

Strukturelle Analyse von Stoffwechselnetzen illustriert am bakteriellen Redox- und Zentralstoffwechsel

Von der Fakultät Maschinenbau der Universität Stuttgart zur Erlangung der
Würde eines Doktors der Ingenieurwissenschaften (Dr. Ing.)
genehmigte Abhandlung

Vorgelegt von

Diplom-Systemwissenschaftler
Steffen Klamt

aus Magdeburg

Hauptberichter: Prof. Dr. Dr. h.c. mult. E. D. Gilles
Mitberichter: Prof. Dr. Stefan Schuster

Tag der mündlichen Prüfung: 22.03.2005

Institut für Systemdynamik und Regelungstechnik
der Universität Stuttgart
2005

Forschungsberichte aus dem Max-Planck-Institut
für Dynamik komplexer technischer Systeme

Band 9

Steffen Klamt

**Strukturelle Analyse von Stoffwechselnetzen
illustriert am bakteriellen
Redox- und Zentralstoffwechsel**

D 93 (Diss. Universität Stuttgart)

Shaker Verlag
Aachen 2005

Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Zugl.: Stuttgart, Univ., Diss., 2005

Copyright Shaker Verlag 2005

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 3-8322-3976-6

ISSN 1439-4804

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen

Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9

Internet: www.shaker.de • eMail: info@shaker.de

Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Max-Planck-Institut für Dynamik komplexer technischer Systeme in Magdeburg.

Die Rahmenbedingungen am Max-Planck-Institut waren ideal. Meinen besonderen Dank richte ich an Prof. Ernst Dieter Gilles für die wissenschaftliche Betreuung und für die angenehme Arbeitsatmosphäre, die er durch den aufgeschlossenen Umgang mit seinen Mitarbeitern schafft. Die wissenschaftliche Freiheit, die mir unter seiner Regie vergönnt war, hatte einen entscheidenden Anteil daran, daß die Arbeit in der vorliegenden Form entstehen konnte.

Ebenfalls in ganz besonderem Maße danke ich dem MItberichter Prof. Stefan Schuster (Universität Jena). Er lenkte damals meine Aufmerksamkeit auf das spannende Gebiet der Strukturanalyse zellulärer Netzwerke und schuf letztlich mit seinen Vorarbeiten überhaupt die Grundlage für wesentliche Inhalte dieser Arbeit. Viele gemeinsame Diskussionen waren für diese Promotion unentbehrlich.

Danken möchte ich auch den Kollegen vom MPI Magdeburg für das jederzeit freundschaftliche und stimulierende Arbeitsklima, was letztlich auch in viele gemeinsame Stunden nach getaner Arbeit mündete. Hervorheben möchte und kann ich hier nur jene, mit denen ich wissenschaftlich direkt zusammengearbeitet habe.

Prof. Dietrich Flockerzi stand mir oft in heiklen Fragen der linearen Algebra zur Seite, wodurch so manches in mathematisch präzisere Form kam. Viele theoretische Betrachtungen über Elementarmoden entstanden in enger Zusammenarbeit mit Dr. Jörg Stelling. Martin Ginkel half mir beim Optimieren von Teilen des Programmcodes im *FluxAnalyzer*. Dr. Hartmut Grammel weckte mein Interesse für die *Rhodospirillen* und war mein ständiger Ansprech- und Diskussionspartner bei Fragen rund um diese faszinierenden Organismen. Viele Daten und Ergebnisse aus Experimenten, die er zusammen mit Dr. Ruxandra Rehner und Andrea Focke durchführte, flossen in die vorliegende Arbeit ein.

Dank der exzellenten Kaffeemaschine von Renate Müller konnte ich jeden Arbeitstag mit einer Tasse frisch gebrühten Kaffees beginnen. Jeder Wissenschaftler weiß, was das bedeutet.

Last but not least danke ich meiner Familie. Die lange Unterstützung und Förderung meiner Eltern brachte mich überhaupt in die Situation, das Unterfangen einer Promotion angehen zu

können. Meinen speziellen Dank richte ich an meine beiden Frauen: An Grit, für Liebe, Rückhalt und Unterstützung – nicht nur während der Promotion. Und an Svenja, weil sie mich mit leuchtenden Kinderaugen und natürlicher Selbstverständlichkeit immer daran erinnert, daß auch die einfachen Dinge des Lebens größte Freude bereiten.

Magdeburg, April 2005

Steffen Klamt

Für Svenja und Grit

Inhaltsverzeichnis

1	EINFÜHRUNG UND ÜBERSICHT	1
2	GRUNDLAGEN STRUKTURELLER ANALYSEN IN STOFFWECHSELNETZEN	5
2.1	Stoffwechselnetze und ihre mathematische Beschreibung	5
2.2	Strukturbasierte Aussagen zur Existenz, Stabilität und Eindeutigkeit von stationären Zuständen	10
2.3	Graphentheoretische Strukturanalysen	11
2.4	Stationäre Flußverteilungen	15
2.4.1	Nullraum der stöchiometrischen Matrix	15
2.4.2	Optimale Flußverteilungen – Flux Balance Analysis	17
2.4.3	Elementarmoden und Metabole Pathwayanalyse	18
2.4.4	Theoretische Eigenschaften der Elementarmoden, konvexer Flußkegel und Vergleich mit extremen Pathways	22
2.4.5	Metabole Flußanalyse	28
2.5	Erhaltungsgleichungen	29
2.6	Zusammenfassende Übersicht über metabolische Strukturanalysen	31
3	ZWEI AUSGEWÄHLTE MODELLSYSTEME DES BAKTERIELLEN ZENTRAL- UND REDOXSTOFFWECHSELS	33
3.1	Organisation des Stoffwechsels in Bakterien	33
3.2	Netzmodell für Stoffwechsel in <i>Escherichia coli</i>	35
3.3	Netzmodell für Stoffwechsel in schwefelfreien Purpurbakterien (<i>Rhodospirillum rubrum</i>)	36
4	FLUXANALYZER: COMPUTERGESTÜTZTE ANALYSE VON STOFFWECHSELNETZEN	41
4.1	Motivation	41
4.2	Konzeption des FluxAnalyzers: Netzwerk-Projekte und interaktive Flußkarten	42
4.3	Konstruktion eines Netzwerk-Projektes	43
4.4	Toolbox: Struktur-, Pathway- und Flußanalyse in Stoffwechselnetzen	46
5	BERECHENBARKEIT VON STATIONÄREN FLÜSSEN IN STOFFWECHSELNETZEN	55
5.1	Berechenbarkeit und Redundanz in metabolischen Flußanalysen	55
5.1.1	Klassifikation von Szenarien: Bestimmtheit und Redundanz	56
5.1.2	Klassifikation von Raten: Spezifische Berechenbarkeit und Redundanz	57
5.1.3	Umsetzung im FluxAnalyzer	59
5.1.4	Beispiele	60
5.2	Verallgemeinerte Berechenbarkeitsanalyse in stationären Reaktionsnetzwerken	61
5.3	Parallelen mit der Beobachtbarkeitsanalyse in der Regelungstechnik	65
5.4	Realisierbarkeit eines Flußszenarios	66
5.5	Anwendungsbeispiel: Redoxbalance im Stoffwechsel der <i>Rhodospirillum rubrum</i>	67
5.5.1	Reduktionsgrad des Substrats, Biomasseausbeute und Netto-CO ₂ -Freisetzung	70
5.5.2	Bedeutung des Calvin-Zyklus für die Redoxbalance	72
5.5.3	Zusammenhang zwischen Calvin-Zyklus und reduktiven/oxidativen TCA	74
5.5.4	Calvin-Zyklus, Transhydrogenase und oxidativer Pentosephosphat-Weg	75
5.5.5	Verallgemeinerte Berechenbarkeitsanalyse für phototrophes Wachstum	76
5.5.6	Chemotrophes Wachstum unter aeroben Bedingungen	77
6	KOMBINATORISCHE ELEMENTARMODENANALYSE	79
6.1	Spezifische Selektion einer Teilmenge von Elementarmoden	79

6.2	Flexibilität des Netzes und Relevanz einer Reaktion für vorgegebene Umweltbedingungen	80
6.2.1	Relative Auftretshäufigkeiten von Reaktionen und Flexibilität eines Szenarios	80
6.2.2	Anwendung auf <i>E. coli</i> : Wachstum auf verschiedenen Substraten	83
6.2.3	Anwendung für Rhodospirillaceae: Phototropher und respirativer Stoffwechsel auf verschiedenen Substraten	87
6.3	Gewichtete Relevanz	90
6.4	Strukturelle Kopplungen zwischen Reaktionen	95
6.5	Kombinatorische Elementarmodenanalyse im FluxAnalyzer	99
7	MINIMALE SCHNITTMENGEN: ELEMENTARE FEHLERMODEN UND EFFIZIENTE EINGRIFFE IN METABOLEN REAKTIONSNETZWERKEN	101
7.1	Definition von minimalen Schritten	101
7.2	Algorithmus zur Berechnung minimaler Schnittmengen	104
7.3	Anwendungen	108
7.3.1	Identifikation von Targets zur Unterdrückung zellulärer Funktionen	108
7.3.2	Fälschungen der Netzstruktur und Mutanten-Phänotyp-Prädiktionen	108
7.3.3	Strukturelle Fragilität und Robustheit	110
7.3.4	Minimale Meßkombinationen	112
7.3.5	Elimination von Erhaltungselationen	114
7.4	Minimale Schnittmengen in der Risikoanalyse und Graphentheorie	114
8	KOMBINATORISCHE KOMPLEXITÄT DER ELEMENTARMODENBESTIMMUNG	117
8.1	Maximal mögliche Anzahl von Elementarmoden	117
8.2	Realistische Netze: $S \ll S_{max}$	119
8.3	Rechenaufwand	122
9	VON DER STRUKTUR ZUR DYNAMIK: MODELLIERUNG DER ELEKTRONENTRANSPORTKETTE UND DER REDOXREGULATION IN PHOTOSYNTHETISCHEN BAKTERIEN	125
9.1	Physiologie und Modellierung der Elektronentransportkette	125
9.1.1	Die Elektronentransportkette der Rhodospirillaceae	125
9.1.2	Operationsweisen der Elektronentransportkette: Elementarmoden	130
9.1.3	Dynamisches Modell der Elektronentransportkette	131
9.2	Physiologie und Modellierung der Redoxregulation	135
9.3	Modellstudien	139
9.3.1	Simulationen für verschiedene Sauerstoffkonzentrationen und Lichtintensitäten	139
9.3.2	Der Redoxzustand des Ubichinons als integrierendes Signal	143
9.3.3	Zusammenfassung zum dynamischen Modell der ETK	146
10	ZUSAMMENFASSUNG / SUMMARY	147
	ANHANG	157
	Anhang A: Stoffwechselnetze <i>Escherichia coli</i> und Rhodospirillaceae	158
	Anhang B: <i>E. coli</i> Deletionsstudien	173
	Anhang C: Elementarmodas mit phototrophem Wachstum der Rhodospirillaceae auf Acetat ohne Einbeziehung des Citramalat-Zyklus	175
	Anhang D: Beispiel zur Berechnung einer gewichteten Relevanz	176
	Anhang E: Theoretische und gemessene Transkriptverhältnisse von Genen des Zentralstoffwechsels in <i>E. coli</i> für die Substratpaare Acetat/Glucose und Glycerin/Glucose	178
	Anhang F: Dynamisches Modell der Elektronentransportkette der Rhodospirillaceae	180
	LITERATURVERZEICHNIS	187

1 Einführung und Übersicht

Leben ist gekennzeichnet durch eine enorme Anzahl von direkten und indirekten Interaktionen und Wechselwirkungen zellulärer Komponenten. Ein ganzheitliches Verständnis von Zellprozessen im Sinne der Systembiologie erfordert deshalb eine Analyse *zellulärer Netzwerke*. Bis auf rein physikalische Vorgänge wie Diffusion oder Konvektion entsprechen die Interaktionen in den Netzen letztlich (bio)chemischen Reaktionen. Dennoch lassen sich zelluläre Netze auf einer höheren Abstraktionsstufe grob in drei Klassen einteilen

- *Stoffwechsel (Metabolismus)*: Stoffaustausch mit der Umgebung, Abbau von Substanzen (Katabolismus) und Aufbau zelleigener Bestandteile (Anabolismus)
- *Signaltransduktion*: Generierung, Weiterleitung und Integration zellulärer Signale
- *Regulatorische (genetische) Netzwerke*: Abruf oder Unterdrückung genetisch verankerter Potentiale

Die Prozesse in diesen drei Ebenen operieren natürlich nicht losgelöst voneinander. Im Gegenteil, Schnittstellen sorgen für einen Austausch von Information und Material (Abb. 1.1). So können bestimmte Zustände oder Flüsse im Stoffwechsel als Signale erfasst und durch Signaltransduktion weitergeleitet und integriert werden. Auf der Stufe der Regulation dienen die Signale zur Anpassung der Expression einzelner oder mehrerer Gene. Damit schließt sich die Rückkopplung zum Stoffwechsel, da über die Genexpression u.a. Enzymkonzentrationen und damit Stoffflüsse variiert werden können.

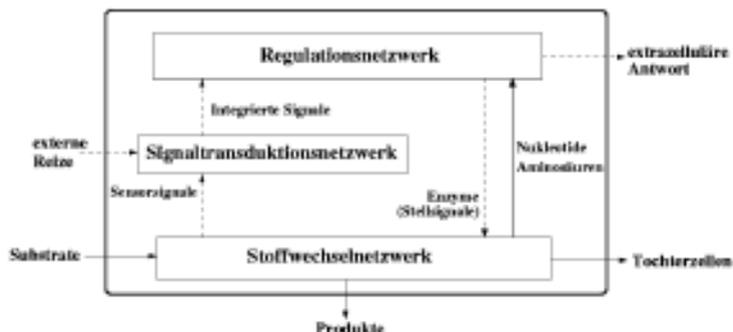


Abb. 1.1: Interaktion von Stoffwechsel-, Signaltransduktions- und Regulationsnetzwerken in der Zelle. Stoffflüsse sind mit vollen und Signale mit Strichellinien gekennzeichnet (nach [47]).

Die meisten mathematischen Modelle zellulärer Systeme basieren auf *Vorwärtsmodellierung*, wo ein gegebenes Modell gemessene Daten und Beobachtungen rekonstruieren soll, während man im Gegensatz dazu bei der *inversen* Modellierung Daten zur Rekonstruktion des Modells

nutzt. Bei der Vorwärtsmodellierung muß die Netzstruktur, also die Karte der Interaktionen, bekannt sein, bevor die Interaktionen in ihrer Dynamik beschrieben werden können.

Da in größeren Netzen nicht nur die Dynamik, sondern auch die topologischen Eigenschaften immer weniger intuitiv erfaßbar sind, haben Methoden zur strukturellen Analyse zellulärer Netze in den letzten Jahren stark an Bedeutung gewonnen. Diese helfen zum einen, allgemeine Leitstrukturen für das Design zellulärer Netze zu identifizieren [5,10]. Zum anderen zeigt sich insbesondere in Stoffwechselnetzen, daß wichtige Aussagen zur Funktionalität [29,148] und Kapazität [164], sowie sogar zur Regulation [148] und zum dynamischen Verhalten [17] allein aus der Netzarchitektur abgeleitet werden können. Nicht zuletzt aufgrund von genom-basierten Rekonstruktionen sind Stoffwechselnetze von ihrer Struktur her derzeit am besten charakterisiert und deshalb ideale Objekte für topologische Analysen.

Ein Großteil der vorliegenden Arbeit beschäftigt sich daher mit der Ausarbeitung neuer Methoden zur Strukturanalyse in Stoffwechselnetzen. Die Ansätze beruhen weitgehend auf einer vertieften Analyse von Elementarmoden [141], mit denen ganz zentrale Eigenschaften eines metabolen Netzes extrahiert werden können. Die eingeführten Methoden werden anhand zweier realer Beispiele des bakteriellen Stoffwechsels illustriert. Das erste Netzmodell repräsentiert den Zentralstoffwechsel von *Escherichia coli*. Da *E. coli* eines der wichtigsten Studienobjekte der Mikrobiologie ist, stehen uns hier auch viele experimentelle Daten und Ergebnisse zur Verfügung. Etwas weniger gut untersucht ist dagegen der Stoffwechsel der schwefelfreien Purpurbakterien (*Rhodospirillaceae*). Eine strukturelle Analyse des Metabolismus dieser fakultativ phototrophen Bakterien ist dennoch sehr reizvoll, nicht zuletzt deshalb, weil diese Bakteriengruppe bzgl. des Stoffwechsels zu den vielseitigsten Organismen gezählt wird. Eine besondere Herausforderung für diese Organismen ist es dabei, den Redoxstoffwechsel den verschiedenen Bedingungen anzupassen. So müssen unter aeroben Bedingungen, wo Respiration mit Sauerstoff überwiegt, reichlich Reduktionsäquivalente zur Verfügung gestellt werden, während letztere bei phototrophen Wachstum (unter anaeroben Bedingungen im Licht) oft im Überschuß vorliegen und deshalb Redoxsenken für eine ausgewogene Balance installiert werden müssen. Stöchiometrische Analysen sollen uns hier ein besseres Verständnis für die zu erbringenden Leistungen und Umstellungen im Stoffwechsel liefern.

Neben der Anpassung des globalen Redoxstoffwechsels müssen die *Rhodospirillaceae* beim Wechsel vom aeroben zum phototrophen Wachstum auch die membranständige Elektronentransportkette (ETK) in ihrer Funktion und Zusammensetzung umstellen. So werden für phototrophes Wachstum photosynthetische Apparate installiert, die unter aeroben Bedingungen nicht ausgebildet sind. Ein Ziel dieser Arbeit war es auch, diesen Umschaltprozeß zu modellieren. Eine funktionale Analyse der ETK über Elementarmoden ist hier ebenfalls hilfreich, für ein tieferes Verständnis dieser Prozesse wird jedoch ein dynamisches Modell benötigt. Dieses Modell soll veranschaulichen, wie Redoxsignale in der ETK generiert und für eine koordinierte Regulation von Photosynthesegenen genutzt werden. Da in diesem Modell sowohl Stoffwechsel als auch Signaltransduktion und Regulation involviert sind, können wir das Zusammenspiel der drei Ebenen zellulärer Netze studieren.

Gliederung

In Kap. 2 beginnen wir mit einem (durchaus kritischen) Überblick über die wichtigsten Methoden zur Strukturanalyse von Stoffwechselnetzen. Dort erarbeitete Grundlagen bilden das theoretische Fundament nachfolgender Kapitel.

In Kap. 3 rekapitulieren wir zunächst die grundlegende Organisation des bakteriellen Stoffwechsels. Dann werden die Besonderheiten und die Modelle der Stoffwechselnetze von *E. coli* und von den *Rhodospirillaceae* erläutert. Wie oben schon erwähnt dienen diese Netzmodelle in späteren Kapiteln zur Illustration.

In Kap. 4 wird mit dem *FluxAnalyzer* eine integrierte graphische Benutzeroberfläche vorgestellt, die für die rechnergestützte Strukturanalyse biochemischer Reaktionsnetze entwickelt wurde. In späteren Kapiteln wird auch gezeigt, wie die jeweils neu entwickelten Methoden im *FluxAnalyzer* integriert worden sind.

In Kap. 5 wenden wir uns der *Berechenbarkeitsanalyse* in metabolen Flußanalysen zu. Wir werden sehen, daß unterbestimmte Stoffwechselnetze, in denen nicht alle stationären Stoffflüsse eindeutig berechnet werden können, u.U. dennoch wertvolle Informationen bereithalten, wie am Redoxstoffwechsel der *Rhodospirillaceae* gezeigt wird. Mit Hilfe der Elementarmoden werden wir außerdem zeigen, wie Meßkombinationen (sofern überhaupt existent) gefunden werden können, mit denen unbekannte Reaktionsraten in einem Stoffwechselnetz berechnet werden können.

Kap. 6 widmet sich der *kombinatorischen Elementarmodenanalyse*, deren Ziel es ist, die Relevanzen einzelner Reaktionen für unterschiedliche Umweltbedingungen zu bestimmen und funktionale Strukturkopplungen zwischen Netzelementen zu quantifizieren. Derartige Untersuchungen werden auch Rückschlüsse auf regulatorische Aspekte erlauben.

In Kap. 7 verändern wir unseren Blickwinkel dahingehend, daß eine spezifische Funktion in einem Reaktionsnetz durch gezielte Manipulation unterdrückt werden soll. Eine geeigneter Ansatz wird mit den *Minimalen Schnittmengen* eingeführt. Mehrere Anwendungen, z.B. Arzneimittelforschung und Robustheitsanalyse in metabolen Netzen werden aufgezeigt.

Ein generelles Problem der Elementarmodenanalyse ist die kombinatorische Komplexität in größeren Netzen, die wir in Kap. 8 näher untersuchen wollen.

Kap. 9 wird schließlich, wie oben bereits erwähnt, ein dynamisches Modell der ETK (einschließlich ihrer Regulation) in den *Rhodospirillaceae* vorgestellt und untersucht.

Einige der in dieser Arbeit vorgestellten Ergebnisse sind – mehr oder weniger stark komprimiert – bereits publiziert worden. Der *FluxAnalyzer* wurde in [77] vorgestellt. Methoden zur Berechenbarkeitsanalyse in metabolen Netzen wurden in [76] publiziert und dort, wie in Kap. 5, am Redoxstoffwechsel der *Rhodospirillaceae* erläutert. Aspekte der kombinatorischen Elementarmodenanalyse (Kap. 6) wurden in [148] eingeführt und am Modell des Zentralstoffwechsel von *E. coli* angewendet. Konzept und Anwendungen der Minimalen Schnittmengen (Kap. 7) wurden in [73] vorgestellt. Aspekte der kombinatorischen Komplexität von Elementarmoden (Kap. 8) sind in [74] besprochen worden. Ein kritischer Vergleich von Elementarmoden und extremen Pathways (Abschnitt 2.4.4) findet sich in [75].

2 Grundlagen struktureller Analysen in Stoffwechselnetzen

2.1 Stoffwechselnetze und ihre mathematische Beschreibung

Die elementaren Bausteine von Stoffwechselnetzen sind – wie in jedem Reaktionsnetz – die *Reaktionen* und die in diesen Reaktionen als Edukte und Produkte auftretenden *Spezies*. Letztere nennt man im Stoffwechsel auch *Metabolite*.

Stoffwechselreaktionen sind i.a. durch folgende Merkmale charakterisiert:

(1) **Stöchiometrie:** Kennzeichnet die beteiligten Edukte (Substrate) und Produkte mit ihrem jeweiligen stöchiometrischen Koeffizienten.

(2) **Reversibilität:** Aufgrund thermodynamischer Grundsätze muß man im allgemeinen Fall alle chemischen Reaktionen als reversibel betrachten. Dennoch können eine ganze Reihe biochemischer Reaktionen als *praktisch irreversibel* eingestuft werden, da sie unter zellulären Bedingungen fast ausschließlich in nur einer Richtung ablaufen. Ein typisches Beispiel ist die CO_2 -Fixierung im Calvin-Zyklus durch die *Ribulosebiphosphat-Carboxylase*, das mengenmäßig häufigste Enzym in der Natur. Wie wir sehen werden, können irreversible Reaktionen die verfügbaren Routen im Stoffwechsel drastisch reduzieren.

(3) **Enzym:** Die meisten biochemischen Reaktionen werden durch ein spezifisches Enzym oder Enzymkomplex katalysiert. Die Aminosäuresequenz dieser Proteine sind in einem (oder mehreren) Gen(en) auf der DNA kodiert. Bei der genom-basierten Rekonstruktion von Stoffwechselnetzen werden vorhandene Reaktionswege durch das katalysierende Enzym identifiziert, dessen Sequenz man im annotierten Genom des jeweiligen Organismus gefunden hat [87]. Ein Enzym wird während einer Reaktion nicht verbraucht, kann aber verschiedene Konformationszustände haben. Außerdem ist die Zuordnung Enzym – Reaktion nicht immer eindeutig. So sind einige Enzyme multifunktional (d.h. sie katalysieren mehrere Reaktionen, z.B. *Transketolase*) und einige Reaktionen können über mehrere, meist sehr ähnliche Enzyme (Isoenzyme) katalysiert werden.

Die folgenden Eigenschaften sind im wesentlichen Eigenschaften des Enzyms, können aber auch als Eigenschaften der biochemischen Reaktion verstanden werden, da sie den Stofffluß maßgeblich bestimmen.

