

Modellierung stationärer Zustände von metabolischen Netzwerken:
Methoden, Anwendungen, Thermodynamik

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Kai Hartmann

aus Peine

Köln, 2006

Berichterstatter: Prof. Dr. D. Schomburg
Prof. Dr. R. Schrader

Tag der mündlichen Prüfung: 12.02.2007

Danke

Prof. Dr. Dietmar Schomburg danke ich für die interessante Aufgabenstellung, die fachlichen Anregungen und Diskussionen sowie für die Möglichkeit, die Arbeit am Institut für Biochemie der Universität zu Köln durchführen zu können.

Prof. Dr. Rainer Schrader danke ich für die fachlichen Diskussionen und die freundliche Übernahme des Zweitgutachtens.

Prof. Dr. Gerd Meyer und PD Dr. Karsten Niefind danke ich für die Übernahme des Vorsitzes bzw. Beisitzes bei meiner Disputation.

Bei der Korrektur dieser Arbeit waren mir Thorsten Fallisch, Silke Schrader und Markus Leber behilflich. Dafür meinen herzlichen Dank.

Während der Promotionszeit haben mir viele Personen aus dem fachlichen Bereich geholfen. Ich bedanke mich für die fachliche Unterstützung bei Thorsten Fallisch, Sebastian Buchinger, Silke Schrader, Christoph Steinbeck, Miguel Rojas, Asad Rahman und Eric von Lieres. Für die Hilfe bei den ersten Schritten der Programmierung bis hin zu fortgeschrittenen Fragen der Softwareentwicklung danke ich Stefan Kuhn (Java) und Ralph Schunk (ISO C++). Für die systemadministrative Betreuung bedanke ich mich bei Peter Krämer, Daniel Wetzler und Michael Zimmermann. Bei der Lunch-Gruppe möchte ich mich für sehr interessante, witzige und teilweise bizarre Gespräche beim Mittagessen bedanken. Der gesamten Arbeitsgruppe danke ich für die gute Atmosphäre und das sehr kollegiale Miteinander.

Für die Unterstützung im privaten Umfeld gilt mein besonderer Dank meiner Verlobten Sabine, die es (fast) immer geschafft hat, mich zu Hause von den Problemen der Arbeit abzulenken. Ein besonderer Dank geht auch an meine Eltern, die mich während des Studiums finanziell und anschließend immer wieder moralisch unterstützt haben. Meinen Freunden und Sportkameraden danke ich sehr für die schöne Zeit außerhalb des Arbeitsalltags, die mir die Kraft gegeben hat, auch Tiefs zu überwinden.

Inhaltsverzeichnis

ABSTRACT	IV
KURZZUSAMMENFASSUNG	V
ABKÜRZUNGSVERZEICHNIS	VII
1 EINLEITUNG	1
1.1 Zur Historie	1
1.2 Kinetische Modelle metabolischer Netzwerke	2
1.3 Modellierungsansätze: Bottom-up und Top-Down	3
1.4 Zielsetzung.....	5
2 THEORETISCHE GRUNDLAGEN	6
2.1 Modellierung von metabolischen Netzwerken	6
2.1.1 Grundlagen der Modellierung	6
2.1.2 Rekonstruktion metabolischer Netzwerke	8
2.1.3 Modellierung der dynamischen Prozesse	9
2.1.4 Annahme eines Fließgleichgewichts.....	10
2.2 Qualität und Quantität experimenteller Daten.....	11
2.2.1 Enzymparameter.....	11
2.2.2 Metabolomanalyse	12
2.2.3 Proteomanalyse	12
2.2.4 Transkriptomanalyse.....	13
2.2.5 Metabolische Stoffflussanalyse.....	13
2.2.6 Konsequenzen für diese Arbeit und weitergehende Annahmen	14
2.3 Stöchiometriebasierte Ansätze.....	16
2.3.1 Conserved Moieties	16
2.3.2 Elementarmoden.....	17
2.3.3 Flux Balance Analysis	18
2.4 Petrinetze	20
2.4.1 Stellen-Transitions-Netze.....	20

2.4.2	Farbige Petrinetze	21
2.5	Berechnung von Fließgleichgewichten.....	24
2.5.1	Lösung algebraischer Gleichungssysteme	24
2.5.2	Lösung von Differentialgleichungssystemen.....	25
2.6	Analysemethoden.....	28
2.6.1	Metabolische Kontrollanalyse	28
2.6.2	Stabilitätsanalyse	29
2.6.3	Biochemische Plausibilität.....	30
3	METHODIK	31
3.1	Reaktionsbilanzierung	31
3.2	Abschätzung der Gleichgewichtskonstanten.....	35
3.2.1	Theorie	35
3.2.2	Berechnung nach Mavrovouniotis.....	38
3.2.3	Berechnung nach Alberty.....	41
3.3	Simulationen ausgewählter Netzwerke	43
3.3.1	Modellierungsprinzipien	43
3.3.2	Modellierung von Conserved Moieties	45
3.3.3	Berechnung der metabolischen Kontrollkoeffizienten.....	46
3.3.4	Artifizielle Modelle	47
3.3.5	Modelle von Stoffwechselwegen.....	50
3.4	Alternative Methoden der Modellierung.....	54
3.4.1	Farbige Petrinetze	54
3.4.2	Lösung des algebraischen Gleichungssystems.....	55
3.5	Software und Datenbanken	56
3.5.1	Eigene Implementierungen	56
3.5.2	Eingesetzte Software	57
3.5.3	Datenbanken	60
4	ERGEBNISSE UND DISKUSSION	62
4.1	Reaktionsbilanzierung	62
4.1.1	Ergebnisse der Reaktionsbilanzierung	62
4.1.2	Diskussion zur Reaktionsbilanzierung	65

4.2	Gibbsenergieberechnung	67
4.2.1	Ergebnisse der Gibbsenergieberechnung	67
4.2.2	Diskussion zur Gibbsenergieberechnung	68
4.3	Simulation des linearen Modells	71
4.3.1	Ergebnisse des linearen Modells	71
4.3.2	Diskussion zum linearen Modell	74
4.4	Simulation der Modelle mit Verzweigung	77
4.4.1	Ergebnisse der Modelle mit Verzweigung	77
4.4.2	Diskussion zu den Modellen mit Verzweigung	78
4.5	Simulation der Glykolyse	79
4.5.1	Ergebnisse des Modells der Glykolyse	79
4.5.2	Diskussion zum Modell der Glykolyse	82
4.6	Simulation des Citratcyclus	84
4.6.1	Ergebnisse des Modells des Citratcyclus	84
4.6.2	Diskussion zum Modell des Citratcyclus	86
4.7	Simulation des Zentralstoffwechsels	88
4.7.1	Ergebnisse des Modells des Zentralstoffwechsels	88
4.7.2	Diskussion zum Modell des Zentralstoffwechsels	90
4.8	Alternative Methoden zum Erreichen eines Fließgleichgewichts	92
4.8.1	Farbige Petrinetze	92
4.8.2	Lösung des algebraischen Gleichungssystems	95
4.9	Zusammenfassende Diskussion	97
5	AUSBLICK	101
6	ANHANG	103
6.1	Inkremete zur Gibbsenergieberechnung nach Mavrovouniotis	103
6.2	Modell Glykolyse: Reaktionen und Parameter	104
6.3	Modell Citratcyclus: Reaktionen und Parameter	105
6.4	Modell Zentralmetabolismus: Reaktionen, Parameter, Biomasse	106
7	LITERATUR	109

Abstract

The simulation of a complete cell is one of the most ambitious aims of theoretical biochemical research. With the advent of Metabolomics, Proteomics and other “Omics” methods and the chance to get comprehensive and quantitative information about the state of the cell, this research field is booming. A new buzzword has been coined for the integration of practical and *in silico* experiments in this manner: Systems Biology.

There are many approaches to achieve this goal. This work deals with the kinetic simulation of metabolic networks, where reaching a steady state is one of the main aspects. Three methods have been evaluated for this: the integration of systems of differential equations, coloured Petri nets and solving algebraic equations. We were able to show that the integration of differential equations is the best method for large networks.

In the process of model-creation we made progress in supporting the network reconstruction process and in the incorporation of thermodynamical parameters into models. Regarding the simulation of models, we have focused on how to identify and deal with the borders of the modeled system and the influence of the borders on the simulation results. Models have been constructed and simulated for the glycolysis, the Krebs cycle and the central carbon metabolism of *Corynebacterium glutamicum* including all incoming and outgoing fluxes that are necessary for the growth of the organism.

Kurzzusammenfassung

Eines der ehrgeizigsten Ziele theoretischer biochemischer Forschung ist die Simulation einer kompletten Zelle. Mit der Aussicht, durch die Metabolom-, Proteom- und weitere systemische Analysen in naher Zukunft umfassende und quantitative Informationen über den Zustand der Zelle zu erhalten, hat dieser Forschungszweig deutlichen Auftrieb bekommen. Für die möglich werdende Integration von laborpraktischem und Computerexperiment auf diese Weise ist ein neues Wort geprägt worden: Systembiologie.

Es existieren viele Ansätze, um dieses Ziel zu erreichen. Diese Arbeit beschäftigt sich mit der kinetischen Simulation von metabolischen Netzwerken. Das Erreichen eines Fließgleichgewichtszustandes stand dabei im Vordergrund. Dazu wurde die Eignung verschiedener Methoden evaluiert: die Integration von Differentialgleichungen, die Simulation farbiger Petrinetze und die Lösung algebraischer Gleichungen. Es konnte gezeigt werden, dass die Integration von Differentialgleichungen die am besten geeignete Methode für große Netzwerke ist.

Bezogen auf die Modellformulierung konnten Erfolge bei der Unterstützung der Netzwerkrekonstruktion und bei der Einarbeitung von thermodynamischen Parametern erzielt werden. Bei der Simulation der Modelle wurde besonderer Wert auf die Behandlung und Identifikation der Systemgrenzen und ihren Einfluss auf die Simulationsergebnisse gelegt. Es wurden Modelle der Glykolyse, des Citratcyclus und des Zentralstoffwechsels von *Corynebacterium glutamicum* mit allen Zu- und Abflüssen, die für das Wachstum der Zelle relevant sind, simuliert.

Abkürzungsverzeichnis

Abkürzungen für Metabolite

13DPG	1,3-Diphosphoglycerat, 3-Phosphoglyceroylphosphat
2PG	2-Phosphoglycerat
3PG	3-Phosphoglycerat
3PGA	3-Phosphoglyceraldehyd
6PGLU	6-Phosphogluconat
ADP	Adenosindiphosphat
AMP	Adenosinmonophosphat
ATP	Adenosintriphosphat
CoA	Coenzym A
DHAP	Dihydroxyacetonphosphat, Glyceronphosphat
E4P	Erythrose-4-phosphat
F16DP	Fruktose-1,6-diphosphat
F6P	Fruktose-6-phosphat
G15L6P	Glucono-1,5-lacton-6-phosphat
G6P	Glukose-6-phosphat
GLC	Glukose
IDP	Inosindiphosphat
ITP	Inosintriphosphat
NAD	Nikotinamidadenindinukleotid
NADH	Nikotinamidadenindinukleotid (reduziert)
NADP	Nikotinamidadenindinukleotidphosphat
NADPH	Nikotinamidadenindinukleotidphosphat (reduziert)
P	Phosphat
PEP	Phosphoenolpyruvat
PP	Diphosphat
PRPP	Phosphoribosyl-diphosphat
R5P	Ribose-5-phosphat
RU5P	Ribulose-5-phosphat
XU5P	Xylulose-5-phosphat

Abkürzungen für Enzyme

fbpa	Fruktose-bisphosphat Aldolase
gapdh	Glyceraldehyd-3-phosphat Dehydrogenase
gk	Glucokinase
gpi	Glukose-6-phosphat Isomerase
pfk	6-Phosphofruktokinase
pgk	Phosphoglycerat Kinase
pgm	Phosphoglycerat Mutase
pk	Pyruvatkinase
pph	Phosphopyruvat Hydratase
tpi	Triosephosphat Isomerase

1 Einleitung

Die Simulation einer kompletten Zelle ist eines der visionären Ziele theoretischer, biochemischer Forschung. Seitdem die Aufklärung ganzer Genome in großem Maßstab durchgeführt wurde und wird, wir uns also in der post-genomischen Ära befinden, rückt dieses Ziel in den Mittelpunkt vieler wissenschaftlicher und wirtschaftlicher Bemühungen. Die Zellsimulation kann als zentrales Gebiet des neuen Fachgebietes Systembiologie verstanden werden. Ungeachtet einer exakten Definition des Begriffes Systembiologie, der erst seit wenigen Jahren in der biologischen Fachwelt verwendet wird, herrscht Einigkeit über zwei Aspekte, die integraler Bestandteil des neuen Gebietes sind: der „Omics“-Aspekt, also die umfassende Sammlung von experimentellen Daten eines Systems, und die mathematische Modellierung, um Vorhersagen zu treffen und diese zu überprüfen, um so das biologische System besser zu verstehen (Ventura *et al.* 2006). Die relevanten „Omics“-Techniken, wie z. B. Metabolomics, Proteomics und Transkriptomics, werden in Abschnitt 2.2 besprochen. Die vorliegende Arbeit beschäftigt sich im Kern mit dem zweiten Aspekt: der Modellierung.

Eine umfassende Beschreibung einer ganzen Zelle müsste sämtliche Interaktionen und ablaufenden Prozesse berücksichtigen, von denen viele unbekannt bzw. nicht quantifiziert sind. Man ist daher zum einen auf vereinfachende Beschreibungen angewiesen. Zum anderen muss man ein Subset der zellularen Prozesse auswählen, das sich beschreiben lässt und im Fokus des Interesses liegt. In der Vergangenheit wurden im Wesentlichen drei Arten von biochemischen Netzwerken modelliert: Signaltransduktionsnetzwerke (Wiley *et al.* 2003; Babu *et al.* 2006), genregulatorische Netzwerke (Ronen *et al.* 2002; Evangelisti und Wagner 2004) und metabolische Netzwerke (Borodina und Nielsen 2005). Wir wollen uns hier mit der kinetischen Beschreibung von metabolischen Netzwerken beschäftigen, also dem dynamischen Zusammenspiel von Metaboliten und Enzymen, die diese Metabolite umsetzen. Nähere Einzelheiten zu metabolischen Netzwerken, deren Rekonstruktion und Modellierung werden wir in Kapitel 2.1 besprechen.

1.1 Zur Historie

Die unseres Wissens nach erste Publikation, die sich konkret mit der Simulation biochemischer Prozesse auseinandersetzt, stammt aus dem Jahr 1943 (Chance 1943). B. Chance beschäftigte sich mit der Kinetik der Peroxidase und zeigte unter anderem zeitabhängige Konzentrationsverläufe im Rahmen der Michaelis-Theorie – noch während des zweiten Weltkrieges und in der Anfangszeit der automatischen Datenverarbeitung.

Einer seiner Mitarbeiter legte später, in den 1960er und 1970er Jahren, mit einer Reihe von Arbeiten den Grundstein für die Simulation größerer biochemischer Systeme: David

Garfinkel. In seinen Pionierarbeiten (Garfinkel *et al.* 1961; Garfinkel *et al.* 1962; Garfinkel *et al.* 1970) entwickelte er unter anderem Modelle für die Glykolyse und trieb die Entwicklung von Methoden und Computerprogrammen voran. Limitierend waren zu dieser Zeit vor allem die mangelnde Rechenleistung der Computer und das Fehlen effizienter Algorithmen. Zur Beseitigung des letztgenannten Schwachpunktes haben vor allem Gear *et al.* beigetragen (Shampine und Gear 1979).

Ein weiterer Meilenstein war die Entwicklung der Metabolic Control Analysis, die zeitgleich und unabhängig voneinander in (Kacser und Burns 1995) und (Heinrich und Rapoport 1974) eingeführt wurde (beim erstgenannten Artikel handelt es sich um eine Neuauflage des Artikels von 1973). Sie erlaubt die Quantifizierung der Kontrolle, die ein Enzym über Stoffflüsse und Konzentrationen in einem Netzwerk hat. Eine der wesentlichen Botschaften der Theorie ist, dass es im Allgemeinen keinen einzelnen geschwindigkeitsbestimmenden Schritt in einem metabolischen Pfad gibt. Eine kurze Einführung wird in Kapitel 2.6.1 gegeben.

Seit Beginn der 1990er Jahre ist die Modellierung metabolischer Netzwerke mit zwei Methoden populär, die auf der Stöchiometrie der Reaktionen basieren: die Flux Balance Analysis (siehe Abschnitt 2.3.3) und die Berechnung der Elementarmoden (siehe Abschnitt 2.3.2). Beide Methoden gehen von der Fließgleichgewichtsannahme aus. Im Fall der Flux Balance Analysis wird eine lineare Optimierung durchgeführt, wodurch man eine Flussverteilung erhält. Entscheidend ist vor allem die Wahl einer sinnvollen Zielfunktion, z. B. die Produktion von Biomasse (Burgard und Maranas 2003). Die Berechnung der Elementarmoden erlaubt eine objektive Bestimmung von metabolischen Pfaden. Bei Anwendung auf größere Netzwerke stößt man allerdings auf Laufzeitprobleme, so dass Heuristiken angewendet werden müssen (Klamt und Stelling 2002).

1.2 Kinetische Modelle metabolischer Netzwerke

Als geeignetes Ziel für Biosimulationen haben sich in der Vergangenheit Erythrocyten-Zellen herausgestellt (Joshi und Palsson 1989; Schuster *et al.* 1989; Mulquiney *et al.* 1999; Jamshidi *et al.* 2001). Es handelt sich um ein experimentell besonders gut untersuchtes System. Biochemische Daten über Enzyme und Metabolite werden seit Jahrzehnten produziert. Des Weiteren haben Erythrocyten keinen Zellkern und keine Gene, so dass keine Genexpression berücksichtigt werden muss (Nakayama *et al.* 2005).

Die Glykolyse von *Saccharomyces cerevisiae* (Bäckerhefe) wurde ebenfalls häufig modelliert (Galazzo und Bailey 1990; Cortassa und Aon 1997; Rizzi *et al.* 1997; Teusink *et al.* 2000). Auch hier ist die Datenlage deutlich günstiger als bei anderen Organismen. Der Grund für das hohe Interesse ist in der industriellen Anwendung der Hefe zu sehen, vor allem als Treibmittel beim Backen und als Alkoholproduzent. Die Optimierung der

Alkoholproduktion durch Metabolic Engineering, also durch gezielte genetische Veränderung oder Veränderung der Wachstumsbedingungen, kann von Computereperimenten mit mathematischen Modellen unterstützt werden.

Der bei weitem am besten untersuchte prokaryotische Organismus ist *Escherichia coli*. Die Daten für die Simulation werden strukturiert gesammelt (Sundararaj *et al.* 2004). Im Jahr 2002 wurde die International *E. coli* Alliance (IECA) gegründet (Holden 2002). Die IECA hat sich zum Ziel gesetzt, die Simulation des Bakteriums innerhalb eines Zeitraums von 10 Jahren zu erreichen. Trotz dieser Bemühungen existieren bislang nur wenige kinetische Modelle. Ein Modell bildet den Pfad der Threonin-Biosynthese (Chassagnole *et al.* 2001), ein anderes den Zentralstoffwechsel (Chassagnole *et al.* 2002) ab. Doch selbst für diese beiden Modelle haben die Autoren eigene kinetische Studien durchgeführt, um zu einem kohärenten Satz von Parametern zu gelangen. Es existiert zwar eine große Menge an Informationen zu *E. coli*, aber die Informationen sind lückenhaft und nicht standardisiert.

Neben den drei genannten Organismen gibt es noch vereinzelte kinetische Modelle anderer Organismen. Dazu zählen der Krebs-Zyklus von *Mycobacterium tuberculosis* (Singh und Ghosh 2006) sowie die Glykolyse in *Lactococcus lactis* (Hoefnagel *et al.* 2002) und *Trypanosoma brucei* (Helfert *et al.* 2001). Eine Datenbank, in die viele der genannten Modelle bereits aufgenommen sind, ist in (Olivier und Snoep 2004) beschrieben. Dort findet man zusätzlich eine Vielzahl von sehr kleinen Modellen.

1.3 Modellierungsansätze: Bottom-up und Top-Down

Die Modellierungsansätze können in zwei gegensätzliche Kategorien eingeteilt werden (siehe Abbildung 1): Bottom-up und Top-down (Katagiri 2003). Zur ersten Kategorie gehören im Kontext metabolischer Netzwerke rein mechanistisch orientierte Ansätze. Das Netzwerk wird zunächst in seine einzelnen Komponenten zerlegt. Diese Komponenten werden einzeln charakterisiert und anschließend wieder zum fertigen Modell zusammengesetzt. Konkret werden die Kinetiken der Enzyme gemessen und die erhaltenen Ratengleichungen assembliert. In einer Studie ist dieser Modellierungsansatz für die Glykolyse von *S. cerevisiae* konsequent angewendet worden (Teusink *et al.* 2000). Bei der Simulation stellte sich heraus, dass die Hälfte der Enzyme annähernd den korrekten Stofffluss aufwies. Bei der anderen Hälfte konnte die Abweichung zwischen *in vivo* Fluss und *in vitro* Rate durch den Unterschied zwischen *in vivo* und *in vitro* Kinetik erklärt werden.

Mit dem Einzug der „Omics“-Techniken (vergleiche Abschnitt 2.2) und den damit einhergehenden großen Datenmengen wird die Top-down-Modellierung möglich. Dabei betrachtet man das Gesamtsystem im Extremfall als „Black Box“ und verwendet nur die experimentellen Daten, in denen man versucht, mit statistischen Verfahren oder

maschinellen Lernverfahren Gesetzmäßigkeiten aufzuspüren. Diese Verfahren können nur so gute Ergebnisse liefern wie die Daten, auf denen sie trainiert sind. Bei einem neuen Datensatz, der beispielsweise unter anderen experimentellen Bedingungen gemessen wurde, kann das Modell schnell versagen.

In der Praxis wird in der Regel eine Kombination aus beiden Ansätzen praktiziert. Die eher wissensbasierten Bottom-up-Ansätze versuchen zunächst, die Reaktionskinetiken so genau wie möglich zu bestimmen. In einem zweiten Schritt werden die Parameter, die nicht bestimmt werden konnten, durch Anpassung an experimentelle Daten geschätzt. Bei einem eher datenbasierten Ansatz werden phänomenologische Ratengleichungen aufgestellt, die nicht notwendigerweise für einzelne Enzyme, sondern für zusammengefasste Stoffwechselwege gelten können. Die Parameter dieser Ratengleichungen werden durch Messung von Schlüsselmetaboliten, z. B. an Verzweigungspunkten im Metabolismus, oder durch das Einbringen von gezielten Störungen (Pulsen) ermittelt (Visser und Heijnen 2003).

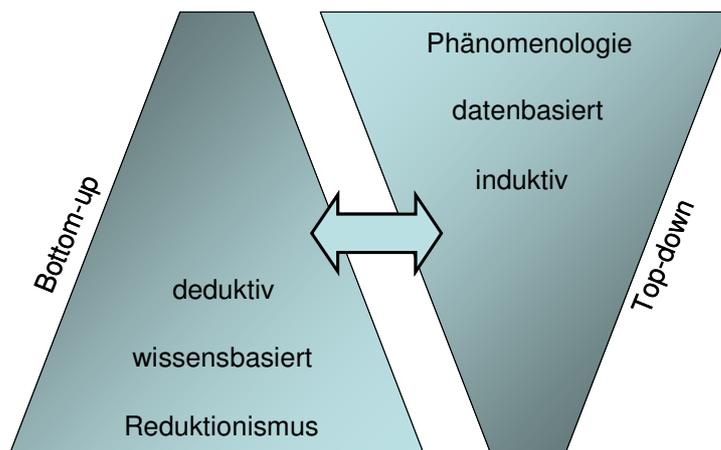


Abbildung 1 Bottom-up- und Top-down-Modellierung im Vergleich.

1.4 Zielsetzung

Das Ziel der vorliegenden Arbeit war die Modellierung von stationären Zuständen metabolischer Netzwerke. Als Modellorganismus wurde das *Corynebacterium glutamicum* (Eggeling und Bott 2005) verwendet, das auch Gegenstand experimenteller Untersuchungen in der Arbeitsgruppe ist.

Für das Bakterium mussten zunächst metabolische Netzwerke rekonstruiert werden. Vor allem die Methodenentwicklung und -anwendung zum automatischen Testen auf Konsistenz der Netzwerke in Bezug auf Stöchiometrie und Systemgrenzen stand dabei im Vordergrund. Eine wesentliche Teilaufgabe der Modellkonstruktion war der sinnvolle Einbau thermodynamischer Parameter, um die Reaktionsrichtung in Abhängigkeit von den Konzentrationen der Metabolite zu berechnen. Die Methoden zur Berechnung der thermodynamischen Parameter mussten dazu implementiert und evaluiert werden.

Im Mittelpunkt der Simulation und Analyse von Fließgleichgewichtszuständen stand die Frage, inwieweit unterschiedliche Modellbeschreibungen, die den gleichen beobachtbaren Zustand erreichen, in Bezug auf die Kontrolle dieses Zustandes äquivalent sind. Außerdem sollte untersucht werden, wie man bei Teilnetzwerken des Metabolismus von *C. glutamicum* überhaupt einen physiologischen Fließgleichgewichtszustand erreichen kann.

Neben der klassischen Methode zur Simulation metabolischer Netzwerke, der Integration von Differentialgleichungen, sollten auch zwei alternative Methoden zur Berechnung von Fließgleichgewichten getestet werden: die farbigen Petrinetze und die Lösung eines algebraischen Gleichungssystems.

Insgesamt sollten hier die Arbeiten geleistet werden, um nach umfassender Aufklärung des Metaboloms von *C. glutamicum* zu evaluieren, ob mit den derzeitigen Methoden und nach jetzigem Kenntnisstand die *in silico* Simulation und Vorhersage der Metabolitkonzentrationen möglich ist.

2 Theoretische Grundlagen

2.1 Modellierung von metabolischen Netzwerken

In diesem Abschnitt werden zunächst einige allgemeine Aussagen über den Modellierungsprozess getroffen und die Verwendung grundlegender Begriffe erläutert. Sodann wird auf das spezielle Problem der metabolischen Netzwerke eingegangen und die Möglichkeiten, das zeitabhängige Verhalten der Variablen in solchen Netzwerken zu beschreiben. Dem Fließgleichgewicht kommt bei der Beschreibung von metabolischen Netzwerken eine besondere Bedeutung zu. Wir werden deshalb im letzten Teil dieses Abschnitts auf die zugrunde liegenden Annahmen besonders eingehen.

2.1.1 Grundlagen der Modellierung

In der Wissenschaft findet das Konzept der Modellierung sehr häufig Anwendung, wenn es darum geht, reale Systeme zu beschreiben (Baumgarten 1996). Zunächst wollen wir die Begriffe System und Modell näher erläutern, wie sie hier verwendet werden. Der Begriff System wird hier für den zu modellierenden Ausschnitt aus der Wirklichkeit benutzt. Das System besteht aus Komponenten, die interagieren können, und kann von seiner Umgebung abgegrenzt werden. Beispiele für Systeme in ihrer Umgebung sind ein Molekül in Lösung oder ein Planet im Sonnensystem. Im Gegensatz dazu ist das Modell eine formale Spezifikation dieses Systems. Es *muss* auf einer Abstraktion beruhen, also eine Vereinfachung oder Idealisierung des Systems darstellen. Außerdem sollte es das Verhalten des Systems im Hinblick auf das Ziel der Analyse ausreichend beschreiben. Dabei muss eine Auswahl der wichtigen Einflussfaktoren (qualitativ) erfolgen und deren quantitativer Einfluss spezifiziert werden.

Nach der Formulierung des Modells ist der zweite Schritt die Untersuchung des Modells mit Analysemethoden und/oder Simulationsmethoden. Das Ergebnis der Untersuchung kann darauf geprüft werden, ob es eindeutig ist. Ist das nicht der Fall, muss gegebenenfalls das Modell neu formuliert werden oder der Ergebnisraum untersucht werden.

In einem dritten Schritt muss das Ergebnis der Untersuchung validiert werden. Zur Validierung stehen verschiedene Möglichkeiten zur Verfügung. Das Ergebnis kann mit dem Experiment verglichen werden. Es kann auf Konsistenz mit einer bestehenden Theorie geprüft werden. Als weitere Option können die Ergebnisse verschiedener Modelle miteinander verglichen werden. Von diesen Optionen ist die erste in der Regel die aussagekräftigste, allerdings nicht immer anwendbar.

Sollte die Validierung wiederholt Anlass zur Veränderung der Modellformulierung geben, spricht man von einer iterativen Verfeinerung des Modells. Dies ist sehr häufig der Fall, da das erste Modell selten eine ausreichende Beschreibung darstellt. Umgekehrt können Modelle auch Experimente provozieren. Wenn sich ein gewünschtes Verhalten zunächst durch Untersuchung des Modells erzeugen lässt, kann das Experiment anschließend zur Überprüfung des Resultats durchgeführt werden. Die Schritte im Modellierungsprozess sind in Abbildung 2 schematisiert.

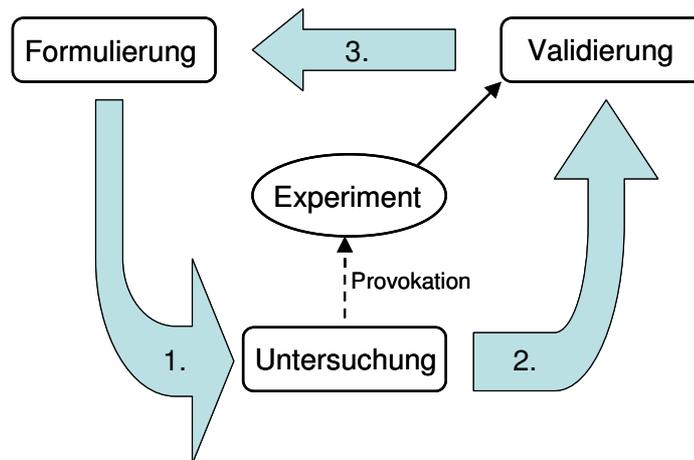


Abbildung 2 Die drei Schritte der Modellierung.

2.1.2 Rekonstruktion metabolischer Netzwerke

Karten von biochemischen Pfaden, wie sie auf den berühmten Postern bzw. im Buch von G. Michal (Michal 1999) illustriert sind, aber auch die Karten im World Wide Web, zu finden bei KEGG (Kanehisa *et al.* 2006) und MetaCyc (Caspi *et al.* 2006), vermitteln einen ersten Eindruck von der Komplexität des metabolischen Netzwerks als Gesamtheit. Die einzelnen Einheiten, die diesen Darstellungen zugrunde liegen, sind Umsetzungen von Metaboliten durch Enzyme, die wiederum untereinander durch die gemeinsamen Metaboliten und Effektoren verknüpft sind. Diese Karten sind allerdings nur oberflächliche Skizzen. Wenn man alle Wechselwirkungen in eine Karte einzeichnen würde und dabei die Vielzahl der noch unbekannteren Interaktionen außer Acht ließe, würde man schnell in einem Chaos von Farben und Verbindungslinien versinken. Doch auch diese Karte wäre für sich wenig aussagekräftig. Allein aus dem Wissen, welche Teile des Ganzen interagieren, können nur qualitative oder bestenfalls semi-quantitative Schlüsse gezogen werden. Für ein umfassendes Verständnis müssen die Relationen zwischen den Teilen quantitativ beschrieben und untersucht werden (Kacser 1986).

Widmen wir uns nun der Struktur eines metabolischen Netzwerks. Konstruiert wird die Struktur für den zu untersuchenden Organismus heute in der Regel aus den Informationen der Genomannotation (Francke *et al.* 2005). Die resultierende Netzwerkstruktur, die man aus Datenbanken wie KEGG oder MetaCyc erhält, muss im nächsten Rekonstruktionsschritt bezüglich des Ziels der Modellierung angepasst werden. Um das Problem zu vereinfachen, wird meist nur ein Teil der Reaktionen verwendet. Falls vorhanden, sollten hier Informationen aus der Proteom- bzw. Transkriptomanalyse einfließen (siehe auch Abschnitt 2.2). Der erhaltene Satz von Reaktionen wird noch um Reaktionen ergänzt, die entweder offensichtliche Lücken in der Genomannotation darstellen oder nicht enzymkatalysiert sind. Eine Anpassung der Struktur kann auch noch nach der Validierung des Modells notwendig werden (vergleiche Abbildung 2).

2.1.3 Modellierung der dynamischen Prozesse

Die Struktur eines metabolischen Netzwerks kann kompakt als stöchiometrische Matrix \underline{S} dargestellt werden. Die Zeilen entsprechen dabei den Metaboliten, die Spalten den Reaktionen. Die Matrixeinträge sind die stöchiometrischen Koeffizienten der Reaktion, wobei Eduktkoeffizienten ein negatives Vorzeichen erhalten (Abbildung 3). Bei externen Reaktionen, z. B. bei Transport durch eine Membran, wird der externe Metabolit häufig weggelassen. Die Matrixnotation hat den zusätzlichen Vorteil, dass das zeitabhängige Verhalten der Metabolitkonzentrationen \underline{c} geschlossen in einer Gleichung angegeben werden kann:

$$\frac{d\underline{c}}{dt} = \underline{S} \cdot \underline{v} \quad (\text{Gleichung 1})$$

Über diese Gleichung gelangen wir von der statischen Struktur zu einem Modell, das die dynamischen Prozesse des metabolischen Netzwerks beschreibt. Die Reaktionsgeschwindigkeiten \underline{v} werden durch Ratengleichungen beschrieben. In Abbildung 3 sind einige bekannte Ratengleichungen aufgeführt. Die einfachste Beschreibung der Reaktionsgeschwindigkeit einer reversiblen Reaktion erfolgt durch chemische Kinetik, die sich aus dem Massenwirkungsgesetz ergibt. Enzymkatalysierte Reaktionen folgen einer komplizierteren Kinetik.

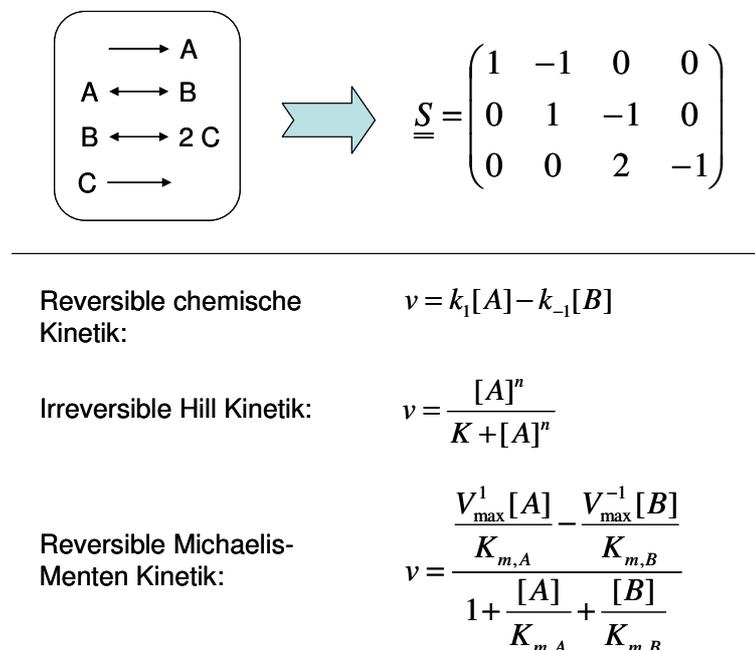


Abbildung 3 Oben: Konstruktion der stöchiometrischen Matrix aus einem Satz von Reaktionen. Unten: Beispiele für Reaktionskinetiken, die bei der Modellierung verwendet werden können.

2.1.4 Annahme eines Fließgleichgewichts

Eine Kerzenflamme ist ein offenes System, das unter ständiger Wärmeentwicklung durch die Verbrennung von Wachs seine Struktur aufrechterhält. Es findet Stoff- und Wärmeaustausch mit der Umgebung statt. Man kann sagen, dass sich die Kerzenflamme im Fließgleichgewicht befindet. Charakteristisch für das Fließgleichgewicht, wie wir es verstehen wollen, ist der konstante Zu- und Abfluss von Materie eines Systems, wobei sich keine Veränderung innerhalb des Systems beobachten lässt. In unserem Beispiel bleibt die Struktur der Kerzenflamme konstant.

Ein metabolisches Netzwerk befindet sich im Fließgleichgewicht, wenn Zu- und Abflüsse sich derart die Waage halten, dass keine Änderungen bei den Konzentrationsvariablen der Moleküle (Metabolite, Enzyme) stattfinden. Es unterscheidet sich damit vom chemischen Gleichgewicht, bei dem jede einzelne Reaktion den Gleichgewichtszustand erreicht hat und keine Zu- und Abflüsse existieren. Das chemische Gleichgewicht kann sich nur in einem geschlossenen System einstellen. Abgegrenzt werden muss in diesem Zusammenhang außerdem der Begriff Homöostase. Dabei handelt es sich um das Bestreben eines Systems, einen bestimmten Zustand auch gegenüber Störungen aufrecht zu erhalten. Im Kontext von metabolischen Netzwerken kann es bedeuten, dass die Konzentrationen der Metabolite durch regulative Prozesse in einem engen Rahmen gehalten werden, wenn sich äußere Einflussfaktoren wie pH-Wert, Temperatur etc. verändern.

Wir wollen nun der Frage nachgehen, unter welchen Bedingungen angenommen werden kann, dass sich das metabolische Netzwerk einer Zelle im Fließgleichgewicht befindet. Bei der Analyse von zeitabhängigen Systemen kann man drei Klassen von Prozessen unterscheiden (Heinrich und Schuster 1996). Die zentrale Klasse beinhaltet die Prozesse, die in der betrachteten Zeitskala ablaufen. Eine zweite Klasse beinhaltet Prozesse, die so langsam sind, dass sie vernachlässigt werden können. Die entsprechenden Variablen können als konstante Parameter behandelt werden. In der dritten Klasse befinden sich die Prozesse, die so schnell sind, dass sie in der betrachteten Zeitskala bereits abgelaufen sind bzw. keinen Einfluss auf das Verhalten haben. Beim metabolischen Netzwerk betrachten wir eine Zeitskala von Millisekunden bis Sekunden. Es kann davon ausgegangen werden, dass genregulatorische Prozesse zu langsam sind, um in dieser Zeitskala Wirkung zu zeigen. Sie gehören also zur zweiten Klasse. Zur dritten Klasse sind Diffusionsprozesse innerhalb eines Kompartiments zu zählen.

2.2 Qualität und Quantität experimenteller Daten

Für eine erfolgreiche Modellierung ist von entscheidender Bedeutung, welche Daten dem Modellierer zur Verfügung stehen. Die Qualität und Quantität der experimentellen Daten bestimmt nicht nur die Formulierung des Modells, sondern vor allem die mögliche Art der Validierung. In diesem Abschnitt wird sowohl die Datenlage im Allgemeinen als auch die Datenlage unter den speziellen Voraussetzungen am biochemischen Institut der Universität zu Köln beleuchtet. Der Fokus liegt auf den Arten von Daten, die für die Modellierung von metabolischen Netzwerken benötigt werden.

2.2.1 Enzymparameter

Die Enzymparameter zählen zu den wichtigsten Bausteinen der kinetischen Modellierung. Am Institut werden große Anstrengungen unternommen, um diese Parameter aus der Literatur zu extrahieren und in die Datenbank BRENDA einzupflegen (Abschnitt 3.5.3.1). Es sind vor allem von Bedeutung:

- die spezifische Aktivität oder die Wechselzahl des Enzyms, um bei Kenntnis der Enzymkonzentration die Grenzgeschwindigkeit V_{\max} zu berechnen,
- die Michaeliskonstanten (K_m -Werte), um die Affinität des Enzyms gegenüber den einzelnen Substraten zu beschreiben,
- die K_i -Werte, um die Affinität des Enzyms gegenüber Inhibitoren zu modellieren.

Die Anzahl der entsprechenden Einträge in der BRENDA-Datenbank sind in Tabelle 1 für alle Organismen, für den Modellorganismus *C. glutamicum* und zum Vergleich für *E. coli* zusammengestellt. Es ist zu beachten, dass die Zahl der verwertbaren Daten für ein Organismusmodell noch deutlich kleiner ist, weil keine Aussagen gemacht werden, ob es sich tatsächlich um den Strang des untersuchten Organismus bzw. um Mutanten handelt oder ob es sich um die Konstante zu einem Substrat oder Inhibitor handelt, der im Modell vorkommt.

Tabelle 1 Anzahl der Einträge für Enzymparameter in der Datenbank BRENDA. Stand: November 2006.

Einträge	Alle	<i>E. coli</i>	<i>C. glutamicum</i>
Enzyme mit spezifischer Aktivität	10933	467	17
Spezifische Aktivitäten	22380	1439	41
Enzyme mit Wechselzahl	2859	267	3
Wechselzahlen	18020	1825	4
Enzyme mit K_m -Wert	13153	601	14
K_m -Werte	68374	5250	52
Enzyme mit K_i -Wert	2646	189	5
K_i -Werte	13282	845	18

2.2.2 Metabolomanalyse

Das Ziel der Metabolomanalyse („Metabolomics“) ist, alle Metabolite in der Zelle und deren Konzentrationen zu einem definierten Zeitpunkt unter definierten Bedingungen zu bestimmen. Die hierfür am häufigsten verwendete Methode ist GC/MS, die Kopplung von Gaschromatographie an ein Massenspektrometer. Man vereint eine etablierte Methode zur Auftrennung von Substanzen mit einer sehr sensitiven Nachweis- und Quantifizierungsmethode.

