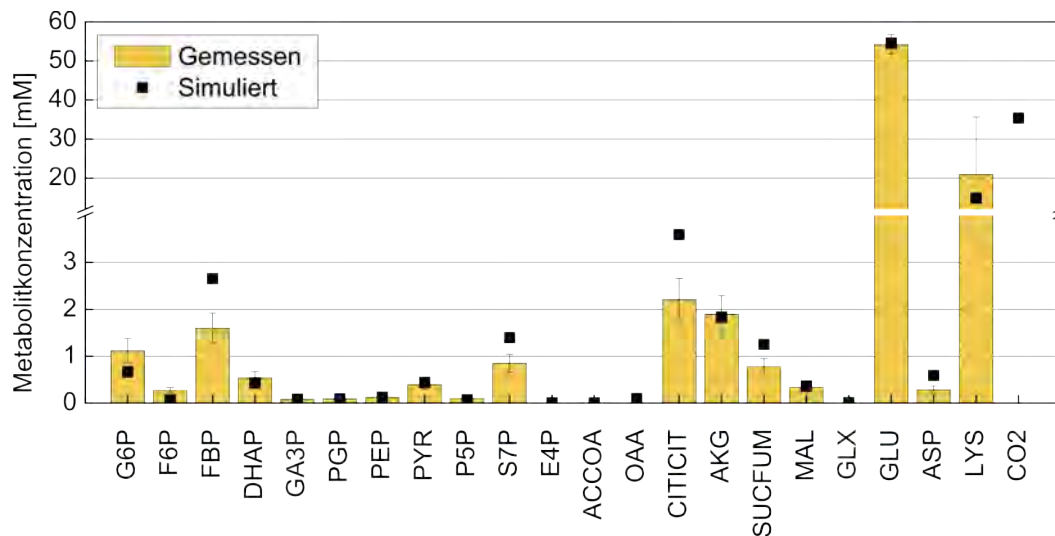


Abbildung 5.8: Geschätzte intrazelluläre Metabolitkonzentrationen auf Basis des isotopisch instationären Modellierungsansatz.

Es wird deutlich, dass die Mehrzahl der gemessenen Markierungsverläufe im Rahmen ihrer Fehlertoleranzen sehr gut durch das Modell beschrieben wird. Dies betrifft sowohl die Metabolite des EMP und PPP, welche durch eine schnelle Markierungsanreicherung gekennzeichnet sind (z.B. G6P, P5P), als auch die Intermediate des TCA und der Lysinbiosynthese, deren Zeitkonstanten deutlich höher sind (z.B. CITICIT, ASP).

Darüber hinaus ist bei den Metaboliten mit hoher Markierungsdynamik bereits erkennbar, dass auch der isotopisch stationäre Punkt am Ende des Experimentes korrekt vorhergesagt wird. Starke Abweichungen zwischen Messung und Simulation treten hingegen für die Massenspuren des PPP-Intermediats S7P auf, wobei hier im Anfangsbereich ein deutliches Überschwingen in den Massenisotopomeren m+2 und m+3 zu erkennen ist.

Im Fall der Metabolitpools werden die gemessenen Konzentrationswerte ebenfalls gut durch das Modell beschrieben. Ausnahmen bilden hier die Intermediate FBP, S7P und CITICIT, welche um bis zu 65% zu hoch simuliert werden.

Intrazelluläre Stoffflussverteilung

Bezüglich der resultierenden Stoffflussverteilung zeigt sich ein teilweise ähnliches Bild wie im isotopisch stationären Ansatz mit gleicher Netzwerkgrundlage (Abb. 5.9). Während das Splitverhältnis zwischen EMP und PPP bestätigt wird, sind deut-

liche Unterschiede im unteren Teil des EMP, im TCA und der Anaplerose erkennbar. Hierbei ist insbesondere der doppelt so hohe Anteil des TCA-Flusses auffällig, welcher durch einen erhöhten Zufluss entlang der Reaktionen Pyruvatkinase (pk) und Pyruvat-Dehydrogenase (pdh) gespeist wird. Als anaplerotische Reaktion scheint hauptsächlich die Pyruvat-Carboxylase (pyc) aktiv zu sein.

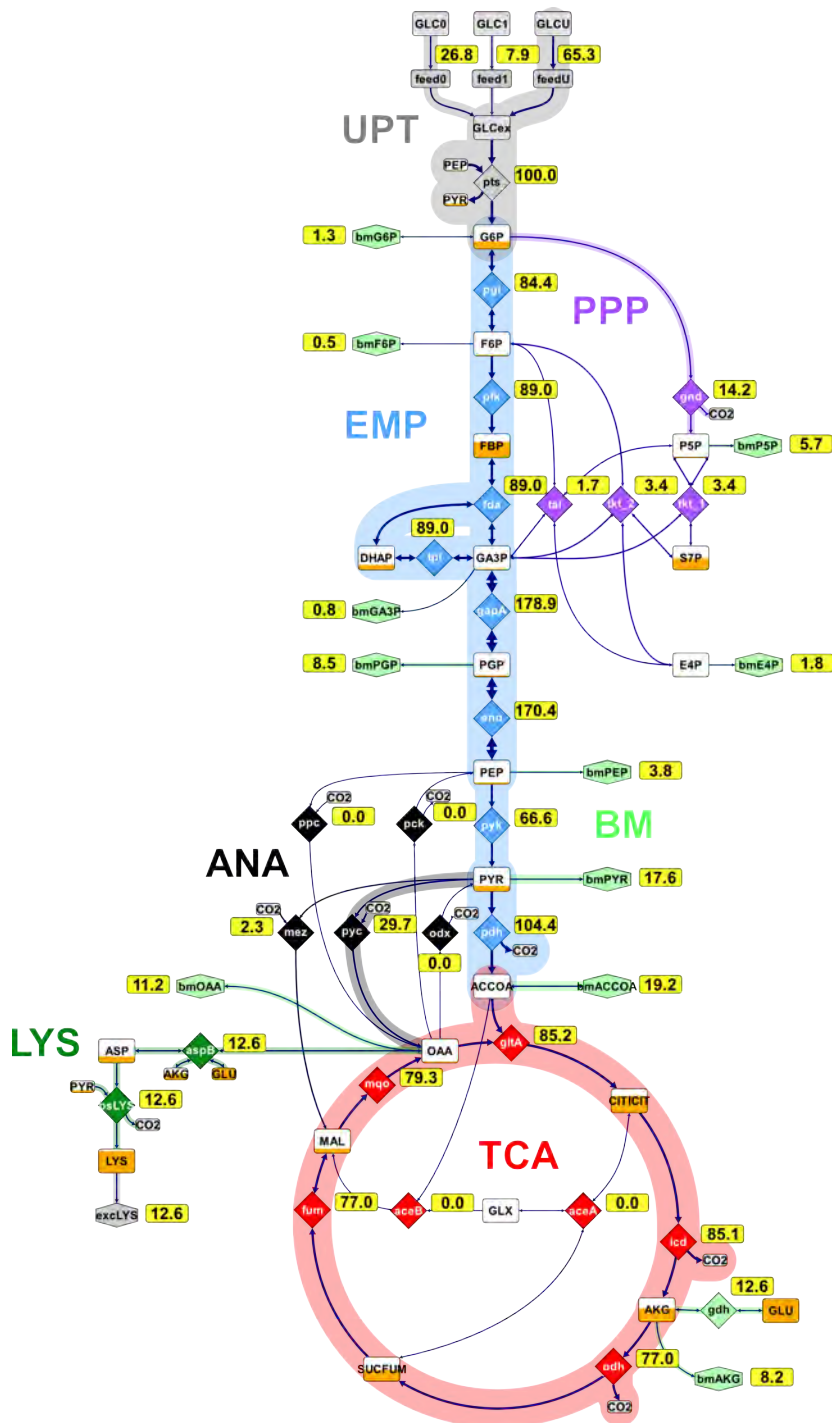


Abbildung 5.9: Verteilung geschätzter intrazellulärer Nettoflüsse für das fokussierte Netzwerk im isotopisch instationären Modellierungsansatz. Angegebene Werte stellen molare Anteile der Glucoseaufnahme dar.

5.7 Statistische Analyse

Die in den vorangegangenen Abschnitten vorgestellten intrazellulären Stoffflussverteilungen sind als optimale Lösungen der Modellanpassungen an die gemessenen Markierungsdaten entstanden. Grundsätzlich stellt sich jedoch die Frage, inwieweit die ermittelten Stoffflüsse statistisch signifikant sind und somit haltbare biologische Aussagen getroffen werden können.

Im Fall von Markierungsnetzwerken gestaltet sich die Beantwortung dieser Frage als keine einfache Aufgabe, da es sich hier in der Regel um nichtlineare Modelle handelt und die etablierten Methoden der linearen Statistik nicht direkt angewendet werden können.

5.7.1 Linearisierungsansatz

Eine Möglichkeit dieses Problem zu beheben, besteht in der linearen Approximation des betrachteten Systems. Entsprechende Verfahren dafür sind auch in der jeweils verwendeten Software 13CFLUX bzw. INST-13CFLUX integriert. Für nähere Erläuterungen zu dieser klassischen Methode sei an dieser Stelle auf den Anhang A.3.1 verwiesen.

Durch Anwendung der linearen Statistik wurden die in Tabelle 5.1 dargestellten Standardabweichungen intrazellulärer Nettoflüsse aller Modellansätze geschätzt. Dabei ist Folgendes festzustellen:

1. Die Flüsse des EMP und PPP sind in beiden Netzwerk- und Messmodellvarianten des isotopisch stationären Ansatzes als auch auf Basis der instationären Modellierung mit guter Genauigkeit bestimmt. Eine Ausnahme bildet hier die Reaktion Pyruvatkinase (pyk).
2. Während innerhalb des TCA im instationären Ansatz für die Mehrheit der Flüsse enge Vertrauensbereiche ermittelt wurden, trifft dies im stationären Fall nur für das erweiterte Netzwerk mit Messmodellvariante 2 (¹³C-MFA ZSW_AS MM2) zu.
3. Fast alle Flüsse um den Knotenpunkt PEP-PYR-MAL-OAA sind statistisch nicht identifizierbar. Lediglich für die Reaktion Pyruvat-Dehydrogenase (pdh) wurde auf Grundlage des instationären Ansatzes ein statistisch signifikanter Flusswert geschätzt.

4. Die Reaktionen der Lysinbiosynthese sind bestimmbar, allerdings kann keine Aussage über das Splitverhältnis der parallelen Reaktionspfade (dapDCEF und ddh) getroffen werden.

Es ist allgemein bekannt, dass die Verwendung des Linearisierungsansatzes nicht unproblematisch ist und bei stark nichtlinearen Modellen sowie großen Messwertfehlern zu einer deutlichen Verzerrung der erzielten statistischen Resultate führen kann (Abb. 5.10).

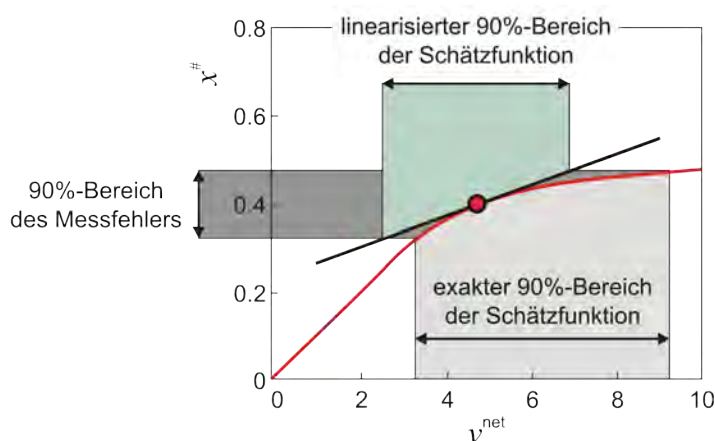


Abbildung 5.10: Schematische Darstellung der Fehlerübertragung bei der linearen Approximation einer nichtlinearen Funktion.

5.7.2 Monte-Carlo-Simulationen

Eine alternative Methode zur Fehlerabschätzung stellt die Monte-Carlo-Simulation dar, bei welcher direkt mit den nichtlinearen Modellgleichungen gerechnet werden kann. Grundsätzlich werden dabei die klassische Monte-Carlo-Simulation (parametrische Bootstrap-Methode) sowie die im Rahmen der Bayes-Statistik verwendete Markov-Chain-Monte-Carlo-Methode unterschieden [vL03]. Ersteres Verfahren fand im Rahmen dieser Arbeit Anwendung und ist im Anhang A.3.2 näher erläutert.

Aufgrund des hohen Rechenaufwandes kann diese Methode derzeit noch nicht für große algebraische Isotopomeren-Gleichungssysteme (isotopisch stationäre Modelle) sowie ebenfalls nicht für mittlere Isotopomeren-Netzwerke auf Basis von ODE-Systemen (isotopisch instationäre Modelle) angewendet werden.

Aus diesem Grund wurden Monte-Carlo-Simulationen ausschließlich für das fokussierte Netzwerk im isotopisch stationären Modellierungsansatz durchgeführt. Ent-

Tabelle 5.1: Schätzwerte intrazellulärer Stoffflüsse auf Basis der zwei Modellierungsansätze (^{13}C -MFA, INST ^{13}C -MFA), Netzwerkvarianten (ZSW, ZSW_AS) sowie Messmodelle (MM1, MM2). Zugehörige Standardabweichungen wurden mittels Linearisierungsansatz bestimmt. Alle Angaben in $\text{mmol g}_{\text{BTM}}^{-1} \text{h}^{-1}$.

Reaktion	^{13}C -MFA ZSW MM1		^{13}C -MFA ZSW_AS MM1		^{13}C -MFA ZSW_AS MM2		INST ^{13}C -MFA ZSW MM1	
	Fluss	Std.abw.	Fluss	Std.abw.	Fluss	Std.abw.	Fluss	Std.abw.
pts	2.00	0.21	2.33	0.22	2.37	0.22	3.03	0.10
pgi	1.68	0.21	0.77	0.25	1.18	0.20	2.56	0.13
pfk	1.72	0.21	1.64	0.19	1.80	0.19	2.70	0.14
fda	1.72	0.21	1.64	0.19	1.80	0.19	2.70	0.14
tpi	1.72	0.21	1.64	0.19	1.80	0.19	2.70	0.14
gapA	3.41	0.43	3.67	0.40	3.88	0.41	5.42	0.28
eno	3.16	0.43	3.37	0.40	3.58	0.41	5.16	0.28
pyk	0.71	> 10	0.93	0.71	1.09	> 10	2.02	2.21
gnd	0.28	0.00	1.52	0.28	1.15	0.19	0.43	0.04
tal	0.00	0.00	0.41	0.09	0.29	0.06	0.05	0.01
tk1_1	0.05	0.00	0.46	0.09	0.34	0.06	0.10	0.01
tk1_2	0.05	0.00	0.46	0.09	0.34	0.06	0.10	0.01
pdh	1.50	> 10	1.34	> 10	1.70	> 10	3.16	0.28
gltA	0.92	0.42	1.10	0.39	1.22	0.40	2.58	0.27
icd	0.92	> 10	1.10	> 10	1.22	0.60	2.58	0.37
odh	0.67	> 10	0.85	> 10	0.97	0.60	2.33	0.37
suc	-	-	0.85	> 10	0.97	> 10	-	-
sdh	-	-	0.85	0.39	0.97	0.40	-	-
fum	0.67	0.30	0.89	0.28	1.01	0.28	2.33	0.19
mgo	1.11	> 10	0.89	> 10	1.89	> 10	2.40	> 10
aceA	0.00	> 10	0.00	> 10	0.00	0.25	0.00	0.19
pyc	0.03	> 10	0.83	> 10	0.00	> 10	0.90	> 10
odx	0.00	> 10	0.00	> 10	0.00	> 10	0.00	> 10
ppc	0.33	> 10	0.00	0.68	0.00	> 10	0.00	2.27
pck	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mez	-0.44	> 10	0.00	> 10	-0.88	> 10	-0.07	> 10
gdh	0.21	0.09	1.36	> 10	1.40	> 10	0.38	0.08
aspB	0.21	0.09	0.56	0.09	0.60	0.09	0.38	0.08
lysC	-	-	0.44	0.09	0.48	0.09	-	-
dapA	-	-	0.32	0.09	0.36	0.09	-	-
dapDCE	-	-	0.00	> 10	0.00	> 10	-	-
ddh	-	-	0.32	> 10	0.36	> 10	-	-
lysA	-	-	0.29	0.09	0.33	0.09	-	-
bsLYS	0.21	> 10	-	-	-	-	0.38	0.09
pqo	-	-	0.34	> 10	0.10	> 10	-	-
pta	-	-	0.42	> 10	0.18	> 10	-	-
outTHF	-	-	29.80	> 10	0.04	0.00	-	-
inTHF	-	-	29.77	> 10	0.01	0.00	-	-

sprechende Ergebnisse sind in Abbildung 5.11 dargestellt. Hierbei fällt Folgendes auf:

1. Alle Histogramme weisen eine deutliche Abweichung von der Normalverteilungsform auf.
2. Für die Flüsse des EMP und PPP wurden vergleichsweise enge Verteilungen gefunden.
3. Die Flüsse des TCA weisen teilweise zweigipflige Verteilungen auf.
4. Innerhalb der Anaplerose konnte nur für das Malat-Enzym (mez) eine Normalverteilung geschätzt werden. Alle anderen Flüsse zeigten eine Gleichverteilung mit starken Tendenzen zu den oberen und unteren Rändern.

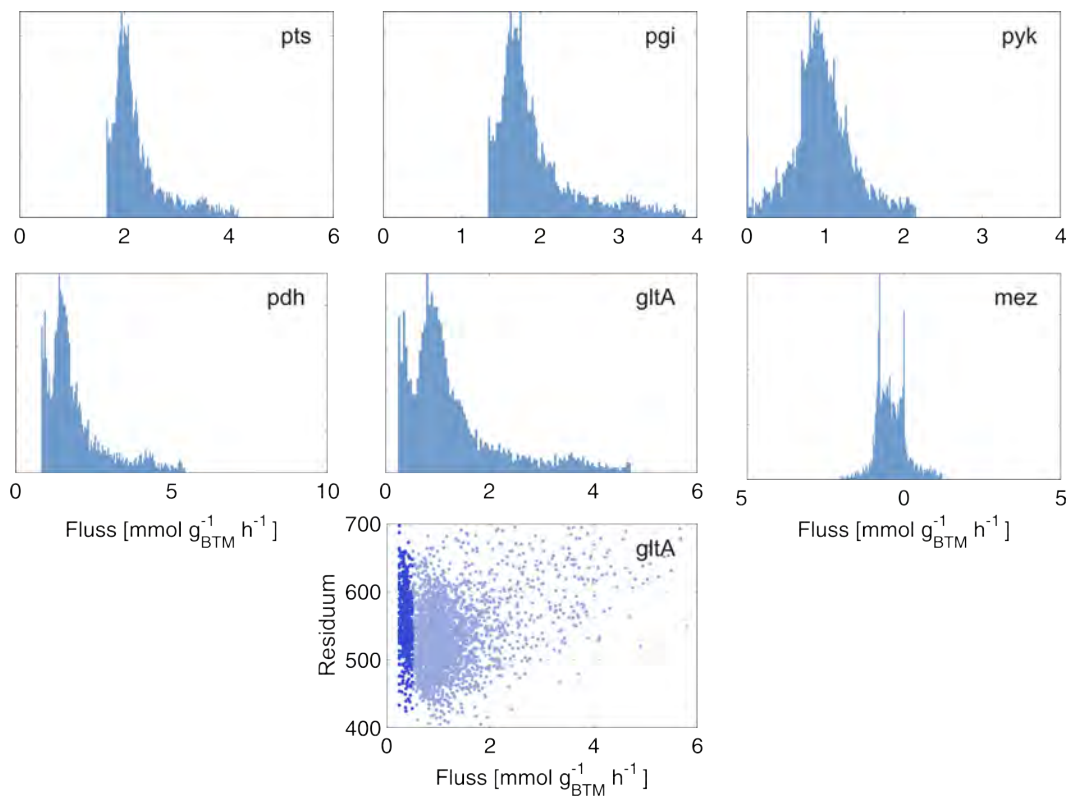


Abbildung 5.11: Histogramme der Bootstrap-Schätzungen ausgewählter intrazellulärer Flüsse des fokussierten Netzwerks im isotopisch stationären Modellierungsansatz.

Daraus lässt sich schlussfolgern, dass einerseits die im zugrunde liegenden Netzwerk vorhandenen Nichtlinearitäten einen nicht vernachlässigbaren Einfluss auf die Schätzung von Fluss-Konfidenzintervallen besitzen. Andererseits deuten die

mehrgipfligen Verteilungen auf das Vorhandensein mehrerer ausgeprägter lokaler Optima hin.

5.7.3 Methodenvergleich

In Tabelle 5.2 sind die Ergebnisse der Schätzung von Fluss-Konfidenzintervallen auf Basis des Linearisierungsansatzes und der Bootstrap-Methode gegenübergestellt. Im letzteren Fall wurden Mittelwerte $\hat{\mu}$ und Standardabweichungen $\hat{\sigma}$ mit Hilfe von Anpassungen entsprechender Normalverteilungsfunktionen an die Histogrammdata geschätzt (vgl. Anhang A.3.2). Darüber hinaus wurden die zugehörigen unsymmetrischen 90%-Konfidenzintervalle $\text{Conf}_{\hat{\theta}}(90)$ ermittelt.

Tabelle 5.2: Vergleich beider statistischer Ansätze zur Schätzung von Konfidenzbereichen ausgewählter Flüsse des fokussierten Netzwerks im isotopisch stationären Modellierungsansatz (^{13}C -MFA ZSW). Alle Angaben in $\text{mmol g}_{\text{BTM}}^{-1} \text{h}^{-1}$.

Reaktion	Linearisierungsansatz			Monte-Carlo-Simulation		
	$\hat{\mu}$	$\hat{\sigma}$	$\text{Conf}_{\hat{\theta}}(90)$	$\hat{\mu}$	$\hat{\sigma}$	$\text{Conf}_{\hat{\theta}}(90)$
feed _{GLC1}	0.14	0.02	[0.11, 0.17]	0.14	0.02	[0.11, 0.27]
feed _{GLCU}	1.49	0.16	[1.23, 1.76]	1.51	0.15	[1.28, 2.62]
feed _{GLC0}	0.37	0.04	[0.30, 0.43]	0.37	0.04	[0.31, 0.66]
pts	2.00	0.21	[1.65, 2.35]	2.02	0.20	[1.73, 3.51]
pgi	1.68	0.21	[1.33, 2.03]	1.70	0.20	[1.41, 3.19]
pfk	1.72	0.21	[1.37, 2.07]	1.74	0.20	[1.44, 3.23]
gapA	3.41	0.43	[2.71, 4.12]	3.46	0.39	[2.86, 6.44]
eno	3.16	0.43	[2.46, 3.86]	3.20	0.39	[2.61, 6.18]
pyk	0.71	> 10	-	0.90	0.16	[0.32, 1.72]
gnd	0.28	0.00	[0.28, 0.28]	0.28	0.00	[0.28, 0.28]
pdh	1.50	> 10	-	1.47	0.38	[0.91, 4.24]
gltA	0.92	0.42	[0.23, 1.61]	0.88	0.38	[0.33, 3.58]
icd	0.92	> 10	-	0.86	0.38	[0.28, 3.48]
odh	0.67	> 10	-	0.62	0.38	[0.03, 3.23]
fum	0.67	0.30	[0.18, 1.16]	0.63	0.27	[0.08, 3.33]
mqo	1.11	> 10	-	1.23	0.49	[0.09, 4.21]
aceA	0.00	> 10	-	0.00	0.00	[0.00, 0.00]
mez	-0.44	> 10	-	-0.50	0.34	[-1.07, 0.54]
exc _{CO2}	2.78	1.27	[0.70, 4.85]	2.66	1.13	[1.00, 10.77]

Die auf Grundlage beider Methoden geschätzten Mittelwerte und Standardabweichungen stimmen für die Flüsse des EMP und PPP sehr gut überein. Darüber hinaus konnten mit Hilfe der Monte-Carlo-Simulationen Standardabweichungen für einen Teil der Flüsse im TCA ermittelt werden, welche auf Grundlage der linearisierten Statistik nicht identifizierbar waren. Im Fall der berechneten 90%-Konfidenzintervalle fällt die bereits angesprochene Unter- bzw. Überschätzung der unteren und oberen

Intervallgrenzen durch die Linearisierung auf (vgl. Abb. 5.10).

5.8 Quantifizierbarkeit intrazellulärer Stoffflüsse

Aus den vorangegangenen Abschnitten wird deutlich, dass unter Berücksichtigung der statistischen Güte aus beiden Modellierungsansätzen (isotopisch (in)stationär) sowie dazugehöriger Netzwerkvarianten (ZSW, ZSW_AS) teilweise beträchtlich unterschiedliche Stoffflussverteilungen resultieren. Neben der Gewinnung dieser relativen Informationen ist das eigentliche Ziel der ^{13}C -MFA jedoch die absolute Quantifizierung intrazellulärer Stoffflüsse. Im Vergleich aller Ansätze sind diesbezüglich folgende Punkte festzuhalten:

1. Die quantitativen Stoffflusswerte sind direkt von der Netzwerkgröße abhängig, d.h. Art und Anzahl der stöchiometrischen Kopplungen. Dies ist anhand der beiden Netzwerkvarianten des stationären Ansatzes erkennbar, bei denen große Unterschiede im fokussierten Netzwerk des Zentralstoffwechsels im Vergleich zum erweiterten Netzwerk mit Aminosäurebiosynthesewegen auftreten (vgl. Tab. 5.1). In beiden Fällen ist das Verhältnis zwischen der Anzahl an Freiheitsgraden und der zur Verfügung stehenden Messinformation ähnlich (Tab. 5.3).
2. Darüber hinaus hat auch die Wahl des Modellierungsansatzes einen erheblichen Einfluss auf die Bestimmung der Stoffflüsse. Zwar wurde eine vergleichbare Verteilung der Stoffflüsse im PPP und oberen Teil des EMP auf Basis des fokussierten Netzwerks gefunden, jedoch sind die zugehörigen Flusswerte in der Lösung des instationären Modells um das 1.5-fache höher. Im TCA beträgt die Abweichung sogar mehr als das Doppelte.

Ursache hierfür sind einerseits die unterschiedlichen Anpassungen der extrazellulären Flüsse, d.h. im zweiten Fall fließt deutlich mehr über die Glucoseaufnahme in das System hinein (Faktor 1.5) und in Form der Produkte CO_2 und Lysin wieder hinaus (Tab. 5.4).

Andererseits ist die verfügbare Messinformation zur Schätzung aller Modellparameter im instationären Ansatz um ein Vielfaches höher, wodurch insgesamt die statistische Sicherheit der Flusschätzung ansteigt (Tab. 5.1 und 5.3).

Damit stellt sich natürlich die grundsätzliche Frage, welche der ermittelten Stoffflusslagen am ehesten der Realität des zugrunde liegenden metabolischen Zustandes der Zelle entspricht. Im folgenden Kapitel wird auf eine mögliche Vorgehens-

weise zur Verifizierung der erhaltenen quantitativen Stoffflussdaten eingegangen.

Tabelle 5.3: Statistische Kenngrößen und zugrunde liegende Messdaten aller Modellansätze.

	¹³ C-MFA ZSW MM1	¹³ C-MFA ZSW_AS MM1	¹³ C-MFA ZSW_AS MM2	INST ¹³ C-MFA ZSW MM1
Metabolitpools	22	66		22
Isotopomere	746	13814		746
Reaktionen	44	113		44
Modellparameter ^a	23	30		45
Probenahmezeitpunkte	5			20
Poolgrößenmessungen	-			22
Markierungsmessungen	65		102	1300
Flussmessungen ^b	16	36		16

^aParameteranzahl = freie Flüsse + Poolgrößen

^beinschließlich Biomasseabflüsse

Tabelle 5.4: Schätzwerte extrazellulärer Raten des Prozessmodells sowie aller Ansätze zur ¹³C-MFA. Alle Angaben in mmol g_{BTM}⁻¹ h⁻¹.

Rate	Prozess	¹³ C-MFA ZSW MM1	¹³ C-MFA ZSW_AS MM1	¹³ C-MFA ZSW_AS MM2	INST ¹³ C-MFA ZSW MM1
feed _{GLC0}	0.29 ± 0.08	0.37 ± 0.04	0.38 ± 0.05	0.40 ± 0.04	0.81 ± 0.04
feed _{GLC1}	0.22 ± 0.06	0.14 ± 0.02	0.20 ± 0.04	0.20 ± 0.02	0.24 ± 0.01
feed _{GLCU}	1.97 ± 0.55	1.49 ± 0.16	1.75 ± 0.17	1.76 ± 0.16	1.98 ± 0.09
exc _{LYS}	0.29 ± 0.09	0.21 ± 0.09	0.25 ± 0.09	0.29 ± 0.09	0.38 ± 0.08
exc _{CO₂}	6.48 ± 1.59	2.78 ± 1.27	4.89 ± 1.29	4.89 ± 1.28	7.92 ± 0.84

Kapitel 6

Interpretation geschätzter Stoffflüsse

Durch die Verwendung eines „genomweiten“ stöchiometrischen Stoffwechselmodells, in welchem die Mehrzahl relevanter biologischer Nebenbedingungen (z.B. NADPH-Bilanzen, ATP-Bedarf usw.) enthalten sind, wurden die zuvor erzeugten quantitativen Fluxom-Daten interpretiert und überprüft.

Dazu wurde ein genomweites Modell für *C. glutamicum* aus der Literatur entnommen und im Rahmen dieser Arbeit für die Durchführung verschiedener Simulationsstudien modifiziert.

6.1 Genomweites Netzwerkmodell von *C. glutamicum*

Die im letzten Jahrzehnt rasant fortschreitende Entwicklung technischer Methoden im Bereich der Molekularbiologie hat dazu geführt, dass eine ganze Reihe von Genomen unterschiedlicher Organismen sequenziert und annotiert wurden. Auf Grundlage dieser Informationen und zusätzlicher experimenteller Befunde wurde anschließend damit begonnen, genomweite stöchiometrische Modelle zu rekonstruieren.

Gegenwärtig existieren bereits genomweite Modelle für eine Vielzahl von Spezies wie *Haemophilus influenzae* [SP00], *Helicobacter pylori* [SCF⁺02], *Escherichia coli* [RVSP03], *Saccharomyces cerevisiae* [FFF⁺03], *Lactococcus lactis* [ONF05], *Streptomyces coelicolor* [BKN05], *Staphylococcus aureus* [HKRP05], *Methanosarcina barkeri* [FSP⁺06], *Lactobacillus plantarum* [TWM⁺06], *Bacillus subtilis* [OPP⁺07] und *Homo sapiens* [DBJ⁺07]. In den aufgeführten Literaturquellen sind ausführliche Informationen zur Vorgehensweise bei der Rekonstruktion solcher Modelle enthalten, weshalb an dieser Stelle nicht näher darauf eingegangen werden soll.

Kürzlich wurde auch für den Organismus *C. glutamicum* ein genomweites metabolisches Netzwerkmodell publiziert [KN09], welches auf dem annotierten Genom des Wildtypstammes ATCC 13032 beruht [KBB⁺03]. Gemäß den Angaben der Autoren wurde dieses Modell gegen Literaturdaten verschiedener experimenteller Zustände (Biomassewachstum, unterschiedliche C-Quellen usw.) validiert.

Netzwerkmodifikationen

Neben der Aufbereitung des publizierten Netzwerks in einer für Simulationszwecke nutzbaren Form, wurde dieses unter Nutzung zusätzlicher Quellen weiterentwickelt. Am Ausgangsmodell wurden dabei folgende Änderungen vorgenommen:

1. Durch einen direkten Vergleich mit dem bisher größten *E. coli*-Netzwerk [FHR⁺07] wurden alle Reaktionen um die Bilanzierung von Protonen (H) und Wasser (H₂O) erweitert.
2. Gleichzeitig wurde die dem *E. coli*-Modell zugrunde liegende Notation übernommen, welche eine Unterscheidung zwischen cytosolischen (z.B. h[c]) und extrazellulären (z.B. h[e]) Komponenten vornimmt. Damit ergeben sich zwangsläufig Konsequenzen in der Beschreibung der Reaktionen, welche an den Transport extrazellulärer Protonen gekoppelt sind (vgl. Tab. 6.1).
3. Reaktionen mit Produkten, die nicht weiter verstoffwechselt werden („Dead-Ends“), wurden aus dem Modell entfernt.
4. Im Modell wurden zusätzliche Reaktionen berücksichtigt, welche die Bildung und Degradation von Glycogen beschreiben [SDSE07].
5. Einige Irreversibilitätsannahmen wurden auf Grundlage anderer Literaturinformationen bzw. des aktuellen Kenntnisstandes geändert.

Das auf diese Weise erstellte genomweite Netzwerkmodell umfasst 399 Metabolite und 454 Reaktionen und ist damit gegenwärtig eines der größten verfügbaren stöchiometrischen Modelle für *C. glutamicum*.

In Abbildung 6.1 ist ein Ausschnitt des erstellten Netzwerkmodells dargestellt. Aufgrund des großen Umfangs des Netzwerks wird auf eine vollständige Darstellung der Reaktionsgleichungen innerhalb dieser Arbeit verzichtet.

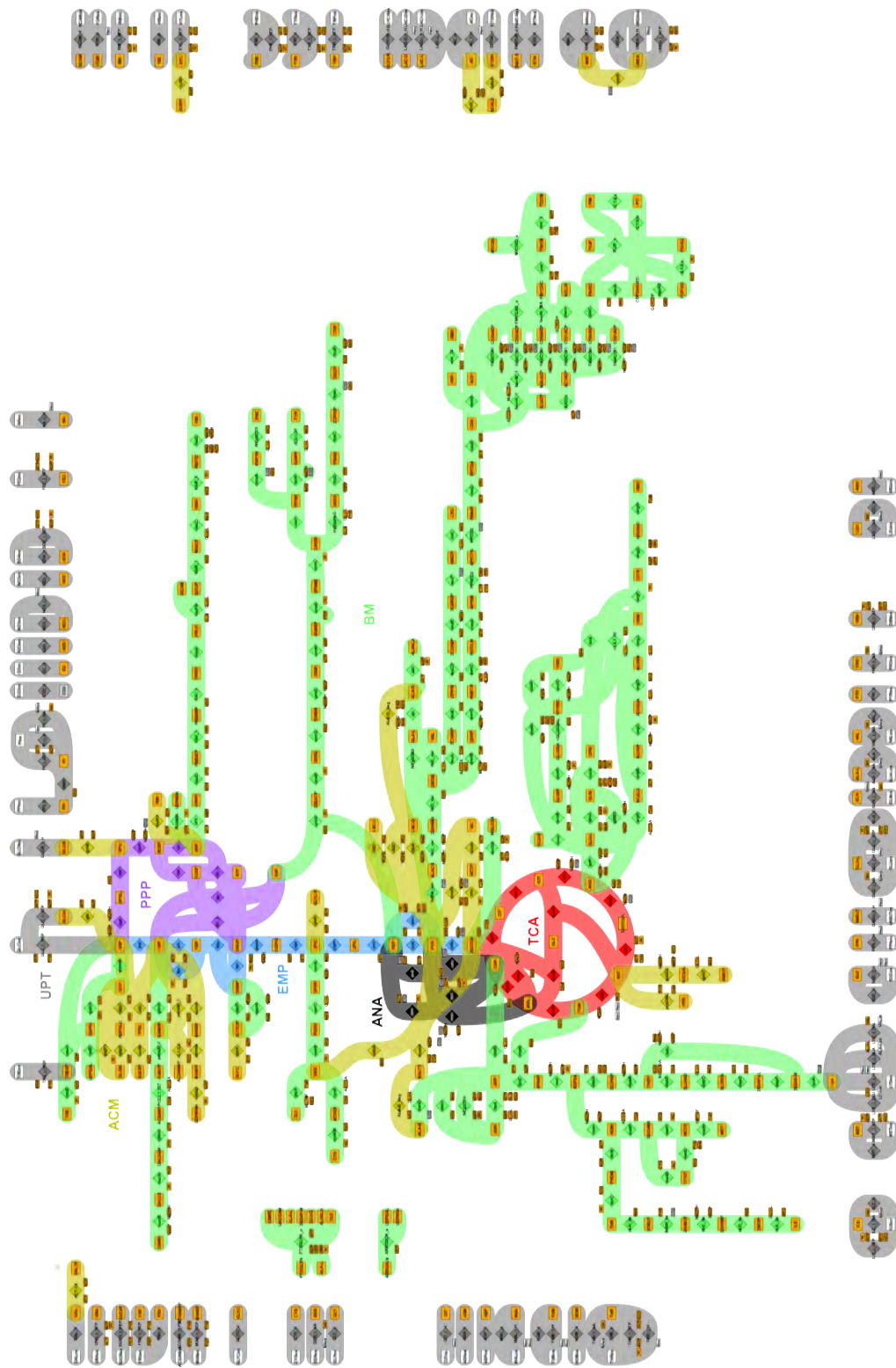


Abbildung 6.1: Ausschnitt des modifizierten genomweiten Netzwerkmodells von *C. glutamicum*.

Tabelle 6.1: Vergleich zwischen originalem (schwarz) und geänderten (rot) Modellansatz anhand ausgewählter Reaktionen.