(4) **Kinetischer Mechanismus und Regulation:** Jede Reaktion unterliegt einem kinetischen Mechanismus, der die Interaktion der beteiligten Spezies und Enzyme, sowie die Geschwindigkeit des Stoffumsatzes beschreibt. Sofern die Metabolite in ausreichend großer Konzentration vorliegen, wird oft das Massenwirkungsgesetz zugrunde gelegt [56]. Da in den meisten Fällen ein Enzym als Katalysator beteiligt ist, setzt sich die Gesamtreaktion aus einer Sequenz einzelner Elementarschritte zusammen, die ihrerseits ebenfalls über das Massenwirkungsge-

setz beschrieben werden können. Versucht man die Reaktionsgeschwindigkeit eines jeden Elementarschrittes mathematisch zu beschreiben, wächst die Anzahl der zu berücksichtigenden Zustände und Parameter sehr schnell. Mit bestimmten Annahmen, wie z.B. Quasistationarität, lassen sich die auftretenden kinetischen Ausdrücke vereinfachen [56,149]. So wird bei einfachen *uni-uni*-Reaktionen (ein Substrat, ein Produkt) oftmals die *Michaelis-Menten-Kinetik* verwendet, die einen hyperbolischen Verlauf der Reaktionsgeschwindigkeit in Abhängigkeit von der Substratkonzentration beschreibt. Allosterische Enzyme mit Kooperativität zeigen dagegen einen sigmoidalen Verlauf [134]. Das kinetische Verhalten von Enzymreaktionen kann durch regulatorische Prozesse gesteuert werden. Auf *metaboler Ebene* können Effektoren (Aktivatoren oder Inhibitoren) die Enzymaktivität verändern, was eine schnelle Adaptation von Stoffflüssen ermöglicht, z.B. wenn ein Produkt hemmend auf die erste Reaktion seines Biosyntheseweges einwirkt. Durch Regulation der *Enzymsynthese* und des *Enzymabbaus* kann die vorhandene Enzymkonzentration in der Zelle justiert und für die langfristige Regulation von Stoffflüssen genutzt werden.

Das mittlerweile berühmte Böhlinger-Poster („Biochemical Pathways“ [92]) stellt die wichtigsten bekannten Stoffwechselfade in graphischer Weise dar (Abb. 2.1). Dieser Stoffwechselkarte können z.B. Stöchiometrie, Reversibilität und Enzyme der Reaktionen entnommen werden. Für einige Reaktionen werden Regulationen auf metaboler Ebene angedeutet.

Das Böhlinger-Poster ist relativ unspezifisch und unterscheidet nur nach Organismenreichen. Wie oben bereits angedeutet, gewinnt die genom-basierte Rekonstruktion von Stoffwechselnetzen – in Kombination mit anderen, experimentellen Nachweismethoden – immer stärker an Bedeutung, da die Anzahl sequenzierter Genome stetig steigt (z.Zt. etwa 130). Rekonstruierte Netze, wenn auch z.T. noch nicht völlig lücken- und fehlerlos, sind in Datenbanken wie KEGG [50,68], WTT [110] oder MetaCyc [70] abgelegt.

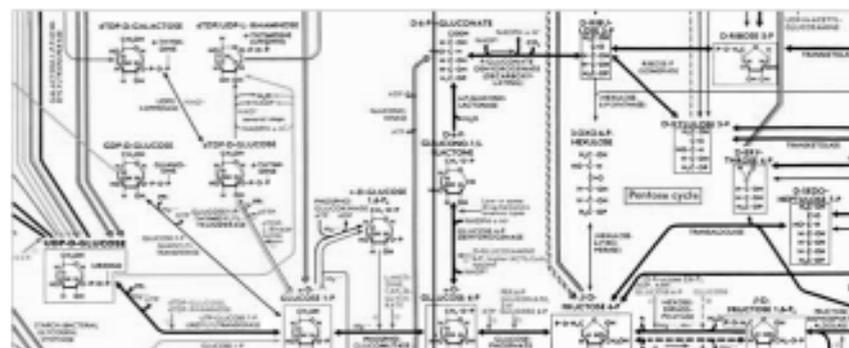


Abb. 2.1: Ausschnitt aus dem Böhlinger-Poster „Biochemical Pathways“

Stehen die oben genannten Eigenschaften (1)-(4) für jede Reaktion eines betrachteten Stoffwechselnetzes zur Verfügung und nehmen wir vereinfachend an, daß die Zelle homogen

durchmischt ist, können wir die zeitliche Konzentrationsänderung einer Spezies i beschreiben durch eine Differentialgleichung [56]:

$$\frac{dx_i}{dt} = \sum_{j=1}^q n_{ij} r_j(t) \quad (2.1)$$

mit folgenden Größen:

- q : Anzahl der Reaktionen
- m : Anzahl der Spezies (Metabolite)
- x_i : Konzentration von Spezies i ($i \in \{1, \dots, m\}$) mit Einheit [mmol/gTM]; wobei gTM für „Gramm Trockenmasse“ steht (= gDW = „gram dry weight“)
- n_{ij} : stöchiometrischer Koeffizient von Spezies i in Reaktion j
- $r_j(t)$: (Netto-)Reaktionsrate der Reaktion j mit Einheit [mmol/(gTM · h)] (molarer Umsatz)

Faßt man die stöchiometrischen Koeffizienten in der stöchiometrischen Matrix \mathbf{N} , die Reaktionsraten im Vektor $\mathbf{r}(t)$, sowie die Konzentrationen im Vektor $\mathbf{x}(t)$ zusammen, so kann man (2.1) auch in Matrix-Notation schreiben:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{N}\mathbf{r}(t) \quad (2.2)$$

Die stöchiometrische Matrix \mathbf{N} hat demnach die Dimension $m \times q$, d.h. in den Spalten stehen die Reaktionen und in den Zeilen die einzelnen Spezies mit ihren stöchiometrischen Koeffizienten (positiv bei Synthese und negativ bei Verbrauch in der jeweiligen Reaktion).

Gleichung (2.2) ist zentral in mathematischen Modellen zellulärer Stoffwechselletze. Sie erscheint zunächst recht simpel. Die Schwierigkeit liegt jedoch in der quantitativen Beschreibung der aktuellen Reaktionsraten $r_j(t)$, die im allgemeinen Fall einer (dem kinetischen Mechanismus entsprechenden) Funktion

$$r_j(t) = f_j(\mathbf{x}(t), \mathbf{p}, t), \quad (2.3)$$

gehörchen, wobei \mathbf{p} die Menge der benötigten kinetischen Parameter darstellt. Die Zeitkomponente t deutet an, daß das System nicht autonom sein muß. Zu beachten ist, daß wir zur Berechnung der Reaktionsraten in (2.3) auch die aktuellen Enzym- und Effektorkonzentrationen benötigen. Genaugenommen müßten wir im Vektor $\mathbf{x}(t)$ also auch die relevanten Enzyme und Effektoren bilanzieren, d.h. ihre Synthese und ihren Abbau berücksichtigen und somit Signaltransduktion und Genexpression in das Modell einbeziehen. Es ist klar, daß die Ge-

samtheit der benötigten Informationen in einem modellierten Bereich des Stoffwechsels fast immer unvollständig vorliegt. Will man dennoch dynamisch modellieren sind folglich vereinfachende Annahmen notwendig, von denen viele aber auch als unkritisch angesehen werden können. So kann man Enzymkonzentrationen beispielsweise als konstant betrachten, wenn man sich nur für einen sehr kurzen Zeitraum nach einer Anregung interessiert. Eine wichtige Erkenntnis der noch jungen Systembiologie ist, daß biologische Systeme in bestimmten Grenzen robust auf innere und äußere Störungen reagieren [72,166]. In plausiblen biologischen Modellen werden deshalb kleinere Abweichungen in den Parameterwerten das dynamische Verhalten des Systems zumindest qualitativ nicht wesentlich beeinflussen. Deshalb können wertvolle Aussagen über das dynamische Verhalten biochemischer Systeme auch mit vereinfachenden Modellannahmen und nicht exakten Parametern gewonnen werden, sofern diese nicht zu stark von den realen Gegebenheiten abweichen.

Dennoch, trotz der gewaltigen Fortschritte in Technologien, die zelluläre Größen erfassen, ist die Wissens- und Datenbasis für ein dynamisches Modell biochemischer Interaktionen noch immer oft gering [112]. Für viele Reaktionen sind die genauen Mechanismen noch nicht aufgeklärt. Außerdem ist über die Regulation von Aktivität und Konzentration der beteiligten Enzyms oft wenig bekannt. Kinetische Parameter stehen relativ selten zur Verfügung und wenn doch, so können *in vitro* gemessene Werte stark von den *in vivo* Werten abweichen.

In Gleichung (2.2) liegt das detaillierte Wissen über Kinetik und Regulation und damit auch der „unsichere Teil“ gerade in den Reaktionsraten $r(t)$ verborgen. Dagegen kann man die Stöchiometrie – abgelegt in der stöchiometrischen Matrix N – in den meisten Fällen als gut bekannt voraussetzen. Zumindest die im Zentralstoffwechsel stattfindenden biochemischen Reaktionen gelten als weitgehend aufgeklärt und mittels oben erwähnter Datenbanken sind organismenspezifische Stoffwechselnetze für eingehende Studien leicht zugänglich [65,87].

Die stöchiometrische Matrix konserviert die Struktur – oder Topologie – eines Stoffwechselnetzes, d.h. die Gesamtheit der beteiligten Metabolite und ihrer Verbindungen untereinander. Die Stöchiometrie des Systems ist damit nicht nur unabhängig von Konzentrationen und kinetischen Parametern, sondern auch von der Zeit und stellt somit eine *Invariante* des betrachteten Systems dar. Die Untersuchung dieser Invarianten ist Gegenstand *struktureller (stöchiometrischer, topologischer) Analysen* in Stoffwechselnetzen. Um die Struktur vollständig zu beschreiben bezieht man meistens auch die Reversibilitäten, d.h. die Richtung(en), der Reaktionen mit ein, die in den meisten Fällen ebenfalls gut bekannt sind. Formal beschrieben stehen also folgende Informationen bei Strukturanalysen von Reaktionsnetzwerken zur Verfügung:

$$\begin{aligned}
 N &: \text{stöchiometrische Matrix mit } m \text{ Zeilen (Metabolite) und } q \text{ Spalten (Reaktionen);} \\
 \mathit{rev} &= \{i_1, i_2, \dots, i_r\}: \text{Menge (Indizes/Namen) der reversiblen Reaktionen;} \\
 \mathit{irrev} &= \{j_1, j_2, \dots, j_s\}: \text{Menge (Indizes/Namen) der irreversiblen Reaktionen;} \\
 w+z &= q, \quad \mathit{rev} \cap \mathit{irrev} = \emptyset
 \end{aligned}
 \tag{2.4}$$

Zwei simple Beispiele von Reaktionsnetzwerken (NB1 und NB2), die später des öfteren der Illustration dienen, sind graphisch und nach Notation (2.4) in Abb. 2.2 bzw. 2.3 dargestellt:

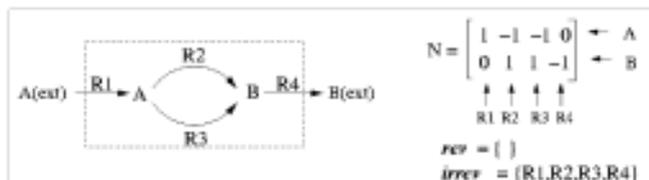


Abb. 2.2: Netz-Beispiel 1 (NB1)

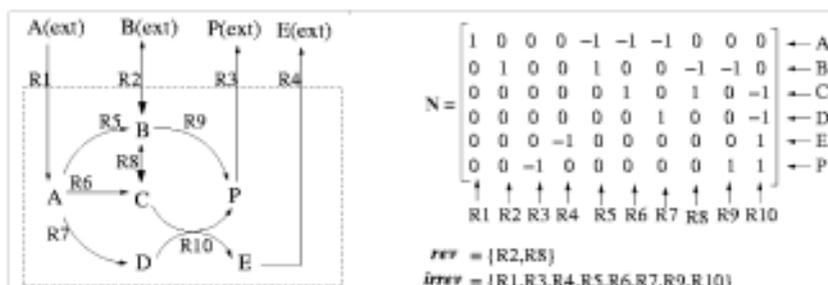


Abb. 2.3: Netz-Beispiel 2 (NB2)

Die stöchiometrische Matrix N für ein gegebenes Reaktionssystem muß nicht eindeutig sein, da die Spalten reversibler Reaktionen mit -1 multipliziert werden können (die Stöchiometrien der reversiblen R2 und R8 in NB2 beziehen sich auf die Richtung des größeren Pfeilkopfes). Dieser Aspekt ist aber ohne Belang für die Ergebnisse struktureller Analysen, genauso wie die Reihenfolge der Metabolite und Reaktionen in N .

Wichtig ist zu bemerken, daß sogenannte *externe Metabolite* bei Strukturanalysen nicht in der stöchiometrischen Matrix berücksichtigt werden [56]. Man kann sie sich quasi als (gepufferte) Senken oder Quellen vorstellen, die außer- oder auch innerhalb der Systemgrenzen liegen können. Insofern erfüllen die in N eingetragenen Reaktionsstöchiometrien nicht immer die Massenbilanz. In NB1 und NB2 liegen die externen Metabolite (mit „ext“ gekennzeichnet) außerhalb der eingezeichneten Systemgrenze und sind folglich nicht in N berücksichtigt.

Wachstumsbedingte Verdünnung und/oder Verbrauch von Metaboliten in der Biomassessynthese können stöchiometrisch leicht durch eine weitere Spalte in N berücksichtigt werden, die dann mit der (Pseudo-)„Reaktion“ Wachstum assoziiert wird (siehe auch Kap. 4).

In den weiteren Abschnitten des Kapitels folgt ein Überblick über die breite Palette an Methoden für die Strukturanalyse metabolischer Netze. Wir werden sehen, daß viele physiologische Aspekte allein aus der Struktur abgeleitet werden können, die wegen des Auftretens der stöchi-

metrischen Matrix in Gl. (2.2) immer auch für das dynamische Verhalten von Bedeutung sind. Die verschiedenen Techniken sollen kritisch beleuchtet werden, um die favorisierte Anwendung des Elementarmodenansatzes in den späteren Kapiteln zu motivieren.

2.2 Strukturbasierte Aussagen zur Existenz, Stabilität und Eindeutigkeit von stationären Zuständen

Steht lediglich die Netzstruktur (2.4) eines Reaktionssystems zur Verfügung, so erscheint es kaum möglich, das Vorhandensein von periodischen Oszillationen oder global stabilen stationären Zuständen vorherzusagen, die charakteristisch für das dynamische Verhalten des Systems (2.2) sind. Daß solche Aussagen dennoch für bestimmte Netzwerke allein aus der Stöchiometrie und ohne Kenntnis kinetischer Parameter ableitbar sind, zeigen die Arbeiten von Feinberg [6,36,37] und die Methode von Clarke [17,18,30]. Beide Ansätze setzen in den meisten Fällen das Massenwirkungsgesetz (MWG) als kinetischen Mechanismus voraus. Außerdem werden nur irreversible Reaktionen betrachtet, d.h. reversible Reaktionen werden in zwei irreversible aufgespalten. Die prinzipiellen Möglichkeiten dieser stöchiometrischen Stabilitätsuntersuchungen sollen hier kurz skizziert und Gemeinsamkeiten mit anderen Strukturanalysen angedeutet werden.

Feinbergs Theorie erlaubt für einige Reaktionsnetzwerke, deren Struktur bestimmte Eigenschaften wie „Deficiency Zero“ oder „schwache Reversibilität“ aufweisen [36], Aussagen über die Existenz und Stabilität von stationären Lösungen und über die Nichtexistenz von periodischen Oszillationen anzugeben. Die wichtigsten Aussagen für solche Netzwerke wurden im „Deficiency Zero Theorem“ festgehalten [36,37]. Dieses Theorem ermöglicht z.B. für das Netz NB1 (Abb. 2.2) die Schlußfolgerung, daß, unabhängig vom gewählten kinetischen Ansatz, keine periodischen Oszillationen auftreten können und es für jeden möglichen Satz an kinetischen Parametern jeweils einen global asymptotisch stabilen stationären Punkt gibt, wo allen Konzentrationen größer Null sind. Eine andere Konsequenz des Theorems ist, daß periodische Prozesse wie der Zellzyklus nur mittels schwach reversibler Reaktionsnetze, die nicht die Deficiency Zero Eigenschaft besitzen, beschrieben werden können [6]. Ähnliche Theoreme wurden von Feinberg auch für Netzwerke mit „Deficiency One“ aufgestellt, die auch Aussagen über Multistationarität beinhalten [36,37].

Clarke's Methode ermöglicht ebenfalls Zusammenhänge zwischen der Struktur eines Reaktionsnetzes (mit MWG) und seiner Dynamik herzustellen. Eine zentrale Rolle in diesen Analysen nehmen die *extreme currents* (extreme Ströme) ein. Extreme Ströme sind bis auf Skalierung eindeutig bestimmte Vektoren, die aus der stöchiometrischen Matrix ableitbar sind. Sie repräsentieren kleinste Teilnetzwerke, in denen ein stationärer Zustand überhaupt möglich ist. Jedem extremen Strom kann z.B. die Eigenschaft „mischend stabil“ oder „mischend semistabil“ oder „mischend instabil“ zugeordnet werden [17]. Eine Anwendung ist der Test auf Instabilität: Instabile Netzwerke weisen mindestens einen mischend instabilen extremen

Strom auf [17,30]. Für NB1 kann man so – im Einklang mit der Feinbergschen Methode – sofort globale Stabilität ableiten, egal welche Parameter für die MWG Kinetik verwendet werden. Wie wir noch sehen werden, sind die Elementarmoden, die uns als zentrales Instrument in späteren Kapiteln dienen werden, den extremen Strömen sehr ähnlich. In Netzen mit nur irreversiblen Reaktionen sind sie sogar identisch (Abschnitt 2.4.4).

Feinberg's und Clarke's Ansätze wurden insbesondere in kleineren (chemischen) Reaktionssystemen erfolgreich angewendet. Für biochemische Reaktionsnetze sind die Ansätze nur bedingt übertragbar und bisher in der Literatur auch kaum eingesetzt worden: Zunächst müssen alle Elementarschritte (auf Enzymebene) einzeln betrachtet werden, deren Reihenfolge, insbesondere bei bi- oder trimolekularen Reaktionen, z.T. noch nicht völlig aufgeklärt ist und auch zu einer großen Anzahl von Zwischenkomplexen führen kann. Desweiteren kann für einige Reaktionen in der Zelle (insbesondere Transportreaktionen) nicht das MWG als Kinetik vorausgesetzt werden. Und schließlich können „Konstanten“ (wie die Enzymmenge) variieren. Da wir in dieser Arbeit ohnehin auf *funktionale* Aspekte von Stoffwechselnetzen fokussieren wollen, werden wir den globalen Stabilitätsanalysen nicht weiter nachgehen.

2.3 Graphentheoretische Strukturanalysen

Graphentheoretische Studien von (insbesondere großskaligen) Stoffwechselnetzen haben in den letzten Jahren einen starken Zuspruch erhalten [65,87,167].

Ein Graph ist ein Tupel $G = (V, E)$, bestehend aus der nichtleeren Menge V der Knoten, sowie einer Menge E der Kanten, die zweielementige Teilmengen aus V sind und die Verbindungen zwischen den Knoten im Graphen repräsentieren [54]. In dieser Notation ist z.B. NB1 (Abb. 2.2) definiert mit $V = \{A_{ext}, A, B, B_{ext}\}$ und $E = \{[A_{ext}, A], [A, B], [A, B], [B, B_{ext}]\}$, d.h. die Metabolite bilden die Menge V der Knoten und die Reaktionen die Menge E der Kanten. Bei *gerichteten* Graphen sind Kanten in nur einer Richtung gültig, d.h. $[A, B] \neq [B, A]$. In NB1 sind alle Kanten gerichtet, da sie für irreversible Reaktionen stehen. Reversible Reaktionen können über zwei gerichtete Kanten abgebildet werden.

Diese Definition zeigt, daß Reaktionsnetze *per se* im allgemeinen keine Graphen sind. In Graphen verbinden Kanten immer genau einen Anfangs- mit einem Endknoten, d.h. sie können nur monomolekulare Reaktionsmechanismen darstellen, wie sie z.B. durch Isomerasen katalysiert werden. An den meisten metabolen Reaktionen sind aber mehrere Substrate bzw. Produkte beteiligt. Stoffwechselnetze sind genau genommen (*gerichtete*) *Hypergraphen* [174]. In diesen verallgemeinerten Graphen können die (Hyper-)Kanten zwei *Mengen* von Knoten verknüpfen, d.h. eine Kante kann mehrere Anfangs- und mehrere Endpunkte haben. NB2 (Abb. 2.3) ist so ein Hypergraph, da Reaktion R10 ein bi-bi-Mechanismus ist. Um graphentheoretische Methoden auch auf Stoffwechselnetze anwenden zu können, muß der Hypergraph folglich in eine Graphenstruktur überführt werden. Die zwei wichtigsten Verfahren sind:

- (1) *Substrat-Graph*: Die Metabolite stellen die Menge der Knoten dar. Zwischen zwei Metaboliten A und B gibt es eine Kante
- (1a) wenn beide an einer Reaktion (als Edukt oder Produkt) beteiligt sind (siehe [167]),
- (1b) oder nur, wenn A Edukt in einer Reaktion ist, wo B Produkt ist (in [87] verwendet).
- (2) *Bipartiter Graph*: In bipartiten Graphen gibt es zwei Klassen von Knoten, da hier sowohl die Metabolite als nun auch die Reaktionen Knoten darstellen. Eine Kante kann nur zwischen einem Metaboliten und einer Reaktion (Edukt-Beziehung) oder zwischen einer Reaktion und einem Metaboliten (Produkt-Beziehung) verlaufen (verwendet in [65]).

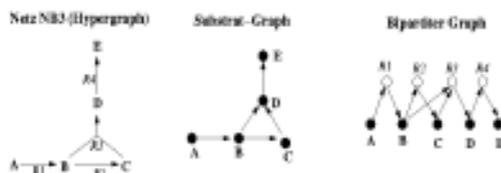


Abb. 2.4: Netz-Beispiel NB3: in der originalen Hypergraphen-Darstellung (links) und in Substrat- bzw. bipartiter Graphen-Darstellung.

Abb. 2.4 zeigt das hypergraphische Reaktionsnetz NB3 und seine graphischen Darstellungen nach Methode (1b) und (2). Nachteil von Methode (1) ist, daß verschiedene Hypergraphen den gleichen Substrat-Graphen haben können. Ersetzt man z.B. Reaktion R3 ($B+C \rightarrow D$) durch die zwei Reaktionen $B \rightarrow D$ und $C \rightarrow D$, ändert das den resultierenden Substrat-Graphen nicht.

Graphenmodelle von rekonstruierten Stoffwechselnetzen verschiedener Organismen wurden auf ihre graphentheoretischen Eigenschaften untersucht [65,87]. Dabei stellte sich heraus, daß alle Netze hinsichtlich dreier Merkmale homologe Strukturen aufweisen:

- **Konnektivität**: Die meisten Metabolitknoten sind nur schwach vernetzt, d.h. die Anzahl k der Kanten, die einen Metaboliten mit anderen verknüpfen (k =Konnektivität), ist im Mittel klein. Das spiegelt den Fakt wider, daß stöchiometrische Matrizen dünn besetzt sind, da an jeder Reaktionen jeweils nur ein Bruchteil aller Spezies teilnimmt. Allerdings stellte man fest, daß die Varianz der Konnektivität viel höher ist als in vergleichbaren, zufällig erstellten Graphen: Viele Metabolite sind zwar schwach verknüpft, einige wenige weisen aber sehr viele Verbindungen auf. Zu diesen zentralen Knoten (engl.: *Hubs*) gehören Co-substrate wie ATP, ADP, NAD(P)H oder CoA, aber auch Gerüstsubstanzen wie Pyruvat oder Glutamat. Die Verteilung der Konnektivität k folgt in metabolen Netzen letztlich einem Potenzgesetz, d.h. der Wahrscheinlichkeit $P(k) \sim k^{-c}$, wobei c netzspezifisch ist [65,87,167]. Graphen mit dieser Eigenschaft nennt man skalenfrei (engl.: *scale-free*, [95]).

- *Cluster*: Stoffwechselnetze sind – wiederum verglichen mit zufälligen Graphen – relativ stark geclustert, d.h. wenn A mit B verbunden ist und B mit C, dann mit relativ hoher Wahrscheinlichkeit auch A mit C. Daher lassen sich Gruppen von Metaboliten finden, in denen letztere stärker untereinander verknüpft sind als mit der Umgebung. Diese Cluster sind daher Ausgangspunkte bei der Modularisierung von Stoffwechselnetzen [42,57].
- Die *Pfad- oder Weglänge* (Abstand) zwischen zwei Knoten A und B ist die kleinste Anzahl an Kanten, entlang derer man von A nach B kommt [54,87]. Im bipartiten und Substrat-Graphen von NB3 wäre die Weglänge von A nach E z.B. 3 (über R1, R3, R4; im bipartiten Graphen zählt man nur jede zweite Kante). Aufgrund der geringen Konnektivität hätte man vermuten können, daß die durchschnittliche Weglänge über alle Paare von Metaboliten relativ groß ist. Tatsächlich erwies sie sich aber als erstaunlich gering, obgleich sie organismenweit variieren können [65,87,167].