Die Aufklärung des Metaboloms ist eine relativ junge Wissenschaft. Der Begriff Metabolom wurde erstmals in (Oliver *et al.* 1998) verwendet. Seitdem versuchen Forschungsgruppen weltweit, die Extraktionsmethodik für verschiedene Organismen zu optimieren und möglichst viele Metabolite zu identifizieren und quantifizieren.

Ein Problem, das für die Modellierung hochrelevant ist, ist die Tatsache, dass man aus dem Experiment keine absoluten, sondern nur relative Konzentrationswerte erhält. Es ist also möglich, die relative Konzentrationsänderung eines Metaboliten von Experiment zu Experiment zu verfolgen. Um die absoluten Konzentrationen zu ermitteln, muss für jeden Metaboliten eine Kalibrierungsgerade erzeugt werden. Dieser Vorgang ist aufwendig und setzt voraus, dass der jeweilige Metabolit zur Verfügung steht.

Am Institut sind bis jetzt (November 2006) 286 Metabolite identifiziert worden (Thielen 2006). Es fehlen allerdings Schlüsselmetabolite wie NADH, ATP und deren Derivate, die für die Modellierung von hoher Bedeutung sind. Als komplementäre Methode wird derzeit LC/MS getestet, bei der die Auftrennung der Substanzen in flüssiger Phase erfolgt. Außerdem wird gerade mit der Arbeit von (Strelkov 2004) fortgefahren, Kalibrierungsgeraden für Metabolite zu erzeugen. Inwieweit sich die Kalibrierungsgeraden von Experiment zu Experiment übertragen lassen, ist noch nicht endgültig geklärt.

2.2.3 Proteomanalyse

Analog zur Metabolomanalyse beschäftigt sich die Proteomanalyse („Proteomics“) mit der Gesamtheit aller Proteine in einer Zelle oder einem anderen biologischen System zu einem bestimmten Zeitpunkt (Kahn 1995; Müllner *et al.* 1998). Im Gegensatz zum Genom, der Gesamtheit aller Gene, ist das Proteom deutlich komplexer. Es gibt mehr Proteine als Gene, da durch alternatives Spleißen (nicht bei Bakterien) und posttranslationale Modifikationen aus einem Gen verschiedene Proteine entstehen können. Die Funktion eines Gens ist durch seine Sequenz festgelegt. Die Funktion der Proteine des Proteoms ist jedoch von der dreidimensionalen Struktur und den Wechselwirkungen von Proteinen untereinander abhängig.

Die Standardmethode zur Analyse des Proteoms ist die 2D-Gelelektrophorese. Dabei werden die Proteine in einer Richtung nach ihrer Ladung, und senkrecht dazu nach dem

Molekulargewicht getrennt. Die Proteine werden anschließend angefärbt. Die Proteinspots können ausgestochen werden, um das entsprechende Protein nach tryptischem Verdau massenspektrometrisch zu charakterisieren. Den umgekehrten Weg geht die Multidimensional Protein Identification Technology (MuDPIT). Hier werden die Proteine zuerst tryptisch verdaut, chromatographisch aufgetrennt und anschließend charakterisiert (Washburn *et al.* 2001). Beide Methoden eignen sich, um die vorhandenen Proteine zu charakterisieren, aber zurzeit nicht für die absolute Quantifizierung.

2.2.4 Transkriptomanalyse

Neben der Genomanalyse, die als Ausgangspunkt für die Rekonstruktion möglicher metabolischer Netzwerke dient, kann auch die Transkriptomanalyse weiterhelfen. Auf der Grundlage von Microarray-Experimenten können die Transkriptionsprodukte der Gene, die mRNA, Aufschluss über die tatsächlich exprimierten Proteine geben (Velculescu *et al.* 1997; Devaux *et al.* 2001). Der Nachteil gegenüber der Proteomanalyse ist, dass man Prozesse, die nach der Translation ablaufen, wie post-translationale Modifikationen, nicht erfasst. Ein großer Vorteil der Transkriptionsanalyse ist allerdings, dass sie sehr leicht genomweit durchgeführt werden kann.

2.2.5 Metabolische Stoffflussanalyse

Die metabolische Stoffflussanalyse (Metabolic Flux Analysis, MFA) wird seit ca. zehn Jahren hauptsächlich im Gebiet Metabolic Engineering eingesetzt (Wiechert 2001). Aus den gemessenen Stoffflüssen können mit Hilfe eines Modells Rückschlüsse gezogen werden, an welchen Stellen im Metabolismus eine Veränderung der Parameter, z. B. durch Genmutation, am ehesten zu einer gewünschten Optimierung der Stoffflüsse führt. Durch Bestimmung der Flüsse können aber auch weitere Informationen für die Modellformulierung und -validierung metabolischer Netzwerke gewonnen werden.

Die Stoffflussanalyse fußt auf mehreren Säulen. Die Kultur wird im Chemostat untersucht, d. h. das System befindet sich im Fließgleichgewicht. Aus dem Material, das dem Chemostat entnommen wird, können über Konzentrationsmessungen der ausgeschleusten Metabolite extrazelluläre Flüsse bestimmt werden. Aus diesen lässt sich aber noch nicht eindeutig auf die intrazellulären Flüsse schließen. Zusätzlich werden daher Bestandteile des Nährmedium ^{13}C -markiert, z. B. Glukose. Durch iterative Anpassung einer Simulation der Isotopenverteilung an die reale Isotopenverteilung in der Biomasse, die mit Massenspektrometrie oder Kernresonanzspektroskopie ermittelt wird, kann die intrazelluläre Flussverteilung berechnet werden (Marx *et al.* 1996; Sauer 2006).

Bislang wurde die Methode nur auf den Zentralstoffwechsel angewendet. Die Erweiterung auf größere Teile des Metabolismus sowie die Bestimmung von Stoffflüssen bei nicht-stationären Experimenten sind Gegenstand aktueller Forschung (Nöh und Wiechert 2006).

2.2.6 Konsequenzen für diese Arbeit und weitergehende Annahmen

Als Ziel der Modellierungen am biochemischen Institut der Universität zu Köln wurde der Modellorganismus *Corynebacterium glutamicum* verwendet. Welche Daten dafür zur Verfügung stehen und welche Konsequenzen sich daraus ergeben wird im Folgenden diskutiert. Wie man in Tabelle 1 (Seite 11) sieht, ist die Anzahl der Enzym-Parameter für *C. glutamicum* äußerst gering. Die Parameter sind aber für eine enzymkinetische Beschreibung der Reaktionen erforderlich. Zusätzlich muss der Reaktionsmechanismus aufgeklärt sein, damit die korrekte Ratengleichung angewendet werden kann. Um zu einer einfacheren Beschreibung zu gelangen, bietet es sich an, chemische Kinetik für die Reaktionen anzunehmen. Die Anzahl der Parameter wird so auf zwei pro Reaktion gesenkt. Durch die Abschätzung der Gleichgewichtskonstanten kann die Zahl sogar auf eins gesenkt werden.

Die experimentelle Abteilung der Arbeitsgruppe von Prof. Schomburg beschäftigt sich mit der Aufklärung des Metaboloms von *C. glutamicum*. Es wurden deutliche Fortschritte bei der Identifizierung der Metabolite gemacht. Von einer vollständigen Identifizierung aller Metabolite, die auch nur an einem Pfad des metabolischen Netzwerks beteiligt sind, ist man allerdings noch weit entfernt. Erste Erfolge beim Einsatz der LC/MS-Technik versprechen, die vorhandenen Lücken zu schließen. Für die Verwendung der Daten in einem kinetischen Modell ist als nächster Schritt die absolute Quantifizierung unbedingt erforderlich.

Als Konsequenz für diese Arbeit ergibt sich die Notwendigkeit, auf rein theoretischer Ebene mögliche Modellierungsmethoden und –prinzipien zu testen und so die Vorarbeit zu leisten, um bei veränderter Datenlage eine kinetische Modellierung zu ermöglichen.

Ungeachtet dieser Tatsache gibt es einige grundlegende Annahmen, die bei der Modellierung gemacht werden müssen, um das Problem zu vereinfachen bzw. überhaupt handhabbar zu machen. Sie sind im Folgenden aufgeführt (siehe auch Abschnitt 2.1.4):

- *Keine Kompartimentierung*: Durch die Wahl des Modellorganismus *C. glutamicum* muss man nur ein Kompartiment behandeln, in dem alle betrachteten internen Prozesse stattfinden.
- *Homogenität*: Die Metabolite sind im Kompartiment homogen verteilt. Es gibt keine Konzentrationsgradienten, weil die Diffusion um Größenordnungen schneller ist als die Reaktion.

- *Durchschnittszellen*: Die realen Zellen sind in Bezug auf Größe, Form, Zustand etc. identisch mit einer Durchschnittszelle. Dies ist nur näherungsweise der Fall. Jede Zelle durchläuft außerdem einen Zellzyklus, der nicht synchronisiert ist.
- *Keine Zell-Zell-Wechselwirkungen*: Die Zellen sind autonom und es gibt keinen Informationsfluss oder andere physikochemische Wechselwirkungen zwischen den Zellen. Man geht davon aus, dass sich die Zellen während der Wachstumsphase am wenigsten gegenseitig beeinflussen und nimmt daher bevorzugt diese Phase zur Modellierung und experimentellen Validierung.
- *Biochemische Systemkenntnis*: Alle biochemischen Faktoren, die das Verhalten des betrachteten Systems quantitativ beeinflussen können, sind bekannt. Dies ist im Allgemeinen nicht der Fall, so dass gegebenenfalls Parameter geschätzt werden müssen.
- *Keine stochastischen Effekte*: Die Anzahl der Moleküle in einer Zelle ist zum Teil so gering, dass stochastische Effekte nicht mehr auszuschließen sind. Betrachtet man nur den Zentralstoffwechsel, kann man davon ausgehen, dass die meisten Metabolite in hohen Konzentrationen vorliegen.

2.3 Stöchiometriebasierte Ansätze

Analysiert man die Stöchiometrie eines metabolischen Netzwerks, so können bereits daraus einige interessante Schlüsse gezogen werden. Wir werden in diesem Abschnitt den Blick auf drei Methoden werfen, die sich in der Systembiologie bewährt haben: Die Berechnung der Conserved Moieties, die Berechnung der Elementarmoden und die Flux Balance Analysis. Die erste und dritte Methode fanden in der vorliegenden Arbeit Anwendung, die zweite wird wegen ihrer wachsenden Bedeutung und zur Vollständigkeit dieses Kapitels besprochen.

2.3.1 Conserved Moieties

In praktisch jedem nicht-trivialen Modell eines metabolischen Netzwerks stößt man auf Erhaltungssätze, die Metabolite betreffen. In Abbildung 4 ist der Ausschnitt aus einem Netzwerk zu sehen. Nimmt man an, dass NAD und NADH in allen Reaktionen gemeinsam als Edukt bzw. Produkt vorkommen, so gilt für die Konzentrationen der beiden Cosubstrate folgender Zusammenhang:

$$NAD + NADH = \textit{konstant} \quad (\text{Gleichung 2})$$

Diese so genannten Conserved Moieties können auch aus mehreren Metaboliten bestehen. Sie haben zwei grundlegende Bedeutungen. Da Gleichung 2 für das Modell allgemein gültig ist, kann sie dazu verwendet werden, die Konzentration eines Metaboliten aus den Konzentrationen der anderen Metaboliten der Conserved Moiety zu berechnen. Die Anzahl der Variablen lässt sich so reduzieren. Die andere Bedeutung besteht in der biochemischen Interpretation. In einer Conserved Moiety gibt es immer eine chemische Substruktur, z. B. ein Atom oder eine Atomgruppe, die bei allen Prozessen unverändert bleibt. Es ist nicht möglich, diese Substruktur vom Rest des Netzwerks aus zu erreichen. NAD kann also in diesem Beispiel nicht aus anderen Komponenten, im Besonderen nicht aus den Bestandteilen des Nährmediums (Glukose, Ammonium, etc.) aufgebaut werden.

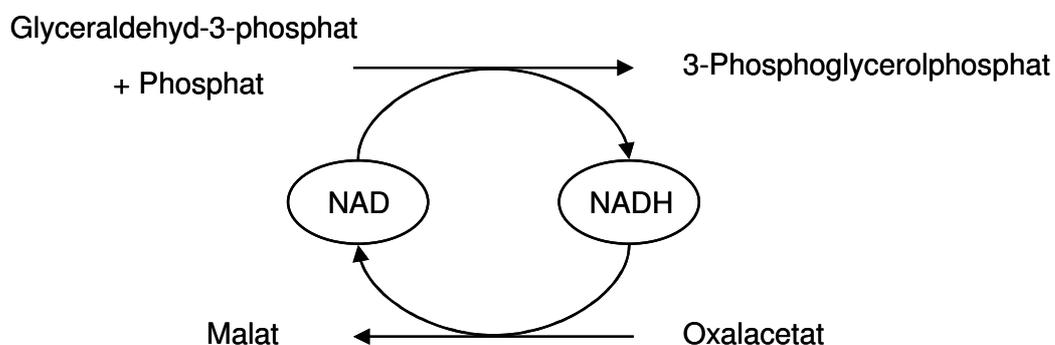


Abbildung 4 Die Conserved Moiety NAD+NADH.

Es bleibt die Frage, wie solche Erhaltungssätze identifiziert werden können. Hat man die Strukturen aller Metabolite und die Konnektivitätsmatrix jeder Reaktion, in der die Atome auf Edukt- und Produktseite einander zugeordnet sind, könnte man eine Erreichbarkeitsanalyse ausgehend von den Metaboliten des Nährmediums durchführen. Die nicht erreichbaren Atomgruppen gehören zu Conserved Moieties. Wiederholt man die Analyse für die übrigen Atome, kann man die einzelnen Conserved Moieties voneinander trennen.

Eine andere Möglichkeit geht von der stöchiometrischen Matrix aus und benötigt keine Atom-Atom-Zuordnung für jede Reaktion. Sind die Zeilen (Metabolite) der stöchiometrischen Matrix \underline{S} linear abhängig, gibt es für die folgende Gleichung nicht-triviale Lösungen für den Vektor \underline{x} :

$$\underline{S}^T \cdot \underline{x} = \underline{0} \quad (\text{Gleichung 3})$$

Die Anzahl der linear unabhängigen Vektoren \underline{x} , die diese Gleichung erfüllen, ist gleich der Anzahl der Zeilen von \underline{S} minus der Anzahl der linear unabhängigen Zeilen bzw. Spalten, also dem Rang von \underline{S} . Diese Vektoren \underline{x} entsprechen den Conserved Moieties, wenn sie nur positive und ganzzahlige Einträge besitzen. Es existieren Methoden, die alle Vektoren berechnen, die diese Voraussetzung erfüllen (Schuster und Höfer 1991; Heinrich und Schuster 1996). Vallabhajosyula *et al.* haben kürzlich festgestellt, dass bei Anwendung auf große Netzwerke bestimmte Algorithmen zur Berechnung verwendet werden müssen, z. B. die Householder Methode, um alle Conserved Moieties berechnen zu können (Vallabhajosyula *et al.* 2006).

2.3.2 Elementarmoden

Das Konzept der Elementarmoden wurde von S. Schuster und C. Hilgetag eingeführt (Schuster und Hilgetag 1994). Eine Elementarmode ist ein minimales Set von Enzymen, das im Fließgleichgewicht arbeiten kann. Die Enzyme werden mit dem relativen Fluss gewichtet, den sie im Fließgleichgewicht tragen müssen. Minimal bedeutet in diesem Zusammenhang, dass es – falls nur diese Enzyme aktiv wären – bei vollständiger Inhibition eines der Enzyme zum Erliegen aller Flüsse im System kommen würde. Die Elementarmoden sind eine alternative und objektive Definition von metabolischen Pfaden (Schuster *et al.* 2000).

Seit der Entwicklung dieses Konzepts hat es einige Veröffentlichungen gegeben, die sich mit der biochemischen Bedeutung der Elementarmoden beschäftigen. In (Stelling *et al.* 2002) wurde der Zusammenhang zwischen Redundanz in Bezug auf Elementarmoden und Robustheit gegenüber genetischen Deletionen des Zentralstoffwechsels von *E. coli* untersucht. In (Schuster und Kenanov 2005) wird die Bedeutung unterschiedlicher salvage

pathways in roten Blutkörperchen unter anderem für die Ausbeute von ATP herausgestellt. Einen guten Überblick über die Anwendungen von Elementarmoden und dem verwandten Konzept der Extreme Pathways geben (Papin *et al.* 2003).

Ein großes Problem bereitet die Berechnung aller Elementarmoden schon für relativ kleine Netzwerke von 50-100 Reaktionen, weil die kombinatorische Explosion zu sehr hohen Laufzeiten führt. Die Parallelisierung und die Entwicklung eines neuen Algorithmus (Urbanczik und Wagner 2005) konnten die Berechnung merklich beschleunigen. Das Programmpaket METATOOL (von Kamp und Schuster 2006) ist die Standardsoftware, um Elementarmoden zu berechnen.

2.3.3 Flux Balance Analysis

Die Flux Balance Analysis, die Anwendung des mathematischen Konzepts der linearen Optimierung auf metabolische Netzwerke im Fließgleichgewicht, wurde erstmals von J. M. Savinell und B. Palsson beschrieben (Savinell und Palsson 1992a, b). Sie ermöglicht die Berechnung von Flüssen \underline{v} im Fließgleichgewicht auf Basis der stöchiometrischen Matrix \underline{S} :

$$\underline{0} = \underline{S} \cdot \underline{v} \quad \text{(Gleichung 4)}$$

Dieses homogene, lineare Gleichungssystem ist bei metabolischen Netzwerken üblicherweise unterbestimmt. Es gibt also unendlich viele Flussverteilungen, die diese Gleichung erfüllen. Um den Raum der Flussverteilungen einzuschränken, werden zusätzliche Nebenbedingungen eingeführt. Reaktionen können in ihrer Richtung beschränkt werden und maximale und minimale Flüsse können pro Reaktion angegeben werden. Um eine diskrete Flussverteilung zu erhalten, führt man eine lineare Optimierung durch. Interessiert beispielsweise das maximal mögliche Zellwachstum auf Minimalmedium unter Glukoselimitierung, so ist die Zielfunktion der Optimierung eine gewichtete Summe der Biomassekomponenten bei begrenztem Glukosezufluss. Als Ergebnis erhält man einen Wert für die Zielfunktion und kann berechnen, wie viel Biomasse pro Glukosemolekül produziert wird. Außerdem erhält man im Idealfall eine einzige Flussverteilung. Sehr häufig ist dies leider nicht der Fall, so dass man Intervalle für die einzelnen Flüsse berechnen muss (Mahadevan und Schilling 2003).

Die Zielfunktion kann beliebig verändert werden, z. B. um die maximal mögliche Produktionsrate einer bestimmten Aminosäure zu berechnen. Weitere Konzepte wie Minimization Of Metabolic Adjustment (Segrè *et al.* 2002) oder Energy Balance Analysis (Beard *et al.* 2002) erweitern die Methodik. Die Veröffentlichungen zum Thema Flux Balance Analysis sind so zahlreich, dass an dieser Stelle auf die einschlägigen

Übersichtsartikel verwiesen werden muss (Edwards und Palsson 1998; Kauffman *et al.* 2003; Lee *et al.* 2006).

2.4 Petrinetze

Als alternatives Konzept zur Modellierung von biochemischen Pfaden sind Petrinetze seit den Arbeiten von Reddy et al. (Reddy *et al.* 1993; Reddy *et al.* 1996) in der Fachliteratur zu finden. Die theoretischen Grundlagen für Petrinetze wurden von Carl Adam Petri in seinem Artikel „Kommunikation mit Automaten“ im Jahr 1962 gelegt (Petri 1962). Die in seiner Arbeit vorgestellten Netze werden auch als Stellen-Transitions-Netze bezeichnet. Bis heute hat es eine Reihe von Erweiterungen der grundlegenden Theorie gegeben, von denen hier nur die farbigen Petrinetze (Colored Petri Net, CPN) vorgestellt werden (Jensen 1992). Eine wertvolle Sammlung von Informationen zu Petrinetz-Varianten und Programmpaketen findet sich bei der Petri Nets World im Internet unter <http://www.informatik.uni-hamburg.de/TGI/PetriNets/> (Stand: November 2006).

2.4.1 Stellen-Transitions-Netze

Zunächst werden die theoretischen Grundlagen in einer kurzen Einführung erläutert. Anhand eines Beispiels wird die Anwendung auf metabolische Netzwerke besprochen. Die Bedeutung von zwei Konzepten, den S- und T-Invarianten, für metabolische Netzwerke und ihre Analogien in der klassischen stöchiometriebasierten Analyse, nämlich Conserved Moieties und Elementarmoden, bilden den Abschluss. Für eine ausführliche Einleitung in die Thematik wird auf (Reisig 1990) verwiesen.

Dem Petrinetz liegt ein bipartiter Graph zugrunde. Die beiden Arten von Knoten des Graphen werden Stellen und Transitionen genannt. Stellen sind mit Transitionen durch gerichtete, gewichtete Kanten verknüpft. Auf den Stellen können sich anonyme Marken befinden. In Abbildung 5 ist die Phosphorylierung von Glukose durch ATP als Petrinetz dargestellt. Die Reaktion wird auf die Transition abgebildet, die üblicherweise als Rechteck dargestellt wird. Die einzelnen Metaboliten werden auf die Stellen abgebildet, die als Ovale oder Kreise gezeichnet werden. Mit einer oder mehreren Marken auf den Stellen können entweder Bedingungen oder konkrete Objekte repräsentiert werden. So kann das Vorhandensein oder die Menge eines Metaboliten modelliert werden.

Bei der Simulation des Petrinetzes werden die Transitionen (Reaktionen) einzeln betrachtet. Eine Transition heißt aktiviert, wenn mindestens so viele Marken auf den Eingangsstellen sind wie von dem entsprechenden Kantengewicht gefordert. Eine aktivierte Transition kann schalten bzw. feuern. Wenn sie feuert, werden die Marken entsprechend dem Kantengewicht von den Eingangsstellen entfernt und Marken entsprechend dem Kantengewicht auf den Ausgangsstellen erzeugt. Die Markierung vor und nach dem Feuern der Transition ist in Abbildung 5 zu sehen. Das Feuern einer Transition kann als Reaktionsereignis interpretiert werden.

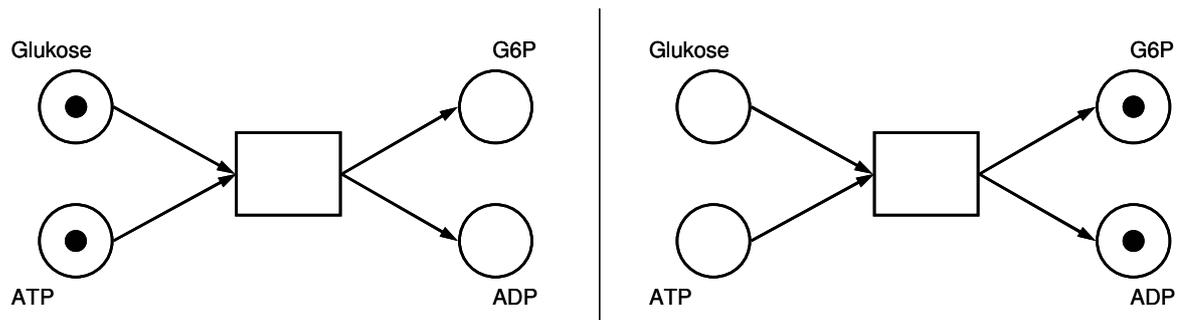


Abbildung 5 Stellen-Transitions-Netz der Glucokinase vor und nach dem Feuern der Transition.

Durch die abwechselnde Verkettung von Stellen und Transitionen werden komplexere Netze von metabolischen Pfaden bzw. Netzwerken aufgebaut. Man kann die sich ändernde Verteilung der Marken auf den Stellen während der Simulation beobachten. Kommt es zu einer Konfliktsituation, d. h. konkurrieren mehrere Transitionen um eine oder mehrere Marken einer bestimmten Stelle, wird zufällig entschieden, welche der Transitionen feuert.

Es gibt eine Reihe interessanter Eigenschaften von Stellen-Transitions-Netzen die hier aus Platzgründen nicht näher besprochen werden können, z. B. Erreichbarkeit und Lebendigkeit. Für die metabolischen Netzwerke besonders bedeutend sind die S- bzw. T-Invarianten. Sie stellen jeweils eine gewichtete Summe von ausgewählten Stellen bzw. Transitionen dar, die konstant (invariant) ist. Bei den S-Invarianten werden die Marken auf den Stellen mit dem Gewicht multipliziert und addiert. Ist die Summe während der Simulation konstant (genauer: bei allen Markierungen des Netzes), handelt es sich um eine S-Invariante. Bei T-Invarianten schalten die Transitionen nacheinander und so häufig, wie es dem Gewicht entspricht. Gelangt man am Ende zur gleichen Situation (Markierung) wie zu Beginn, handelt es sich um eine T-Invariante. Ein Vergleich mit den Abschnitten 2.3.1 und 2.3.2 lässt deutlich werden, dass es sich bei den S-Invarianten um Conserved Moieties, bei den T-Invarianten um Elementarmoden handelt. Der Vergleich wurde in (Zevedei-Oancea und Schuster 2003) beschrieben. Im Detail gibt es einige Unterschiede zwischen den Konzepten, die kürzlich von A. Sackmann *et al.* aufgearbeitet worden sind (Sackmann *et al.* 2006).

2.4.2 Farbige Petrinetze

Die Grundsätze der Petrinetz-Theorie wurden im vorangegangenen Kapitel besprochen. Im Folgenden werden die Erweiterungen der farbigen Petrinetze (Jensen 1992) besprochen.

Im Gegensatz zu Stellen-Transitions-Netzen sind die Marken bei farbigen Petrinetzen nicht anonym. Jeder Marke wird ein Datentyp zugeordnet. Als Datentypen kommen zum Beispiel ganze Zahlen (Integer), Fließkommazahlen (Float) oder selbst definierte

Datentypen in Frage. Der Informationsgehalt der Marken wird auf diese Weise erhöht. Dadurch gelingen bei vielen Netzen eine kompaktere Darstellung und häufig eine effizientere Simulation. Im Prinzip kann jedoch jedes farbige Petrinetz in ein Stellen-Transitions-Netz übersetzt werden.

Einhergehend mit den unterscheidbaren Marken müssen auch die Kantengewichte bzw. Kantenbeschriftungen zwischen den Datentypen der Marken unterscheiden können. Die Beschriftung der Eingangskanten spezifiziert, wie viele Marken von welchen Datentypen für die Aktivierung der Transition benötigt werden. Feuert eine Transition, wird ein Codesegment ausgeführt, das mit der Transition verbunden ist. Aus den Werten, die die Eingangsmarken tragen, und der Anzahl der Eingangsmarken werden die Werte für die Ausgangsmarken und deren Anzahl berechnet. Auf den Ausgangsstellen werden entsprechend der Kantenbeschriftung neue Marken erzeugt.

Genrich et al. haben farbige Petrinetze für die Simulation eines metabolischen Pfades verwendet (Genrich *et al.* 2001). Die Zuordnung von Metaboliten zu Stellen und Reaktionen zu Transitionen bleibt dabei erhalten. Auf jeder Stelle befindet sich eine Marke vom Typ Fließkommazahl, die die Konzentration des Metaboliten repräsentiert. Die Stellen sind mit den Transitionen jeweils mit einer Eingangs- und einer Ausgangskante verknüpft, so dass die Marke nach dem Feuern der Transition wieder auf die gleiche Stelle, aber im Regelfall mit verändertem Konzentrationswert, zurückgelegt wird. Im Codesegment der Transition werden die Umsatzrate und die neuen Konzentrationen berechnet. Enzyme und Inhibitoren können auf gleiche Weise modelliert werden (Abbildung 6). Ein Modell der Glykolyse wurde nach diesem Prinzip von Genrich *et al.* simuliert.

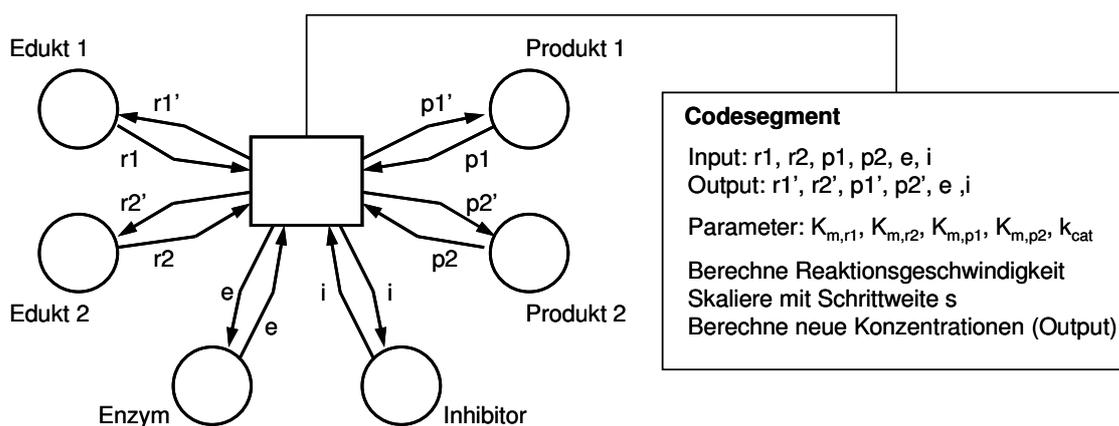


Abbildung 6 Farbiges Petrinetz einer Reaktion mit zwei Edukten, zwei Produkten und einem Inhibitor.

Die hybriden Netzwerke (Nagasaki *et al.* 2004), die sowohl diskrete und anonyme als auch farbige Marken verwenden, können als Untergruppe der farbigen Petrinetze aufgefasst werden. Diskrete Marken können im Formalismus der farbigen Petrinetze leicht durch Wahl eines geeigneten Datentyps repräsentiert werden.

2.5 Berechnung von Fließgleichgewichten

Man kann sich verschiedener Verfahren zur Berechnung des Fließgleichgewichts eines metabolischen Netzwerks bedienen. Ein Verfahren, das bislang im biochemischen Kontext eher selten verwendet wurde, ist die Simulation von farbigen Petrinetzen, die im vorangegangenen Kapitel besprochen wurde. Hier sollen zwei weitere Methoden vorgestellt werden, die weit häufiger zum Einsatz kommen: die Lösung nicht-linearer algebraischer Gleichungen und die Integration von Differentialgleichungen. Ausgangspunkt für beide Methoden ist wiederum das System von Differentialgleichungen, das ein metabolisches Netzwerk beschreibt (vergleiche Abschnitt 2.1.3):

$$\frac{dc}{dt} = \underline{S} \cdot \underline{v} \quad (\text{Gleichung 5})$$

2.5.1 Lösung algebraischer Gleichungssysteme

Geht man von der Bedingung für das Fließgleichgewicht aus, wird aus dem System von Differentialgleichungen ein System von algebraischen Gleichungen:

$$\underline{0} = \underline{S} \cdot \underline{v} \quad (\text{Gleichung 6})$$

Eine Lösung kann mit Verfahren der nicht-linearen Optimierung gefunden werden. Im Gegensatz zur Flux Balance Analysis, die auf derselben Gleichung beruht, ist das Problem nicht-linear, weil die Flüsse \underline{v} durch Ratengleichungen ersetzt werden. Wir suchen also nicht nach Flüssen, sondern nach Konzentrationen, die bestimmte Optimalitätskriterien erfüllen.

Für die nicht-lineare Optimierung stehen etablierte Algorithmen zur Verfügung. Häufig ist man mit dem Problem konfrontiert, dass die Algorithmen nur ein lokales Minimum finden. Für die Suche nach dem globalen Minimum werden diese Algorithmen meist mit anderen Algorithmen kombiniert, z. B. das Conjugate Gradient Verfahren (lokal) mit dem Simulated Annealing (global).

Für unsere Zwecke ist es zunächst ausreichend, *eine* Lösung für das Gleichungssystem zu finden. Die Lösung sollte allerdings bestimmte Nebenbedingungen erfüllen. Alle Metabolit-Konzentrationen müssen in einem physiologischen Bereich sein. Diesen Nebenbedingungen kann zusätzlich bei der Stellung des Optimierungsproblems Rechnung getragen werden, indem obere und untere Grenzen für die Konzentrationen angegeben werden.

2.5.2 Lösung von Differentialgleichungssystemen

Das Differentialgleichungssystem (Gleichung 5) beschreibt das zeitabhängige Verhalten des metabolischen Netzwerks. Durch Integration des Gleichungssystems kann man einen stationären Zustand erreichen. Nur in sehr einfachen Spezialfällen ist eine Differentialgleichung analytisch lösbar. In der Regel muss auf numerische Integrationsverfahren zurückgegriffen werden. Wir werden in diesem Abschnitt auf die Probleme numerischer Verfahren eingehen und nach einer kurzen Einführung ein Verfahren besprechen, das sich bei der Beschreibung von biochemischen Reaktionsmodellen in der Vergangenheit bewährt hat: das Prädiktor-Korrektor-Verfahren (Shampine und Gear 1979; Hindmarsh *et al.* 2005).

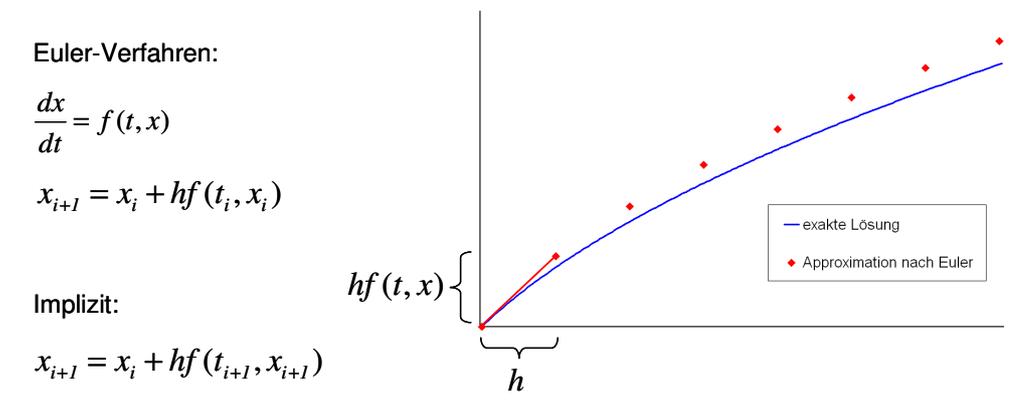
Bei der Lösung von Differentialgleichungen handelt es sich im Allgemeinen um ein Anfangswertproblem. Zusätzlich zu den Differentialgleichungen benötigt man geeignete Startwerte für die Variablen (Konzentrationen), um das Problem vollständig zu beschreiben. Ausgehend von den Startwerten wird das Gleichungssystem iterativ gelöst. Das einfachste Lösungsverfahren ist die Euler-Methode, die zur Darstellung der Prinzipien im Folgenden kurz erläutert wird.

Wir betrachten ein System mit einer Variable x und einer Differentialgleichung (Abbildung 7). Es wird eine Schrittweite h und ein Startwert x_0 an der Stelle t_0 festgelegt. Ausgehend vom Startwert wird iterativ der neue Funktionswert x_{i+1} berechnet, indem ein Schritt entlang der Tangenten gemacht wird, deren Steigung durch die Differentialgleichung gegeben ist. Die tatsächliche Funktion wird so diskretisiert und schrittweise linear angenähert. Je kleiner man die Schrittweite h wählt, desto genauer ist die Annäherung, desto höher ist aber auch der Rechenaufwand. Im Regelfall wird die Schrittweite variabel angepasst. Dies setzt allerdings eine Möglichkeit voraus, den Fehler in jedem Schritt abschätzen zu können, was angesichts der Unkenntnis des realen Wertes nicht exakt möglich ist. Anhand einer spezifizierten Obergrenze für den Fehler wird jeweils die aktuelle Schrittweite berechnet.

Das Euler-Verfahren ist weder effizient noch genau. Verbesserungen lassen sich dadurch erreichen, dass die Differentialgleichung nicht nur am aktuellen Punkt, sondern zusätzlich an mehreren Stellen, z. B. bei $t_i + \frac{1}{2}h$ ausgewertet wird. Zur Berechnung des nächsten Funktionswerts werden die berechneten Werte gewichtet aufsummiert. Diese Verfahren heißen Runge-Kutta-Verfahren. Bei einer anderen Art der Verbesserung beruht die Berechnung des nächsten Funktionswerts nicht nur auf dem aktuellen Wert, sondern auch auf zuvor berechneten Werten. Diese Verfahren bezeichnet man als lineare Mehrschrittverfahren. Die Anzahl der Auswertungen (Runge-Kutta) bzw. die Anzahl einbezogenen Funktionswerte (Mehrschrittverfahren) bestimmt die Ordnung des

Verfahrens. Sowohl die Ordnung also auch die Schrittweite werden bei den meisten Implementierungen automatisch angepasst.

Für das weitere Verständnis ist der Unterschied zwischen impliziten und expliziten Verfahren von Bedeutung. In Abbildung 7 sind die Iterationsformeln für das explizite und implizite Euler-Verfahren zu sehen. Wie man erkennen kann, muss im zweiten Fall die Steigung an der Stelle des neuen Funktionswerts berechnet werden.



Runge-Kutta-Verfahren:

$$x_{i+1} = x_i + h \sum_{j=1}^n b_j k_j \quad \text{mit} \quad \begin{array}{l} b_j: \text{Gewichte} \\ k_j: \text{Ableitungen} \end{array}$$

Lineare Mehrschritt-Verfahren:

$$\sum_{j=0}^{m_1} a_{m_1-j} x_{i+1-j} = h \sum_{j=0}^{m_2} b_{m_2-j} f_{i+1-j} \quad \text{mit} \quad \begin{array}{l} a_p, b_j: \text{Gewichte} \\ f_j: \text{Ableitungen} \end{array}$$

Abbildung 7 Verfahren zur Integration von Differentialgleichungssystemen.

Steifheit

Ein Phänomen, mit dem man bei der kinetischen Modellierung von nicht-trivialen metabolischen Netzwerken mit Differentialgleichungen fast immer zu tun hat, ist Steifheit (Shampine und Gear 1979). Sie äußert sich in Instabilität und/oder Ineffizienz wegen zu geringer Schrittweiten. Die Lösung kann numerisch instabil werden, wenn sich z. B. Rundungsfehler, die in jedem Schritt gemacht werden, aufsummieren. Die Ursache des Phänomens liegt in stark unterschiedlichen Zeitskalen bzw. Geschwindigkeiten. In unserem konkreten Fall sind es die Reaktionsgeschwindigkeiten, die bei enzymkatalysierten Reaktionen stark unterschiedlich sein können. Ob und wie steif ein Problem ist, hängt von den Anfangsbedingungen und den Differentialgleichungen ab.

Wie bereits erwähnt wurde, haben sich Prädiktor-Korrektor-Verfahren, die auf den linearen Mehrschrittverfahren aufbauen, bei der Lösung von reaktionskinetischen Problemen in der Vergangenheit bewährt. Das grundsätzliche Vorgehen soll hier kurz skizziert werden. Dafür wird das Mehrschrittverfahren mit $m_2=0$ verwendet (siehe Abbildung 7), das auch als Backward Differentiation Formula (BDF) bezeichnet wird.

Der Funktionswert für den nächsten Schritt wird zunächst mit einem Prädiktor, der auf einer expliziten BDF beruht, berechnet. Dieser Funktionswert dient als Startwert für den Korrektor, eine implizite BDF, die so lange Iterationen durchführt, bis ein bestimmtes Fehlerkriterium erfüllt ist. Erst dann wird der Schritt akzeptiert und der nächste Schritt ausgeführt. Die effiziente Berechnung des linearen Gleichungssystems im Korrektorschritt ist die Grundvoraussetzung für eine schnelle Lösung des Problems. Die Differenz zwischen den Funktionswerten des Prädiktors und Korrektors kann in jedem Schritt als Abschätzung für den Fehler verwendet werden. Anhand dieser Fehlerabschätzung kann die Schrittweite, falls erforderlich, verändert werden.

2.6 Analysemethoden

2.6.1 Metabolische Kontrollanalyse

Die metabolische Kontrollanalyse (Metabolic Control Analysis, MCA) berechnet, wie die Kontrolle über Stoffflüsse und Konzentrationen im Fließgleichgewicht auf die Enzyme eines metabolischen Netzwerks verteilt sind. Die Kontrolle wird quantifiziert, indem ein Enzymparameter, z. B. die Konzentration des Enzyms, gestört wird. Das System geht in einen neuen Fließgleichgewichtszustand über. Daraufhin wird gemessen, wie stark die Auswirkung dieser Störung auf die Stoffflüsse und Konzentrationen ist. Die metabolische Kontrollanalyse wurde von Kacser & Burns und Heinrich & Rapoport eingeführt (vergleiche Abschnitt 1.1). Wegen der fundamentalen Bedeutung für die Deutung von Fließgleichgewichtszuständen wird im Folgenden eine kurze Einleitung in die Theorie gegeben. In Abschnitt 3.3.3 wird besprochen, wie die Kontrollkoeffizienten in dieser Arbeit verwendet werden.