Reaktion	Stöchiometrische Gleichung
Succinat-Dehydrogenase (sdh)	SUCC + MK → FUM + MKH2 $\text{suc}[c] + \text{mq}[c] + 2 \text{h}[e] \rightleftharpoons \text{fum}[c] + \text{mqh2}[c] + 2 \text{h}[c]$
Elektronentransport (cyto-bd-complex)	MKH2 + 0.5 O2 + 2 H_PO → MK + 2 H_POxt $\text{mqh2}[c] + 0.5 \text{o2}[c] + 2 \text{h}[c] \rightarrow \text{mq}[c] + 2 \text{h}[e] + \text{h2o}[c]$
ATP-Synthese (ATPase-complex)	ADP + PI + 4 H_POxt → ATP + 4 H_PO $\text{adp}[c] + 4 \text{h}[e] + \text{pi}[c] \rightarrow \text{atp}[c] + \text{h2o}[c] + 3 \text{h}[c]$
Lysintransport (lysE)	LYS + 2 H_transport_xt → LYSxt + 2 H_transport $\text{lys}[c] + 2 \text{h}[e] \rightarrow \text{lys}[e] + 2 \text{h}[c]$
Hilfsreaktion (Proton-ATP)	H_transport_xt + ADP + PI ⇌ H_transport + ATP $_{-a}$

^aDiese Reaktion kann entfallen, da im geänderten Netzwerk eine direkte Verbindung zwischen extrazellulären Protonen und der oxidativen Phosphorylierung (Atmungskette) besteht.

6.2 Flussbilanzanalyse

In den nachfolgenden Simulationsstudien wurde die Methode der FBA zur Lösung des genomweiten *C. glutamicum*-Modells genutzt, wobei als Optimierungskriterium entweder Biomassewachstum oder Lysinbildung definiert wurde (vgl. Abschn. 3.3.4). Alle Simulationen wurden mit MATLAB (Mathworks, R2008b) unter Verwendung der COBRA-Toolbox sowie des Lösungsverfahrens GLPK [BFM⁺07] durchgeführt.

6.2.1 Vergleich optimaler Stoffflusslagen

Um der Frage nachzugehen, ob die mittels ¹³C-MFA ermittelten Stoffflussverteilungen aus Kapitel 5 im Kontext des genomweiten Netzwerks sinnvolle Lösungen darstellen, wurde das Modell unter Festlegung einiger Nebenbedingungen sowie des Biomassewachstums als Zielfunktion optimiert.

In einem ersten Schritt wurde der stationäre Lösungsraum untersucht, der sich in Abhängigkeit der gemessenen extrazellulären Flüsse ergab (Abb. 6.2). Auf Grundlage des Netzwerkes für den Stamm DM1730, d.h. Deletion im Enzym PEP-Carboxykinase (pck), wurden der Reihe nach weitere Deletionen in Reaktionen eingeführt, für die auch im Rahmen der Stoffflussanalyse je nach Modellvariante kein Fluss bestimmt wurde.

Da in allen Fällen dieselbe optimale Wachstumsrate von $\mu = 0.195 \text{ mmol g}_{\text{BTM}}^{-1} \text{ h}^{-1}$ berechnet wurde und diese im Rahmen der Messgenauigkeit der experimentell er-

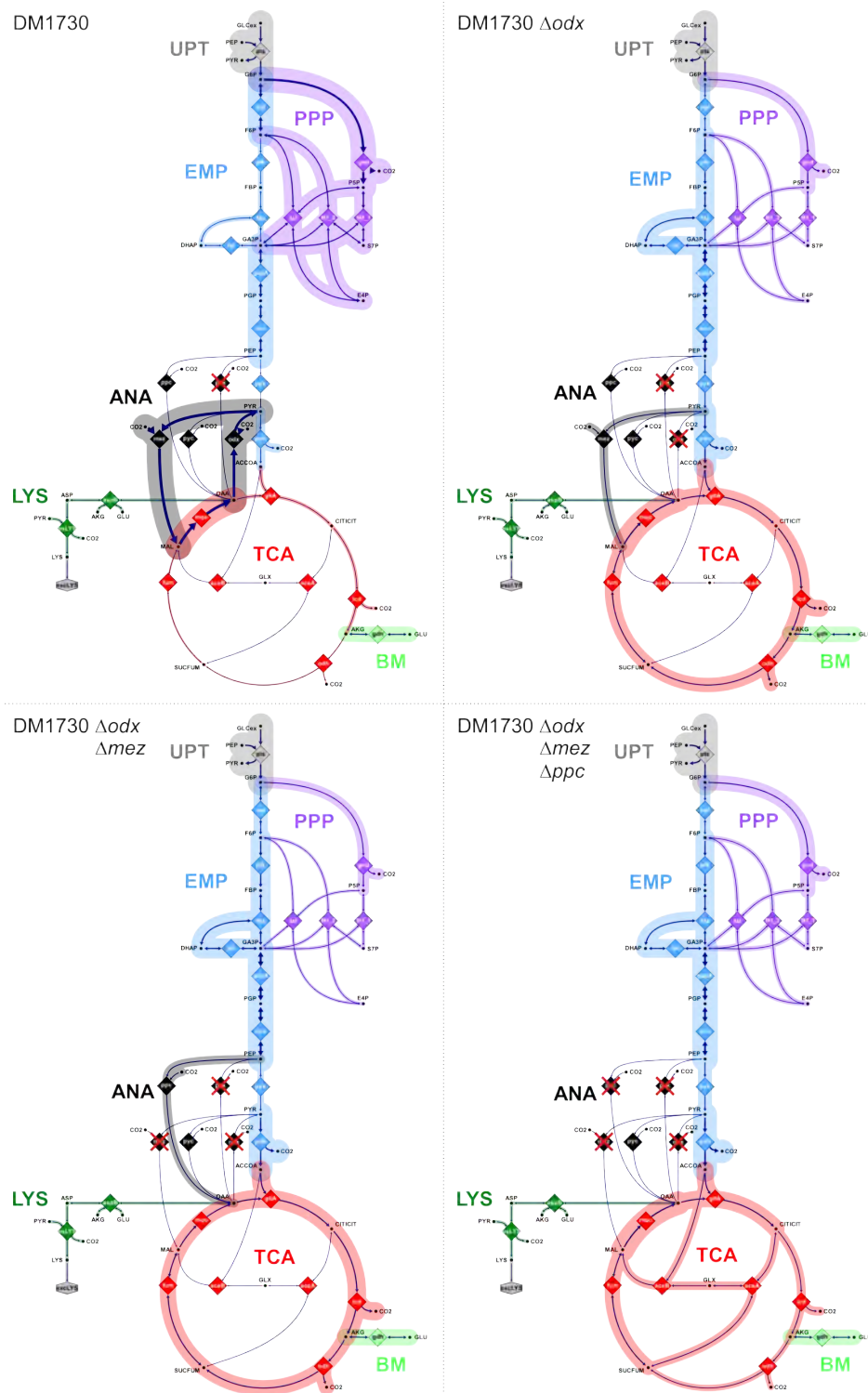


Abbildung 6.2: Optimale Stoffflusslagen verschiedener Deletionsmutanten von *C. glutamicum* DM1730 auf Basis des genomweiten Netzwerkmodells. Als Zielkriterium wurde das Biomassewachstum gewählt. Die extrazellulären Flüsse (Glucoseaufnahme, Lysin- und CO_2 -Produktion) wurden in allen Fällen auf die gemessenen Raten des Markierungsexperimentes fixiert. Die Bildung weiterer Nebenprodukte wurde ausgeschlossen.

mittelten Rate von $\mu = 0.198 \pm 0.028 \text{ mmol g}_{\text{BTM}}^{-1} \text{ h}^{-1}$ entspricht, folgt daraus zunächst:

- Das genomweite Netzwerk macht sinnvolle Prädiktionen im Hinblick auf das exponentielle Biomassewachstum bei Kenntnis der extrazellulären Flüsse.
- Wie bereits angedeutet, gibt es mehrere Varianten intrazellulärer Stoffflussverteilungen bei gleichen Kohlenstoffflüssen über die Systemgrenzen. Dies liegt vor allem an der Vielzahl paralleler Reaktionswege (z.B. Anaplerose).
- Es existiert eine indirekte Proportionalität zwischen den Flüssen des PPP und TCA. Da insbesondere der Fluss in den PPP via 6-Phospho-Gluconat-Dehydrogenase (gnd) in beiden Modellierungsansätzen der ^{13}C -MFA sehr gut bestimmt wurde, besteht hier ein Ansatzpunkt für einen Plausibilitätstest.
- Da keine metabolische Regulation im Modell enthalten ist, werden hier biologisch fragwürdige Lösungen erzeugt (z.B. hohe Aktivität des Glyoxylatshunts bei Wachstum auf Glucose).

Auf Basis dieser Ergebnisse wurden anschließend gezielt mögliche Phänotypen ermittelt, indem jeweils ein Teil der in Abhängigkeit beider Modellierungsansätze sowie Netzwerk- und Messmodellvarianten ermittelten Stoffflüsse im Rahmen ihrer statistischen Genauigkeit variiert wurde.

Für den Fall, dass eine mittels ^{13}C -MFA gefundene Stoffflusslage unter Berücksichtigung gewisser Toleranzen nicht innerhalb des zugehörigen stationären Lösungsraumes des genomweiten Modells liegt, wäre diese dann unter den gegebenen Voraussetzungen zu verwerfen.

Im Ergebnis dieser Simulationsstudien lassen sich folgende Schlussfolgerungen für die geschätzten Stoffflusslagen der ^{13}C -MFA ziehen (Abb. 6.3):

1. Die auf Grundlage des isotopisch stationären Ansatzes im fokussierten Netzwerkmodell erzeugte Stoffflusslage stellt keinen gültigen Phänotyp dar, d.h. Biomassewachstum ist unter diesen Bedingungen ($\mu = 2.00 \pm 0.21$, $\text{gltA} = 0.92 \pm 0.42 \text{ mmol g}_{\text{BTM}}^{-1} \text{ h}^{-1}$) nicht möglich. Darüber hinaus werden bei höheren Flüssen über die Citrat-Synthase deutlich geringere Wachstumsraten ($\mu \leq 0.157 \text{ mmol g}_{\text{BTM}}^{-1} \text{ h}^{-1}$) simuliert als im Markierungsexperiment ermittelt ($\mu = 0.198 \pm 0.028 \text{ mmol g}_{\text{BTM}}^{-1} \text{ h}^{-1}$). Somit ist dieses Stoffflussanalyse-Ergebnis sehr fragwürdig und sollte verworfen werden.

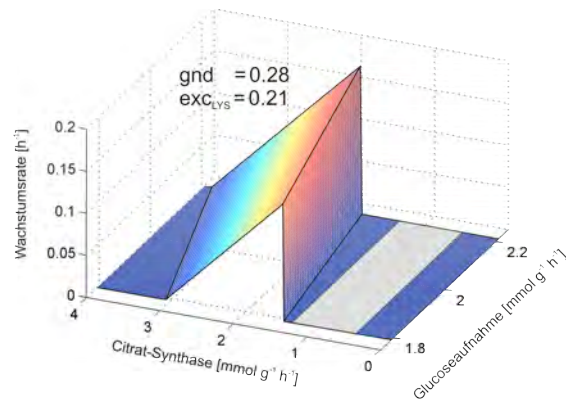
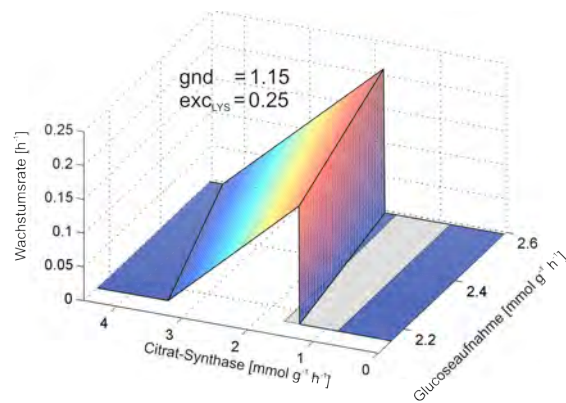
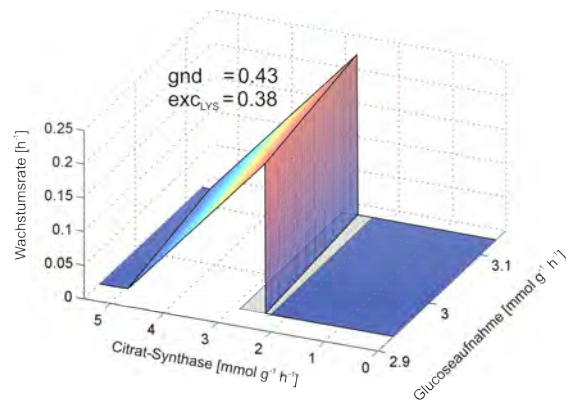
(a) ¹³C-MFA ZWS MM1(b) ¹³C-MFA ZWS_AS MM2(c) INST-¹³C-MFA ZWS MM1

Abbildung 6.3: Simulierte Phänotypen für *C. glutamicum* DM1730 auf Basis des genomweiten Netzwerkmodells. In allen Fällen wurde der Fluss in den PPP (g_{nd}) sowie die Lysinbildungsrate (exc_{LYS}) auf die geschätzten Werte der ¹³C-MFA fixiert (a-c). Die Glucoseaufnahmerate (pts) wurde innerhalb ihres Konfidenzbereiches variiert, für den Fluss in den TCA über Citrat-Synthase ($gltA$) wurde der zugehörige Wachstumsbereich abgedeckt. Die grauen Bereiche kennzeichnen die ermittelten Konfidenzbereiche für den Citrat-Synthase-Fluss auf Basis des Linearisierungsansatzes.

2. Ein etwas anderes Bild ergibt sich für die Lösungen des erweiterten Netzwerks, die im Rahmen ihrer statistischen Spannweite (insbesondere auch unter Berücksichtigung eines möglicherweise höheren PPP-Flusses, $gnd_{\max} = 1.34 \text{ mmol g}_{\text{BTM}}^{-1} \text{ h}^{-1}$) phänotypisch zulässig sind.
3. Auch im Fall des isotopisch instationären Ansatzes liegt der Großteil möglicher Stoffflusslagen im gültigen Bereich.

Letztendlich kann auf diese Weise eine Lösung der ^{13}C -MFA ausgeschlossen werden, jedoch verbleiben zwei sehr verschiedene Stoffflussverteilungen als mögliche Beschreibungen des metabolischen Zustandes der Zelle.

6.2.2 Maximierung der Lysinproduktion

Unter der Maßgabe der Richtigkeit ermittelter intrazellulärer Stoffflusswerte kann mit Hilfe der Flussbilanzanalyse untersucht werden, inwieweit sich die Lysinproduktion unter den gegebenen Umständen (Substrataufnahme, Wachstum) theoretisch steigern lässt.

In Abbildung 6.4 sind die Ergebnisse einer solchen Simulationsstudie dargestellt, wobei zu Vergleichszwecken die Stoffflusswerte aus dem isotopisch instationären Ansatz herangezogen wurden.

Es wird deutlich, dass die Lysinproduktionsleistung im experimentellen Zustand weit vom theoretischen Optimum entfernt liegt. Aufgrund der bereits erwähnten indirekten Proportionalität zwischen dem PPP und TCA-Fluss führen alle Kombinationen aus höherem Fluss in den PPP und gleichzeitig niedrigerem Fluss in den TCA in Richtung Maximierung des Lysinertrages.

Dabei spielt natürlich die jeweils zugrunde liegende optimale Stoffflussverteilung eine wesentliche Rolle (Abb. 6.4, oben). Ein hoher Fluss in den PPP ist gleichzeitig an eine hohe NADPH-Bildung gekoppelt. Ein Teil des NADPH wird für die Aminosäuresynthese (einschließlich Lysin) benötigt, jedoch bleibt ein erheblicher Überschuss bestehen, der wieder recycled werden muss.

Dies wäre prinzipiell über das Malat-Enzym (mez) möglich, wodurch gleichzeitig eine hohe Substratmenge in Form von Malat für das Atmungskettenenzym Menachinon-Oxireductase (mqo) zur Verfügung stünde. Die damit erzeugte Energiemenge reicht aus, um die Wachstumsanforderungen zu erfüllen ohne dass ein zusätzlicher Fluss über den TCA notwendig ist. Die Reaktion Oxalacetat-Decarboxylase (odx) schließt den Kohlenstoffkreislauf, bei dem durch den gleichzeitigen Ein- und

Ausbau von CO₂ kein Kohlenstoff verloren geht. Insgesamt könnte somit unter diesen Wachstumsbedingungen der Lysinertag nahezu verdreifacht werden.

Allerdings stellt sich die biologische Realität aufgrund der vielfältigen Regulationen innerhalb der Zelle (vgl. Kapitel 3.1) oft deutlich anders dar und Prädiktionen in der obigen Form erweisen sich in der Praxis als nicht umsetzbar.

Um diesbezüglich ein realitätsgetreueres Abbild des biologischen Systems zu erhalten, sollten kinetische Modelle formuliert werden, welche auch eine Beschreibung von Regulationsmechanismen beinhalten. Insofern diese Modelle wiederum mit Hilfe von Messdaten validierbar sind, können weitaus genauere Vorhersagen über potentielle Stammkonstrukte zur Produktionssteigerung getroffen werden.

Auf eine solche Vorgehensweise wird im nachfolgenden Kapitel für den Organismus *C. glutamicum* näher eingegangen.

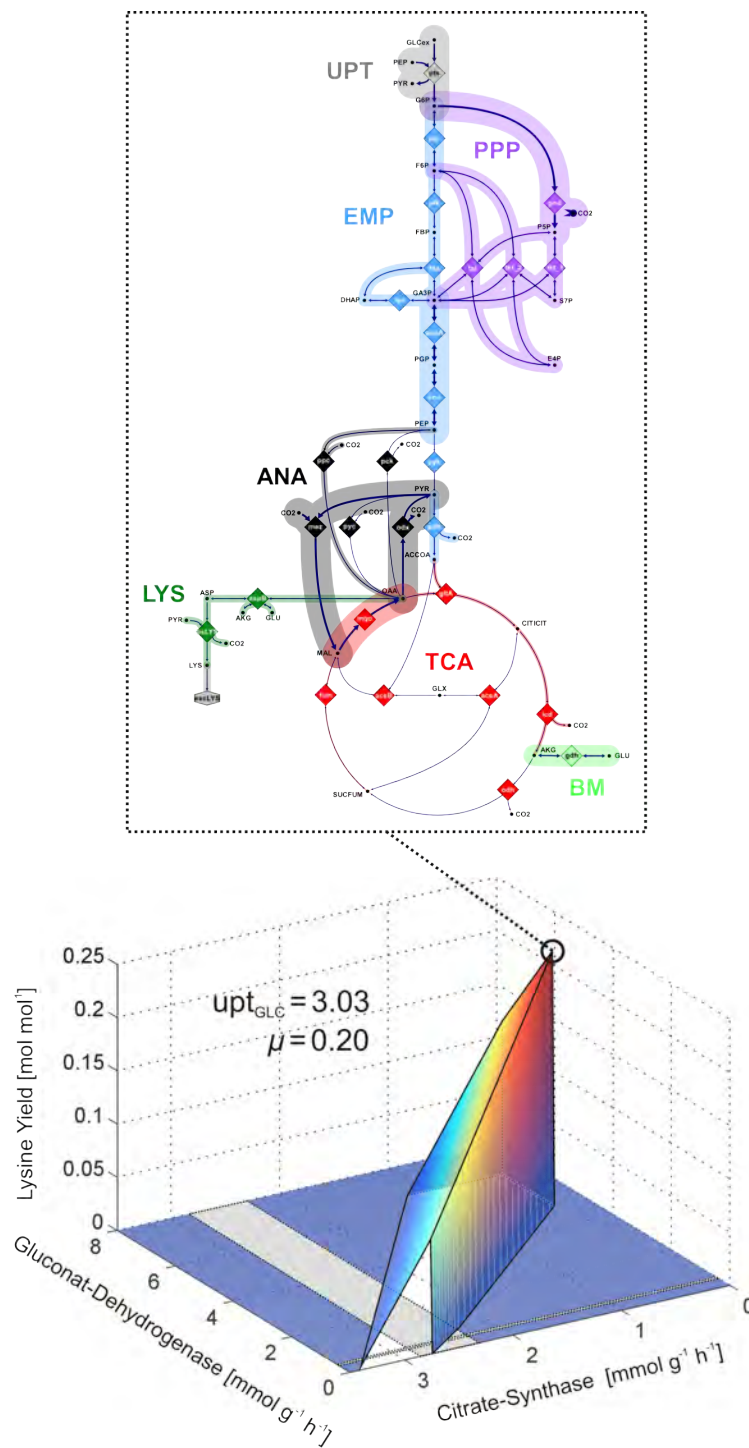


Abbildung 6.4: Maximierung der Lysinproduktion für *C. glutamicum* DM1730. Unten: Maximale Lysinertäge bei verschiedenen Kombinationen von Zuflüssen in den PPP und TCA sowie einer festen Substrataufnahme- und Wachstumsrate. Die grauen Bereiche kennzeichnen die aus der instationären ¹³C-MFA ermittelten Konfidenzbereiche des PPP und TCA Zuflusses. Oben: Optimale Stoffflussverteilung im Fall des theoretisch maximalen Lysinertäges.

Teil III

Integrative mechanistische Modellierung

Kapitel 7

Gewinnung quantitativer Multi-Omics-Daten

Mit dem Ziel einer integrativen Analyse von Multi-Omics-Daten wurden verschiedene *C. glutamicum*-Stämme im Rahmen von kontinuierlichen Kultivierungen (Chemostat) untersucht. Hierbei wurden Proben zur Analyse des Transkriptoms, Proteoms und Metaboloms generiert. Ein Teil der daraus resultierenden Datensätze bildete später die Grundlage zur Validierung eines vertikalen Netzwerkmodells für den TCA von *C. glutamicum*.

In den nachfolgenden Abschnitten wird zunächst auf die experimentelle und analytische Vorgehensweise zur Gewinnung dieser Daten eingegangen. Anschließend werden sie im Hinblick auf Stamm-, Wachstums- und Produktionscharakteristika miteinander verglichen.

7.1 Bioprozess- und Analysemethoden

7.1.1 Stämme und Kultivierungsbedingungen

In Tabelle 7.1 sind die mit den verschiedenen *C. glutamicum*-Stämmen durchgeführten Chemostat-Experimente sowie resultierende Omics-Datensätze aufgeführt. Neben dem Wildtyp und dem bereits erwähnten Stamm DM1730 (Kapitel 5.3.1) wurde ein zweiter Lysinproduzent (DM1933) untersucht, welcher durch weiterführendes Stammdesign seitens des Industriepartners Evonik Degussa GmbH erzeugt wurde. Die Herstellung von Kryo- und Vorkulturen sowie die verwendete Medienzusammensetzung entspricht der in Abschnitt 5.3.1 beschriebenen Art und Weise. Die Kul-

tivierungen der Stämme WT und DM1730 erfolgten im 20 l-Bioreactor (ATGU 503, Infors AG, Schweiz) mit einem Arbeitsvolumen von 6.8 l. Der Stamm DM1933 wurde hingegen in einer 1.5 l-Sixfors-Anlage (SixFors, Infors AG, Schweiz) mit Arbeitsvolumina von 1 l in jeweils 3 technischen Replika je Wachstumsrate kultiviert.

Die Kultivierung wurde im Batchbetrieb gestartet und danach, vor dem Einsetzen einer Glucoselimitation, in den Chemostatbetrieb umgeschaltet. Das Kulturvolumen wurde dabei durch einen konstanten Zufluss frischen Mediums mit 30 g l⁻¹ Glucose und einem Abfluss von Biomasse und Produkt konstant gehalten. Die Regelung und Erfassung prozesstechnischer Größen (*pH*, *T* usw.) erfolgte analog zum ¹³C-Markierungsexperiment.

Nach ca. 4-5 Verweilzeiten (Einstellung des stationären Betriebes) wurden Proben zur Analyse des Transkriptoms, Proteoms und Metaboloms entnommen. Für die Transkriptomanalyse wurden jeweils 5 ml Probe direkt in flüssigen Stickstoff gefroren. Zur Gewinnung von Rohextrakten für Enzymaktivitätsmessungen wurden 5 ml Probe abzentrifugiert, in 5 ml Puffer gewaschen, erneut zentrifugiert und anschließend das Pellet in flüssigem Stickstoff gefroren.

Für die Gewinnung quantitativer Metabolom-Daten wurden je Kultivierung 4 Proben à 5 ml in 15 ml kalter Methanollösung (-50°C, 60% v v⁻¹) gequencht und danach bei -25°C abzentrifugiert. Alle Proben wurden anschließend bei -80°C gelagert.

Tabelle 7.1: Übersicht der Chemostatexperimente verschiedener *C. glutamicum*-Stämme zur Generierung quantitativer Multi-Omics-Datensätze.

Stamm	Genotyp	μ [h ⁻¹]	Transkriptom	Proteom	Metabolom
WT	Wildtyp ATCC13032 (Biotin-auxotroph)	0.1	x	x	x
		0.2	x	x	-
DM1730	$\Delta pck pyc(P458S) hom(V59A)$ <i>lysC(T311I)</i>	0.1	x	x	x
		0.2	x	x	x
DM1933	$\Delta pck pyc(P458S) hom(V59A)$ 2x[<i>lysC(T311I) asd dapA</i> <i>dapB ddh lysA lysE</i>]	0.05 ^a	-	x	x
		0.1 ^a	-	x	x
		0.2 ^a	-	x	x

^ajeweils in 3 technischen Replikaten

7.1.2 Genomweite Transkriptom-Daten

Die im Zuge der Chemostatkultivierungen mit WT und DM1730 erzeugten Proben zur Transkriptomanalyse wurden an den Industriepartner Evonik Degussa GmbH weitergeleitet, wo mit Hilfe von Affymetrix Chips (Affymetrix, Inc.) genomweite

normierte mRNA-Daten generiert wurden. Das zugehörige experimentelle Setup ist in Abbildung 7.1 dargestellt.

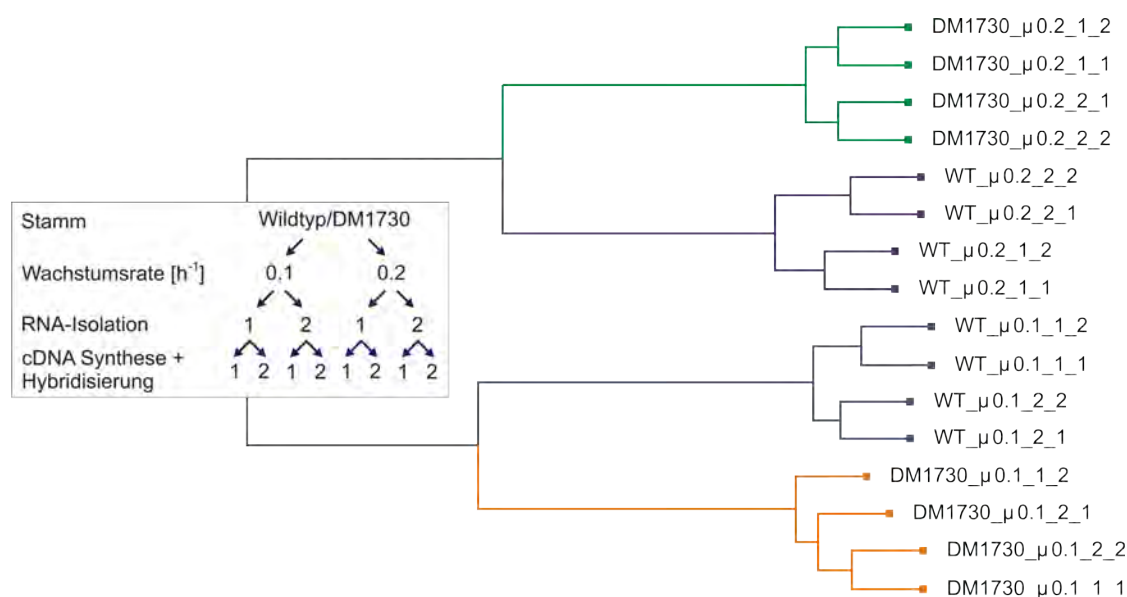


Abbildung 7.1: Analyse des Transkriptoms mittels Affymetrix Chips. Links: Je Stamm und Wachstumsrate wurden 2 unabhängige RNA-Isolierungen durchgeführt. Anschließend wurde jede RNA in 2 unabhängigen Ansätzen in cDNA umgeschrieben und hybridisiert. Somit ergeben sich insgesamt 16 Chips, welche Daten über absolute Transkriptmengen der analysierten Gene liefern. Rechts: Hierarchische Clusteranalyse der Transkriptom-Daten. Neben der Bezeichnung des Stammes und der Wachstumsrate kennzeichnen die nachfolgenden Ziffern die unabhängigen RNA-Isolierungen und cDNA-Synthesen.

Normiert bedeutet in diesem Zusammenhang, dass keine direkten mRNA-Konzentrationen gemessen wurden, sondern auf die gesamte RNA normierte Intensitäten einzelner mRNA-Transkripte.

Aus den 4 durchgeführten Experimenten resultierten bei einer Gesamtheit von 3571 getesteten Gensequenzen insgesamt 14284 normierte mRNA-Intensitätsmessungen.

Qualität der Transkriptom-Daten

Zur weiterführenden statistischen Auswertung der Daten wurden zunächst aus der Gesamtheit aller getesteten Gensequenzen 3045 extrahiert, welche über eine Annotation verfügen. Aus dieser Untermenge wurden wiederum nur diejenigen Gene verwendet, für die eine Transkription in allen 16 Chips nachgewiesen wurde. Diese zusätzliche Einschränkung ist notwendig, da im Falle eines fehlenden Transkriptionswertes keine genaue Aussage darüber getroffen werden kann, ob es sich um eine

Falschmessung oder um einen nicht nachweisbaren Wert (unteres Detektionslimit) handelt.

Für die somit erhaltenen Daten wurde eine hierarchische Clusteranalyse unter Nutzung der Software Genedata Expressionist (Genedata GmbH, Martinsried) durchgeführt (vgl. Abb. 7.1, rechts). Es wird deutlich, dass mit Ausnahme des Experiments *DM1730_μ 0.1* alle Ansätze des gleichen Stammes und gleicher Wachstumsrate in Abhängigkeit der RNA-Isolierung und darüber hinaus bezüglich der jeweiligen cDNA-Synthese clustern. Dies beweist eine gute Reproduzierbarkeit in der Herstellung eines Affymetrix-Chips ausgehend von einer einzelnen Transkriptomprobe.

Da keine Mehrfachmessungen aus technischen Replika zur Verfügung standen, wurde für die weiterführende Auswertung der Transkriptomdaten entsprechend den Angaben des Kooperationspartners Evonik ein Messfehler von 10% angenommen.

7.1.3 Proteom-Daten aus Enzymaktivitätsmessungen

In allen Chemostat-Experimenten wurden Proben zur Messung von Enzymaktivitäten des TCA und der Anaplerose genommen. Entsprechende Messungen erfolgten hierbei von den Kooperationspartnern des SysMAP-Projektes (vgl. Abschn. 1.3). Für nähere Informationen in Bezug auf die verwendeten Protokolle der Enzymassays wird an dieser Stelle auf die Literatur verwiesen [EB05].

Eine Einheit der Enzymaktivität (Unit, U) ist als die Enzymmenge definiert, welche einen Stoffumsatz von $1 \mu\text{mol min}^{-1}$ katalysiert. Zum Zwecke der Vergleichbarkeit wurden alle Aktivitäten auf die gleichzeitig ermittelte Gesamtproteinmenge der jeweiligen Probe bezogen.

Insgesamt standen damit 13 Datensätze mit Enzymaktivitätsmessungen von 16 Enzymen des Zentralstoffwechsels und damit 208 Einzelmessungen zur Verfügung.

Im Fall der Experimente mit dem Stamm DM1933 konnten für die einzelnen Aktivitätsdaten Mess-Standardabweichungen aus den 3 technischen Replikaten geschätzt werden. Da hier sehr starke Unterschiede in der Reproduzierbarkeit auftraten ($0.05 \leq \hat{\sigma}_{\text{rel}} \leq 0.94$), wurde für alle anderen experimentellen Daten weiterführend ein Messfehler von 20% angenommen.

7.1.4 Metabolom-Daten intra- und extrazellulärer Metabolite

In Anlehnung an die Proteom-Daten wurden in jedem Experiment das mittels LC-MS/MS messbare Spektrum intrazellulärer Metabolite sowie die durch HPLC quantifizierbaren organischen Säuren des Überstandes erfasst. Die damit verbundene Probenaufarbeitung und Analytik wurde bereits in Kapitel 5.3.1 erläutert.

Auf diese Weise wurden 12 Datensätze mit intra- und extrazellulären Konzentrationen von 45 Intermediaten des ZSW generiert. Somit wurden auf Ebene des Metaboloms 540 Einzelmessungen ausgewertet.

Neben den technischen Replikaten mit DM1933 standen in allen Experimenten 6 unabhängige Probenreplikate für die Schätzung von Standardabweichungen der Konzentrationsmessungen zur Verfügung.

7.1.5 Fluxom-Daten extrazellulärer Flüsse

In allen Experimenten wurden auf Grundlage der Metabolom-Daten die Aufnahme- und Produktbildungsraten extrazellulärer Metabolite ermittelt.

Innerhalb der Chemostat-Experimente waren diesbezüglich die Glucoseaufnahme- sowie die Lysin- und CO₂-Bildungsraten relevant, wodurch insgesamt 39 extrazelluläre Flussmessungen erhalten wurden.

Wie bei den Proteom-Daten konnten nur im Fall der unabhängigen experimentellen Ansätze mit DM1933 vertrauenswürdige Messfehler ermittelt werden. In den anderen Fällen wurden jeweils Standardabweichungen von 5% für die Raten angenommen.

7.2 Vergleichende Analysen

Auf Grundlage der verfügbaren Multi-Omics-Datensätze wurden die Stämme hinsichtlich stamm- und wachstumsspezifischer Unterschiede sowie im Hinblick auf die Lysinproduktion charakterisiert.

7.2.1 Stamm- und wachstumsspezifische Unterschiede

Transkriptom

Aus der Gesamtmenge der extrahierten Transkriptionsdaten für WT und DM1730 wurden zunächst die Gene herausgefiltert, welche signifikante Unterschiede, d.h.

um den Faktor 2 höhere oder niedrigere Werte, zwischen beiden Stämmen aufweisen. Damit ergab sich je eine Teilmenge für die Wachstumsrate $\mu = 0.1 \text{ h}^{-1}$ bzw. $\mu = 0.2 \text{ h}^{-1}$ (Abb. 7.2, links). Aus der Schnittmenge beider Teilmengen ergaben sich 6 Gene, welche einen ausschließlich stammspezifischen Unterschied zeigten.

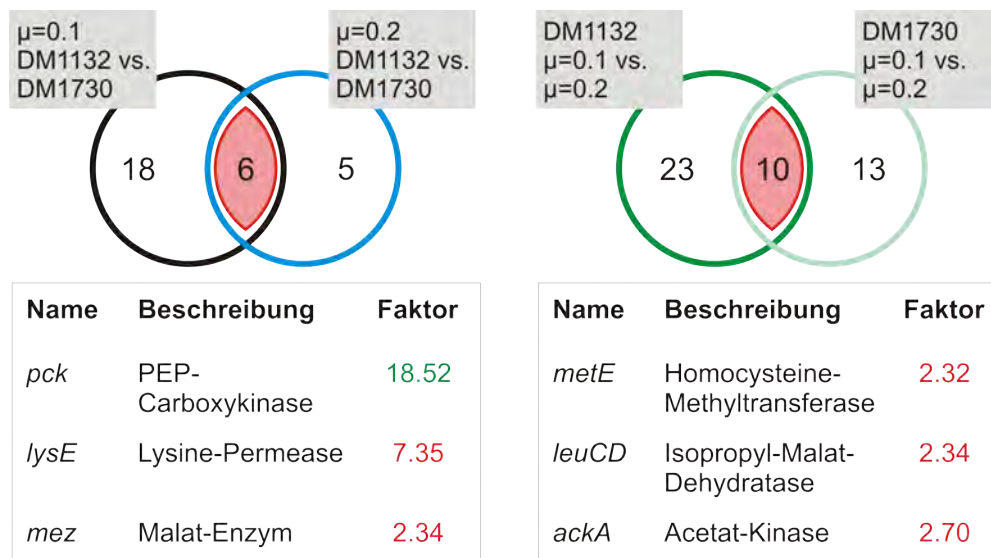


Abbildung 7.2: Venn-Diagramme für die Teilmengen an Genen mit signifikanter Transkriptionsänderung im Vergleich der 4 Ansätze. Links: Stammspezifische Unterschiede. Rechts: Wachstumsabhängige Unterschiede. In den unteren Tabellen ist der Teil der Gene aufgelistet, welcher sich aus den jeweiligen Schnittmengen ergibt und für den darüber hinaus eine entsprechende Funktion zugeordnet werden konnte. Der zugehörige Zahlenwert kennzeichnet dabei den Faktor, um den einzelne Gene herauf- bzw. herunterreguliert sind (rot bzw. grün).

Hierbei fällt sofort der deutlich niedrigere Transkriptionswert des *pck*-Gens in DM1730 auf, der auf die zugehörige Deletion zurückzuführen ist. Neben einer 7-fach höheren Transkription des Gens für die Lysin-Permease (*lysE*) zeigte auch das Gen für das Malat-Enzym (*mez*) eine Verdopplung der Transkription. Insbesondere im Hinblick auf die Ergebnisse der Stoffflussanalyse mit DM1730 (vgl. Kap. 5) stellt sich damit die Frage, inwieweit die Deletion von *pck* eventuell zu einer erhöhten Aktivität des Malat-Enzyms führt.