Mit diesen drei charakteristischen Merkmalen lassen sich Stoffwechselnetze über alle Organismenreiche hinweg als *small world* Graphen klassifizieren [65,87,95,167]. So eine globale Graphenstruktur kennt man z.B. auch vom Internet oder soziologischen Netzwerken. Sie ist gerade gekennzeichnet durch nur wenige, stark vernetzte Knoten, geclusterte Bereiche und trotzdem relativ kurze Distanzen zwischen den Knoten. Zwei wichtige Konsequenzen für den Stoffwechsel wurden daraus in [167] formuliert: Kurze Weglängen ermöglichen eine schnelle Weiterleitung von (und dann Antwort auf) Störungen, die z. B. durch hohe oder niedrige Metabolitkonzentrationen entstehen können. Die zweite Schlußfolgerung besteht darin, daß Stoffwechselnetze in der Evolution sehr wahrscheinlich sukzessive gewachsen sind, und zwar so, daß neue Kanten (Reaktionen) vornehmlich an die Hubs angesetzt wurden. Stoffwechselwege, in denen viele Hubs auftreten (wie z.B. Glykolyse oder Krebs-Zyklus) sind damit sehr wahrscheinlich evolutionär am ältesten.

Obwohl diese Arbeiten zu einem tieferen Verständnis des globalen Aufbaus von Stoffwechselnetzen beitrugen, bringt der graphische Ansatz einige technische Probleme mit sich, die letztlich auf der Berechnung der Pfadlängen beruhen. So ist das Phänomen der kurzen Pfade in metabolen Graphenmodellen sehr leicht darauf zurückzuführen, daß Cosubstrate wie ATP oder NAD(P)H als „Abkürzungen“ wirken und eigentlich „weit entfernte“ Metabolite über kurze Wege verbinden. Dies führt zu der etwas überraschenden Tatsache, daß zwei im Kohlenstoffgerüst völlig unterschiedliche Metabolite wie Fructose-6-phosphat (F6P, in der Glykolyse lokalisiert) und Succinat (im Citrat-Zyklus (TCA)) gerade mal den Abstand 2 haben (Abb. 2.5): Mittels der Phosphofruktokinase kann F6P mit ATP zu Fructose-1,6-bisphosphat phosphoryliert werden. Als Nebenprodukt entsteht dabei auch ADP, das neben Succinyl-CoA ein Substrat der Succinat-CoA-Synthetase ist, die daraus Succinat und ATP synthetisiert. Hier ist ADP der verbindende Knoten, der zwar tatsächlich einen *kinetischen* Link darstellt [167], allerdings scheint eine Pfadlänge von 2 kein adäquater Wert für die Distanz der beiden Metabolite zu sein, da die Synthese von Succinat aus F6P (oder umgekehrt) natürlich weitaus mehr

Reaktionen benötigt. In [87] versuchte man dieses Problem zu umgehen, indem man Cosubstrate für die Berechnung von Pfadlängen nicht berücksichtigte. In der Tat wurden die ermittelten Abstände zwischen den Metaboliten größer. Allerdings ist die (manuelle) Zuweisung des Kriteriums „Cosubstrat“ letztlich subjektiv. So kann auch Glutamat in vielen Reaktionen als Cosubstrat (Übertragung einer Amino-Gruppe) verstanden werden.

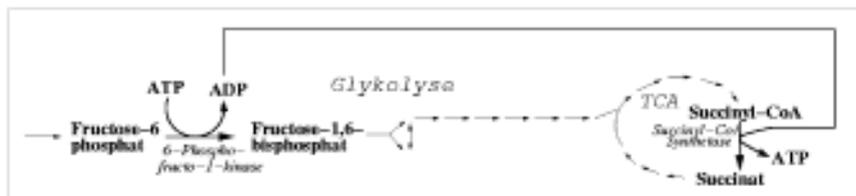


Abb. 2.5: Effekt der kurzen Wege in Graphenmodellen von Stoffwechselnetzen: Der Abstand zwischen Fructose-6-phosphat und Succinat ist in Graphen, aufgrund der „Abkürzung“ über ADP, nur 2.

Diese konzeptionellen Schwierigkeiten in Graphen können letztlich darauf zurückgeführt werden, daß die hypergraphische Natur des Stoffwechsels, d.h. daß mehrere Substrate und Produkte an einer Reaktion beteiligt sein können, nicht adäquat berücksichtigt wird. Das verdeutlicht auch Netz NB3 (Abb. 2.4): die Pfadlänge von A nach E ist 3 (R1, R3, R4), obwohl ganz offensichtlich auch die vierte Reaktion (R2) essentiell für die Produktion von E aus A ist, da C als Zwischenprodukt für die Synthese von E benötigt wird. Man kann sogar noch weitergehen. Entfernt man in NB3 die Reaktion R2, so hat das im resultierenden Graphen keinen Einfluß auf die Pfadlänge von A nach E und es entsteht (sowohl im bipartiten als auch im Substrat-Graphen) sogar der falsche Eindruck, daß die Synthese von E noch immer möglich ist (Abb. 2.6). Eine Pfadanalyse in Graphen hat daher nur eine begrenzte Aussagefähigkeit für die Funktionalität des Originalnetzes. Allerdings läßt sich aussagen: Existiert im Graphen kein Pfad zwischen zwei Metaboliten A und B, dann kann B auch im Hypergraphen nicht aus A produziert werden.

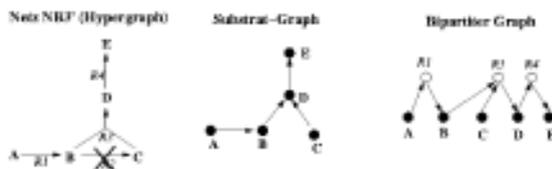


Abb. 2.6: Der resultierende bipartite bzw. Substrat-Graph nach Entfernen der Reaktion R2 aus NB3.

Die folgenden Ansätze setzen sich mit der stöchiometrischen Matrix auseinander und operieren deshalb direkt auf der hypergraphischen Struktur des Stoffwechsels.

2.4 Stationäre Flußverteilungen

Beobachtet man die Metabolitkonzentrationen in der Zelle, so stellt man fest, daß diese sich weitgehend im Fließgleichgewicht (steady-state) befinden [149]. Das ist darin begründet, daß die relativen Umsätze, d.h. die Verhältnisse zwischen Reaktionsraten und Metabolitkonzentrationen, zumindest im Zentralstoffwechsel sehr hoch sind [149,156]. So wird der gesamte ATP-Pool in einer Bakterienzelle in nur wenigen Sekunden umgesetzt. Deshalb pegeln sich die Konzentrationen auch nach Anregungen relativ schnell auf u.U. neuen Niveaus ein (im Sekunden- bis Minutenbereich [156]). Daher bietet es sich an, die Ableitungen der Konzentrationen in Gl. (2.2) approximativ auf Null zu setzen, d.h. Quasistationarität anzunehmen:

$$\mathbf{0} = \mathbf{N}\mathbf{r} \quad (2.5)$$

In (2.5) betrachten wir die (Netto-)Reaktionsraten als die unbekannt Variablen, deren implizite Abhängigkeit von den Konzentrationen und kinetischen Parametern (2.3) nicht bekannt sein muß. Das Differentialgleichungssystem wurde so auf ein homogenes, lineares Gleichungssystem reduziert. Die erhaltenen Massenbilanzgleichungen sind die wohl wichtigsten Nebenbedingungen in metabolen Strukturanalysen und fordern für jeden Metaboliten, daß die Summe der produzierenden Flüsse gleich der Summe der konsumierenden ist. Sinngemäß entspricht das dem 1. Kirchhoffschen Gesetz (Knotenregel) in der Elektrotechnik [28,165].

In einigen Organismen konnten für bestimmte Bedingungen periodische Oszillationen nachgewiesen werden (z.B. [124,171]). Obwohl nur selten erwähnt, hat die Bilanzgleichung auch in diesem Fall eine maßgebliche Bedeutung, denn dann müssen die Mittelwerte \bar{r} der Reaktionsraten ebenfalls (2.5) erfüllen (T = Periodendauer):

$$\mathbf{0} = \mathbf{x}(t_0 + T) - \mathbf{x}(t_0) = \int_{t_0}^{t_0+T} \mathbf{N}\mathbf{r}(t) dt = \mathbf{N} \int_{t_0}^{t_0+T} \mathbf{r}(t) dt = \mathbf{N}(T\bar{\mathbf{r}}) = T\mathbf{N}\bar{\mathbf{r}} \rightarrow \mathbf{0} = \mathbf{N}\bar{\mathbf{r}} \quad (2.6)$$

2.4.1 Nullraum der stöchiometrischen Matrix

In der linearen Algebra bezeichnet man die Lösungsmenge von (2.5) den *Kern* oder (*rechten*) *Nullraum* von \mathbf{N} [56,150]. Die Dimension des aufgespannten Nullraums, d.h. die Anzahl der Freiheitsgrade von (2.5), ist $q - \text{Rang}(\mathbf{N})$. Lediglich wenn $q = \text{Rang}(\mathbf{N})$ gilt, ist der Nullvektor die einzige (nulldimensionale) Lösung für \mathbf{r} . In diesem Fall kann das Netz keine nichttriviale stationäre Flußverteilung erzeugen. Da in den meisten Stoffwechselnetzen die Anzahl Reaktionen (q) größer als die Anzahl Metabolite (m) ist, ergibt sich gewöhnlich ein mehrdimensionaler Nullraum. Mit der Kernmatrix \mathbf{K} gibt man eine Basis des Nullraums an, die dann

$$\mathbf{0} = \mathbf{N}\mathbf{K} \quad (2.7)$$

erfüllt, wobei die Matrizen \mathbf{K} und $\mathbf{0}$ die Dimension $q \times (q - \text{Rang}(\mathbf{N}))$ bzw. $m \times (q - \text{Rang}(\mathbf{N}))$ haben. Eine mögliche Kernmatrix von Netz NB1 ($q=4, m=2; \text{Rang}(\mathbf{N})=2$) ist folgende:

$$\mathbf{K} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 0 \\ 1 & 1 \end{pmatrix} \quad (2.8)$$

Alle stationären Flußverteilungen \mathbf{r} sind Linearkombinationen \mathbf{a} der Spalten in \mathbf{K} :

$$\mathbf{r} = \mathbf{K} \mathbf{a} \quad (2.9)$$

Die Kernmatrix ist leicht berechenbar und erlaubt einige strukturelle Aussagen:

Blockierte Reaktionen: Über Nullzeilen in \mathbf{K} können „blockierte“ Reaktionen identifiziert werden, deren Raten im Fließgleichgewicht *immer* Null sind (in [56] „strikt balanciert“ Reaktionen genannt). Ein einfaches Beispiel ist eine Reaktion mit Beteiligung eines „dead-end“-Metaboliten (nur an dieser einen Reaktion beteiligt), es können aber auch weitaus komplexere und weniger intuitive Fälle auftreten. Blockierte Reaktionen sind praktisch ohne funktionellen Nutzen und weisen deshalb auf eventuelle strukturelle Inkonsistenzen hin.

Enzyme Subsets: Mittels Kernmatrix können auch Enzyme Subsets (ESs) detektiert werden. Jedes ES ist eine Gruppe von Reaktionen, die im Fließgleichgewicht immer gemeinsam und in einem fixen Verhältnis ablaufen [116]. Ist die Rate einer Reaktion R_x Null, so folglich auch alle Raten der Reaktionen, die mit R_x in einem ES sind. Man kann die ES leicht über die Kernmatrix identifizieren: die entsprechenden Zeilen in \mathbf{K} von je zwei Reaktionen aus einem ES unterscheiden sich nur um ein Vielfaches. Typische Beispiele von ESs sind die Reaktionen in einem linearen (unverzweigten) Reaktionspfad. Es gibt aber auch andere Beispiele. So bilden $\{R_1, R_4\}$ in NB1 ein ES und ihre stationären Reaktionsraten haben immer ein Verhältnis von 1 (siehe (2.8)), sind also immer identisch. In NB2 bilden $\{R_4, R_7, R_{10}\}$ das einzige ES. Da die Reaktionen eines ESs immer gemeinsam operieren, ist es sehr wahrscheinlich, daß die Gene der katalysierenden Enzyme einem gemeinsamen Regulationsmuster unterworfen sind [75,143]. In Strukturanalysen können alle Reaktionen eines ESs zu einer Gesamtreaktion zusammengefaßt werden, so z.B. R_1 und R_4 in NB1 zu $B \rightarrow A$. Tauscht man R_1 und R_4 in \mathbf{N} durch diese Reaktion aus, reduziert man damit die Anzahl der Reaktionen.

Die Analyse des Lösungsraums mittels Kernmatrix hat aber einige Nachteile: zum einen ist \mathbf{K} i.a. nicht eindeutig und zum anderen können Restriktionen wie z.B. irreversible Reaktionen (die keine negativen Raten aufweisen dürfen) nicht direkt berücksichtigt werden. So ist der negative Wert für R_3 in (2.8) „physiologisch“ nicht sinnvoll, da R_3 irreversibel ist.

2.4.2 Optimale Flußverteilungen – Flux Balance Analysis

Optimale Flußverteilungen im Stoffwechsel sind sowohl physiologisch als auch biotechnologisch von Bedeutung. Die Maximierung der Wachstumsrate (oder Biomasseausbeute) scheint, zumindest bei einfacheren Organismen wie *Escherichia coli*, ein maßgebliches „natürliches“ Kriterium für das Justieren metaboler Flüsse zu sein [27,61]. Maximierung der Ausbeute biochemischer Produkte ist dagegen zentrales Ziel des *Metabolic Engineering*, wofür genetische und andere Manipulationen im Stoffwechsel vorgenommen werden [97,149,164]. Die Bestimmung einer optimalen Flußverteilung in Reaktionsnetzen ist gewöhnlich eine lineare Optimierungsaufgabe mit folgender allgemeiner Gestalt:

$$\begin{aligned} \text{Zielfunktion: } & \mathbf{c}^T \mathbf{r} = c_1 r_1 + c_2 r_2 + \dots + c_n r_n = \max! \\ \text{Nebenbedingungen: } & \mathbf{N} \mathbf{r} = \mathbf{0} \\ & \mathbf{A} \mathbf{r} = \mathbf{b} \\ & \alpha_i \leq r_i \leq \beta_i \end{aligned} \tag{2.10}$$

Als notwendige Nebenbedingung muß die zu ermittelnde optimale Lösung \mathbf{r} die Massenbilanz (2.5) erfüllen, also im Nullraum liegen. Weitere Gleichungen können mittels Matrix \mathbf{A} und Vektor \mathbf{b} definiert werden, z.B. kann man so die Aufnahmemenge eines Substrats festlegen. Matrix \mathbf{A} und \mathbf{N} können auch zusammengelegt werden. Ein Vorteil der linearen Optimierung ist, daß auch untere (α_i) und obere (β_i) Grenzen der Reaktionsraten als Nebenbedingungen angegeben und somit insbesondere evtl. Irreversibilitäten berücksichtigt werden können ($\alpha_i = 0$ für irreversible Reaktionen). Der Vektor \mathbf{c} definiert die Koeffizienten der linearen Zielfunktion. Soll z.B. Rate r_1 maximiert werden, wäre $c_1=1$ und alle anderen c_k setzt man auf Null. Mit Methoden der linearen Optimierung kann (2.10) problemlos gelöst werden, vorausgesetzt die Nebenbedingungen sind konsistent [8].

Der skizzierte Ansatz der linearen Optimierung von Stoffflüssen in metabolen Netzen wird in der Fachliteratur auch als *Flux Balance Analysis* (FBA) bezeichnet und wurde in zahlreichen Applikationen für oben beschriebene Zwecke eingesetzt [9,26,27,28,29,61,71,118,164]. Mit FBA kann man die „extremen“ Fähigkeiten eines Stoffwechselnetzes studieren, einschließlich der Frage, ob eine bestimmte Funktion *überhaupt* realisiert werden kann. Dazu optimiert man einfach bzgl. der Rate der entsprechenden Reaktion und überprüft dann, ob diese Rate in der Optimierung einen Wert größer Null erreicht hat oder nicht. So läßt sich z.B. auch testen, ob eine Mutante (mit einer fehlenden Reaktion) noch wachsen kann [29].

FBA findet grundsätzlich nur eine spezifische (optimale) Lösung. Der Lösungsraum der realisierbaren Flußverteilungen, d.h. die gesamte Flexibilität und Funktionalität des Netzes, läßt sich mit FBA also nicht untersuchen (siehe auch Kapitel 6). Suboptimale – aber u.U. dennoch wichtige – Lösungen werden ebenfalls nicht betrachtet. Außerdem ist zwar der optimale Wert der Zielfunktion in (2.10) eindeutig, i.a. jedoch nicht die optimale Lösung an sich, auch wenn

der Lösungsraum im Vergleich zu (2.5) eingeschränkt ist. Wollen wir z.B. in NB2 die Ausbeute von P bei Wachstum auf Substrat A maximieren (B nicht angeboten), so wäre

- \mathbf{N} wie in Abb. 2.3 gegeben
- $\mathbf{c}^T = (0,0,1,0,0,0,0,0,0)$
- $\alpha_1 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = \alpha_{10} = 0$; $\alpha_2 = \alpha_9 = -\text{Inf}$; $\beta_i = \text{Inf} \quad \forall i = 1 \dots 10$
- $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$; $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (2.11)

($\text{Inf} = \text{unbeschränkt}$). Alternativ zur ersten Zeile in \mathbf{A} und \mathbf{b} hätte man auch $\beta_1 = 1$ setzen und statt der zweiten Zeile R2 aus dem Netzwerk (und \mathbf{N}) entfernen können. Man sieht leicht, daß in NB2 in jeder bzgl. (2.11) optimalen Lösung aus einem Mol A ein Mol P produziert wird. Abb. 2.7 zeigt drei optimale Lösungen. Die dritte (rechts) wurde durch eine Linearkombination der ersten beiden erzeugt. Es gibt sogar unendlich viele optimale Lösungen, denn jede konvexe Linearkombination $\mathbf{v} = \lambda_1 \mathbf{v}^1 + \lambda_2 \mathbf{v}^2$ ($\lambda_1, \lambda_2 \geq 0$; $\lambda_1 + \lambda_2 = 1$) der ersten beiden Optima ist ebenfalls eine optimale Lösung (wie Abb. 2.7 rechts). Selbst wenn also bekannt ist, daß eine Zelle optimal wächst, ist damit die realisierte Flußverteilung nicht eindeutig identifiziert.

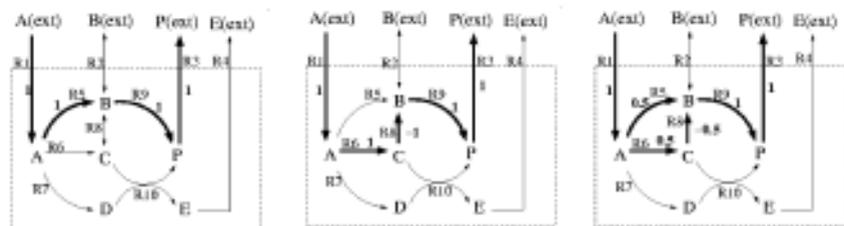


Abb. 2.7: Optimale Lösungen für (2.11) zur Produktion von P aus einem Mol A in NB2.

2.4.3 Elementarmoden und Metabole Pathwayanalyse

Metabole Pathwayanalyse ist die Beschreibung und Analyse von Reaktionssequenzen im Stoffwechsel, denen man eine physiologische oder biotechnologische (ökonomische) Bedeutung zuordnet [75,133,140,141]. Intuitiv stellt man sich unter einem Pathway – mittlerweile ist diese Bezeichnung für „Pfad“ oder „Route“ auch im Deutschen geläufig – eine Reihe aufeinanderfolgender Reaktionen vor, die Ausgangssubstanz(en) in Produkt(c) überführen. In diesem Sinne sollte ein Pathway einen funktionalen und zusammenhängenden Teilbereich des Originalnetzes beschreiben, wie es z.B. bei der bekannten Glykolyse oder dem Citrat-Zyklus der Fall ist. Wie aber kann man Pathways *mathematisch* definieren und dabei den intuitiven und physiologischen Kriterien genügen?

Ein vielversprechender Ansatz ist mit den elementaren Flußmoden (oder Elementarmoden, abgek. EMen) konzipiert worden [56,139,140,141,142], wobei auch Ideen früherer Arbeiten integriert wurden [82,89,146]. Ein Elementarmodus (EM) kann durch einen Vektor \mathbf{e} dargestellt werden, der einer Flußverteilung im Netz entspricht, also q Reaktionsraten enthält. Der Pathway, der durch \mathbf{e} repräsentiert wird, ergibt sich aus den Indizes der tatsächlich verwendeten Reaktionen $\mathbf{R}(\mathbf{e})$, d.h.

$$\mathbf{R}(\mathbf{e}) = \{i: e_i \neq 0\}. \quad (2.12)$$

Wir sagen auch $\mathbf{R}(\mathbf{e})$ enthält die beteiligten (involvierten) Reaktionen von \mathbf{e} . Um tatsächlich ein EM zu sein, muß \mathbf{e} folgende drei Kriterien erfüllen:

- (B1) $\mathbf{N} \mathbf{e} = \mathbf{0}$ (\mathbf{e} ist eine stationäre Flußverteilung, d.h. liegt im Nullraum von \mathbf{N})
- (B2) $e_i \geq 0$ für alle $i \in \text{irrev}$ (irreversible Reaktionen dürfen nur in die thermodynamisch mögliche Richtung laufen)
- (B3) Unzerlegbarkeit: es gibt keine nichttriviale Flußverteilung \mathbf{v} die (B1) und (B2) erfüllt und deren Menge an beteiligten Reaktionen $\mathbf{R}(\mathbf{v})$ eine echte Teilmenge von $\mathbf{R}(\mathbf{e})$ ist, d.h. $\mathbf{R}(\mathbf{v}) \subsetneq \mathbf{R}(\mathbf{e})$.

$$(2.13)$$

(B1) sorgt für einen „echten“ hypergraphischen Ansatz, denn im Gegensatz zu graphentheoretischen Pfaden wird durch das Fließgleichgewicht in \mathbf{e} sichergestellt, daß alle Metabolite, die in den beteiligten Reaktionen des Pathways konsumiert bzw. produziert werden, in der entsprechenden Menge auch synthetisiert bzw. verbraucht werden. Insofern ist ein EM \mathbf{e} zwar eindeutig über $\mathbf{R}(\mathbf{e})$ definiert, jedoch sind die Verhältnisse der Reaktionsraten in \mathbf{e} wegen (B1) ebenfalls von Bedeutung. (B2) grenzt den zulässigen Bereich des Nullraums von \mathbf{N} bzgl. der irreversiblen Komponenten ein. Die Unzerlegbarkeit (B3) ist das charakteristische Merkmal der EMen und sorgt dafür, daß der Pathway sich nicht weiter zerlegen läßt. Dadurch repräsentiert jeder EM \mathbf{e} ein Teilnetz (gegeben durch $\mathbf{R}(\mathbf{e})$), das gerade noch eine Funktion übernehmen kann. Denn wird in \mathbf{e} nur eine beteiligte Reaktion $i \in \mathbf{R}(\mathbf{e})$ entfernt, läßt sich im verbleibenden Teilnetz $\mathbf{R}(\mathbf{e}) \setminus \{i\}$ keine nichttriviale, ausgeglichene Flußverteilung mehr realisieren. (B3) wird auch als „genetische Unabhängigkeit“ bezeichnet [146], weil kein EM auf der Basis der involvierten Reaktionen (Enzymen, Genen) eines anderen EM operieren kann.

Mit den EMen wird der Nullraum von \mathbf{N} , der i.a. unendlich viele (und z.T. thermodynamisch nicht realisierbare) Lösungen enthält, auf eine *endliche Menge von zulässigen und qualitativ unterschiedlichen Lösungen* abgebildet. Die EMen sind für eine gegebene Netzstruktur (2.4) bis auf positive Skalierungsfaktoren eindeutig bestimmt und lassen sich mit einem dem Gaußschen Algorithmus ähnlichen Schema in endlich vielen Iterationen berechnen [141,142]. E soll die Menge der EMen bezeichnen. Die Anzahl $|E|$ der EMen, die meistens viel größer als

die Dimension des Nullraums ist (d.h. die EMen sind i.a. linear abhängig), kann in sehr großen Netzwerken stark ansteigen. In Kap. 8 führen wir dazu eine Komplexitätsanalyse durch.