Die metabolische Kontrollanalyse ist eine phänomenologische Analysetechnik. Die grundlegenden Aussagen gelten unabhängig vom Mechanismus der Enzyme bzw. der Ratengleichung. Die Enzymkinetik wird also als „Black Box“ betrachtet. Für jede Reaktion im metabolischen Netzwerk kann ein Kontrollkoeffizient bezüglich des Stoffflusses durch jede Reaktion (Gleichung 7) und bezüglich der Konzentration jedes Metaboliten des Netzwerks (Gleichung 8) angegeben werden:

$$C_i^r = \frac{\partial J_r / J_r}{\partial E_i / E_i} \quad \text{Flusskontrollkoeffizient} \quad (\text{Gleichung 7})$$

$$C_i^k = \frac{\partial S_k / S_k}{\partial E_i / E_i} \quad \text{Konzentrationskontrollkoeffizient} \quad (\text{Gleichung 8})$$

Hierbei steht C für einen Kontrollkoeffizienten, J_r für den Fluss durch die Reaktion r , S_k für die Konzentration des Metaboliten k und E_i für die Enzymkonzentration des Enzyms i .

Die Kontrollkoeffizienten gelten für das Netzwerk als Ganzes, sind also eine systemische Eigenschaft. Für sie existieren zwei grundlegende Summations-Theoreme:

$$\sum_{i=1}^n C_i^r = 1 \quad \text{Fluss-Summations-Theorem} \quad (\text{Gleichung 9})$$

$$\sum_{i=1}^n C_i^k = 0 \quad \text{Konzentrations-Summations-Theorem} \quad (\text{Gleichung 10})$$

Gleichung 9 bedeutet, dass die Summe der Flusskontrollkoeffizienten aller Reaktionsschritte bezüglich eines Flusses immer eins ergeben muss. Die Summe der Konzentrationskontrollkoeffizienten aller Reaktionsschritt bezüglich der Konzentration eines betrachteten Metaboliten muss immer gleich null sein (Gleichung 10). Diese Eigenschaften der Kontrollkoeffizienten machen die globale Bedeutung und Interkonnektivität deutlich.

2.6.2 Stabilitätsanalyse

Hat man einen Fließgleichgewichtszustand gefunden, kann man sich die Frage stellen, ob dieser Zustand stabil ist. Für die Stabilität von stationären Zuständen gibt es definierte Kriterien, die auf einer Analyse der Jacobi-Matrix des Differentialgleichungssystems beruhen. Die Jacobi-Matrix enthält die partiellen Ableitungen der rechten Seite der Differentialgleichungen nach allen Variablen (Konzentrationen). Aus den Eigenwerten der Matrix, der Spur und der Determinante kann auf die Stabilität geschlossen werden (Abbildung 8). Für eine Herleitung der Stabilitätskriterien sei hier auf (Heinrich *et al.* 1977) verwiesen.

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{pmatrix}$$

f_1, f_2, f_3 : Rechte Seite der Differentialgleichungen
 x_1, x_2, x_3 : Konzentrationen
 Spur: $\text{tr} = \Sigma$ ■
 Determinante: $\Delta = \det(J)$
 Eigenwerte λ_i

Stabilitätsbedingungen: $\text{tr} < 0$, $\Delta > 0$, $\text{Realteil}(\lambda_i) < 0$

Abbildung 8 Bedingungen für die Stabilität eines stationären Zustands für ein System aus drei Differentialgleichungen und drei Konzentrationsvariablen.

2.6.3 Biochemische Plausibilität

Es ist nicht immer sinnvoll oder methodisch möglich, das gesamte vorhandene biochemische Wissen a priori in die Modellformulierung einzubringen. In diesen Fällen bietet es sich an, die Resultate mit dem nicht eingebrachten Wissen a posteriori auf Plausibilität zu überprüfen. Für die meisten Variablen können biochemisch plausible Ober- bzw. Untergrenzen angegeben werden.

Die Konzentrationen von Metaboliten bewegen sich intrazellulär in der Regel unterhalb von 10 mmol/l. Für die Anzahl der Moleküle in einer *Escherichia coli* Zelle findet man für ein Teilvolumen von 10^{-18} Liter in (Goodwell 1993) folgende ungefähre Angaben: 450 Proteine, 30 Ribosomen, 340 tRNA-Moleküle, 30000 kleine organischen Moleküle und 50000 Ionen. Übertragen auf das durchschnittliche Volumen einer Zelle des verwendeten Modellorganismus *C. glutamicum* (ca. 10^{-16} Liter) würde man in etwa die 100-fache Menge erwarten. Da Modelle meist nur von Teilen der Zelle angefertigt werden, muss die Zahl der Moleküle im Modell deutlich geringer sein.

Trifft man die konservativen Annahmen, dass 80 % der Proteine Enzyme sind und keine einzelne Enzymklasse mehr als ein Zehntel der Gesamtenzymzahl stellt, kommt man auf eine Obergrenze von 3000 Enzymen pro Enzymklasse. Ist die Enzymkonzentration nicht bekannt, aber die Wechselzahl eines Enzyms, kann daraus auf den maximalen Stofffluss geschlossen werden. So ergibt sich für eine Wechselzahl von 10 Molekülen pro Sekunde bei einem Zellvolumen von 10^{-16} Liter ein maximal möglicher Fluss von 0,5 mM/s.

3 Methodik

3.1 Reaktionsbilanzierung

Hinter den Reaktionsgleichungen, die zur Beschreibung von enzymkatalysierten Reaktionen üblicherweise verwendet werden, verbirgt sich mehr, als man auf den ersten Blick vermuten würde. Ein bedeutender Anteil der Reaktionen, die man in der Datenbanken KEGG (Kanehisa *et al.* 2006) findet, weist keine korrekte Massenbilanz auf. Neben dem einfachen Fall, dass die Reaktion tatsächlich falsch in die Datenbank eingepflegt wurde, ist vor allem ein Effekt verantwortlich: es handelt sich bei den jeweiligen Substraten nicht um eine, sondern um mehrere Spezies.

Als Beispiel betrachten wir zunächst die Reaktion von Pyruvat zu Phosphoenolpyruvat, die durch das Enzym Pyruvatkinase katalysiert wird (Abbildung 9). Durch die neutrale Schreibweise der Metabolite wird die Tatsache, dass es sich um ein Substanzgemisch handelt, verschleiert. Da die wirkliche Zusammensetzung der intrazellulären Lösung von einer Reihe von variablen Faktoren abhängig ist, wie z. B. Temperatur, pH-Wert und Ionenstärke, ist eine solche Vereinfachung zulässig und beabsichtigt. Dies gilt insbesondere für Datenbanken, die nicht auf Reaktionen spezialisiert sind. Bei der Verwendung der Reaktionsgleichungen zur Generierung eines Modells ist man allerdings auf ihre korrekte Bilanzierung angewiesen. Für den automatischen Bilanzierungsprozess bietet es sich an, die jeweils vorherrschenden Spezies beim gewählten Standardzustand zu verwenden.

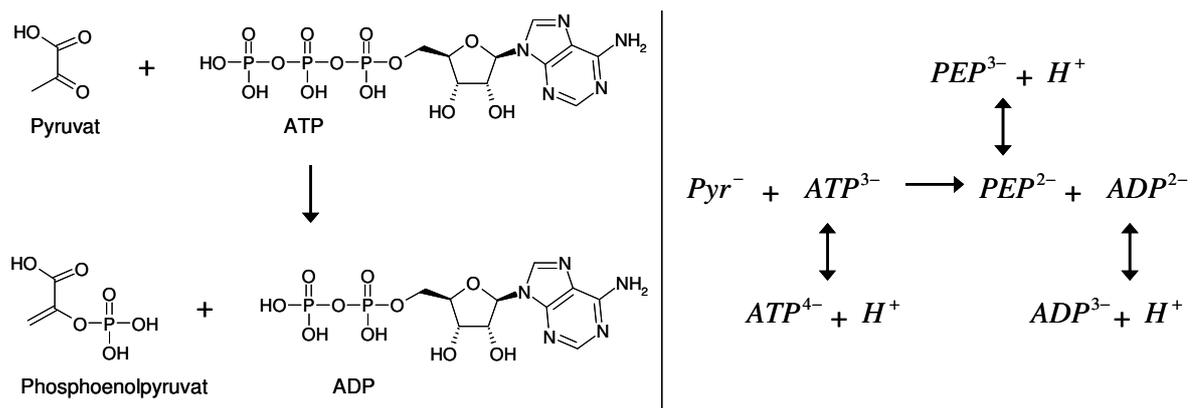


Abbildung 9 Links: Die Edukte und Produkte enzymkatalysierter Reaktionen werden in biochemischen Fachbüchern und Datenbanken häufig als neutrale Spezies dargestellt. Rechts: Tatsächlich liegt immer ein Gemisch verschiedener Spezies vor. Zur Reaktionsbilanzierung sollte man sich auf eine konkrete Reaktion konzentrieren, z. B. mit den vorherrschenden Spezies bei gegebenem Standardzustand.

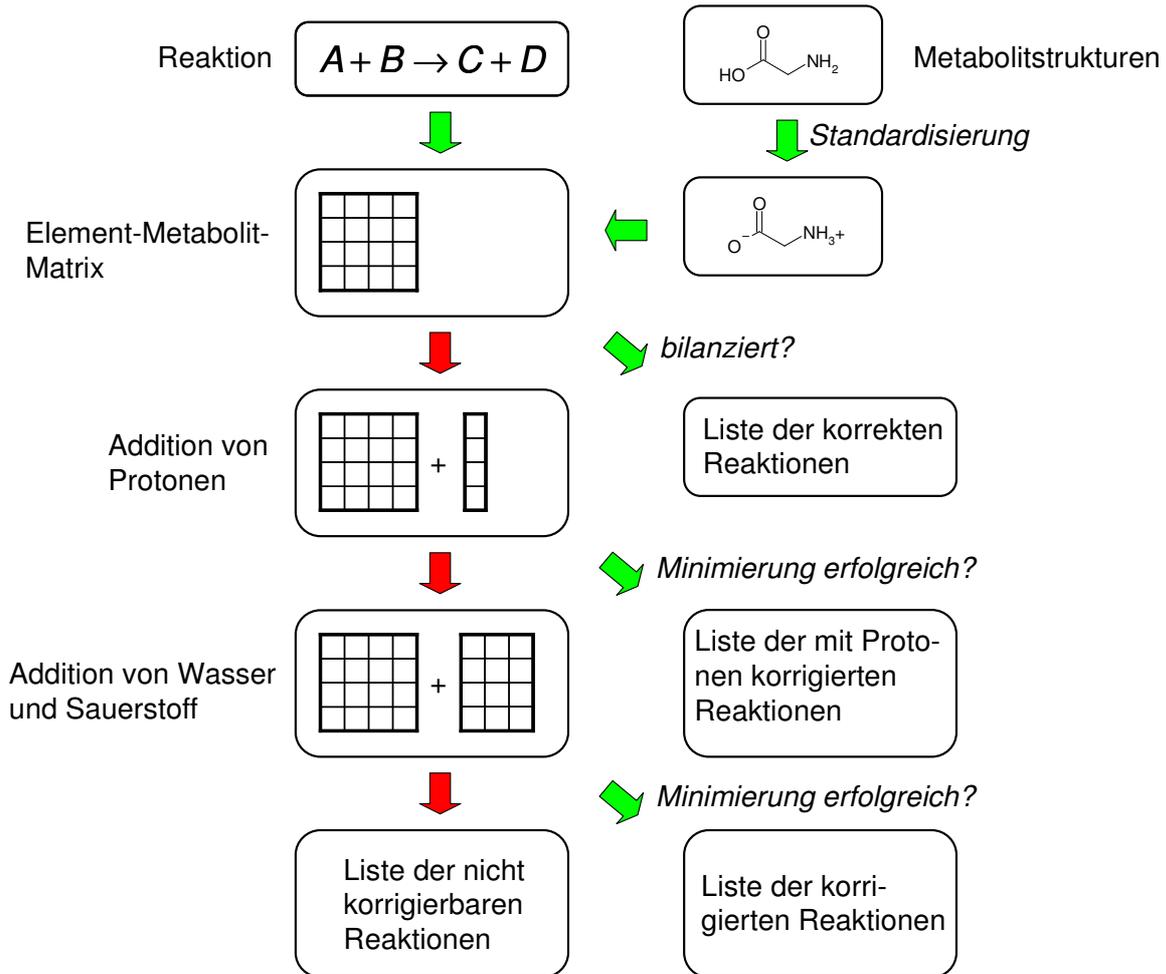


Abbildung 10 Übersicht über den Ablauf der Reaktionsbilanzierung.

Die Reaktionsbilanzierung folgt dem Schema in Abbildung 10. Man startet mit den Strukturen der einzelnen Metabolite sowie den Stöchiometrien der Reaktionen. Als erstes werden die Metabolite modifiziert, so dass der Protonierungsgrad dem der vorherrschenden Spezies bei pH 7 und 298,15 K entspricht. Details dazu finden sich in Abschnitt 3.2.1. Anschließend werden die Summenformeln der Metabolite inklusive Formalladung berechnet. Ist die Element- und Ladungsbilanz korrekt, wird die Reaktion nicht verändert. Ist sie unkorrekt, wird zunächst versucht, auf beiden Seiten der Reaktion mit Protonen auszugleichen. Dies soll mögliche Fehler in der Bilanz beseitigen, die durch den veränderten Protonierungsgrad der Metabolite entstehen können. Sollte es nicht zum Erfolg führen, wird die alte Stöchiometrie verworfen. Auf beiden Seiten der Reaktion werden die Metabolite Wasser, Sauerstoff und Proton hinzugefügt. Die Protonen sollen mögliche Ladungsdifferenzen bilanzieren, Wasser und Sauerstoff werden für die Elementbilanzen von Wasserstoff und Sauerstoff bzw. die Anzahl der Redoxäquivalente benötigt. Es wird versucht, eine Kombination von ganzzahligen stöchiometrischen Faktoren ν zu berechnen, die eine minimale Summe ergibt. Die Summe wird zusätzlich

gewichtet: die neu hinzugefügten Metabolite j bekommen ein höheres Gewicht ω , so dass diese nur hinzugefügt werden, wenn es notwendig wird:

$$\sum_i v_i + \sum_j \omega v_j \rightarrow \min \quad (\text{Gleichung 11})$$

Zur Berechnung der Koeffizienten wird die lineare ganzzahlige Optimierung verwendet. Die Details werden im Folgenden anhand eines einfachen Beispiels erläutert.

In der Datenbank KEGG findet man unter dem Eintrag R00937 die Oxidation von Tetrahydrofolat zu Folat, die durch NAD erfolgt. Die vorgeschlagene Reaktionsgleichung lautet:



Im ersten Schritt wird die Element-Metabolit-Matrix für die Reaktion erzeugt, wobei die Edukte ein negatives Vorzeichen erhalten. Die Spalten entsprechen dabei den Summenformeln der Metabolite:

Tabelle 2 Element-Metabolit-Matrix für die Oxidation von Tetrahydrofolat (THF).

	THF	NAD	Folat	NADH
O	-6	-14	6	14
H	-21	-26	17	27
C	-19	-21	19	21
N	-7	-7	7	7
P	0	-2	0	2
Ladung	2	1	-2	-2

Zu dieser Matrix werden weitere Spalten hinzugefügt für die Metabolite, die zur Bilanzierung benötigt werden. Als Standard werden Protonen, Wasser und Sauerstoff verwendet (siehe oben). Die Metabolite müssen mit negativem und positivem Vorzeichen eingefügt werden, da sie sowohl auf Edukt- als auch auf Produkt-Seite auftreten können. Für die lineare Optimierung wird den Original-Metaboliten ein Gewicht von eins zugeordnet, die hinzugefügten Metabolite bekommen ein Gewicht von zehn. Zusätzlich müssen untere und obere Schranken für die Stöchiometrien, die optimiert werden sollen, angegeben werden. Für die Original-Metabolite wird eine untere Schranke von eins gewählt, es sei denn, der Metabolit ist identisch mit einem hinzugefügten Metaboliten. In diesem Fall wird die untere Schranke auf null gesetzt. Dies ist auch die untere Schranke von allen zugefügten Metaboliten. Die obere Schranke für alle stöchiometrischen Faktoren ist 20.

Die lineare ganzzahlige Optimierung sucht innerhalb dieser Beschränkungen nach einer minimalen Lösung der gewichteten Summe der Koeffizienten (Gleichung 11). Für das betrachtete Beispiel liefert es die korrekte Lösung (Tabelle 3). Es werden zwei Äquivalente

NAD benötigt, um Tetrahydrofolat zu Folat zu oxidieren. Zusätzlich werden zwei Protonen auf Produktseite addiert.

Tabelle 3 Eingabeparameter für die Optimierung der stöchiometrischen Faktoren und deren Ergebnis.

	THF	NAD	Folat	NADH	H ⁺	H ⁺	H ₂ O	H ₂ O	O ₂	O ₂
O	-6	-14	6	14	0	0	-1	1	-2	2
H	-21	-26	17	27	-1	1	-2	2	0	0
C	-19	-21	19	21	0	0	0	0	0	0
N	-7	-7	7	7	0	0	0	0	0	0
P	0	-2	0	2	0	0	0	0	0	0
Ladung	2	1	-2	-2	-1	1	0	0	0	0
Gewicht	1	1	1	1	10	10	10	10	10	10
Grenzen	1-20	1-20	1-20	1-20	0-20	0-20	0-20	0-20	0-20	0-20
Ergebnis	1	2	1	2	0	2	0	0	0	0

Details und Implementierung

Um die Wirksamkeit des Verfahrens zu überprüfen, wurde es auf alle Reaktionen der Datenbank KEGG Ligand (Kanehisa *et al.* 2006) angewendet (Stand: September 2006). Als erstes wurde der Protonierungsgrad der Metabolite mit dem im Gibbspredictor (siehe Abschnitt 3.5.1.3) verwendeten Verfahren korrigiert. Die Reaktionen und Metabolite wurden mit einem Programm, das auf der Bibliothek CMP (siehe Abschnitt 3.5.1.2) basiert, prozessiert. Für die lineare Optimierung wurde die C-Bibliothek Ipsolve (siehe Punkt 3.5.2.8) verwendet, die in CMP integriert wurde. Die Ergebnisse der Optimierung für alle Reaktionen der Datenbank KEGG Ligand sind in Abschnitt 4.1.1 zusammengetragen.

3.2 Abschätzung der Gleichgewichtskonstanten

3.2.1 Theorie

Der Bestimmung der Gleichgewichtskonstanten K_{eq} kommt bei der Beschreibung metabolischer Netzwerke eine hohe Bedeutung zu. Sie dienen nicht nur zur Ermittlung der transienten Reaktionsrichtung einer Reaktion, sondern können auch dazu verwendet werden, die Anzahl der Parameter einer Ratengleichung um eins zu verringern. Im einfachsten Fall, der chemischen Kinetik, erhält man folgende Ratengleichung:

$$v = -k_1 \prod_{\text{Edukte } i} c_i + k_{-1} \prod_{\text{Produkte } j} c_j = -k_1 \prod_{\text{Edukte } i} c_i + \frac{k_1}{K_{eq}} \prod_{\text{Produkte } j} c_j \quad (\text{Gleichung 13})$$

Hier sind k_1 und k_{-1} die kinetischen Konstanten der Hin- bzw. Rückreaktion, K_{eq} die Gleichgewichtskonstante und c_i bzw. c_j die Konzentrationen der beteiligten Metabolite.

Die größte publizierte Sammlung von Gleichgewichtskonstanten für biochemische Reaktionen ist von R. N. Goldberg, Y. B. Tewari und T. N. Bhat zusammengestellt worden (Goldberg *et al.* 2004). Die Daten sind in einer Datenbank gespeichert, auf die über ein Webinterface zugegriffen werden kann (siehe Abschnitt 3.5.3.4). Soweit vorhanden, sind die experimentellen Bedingungen wie Pufferkonzentrationen, Temperatur, pH-Wert, Ionenstärke, Methode und Literaturquelle festgehalten. Außerdem wurden die Datensätze je nach experimenteller Methode und methodischer Prozedur in vier Qualitätsklassen von A (beste) bis D (schlechteste) eingeteilt.

Ein alternativer Zugang zu Gleichgewichtskonstanten erfolgt über Standard Gibbsenergien von Reaktionen $\Delta_R G^{\circ}$. Aus der Beziehung

$$\Delta_R G' = \Delta_R G^{\circ} + RT \ln Q \quad \text{mit} \quad Q = \prod_{\text{Metabolite}} c^{\nu} \quad (\text{Gleichung 14})$$

ergibt sich für den Fall des thermodynamischen Gleichgewichts, also $\Delta_R G' = 0$:

$$K_{eq} = Q = e^{-\frac{\Delta_R G^{\circ}}{RT}} \quad (\text{Gleichung 15})$$

Dabei stehen $\Delta_R G'$ für die aktuelle Gibbsenergie der Reaktion, R für die Gaskonstante, T für die Temperatur, Q für den Reaktionskoeffizienten, c für die Konzentrationen der Metabolite, ν für die stöchiometrischen Faktoren und K_{eq} für die Gleichgewichtskonstante.

Über die Gibbsenergie der Reaktion können eine Reihe fundamentaler physikochemischer Gesetze zueinander in Beziehung gesetzt werden. Neben Gleichung 14 gibt es zwei weitere Beziehungen, die im biochemischen Kontext relevant sind:

$$\begin{aligned} \Delta_R G' &= \Delta_R H' - T \Delta_R S' \\ \Delta_R G' &= -nF \Delta E' \end{aligned} \quad (\text{Gleichungen 16 und 17})$$

Die erste Gleichung ist eine unmittelbare Folge des Zweiten Hauptsatzes der Thermodynamik: Eine Reaktion kann nur dann spontan ablaufen, wenn die Gibbsenergie des Systems abnimmt. Dies kann grundsätzlich entweder durch eine Zunahme der Entropie erfolgen, oder durch eine Abnahme der Enthalpie des Systems (was mit einer Entropiezunahme in der Umgebung einhergeht). Eine Kombination beider Effekte bzw. eine Kompensation, insofern der negative Term überwiegt, sind ebenfalls möglich.

Gleichung 17 verknüpft die Gibbsenergie mit dem Zellpotential einer elektrochemischen Zelle. Es wird häufig verwendet, um Membranpotentiale zu beschreiben, die elektrische Spannung zwischen Innen- und Außenseite einer Biomembran. Das Membranpotential kann auch für einzelne Ionen angegeben werden, z. B. Wasserstoff-Ionen (Protonenmotorische Kraft) oder Kalium-Ionen (Kaliumpotential).

Kehren wir nun zurück zu Gleichung 15, die es ermöglicht, bei Kenntnis der Standard Gibbsenergie einer Reaktion deren Gleichgewichtskonstante zu berechnen. Die Standard Gibbsenergie der Reaktion kann aus den Bildungs-Gibbsenergien G_f° der beteiligten Metabolite leicht errechnet werden:

$$\Delta_R G^{\circ} = \sum_i \nu_i G_{f,i}^{\circ} \quad (\text{Gleichung 18})$$

Hier ist ν_i der stöchiometrische Faktor des Metaboliten i in der betrachteten Reaktion. In Abschnitt 3.2.2 wird ein Verfahren vorgestellt, um Bildungs-Gibbsenergien abzuschätzen. Bei Anwendung der Formel ist zu berücksichtigen, dass es sich bei den Bildungsenergien um aufsummierte Energien, eine so genannte apparente Eigenschaft, handelt. Apparente Eigenschaften werden im folgenden Abschnitt behandelt.

Apparente Eigenschaften

Zu den meisten Metaboliten existieren mehrere verschiedene Spezies in wässriger Lösung. Abhängig vom pK_a -Wert eines Metaboliten und dem pH-Wert der Lösung liegt der Metabolit beispielsweise zu bestimmten Anteilen als Säure bzw. konjugierte Base vor (vergleiche Abbildung 9). Aber auch andere Spezies können eine Rolle spielen, z. B. Addukte von Magnesium-Ionen an die Phosphatgruppen von ATP oder unterschiedliche Komplexe bei Eisen-Ionen. Bei der Berechnung einer Eigenschaft eines bestimmten Metaboliten müssen alle Spezies berücksichtigt werden, die für die Berechnung im Rahmen der angestrebten Genauigkeit relevant sind. Das gilt zum einen für die Spezies, die zu großen Anteilen in Lösung vorliegen. Zum anderen können aber auch Spezies einen Einfluss haben, die nur einen geringen Anteil ausmachen, bei denen aber der absolute Wert groß ist gegenüber dem Wert der anderen Spezies.

Die Anteile der unterschiedlich protonierten Spezies bei gegebenem pH-Wert kann anhand der pK_a -Werte des betrachteten Metaboliten berechnet werden. Andere Einflussfaktoren

werden zunächst nicht berücksichtigt. Zur Berechnung des pK_a -Werts wird die Bibliothek JChem von ChemAxon verwendet (siehe Abschnitt 3.5.2.7).

Transformierte Gibbsenergien

In manchen Fällen ist es zweckmäßig oder notwendig, konstante Terme eines Ausdrucks in den Gibbsenergie-Term einzubeziehen. Man gelangt so zu transformierten Gibbsenergien (Wadsö *et al.* 1976). Ein häufig auftretendes Beispiel ergibt sich bei gepuffertem pH-Wert und sehr hohen Wasser-Konzentrationen, die wie folgt in die Gibbsenergie einer Reaktion eingerechnet werden können:

$$\Delta_R G' = \Delta_R G + \nu_{H^+} RT \ln c_{H^+} + \nu_{H_2O} RT \ln c_{H_2O} \quad (\text{Gleichung 19})$$

Auch Aktivitätskoeffizienten können auf diese Weise berücksichtigt werden:

$$\begin{aligned} \Delta_R G &= \Delta_R G^0 + RT \ln \prod_i a_i^{\nu_i} = \Delta_R G^0 + RT \ln \prod_i \gamma_i^{\nu_i} + RT \ln \prod_i c_i^{\nu_i} \\ &= \Delta_R G^0 + RT \ln \prod_i c_i^{\nu_i} \end{aligned} \quad (\text{Gleichung 20})$$

Dabei ist ν_i der stöchiometrische Faktor, c_i die Konzentration und a_i die Aktivität des Metaboliten i . R steht für die universelle Gaskonstante und T für die Temperatur.

Berücksichtigung der Aktivitätskoeffizienten

Die biologische Zelle ist eine Mischung aus einer Vielzahl chemischer Verbindungen. Um den Begriff Aktivität zu erklären, müssen zunächst die Begriffe *reale* und *ideale* Mischung erläutert werden. Ist die Wechselwirkungsenergie zwischen den Molekülen der verschiedenen Komponenten einer Mischung gleich dem arithmetischen Mittel der Wechselwirkungsenergien zwischen gleichartigen Molekülen, spricht man von idealem Verhalten. Dies ist nur dann der Fall, wenn sich die Eigenschaften der einzelnen Komponenten in Lösung sehr ähnlich sind oder wenn die Wechselwirkungsenergie sehr gering ist, zum Beispiel bei extrem (ideal) verdünnten Lösungen. Beide Voraussetzungen sind im biologischen System im Allgemeinen nicht erfüllt, so dass wir es mit einer nicht-idealen bzw. realen Mischung zu tun haben (Wedler 1997).

Die Abweichung einer realen Mischung vom idealen Verhalten kann berücksichtigt werden, wenn man die Konzentration der Metaboliten durch die Aktivität ersetzt. Die Aktivität a_i ist proportional zur Konzentration c_i :

$$a_i = \gamma_i c_i \quad (\text{Gleichung 21})$$

Der Proportionalitätsfaktor ist der Aktivitätskoeffizient γ_i . Bei idealem Verhalten gilt $\gamma_i = 1$.

Die klassischen experimentellen Verfahren zur Bestimmung von Aktivitätskoeffizienten, z. B. Dampfdruckmessungen oder potentiometrische Verfahren, sind bei Mischungen mit vielen Komponenten sehr aufwendig bzw. nicht anwendbar. Es bietet sich an, die

Koeffizienten abzuschätzen. Die erweiterte Debye-Hückel-Theorie liefert hierzu einen praktikablen Ansatz. Einzig die Ladung z_i des jeweiligen Metabolite muss dafür bekannt sein:

$$\ln \gamma_i = -\alpha \frac{z_i^2 \sqrt{I}}{1 + B\sqrt{I}} \quad (\text{Gleichung 22})$$

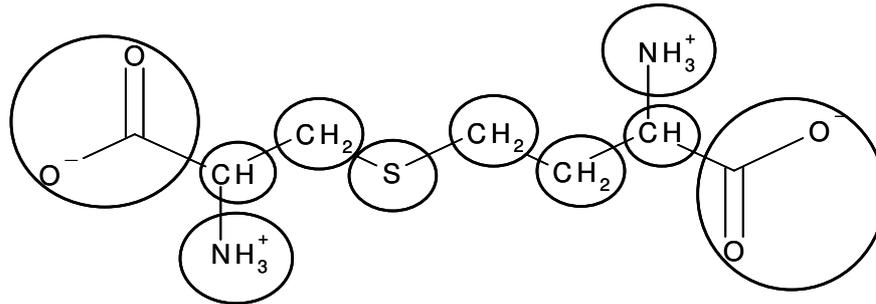
Die beiden Konstanten sind $\alpha = 0,510651 \text{ mol}^{1/2} \text{ l}^{-1/2}$ und $B = 1,6 \text{ mol}^{-1/2} \text{ l}^{1/2}$ bei einer Temperatur von 298,15 K in wässriger Lösung. Die Gleichung gilt nur für geringe Ionenstärken, $I \leq 0,01 \text{ mol/kg}$ (Atkins 1996) bzw. $I \leq 0,35 \text{ mol/kg}$ (Alberty 2006). Im Cytosol herrscht eine durchschnittliche Ionenstärke von 0,25 mol/kg (Alberty 2006), so dass laut Alberty Gleichung 22 anwendbar ist.

Die Ladung der Metabolite ist diejenige der vorherrschenden Spezies bei pH 7. Sollte der Metabolit einen pK_a -Wert in der Nähe des pH-Wertes besitzen, müssen die einzelnen Spezies anteilig berücksichtigt werden. Die Anteile ergeben sich analog zur zweiten Gleichung in Abbildung 12 auf Seite 42.

3.2.2 Berechnung nach Mavrovouniotis

M. Mavrovouniotis hat ein Verfahren entwickelt, das die Abschätzung von Bildungs-Gibbsenergien für Moleküle erlaubt (Mavrovouniotis 1990, 1991). Der grundsätzliche Prozess läuft wie folgt ab: Ein Molekül wird in funktionelle Gruppen aufgespalten. Jeder Gruppe ist ein Beitrag zur Gibbsenergie zugeordnet. Die Summe dieser Beiträge ergibt die Bildungs-Gibbsenergie des Moleküls (Abbildung 11).

Die einzelnen Beiträge wurden aus einem Datensatz von Gleichgewichtskonstanten enzymkatalysierter Reaktionen, der aus verschiedenen Quellen zusammengestellt wurde, durch multiple lineare Regression ermittelt. Die Daten sind für eine verdünnte wässrige Lösung bei pH 7 und 25 °C als Standardzustand ausgewählt worden. Das Verfahren liefert daher nur Abschätzungen für diese Standardbedingungen. Aus den abgeschätzten Werten für die Bildungs-Gibbsenergie der Edukte und Produkte sowie der Stöchiometrie kann die Standard Gibbsenergie der Reaktion und daraus die Gleichgewichtskonstante näherungsweise berechnet werden (Gleichungen 18 und 15).



Gruppe	Anzahl	Beitrag [kcal/mol]	Gesamt [kcal/mol]
Grundbeitrag	1	-23,6	-23,6
-S-	1	9,5	9,5
-NH ₃ ⁺	2	4,3	8,6
-COO ⁻	2	-72,0	-144,0
>CH ₂	3	1,7	5,1
>CH-	2	-4,8	-9,6
Bildungs-Gibbsenergie			-153,6

Abbildung 11 Berechnung der Gibbsenergie eines Metaboliten nach der Inkrementmethode von Mavrovouniotis.

Für insgesamt 60 funktionelle Gruppen stehen Energiebeiträge zur Verfügung (Anhang 6.1). Damit werden fast alle Moleküle im biochemischen Kontext abgedeckt, die die Elemente Schwefel, Phosphor, Stickstoff, Sauerstoff und Kohlenstoff enthalten. Zusätzlich gibt es noch einige Korrekturterme, z. B. für Amid-Gruppen oder Dreiringe.

Veränderung des Protonierungsgrads

Bevor die Methode angewendet werden kann, muss das betreffende Molekül in den Zustand bei Standardbedingungen überführt werden. Um die vorherrschende Spezies bei pH 7 und 298.15 K zu ermitteln, müssen die pK_a-Werte und die zugehörigen aciden Wasserstoffe bekannt sein. Beide Informationen liefert das pK_a-Plugin der Software JChem (siehe Abschnitt 3.5.2.7). Hat man die vorherrschende Spezies identifiziert, müssen noch einige Modifizierungen vorgenommen werden, weil die Gruppen sonst nicht korrekt erkannt werden.

Erweiterungen

Die Methode funktioniert schlecht bei anorganischen Molekülen und kleinen organischen Molekülen. Damit auch für diese Moleküle eine Bildungs-Gibbsenergie ermittelt werden kann, wurden die Strukturen mit den zugehörigen Energien gespeichert (Tabelle 4). Die Energien stammen aus (Mavrovouniotis 1990) und (Weast 1989). Sollte das Molekül exakt

mit einem gespeicherten Molekül übereinstimmen, wird die entsprechende Energie ausgegeben, anstatt die Berechnung zu starten.

Tabelle 4 Standard Gibbsenergien von kleinen Molekülen und für Metabolite, die als Teil von Metabolitpaaren auftreten.

Metabolit	Gibbsenergie [kcal/mol]
Wasser	-56,6
Ammonium-Ion	-18,1
Nitrit	-7,0
Nitrat	-27,5
Sulfit	-118,4
Phosphat	-59,6
Diphosphat	-57,4
Bicarbonat	-140,2
Formaldehyd	-36,1
Format	-84,9
Methan	-8,2
Oxalat	-159,8
Sulfat	-178,0
Wasserstoffperoxid	-32,0
Proton	-9,5
CO ₂	-93,1
NAD (mit NADH)	-132,44
NADP (mit NADPH)	-130,44
CoA (mit substituiertem CoA)	-115,7

In (Mavrovouniotis 1990) sind auch Korrekturterme für einige in Reaktionen häufig auftretende Metabolitpaare angegeben. Diese sind im einzelnen NAD/NADH, NADP/NADPH und CoA/substituiertes CoA. Da die Paare meist zusammen auftreten, reicht es aus, die Gibbsenergie eines der beiden Metabolite zu modifizieren. Für den Fall, dass sie nicht zusammen auftreten, lässt sich diese Korrektur abschalten.

Vinyloge Carbonsäuren wurden von der ursprünglichen Methode nicht behandelt. Man kann näherungsweise annehmen, dass die Bildungs-Gibbsenergie ähnlich zu der Gibbsenergie einer Säuregruppe ist. Weil negativ geladener Sauerstoff in dieser Methode nur bei Säuregruppen vorkommt, war es möglich, die Energie für die Gruppe O⁻ indirekt aus der Differenz einer Säuregruppe und einer Carbonylgruppe zu berechnen.

Die Ionenstärke wird im Rahmen der erweiterten Debye-Hückel-Theorie berücksichtigt (siehe Gleichung 22). Dabei werden nicht nur die vorherrschende Spezies, sondern alle relevanten Spezies und ihre Formalladung anteilig einbezogen.

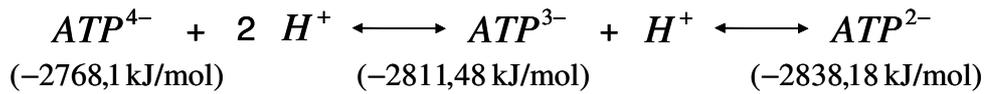
3.2.3 Berechnung nach Alberty

Die Berechnung von Gleichgewichtskonstanten nach Alberty unterscheidet sich in grundsätzlicher Weise von der Methode nach Mavrovouniotis (Abschnitt 3.2.2). Die Grundlage bildet ein Datensatz mit Bildungs-Gibbsenergien für 199 Metabolite (z. B. ATP) und insgesamt 286 Spezies (z. B. ATP^{4-} , HATP^{3-}). Der Datensatz ist in (Alberty 2006) veröffentlicht.

Es ist mit dem Verfahren möglich, die Gibbsenergie der Reaktion im pH-Bereich 5-9 und bei Ionenstärken bis zu 0,35 mol/kg zu bestimmen. Ausgehend von den Bildungs-Gibbsenergien der einzelnen Spezies werden für vorgegebenen pH-Wert und Ionenstärke für jeden Metaboliten zunächst die pK_a -Werte und daraus die Anteile der Spezies im Gleichgewicht ermittelt. Daraufhin wird die transformierte Gibbsenergie, korrigiert mit pH-Wert und Ionenstärke, der einzelnen Spezies berechnet. Da die Gibbsenergien der einzelnen Spezies im Gleichgewicht identisch sein müssen, reicht es aus, die Gibbsenergie einer Spezies mit ihrem entsprechenden Anteil zu korrigieren, um die Bildungs-Gibbsenergie des Metaboliten zu erhalten.

Der Ablauf der Berechnung wird in Abbildung 12 am Beispiel von ATP vorgestellt, wobei aus Gründen der Übersichtlichkeit der Vorgang leicht verändert dargestellt ist und eine Ionenstärke von null angenommen wird. Die Ionenstärke-Korrektur erfolgt im Rahmen der erweiterten Debye-Hückel-Theorie (Gleichung 22).

Mit der so erhaltenen apparenten Bildungs-Gibbsenergie kann die Standard Gibbsenergie der Reaktion berechnet werden. Die Protonen müssen nicht mehr berücksichtigt werden, da die Änderung der Protonenanzahl schon implizit in den Bildungs-Gibbsenergien enthalten ist.



$$1. \quad pK_a = - \left(\frac{G_{f,\text{deprotoniert}}^0 - G_{f,\text{protoniert}}^0}{RT \cdot \ln(10)} \right)$$

$$2. \quad x_i = \frac{10^{-pH \cdot i} \cdot \prod_{k=1}^i 10^{pK_{a,k}}}{1 + \sum_{j=1}^{n-1} \prod_{k=1}^j 10^{pK_{a,k} - pH}} \quad \text{mit } \begin{array}{l} i = 0, \dots, n-1 \\ n = \text{Anzahl der Spezies} \end{array}$$

$$3. \quad G_{f,i}^0 = G_f^0 + n_{\text{H}^+} \cdot RT \ln(10^{-pH}) + RT \ln(x_i)$$

$$pK_{a,1} = 7,60$$

$$pK_{a,2} = 4,68$$

$$x_0 = 0,200$$

$$x_1 = 0,796$$

$$x_2 = 0,004$$

$$G_{f,ATP}^0 = -2852,0 \text{ kJ/mol}$$

Abbildung 12 Berechnung der Gibbsenergie einer Reaktion nach Alberty. 1. Berechnung des pK_a -Wertes aus den Gibbsenergien der einzelnen Spezies. 2. Kalkulation der Anteile der Spezies. 3. Berechnung der apparenten Bildungs-Gibbsenergie.

3.3 Simulationen ausgewählter Netzwerke

In diesem Kapitel werden einige metabolische Netzwerke vorgestellt und die Analyse und Simulation mit verschiedenen Methoden besprochen. Wir beginnen mit sehr simplen, artifiziellen Modellen und gehen dann zu komplexeren Modellen über, die einem Teilnetzwerk des bakteriellen Gesamtmetabolismus entsprechen. Die höhere Komplexität kann sich hierbei sowohl auf die Größe und Vernetzung des Netzwerks beziehen, als auch auf die mathematische Beschreibung der einzelnen Prozesse.

Grundsätzlich wurde die Simulation per Integration eines Differentialgleichungssystems durchgeführt. Zusätzlich wurden die Simulation eines farbigen Petrinetzes und die Lösung des algebraischen Gleichungssystems getestet.

Als erstes wird ein sehr einfaches, lineares Modell beschrieben. Anhand dieses Beispiels kann man sehr gut illustrieren, welche Einflüsse verschiedene Zufluss- und Abflussfunktionen haben. Außerdem wird die metabolische Kontrollanalyse (vergleiche Abschnitte 2.6.1 und 3.3.3) und die Aussagen, die daraus für das Verhalten des Netzwerkes geschlossen werden können, angewendet. Der Einfluss unterschiedlicher Ratengleichungen kann gleichfalls nachvollziehbar dargelegt werden. Anhand eines verzweigten Modells wird gezeigt, wie sich getrennt und gemeinsam modellierte Abflüsse auswirken.