Bezüglich der wachstumsabhängigen Unterschiede fällt auf, dass lediglich ein paar Gene der Leucin- und Methionin-Biosynthese bei der höheren Wachstumsrate stärker transkribiert wurden (Abb. 7.2, rechts). Hingegen der ursprünglichen Erwartungshaltung zeigten sämtliche ribosomale Gene keinen signifikanten Wachstumseffekt, d.h. deren Transkription wurde nicht mit zunehmender Wachstumsrate erhöht. Ein Blick auf die absoluten Transkriptionswerte dieser Gene zeigte, dass die-

se in beiden Fällen schon vergleichsweise hoch waren. Daraus ließe sich einerseits schlussfolgern, dass der jeweilige Organismus im Chemostat unter substratlimitierten Bedingungen generell über einen Proteinsyntheseapparat verfügt, welcher auf eine maximale Leistung reguliert ist. Andererseits besteht auch die Möglichkeit, dass die Regulation auf Ebene des Transkriptoms wesentlich feiner erfolgt und somit diese Änderungen bei dem hier gewählten Signifikanzniveau (Faktor 2) nicht erfasst wurden.

Proteom

Ein Vergleich der Enzymaktivitäten zwischen allen drei Stämmen, unabhängig von der Wachstumsrate, zeigte keine wesentlichen Unterschiede. Einzige Ausnahme bildete hier die Isocitrat-Dehydrogenase (*icd*), welche bei den Wachstumsraten 0.1 h^{-1} und 0.2 h^{-1} eine ca. 7-fach niedrigere Aktivität im Stamm DM1933 aufwies.

Ein ähnliches Bild ergab sich bei der Untersuchung des Einflusses der Wachstumsrate auf die Enzymaktivitäten. Hierbei wurde lediglich für die Fumarase (*fum*) eine deutliche Erhöhung der Enzymaktivität mit steigender Wachstumsrate unabhängig vom Stammkonstrukt beobachtet. In Abbildung 7.3 sind beispielhaft die Daten für den Stamm DM1933 dargestellt.

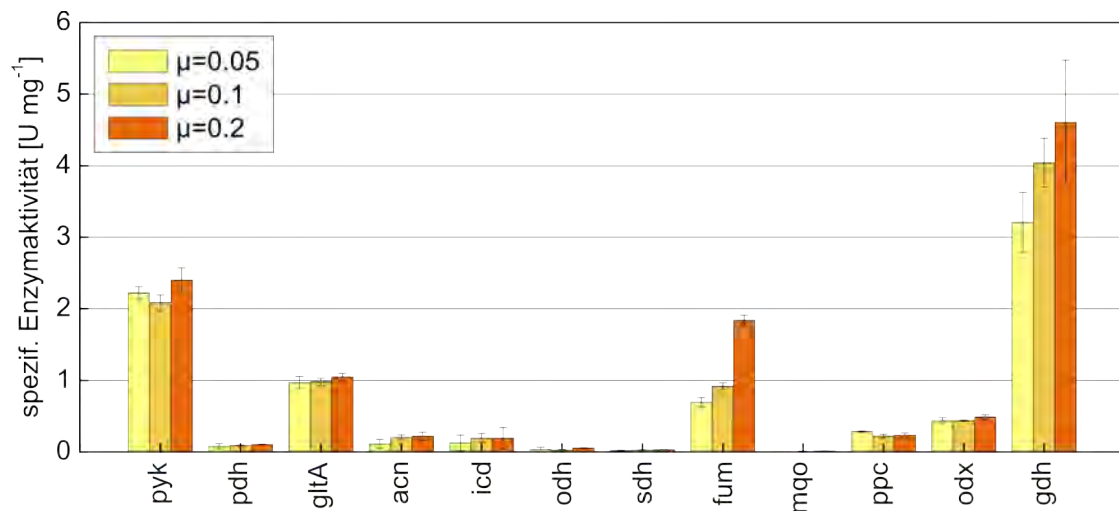


Abbildung 7.3: Vergleich von Enzymaktivitätsmessungen für den Stamm DM1933 bei verschiedenen Wachstumsraten in kontinuierlicher Kultivierung. Die Fehlerbalken basieren auf 3 technischen Replikaten.

Metabolom

Im Fall der ermittelten Metabolom-Daten war ein deutlicher Abfall der Konzentrationen aller Intermediate entlang der Stämme DM1730 und DM1933 zu beobachten (Abb. 7.4). Möglicherweise spielen hier die zunehmenden genetischen Modifikationen eine Rolle.

In Bezug auf die verschiedenen Wachstumsraten waren deutliche Unterschiede in allen Intermediaten erkennbar, diese unterlagen jedoch keinem systematischen Trend.

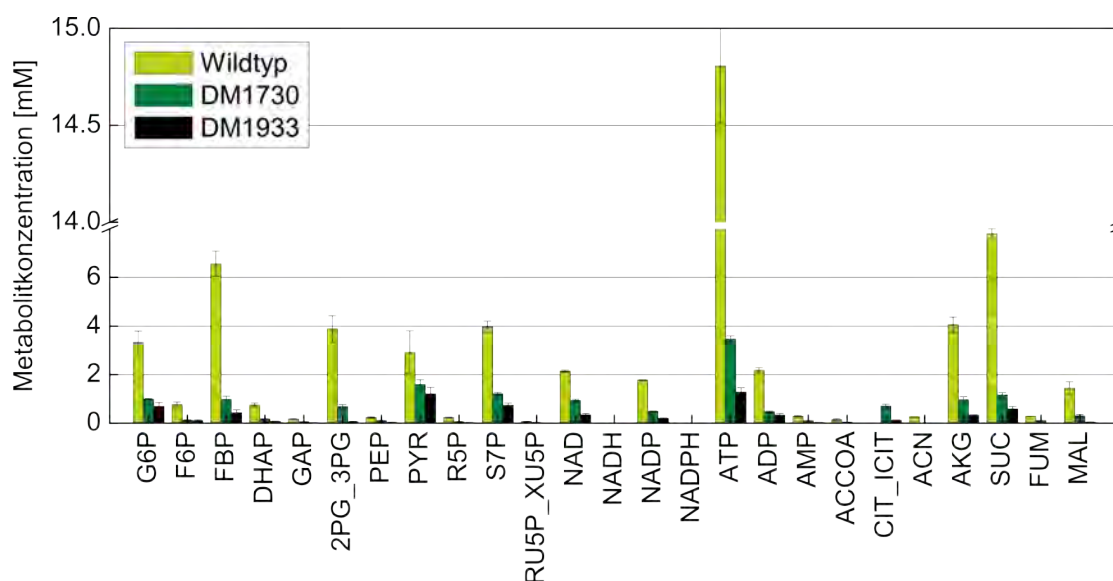


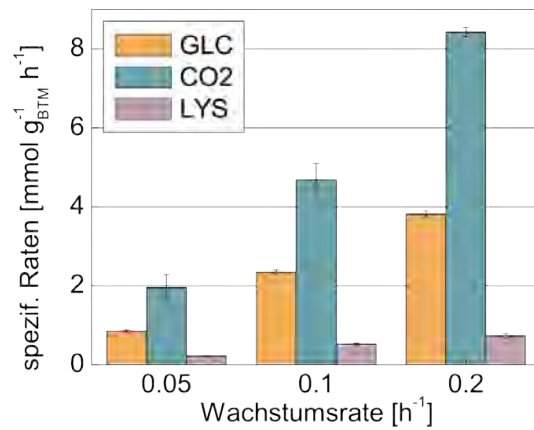
Abbildung 7.4: Vergleich von Metabolitkonzentrationen verschiedener *C. glutamicum*-Stämme bei der Wachstumsrate 0.1 h^{-1} in kontinuierlicher Kultivierung. Die Fehlerbalken basieren auf 6-fach-Bestimmungen eines Ansatzes (WT, DM1730) bzw. auf 4-fach-Bestimmungen in 3 technischen Replika (DM1933).

Produktbildung

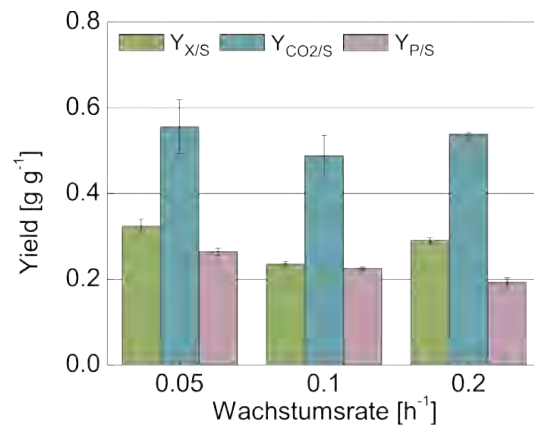
Im Hinblick auf die spezifischen Bildungsraten und Erträge sind deutliche Unterschiede im Stamm DM1933 im Vergleich zu seinem Vorgänger DM1730 erkennbar. Durch eine 1.5-fach höhere Substrataufnahmerate ist bei vergleichbarer CO_2 -Bildung die Lysinproduktion um den Faktor 2.5 höher als im Stamm DM1730.

Wie zu erwarten, stiegen bei beiden Stämmen mit zunehmender Wachstumsrate die Produktbildungsraten, während der substratspezifische Lysinertrag ($Y_{P/S}$) sank (beispielhaft für DM1933 in Abb. 7.5 dargestellt).

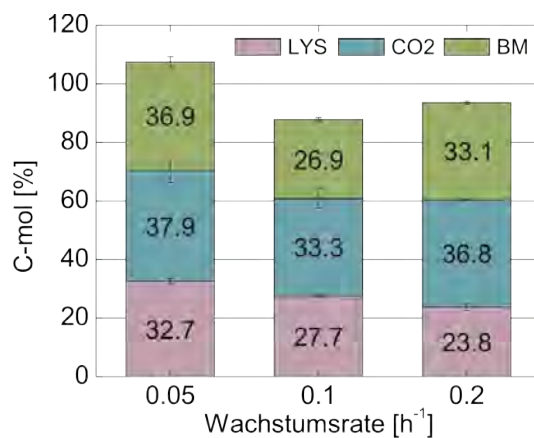
Im Gegensatz zu DM1730 ist im Stamm DM1933 jedoch nicht ein erhöhter Biomasseertrag ($Y_{X/S}$) die Ursache. Dementsprechend waren in diesem Fall auch die C-Bilanzen bei $\mu = 0.1, 0.2 \text{ h}^{-1}$ nicht geschlossen. Da hier keine Nebenproduktbildung in Form organischer Säuren (Acetat, Glutamat usw.) festgestellt werden konnte, besteht die Vermutung, dass nicht erfasste extrazelluläre Speicherstoffe, wie beispielsweise Trehalose, gebildet wurden [WKM03].



(a) Biomassespezifische Aufnahme- und Produktbildungsrate



(b) Substratspezifische Erträge



(c) Kohlenstoffbilanzen

Abbildung 7.5: Vergleich spezifischer Kenngrößen für den Stamm DM1933 bei verschiedenen Wachstumsraten. Die jeweiligen Fehlerbalken basieren auf 3 technischen Replikaten.

Kapitel 8

Vertikales Netzwerkmodell des Zitratzyklus

Aus den vorangegangenen Abschnitten wird deutlich, dass zwar eine Vielzahl quantitativer Messdaten entlang der verschiedenen Regulationsebenen mikrobieller Stoffwechselnetzwerke gewonnen werden können, sich die Interpretation dieser Daten ohne den zugehörigen Netzwerkkontext allerdings als schwierig und teilweise erfolglos gestaltet.

Es ist daher zwingend erforderlich, modellgestützte Auswertungsmethoden zu entwickeln, welche die notwendigen Zusammenhänge zwischen Multi-Omics-Messdaten und den zugrunde liegenden *in vivo*-Stoffwechselprozessen herstellen.

In den nachfolgenden Abschnitten wird ein solcher integrativer Modellierungsansatz vorgestellt. Aufgrund der bereits in der Einleitung abgeleiteten Zielsetzungen liegt der Schwerpunkt der Modellierung hierbei auf dem TCA und der Anaplerose von *C. glutamicum*.

8.1 Vertikale Modellierung

Mit dem Ziel einer quantitativen Beschreibung der erhobenen Multi-Omics-Daten kommen bezüglich der zu wählenden Modellierungsstrategie vor allem dynamische Systemansätze in Frage (vgl. 3.3 und 3.4.2). Damit besteht die Möglichkeit einer mechanistischen Modellierung der verschiedenen Omics-Ebenen (Genom, Proteom, Metabolom), wodurch auch das stark nichtlineare Verhalten aller Regulationsprozesse abgebildet werden kann.

Aufgrund der noch vielfach unbekannt komplexen Wechselwirkungsmechanis-

men innerhalb und zwischen den verschiedenen Regulationsebenen können gegenwärtig keine mechanistischen „Ganzzellmodelle“ formuliert werden, welche alle relevanten Zustandsgrößen umfassen und darüber hinaus prädiktiv sind.

Damit stellt sich natürlich die grundlegende Frage, welche Systemgrenzen in einem fokussierten Modell angenommen werden sollten.

8.1.1 Massenbilanzen von mRNA, Proteinen und Metaboliten

Werden in einem ersten Vereinfachungsschritt alle mRNAs, Proteine, Enzyme und Metabolite als zeitabhängige Zustandsgrößen sowie alle Gene, RNA-Polymerasen, Transkriptions- bzw. Translationsfaktoren, Ribosomen, Nukleasen und Proteasen als externe, d.h. außerhalb der Systemgrenzen liegende, zeitabhängige Variablen angenommen (Abb. 8.1) können zunächst folgende Massenbilanzen formuliert werden:

$$\frac{dmRNA_i}{dt} = v_{tc}(G_i, \mathbf{RNAP}, \mathbf{TKF}, \mathbf{c}) - v_{dg,mRNA}(mRNA_i, \mathbf{NK}) - \mu \cdot mRNA_i \quad (8.1)$$

Die Konzentration der i -ten $mRNA$ hängt einerseits von der Transkriptionsrate v_{tc} ab, welche durch die Konzentrationen an i -tem Gen G_i , RNA-Polymerasen \mathbf{RNAP} , Transkriptionsfaktoren \mathbf{TKF} und Effektormolekülen \mathbf{c} bestimmt wird. Andererseits nimmt die Menge der $mRNA$ pro Zellvolumen über die Degradation $v_{dg,mRNA}$ durch Nukleasen \mathbf{NK} und einer Verdünnung infolge von Zellwachstum ab.

$$\begin{aligned} \frac{dP_i}{dt} &= v_{tl}(mRNA_i, \mathbf{RB}, \mathbf{TLF}, \mathbf{c}) - v_{dg,P}(P_i, \mathbf{PT}) - \mu \cdot P_i \\ \frac{dE_j}{dt} &= v_{ic}(P_{ij}, \mathbf{c}) \end{aligned} \quad (8.2)$$

Die Translationsrate v_{tl} der i -ten $mRNA$ in das i -te Protein P_i wird durch die Konzentration an Ribosomen \mathbf{RB} sowie Translationsfaktoren \mathbf{TLF} und Effektormolekülen bestimmt. Die Rate $v_{dg,P}$ kennzeichnet wiederum die Degradation des i -ten Proteins durch Proteasen \mathbf{PT} . Ein bestimmter Teil der synthetisierten Proteine wird durch reversible Interkonversion mit der Rate v_{ic} in aktive Enzyme E überführt.

$$\frac{dc_k}{dt} = \sum_{j=1}^n v_{ct}(E_j, \mathbf{c}) - \mu \cdot c_k \quad (8.3)$$

Die zeitliche Konzentrationsänderung des k -ten Metaboliten c_k hängt von mehreren enzymkatalysierten Reaktionen v_{ct} ab, die ihrerseits durch das jeweilige Verhältnis

Edukt zu Produkt und Modulationen der Enzymaktivitäten durch Effektoren bestimmt werden.

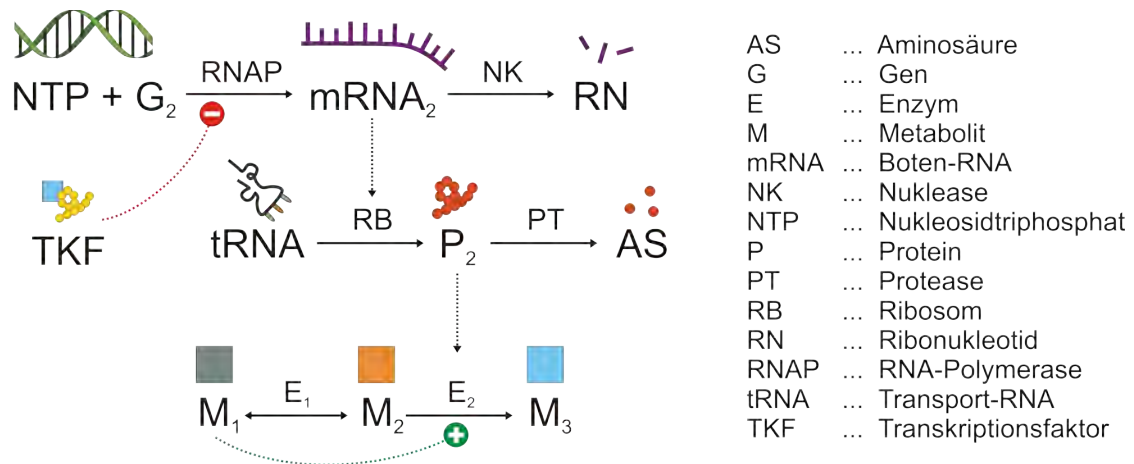


Abbildung 8.1: Schematische Darstellung der Wirkungskette vom Gen zum metabolischen Fluss. In einer Zwei-Schritt-Reaktion werden die Metabolite M_1 bis M_3 ineinander umgewandelt. Die zweite Reaktion wird dabei durch das mit dem Protein P_2 identische Enzym E_2 katalysiert. P_2 wird an den Ribosomen durch Verknüpfung entsprechender Aminosäuren gebildet und durch Proteasen wieder zersetzt (Degradation). Die für die Translation von P_2 notwendige Boten-RNA $mRNA_2$ wird durch RNA-Polymerase synthetisiert und durch Nukleasen degradiert. Der Metabolit M_1 wirkt als Aktivator auf das Enzym E_2 . Darüber hinaus kann M_3 als Effektor eines Transkriptionsfaktors fungieren, wodurch eine Feedback-Inhibierung der Transkription von $mRNA_2$ resultiert. Das Schema berücksichtigt keine Regulation auf Ebene der Translation sowie Interkonversion von Proteinen.

8.2 Modellierung genetischer Regulation

Im Bereich des TCA von *C. glutamicum* sind mittlerweile eine Reihe von Faktoren bekannt, welche die Enzymsynthese auf genetischer Ebene regulieren [KWB05, WKB05, CGS⁺06, CAF⁺07] oder eine Post-translationale Kontrolle ausüben [NKS⁺06]. Allerdings fehlen in den meisten Fällen die Informationen über mögliche Metaboliteffektoren, welche die inaktiven in aktive Regulatorproteine verwandeln.

Darüber hinaus ist die Mehrzahl der externen Variablen, welche die Transkription, Translation und Interkonversion bestimmen (Gl. 8.1 und 8.2) derzeit noch nicht quantitativ erfassbar. Deshalb ließe sich nur ein geringer Teil der damit verbundenen Modellparameter für den *in vivo* relevanten Fall statistisch identifizieren. Aus diesem Grund wurden für die hier vorgestellte, datengetriebene Modellierung die-

ser Prozesse notwendigerweise eine Reihe von vereinfachenden Annahmen getroffen.

8.2.1 Vereinfachende Annahmen

1. Eine direkte Regulation der Transkription und Translation wird nicht berücksichtigt, d.h. weder die Transkriptions- noch die Translationsrate sind abhängig von Regulationsfaktoren (**TKF**, **TLF**) und Effektormolekülen (**c**). Damit gilt: $v_{tc} = f(G_i, \mathbf{RNAP})$ und $v_{tl} = f(mRNA_i, \mathbf{RB})$. Gleichzeitig gibt es damit keine Rückkopplungen zwischen der Metabolom- und Genomebene (vgl. Gl. 8.4).
2. Eine post-translationale Modifikation von Enzymaktivitäten wird nicht berücksichtigt, d.h. alle synthetisierten Proteine stellen aktive Enzyme dar ($P_i = E_j$).
3. Die Elongationsrate ist für alle mRNA-Synthesen und Peptidbindungen gleich groß (vgl. Tab. Anhang C.1).
4. Die Degradation von mRNA und Proteinen unterliegt keiner Regulation und ist direkt proportional zur jeweiligen Bilanzgröße ($\mathbf{v}_{dg,mRNA} = k_{dg,mRNA} \cdot \mathbf{mRNA}$, $\mathbf{v}_{dg,P} = k_{dg,P} \cdot \mathbf{P}$). Die Terme $k_{dg,mRNA}$ und $k_{dg,P}$ stellen dabei Degradationskonstanten dar, welche für alle mRNAs bzw. Enzyme gleich groß sind (vgl. Tab. Anhang C.1).

8.2.2 Phänomenologische Beziehungen

Aus experimentellen Untersuchungen mit *E. coli* ist seit längerem bekannt, dass die Geschwindigkeit der Transkription (insbesondere rRNA) und Translation eng mit der Wachstumsrate verbunden ist (vgl. Abb. 8.2). Die Untersuchung der dabei zugrunde liegenden Regulationsmechanismen ist Gegenstand andauernder Forschung [BD08].

Eine Möglichkeit der vereinfachten mathematischen Beschreibung solcher komplexer Mechanismen besteht in der Formulierung von phänomenologischen Beziehungen zwischen messbaren bzw. mit gewisser Sicherheit abschätzbaren Größen.

Für die Abhängigkeit der Transkriptionsrate v_{tc} sowie der Translationsrate v_{tl} von der spezifischen Wachstumsrate μ wurden folgende Beziehungen aufgestellt [AN99,

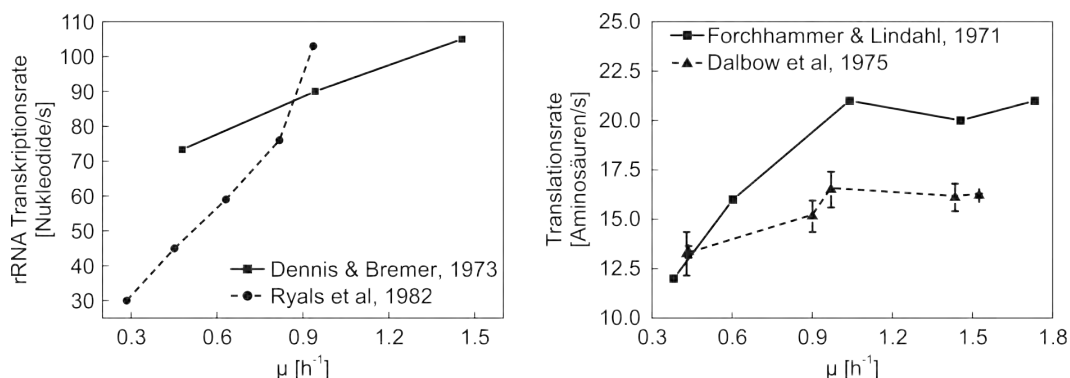


Abbildung 8.2: Zusammenhang zwischen der Transkriptions- bzw. Translationsrate und der spezifischen Wachstumsrate in verschiedenen *E. coli*-Experimenten.

LB00]:

$$v_{tc} = \frac{v_{eg,mRNA}}{d_{RNAP}(\mu)} \cdot GP \quad (8.4)$$

$$v_{tl} = \frac{v_{eg,P} \cdot \beta(\mu)}{d_{RB}(\mu)} \cdot mRNA$$

Hierbei bezeichnen $v_{eg,mRNA}$ und $v_{eg,P}$ die Raten der mRNA-Elongation pro aktiver RNAP und der Peptid-Elongation pro aktivem Ribosom. Der Term GP kennzeichnet die Konzentration an Genen mit gebundener RNAP. Während v_{tc} stark von μ abhängt, bleibt v_{tl} oberhalb einer Verdopplungszeit $\tau = \mu / \ln(2)$ von einer Stunde konstant. Für $\tau < 1h$ ist hingegen ein linearer Abfall zu beobachten [FL71,DY75]:

$$\beta(\mu) = \begin{cases} 1 & \mu > \ln(2) \\ \mu / \ln(2) & \mu < \ln(2) \end{cases} \quad (8.5)$$

Die Terme $d_{RNAP}(\mu)$ und $d_{RB}(\mu)$ kennzeichnen den mittleren molekularen Abstand zwischen transkribierenden RNAP-Molekülen sowie translatierenden Ribosomen auf den jeweiligen Templaten (DNA, mRNA). Auf der Basis experimenteller Daten für *E. coli* wurden diese Abstände wie folgt geschätzt [DEB04]:

$$d_{RNAP}(\mu) = 233 \cdot \mu^{-2} + 78 \text{ NT} \quad (8.6)$$

$$d_{RB}(\mu) = 82.5 \cdot \mu^{-1} + 145 \text{ NT} \quad (8.7)$$

8.3 Modellierung metabolischer Regulation

Im Gegensatz zum Transkriptom und Proteom ist die Modellierung im Bereich der metabolischen Ebene in der Systembiologie bereits gut etabliert (vgl. Kap. 3.3).

Die Stoffwechselwege des TCA und der Anaplerose von *C. glutamicum* sind gut untersucht, so dass eine Vielzahl von Literaturquellen ausgewertet werden konnte, um ein detailliertes Regulationsnetzwerk auf metabolischer Ebene aufzubauen (Abb. 8.3).

Aufgrund der fokussierten Betrachtungsweise sind aber auch in diesem Teilnetzwerk eine Reihe von Annahmen notwendig.

8.3.1 Vereinfachende Annahmen

1. Die einzelnen Reaktionen des EMP und des PPP werden nicht im Detail modelliert, sondern in einer stationären, globalen Massenbilanz erfasst (Gl. 8.8). Ähnliches gilt für den Lysin-Biosyntheseweg. Hier werden alle Teilreaktionen in einem einzigen kinetischen Ausdruck zusammengefasst (vgl. Abb. 8.3, bs-LYS).
2. Der Verbrauch an Vorläufern für die Biomassesynthese wird aus Literaturdaten geschätzt (vgl. Anhang Tab. B.2) und ist für alle betrachteten Stämme gleich groß und proportional zur Wachstumsrate ($v_{\text{bm}} = Y_{\text{IM}/X} \cdot \mu$).
3. Die Reaktionen des Glyoxylatshunts (aceA, aceB, vgl. Abb. 5.4) sind bei Wachstum auf Glucose inaktiv [GWS⁺03].
4. Der Einfluss der Menaquinone (MQ) als Elektronen-Akzeptoren der Oxireduktase-Reaktionen (sdh, mqo) wird vernachlässigt.
5. Die Reaktion der PEP-Carboxykinase (pck) wird als ATP-abhängige Phosphorylierung angenommen.

8.3.2 Reaktionskinetiken

Im Kapitel 3.3.5 zur kinetischen Modellierung metabolischer Netzwerke wurden neben dem klassischen mechanistischen Ansatz nach Michaelis-Menten auch eine Reihe verschiedener approximativer Modellansätze eingeführt. Da für die Enzyme des TCA und der Anaplerose von *C. glutamicum* in der überwiegenden Zahl der

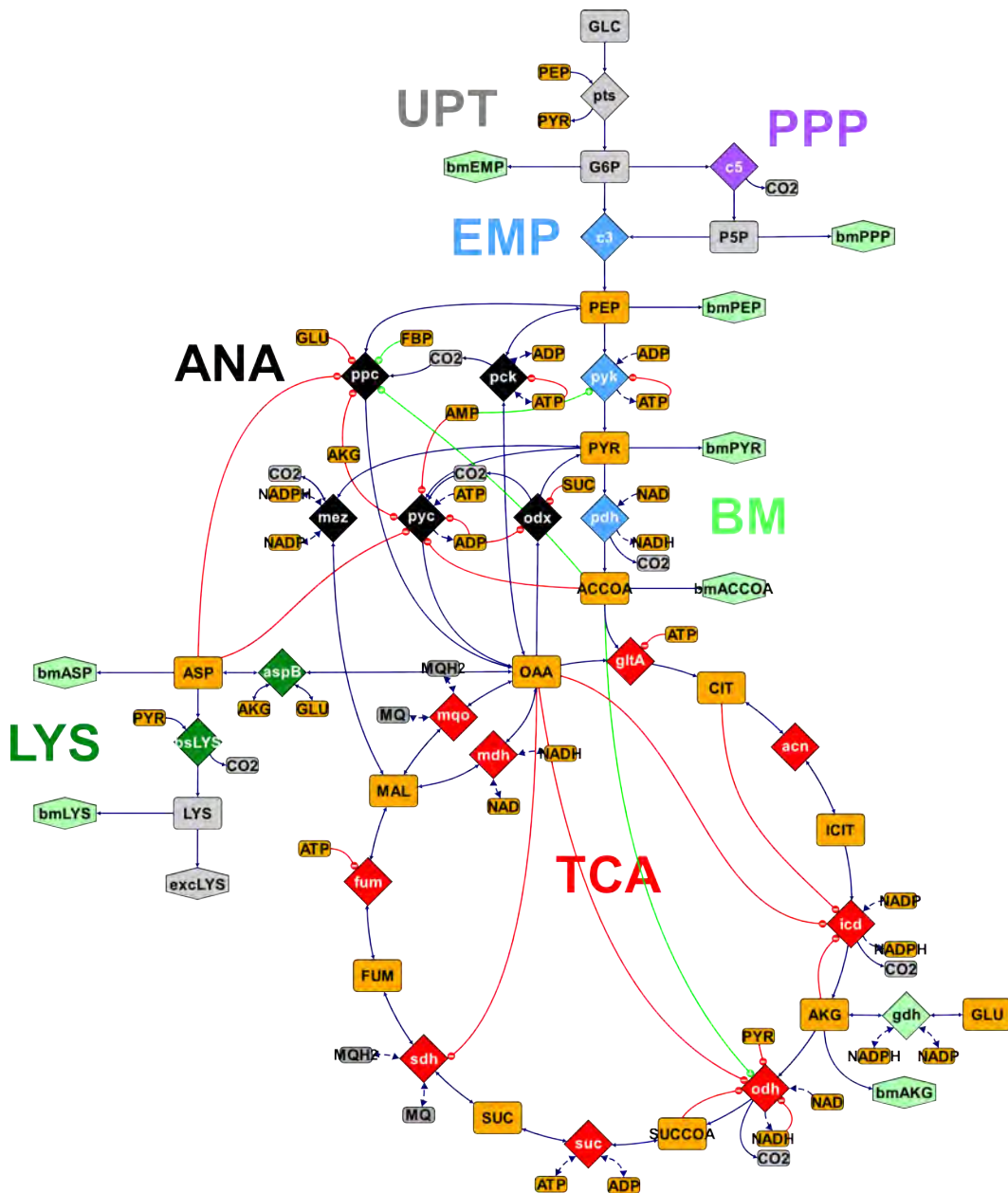


Abbildung 8.3: Metabolisches Teilmodell des vertikalen Netzwerks mit den Reaktionen des TCA und der Anaplerose von *C. glutamicum*. Die orange gekennzeichneten Metabolitpools stellen hierbei die Zustandsvariablen des Modells dar. Alle modellierten regulatorischen Wechselwirkungen durch Inhibitoren oder Aktivatoren sind durch entsprechende rote bzw. grüne Kreise gekennzeichnet. Annahmen bezüglich der Irreversibilitäten einiger Reaktionen sind durch die jeweiligen Pfeile in physiologischer Nettorichtung gekennzeichnet.

Fälle die Enzymmechanismen nicht bekannt sind, wurde auf einen approximativen Ansatz für die Beschreibung aller Reaktionskinetiken zurückgegriffen.

In voran gegangenen Untersuchungen hat sich gezeigt, dass der Convenience-Ansatz (vgl. 3.3.5) bei geringer Parameteranzahl gute Approximationseigenschaften aufweist und hier darüber hinaus keine Fehler bei sehr niedrigen Metabolitkonzentrationen auftreten [HNW08].

Desweiteren beruht dieser Ansatz auf einem mechanistischem Konzept, wodurch für eine Vielzahl der kinetischen Parameter die in der Literatur verfügbaren Informationen über gemessene *in vitro*-Werte im Zuge der Parameterschätzung genutzt werden können (vgl. Abschn. 9.2).

Eine Auflistung der durch Convenience-Kinetiken beschriebenen enzymatischen Reaktionen sowie aller im Modell berücksichtigten regulatorischen Wechselwirkungen durch Modulationen ist in Tabelle C.3 des Anhangs zu finden.

8.3.3 Stöchiometrische Gleichungen

Lokale Bilanzen

Für die im Fokus der Modellierung stehenden Reaktionen des TCA und der Anaplerose werden entsprechend Gl. 8.3 lokale Massenbilanzen um einzelne Metabolitpools aufgestellt. Auf diese Weise können die nichtlineare Kinetiken zur Beschreibung der metabolischen Regulationsprozesse in das Netzwerk integriert werden.

Globale Bilanzen

Mit Hilfe einer globalen Massenbilanz werden alle übrigen Stoffwechselabschnitte des Zentralstoffwechsels erfasst und in die Stöchiometrie des gesamten metabolischen Netzwerks eingebunden.

Für die Reaktionen des EMP und PPP wurde folgende Gesamtbilanz um den G6P-Knoten formuliert:

$$0 = 6 \cdot v_{\text{pts}} - 3 \cdot v_{\text{c3}} - v_{\text{c5}} - 6 \cdot (Y_{\text{EMP}/X} + Y_{\text{PPP}/X}) \cdot \mu \quad (8.8)$$

Die Ausbeutekoeffiziente $Y_{\text{EMP}/X}$ und $Y_{\text{PPP}/X}$ beinhalten den Verbrauch an G6P, F6P, GA3P und 3PG bzw. R5P und E4P für die Biomassesynthese (vgl. Tab. Anhang

C.2) und wurden wie folgt ermittelt:

$$\begin{aligned}
 Y_{\text{EMP}/X} &= Y_{\text{G6P}/X} + Y_{\text{F6P}/X} + \frac{1}{2} \cdot (Y_{\text{GA3P}/X} + Y_{\text{3PG}/X}) \\
 Y_{\text{PPP}/X} &= \frac{5}{6} \cdot Y_{\text{R5P}/X} + \frac{2}{3} \cdot Y_{\text{E4P}/X}
 \end{aligned}
 \tag{8.9}$$

Die stöchiometrischen Koeffizienten ergeben sich aus der Vereinfachung des gesamten algebraischen Gleichungssystems, welches durch entsprechende Massenbilanzierung um die Reaktionsknoten des EMP und PPP entsteht. Eine Herleitung der Gleichungen 8.8 und 8.9 ist im Anhang C.3 aufgeführt.

Zur Einbindung extrazellulärer Flussmessungen für die Parameterschätzung werden Beziehungen zwischen den Aufnahme- bzw. Produktionsraten und den intrazellulären Stoffflüssen benötigt.

Substrataufnahme Unter der Annahme, dass als einzige C-Quelle Glucose über das Phosphotransferasesystem (pts) aufgenommen wird, folgt unter Berücksichtigung von Gl. 8.8:

$$\text{upt}_{\text{GLC}} = v_{\text{pts}}
 \tag{8.10}$$

Lysinbildung: Durch eine vereinfachte Beschreibung des Syntheseweges wird die Lysinproduktionsrate aus folgender Bilanz um den Lysin-knoten erhalten:

$$\text{exc}_{\text{LYS}} = v_{\text{bsLYS}} - Y_{\text{LYS}/X} \cdot \mu
 \tag{8.11}$$

CO₂-Bildung: Die CO₂-Produktionsrate ergibt sich aus der Nettobilanz aller CO₂-bildenden und -verbrauchenden Reaktionen:

$$\text{exc}_{\text{CO}_2} = v_{\text{c5}} + v_{\text{pdh}} + v_{\text{icd}} + v_{\text{odh}} + v_{\text{pck}} + v_{\text{mez}} + v_{\text{odx}} + v_{\text{bsLYS}} + Y_{\text{CO}_2/X} \cdot \mu - v_{\text{pyc}} - v_{\text{ppc}}
 \tag{8.12}$$

Verfügbarkeit von NADPH Die durch NADPH gespeicherten Reduktionsäquivalente werden für die Biomasse- und Lysin-synthese sowie den Erhaltungsstoffwechsel benötigt. Eine Regeneration von NADPH kann im PPP durch Glucose-6-phosphat-Dehydrogenase (zwf) und Gluconat-Dehydrogenase (gnd), im TCA durch Isocitrat-Dehydrogenase (icd) sowie in der Anaplerose durch das Malat-Enzym (mez)

stattfinden (siehe auch Abb. 5.3):

$$v_{\text{NADPH}} = 2 \cdot v_{\text{c5}} + v_{\text{icd}} + v_{\text{mez}} - 4 \cdot v_{\text{bsLYS}} - Y_{\text{NADPH/X}} \cdot \mu \geq 0 \quad (8.13)$$

Aufgrund der Ungleichheit dient diese Massenbilanz nicht zur Reduktion eines Freiheitsgrades, sondern hilft die Sinnhaftigkeit der geschätzten Stoffflusslage zu überprüfen.

8.4 Modellimplementierung mit Modelica

Im Kapitel 4 wurde ausführlich auf die Anwendung von Modelica zur Modellierung biochemischer Netzwerke eingegangen. Auf der Grundlage der dabei entwickelten Modelica-Komponenten zur grafischen Netzwerkmodellierung unter Dymola wurde das in Abbildung 8.4 dargestellte vertikale Netzwerkmodell implementiert.