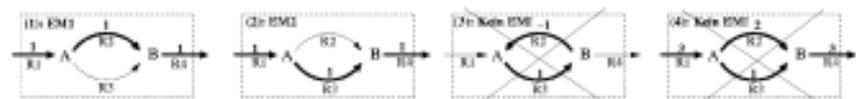


Abb. 2.8: Die beiden Elementarmoden (links) in Netz NB1 und zwei Flußverteilungen, die keinen Elementarmodus darstellen. Die beteiligten Reaktionen und Raten sind hervorgehoben.

Im Netz NB1 gibt es gerade zwei EMen, die in Abb. 2.8 dargestellt sind (links). EM1 ist z.B. ein EM, weil die internen Metabolite A und B in der Flußverteilung balanciert sind (B1), die drei benutzten irreversiblen Reaktionen $R(EM1) = \{R1, R2, R4\}$ vorwärts laufen (B2) und mit keiner echten Teilmenge dieser drei Reaktionen eine stationäre Flußverteilung außer der

trivialen möglich ist (B3). In Vektorschreibweise schreibt man EM1 als $e^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, wobei ein

beliebiges positives Vielfache von e^1 auch möglich wäre. Die dritte Flußverteilung in Abb. 2.8 ist kein EM, da Bedingung (B2) für R2 verletzt ist. Die vierte Flußverteilung ist kein EM, weil Bedingung (B3) nicht erfüllt ist: R2 oder R3 können entfernt werden, und es gäbe noch immer eine realisierbare stationäre Flußverteilung.

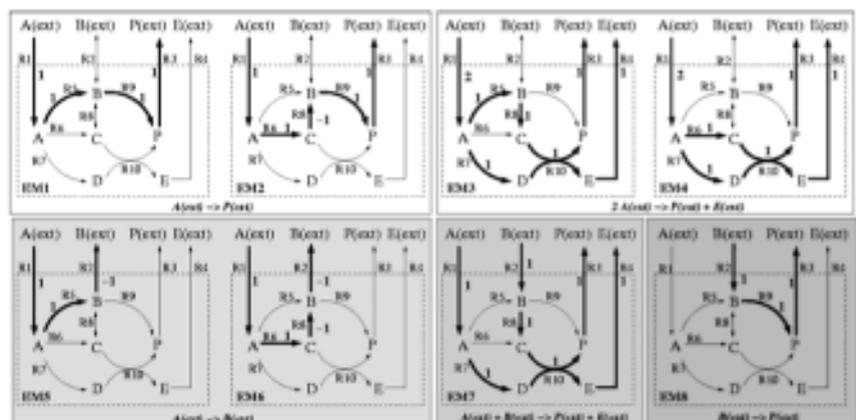


Abb. 2.9: Die acht Elementarmoden von Netz NB2, sortiert nach Nettoumsatz der externen Spezies.

Die systematische Berechnung und Analyse von EMen ermöglicht eine enorme Vielfalt an strukturellen und funktionalen Aspekten in (metabolen) Netzwerken zu detektieren und phy-

siologisch zu bewerten. Die ursprünglichste Anwendung der EMen ist die *Identifikation von Pathways, einschl. von Zyklen*, was zunächst anhand von Netzen mit moderater Größe demonstriert wurde [14,25,101,127,139,140,141,142,143,162]. Die „klassischen“ Stoffwechselwege wie Glykolyse, Citrat-Zyklus oder auch Calvin-Zyklus entsprechen jeweils einem EM (sofern Cometabolite wie ADP/ATP oder NAD/NADH als extern betrachtet werden). Zusätzlich detektiert man aber auch solche Routen, denen man bisher –möglicherweise zu Unrecht– keine Bedeutung zumaß. Andererseits kann man in Fällen, in denen es keinen EM gibt, hypothetische Routen von vornherein ausschließen [117]. Außerdem können sogenannte „futile cycles“ erkannt werden, bei denen im Nettoumsatz nur Energie verbraucht wird (Bsp.: Mini-zyklus aus Fructose-1,6-Bisphosphatase und 6-Phosphofruktokinase). Zyklen ohne jeglichen Verbrauch (externer) Energie deuten auf thermodynamische Unstimmigkeiten hin [119]. Bereits in [140] wurde gezeigt, daß man bestimmte *physiologische Effekte* des Entfernens oder Hinzufügens von Reaktionen (d.h. ihrer Enzyme bzw. Gene) durch EMen testen kann.

Abb. 2.9 zeigt die EMen in NB2 sortiert nach ihrem Nettoumsatz. Ein Nettoumsatz kann wegen (B1) in (2.13) nur für externe Metabolite auftreten. Selbst in diesem einfachen Netz ist es z.B. nicht trivial, alle genetisch unabhängigen Routen zu erkennen, mit denen exklusiv aus Substrat A das Produkt P synthetisiert werden kann (EM1-4). Die EMen lassen auch darauf schließen, daß bei exklusivem Wachstum auf B das Produkt E nicht synthetisiert werden kann. Zyklen gibt es in NB2 keine, wäre R6 reversibel, so würden [R5,R6,R8] einen solchen darstellen.

EMen ermöglichen die *Identifikation aller optimalen und suboptimalen Routen* bzgl. der molaren Ausbeute eines Produktes aus einem Substrat [84,140,141]. Betrachten wir A als exklusives Substrat und P als Produkt, so erkennen wir in NB2 mit EM1 und EM2 zwei optimale Routen (mit Ausbeute $P/A=1$), sowie zwei suboptimale ($P/A=0.5$, EM3 und EM4). Bereits in Abschnitt 2.4.2 sind wir der gleichen Aufgabenstellung mittels linearer Optimierung nachgegangen. Während man mit FBA grundsätzlich nur eine der optimalen Lösungen findet, liefern die EMen die *qualitativ* unterschiedlichen optimalen Lösungen (EM1 und EM2 entsprechen den ersten beiden Lösungen in Abb. 2.7), aus deren nichtnegativen Linearkombinationen sich alle möglichen optimalen Flußverteilungen ableiten lassen (wie in Abb. 2.7 rechts).

Weitere Anwendungen der Pathwayanalyse auf der Basis von EMen (und den *extremen Pathways* – einem abgeleiteten Konzept (s.u.)) wurden jüngst erschlossen und in größeren Netzen, die nun z.T. auch Biomassensynthese berücksichtigten, eingesetzt [113,114,115,120,145,148]. Dabei wird immer stärker deutlich, welche fundamentale Bedeutung die EMen für Stoffwechselnetze haben. Neue Anwendungsfelder, nämlich *kombinatorische Pathwayanalyse*, *Berechenbarkeit von Reaktionsraten* und *Minimale Schnittmengen* (elementare Fehlermoden) werden in dieser Arbeit in den Kapiteln 5-7 vorgestellt und für Fragestellungen des bakteriellen Redox- und Zentralstoffwechsels angewendet.

2.4.4 Theoretische Eigenschaften der Elementarmoden, konvexer Flußkegel und Vergleich mit extremen Pathways

In diesem Abschnitt wollen wir uns mit einigen wichtigen mathematischen Eigenschaften der EMen beschäftigen, auf die wir später z.T. noch zurückgreifen werden. Insbesondere soll dabei der Bezug zu konvexen polyhedralen Kegeln erläutert werden. Der Zusammenhang zwischen den extremen Strömen (Abschnitt 2.2) und konvexen Basen (Kegelerzeuger), sowie einem den EMen ähnlichen Konzept, den *extremen Pathways*, soll hergestellt werden. Beweise werden skizziert, falls sie nicht in der Literatur gefunden worden sind.

(2.14) Es kann (selten) passieren, daß ein EM sowohl in der einen als auch in der anderen Richtung realisierbar ist. In diesem Fall betrachten wir für jede Richtung einen EM.

(2.15) Die Bedingungen (B1)-(B3) in (2.13) sorgen dafür, daß das aus den Reaktionen $R(e)$ eines EM e aufgespannte Teilnetz vollständig verbunden und nicht separierbar ist. Man kann sich leicht überzeugen, daß diese Aussage in Abb. 2.9 erfüllt ist. Bzgl. der an e beteiligten Reaktionen $R(e)$ und involvierten Metabolite $M(e) = \{i: N_{ij} \neq 0 \text{ für mind. ein } j \in R(e)\}$ bedeutet das, es läßt sich keine Zerlegung in je zwei disjunkte (nichtleere) Teilmengen $M1 \cup M2 = M(e)$ und $R1 \cup R2 = R(e)$ finden, so daß in den Reaktionen von $R1$ nur Metabolite aus $M1$ und in den Reaktionen von $R2$ nur Metabolite aus $M2$ involviert sind. Wäre es doch möglich, so können wir leicht den Vektor \hat{e} konstruieren mit $\hat{e}_i = e_i$ (falls $i \in R1$) bzw. $\hat{e}_i = 0$ (sonst). \hat{e} erfüllt dann (B1) und (B2) in (2.13) und verletzt damit (B3) für $e: R(\hat{e}) = R1 \subset R(e)$.

(2.16) Entfernt man eine Reaktion i aus dem Netz, so bilden die EMen aus dem ursprünglichen Netz, in denen $r_i=0$ ist, die neue Menge der EMen (in den verbleibenden EMen ist noch die i -te Komponente e_i zu streichen) [140]. So bilden EMen 1, 4, 5, 8 die Menge der verbleibenden EMen, wenn man in NB2 die Reaktion R8 streicht (Abb. 2.9). Damit können die EMen von „Mutanten“, bei denen jeweils nur eine Untermenge der Reaktionen zur Verfügung steht, sehr leicht aus dem „Wildtyp“ abgeleitet werden.

(2.17) Eine ähnliche Erhaltungsrelation gilt, wenn eine ursprünglich reversible Reaktion i nun als irreversibel deklariert wird: die neue Menge der EMen konstituiert sich aus all jenen EMen des ursprünglichen Netzes, wo $r_i \geq 0$ ist [142]. Setzt man z.B. R8 in NB2 auf irreversibel, entfallen EM2 und EM5 und alle anderen bilden die neue Menge an EMen.

(2.18) Aussage (2.17) hat eine wichtige praktische Bedeutung: Wir nehmen an, Reaktion i sei in einem Reaktionsnetz reversibel (Bsp.: R8 in NB2). Die Menge der resultierenden EMen im Netz kennzeichnen wir mit E_{rev} . Deklarieren wir nun Reaktion i als irreversibel in Vorwärtsrichtung resultiert die Menge E_{fwd} der EMen (im Bsp.: {EM1,3,4,5,7,8}), wobei wegen (2.17) $E_{fwd} \subseteq E_{rev}$ gilt. Nun sei Reaktion i irreversibel in Rückrichtung. Dazu multipliziert man die i -te Spalte in N mit -1 und setzt die Reaktion auf irreversibel. Nach Berechnung der EMen

substituiert man zurück, indem die i -te Komponente in den erhaltenen EMen mit -1 multipliziert wird, da sie die Rückrichtung darstellt. So erhalten wir $E_{\text{rück}}$ (im Bsp.: $\{EM1,2,4,5,6,8\}$) und wiederum gilt $E_{\text{rück}} \subseteq E_{\text{rev}}$. Da also kein neuer EM in den Teilmengen E_{hin} oder $E_{\text{rück}}$ auftritt und andersherum jeder EM $e \in E_{\text{rev}}$ entweder in E_{hin} (falls e_i positiv) oder in $E_{\text{rück}}$ (falls e_i negativ) oder in beiden dieser Mengen (falls $e_i = 0$) enthalten sein muß, gilt: $E_{\text{hin}} \cup E_{\text{rück}} = E_{\text{rev}}$. Wir können also die EMen auch in zwei unabhängigen Schritten bestimmen, nämlich durch getrenntes Berücksichtigen der Hin- und Rückrichtung einer reversiblen Reaktion. Das ist natürlich auch rekursiv für alle reversiblen Reaktionen möglich und wäre ein möglicher Ansatz für eine parallele Berechnung von EMen.

(2.19) Eine weitere Schlußfolgerung ergibt sich daraus für den Fall, daß man eine reversible Reaktion i von vornherein in zwei irreversible aufspaltet. Dazu nimmt man die Stöchiometrie von Reaktion i (i -te Spalte in \mathbf{N}), multipliziert sie mit -1 und fügt sie als weitere Spalte (Rückreaktion von i) in \mathbf{N} ein. So erhält man ein modifiziertes System (2.4)' mit \mathbf{N}' und \mathbf{r}' um jeweils eine Dimension größer ($q' = q + 1$), sowie $\text{rev}' = \text{rev} \setminus \{i\}$ und $\text{irrev}' = \text{irrev} \cup \{i, q+1\}$. In NB2' (Abb. 2.10) haben wir als Beispiel die reversible Reaktion R8 in NB2 in zwei irreversible (R8 und R8r) aufgespalten.

Die Menge der EMen E' im Netz mit aufgespalteener Reaktion i läßt sich nun leicht in zwei Schritten aus den ursprünglichen EMen E ableiten (vgl. EMen in Abb. 2.9 und 2.10):

- (1) Jeder EM $e \in E$ läßt sich eindeutig zu einem äquivalenten $e' \in E'$ übertragen. Dazu hängt man zunächst die zusätzliche Komponente e_{q+1}' für die Rückreaktion von i an und initialisiert sie mit Null. Wenn in e ein negativer Wert für Reaktion i auftritt, so wird dieser als positiver Wert in der hinzugekommenen Komponente e_{q+1}' für die Rückreaktion eingetragen und e_i' auf Null gesetzt. Alles andere bleibt unverändert. Formal geschrieben:
 - (i) ist $e_i = 0$ so ist $e_i' = e_{q+1}' = 0$; für alle anderen $j \neq i, q+1$ gilt $e_j' = e_j$
 - (ii) ist $e_i > 0$ so ist $e_i' = e_i$ und $e_{q+1}' = 0$; für alle anderen $j \neq i, q+1$ gilt $e_j' = e_j$
 - (iii) ist $e_i < 0$ so ist $e_i' = 0$ und $e_{q+1}' = -e_i$; für alle anderen $j \neq i, q+1$ gilt $e_j' = e_j$
 Entsprechend sind EM1'-EM8' in Abb. 2.10 die jeweils äquivalenten EMen von Abb. 2.9.
- (2) Für die aufgespaltene Reaktion fügt man in E' einen 2er-Zyklus \mathbf{x}' hinzu, in dem Hin- und Rückreaktion im gleichen Verhältnis beteiligt sind: $z_i' = z_{q+1}' = 1$ und sonst $z_j' = 0$. Dieser Zyklus ist trivialerweise ein EM in E' . In NB2' ergibt sich folglich EM9'.

Die EMen E und E' sind damit bis auf den in E' hinzugekommenen 2er-Zyklus äquivalent – nur in anderen Dimensionen dargestellt. Wie man E' aus E erhält ist oben beschrieben. Für die Rückrichtung zerlegt man E' in drei (nicht unbedingt disjunkte) Teilmengen:

$$E1' = \{e' : e_{q+1}' = 0\}, E2' = \{e' : e_i' = 0\} \text{ und } E3' = \{e' : e_i' > 0, e_{q+1}' > 0\}.$$

$E1'$ entspricht den EMen von E ohne Verwendung von i in Rückrichtung, d.h. die EMen in $E1'$ sind denen der Menge E_{hin} von (2.18) äquivalent, nur mit zusätzlicher Null-Komponente bei e_{q+1}' . Analog entspricht $E2'$ den EMen in E , wo Reaktion i nicht in Vorwärtsrichtung

involviert ist, d.h. die EMen in $E2'$ sind denen von $E_{\text{rück}}$ äquivalent, nur mit zusätzlicher Null-Komponente bei e_1' . Nach (2.18) ist $E_{\text{fhh}} \cup E_{\text{rück}} = E$, d.h. E ist damit äquivalent in E' (mit jeweils zusätzlicher Null-Komponente für Rück- bzw. Hinreaktion) enthalten. Es bleibt zu klären, welche EMen in $E3'$ zusätzlich zu E auftreten. Per Definition enthält $E3'$ EMen, in denen Hin- und Rückreaktion von i auftreten. $E3'$ kann aber als einziges Element nur den trivialen 2er-Zyklus enthalten. In keinem weiteren EM können Hin- und Rückreaktion simultan beteiligt sein, da ein solcher wegen diesem 2er-Zyklus sofort zerlegbar wäre.

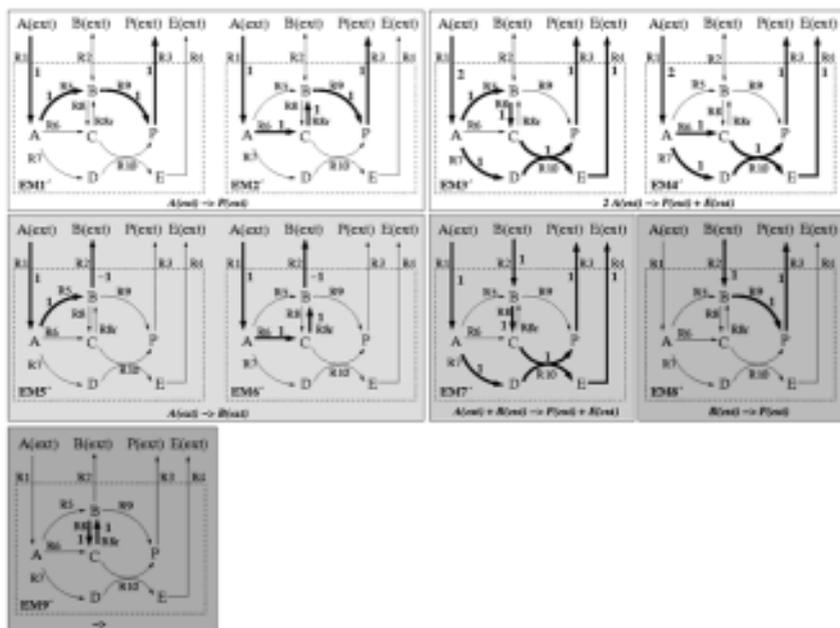


Abb. 2.10: Die neun Elementarmoden von Netz NB2'; in dem im Vergleich zu NB2 die reversible Reaktion R8 in zwei irreversible (R8 und R8r) aufgespalten wurde. Sortierung wie in Abb. 2.9.

Spaltet man (sukzessive) mehrere oder alle reversiblen Reaktionen auf, ist das Ergebnis analog: Die EMen im Netz mit aufgespaltenen Reaktionen sind den ursprünglichen äquivalent, bis auf die hinzukommenden 2er-Zyklen, deren physiologische Interpretation trivial ist. Wir kommen im nächsten Abschnitt auf diese wichtige Eigenschaft der EMen zurück.

Konvexer Kegel und konvexe Basen

Alle zulässigen stationären Flußverteilungen r eines Reaktionsnetzes erfüllen wie die EMen Bedingungen (B1) und (B2) von (2.13), d.h. sie liegen im Nullraum und ihre irreversiblen Komponenten sind nur im nichtnegativen Halbraum lokalisiert:

$$\mathbf{0} = \mathbf{N}\mathbf{r}, \quad r_i \geq 0 \text{ für alle } i \in \text{irrev} \quad (2.20)$$

Der Lösungsbereich dieses Gleichungs-/Ungleichungssystems bildet einen konvexen polyhedralen Kegel (engl.: *cone*) [56,126,139,142]. In metabolischen Netzen nennt man diesen Kegel auch „Flußkegel“ (engl.: *flux cone*) F :

$$F = \{\mathbf{r} \in \mathbb{R}^r : \mathbf{0} = \mathbf{N}\mathbf{r}, \quad r_i \geq 0 \quad i \in \text{irrev}\} \quad (2.21)$$

Ein Teilgebiet der Mathematik beschäftigt sich mit dem Studium solcher konvexen Mengen und Kegel [126]. Ein konvexer polyhedraler Kegel ist abgeschlossen bzgl. nichtnegativer Linearkombinationen und kann mittels einer Menge G von Kegelerzeugern generiert werden [56,126,142]. G wird auch konvexe Basis genannt [116] und ist eine nicht weiter reduzierbare Menge von Vektoren \mathbf{g} , die F durch nichtnegative Linearkombinationen generieren:

$$F = \{\mathbf{r} : \mathbf{r} = \sum \gamma_j \mathbf{g}^j, \quad \gamma_j \geq 0, \mathbf{g}^j \in G\} \quad (2.22)$$

Sofern mit $\mathbf{g} \in G$ automatisch $-\mathbf{g} \notin G$ folgt, ist der Kegel F gepunktet (engl.: *pointed*) und G (bis auf Skalierung) eindeutig [56,126,142]. Die Elemente in G sind dann die Kanten (auch: *extreme rays*) des aufgespannten konvexen Kegels, wie man es z.B. von einer Pyramide her kennt. Gilt obige Forderung nicht (vgl. mit (2.14)), ist der Kegel nicht gepunktet und G nicht eindeutig. Die Anzahl der konvexen Basisvektoren kann, wie bei den EMen, die Dimension des Nullraums von \mathbf{N} deutlich übersteigen, selten aber auch geringer sein.

Der Algorithmus zur Berechnung einer konvexen Basis ist ähnlich dem für die EMen [116]. G steht mit der Menge E der EMen in folgendem Zusammenhang: Es kann immer ein G gefunden werden, so daß $G \subseteq E$. Deshalb kann man den Kegel F und damit die in ihm liegenden, i.a. unendlich vielen realisierbaren stationären Flußverteilungen auch aus nichtnegativen Linearkombination der EMen generieren:

$$F = \{\mathbf{r} : \mathbf{r} = \sum_{j=1}^{|E|} \gamma_j \mathbf{e}^j, \quad \gamma_j \geq 0\} \quad (2.23)$$

Es gibt Fälle, in denen die EMen nicht einer Kante des Kegels entsprechen und durch positive Linearkombinationen anderer geschrieben werden können, z.B. $EM1' = EM5' + EM8'$ in Abb. 2.10. Insofern ist die Menge der EMen i.a. nicht minimal bzgl. der Beschreibung des Kegels F , wie es die konvexen Basen sind. Für die konvexe Basis und auch für die EMen gilt, daß die Zerlegung einer stationären Flußverteilung in (2.22) bzw. (2.23) i.a. nicht eindeutig ist.

Abb. 2.11 veranschaulicht noch einmal die geometrische Interpretation. Dort betrachten wir ein einfaches Netz mit einem Metaboliten und drei Reaktionen, wobei R2 reversibel ist. Der Flußkegel F ist im Raum der Reaktionsraten in grau dargestellt. Er ist nach oben hin offen,

zweidimensional und gepunktet. Es gibt drei EMen, aber nur EM1 und EM3 bilden die hier eindeutige Menge G der Kegelerzeuger (Kanten) von F . EM2 liegt im Inneren des Kegels und kann über die anderen dargestellt werden: $EM2=EM1+EM3$. Dennoch ist EM2 von Bedeutung: Will man z.B. alle optimalen und suboptimalen Routen von $S1(ext)$ zu $S3(ext)$ finden, so ist EM2 der einzige EM, der dafür in Frage kommt. Aus diesem und ähnlichen Gründen (siehe Besprechung *extreme Pathways* unten) ist es wichtig, *alle* genetisch unabhängigen Pathways zu finden. Die konvexen Basen, als Teilmenge der EMen, reichen für solche Studien i.a. nicht aus, auch wenn alle EMen aus ihnen konstruierbar sind.

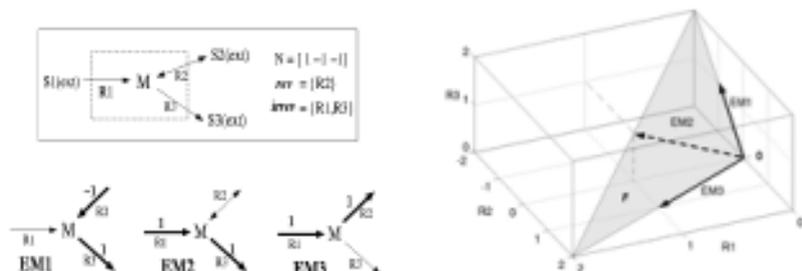


Abb. 2.11: Beispiel für konvexen Kegel, konvexe Basis und Elementarmoden in einem Reaktionsnetz.