Als erstes Modell eines Teilnetzwerks wird die Glykolyse des *Corynebacterium glutamicum* behandelt. Die Struktur des Netzwerks ist dabei auf andere Organismen übertragbar. Die Werte einiger Parameter sind an *C. glutamicum* angepasst. Als Beispiel für ein zyklisches Netzwerk wird der Citratcyclus behandelt. Den Abschluss bildet ein Modell des Zentralstoffwechsels.

3.3.1 Modellierungsprinzipien

Alle Modellierungen sollen auf das Erreichen eines realitätsnahen Fließgleichgewichtszustandes hinauslaufen. Die wesentlichen Probleme bezüglich der Datenlage wurden in Abschnitt 2.2 beschrieben. Als Konsequenz daraus ergibt sich die Notwendigkeit, die Anzahl der unbekannt Parameter so gering wie möglich zu halten. Die einfachste kinetische Beschreibung der Reaktionen erfolgt über chemische Kinetik, die als Parameter nur die kinetischen Konstanten der Hin- und Rückreaktion beinhaltet. Über die Gleichgewichtskonstante der Reaktion lassen sich die beiden Konstanten miteinander in Beziehung setzen, so dass nur ein unbekannter Parameter pro Reaktion übrig bleibt. Eine Beziehung dieser Art gibt es nicht nur für chemische Kinetik. Für jede Ratengleichung lässt sich eine solche Relation finden (Abbildung 13). Zur Berechnung der Gleichgewichtskonstanten wurde das Verfahren nach Mavrovouniotis verwendet.

$$v = k_1 \left(e - \frac{p}{K_{eq}} \right) \quad \text{Chemische Kinetik}$$

$$v = \frac{V_f \cdot \frac{e}{K_e} \left(1 - \frac{Q}{K_{eq}} \right)}{1 + \frac{e}{K_e} + \frac{p}{K_p}} \quad \text{Michaelis-Menten (uni-uni)}$$

$$v = \frac{V_f \cdot \frac{e_1 e_2}{K_{e_1} K_{e_2}} \left(1 - \frac{Q}{K_{eq}} \right)}{\left(1 + \frac{e_1}{K_{e_1}} + \frac{p_1}{K_{p_1}} \right) \cdot \left(1 + \frac{e_2}{K_{e_2}} + \frac{p_2}{K_{p_2}} \right)} \quad \text{Michaelis-Menten (bi-bi)}$$

Abbildung 13 Parameterreduktion bei reversiblen Ratengleichungen durch die Gleichgewichtskonstante K_{eq} . e, p : Metabolitkonzentrationen; k : Geschwindigkeitskonstante; Q : Massenwirkungsgesetz; V_f : Maximalgeschwindigkeit der Hinreaktion; K_x : Michaelis-Konstante von Metabolit x .

Das Auffinden der Metabolite, die die Systemgrenzen bilden, ist der erste wichtige Schritt der Modellierung. Üblicherweise fließen die Kenntnisse über Zu- und Abflüsse bereits in die Modellformulierung ein. Bei einer Bakterienzelle, die auf Minimalmedium wächst, sind die Zuflussstellen durch die Bestandteile des Minimalmediums gegeben. Die Abflussstellen müssen aus biochemischem Wissen abgeleitet werden. Man kann die Precursor der Substanzen, die für das Wachstum der Zelle notwendig und Teil des Modells sind, identifizieren und quantifizieren. Als Precursor für den Proteinanteil können die einzelnen Aminosäuren herangezogen werden. Die Abflüsse der Aminosäuren können theoretisch in statistisch ermitteltem Verhältnis zueinander in Beziehung gesetzt werden. Analoges gilt für den Aufbau von DNA und RNA aus Nucleotiden oder den Aufbau der Zellwand aus Peptidoglycanen und weiteren Bestandteilen wie Lipiden oder Proteinen. Neben den bekannten gibt es möglicherweise einige unbekannte Systemgrenzen. Die Identifikation der losen Enden des Netzwerks, also der Metabolite, die nur in einer Reaktion vorkommen, kann einen ersten Anhaltspunkt geben. Die Berechnung der Conserved Moieties (vergleiche Abschnitt 2.3.1) kann weitere Systemgrenzen aufzeigen. Die Probleme, die auftreten können, wenn man die Conserved Moieties nicht berücksichtigt, werden in Abschnitt 3.3.2 diskutiert. Die Konzentrationen der Metabolite, die lose Enden und Conserved Moieties darstellen, werden bei der Simulation in den meisten Fällen fixiert.

Als zusätzlicher Schnelltest der prinzipiellen Funktionsfähigkeit des Modells kann eine Flux Balance Analysis durchgeführt werden. Dies ist allerdings kein Garant für eine erfolgreiche kinetische Simulation, sondern kann nur Modelle ausschließen, für die keine Flussverteilung realisierbar ist.

Soweit nicht anders angegeben, wurden die Simulationen durch Integration von Differentialgleichungssystemen mit Copasi (Abschnitt 3.5.2.5) durchgeführt. Einen zentralen Punkt in der Analyse der Fließgleichgewichtszustände nahm die metabolische Kontrollanalyse ein. Die Kontrollkoeffizienten wurden ebenfalls mit Copasi berechnet.

3.3.2 Modellierung von Conserved Moieties

Die Conserved Moieties (vergleiche Abschnitt 2.3.1, Seite 16) müssen im Modellierungsprozess besonders behandelt werden. Es gibt mehrere Möglichkeiten, mit den Erhaltungssätzen für molekulare Teilstrukturen umzugehen. Im Grundsatz handelt es sich bei den entsprechenden Metaboliten um Systemgrenzen. Wenn das Modell vollständig wäre, müssten alle molekularen Teilstrukturen aus dem Nährmedium aufgebaut werden können bzw. direkt durch Transportprozesse in die Zelle gelangen. Behandelt man die Metabolite nicht gesondert, kann es dazu führen, dass die Conserved Moieties die Flussverteilung im System entscheidend beeinflussen oder im schlimmsten Fall das Erreichen eines stationären Zustands unmöglich machen. Als einfaches Beispiel für den letzteren Fall kann das Metabolit-Paar ATP/ADP in der Glykolyse herangezogen werden. Da die Glykolyse in der Nettobilanz zwei Moleküle ADP zu ATP phosphoryliert, kann kein Fließgleichgewicht erreicht werden, wenn die ADP- und ATP-Konzentrationen als Variablen betrachtet werden. Nehmen wir an, dass die Glykolyse in ein größeres metabolisches Netzwerk eingebettet wird, das in einer anderen Reaktion ATP verbraucht. In diesem Fall müsste der Fluss durch diese Reaktion exakt doppelt so hoch sein wie der Fluss durch die Glykolyse. Bei mehreren Stellen mit ATP-Verbrauch ist der Zusammenhang komplizierter, aber dennoch kann das Systemverhalten maßgeblich von dieser Conserved Moiety beeinflusst werden.

Um dies zu verhindern, bieten sich mehrere Lösungsmöglichkeiten an, die alle zum gleichen Verhalten führen. Die Konzentration kann auf einen bestimmten Wert fixiert werden, so dass in unserem Beispiel das ATP/ADP-Verhältnis konstant ist. Es kann eine externe Reaktion für jeden Metaboliten eingerichtet werden, die schnell gegenüber den anderen Reaktionen im System ist und für eine konstante Konzentration der Metabolite sorgt. Als dritte Option kann eine schnelle Gleichgewichtsreaktion zwischen den Metaboliten eingerichtet werden, die für die Regeneration sorgt. Durch die Lage des Gleichgewichts können die Konzentrationen eingestellt werden. Würde man den Fluss genau kennen, wie es in der Glykolyse der Fall ist, könnte man auch einen konstanten Zu- und Abfluss für beide Metabolite bzw. zwischen den Metaboliten einrichten.

Soweit nicht anders beschrieben, wurde immer die erste Methode gewählt, da sie das Modell nicht zusätzlich vergrößert und es keine Unterschiede bezüglich des

Systemverhaltens zwischen den einzelnen Behandlungsmethoden für Conserved Moieties gibt.

3.3.3 Berechnung der metabolischen Kontrollkoeffizienten

Die Theorie zur metabolischen Kontrollanalyse wurde bereits in Abschnitt 2.6.1 behandelt. Im Folgenden wird erläutert, wie die Koeffizienten gedeutet und in der vorliegenden Arbeit dargestellt werden. Die Berechnung der Kontrollkoeffizienten erfolgte mit Copasi (vergleiche Abschnitt 3.5.2.5).

Die Kontrollkoeffizienten eines Systems können kompakt in einer einzigen Tabelle dargestellt werden (siehe Abbildung 14). Die Metabolite bzw. Stoffflüsse stehen in den Zeilen. In den Spalten stehen die Kontrollkoeffizienten und die Konzentrationen und Flüsse im Fließgleichgewicht. Betrachten wir die erste Zeile aus dem Tabellenrumpf in Abbildung 14. Die Konzentration von Metabolit M1 im Fließgleichgewicht ist in der letzten Spalte angegeben. Eine Veränderung der Konzentration des Input-Enzyms hat den größten Effekt, da der Kontrollkoeffizient den betragsmäßig höchsten Wert aufweist. Ein positives Vorzeichen bedeutet, dass bei einer Erhöhung der Enzymkonzentration auch die Metabolitkonzentration erhöht wird. Die anderen drei Reaktionen haben jeweils gleichen Einfluss auf die Konzentration von M1. Eine Erhöhung bewirkt eine Verringerung der Metabolitkonzentration. Analoges gilt für die anderen Metabolite sowie die Stoffflüsse.

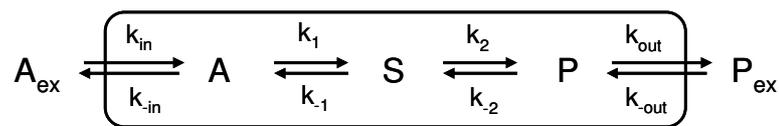
		Kontrollkoeffizienten aus der Metabolic Control Analysis			Konzentrationen und Flüsse im Fließgleichgewicht	
		Input	R 1	R 2	Output	Konz/Fluss*
Metabolite	M 1	0.247	-0.082	-0.082	-0.082	0.7525
	M 2	0.245	0.245	-0.245	-0.245	0.505
	M 3	0.240	0.240	0.240	-0.721	0.2575
Stoffflüsse	Fluss 1	0.250	0.250	0.250	0.250	2.48E-15
	Fluss 2	0.250	0.250	0.250	0.250	2.48E-15

Abbildung 14 Tabelle mit Kontrollkoeffizienten der metabolischen Kontrollanalyse.
*Konzentrationen in mM, Flüsse in mmol/s.

3.3.4 Artificielle Modelle

3.3.4.1 Lineares Modell

Dieses Modell besteht aus zwei externen und drei internen Metaboliten und vier Reaktionen, einer Zufluss-, einer Abfluss- und zwei internen Reaktionen. Die Ratengleichungen der einzelnen Reaktionen beruhen auf chemischer Kinetik. Der Referenzzustand für das System im Fließgleichgewicht inklusive kinetischer Parameter, Stoffflüsse und Konzentrationen findet sich in Abbildung 15.



Parameter / Variable	Wert
$k_{in,1,2,out}$ und $k_{-in,-1,-2,-out}$	1 s^{-1}
$[A_{ex}]$	1 mM (fixiert)
$[A]$, $[S]$, $[P]$	0,1 mM (initial)
$[P_{ex}]$	0,01 mM (fixiert)
$[A]_{ss}$	0,75 mM
$[S]_{ss}$	0,51 mM
$[P]_{ss}$	0,26 mM
Stofffluss _{ss}	$2,48e-15 \text{ mmol/s}$

Abbildung 15 Lineares Netzwerk mit zwei internen Reaktionen. Ergebnis der Berechnung des Fließgleichgewichts (ss = steady state) bei den angegebenen Parametern.

Die einzelnen Parameter wurden variiert, um den Einfluss auf die Konzentrationen und den Fluss im Fließgleichgewicht zu erfassen. Bei jeder Variation wurden außerdem die Fluss- und Metabolit-Kontrollkoeffizienten berechnet, um zu zeigen, wie sich die Kontrolle von Experiment zu Experiment verändert.

Für die Zufluss- und Abflussreaktionen stehen zwei grundsätzlich verschiedene Typen von Ratengleichungen zur Verfügung. Der Fluss kann irreversibel auf einen konstanten Wert gesetzt werden, oder es wird eine reversible Ratengleichung, z. B. chemische Kinetik, verwendet. In beiden Fällen muss die Konzentration des externen Metaboliten fixiert werden. Da es bei diesem Modell zwei externe Reaktionen und zwei Typen von Ratengleichungen gibt, gibt es 4 Kombinationsmöglichkeiten, die es zu testen galt.

Die Ratengleichungen der beiden internen Reaktionen wurden außerdem durch reversible Michaelis-Menten-Kinetik ausgetauscht. Der Effekt von Parametervariationen in diesen Fällen wurde ebenfalls untersucht. Dabei wurde jeweils ein Parameter verändert und anschließend der freie Parameter der chemischen Kinetik, die Enzymkonzentration, so

verändert, dass man annähernd das gleiche Ergebnis erhält. Die Parameter des Referenzzustandes sind in Tabelle 5 zu sehen.

Tabelle 5 Parameterwerte der beiden internen Reaktionen bei Michaelis-Menten-Kinetik.

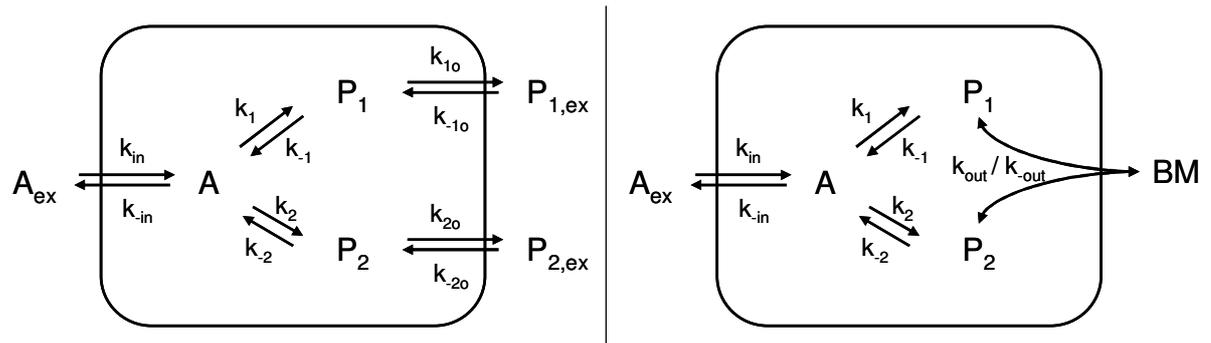
Parameter	Reaktion R ₁	Reaktion R ₂
$V_{\max, \text{hin}}$	2 mM/s	2 mM/s
$V_{\max, \text{rück}}$	2 mM/s	2 mM/s
$K_{m, \text{Edukt}}$	1 mM/s	1 mM/s
$K_{m, \text{Produkt}}$	1 mM/s	1 mM/s

Als letztes sollte untersucht werden, wie sich unterschiedliche Gleichgewichtskonstanten auswirken. Die Gleichgewichtskonstanten sind zwar unveränderliche Parameter der Reaktion, aber Abschätzungsverfahren liefern nur ungefähre Werte, so dass mit dieser Untersuchung der Einfluss von Fehlern im Abschätzungsverfahren deutlich gemacht werden kann.

3.3.4.2 Modelle mit Verzweigungen

An den Modellen mit Verzweigung wurde untersucht, wie sich die Einführung eines artifiziellen Metaboliten „Biomasse“, der das Zellwachstum darstellen soll, auswirkt. Dazu wurden ein Modell mit zwei Abflüssen und ein Modell mit einem gemeinsamen Abfluss zur Biomasse erstellt. Das Modell und die anfänglichen Parameterwerte sind in Abbildung 16 zu sehen.

Im Verlauf der Untersuchungen wurde der Wert von k_1 zunächst um den Faktor 100 auf 100 s^{-1} erhöht. Um bei beiden Modellen wieder das gleiche Verhalten zu provozieren, wurden anschließend k_1 auf $0,5 \text{ s}^{-1}$ und k_{-1} auf $0,005 \text{ s}^{-1}$ erniedrigt, was einer Division beider Parameterwerte durch 200 entspricht.



Parameter / Variable	Wert
$k_{in,1,2,1o,2o}$ und $k_{-in,-1,-2,-1o,-2o}$	1 s^{-1}
k_{out}	$4,4 \text{ mM}^{-1}\text{s}^{-1}$
k_{-out}	$4,4 \text{ s}^{-1}$
$[A_{ex}]$	1 mM (fixiert)
$[P_{1,ex}], [P_{2,ex}], [BM]$	0,01 mM (fixiert)
$[A], [P_1], [P_2]$	0,1 mM (initial)
$[A]_{ss}$	0,505 mM
$[P_1]_{ss}$	0,258 mM
$[P_2]_{ss}$	0,258 mM
Zufluss _{ss}	$4,95\text{e-}16 \text{ mmol/s}$

Abbildung 16 Modelle mit Verzweigung (oben) und Werte der Parameter- und Variablen im Referenzzustand (unter).

3.3.5 Modelle von Stoffwechselwegen

3.3.5.1 Glykolyse

Die Glykolyse ist ein beliebtes Ziel für Modellierungen. Die Datenbank für SBML-Modelle (Le Novère *et al.* 2006) hat sieben Einträge für Glykolysemodelle, wobei die Anzahl der modellierten internen Reaktionen von 10 bis 14 variiert (Stand: November 2006). Die Glukoseaufnahme und -verstoffwechslung spielt bei sehr vielen Organismen eine große Rolle. In der Vergangenheit hat man sich besonders für die Glykolyse der Hefe interessiert, da nur noch wenige Schritte daran gekoppelt werden müssen, um die alkoholische Gärung zu modellieren (Teusink *et al.* 2000).

Das Ziel der hier durchgeführten Modellierung ist die Glykolyse des *Corynebacterium glutamicum*. Das Modell besteht aus zehn enzymkatalysierten Reaktionen. Es startet mit der Phosphorylierung von Glukose durch die Glucokinase. Der letzte Schritt ist die Umsetzung von Phosphoenolpyruvat zu Pyruvat durch die Pyruvatkinase. Da die Rate der Glukoseaufnahme bekannt ist, wird eine artifizielle Reaktion mit konstanter Zuflussrate für Glukose in das Modell eingefügt. Die Konzentrationen der Metabolite, die nicht an der Umsetzung des Kohlenstoffgerüsts von Glukose beteiligt sind, werden konstant gehalten. Die Metabolite sind in Abbildung 17 kreisförmig dargestellt. Es bleiben zehn freie Konzentrationsvariablen übrig, deren zugehörige Metabolite oval dargestellt sind. Die anfänglichen Parameterwerte sind in Anhang 6.1 zusammengetragen. Die kinetischen Konstanten wurden so geschätzt, dass die Konstanten der Hinreaktion eine ähnliche Größenordnung haben. Nur so konnte gewährleistet werden, dass überhaupt ein Fließgleichgewicht erreicht werden kann. Die kinetische Konstante der Rückreaktion wurde aus der Gleichgewichtskonstante berechnet. Die Werte der fixierten Konzentrationen wurden teilweise aus der Literatur übernommen (Anhang 6.1).

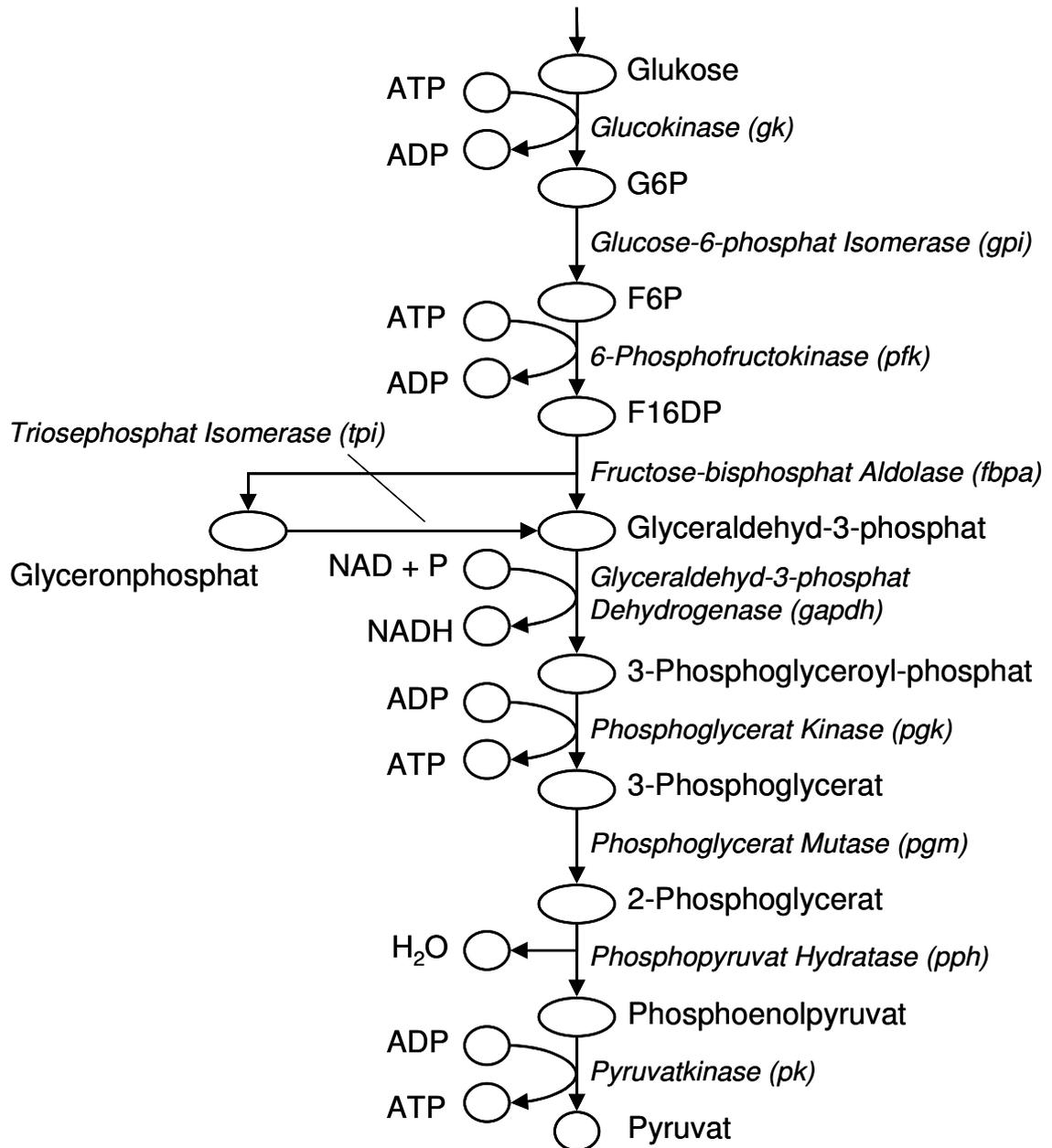


Abbildung 17 Das Modell der Glykolyse. Metabolite, deren Konzentrationen im Simulationsverlauf fixiert sind, sind kreisförmig dargestellt, die freien Konzentrationsvariablen oval.

3.3.5.2 Modell des Citratcyclus

Die Erstellung des Modells gelang analog zur Erstellung des Glykolyse-Modells. Die beiden Modelle unterscheiden sich aber in zwei wesentlichen Punkten: dies ist ein zyklisches Modell und es gibt mehrere Zu- und Abflüsse. Der größte Zufluss stammt aus dem Pyruvat der Glykolyse, weshalb die Zuflussrate der Abflussrate aus dem Glykolyse-Modell entspricht. Außerdem bildet Fumarat einen weiteren Zufluss. Bei verschiedenen anabolen Prozessen entsteht Fumarat als Nebenprodukt. Die Metabolite des Citratcyclus sind Ausgangspunkt für viele Biosynthesepfade. Ausgehend von 2-Oxoglutarat beginnt die Synthese von Glutamat, Glutamin, Prolin und Arginin. Oxalacetat wird unter anderem für die Synthese von Pyrimidinbasen, Aspartat, Asparagin, Threonin, Lysin, Methionin und Isoleucin benötigt. Bei Succinyl-CoA startet die Biosynthese von Porphyrinen. Die entsprechenden Abflüsse wurden im Modell eingerichtet (Abbildung 18).

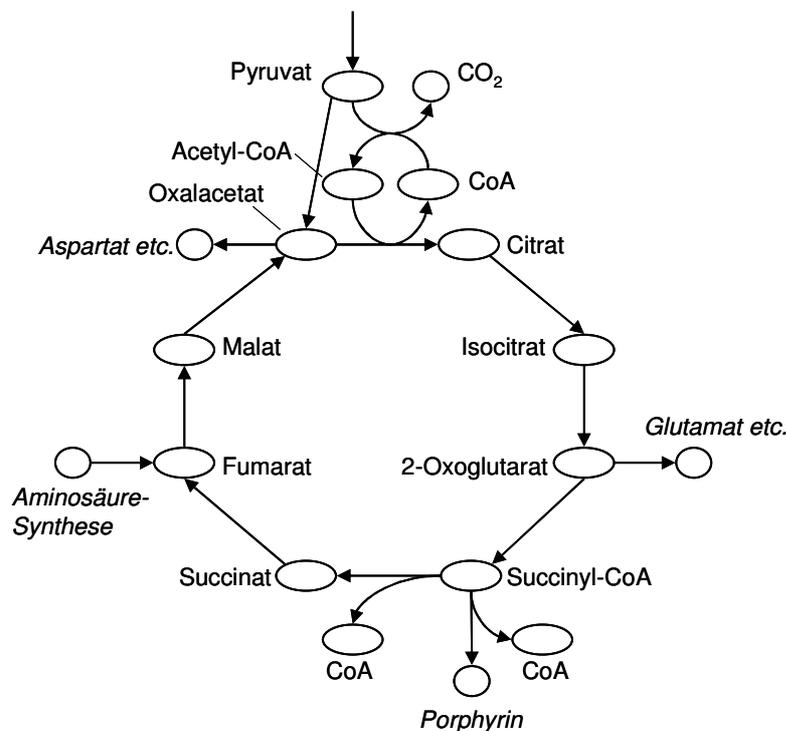


Abbildung 18 Schematische Darstellung des Citratcyclus-Modells. Freie Konzentrationen sind durch Ovale repräsentiert. Metabolite, die einen Zu- oder Abfluss bilden, sind kreisförmig gezeichnet. Aus Gründen der Übersichtlichkeit werden Enzymnamen und Nebenmetabolite nicht gezeigt.

3.3.5.3 Modell des Zentralstoffwechsels

Das Modell des Zentralstoffwechsels wurde erstellt, indem die Modelle Glykolyse und Citratcyclus vereinigt wurden. Ergänzt wurden die Reaktionen des Pentosephosphatwegs,

der bei *C. glutamicum* für Biosyntheseprozesse und zur Regeneration des Reduktionsmittels NADPH stark aktiv ist.

Zunächst wurden die Zu- und Abflüsse aus den vereinigten Modellen verwendet, um die kinetischen Konstanten so anzupassen, dass die Flussrichtung in allen Pfaden sinnvoll wurde. Für den Pentosephosphatweg wurde ein zusätzlicher Abfluss von 5-Phosphoribosyl-diphosphat eingerichtet. Dies ist ein zentraler Metabolit für die Biosynthese von Aminosäuren, Nucleotiden und Bestandteilen der Zellhülle. Die Modellparameter finden sich in Anhang 6.4.

Um zu testen, ob das Modell auch bei realistischen Zu- und Abflüssen Ergebnisse liefert, wurden diese aus der Biomassezusammensetzung berechnet. Grundlage bildet die Biomasse, wie sie in (Fallisch 2006) definiert wurde. Aus dieser Zusammensetzung lässt sich zurückrechnen, wie hoch der Verbrauch an Metaboliten des Zentralstoffwechsels ist, wenn man für jeden Metabolit der Biomasse die Synthesebausteine, die im Modell vorkommen, bestimmt. Diese Zusammenstellung und die verwendete Biomassegleichung befinden sich in Anhang 6.4. Es ergeben sich die Zu- und Abflüsse laut Tabelle 6. Dabei wurde zugrunde gelegt, dass 35,4 % des einfließenden Kohlenstoffs in Biomasse umgewandelt werden (Wendisch *et al.* 2000). Es ergibt sich für diese Biomassezusammensetzung ein Verhältnis von Glukose-Zufluss zu Biomasse-Abfluss von 1 zu 0,0584 (Fallisch 2006). Der Glukose-Zufluss wird, wie beim Modell der Glykolyse, auf 0,771 mM/s fixiert. Das Modell besteht nach diesen Veränderungen aus 46 Reaktionen, wovon 15 extern sind.

Tabelle 6 Zu- und Abflussraten des Zentralstoffwechsel-Modells bei fixierter Biomasseproduktion.

Metabolit	Zu-/Abflussrate [mM/s]
Glukose	7,71E-1
5-Phosphoribosyl diphosphat	-5,0389E-2
2-Oxoglutarat	-4,8039E-2
3-Phosphoglycerat	-6,8530E-2
3-Phosphoglyceraldehyd	2,2513E-3
Erythrose-4-phosphat	-1,5759E-2
Fruktose-6-phosphat	-3,1924E-3
Fumarat	2,1838E-2
Glukose-6-phosphat	-7,6410E-3
Oxalacetat	-1,1152E-1
Phosphoenolpyruvat	-3,3819E-2
Pyruvat	-1,1134E-1
Ribulose-5-phosphat	-1,0581E-3
Seduheptulose-7-phosphat	-1,0581E-3
Glyceronphosphat	-5,8084E-3
Acetyl-CoA	-1,5143E-1
CoA	1,5143E-1

3.4 Alternative Methoden der Modellierung

3.4.1 Farbige Petrinetze

Die Simulation der farbigen Petrinetze wird mit dem selbst erstellten Programm MPN durchgeführt (siehe Abschnitt 3.5.1.1). Die Grundlagen zum Ablauf einer Simulation mit farbigen Petrinetzen sind in Abschnitt 2.4.2 diskutiert. Hier werden Simulationsablauf und Implementierungsdetails erläutert.

Dem Programm liegt ein objektorientiertes Design zugrunde. Die Stellen, Transitionen und Kanten des Petrinetzes, wie auch das Petrinetz selbst, sind einzelne Klassen. Die Simulation wird in der Klasse Simulator durchgeführt. Der Simulator kann mit verschiedenen Codefragmenten ausgeführt werden, die die Ratengleichungen für den Umsatz der einzelnen Reaktionen enthalten. Die Ratengleichungen können mit einer global einstellbaren Schrittweite skaliert werden. Diese Schrittweite hat Einfluss auf die Genauigkeit und Geschwindigkeit der Simulation.

Die Simulation läuft in Zyklen ab, wobei in einem Zyklus jede Transition (Reaktion) ein Mal ausgeführt wird. Die Reaktionen werden vor jedem Zyklus neu gemischt, so dass sie in zufälliger Reihenfolge durchlaufen werden. Das Codesegment, das mit dem Simulator verknüpft wurde, wird für jede Reaktion ausgeführt. Die Konzentrationen der Reaktionspartner werden direkt aktualisiert.

Vor und nach jedem Reaktionsereignis sowie vor und nach jedem Zyklus können vom Simulator beliebig viele weitere Codeabschnitte ausgeführt werden, die vorher registriert wurden. So wurden die Ausgabe der Ergebnisse nach jedem Zyklus oder Veränderungen wie Zu- und Abflüsse realisiert.

Die Ausgabe der Simulationsergebnisse erfolgt in eine Textdatei oder wahlweise on-the-fly über das Programm Grace (Abschnitt 3.5.2.2). Die Textausgabe kann auch nachträglich durch Grace angezeigt, in eine Tabellenkalkulation eingelesen oder anderweitig weiterverarbeitet werden.

3.4.2 Lösung des algebraischen Gleichungssystems

Wie wir in Abschnitt 2.5.1 gesehen haben, ist die Berechnung der Konzentrationen im Fließgleichgewicht prinzipiell auch durch die Lösung eines nicht-linearen algebraischen Gleichungssystems möglich. Weil wir ausschließlich chemische Kinetik verwenden, handelt es sich bei der rechten Seite der Gleichungen um Polynome:

$$0 = \sum_{i_1, \dots, i_n} a_{i_1, \dots, i_n} c_1^{i_1} \dots c_n^{i_n} \quad (\text{Gleichung 23})$$

Es bietet sich an, ein für diese Form der Gleichungen spezialisiertes Lösungsverfahren zu wählen. Das Computer-Algebra-System MuPAD (siehe Abschnitt 3.5.2.9) bietet eine Funktion an, die die gemeinsamen Nullstellen von Polynomen berechnet: `numeric::polysysroots`. Um die Gleichungen für die vorhandenen Modellsysteme automatisch zu erstellen, wurde das Programm `createMupadInput` erstellt, das Teil der Skriptsammlung von MPN (siehe Punkt 3.5.1.1) ist. Das Programm liest die Dateien im MPN-eigenen Format ein und erzeugt eine Datei, die in MuPAD eingelesen und direkt ausgeführt werden kann. Als Ergebnis erhält man entweder eine leere Lösungsmenge, eine bzw. mehrere explizite Lösungen oder die Lösungsmenge in symbolischer Ausgabe. Getestet wurde das Glykolyse-Modell aus Abschnitt 3.3.5.1 und das Modell des Zentralstoffwechsels (Punkt 3.3.5.3).

3.5 Software und Datenbanken

3.5.1 Eigene Implementierungen

Alle hier beschriebenen Programme und Bibliotheken sind unter der Open Source Lizenz GNU Lesser General Public License (GNU LGPL) lizenziert. Diese Lizenz ermöglicht die Einbindung in freie und proprietäre Software. Die Programme sind zur Drucklegung unter <http://biotools.uni-koeln.de/> erhältlich.

3.5.1.1 Metabolic Petri Net

Das Programmpaket MPN, ursprünglich die Abkürzung für Metabolic Petri Net, führt Simulationen von metabolischen Netzwerken durch. Entgegen dem Namen basiert die Simulation nicht nur auf Petrinetzen, sondern es ist auch die Lösung von Differentialgleichungssystemen und Algebrodifferentialgleichungssystemen möglich. Alle Methoden verwenden das gleiche Datenformat für die Eingabeparameter. Die Einleseroutinen greifen teilweise auf die Bibliothek CMP zurück (siehe Punkt 3.5.1.2). Die Ausgabe der Netzwerkstruktur ist auch in SBML (vergleiche Abschnitt 3.5.2.4) möglich. Das Programm wurde in C++ implementiert.

Die Simulation der Petrinetze wurde ohne Verwendung weiterer Bibliotheken implementiert. Für weitere Einzelheiten sei hier auf Abschnitt 3.4.1 verwiesen. Die Funktionalität zur Integration von (Algebro-)Differentialgleichungssystemen wird von der Bibliothek Sundials (Abschnitt 3.5.2.1) bereitgestellt. Für die einzelnen Komponenten wurden Wrapperklassen geschrieben. Für verschiedene Ratengleichungen wurden einzelne Funktionen implementiert. Die Ausgabe der Simulationsergebnisse erfolgt analog zur der Petrinetz-Simulation (siehe Abschnitt 3.4.1).

Eine zusätzliche Visualisierung der Ergebnisse ist über ein Plugin für Cupe (Abschnitt 3.5.2.3) möglich. Dieses Plugin liegt zurzeit als funktionsfähiger Prototyp vor. Die Ergebnisse, also Flüsse und Konzentrationen, werden nicht als Zahlen, sondern kodiert in Farbe und Größe der Komponenten angezeigt. Viele Simulationsparameter (Startkonzentrationen, Zu- und Abflüsse etc.) können über die graphische Schnittstelle eingestellt werden.

3.5.1.2 CUBIC Metabolome Project

Das CUBIC Metabolome Project (CMP) ist eine C++ Bibliothek, die Funktionalität zur Behandlung von metabolischen Netzwerken zur Verfügung stellt. Die Bibliothek wurde gemeinsam mit T. Fallisch entwickelt und wurde bereits in (Fallisch 2006) beschrieben. Neben den notwendigen Datenstrukturen, die die Komponenten abbilden (Metabolisches Netzwerk, Reaktionen, Metabolite, Atome) erlaubt sie unter anderem, Netzwerke in

verschiedenen Formaten ein- und auszulesen, Reaktionen stöchiometrisch zu bilanzieren und eine Flux Balance Analysis durchzuführen. Für einen Teil der Funktionalität werden die Bibliotheken libSBML (Punkt 3.5.2.4) und Ipsolve (Abschnitt 3.5.2.8) benötigt.

3.5.1.3 Gibbspredictor

Mit dem Gibbspredictor ist es möglich, die freien Standardbildungsenthalpien von Metaboliten und die Standard Gibbsenergien von Reaktionen in wässriger Lösung abzuschätzen. Es wurde ein Verfahren in Java implementiert, das in (Mavrovouniotis 1990, 1991) beschrieben wurde. Das Programm verwendet die Bibliotheken Chemistry Development Kit (CDK, Punkt 3.5.2.6) und JChem (Abschnitt 3.5.2.7).

Als Eingabe wird nur eine Datei benötigt, die die Struktur des Metaboliten bzw. die Metabolitstrukturen aller Reaktionspartner inklusive Stöchiometrie benötigt. Die Metabolite und Reaktionen können in allen Dateiformaten vorliegen, die durch das Chemistry Development Kit eingelesen werden können. Als Ausgabe erhält man die Gibbsenergie des Metaboliten bzw. der Reaktion in kcal/mol.

Dieser automatisierte Prozess ist nur möglich, weil die ursprüngliche Methode durch die Standardisierung der Metabolite auf pH 7 und 298,15 K ergänzt wurde. Zur Korrektur des Ionisierungsgrades wird das pKa Plugin von JChem benötigt, um die Position und den pK_a-Wert der aciden Protonen zu bestimmen.

Zusätzlich kann eine Korrektur der Gibbsenergie vorgenommen werden, wenn die Ionenstärke der Lösung bekannt ist. Die Korrektur erfolgt im Rahmen der erweiterten Debye-Hückel-Regel und berücksichtigt die unterschiedlich ionisierten Spezies des Metaboliten bei einem pH-Wert von 7.

3.5.2 Eingesetzte Software

3.5.2.1 Sundials

Mit der Suite Sundials (**SU**ite of **N**onlinear and **D**ifferential/**AL**gebraic equation **S**olvers) können nichtlineare Gleichungssysteme gelöst werden (Hindmarsh *et al.* 2005). Die Suite besteht aus vier einzelnen Bibliotheken zur Lösung von gewöhnlichen Differentialgleichungen (cvode), von gewöhnlichen Differentialgleichungen mit Sensitivitätsanalyse (cvodes), von Algebrodifferentialgleichungen (ida) und von algebraischen Gleichungen (kinsol). Alle Bibliotheken sind in ANSI-C implementiert. In jeder Bibliothek stehen mehrere Algorithmen zur Auswahl. Zur Lösung von gewöhnlichen Differentialgleichungen hat man beispielsweise die Wahl zwischen Algorithmen, die auf dicht besetzte Matrizen, Bandmatrizen oder dünn besetzte Matrizen optimiert sind. Der modulare Aufbau der einzelnen Bibliotheken erleichtert die Inkludierung in andere

Programme. Verwendet wurde die Version 2.2.0. Die Bibliotheken sind unter einer Open Source Lizenz erhältlich (LGPL).

3.5.2.2 Grace

Grace ist ein Programm zur zweidimensionalen Darstellung von numerischen Daten. Es läuft unter praktisch allen Unix-Varianten. Portierungen existieren für Windows 95-XP. Die Daten, die dargestellt werden sollen, können in verschiedenen Dateiformaten übergeben werden. Außerdem gibt es eine Schnittstelle in C, über die andere Programme auf Grace zugreifen können. Verwendet wurde die Version 2.5.19. Das Programm ist unter einer Open Source Lizenz (GPL) erhältlich unter <http://plasma-gate.weizmann.ac.il/Grace/>.

3.5.2.3 Cupe

Der CUBIC Pathway Editor (Cupe) wurde innerhalb der Arbeitsgruppe von Prof. Schomburg durch R. Schunk in der Programmiersprache C++ entwickelt. Der Editor ermöglicht neben dem Zeichnen von Reaktionsnetzwerken auch die Erstellung automatischer Layouts. Dafür werden unter anderem hierarchische, kräftebasierte und Planarisierungsalgorithmen benutzt. Eine Anbindung an eine Hauseigene Datenbank mit Reaktionen und Molekülstrukturen, die auf den Datenbanken BRENDA (Abschnitt 3.5.3.1) und KEGG (Abschnitt 3.5.3.2) basiert, ermöglicht die schnelle Organismus- und Pfad-spezifische Auswahl eines Reaktionsdatensatzes. Durch die Möglichkeit, Graphen hierarchisch zu clustern, gelingt eine informationsreduzierte und übersichtlichere Darstellung. Über eine Plugin-Schnittstelle ist es möglich, eigene Erweiterungen zu schreiben. Verwendet wurde die Version 1.7.0. Das Programm ist unter einer Open Source Lizenz (GPL) vom Autor erhältlich.