Hierbei sind folgende Dinge anzumerken:

1. Der Aufbau des metabolischen Netzwerks erfolgt im Hinblick auf eine bevorzugte visuelle Darstellung mit Metabolitknoten und Reaktionskanten. Beide Elemente sind ungerichtet über Konnektoren für Substrate und Produkte miteinander verbunden (vgl. Def. Abb. 4.3 und 4.4).
2. Einzelne Reaktionskomponenten verfügen über zusätzliche Schnittstellen zur Einbindung von Cometaboliten, welche externe Variablen darstellen und nicht über Massenbilanzen beschreibbar sind.
3. Die Wirkung von Effektormetaboliten auf Reaktionen wird durch gerichtete rote (Inhibitoren) bzw. grüne (Aktivatoren) Verbindungen definiert (vgl. Def. Abb. 4.5).
4. Aufgrund der vereinfachten Annahme fehlender Regulation auf Ebene des Genoms und Proteoms werden die Prozesse Transkription und Translation (vgl. Kap. 8.2) in eine gemeinsame Modellkomponente (*Genexpression*) integriert. Die für jedes Gen spezifizierten Expressionsmodelle liefern als Ausgangsvariable eine bestimmte Menge aktiven Enzyms und sind deshalb gerichtet mit den Reaktionskomponenten verbunden.

Anhand der Abbildung 8.4 wird deutlich, dass mit den definierten Komponenten ein grafischer Netzwerkaufbau auf einfache Art und Weise möglich ist, sich jedoch

aufgrund der damit verbundenen Darstellung der Netzwerkzusammenhang nicht mehr intuitiv erschließt.

Insbesondere zur anschließenden Interpretation der umfangreichen Simulationsdaten im Netzwerkkontext ist daher die Anwendung von speziellen Werkzeugen zur Netzwerkvisualisierung zwingend erforderlich [DNNW09].

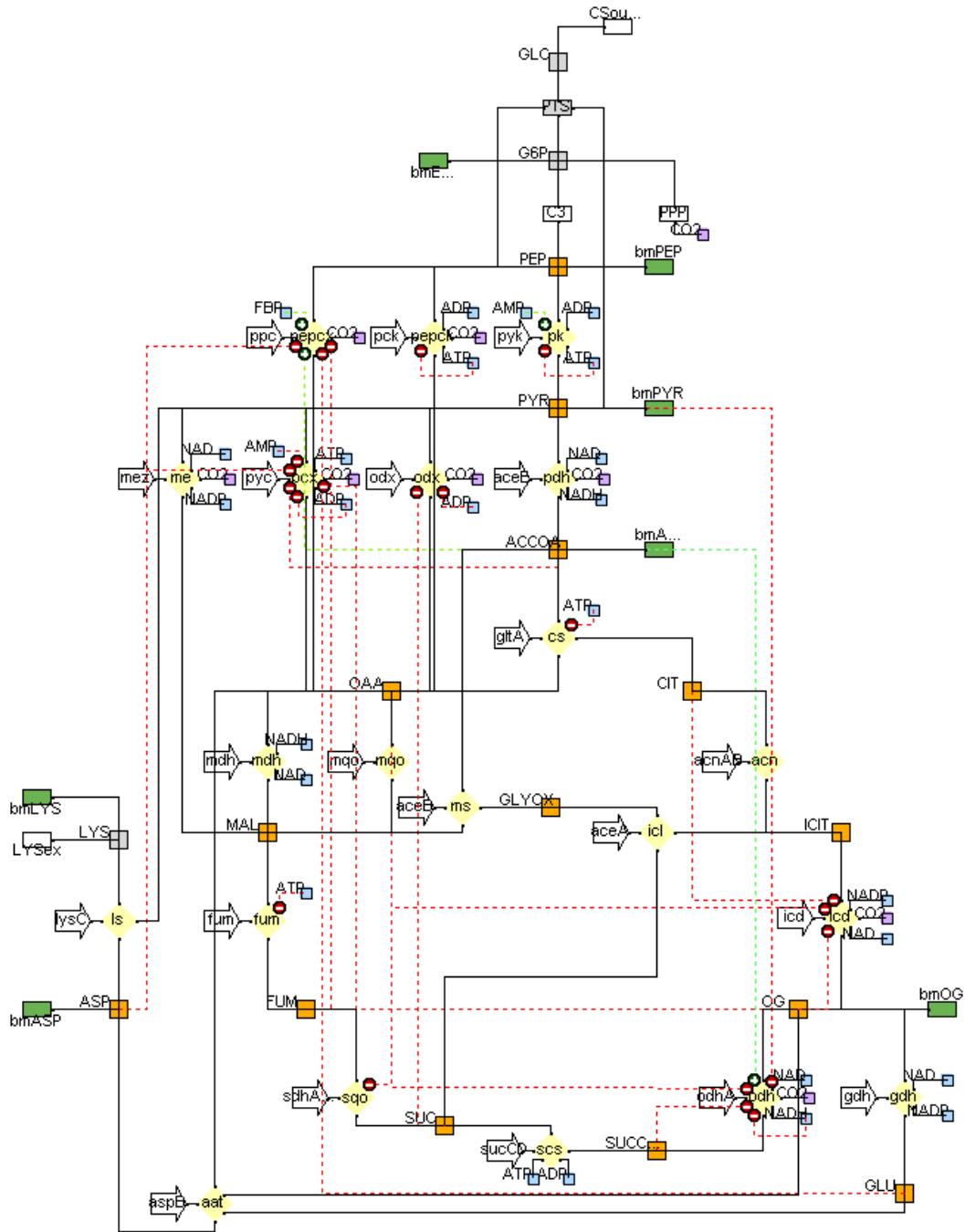


Abbildung 8.4: Implementierung des vertikalen Netzwerkmodells mit Dymola.

Kapitel 9

Validierung des vertikalen Netzwerkmodells

Nachdem die Formulierung und Implementierung des vertikalen Netzwerkmodells für *C. glutamicum* dargestellt wurde, befasst sich das folgende Kapitel mit der Integration der Multi-Omics-Messdaten zur Modellvalidierung. Neben der Parameterschätzung beinhaltet dies auch eine Sensitivitätsanalyse der Modellparameter und -variablen sowie die Interpretation der Simulationsergebnisse im Netzwerkkontext.

9.1 Integration von Multi-Omics-Daten

Wie bereits im Abschnitt 7.1 erläutert, wurden im Rahmen kontinuierlicher Kultivierungen verschiedene *C. glutamicum*-Stämme hinsichtlich des Transkriptoms (mRNA-Mengen mittels Affymetrix-Chips), Proteoms (Enzymaktivitätsmessungen), Metaboloms (Metabolitkonzentrationen mittels LC/MS) und Fluxoms (ermittelte extrazelluläre Raten) analysiert.

Zum Zweck der Modellvalidierung wurden aus der Gesamtheit der Experimente die drei vollständigen Multi-Omics-Datensätze mit dem Lysinproduzenten DM1730 ($\mu = 0.1, 0.2 \text{ h}^{-1}$) sowie dem Wildtyp ($\mu = 0.1 \text{ h}^{-1}$) ausgewählt.

Für die Integration dieser Daten in den vertikalen Modellansatz müssen eine Reihe weiterer Annahmen getroffen werden:

1. Zum Zeitpunkt des Experiments sind alle Zustandsgrößen in einem stationären Zustand, d.h. es finden keine zeitlichen Änderungen der mRNA-, Enzym- und Metabolitpools statt. Zur Beschreibung dieses Zustandes wird das differentielle in ein algebraisches Gleichungssystem überführt und es gilt:

$$d\mathbf{mRNA}/dt = d\mathbf{E}/dt = d\mathbf{c}/dt = 0.$$

- Die auf der Basis von Affymetrix-Chips erzeugten mRNA-Intensitäten sind direkt proportional zur Konzentration der jeweiligen mRNA. Eine Transformation zwischen Messgrößen und simulierten Werten erfolgt mittels linearer Regression:

$$\widehat{\mathbf{mRNA}}_i = \omega_i \cdot \mathbf{mRNA}_i \quad (9.1)$$

Hierbei bezeichnen $\widehat{\mathbf{mRNA}}_i$ und \mathbf{mRNA}_i die Vektoren der simulierten mRNA-Konzentrationen bzw. gemessenen mRNA-Intensitäten des i -ten Experimentes. Der Skalierungsfaktor ω_i kennzeichnet die unbekannte Gesamtmenge an RNA.

- Im Fall von Enzymkomplexen (z.B. pdh) stellt der Mittelwert der Transkriptmengen aller Gene, welche für die einzelnen Teilenzyme kodieren, ein repräsentatives Maß dar.
- Die mit Hilfe von Enzymassays ermittelten Enzymaktivitäten sind direkt proportional zur Konzentration des jeweiligen aktiven Enzyms und stellen somit maximale Reaktionsgeschwindigkeiten dar. Für die in physiologischer Richtung mit maximaler Geschwindigkeit ablaufende Reaktion gilt demnach mit Gl. 3.21:

$$v_{\max}^{\rightarrow} = k^{\rightarrow} \cdot E \quad (9.2)$$

Für eine Umrechnung der spezifischen Enzymaktivitäten in U mg^{-1} auf biomassenspezifische Stoffflüsse in $\text{mmol g}_{\text{BTM}}^{-1} \text{h}^{-1}$ wurde mit einem konstanten Proteinanteil von 50% der Biotrockenmasse gerechnet [MdGW⁺96].

9.2 Gekoppelte Parameterschätzung

Aufgrund der Vielzahl an zu schätzenden Modellparametern kann sich die Suche nach einem globalen Optimum der Zielfunktion als schwierig erweisen (vgl. Anhang A.1). Eine generelle Möglichkeit zur Eingrenzung dieses Problems besteht in der Definition von sinnvollen Startwerten und engen Schätzgrenzen aller Modellparameter. Allerdings sind genau diese Informationen in Bezug auf enzymkinetische Parameter nur begrenzt verfügbar. Darüber hinaus ist es fraglich, inwie-

weit *in vitro* gemessene Werte die *in vivo*-Realität auch nur näherungsweise abbilden [Mat93, RPKW98].

Eine Alternative bietet die Formulierung von Nebenbedingungen für das Schätzproblem, welche aus dem biochemischen Kontext abgeleitet werden können. Hinsichtlich der Modellparameter lässt sich eine Unterteilung in „organismenspezifische“ und „experimentenspezifische“ Parameter treffen. Als organismenspezifische Modellparameter werden alle kinetischen Parameter bezeichnet, welche Enzymeigenschaften beschreiben (z.B. k^{\rightarrow} , K_m usw.). Diese sind, insofern keine genetische Modifikation (z.B. Beseitigung der Inhibierung) vorliegt, für einen bestimmten Organismus als konstant zu betrachten. Hingegen umfassen experimentenspezifische Modellparameter die von den experimentellen Bedingungen abhängigen Systemgrößen, welche nicht durch Massenbilanzen erfasst wurden (z.B. GP , c_{FBP} , c_{AMP} usw.) oder zur Lösung der Stöchiometrie dienen (c_{ACCOA} , c_{AKG}).

Unter Berücksichtigung der obigen Aspekte wurden für die Parameterschätzung folgende Festlegungen getroffen:

1. Sowohl die in der Literatur verfügbaren Informationen als auch aktuelle Ergebnisse der SysMAP-Kooperationspartner zu *in vitro*-Werten von Affinitätsparametern K_m und Gleichgewichtskonstanten K_{eq} wurden genutzt, um Startwerte für die Optimierung zu definieren und gleichzeitig den Schätzbereich auf $\pm 50\%$ des *in vitro*-Wertes zu begrenzen (vgl. Anhang Tab. C.5).
2. Für alle anderen Parameter wurden je nach biochemischem Zusammenhang entsprechend weite Schätzgrenzen definiert.
3. Sämtliche Modellparameter wurden im Zuge einer parallelen Optimierung geschätzt. Dies bedeutet, dass eine gleichzeitige Anpassung mehrerer Modelle an die zur Verfügung stehenden Multi-Omics-Datensätze erfolgte (Abb. 9.1). Die Modelle sind über die organismenspezifischen Parameter miteinander gekoppelt und unterscheiden sich in der Wertzuweisung der experimentenspezifischen Parameter sowie den jeweils definierten Wachstumsraten μ und damit verbundenen Biomasseabflüssen.

Unter Berücksichtigung der drei vertikalen Datensätze müssen somit 87 experimentenspezifische und 103 organismenspezifische Modellparameter geschätzt werden.

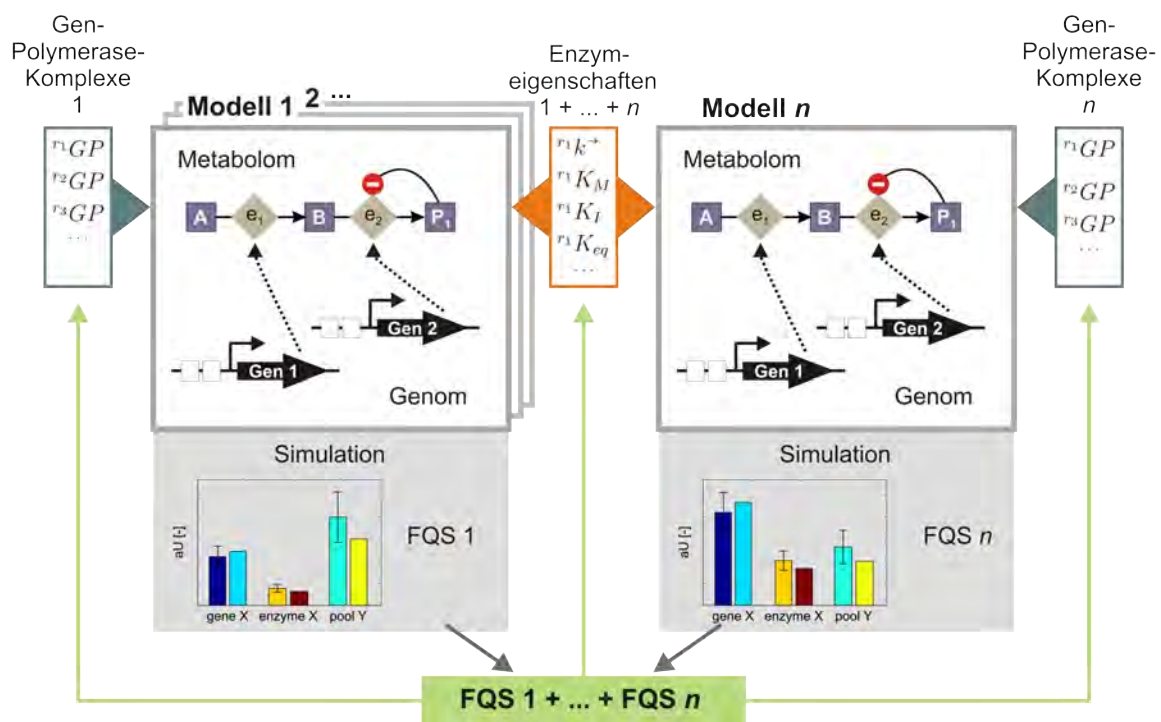


Abbildung 9.1: Prinzip der parallelen Optimierung von gekoppelten Modellen verschiedener Datensätze. Organismenspezifische Parameter (orange) werden in Abhängigkeit aller Datensätze geschätzt. Hingegen erfolgt die Identifizierung experimentenspezifischer Parameter (blau) in Bezug auf jeden einzelnen Datensatz getrennt. Sämtliche Abweichungen zwischen Messungen und Simulationen werden addiert und bilden eine gemeinsame Fehlerquadratsumme (FQS) aller gekoppelten Modelle.

9.3 Modellanpassungen

In den Tabellen 9.1 bis 9.4 sind die experimentellen Daten der gemessenen extrazellulären Raten, Transkripte, Enzymaktivitäten und Metabolitkonzentrationen den zugehörigen simulierten Werten gegenübergestellt.

Grundsätzlich ist die Beschreibung der Omics-Daten durch die im Zuge der Parameterschätzung gefundene optimale Lösung des vertikalen Netzwerkmodells zufriedenstellend, da in den meisten Fällen die richtige Größenordnung getroffen wurde. Ein Vergleich zwischen den Modellanpassungen an die unterschiedlichen Daten legt folgende Schlüsse nahe:

1. Insgesamt werden die Enzymaktivitätsmessungen am schlechtesten vom Modell wiedergegeben, gefolgt von den Metabolitkonzentrationen, Transkriptmengen und extrazellulären Raten.

2. Einzelne Daten bestimmter mRNAs (*pyk*, *icd*, *pyc*, *lysC*), Enzyme (*icd*) und Metabolite (SUC) sind über alle Experimente hinweg gut beschrieben. Jedoch sind bei einer Vielzahl der Datenpunkte in jeweils einem oder zwei der Experimente die Abweichungen zwischen Messung und Simulation sehr hoch ($> 100\%$).
3. Eine Ausnahme bilden die Messdaten der mRNAs für *acn*, *odh*, *sdh*, *mdh* und *gdh*, welche in allen Fällen deutlich niedriger simuliert werden. Interessanterweise sind die Aktivitäten der zugehörigen Enzyme passend bzw. zu hoch simuliert. Mögliche Ursachen sind hierfür unberücksichtigte regulatorische Effekte auf translatorischer oder post-translatorischer Ebene, die bei höherer Transkriptionsaktivität gleichzeitig zu niedrigeren Mengen an aktiven Enzym führen können.
4. Im Hinblick auf die Metabolom-Daten fällt auf, dass einige Metabolite wie PEP, PYR, SUC und AKG weitaus höher gemessen worden sind als es die Simulationen zeigen. Ein umgekehrtes Bild ergibt sich hingegen für die Cometabolite (AMP, ADP, ATP, NAD(H), NADP(H)). Im letzteren Fall könnte das Metabolit-Leakage eine Rolle spielen [BKL⁺07, WWR⁺08], d.h. aufgrund des Quenchingverlustes könnten die experimentell ermittelten Konzentrationsbereiche weit unterhalb der tatsächlichen *in vivo*-Werte liegen.

Eine Auflistung aller geschätzten Modellparameter sind in den Tabellen C.4 und C.5 des Anhangs zu finden. Die Mehrzahl dieser Parameter liegt dabei verteilt zwischen den festgelegten Bereichsgrenzen. Es gibt allerdings auch ein Reihe von Parametern, deren Werte gegen die obere oder untere Schranke gelaufen sind. Hierbei stellt sich natürlich die grundsätzliche Frage, inwieweit diese Parameter innerhalb der gefundenen optimalen Lösungsräume sensitiv und damit auf Grundlage der vorhandenen Messdaten bestimmbar sind. Einen Aufschluss darüber kann die Sensitivitätsanalyse geben, auf die im Abschnitt 9.5 näher eingegangen wird.

9.4 Simulationsergebnisse

In den Abbildungen 9.2 und 9.3 sind die auf Basis der Modellanpassungen resultierenden Stoffflussverteilungen für *C. glutamicum* Wildtyp und DM1730 dargestellt. Neben den Poolgrößen und Flüssen sind auch die quantifizierten regulatorischen Wechselwirkungen visualisiert. Dabei bildet das Quantifizierungsmaß für die

Tabelle 9.1: Vergleich zwischen gemessenen und simulierten extrazellulären Raten. Alle Angaben in $\text{mmol g}_{\text{BTM}}^{-1} \text{h}^{-1}$.

Rate	WT $\mu = 0.1 \text{ h}^{-1}$			DM1730 $\mu = 0.1 \text{ h}^{-1}$			DM1730 $\mu = 0.2 \text{ h}^{-1}$		
	gem.	sim.	Abw. [%] ^a	gem.	sim.	Abw. [%]	gem.	sim.	Abw. [%]
upt _{GLC}	1.57	1.56	-1.1	1.49	1.54	3.3	2.72	2.74	0.5
excl _{LYS}	0.00	0.01	851.5	0.19	0.19	-3.7	0.28	0.28	0.8
excc _{CO₂}	4.36	5.56	27.5	4.53	4.39	-3.1	8.69	7.27	-16.3

^a Auf Messdaten normierte Abweichungen zwischen gemessenen und simulierten Werten.

Effektor-Wechselwirkung die relative Stärke der momentanen Inhibierung oder Aktivierung eines einzelnen Effektors ab. Für ausführliche Erläuterungen zur Berechnung dieser Kenngröße für mechanistische Enzymkinetiken mit unterschiedlicher Art und Anzahl von Effektoren sei an dieser Stelle auf die Literatur verwiesen [NWQW07].

In Tabelle 9.5 sind die zugehörigen quantitativen Werte aller intrazellulären Stoffflüsse aufgeführt. Daneben sind auch potentielle Flusswerte für die unter metabolischer Kontrolle liegenden Reaktionen für den nicht regulierten Zustand enthalten. Damit lassen sich einerseits Aussagen ableiten, wie stark die betreffenden Flüsse in der kombinierten Wirkung von Aktivatoren und Inhibitoren auf metabolischer Ebene reguliert sind. Andererseits kann bei geringen Flusswerten differenziert werden, ob starke Inhibierungen, geringe Enzymaktivitäten oder niedrige Substrataffinitäten eine maßgebliche Rolle spielen.

Durch den Vergleich der verschiedenen Stoffflusskarten und Modellanpassungen lässt sich eine Reihe stammspezifischer Effekte feststellen, wodurch unmittelbar Hypothesen über die *in vivo*-Regulation aufgestellt werden können.

Folgende Gemeinsamkeiten zwischen dem Wildtyp und dem Lysinproduzenten DM1730 von *C. glutamicum* sind auffällig:

- Sehr hoher Fluss in den PPP (c5), welcher weit über den ermittelten Werten der Stoffflussanalyse mit DM1730 (vgl. Kap. 5) bzw. vergleichbaren Literaturangaben [MdGW⁺96, MSdG⁺97, PMdG⁺01, WH02, DEMM⁺03, DEMW⁺04, KSK⁺04, KHZW04, YWH06, BKW08] liegt. Demgegenüber ist der Fluss entlang des EMP (c3) und TCA (gltA, acn usw.) deutlich niedriger.
- Trotz hoher Enzymaktivitäten ($v_{\text{max}} > 29 \text{ mmol g}_{\text{BTM}}^{-1} \text{h}^{-1}$) ist kein Fluss über das Enzym Pyruvat-Kinase (pk) beobachtbar. Im Wildtyp ist dies bedingt durch die Kombination von niedriger PEP-Konzentration ($c_{\text{PEP}} < 0.05 \text{ mM}$) und

Tabelle 9.2: Vergleich zwischen gemessenen und simulierten mRNA-Daten. Die grau unterlegten Einträge beinhalten die gemessenen Transkriptmengen einzelner Gene, welche für Enzymkomplexe kodieren. Die zugehörigen simulierten Mittelwerte sind in der jeweils untersten Zeile aufgeführt. Alle Angaben in $\text{mmol mmol}_{\text{RNA}}^{-1}$.

mRNA	WT $\mu = 0.1 \text{ h}^{-1}$			DM1730 $\mu = 0.1 \text{ h}^{-1}$			DM1730 $\mu = 0.2 \text{ h}^{-1}$		
	gem.	sim.	Abw. [%]	gem.	sim.	Abw. [%]	gem.	sim.	Abw. [%]
pyk	3195	3388	6	3186	3626	14	2944	2467	-16
aceE	4594			3991			3602		
lpd	2154	737	-78	2210	3446	11	1863	1759	-36
gltA	3637	3373	-7	2831	3213	13	3235	4038	25
acn	2790	561	-80	4234	3383	-20	3830	187	-95
icd	1723	1627	-6	1646	1693	3	1776	1679	-5
odhA	2706			2006			2367		
lpd	2154			2210			1863		
sucB	2242	1620	-32	2258	1320	-39	1897	1438	-30
sucC	2065			3783			2187		
sucD	1442	2080	19	2470	2497	-20	1637	1541	-19
sdhA	2584			2795			3561		
sdhB	1994			2135			2945		
sdhCD	3160	1356	-47	3104	2121	-21	3734	2202	-35
fum	2220	2799	26	1977	1576	-20	1828	1147	-37
mdh	2019	0	-100	1743	1	-100	1676	1	-100
mgo	946	1234	30	827	1517	83	1103	971	-12
pyc	915	825	-10	629	601	-4	978	811	-17
ppc	1478	859	-42	1334	1398	5	1237	465	-62
pck	3271	4198	28	212	246	16	178	115	-36
mez	1177	1655	41	2543	3264	28	2250	3870	72
gdh	1830	112	-94	1958	605	-69	1746	295	-83
aspB	667	654	-2	680	697	3	694	453	-35
lysC	1195	1127	-6	1177	1114	-5	1255	1174	-6
ω_i^{-1a}		3413			4927			10796	

^aSkalierungsfaktor w aus Gl. 9.1.

niedriger Affinität des Enzyms gegenüber PEP ($K_{\text{PEP}} = 2.8 \text{ mM}$). Im Stamm DM1730 ist die Ursache in der starken Inhibierung durch ATP ($K_{\text{I,ATP}} = 0.43 \text{ mM}$) zu finden.

- Starke Regulation der TCA-Flüsse Citrat-Synthase (cs), α -Ketoglutarat-Dehydrogenase (odh) und Fumarase (fum) mit jeweiligen Inhibierungen durch ATP, SUCCOA und ATP.
- Kein Fluss in Vorwärts- und Rückwärtsrichtung über das Enzym Malat-Dehydrogenase (mdh). Da keine Regulation vorliegt und das Enzym Aktivität in der thermodynamisch bevorzugten Richtung (MAL \rightarrow OAA) zeigt ($v_{\text{max}}^{\rightarrow} > 0.7 \text{ mmol g}_{\text{BTM}}^{-1} \text{ h}^{-1}$), ist die sehr niedrige Affinität gegenüber MAL

Tabelle 9.3: Vergleich zwischen gemessenen und simulierten Enzymaktivitäten. Alle Angaben in $\text{mmol g}_{\text{BTM}}^{-1} \text{h}^{-1}$.

Enzym	WT $\mu = 0.1 \text{ h}^{-1}$			DM1730 $\mu = 0.1 \text{ h}^{-1}$			DM1730 $\mu = 0.2 \text{ h}^{-1}$		
	gem.	sim.	Abw. [%]	gem.	sim.	Abw. [%]	gem.	sim.	Abw. [%]
pyk \rightarrow^a	-	42.01	-	-	31.14	-	-	29.44	-
pdh \rightarrow	1.17	4.83	313	-	15.63	-	1.47	11.09	654
gltA \rightarrow	21.54	15.15	-30	-	9.99	-	20.64	17.46	-15
acn \leftarrow	6.57	71.49	988	15.90	298.74	1779	11.25	22.90	104
icd \rightarrow	40.44	45.57	13	48.30	32.85	-32	37.17	45.28	22
odh \rightarrow	0.83	2.70	226	1.42	1.53	7	0.68	2.31	240
scs \leftarrow	2.25	3.39	51	1.38	2.82	105	2.28	2.42	6
sdh \rightarrow	0.42	0.78	85	0.27	0.84	211	0.33	1.21	267
fum \rightarrow	32.70	122.43	274	-	47.77	-	51.00	48.32	-5
mdh \rightarrow	1.05	0.73	-30	2.94	1.69	-43	1.98	2.10	6
mgo \rightarrow	0.09	0.43	381	0.09	0.37	309	0.09	0.33	264
pyc \rightarrow	-	47.21	-	-	23.81	-	-	44.65	-
ppc \rightarrow	7.65	31.19	308	6.93	35.15	407	7.77	16.25	109
odx \rightarrow	10.44	42.98	312	12.60	9.17	-27	16.32	31.44	93
pck \leftarrow	0.69	0.01	-99	0.00	0.00	-47	0.00	0.00	21
mez \rightarrow	0.15	0.44	193	0.15	0.60	301	0.33	0.99	200
gdh \rightarrow	100.50	117.70	17	130.50	442.55	239	156.60	299.52	91
gdh \leftarrow	15.00	0.00	-100	15.90	0.01	-100	15.90	0.00	-100

^aRichtung der Aktivitätsmessung im jeweiligen Assay im Vergleich zur definierten Nettorichtung im Modell.

($K_{\text{PEP}} = 29.9 \text{ mM}$) dafür verantwortlich.

- Hoher Fluss über das anaplerotische Enzym Pyruvat-Carboxylase (pyc), wodurch der TCA mit C4-Körpern aufgefüllt wird. Dennoch wird pyc stark durch ACCOA, AMP und ADP inhibiert, weshalb der theoretisch maximale Fluss um den Faktor 30 im Wildtyp bzw. 10 im Typ DM1730 höher ist.
- Das Enzym Oxalacetat-Decarboxylase (odx) weist eine hohe Aktivität ($v_{\text{max}}^{\rightarrow} > 10 \text{ mmol g}_{\text{BTM}}^{-1} \text{h}^{-1}$) auf, allerdings ist die Affinität zu OAA ($c_{\text{OAA}} < 0.05 \text{ mM}$) sehr gering ($K_{\text{OAA}} = 1.99 \text{ mM}$). Die gleichzeitig starke Inhibition durch SUC und ADP spielt demnach eine untergeordnete Rolle.

Darüber hinaus sind folgende Unterschiede zwischen dem Wildtyp und dem Lysinproduzenten DM1730 wesentlich:

- Bezogen auf die Glucoseaufnahme ist der Fluss in den PPP (c5) und TCA (gltA) im Stamm DM1730 niedriger als im Wildtyp. Dementsprechend fließt ein größerer Anteil entlang des EMP (c3) bis zum ersten Knotenpunkt PEP (vgl. Tab. 9.5)

Tabelle 9.4: Vergleich zwischen gemessenen und simulierten Metabolitkonzentrationen. Alle Angaben in mM.

Metabolit	WT $\mu = 0.1 \text{ h}^{-1}$			DM1730 $\mu = 0.1 \text{ h}^{-1}$			DM1730 $\mu = 0.2 \text{ h}^{-1}$		
	gem.	sim.	Abw. [%]	gem.	sim.	Abw. [%]	gem.	sim.	Abw. [%]
PEP	0.24	3.3e-8	-100	0.10	0.02	-84	0.06	0.03	-58
PYR	2.91	0.51	-82	1.59	0.12	-93	1.00	0.28	-72
ACCOA	0.13	0.31	137	0.04	0.18	345	0.06	0.31	423
CIT	-	0.26	-	-	0.05	-	-	0.88	-
ICIT	-	1.9e-4	-	-	2.1e-4	-	-	2.2e-4	-
CITICIT ^b	-	0.26	-	0.70	0.05	-93	0.55	0.88	59
AKG	4.06	0.15	-96	0.96	0.26	-73	0.46	0.28	-40
SUCCOA	-	30.94	-	-	25.96	-	-	9.98	-
SUC	7.81	4.43	-43	1.16	1.24	7	1.30	1.08	-17
FUM	-	0.23	-	-	0.06	-	-	0.14	-
MAL	1.44	0.84	-41	0.30	0.21	-31	0.24	0.50	109
OAA	-	3.6e-4	-	-	4.5e-4	-	-	2.3e-4	-
ASP	1.04	0.01	-99	1.06	0.27	-74	-	0.15	-
FBP	6.55	6.59	1	0.97	0.84	-14	0.65	2.49	283
AMP	0.28	0.91	226	0.10	0.27	166	0.48	0.23	-51
ADP	2.18	3.82	75	0.47	1.02	116	1.41	1.45	3
ATP	14.81	15.58	5	3.47	12.41	258	2.59	6.88	166
NAD	2.15	2.26	5	0.93	1.60	72	1.33	1.57	18
NADH	0.00 ^a	0.04	1202	0.00 ^a	5.0e-3	81	0.00 ^a	0.63	22897
NADP	1.77	1.74	-2	0.48	1.19	148	0.62	0.87	40
NADPH	0.00 ^a	0.03	966	0.00 ^a	0.01	332	0.01	0.05	444

^aMesswert unterhalb der Nachweisgrenze.

^bSummenpool aus CIT und ICIT.

- Die damit verbundene Verfügbarkeit von PEP im DM1730 bedingt das Zuschalten des zweiten anaplerotischen Enzyms PEP-Carboxylase (ppc). Dieses Enzym wird zwar durch GLU und ASP inhibiert, unterliegt aber gleichzeitig einer starken Aktivierung durch ACCOA. Die resultierende Wirkung ist somit eine schwache Aktivierung.
- Als Folge des stärkeren Zuflusses in OAA durch pyc und des gleichzeitig verringerten Abflusses in Richtung TCA (gltA) wird die vermehrte Lysinsynthese im DM1730 ermöglicht.
- Der geringe Fluss über die PEP-Carboxykinase (pck) im Wildtyp in gluconeogenetischer Richtung (OAA→PEP) resultiert aus der starken Inhibierung durch ATP.
- Der fehlende pck-Fluss im DM1730 resultiert aus der fehlenden Enzymaktivität (simulierte Proteinmenge $c_{\text{pck}} < 2.3 \text{ mM}$, Wildtyp: $c_{\text{pck}} = 56.5 \text{ mM}$) auf-

Tabelle 9.5: Schätzwerte intrazellulärer Stoffflüsse des vertikalen Netzwerkmodells.

	Fluss [mmol g _{BTM} ⁻¹ h ⁻¹]			Anteil der Glucoseaufnahme [%]			unregulierter Fluss [%]		
	WT	DM1730		WT	DM1730		WT	DM1730	
μ [h ⁻¹]	0.1	0.1	0.2	0.1	0.1	0.2	0.1	0.1	0.2
c5	3.37	2.57	4.68	216.6	167.0	171.3	-	-	-
c3	1.61	1.84	3.15	103.4	119.7	115.1	-	-	-
pyk	0.00	0.01	0.02	0.0	0.4	0.6	0.0	10.6	9.1
pdh	0.95	0.83	1.31	61.2	53.8	47.8	-	-	-
gltA	0.66	0.53	0.72	42.3	34.7	26.3	86.3	63.5	38.3
acn	0.66	0.53	0.72	42.3	34.7	26.3	-	-	-
icd	0.66	0.53	0.72	42.3	34.7	26.3	43.5	35.6	28.3
odh	0.53	0.41	0.47	34.3	26.6	17.1	119.5	75.9	65.3
suc	0.53	0.41	0.47	34.3	26.6	17.1	-	-	-
sdh	0.53	0.41	0.47	34.3	26.6	17.1	34.3	26.6	17.1
fum	0.53	0.41	0.47	34.3	26.6	17.1	437.1	275.6	106.1
mgo	0.43	0.37	0.33	27.8	23.9	12.0	-	-	-
mdh	0.00	0.00	0.00	0.0	0.0	0.0	-	-	-
pck	0.00	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.0
ppc	0.00	0.24	0.29	0.0	15.8	10.7	0.0	13.1	5.3
pyc	0.41	0.28	0.73	26.2	18.4	26.6	790.3	116.0	261.9
mez	0.10	0.04	0.14	6.5	2.7	5.2	-	-	-
odx	0.00	0.00	0.00	0.1	0.1	0.1	0.5	0.1	0.1
gdh	0.18	0.36	0.63	11.6	23.3	22.9	-	-	-
aspB	0.18	0.36	0.63	11.6	23.3	22.9	-	-	-
bsLYS	0.03	0.21	0.32	1.9	13.5	11.8	-	-	-

grund der eingebrachten Gendelektion, welche durch das Modell korrekt vorhergesagt wird.

Zusammenfassend ist festzuhalten, dass *C. glutamicum* einer starken Regulation auf metabolischer Ebene im Bereich des TCA und der Anaplerose unterliegt. Die teilweise geringen Flüsse über einige Enzyme sind jedoch in vielen Fällen durch die Verfügbarkeit der Substrate und die sehr niedrigen Affinitäten zu diesen verursacht.

9.5 Sensitivitätsanalyse

Wie in den vorangegangenen Abschnitten erläutert, können auf der Grundlage eines mechanistischen Stoffwechselmodells eine Vielzahl von Wechselwirkungen zwischen den Systemvariablen und -parametern identifiziert werden, die gleichzeitig zu einem verbesserten Verständnis des untersuchten Organismus führen. Darüber hinaus können diese Modelle auch zur Vorhersage von gezielten molekularbiologischen Eingriffen (Metabolic Engineering) für die Verbesserung der Produktionsleistungen des jeweiligen Mikroorganismus genutzt werden.

Zu diesem Zweck wurden in der Vergangenheit eine Reihe verschiedener mathematischer Analysemethoden für biochemische Netzwerke entwickelt, welche in der Literatur oft unter dem Schlagwort „Metabolic Control Analysis“ (MCA) zusammengefasst sind. Die ersten Ansätze der MCA waren dabei ausschließlich auf die Verwendung experimenteller Daten beschränkt. Es wurden Quantifizierungsmaße, sogenannte Kontrollkoeffizienten, entwickelt, um den Einfluss von Änderungen in den Enzymmengen auf Flüsse und intrazelluläre Metabolite zu untersuchen [KB73,HR74].

Die Vorhersagekraft dieser experimentell ermittelten Koeffizienten ist jedoch teilweise stark eingegrenzt. In der Regel treten infolge von genetischen Manipulationen große Änderungen auf, die nicht nur auf die Menge an aktivem Enzym beschränkt sind, sondern auch die Substrataffinität und die regulatorischen Einflüsse von Effekten betreffen.

Aus diesem Grund stellt die Ermittlung von Kontrollkoeffizienten mit Hilfe eines strukturierten Stoffwechselmodells eine sinnvolle Alternative dar [MAR97]. Dabei werden, wie im Fall der linearen Regression (vgl. Kap. 5.7.1), auch Informationen über die Sensitivitäten einzelner Modellparameter benötigt.

Für die Berechnung von Sensitivitätsfunktionen auf Basis des vertikalen Netzwerkmodells wurde das bereits vorgestellte Werkzeug ADModelica genutzt (vgl. Abschn. 4.4).