Für manche Zwecke ist es bequemer, wenn in Ungleichungssystemen wie (2.20) alle Variablen $r_i \geq 0$ sind. Von der linearen Optimierung kennt man den Trick, eine freie Variable y als eine Differenz aus zwei nichtnegativen Variablen zu schreiben: $y = y^+ - y^-$ mit $y^+, y^- \geq 0$. Das erhöht zwar die Dimension, beschränkt aber das Geschehen vollständig auf den positiven Orthanten. Exakt der gleiche Vorgang ist das Aufspalten einer reversiblen Reaktion in zwei irreversible: die ursprünglich freie Nettoreaktionsrate r_i wird nun als Summe zweier nichtnegativer Raten geschrieben, je eine für Hin- und Rückreaktion: $r_i = r_i^+ - r_i^-$ mit $r_i^+, r_i^- \geq 0$. Spalten wir alle reversiblen Reaktionen wie oben erläutert auf, erhalten wir letztlich für (2.22):

$$\mathbf{0} = \mathbf{N}' \mathbf{r}' \quad , \quad r'_i \geq 0 \quad \text{für alle } i \quad (2.24)$$

In der linearen Optimierung nennt man (2.24) die *Standardform* von (2.20) [8]. Der mit (2.24) aufgespannte Kegel F' liegt ausschließlich im nichtnegativen Bereich. Er ist deshalb gepunktet und hat eine eindeutige, die Kanten von F' darstellende, konvexe Basis G' . Mit einer Normierung der konvexen Basisvektoren (Beschränkung des Kegels auf ein Polytop), entsprechen diese in der linearen Optimierung genau der Menge der *zulässigen Basislösungen*, die völlig unabhängig von der linearen Zielfunktion sind und die Ecken des entstandenen Polytops darstellen [8,40]. Mehr noch, da an (2.24) nur irreversible Reaktionen teilnehmen, sind die Kanten des Kegels F' auch den *extremen Strömen* aus der Clarkeschen Theorie äquivalent

(Abschnitt 2.2), in der grundsätzlich nur Netze vom Typ (2.24) betrachtet werden können. Auch die Menge E' der EMen von (2.24) ist mit den konvexen Basen (und damit auch den extremen Strömen und zulässigen Basislösungen) identisch, und jeder EM entspricht somit auch einer Kante des Kegels F' . Letzteres ergibt sich aus der Tatsache, daß in (2.24) kein EM durch einen anderen dargestellt werden kann, da alle Komponenten in den EMen nichtnegativ sind [139]. Aus (2.19) folgt, daß die ursprüngliche Menge E der EMen von (2.20) der von (2.24) bis auf die 2er-Zyklen der ursprünglich reversiblen Reaktionen äquivalent ist.

Damit haben wir folgendes Resultat: Die EMen vom Originalnetz (2.20) sind, bis auf die hinzukommenden 2er-Zyklen, *den EMen, der konvexen Basis (Kegelerzeuger), den zulässigen Basislösungen und extremen Strömen im Netz mit aufgespaltenen reversiblen Reaktionen äquivalent*. Sie werden in (2.24) lediglich in einer größeren Dimension mit ausschließlich nichtnegativen Komponenten repräsentiert. Diese Tatsache ist in der Literatur bisher kaum dargestellt worden, bedeutet aber, daß die entsprechenden Anwendungen – z.B. der extremen Ströme – leicht auf die EMen übertragen werden können. Andererseits können die extremen Ströme bei einem Netz mit reversiblen Reaktionen leicht aus dessen EMen abgeleitet werden. Leider läßt sich selbst für das kleine Netz in Abb. 2.11 der zugehörige Kegel F' für das System mit Aufspaltung der reversiblen Reaktion R2 nicht mehr darstellen, da dieser als dreidimensionales Gebilde im vierdimensionalen Raum auftritt. Man kann aber leicht überprüfen, daß als einziger EM nur der 2er-Zyklus von $R2_{\text{hin}}$ und $R2_{\text{rück}}$ hinzukommt und kein EM mehr durch einen anderen darstellbar ist und deshalb $E' = G'$ ist.

Extreme Pathways

Mit obigen Vorüberlegungen können wir nun den Zusammenhang zwischen den EMen und den extremen Pathways (EPs) herstellen. Die EPs sind als ein alternatives Konzept zur Beschreibung biochemischer Pathways entwickelt worden [132]. Dazu muß in einem gegebenen Reaktionsnetz (2.4) jede Reaktion zunächst als intern oder als Austauschreaktion (*exchange flux*) deklariert werden. Bei letzteren überschreiten Metabolite die Systemgrenze. Sie entsprechen meistens Transportreaktionen, wobei jeder Metabolit an höchstens einer Austauschreaktion beteiligt sein darf. Alle internen Reaktionen müssen in zwei irreversible Reaktionen aufgespalten werden, reversible Austauschreaktionen bleiben dagegen unverändert. Bzgl. NB2 entspricht das in Abb. 2.9 verwendete NB2' genau diesen Anforderungen, wobei R1-R4 Austauschreaktionen wären (und R2 deshalb nicht aufgespalten wird). Die EPs sind nun definiert als die konvexen Basen in diesem umkonfigurierten Netz und sind deshalb immer eine Teilmenge der EMen dieses (umkonfigurierten) Netzes. Wegen (2.19) können sie, bis auf die bedeutungslosen 2er-Zyklen der aufgespaltenen Reaktionen, auch als eine Teilmenge der EMen im ursprünglichen Netz dargestellt werden, dort aber in einer kleineren Dimension. Im Netz NB2' bilden die EM4'-EM9' die Menge der EPs. EM1-EM3 liegen also im aufgespannten Kegel nicht auf einer Kante (sind durch andere EMen darstellbar) und „fallen deshalb durch das Raster“: $EM1' = EM5' + EM8'$; $EM2' = EM6' + EM8'$ und $EM3' = EM5' + EM7'$.

Durch das Auslassen von EMen bleiben einige genetisch unabhängige Routen des Netzes unberücksichtigt, was Schwierigkeiten bei vielen Anwendungen nach sich zieht:

- Fragen wie „Gibt es in $NB2'$ eine nicht zerlegbare Route von A über B nach P?“ können i.a. nicht mehr sicher beantwortet werden. Die EPs in $NB2'$ enthalten z.B. keine solche Route, obwohl mit EM1 und EM2 offensichtlich welche existieren.
- In der Menge der EPs können optimale Lösungen verlorengehen. In $NB2'$ sind EM1' und EM2' die optimalen Moden für die Synthese von P aus A. Beide sind keine EPs. Daher ist die Behauptung in [114,120] inkorrekt, die Menge der EPs enthalte alle optimalen Pfade.
- Die Erhaltungseigenschaften (2.16) und (2.17) gelten für EPs i.a. nicht. Entfernt man eine Reaktion oder setzt eine reversible Reaktion auf irreversibel, kann es passieren, daß EMen des Netzes, die vorher keinen EP darstellten, nun als EP betrachtet werden (z.B. EM1 wird zu einem EP, wenn R2 in $NB2'$ entfernt wird). Die Menge der EPs muß daher neu berechnet werden. Aus diesem Grund kann die Lebensfähigkeit von Mutanten (mit entfernten Reaktionen) nicht aus den EPs abgeleitet werden.
- Weitere strukturelle Eigenschaften können mit den EPs nicht gefiltert werden: Sowohl die längste (EM1) als auch die kürzeste Route (EM2) von A nach P werden über EPs nicht identifiziert (gemessen an Anzahl der beteiligten Reaktionen). Die in den Kapiteln 5-7 einzuführenden Methoden benötigen ebenfalls *alle* genetisch unabhängigen Routen.

Insofern kann die Menge der EPs i.a. nicht als eine geeignete Approximation der Menge der EMen angesehen werden. Allerdings kommt es in der Realität sehr häufig vor, daß die Austauschreaktionen irreversibel sind (man betrachtet also den Fall Substrat *oder* Produkt und nicht beides gleichzeitig). Dann liegt das Netz wie in (2.24) vor, da nun *alle* Reaktionen irreversibel sind. Wie wir gezeigt haben, sind die EMen in so einem Netz den konvexen Basen äquivalent und damit nun auch den EPs. Solche Fälle liegen z.B. in den Arbeiten [114, 120] vor. Es gibt auch andere Spezialfälle, in denen EMen und EPs zusammenfallen, z.B. wenn alle reversiblen Austauschreaktionen nur solche Metabolite ex/importieren, die jeweils nur noch an einer weiteren internen irreversiblen Reaktion beteiligt sind. Auch wenn zusätzliche Austauschreaktionen eingeführt werden [131], sind die EPs mit den EMen identisch [138]. Eine Zusammenfassung der Gegenüberstellung von EPs und EMen wurde in [75] publiziert.

2.4.5 Metabole Flußanalyse

Nullraum- und Pathwayanalyse widmen sich dem Studium des *gesamten* Raumes bzw. aller minimalen Realisierungen von stationären Flußverteilungen. Bei der metabolen Flußanalyse (MFA) geht es darum, eine konkrete Ausprägung von metabolen Reaktionsraten unter definierten steady-state Bedingungen zu bestimmen. In der Literatur finden sich viele Beispiele von Flußanalysen, die aus biotechnologischem oder rein biologischem Interesse durchgeführt worden sind (z.B. [4,98,149,158]).

Wir hatten bereits erwähnt, daß die durch die Quasistationaritätsannahme abgeleitete Massenbilanz (2.5) gewöhnlich unterbestimmt, d.h. $q > \text{Rang}(N)$ ist. Ansonsten würde auch keine sta-

tionäre Flußverteilung (außer der trivialen) existieren. Deshalb werden weitere Informationen zum Einschränken des Lösungsraumes benötigt, so daß im optimalen Fall eine eindeutige Lösung übrigbleibt. Dazu mißt man extrazelluläre Raten (Aufnahme und Exkretion von Substraten bzw. Produkten) bspw. während kontinuierlicher Fermentationen, wo sich die wachsenden Kulturen im Fließgleichgewicht befinden. Manchmal können zusätzlich auch intrazelluläre Raten auf Null gesetzt werden (z.B. für Respiration unter anaeroben Bedingungen). Um bekannte Reaktionsraten zu berücksichtigen, partitioniert man die stöchiometrische Matrix und den Ratenvektor – gegebenenfalls durch problemloses Vertauschen von Spalten in \mathbf{N} bzw. Komponenten in \mathbf{r} – in einen bekannten (Index b) und unbekanntem Teil (Index u):

$$\mathbf{r} = \begin{pmatrix} \mathbf{r}_b \\ \mathbf{r}_u \end{pmatrix}, \quad \mathbf{N} = (\mathbf{N}_b, \mathbf{N}_u) \quad (2.25)$$

Bekanntem und unbekanntem Anteil kann man in (2.5) nun leicht separieren:

$$\begin{aligned} \mathbf{0} &= \mathbf{N}\mathbf{r} = \mathbf{N}_b\mathbf{r}_b + \mathbf{N}_u\mathbf{r}_u \\ \mathbf{N}_u\mathbf{r}_u &= -\mathbf{N}_b\mathbf{r}_b \end{aligned} \quad (2.26)$$

(2.26) ist nun ein *inhomogenes* lineares Gleichungssystem, das weniger Unbekannte enthält als das homogene (2.5). Sofern \mathbf{N}_u invertierbar ist (dann ist eindeutige und konsistente Lösung für \mathbf{r}_u vorhanden) oder zumindest eine eindeutige Kleinste-Quadrate-Lösung für \mathbf{r}_u vorliegt, können die bekannten algebraischen Lösungsmethoden angewendet werden, wie sie speziell für die MFA in [149,159,160] beschrieben worden sind.

Allerdings passiert es nicht selten, daß (2.26) unterbestimmt bleibt. In der Literatur über MFA und selbst in Standardwerken zur linearen Algebra finden sich kaum praktische Hinweise, wie unterbestimmte Gleichungssysteme zu behandeln sind. Wie wir an einem realen Beispiel zeigen werden, kann es nämlich durchaus passieren, daß trotz globaler Unterbestimmtheit des Systems, einige unbekannte Raten dennoch eindeutig berechenbar sein können und somit das mögliche Verhalten des Netzes an einigen Stellen klar eingeschränkt ist. In Kap. 5 wird systematisch dargestellt, welche Fälle in (2.26) auftreten können und wie mit ihnen zu verfahren ist, wobei insbesondere die Analyse der Berechenbarkeit einzelner Raten herausgestellt wird. Wir werden diese Berechenbarkeitsanalyse verallgemeinern und zeigen, wie man anhand der Elementarmoden feststellen kann, ob eine Rate in einem Netz überhaupt über MFA bestimmbar ist oder nicht und wie man gegebenenfalls Kombinationen von Messungen findet, die diese Rate berechenbar machen.

2.5 Erhaltungsgleichungen

Erhaltungsgleichungen (ERen) sind ein sehr schönes Beispiel dafür, wie Struktureigenschaften die Dynamik eines Reaktionsnetzes beeinflussen können. Jede ER kennzeichnet eine *gewicht-*

tere Summe der Metabolitkonzentrationen, die im System zu jedem Zeitpunkt t , also auch im nichtstationären Zustand, konstant sind. Typisches Beispiel in metabolen Netzen ist die konstante Summe der Konzentrationen von $[NAD]+[NADH] = S$ (falls Biosynthese der Moleküle NAD bzw. NADH nicht betrachtet wird). Da in jeder Reaktion, wo NADH produziert wird, ein NAD-Molekül verbraucht wird (und umgekehrt), ändert sich die Summe ihrer Konzentrationen nicht. Formal erfüllt damit jede ER y folgende Relation [56]:

$$y^T x(t) = S = \text{Konstant} \quad (2.27)$$

Vektor y enthält für jeden Metaboliten einen Koeffizienten. Abgeleitet ergibt (2.27) mit (2.2)

$$(y^T x(t))' = y^T x'(t) = y^T N r(t) = 0 \quad (2.28)$$

Soll (2.28) für einen beliebigen Ratenvektor $r(t)$ gelten, muß $y^T N = 0^T$ sein, wofür wir auch

$$N^T y = 0 \quad (2.29)$$

schreiben können. Eine ER ist also durch linear abhängige Spalten der Transponierten von N , d.h. also durch linear abhängige Zeilen von N gegeben. Die Lösungen liegen im Nullraum der Transponierten von N . Dieser wird auch *linker* Nullraum von N genannt und hat eine Dimension von $m - \text{Rang}(N)$ und damit, wegen der fast immer geltenden Relation $m < g$, meistens eine viel kleinere Dimension als der Nullraum von N in (2.5). Analog zur Kernmatrix K von N kann man eine Basis des Nullraums von N^T in einer Matrix anordnen, aus denen alle möglichen EREN konstruiert werden können (2.9). Gilt $\text{Rang}(N) = m$, gibt es überhaupt keine EREN, wie in unseren Netzbeispielen NB1 und NB2. Da in (2.29) wiederum ein Nullraum untersucht wird, läßt sich das Konzept der Elementarmoden in ein ganz analoges Konzept der *elementaren Erhaltungsrelationen* übertragen [56], indem man (B.1) in (2.13) durch (2.29) ersetzt. Elementare ER lassen sich nicht in kleinere EREN zerlegen und sind deshalb minimal bzgl. der beteiligten Metabolite. Vorzeichenbedingungen (2.13-(B2)) müssen hier nicht geltend gemacht werden, es sei denn man ist an *nichtnegativen* elementaren EREN interessiert.

Zur Illustration betrachten wir ein simples Netz mit vier Metaboliten und nur einer Reaktion: $A + B \rightarrow C + D$. Damit ist $N^T = (-1 \ -1 \ 1 \ 1)$ und die Dimension des aufgespannten Raumes der EREN drei. Eine ER wäre: $ER1 = [A]+[B]-[C]-[D] = S1$. Diese ist aber nicht *elementar*, da $ER2 = [A]+[C] = S2$ nur eine Teilmenge der Metabolite von ER1 enthält. ER2 läßt sich dagegen nicht weiter zerlegen und ist deshalb eine elementare ER. In diesem einfachen System ergibt jede Kombination von zwei der vier Metaboliten eine elementare ER, d.h. es gibt insgesamt 6, wovon 4 nichtnegativ sind. Theoretische Eigenschaften elementarer EREN, sowie die geometrische Interpretation in konvexen Kegeln sind ganz analog zu den EMen [56].

Da EREN sowohl für die strukturelle als auch dynamische Modellierung letztlich redundante Informationen darstellen ist es sinnvoll, die EREN vor weiteren Analysen zu entfernen, indem

man bestimmte m -Rang(N) viele abhängige Zeilen (d.h. Metabolite) aus N entfernt. Kennt man in obigem Beispiel die Anfangskonzentrationen aller Metabolite, so ist es völlig ausreichend, in einem reduzierten System nur noch die Konzentration von A explizit zu modellieren. Die Konzentrationen von B, C und D sind dann zu jedem Zeitpunkt eindeutig aus den ERen ableitbar. Manchmal ist leicht zu sehen, welche Zeile(n) man zur Eliminierung der ER(en) entfernen muß. Tritt z.B. genau eine ER auf, entfernt man die Zeile eines der beteiligten Metabolite. In größeren Netzen kann das aber wesentlich komplizierter werden. Welche Kombinationen von zu entfernenden Zeilen in N alle ERen aufbrechen, kann man bei mehrfachen ERen über *Minimale Schnittmengen* bestimmen, wie in Kap. 7 gezeigt wird.

2.6 Zusammenfassende Übersicht über metabolische Strukturanalysen

Abb. 2.12 zeigt noch einmal eine Übersicht der strukturellen Analysemethoden in Stoffwechselnetzen, aufsteigend sortiert nach dem Umfang der jeweils eingehenden Informationen bzw. Annahmen. Die auf Quasistationarität beruhenden Methoden der Pathwayanalyse, Flux Balance Analysis und metabolischen Flußanalyse werden auch als „constraint-based modeling“ bezeichnet, da sie den Raum der möglichen Flußverteilungen (unterschiedlich stark) durch stöchiometrische, thermodynamische, stationäre, kapazitive, optimale und gemessene Randbedingungen einschränken [121,122]. Erste Ansätze zur Einbeziehung regulatorischer Randbedingungen in die Strukturanalyse gibt es bereits [22]. Die Flußanalyse schränkt den Lösungsraum am weitesten (u.U. auf eine spezifische Flußverteilung) ein, obwohl Reversibilitäten und Kapazitätsgrenzen nur indirekt berücksichtigt werden (z.B. für Konsistenzkontrollen). Strukturmodelle des Stoffwechsels sind – gemessen an den berücksichtigten Komponenten – die bei weitem größten zellulären Modelle überhaupt [121,122].

Methode/ Strukturelle Eigenschaften	Nebenbedingungen (Constraints)				
	Stöchiometrie N	Thermodynamik $r_i \geq 0$	Stationarität $N r = 0$	Kapazitäten $\alpha_i \leq r_i \leq \beta_i$	Messungen $N_s r_s = -N_s r_s$
<i>Erhaltungsgleichungen</i>	■				
<i>Graphentheoretische Strukturanalyse</i>	■	■			
<i>Stabilitätsuntersuchungen (Feinberg, Clarke)</i>	■	■	■		
<i>Analyse der Kernmatrix (Enzyme subsets, ...)</i>	■	■	■		
<i>Elementarmodenanalyse (Metabolic Pathwayanalyse)</i>	■	■	■		
<i>Optimierte Flußverteilungen (Flux Balance Analysis, FBA)</i>	■	■	■	■	■
<i>Metabolische Flußanalyse</i>	■	■	■	■	■

Abb. 2.12: Systematik der Strukturanalysen in Stoffwechselnetzen

3 Zwei ausgewählte Modellsysteme des bakteriellen Zentral- und Redoxstoffwechsels

3.1 Organisation des Stoffwechsels in Bakterien

Das Reich der Bakterien, aufgeteilt in die beiden Hauptgruppen der *Eubacteria* und *Archaeobacteria*, ist gekennzeichnet durch eine große Vielfalt unterschiedlicher Stoffwechselwege. Für fast alle Umwelt- und Nährstoffbedingungen finden sich mehr oder weniger spezialisierte Arten, die mit ihrem Stoffwechselschema die notwendigen Voraussetzungen zum Überleben haben. Bakterien sind Einzeller und müssen deshalb zum Wachstum alle Bestandteile vollständig in eben dieser einen Zelle synthetisieren, d.h. eine Aufgabenteilung wie bei Mehrzellern ist nicht möglich. Als Prokaryonten weisen sie außerdem keine oder nur eine sehr geringe Kompartimentierung auf, d.h. der Metabolismus ist in den Zellen räumlich kaum strukturiert.

Der Stoffwechsel läßt sich in allen Bakterienzellen in Katabolismus und Anabolismus einteilen (Abb. 3.1). Substrate, Mineralien und evtl. weitere benötigte Substanzen aus der Umgebung gelangen durch Diffusion oder aktiven Transport in das Zellinnere. Aufgenommene organische Verbindungen wie Zucker oder organische Säuren gehen in den katabolen, also abbauenden Stoffwechsel ein. Dessen Aufgabe ist insbesondere die Zerlegung des Substrates in kleinere Moleküle mit 2-6 Kohlenstoffatomen. Die Abbauewege führen in den dicht verknüpften Zentral- oder Intermediärstoffwechsel (Amphibolismus), der die bekannten Haupttrouten Glykolyse (bzw. Gluconeogenese), Tricarbonsäure-Zyklus (TCA) und Pentosephosphat-Weg enthält und gewissermaßen katabole *und* anabole Aufgaben übernimmt. In [94] werden katabole und amphibole Reaktionen adäquater auch unter dem Begriff *fueling reactions* (Auffüllreaktionen) zusammengefaßt, da sie letztlich dem Auffüllen der Energie-, Redox- und Präkursoren-Pools dienen. Die (nur) 12 Präkursoren sind die zentralen Kohlenstoffgerüste und Ausgangspunkte anaboler Synthesereaktionen. Energie und Redoxkraft für die Biosynthese werden meistens in den universellen „Währungen“ ATP bzw. NADPH bereitgestellt. Eine geringe Menge von ATP kann über Substratphosphorylierung gewonnen werden, viel effizienter aber ist die Elektronentransportphosphorylierung (ETP), bei der protonenmotorische Kraft generiert und dann zur Phosphorylierung von ADP zu ATP genutzt wird. Eine Form der ETP ist aerobe oder anaerobe Respiration, bei der aeroben Variante ist Sauerstoff der finale Elektronenakzeptor. Die für die Respiration benötigten Elektronen werden gewöhnlich über den Cofaktor NADH bereitgestellt, der durch Oxidation eines Teils des aufgenommenen Substrates – vornehmlich im TCA – generiert wird. Die Photosynthese ist eine andere Form der ETP und dient ebenfalls der Bereitstellung von ATP (bei Cyanobakterien auch von Redoxkraft).

Im anabolen Stoffwechsel werden aus den bereitgestellten Präkursoren, sowie Energie- und Redoxäquivalenten in zumeist längeren, aber nur schwach verzweigten Reaktionssequenzen die ca. 75 „building blocks“ (Monomere) synthetisiert. Dazu gehören z.B. die Amino- und

Fettsäuren und die Nucleotide. Nebenprodukte wie ADP und NADP, aber auch NADH und Kohlenstoffverbindungen wie Fumarat oder Succinat, fließen in den Zentralstoffwechsel zurück. Aus den Monomeren werden, unter weiterem Verbrauch von Energie, in Polymerisationsreaktionen die Makromoleküle (Polymere) synthetisiert, die schließlich noch in die Biomasse integriert werden („Assembly“; z.B. Einbau von Proteinen in die Zellmembran). Die wichtigsten Makromoleküle sind Proteine, DNA, RNA, Lipide, Lipopolysaccharide, Murein (Peptidoglykan) und Speicherstoffe wie Glycogen oder Polyphosphat (Abb. 3.1). Mit einem Anteil von 96% bilden sie das Gros des Trockengewichts der Zelle. In [94] findet man eine systematische Aufstellung von Stoff- und Energiebilanzen bzgl. der Monomer- und Polymer-synthesen, die Grundlage vieler Stoff- und Flußbilanzen in metabolischen Netzwerken ist.

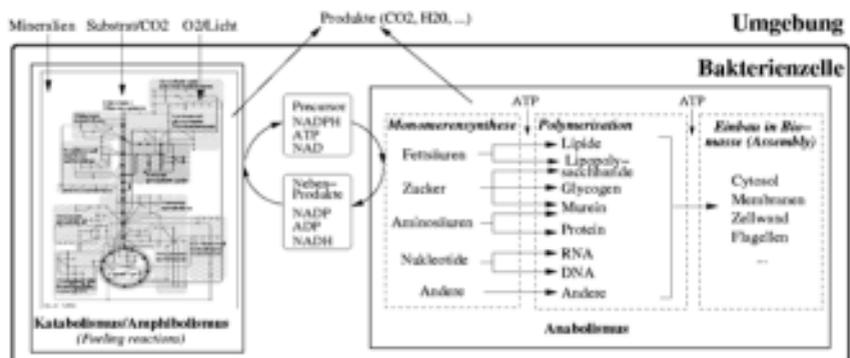


Abb. 3.1: Die Organisation des Stoffwechsels in Bakterienzellen.