3.5.2.4 SBML

Die Systems Biology Markup Language (SBML) ist ein Dateiformat für Modelle biochemischer Netzwerke (Hucka *et al.* 2003). Es ist das international am weitesten verbreitete Format für metabolische Netzwerke, Signaltransduktionsnetzwerke und regulative Netzwerke. Über 100 Softwareprodukte unterstützen das Format, so dass der Austausch von Modellen zwischen verschiedenen Programmen möglich wird. Das Format wird ständig weiterentwickelt. Level 2 Version 2 ist aktuell. Zurzeit wird über die dritte Version, SBML Level 3, diskutiert. Als XML-basiertes Dateiformat ist SBML leicht erweiterbar. So erweitern z. B. Layout-Extensions zur Speicherung des graphischen Layouts von Netzwerken die Funktionalität (Deckard *et al.* 2006; Gauges *et al.* 2006). Zusätzlich zur Spezifikation des Dateiformats werden Bibliotheken in verschiedenen Programmiersprachen angeboten (unter anderem für Java und C), die die Erzeugung und das Einlesen entsprechender Dateien erleichtern.

3.5.2.5 Copasi

Der Complex Pathway Simulator (Copasi) ist ein Programm zur Analyse und Simulation von biochemischen Netzwerken (Hoops *et al.* 2006). Es läuft unter Linux, Sun Solaris, Windows und Mac OS X. Copasi bietet eine Reihe von Funktionen wie stochastische und deterministische Simulation, Berechnung von Fließgleichgewichten, Metabolic Control Analysis und die Berechnung von Elementarmoden und Conserved Moieties. Weiterhin gibt es die Möglichkeit, Parameterbereiche zu scannen bzw. Parameter mit verschiedenen Algorithmen zu optimieren. Man hat die Auswahl aus einer Reihe von Ratengleichungen, die den Reaktionen einzeln zugeordnet werden können. Verwendet wurde die Version 4 (Build 18). Das Programm ist für akademische Zwecke frei verfügbar. Der Source Code ist ebenfalls verfügbar.

3.5.2.6 Chemistry Development Kit

Das Chemistry Development Kit (CDK) ist eine Java-Bibliothek für strukturelle Chemo- und Bioinformatik (Steinbeck *et al.* 2003). Die Bibliothek stellt Datenstrukturen für Moleküle, Atome, Bindungen, Reaktionen, Polymere und vieles andere mehr bereit. Zu den Funktionen zählen der Import- und Export von strukturellen Dateien in über 20 Formaten, die Aromatizitätsdetektion, das Layout und Rendering von zwei- und dreidimensionalen Strukturen, die Berechnung von QSAR-Deskriptoren (Steinbeck *et al.* 2006), die Bewertung der Ähnlichkeit von Strukturen und die Isomorphismus- und Substruktursuche. Eine Reihe weiterer Programme verwenden das Chemistry Development Kit, unter anderem der Editor für zweidimensionale Molekülstrukturen JChemPaint, das Programm für die Darstellung von dreidimensionalen Strukturen Jmol oder die Datenbank für NMR Spektren NMRShiftDB. Verwendet wurde die CDK Version 20060714. Die Bibliothek ist unter einer Open Source Lizenz (LGPL) erhältlich.

3.5.2.7 JChem

JChem ist eine Java-Bibliothek für strukturelle Chemoinformatik, die von der Firma ChemAxon entwickelt wird (<http://www.chemaxon.com/>). Neben der Manipulation chemischer Strukturen und der Substruktur- und Ähnlichkeitssuche, die zum Kern von JChem gehören, wird ein großer Teil der Funktionalität über Plugins realisiert. Dazu gehören die Elementaranalyse, Berechnung von pK_a -Werten, Verteilungskoeffizienten, partiellen Ladungen, Isomeren, Moleküloberflächen oder von Wasserstoff-Donoren und Akzeptoren. Bestandteil von JChem sind außerdem die Programme MarvinSketch zum Zeichnen von Strukturen und MarvinView zur Anzeige, die beide auch als Java-Applet weit verbreitet sind. Verwendet wurde die Version 3.1.6. Für die Bibliothek gibt es eine zeitlich begrenzte, aber kostenlose akademische Lizenz.

3.5.2.8 Lpsolve

Lpsolve ist eine C-Bibliothek für lineare und ganzzahlige lineare Optimierung (<http://lpsolve.sf.net/>). Die lineare Optimierung benutzt den verbesserten Simplex-Algorithmus. Die ganzzahlige Optimierung wendet den Brach-and-Bound-Algorithmus an. Verwendet wurde lpsolve in der Version 5.5.0.8. Die Bibliothek ist unter einer Open Source Lizenz (LGPL) verfügbar.

3.5.2.9 MuPAD

MuPAD ist ein Computer-Algebra-System (<http://www.mupad.de/>). Mit dem Programm sind die Manipulation von symbolischen Formeln und das Rechnen mit exakten und symbolischen Größen möglich. MuPAD erlaubt numerische Rechnungen mit beliebiger Genauigkeit. Zum Funktionsumfang gehören Berechnungen in den Bereichen lineare Algebra, lineare Optimierung, Differentialgleichungen, Zahlentheorie, Statistik und anderen. Das Programm wird durch die Firma SciFace vertrieben und ist kommerziell. Verwendet wurde Version 4.0 pro.

3.5.3 Datenbanken

3.5.3.1 BRENDA

Die Braunschweiger Enzymdatenbank (BRENDA) enthält umfassende Informationen zu Enzymen (Schomburg *et al.* 2004). Zu jedem der Enzyme sind neben der katalysierten Reaktion und Informationen zu funktionellen Enzymparametern über 50 weitere Datenfelder aufgenommen. Darunter sind Datenfelder zu Fragen der Nomenklatur, der Aufreinigung, der Proteinstabilität, der Proteinstruktur, Organismus-spezifische Daten und relevante Informationen zu Krankheiten. Das Sammeln der Informationen geschieht durch manuelle Extraktion aus der Literatur, wobei jeder Datensatz durch einen Biologen oder Chemiker kontrolliert wird. Die entsprechenden Literaturstellen sind Bestandteil der Datensätze. Neben der manuellen Extraktion gibt es optional zusätzliche Daten, die mittels Datamining gewonnen wurden. Die Datenbank ist über ein Webinterface zugänglich, das ausführliche Suchoperationen von der einfachen Suche nach EC-Nummer oder Enzymname über die Suche nach Substrukturen, Sequenzen oder im Genom bis hin zur detaillierten Suche nach einzelnen Werten erlaubt. BRENDA ist über das Webinterface für nicht-kommerzielle Zwecke frei verfügbar. Für kommerzielle Nutzer gibt es die Möglichkeit, die Datenbank zu erwerben.

3.5.3.2 KEGG

Die Kyoto Encyclopedia of Genes and Genomes (KEGG) beschäftigt sich mit dem Genom und der Funktion von Genen (Kanehisa *et al.* 2006). Das System besteht aus vier

Komponenten, der Datenbank über metabolische Pfade (Pathway), der Gendatenbank (Genes), der Datenbank über Reaktionen und Metabolite (Ligand) und der KEGG Ontologie (Brite). Den hohen Bekanntheitsgrad hat die Pathway-Datenbank durch die über 300 von Hand gezeichneten Übersichtskarten von metabolischen Pfaden erlangt. In der Genes-Datenbank sind die annotierten Gene von ca. 350 Organismen enthalten. Die vier Datenbanken sind untereinander verknüpft und über ein Webinterface zugänglich. Für akademische Zwecke sind die Daten auch in Dateiform oder über Web Service abrufbar.

3.5.3.3 MetaCyc

In der Datenbank MetaCyc sind Informationen zu metabolischen Pfaden abgelegt (Caspi *et al.* 2006). Die Daten über Reaktionen und Moleküle werden aus experimenteller Literatur extrahiert und von Biochemikern und Molekularbiologen kontrolliert. Die Datenbank enthält Informationen über Aktivatoren, Inhibitoren, Substratspezifität und kinetische Konstanten. Die meisten Pfade der über 600 Organismen stammen aus Mikroorganismen und Pflanzen. Über die Software Pathway Tools werden die Pfade visualisiert und abgefragt. Diese Software ist als eigenständige Desktop-Anwendung und über ein Webinterface zugänglich.

3.5.3.4 NIST Datenbank zur Thermodynamik biochemischer Reaktionen

Die NIST Standard Reference Database 74 vom National Institute of Standards and Technology (NIST) besteht aus thermodynamischen Daten von enzymkatalysierten Reaktionen (Goldberg *et al.* 2004). Die enthaltenen apparenten Gleichgewichtskonstanten und kalorimetrisch bestimmten Reaktionsenthalpien sind experimentelle Daten, die aus ca. 1000 Literaturstellen extrahiert wurden. Die Datensätze umfassen ca. 400 Reaktionen. Auf die Datenbank kann über ein Webinterface zugegriffen werden.

4 Ergebnisse und Diskussion

4.1 Reaktionsbilanzierung

4.1.1 Ergebnisse der Reaktionsbilanzierung

Nach der Auswahl der Reaktionen, die im Modell berücksichtigt werden sollen, ist die Reaktionsbilanzierung der zweite Schritt im Modellierungsprozess. Dieser Schritt verfolgt zwei Ziele. Es werden diejenigen Reaktionen identifiziert, die keine korrekte Massen- und Ladungsbilanz aufweisen. Anschließend werden diese Reaktionen – soweit möglich – automatisch bilanziert. Ist dies nicht möglich, müssen sie manuell korrigiert werden. Das allgemeine Ablaufschema ist im Abschnitt Methoden abgebildet (Abbildung 10, Seite 32).

Als Testdatensatz für die Bilanzierung wurden die Reaktionen aus der Datenbank KEGG Ligand verwendet (Kanehisa *et al.* 2006). Die Datenbank enthält insgesamt 6580 Reaktionseinträge, wovon 5619 ein oder mehrere Enzyme zugeordnet sind. Für die Bilanzierung kamen nur die Reaktionen in Frage, bei denen für alle Metabolite die Strukturdateien und somit die Summenformeln vorhanden waren (Tabelle 7).

Tabelle 7 Reaktionseinträge in der Datenbank KEGG Ligand und Angaben über die Korrekturversuche.

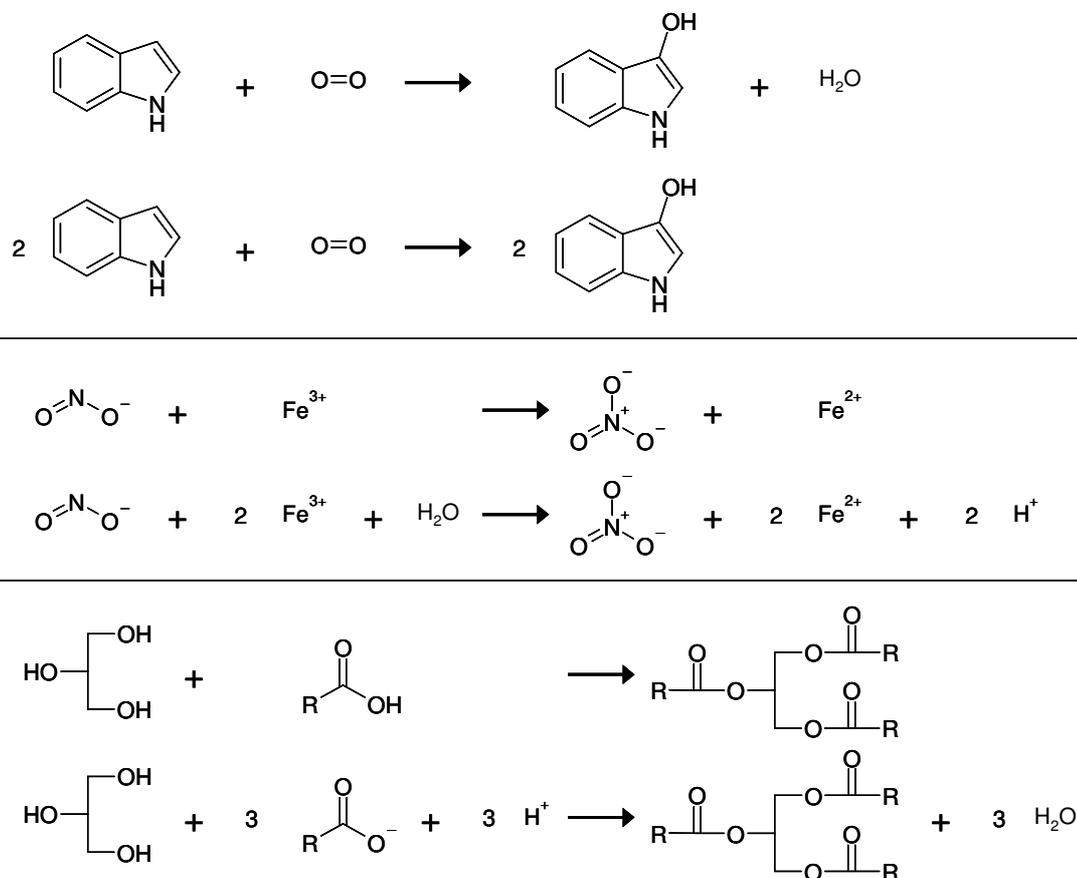
Reaktionseinträge	6811
Enzymkatalysierte Reaktionen	5619
Reaktionen mit Strukturdateien, davon	5757 [= 100 %]
... korrekte Reaktionen	2858 [= 49,6 %]
... mit Protonen korrigierbare Reaktionen	2038 [= 35,4 %]
... sonstige korrigierbare Reaktionen	473 [= 8,2 %]
... nicht korrigierbare Reaktionen	388 [= 6,7 %]

Um den Erfolg der Bilanzierung zu bewerten, richten wir den Blick auf die sonstigen korrigierbaren Reaktionen. Diese Gruppe lässt sich weiter aufteilen in Reaktionen mit falscher Redox-Bilanz, bei denen das Oxidationsmittel Sauerstoff zur Reaktion hinzugefügt werden musste, und Reaktionen mit korrekter Redox-Bilanz. Die letztgenannte Gruppe ist besonders interessant, da sich die Stöchiometrie der ursprünglichen Metabolite verändert. Hierfür wurden die automatisch korrigierten Reaktionen einzeln manuell auf Richtigkeit geprüft (Tabelle 8).

Tabelle 8 Statistik über die sonstigen korrigierbaren Reaktionen (vergleiche Tabelle 7).

Korrigierbare Reaktionen mit falscher Redoxbilanz	329
Korrigierbare Reaktionen mit richtiger Redoxbilanz, davon	144
... nur durch Wasser- und Protonenausgleich	52
... mit Veränderung der ursprünglichen Stöchiometrie, davon	92 [= 100 %]
... Strukturdatei falsch / falsch eingelesen	14 [= 15,2 %]
... richtig korrigiert	66 [= 71,7 %]
... falsch korrigiert	12 [= 13,0 %]

Bei den falschen bzw. falsch eingelesenen Strukturdateien trifft man auf drei Probleme. Es kann sich um Polymere handeln, die von der Bibliothek CDK (Steinbeck *et al.* 2003) nicht korrekt eingelesen werden können, die Strukturdateien können schlicht falsch sein oder die Strukturdateien sind unzureichend spezifiziert. Bei den falsch korrigierten Reaktionen lassen sich zwei Fälle unterscheiden. Entweder die Reaktion ist grundsätzlich falsch oder unvollständig oder die Stöchiometrie wurde falsch korrigiert, obwohl mit den vorhandenen Metaboliten die richtige Lösung möglich wäre. Beispiele zu den einzelnen Fällen finden sich in Abbildung 19, Abbildung 20 und Abbildung 21.

**Abbildung 19** Durch das Verfahren korrekt bilanzierte Reaktionen (KEGG Referenz R02339, R01106, R01350). Jeweils oben vor der Korrektur, unten nach der Korrektur.

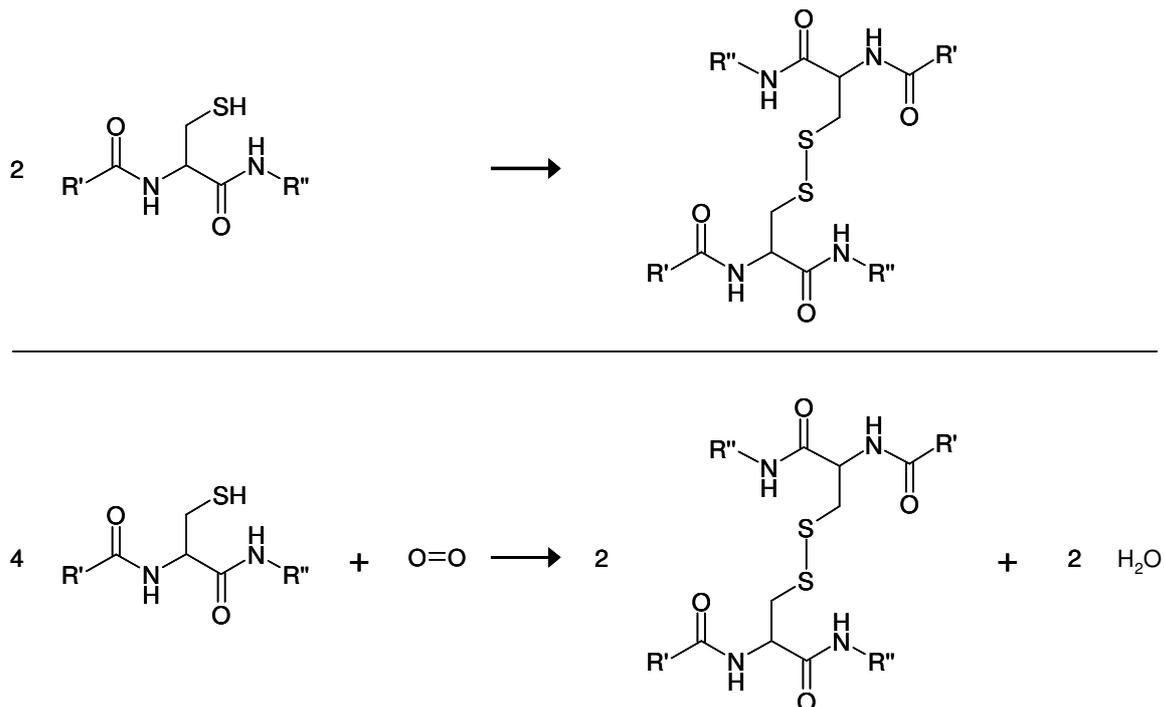


Abbildung 20 Bei Addition von Sauerstoff ist die resultierende Korrektur zwar chemisch richtig, entspricht aber nicht unbedingt der korrekten Reaktion (KEGG Referenz R00500). Oben: ursprüngliche Reaktion, unten: korrigierte Reaktion.

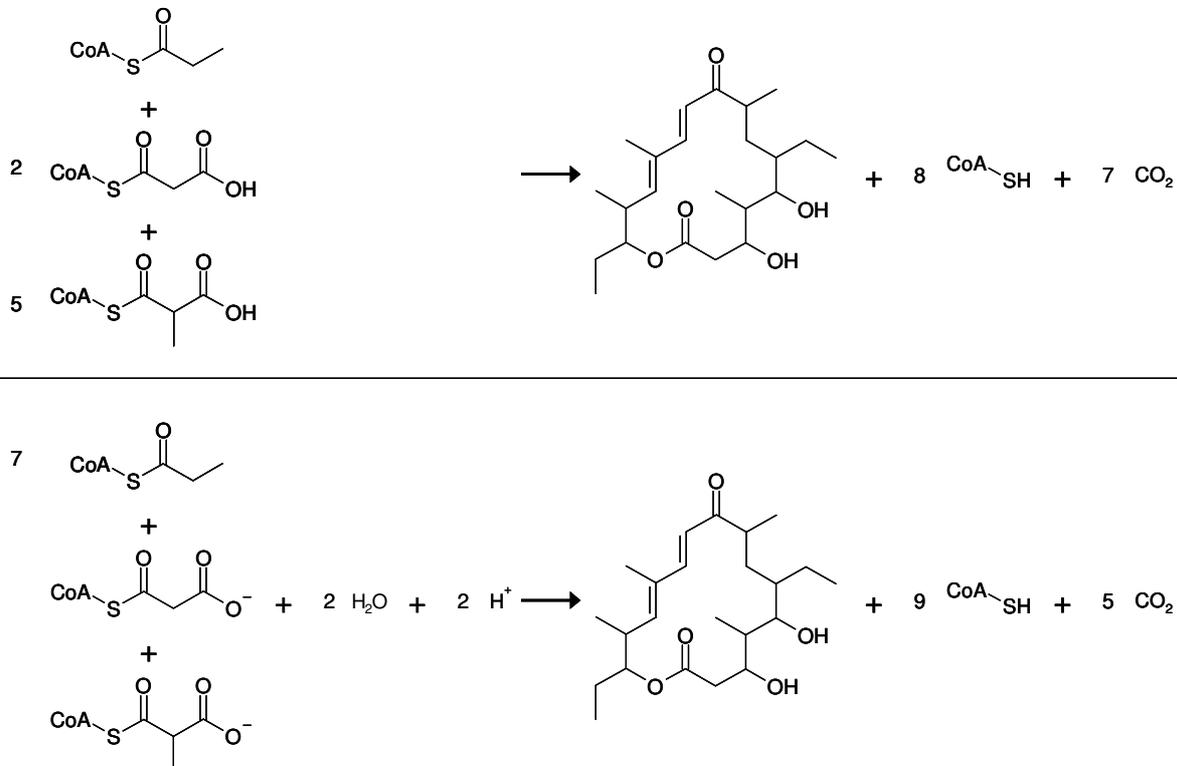


Abbildung 21 Gibt es mehrere Kombinationsmöglichkeiten, findet das Verfahren nicht immer die richtige Stöchiometrie (KEGG Referenz R06448). Oben: ursprüngliche Reaktion, unten: korrigierte Reaktion.

4.1.2 Diskussion zur Reaktionsbilanzierung

Die hier vorgestellte Methode zur Reaktionsbilanzierung erfüllt größtenteils die gesteckten Ziele. Es gelingt eine Auftrennung der Reaktionen in vier verschiedene Gruppen (Tabelle 7 & Tabelle 8), die im Folgenden definiert werden. Die erste Gruppe besteht aus den korrekten bzw. leicht zu korrigierenden Reaktionen. Eine Reaktion ist leicht zu korrigieren, wenn unter Erhalt der ursprünglichen Stöchiometrie Wasser oder Protonen addiert werden. Die zweite Gruppe enthält die Reaktionen, die wissentlich falsch korrigiert wurden, bei denen aber eine Hilfestellung für die manuelle Korrektur gegeben wird. Dies gilt für die Addition von Sauerstoff, der eigentlich nicht an der Reaktion beteiligt ist und als Oxidationsäquivalent angesehen werden muss. In diesem Fall muss angenommen werden, dass reduktive Cosubstrate wie NADH fehlen oder dass ein Fehler bei den Molekülstrukturen vorliegt. Die dritte Gruppe wird von den nicht korrigierbaren Reaktionen gebildet. Hier muss die Reaktion in jedem einzelnen Fall manuell überprüft werden.

In der letzten Gruppe finden sich die Reaktionen, die schwer zu korrigieren sind. Die Stöchiometrie der ursprünglichen Metabolite muss verändert werden, damit die Reaktion bilanziert werden kann. Zusätzlich werden meist Wasser und/oder Protonen addiert. Wie Tabelle 8 zeigt, wird der überwiegende Teil (71,7 %) der Reaktionen in dieser Gruppe richtig korrigiert. 15,2 % der falschen Korrekturen beruhen auf falschen bzw. falsch eingelesenen Strukturen. Wenn man die letzteren vorher ausfiltern würde, ließe sich der Anteil der richtig korrigierten Reaktionen auf bis zu 84,6 % steigern. Zur Identifizierung wäre ein Eingriff in den Einleseprozess notwendig.

Exemplarisch sind einige richtig korrigierte Reaktionen in Abbildung 19 aufgelistet. In allen Fällen wurde die ursprüngliche Stöchiometrie korrigiert. In der ersten Reaktion wurde zusätzlich ein Metabolit entfernt. Dies ist nur zulässig, wenn sich der Metabolit in der Liste der addierten Metabolite befindet. An der zweiten Reaktion ist Eisen in unterschiedlichen Oxidationsstufen beteiligt. Hier sind häufig Fehler zu beobachten, die jedoch ausgeglichen werden können. Das dritte Beispiel zeigt, dass auch unbekannte Reste beim Bilanzieren keine Probleme verursachen.

In Abbildung 20 sehen wir eine typische Reaktion, die unter Sauerstoff-Addition ausgeglichen wurde (R00500). Der Reaktion fehlt ein Oxidationsmittel, das vermutlich absichtlich weggelassen wurde, da es unbekannt ist. Die in KEGG Ligand zu dieser Reaktion gehörende unvollständige EC-Nummer 1.8.4.- deutet ebenfalls darauf hin.

Die Grundvoraussetzungen für eine erfolgreiche automatische Korrektur sind, dass die beteiligten Moleküle eine korrekte Summenformel besitzen und alle an der Reaktion beteiligten Metabolite bekannt sind. Bereits die Addition der Spezies Wasser und Proton ist ein Eingriff, der begründet werden muss. Die Ergebnisse zeigen deutlich, dass diese Metabolite häufig tatsächlich fehlen und die Korrektur erfolgreich ist.

Die Schwierigkeiten der hier vorgestellten Methode sind inhärent. Die Minimierung der Summe der stöchiometrischen Faktoren kann zu falschen Korrekturen führen, wie wir an dem Beispiel in Abbildung 21 sehen. Die Massenbilanz der Ausgangsreaktion ist nicht korrekt, wie man anhand der Anzahl der Sauerstoffatome auf Edukt- und Produktseite leicht nachrechnen kann. Die korrigierte Reaktion ist zwar bezüglich Massen- und Ladungsbilanz korrekt, die stöchiometrischen Koeffizienten weichen jedoch stark von denen der ursprünglichen Reaktion ab. Hierbei gibt es mehrere mögliche Kombinationen der Koeffizienten. Die minimale Summe muss nicht notwendigerweise die richtige Kombination sein. Eine Beschränkung, die nur eine bestimmte Abweichung jedes Koeffizienten vom Ursprungswert zulässt, könnte das Problem wahrscheinlich eindämmen. Eine zweite denkbare falsche Korrektur konnte nicht beobachtet werden. Es ist möglich, dass zwei für sich genommen vollständige Reaktionen aneinander gekoppelt sind. Ein prominentes Beispiel ist die Phosphorylierung von ADP, die durch die protonentranslozierende ATP-Synthase katalysiert wird. Der Protonentransport wird an die ATP-Synthase gekoppelt – zwei Prozesse, die auch unabhängig voneinander ablaufen könnten, wenn man den energetischen Aspekt unberücksichtigt lässt. Ein experimentell bestimmtes Verhältnis von H^+ zu ATP ist 4:1 (Turina *et al.* 2003). Die Koeffizientenminimierung würde in diesem Fall das Verhältnis fälschlicherweise auf 1:1 reduzieren.

Die Validierung der Methode erfolgte anhand der Datenbank KEGG Ligand. Sie erlaubt umgekehrt Rückschlüsse über die Qualität der Datenbank. Es gibt eine Reihe von Reaktionen, bei denen nicht für alle Metabolite Strukturdateien vorhanden sind (17,5 %). Einen großen Anteil daran machen generische Metabolite wie „Donor“, „Akzeptor“ oder „Peptid“ aus. Betrachtet man nur die Reaktionen mit kompletter Strukturinformation und lässt fehlende Protonen außer Acht, sind ca. $\frac{3}{4}$ korrekt, wie oben bereits bemerkt wurde. Bei den unkorrekten Reaktionen treten sehr häufig Fehler in der Ladungs- oder Redoxbilanz auf. Das legt die Vermutung nahe, dass bei Erstellung der Reaktionsdatensätze nur auf die Massenbilanz der Schweratome geachtet wurde.

Eine Anwendung auf andere Datensätze könnte weitere Anpassungen im Prozess nötig machen, die auf den jeweiligen Datensatz und dessen typische Fehler zugeschnitten sind. Die Liste der Moleküle, die für die Bilanzierung zur Reaktion addiert werden, kann nach Bedarf verkleinert oder erweitert werden. Der grundsätzliche Mechanismus, die Minimierung der Summe der stöchiometrischen Faktoren (siehe Abschnitt 3.1, Seite 31), bleibt mit den dargestellten Einschränkungen ein wirkungsvoller Korrekturmechanismus.

4.2 Gibbsenergieberechnung

4.2.1 Ergebnisse der Gibbsenergieberechnung

In Abschnitt 3.2.1 wurde gezeigt, wie aus der Standard Gibbsenergie einer Reaktion deren Gleichgewichtskonstante berechnet werden kann. Außerdem wurden zwei verschiedene Methoden besprochen, mit denen sich Standard Gibbsenergien von Reaktionen bestimmen lassen. Beide Methoden ermitteln zunächst die Bildungs-Gibbsenergien der Metabolite. Aus diesen wird durch Aufsummierung unter Berücksichtigung der Stöchiometrie die Standard Gibbsenergie der Reaktion berechnet.

Das Verfahren nach Mavrovouniotis ist inkrementbasiert. Ein Metabolit wird in Gruppen aufgespalten und jeder Gruppe wird ein Inkrement zugeordnet. Die Inkremente werden anschließend aufsummiert. Das Verfahren nach Alberty beruht auf den Bildungs-Gibbsenergien der einzelnen Spezies (z. B. ATP^{4-} , HATP^{3-} , $\text{H}_2\text{ATP}^{2-}$). Aus diesen wird die apparente Bildungs-Gibbsenergie des Metaboliten (z. B. ATP) berechnet. Details zu den Methoden finden sich in den Abschnitten 3.2.2 und 3.2.3.

Zur Bewertung des Spektrums an Reaktionen, die mit den beiden Methoden vorhergesagt werden können, wurde der Datensatz der korrekten und korrigierten Reaktionen aus der Datenbank KEGG Ligand verwendet (vergleiche Abschnitt 3.5.3.2). Der Anteil, der mit der jeweiligen Methode vorhergesagt werden kann, ist in Tabelle 9 zu sehen.

Tabelle 9 Anzahl der Reaktionen in der Datenbank KEGG/Ligand und Anzahl der Reaktionen, für die die Standard Gibbsenergie berechnet werden kann.

	Alle Reaktionen	Reaktionen des Zentralmetabolismus
Insgesamt	5369	105
Mavrovouniotis	3977 [= 74,1 %]	99 [= 94,3 %]
Alberty	344 [= 6,4 %]	39 [= 37,1 %]

Zur Beurteilung der Qualität der Vorhersage werden die Reaktionen verwendet, die durch beide Methoden abgedeckt werden und für die ein entsprechender Eintrag in der NIST Datenbank für thermodynamische Daten enzymkatalysierter Reaktionen (Abschnitt 3.5.3.4) existiert. Insgesamt erfüllen 30 Reaktionen diese Kriterien (Tabelle 9). Für die Methode nach Alberty erhält man einen mittleren Fehler von 4,4 kJ/mol, für die Methode nach Mavrovouniotis einen mittleren Fehler von 19,6 kJ/mol.

Tabelle 10 Experimentelle Daten für Standard Gibbsenergien aus der Datenbank vom NIST (siehe Abschnitt 3.5.3.4) im Vergleich mit berechneten Werten nach Alberty und Mavrovouniotis.

Nr.	Reaktion	$\Delta_R G_{\text{exp}}$	$\Delta_R G_{\text{Alberty}}$	$\Delta_R G_{\text{Mavro}}$
1	ATP + Pyruvat + H ₂ O = AMP + PEP + P	17,5	21,1	5,3
2	Xylitol + NADP = Xylulose + NADPH	20,1	23,5	16,1
3	2 Acetyl-CoA = CoA + Acetoacetyl-CoA	27,4	27,5	36,6
4	Fructose-1,6-diphosphat = Glyceron-P + Glyceraldehyd-3-P	23,4	17,3	25,0
5	ATP = cAMP + PP	6,8	7,7	-40,0
6	Glyceraldehyd-3-P = Glyceron-P	-7,7	-7,7	-1,0
7	Galactose + ATP = Galactose-1-P + ADP	-7,9	-24,4	0,5
8	Pyruvat + ATP + CO ₂ = Oxaloacetat + ADP + P	-0,8	-8,8	-6,6
9	Methylmalat = Methylmaleate + H ₂ O	5,8	5,8	5,6
10	Ribose-5-P = Ribulose-5-P	3,3	8,1	3,4
11	Mannose = Fructose	-2,7	-5,5	-12,0
12	Aspartat = Fumarat + NH ₃	11,4	12,2	2,3
13	Xylose = Xylulose	0,8	4,3	-12,0
14	Mannose-6-P = Mannose-1-P	-0,5	-5,3	12,8
15	Ethanol + NAD = Acetaldehyd + NADH	22,9	22,9	15,7
16	Xylulose = Lyxose	3,6	3,0	12,0
17	Glucose-1-P = Glucose-6-P	-3,2	-7,0	-12,8
18	Oxaloacetat + ITP = PEP + IDP + CO ₂	1,3	1,4	30,8
19	Adenosin + H ₂ O = Adenin + Ribose	-9,8	-6,9	13,1
20	Adenosin + P = Adenin + Ribose-1-P	12,9	17,1	43,7
21	Malat = Maleat + H ₂ O	18,9	13,4	3,6
22	2-Phosphoglycerat = PEP + H ₂ O	-3,4	-3,3	-4,9
23	Isocitrat + NADP = 2-Oxoglutarat + CO ₂ + NADPH	0,4	4,4	-1,5
24	Glucose-6-P = Fructose-6-P	3,0	3,2	-12,0
25	ATP + Acetat + CoA = AMP + PP + Acetyl-CoA	-6,0	-6,0	55,8
26	Serin + Pyruvat = Alanin + Hydroxypyruvat	-3,6	-3,6	0,0
27	Glycerat + NAD = Hydroxypyruvat + NADH	32,6	31,9	11,9
28	Glycerol + NAD = Glyceron + NADH	23,9	25,2	11,3
29	Alanin + 2-Oxoglutarat = Pyruvat + Glutamat	1,0	-3,5	8,8
30	Glutamat + NAD + H ₂ O = 2-Oxoglutarat + NH ₃ + NADH	33,4	37,0	2,3

4.2.2 Diskussion zur Gibbsenergieberechnung

Die Methoden nach Alberty und nach Mavrovouniotis unterscheiden sich in mehreren Punkten. Die grundsätzlichen methodischen Unterschiede wurden im vorherigen Abschnitt kurz beschrieben. Ein weiterer Unterschied sind die Datensätze, auf denen die beiden Methoden aufbauen. Während Alberty den Datensatz vom NIST zur Verfügung hatte (Goldberg *et al.* 2004), hat Mavrovouniotis auf eine Reihe von älteren Datensammlungen zurückgegriffen, unter anderem (Barman 1969; Benson 1976; Edsall und Gutfreund 1983). Inwieweit sich die verwendeten Datensätze überlappen, soll im Rahmen dieser Dissertation nicht beantwortet werden.

Ein Blick auf die Ergebnisse zeigt die Stärken und Schwächen beider Methoden. Mavrovouniotis bietet eine Methode an, die für Metabolite und Reaktionen im biochemischen Kontext breit anwendbar ist. Für ca. $\frac{3}{4}$ der untersuchten Reaktionen ist eine Berechnung möglich. Beschränkt man sich auf die Reaktionen des Zentralmetabolismus, sind fast 95 % berechenbar. Die Methode nach Alberty hingegen liefert für den gesamten Datensatz nur ein Ergebnis für 6,4 % der Reaktionen, was angesichts der sehr beschränkten Anzahl an Metaboliten (199) nicht verwundert. Von einer vollständigen thermodynamischen Beschreibung eines metabolischen Netzwerks ist man hier weit entfernt. Bei der Genauigkeit der Methoden ergibt sich ein umgekehrtes Bild. Aus den Daten in Tabelle 10 lässt sich für die Methode nach Alberty ein mittlerer Fehler von 4,4 kJ/mol pro Reaktion bestimmen. Der mittlere Fehler der anderen Methode ist mit 19,6 kJ/mol über viermal so hoch.

Aus der Fehlerverteilung in (Mavrovouniotis 1990) kann ein mittlerer Fehler von ca. 9,2 kJ/mol pro Metabolit berechnet werden. Leider trifft der Autor keine Aussage über den mittleren Fehler bei Anwendung auf Reaktionen. Abschätzungen aus der Fehlerfortpflanzung auf Basis des mittleren Fehlers der Metaboliten ergeben für die Reaktionen in Tabelle 10 einen mittleren Fehler von 14,0 kJ/mol. Die verbleibende Differenz von 5,6 kJ/mol ist mutmaßlich darauf zurückzuführen, dass die Anwendung auf einen Datensatz erfolgt, der nicht zur Berechnung der Inkremente verwendet wurde. Es ist allgemein bekannt, dass Inkrementmethoden bei Anwendung auf fremde Datensätze häufig schlechtere Ergebnisse liefern. Eine 10-fache Kreuzvalidierung konnte laut (Mavrovouniotis 1990) bei Berechnung der Inkremente nicht vollständig durchgeführt werden, da für einige Inkremente nur eine sehr geringe Anzahl von korrespondierenden Datenpunkten vorhanden war.

Dass die Inkrementmethode nach Mavrovouniotis der direkten Methode nach Alberty bezüglich der Genauigkeit unterlegen ist, hat neben den eben genannten Gründen (Anwendung auf fremden Datensatz, kleiner Ausgangsdatsatz) noch weitere Gründe. Bei der Berechnung der Inkremente ist nicht die tatsächliche Protonendifferenz in Lösung bei gegebenem pH-Wert verwendet worden, sondern die Protonendifferenz der Referenzreaktion mit den vorherrschenden Metaboliten bei pH 7 (vergleiche Abschnitt 3.2.1, Seite 35). Im Besonderen ist letztere immer ganzzahlig, während erstere dies im Regelfall nicht ist. Eine Differenz von 0,5 Protoneneinheiten macht aber bei pH 7 schon eine Gibbsenergie-Differenz von ca. 20 kJ/mol aus. Des Weiteren sind weder lang reichende Wechselwirkungen (Konjugation, Molekülorbitale) noch räumliche Wechselwirkungen durch die Methode erfasst. Durch die Anwendung von Deskriptoren, die diese Wechselwirkungen implizit berücksichtigen, wäre eine Verbesserung des

Inkrementensystems möglich – vorausgesetzt, die Datenlage verbessert sich qualitativ und quantitativ.

Einige Probleme, mit denen beide Methoden konfrontiert sind, sind die fehlende Standardisierung der Messbedingungen (Puffer, ionische Zusätze) und die mangelnde Fehlerbestimmung. Häufig sind Gleichgewichtskonstanten nur ein Abfallprodukt von kinetischen oder anderen Messungen, so dass auf die Vermeidung von systematischen Fehlern bei der Bestimmung der Gleichgewichtskonstante oder auf die Bestimmung des Messfehlers kein besonderer Wert gelegt wird. In der Datenbank vom NIST (vergleiche Abschnitt 3.5.3.4) fehlen solche Angaben völlig.

In (Ould-Moulaye *et al.* 1999) wurde die Verteilung der Spezies von ATP in Lösung mit Magnesium-Ionen genauer untersucht und herausgefunden, dass auch Magnesium-ATP-Addukte einen Einfluss auf die apparente Gibbsenergie haben. Die Autoren schlagen vor, ähnlich dem pH-Wert einen pMg-Wert bei der Messung anzugeben. Dies ist aber nur ein Teilaspekt der generellen Forderung, die Messbedingungen so weit wie möglich zu standardisieren.

Eine weitere mögliche Fehlerquelle der Methode nach Mavrovouniotis ist die Abschätzung des pK_a -Wertes und der daraus resultierende Protonierungsgrad. Laut (Szegezdi und Csizmadia 2003) hat diese Abschätzung einen Korrelationskoeffizienten von 0,98. Der betrachtete Datensatz bestand hauptsächlich aus einfach ionisierten Molekülen, während biochemische Moleküle häufig mehrfach ionisiert sind. Weitere Untersuchungen hierzu sind nicht bekannt.

4.3 Simulation des linearen Modells

4.3.1 Ergebnisse des linearen Modells

Anhand eines linearen Modells mit zwei internen Reaktionen (vergleiche Abbildung 15, Seite 47) wurde zunächst das Verhalten bei verschiedenen Zufluss- und Abflussfunktionen untersucht. Im Fokus der Beobachtung standen bei allen Untersuchungen die Veränderungen der freien Konzentrationsvariablen, der Flussrate und der Koeffizienten der metabolischen Kontrollanalyse im Fließgleichgewicht.

Die vier Kombinationsmöglichkeiten für die Zufluss- und Abflussfunktionen sind:

- Zuflussrate und Abflussrate nach reversibler chemischer Kinetik
- Konstante Zuflussrate & konstante Abflussrate
- Konstante Zuflussrate und Abflussrate nach reversibler chemischer Kinetik
- Zuflussrate nach reversibler chemischer Kinetik und konstante Abflussrate

Die konstanten Zufluss- und Abflussraten müssen bei Punkt b übereinstimmen, sonst kann kein Fließgleichgewicht erreicht werden. Die konstanten Raten werden jeweils auf den gleichen Wert gesetzt, den der Fluss des Referenzsystems von Punkt a hat. Es ergeben sich Unterschiede bei den Konzentrationen und Kontrollkoeffizienten, die in Abbildung 22 aufgelistet sind.