In Abbildung 9.4 ist eine Matrix von Kontrollkoeffizienten aller Metabolitkonzentrationen des vertikalen Netzwerkmodells in visueller Form dargestellt [NWH⁺08]. Per Definition haben die experimentspezifischen Modellparameter nur einen Einfluss auf die Werte der Systemvariablen, die sich in Abhängigkeit der jeweils experimentell beobachteten metabolischen Stationarität ergeben. Darüber hinaus werden die externen Poolgrößen (ACCOA, AKG, AMP, NAD usw.) ausschließlich durch die Wertzuweisung des jeweiligen Modellparameters bestimmt (Einzeleinträge in Matrixzeilen). Hingegen beeinflussen die organismenspezifischen Parameter alle möglichen stationären Systemzustände.

Desweiteren ist zu erkennen, dass in den stationären Zuständen, welche die Experimente für den Stamm DM1730 charakterisieren, nur etwa ein Drittel aller experimentspezifischen Parameter (vgl. Tab. C.4) eine signifikante Sensitivität in Bezug auf die Poolgrößen zeigt. Auch sind es bei beiden Wachstumsraten nahezu dieselben Parameter, die einen vergleichbaren Einfluss ausüben. Diese Untermenge an sensitiven Zustandsparametern ist ebenfalls für das Wildtypexperiment maßgeblich, jedoch sind hier noch eine Reihe anderer Parameter relevant.

Vergleichbare Muster zwischen den einzelnen stationären Zuständen ergeben sich auch für die organismenspezifischen Parameter. Allerdings ist auch hier anzumerken, dass ca. die Hälfte aller Parameter eine sehr geringe Sensitivität besitzt. Dies betrifft insbesondere die Parameter, welche im Zuge der Optimierung gegen die untere oder obere Schätzgrenze gelaufen sind.

Insgesamt sind somit lediglich 102 der 190 Modellparameter im Bereich der untersuchten metabolisch stationären Zustände sensitiv. Für die übrigen Parameter lässt sich mit hoher Wahrscheinlichkeit ableiten, dass diese auf Grundlage der vorhandenen Messdaten bzw. im abgedeckten Werteraum abhängiger Zustandsvariablen nicht identifizierbar sind.

9.5.1 Fokus Lysinbiosynthese

Um der Fragestellung nachzugehen, welche Modellparameter einen nennenswerten Einfluss auf die Lysinbiosynthese besitzen, wurden nachfolgend gezielt die Kontrollkoeffizienten für das simulierte Experiment von *C. glutamicum* DM1730 bei $\mu = 0.2 \text{ h}^{-1}$ untersucht (Abb. 9.5). Insbesondere sind hierbei diejenigen Parameter wichtig, welche die Konzentrationen an Gen-Polymerasekomplexen (*GP*) symbolisieren (vgl. Modellansatz nach Gl. 8.4). Eine Veränderung in diesen Parametern ist demnach gleichbedeutend mit einer erhöhten bzw. verringerten Expression betreffender Gene. Darüber hinaus sind auch die externen Parameter interessant, da diese Zustandsvariablen des Modells beschreiben, die nicht durch Massenbilanzen im Modell integriert sind.

Für eine Steigerung der Lysinproduktion sind neben der Verfügbarkeit der Lysinprecursor PYR und ASP auch die Wechselwirkungen aufgrund vorhandener metabolischer Regulation zu berücksichtigen.

Auf Basis der Kontrollanalyse können folgende Vorhersagen getroffen werden:

1. **Citrat-Synthase (gltA)** $\downarrow \leftrightarrow$ geringere Umsetzung an OAA im TCA und damit höhere Verfügbarkeit für die Bildung von ASP via Aspartat-Aminotransferase (aspB)
2. **Pyruvat-Dehydrogenase (pdh)** $\downarrow \leftrightarrow$ geringerer Verbrauch an PYR in Richtung TCA und damit höhere Verfügbarkeit für die Lysinbiosynthese (bsLYS), in der ASP umgesetzt wird (negativer Einfluss)
3. **AKG-Dehydrogenase (odh)** $\uparrow \leftrightarrow$ da Enzym starker Regulation unterliegt, bewirkt ein Anstieg der aktiven Enzymmenge eine Erhöhung des Flusses entlang

des TCA bis in den Precursor ASP (negativer Effekt auf GLU). Gleichzeitig muss der höhere stationäre TCA-Fluss durch einen vermehrten Zufluss durch den EMP realisiert werden, wodurch der positive Einfluss auf die Metabolitpools PEP und PYR erklärt wird.

4. **GLU-Dehydrogenase (gdh)** $\uparrow \leftrightarrow$ erhöhte Bildung von GLU, welcher als Aminogruppendonor der Aspartat-Aminotransferase (aspB) fungiert

Daneben stecken in den Sensitivitäten eine Vielzahl von Informationen über Wirkungsketten im Zusammenhang mit regulatorischen Effekten, wie beispielsweise

- $gdh \uparrow \leftrightarrow GLU \uparrow \leftrightarrow$ Inhibierung der PEP-Carboxylase ($ppc \uparrow \leftrightarrow PEP \uparrow$)
- $AKG, NADPH \uparrow$ bzw. $NADP \downarrow \leftrightarrow gdh \uparrow \leftrightarrow$ s.o.

Insbesondere im Fall des simulierten Wildtyp-Experiments bildet das vertikale Modell damit auch eine gute Ausgangsbasis für die Optimierung anderer Stoffwechselprodukte mit z.B. PEP oder AKG als Precursormolekülen.

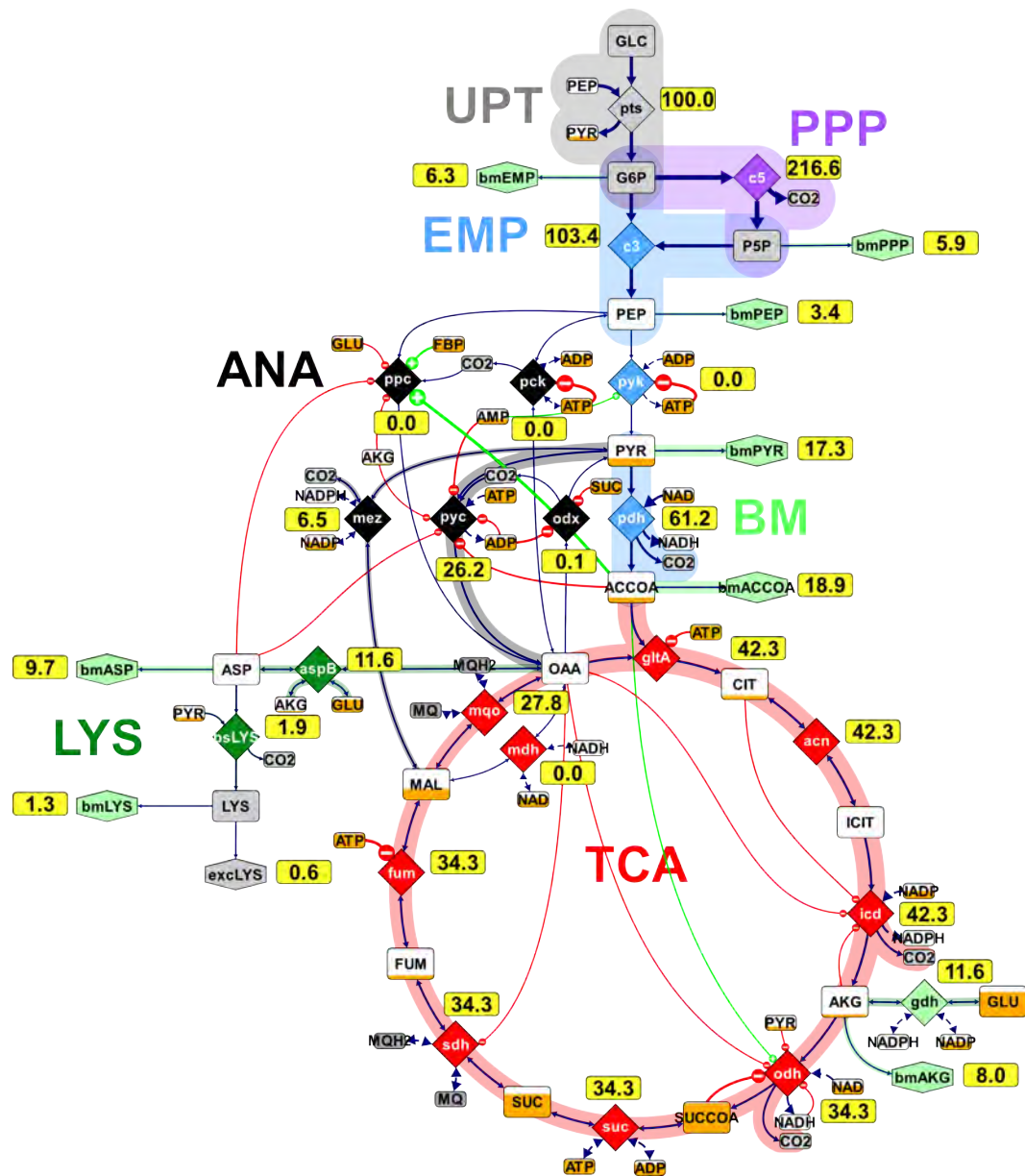


Abbildung 9.2: Geschätzte intrazelluläre Stoffflussverteilung für *C. glutamicum* Wildtyp bei einer Wachstumsrate von $\mu = 0.1 \text{ h}^{-1}$ auf Basis des vertikalen Netzwerkmodells. Die angegebenen Werte stellen die molaren Anteile der Glucoseaufnahme dar. Neben den Flüssen sind auch die simulierten Messungen der Poolgrößen über den Füllstand der Boxen bzw. die quantifizierten regulatorischen Wechselwirkungen über die Größe der Effektorsymbole visualisiert.

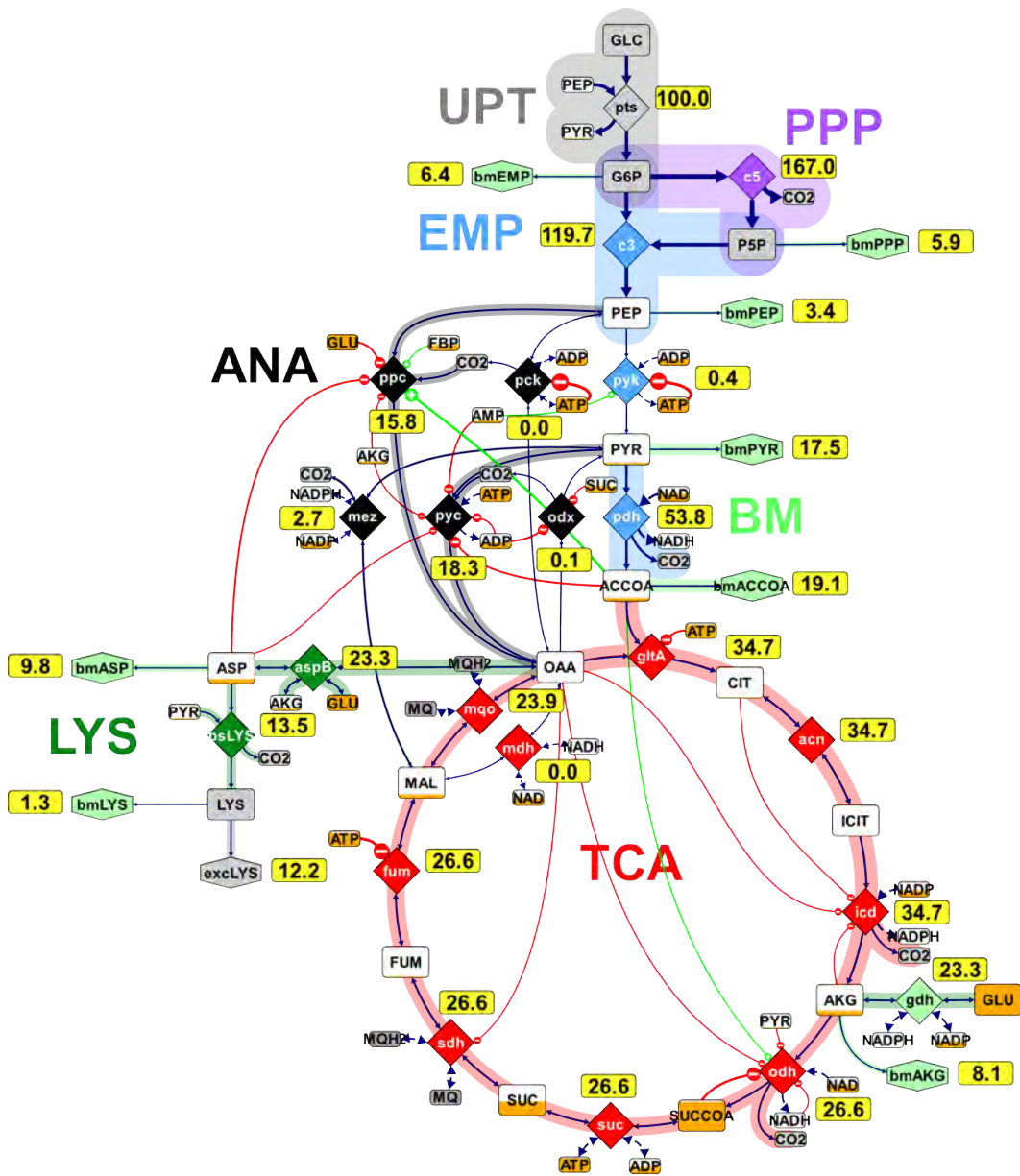


Abbildung 9.3: Geschätzte intrazelluläre Stoffflussverteilung für *C. glutamicum* DM1730 bei einer Wachstumsrate von $\mu = 0.1 \text{ h}^{-1}$ auf Basis des vertikalen Netzwerkmodells. Angegebene Werte stellen molare Anteile der Glucoseaufnahme dar.

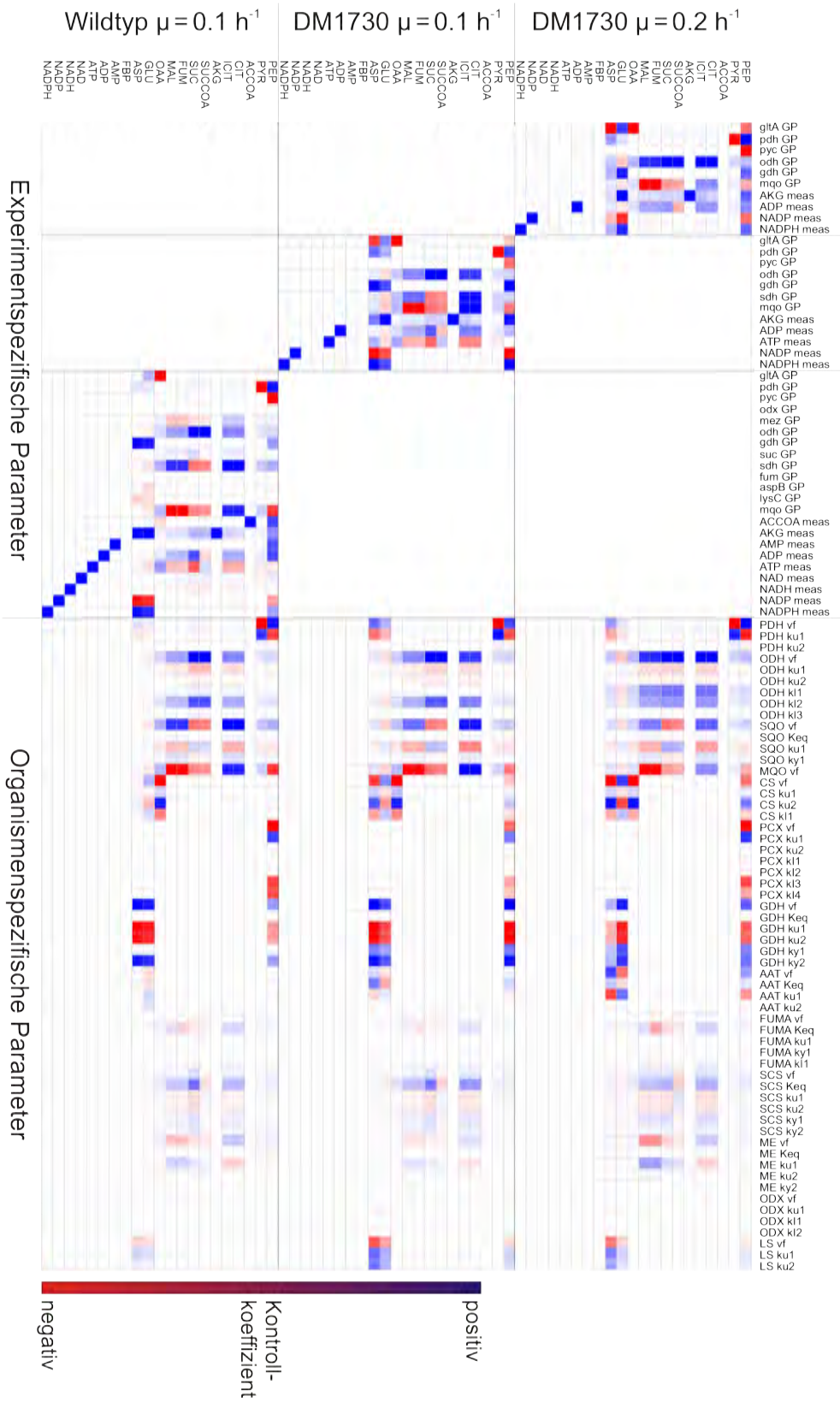
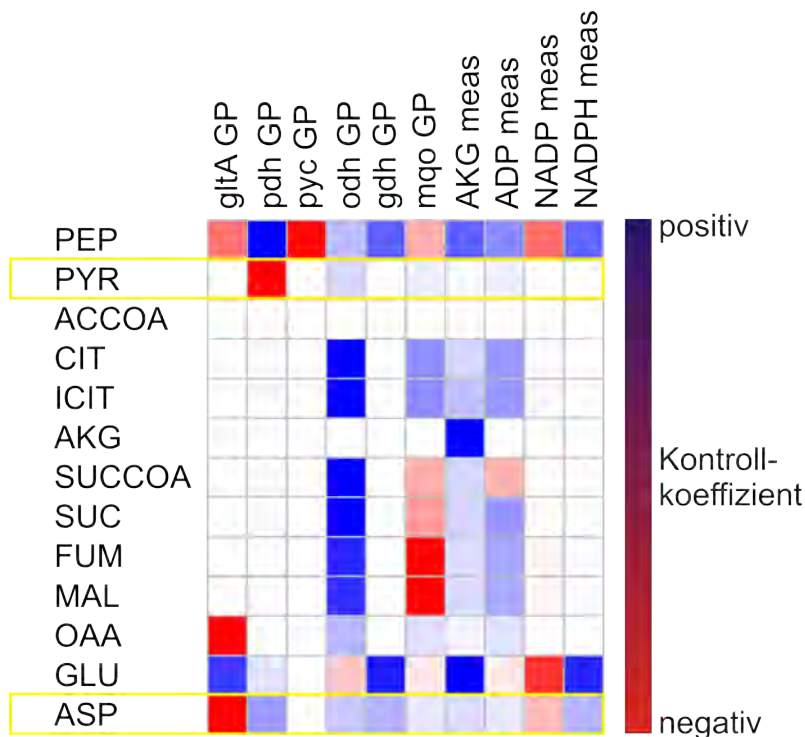
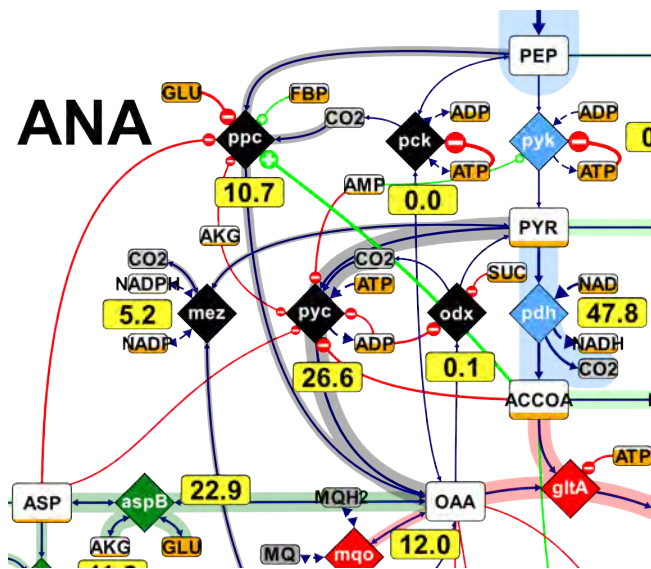


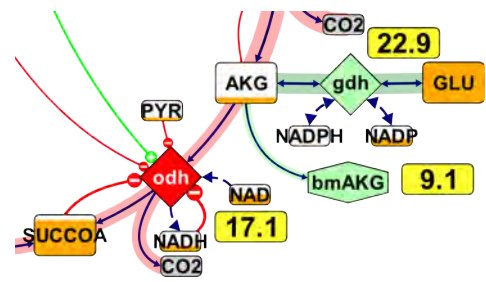
Abbildung 9.4: Visualisierte Matrix ausgewählter Konzentrations-Kontrollkoeffizienten des vertikalen Netzwerkmodells. Für die Visualisierung wurden die Matrixeinträge mit Hilfe der Zellennorm ($\|x_i\|_\infty = \max_j (C_{ij}^M)$) $i = 1 \dots m, j = 1 \dots n$ skaliert. Modellparameter, die keine bzw. eine vernachlässigbare Sensitivität gegenüber allen Poolgrößen zeigten, wurden entfernt.



(a) Kontrollkoeffizienten experimentspezifischer Parameter.



(b) Ausschnitt der simulierten Stoffflussverteilung um den PEP-PYR-OAA-Knotenpunkt.



(c) Ausschnitt der simulierten Stoffflussverteilung um den AKG-GLU-Knotenpunkt.

Abbildung 9.5: Untersuchung des Einflusses experimentspezifischer Parameter auf die Lysinbiosynthese für das simulierte Experiment von DM1730 bei einer Wachstumsrate von $\mu = 0.2 \text{ h}^{-1}$.

Kapitel 10

Zusammenfassung, Diskussion und Ausblick

Gegenstand dieser Arbeit war die Entwicklung und Anwendung von Methoden zur modellgestützten Auswertung von Multi-Omics-Daten. Als Modellsystem diente dabei der gut untersuchte Organismus *C. glutamicum*, wobei speziell die Modellierung des Zentralstoffwechsels (ZSW) im Mittelpunkt des Interesses stand.

Auf der Grundlage von Experimenten mit unterschiedlichen *C. glutamicum*-Stämmen (Wildtyp, Lysinproduzenten) wurden durch Anwendung verschiedener Omics-Technologien Transkriptom-, Proteom-, Metabolom- und Fluxom-Daten integrativ ausgewertet.

Um aus den experimentellen Rohdaten die für eine Modellierung notwendigen konsistenten Nutzdaten zu gewinnen, sind eine Reihe von Schritten der Präprozessierung notwendig. Dies betrifft insbesondere die Ermittlung metabolischer Stoffflussdaten, da hier bereits modellgestützte Methoden zum Einsatz kommen.

Den ersten Schwerpunkt dieser Arbeit bildete damit die Untersuchung zweier verschiedener Ansätze zur ^{13}C -Stoffflussanalyse im Zentralstoffwechsel von *C. glutamicum*.

10.1 Gewinnung quantitativer Fluxom-Daten

Hierbei wurde der Organismus *C. glutamicum* erstmalig im Rahmen eines isotopisch instationären Markierungsexperiments im Fed-Batch-Ansatz untersucht. Auf diese Weise wurden sowohl isotopisch transiente als auch stationäre Markierungsanreicherungen von Zentralstoffwechselmetaboliten mittels LC-MS-Analytik erzeugt.

Dieser Messdatensatz wurde durch Konzentrationsmessungen der intrazellulären Metabolite sowie durch extrazelluläre Flussmessungen komplettiert, welche auf der Basis eines Prozessmodells geschätzt wurden.

Instationäres Markierungsexperiment

Aus dem durchgeführten Markierungsexperiment können zunächst folgende Schlussfolgerungen in Bezug auf die Markierungsdynamik im untersuchten Organismus sowie die Qualität der erhaltenen Rohdaten gezogen werden:

- Im Hinblick auf die Zeitkonstanten der Markierungsanreicherungen von Intermediaten des EMP, PPP, TCA und der Aminosäurebiosynthesewege sind deutliche Unterschiede erkennbar. Von besonderer Bedeutung ist dabei die stark verzögerte Dynamik im TCA, hier erfolgte ein Markierungseintrag erst nach ca. 5 Sekunden. Dieser Effekt scheint nicht organismenspezifisch zu sein, da er bereits auch im ersten Experiment dieser Art mit *E. coli* beobachtet wurde [NGL⁺07].
- Mit Ausnahme einiger Aminosäuren konnten für eine Vielzahl intrazellulärer Metabolite reproduzierbare Massenisotopomeranteile für die instationäre als auch stationäre ¹³C-MFA ermittelt werden. Hierbei sei angemerkt, dass die Auswertung von Massenspektren nach wie vor manuell geschieht und insbesondere im Fall zeitlich hochaufgelöster transienter Isotopendaten einen sehr zeitintensiven Arbeitsschritt darstellt. Im konkreten Fall wurden 33 Proben hinsichtlich der Spektren von 357 Massenisotopomeren (einschließlich Fragmentierungen) analysiert und damit insgesamt 12111 Einzelmessungen generiert.
- Für die Berechnung von Mess-Standardabweichungen der stationären Markierungsdaten konnten mehrere Verlaufsproben am Ende des Experimentes genutzt werden. Aufgrund des hohen Aufwandes bei der Erzeugung transienter Messdaten war eine Auswertung von Probenreplikaten in diesem Fall nicht möglich. Alternativ erfolgte hier eine Schätzung zugehöriger Standardabweichungen mit Hilfe einer Fehlerübertragungsrechnung auf Basis der ermittelten Fehler einzelner stationärer Massenspektren.
- Mit Hilfe des formulierten Prozessmodells, welches den Fed-Batch-Verlauf auf makroskopischer Ebene abbildet, wurden die extrazellulären Flussmessungen

mit realistischen Konfidenzintervallen geschätzt. Somit konnte auf eine fehleranfällige Dateninterpolation und stückweise Raten-Approximation verzichtet werden.

Insgesamt konnte damit eine Datenbasis geschaffen werden, die einen direkten Vergleich beider Methoden zur ^{13}C -MFA ermöglichte.

Vergleich isotopisch stationärer und instationärer MFA

Die im Zuge des Markierungsexperiments generierten Rohdaten wurden auf der Basis von isotopisch stationären und instationären Modellierungsansätzen unter Verwendung verschiedener Netzwerk- und Messkonfigurationen ausgewertet. Zur Bestimmung von Konfidenzintervallen der berechneten Stoffflüsse kamen Verfahren der linearen und nichtlinearen Statistik zum Einsatz. Unter dem vorrangigen Ziel der Gewinnung vertrauenswürdiger quantitativer Fluxom-Daten sind die erhaltenen Ergebnisse wie folgt zu bewerten:

- Es wurde eine signifikante Abhängigkeit der ermittelten Stoffflussverteilungen von der Art des zugrunde liegenden stöchiometrischen Netzwerks im Fall des isotopisch stationären Modellansatzes festgestellt. Hierbei ergaben sich große Unterschiede in den quantifizierten Stoffflüssen zwischen einem fokussierten Netzwerk des ZSW sowie einem erweiterten Modell mit Aminosäurebiosynthesewegen. Die Wahl der Messkonfiguration, d.h. ausschließlich Markierungsmessungen von ZSW-Intermediaten bzw. zusätzliche Hinzunahme von Aminosäuremessungen, hatte demgegenüber einen geringen Einfluss. Da in beiden Fällen das Verhältnis zwischen Anzahl an Freiheitsgraden (23 bzw. 30 Flussparameter) und der zur Verfügung stehenden Messinformation (65 bzw. 102 Markierungsmessungen) vergleichbar war, liegt der Schluss nahe, dass einerseits die zusätzlichen stöchiometrischen Kopplungen im erweiterten Netzwerk einen wesentlichen Einfluss auf die Schätzung der Stoffflüsse besitzen. Andererseits existieren im vorliegenden Fall mindestens zwei verschiedene stationäre Stoffflusslagen, welche die Messdaten in gleicher Güte reproduzieren können.
- Darüber hinaus waren die ermittelten quantitativen Stoffflusswerte von der Art des Modellierungsansatzes (isotopisch (in)stationär) abhängig. Obwohl teilweise vergleichbare Verteilungen der Flüsse in einzelnen Stoffwechselabschnitten (z.B. EMP, PPP) gefunden wurden, sind die zugehörigen Flusswerte

im Rahmen ihrer statistischen Genauigkeit teilweise um den Faktor 2 höher.

Eine Ursache dafür sind die verhältnismäßig hohen Messfehler der geschätzten extrazellulären Raten für die Glucoseaufnahme und CO₂-Produktion ($\hat{\sigma}_{\text{GLC}} = 0.69$, $\hat{\sigma}_{\text{CO}_2} = 1.59$), welche zu deutlich unterschiedlichen Anpassungen dieser Messdaten innerhalb beider Ansätze führte. Dadurch ist u.a. im Bereich des EMP und PPP die quantitative Stoffflusslage nach oben bzw. unten verschoben, während das Split-Verhältnis am G6P-Knotenpunkt vergleichbar ist.

- Die beiden eingesetzten Verfahren zur Schätzung von Fluss-Konfidenzintervallen (Linearisierungsansatz bzw. Monte-Carlo-Simulationen) zeigten eine gute Übereinstimmung in den Ergebnissen des fokussierten Netzwerkmodells im isotopisch stationären Ansatz. Darüber hinaus bildet die Monte-Carlo-Methode auch die Nichtlinearitäten des Isotopomermodells (bimolekulare Reaktionen) ab, wodurch für einen Teil der TCA-Flüsse zusätzliche Standardabweichungen ermittelt werden konnten. Aufgrund der hohen Rechenintensität kann diese Methode derzeit nicht für das erweiterte Netzwerk sowie das im instationären Ansatz vorliegende Differentialgleichungsmodell angewendet werden.

Letztendlich wurden auf Basis der hier durchgeführten ¹³C-Stoffflussanalysen keine eindeutigen Ergebnisse in Bezug auf die quantitativen Stoffflüsse von *C. glutamicum* unter den gewählten experimentellen Rahmenbedingungen und verwendeten Modellierungsansätzen erhalten.

Interpretation geschätzter Stoffflüsse

Um der Frage nachzugehen, welche der ermittelten Stoffflusslagen am ehesten der Realität des zugrunde liegenden metabolischen Zustandes der Zelle entspricht, wurde ein publiziertes genomweites Stoffwechselmodell für *C. glutamicum* weiterentwickelt und für die Interpretation der erhaltenen Stoffflussverteilungen auf Grundlage von Flussbilanzanalysen genutzt. Hierbei ist Folgendes festzuhalten:

- Das erstellte genomweite Netzwerkmodell umfasst 399 Metabolite und 454 Reaktionen und ist damit gegenwärtig eines der größten verfügbaren stöchiometrischen Modelle für *C. glutamicum*. Im Vergleich dazu umfasst das fokussierte Netzwerk zur ¹³C-MFA gerade einmal 22 Metabolite und 44 Reaktionen.

- Durch den Vergleich simulierter Phänotypen mit den zuvor berechneten Stoffflusslagen konnte die Lösung der ^{13}C -MFA auf Basis des isotopisch stationären Ansatzes im fokussierten Netzwerk ausgeschlossen werden.

Da die verfügbare Messinformation zur Schätzung aller Modellparameter im instationären Ansatz bedeutend höher war und dadurch auch die statistische Sicherheit der Fluss schätzung anstieg, wird die hieraus resultierende Lösung als am wahrscheinlichsten angesehen.

10.2 Integrative mechanistische Modellierung

Die modellgestützte Auswertung von Multi-Omics-Daten bildete den zweiten Schwerpunkt innerhalb dieser Arbeit. Dabei wurde ein neuartiges Modellierungskonzept entwickelt, welches auf einem vertikalen Netzwerkansatz beruht. Die Implementierung eines fokussierten Modells des Zitratzyklus und der Anaplerose von *C. glutamicum* erfolgte mit Hilfe der Modellierungssprache Modelica. Auf Grundlage der zuvor erzeugten experimentellen Daten mit verschiedenen *C. glutamicum*-Stämmen wurde anschließend eine Validierung dieses Modells durchgeführt.

Im Ergebnis dieser Arbeiten können nachfolgende Sachverhalte zusammengefasst werden.

Gewinnung quantitativer Multi-Omics-Datensätze

- Im Zuge von insgesamt 13 kontinuierlichen Kultivierungsexperimenten wurde der Wildtyp (WT) und zwei Lysinproduzenten (DM1730, DM1930) bei verschiedenen Wachstumsraten ($\mu = 0.05, 0.1, 0.2 \text{ h}^{-1}$) untersucht. Dabei wurden eine Vielzahl quantitativer Multi-Omics-Daten generiert. Diese umfassen:
 - genomweite Transkriptom-Daten (aus 4 Experimenten), d.h. bei einer Gesamtheit von 3571 getesteten Gensequenzen standen 14284 normierte mRNA-Intensitätsmessungen zur Verfügung,
 - Proteom-Daten von 16 Enzymen des ZSW und damit 208 Enzymaktivitätsmessungen,
 - Metabolom-Daten (aus 12 Experimenten) von 45 intra- und extrazellulären Intermediaten des ZSW und somit ca. 540 Metabolitkonzentrationen und schließlich

- Fluxom-Daten in Form von 3 relevanten Aufnahme- bzw. Produktbildungsraten, d.h. 39 extrazelluläre Flüsse.

Zwar ergaben sich aus der Kombination der verfügbaren Omics-Messdaten lediglich 3 komplette vertikale Datensätze, dennoch konnte damit insgesamt eine konsistente Datenbasis für Modellierungszwecke geschaffen werden.

Vertikaler Modellierungsansatz

- Für die Beschreibung der Multi-Omics-Datensätze wurde ein mechanistischer Modellierungsansatz entwickelt, welcher entlang der Regulationsebenen einer Zelle (Genom, Proteom, Metabolom) formuliert ist.
- Mit der Fokussierung auf den Zitratzyklus und der Anaplerose von *C. glutamicum* konnte unter Einbezug des aktuellen Wissensstandes dieses Stoffwechselweges ein detailliertes Regulationsnetzwerk auf metabolischer Modellebene aufgebaut werden.
- Hingegen konnten infolge fehlender Informationen über die Regulation der Genexpression, d.h. insbesondere das Wissen über Transkriptionsfaktoren sowie zugehöriger Metaboliteffektoren, keine mechanistischen Modelle der genetischen Regulation und der entsprechenden Rückkopplungen zwischen der Metabolom- und Genomebene formuliert werden. Stattdessen wurden phänomenologische Beziehungen für die Modellierung der Transkription und Translation verwendet, die aufgrund der Wachstumsratenabhängigkeit zumindest einen teilweise regulatorischen Charakter besitzen.

Angesichts der notwendigen vereinfachenden Annahmen stellt das gegenwärtige vertikale Modell nur bedingt ein realistisches Abbild der gesamtzellulären Regulationsprozesse dar. Gleichzeitig ist es jedoch in Bezug auf die Multi-Omics-Daten auf einem sinnvollen Detaillierungsniveau gehalten, so dass eine Auswertung dieser Daten im biologischen Kontext erfolgen kann.

Technische Implementierung und Validierung

- Für die Modellierung biochemischer Netzwerke mit Hilfe von Modelica wurde eine geeignete Komponentenbibliothek entwickelt [WNE09]. Damit konnte eine Implementierung des vertikalen Netzwerkmodells unter Nutzung der Objektorientierung dieser Modellierungssprache erfolgen.

- Mit dem Ziel der Modellvalidierung wurden die verschiedenen Multi-Omics-Datensätze mit Hilfe einer parallelen Optimierungsmethode von gekoppelten Einzelmodellen integriert. Die einzelnen Modelle sind dabei über die organismenspezifischen Parameter (z.B. Enzymeigenschaften) miteinander gekoppelt und unterscheiden sich in der Wertzuweisung der experimentsspezifischen Parameter (z.B. externe Systemvariablen).
- Durch die Verwendung phänomenologischer und approximativer mechanistischer Ansätze zur Kinetikbeschreibung enthält das Modell damit 29 experiment- und 103 organismenspezifische Parameter. Für die Validierung standen aus den drei vertikalen Datensätzen 145 Messdaten zur Verfügung, wodurch das Netzwerk insgesamt unterbestimmt war. Ein Teil der Modellparameter war somit auf der gegenwärtigen Messdatengrundlage nicht identifizierbar.
- In Bezug auf die Modellanpassungen wurden im überwiegenden Teil der Messungen die richtige Größenordnung durch das Modell getroffen. Allerdings wurden auch erhebliche Abweichungen zwischen simulierten und gemessenen Multi-Omics-Daten gefunden. Dies trifft vor allem für die Enzymaktivitätsmessungen und Teile der Metabolitkonzentrationsdaten zu.

Inwieweit hier die Ursache in der Modellstruktur oder in der Genauigkeit der Daten selbst liegt, kann gegenwärtig nicht mit Sicherheit beurteilt werden. Es liegt jedoch die Vermutung nahe, dass zumindest der durch ein Zell-Leakage verursachte Metabolitverlust beim Quenchen der Kultivierungsproben eine nicht unerhebliche Rolle spielt.