Die anabolen Stoffwechselwege sind in den meisten Bakterien identisch, während eine große Diversität bzgl. der Bereitstellung von Energie, Redoxkraft und organischen Gerüstsubstanzen im katabolen/amphibolen Stoffwechsel vorhanden ist. Je nachdem, auf welcher Grundlage die drei genannten Komponenten generiert werden, lassen sich alle in der Natur vorkommenden Stoffwechselstrategien nach folgenden drei Hauptkriterien klassifizieren [83]:

- **Bereitstellung organischer Verbindungen**
 - *heterotroph*: aus organischen Verbindungen (Substraten), wie Glucose oder Succinat
 - *autotroph*: aus anorganischen Verbindungen (z.B. CO₂)
- **Bereitstellung von Energie (ATP)**
 - *chemotroph*: aus chemischer Energie (z.B. gespeichert in Substraten; hierzu gehören aerobe und anaerobe Respiration, sowie Gärungen)
 - *phototroph*: aus Lichtenergie mittels Photosynthese
- **Bereitstellung von Elektronen (Redoxkraft, NAD(P)H)**
 - *lithotroph*: aus anorganischen Verbindungen wie H₂O oder H₂
 - *organotroph*: aus organischen Verbindungen (Substraten), wie Glucose oder Succinat

Demnach ist der Stoffwechsel von Tieren und Menschen chemoorganoheterotroph (die organische Nahrung ist zugleich Kohlenstoff-, Energie- und Elektronenspender), der von Pflanzen gewöhnlich photolithoautotroph. Im Reich der Bakterien sind für die drei Kriterien jeweils beide Ausprägungen bekannt, jedoch nur selten, wie in der Familie der *Rhodospirillaceae* (Abschnitt 3.3), sind alle Varianten in einem Organismus realisierbar.

3.2 Netzmodell für Stoffwechsel in *Escherichia coli*

Das Darmbakterium *Escherichia coli* aus der Gruppe der *Enterobacteriaceae* ist der wohl wichtigste prokaryotische Modellorganismus überhaupt. Seit Jahrzehnten wird *E. coli* von Mikrobiologen und Genetikern studiert. Sowohl physiologisch als auch genetisch (das Genom ist vollständig sequenziert; ca. 4400 Gene) ist dieser Organismus deshalb gut charakterisiert. Daher wundert es nicht, daß *E. coli* nun auch in der Systembiologie häufig als Studienobjekt für *in silico* Analysen herangezogen wird [79,122], da die Wissens- und Datenbasis für mathematische Modelle relativ gut ist.

Aus diesem Grunde wurde in dieser Arbeit der Zentralstoffwechsel von *E. coli* als eines von zwei exemplarischen Metabolismen ausgewählt, anhand derer verschiedene Methoden der Strukturanalysen illustriert und angewendet werden sollen. Neben der gut charakterisierten Topologie des Stoffwechselnetzes von *E. coli* werden wir auch auf Genexpressionsmessungen und auf untersuchte Phänotypen von Mutanten zurückgreifen können, was eine Überprüfung von Prädiktionen ermöglicht.

E. coli kann eine breite Palette von Substraten, insbesondere Zuckern, metabolisieren. Nach den obigen drei Kriterien kann es ausschließlich chemoorganoheterotroph wachsen, allerdings, als fakultativ anaerobes Bakterium, sowohl mittels Gärung (ohne Sauerstoff) oder auch respirativ. Das hier betrachtete Netz umfaßt 110 Reaktionen und 89 Metabolite und ist auf den (katabolen und amphibolen) Zentralstoffwechsel fokussiert. Letzterer ist für Strukturanalysen besonders interessant, da im Vergleich zu den oft linearen anabolen Stoffwechselwegen eine viel stärkere Verknüpfung der beteiligten Metabolite vorliegt und daher kaum intuitive Aussagen möglich sind. Darum wurde der Zentralstoffwechsel detailliert modelliert, auch mit weniger geläufigen Reaktionen wie die von Pyruvatoxidase oder Aspartase. Anabole Reaktionssequenzen wurden mittels [94] zusammengestellt und sind größtenteils in kumulierten Stöchiometrien zusammengefaßt (insbesondere Enzyme Subsets, Abschnitt 2.4.1). Die Einbeziehung der anabolen Syntheserouten ist dennoch von Bedeutung: Sie repräsentieren den Abfluß von Monomeren in die Biomasse und erlauben somit, Wachstum quantitativ im Modell zu berücksichtigen.

Eine detaillierte Auflistung der betrachteten Metabolite und Reaktionen findet man in Anhang A.1 bzw. A.2. Wir werden uns auf die dort verwendeten Symbole beziehen. Eine graphische Darstellung des katabolen/amphibolen Stoffwechselnetzes zeigt die in Abb. 4.9 dargestellte

interaktive Flußkarte im *FluxAnalyzer*. (Die erfaßten anabolen Reaktionswege sind weitgehend mit denen in Abb. 4.8 identisch.) Berücksichtigte Stoffwechselfade sind:

- Haupttrouten: Glykolyse, TCA, oxidativer + nichtoxidativer Pentosephosphat-Weg
- 5 anaplerotische Reaktionen (Glyoxalat-Zyklus (2), PEP-Carboxykinase, PEP-Carboxylase, Malat-Enzym)
- Pyruvat-Oxidase, Aspartase, Transhydrogenase
- Respiration (via NADH-Dehydrogenase und Oxidase)
- 6 (Pseudo-) Reaktionen zur Aufnahme von 4 Substraten: Glucose (via PTS oder Hexokinase), Succinat, Acetat, Glycerin (via NAD- oder Ubichinon-abhängige Dehydrogenase)
- 5 (Pseudo-) Reaktionen zur Ausscheidung der 5 wichtigsten Produkte: Acetat, Lactat, Ethanol, Formiat, CO₂
- Aufnahme von Ammonium und Sulfat (für Biosynthese)
- Verbrauch von ATP für Erhaltungs- (Maintenance) und andere, nicht bilanzierte Prozesse
- kumulierte Biomassesynthesereaktion (μ) laut Biomassezusammensetzung

Die Reaktionsliste (Anhang A.2) berücksichtigt nur die als intern (balanciert) betrachteten Metabolite. Redundante Erhaltungsrelationen wurden von vornherein ausgeschlossen. Daher tauchen z.B. die komplementären Partner von ATP und NAD(P)H (ADP bzw. NAD(P)) im Modell nicht auf (Abschnitt 2.5). Die für μ verwendete Zusammensetzung der Biomasse – bestehend aus 7 berücksichtigten Makromolekülen – kann ebenfalls der Reaktionsliste entnommen werden. Moderate Verschiebungen in der Zellzusammensetzung, die z.B. durch eine veränderte Wachstumsrate auftreten können, haben hier keinen wesentlichen Einfluß auf die Ergebnisse der Strukturanalysen (nicht gezeigt).

Insgesamt 57 der 110 Reaktionen entsprechen realen (nicht zusammengefaßten) Reaktionen und können daher Genprodukten (Enzymen, Transportern) bzw. Enzymkomplexen zugeordnet werden (Anhang A). Isoenzyme mit identischem stöchiometrischem Umsatz von Metaboliten werden im Netzmodell mit nur *einer* gemeinsamen Reaktion berücksichtigt. Wir führen also keine parallelen Reaktionen für jedes Isoenzym ein, sondern aggregieren die jeweiligen Flüsse in einer Reaktion(rate). Wird eine Reaktion virtuell aus dem Netz entfernt („Mutante“), ist damit also die gleichzeitige Stilllegung der Aktivitäten all ihrer Isoenzyme gemeint.

3.3 Netzmodell für Stoffwechsel in schwefelfreien Purpurbakterien (*Rhodospirillaceae*)

Die Bakterien aus der Gruppe der *Rhodospirillaceae* (schwefelfreie Purpurbakterien; abgek. SPB) gehören bzgl. der realisierbaren Stoffwechselmodi zu den vielseitigsten Organismen überhaupt [90,155]. Nach der in Abschnitt 3.1 gegebenen Klassifikation können sie als fakultativ photosynthetische Bakterien sowohl phototroph als auch chemotroph wachsen. Chemotrophes Wachstum erfolgt vornehmlich organoheterotroph, entweder (aerob) mit Respiration über Sauerstoff bzw. (anaerob, ohne Licht) mittels Gärung oder anaerober Respiration

(mehrere Elektronenakzeptoren sind möglich). Steht anaerob Licht zur Verfügung erfolgt phototrophes Wachstum, welches je nach Bedingungen organoheterotroph (Kohlenstoff und Elektronen kommen hauptsächlich aus organischem Substrat) oder lithoautotroph (Fixierung von Kohlendioxid mittels Calvin-Zyklus; anorganische Substanzen wie Wasserstoff dienen als Elektronenspender) erfolgen kann [155]. Daneben beherrschen die SPB auch andere Stoffwechselwege wie die energieaufwendige Fixierung von Stickstoff mittels Nitrogenase.

Wir werden hier auf das für SPB typische (photo- oder chemo-) heterotrophe Wachstum fokussieren, wo bevorzugt organische Säuren wie Malat oder Succinat oder – mit meistens geringeren Wachstumsraten – auch diverse Zucker als Kohlenstoff- und Elektronenquelle verwendet werden. Sauerstoff ist das entscheidende Signal für die Regulation, ob der Stoffwechsel chemo- oder phototroph auszurichten ist (Kap. 9). Im aeroben Milieu ist chemoheterotrophes Wachstum mit Sauerstoff als Elektronenakzeptor der klar dominante Operationsmodus [90]. Beim Wechsel zu anaeroben Bedingungen erfolgt eine Umschaltung zum phototrophen Stoffwechsel, wobei folgende wesentliche Umstellungen notwendig sind:

- **Energiestoffwechsel:** Respirative Elektronentransportphosphorylierung wird durch eine photosynthetische ersetzt [90]. Für die Photosynthese müssen photosynthetische Membranen (Chromatophoren) und in ihnen Lichtsammelkomplexe (LH) und Reaktionszentren (RC) installiert werden. RC und LH sind große Proteinkomplexe mit eingebauten Pigmenten wie Bacteriochlorophyll und Carotinoiden.
- **Kohlenstoff- und Redoxstoffwechsel:** Da nun Sauerstoff nicht mehr als finaler Elektronenakzeptor zur Verfügung steht, werden viel weniger Reduktionsäquivalente (insbes. NADH) benötigt, zum Teil muß sogar überschüssige Reduktionskraft beseitigt werden. Dazu werden bestimmte Stoffwechselwege neu installiert (z.B. Calvin-Zyklus, evtl. reduktiver TCA) bzw. abgeschaltet (z.B. oxidativer TCA) [155].

Bezüglich der Photosynthese gibt es drei entscheidende Unterschiede zwischen SPB auf der einen und Pflanzen und Cyanobakterien auf der anderen Seite [134]: SPB haben nur ein Photosystem mit dem sie nicht in der Lage sind, Wasser als Elektronendonator zu verwenden und dabei zu spalten. Deshalb erfolgt die Photosynthese (i) anoxygen, d.h. ohne Sauerstofffreisetzung, (ii) in einem zyklischen Elektronenfluß, in dem (iii) keine Bereitstellung von Reduktionsäquivalenten wie NAD(P)H erfolgt. Elektronen können allerdings (z.B. auf der Stufe von Succinat) in die Elektronentransportkette eingebracht und dann energetisch bergauf über NADH wieder ausgeschleust werden („reverse electron flow“).

Aufgrund der besonderen Bedeutung widmet sich Kap. 9 der detaillierten Beschreibung und dynamischen Modellierung der oxidativen bzw. photosynthetischen Elektronentransportphosphorylierung der SPB, was über eine rein strukturelle Betrachtung der Stoff- und Elektronenflüsse hinausgehen wird. Dabei werden wir insbesondere die Regulation der Expression photosynthetischer Komponenten studieren, die u.a. durch Aktivitäten und Redoxzustände von Komponenten in der Elektronentransportkette vermittelt wird [103].

Zuvor werden wir den Zentralstoffwechsel in diesen Bakterien untersuchen, in dem die Elektronentransportkette nur einen Bestandteil von vielen darstellt. Der Stoffwechsel der SPB ist längst nicht so gut aufgeklärt und verstanden wie z.B. in *E. coli*. Gerade deshalb sollte eine metabolische Strukturanalyse zu einem besseren Gesamtverständnis der Operationsweisen des Stoffwechsels der SPB beitragen. Insbesondere wollen wir der Frage nachgehen, wie der Redoxstoffwechsel durch Umverteilungen von Flüssen den jeweiligen Lebensbedingungen angepaßt und damit die Balance der Reduktionsäquivalente NAD(P)H sichergestellt werden kann. Die Anpassungsfähigkeit des Redoxstoffwechsels ist gerade eine wesentliche Voraussetzung für die metabolische Vielseitigkeit der SPB [155].

Beim Zusammenstellen des Stoffwechselnetzmodells wurde auf Publikationen der drei wichtigsten Vertreter der *Rhodospirillaceae* zurückgegriffen, nämlich *Rhodobacter capsulatus*, *Rhodobacter sphaeroides* und *Rhodospirillum rubrum*. Letztere Spezies war Gegenstand vieler metabolischer Studien, während die beiden ersten hinsichtlich der redoxgesteuerten Regulation der Synthese photosynthetischer Komponenten eingehend untersucht worden sind [103].

Insgesamt besteht das konstruierte Netzmodell für die *Rhodospirillaceae* aus 121 Reaktionen und 97 Metaboliten (Anhang A.1 und A.3). Der Bereich des Zentralstoffwechsels ist in Abb. 4.7 zu sehen und sollte die wichtigsten identifizierten Wege der SPB enthalten (jede berücksichtigte Reaktion wurde in zumindest einer der drei Arten nachgewiesen):

- Glykolyse, oxidativer TCA, oxidativer + nichtoxidativer Pentosephosphat-Weg
- 3 „typische“ anaplerotische Reaktionen (PEP-Carboxykinase, Malat-Enzym, Pyruvat-Carboxylase)
- Citramalat-Zyklus: *R. sphaeroides* und *R. rubrum* besitzen keinen vollständigen Glyoxalat-Zyklus, da ihnen das Enzym Isocitrat-Lyase fehlt. In [7,64] wurde ein alternativer Citramalat-Zyklus vorgeschlagen, der das beobachtete Wachstum auf Acetat erklären kann. Dieser hypothetische Pathway wurde im Netz berücksichtigt.
- Transhydrogenase, Fumarat-Reductase
- Respiration (über NADH-Dehydrogenase, bc₁-Komplex und Cytochrom-Oxidase)
- zyklische Photosynthese (via bc₁-Komplex und Reaktionszentrum)
- 5 repräsentative Substrate: Fructose (Aufnahme via PTS), Succinat, Acetat, Malat, Propionat
- Ausscheidung der 3 wichtigsten Produkte: Acetat, Formiat, Succinat
- CO₂ kann sowohl aufgenommen (dann Nettofixierung) als auch ausgeschieden werden.
- Hydrogenase: Aufnahme von Wasserstoff für lithoautotrophes Wachstum (mit CO₂)
- Calvin-Zyklus (reduktiver Pentosephosphat-Weg): vier zusätzliche Reaktionen ergänzen den Pentosephosphat-Weg und einen Teil der Glykolyse zum Calvin-Zyklus. Das Enzym Rubisco (Ribulosebiphosphat-Carboxylase/Oxygenase) katalysiert dabei den CO₂-fixierenden Schritt.

- 3 zusätzliche Reaktionen ergänzen den oxidativen TCA zum rückwärts laufenden und CO₂-fixierenden *reduktiven TCA* [11]. Sie werden katalysiert durch Pyruvat- und α -Ketoglutarat-Synthase, sowie Citrat-Lyase. Die ersten beiden Reaktionen benötigen reduziertes Ferredoxin, das hier über NADH:Ferredoxin-Oxidoreductase bereitgestellt wird.
- Die Nitrogenase ermöglicht Fixierung von Stickstoff. Sie ist deaktiviert unter aeroben Bedingungen oder bei „starker“ Stickstoffquelle wie Ammonium.
- Aufnahme von Ammonium und Sulfat (Stickstoff/Schwefel für Biosynthese)
- Verbrauch von ATP für Erhaltungs- (Maintenance) und andere, nicht bilanzierte Prozesse
- kumulierte Biomassesynthesereaktion (μ) laut Biomassezusammensetzung

Mittlerweile sind die Genome von *R. sphaeroides* und *R. rubrum* sequenziert. Sequenzen und vorläufige Ergebnisse der Annotation sind auf der Internetseite www.jgi.doe.gov des Joint Genome Institutes (JGI) zugänglich. Dort kann man sich auch die über die Annotation rekonstruierten Stoffwechselwege anzeigen. Die meisten der oben angegebenen Routen, die ursprünglich aus biochemischen Analysen abgeleitet wurden, konnten auch im Genom verankert wiedergefunden werden. Ausnahmen bilden einige Enzyme des Citramalat-Zyklus, bei denen unklar bleibt, ob sie tatsächlich nicht vorhanden sind oder unerkannt blieben. Einige dieser Enzyme sind offensichtlich noch nicht einmal benannt (sie konnten weder in der Literatur noch in Datenbanken gefunden werden), geschweige denn in ihrer Sequenz bekannt. Zwei sehr spezifische Ausprägungen (Glucose-6-Phosphat-Dehydrogenase fehlt in *R. rubrum*; Pyruvat- und α -Ketoglutarat-Synthase ist nur in *R. rubrum* vorhanden), die sowohl biochemisch als auch genetisch belegt sind, werden wir in Kap. 5 noch genauer diskutieren.

Anabole Stoffwechselwege wurden – wie bei *E. coli* – vornehmlich als zusammengefaßte Reaktionssequenzen aufgenommen. Die Stöchiometrien der Monomersynthese wurden weitgehend von *E. coli* übernommen (Ausnahme: Fettsäuren nach [62]), ebenso die Monomerzusammensetzung der Makromoleküle mit Ausnahme der DNA (63% GTP/CTP-Anteil [63]). Zusätzlich zu den 7 Makromolekülen in *E. coli* wurden noch die häufig auftretenden Speicherstoffe Polyphosphat (ca. 2% der Trockenmasse) und PHB (Polyhydroxybutyrat, ca. 5% [41]) berücksichtigt, sowie die u.U. beträchtliche Menge an Bacteriochlorophyll unter phototrophen Bedingungen (ca. 1%; gemittelt nach [49]). Durch Bildung photosynthetischer Membranen ist der Lipidanteil in der Biomasse unter phototrophen Bedingungen etwa 50-70% höher als unter aeroben Bedingungen [155] und beträgt damit etwa 16%. Die prozentualen Anteile der anderen Makromoleküle in der Biomasse (RNA, DNA, Protein, Lipopolysaccharide, Peptidoglycane, Glycogen) wurden in etwa im gleichen Verhältnis wie bei *E. coli* [94] festgelegt, sie sind jedoch aufgrund der hinzugekommenen Komponenten insgesamt reduziert. Abb. 4.8 zeigt die kompakte Darstellung des Anabolismus der *Rhodospirillaceae* im *FluxAnalyzer*.

4 FluxAnalyzer: Computergestützte Analyse von Stoffwechselnetzen

4.1 Motivation

Systembiologie – die synergistische Kombination von Theorie, Computersimulation und Experiment – erfordert per Definition den Einsatz von Rechnern. Längst zielt dabei die Entwicklung und der Einsatz von Modellierungs- und Simulationswerkzeugen nicht nur auf das „bloße Rechnen“ ab. Neben der numerischen Behandlung zellulärer Modelle hat die Bereitstellung benutzerfreundlicher Oberflächen einen enorm wichtigen Stellenwert erlangt, soll doch die *in silico* Analyse virtueller Zellen (oder Zellkomponenten) nicht nur dem erfahrenen Modellierer, sondern auch dem Biologen möglich sein. Ein weiterer wichtiger Punkt ist die Integration verschiedener Methoden und heterogener Daten. Aufgrund der vielfältigen Modellierungsmethoden (deterministisch vs. probabilistisch, stationär vs. dynamisch, inverse vs. Vorwärtsmodellierung, datengetrieben vs. wissensbasiert, quantitativ vs. qualitativ) erscheint es zur Zeit unmöglich, eine Plattform für alle denkbar möglichen Modelle der Systembiologie anzubieten. Folgerichtig existiert daher eine breite Vielfalt an Rechnerwerkzeugen, denen mit der neugeschaffenen *Systems Biology Markup Language* (SBML, [59]) die Möglichkeit des Austauschs von Modelldefinitionen zur Verfügung steht.

Verbreitete Rechnerwerkzeuge zur Modellierung metabolischer Systeme sind ProMol/Diva [48,80], GEPASI [91], JARNAC [129], E-Cell [157], Virtual Cell [85], StochSim [81], ScrumPy [117] und METATOOL [116]. Mit Ausnahme von METATOOL und ScrumPy sind Methoden zur Strukturanalyse in Stoffwechselnetzen in diesen Programmen kaum berücksichtigt (GEPASI und JARNAC nutzen einige Routinen von METATOOL). METATOOL wurde mittlerweile in verschiedenen Anwendungen genutzt und widmet sich insbesondere der Bestimmung von Elementarmoden und konvexen Basen, sowie einiger anderer topologischer Merkmale, wie Enzyme Subsets, Erhaltungsrelationen und Kernmatrix [116]. Ähnliches leistet ScrumPy. Methoden zur metabolen Flußanalyse, zur Berechnung optimaler Flußverteilungen und weitergehende Funktionalitäten zur Auswertung der Elementarmoden sind noch nicht enthalten. Die Eingabe des Netzmodells, als auch die Ausgabe berechneter Ergebnisse erfolgt über ASCII-Dateien (bzw. in ScrumPy auf Kommandozeile). Damit gibt es bisher praktisch keine Modellierungssoftware, die die unterschiedlichen Methoden zur Analyse stationärer Reaktionsnetzwerke in einer benutzerfreundlichen und interaktiven Umgebung integriert.

Um diese Lücke zu schließen wurde der *FluxAnalyzer* entwickelt. Zielsetzung war und ist es, eine möglichst vollständige Bibliothek an Funktionen für die Strukturanalyse metaboler Netze bereitzustellen, die in einer interaktiven Umgebung mit graphischen Visualisierungen und ansprechendem Bedienkomfort zur Verfügung steht.

4.2 Konzeption des *FluxAnalyzers*: Netzwerk-Projekte und interaktive Flusskarten

Der *FluxAnalyzer* ist ein Paket (Sammlung ausführbarer Dateien) für das kommerzielle Programm MATLAB® (Mathworks, Inc). MATLAB ist eine Entwicklungsumgebung, die numerische Berechnungen, Graphiken und Visualisierungen, sowie eine intuitive Programmiersprache für Anwendungen der Ingenieur- und Naturwissenschaften umfaßt. Trotz des kommerziellen Status ist MATLAB sowohl an akademischen Instituten als auch in der Industrie weit verbreitet. Der große Vorteil, MATLAB als Umgebung des *FluxAnalyzers* zu verwenden, liegt zum einen in den mächtigen Möglichkeiten, Benutzeroberflächen schnell und flexibel zusammenzubauen, und zum anderen in der Fülle verfügbarer numerischer Funktionen, die bei der Programmierung eigener Routinen leicht mit eingebunden werden können.