(a)	Input	R1	R2	Output	Konz/Fluss*
A	0.247	-0.082	-0.082	-0.082	0.7525
S	0.245	0.245	-0.245	-0.245	0.505
P	0.240	0.240	0.240	-0.721	0.2575
Fluss	0.250	0.250	0.250	0.250	2.48E-15

(b)	Input	R1	R2	Output	Konz/Fluss*
A	0	0	0	0	0.86118
S	0	0	0	0	0.61368
P	0	0	0	0	0.36618
Fluss	0	0	0	0	2.48E-15

(c)	Input	R1	R2	Output	Konz/Fluss*
A	0.987	-0.329	-0.329	-0.329	0.7525
S	0.980	0	-0.490	-0.490	0.505
P	0.961	0	0	-0.961	0.2575
Fluss	1	0	0	0	2.48E-15

(d)	Input	R1	R2	Output	Konz/Fluss*
A	0.329	0	0	-0.329	0.7525
S	0.490	0.490	0	-0.980	0.505
P	0.961	0.961	0.961	-2.883	0.2575
Fluss	0	0	0	1	2.48E-15

Abbildung 22 Kontrollkoeffizienten-Matrix für das lineare System bei unterschiedlichen Zu- und Abflussmethoden (siehe Text). Zur Deutung der Darstellung siehe Abschnitt 3.3.3 auf Seite 46.

*Konzentrationen in mM, Flüsse in mmol/s.

Im nächsten Versuch wurden die Ratengleichungen der beiden internen Reaktionen durch reversible Michaelis-Menten Kinetiken (Abbildung 23 links) ausgetauscht. Die Michaelis-Konstanten und die Enzymkonzentration wurden für die erste interne Reaktion (R1)

variiert. Zum Vergleich wurden die gleichen Versuche parallel mit chemischer Kinetik durchgeführt (Abbildung 23 rechts).

Von besonderem Interesse für diese Arbeit ist der Einfluss der Gleichgewichtskonstanten auf das Fließgleichgewicht. Die Gleichgewichtskonstanten sind zwar unveränderliche Parameter der Reaktionen, aber die durchgeführten Abschätzungen zur Bestimmung der Gibbsenergie und damit der Gleichgewichtskonstanten sind fehlerbehaftet. Die Resultate bei Verschiebung des Gleichgewichts sind in Abbildung 24 dargestellt.

Michaelis-Menten Kinetik						Chemische Kinetik					
(a)	Input	R1	R2	Output	Konz/Fluss*	(b)	Input	R1	R2	Output	Konz/Fluss*
A	0,267	-0,081	-0,081	-0,105	0,7512	A	0,247	-0,082	-0,082	-0,082	0,7525
S	0,230	0,292	-0,227	-0,295	0,4744	S	0,245	0,245	-0,245	-0,245	0,505
P	0,185	0,235	0,235	-0,656	0,2588	P	0,240	0,240	0,240	-0,721	0,2575
Fluss	0,193	0,245	0,245	0,318	2,49E-15	Fluss	0,250	0,250	0,250	0,250	2,48E-15
(c)	Input	R1	R2	Output	Konz/Fluss*	(d)	Input	R1	R2	Output	Konz/Fluss*
A	0,354	-0,002	-0,163	-0,189	0,6716	A	0,327	-0,002	-0,163	-0,163	0,6711
S	0,353	0,004	-0,165	-0,192	0,6678	S	0,327	0,003	-0,165	-0,165	0,6678
P	0,268	0,003	0,323	-0,595	0,3384	P	0,322	0,003	0,322	-0,648	0,3389
Fluss	0,277	0,003	0,333	0,387	3,28E-15	Fluss	0,332	0,003	0,332	0,332	3,29E-15
(e)	Input	R1	R2	Output	Konz/Fluss*	(f)	Input	R1	R2	Output	Konz/Fluss*
A	0,010	-0,009	0,000	0,000	0,9904	A	0,010	-0,009	0,000	0,000	0,9904
S	0,003	0,582	-0,200	-0,385	0,0246	S	0,006	0,639	-0,323	-0,323	0,0292
P	0,002	0,477	0,004	-0,482	0,0196	P	0,005	0,476	0,005	-0,485	0,0196
Fluss	0,001	0,980	0,002	0,017	9,59E-17	Fluss	0,010	0,971	0,010	0,010	9,61E-17
(g)	Input	R1	R2	Output	Konz/Fluss*	(h)	Input	R1	R2	Output	Konz/Fluss*
A	0,122	-0,090	-0,012	-0,019	0,8795	A	0,124	-0,088	-0,018	-0,018	0,8763
S	0,117	0,690	-0,313	-0,494	0,2113	S	0,120	0,601	-0,360	-0,360	0,2575
P	0,103	0,607	0,083	-0,792	0,1305	P	0,116	0,578	0,116	-0,810	0,1338
Fluss	0,111	0,657	0,090	0,142	1,20E-15	Fluss	0,125	0,625	0,125	0,125	1,24E-15
(i)	Input	R1	R2	Output	Konz/Fluss*	(j)	Input	R1	R2	Output	Konz/Fluss*
A	0,111	-0,077	-0,013	-0,021	0,8907	A	0,109	-0,082	-0,014	-0,014	0,8907
S	0,097	0,653	0,287	-0,463	0,1910	S	0,106	0,640	-0,373	-0,373	0,2285
P	0,085	0,578	0,097	-0,760	0,1193	P	0,101	0,613	0,101	-0,815	0,1193
Fluss	0,093	0,631	0,106	0,170	1,09E-15	Fluss	0,110	0,669	0,110	0,110	1,09E-15

Abbildung 23 Kontrollkoeffizienten-Matrix des linearen Systems bei chemischer und Michaelis-Menten Kinetik. Die Enzymparameter wurden wie folgt variiert: (a) & (b) Referenzzustand; (c) & (d) Enzymkonzentration von $R1 \cdot 100$; (e) & (f) Enzymkonzentration von $R1 / 100$; (g) K_m -Werte bei $R1 \cdot 10$, (h) Enzymkonzentration von $R1 / 5$; (i) $R1 \cdot K_{m,s} \cdot 10$, (j) Enzymkonzentration von $R1 \cdot 0,165$. Zur Deutung der Darstellung siehe Abschnitt 3.3.3 auf Seite 46. * Konzentrationen in mM, Flüsse in mmol/s.

(a)	Input	R1	R2	Output	Konz/Fluss*
A	0.968	-0.323	-0.323	-0.323	0.0292
S	0.966	0.010	-0.488	-0.488	1.9516
P	0.961	0.010	0.010	-0.980	0.9808
Fluss	0.971	0.010	0.010	0.010	9.71E-15

(b)	Input	R1	R2	Output	Konz/Fluss*
A	0.495	-0.485	-0.005	-0.005	0.5050
S	0.490	0.490	-0.490	-0.490	0.0100
P	0.485	0.485	0.005	-0.975	0.5050
Fluss	0.495	0.495	0.005	0.005	4.95E-15

(c)	Input	R1	R2	Output	Konz/Fluss*
A	nicht berechenbar				1.0000
S	nicht berechenbar				0.0100
P	nicht berechenbar				0.0100
Fluss	nicht berechenbar				0

(d)	Input	R1	R2	Output	Konz/Fluss*
A	0.247	-0.082	-0.082	-0.082	0.7525
S	0.245	0.245	-0.245	-0.245	0.00505
P	0.240	0.240	0.240	-0.721	0.2575
Fluss	0.250	0.250	0.250	0.250	2.48E-15

Abbildung 24 Kontrollkoeffizienten-Matrix bei Verschiebung des Gleichgewichts. (a) $R1.k_{hin} * 100$, (b) $R2.k_{hin} * 100$, (c) $R1.k_{rück} * 100$, (d) $R1.k_{rück} * 100$ & $R2.k_{hin} * 100$. Zur Deutung der Darstellung siehe Abschnitt 3.3.3 auf Seite 46. *Konzentrationen in mM, Flüsse in mmol/s.

4.3.2 Diskussion zum linearen Modell

Das lineare Modell mit zwei internen Reaktionen ist beinahe das einfachste Modell für ein System im Fließgleichgewicht. Auf der einen Seite sind die Ergebnisse ohne direkten praktischen Bezug. Auf der anderen Seite lässt sich gerade an einfachen Modellen das Verhalten bei Veränderungen der Parameter besonders gut interpretieren. Die Schlüsse, die hier gezogen werden, sind auch für komplexere Modelle von Bedeutung. Ohne Verständnis der fundamentalen Prinzipien ist die Interpretation des Verhaltens größerer Modelle kaum möglich.

Für dieses und alle weiteren Modelle musste zuerst geklärt werden, wie sich verschiedene Zufluss- und Abflussfunktionen auswirken. Wir werden bei anderen Modellen bei Bedarf auf diesen Punkt zurückkommen. Die Ergebnisse der vier Kombinationsmöglichkeiten von konstanter Konzentration in Verbindung mit chemischer Kinetik und konstantem Fluss sind in Abbildung 22 zu sehen.

Es gelingt in allen Fällen, identische Fluss- und Konzentrationswerte zu erhalten. Der Fall b bildet keine Ausnahme: für die drei Variablen (Konzentrationen) gibt es hier allerdings bei konstanter Zu- und Abflussrate nur zwei linear unabhängige Bestimmungsgleichungen, also unendlich viele Konzentrationsverteilungen. Bei geeignetem Startwert erreicht man auch hier die gleichen Konzentrationen im Fließgleichgewicht wie bei a, b und d. Wegen der Unterbestimmtheit konnten keine Kontrollkoeffizienten berechnet werden.

Obwohl Konzentrationen und Flüsse identisch sind, unterscheiden sich die Kontrollkoeffizienten bei a, c und d deutlich. Die Änderung eines Parameters (Enzymkonzentration) führt somit zur Einstellung von unterschiedlichen neuen Gleichgewichtszuständen. Im Besonderen verlieren in Fall c (konstanter Zufluss) die Reaktionen sämtliche Kontrolle über die Konzentrationen der Metabolite, die im Stoffwechsel folgen. Umgekehrt verhält es sich in Fall d (konstanter Abfluss). Die Reaktionen haben keine Kontrolle über die Konzentrationen der Vorläufermetabolite.

Welche Antwort des Systems auf eine Veränderung ist die richtige? In einem realen biochemischen System müssen zwei Punkte beachtet werden. Prinzipiell sind alle Reaktionen reversibel. Das gilt auch für Transportvorgänge. Deren Geschwindigkeit hängt von der Konzentration des intrazellulären und des extrazellulären Metaboliten ab, was im Fall von konstanten Zu- und Abflüssen nicht der Fall ist. Zweitens ist es durchaus möglich, dass eine interne Reaktion für den Fluss durch das System limitierend ist. Die Kontrolle über den Fluss sowie die Kontrolle über die Konzentrationen sollte nicht verloren gehen. Bleiben die Transportreaktionen reversibel, kann es vorkommen, dass der experimentell bestimmte Fluss nicht mit dem Modellfluss übereinstimmt. Dann liegt das Problem bei der Modellformulierung (falsche Parameter oder fehlende Komponenten). Es sei darauf

hingewiesen, dass auch eine Anpassung der kinetischen Parameter der Zu- und Abflussreaktionen möglich ist.

Ein Problem, das bei konstanten Abflüssen vorkommen kann, ist das Auftreten von negativen Konzentrationen im Simulationsverlauf, wenn die Abflüsse eines Metaboliten temporär größer sind als die Zuflüsse. Ein solches Problem kann bei konstanten Zuflüssen nicht auftreten. Die Zu- und Abflussraten sind häufig aus dem Experiment bekannt, so dass man trotz des möglichen Kontrollverlusts in der Praxis die Zu- und Abflussrate für einige Metabolite fixiert und negative transiente Konzentrationen in Kauf nimmt.

Wir wollen nun die chemische Kinetik mit der reversiblen Michaelis-Menten Kinetik vergleichen. Dabei steht die Frage im Vordergrund, in welchen Fällen man die Michaelis-Menten Kinetik durch chemische Kinetik annähern kann. Es wird jeweils ein Parameter der Michaelis-Menten-Kinetik variiert und anschließend versucht, den gleichen Zustand mit chemischer Kinetik zu erreichen, indem die Enzymkonzentration verändert wird. Die Gleichgewichtskonstante ist in allen Fällen gleich. In Abbildung 23a und b kann man erkennen, dass bei den gewählten Parameterwerten eine recht gute Annäherung möglich ist. Das gilt für sowohl für die Konzentrationen und den Stofffluss als auch für die Kontrollkoeffizienten. Eine Erhöhung der Enzymkonzentrationen um den Faktor 100 bei Reaktion R1 bei sonst gleichen Parametern kann gut durch chemische Kinetik angenähert werden (Abbildung 23c-d). Reaktion R1 ist praktisch im Gleichgewicht und hat nur noch wenig Kontrolle über Flüsse und Konzentrationen. Auch eine Verringerung der Enzymkonzentrationen, die mit einer Erhöhung der Kontrolle einhergeht, oder eine Erhöhung der K_m -Werte lassen sich kompensieren (Abbildung 23e-j). Im Detail sieht man jedoch Unterschiede. In allen Fällen außer c und d lässt sich die Konzentration von S nicht erreichen. Die Konzentrationen von A und P können nur deshalb gut reproduziert werden, weil sie zum größten Teil durch die Input- und Output-Reaktionen kontrolliert werden. Die Kontrolle über die Konzentration von S liegt jedoch zu einem beträchtlichen Teil bei den beiden internen Reaktionen. Die Kontrollkoeffizienten zeigen in allen Fällen, dass das System auf eine Parameterveränderung zwar qualitativ ähnlich, aber nicht quantitativ identisch reagiert.

Als letztes widmen wir uns dem Einfluss unterschiedlicher Gleichgewichtskonstanten. Die Veränderung dieser Konstanten wird durch die Anhebung einer oder mehrerer kinetischer Parameter erreicht. Zur Realisierung der Zu- und Abflüsse werden nur externe Konzentrationen, also keine Flüsse, fixiert. Es zeigt sich das erwartete Verhalten: eine Gleichgewichtsverschiebung einer Reaktion in Richtung Produkt erniedrigt die Konzentrationen der vorherigen und erhöht die Konzentration der folgenden Metabolite. Der Fluss wird größer (Abbildung 24a). Interessant ist, dass die Kontrolle über den Fluss auf die vorherigen Reaktionen übergeht. Die Steigerung des Flusses ist umso höher, je

früher im Pfad das Gleichgewicht verschoben wird, wie ein Vergleich mit Abbildung 24b zeigt. Im Fall c kommt der Fluss vollständig zum Erliegen, weil das eingestellte Konzentrationsgefälle zwischen den externen Metaboliten durch die veränderte Gleichgewichtslage ausgeglichen wird. Durch die positive Gibbsenergie der Reaktion entsteht eine „Energiebarriere“, die allerdings durch eine negative Gibbsenergie einer anderen Reaktion wieder kompensiert werden kann, wie Abbildung 24d zeigt.

Es wird festgehalten, dass die Gleichgewichtskonstanten der Reaktionen einen großen Einfluss auf den Fluss und vor allem auf die Konzentrationen im Fließgleichgewicht haben. Auch die Kontrolle über Fluss und Konzentrationen kann sich je nach Gleichgewichtslage stark verschieben.

4.4 Simulation der Modelle mit Verzweigung

4.4.1 Ergebnisse der Modelle mit Verzweigung

In Abbildung 16 auf Seite 49 sind die beiden Modelle mit Verzweigung abgebildet. Sie unterscheiden sich nur in den Abfluss-Reaktionen. Die Metaboliten P_1 und P_2 werden entweder direkt ausgeschleust oder reagieren gemeinsam im Verhältnis 1:1 zu einem artifiziellen Metaboliten Biomasse (BM). Verschiedene Parameterwerte werden mit beiden Modellen parallel getestet.

Abfluss zu Biomasse					
(a)	Input	R1	R2	Output	Konz/Fluss*
A	0.411	-0.142	-0.142	-0.126	0.5049
P1	0.247	0.543	-0.419	-0.370	0.2574
P2	0.247	-0.419	0.543	-0.370	0.2574
Fluss	0.581	0.145	0.145	0.128	2.48E-16

(c)	Input	R1	R2	Output	Konz/Fluss*
A	0.661	0.000	-0.657	-0.004	0.3351
P1	0.661	0.010	-0.667	-0.005	33.1728
P2	-0.072	-0.010	0.959	-0.877	0.0026
Fluss	0.667	0.000	0.331	0.002	3.32E-16

(e)	Input	R1	R2	Output	Konz/Fluss*
A	0.497	-0.492	-0.003	-0.002	0.5013
P1	-0.069	1.890	-0.980	-0.841	0.2646
P2	0.494	-1.469	0.982	-0.007	0.2520
Fluss	0.500	0.495	0.003	0.002	2.49E-16

Zwei Abfluss-Reaktionen						
(b)	Input	R1	Output1	R2	Output2	Konz/Fluss*
A	0.490	-0.123	-0.123	-0.123	-0.123	0.5050
P1	0.481	0.360	-0.601	-0.120	-0.120	0.2575
P2	0.481	-0.120	-0.120	0.360	-0.601	0.2575
Fluss1	0.500	0.375	0.375	-0.125	-0.125	2.48E-16
Fluss2	0.500	-0.125	-0.125	0.375	0.375	2.48E-16

(d)	Input	R1	Output1	R2	Output2	Konz/Fluss*
A	0.971	-0.483	-0.483	-0.002	-0.002	0.0196
P1	0.966	0.014	-0.975	-0.002	-0.002	0.9856
P2	0.643	-0.320	-0.320	0.161	-0.164	0.0148
Fluss1	0.976	0.015	0.015	-0.002	-0.002	9.76E-16
Fluss2	1.981	-0.985	-0.985	0.495	0.495	4.81E-18

(f)	Input	R1	Output1	R2	Output2	Konz/Fluss*
A	0.494	-0.248	-0.001	-0.123	-0.123	0.5032
P1	0.475	0.718	-0.958	-0.118	-0.118	0.2603
P2	0.485	-0.243	-0.001	0.360	-0.601	0.2566
Fluss1	0.249	0.747	0.004	-0.125	-0.125	2.50E-16
Fluss2	0.250	-0.249	-0.001	0.375	0.375	2.47E-16

Abbildung 25 Kontrollkoeffizienten des verzweigten Modells bei Abfluss zu Biomasse bzw. Abfluss zu zwei Reaktionen. (a) & (b) Ausgangssituation, (c) & (d) $R1 \cdot k_1 \cdot 100$, (e) & (f) $R1 \cdot k_1 \cdot 0,5$ & $R1 \cdot k_1 \cdot 0,005$. Zur Deutung der Darstellung siehe Abschnitt 3.3.3 auf Seite 46. *Konzentrationen in mM, Flüsse in mmol/s.

4.4.2 Diskussion zu den Modellen mit Verzweigung

Bei den meisten Modellen biochemischer Systeme existieren mehrere Zu- und Abflüsse. Die Versuche mit den Modellen von Abbildung 16 (Seite 49) wurden unternommen, um die Unterschiede zwischen den beiden Modellierungskonzepten für die Abflüsse „mit gemeinsamer Biomasse“ und „ohne gemeinsame Biomasse“ zu untersuchen.

In Abbildung 25a und b sind die Modellparameter so eingestellt, dass die Konzentrationen und Flüsse praktisch identisch sind. Die Gleichgewichtskonstanten aller Reaktionen sind gleich eins. Da das Modell symmetrisch ist, werden beide Pfade zu gleichen Teilen genutzt. Bereits hier sieht man jedoch den Unterschied beider Konzepte anhand der Kontrollkoeffizienten. Würde man beispielsweise die Enzymkonzentration von R_1 im Modell ohne Biomasse erhöhen, würde der Fluss durch Pfad 1 erwartungsgemäß steigen, während der Fluss durch Pfad 2 geringer werden würde. Bei dem Modell mit Biomasse muss der Fluss durch beide Pfade gleich bleiben und würde zu gleichen Teilen angehoben werden.

Es ist sehr unwahrscheinlich, dass die Gleichgewichtskonstanten bei zwei Pfaden identisch sind. Zum einen, weil sich die Bildungs-Gibbsenergien der Produkte im Regelfall unterscheiden, zum anderen deshalb, weil die Reaktionen an unterschiedliche Cosubstrate gekoppelt sind. Das Gleichgewicht der Reaktion R_1 wird deshalb zugunsten des Produkts verlagert, indem die Geschwindigkeitskonstante der Hinreaktion um den Faktor 100 erhöht wird (Fall c und d). Die Modelle reagieren wie erwartet mit einer entsprechenden Flussteigerung. Beachtlich ist die große Konzentrationsänderung von P_1 und P_2 im Biomasse-Modell. Der Zwang, auch den Fluss durch Pfad 2 zu steigern, führt zu einer vergleichsweise geringen Steigerung des Flusses durch Pfad 1. Daher muss das Gleichgewichtsgefälle über die Konzentrationsverhältnisse ausgeglichen werden.

Wir betrachten hier nur eine Reaktion und ein Gleichgewichtsgefälle. Hat man ein Modell mit vielen gekoppelten Reaktionen und Verzweigungen, können die Konzentrationen leicht in unphysiologische Bereiche verschoben werden. Es ist daher anzunehmen, dass eine biochemische Zelle eine andere Strategie hat, um die Verhältnisse der Flüsse aufeinander abzustimmen. In Abbildung 25e und f wird eine andere Möglichkeit gezeigt, das Gleichgewichtsgefälle zu kompensieren: die Verringerung der Enzymkonzentration von R_1 . Verringert man die Enzymkonzentration, also die Geschwindigkeitskonstanten, im Modell mit Biomasse so weit, dass die Flussrate wieder der ursprünglichen Flussrate entspricht, sind auch die Konzentrationswerte wieder in einem engen Bereich. Die gleichen Änderungen führen beim Modell ohne gemeinsame Biomasse zu einer identischen Fluss- und Konzentrationsverteilung.

Wirken äußere Zwänge auf das Modell, wie bei einer gemeinsamen Biomasse, müssen die internen Parameter (hier: effektive Enzymkonzentrationen) trotzdem so angepasst

werden, dass die Variablen (Metabolitkonzentrationen) in einem physiologischen Bereich bleiben. Die biochemische Zelle kann dies beispielsweise durch Feedback-Hemmung der entsprechenden Enzyme erreichen.

4.5 Simulation der Glykolyse

4.5.1 Ergebnisse des Modells der Glykolyse

Die Glykolyse von *C. glutamicum* ist das erste reale System, das modelliert wird. Die Werte für die Parameter sind so realistisch wie möglich gewählt worden (siehe Anhang 6.1). Untersucht wurden Fließgleichgewichtszustände in Abhängigkeit von den freien Parametern und von den Anfangsbedingungen. Das Modell ist in Abbildung 17 auf Seite 51 zu sehen.

Bei konstanter Zuflussrate von Glukose und konstanten Konzentrationen der Metabolite, die an den Systemgrenzen liegen, wurden die freien Parameter der Reaktionen variiert. Dies geschah möglichst zielgerichtet (Abbildung 26), um zu physiologischen Konzentrationen zu gelangen (Tabelle 11).

Die Integration der Differentialgleichungen zeigt die unterschiedlichen Wege in das Fließgleichgewicht bei verschiedenen Anfangsbedingungen (Abbildung 27 und Abbildung 28).

Tabelle 11 Konzentrationsveränderung in der Glykolyse im Fließgleichgewicht durch die Anpassung der kinetischen Konstanten (vergleiche Abbildung 26).

Metabolit	Konzentration vor Anpassung [mM]	Konzentration nach Anpassung [mM]
Glukose	0,012	0,012
G6P	0,050	0,009
F6P	2,810	0,012
F16DP	15850,800	0,928
DHAP	0,062	0,005
3PGA	0,039	<0,001
13DPG	0,063	0,030
3PG	4,461	0,579
2PG	0,107	0,012
PEP	0,255	0,255

(a)	gk	gpi	pfk	fbpa	tpi	gapdh	pgk	pgm	pph	pk	Konzentration [mM]
GLC	-9,99E-01	-1,38E-04	-2,55E-06	-2,39E-09	-8,78E-05	-1,12E-03	-5,36E-07	-9,68E-10	-4,23E-08	-4,78E-10	0,012
G6P	0	-1,86E-01	-3,43E-03	-3,20E-06	-1,18E-01	-1,50E+00	-7,20E-04	-1,30E-06	-5,68E-05	-6,42E-07	0,050
F6P	0	0	-4,21E-03	-3,93E-06	-1,45E-01	-1,85E+00	-8,84E-04	-1,60E-06	-6,98E-05	-7,88E-07	2,810
<i>F16DP</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>-3,95E-06</i>	<i>-1,45E-01</i>	<i>-1,85E+00</i>	<i>-8,88E-04</i>	<i>-1,60E-06</i>	<i>-7,01E-05</i>	<i>-7,91E-07</i>	15850,800
DHAP	0	0	0	0	-1,45E-01	-8,54E-01	-4,09E-04	-7,39E-07	-3,23E-05	-3,65E-07	0,062
3PGA	0	0	0	0	0	-9,99E-01	-4,79E-04	-8,64E-07	-3,78E-05	-4,27E-07	0,039
13DPG	0	0	0	0	0	0	-9,25E-01	-1,67E-03	-7,30E-02	-8,24E-04	0,063
3PG	0	0	0	0	0	0	0	-2,21E-02	-9,67E-01	-1,09E-02	4,461
2PG	0	0	0	0	0	0	0	0	-9,89E-01	-1,12E-02	0,107
PEP	0	0	0	0	0	0	0	0	0	-1,00E+00	0,255

↓ $\cdot 10^3$

(b)	gk	gpi	pfk	fbpa	tpi	gapdh	pgk	pgm	pph	pk	Konzentration [mM]
<i>F16DP</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>-1,74E-02</i>	<i>-9,74E-01</i>	<i>-6,53E-01</i>	<i>-3,13E-01</i>	<i>-5,65E-04</i>	<i>-2,47E-02</i>	<i>-2,79E-04</i>	3,591
3PG	0	0	0	0	0	0	0	-2,21E-02	-9,67E-01	-1,09E-02	4,461

↓ $\cdot 2$ ↓ $\cdot 2$ ↓ $\cdot 2$

(c)	gk	gpi	pfk	fbpa	tpi	gapdh	pgk	pgm	pph	pk	Konzentration [mM]
<i>F16DP</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>-6,47E-02</i>	<i>-9,27E-01</i>	<i>-6,06E-01</i>	<i>-2,90E-01</i>	<i>-1,05E-03</i>	<i>-4,58E-02</i>	<i>-5,17E-04</i>	0,968
<i>3PG</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>-2,21E-02</i>	<i>-9,67E-01</i>	<i>-1,09E-02</i>	4,461

↓ $\cdot 10$

(d)	gk	gpi	pfk	fbpa	tpi	gapdh	pgk	pgm	pph	pk	Konzentration [mM]
<i>F16DP</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>-6,75E-02</i>	<i>-9,24E-01</i>	<i>-6,32E-01</i>	<i>-3,03E-01</i>	<i>-1,09E-03</i>	<i>-4,78E-03</i>	<i>-5,39E-04</i>	0,928
<i>3PG</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>-1,70E-01</i>	<i>-7,45E-01</i>	<i>-8,41E-02</i>	0,579

Abbildung 26 Kontrollkoeffizienten und Konzentrationen bei der Veränderung von einzelnen Enzymkonzentrationen der Glykolyse im Fließgleichgewicht. Die Geschwindigkeitskonstanten der Hin- und Rückreaktionen wurden mit folgenden Faktoren multipliziert: (a) Ausgangslage, (b) gapdh mit 1000, (c) zusätzlich tpi, gapdh und pgk mit 2, (d) zusätzlich peph mit 10. Zur Deutung der Darstellung siehe Abschnitt 3.3.3 auf Seite 46.

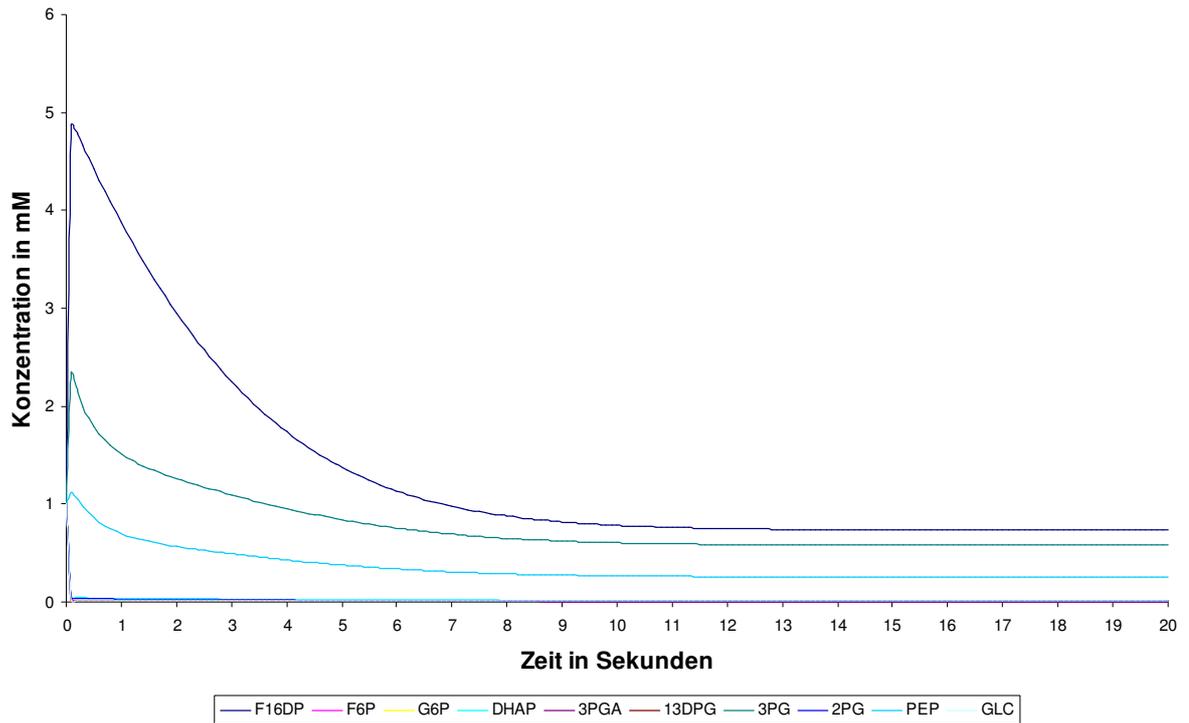


Abbildung 27 Simulation der Glykolyse bis zum Erreichen des Fließgleichgewichts. Die Startkonzentrationen aller Metabolite sind gleich 1,0 mM.

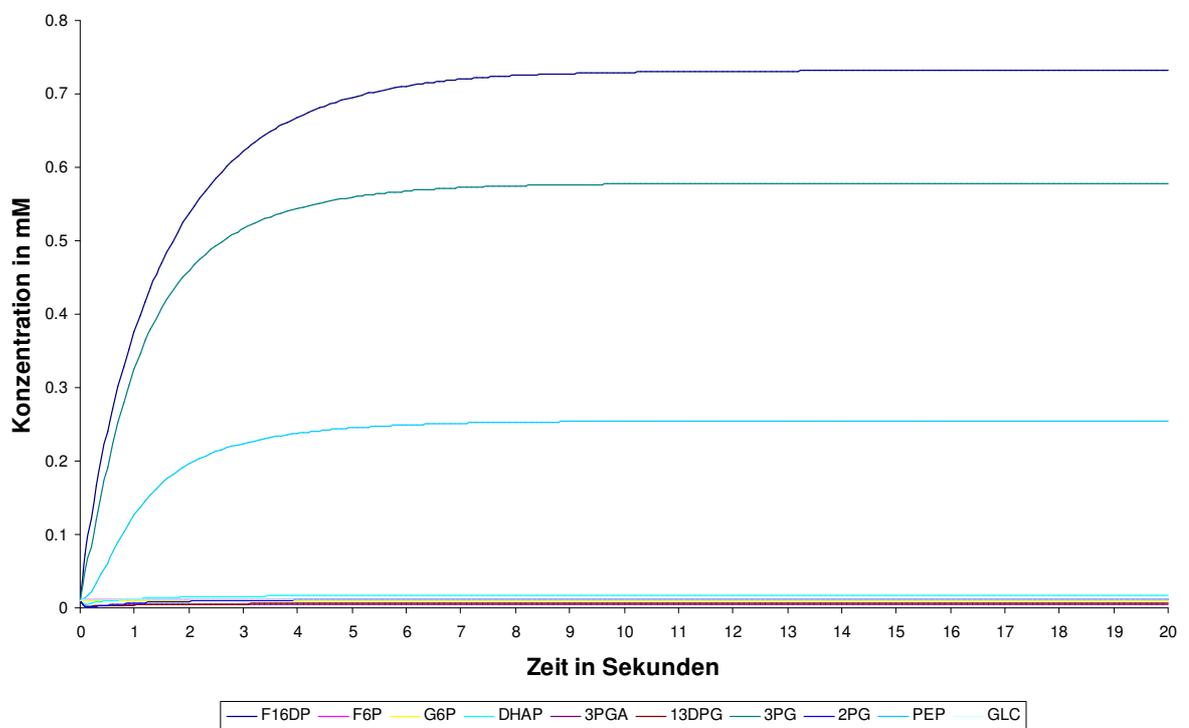


Abbildung 28 Simulation der Glykolyse bis zum Erreichen des Fließgleichgewichts. Die Startkonzentrationen aller Metabolite sind gleich 0,01 mM.

4.5.2 Diskussion zum Modell der Glykolyse

Die artifiziellen Modelle haben bereits einige grundlegende Modellierungsprinzipien für metabolische Netzwerke aufgezeigt. Im Folgenden wird anhand des Glykolyse-Modells erläutert, wie man – ausgehend von einem beliebigen Fließgleichgewichtszustand – durch Veränderung der kinetischen Konstanten zu einem physiologisch sinnvollen Fließgleichgewichtszustand gelangt. Die Simulation mit verschiedenen Anfangskonzentrationen kann Aufschlüsse über die Eindeutigkeit des Fließgleichgewichtszustandes geben.

Wie man in Abbildung 26a sieht, kontrollieren die Reaktionen in diesem weitgehend linearen Netzwerk nur die vorhergehenden Konzentrationen (vergleiche Abschnitt 4.3). Die meisten Reaktionen lagen bereits in einem physiologisch sinnvollen Bereich, mit Ausnahme der Konzentration von Fruktose-1,6-diphosphat (Abbildung 26a). Ziel musste es also sein, durch gemeinsame Skalierung der kinetischen Konstanten von Hin- und Rückreaktionen, also bei konstanter Gleichgewichtslage, diese Konzentration zu verringern. Dabei helfen die Kontrollkoeffizienten, die anzeigen, dass die Kontrolle über diese Konzentration im Wesentlichen bei der Glyceraldehyd-3-phosphat Dehydrogenase liegt. Allerdings hat dieses Enzym auch eine nicht zu vernachlässigende Kontrolle über andere Konzentrationen, so dass eine Veränderung an dieser Stelle Nebenwirkungen hat. Nach der Skalierung um den Faktor 1000 ist die Fließgleichgewichtskonzentration von Fruktose-1,6-diphosphat beträchtlich gesunken. Dieser Zustand ist physiologisch durchaus möglich. Wir wollen aber noch einen Schritt weiter gehen und alle Konzentrationen in den Bereich unter 1 mM verändern. In Abbildung 26b kann man erkennen, dass die Konzentration von Fruktose-1,6-diphosphat nach der Veränderung nicht mehr am stärksten von der Glyceraldehyd-3-phosphat Dehydrogenase, sondern zu ungefähr gleichen Teilen von der Triosephosphat Isomerase, der Phosphoglyceraldehyd Dehydrogenase und der Phosphoglyceratkinase kontrolliert wird. Diese Wechsel der Kontrolle sind typisch, wenn eine Enzymkonzentration oder Konstante eines Enzyms, das die Kontrolle besitzt, stark verändert wird. Eine Skalierung bei den genannten Enzymen mit dem Faktor 2 bringt den gewünschten Effekt. Die gleiche Prozedur kann auch mit 3-Phosphoglycerat und Phosphoenolpyruvat Hydratase durchgeführt werden.

Man gelangt so zu dem Endzustand, der in Tabelle 11 zu sehen ist. Ein Vergleich mit dem Anfangszustand zeigt, dass die Konzentrationen der meisten Metabolite von den drei Skalierungsschritten betroffen sind. Ausnahmen bilden nur der Anfangs- und Endpunkt des Pfades, weil die Glukose-Konzentration hauptsächlich durch die Glucokinase kontrolliert wird, während die Phosphoenolpyruvat-Konzentration ausschließlich durch die Pyruvatkinase bestimmt wird. Der Einfluss auf die dazwischen liegenden Metabolite ist unterschiedlich stark ausgeprägt. In allen Fällen findet eine Verringerung der

Konzentrationen im Fließgleichgewicht statt, weil alle Kontrollkoeffizienten negativ sind. Dies ist wiederum nur deshalb der Fall, weil ein konstanter Zufluss eingerichtet wurde. Bei höherer Enzymkonzentration wird eine geringere Konzentration der Metabolite benötigt, um den gleichen Stofffluss zu erzeugen.

Die exemplarisch in Abbildung 27 und Abbildung 28 gezeigten Simulationsverläufe zeigen, dass der Fließgleichgewichtszustand durch die gegebenen Parameter vollständig determiniert ist. Neben den gezeigten wurden fünfzig weitere, zufällig erzeugte Anfangswerte im Bereich 0,01 mM bis 10 mM getestet. Der Weg, auf dem das Fließgleichgewicht erreicht wird, unterscheidet sich, der Endzustand jedoch nicht. Sind die Konzentrationen im gesamten Pfad im Vergleich zum Fließgleichgewicht relativ hoch, kann es vorübergehend zu einer Anhäufung bestimmter Metabolite kommen (Abbildung 27). Bei relativ geringen Startkonzentrationen trifft man auf den umgekehrten Fall, dass Konzentrationen transient stark erniedrigt werden können und der Endzustand „von unten“ erreicht wird (Abbildung 28). Dies kann dazu führen, dass im Verlauf einige Konzentrationen negative Werte erreichen. Dass es für dieses Modell tatsächlich nur einen Fließgleichgewichtszustand gibt, werden wir in Abschnitt 4.8.2 sehen.

Der Versuch, dieses Modell mit konstanter Zu- und Abflussrate zu simulieren, führte nicht zu einem Fließgleichgewichtszustand. Den Grund dafür werden wir ebenfalls in Abschnitt 4.8.2 kennenlernen.

4.6 Simulation des Citratcyclus

4.6.1 Ergebnisse des Modells des Citratcyclus

In diesem Abschnitt sind die Ergebnisse der Citratcyclus-Modellierung zusammengestellt. Die Details zum Modell befinden sich in Abschnitt 3.3.5.2 (Seite 52) und in Anhang 6.3. Das Fließgleichgewicht in Abhängigkeit von unterschiedlichen Zu- und Abflussmethoden und die Auswirkungen auf den Modellierungsprozess wurden hier untersucht.

In Tabelle 12 befinden sich die Flusskontrollkoeffizienten, in Tabelle 13 die Konzentrationskontrollkoeffizienten. In einem Fall sind die Zu- und Abflüsse im Citratcyclus jeweils durch konstante Konzentrationen der externen Metabolite realisiert, während im anderen Fall die Flüsse explizit vorgegeben werden. Die zugehörige Flussverteilung ist in Abbildung 29 links zu sehen. Die Fluss- und Konzentrationsverteilung ist für beide Modelle identisch. Wird der Verbrauch von 2-Oxoglutarat im Fall der fixierten Flüsse zurückgefahren, ergibt sich eine neue Flussverteilung. Sie ist in Abbildung 29 rechts schematisiert.

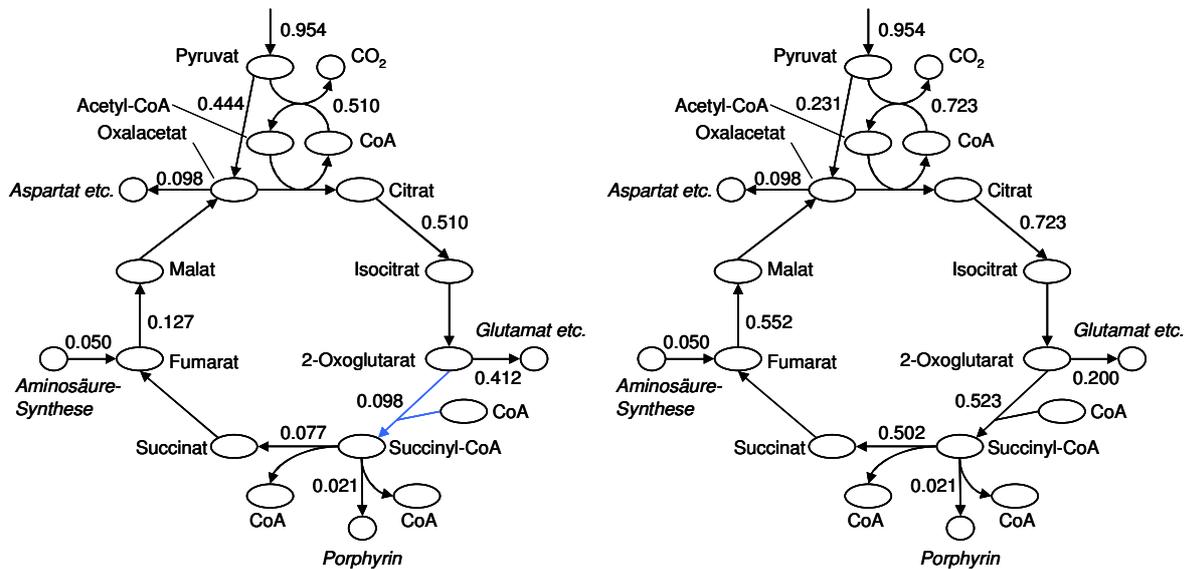


Abbildung 29 Flussverteilung im Citratcyclus bei unterschiedlichem Verbrauch von 2-Oxoglutarat. Links: hoher Verbrauch. Rechts: geringerer Verbrauch. Die Reaktion der 2-Oxoglutarat Dehydrogenase ist blau gezeichnet. Flüsse in mM/s.