Analyse der Simulationsergebnisse

- Durch den Vergleich der resultierenden Simulationsergebnisse konnten stamm- und wachstumspezifische Effekte im untersuchten Organismus *C. glutamicum* festgestellt werden. Insbesondere unter Verwendung neuartiger Visualisierungstechniken [NWQW07, DNNW09] konnten die komplexen Zusammenhänge im biochemischen Netzwerk vereinfacht dargestellt und somit leichter Hypothesen über die *in vivo*-Regulation abgeleitet werden.
- Auf der Grundlage einer Sensitivitätsanalyse erfolgte die Prädiktion geeigneter Stellen für molekularbiologische Eingriffe zur Verbesserung der Lysin-Produktbildung, d.h. Verstärkung oder Verminderung der Expression bestimmter Zielgene.

Inwiefern diese Modellprädiktionen einer praktischen Überprüfung standhalten, werden hingegen erst zukünftige Untersuchungen zeigen.

10.3 Ausblick

Aus den hier dargestellten Ergebnissen und deren Diskussion ergeben sich eine Reihe von interessanten Frage- und Zielstellungen für weiterführende Arbeiten.

Im Hinblick auf die Gewinnung aussagekräftiger Fluxom-Daten müssen zunächst einige grundlegende Untersuchungen erfolgen. Diese betreffen:

- Die Durchführung von stationären und instationären Markierungsexperimenten auf der Basis technischer Replikate mit verschiedenen Organismen bzw. Stämmen.
- Im Fall des isotopisch stationären Ansatzes ist das kontinuierliche Prozessverfahren zu bevorzugen (Chemostat), da somit die metabolische Stationarität einfacher gewährleistet werden kann. Zu einem späteren Zeitpunkt ist es jedoch unbedingt notwendig, dass auch entsprechende Versuche im Fed-Batch-Ansatz erfolgen, da dies den industriell vorherrschenden Prozess darstellt.
- Für die isotopisch instationäre Stoffflussanalyse ist das Fed-Batch-Verfahren bereits etabliert, allerdings wäre hier auch eine Anwendung im kontinuierlichen Verfahren interessant. Auf diese Weise könnten in einem Experiment die An- und Abreicherung von ^{13}C -markierten Intermediaten durch Zu- und Abschalten des Markierungszuflusses verfolgt werden.
- Es hat sich erwiesen, dass zur Minimierung des hohen Zeitaufwands bei der Auswertung der Massenspektren die Entwicklung eines (halb)-automatisierten Auswertungsverfahrens dringend erforderlich ist.

Mit Hilfe reproduzierbarer, quantitativer Stoffflussdaten könnten die bereits bestehenden Multi-Omics-Datensätze um die Fluxom-Ebene erweitert werden.

Auch im Bereich der mechanistischen Modellierung von Multi-Omics-Daten ist eine Reihe weiterer Zielsetzungen naheliegend:

- Essentiell ist zunächst die experimentelle Überprüfung der Prädiktionen durch das vertikale Netzwerkmodell, z.B. durch Erzeugung und experimentelle Untersuchung der vorgeschlagenen genetisch veränderten Stämme.

- Gleichzeitig kann durch die Bereitstellung weiterer Multi-Omics-Datensätze die Parameterbestimmtheit und damit die Aussagekraft des Modells gesteigert werden.
- Mit der Anbindung des fokussierten mechanistischen Netzwerkmodells an das genomweite Stoffwechselmodell von *C. glutamicum* können ähnlich wie im hier gezeigten Fall der ^{13}C -MFA zusätzliche Nebenbedingungen in die Parameterschätzung einfließen und damit der Raum möglicher Lösung deutlich reduziert werden.
- Interessant ist auch die direkte Kopplung an ein ^{13}C -Stoffwechselmodell des Zentralstoffwechsels für die integrative Auswertung von ^{13}C -Markierungsdaten auf Modelica-Ebene. Die dafür notwendigen Vorarbeiten zur automatischen Generierung der komplexen Gleichungssysteme in Modelica wurden bereits durchgeführt [Til08, TNE⁺09].

Anhang A

Methoden zur Systemanalyse

Im Folgenden werden die grundlegenden Methoden der Systemanalyse in kompakter Form erläutert, welche im Rahmen der vorliegenden Arbeit zum Einsatz kamen. Mit dem Ziel einer möglichst allgemeingültigen Formulierung wird hierbei der im Abschnitt 3.3 eingeführte Modellierungsansatz für die Beschreibung metabolisch und isotopisch instationärer Zustände als Beispiel genutzt:

$$\frac{d\mathbf{c}_{\text{IM}}^{\#}}{dt} = \mathbf{f}(\mathbf{v}^{\#}, \mathbf{c}_{\text{IM}}^{\#}, \mathbf{c}_{\text{S}}^{\#}, \boldsymbol{\alpha}) - \mu \cdot \mathbf{c}_{\text{IM}}^{\#}, \quad \mathbf{c}_{\text{IM}}^{\#}(0) = \mathbf{c}_{\text{IM},0}^{\#} \quad (\text{A.1})$$

Die Beschreibung der Reaktionsraten $\mathbf{v}^{\#}$ erfolgt mit Hilfe kinetischer Modelle (vgl. Abschn. 3.3.5), die eine Vielzahl zu schätzender Parameter $\boldsymbol{\alpha}$ enthalten können.

A.1 Parameterschätzung

Bei einer Parameterschätzung wird untersucht, welchen Wert alle Modellparameter annehmen müssen, um einen vorgegebenen Verlauf experimenteller Daten bestmöglich wiederzugeben. Die Problemstellung wird dabei als Minimierung einer Zielfunktion $\kappa(\boldsymbol{\alpha}, \mathbf{c}_{\text{IM},0}^{\#})$ formuliert, welche von den kinetischen Modellparametern $\boldsymbol{\alpha}$ und Anfangskonzentrationen $\mathbf{c}_{\text{IM},0}^{\#}$ abhängt (Abb. A.1):

$$\begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\mathbf{c}}_{\text{IM},0}^{\#} \end{pmatrix} = \arg \min_{\boldsymbol{\alpha}, \mathbf{c}_{\text{IM},0}^{\#}} \kappa(\boldsymbol{\alpha}, \mathbf{c}_{\text{IM},0}^{\#}) \quad (\text{A.2})$$

Häufig wird die gewichtete Fehlerquadratsumme (FQS) als Kriterium genutzt, bei der die Abweichungen zwischen gemessenen $\mathbf{c}_{\text{IM}}^{\#}(\boldsymbol{\xi})$ und simulierten $\hat{\mathbf{c}}_{\text{IM}}^{\#}(\boldsymbol{\xi})$ Werten

zu den Zeitpunkten $\xi = (t_1, \dots, t_n)^T$ minimiert werden:

$$\kappa(\alpha, \mathbf{c}_{\text{IM},0}^\#) = \sum_{i=1}^n \|\mathbf{c}_{\text{IM}}^\#(t_i) - \hat{\mathbf{c}}_{\text{IM}}^\#(t_i)\|_{\sum \mathbf{c}_{\text{IM}}^\#(t_i)}^2 \quad (\text{A.3})$$

Die Gewichtung erfolgt auf Basis der Messwertvarianzen $\sum \mathbf{c}_{\text{IM}}^\#(t_i)$, wodurch Messwerte mit kleinem Fehler einen höheren Einfluss auf die FQS erhalten als ungenaue Messwerte.

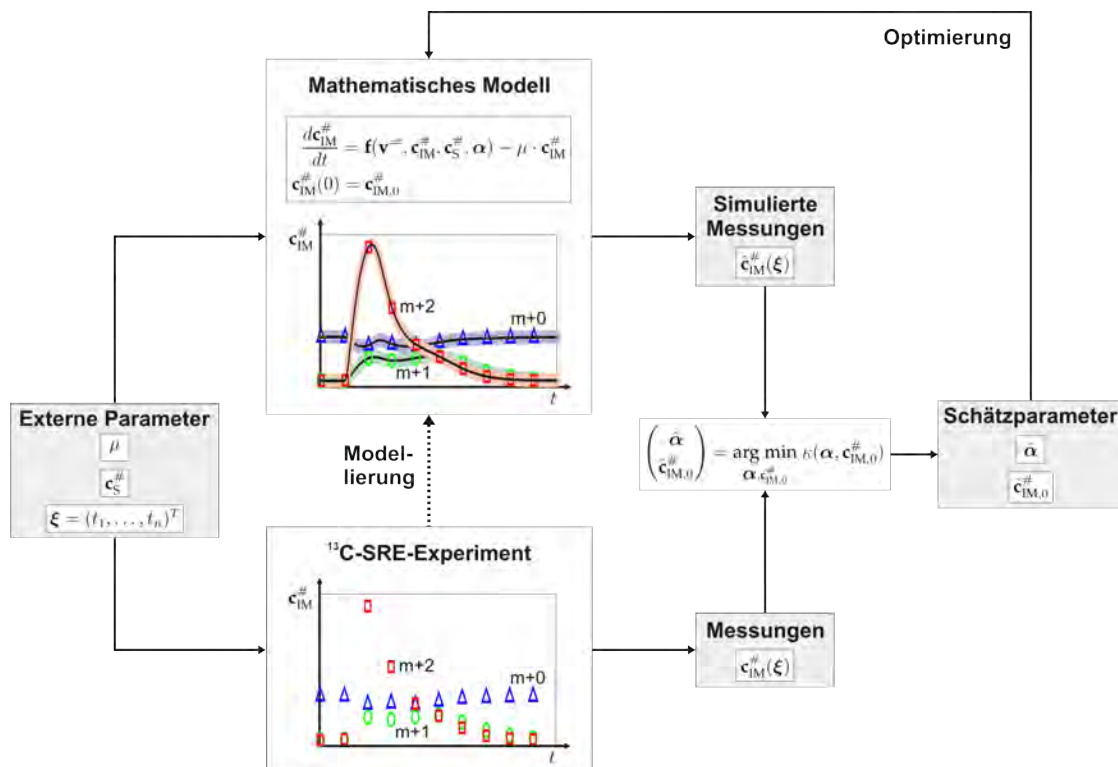


Abbildung A.1: Schematische Darstellung der Parameterschätzung am Beispiel der Modellierung von metabolisch und isotopisch instationären Experimenten.

Für Optimierungsprobleme dieser Art wurden in der Vergangenheit eine ganze Reihe unterschiedlicher Verfahren entwickelt. Grundsätzlich unterscheidet man hierbei ableitungsgestützte Verfahren, wie z.B. das Gradienten- oder Newton-Verfahren [Deu95] bzw. ableitungsfreie Methoden, wie z.B. die Simplex-Methode [MF02] oder genetische Algorithmen [Bey01].

A.2 Sensitivitätsanalyse

Mit Hilfe einer Sensitivitätsanalyse wird das Verhalten von Ausgangsgrößen eines Modells (z.B. simulierte Zustandsvariablen) in Bezug auf Veränderungen von verschiedenen Eingangsgrößen (z.B. Modellparameter) untersucht. Prinzipiell lassen sich die verschiedenen Methoden dabei in „lokale“ und „globale“ Ansätze einteilen, allerdings ist hierbei zu beachten, worauf der jeweilige Terminus bezogen ist.

Liegt der Sensitivitätsanalyse ein bestimmter Referenzpunkt (i.d.R. optimaler Parametersatz) zugrunde, dann wird mit Hilfe einer lokalen Sensitivitätsfunktion (z.B. mittels Finite-Differenzen-Approximation) der jeweils untersuchte Parameter variiert, während alle anderen konstant gehalten werden. In diesem Fall würden bei einem globalen Verfahren alle Parameter gleichzeitig geändert werden (z.B. Monte-Carlo-Simulation). Demgegenüber existieren mittlerweile globale Methoden, welche Sensitivitätsinformationen im gesamten Parameterraum, d.h. auch außerhalb des geschätzten, lokalen Referenzpunktes, bereitstellen [SC00].

Bei der Berechnung lokaler Sensitivitäten werden i.d.R. partielle Ableitungen 1. Ordnung ermittelt:

$$S_{i,j}(t) = \frac{\partial c_{\text{IM},i}^{\#}(t, \alpha)}{\partial \alpha_j} \quad \forall i = 1 \dots m, \forall j = 1 \dots p \quad (\text{A.4})$$

Die resultierende Sensitivitätsmatrix $\mathbf{S}(t)$ erlaubt Aussagen darüber zu treffen, ob und wie stark die Änderung eines Parameters α_j die Änderung einer Zustandsgröße $c_{\text{IM},i}^{\#}$ bedingt. Negative Werte dieser Matrix ($S_{i,j}(t) < 0$) bedeuten demnach, dass eine Erhöhung des Parameters j den Wert des Intermediatpools i verringert.

Die partiellen Ableitungen können auf verschiedenen Wegen berechnet werden.

A.2.1 Berechnung lokaler Sensitivitäten

Finite-Differenzen-Approximation

Eine Methode basiert auf der Approximation der Ableitungen durch zentrale Differenzenquotienten. Dabei wird jeweils der j -te Parameter um $\Delta\alpha$ ausgelenkt, während alle anderen Parameter konstant gehalten werden:

$$S_{i,j}(t) \approx \frac{c_{\text{IM},i}^{\#}(t, \alpha_j + \Delta\alpha_j, \alpha_{k=1\dots p \neq j}) - c_{\text{IM},i}^{\#}(t, \alpha_j - \Delta\alpha_j, \alpha_{k=1\dots p \neq j})}{2 \cdot \Delta\alpha_j} \quad (\text{A.5})$$

Die Methode ist leicht zu implementieren, aber die Genauigkeit der errechneten Werte hängt stark von der Parameteränderung $\Delta\alpha$ ab. Diese muss klein genug sein, um die Annahme der lokalen Linearität des nichtlinearen Modells zu erfüllen. Hingegen ist bei zu kleinen Parameteränderungen die Differenz zwischen den Modelllösungen vor und nach der Auslenkung zu gering und der damit verbundene Rundungsfehler zu hoch.

Automatische Differentiation

Ein wesentlich genauerer Ansatz für die Bestimmung lokaler Sensitivitäten basiert auf der Methode der Automatische Differentiation (AD), bei der auf effiziente Art und Weise exakte Ableitungen berechnet werden können [GW09]. Die grundlegende Idee der AD ist dabei, dass jede Funktionsauswertung in einer Abfolge von Zwischenschritten (elementaren Umwandlungen) abläuft. Durch eine entsprechende Zerlegung können somit Folgen von Funktionen und Zwischenwerten erhalten werden, die nur von wenigen Parametern abhängen (Abb. A.2).

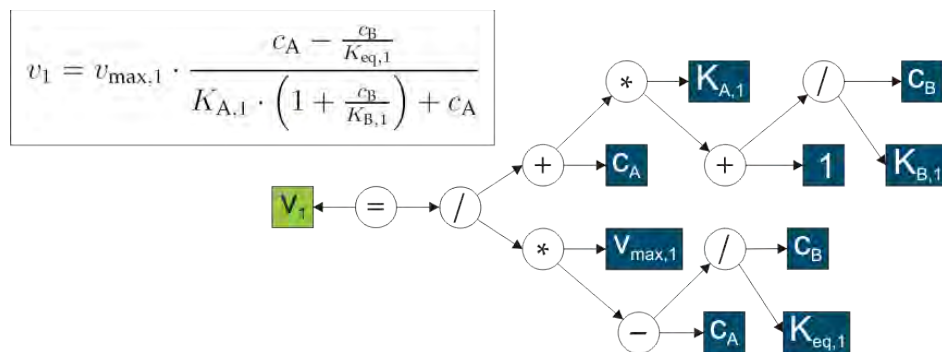


Abbildung A.2: Prinzipschema der Automatischen Differentiation. Die kinetische Beziehung wird dabei in elementare Umwandlungsschritte zerlegt.

Im Rahmen dieser Arbeit wurde das speziell für die Modellierungssprache Modelica entwickelte Werkzeug ADModelica [ENW08] eingesetzt. Dabei wird auf der Basis eines bestimmten Modells neuer Modelica-Code erzeugt, welcher im konkreten Fall die gesuchten Ableitungen enthält.

A.2.2 Kontrollkoeffizienten

Die unter Verwendung der oben vorgestellten Methoden berechneten Parametersensitivitäten sind skalensvariant, d.h. sie sind von der Größe der Intermediatkonzentration bzw. des Parameters abhängig. Um dieses Problem zu beheben und eine

bessere Vergleichsbasis zwischen den verschiedenen Sensitivitätsfunktionen zu erhalten, müssen diese in dimensionslose Größen skaliert werden:

$$\mathbf{C}^M(t) = (\text{diag } \mathbf{c}_{\text{IM}}^\#(t))^{-1} \cdot \mathbf{S}(t) \cdot \text{diag } \boldsymbol{\alpha} \quad (\text{A.6})$$

Diese relativen Sensitivitäten werden auch als Konzentrations-Kontrollkoeffizienten bezeichnet. Sie beschreiben die relative Änderung einer Poolgröße $c_{\text{IM}}^\#(t)$ als Folge einer relativen, infinitesimal kleinen Änderung eines Parameters α .

A.3 Fehlerübertragung

Mit Hilfe der Fehlerübertragung lässt sich bestimmen, wie sich im Zuge der Parameterschätzung die Fehler in den Messdaten auf die geschätzten Modellparameter übertragen. Anhand der resultierenden Konfidenzintervalle können dann Aussagen über die Identifizierbarkeit einzelner Parameter in Abhängigkeit der Messdaten getroffen werden.

A.3.1 Linearisierungsansatz

Im Fall nichtlinearer Modellstrukturen ist unter der Annahme, dass der geschätzte Parametervektor $\hat{\boldsymbol{\alpha}}$ mit Gl. A.2 nah genug an der wahren Lösung $\boldsymbol{\alpha}^*$ liegt, die Anwendung der linearisierten Regressionstheorie möglich [CH88]. Die gesuchte Parameter-Kovarianzmatrix kann dabei unter Nutzung der Sensitivitätsinformationen $\partial \hat{\mathbf{c}}_{\text{IM}}^\#(\boldsymbol{\xi}) / \partial \hat{\boldsymbol{\alpha}}$ und den ermittelten Kovarianzen der Messungen $\text{Cov}(\mathbf{c}_{\text{IM}}^\#(\boldsymbol{\xi}))$ folgendermaßen geschätzt werden:

$$\text{Cov}(\hat{\boldsymbol{\alpha}}) = \left(\left(\frac{\partial \hat{\mathbf{c}}_{\text{IM}}^\#(\boldsymbol{\xi})}{\partial \hat{\boldsymbol{\alpha}}} \right)^T \cdot \text{Cov}(\mathbf{c}_{\text{IM}}^\#(\boldsymbol{\xi}))^{-1} \cdot \frac{\partial \hat{\mathbf{c}}_{\text{IM}}^\#(\boldsymbol{\xi})}{\partial \hat{\boldsymbol{\alpha}}} \right)^{-1} \quad (\text{A.7})$$

Entsprechende Konfidenzintervalle für die geschätzten Parameter können dann aus den Diagonalelementen der Kovarianzmatrix ermittelt werden:

$$\hat{\sigma}(\hat{\alpha}_i) = \sqrt{\text{Cov}_{i,i}(\hat{\boldsymbol{\alpha}})} \quad (\text{A.8})$$

A.3.2 Monte-Carlo-Simulation

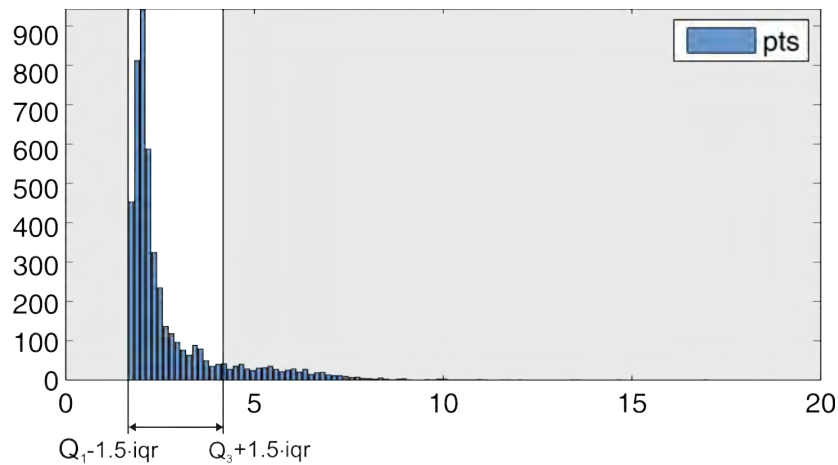
Die klassische Monte-Carlo-Simulation stellt eine weitere Methode dar, Konfidenzintervalle für Parameter zu schätzen. Ihr Vorteil gegenüber dem Linearisierungsansatz ist, dass die Fehlerübertragung direkt mit dem nichtlinearen Modell gerechnet werden kann.

Innerhalb dieser Arbeit wurde eine Bootstrap-Methode verwendet, wobei folgende Schritte für eine festgelegte Anzahl an Iterationen (hier $n = 5000$) durchlaufen werden:

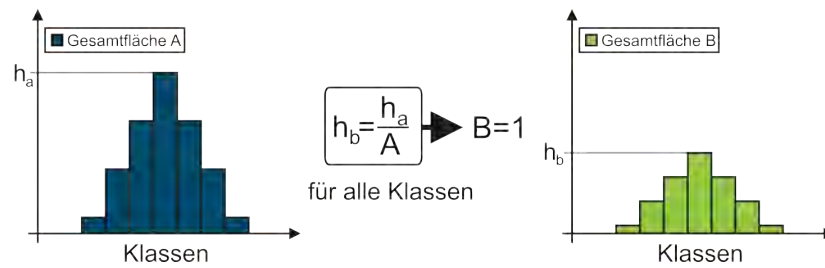
1. Die gemessenen Zustandsvariablen werden im Rahmen ihrer geschätzten Mess-Standardabweichungen unter der Annahme normalverteilter Messfehler zufallsbedingt verrauscht.
2. Die Startwerte der zu optimierenden Parameter werden vor jeder Iteration um einen bestimmten Betrag (relativ oder absolut) verrauscht. Damit soll gewährleistet werden, dass der Optimierungsalgorithmus in ein stabiles Optimum konvergiert.
3. Optimierung nach Gl. A.3 und Ermittlung des Kleinsten-Quadrate-Schätzers $\hat{\alpha}$.

Bei der anschließenden Auswertung der durch die MC-Simulation erhaltenen Histogramm-Daten werden folgende Schritte durchgeführt (vgl. Abb. A.3):

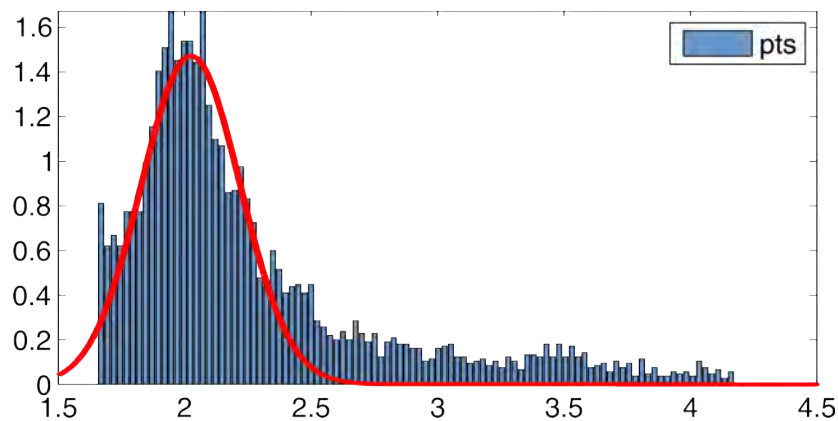
1. Eliminierung von Ausreißern über Boxplots.
2. Normierung der Histogramme.
3. Ermittlung von 90%-Konfidenzintervallen $\text{Conf}_{\hat{\theta}}(90)$.
4. Fit der Daten durch eine Normverteilungsfunktion $N(\mu, \sigma^2)$ zur Schätzung von Mittelwerten $\hat{\mu}$ und Standardabweichungen $\hat{\sigma}^2$.



(a) Ausreißertest. Der Interquartilsbereich (iqr) ergibt sich aus der Differenz des unteren und oberen Quartils ($Q_1 - Q_3$). Alle Werte innerhalb des grauen Bereiches werden aus dem Datenvektor entfernt.



(b) Normierung der Histogramme durch Division der Höhen, welche den Häufigkeiten einzelner Klassen entsprechen, durch die Gesamtfläche A.



(c) Anpassung der normierten Daten durch eine Normalverteilungsfunktion.

Abbildung A.3: Auswertung der Daten auf Basis von Monte-Carlo-Simulationen.

Anhang B

^{13}C -MFA von *C. glutamicum* DM1730

B.1 Prozessmodell

Tabelle B.1: Schätzwerte der Parameter des Prozessmodells auf der Basis von Monte-Carlo-Simulationen.

Parameter	Schätzwert	Std.abw.
$c_{S,0}$	11.42	2.08
$c_{X,0}$	0.73	0.21
$c_{P,0}$	0.19	0.07
$Y_{X,S}$	0.50	0.18
$Y_{P,S}$	0.83	0.63
μ_{\max}	0.20	0.02
π_{\max}	0.04	0.01
$K_{X,S}$	<0.01	-
$K_{P,S}$	<0.01	-
m_S	<0.01	-

B.2 Biomasseabflüsse

B.3 Modelle zur ^{13}C -MFA

B.4 Modellanpassungen der ^{13}C -MFA

Tabelle B.2: Bedarf an Precursoren für die Biomassesynthese nach [MdGW⁺96].

Precursor	AS	Bedarf $Y_{IM/X}^{-1}$ [$\mu\text{mol g}_{\text{BTM}}^{-1}$]	Biomasseabfluss [$\text{mmol g}_{\text{BTM}}^{-1} \text{h}^{-1}$]
G6P		205	0.041 ± 0.006
F6P		71	0.014 ± 0.002
GA3P		129	0.026 ± 0.004
R5P ^a		754	0.149 ± 0.021
	HIS	71	0.014 ± 0.002
E4P		-	
PEP ^b		52	0.010 ± 0.002
	PHE	133	0.026 ± 0.004
	TYR	81	0.016 ± 0.002
	TRP	54	0.011 ± 0.002
PGP		620	0.123 ± 0.018
	SER	225	0.045 ± 0.006
	GLY	361	0.072 ± 0.010
	CYS	87	0.017 ± 0.003
PYR ^c		83	0.016 ± 0.002
	ALA	606	0.120 ± 0.017
	VAL	284	0.056 ± 0.008
	LEU	440	0.087 ± 0.012
ACCOA ^d		2500	0.450 ± 0.071
OAA		340	0.067 ± 0.010
	ASP	199.5	0.040 ± 0.006
	ASN	199.5	0.040 ± 0.006
	THR	275	0.055 ± 0.008
	ILE	202	0.040 ± 0.006
	MET	146	0.029 ± 0.004
	LYS	202	0.040 ± 0.006
	DAP	146	0.029 ± 0.004
AKG		87	0.017 ± 0.003
	GLU	610	0.121 ± 0.017
	GLN	196	0.039 ± 0.006
	PRO	170	0.034 ± 0.005
	ARG	189	0.037 ± 0.005

^aR5P ist auch für die Synthese von TRP notwendig.

^bFür die Bildung aromatischer Aminosäuren werden 2 Moleküle PEP benötigt.

^cFür die Bildung verzweigt-kettiger Aminosäuren werden 2 Moleküle PYR benötigt. PYR ist auch für die Synthese von ILE, LYS und DAP notwendig.

^dACCOA ist auch für die Synthese von LEU notwendig.

Tabelle B.3: Reaktionen des fokussierten Modells des Zentralstoffwechsels von *C. glutamicum*.

Reaktion	Stöchiometrie und C-Atomtransitionen
pts	GLCex (#ABCDEF) + PEP (#abc) → G6P (#ABCDEF) + PYR (#abc)
pgi	G6P (#ABCDEF) ⇌ F6P (#ABCDEF)
pfk	F6P (#ABCDEF) → FBP (#ABCDEF)
fda	FDP (#ABCDEF) ⇌ DHAP (#CBA) + G3P (#DEF)
tpi	DHAP (#ABC) ⇌ G3P (#ABC)
gapA	G3P (#ABC) ⇌ PGP (#ABC)
eno	PGP (#ABC) ⇌ PEP (#ABC)
pyk	PEP (#ABC) → PYR (#ABC)
gnd	G6P (#ABCDEF) → P5P (#BCDEF) + CO ₂ (#A)
tk1_1	P5P (#ABCDE) + P5P (#abcde) ⇌ S7P (#ABabcde) + G3P (#CDE)
tal	P5P (#ABCDE) + E4P (#abcd) ⇌ G3P (#CDE) + F6P (#ABabcd)
tk1_2	G3P (#ABC) + S7P (#abcdefg) ⇌ E4P (#defg) + F6P (#abcABC)
pdh	PYR (#ABC) → ACCOA (#BC) + CO ₂ (#A)
gltA	OAA (#ABCD) + ACCOA (#ab) → CITICIT (#DCBbaA)
icd	CITICIT (#ABCDEF) → AKG (#ABCDE) + CO ₂ (#F)
odh	AKG (#ABCDE) ⇌ SUCCOA (#BCDE) + CO ₂ (#A)
suc	SUCCOA (#ABCD) ⇌ SUC (#ABCD)
sdh	SUC (#ABCD) ⇌ FUM (#ABCD)
fum	FUM (#ABCD) ⇌ MAL (0.5 #ABCD + 0.5 #DCBA)
mco1/mdh	MAL (#ABCD) ⇌ OAA (#ABCD)
aceA	CITICIT (#ABCDEF) ⇌ GLX (#AB) + SUC (#FCDE)
aceB	GLX (#AB) + ACCOA (#ab) → MAL (#ABba)
pyc/odx	PYR (#ABC) + CO ₂ (#a) ⇌ OAA (#ABCa)
ppc	PEP (#ABC) + CO ₂ (#a) → OAA (#ABCa)
mez	MAL (#ABCD) ⇌ PYR (#ABC) + CO ₂ (#D)
gdh	AKG (#ABCDE) ⇌ GLU (#ABCDEF)
aspC	OAA (#ABCD) + GLU (#abcde) ⇌ ASP (#ABCD) + AKG (#abcde)
bsLYS	ASP (#ABCD) + PYR (#abc) → LYS (0.5 #ABCDcb + 0.5 #abcDCB) + CO ₂ (0.5 #a + 0.5 #A)
excCO2	CO ₂ (#A) → CO ₂ ex (#A)
excLYS	LYS (#ABCDEF) → LYSex (#ABCDEF)

Tabelle B.4: Zusätzliche Reaktionen des erweiterten Modells mit Aminosäurebiosynthese wegen von *C. glutamicum*.

Reaktion	Stöchiometrie und C-Atomtransitionen
glnA	GLU (#ABCDE) → GLN (#ABCDEF)
proC	GLU (#ABCDE) → PRO (#ABCDEF)
argDJ	GLU (#ABCDE) + GLU (#abcde) → ORN (#ABCDE) + AKG (#abcde)
carA	GLN (#ABCDE) + CO ₂ (#a) → GLU (#ABCDE) + CPB (#a)
argF	ORN (#ABCDE) + CBP (#a) → CIL (#ABCDEa)
argGH	CIL (#ABCDEF) + ASP (#abcd) → ARG (#ABCDEF) + FUM (0.5 #abcd + 0.5 #dcba)
ltsA	ASP (#ABCD) + GLN (#abcde) → ASN (#ABCDE) + GLU (#abcde)
lysC	ASP (#ABCD) → ASPSA (#ABCD)
dapA	ASPSA (#ABCD) + PYR (#abc) → THDP (#ABCDcba)
dapD	THDP (#ABCDEF) + SUCCOA (#abcd) → SAP (#ABCDEFGabcd)
dapC	SAP (#ABCDEF) + GLU (#abcde) → SDAP (#ABCDEF) + AKG (#abcde)
dapE	SDAP (#ABCDEF) → DAP (0.5 #ABCDEF + 0.5 #GFEDCBA) + SUC (0.5 #HIJK + 0.5 #KJIH)
ddh	THDP (#ABCDEF) → DAP (#ABCDEF)
lysA	DAP (#ABCDEF) → LYS (#ABCDEF) + CO ₂ (#G)
hom	ASPSA (#ABCD) → HOM (#ABCD)
metX	HOM (#ABCD) + ACCOA (#ab) → SUCHMS (#ABCDab)
metB	ASP (#ABCDEF) + CYS (#abc) → CYST (#ABCDEF) + AC (#EF)
metC	CYST (#ABCDEF) → HCYS (#ABCD) + PYR (#EFG)
metH	HCYS (#ABCD) + THF (#a) → MET (#ABCDa)
thrB	HOM (#ABCD) → THR (#ABCD)
ilvBN_ile	THR (#ABCD) + PYR (#abc) → 3MOP (#ABbcCD) + CO ₂ (#a)
ilvE_ile	3MOP (#ABCDEF) + GLU (#abcde) → ILEU (#ABCDEF) + AKG (#abcde)
serC	PGP (#ABC) + GLU (#abcde) → SER (#ABC) + AKG (#abcde)
glyA	SER (#ABC) → GLY (#AB) + THF (#C)
cysEKM	SER (#ABC) + ACCOA (#ab) → CYS (#ABC) + AC (#ab)
NCgl0388	PYR (#ABC) + VAL (#abcde) → ALA (#ABC) + 3MOB (#abcde)
ilvBN_val	PYR (#ABC) + PYR (#abc) → 3MOB (#ABbcC) + CO ₂ (#a)
ilvE_val	3MOB (#ABC) + GLU (#abcde) → VAL (#ABCDE) + AKG (#abcde)
leuA	3MOB (#ABCDE) + ACCOA (#ab) → 3C4MOP (#abBCDEA)
leuP	3C4MOP (#ABCDEF) → 4MOP (#ABCDEF) + CO ₂ (#G)
ilvE	4MOP (#ABCDEF) + GLU (#abcde) → LEU (#ABCDEF) + AKG (#abcde)
aroG	PEP (#ABC) + E4P (#abcd) → SKM5P (#ABCabcd)
aroA	PEP (#ABC) + SKM5P (#abcdefg) → CHOR (#ABCabcdefg)
pheA	CHOR (#ABCDEF) → PHPYR (#ABCEFGHI) + CO ₂ (#D)
NCgl0215	PHPYR (#ABCDEF) + GLU (#abcde) → PHE (#ABCDEF) + AKG (#abcde)
tyrA	CHOR (#ABCDEF) → 34HPP (#ABCEFGHI) + CO ₂ (#D)
r2.6.1.57	34HPP (#ABCDEF) + GLU (#abcde) → PHE (#ABCDEF) + AKG (#abcde)
trpE	CHOR (#ABCDEF) → NANTH (#DEFGHI) + PYR (#ABC)
trpG	NANTH (#ABCDEF) + GLN (#abcde) → ANTH (#ABCDEF) + GLU (#abcde)
trpDA	ANTH (#ABCDEF) + P5P (#abcde) → 3IG3P (#BCDEF) + CO ₂ (#A)
trpB	3IG3P (#ABCDEF) + SER (#abc) → TRP (#ABCDEF) + G3P (#abc)
hisG	P5P (#ABCDE) + THF (#a) → PRB (#aEDCBA)
hisFHC	PRB (#ABCDEF) + GLN (#abcde) → HIS (#ABCDEF) + AKG (#abcde)
pqo	PYR (#ABC) → AC (#BC) + CO ₂ (#A)
pta	AC (#AB) → ACCOA (#AB)
inTHF	THFin (#A) → THF (#A)
outTHF	THF (#A) → THFout (#A)

Tabelle B.5: Gemessene und simulierte Massenisotopomeranteile intrazellulärer Metabolite auf Basis der stationären ^{13}C -MFA. Vergleich zwischen verschiedenen Netzwerkvarianten und Messmodellen: ZSW) Fokussiertes Modell des Zentralstoffwechsels; ZSW_AS) Erweitertes Modell mit Aminosäurebiosynthesewegen; MM1) Messungen von ZSW-Intermediaten; MM2) Zusätzliche Messungen in freien Aminosäuren.