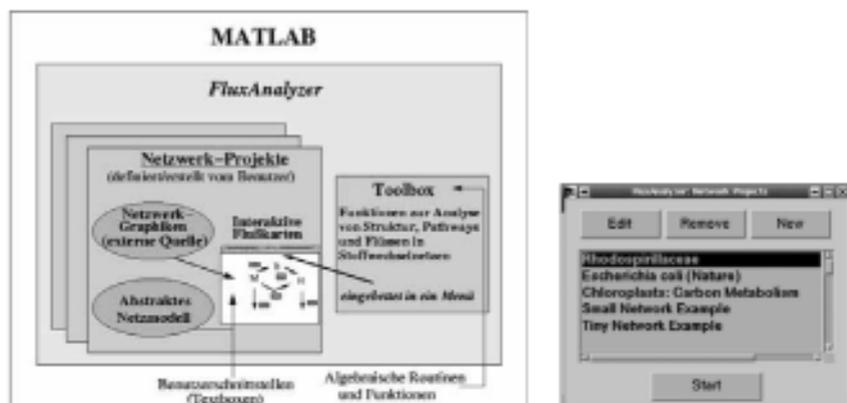


Abb. 4.1 (links): Konzeptioneller Aufbau des *FluxAnalyzers*

Abb. 4.2 (rechts): *Network Project Manager*

Der logische Aufbau des *FluxAnalyzers* ist in Abb. 4.1 zu sehen und untergliedert sich in zwei Hauptteile: (i) benutzerdefinierte Netzwerk-Projekte, die mit einem *Network Project Manager* verwaltet werden (Abb. 4.2) und (ii) eine umfangreiche Toolbox mit Funktionen für die Strukturanalyse. Jedes Netzwerk-Projekt besteht wiederum aus zwei Komponenten, die vom Benutzer erstellt bzw. zur Verfügung gestellt werden müssen:

- ein abstraktes Modell der Netzstruktur
- eine (oder mehrere) graphische Visualisierung(en) des Netzes („Netzarten“)

Die symbolische Definition des Netzmodells wird im *FluxAnalyzer* durch einen *Network Composer* unterstützt (s.u.). Die graphische Repräsentation des Netzes muß dagegen direkt vom Benutzer bereitgestellt werden. Es gibt mehrere Gründe, warum auf ein automatisches Erstellen der Netzgraphik (z.B. aus dem Netzmodell) verzichtet wurde. Algorithmen zum automatischen Zeichnen von Hypergraphen auf der Basis einer Netzbeschreibung sind komplex

und führen, insbesondere bei größeren Stoffwechselnetzen, nur selten zu einer Darstellung, die den Vorstellungen des Anwenders entspricht. Deshalb bietet der *FluxAnalyzer* dem Benutzer die Möglichkeit, die Netzgraphiken selbst zu erstellen und dann in ein Netzwerk-Projekt zum Zwecke der Visualisierung einzubinden. Diese Herangehensweise läßt insbesondere die Optionen zu, das gesamte Netz auf mehrere Graphiken verteilt darzustellen und beliebige Anmerkungen, Beschriftungen oder Hervorhebungen in die Graphik aufzunehmen. Alternativ kann auch auf andere Quellen zurückgegriffen werden. So ist das Einbinden metaboler Karten z.B. von BioPath [biopath.fmi.uni-passau.de] oder KEGG [68] möglich. Nach dem Starten eines Netzwerk-Projektes im *FluxAnalyzer* wird jede Netzgraphik als Hintergrund in ein separates MATLAB-Fenster geladen. Die Verknüpfung der Netzkarten mit dem eigentlichen Netzmodell erfolgt über kleine Benutzerschnittstellen (Textboxen), die jeweils einem Netzelement (z.B. einer Reaktion) zugeordnet sind und auf der Netzkarte (z.B. am entsprechenden Reaktionspfeil) plaziert werden können. Sie ermöglichen die Eingabe von Daten vom Benutzer bzw. die Ausgabe berechneter Resultate in der unterlegten Visualisierung. Diese Kombination von Netzmodell, Netzkarte und Interaktion bezeichnen wir hier als *interaktive Flußkarte*, da in den meisten Anwendungen (metabole) Flüsse die Ein- und Ausgabegrößen in der Netzgraphik sind.

Die Funktionen der Toolbox stehen jedem Netzwerk-Projekt nach Laden der interaktiven Flußkarten zur Verfügung und können vom Benutzer bequem von einem Menü aus gestartet werden (Abb. 4.7 und 4.9: Menüeintrag „FluxAnalyzer“).

Die in den Abschnitten 3.2 bzw. 3.3 vorgestellten Stoffwechselnetzmodelle wurden im *FluxAnalyzer* implementiert. Abb. 4.7 und 4.8 zeigen die zwei interaktiven Flußkarten (aufgeteilt in Zentralstoffwechsel und Anabolismus), die zum Netzwerk-Projekt der *Rhodospirillaceae* gehören. Das Netzwerk-Projekt von *E. coli* wird im *FluxAnalyzer* ebenfalls über eine katabole (Abb. 4.9) und eine anabole Flußkarte (ähnlich der in Abb. 4.8) repräsentiert.

4.3 Konstruktion eines Netzwerk-Projektes

Nach der Initialisierung eines neuen Netzwerk-Projektes im *Project Manager*, bei der die einzubindenden Netzkarten und allgemeine Parameter wie Größe, Schrift und Farben der Textboxen definiert werden, kann der Benutzer mit dem Zusammenbau seines Netzes beginnen. Um der globalen Organisation des zellulären Stoffwechsels Rechnung zu tragen, stehen dafür im *FluxAnalyzer* vier verschiedene Typen von Netzelementen zur Verfügung, mit denen beliebige (insbesondere metabole) Reaktionsnetzwerke konstruiert werden können:

- *Metabolite*: sind die freien, niedrigmolekularen Spezies der Zelle, die im Katabolismus, Amphibolismus oder bei der Monomersynthese beteiligt sind. Dazu gehören z.B. ATP, Pyruvat oder Aminosäuren wie Tryptophan. Für jeden internen Metaboliten (s.u.) wird in der stöchiometrischen Matrix eine Zeile reserviert.

- *Reaktionen*: sind die biochemischen Reaktionen (Stoffwandler). Jede Reaktion belegt mit ihrer spezifischen Stöchiometrie eine Spalte in **N**.
- *Makromoleküle*: sind die wesentlichen makromolekularen Bestandteile der Biomasse, wie z.B. Proteine oder RNA.
- *Assembly-Routen*: repräsentieren den Abfluß (Einbau) eines Metaboliten in eine Biomassekomponente (vgl. Abb. 3.1).

Elementtyp	Metabolit	Reaktion	Makromolekül	Assembly-Route
Attribute	<ul style="list-style-type: none"> • Name • Symbol • External-Flag 	<ul style="list-style-type: none"> • symbolischer Name • Reaktionsgleichung • Default-Wert • min. Reaktionsrate • max. Reaktionsrate • Koeffizient in linearer Zielfunktion • Varianz von Messungen • Textbox-Parameter 	<ul style="list-style-type: none"> • Name • Symbol • Default-Konzentration • kumulative Synthesgleichung • Textbox-Parameter 	<ul style="list-style-type: none"> • Name • beteiligter Metabolit • beteiligtes Makromolekül • Textbox-Parameter
Textbox	keine	Reaktionsrate [nmol(gTM ⁻¹ h)]	relativer Biomasseanteil [g/gTM]	Abfluß eines Metaboliten in Synthese eines Makromoleküls [nmol(gTM ⁻¹ h)]

Tab. 4.1: Typen von Netzwerkelementen im FluxAnalyzer und ihre Eigenschaften

Die Konstruktionszentrale im FluxAnalyzer ist der *Network Composer*, mit dem neue Instanzen eines Elementtyps deklariert oder bestehende modifiziert oder entfernt werden können (Abb. 4.3). Jeder Elementtyp besitzt eine Menge von Attributen (Tab. 4.1), die bei der Deklaration eines neuen Elements dieses Typs zu definieren sind. Dazu stehen spezifische Eingabemasken zur Verfügung. Abb. 4.4 zeigt die Maske zur Deklaration eines neuen Elements des Typs *Reaktion*. Zu den Attributen einer *Reaktion* gehört z.B. eine symbolische Reaktionsgleichung, aus der dann intern die stöchiometrische Matrix **N** generiert werden kann. Auf andere Parameter wird in bestimmten Algorithmen zurückgegriffen, z.B. legt die *minimale Reaktionsrate* u.a. auch fest, ob die Reaktion reversibel ist.

Mit Ausnahme der *Metabolite* ist jedem Netzelement genau eine Textbox zugeordnet, die an einer geeigneten Stelle auf einer Flußkarte positioniert werden kann. Die Inhalte der Box beziehen sich dann immer auf das korrespondierende Netzelement und stehen z.B. für Reaktionsraten (Typ *Reaktionen*) oder für den relativen Anteil an der Biomasse (Typ *Makromoleküle*) oder für Abflüsse in die Biomasse (Typ *Assembly-Routen*). Auf Textboxen für *Metabolite* wurde konzeptionell verzichtet. Die meisten stöchiometrischen Analysen, insbesondere solche, die auf der Quasistationaritätsannahme beruhen (Abschnitt 2.4), implizieren Ein- bzw. Ausgaben von *Flüssen* (Reaktionsraten). Textboxen für *Metabolite* wären nur in einigen we-

nigen Anwendungen (z.B. Erhaltungsrelationen) nützlich, würden aber die Flußkarten erheblich unübersichtlicher gestalten. Entsprechende Ausgaben für Metabolite erfolgen dann im Kommandofenster von MATLAB.

Die gewünschte Position einer Textbox auf den interaktiven Flußkarten kann vom Benutzer bequem über ein Fadenkreuz definiert werden („Get x/y-Pos“ in Abb. 4.4). Neben der Position kann auch die Editierbarkeit festgelegt werden. In einigen Fällen kann es durchaus sinnvoll sein, Textboxen nur als Ausgabeschnittstelle fungieren zu lassen.

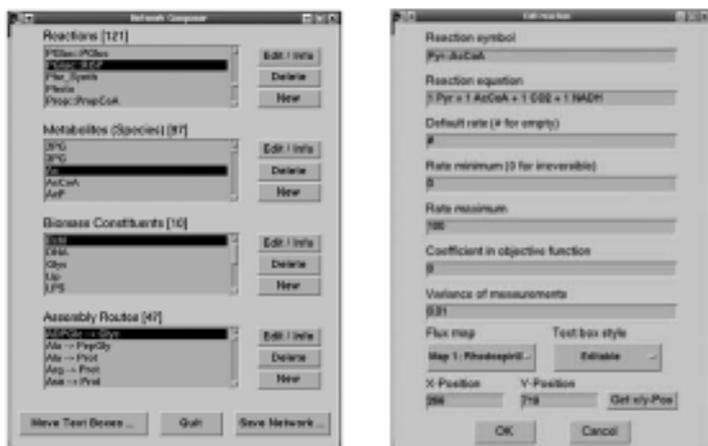


Abb. 4.3 (links): Network Composer im FluxAnalyzer

Abb. 4.4 (rechts): Eingabemaske zur Deklaration einer Reaktion

Eigentlich erlauben die beiden Elementtypen *Reaktionen* und *Metabolite* den Zusammenbau beliebiger Netze. Die Elementtypen *Makromoleküle (MME)* und *Assembly-Routen* wurden zusätzlich bereitgestellt, um Biomassessynthese auf der Basis einer vorgegebenen Zellzusammensetzung adäquat berücksichtigen zu können (vgl. Abb. 3.1). Der kumulierte Abfluß von Metaboliten in die Biomasse wird im FluxAnalyzer über eine „Reaktion“ *mse* identifiziert, deren Rate gerade die Wachstumsrate μ ist. Die Stöchiometrie von *mse*, d.h. der Bedarf an Metaboliten zur Produktion eines Gramms Trockenbiomasse (gTM), kann aufgrund einer variablen Zellzusammensetzung auch für dieselbe Spezies variieren. Eine feste Stöchiometrie wäre deshalb nur wenig flexibel und müßte bei Veränderungen in der Zusammensetzung der Biomasse jedes Mal mühsam neu zusammengestellt werden. *MME* erleichtern diesen Schritt: Elemente diesen Typs sind Biomassebestandteile, für die eine (symbolische) Synthesegleichung definiert werden kann (Tab. 4.1), die den Bedarf an Monomeren, Energie, Reduktionskraft usw. für dessen Zusammenbau angibt ([$\mu\text{mol Metabolit} / \text{g MM}$]). Intern werden diese Stöchiometrien in einer Matrix Γ gespeichert (Zeilen: *Metabolite*; Spalten: *MME*). Auf den interaktiven Flußkarten kann der Benutzer in den Textboxen der *MME* den jeweiligen Anteil an der Biomasse eingeben ([$\text{g MM} / \text{g TM}$]). Abb. 4.8 zeigt z.B., wie die Zellzusammensetzung

für die *Rhodospirillaceae* eingegeben werden kann. Wird eine Berechnung im *FluxAnalyzer* gestartet, werden im Vektor **ma** zunächst die aktuell definierten Konzentrationen der *MMe* abgelegt. Das Produkt $\Gamma \cdot \mathbf{ma}$ ergibt dann die kumulierte Stöchiometrie der Biomassessynthese, die in der entsprechenden Spalte von **mue** in **N** eingetragen wird. Als einzige Spalte in **N** weist **mue** dann die Einheit [nmol Metabolit / gTM] auf. Dafür hat entsprechend die zugeordnete Wachstumsrate μ im Reaktionsratenvektor **r** als einzige Rate die Einheit h^{-1} , so daß das Produkt $\mathbf{N} \cdot \mathbf{r}$ für alle Zeilen die gleiche Einheit [nmol Metabolit / gTM · h] ergibt (Gl. (2.1)). Letztlich erlaubt diese Vorgehensweise bei Bedarf auch die durch Wachstum erzeugten *Verdünnungsterme* von Metaboliten zu berücksichtigen, indem die Konzentrationen der Metabolite (die i.a. aber sehr klein sind) ebenfalls als Bestandteile der Biomasse aufgefaßt werden.

Assembly-Routen sind keine „echten“ Strukturelemente, sondern ermöglichen nur die quantitative Ausgabe des Abflusses (Einbaus) eines Metaboliten *i* in eine Biomassekomponente *j* in den Flußkarten. Dieser Abfluß berechnet sich nach obigen Notationen aus dem (skalaren) Produkt $\Gamma_{ij} \cdot \mathbf{ma}_j \cdot \mu$ und entspricht genau einem Pfeil in Abb. 3.1, der von einem Monomer in die Polymerisations- und nachfolgenden Assemblyprozesse eines *MMs* führt. *Assembly-Routen* sind in Abb. 4.8 daran zu erkennen, daß ihre Textboxen keinen Rahmen haben. Die Aminosäure Alanin (Symbol *Ala*, Mitte rechts) ist z.B. an zwei *Assembly-Routen* beteiligt, da sie in die Synthese der Peptidoglykane und der Proteine eingeht. Die Summe beider Abflüsse in diesen Routen muß entsprechend vom Netzwerk generiert werden (hier: aus Pyruvat in Reaktion *Ala_Synth*, Anhang A.3). Es sei nochmals darauf hingewiesen, daß sich die Raten der *Assembly-Routen* implizit aus der Biomassezusammensetzung und der Wachstumsrate μ ergeben. Deshalb sind *Assembly-Routen* intern nicht in der Matrix **N** und ihre Raten nicht im Vektor **r** enthalten und ihre assoziierten Textboxen als nicht editierbar konfiguriert.

Wo die Grenze zwischen dem modellierten Stoffwechsellnetz und den Abflußkanälen in die Biomasse gezogen wird, hängt letztlich von der gewünschten Feinkörnigkeit des Netzmodells ab. In vielen Stoffwechsellmodellen werden bereits die Präkursoren stöchiometrisch in die Biomassessynthesgleichung aufgenommen, so daß die Synthese von Monomeren dann nur implizit darin enthalten ist. Der *FluxAnalyzer* unterstützt beide Varianten und natürlich auch den Fall, wo Biomassessynthese überhaupt nicht berücksichtigt werden soll.

Die Netzstruktur mit den deklarierten Elementen und deren Attributen kann jederzeit gespeichert werden. Ein manuelles Editieren der erzeugten ASCII-Dateien ist ebenfalls möglich.

4.4 Toolbox: Struktur-, Pathway- und Flußanalyse in Stoffwechsellnetzen

Hat der Benutzer sein Netzwerk-Projekt erstellt, kann er sämtliche Funktionen der Toolbox über ein entsprechendes Pulldown-Menü (Abb. 4.5) abrufen und damit das Netzmodell analysieren. Menü und interaktive Flußkarten ermöglichen auch Nutzern ohne detaillierte Kenntnisse der zugrundeliegenden Theorie, die Funktionalitäten anzuwenden. Nach Aufruf einer

Funktion startet der *FluxAnalyzer* gewöhnlich mit dem Auslesen benutzerdefinierter Werte aus den Textboxen. Nach den Berechnungen werden die Resultate in den Flußkarten und z.T. in separaten Fenstern ausgegeben. Zur besseren Illustration und Visualisierung werden unterschiedliche Farben für die Textboxen eingesetzt (in Abb. 4.7–4.9: Graustufen).

Die implementierten Funktionen (oder Skripte) wurden mit der MATLAB-Programmiersprache erstellt und sind entsprechend in MATLAB ausführbar. Insbesondere bei Matrizenoperationen (z.B. Berechnung einer Pseudoinversen) konnte auf in MATLAB zur Verfügung gestellte Funktionen zurückgegriffen werden. Einige spezielle Routinen, bspw. zur Berechnung von Elementarmoden, wurden dagegen vollständig selbst implementiert.



Abb. 4.5: Menüfunktionen im *FluxAnalyzer*

Abb. 4.5 zeigt das *FluxAnalyzer*-Menü mit seinen (nach Themenbereichen geordneten) Funktionen, die kurz besprochen werden sollen. Auf einige Funktionalitäten gehen wir in späteren Kapiteln, wo die entsprechenden theoretischen Methoden ausgearbeitet werden, genauer ein.

Bearbeiten/Ansehen der Netzstruktur

Hierzu gehört insbesondere der oben beschriebene *Network Composer* (Abb. 4.3), die zentrale Konstruktionsstelle des Netzmodells. *Constraints* (min. und max. Rate einer Reaktion) können alternativ (und schneller) auch direkt über die Textboxen der Reaktionen editiert werden.

Laden/Speichern/Verwalten von Szenarien

Ein Szenario ergibt sich aus den Werten, die vom Benutzer in den Textboxen eingetragen worden sind. Ein Szenario kann z.B. eine spezifische Konstellation für die Flußanalyse darstellen, die durch den Satz gemessener Reaktionsraten, sowie durch die vorgegebene Biomass-

sezusammensetzung definiert ist. Die Funktionen dieser Menügruppe ermöglichen das Laden, Setzen, Speichern oder Löschen eines Szenarios.

Flux Clipboard

Eine recht nützliche Einrichtung ist das *Flux Clipboard*, das ein (temporäres) Zwischenspeichern aktueller Resultate – bspw. einer Flußverteilung – ermöglicht. Die Zwischenergebnisse können später wiederverwendet, d.h. in die Textboxen zurückgelesen werden. Noch wertvoller ist die Möglichkeit, die aktuelle Flußverteilung mit der im Flux Clipboard abgelegten „arithmetisch“ zu kombinieren. So können z.B. Differenzen oder Quotienten zweier Flußverteilungen berechnet und in den Flußkarten ausgegeben werden. Damit lassen sich physiologisch unterschiedliche Zustände der Zelle effizient miteinander vergleichen.

Topologische/Strukturelle Analysen

Dieser Block stellt einen Großteil der in Kapitel 2.3-2.5 beschriebenen Methoden zur Verfügung. Die Bestimmung *einfacher Struktureigenschaften* ist insbesondere nach dem Editieren der Netzstruktur nützlich, da sich damit (ungewollte) strukturelle Inkonsistenzen schnell ausfindig machen lassen. Dazu gehören z.B. die Bestimmung von *strikt balancierten Raten*, sowie von („Sackgassen-“)Metaboliten, die nur produziert oder nur konsumiert werden können (Abschnitt 2.4.1). Desweiteren detektiert diese Funktion die *Enzyme Subsets* und *Isoenzyme*. Auftretende Erhaltungrelationen können in Form der *elementaren Erhaltungrelationen* (Abschnitt 2.5) bestimmt werden.

Die *stöchiometrische Matrix* kann in übersichtlicher graphischer Form ausgegeben werden, inkl. der Konnektivität der Metabolite (siehe Beispiel in [77]). Graphentheoretische Eigenschaften können anhand eines *Konnektivitätshistogramms*, sowie über die Berechnung von (*kürzesten*) *Pfadlängen* zwischen den Metaboliten studiert werden (Abschnitt 2.3). Dabei kann zwischen gerichtetem (Reversibilitäten berücksichtigt) und ungerichtetem Substrat-Graphen-Modell gewählt werden. Die Resultate können als Abstandsmatrix graphisch inspiziert oder in ASCII-Format exportiert werden. Mittlere Pfadlänge und maximaler Abstand werden ebenfalls ermittelt.

Die umfangreiche Sammlung an Algorithmen und Funktionalitäten für die *Elementarmoden- und Pathwayanalyse* nimmt einen ganz zentralen Platz im *FluxAnalyzer* ein. Vor der Berechnung der EMen können vom Benutzer Spezifikationen angegeben werden (Abb. 4.6 links):

- Ausschluß bestimmter Reaktionen (Betrachtung eines Teilnetzes)
- alle Reaktionen können als reversibel betrachtet werden (siehe Abschnitt 5.2)
- Berechnung einer konvexen Basis statt (aller) EMen; ermöglicht auch die Berechnung der extremen Pathways (Abschnitt 2.4.4)
- Isoenzyme werden auf Wunsch nur einmal im Netz berücksichtigt
- Verwendung der (schnellen) MEX-Version des Algorithmus

Die Berechnung von EMen ist insbesondere in größeren Netzen aufgrund der kombinatorischen Komplexität eine diffizile Aufgabe (Kap. 8). Die erste, vierte und fünfte der obigen Optionen sind deshalb dazu gedacht, die Komplexität des Netzes u.U. zu reduzieren bzw. eine besonders effiziente (MEX-)Version des Algorithmus zu verwenden (s.a.). Die Prozedur im *FluxAnalyzer* orientiert sich an dem iterativen Schema nach Schuster et al. [141] (ein Tableau wird mit Operationen ähnlich derer im Gaußschen Algorithmus bearbeitet). Speicherverwaltung und numerische Operationen wurden so effizient wie möglich programmiert. Der Algorithmus startet zunächst mit einer Vorverarbeitung, bei der das betrachtete Netz auf ein kleinstmögliches abgebildet wird, das eine äquivalente Menge von EMen aufweist (nach der Berechnung erfolgt die Rücktransformation). Dazu werden Reaktionen eines Enzym Subsets zu einer Gesamtreaktion zusammengefaßt (siehe [141] und Abschnitt 2.4.1), außerdem werden strikt balancierte Reaktionen (Abschnitt 4.2.1) und nicht partizipierende Metabolite aus dem Netz entfernt. Die Metabolite werden im Tableau aufsteigend nach ihrer Konnektivität sortiert, was sich erfahrungsgemäß positiv auf die Rechenzeit auswirkt.

Danach wird der eigentliche Algorithmus gestartet, der in endlich vielen Iterationen die EMen liefert. Als zeitintensivster Schritt überhaupt erweisen sich die Tests auf Unzerlegbarkeit der vorläufigen EMen (Eigenschaft (B3) in (2.13)). Für jedes Paar von vorläufigen EMen e^i und e^j muß dabei überprüft werden, ob die Menge $R(e^i)$ der in e^i involvierten Reaktionen eine Teilmenge von $R(e^j)$ ist (der Test in die umgekehrte Richtung erfolgt ebenfalls):

$$R(e^i) \subseteq R(e^j) ? \quad (4.1)$$

Gilt (4.1), dann ist e^i nicht elementar oder ein Duplikat und wird verworfen. Für eine schnelle Durchführung dieser Tests wird im *FluxAnalyzer* zusätzlich zu jedem EM e^i noch ein binäres Muster bm^i abgespeichert, das für jede beteiligte Reaktion in e^i eine Eins und für jede andere eine Null aufweist. Diese Muster werden (speichersparend) bitweise abgelegt und ermöglichen nun sehr schnelle binäre Operationen. So entspricht (4.1) dem (binären) Test:

$$(bm^i \text{ OR } bm^j) = bm^j ? \quad (4.2)$$

Einige Subroutinen des Algorithmus wurden zusätzlich als externe C-Funktionen programmiert, die in MATLAB über das bereitgestellte MEX-Interface eingebunden werden können. Gerade bei den hier auftretenden relativ einfachen Schleifen verbunden mit großem Speicherdurchsatz wirkt sich diese Option sehr vorteilhaft aus. In größeren Netzen ist die Berechnung der EMen bei Verwendung der C-Subroutinen um einen Faktor zwischen 4 und 8 schneller.

Da der Algorithmus – im Gegensatz zur Integer-Arithmetik in METATOOL [116] – auf endlich genauen Gleitkommazahlen operiert, kann der Benutzer mit der Variablen *epsilon* die kleinste Zahl größer Null definieren. Außerdem wird zur Vermeidung schlecht skalierter Zeilenoperationen jeder vorläufige EM normalisiert (auf seine größte absolute Komponente).

Die Berücksichtigung der oben genannten Punkte macht die Berechnung der EMen im *FluxAnalyzer* recht effizient und robust. Komplexe Netze mit mehr als einer halben Million EMen konnten behandelt werden. Beispiele für Berechnungen von EMen inkl. des dabei benötigten Zeitaufwandes finden sich in Kap. 8.