Tabelle 12 Flüsse und Flusskontrollkoeffizienten von 2-Oxoglutarat Dehydrogenase bei verschiedenen Zu- und Abflussmethoden des Citratcyclus. Bei mehreren Reaktionen mit gleichen Flüssen wurde eine Reaktion ausgewählt.

Enzym / Reaktion	Fluss [mM/s]	Kontrollkoeffizient (fixierte Konzentration)	Kontrollkoeffizient (fixierter Fluss)
Zufluss Pyruvat (immer fixiert)	0.954	0	0
Pyruvat Dehydrogenase	0.510	0.092	0
2-Oxoglutarat Dehydrogenase	0.098	1.151	0
Succinyl-CoA Synthetase	0.077	1.296	0
Fumarase	0.127	0.785	0
Anaplerotische Reaktion	0.444	-0.105	0
Abfluss 2-Oxoglutarat	0.412	-0.158	0
Abfluss Succinyl-CoA	0.021	0.608	0
Zufluss Fumarat	0.050	0.000	0
Abfluss Oxalacetat	0.098	0.100	0

Tabelle 13 Konzentrationen und Konzentrationskontrollkoeffizienten von 2-Oxoglutarat Dehydrogenase bei verschiedenen Zu- und Abflussmethoden des Citratcyclus.

Metabolit	Konzentration [mM]	Kontrollkoeffizient (fixierte Konzentration)	Kontrollkoeffizient (fixierter Fluss)
Pyruvat	0.074	-0.105	0
Acetyl-CoA	2.075	-0.006	0
CoA	4.31E-3	0.197	0
Citrat	0.141	0.091	-2.23E-3
Isocitrat	0.128	0.091	-2.51E-3
2-Oxoglutarat	0.141	-0.046	-1
Succinyl-CoA	0.021	0.599	-3.81E-9
Succinat	1.831	0.246	-4.66E-9
Fumarat	3.79E-4	0.235	-4.71E-9
Malat	0.063	0.184	-5.14E-9
Oxalacetat	6.15E-5	0.098	-5.88E-9

4.6.2 Diskussion zum Modell des Citratcyclus

Der Citratcyclus ist eines der Herzstücke des Metabolismus. Er wird bei Glukosefütterung von *C. glutamicum* aus der Glykolyse gespeist und ist Ausgangspunkt für die Synthese einer Reihe von Aminosäuren und anderer funktionell wichtiger Substanzen, wie z. B. der Porphyrine. Wir gehen davon aus, dass sich der Organismus in der Wachstumsphase befindet. Daher reicht es näherungsweise aus, neben der Oxidation des Pyruvats aus der Glykolyse zur Energiegewinnung, nur die anabolen Prozesse zu berücksichtigen.

Für den Modellierungsprozess stellen sich neue Herausforderungen. Die zyklische Natur des Pfades bietet im Zusammenhang mit verschiedenen Zu- und Abflüssen eine Reihe von Justierungsmöglichkeiten für die Stoffflüsse und Konzentrationen. Die Problematik einer Conserved Moiety (Kapitel 3.3.2, Seite 45), nämlich der Erhaltungssatz für Coenzym A, wird im Folgenden ebenfalls diskutiert.

Im Gegensatz zum Modell der Glykolyse (Abschnitt 4.5) war es möglich, einen stabilen Fließgleichgewichtszustand auch bei expliziter Vorgabe konstanter Zu- und Abflussraten zu erreichen. Die Parameter reichten hier also aus, um den Zustand ohne Fixierung zusätzlicher Konzentrationsvariablen eindeutig zu bestimmen.

Der Unterschied zwischen konstanter Konzentration und konstantem Fluss bei der Modellierung der Zu- und Abflüsse wurde schon in Abschnitt 4.4 besprochen. Dort hatten wir gesehen, dass eine identische Fluss- und Konzentrationsverteilung mit beiden Methoden zu erreichen ist. Dies bestätigt sich hier für ein komplizierteres Modell. Es ist einleuchtend, dass in einem System, bei dem alle Flüsse durch die externen Reaktionen fixiert sind, die internen Reaktionen keine Kontrolle mehr über den Fluss haben (Tabelle 12). Das kann aber auch einen drastischen Verlust bzw. eine drastische Erhöhung der Kontrolle über die Konzentrationen bedeuten, wie man am Beispiel der 2-Oxoglutarat Dehydrogenase in Tabelle 13 sieht. Wenn der Parameter einer internen Reaktion verändert wird und, im Gegensatz zur Situation bei fixierten externen Konzentrationen, das System nicht mit einer Flussveränderung reagieren kann, muss es mit einer Konzentrationsveränderung reagieren. Im konkreten Fall der 2-Oxoglutarat Dehydrogenase steigt die Kontrolle über 2-Oxoglutarat, während die Kontrolle über alle anderen Metabolitkonzentrationen stark abnimmt.

Ein Vorteil der Fixierung der externen Flüsse ist allerdings, dass man aufgrund der direkten Kontrolle sehr viel leichter verschiedene Zu- und Abflüsse testen kann. Bei fixierten Konzentrationen ist es sehr viel komplizierter, weil sich Parameterveränderungen bei einer externen Reaktion nicht-linear auf die Flüsse durch die anderen externen Reaktionen auswirken. Es wurde deshalb das Verhalten bei verringertem Abfluss von 2-Oxoglutarat bei fixierten Flüssen untersucht. Wie man in Abbildung 29 rechts sieht, führt

dies zu einer Zunahme des zyklischen Flusses. Das System kann bei geringerer Aktivität der anabolen Prozesse also den Katabolismus aktivieren, hier die Energiegewinnung in Form von Reduktionsäquivalenten und ATP durch erhöhten Fluss durch den Citratcyclus.

Nun wollen wir noch kurz auf das eingangs erwähnte Problem der Conserved Moiety eingehen:

Succinyl-CoA + Acetyl-CoA + CoA = konstant

Während beim Modell der Glykolyse alle Conserved Moieties durch Fixierung der Konzentrationen beseitigt wurden, würde das gleiche Vorgehen in diesem Modell dazu führen, dass alle Produkte der Pyruvat Dehydrogenase – also Acetyl-CoA, NAD und CO₂ – in der Konzentration fixiert worden wären. Das gleiche gilt für die Reaktionen, an denen Succinyl-CoA beteiligt ist. Dies würde wiederum zu einer Entkopplung der Flüsse durch das System führen. In Abschnitt 3.3.2 wurde diskutiert, welche Folgen die Behandlung der Konzentrationen einer Conserved Moiety als Variablen auf das Systemverhalten haben kann. Da diese Probleme hier nicht auftreten, wurden die CoA-Derivate als freie Variablen modelliert. Auf die Conserved Moiety musste trotzdem bei der Modellierung Rücksicht genommen werden. Würde man den Abfluss von Succinyl-CoA wie alle anderen externen Reaktionen modellieren, könnte kein stationärer Zustand erreicht werden, weil der Conserved Moiety mit konstanter Rate ein Metabolit entzogen werden würde. Erst die Kopplung des Abflusses an die Produktion von CoA kann dieses Problem beseitigen. Aus ungeklärter Ursache konnte bei Verringerung der Konzentrationssumme der CoA-Derivate unter 0,35 mM kein Fließgleichgewicht erreicht werden.

4.7 Simulation des Zentralstoffwechsels

4.7.1 Ergebnisse des Modells des Zentralstoffwechsels

Das Modell des Zentralstoffwechsels beinhaltet die Reaktionen der Glykolyse, des Pentosephosphatwegs und des Citratcyclus. Aus der Biomassezusammensetzung wurden die Precursor bestimmt, die Teil des Modells sind, und entsprechende Zu- und Abflüsse eingerichtet. Der Modellierungsprozess wird genauer in Abschnitt 3.3.5.3 beschrieben. Die Flüsse im Fließgleichgewicht sind in Abbildung 30 zu sehen. Die zugehörigen Konzentrationen finden sich in Tabelle 14.

Tabelle 14 Konzentrationen des Zentralstoffwechsels im Fließgleichgewicht.

Metabolit	Konzentration [μM]
Glukose	11.8
Glukose-6-phosphat (G6P)	6.9
Fruktose-6-phosphat (F6P)	1.1
Fruktose-1,6-diphosphat (F16DP)	63.6
Glyceronphosphat	35.3
3-Phosphoglyceraldehyd	25.0
3-Phosphoglyceroylphosphat	0.7
3-Phosphoglycerat	163.1
2-Phosphoglycerat	1.9
Phosphoenolpyruvat	213.9
Pyruvat	26.6
Acetyl-CoA	1108.0
CoA	24.1
Citrat	240.2
Isocitrat	218.3
2-Oxoglutarat	214.1
Succinyl-CoA	67.9
Succinat	509.6
Fumarat	0.1
Malat	18.3
Oxalacetat	<0.1
Glucono-1,5-lacton-6-phosphat (G15L6P)	4.7
6-Phosphogluconat (6PGLU)	47.3
Ribulose-5-phosphat (RU5P)	44.9
Xylulose-5-phosphat (XU5P)	42.8
Ribose-5-phosphat (R5P)	140.7
Sedoheptulose-7-phosphat	1.8
Erythrose-4-phosphat (E4P)	4.4
Phosphoribosyl-diphosphat (PRPP)	8305.7

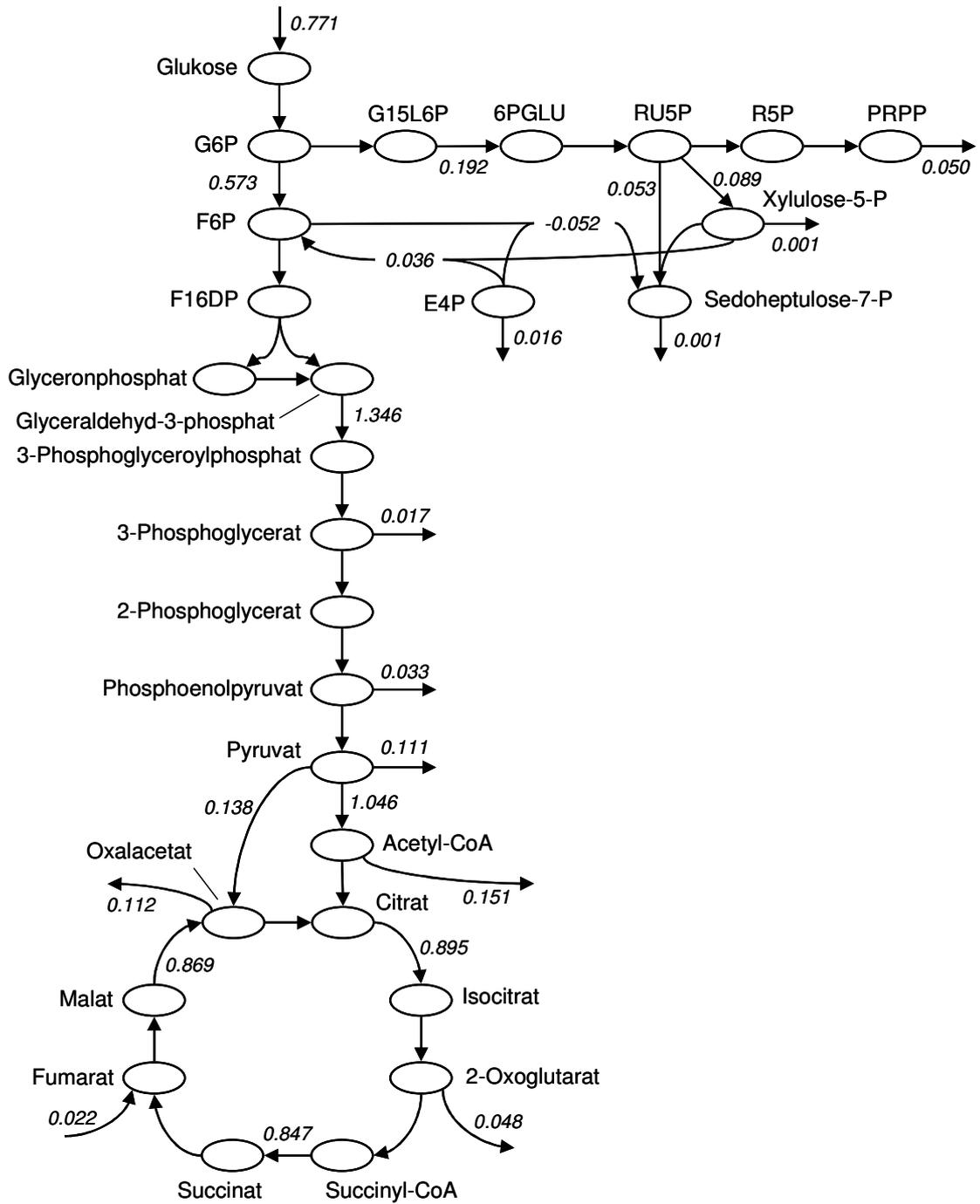


Abbildung 30 Modell des Zentralstoffwechsels mit Flüssen im Fließgleichgewicht in mM/s.

4.7.2 Diskussion zum Modell des Zentralstoffwechsels

Im Zentralstoffwechsel herrschen im Allgemeinen die höchsten Stoffflüsse und es werden bei vielen Metaboliten im Vergleich zu anderen Pfaden hohe Konzentrationen gemessen. Der Zugang zu experimentellen Daten ist leichter und präziser möglich als bei anderen Stoffwechselwegen. Für die Modellierung ist er deshalb von besonderem Interesse.

Beim Aufbau des Modells wurde versucht, so viele experimentelle Daten wie möglich zu verwenden. Die angenommene Biomassezusammensetzung soll ein realistisches Bild für die Metabolite, die für das Zellwachstum benötigt werden, und deren Verhältnis liefern. Die Glukose-Zuflussrate und die Konzentrationen von fixierten Metaboliten wurden aus der Literatur übernommen, um zumindest eine vernünftige Größenordnung bzw. ein vernünftiges Mengenverhältnis der Cosubstrate zueinander zu gewährleisten. Als Folge erhält man plausible eine Fluss- und Konzentrationsverteilung.

Die interne Flussverteilung ist durch die fixierten Zu- und Abflüsse im Wesentlichen vorgegeben. Es wären prinzipiell noch andere globale Flussverteilungen möglich, wie z. B. ein umgekehrter Fluss durch den Pentosephosphatweg. Diese werden aber durch die Beschränkungen, die durch die Gibbsenergien vorgegeben werden, nur bei ungünstigen Konzentrationsverhältnissen möglich. Die Folge ist eine realistische globale Flussverteilung mit Verstoffwechslung von Glukose zur Energiegewinnung und gleichzeitig ablaufendem Zellwachstum.

Die meisten Konzentrationen bewegen sich im micro- bis millimolaren Bereich. Es gibt jedoch einige Ausnahmen. Für die ungewöhnlich hohe Konzentration von Phosphoribosyl-diphosphat (PRPP) gibt es zwei Erklärungsmöglichkeiten. PRPP kommt in nur einer internen Reaktion vor. Die abgeschätzte Gleichgewichtskonstante für diese Reaktion könnte zu gering sein. Des Weiteren ist nur an der Reaktion des Netzwerk AMP beteiligt. Wenn die Konzentration dieses Metaboliten zu gering gewählt wurde, wird das Gleichgewicht, wie bei einer zu kleinen Gleichgewichtskonstante, fälschlicherweise zugunsten von PRPP verschoben. Das Verhältnis von Acetyl-CoA zu CoA von ca. 46 zu 1 ist ebenso wie die Konzentration von Acetyl-CoA deutlich zu hoch. Eine Reduktion der Aktivität der Pyruvatdehydrogenase sollte zu einer Verminderung der Konzentration führen. Es wurde versucht, die Aktivität des Enzyms und den CoA-Pool weiter zu reduzieren. Beides führte zum Zusammenbruch des Fließgleichgewichts. Offenbar fehlen hier die korrekten regulativen Prozesse und/oder weitere Prozesse, die an Acetyl-CoA gekoppelt sind.

Die geringe Konzentration von Oxalacetat konnte auch durch Veränderung der Geschwindigkeitskonstanten der Reaktionen (bei gleicher Gleichgewichtslage) nicht weiter erhöht werden, da die Enzyme, die Kontrolle über die Konzentration haben, auch starke Kontrolle über eine Reihe anderer Konzentrationen im Modell besitzen. Die Konzentration

scheint also unter allen Umständen sehr gering zu sein. Dies stimmt mit allen bisher im Hause gemessenen Metabolitprofilen von *C. glutamicum* überein. Es konnte bislang intrazellulär noch kein Oxalacetat nachgewiesen werden (Buchinger 2006).

4.8 Alternative Methoden zum Erreichen eines Fließgleichgewichts

In diesem Abschnitt werden die Vor- und Nachteile weiterer Methoden zum Erreichen eines Fließgleichgewichts diskutiert. Auf eine Trennung zwischen der Präsentation der Ergebnisse und der Diskussion wird hier verzichtet.

4.8.1 Farbige Petrinetze

Als Alternative zur zeitabhängigen Simulation durch Integration von Differentialgleichungen wurde die Simulation farbiger Petrinetze getestet. Eine kurze Einführung zu Petrinetzen befindet sich in Abschnitt 2.4. Die Details der Implementierung sind in Abschnitt 3.4.1 beschrieben.

Eines der wesentlichen Merkmale von Petrinetzen ist, dass Ereignisse nacheinander ausgeführt werden. Im Gegensatz zur Integration von Differentialgleichungssystemen, bei denen die einzelnen Gleichungen gleichzeitig gelöst werden, werden die Reaktionsereignisse nacheinander abgearbeitet. Zwei Prozesse können jedoch nebenläufig (kausal unabhängig) ablaufen.

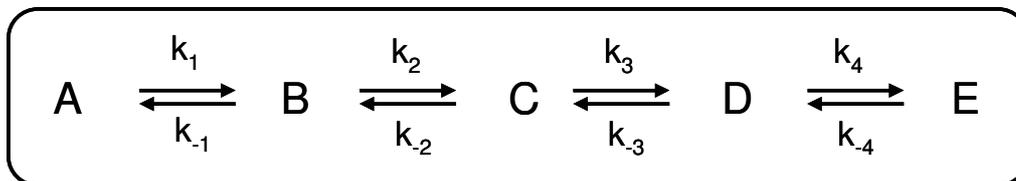


Abbildung 31 Lineares Reaktionsmodell.

Wir beschäftigen uns hier mit einem geschlossenen, linearen System (Abbildung 31) und wollen den Gleichgewichtszustand des Systems finden. Die Reaktionsgeschwindigkeit gehorcht chemischer Kinetik. Die Geschwindigkeitskonstanten der einzelnen Reaktionen sind in Tabelle 15 zusammengestellt.

Die zweite Reaktion, die B und C ineinander umwandelt, wurde bei drei verschiedenen Geschwindigkeiten simuliert. In allen drei Fällen wurde die Gleichgewichtskonstante der Reaktion nicht verändert, so dass der gleiche Endzustand erreicht werden sollte. Tatsächlich erhält man für beide Methoden und alle Versuche praktisch den gleichen Kurvenverlauf und den gleichen Endzustand. Das Gleichgewicht wird jeweils nach ca. 5000 Sekunden Simulationszeit erreicht. Exemplarisch ist der zeitliche Verlauf in Abbildung 32 (Seite 94) zu sehen.

Tabelle 15 Parameter und Simulationsergebnisse des Modells aus Abbildung 31 bei drei verschiedenen Versuchen.

Konstante	Versuch 1	Versuch 2	Versuch 3
k_1	1 Ms^{-1}	1 Ms^{-1}	1 Ms^{-1}
k_{-1}	1 Ms^{-1}	1 Ms^{-1}	1 Ms^{-1}
k_2	2 Ms^{-1}	200 Ms^{-1}	200000 Ms^{-1}
k_{-2}	1 Ms^{-1}	100 Ms^{-1}	100000 Ms^{-1}
k_3	1 Ms^{-1}	1 Ms^{-1}	1 Ms^{-1}
k_{-3}	1 Ms^{-1}	1 Ms^{-1}	1 Ms^{-1}
k_4	0.001 Ms^{-1}	0.001 Ms^{-1}	0.001 Ms^{-1}
k_{-4}	0.001 Ms^{-1}	0.001 Ms^{-1}	0.001 Ms^{-1}
Simulationszeit bis zum Gleichgewicht	5000s	5000s	5000s
Laufzeit (DGL)	<1 Sekunde	<1 Sekunde	<1 Sekunde
Laufzeit (Petrietz)	<1 Sekunde	1:13 min	20:44:36 h
Schrittweite Petrietz	0.1 Sekunden	1e-3 Sekunden	1e-6 Sekunden

Bei der Petrietz-Simulation wurde bei jedem Versuch die größtmögliche Schrittweite gewählt (vergleiche Kapitel 3.4.1, Seite 54). Wenn eine Simulation abbrach, ist sie mit geringerer Schrittweite wiederholt worden. Die Schrittweiten und Laufzeiten der Simulation finden sich ebenfalls in oben stehender Tabelle. Wie man sieht, verlängert sich die Laufzeit von Versuch zu Versuch drastisch. Für die notwendige Schrittweitenreduzierung sind die stark unterschiedlichen Reaktionsgeschwindigkeiten verantwortlich. Sie umfassen im dritten Versuch 8 Größenordnungen. Bei einer größeren Schrittweite würde bereits der erste diskrete Schritt für die schnelle Reaktion zu einem großen Ausschlag mit einem entsprechend großen Fehler führen, der in den folgenden Schritten nicht mehr kompensiert werden kann. Bei den langsamen Reaktionen gibt es bei kleinen Schrittweiten kaum Veränderung, so dass entsprechend viele Schritte gemacht werden müssen, was zu Lasten der Laufzeit geht. Das System ist steif (vergleiche Abschnitt 2.5.2, Seite 25).

Es wurde versucht, die Schrittweite zu erhöhen, sobald die schnelle Reaktion sich in einem Quasigleichgewicht befindet und sich die Konzentrationen der Reaktionspartner, B und C, nur noch geringfügig ändern. Wohin das führt, sieht man in Abbildung 33. Die Fehler in jedem Schritt addieren sich und das System gerät schnell außer Kontrolle. Der Grund ist die immer noch vorhandene Kopplung an die langsamen Reaktionen, die bewirken, dass sich die Konzentrationen der beteiligten Metabolite langsam verändern.

Auch wenn Systeme nicht steif sind, sind die Petrietze in der vorliegenden Implementierung den Integratoren für Differentialgleichungssysteme unterlegen, weil keine dynamische Anpassung der Schrittweite erfolgt. Die Algorithmen, mit denen Schrittweitenanpassung und die Lösung steifer Systeme effizient möglich wird, sind in diversen Implementierungen frei verfügbar. Eine Übertragung auf die Simulation mit

farbigen Petrinetzen wäre möglich, jedoch mit erheblichem Zeitaufwand verbunden. Als einziger Unterschied würde dann noch die Auslösung der Reaktionsereignisse übrig bleiben (sequentiell gegenüber parallel). Es ist zu erwarten, dass dieser Unterschied entweder nicht ins Gewicht fällt oder die Laufzeit zu Ungunsten der farbigen Petrinetze verschieben könnte, weil die sequentielle Ausführung einen zusätzlichen Störfaktor darstellt.

Um die Simulation mit farbigen Petrinetzen auch an realistischen Modellen in der Praxis zu testen, wurden die Modelle Glykolyse und Zentralstoffwechsel mit dem Petrinetz simuliert. Aufgrund der angesprochenen Problematik konnte bei akzeptabler Schrittweite kein Fließgleichgewichtszustand erreicht werden.

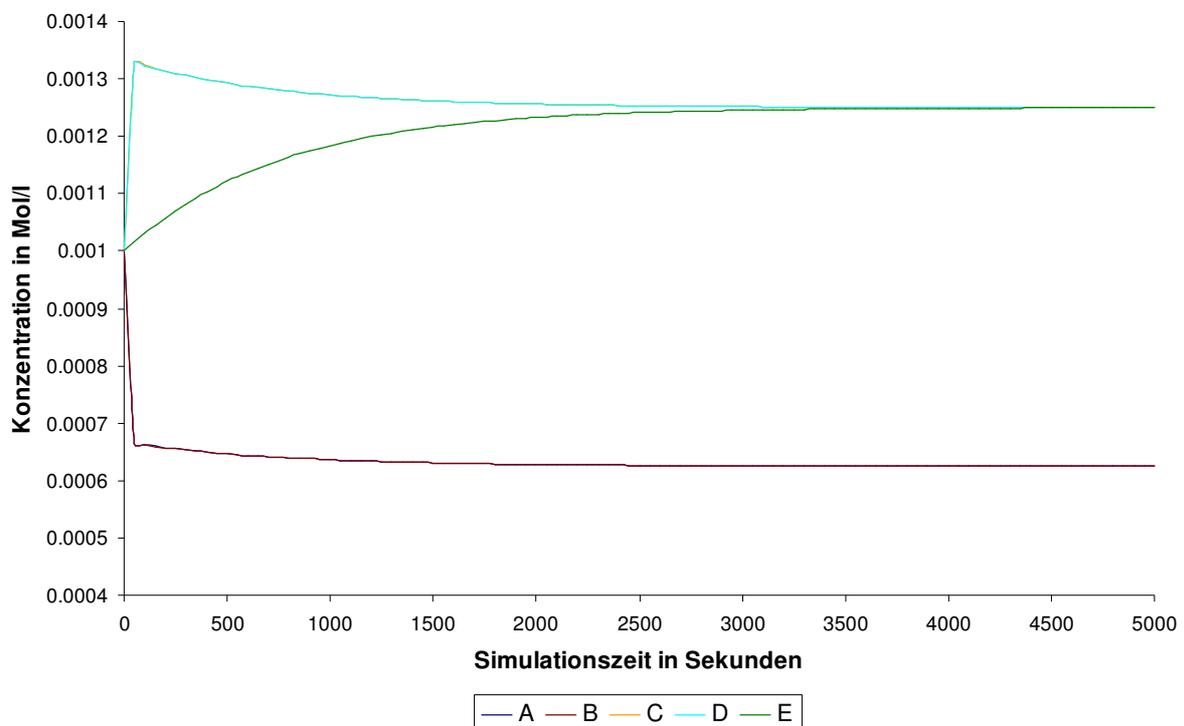


Abbildung 32 Simulationsverlauf für das lineare System aus Abbildung 31.



Abbildung 33 Petrinetzsimulation des linearen Modells bei zu großer Schrittweite.

4.8.2 Lösung des algebraischen Gleichungssystems

Ist man nur an den Werten der Variablen im Fließgleichgewicht und nicht an dem transienten Verlauf interessiert, kann man auch versuchen, diese ohne den Umweg über die Lösung des Differentialgleichungssystems zu berechnen. Ein Verfahren ist bereits in Abschnitt 2.5.1 (Seite 24) beschrieben worden. Die methodischen Details finden sich in Abschnitt 3.4.2 auf Seite 55.

Die Berechnung der Konzentrationen der Glykolyse im Fließgleichgewicht ist zunächst mit aufeinander abgestimmter Zu- und Abflussrate durchgeführt worden. Als Ergebnis erhält man eine Lösungsmenge (Tabelle 16). Das System ist also unterbestimmt und kann durch Fixierung einer beliebigen Konzentration eindeutig gelöst werden.

Die symbolische Lösung der algebraischen Gleichungen hat den Vorteil, dass man auf den ersten Blick sieht, wie viele Freiheitsgrade das System hat. Eine Integration der Differentialgleichungen findet bei unterbestimmten Systemen unter Umständen keinen Fließgleichgewichtszustand, ohne dabei nähere Informationen zu den Gründen zu liefern.

In Tabelle 16 in der letzten Spalte wird die Pyruvatkonzentration exemplarisch auf $1e-6$ mol/l fixiert, um eine eindeutige Lösung zu erhalten. Diese Lösung ist bei allen physiologischen Pyruvatkonzentrationen im Wesentlichen durch das absolute Glied der Formel festgelegt, also unabhängig von der Pyruvatkonzentration. Wählt man einen anderen Metabolit, um dessen Konzentration zu fixieren, so hat diese Konzentration

maßgeblichen Einfluss auf den Fließgleichgewichtszustand. Nur durch die Wahl des letzten Metaboliten im Pfad konnte diese Unabhängigkeit erreicht werden.

Es wurde auch versucht, das Fließgleichgewicht für das Modell des Zentralstoffwechsels auf diese Weise zu berechnen. Die Berechnung wurde nach 30 Stunden ohne Lösung abgebrochen. Das System ist offensichtlich zu kompliziert, um es auf diese Weise zu berechnen. Im Vergleich kam man durch Integration der Differentialgleichungen in wenigen Sekunden zu einer Lösung. Diese Erfahrung floss auch in die Entwicklung von Copasi (Hoops *et al.* 2006) ein. Das Programm versucht Gleichgewichte zunächst mit einem Newton-Verfahren das algebraische System zu lösen. Schlägt dies fehl, wird auf die Integration der Differentialgleichungen umgeschaltet.

Tabelle 16 Symbolische Berechnung des Fließgleichgewichts der Glykolyse.

Metabolit	Konzentrationsformel [mol/l]	Konzentration [mol/l] PYR=1e-6
GLC	$7,17e-26 \cdot \text{PYR}^2 + 1,81e-18 \cdot \text{PYR} + 1,18e-5$	1,18e-5
G6P	$4,06e-22 \cdot \text{PYR}^2 + 1,03e-14 \cdot \text{PYR} + 9,45e-6$	9,45e-6
F6P	$2,81e-20 \cdot \text{PYR}^2 + 7,09e-13 \cdot \text{PYR} + 1,20e-5$	1,20e-5
F16DP	$1,59e-16 \cdot \text{PYR}^2 + 4,02e-9 \cdot \text{PYR} + 9,28e-4$	9,28e-4
DHAP	$1,82e-13 \cdot \text{PYR} + 4,57e-6$	4,57e-6
3PGA	$1,32e-13 \cdot \text{PYR} + 2,87e-8$	2,87e-8
13DPG	$4,15e-10 \cdot \text{PYR} + 2,96e-5$	2,96e-5
3PG	$3,91e-7 \cdot \text{PYR} + 5,79e-4$	5,79e-4
2PG	$9,56e-9 \cdot \text{PYR} + 1,17e-5$	1,17e-5
PEP	$2,04e-6 \cdot \text{PYR} + 2,55e-4$	2,55e-4

4.9 Zusammenfassende Diskussion

Der Weg von einem Reaktionsdatensatz zu einem funktionsfähigen und aussagekräftigen kinetischen Modell besteht aus vielen Schritten. In dieser Arbeit sind einige Schritte in die richtige Richtung gemacht worden. Zunächst ging es darum, dem Modellierer bei der Konsistenzprüfung des Modells Hilfsmittel an die Hand zu geben. Die automatische Reaktionsbilanzierung ist eines dieser Hilfsmittel (Abschnitt 4.1, Seite 62). Damit ist man in der Lage, korrekte von unkorrekten Reaktionen zu trennen. Die unkorrekten Reaktionen werden automatisch korrigiert oder es wird nach Möglichkeit ein Korrekturvorschlag gemacht. Sehr wichtig bei diesem Prozess ist, dass es möglichst keine falsch positiven Korrekturen gibt. Dies kann man durch Post-Filter gewährleisten, die ungewöhnliche hohe Abweichungen von der Originalstöchiometrie bestrafen. Möglicherweise bringt der Ansatz, die Abweichung von der ursprünglichen Stöchiometrie so gering wie möglich zu halten, also diese Abweichung zu minimieren, noch bessere Ergebnisse.

Sind die Reaktionen einzeln korrekt, muss als nächstes die Konsistenz des gesamten Reaktionssatzes bezüglich der Systemgrenzen überprüft werden. Als erstes werden dazu alle losen Enden identifiziert. Außerdem werden die Conserved Moieties berechnet, die weitere Grenzen des Systems bilden (vergleiche Abschnitt 3.3.2, Seite 45). Die losen Enden und Conserved Moieties müssen entweder durch zusätzliche Reaktionen in das Netzwerk eingebunden oder gesondert behandelt werden. Im ersten Fall kann das Netzwerk deutlich komplizierter werden, was Probleme bei der Laufzeit und dem Auffinden eines Fließgleichgewichtszustandes nach sich zieht. Im zweiten Fall muss man darauf achten, dass das Verhalten des Systems nicht durch eine unkorrekte Behandlung der Systemgrenzen verfälscht wird. Man sollte hier auf biochemisches Wissen setzen, also die Konzentrationen aus dem Experiment nach Möglichkeit in das Modell einbringen.

Hat man die Struktur des Modells fertig gestellt, müssen anschließend die kinetischen Gleichungen und darin enthaltene Parameter bestimmt werden. Die Gleichgewichtskonstante, die in Abhängigkeit von den Konzentrationsvariablen die Reaktionsrichtung bestimmt, ist einer dieser Parameter. Sie kann aus der Standard Gibbsenergie der Reaktion leicht berechnet werden, die wiederum abgeschätzt werden kann. Zwei Methoden, die eine genauer (Methode nach Alberty), die andere breiter anwendbar (Methode nach Mavrovouniotis), wurden hier evaluiert (Abschnitt 4.2.2, Seite 68). Schon der mittlere Fehler der exakteren Methode nach Alberty ist mit 4,4 kJ/mol als hoch zu bewerten. Man muss bedenken, dass eine Änderung der Standard Gibbsenergie der Reaktion um 5,8 kJ/mol wegen des exponentiellen Zusammenhangs bereits eine Änderung der Gleichgewichtskonstante um den Faktor 10 ausmacht. Die verwendeten Werte der Methode nach Mavrovouniotis weichen somit im Mittel ungefähr um den Faktor

1000 von den realen Werten ab. Sie können also nur als grobes Maß angesehen werden. Es würde nahe liegen, grundsätzlich die Gleichgewichtskonstanten der exakteren Methode zu verwenden und diese bei Bedarf durch die Konstanten der ungenaueren Methode zu ersetzen. Leider ist aufgrund der unterschiedlichen Konzepte eine Kombination nicht möglich, ohne gegen thermodynamische Grundsätze zu verstoßen. Bei der Gibbsenergie handelt es sich um eine Zustandsfunktion. Eine Mischung beider Methoden würde diesen Grundsatz verletzen, weil die Gibbsenergie über einen geschlossenen Kreislauf nicht notwendigerweise null wäre.

Verwendet man chemische Kinetik als Ratengleichung, wozu man angesichts der Datenlage nur wenige Alternativen hat, bleibt pro Reaktion ein unbekannter Parameter übrig. Über diese Parameter kann das Systemverhalten beeinflusst werden. Dies ist der schwierigste Teil der Modellierung, da das gesamte System auf eine Parameterveränderung nicht-linear reagiert. Das erste Ziel ist daher, überhaupt eine Parameterkombination zu finden, mit der ein stabiles Fließgleichgewicht erreicht werden kann. Wir werden am Ende nochmals auf diesen Punkt zurückkommen. Hat man einen Fließgleichgewichtszustand gefunden, können die Parameter mit Hilfe der Kontrollkoeffizienten (vergleiche Abschnitt 3.3.3, Seite 46) so verändert werden, dass ausgewählte Konzentrationsvariablen den gewünschten Wert erhalten.

Um das Verhalten von metabolischen Netzwerken näher zu untersuchen, wurden zunächst kleine artifizielle Modelle formuliert. Aus den Untersuchungen des linearen Modells konnten bereits wichtige Schlüsse gezogen werden. Die unterschiedlichen Zu- und Abflussfunktionen führen auch bei gleichem makroskopisch beobachtbarem Zustand, also den gleichen Stoffflüssen und Konzentrationen im Fließgleichgewicht, zu einem unterschiedlichen Kontrollverhalten. Bei einer Parameterveränderung würde sich also ein jeweils anderer Fließgleichgewichtszustand einstellen. Man kann davon ausgehen, dass auch die externen Reaktionen prinzipiell reversibel sind, weshalb mutmaßlich die Modellierung mit reversibler Kinetik und konstanten Konzentrationen der externen Metabolite das Verhalten bei Veränderungen korrekt beschreibt. Mit dem linearen Modell wurde außerdem versucht, die Michaelis-Menten-Kinetik durch chemische Kinetik anzunähern. Dies kann prinzipiell nur dann funktionieren, wenn keine Sättigungseffekte eintreten. Dies ist entweder bei geringen Substratkonzentrationen der Fall oder bei hohen Umsatzraten, also hohem V_{\max} (vergleiche Abbildung 13, Seite 44). Die Kontrollkoeffizienten unterscheiden sich in jedem Fall, so dass aus einer Parameterveränderung wiederum unterschiedliches Verhalten resultiert.

Beim Vergleich der Modelle mit Verzweigung (Abschnitt 3.3.4.2 auf 48 und Abschnitt 4.4 auf Seite 77) sollte festgestellt werden, welchen Effekt die Einführung eines künstlichen Metaboliten hat, der im Modell den Aufbau von Biomasse repräsentieren soll. Die

Tatsache, dass in beiden Pfaden der gleiche Stofffluss herrschen muss, führt zu einem zusätzlichen Zwang. Da das Modell bei Veränderung der Enzymparameter nicht mit einer Flussänderung reagieren kann, muss es die Metabolitkonzentrationen stärker verändern, um sich auf die neue Situation einzustellen. Erst eine weitere Anpassung der Enzymparameter einer internen Reaktion (R1) sorgt dafür, dass nach der Störung in beiden Fällen der gleiche neue Fließgleichgewichtszustand hergestellt wird.

Während des Modellierungsprozesses der realen Systeme ist klar geworden, dass man auf gute Schätzungen für die Parameter angewiesen ist, um einen Fließgleichgewichtszustand erreichen. Es lässt sich nicht a priori feststellen, ob zu einem Parametersatz ein Fließgleichgewichtszustand existiert. Als guter Startpunkt hat es sich bewährt, bei Kenntnis der Reaktionsrichtung die Geschwindigkeitskonstanten der Hinreaktionen in die gleiche Größenordnung zu bringen. Weitere Variationen der Parameter sind praktisch immer notwendig (siehe Abschnitt 4.5, Seite 79). Da man mit wachsender Modellgröße immer mehr „Stellschrauben“ hat, wird es schwierig, das System durch manuelle Korrekturen ins Fließgleichgewicht zu bringen. Eine nicht-lineare Optimierung der Geschwindigkeitskonstanten könnte hier weiterhelfen. Angesichts der üblicherweise zunehmenden Steifheit von größeren Systemen sind die einzelnen Berechnungen des Fließgleichgewichts allerdings teuer – selbst bei Algorithmen, die auf steife Probleme spezialisiert sind.

Die anderen getesteten Methoden zur Berechnung des Fließgleichgewichts sind keine echte Alternative. Die Simulation farbiger Petrinetze kann in dieser Form nicht mit steifen Problemen umgehen und eignet sich deshalb nicht für biochemische Netzwerke (Abschnitt 4.8.1, Seite 92). Die etablierten Analysemethoden, die die Petrinetz-Theorie bietet, wie Erreichbarkeitsanalyse, Invariantenanalyse etc., können aufgrund des großen Zustandsraums bei farbigen Petrinetzen mit kontinuierlichen Variablen nicht eingesetzt werden. Möchte man biochemische Netzwerke mit Petrinetzen analysieren, sollte man sich nach derzeitigem Kenntnisstand auf die stöchiometrische Analyse mit Stellen-Transitions-Netzen oder farbigen Netzen mit nicht-kontinuierlichen Variablen beschränken, um die mathematischen Stärken der Petrinetztheorie ausnutzen zu können. Auch das getestete Lösungsverfahren für das nicht-lineare algebraische Gleichungssystem (siehe Abschnitt 4.8.2, Seite 95) eignet sich nicht für größere Systeme.

Das unterschiedliche Verhalten bei Parameterveränderungen zog sich wie ein roter Faden durch alle durchgeführten Computereperimente. Das eigentliche Ziel sind Modelle, die nicht nur einen bestimmten Zustand beschreiben, sondern bei Veränderungen am Modell auch andere Zustände vorhersagen können. Dazu müssen, wie wir gesehen haben, die korrekten Ratengleichungen sowohl für interne als auch für externe Reaktionen verwendet

werden. Eine vereinfachte Beschreibung durch chemische Kinetik oder konstante Raten kann dies nicht leisten.