	G6P	ZSW MM1	ZSW_AS MM1	ZSW_AS MM2	F6P	ZSW MM1	ZSW_AS MM1	ZSW_AS MM2
m+0	0.088 ± 0.003	0.106	0.099	0.104	0.095 ± 0.006	0.104	0.098	0.102
m+1	0.068 ± 0.004	0.062	0.071	0.074	0.074 ± 0.008	0.061	0.070	0.074
m+2	0.039 ± 0.004	0.055	0.042	0.041	0.026 ± 0.009	0.056	0.043	0.042
m+3	0.052 ± 0.001	0.048	0.049	0.048	0.052 ± 0.006	0.049	0.050	0.049
m+4	0.060 ± 0.002	0.055	0.063	0.061	0.067 ± 0.007	0.056	0.064	0.062
m+5	0.079 ± 0.002	0.074	0.076	0.078	0.087 ± 0.008	0.075	0.078	0.079
m+6	0.615 ± 0.005	0.601	0.600	0.594	0.598 ± 0.011	0.598	0.597	0.591
$\bar{\kappa}^a$		10.9	4.5	7.9		3.0	0.8	1.0
	S7P				FBP			
m+0	0.020 ± 0.003	0.012	0.020	0.020	0.079 ± 0.019	0.075	0.075	0.078
m+1	0.019 ± 0.003	0.020	0.027	0.027	0.053 ± 0.015	0.045	0.052	0.054
m+2	0.080 ± 0.004	0.058	0.075	0.073	0.026 ± 0.011	0.044	0.036	0.035
m+3	0.077 ± 0.006	0.093	0.083	0.086	0.133 ± 0.028	0.139	0.134	0.136
m+4	0.112 ± 0.008	0.139	0.097	0.104	0.067 ± 0.017	0.070	0.071	0.071
m+5	0.195 ± 0.009	0.147	0.143	0.147	0.085 ± 0.020	0.082	0.076	0.078
m+6	0.105 ± 0.011	0.147	0.147	0.149	0.557 ± 0.112	0.545	0.557	0.548
m+7	0.391 ± 0.015	0.383	0.409	0.393				
$\bar{\kappa}$		12.3	7.9	7.1		0.5	0.2	0.1
	P5P				E4P			
m+0	0.062 ± 0.044	0.054	0.118	0.111	0.187 ± 0.013	0.179	0.173	0.175
m+1	0.027 ± 0.020	0.048	0.048	0.049	0.045 ± 0.004	0.080	0.065	0.068
m+2	0.155 ± 0.017	0.155	0.072	0.082	0.015 ± 0.004	0.002	0.021	0.023
m+3	0.165 ± 0.031	0.108	0.070	0.082	0.074 ± 0.004	0.070	0.081	0.085
m+4	0.114 ± 0.012	0.106	0.107	0.113	0.678 ± 0.008	0.669	0.660	0.649
m+5	0.478 ± 0.025	0.529	0.585	0.563				
$\bar{\kappa}$		1.6	9.2	6.7		17.8	7.1	11.1
	GA3P				DHAP			
m+0	0.229 ± 0.034	0.186	0.192	0.196	0.209 ± 0.003	0.185	0.191	0.194
m+1	0.025 ± 0.012	0.060	0.056	0.059	0.057 ± 0.004	0.061	0.057	0.060
m+2	0.023 ± 0.013	0.066	0.051	0.053	0.052 ± 0.005	0.067	0.052	0.055
m+3	0.723 ± 0.034	0.688	0.701	0.692	0.682 ± 0.010	0.687	0.700	0.691
$\bar{\kappa}$		5.5	3.3	3.8		15.7	7.9	5.4
	PGP				PEP			
m+0	0.216 ± 0.044	0.186	0.192	0.196	0.252 ± 0.051	0.186	0.192	0.196
m+1	0.050 ± 0.014	0.060	0.056	0.059	0.055 ± 0.015	0.060	0.056	0.059
m+2	0.045 ± 0.013	0.066	0.051	0.053	0.031 ± 0.012	0.066	0.051	0.053
m+3	0.689 ± 0.138	0.688	0.701	0.692	0.661 ± 0.133	0.688	0.701	0.692
$\bar{\kappa}$		0.9	0.2	0.3		2.6	1.1	1.2
	SUC				PYR			
m+0	0.089 ± 0.021	0.032	0.031	0.035	0.019 ± 0.043	0.186	0.189	0.191
m+1	0.094 ± 0.006	0.092	0.091	0.099	0.054 ± 0.009	0.060	0.059	0.062
m+2	0.215 ± 0.033	0.175	0.178	0.171	0.066 ± 0.007	0.066	0.062	0.060
m+3	0.230 ± 0.010	0.229	0.229	0.216	0.860 ± 0.032	0.688	0.690	0.687
m+4	0.373 ± 0.042	0.472	0.471	0.480				
$\bar{\kappa}$		3.0	2.9	3.6		11.1	11.1	11.8

^aMittelwert des Residuumanteils aller Massenspuren der betreffenden Messgruppe.

Tabelle B.6: Gemessene und simulierte Massenisotopomeranteile freier Aminosäuren auf Basis der stationären ¹³C-MFA.

	TYR	ZSW_AS MM2	VAL	ZSW_AS MM2	GLN	ZSW_AS MM2
m+0	0.114 ± 0.015	0.007	0.012 ± 0.004	0.041	0.137 ± 0.014	0.015
m+1	0.026 ± 0.002	0.007	0.040 ± 0.006	0.026	0.046 ± 0.001	0.037
m+2	0.022 ± 0.003	0.030	0.151 ± 0.003	0.154	0.152 ± 0.007	0.112
m+3	0.047 ± 0.001	0.049	0.196 ± 0.005	0.196	0.204 ± 0.005	0.235
m+4	0.052 ± 0.002	0.061	0.090 ± 0.001	0.090	0.201 ± 0.003	0.210
m+5	0.105 ± 0.004	0.127	0.510 ± 0.004	0.492	0.260 ± 0.018	0.391
m+6	0.143 ± 0.004	0.154				
m+7	0.130 ± 0.008	0.146				
m+8	0.108 ± 0.003	0.095				
m+9	0.252 ± 0.010	0.324				
$\bar{\kappa}$		28.5		13.4		50.5
	MET		ASP		SER	
m+0	0.198 ± 0.051	0.011	0.090 ± 0.010	0.036	0.290 ± 0.029	0.166
m+1	0.008 ± 0.010	0.057	0.113 ± 0.008	0.106	0.101 ± 0.011	0.087
m+2	0.044 ± 0.052	0.122	0.153 ± 0.015	0.159	0.085 ± 0.018	0.102
m+3	0.199 ± 0.022	0.176	0.238 ± 0.008	0.215	0.524 ± 0.039	0.645
m+4	0.273 ± 0.006	0.297	0.406 ± 0.014	0.484		
m+5	0.278 ± 0.012	0.337				
$\bar{\kappa}$		13.5		14.4		7.7

Anhang C

Vertikales Netzwerkmodell von *C. glutamicum*

C.1 Modellkonstanten

Tabelle C.1: Angenommene Werte der Elongationsraten von mRNA-Synthese und Peptidbindung sowie der Degradationsraten von mRNA und Proteinen.

Konstante	Wert	Einheit
$v_{eg,mRNA}$	2400	NT RNAP ⁻¹ min ⁻¹
$v_{eg,P}$	3600	NT RB ⁻¹ min ⁻¹
$k_{dg,mRNA}$	27.6	h ⁻¹
$k_{dg,P}$	0.6	h ⁻¹

Tabelle C.2: Bedarf an Precursoren für die Biomassesynthese nach [MSdG⁺97].

Precursor	Bedarf $Y_{IM/X}$ [mmol g _{BIM} ⁻¹]
G6P	0.21
F6P	0.07
R5P	0.88
E4P	0.27
GA3P	0.13
3PG	1.29
CO ₂	3.42
NADPH	14.9

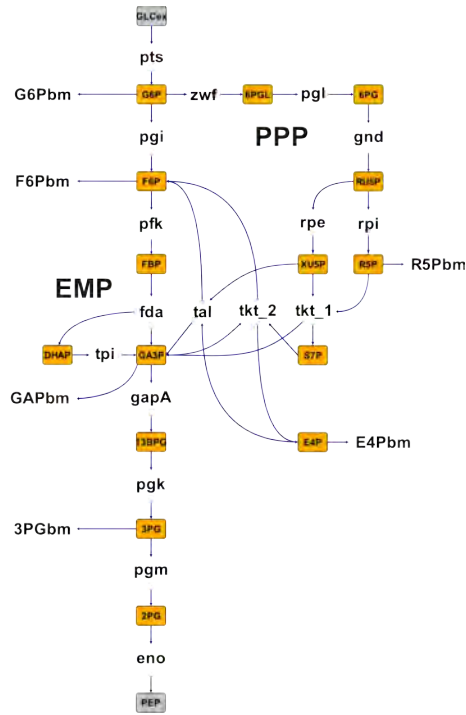
Precursor	Bedarf $Y_{IM/X}$ [mmol g _{BIM} ⁻¹]
PEP	0.53
PYR	2.69
ACCOA	2.94
AKG	1.25
ASP	1.51
LYS	0.20

C.2 Modellreaktionen

Tabelle C.3: Reaktionen des vertikalen Netzwerkmodells von *C. glutamicum*.

Enzym	Katalysierte Reaktion	Aktivator	Inhibitor	Referenz
pyk	$\text{PEP} + \text{ADP} \rightarrow \text{PYR} + \text{ATP}$	AMP	ATP	[JGLS94]
pdh	$\text{PYR} + \text{NAD} + \text{COA} \rightarrow \text{ACCOA} + \text{NADH} + \text{CO}_2$	-	-	
gltA	$\text{ACCOA} + \text{OAA} + \text{H}_2\text{O} \rightarrow \text{CIT} + \text{COA}$	-	ATP, (ACN), (ICIT)	[ETSE ⁺ 94]
acn	$\text{CIT} \rightleftharpoons \text{ICIT}$	-	-	
icd	$\text{ICIT} + \text{NADP} \rightarrow \text{AKG} + \text{NADPH} + \text{CO}_2$	-	OAA, CIT, AKG, (GLX)	[ERS95]
odh	$\text{AKG} + \text{NAD} + \text{COA} \rightarrow \text{SUCCOA} + \text{NADH} + \text{CO}_2$	ACCOA	SUCCOA, NADH, PYR, OAA, (ACN), (NADPH)	[SU80]
suc	$\text{SUCCOA} + \text{ADP} + \text{P}_i \rightleftharpoons \text{SUC} + \text{ATP} + \text{COA}$	-	-	
sdh	$\text{SUC} + \text{MQ} \rightleftharpoons \text{FUM} + \text{MQH}_2$	-	OAA	[MvdRDY00]
fum	$\text{FUM} + \text{H}_2\text{O} \rightleftharpoons \text{MAL}$	-	ATP	[GWO06]
mqo	$\text{MAL} + \text{MQ} \rightleftharpoons \text{OAA} + \text{MQH}_2$	-	-	
mdh	$\text{OAA} + \text{NADH} \rightleftharpoons \text{MAL} + \text{NAD}$	-	-	
ppc	$\text{PEP} + \text{CO}_2 + \text{H}_2\text{O} \rightarrow \text{OAA} + \text{P}_i$	ACCOA, FBP	ASP, AKG, GLU	[OS69, MS85, DDLEG04]
pyc	$\text{PYR} + \text{ATP} + \text{CO}_2 \rightarrow \text{OAA} + \text{ADP} + \text{P}_i$	-	ADP, AMP, ACCOA, AKG, (OAA), (ASP)	[PWWP ⁺ 97, KJAS02]
pck	$\text{PEP} + \text{GDP} + \text{CO}_2 \rightleftharpoons \text{OAA} + \text{GTP}$	-	ATP	[JS93]
mez	$\text{MAL} + \text{NADP} \rightleftharpoons \text{PYR} + \text{NADPH} + \text{CO}_2$	-	(OAA), (GLU)	[GBLG00]
odx	$\text{OAA} \rightarrow \text{PYR} + \text{CO}_2$	-	ADP, SUC	[JS95]
gdh	$\text{AKG} + \text{NADPH} + \text{NH}_3 \rightleftharpoons \text{GLU} + \text{NADP} + \text{H}_2\text{O}$	-	-	
aspB	$\text{OAA} + \text{Glu} \rightleftharpoons \text{ASP} + \text{AKG}$	-	-	
bsLYS	$\text{ASP} + \text{PYR} \rightarrow \text{LYS}$	-	-	

C.3 Globale Bilanzen



$$\begin{aligned}
 \text{G6P} : 0 &= v_{\text{pts}} - v_{\text{pgi}} - v_{\text{zwf}} - v_{\text{G6Pbm}} \\
 \text{F6P} : 0 &= v_{\text{pgi}} - v_{\text{pfk}} + v_{\text{tal}} + v_{\text{tkt}_2} - v_{\text{F6Pbm}} \\
 \text{FBP} : 0 &= v_{\text{pfk}} - v_{\text{fda}} \\
 \text{DHAP} : 0 &= v_{\text{fda}} - v_{\text{tpi}} \\
 \text{GA3P} : 0 &= v_{\text{fda}} + v_{\text{tpi}} - v_{\text{gapA}} + v_{\text{tal}} - v_{\text{tkt}_2} + v_{\text{tkt}_1} - v_{\text{GAPbm}} \\
 \text{13BPG} : 0 &= v_{\text{gapA}} - v_{\text{pgk}} \\
 \text{3PG} : 0 &= v_{\text{pgk}} - v_{\text{pgm}} - v_{\text{3PGbm}} \\
 \text{2PG} : 0 &= v_{\text{pgm}} - v_{\text{eno}} \\
 \text{6PGL} : 0 &= v_{\text{zwf}} - v_{\text{pgl}} \\
 \text{6PG} : 0 &= v_{\text{pgl}} - v_{\text{gnd}} \\
 \text{RU5P} : 0 &= v_{\text{gnd}} - v_{\text{rpe}} - v_{\text{rpi}} \\
 \text{XU5P} : 0 &= v_{\text{rpe}} - v_{\text{tkt}_1} - v_{\text{tal}} \\
 \text{R5P} : 0 &= v_{\text{rpi}} - v_{\text{tkt}_1} - v_{\text{R5Pbm}} \\
 \text{S7P} : 0 &= v_{\text{tkt}_1} - v_{\text{tkt}_2} \\
 \text{E4P} : 0 &= v_{\text{tkt}_2} - v_{\text{tal}} - v_{\text{E4Pbm}}
 \end{aligned}$$

Abbildung C.1: Netzwerkausschnitt mit den Reaktionen der Glykolyse (EMP) und des Pentose-Phosphatweges (PPP) von *C. glutamicum* (links). Massenbilanzierung um die Reaktionsknoten des EMP und PPP unter der Annahme eines stationären Zustandes ($dc/dt = 0$) führt auf ein algebraisches Gleichungssystem (rechts).

Durch Lösung und Vereinfachung obigen Gleichungssystems lässt sich folgende Bilanz um den G6P-Knoten ableiten:

$$0 = v_{\text{pts}} - \frac{1}{2}(v_{\text{eno}} + v_{\text{GA3Pbm}} + v_{\text{3PGbm}}) - \frac{1}{6}v_{\text{gnd}} - v_{\text{G6Pbm}} - v_{\text{F6Pbm}} - \frac{5}{6}v_{\text{R5Pbm}} - \frac{2}{3}v_{\text{E4Pbm}}$$

Mit der Beziehung $v_{\text{bm}} = Y_{\text{IM}/X} \cdot \mu$ sowie den allgemeinen Bezeichnungen $v_{\text{eno}} = v_{\text{c3}}$ und $v_{\text{gnd}} = v_{\text{c5}}$ ergeben sich dann die Gleichungen 8.8 und 8.9:

$$\begin{aligned}
 0 &= 6 \cdot v_{\text{pts}} - 3 \cdot v_{\text{c3}} - v_{\text{c5}} - 6 \cdot (Y_{\text{EMP}/X} + Y_{\text{PPP}/X}) \cdot \mu \\
 Y_{\text{EMP}/X} &= Y_{\text{G6P}/X} + Y_{\text{F6P}/X} + \frac{1}{2} \cdot (Y_{\text{GA3P}/X} + Y_{\text{3PG}/X}) \\
 Y_{\text{PPP}/X} &= \frac{5}{6} \cdot Y_{\text{R5P}/X} + \frac{2}{3} \cdot Y_{\text{E4P}/X}
 \end{aligned}$$

C.4 Parameterschätzungen

Tabelle C.4: Geschätzte Werte aller experimentspezifischen Parameter des vertikalen Netzwerkmodells. Alle Angaben in mM.

Reaktion	Parameter	WT $\mu = 0.1 h^{-1}$	DM1730 $\mu = 0.1 h^{-1}$	DM1730 $\mu = 0.2 h^{-1}$
pyk	GP	4.465	3.310	0.260
pdh	GP	0.972	3.145	0.186
gltA	GP	4.445	2.932	0.426
acn	GP	0.739	3.088	0.020
icd	GP	2.143	1.545	0.177
odh	GP	2.135	1.205	0.152
suc	GP	2.740	2.280	0.163
sdh	GP	1.786	1.936	0.232
fum	GP	3.688	1.439	0.121
mdh	GP	0.000	0.001	0.000
mgo	GP	1.626	1.384	0.103
ppc	GP	1.132	1.277	0.049
pyc	GP	1.087	0.548	0.086
pck	GP	5.532	0.224	0.012
odx	GP	0.287	0.061	0.017
mez	GP	2.180	2.979	0.409
gdh	GP	0.147	0.553	0.031
aspB	GP	0.862	0.637	0.048
lysC	GP	1.485	1.017	0.124
	ACCOA	0.308	0.178	0.314
	AKG	0.146	0.261	0.278
	FBP	6.594	0.836	2.488
	AMP	0.913	0.266	0.234
	ADP	3.818	1.015	1.447
	ATP	15.578	12.411	6.877
	NAD	2.264	1.603	1.569
	NADP	1.738	1.189	0.865
	NADH	0.036	0.005	0.631
	NADPH	0.029	0.012	0.054

Tabelle C.5: Geschätzte Werte aller organismenspezifischen Parameter des vertikalen Netzwerkmodells.

Reaktion	Parameter	<i>In vitro</i> Messwert	untere Grenze	obere Grenze	<i>In vivo</i> Schätzwert	Referenz
pyk	k_{for}	-	1e-6	1e3	0.133	
	K_{PEP}	< 2.8	1.2	2.8	2.791	[JGLS94]
	K_{ADP}	0.07	0.04	0.11	0.092	[JGLS94]
	$K_{A,AMP}$	< 2.0	0.5	2.0	1.975	[SE05a]
	$K_{I,ATP}$	0.4	0.1	1.0	0.431	[SE05a]
pdh	k_{for}	-	1e-6	1e3	0.070	
	K_{Pyr}	1.7	0.85	2.55	1.970	[SE05b]
	K_{NAD}	0.08	0.04	0.12	0.109	[STS84]
glcA	k_{for}	-	1e-6	1e3	0.048	
	K_{AcCoA}	0.065	0.033	0.098	0.061	[RE07]
	K_{OAA}	0.002	0.001	0.003	0.003	[ETSE+94]
	$K_{I,ATP}$	10.0	5.00	15.00	14.999	[ETSE+94]
acn	k_{for}	-	1e-6	1e3	24.492	
	K_{eq}	0.072	0.036	0.108	0.296	Therm. Daten
	K_{Cit}	-	1e-6	5e2	499.996	
	$K_{I,Cit}$	-	1e-6	5e2	8.253	
icd	k_{for}	-	1e-6	1e3	0.300	
	$K_{I,Cit}$	0.012	0.006	0.018	0.012	[ERS95]
	K_{NADP}	0.024	0.012	0.036	0.012	[ERS95]
	$K_{I,Cit}$	15.0	7.500	22.500	16.784	Eikmanns unveröff.
	$K_{I,OAA}$	10.0	5.000	15.000	11.552	Eikmanns unveröff.
odh	k_{for}	-	1e-6	1e3	0.018	
	K_{AKG}	0.080	0.040	0.120	0.055	[SU80]
	K_{NAD}	0.086	0.043	0.129	0.122	[SU80]
	$K_{A,AcCoA}$	-	1e-6	5e2	499.996	
	$K_{I,NADH}$	-	1e-6	5e2	0.523	
	$K_{I,SucCoA}$	-	1e-6	5e2	14.384	
	$K_{I,Pyr}$	-	1e-6	5e2	14.449	
suc	k_{for}	-	1e-6	1e3	2.840	
	K_{eq}	0.453	0.113	0.793	0.720	Therm. Daten
	K_{SucCoA}	-	1e-6	5e2	34.012	
	K_{ADP}	-	1e-6	5e2	0.630	
	K_{Suc}	-	1e-6	5e2	0.013	
	K_{ATP}	-	1e-6	5e2	7.052	
sdh	k_{for}	-	1e-6	5e2	0.006	
	K_{eq}	1.687	0.844	2.531	1.660	Therm. Daten
	K_{Suc}	-	1e-6	5e2	1.025	
	K_{Fum}	-	1e-6	5e2	0.298	
	$K_{I,OAA}$	-	1e-6	5e2	1.099	

Fortsetzung nächste Seite

ANHANG C. VERTIKALES NETZWERKMODELL VON C. GLUTAMICUM

Reaktion	Parameter	<i>In vitro</i> Messwert	untere Grenze	obere Grenze	<i>In vivo</i> Schätzwert	Referenz
fum	k_{for}	-	1e-6	1e3	0.468	Therm. Daten [GWO06] [GWO06] [GWO06]
	K_{eq}	4.190	2.095	6.285	4.133	
	K_{Fum}	3.8e-6	1.9e-6	5.7e-6	1.9e-6	
	K_{Mal}	6.3e-6	3.2e-6	9.5e-6	9.5e-6	
	$K_{I,ATP}$	< 1.6	0.3	1.6	1.326	
mdh	k_{for}	-	1e-6	1e3	21.433	Therm. Daten [JGLS94]
	K_{eq}	1e5	5.0e4	1.5e5	1.5e5	
	K_{OAA}	0.25	0.1	0.25	0.246	
	K_{NADH}	-	1e-6	5e2	0.626	
	K_{Mal}	-	1e-6	5e2	29.888	
	K_{NAD}	-	1e-6	5e2	0.008	
mqo	k_{for}	-	1e-6	1e3	0.004	Therm. Daten
	K_{eq}	1805.182	902.59	2707.77	1667.905	
	K_{Mal}	-	1e-6	5e2	1.35e-5	
	K_{OAA}	-	1e-6	5e2	30.180	
ppc	k_{for}	-	1e-6	1e3	0.388	[PW96] [SE05a] [SE05a] [MS85, JGLS94] [MS85] [SE05a]
	K_{PEP}	< 3.0	0.8	3.0	2.787	
	$K_{A,FBP}$	0.2	0.05	1.0	0.133	
	$K_{A,AcCoA}$	0.1	0.04	0.2	0.157	
	$K_{I,Asp}$	0.044	0.04	1.5	0.160	
	$K_{I,AKG}$	4.7	2.35	7.05	4.182	
	$K_{I,Glu}$	40.0	20.0	60.0	48.207	
pyc	k_{for}	-	1e-6	1e3	0.612	[PWWP+97] Patent ADM; WO 02/31158 Patent ADM; WO 02/31158 [PWWP+97] [PWWP+97] [PWWP+97] [PWWP+97]
	K_{Pyr}	1.3	0.3	1.5	1.428	
	K_{ATP}	0.35	0.18	0.53	0.233	
	$K_{I,Asp}$	-	5.0	10.0	9.370	
	$K_{I,AKG}$	10.0	5.0	15.0	7.244	
	$K_{I,AcCoA}$	0.11	0.06	0.17	0.109	
	$K_{I,AMP}$	0.75	0.38	1.13	0.434	
	$K_{I,ADP}$	2.6	1.30	3.90	2.563	
pck	k_{for}	-	1e-6	1e3	3.16e-5	Therm. Daten [JS93] [JS93] [JS93] [JS93] [JS93]
	K_{eq}	0.548	0.27	0.82	0.811	
	K_{OAA}	1.0	0.5	1.5	1.499	
	$K_{GDP/IDP}$	0.5	0.25	0.75	0.739	
	K_{PEP}	0.7	0.35	1.05	1.015	
	$K_{GTP/ITP}$	0.4	0.20	0.60	0.598	
	$K_{I,ATP}$	0.1	0.05	0.15	0.116	
odx	k_{for}	-	1e-6	1e3	2.109	[JGLS94] [JS95] [JS95]
	K_{OAA}	2.1	1.05	3.15	1.988	
	$K_{I,Suc}$	2.8	1.40	4.20	4.185	
	$K_{I,ADP}$	1.2	0.60	1.80	1.797	
mez	k_{for}	-	1e-6	1e3	0.003	Therm. Daten [GBLG00] [GBLG00] [GBLG00] [GBLG00]
	K_{eq}	1.657	0.829	2.486	2.602	
	K_{Mal}	3.8	1.90	5.70	2.680	
	K_{NADP}	0.08	0.04	0.12	0.053	
	K_{Pyr}	13.4	6.70	20.10	20.100	
	K_{NADPH}	0.06	0.03	0.09	0.090	

Fortsetzung nächste Seite

Reaktion	Parameter	<i>In vitro</i> Messwert	untere Grenze	obere Grenze	<i>In vivo</i> Schätzwert	Referenz
gdh	k_{for}	-	1e-6	1e3	11.294	Therm. Daten [SO70] [SO70] [SO70] [SO70]
	K_{eq}	5e6	3e6	7.5e6	7.5e6	
	K_{AKG}	5.72	2.86	8.58	5.113	
	K_{NADPH}	0.027	0.014	0.041	0.014	
	K_{Glu}	100.0	50.0	150.0	149.997	
aspB	K_{NAPD}	0.033	0.017	0.050	0.050	Therm. Daten [SMO82] [SMO82]
	k_{for}	-	1e-6	1e3	10.940	
	K_{eq}	1.657	0.829	2.486	1.224	
	K_{OAA}	0.11	0.06	0.17	0.123	
	K_{Glu}	24.0	12.0	36.0	29.448	
bsLYS	K_{AKG}	-	1e-6	5e2	499.982	
	K_{Asp}	-	1e-6	5e2	499.984	
	k_{for}	-	1e-6	1e3	0.121	
	K_{Asp}	-	1e-6	5e2	1.996	
	K_{Pyr}	-	1e-6	5e2	0.481	

Literaturverzeichnis

- [AEZ⁺06] ALTINTAS, M. M., C. K. EDDY, M. ZHANG, J. D. MCMILLAN und D. S. KOMPALA: *Kinetic Modeling to Optimize Pentose Fermentation in Zymomonas mobilis*. *Biotechnol Bioeng*, 94(2):273–295, 2006.
- [AFN04] AKESSON, M., J. FÖRSTER und J. NIELSEN: *Integration of gene expression data into genome-scale metabolic models*. *Metab Eng*, 6(4):285–293, 2004.
- [AKS07a] ANTONIEWICZ, M. R., J. K. KELLEHER und G. STEPHANOPOULOS: *Accurate assessment of amino Acid mass isotopomer distributions for metabolic flux analysis*. *Anal Chem*, 79(19):7554–9, 2007.
- [AKS07b] ANTONIEWICZ, M. R., J. K. KELLEHER und G. STEPHANOPOULOS: *Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions*. *Metab Eng*, 9(1):68–86, 2007.
- [AN99] AGGER, T. und J. NIELSEN: *Genetically structured modeling of protein production in filamentous fungi*. *Biotechnol Bioeng*, 66(3):164–70, 1999.
- [AP98] ASCHER, U. M. und L. R. PETZOLD: *Computer Methods for Ordinary Differential Equations and Differential Algebraic Equations*. *SIAM J. Sci. Comput.*, 1998.
- [Azu02] AZUAJE, F.: *A cluster validity framework for genome expression data*. *Bioinformatics*, 18(2):319–320, 2002.
- [BA03] BOLSHAKOVA, N. und F. AZUAJE: *Cluster validation techniques for genome expression data*. *Signal Processing*, 83:825–833, 2003.
- [BBS⁺03] BENDT, A. K., A. BURKOVSKI, S. SCHAFFER, M. BOTT, M. FARWICK und T. HERMANN: *Towards a phosphoproteome map of Corynebacterium glutamicum*. *Proteomics*, 3(8):1637–1646, 2003.

- [BCE⁺09] BÜSCHER, J. M., D. CZERNIK, J. C. EWALD, U. SAUER und N. ZAMBONI: *Cross-Platform Comparison of Methods for Quantitative Metabolomics of Primary Metabolism*. *Anal Chem*, 81(6):2135–2143, 2009.
- [BD08] BREMER, H. und P. DENNIS: *Feedback control of ribosome function in Escherichia coli*. *Biochimie*, 90(3):493–9, 2008.
- [Bey01] BEYER, H.: *The Theory of Evolution Strategies*. Springer, Berlin, 2001.
- [BFK⁺06] BETTENBROCK, KATJA, SOPHIA FISCHER, ANDREAS KREMLING, KNUT JAHREIS, THOMAS SAUTER und ERNST-DIETER GILLES: *A quantitative approach to catabolite repression in Escherichia coli*. *J Biol Chem*, 281(5):2578–2584, Feb 2006.
- [BFM⁺07] BECKER, S. A., A. M. FEIST, M. L. MO, G. HANNUM, B. Ø. PALSSON und M. J. HERRGARD: *Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox*. *Nat Protoc*, 2(3):727–738, 2007.
- [Bis02] BISSWANGER, H.: *Enzyme Kinetics, Principles and Methods*. Wiley-VCH, 2002.
- [BKL⁺07] BOLTEN, C. J., P. KIEFER, F. LETISSE, J. C. OORTAIS und C. WITTMANN: *Sampling for metabolome analysis of microorganisms*. *Anal Chem*, 79(10):3843–3849, 2007.
- [BKN05] BORODINA, I., P. KRABBen und J. NIELSEN: *Genome-scale analysis of Streptomyces coelicolor A3(2) metabolism*. *Genome Res*, 15(6):820–829, 2005.
- [BKW08] BECKER, J., C. KLOPPROGGE und C. WITTMANN: *Metabolic responses to pyruvate kinase deletion in lysine producing Corynebacterium glutamicum*. *Microb Cell Fact*, 7:8, 2008.
- [BMKB07] BERTH, M., F. M. MOSER, M. KOLBE und J. BERNHARDT: *The state of the art in the analysis of two-dimensional gel electrophoresis images*. *Appl Microbiol Biotechnol*, 76(6):1223–1243, 2007.
- [BN08] BORUKHOV, S. und E. NUDLER: *RNA polymerase: the vehicle of transcription*. *Trends Microbiol*, 16(3):126–34, 2008.

- [BO84] BAILEY, J. E. und A. OLIIS: *Biochemical Engineering Fundamentals*. McGraw Hill, New York, 1984.
- [Bus00] BUSTIN, S. A.: *Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays*. J Mol Endocrinol, 25(2):169–93, 2000.
- [CAF⁺07] CRAMER, ANNETTE, MARC AUCHTER, JULIA FRUNZKE, MICHAEL BOTT und BERNHARD J EIKMANN: *RamB, the transcriptional regulator of acetate metabolism in Corynebacterium glutamicum, is subject to regulation by RamA and RamB*. J Bacteriol, 189(3):1145–1149, Feb 2007.
- [Cel06] CELLIER, F. E.: *Continuous System Modeling*. Springer-Verlag New York, Inc., NY, 2006.
- [CGS⁺06] CRAMER, ANNETTE, ROBERT GERSTMEIR, STEFFEN SCHAFFER, MICHAEL BOTT und BERNHARD J EIKMANN: *Identification of RamA, a novel LuxR-type transcriptional regulator of genes involved in acetate metabolism of Corynebacterium glutamicum*. J Bacteriol, 188(7):2554–2567, Apr 2006.
- [CH88] CHATTERJEE, S. und A.S. HADI: *Sensitivity Analysis in Linear Regression, Probability and Mathematical Statistics*. John Wiley & Sons, New York, 1988.
- [CMB⁺07] COSTENOBLE, R., D. MULLER, T. BARL, W. M. VAN GULIK, W. A. VAN WINDEN, M. REUSS und J. J. HEIJNEN: *¹³C-Labeled metabolic flux analysis of a fed-batch culture of elutriated Saccharomyces cerevisiae*. FEMS Yeast Res, 7(4):511–526, 2007.
- [CNRS⁺02] CHASSAGNOLE, C., N. NOISOMMIT-RIZZI, J. W. SCHMID, K. MAUCH und M. REUSS: *Dynamic modeling of the central carbon metabolism of Escherichia coli*. Biotechnol Bioeng, 79(1):53–73, 2002.
- [CXCK08] COVERT, MARKUS W, NAN XIAO, TIFFANY J CHEN und JONATHAN R KARR: *Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli*. Bioinformatics, 24(18):2044–2050, Sep 2008.

- [DB73] DENNIS, P. P. und H. BREMER: *Regulation of ribonucleic acid synthesis in Escherichia coli B-r: an analysis of a shift-up. 1. Ribosomal RNA chain growth rates.* J Mol Biol, 75(1):145–59, 1973.
- [DBJ⁺07] DUARTE, N. C., S. A. BECKER, N. JAMSHIDI, I. THIELE, M. L. MO, T. D. VO, R. SRIVAS und B. Ø. PALSSON: *Global reconstruction of the human metabolic network based on genomic and bibliomic data.* Proc Natl Acad Sci U S A, 104(6):1777–1782, 2007.
- [DDLEG04] DELAUNAY, STÉPHANE, PASCALE DARAN-LAPUJADE, JEAN-MARC ENGASSER und JEAN-LOUIS GOERGEN: *Glutamate as an inhibitor of phosphoenolpyruvate carboxylase activity in Corynebacterium glutamicum.* J Ind Microbiol Biotechnol, 31(4):183–188, May 2004.
- [DEB04] DENNIS, P. P., M. EHRENBERG und H. BREMER: *Control of rRNA synthesis in Escherichia coli: a systems biology approach.* Microbiol Mol Biol Rev, 68(4):639–68, 2004.
- [DEMM⁺03] DRYSCH, A., M. EL MASSAOUDI, C. MACK, R. TAKORS, A. A. DE GRAAF und H. SAHM: *Production process monitoring by serial mapping of microbial carbon flux distributions using a novel Sensor Reactor approach: II-(13)C-labeling-based metabolic flux analysis and L-lysine production.* Metab Eng, 5(2):96–107, 2003.
- [DEMW⁺04] DRYSCH, A., M. EL MASSAOUDI, W. WIECHERT, A. A. DE GRAAF und R. TAKORS: *Serial flux mapping of Corynebacterium glutamicum during fed-batch L-lysine production using the sensor reactor approach.* Biotechnol Bioeng, 85(5):497–505, 2004.
- [Deu95] DEUFLHARD, P.: *Newton Methods for Highly Nonlinear Problems: Theory, Algorithms, Code.* Academic Press, 1995.
- [dJ02] JONG, H. DE: *Modeling and simulation of genetic regulatory systems: a literature review.* J Comput Biol, 9(1):67–103, 2002.
- [DNNW09] DROSTE, P., S. NOACK, K. NÖH und W. WIECHERT: *Customizable Visualization of Multi-Omics Data in the Context of Biochemical Networks.* In: *Proceedings of the 13th International Conference on Information Visualization, Barcelona, Spain, 2009.*

-
- [Dre01] DREW, D. A.: *A mathematical model for prokaryotic protein synthesis*. Bull Math Biol, 63(2):329–351, 2001.
- [DY75] DALBOW, D. G. und R. YOUNG: *Synthesis time of beta-galactosidase in Escherichia coli B/r as a function of growth rate*. Biochem J, 150(1):13–20, 1975.
- [EB05] EGGELING, L. und M. BOTT (Herausgeber): *Handbook of Corynebacterium glutamicum*. Taylor & Francis Group, LLC, 2005.
- [ECP02] EDWARDS, J. S., M. COVERT und B. PALSSON: *Metabolic modelling of microbes: the flux-balance approach*. Environ Microbiol, 4(3):133–140, 2002.
- [EGS⁺02] EGELHOFER, V., J. GOBOM, H. SEITZ, P. GIAVALISCO, H. LEHRACH und E. NORDHOFF: *Protein identification by MALDI-TOF-MS peptide mapping: a new strategy*. Anal Chem, 74(8):1760–1771, 2002.
- [Elm93] ELMQVIST, H.: *Object-oriented modeling and automatic formula manipulation in Dymola*. In: SOCIETY, SCANDINAVIAN SIMULATION (Herausgeber): SIMS '93 Kongsberg, Norway, 1993.
- [EMSD⁺03] EL MASSAOUDI, M., J. SPELTHAHN, A. DRYSCH, A. DE GRAAF und R. TAKORS: *Production process monitoring by serial mapping of microbial carbon flux distributions using a novel sensor reactor approach: I-Sensor reactor system*. Metab Eng, 5(2):86–95, 2003.
- [ENW08] ELSHEIKH, A., S. NOACK und W. WIECHERT: *Sensitivity Analysis of Modelica Applications via Automatic Differentiation*. In: 6th International Modelica Conference, Bielefeld, Germany, 2008.
- [ERP02] EDWARDS, JEREMY S, RAMPRASAD RAMAKRISHNA und BERNHARD O PALSSON: *Characterizing the metabolic phenotype: a phenotype phase plane analysis*. Biotechnol Bioeng, 77(1):27–36, Jan 2002.
- [ERS95] EIKMANN, B. J., D. RITTMANN und H. SAHM: *Cloning, sequence analysis, expression, and inactivation of the Corynebacterium glutamicum icd gene encoding isocitrate dehydrogenase and biochemical characterization of the enzyme*. J Bacteriol, 177(3):774–782, Feb 1995.

- [ETSE⁺94] EIKMANN, B. J., N. THUM-SCHMITZ, L. EGGELING, K. U. LÜDTKE und H. SAHM: *Nucleotide sequence, expression and transcriptional analysis of the Corynebacterium glutamicum gltA gene encoding citrate synthase*. Microbiology, 140 (Pt 8):1817–1828, Aug 1994.
- [FFF⁺03] FÖRSTER, J., I. FAMILI, P. FU, B. Ø. PALSSON und J. NIELSEN: *Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network*. Genome Res, 13(2):244–253, 2003.
- [FHR⁺07] FEIST, A. M., C. S. HENRY, J. L. REED, M. KRUMMENACKER, A. R. JOYCE, P. D. KARP, L. J. BROADBELT, V. HATZIMANIKATIS und B. Ø. PALSSON: *A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information*. Mol Syst Biol, 3:121, 2007.
- [FL71] FORCHHAMMER, J. und L. LINDAHL: *Growth rate of polypeptide chains as a function of the cell growth rate in a mutant of Escherichia coli 15*. J Mol Biol, 55(3):563–8, 1971.
- [FLNP00] FRIEDMAN, N., M. LINIAL, I. NACHMAN und D. PE'ER: *Using Bayesian networks to analyze expression data*. J Comput Biol, 7(3-4):601–620, 2000.
- [FSP⁺06] FEIST, A. M., J. C. M. SCHOLTEN, B. Ø. PALSSON, F. J. BROCKMAN und T. IDEKER: *Modeling methanogenesis with a genome-scale metabolic reconstruction of Methanosarcina barkeri*. Mol Syst Biol, 2:2006.0004, 2006.
- [FTM03] FUNAHASHI, A., N. TANIMURA und M. ET AL. MOROHASHI: *Cell-Designer: a process diagram editor for gene-regulatory and biochemical networks*. BIOSILICO, 1:159–162, 2003.
- [FZS04] FISCHER, E., N. ZAMBONI und U. SAUER: *High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived 13C constraints*. Anal Biochem, 325(2):308–316, 2004.
- [GBLG00] GOURDON, P., M. F. BAUCHER, N. D. LINDLEY und A. GUYON-VARCH: *Cloning of the malic enzyme gene from Corynebacterium glutamicum and role of the enzyme in lactate metabolism*. Appl Environ Microbiol, 66(7):2981–2987, Jul 2000.

- [GH03] GRUNDY, F. J. und T. M. HENKIN: *The T box and S box transcription termination control systems*. Front Biosci, 8:d20–31, 2003.
- [GW09] GRIEWANK, A. und A. WALTHER: *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. SIAM, 2009.
- [GWO06] GENDA, T., S. WATABE und H. OZAKI: *Purification and characterization of fumarase from Corynebacterium glutamicum*. Biosci Biotechnol Biochem, 70(5):1102–9, 2006. 0916-8451 (Print) Journal Article.
- [GWS⁺03] GERSTMEIR, ROBERT, VOLKER F WENDISCH, STEPHANIE SCHNICK, HONG RUAN, MIKE FARWICK, DIETER REINSCHIED und BERNHARD J EIKMANN: *Acetate metabolism and its regulation in Corynebacterium glutamicum*. J Biotechnol, 104(1-3):99–122, Sep 2003.
- [Hei05] HEIJNEN, J. J.: *Approximative kinetic formats used in metabolic network modeling*. Biotechnol Bioeng, 91(5):534–45, 2005.
- [HF07] HELLMAN, L. M. und M. G. FRIED: *Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions*. Nat Protoc, 2(8):1849–1861, 2007.
- [HFS⁺03] HUCKA, M., A. FINNEY, H. M. SAURO, H. BOLOURI, J. C. DOYLE, H. KITANO, A. P. ARKIN, B. J. BORNSTEIN, D. BRAY, A. CORNISH-BOWDEN, A. A. CUELLAR, S. DRONOV, E. D. GILLES, M. GINKEL, V. GOR, II GORYANIN, W. J. HEDLEY, T. C. HODGMAN, J. H. HOFMEYR, P. J. HUNTER, N. S. JUTY, J. L. KASBERGER, A. KREMLING, U. KUMMER, N. LE NOVERE, L. M. LOEW, D. LUCIO, P. MENDES, E. MINCH, E. D. MJOLSNESS, Y. NAKAYAMA, M. R. NELSON, P. F. NIELSEN, T. SAKURADA, J. C. SCHAFF, B. E. SHAPIRO, T. S. SHIMIZU, H. D. SPENCE, J. STELLING, K. TAKAHASHI, M. TOMITA, J. WAGNER und J. WANG: *The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models*. Bioinformatics, 19(4):524–531, 2003.
- [HKRP05] HEINEMANN, M., A. KÜMMEL, R. RUINATSCHA und S. PANKE: *In silico genome-scale reconstruction and validation of the Staphylococcus aureus metabolic network*. Biotechnol Bioeng, 92(7):850–864, 2005.

- [HMC⁺01] HOLTER, N. S., A. MARITAN, M. CIEPLAK, N. V. FEDOROFF und J. R. BANAVAR: *Dynamic modeling of gene expression data*. PNAS, 98(4):1693–1698, 2001.
- [HMM⁺00] HOLTER, N. S., M. MITRA, A. MARITAN, M. CIEPLAK, J. R. BANAVAR und N. V. FEDOROFF: *Fundamental patterns underlying gene expression profiles: simplicity from complexity*. PNAS, 97(15):8409–8414, 2000.
- [HNW08] HADLICH, F., S. NOACK und W. WIECHERT: *Translating biochemical network models between different kinetic formats*. Metab Eng, 11(2):87–100, 2008.
- [HPB⁺01] HERMANN, T., W. PFEFFERLE, C. BAUMANN, E. BUSKER, S. SCHAFER, M. BOTT, H. SAHM, N. DUSCH, J. KALINOWSKI, A. PÜHLER, A. K. BENDT, R. KRÄMER und A. BURKOVSKI: *Proteome analysis of Corynebacterium glutamicum*. Electrophoresis, 22(9):1712–1723, 2001.
- [HR74] HEINRICH, R. und T. A. RAPOPORT: *A linear steady-state treatment of enzymatic chains. General properties, control and effector strength*. Eur J Biochem, 42(1):89–95, 1974.
- [HSG⁺06] HOOPS, S., S. SAHLE, R. GAUGES, C. LEE, J. PAHLE, N. SIMUS, M. SINGHAL, L. XU, P. MENDES und U. KUMMER: *COPASI—a COmplex Pathway Simulator*. Bioinformatics, 22(24):3067–3074, 2006.
- [HSG⁺07] HUEGE, J., R. SULPICE, Y. GIBON, J. LISEC, K. KOEHL und J. KOPKA: *GC-EI-TOF-MS analysis of in vivo carbon-partitioning into soluble metabolite pools of higher plants by monitoring isotope dilution after ¹³CO₂ labelling*. Phytochemistry, 68(16-18):2258–2272, 2007.
- [JGLS94] JETTEN, M. S., M. E. GUBLER, S. H. LEE und A. J. SINSKEY: *Structural and functional analysis of pyruvate kinase from Corynebacterium glutamicum*. Appl Environ Microbiol, 60(7):2501–2507, Jul 1994.
- [JR89] JANSON, J.C. und L. RYDEN (Herausgeber): *Protein Purification: Principles, High Resolution Methods, and Applications*. Wiley-VCH, New York, 1989.
- [JS93] JETTEN, M. S. und A. J. SINSKEY: *Characterization of phosphoenolpyruvate carboxykinase from Corynebacterium glutamicum*. FEMS Microbiol Lett, 111:183–188, 1993.

- [JS95] JETTEN, M. S. und A. J. SINSKEY: *Purification and properties of oxaloacetate decarboxylase from Corynebacterium glutamicum*. *Antonie Van Leeuwenhoek*, 67(2):221–7, 1995. 0003-6072 (Print) Comparative Study Journal Article Research Support, Non-U.S. Gov't.
- [Kau69] KAUFFMAN, S.: *Homeostasis and differentiation in random genetic control networks*. *Nature*, 224(5215):177–178, 1969.
- [KB73] KACSER, H. und J. A. BURNS: *The control of flux*. *Symp Soc Exp Biol*, 27:65–104, 1973.
- [KBB⁺03] KALINOWSKI, J., B. BATHE, D. BARTELS, N. BISCHOFF, M. BOTT, A. BURKOVSKI, N. DUSCH, L. EGGELING, B. J. EIKMANN, L. GAI-GALAT, A. GOESMANN, M. HARTMANN, K. HUTHMACHER, R. KRÄ-MER, B. LINKE, A. C. MCHARDY, F. MEYER, B. MÖCKEL, W. PFEFFERLE, A. PÜHLER, D. A. REY, C. RÜCKERT, O. RUPP, H. SAHM, V. F. WENDISCH, I. WIEGRÄBE und A. TAUCH: *The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins*. *J Biotechnol*, 104(1-3):5–25, Sep 2003.
- [KG01] KREMLING, A. und E. D. GILLES: *The organization of metabolic reaction networks. II. Signal processing in hierarchical structured functional units*. *Metab Eng*, 3(2):138–50, 2001.
- [KHZW04] KIEFER, P., E. HEINZLE, O. ZELDER und C. WITTMANN: *Comparative metabolic flux analysis of lysine-producing Corynebacterium glutamicum cultured on glucose or fructose*. *Appl Environ Microbiol*, 70(1):229–239, 2004.
- [KJAS02] KOFFAS, MATTHEOS A G, GYOO YEOL JUNG, JUAN C AON und GREGORY STEPHANOPOULOS: *Effect of pyruvate carboxylase overexpression on the physiology of Corynebacterium glutamicum*. *Appl Environ Microbiol*, 68(11):5422–5428, Nov 2002.
- [KN09] KJELDSSEN, K. R. und J. NIELSEN: *In silico genome-scale reconstruction and validation of the Corynebacterium glutamicum metabolic network*. *Biotechnol Bioeng*, 102(2):583–597, 2009.

- [KR03] KOOLMAN, J. und K. H. ROHM: *Taschenatlas der Biochemie*, Band 3. Georg Thieme, Stuttgart, 2003.
- [KRA07] KACZANOWSKA, M. und M. RYDEN-AULIN: *Ribosome biogenesis and the translation process in Escherichia coli*. *Microbiol Mol Biol Rev*, 71(3):477–94, 2007.
- [KSK⁺04] KRÖMER, J. O., O. SORGENFREI, K. KLOPPROGGE, E. HEINZLE und C. WITTMANN: *In-depth profiling of lysine-producing Corynebacterium glutamicum by combined analysis of the transcriptome, metabolome, and fluxome*. *J Bacteriol*, 186(6):1769–1784, 2004.
- [KTY98] KOH, B. T., R. B. TAN und M. G. YAP: *Genetically structured mathematical modeling of trp attenuator mechanism*. *Biotechnol Bioeng*, 58(5):502–9, 1998.
- [KWB05] KRUG, ANDREAS, VOLKER F WENDISCH und MICHAEL BOTT: *Identification of AcnR, a TetR-type repressor of the aconitase gene acn in Corynebacterium glutamicum*. *J Biol Chem*, 280(1):585–595, Jan 2005.
- [KYB⁺04] KAO, K. C., Y. L. YANG, R. BOSCOLO, C. SABATTI, V. ROYCHOWDHURY und J. C. LIAO: *Transcriptome-based determination of multiple transcription regulator activities in Escherichia coli by using network component analysis*. *PNAS*, 101(2):641–646, 2004.
- [LB00] LEE, S. B. und J. E. BAILEY: *Analysis of growth rate effects on productivity of recombinant Escherichia coli populations using molecular mechanism models*. *Biotechnol Bioeng*, 67(6):805–12, 2000.
- [LGT⁺07] LUO, B., K. GROENKE, R. TAKORS, C. WANDREY und M. OLDIGES: *Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography-mass spectrometry*. *J Chromatogr A*, 1147(2):153–164, 2007.
- [LH01] LEITOLD, A. und K. M. HANGOS: *Structural solvability analysis of dynamic process models*. *Computers & Chemical Engineering*, 25:1633–1646, 2001.

- [LK06] LIEBERMEISTER, W. und E. KLIPP: *Bringing metabolic networks to life: convenience rate law and thermodynamic constraints*. Theor Biol Med Model, 3:41, 2006. 1742-4682 (Electronic) Journal Article Research Support, Non-U.S. Gov't.
- [LPC⁺06] LIPNIACKI, T., P. PASZEK, A. M. CZOCHRA, A. R. BRASIER und M. KIMMEL: *Transcriptional stochasticity in gene expression*. J Theor Biol, 238(2):348–367, 2006.
- [LPDG00] LEE, S., C. PALAKORNKULE, M. M. DOMACH und I. E. GROSSMANN: *Recursive MILP model for finding all the alternate optima in LP models for metabolic networks*. Computers & Chemical Engineering, 24(2-7):711 – 716, 2000.
- [LPK⁺07] LEE, KWANG HO, JIN HWAN PARK, TAE YONG KIM, HYUN UK KIM und SANG YUP LEE: *Systems metabolic engineering of Escherichia coli for L-threonine production*. Mol Syst Biol, 3:149, 2007.
- [MAR97] MAUCH, K., S. ARNOLD und M. REUSS: *Dynamic sensitivity analysis for metabolic systems*. Chem Eng Sci, 52(15):2589–2598, 1997.
- [Mat93] MATHEWS, C. K.: *The cell-bag of enzymes or network of channels?* J Bacteriol, 175(20):6377–6381, 1993.
- [McC85] MCCLURE, W. R.: *Mechanism and control of transcription initiation in prokaryotes*. Annu Rev Biochem, 54:171–204, 1985.
- [MdGW⁺96] MARX, A., A. A. DE GRAAF, W. WIECHERT, L. EGGELING und H. SAHM: *Determination of the fluxes in the central metabolism of Corynebacterium glutamicum by nuclear magnetic resonance spectroscopy combined with metabolite balancing*. Biotechnology and Bioengineering, 49(2):111–129, 1996.
- [MF02] MICHALEWICZ, Z. und D.B. FOGEL: *How to Solve It: Modern Heuristics*. 2002.
- [MGL96] MAFFEZZONI, C., R. GIRELLI und P. LLUKA: *Generating efficient computational procedures from declarative models*. Simul. Pr. Theory, 4:303–317, 1996.

- [MHOT06] MAGNUS, J. B., D. HOLLWEDEL, M. OLDIGES und R. TAKORS: *Monitoring and modeling of the reaction dynamics in the valine/leucine synthesis pathway in Corynebacterium glutamicum*. Biotechnol Prog, 22(4):1071–1083, 2006.
- [MRDM⁺07] MASHEGO, M. R., K. RUMBOLD, M. DE MEY, E. VANDAMME, W. SOETAERT und J. J. HEIJNEN: *Microbial metabolomics: past, present and future methodologies*. Biotechnol Lett, 29(1):1–16, 2007.
- [MS85] MORI, M. und I. SHIIO: *Synergistic inhibition of phosphoenolpyruvate carboxylase by aspartate and 2-oxoglutarate in Brevibacterium flavum*. J Biochem, 98(6):1621–1630, Dec 1985.
- [MS93] MATTSSON, S. E. und G. SÖDERLIND: *Index reduction in differential-algebraic equations using dummy derivatives*. SIAM J. Sci. Comput., 14:677–692, 1993.
- [MSdG⁺97] MARX, A., K. STRIEGEL, A. A. DE GRAAF, H. SAHM und L. EGGE-LING: *Response of the central metabolism of Corynebacterium glutamicum to different flux burdens*. Biotechnology and Bioengineering, 56(2):168–180, 1997.
- [Mur87] MUROTA, K.: *Systems Analysis by Graphs and Matroids*. Springer, Berlin, 1987.
- [MvdRDY00] MOLENAAR, D., M. E. VAN DER REST, A. DRYSCH und R. YÜCEL: *Functions of the membrane-associated and cytoplasmic malate dehydrogenases in the citric acid cycle of Corynebacterium glutamicum*. J Bacteriol, 182(24):6884–6891, 2000.
- [NF05] NILSSON, E. L. und P. FRITZSON: *Biochemical and Metabolic Modeling and Simulation with Modelica*. In: *BioMedSim Linköping, Sweden, 2005*.
- [NGL⁺07] NÖH, K., K. GRÖNKE, B. LUO, R. TAKORS, M. OLDIGES und W. WIECHERT: *Metabolic flux analysis at ultra short time scale: isotopically non-stationary ¹³C labeling experiments*. J Biotechnol, 129(2):249–267, 2007.
- [Nöh06] NÖH, K.: *Isotopically Non-Stationary ¹³C-Labeling Experiments under Metabolic Steady State Conditions: Mathematical Modeling, Simulation, Statistical Evaluation and Experimental Design*. Dissertation, Universität Siegen, 2006.

- [NKB08] NOACK, S., W. KLÖDEN und T. BLEY: *Modeling synchronous growth of bacterial populations in phased cultivation*. Bioprocess Biosyst Eng, 31(5):435–443, Aug 2008.
- [NKS⁺06] NIEBISCH, A., A. KABUS, C. SCHULTZ, B. WEIL und M. BOTT: *Corynebacterial protein kinase G controls 2-oxoglutarate dehydrogenase activity via the phosphorylation status of the OdhI protein*. J Biol Chem, 281(18):12300–12307, 2006.
- [NM04] NUDLER, E. und A. S. MIRONOV: *The riboswitch control of bacterial metabolism*. Trends Biochem Sci, 29(1):11–7, 2004.
- [Nud06] NUDLER, E.: *Flipping riboswitches*. Cell, 126(1):19–22, 2006.
- [NWH⁺08] NOACK, S., S. A. WAHL, M. HAUNSCHILD, E. QELI, B. FREISLEBEN und W. WIECHERT: *Visualizing regulatory interdependencies and parameter sensitivities in biochemical network models*. Mathematics and Computers in Simulation, 79(4):991–998, 2008.
- [NWQW07] NOACK, S., S. A. WAHL, E. QELI und W. WIECHERT: *Visualizing regulatory interactions in metabolic networks*. BMC Biol, 5:46, 2007.
- [NWW06] NÖH, K., S. A. WAHL und W. WIECHERT: *Computational tools for isotopically instationary ¹³C labeling experiments under metabolic steady state conditions*. Metab Eng, 8(6):554–577, 2006.
- [OLP⁺07] OLDIGES, M., S. LÜTZ, S. PFLUG, K. SCHROER, N. STEIN und C. WIENDAHL: *Metabolomics: current state and evolving methodologies and tools*. Appl Microbiol Biotechnol, 76(3):495–511, 2007.
- [ONF05] OLIVEIRA, A. P., J. NIELSEN und J. FÖRSTER: *Modeling Lactococcus lactis using a genome-scale flux model*. BMC Microbiol, 5:39, 2005.
- [OPP⁺07] OH, Y. K., B. Ø. PALSSON, S. M. PARK, C. H. SCHILLING und R. MAHADEVAN: *Genome-scale reconstruction of metabolic network in Bacillus subtilis based on high-throughput phenotyping and gene essentiality data*. J Biol Chem, 282(39):28791–28799, 2007.
- [OS69] OZAKI, H. und I. SHIIO: *Regulation of the TCA and glyoxylate cycles in Brevibacterium flavum. II. Regulation of phosphoenolpyruvate carboxylase and pyruvate kinase*. J Biochem, 66(3):297–311, Sep 1969.

- [OT05] OLDIGES, M. und R. TAKORS: *Applying metabolic profiling techniques for stimulus-response experiments: chances and pitfalls*. Adv Biochem Eng Biotechnol, 92:173–196, 2005.
- [Pan98] PANTELIDES, C. C.: *The consistent initialization of differential-algebraic systems*. SIAM J. Sci. and Stat. Comput., 9:213–231, 1998.
- [PdGE⁺00] PETERSEN, S., A. A. DE GRAAF, L. EGGELING, M. MOLLNEY, W. WIECHERT und H. SAHM: *In vivo quantification of parallel and bi-directional fluxes in the anaplerosis of Corynebacterium glutamicum*. J Biol Chem, 275(46):35932–35941, 2000.
- [Pet00] PETERSEN, S.: *Untersuchungen zur in vivo Aktivität anaplerotischer Stoffwechselwege in Corynebacterium glutamicum mittels 13C-Markierungstechnik*. Dissertation, Düsseldorf, 2000.
- [PMBM03] PFEFFERLE, W., B. MÖCKEL, B. BATHE und A. MARX: *Biotechnological manufacture of lysine*. Adv Biochem Eng Biotechnol, 79:59–112, 2003.
- [PMdG⁺01] PETERSEN, S., C. MACK, A. A. DE GRAAF, C. RIEDEL, B. J. EIKMANNNS und H. SAHM: *Metabolic consequences of altered phosphoenolpyruvate carboxykinase activity in Corynebacterium glutamicum reveal anaplerotic regulation mechanisms in vivo*. Metab Eng, 3(4):344–361, 2001.
- [PREF01] PE'ER, D., A. REGEV, G. ELIDAN und N. FRIEDMAN: *Inferring sub-networks from perturbed expression profiles*. Bioinformatics, 17:215–224, 2001.
- [PW96] PETERS-WENDISCH, P. G.: *Anaplerotische Reaktionen in Corynebacterium glutamicum: Untersuchungen zur Bedeutung der PEP-Carboxylase und der Pyruvat-Carboxylase im Zentralstoffwechsel und bei der Aminosäure-Produktion*. Dissertation, Universität Düsseldorf, 1996.
- [PWWP⁺97] PETERS-WENDISCH, P. G., V. F. WENDISCH, S. PAUL, B. J. EIKMANNNS und H. SAHM: *Pyruvate carboxylase as an anaplerotic enzyme in Corynebacterium glutamicum*. Microbiology, 143:1095–1103, 1997.
- [RBTR97] RIZZI, M., M. BALTES, U. THEOBALD und M. REUSS: *In Vivo Analysis of Metabolic Dynamics in Saccharomyces cerevisiae: II. Mathematical Model*. Biotechnol Bioeng, 55(4):592–608, 1997.

- [RE07] RADMACHER, E. und L. EGGELING: *The three tricarboxylate synthase activities of Corynebacterium glutamicum and increase of L-lysine synthesis*. Appl Microbiol Biotechnol, 76(3):587–95, 2007. 0175-7598 (Print) Journal Article.
- [REMP01] RAMAKRISHNA, R., J. S. EDWARDS, A. MCCULLOCH und B. O. PALSSON: *Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints*. Am J Physiol Regul Integr Comp Physiol, 280(3):695–704, 2001.
- [RLB82] RYALS, J., R. LITTLE und H. BREMER: *Temperature dependence of RNA synthesis parameters in Escherichia coli*. J Bacteriol, 151(2):879–87, 1982.
- [RPKW98] ROHWER, J. M., P. W. POSTMA, B. N. KHOLODENKO und H. V. WESTERHOFF: *Implications of macromolecular crowding for signal transduction and metabolite channeling*. Proc Natl Acad Sci U S A, 95(18):10547–10552, 1998.
- [RSSK08] RICHARDS, J., T. SUNDERMEIER, A. SVETLANOV und A. W. KARZAI: *Quality control of bacterial mRNA decoding and decay*. Biochim Biophys Acta, 2008.
- [RVSP03] REED, J. L., T. D. VO, C. H. SCHILLING und B. Ø. PALSSON: *An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR)*. Genome Biol, 4(9):R54, 2003.
- [RZ06] ROUSSEL, MARC R und RUI ZHU: *Stochastic kinetics description of a simple transcription model*. Bull Math Biol, 68(7):1681–1713, Oct 2006.
- [Sau06] SAUER, U.: *Metabolic networks in motion: 13C-based flux analysis*. Mol Syst Biol, 2:62, 2006. 1744-4292 (Electronic) Journal Article Review.
- [Sav91] SAVAGEAU, M. A.: *Metabolite channeling: implications for regulation of metabolism and for quantitative description of reactions in vivo*. J Theor Biol, 152(1):85–92, 1991.
- [SBB00a] SMOLEN, P., D. A. BAXTER und J. H. BYRNE: *Mathematical modeling of gene networks*. Neuron, 26(3):567–580, Jun 2000.

- [SBB00b] SMOLEN, P., D. A. BAXTER und J. H. BYRNE: *Modeling transcriptional control in gene networks-methods, recent results, and future directions*. Bull Math Biol, 62(2):247–292, Mar 2000.
- [SC00] SALTELLI, A. und K. CHAN: *Sensitivity Analyses*. John Wiley & Sons, Inc., 2000.
- [SCF⁺02] SCHILLING, C. H., M. W. COVERT, I. FAMILI, G. M. CHURCH, J. S. EDWARDS und B. Ø. PALSSON: *Genome-scale metabolic model of Helicobacter pylori 26695*. J Bacteriol, 184(16):4582–4593, 2002.
- [Sch91] SCHRUMPF, B.: *Lysinbildung mit Corynebacterium glutamicum: Analyse des Metabolitflusses durch Bestimmung zellinterner Aminosäurekonzentrationen und enzymatische Untersuchungen*. Dissertation, Universität Düsseldorf, 1991.
- [Sch08] SCHWENDER, J.: *Metabolic flux analysis as a tool in metabolic engineering of plants*. Curr Opin Biotechnol, 19(2):131–137, 2008.
- [SDSE07] SEIBOLD, G., S. DEMPFF, J. SCHREINER und B. J. EIKMANN: *Glycogen formation in Corynebacterium glutamicum and role of ADP-glucose pyrophosphorylase*. Microbiology, 153(Pt 4):1275–1285, 2007.
- [SE05a] SAUER, UWE und BERNHARD J EIKMANN: *The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria*. FEMS Microbiol Rev, 29(4):765–794, Sep 2005.
- [SE05b] SCHREINER, MARK E und BERNHARD J EIKMANN: *Pyruvate:quinone oxidoreductase from Corynebacterium glutamicum: purification and biochemical characterization*. J Bacteriol, 187(3):862–871, Feb 2005.
- [Seg75] SEGEL, I. H.: *Enzyme Kinetics*. John Wiley & Sons, Inc., 1975.
- [SELP00] SCHILLING, C. H., J. S. EDWARDS, D. LETSCHER und B. Ø. PALSSON: *Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems*. Biotechnol Bioeng, 71(4):286–306, 2000.
- [SFL09] SCHÄDEL, F. und E. FRANCO-LARA: *Rapid sampling devices for metabolic engineering applications*. Appl Microbiol Biotechnol, 83(2):199–208, 2009.

- [SM01] SANTILLAN, M. und M. C. MACKEY: *Dynamic regulation of the tryptophan operon: a modeling study and comparison with experimental data.* Proc Natl Acad Sci U S A, 98(4):1364–1369, 2001.
- [SM09] STAMATAKIS, M. und N. V. MANTZARIS: *Comparison of deterministic and stochastic models of the lac operon genetic network.* Biophys J, 96(3):887–906, 2009.
- [SMK⁺06] SHIRAI, T., K. MATSUZAKI, M. KUZUMOTO, K. NAGAHISA, C. FURUSAWA, S. SHIOYA und H. SHIMIZU: *Precise metabolic flux analysis of coryneform bacteria by gas chromatography-mass spectrometry and verification by nuclear magnetic resonance.* J Biosci Bioeng, 102(5):413–424, 2006.
- [SMO82] SHIIO, I., M. MORI und H. OZAKI: *Amino acid aminotransferases in an amino acid producing bacterium, Brevibacterium flavum.* Agric. Biol. Chem., 46:2967–2977, 1982.
- [SMR⁺04] SCHMID, JOACHIM W, KLAUS MAUCH, MATTHIAS REUSS, ERNST D GILLES und ANDREAS KREMLING: *Metabolic design based on a coupled gene expression-metabolic network model of tryptophan production in Escherichia coli.* Metab Eng, 6(4):364–377, Oct 2004.
- [SMR08] SCHAUB, J., K. MAUCH und M. REUSS: *Metabolic flux analysis in Escherichia coli by integrating isotopic dynamic and isotopic stationary ¹³C labeling data.* Biotechnol Bioeng, 99(5):1170–1185, 2008.
- [SO70] SHIIO, I. und H. OZAKI: *Regulation of nicotinamide adenine dinucleotide phosphate-specific glutamate dehydrogenase from Brevibacterium flavum, a glutamate-producing bacterium.* J Biochem, 68(5):633–47, 1970. 0021-924X (Print) Journal Article.
- [SP92] SAVINELL, J. M. und B. Ø. PALSSON: *Optimal selection of metabolic fluxes for in vivo measurement. I. Development of mathematical methods.* J Theor Biol, 155(2):201–214, 1992.
- [SP00] SCHILLING, C. H. und B. Ø. PALSSON: *Assessment of the metabolic capabilities of Haemophilus influenzae Rd through a genome-scale pathway analysis.* J Theor Biol, 203(3):249–283, 2000.

- [SS96] SOMOGYI, R. und C. SNIEGOSKI: *Modeling the complexity of gene networks: understanding multigenic and pleiotropic regulation*. Complexity, 1:45–63, 1996.
- [Ste99] STEPHANOPOULOS, G.: *Metabolic fluxes and metabolic engineering*. Metab Eng, 1(1):1–11, 1999.
- [STS84] SHIIO, I., Y. TORIDE und S. SUGIMOTO: *Production of lysine by pyruvate dehydrogenase mutants of Brevibacterium flavum*. Agric. Biol. Chem., 48:3091–3098, 1984.
- [Stu02] STULKE, J.: *Control of transcription termination in bacteria by RNA-binding proteins that modulate RNA structures*. Arch Microbiol, 177(6):433–40, 2002.
- [SU80] SHIIO, I. und K. UJIGAWA: *Presence and regulation of α -ketoglutarate dehydrogenase complex in a glutamate-producing bacterium, Brevibacterium flavum*. Agric. Biol. Chem., 44:1897–1904, 1980.
- [SV82] SAVAGEAU, M. A. und E. O. VOIT: *Power-law approach to modeling biological systems. I. Theory*. Journal of Fermentation Technology, 60(3):221–228, 1982.
- [Szy95] SZYPERSKI, T.: *Biosynthetically directed fractional ^{13}C -labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism*. Eur J Biochem, 232(2):433–448, 1995.
- [Szy98] SZYPERSKI, T.: *^{13}C -NMR, MS and metabolic flux balancing in biotechnology research*. Q Rev Biophys, 31(1):41–106, 1998.
- [TAKP07] TROEIN, C., D. AHREN, M. KROGH und C. PETERSON: *Is transcriptional regulation of metabolic pathways an optimal strategy for fitness?* PLoS ONE, 2(9):e855, 2007. 1932-6203 (Electronic) Journal Article.
- [TCN03] TYSON, J. J., K. C. CHEN und B. NOVAK: *Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell*. Curr Opin Cell Biol, 15(2):221–31, 2003.
- [TFS⁺08] THIELE, B., K. FÜLLNER, N. STEIN, M. OLDIGES, A.J. KUHN und D. HOFMANN: *Analysis of amino acids without derivatization in barley extracts by LC-MS-MS*. Anal Bioanal Chem, 391(7):2663–2672, 2008.

- [Til08] TILLACK, J.: *Kinetikbasierte Modellierung und Simulation isotopisch in-stationärer metabolischer Netzwerke*. Diplomarbeit, Universität, Magdeburg, 2008.
- [TMM⁺09] TANG, Y. J., H. G. MARTIN, S. MYERS, S. RODRIGUEZ, E. E. K. BAIDOO und KEASLING J. D.: *Advances in analysis of microbial metabolic fluxes via (13)C isotopic labeling*. *Mass Spectrom Rev*, 28(2):362–375, 2009.
- [TNE⁺09] TILLACK, J., S. NOACK, A. ELSHEIHK, K. NÖH und W. WIECHERT: *A Software Framework for Modeling and Simulation of Dynamic Metabolic and Isotopic Systems*. In: *Proceedings of 6th International Conference on Mathematical Modeling, Argesim, Vienna, Austria*, 2009.
- [TWM⁺06] TEUSINK, B., A. WIERSMA, D. MOLENAAR, C. FRANCKE, W.M. DE VOS, R.J. SIEZEN und E.J. SMID: *Analysis of growth of Lactobacillus plantarum WCFS1 on a complex medium using a genome-scale metabolic model*. *J Biol Chem*, 281(52):40041–40048, 2006.
- [VBMA⁺05] VILLAS-BÔAS, S. G., S. MAS, M. AKESSON, J. SMEDSGAARD und J. NIELSEN: *Mass spectrometry in metabolome analysis*. *Mass Spectrom Rev*, 24(5):613–646, 2005.
- [VBRR99] VASEGHI, S., A. BAUMEISTER, M. RIZZI und M. REUSS: *In Vivo Dynamics of the Pentose Phosphate Pathway in Saccharomyces cerevisiae*. *Metab Eng*, 1:128–140, 1999.
- [vL03] LIERES, E. VON: *Integrative Modellierung und Auswertung stationärer Datensätze für das Metabolic Engineering*. Dissertation, Universität Siegen, 2003.
- [VP94] VARMA, A. und B. Ø. PALSSON: *Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use*. *Nat Rev Biotechnol*, 12:994–998, 1994.
- [VSM⁺04] VISSER, DIANA, JOACHIM W SCHMID, KLAUS MAUCH, MATTHIAS REUSS und JOSEPH J HEIJNEN: *Optimal re-design of primary metabolism in Escherichia coli using linlog kinetics*. *Metab Eng*, 6(4):378–390, Oct 2004.

- [vWHV02] WINDEN, W. A. VAN, J. J. HEIJNEN und P. J. T. VERHEIJEN: *Cumulative bondomers: a new concept in flux analysis from 2D [13C,1H] COSY NMR data*. *Biotechnol Bioeng*, 80(7):731–745, 2002.
- [vWvDR⁺05] WINDEN, W. A. VAN, J. C. VAN DAM, C. RAS, R. J. KLEIJN, J. L. VINKE, W. M. VAN GULIK und J. J. HEIJNEN: *Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of (13)C-labeled primary metabolites*. *FEMS Yeast Res*, 5(6-7):559–568, 2005.
- [Wah07] WAHL, S. A.: *Methoden zur integrierten Analyse metabolischer Netzwerke unter stationären und instationären Bedingungen*. Dissertation, Universität, Siegen, 2007.
- [WdG97] WIECHERT, W. und A. A. DE GRAAF: *Bidirectional reaction steps in metabolic networks: I. Modeling and simulation of carbon isotope labeling experiments*. *Biotechnology and Bioengineering*, 55(1):101–117, 1997.
- [WdGSE00] WENDISCH, V. F., A. A. DE GRAAF, H. SAHM und B. J. EIKMANN: *Quantitative determination of metabolic fluxes during coutilization of two carbon sources: comparative analyses with *Corynebacterium glutamicum* during growth on acetate and/or glucose*. *J Bacteriol*, 182(11):3088–3096, 2000.
- [WFM⁺98] WEN, X., S. FUHRMAN, G. S. MICHAELS, D. B. CARR, S. SMITH, J. L. BARKER und R. SOMOGYI: *Large-scale temporal gene expression mapping of central nervous system development*. *PNAS*, 95(1):334–339, 1998.
- [WH02] WITTMANN, C. und E. HEINZLE: *Genealogy profiling through strain improvement by using metabolic network analysis: metabolic flux genealogy of several generations of lysine-producing corynebacteria*. *Appl Environ Microbiol*, 68(12):5843–5859, 2002.
- [WHOW06] WAHL, S. A., M. D. HAUNSCHILD, M. OLDIGES und W. WIECHERT: *Unravelling the regulatory structure of biochemical networks using stimulus response experiments and large-scale model selection*. *Syst Biol*, 153(4):275–285, 2006.

- [Wie95] WIECHERT, W.: *Metabolische Kohlenstoff-Markierungsexperimente: Modellierung, Simulation, Analyse, Datenauswertung*. Habilitationsschrift, Universität Bonn, 1995.
- [WKB05] WENNERHOLD, J., A. KRUG und M. BOTT: *The AraC-type regulator RipA represses aconitase and other iron proteins from Corynebacterium under iron limitation and is itself repressed by DtxR*. J Biol Chem, 280(49):40500–40508, 2005.
- [WKM03] WOLF, A., R. KRÄMER und S. MORBACH: *Three pathways for trehalose metabolism in Corynebacterium glutamicum ATCC13032 and their significance in response to osmotic stress*. Mol Microbiol, 49(4):1119–1134, 2003.
- [WMI+99] WIECHERT, W., M. MÖLLNEY, N. ISERMANN, M. WURZEL und A. A. DE GRAAF: *Bidirectional reaction steps in metabolic networks: III. Explicit solution and analysis of isotopomer labeling systems*. Biotechnology and Bioengineering, 66(2):69–85, 1999.
- [WMPdG01] WIECHERT, W., M. MOLLNEY, S. PETERSEN und A. A. DE GRAAF: *A universal framework for ¹³C metabolic flux analysis*. Metab Eng, 3(3):265–83, 2001. 1096-7176 (Print) Journal Article Research Support, Non-U.S. Gov't.
- [WMS+04] WAHL, S. A., M. E. MASSAOUDI, D. SCHIPPER, W. WIECHERT und R. TAKORS: *Serial ¹³C-based flux analysis of an L-phenylalanine-producing E. coli strain using the sensor reactor*. Biotechnol Prog, 20(3):706–714, 2004.
- [WN05] WIECHERT, W. und K. NÖH: *From stationary to instationary metabolic flux analysis*. Adv Biochem Eng Biotechnol, 92:145–172, 2005.
- [WNE09] WIECHERT, W., S. NOACK und A. ELSHEIHK: *Modeling Languages for Biochemical Network Simulation: SBML versus Modelica*. ABB (submitted), 2009.
- [WO03] WEITZKE, E. L. und P. J. ORTOLEVA: *Simulating cellular dynamics through a coupled transcription, translation, metabolic model*. Comput Biol Chem, 27(4-5):469–80, 2003.
- [WvD87] WESTERHOFF, H. V. und K. VAN DAM: *Thermodynamics and control of biological free energy transduction*. Elsevier, Amsterdam, 1987.

- [WWR⁺08] WELLERDIEK, M., D. WINTERHOFF, W. REULE, J. BRANDNER und M. OLDIGES: *Metabolic quenching of Corynebacterium glutamicum: efficiency of methods and impact of cold shock*. Bioprocess Biosyst Eng, 2008.
- [WYH⁺08] WANG, Z., F. YANG, D. W. C HO, S. SWIFT, A. TUCKER und X. LIU: *Stochastic dynamic modeling of short gene expression time-series data*. IEEE Trans Nanobioscience, 7(1):44–55, 2008.
- [XZD97] XIU, ZHI-LONG, AN-PING ZENG und WOLF-DIETER DECKWER: *Model analysis concerning the effects of growth rate and intracellular tryptophan level on the stability and dynamics of tryptophan biosynthesis in bacteria*. Journal of Biotechnology, 58(2):125 – 140, 1997.
- [YM03] YILDIRIM, N. und M. C. MACKEY: *Feedback regulation in the lactose operon: a mathematical modeling study and comparison with experimental data*. Biophys J, 84(5):2841–2851, 2003.
- [YWH06] YANG, T. H., C. WITTMANN und E. HEINZLE: *Respirometric ¹³C flux analysis-Part II: in vivo flux estimation of lysine-producing Corynebacterium glutamicum*. Metab Eng, 8(5):432–446, 2006.
- [ZFL09] ZEISER, S., U. FRANZ und V. LIEBSCHER: *Autocatalytic genetic networks modeled by piecewise-deterministic Markov processes*. J Math Biol, 2009.
- [ZFRS09] ZAMBONI, N., S. M. FENDT, M. RÜHL und U. SAUER: *¹³C-based metabolic flux analysis*. Nat Protoc, 4, 2009.