Trotzdem ist es möglich, auf diese Weise physiologisch sinnvolle und stabile Fließgleichgewichtszustände zu erreichen. Inwieweit diese mit dem Experiment korrelieren, kann nur durch die Durchführung entsprechender Experimente nachgewiesen werden.

5 Ausblick

In der vorliegenden Arbeit ist es gelungen, eine kinetische Simulation verschiedener metabolischer Netzwerke ohne Kenntnis der kinetischen Parameter durchzuführen. Die Abschätzung der thermodynamischen Parameter und die Verwendung von Daten aus der Primärliteratur führten zu einem sinnvollen Fließgleichgewichtszustand. Bei der Beschreibung der einzelnen Reaktionen durch chemische Kinetik bleibt ein freier Parameter pro Reaktion übrig, der abgeschätzt bzw. verändert werden musste. Über die Variation dieser Parameter lassen sich unendlich viele weitere Fließgleichgewichtszustände erzeugen, die erheblich voneinander abweichen können.

Zurzeit werden Anstrengungen am Institut unternommen, um die Konzentrationen möglichst vieler Metabolite absolut zu quantifizieren. Bei Kenntnis der Metabolitkonzentrationen aus dem Experiment sollte dieser freie Parameter genutzt werden, um das Modell an die experimentellen Daten anzupassen. Für diese Anpassung ist eine nicht-lineare Optimierung notwendig. P. Mendes und D. Kell haben in (Mendes und Kell 1998) 12 Algorithmen evaluiert, die für diese Prozedur bei Anwendung auf biochemische Netzwerke geeignet schienen. Aus den Untersuchungen lässt sich schließen, dass man immer mehrere Algorithmen für ein gegebenes Problem testen sollte, weil nicht a priori festgestellt werden kann, welches der am besten geeignete Algorithmus ist.

Wie ausführlich dargelegt wurde, kann man ohne korrekte kinetische Beschreibung der einzelnen Reaktionen im Bottom-up-Ansatz (siehe Abschnitt 1.3) keine prädiktiven Computerexperimente durchführen. Bei Veränderung eines Parameters bekommt man einen falschen neuen Fließgleichgewichtszustand. Wenn die Kontrollkoeffizienten nicht mit denen des realen Systems übereinstimmen, muss das Modell verändert werden. Dazu gibt es mehrere Möglichkeiten. Die entsprechenden Konstanten (K_m -Werte) können experimentell bestimmt werden, wobei beachtet werden muss, dass es Unterschiede zwischen dem *in vitro* und *in vivo* Verhalten geben kann (Teusink *et al.* 2000). Mit dem V_{max} -Wert würde wiederum ein unbekannter Parameter übrig bleiben, der angepasst werden kann, wenn die Enzymkonzentration oder die Enzymaktivität nicht bekannt ist. Eine andere Möglichkeit besteht darin, anstelle von korrekten Ratengleichungen phänomenologische Ratengleichungen zu verwenden und über viele Messungen bei unterschiedlichen Bedingungen die Parameter so anzupassen, dass die korrekte Antwort des Systems auf eine Störung erfolgt.

Eine Erweiterung des Netzwerks durch quantifizierte Effekte von Inhibitoren bzw. Aktivatoren der Enzyme wäre ein weiterer Schritt in Richtung von Modellen, die für realistische Vorhersagen verwendet werden können. Die Feedback-Inhibition ist zum

Beispiel ein von der Zelle häufig genutzter Mechanismus, um bei veränderten Stoffflüssen die Konzentrationen dahingehend zu kontrollieren, dass sie in physiologischen Bereichen bleiben (Cornish-Bowden und Cárdenas 2005).

Ein Aspekt, der hier unberücksichtigt bleiben konnte, ist das Auftreten von Oszillationen. Oszillierende Metabolitkonzentrationen können häufig auftreten, wenn regulative Prozesse Teil des Modells sind. In welchen Fällen es zu Oszillationen kommt und wie stark sie sind, wurde unlängst in (Reijenga *et al.* 2005) untersucht. Ein verwandtes Phänomen ist das Auftreten von multiplen stationären Zuständen. Das Auftreten beider Phänomene in Abhängigkeit von den Modellparametern kann durch Bifurkationsanalyse aufgeklärt werden (Chickarmane *et al.* 2005).

Stochastische Effekte sollten bei geringen Teilchenzahlen, wie sie in der Zelle auftreten können, berücksichtigt werden. Leider sind die zur Verfügung stehenden Algorithmen langsam gegenüber deterministischen Algorithmen, obgleich es in den letzten Jahren deutlich Fortschritte gegeben hat (Gillespie 2006). Die Kombination von deterministischer Simulation bei hohen Teilchenzahlen und stochastischer Simulation bei geringen Teilchenzahlen, also eine hybride Simulation, kann hier teilweise Abhilfe schaffen (Griffith *et al.* 2006).

Auf dem Weg zur virtuellen Zelle sind noch viele Fragen unbeantwortet. Nur eine ständige Weiterentwicklung der Algorithmen, vor allem aber das Vorantreiben der experimentellen Methodenentwicklung und die Produktion von kohärenten Datensätzen werden langfristig zum Erfolg führen.

6 Anhang

6.1 Inkremente zur Gibbsenergieberechnung nach Mavrovouniotis

Tabelle 17 Gruppenbeiträge zur Gibbsenergie nach (Mavrovouniotis 1991). Neben den gezeigten Beiträgen müssen noch Korrekturfaktoren für das gesamte Molekül berücksichtigt werden.

Substruktur	Inkrement [kcal/mol]	Substruktur	Inkrement [kcal/mol]
$-\text{SO}_3^{1-}$	-105,8	$-\text{OH}$ (primär)	-29,3
$-\text{SH}$	13,4	$-\text{OH}$ (sekundär)	-32,0
$-\text{S}-\text{S}-$	5,8	$-\text{OH}$ (tertiär)	-30,5
$-\text{S}-$	9,5	$-\text{COO}^{1-}$	-72,0
$-\text{O}-\text{PO}_2^{1-}$ (Ring)	14,8	$-\text{CO}-\text{O}-$	-73,6
$-\text{CO}-\text{OPO}_3^{2-}$	-72,5	$>\text{CO}$	-27,2
$-\text{OPO}_3^{2-}$ (primär)	-29,5	$-\text{O}-$	-22,5
$-\text{OPO}_3^{2-}$ (sekundär)	-30,0	$-\text{O}^{1-}$	-44,8
$-\text{OPO}_3^{2-}$ (tertiär)	-25,7	$>\text{C}=\text{C}$ (2 Benzenringe)	2,5
$-\text{PO}_3^{2-}$	9,5	$>\text{C}=\text{C}$ (2 Ringe)	16,8
$-\text{O}-\text{PO}_2^{1-}-\text{O}-$	-29,8	$>\text{C}=\text{C}$ (Benzenring & Ring)	6,0
$-\text{O}-\text{PO}_2^{1-}-$	-5,2	$-\text{CH}=\text{C}$ (Benzenring)	8,4
$-\text{N}<$ (2 Ringe)	18,9	$>\text{C}=\text{C}$ (Benzenring)	1,5
$>\text{N}^{1+}=\text{C}$ (Ring)	0,4	$>\text{C}=\text{C}$ (Einfachbindungen in Ring)	22,8
$=\text{N}-$ (Ring)	10,4	$-\text{CH}=\text{C}$ (Ring)	9,6
$>\text{NH}$ (Ring)	9,5	$>\text{C}=\text{C}$ (Doppelbindung in Ring)	8,2
$-\text{N}<$ (Ring)	7,6	$\equiv\text{CH}$	35,7
$-\text{NH}_2$	10,3	$\equiv\text{C}-$	24,0
$-\text{NH}_3^{1+}$	4,3	$-\text{CH}=\text{C}$	11,1
$\equiv\text{N}$	14,9	$=\text{CH}_2$	18,4
$=\text{NH}_2^{1+}$	0,4	$>\text{C}=\text{C}$	5,0
$=\text{NH}$	13,6	$>\text{CH}-$ (2 Ringe)	-0,9
$>\text{NH}_2^{1+}$	6,9	$>\text{C}<$ (2 Ringe)	-12,0
$>\text{NH}$	7,6	$>\text{CH}_2$ (Ring)	6,1
$>\text{N}-$	6,3	$>\text{CH}-$ (Ring)	-2,2
$>\text{NH}^{1+}-$	8,6	$>\text{C}<$ (Ring)	-12,8
$-\text{O}-\text{CO}-$	-54,6	$-\text{CH}_3$	7,9
$>\text{CO}$	-27,4	$>\text{CH}_2$	1,7
$-\text{O}-$	-24,3	$>\text{CH}-$	-4,8
$-\text{CH}=\text{O}$	-17,8	$>\text{C}<$	-12,8
$-\text{OH}$ (Phenyl)	-31,8		

6.2 Modell Glykolyse: Reaktionen und Parameter

Tabelle 18 Reaktionen des Glykolysemodells.

Enzym	Reaktion
Glucokinase	Glukose + ATP → G6P + ADP
Glucose-6-phosphat Isomerase	G6P → F6P
6-Phosphofruktokinase	F6P + ATP → F16DP + ADP
Fructose-bisphosphat Aldolase	F16DP → Glyceronphosphat + Glyceraldehyd-3-phosphat
Triosephosphat Isomerase	Glyceronphosphat → Glyceraldehyd-3-phosphat
Glyceraldehyd-3-phosphat Dehydrogenase	Glyceraldehyd-3-phosphat + NAD + Phosphat → 3-Phosphoglycerolphosphat + NADH
Phosphoglycerat Kinase	3-Phosphoglyceroylphosphat + ADP → 3-Phosphoglycerat + ATP
Phosphoglycerat Mutase	3-Phosphoglycerat → 2-Phosphoglycerat
Phosphoenolpyruvat Hydratase	2-Phosphoglycerat → Phosphoenolpyruvat + H ₂ O
Pyruvatkinase	Phosphoenolpyruvat + ADP → Pyruvat + ATP

Tabelle 19 Parameter und fixierte Konzentrationen des Glykolysemodells.

Parameter	Wert	Referenz
Zellvolumen	1,25e-16 l	(Krömer 2006)
Zuflussrate Glukose	0,771 mM/s	(Zittrich und Krämer 1994)
Glucokinase	$k_1=43,5 \text{ s}^{-1}\text{mM}^{-1}$, $k_{-1}=2,30 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
Glucose-6-phosphat Isomerase	$k_1=83,1 \text{ s}^{-1}$, $k_{-1}=1,20 \text{ s}^{-1}$	Abgeschätzt
6-Phosphofruktokinase	$k_1=43,5 \text{ s}^{-1}\text{mM}^{-1}$, $k_{-1}=2,30 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
Fructose-bisphosphat Aldolase	$k_1=12,3 \text{ s}^{-1}$, $k_{-1}=8,12\text{e}+4 \text{ s}^{-1}$	Abgeschätzt
Triosephosphat Isomerase	$k_1=8,51 \text{ s}^{-1}$, $k_{-1}=117 \text{ s}^{-1}$	Abgeschätzt
Glyceraldehyd-3-phosphat Dehydrogenase	$k_1=10,0 \text{ s}^{-1}\text{mM}^{-2}$, $k_{-1}=12,7 \text{ s}^{-1}\text{mM}^{-2}$	Abgeschätzt
Phosphoglycerat Kinase	$k_1=53,2 \text{ s}^{-1}\text{mM}^{-1}$, $k_{-1}=1,88 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
Phosphoglycerat Mutase	$k_1=15,6 \text{ s}^{-1}$, $k_{-1}=64,0 \text{ s}^{-1}$	Abgeschätzt
Phosphoenolpyruvat Hydratase	$k_1=14,6 \text{ s}^{-1}$, $k_{-1}=6,84 \text{ s}^{-1}$	Abgeschätzt
Pyruvatkinase	$k_1=12,1 \text{ s}^{-1}\text{mM}^{-1}$, $k_{-1}=8,25\text{e}-2 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
[ATP]	1,5 mM	(Moritz <i>et al.</i> 2000)
[ADP]	0,5 mM	(Moritz <i>et al.</i> 2000)
[NADH]	0,1 mM	(Dominguez <i>et al.</i> 1998)
[NAD]	0,4 mM	(Dominguez <i>et al.</i> 1998)
[P]	1,0 mM	Abgeschätzt
[Pyruvat]	0,001 mM	Abgeschätzt

6.3 Modell Citratcyclus: Reaktionen und Parameter

Tabelle 20 Reaktionen des Citratcyclusmodells.

Enzym	Reaktion
Pyruvat Dehydrogenase	Pyruvat + CoA + NAD -> Acetyl-CoA + HCO ₃ + NADH
Citrat Synthase	Acetyl-CoA + Oxalacetat -> Citrat + CoA
Aconitase	Citrat -> Isocitrat
Isocitrat Dehydrogenase	Isocitrat + NADP -> 2-Oxoglutarat + HCO ₃ + NADPH
2-Oxoglutarat Dehydrogenase	2-Oxoglutarat + CoA + NAD -> Succinyl-CoA + HCO ₃ + NADH
Succinyl-CoA Synthetase	Succinyl-CoA + ADP + Phosphat -> Succinat + CoA + ATP
Succinat Dehydrogenase	Succinat + FAD -> Fumarat + FADH ₂
Fumarase	Fumarat -> Malat
Malat Dehydrogenase	Malat + NAD -> Oxalacetat + NADH
Pyruvat Carboxylase	Pyruvat + ATP + HCO ₃ -> Oxalacetat + ADP + P

Tabelle 21 Parameter und fixierte Konzentrationen des Citratcyclusmodells.

Parameter	Wert	Referenz
Zellvolumen	1.25e-16 l	(Krömer 2006)
Zuflussrate Pyruvat	1,442 mM/s	Aus Glykolyse
Pyruvat Dehydrogenase	$k_1=4000 \text{ s}^{-1}\text{mM}^{-2}$, $k_{-1}=1,08\text{e-}3 \text{ s}^{-1}\text{mM}^{-2}$	Abgeschätzt
Citrat Synthase	$k_1=4000 \text{ s}^{-1}\text{mM}^{-1}$, $k_{-1}=2,05\text{e-}3 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
Aconitase	$k_1=40,0 \text{ s}^{-1}$, $k_{-1}=40,0 \text{ s}^{-1}$	Abgeschätzt
Isocitrat Dehydrogenase	$k_1=40,0 \text{ s}^{-1}\text{mM}^{-1}$, $k_{-1}=0.454 \text{ s}^{-1}\text{mM}^{-2}$	Abgeschätzt
2-Oxoglutarat Dehydrogenase	$k_1=400 \text{ s}^{-1}\text{mM}^{-1}$, $k_{-1}=1,08\text{e-}4 \text{ s}^{-1}\text{mM}^{-2}$	Abgeschätzt
Succinyl-CoA Synthetase	$k_1=40,0 \text{ s}^{-1}\text{mM}^{-2}$, $k_{-1}=29,0 \text{ s}^{-1}\text{mM}^{-2}$	Abgeschätzt
Succinat Dehydrogenase	$k_1=40,0 \text{ s}^{-1}\text{mM}^{-1}$, $k_{-1}=1,91\text{e+}5 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
Fumarase	$k_1=4000 \text{ s}^{-1}$, $k_{-1}=22,0 \text{ s}^{-1}$	Abgeschätzt
Malat Dehydrogenase	$k_1=40,0 \text{ s}^{-1}\text{mM}^{-1}$, $k_{-1}=1,44\text{e+}5 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
Pyruvat Carboxylase	$k_1=40,0 \text{ s}^{-1}\text{mM}^{-2}$, $k_{-1}=1,65 \text{ s}^{-1}\text{mM}^{-2}$	Abgeschätzt
[ATP]	1,5 mM	(Moritz <i>et al.</i> 2000)
[ADP]	0,5 mM	(Moritz <i>et al.</i> 2000)
[NADH]	0,1 mM	(Dominguez <i>et al.</i> 1998)
[NAD]	0,4 mM	(Dominguez <i>et al.</i> 1998)
[P]	1,0 mM	Abgeschätzt
[NADPH]	0,2 mM	(Moritz <i>et al.</i> 2002)
[NADP]	0,1 mM	(Moritz <i>et al.</i> 2002)
[FADH ₂]	0,1 mM	Abgeschätzt
[FAD]	0,1 mM	Abgeschätzt
[HCO ₃]	0,1 mM	Abgeschätzt

6.4 Modell Zentralmetabolismus: Reaktionen, Parameter, Biomasse

Tabelle 22 Zusätzliche Reaktionen des Zentralmetabolismus (neben Tabelle 18, Tabelle 20).

Enzym	Reaktion
Glukose-6-phosphat Dehydrogenase	$G6P + NADP \rightarrow G15L6P + NADPH$
6-Phosphogluconolactonase	$G15L6P \rightarrow 6PGLU$
Phosphogluconate Dehydrogenase	$6PGLU + NADP \rightarrow RU5P + HCO_3 + NADPH$
Ribulosephosphat Epimerase	$RU5P \rightarrow XU5P$
Ribose-5-phosphat Isomerase	$RU5P \rightarrow R5P$
Transketolase 1	$RU5P + XU5P \rightarrow SHEP7P + 3PGA$
Transketolase 2	$E4P + XU5P \rightarrow F6P + 3PGA$
Transaldolase	$E4P + F6P \rightarrow SHEP7P + 3PGA$
Ribosephosphat Diphosphokinase	$R5P + ATP \rightarrow PRPP + AMP$

Tabelle 23 Zusätzliche Parameter des Zentralstoffwechselmodells (neben Tabelle 19, Tabelle 21).

Parameter	Wert	Referenz
Glukose-6-phosphat Dehydrogenase	$k_1=400 \text{ s}^{-1}\text{mM}^{-1}, k_{-1}=93,6 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
6-Phosphogluconolactonase	$k_1=40,0 \text{ s}^{-1}, k_{-1}=3.53\text{e-}10 \text{ s}^{-1}$	Abgeschätzt
Phosphogluconate Dehydrogenase	$k_1=40,0 \text{ s}^{-1}\text{mM}^{-1}, k_{-1}=43,4 \text{ s}^{-1}\text{mM}^2$	Abgeschätzt
Ribulosephosphat Epimerase	$k_1=40,0 \text{ s}^{-1}, k_{-1}=40,0 \text{ s}^{-1}$	Abgeschätzt
Ribose-5-phosphat Isomerase	$k_1=40 \text{ s}^{-1}, k_{-1}=12,4 \text{ s}^{-1}$	Abgeschätzt
Transketolase 1	$k_1=40,0 \text{ s}^{-1}\text{mM}^{-1}, k_{-1}=57,5 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
Transketolase 2	$k_1=400 \text{ s}^{-1}\text{mM}^{-1}, k_{-1}=146 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
Transaldolase	$k_1=40,0 \text{ s}^{-1}\text{mM}^{-1}, k_{-1}=114 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
Ribosephosphat Diphosphokinase	$k_1=40,0 \text{ s}^{-1}\text{mM}^{-1}, k_{-1}=2,02 \text{ s}^{-1}\text{mM}^{-1}$	Abgeschätzt
[AMP]	0,1 mM	Abgeschätzt

Tabelle 24 Biomassezusammensetzung, die als Grundlage für Tabelle 25 verwendet wurde.

0,58 GLY + 0,2776 GLU + 0,25 GLN + 0,17 PHE + 0,13 TYR + 0,377 SER + 0,05 TRP + 0,22 ASP + 0,22 ASN + 0,32 LYS + 0,09 CYS + 0,14 MET + 0,24 THR + 0,27 ILE + 0,2 PRO + 0,28 ARG + 0,5352 ALA + 0,4 VAL + 0,42 LEU + 0,09 HIS + 0,185 ATP + 0,2 GTP + 0,13 UTP + 0,12 CTP + 0,02 DATP + 0,02 DGTP + 0,02 DCTP + 0,02 DTTP + 0,129 GL3P + 0,258 C160 + 0,1697 UDPG + 0,0235 C140 + 0,0235 C141 + 0,0235 CMPKDO + 0,0235 ADPHEP + 0,0433 UDPNAG + 0,0276 UDPNAM + 0,0276 DAPIM + 0,0593 ORN -> Biomasse

Tabelle 25 Metabolite mit Zu- und Abflüssen und deren Äquivalente in der Biomasse.

Biomasse-Metabolit	Biosyntheseäquivalente aus dem Modell
HIS	PRPP (- AICAR->IMP)
GLU	2OG
GLN	2OG
PRO	2OG
ARG	2OG + OA + ACCOA - COA - FUM
ALA	PYR
SER	3PG
GLY	3PG
CYS	3PG
ASP	OA
ASN	OA
THR	OA
LYS	OA + PYR
MET	OA + ACCOA - COA
VAL	2*PYR
ILE	OA
LEU	2*PYR + ACCOA - COA
PHE	2*PEP + E4P
TYR	2*PEP + E4P
TRP	2*PEP + E4P + PRPP + 3PG - PYR - 3PGA
dTTP	PRPP + 3PG
(d)A(X)P	PRPP + 3PG + OA - FUM
(d)G(X)P	PRPP + 3PG
(d)U(X)P	PRPP + OA
(d)C(X)P	PRPP + OA
CMPKDO	RU5P + PEP + OA + PRPP
C140, C141	7*ACCOA - 7*COA
C160	8*ACCOA - 8*COA
ADPHEP	SHEP7P
UDPNAG	F6P + OA + PRPP + ACCOA - COA
UDPNAM	PEP + F6P + OA + PRPP + ACCOA - COA
DAPIM	OA + PYR
ORN	2PG + ACCOA - COA
GL3P	DHAP
UDPG	PRPP + OA + G6P

7 Literatur

- Alberty, R. A. (2006). Biochemical Thermodynamics: Applications of Mathematica, John Wiley & Sons, Inc.
- Atkins, P. W. (1996). Physikalische Chemie, Wiley-VCH.
- Babu, C. V. S., E. J. Song, *et al.* (2006). "Modeling and simulation in signal transduction pathways: a systems biology approach." Biochimie **88**(3-4): 277-283.
- Barman, T. E. (1969). Enzyme Handbook, Springer-Verlag, New York.
- Baumgarten, B. (1996). Petri-Netze - Grundlagen und Anwendungen. Heidelberg, Spektrum Akademischer Verlag.
- Beard, D. A., S.-d. Liang, *et al.* (2002). "Energy balance for analysis of complex metabolic networks." Biophys J **83**(1): 79-86.
- Benson, S. W. (1976). Thermochemical Kinetics, John Wiley & Sons, New York.
- Borodina, I. und J. Nielsen (2005). "From genomes to in silico cells via metabolic networks." Curr Opin Biotechnol **16**(3): 350-355.
- Buchinger, S. (2006). Persönliche Kommunikation.
- Burgard, A. P. und C. D. Maranas (2003). "Optimization-based framework for inferring and testing hypothesized metabolic objective functions." Biotechnol Bioeng **82**(6): 670-677.
- Caspi, R., H. Foerster, *et al.* (2006). "MetaCyc: a multiorganism database of metabolic pathways and enzymes." Nucleic Acids Res **34**(Database issue): D511-D516.
- Chance, B. (1943). "The Kinetics of the Enzyme-Substrate Compound of Peroxidase." J Biol Chem **151**: 553-577.
- Chassagnole, C., N. Noisommit-Rizzi, *et al.* (2002). "Dynamic Modeling of the Central Carbon Metabolism of Escherichia coli." Biotech Bioeng **79**: 53-73.
- Chassagnole, C., B. Raïs, *et al.* (2001). "An integrated study of threonine-pathway enzyme kinetics in Escherichia coli." Biochem J **356**(Pt 2): 415-423.
- Chickarmane, V., S. R. Paladugu, *et al.* (2005). "Bifurcation discovery tool." Bioinformatics **21**(18): 3688-3690.
- Cornish-Bowden, A. und M. L. Cárdenas (2005). "Systems biology may work when we learn to understand the parts in terms of the whole." Biochem Soc Trans **33**(Pt 3): 516-519.
- Cortassa, S. und M. A. Aon (1997). "Distributed control of the glycolytic flux in wild-type cells and catabolite repression mutants of Saccharomyces cerevisiae growing in carbon-limited chemostat cultures." Enz Microb Technol **21**: 596-692.
- Deckard, A., F. T. Bergmann, *et al.* (2006). "Supporting the SBML layout extension." Bioinformatics **22**(23): 2966-2967.

- Devaux, F., P. Marc, *et al.* (2001). "Transcriptomes, transcription activators and microarrays." FEBS Lett **498**(2-3): 140-144.
- Dominguez, H., C. Rollin, *et al.* (1998). "Carbon-flux distribution in the central metabolic pathways of *Corynebacterium glutamicum* during growth on fructose." Eur J Biochem **254**(1): 96-102.
- Edsall, J. T. und H. Gutfreund (1983). Biothermodynamics, John Wiley & Sons, New York.
- Edwards, J. S. und B. O. Palsson (1998). "How will bioinformatics influence metabolic engineering?" Biotechnol Bioeng **58**(2-3): 162-169.
- Eggeling, L. und M. Bott (2005). Handbook of Corynebacterium Glutamicum, CRC Pr Inc.
- Evangelisti, A. M. und A. Wagner (2004). "Molecular evolution in the yeast transcriptional regulation network." J Exp Zool B Mol Dev Evol **302**(4): 392-411.
- Fallisch, T. (2006). Uniting In Silico and In Vivo Systems Biology: a New Concept to Approximate Theory to Real-Life Flux Distributions, Universität zu Köln.
- Francke, C., R. J. Siezen, *et al.* (2005). "Reconstructing the metabolic network of a bacterium from its genome." Trends Microbiol **13**(11): 550-558.
- Galazzo, J. L. und J. E. Bailey (1990). "Fermentation pathway kinetics and metabolic flux control in suspended and immobilized *Saccharomyces cerevisiae*." Enz Microbiol Technol **12**: 162-172.
- Garfinkel, D., L. Garfinkel, *et al.* (1970). "Computer applications to biochemical kinetics." Annu Rev Biochem **39**: 473-498.
- Garfinkel, D., W. Polk, *et al.* (1962). "Simulation and analysis of biochemical systems: III. analysis and pattern recognition." Commun. ACM **5**(2): 115-118.
- Garfinkel, D., J. D. Rutledge, *et al.* (1961). "Simulation and analysis of biochemical systems: I. representation of chemical kinetics." Commun ACM **4**: 559-562.
- Gauges, R., U. Rost, *et al.* (2006). "A model diagram layout extension for SBML." Bioinformatics **22**(15): 1879-1885.
- Genrich, H., R. Küffner, *et al.* (2001). "Executable Petri net models for the analysis of metabolic pathways." Int J STTT **3**: 394-404.
- Gillespie, D. T. (2006). "Stochastic Simulation of Chemical Kinetics." Annu Rev Phys Chem.
- Goldberg, R. N., Y. B. Tewari, *et al.* (2004). "Thermodynamics of enzyme-catalyzed reactions-a database for quantitative biochemistry." Bioinformatics **20**(16): 2874-2877.
- Goodwell, D. S. (1993). The Machinery of Life. Berlin, Springer-Verlag.
- Griffith, M., T. Courtney, *et al.* (2006). "Dynamic partitioning for hybrid simulation of the bistable HIV-1 transactivation network." Bioinformatics **22**(22): 2782-2789.
- Heinrich, R., S. M. Rapoport, *et al.* (1977). "Metabolic regulation and mathematical models." Prog Biophys Mol Biol **32**(1): 1-82.

- Heinrich, R. und T. A. Rapoport (1974). "A linear steady-state treatment of enzymatic chains. General properties, control and effector strength." Eur J Biochem **42**(1): 89-95.
- Heinrich, R. und S. Schuster (1996). The Regulation Of Cellular Systems, Springer.
- Helfert, S., A. M. Estévez, *et al.* (2001). "Roles of triosephosphate isomerase and aerobic metabolism in *Trypanosoma brucei*." Biochem J **357**(Pt 1): 117-125.
- Hindmarsh, A. C., P. N. Brown, *et al.* (2005). "SUNDIALS: Suite of Nonlinear and Differential/Algebraic Equation Solvers." ACM Transactions on Mathematical Software **31**: 363-396.
- Hoefnagel, M. H. N., A. van der Burgt, *et al.* (2002). "Time dependent responses of glycolytic intermediates in a detailed glycolytic model of *Lactococcus lactis* during glucose run-out experiments." Mol Biol Rep **29**(1-2): 157-161.
- Holden, C. (2002). "Cell biology. Alliance launched to model *E. coli*." Science **297**(5586): 1459-1460.
- Hoops, S., S. Sahle, *et al.* (2006). "COPASI - a COmplex PATHway Simulator." Bioinformatics **22**: 3067-3074.
- Hucka, M., A. Finney, *et al.* (2003). "The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models." Bioinformatics **19**(4): 524-531.
- Jamshidi, N., J. S. Edwards, *et al.* (2001). "Dynamic simulation of the human red blood cell metabolic network." Bioinformatics **17**(3): 286-287.
- Jensen, K. (1992). Coloured Petri nets. Basic concepts, analysis methods and practical use. Berlin, Heidelberg, New York, Springer-Verlag.
- Joshi, A. und B. O. Palsson (1989). "Metabolic dynamics in the human red cell. Part I-A comprehensive kinetic model." J Theor Biol **141**(4): 515-528.
- Kacser, H. (1986). The Organization of Cell Metabolism. G. R. Welch und J. S. Clegg. New York, Plenum Press: 327-337.
- Kacser, H. und J. A. Burns (1995). "The control of flux." Biochem Soc Trans **23**(2): 341-366.
- Kahn, P. (1995). "From genome to proteome: looking at a cell's proteins." Science **270**(5235): 369-370.
- Kanehisa, M., S. Goto, *et al.* (2006). "From genomics to chemical genomics: new developments in KEGG." Nucleic Acids Res **34**(Database issue): D354-D357.
- Katagiri, F. (2003). "Attacking complex problems with the power of systems biology." Plant Physiol **132**(2): 417-419.
- Kauffman, K. J., P. Prakash, *et al.* (2003). "Advances in flux balance analysis." Curr Opin Biotechnol **14**(5): 491-496.
- Klamt, S. und J. Stelling (2002). "Combinatorial complexity of pathway analysis in metabolic networks." Mol Biol Rep **29**(1-2): 233-6.

- Krömer, J. O. (2006). Systembiotechnologische Studien an *Corynebacterium glutamicum* zur Charakterisierung der Methioninsynthese, Universität des Saarlandes.
- Le Novère, N., B. Bornstein, *et al.* (2006). "BioModels Database: A Free, Centralized Database of Curated, Published, Quantitative Kinetic Models of Biochemical and Cellular Systems." Nucleic Acids research **34**: D689-D691.
- Lee, J. M., E. P. Gianchandani, *et al.* (2006). "Flux balance analysis in the era of metabolomics." Brief Bioinform **7**(2): 140-150.
- Mahadevan, R. und C. H. Schilling (2003). "The effects of alternate optimal solutions in constraint-based genome-scale metabolic models." Metab Eng **5**(4): 264-276.
- Marx, A., A. A. de Graaf, *et al.* (1996). "Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing." Biotechnol Bioeng **49**(2): 111-129.
- Mavrovouniotis, M. L. (1990). "Group contributions for estimating standard Gibbs energies of formation of biochemical-compounds in aqueous-solution." Biotechnol Bioeng **36**: 1070–1082.
- Mavrovouniotis, M. L. (1991). "Estimation of standard Gibbs energy changes of biotransformations." J Biol Chem **266**(22): 14440-5.
- Mendes, P. und D. Kell (1998). "Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation." Bioinformatics **14**(10): 869-83.
- Michal, G. (1999). Biochemical Pathways. Biochemie-Atlas, Spektrum Akademischer Verlag.
- Moritz, B., K. Striegel, *et al.* (2002). "Changes of pentose phosphate pathway flux in vivo in *Corynebacterium glutamicum* during leucine-limited batch cultivation as determined from intracellular metabolite concentration measurements." Metab Eng **4**(4): 295-305.
- Moritz, B., K. Striegel, *et al.* (2000). "Kinetic properties of the glucose-6-phosphate and 6-phosphogluconate dehydrogenases from *Corynebacterium glutamicum* and their application for predicting pentose phosphate pathway flux in vivo." Eur J Biochem **267**(12): 3442-3452.
- Müllner, S., T. Neumann, *et al.* (1998). "Proteomics-a new way for drug target discovery." Arzneimittelforschung **48**(1): 93-95.
- Mulquiney, P. J., W. A. Bubba, *et al.* (1999). "Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: in vivo kinetic characterization of 2,3-bisphosphoglycerate synthase/phosphatase using ¹³C and ³¹P NMR." Biochem J **342 Pt 3**: 567-580.
- Nagasaki, M., A. Doi, *et al.* (2004). "A versatile petri net based architecture for modeling and simulation of complex biological processes." Genome Inform **15**(1): 180-197.
- Nakayama, Y., A. Kinoshita, *et al.* (2005). "Dynamic simulation of red blood cell metabolism and its application to the analysis of a pathological condition." Theor Biol Med Model **2**(1): 18.

- Nöh, K. und W. Wiechert (2006). "Experimental design principles for isotopically instationary ^{13}C labeling experiments." Biotechnol Bioeng **94**(2): 234-251.
- Oliver, S. G., M. K. Winson, *et al.* (1998). "Systematic functional analysis of the yeast genome." Trends Biotechnol **16**(9): 373-378.
- Olivier, B. G. und J. L. Snoep (2004). "Web-based kinetic modelling using JWS Online." Bioinformatics **20**(13): 2143-2144.
- Ould-Moulaye, C. B., C. G. Dussap, *et al.* (1999). "Estimation of Gibbs energy changes of central metabolism reactions." Biotechnology Techniques **13**: 187-193.
- Papin, J. A., N. D. Price, *et al.* (2003). "Metabolic pathways in the post-genome era." Trends Biochem Sci **28**(5): 250-258.
- Petri, C. A. (1962). Kommunikation mit Automaten, Institut für Instrumentelle Mathematik, Universität Bonn.
- Reddy, V. N., M. N. Liebman, *et al.* (1996). "Qualitative analysis of biochemical reaction systems." Comput Biol Med **26**(1): 9-24.
- Reddy, V. N., M. L. Mavrovouniotis, *et al.* (1993). "Petri net representations in metabolic pathways." Proc Int Conf Intell Syst Mol Biol **1**: 328-336.
- Reijenga, K. A., Y. M. G. A. van Megen, *et al.* (2005). "Yeast glycolytic oscillations that are not controlled by a single oscillator: a new definition of oscillator strength." J Theor Biol **232**(3): 385-398.
- Reisig, W. (1990). Petrinetze - Eine Einführung. Berlin, Springer.
- Rizzi, M., M. Baltes, *et al.* (1997). "In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model." Biotechnol Bioeng **55**: 592-608.
- Ronen, M., R. Rosenberg, *et al.* (2002). "Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics." Proc Natl Acad Sci U S A **99**(16): 10555-10560.
- Sackmann, A., M. Heiner, *et al.* (2006). "Application of Petri net based analysis techniques to signal transduction pathways." BMC Bioinformatics **7**(1): 482.
- Sauer, U. (2006). "Metabolic networks in motion: ^{13}C -based flux analysis." Mol Syst Biol **2**: 62.
- Savinell, J. M. und B. O. Palsson (1992a). "Network analysis of intermediary metabolism using linear optimization. I. Development of mathematical formalism." J Theor Biol **154**(4): 421-454.
- Savinell, J. M. und B. O. Palsson (1992b). "Network analysis of intermediary metabolism using linear optimization. II. Interpretation of hybridoma cell metabolism." J Theor Biol **154**(4): 455-473.
- Schomburg, I., A. Chang, *et al.* (2004). "BRENDA, the enzyme database: updates and major new developments." Nucleic Acids Res **32**(Database issue): D431-D433.

- Schuster, R., G. Jacobasch, *et al.* (1989). "Mathematical modelling of metabolic pathways affected by an enzyme deficiency. Energy and redox metabolism of glucose-6-phosphate-dehydrogenase-deficient erythrocytes." *Eur J Biochem* **182**(3): 605-612.
- Schuster, S., D. A. Fell, *et al.* (2000). "A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks." *Nat Biotechnol* **18**(3): 326-32.
- Schuster, S. und C. Hilgetag (1994). "On Elementary Flux Modes in Biochemical Reactions Systems at Steady State." *J Bio Syst* **2**(2): 165-82.
- Schuster, S. und T. Höfer (1991). "Determining All Extreme Semi-positive Conservation Relations in Chemical Reaction Systems: A Test Criterion for Conservativity." *J Chem Soc Faraday Trans* **87**: 2561-2566.
- Schuster, S. und D. Kenanov (2005). "Adenine and adenosine salvage pathways in erythrocytes and the role of S-adenosylhomocysteine hydrolase. A theoretical study using elementary flux modes." *FEBS J* **272**(20): 5278-5290.
- Segrè, D., D. Vitkup, *et al.* (2002). "Analysis of optimality in natural and perturbed metabolic networks." *Proc Natl Acad Sci U S A* **99**(23): 15112-7.
- Shampine, L. F. und C. W. Gear (1979). "A User's View of Solving Stiff Ordinary Differential Equations." *SIAM Review* **21**: 1-17.
- Singh, V. K. und I. Ghosh (2006). "Kinetic modeling of tricarboxylic acid cycle and glyoxylate bypass in *Mycobacterium tuberculosis*, and its application to assessment of drug targets." *Theor Biol Med Model* **3**: 27.
- Steinbeck, C., Y. Han, *et al.* (2003). "The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics." *J Chem Inf Comput Sci* **43**(2): 493-500.
- Steinbeck, C., C. Hoppe, *et al.* (2006). "Recent developments of the chemistry development kit (CDK) - an open-source java library for chemo- and bioinformatics." *Curr Pharm Des* **12**(17): 2111-2120.
- Stelling, J., S. Klamt, *et al.* (2002). "Metabolic network structure determines key aspects of functionality and regulation." *Nature* **420**(6912): 190-3.
- Strelkov, S. (2004). Entwicklung und Anwendung einer Methode zur Metabolomanalyse von *Corynebacterium glutamicum*, Universität zu Köln.
- Sundararaj, S., A. Guo, *et al.* (2004). "The CyberCell Database (CCDB): a comprehensive, self-updating, relational database to coordinate and facilitate in silico modeling of *Escherichia coli*." *Nucleic Acids Res* **32 Database issue**: D293-5.
- Szegezdi, J. und F. Csizmadia (2003). New method for pKa estimation. eCheminformatics Proceedings.
- Teusink, B., J. Passarge, *et al.* (2000). "Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry." *Eur J Biochem* **267**(17): 5313-5329.
- Thielen, B. (2006). Persönliche Kommunikation.

- Turina, P., D. Samoray, *et al.* (2003). "H⁺/ATP ratio of proton transport-coupled ATP synthesis and hydrolysis catalysed by CF₀F₁-liposomes." EMBO J **22**(3): 418-426.
- Urbanczik, R. und C. Wagner (2005). "Functional stoichiometric analysis of metabolic networks." Bioinformatics **21**(22): 4176-4180.
- Vallabhajosyula, R. R., V. Chickarmane, *et al.* (2006). "Conservation analysis of large biochemical networks." Bioinformatics **22**(3): 346-353.
- Velculescu, V. E., L. Zhang, *et al.* (1997). "Characterization of the yeast transcriptome." Cell **88**(2): 243-251.
- Ventura, B. D., C. Lemerle, *et al.* (2006). "From in vivo to in silico biology and back." Nature **443**(7111): 527-533.
- Visser, D. und J. J. Heijnen (2003). "Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics." Metab Eng **5**(3): 164-76.
- von Kamp, A. und S. Schuster (2006). "Metatool 5.0: fast and flexible elementary modes analysis." Bioinformatics **22**(15): 1930-1931.
- Wadsö, I., H. Gutfreund, *et al.* (1976). "Recommendations for measurement and presentation of biochemical equilibrium data. Prepared by the Interunion Commission on Biothermodynamics." J Biol Chem **251**(22): 6879-6885.
- Washburn, M. P., D. Wolters, *et al.* (2001). "Large-scale analysis of the yeast proteome by multidimensional protein identification technology." Nat Biotechnol **19**(3): 242-247.
- Weast, R. C., Ed. (1989). CRC Handbook of Chemistry & Physics, CRC Pr I Llc, Boca Raton, Florida, USA.
- Wedler, G. (1997). Lehrbuch der Physikalischen Chemie, Wiley-VCH.
- Wendisch, V. F., A. A. de Graaf, *et al.* (2000). "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.
- Wiechert, W. (2001). "13C metabolic flux analysis." Metab Eng **3**(3): 195-206.
- Wiley, H. S., S. Y. Shvartsman, *et al.* (2003). "Computational modeling of the EGF-receptor system: a paradigm for systems biology." Trends Cell Biol **13**(1): 43-50.
- Zevedei-Oancea, I. und S. Schuster (2003). "Topological analysis of metabolic networks based on Petri net theory." In Silico Biol **3**(3): 323-345.
- Zittrich, S. und R. Krämer (1994). "Quantitative discrimination of carrier-mediated excretion of isoleucine from uptake and diffusion in *Corynebacterium glutamicum*." J Bacteriol **176**(22): 6892-6899.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Dietmar Schomburg betreut worden.

Teilpublikationen: keine

Köln, den 07.12.2006