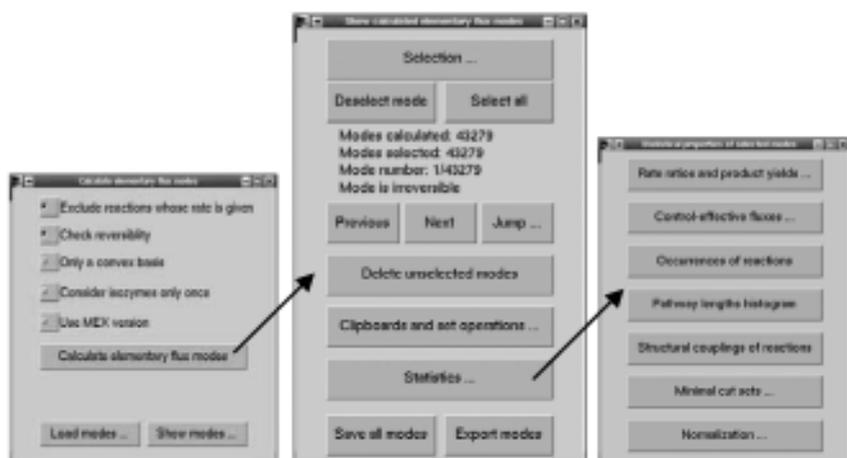


Abb. 4.6: Elementarmoden im FluxAnalyzer: Starten der Berechnung (links), Control Panel (Mitte) und statistische Funktionen (rechts).

Sind die EMen berechnet, steht eine Reihe von Methoden für ihre Auswertung zur Verfügung (Abb. 4.6 Mitte und rechts). Zunächst kann jeder EM in den Flußkarten – mit entsprechend farblicher Gestaltung der Textboxen – angezeigt werden (wie in Abb. 4.9). In Netzen mit mehreren Tausend Pathways ist das eine müßige Aufgabe. Eine hilfreiche Option ist es daher, Teilmengen der EMen nach bestimmten Kriterien zu selektieren („Selection-Tool“):

- Reaktionen, die im EM enthalten sein müssen / nicht enthalten sein dürfen
- Metabolite, die am EM beteiligt sein müssen / nicht beteiligt sein dürfen
- minimale/maximale Anzahl involvierter Reaktionen im EM (Pathwaylänge)
- (reversible) Reaktionen, die im EM nur vorwärts / nur rückwärts laufen dürfen

So können z.B. „alle EMen, in denen Biomasse produziert und Reaktion Rx nicht verwendet wird“, selektiert werden. Mit dem „Selection Clipboard“ kann man diese Selektionen mittels Mengenoperationen (wie Schnitt, Vereinigung, Differenz) auch kombinieren und verfeinern. Das ermöglicht schließlich auch kompliziertere Auswahlen wie „alle EMen, in denen Reaktion Rx oder Ry involviert ist“.

Der Vorteil einer Selektion liegt auch darin, daß statistische und kombinatorische Analysen (Abb. 4.6 rechts) nur auf der getroffenen Auswahl operieren. Dazu gehören z.B. die Auftrittshäufigkeiten der Reaktionen in den EMen (Kap. 6) oder das Aufspüren (aller) optimalen

EMen bzgl. eines vorgegebenen Kriteriums (Abschnitt 2.4.3). Insbesondere kann hier die Berechnung der *Minimalen Schnittmengen* gestartet werden (Kap. 7). Auf diese und andere Funktionalitäten geben wir an entsprechender Stelle noch ein.

Berechnete EMen können gespeichert oder für weitergehende Studien exportiert werden.

Metabole Flußanalyse und Bestimmung optimaler Flußverteilungen

Nach der Eingabe eines Szenarios in den Flußkarten (mit gemessenen oder bekannten Reaktionsraten) kann der Benutzer die metabole Flußanalyse starten. Diese Routine verfährt nach Prinzipien, die in Kap. 5 erläutert werden, und ermöglicht auch die Betrachtung unterbestimmter und redundanter Systeme. Sensitivitäten einzelner Reaktionsraten können bestimmt und Inkonsistenzen vorgegebener Raten können ausfindig gemacht werden.

Vorgegebene Reaktionsraten können auch als Nebenbedingung für die Optimierung einer Flußverteilung dienen (Abschnitt 2.4.2). Die min./max. erlaubten Werte für die Reaktionsraten, sowie die Koeffizienten der linearen Zielfunktion (2.10) werden über die entsprechenden Attribute der Reaktionen (Tab. 4.1) definiert. Damit können optimale Flußverteilungen für *beliebige lineare Zielfunktionen und Nebenbedingungen* berechnet werden.

Sonstige Funktionen

Aktuelle Werte können in einem Säulen-Diagramm dargestellt werden, was eine schnelle Identifikation von extremalen Werten ermöglicht. Flußkarten (inkl. der eingetragenen Werte) können gespeichert oder gedruckt werden. Die stöchiometrische Matrix kann für andere Werkzeuge (in ASCII-Format) exportiert werden. Für METATOOL steht ein Konverter zur Verfügung der ein Netzwerk sowohl ex- als auch importiert. ProMot/DIVA unterstützt den Export von Reaktionsnetzwerken im *FluxAnalyzer*-Format.

Mittelfristig ist die Entwicklung einer SBML-Schnittstelle [59] wünschenswert, die einen Austausch von Netzmodellen mit anderen Simulatoren und Modellierungswerkzeugen ermöglicht. Die Tatsache, daß der *FluxAnalyzer* von drei kommerziellen Unternehmen genutzt wird und von mehr als 150 Forschergruppen weltweit angefordert wurde, unterstreicht den Bedarf an Rechnerwerkzeugen, die detaillierte Analysen von Stoffwechselnetzen auf Basis ihrer Struktur unterstützen. Anwendungen des *FluxAnalyzers* finden sich in dieser Arbeit und in [12,73,76,80,148].

Prinzipiell lassen sich mit dem *FluxAnalyzer* alle Typen von stöchiometrischen Netzen untersuchen, so daß sich z.B. auch Studien durchführen lassen, die aus der Verfahrenstechnik bekannt sind [55,153].

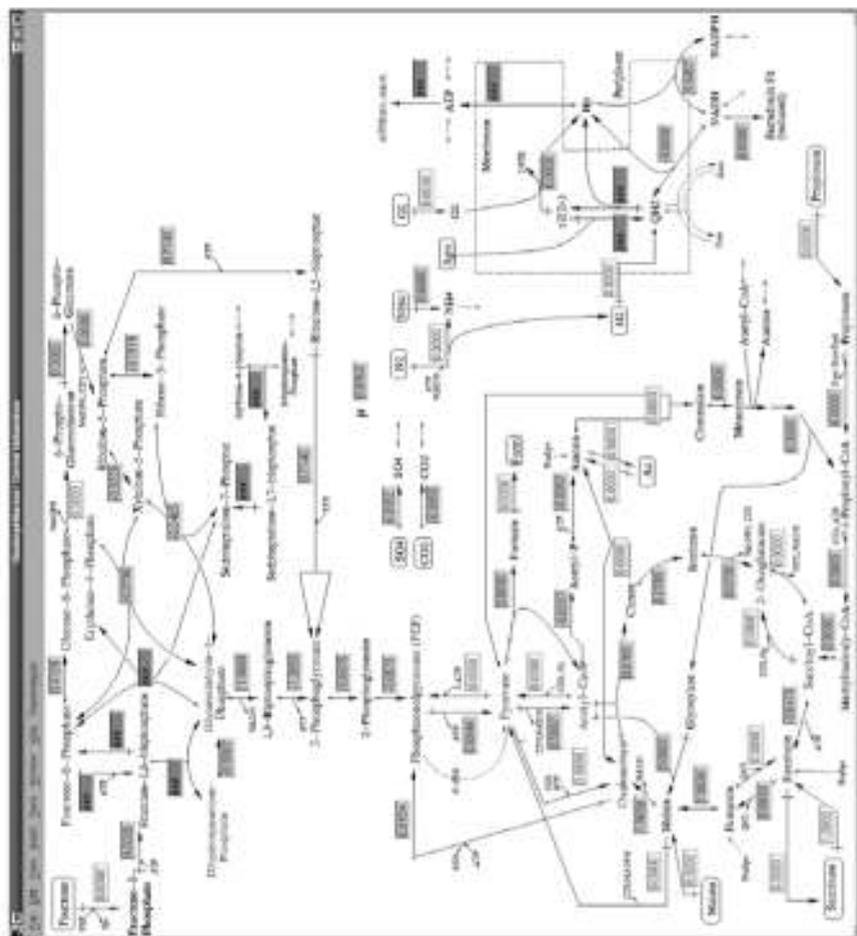


Abb. 4.7: Interaktive Flusskarte im FluxAnalyzer. Zeigt den Zentralstoffwechsel der schwefelfreien Purpurbakterien (Rhodospirillaceae) wie beschrieben in Abschnitt 3.3. und Anhang A.3. Jede Textbox ist hier jeweils einem Netzelement vom Typ „Reaktion“ zugeordnet. Wiedergegeben wird gerade das Flusszenario S2 vom Abschnitt 5.5.2. Weiße Boxen stehen für vom Benutzer vorgegebene Reaktionsraten, hellgraue zeigen berechnete Raten und dunkelgraue Boxen mit Eintrag „###“ stehen für unbekannte (nicht definierte und nicht berechenbare) Raten. Irreversible Reaktionen sind mit senkrechtem Strich am Anfang des Reaktionspfeils angedeutet. Das Menü verbirgt sich hinter dem Menüeintrag „FluxAnalyzer“.

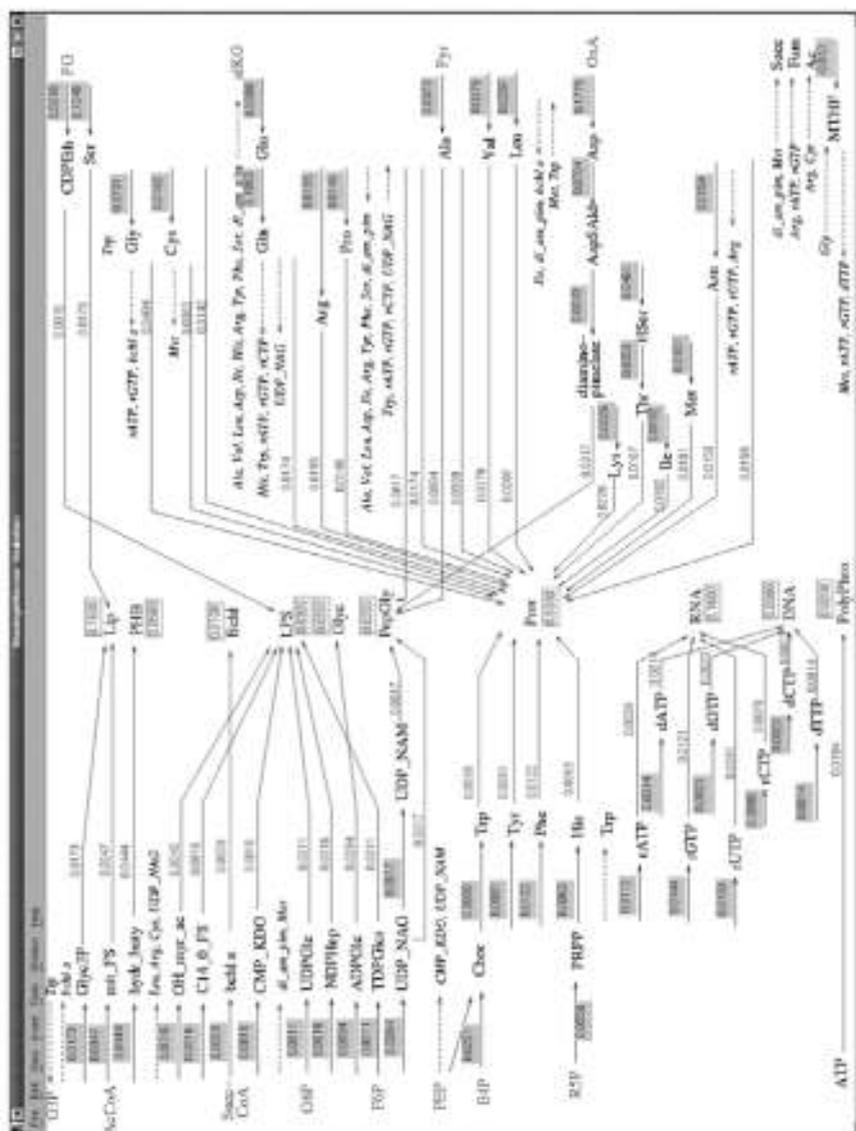


Abb. 4.8: Interaktive Fluxkarte im FluxAnalyzer: Anaboler Stoffwechsel der Rhodospirillaceae (vgl. Abschnitt 5.3 und Anhang A.3). In der Mitte der Fluxkarte kann die Biomassezusammensetzung aus den 10 betrachteten Makromolekülen vom Benutzer definiert werden (weiße Textboxen). Zahlenwerte ohne Rahmen gehören zu Assembly-Routen, Textboxen mit grauem Hintergrund sind jeweils mit Netz-elementen vom Typ Reaktion verknüpft. Die Textboxen sowohl der Assembly-Routen, als auch der anabolen Reaktionen wurden als „nicht editierbar“ konfiguriert, da ihre Raten nur von der Wachstumsrate μ und der Biomassezusammensetzung abhängen. Die hier gezeigten Zahlenwerte gehören wie Abb. 4.7 zum Szenario S2 von Abschnitt 5.5.2.

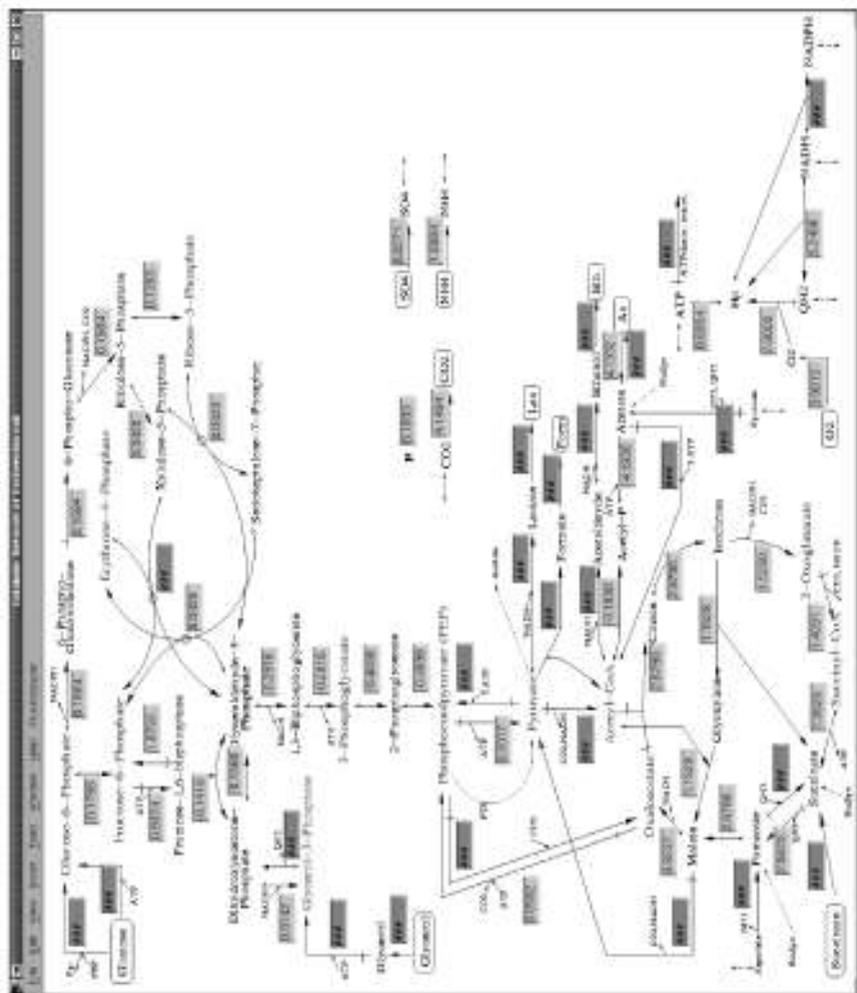


Abb. 4.9: Interaktive Flusskarte des Modells des (katabolen/amphibolen) Zentralstoffwechsels von *Escherichia coli* nach Abschnitt 3.2 und Anhang A.2. Alle Textboxen sind hier jeweils mit einem Netzelement vom Typ Reaktion zugeordnet. Gezeigt wird gerade ein Elementarmodus für Wachstum auf Acetat (nicht involvierte Reaktionen haben dunkle Textboxen). Siehe auch Anmerkungen zu Abb. 4.7. Die Karte zum anabolen Stoffwechsel von *E. coli* ähnelt der in Abb. 4.8.

5 Berechenbarkeit von stationären Flüssen in Stoffwechselnetzen

5.1 Berechenbarkeit und Redundanz in metabolen Flußanalysen

In Abschnitt 2.4.5 hatten wir die Grundlagen der metabolen Flußanalyse (MFA) auf Basis der steady-state Massenbilanzgleichungen der Metabolite bereits dargelegt. Bekannte oder gemessene Reaktionsraten reduzieren den Lösungsbereich für eine spezifische stationäre Flußverteilung und überführen das ursprünglich homogene Gleichungssystem in ein inhomogenes:

$$\text{Bilanzgleichungen der Metabolite im steady-state:} \quad \mathbf{0} = \mathbf{N}\mathbf{r} \quad (5.1)$$

$$\begin{aligned} \text{Flußszenario: bekannter (Index } b) \text{ und unbekannter (} u) \text{ Anteil:} \quad \mathbf{0} &= \mathbf{N}_b \mathbf{r}_b + \mathbf{N}_u \mathbf{r}_u \\ \mathbf{N}_u \mathbf{r}_u &= -\mathbf{N}_b \mathbf{r}_b \end{aligned} \quad (5.2)$$

Linke bzw. rechte Seite in (5.2) sollen als Nullvektoren mit m Elementen definiert sein, falls keine Rate unbekannt bzw. bekannt ist. Gl. (5.2) ist zentral für die MFA. Sofern \mathbf{N}_u quadratisch und invertierbar ist oder zumindest vollen Rang hat, sind die Lösungsverfahren aus der linearen Algebra gut bekannt. Damit sind aber nicht alle Fälle abgedeckt. Eine systematische Darstellung, wie mit (5.2) zu verfahren ist, wurde in [159] gegeben. Die Vorgehensweise wird in diesem Abschnitt aufgegriffen und z.T. vereinfacht und modifiziert, insbesondere bzgl. der spezifischen Berechenbarkeit einer Rate. Folgende Variablen werden verwendet:

- m = Anz. Metabolite (Zeilen in \mathbf{N} , \mathbf{N}_b und \mathbf{N}_u), q = Anz. Reaktionen (Spalten in \mathbf{N})
- u = Anz. unbekannter Raten = Anz. Spalten in \mathbf{N}_u = Anz. Elemente in \mathbf{r}_u
- $b = q - u$ = Anz. bekannter Raten = Anz. Spalten in \mathbf{N}_b = Anz. Elemente in \mathbf{r}_b

Bevor wir ein gegebenes Szenario (5.2) hinsichtlich Bestimmtheit und Redundanz klassifizieren, müssen wir noch klären, was wir hier unter einer *Lösung* für \mathbf{r}_u verstehen. Im Spezialfall, wo Matrix \mathbf{N}_u quadratisch, regulär ($\text{Rang}(\mathbf{N}_u) = u = m$) und invertierbar ist, können alle Raten in \mathbf{r}_u eindeutig mit der dann existierenden Inversen von \mathbf{N}_u berechnet werden:

$$\mathbf{r}_u = -\mathbf{N}_u^{-1} \mathbf{N}_b \mathbf{r}_b \quad (5.3)$$

Leider ergibt sich diese Konstellation bei der MFA eher selten bzw. wird durch entsprechende Annahmen „erzwingen“. Im allgemeinen Fall können wir nur eine *Kleinste-Quadrate-Lösung* (KQL) berechnen, die auch von der linearen Ausgleichsrechnung bekannt ist. Ist \mathbf{r}_u eine KQL, so minimiert sie den Abstand zwischen linker und rechter Seite in (5.2), d.h. wir fordern

$$\|\mathbf{N}_b \mathbf{r}_b + \mathbf{N}_u \mathbf{r}_u\|_2 = \min! \quad (5.4)$$

Die Lösung für (5.4) muß nicht eindeutig sein, d.h. mehrere KQLen können existieren (was in unterbestimmten Systemen der Fall ist, s.u.). Eine spezielle KQL liefert folgender Ansatz:

$$\mathbf{r}_a = -\mathbf{N}_a^{\#} \mathbf{N}_b \mathbf{r}_b \quad (5.5)$$

Hierbei ist $\mathbf{N}_a^{\#}$ die *(u,b)-Moore-Penrose Pseudoinverse* von \mathbf{N}_a , die sich aus der Singulärwertzerlegung von \mathbf{N}_a berechnen läßt [99,156]. Diese verallgemeinerte Inverse existiert für jede Matrix und ist damit allgemeiner als jene, die in [149] verwendet wurde¹. Im oben erwähnten Spezialfall fällt sie mit der Inversen von \mathbf{N}_a zusammen. Sofern die über Gl. (5.5) ermittelte KQL das Szenario (5.2) exakt löst, ist das Szenario (5.2) *konsistent*, andernfalls *inkonsistent*.

5.1.1 Klassifikation von Szenarien: Bestimmtheit und Redundanz

Das Szenario oder System (5.2) kann durch die beiden globalen Merkmale *Bestimmtheit* und *Redundanz* klassifiziert werden. Beide sind durch den Rang der Matrix \mathbf{N}_a bestimmt:

Bestimmtheit (globale Berechenbarkeit)

- *unterbestimmt*: $\text{Rang}(\mathbf{N}_a) < u$; es gibt mehrere KQLen, d.h. nicht alle Komponenten von \mathbf{r}_a sind über (5.4) eindeutig bestimmt. Das System hat $u - \text{Rang}(\mathbf{N}_a)$ viele Freiheitsgrade.
- *bestimmt*: $\text{Rang}(\mathbf{N}_a) = u$; es gibt nur genau eine KQL und damit sind alle Raten von \mathbf{r}_a eindeutig bestimmt.

Redundanz

- *redundant*: $\text{Rang}(\mathbf{N}_a) < m$; es bestehen lineare Abhängigkeiten zwischen einigen Zeilen in \mathbf{N}_a . Es lassen sich gewisse $(m - \text{Rang}(\mathbf{N}_a))$ Gleichungen (Zeilen) aus (5.2) entfernen, ohne daß dabei die Anzahl der Freiheitsgrade erhöht würde.
- *nichtredundant*: $\text{Rang}(\mathbf{N}_a) = m$; es bestehen keine linearen Abhängigkeiten zwischen Zeilen in \mathbf{N}_a . Das Entfernen einer beliebigen Gleichung (Zeile) im System (5.2) würde die Anzahl der Freiheitsgrade erhöhen.

Überbestimmt ist ein häufig gebrauchter Begriff in der Theorie linearer Gleichungssysteme, der aber leicht irreführend sein kann. Man könnte meinen, ein redundantes System wäre überbestimmt. Bestimmtheit und Redundanz sind aber unabhängige Eigenschaften und können in beliebiger Kombination der $2 \times 2 = 4$ Möglichkeiten auftreten (Tab. 5.1), d.h. das System kann durchaus redundant, aber dennoch unterbestimmt sein. Am ehesten läßt sich überbestimmt wohl als Spezialfall auffassen, wo das System bestimmt und redundant ist (Fall 4, Tab. 5.1).

Bzgl. der *Konsistenz* eines Systems läßt sich sagen, daß nur redundante Systeme inkonsistent sein können (und es in der Praxis auch häufig sind). Spezialfälle von Redundanzen sind Erhaltungsgleichungen (Abschnitt 2.5), die bereits ohne vorgegebene Raten, d.h. in Matrix \mathbf{N} in Gl. (5.1) auftreten. Sie sind immer konsistent.

¹ Dort ist $\mathbf{N}_a^{\#} = (\mathbf{N}_a^T \mathbf{N}_a)^{-1} \mathbf{N}_a^T$. Dann existiert $\mathbf{N}_a^{\#}$ aber nur wenn $\text{Rang}(\mathbf{N}_a) = u$ ist, was in unterbestimmten Systemen nicht gilt.

		Redundanz	
		nichtredundant $\text{Rang}(N_s) = n$	redundant $\text{Rang}(N_s) < n$
Bestimmtheit	unterbestimmt $\text{Rang}(N_s) < n$	Fall 1 <ul style="list-style-type: none"> nicht alle Raten in r_s sind eindeutig bestimmt (mit Freiheitsgraden) System ist garantiert konsistent 	Fall 2 <ul style="list-style-type: none"> nicht alle Raten in r_s sind eindeutig bestimmt (mit Freiheitsgraden) System kann inkonsistent sein
	bestimmt $\text{Rang}(N_s) = n$	Fall 3 <ul style="list-style-type: none"> alle Raten in r_s eindeutig bestimmt System ist garantiert konsistent Pseudoinverse = Inverse (von N_s) = exakt determiniertes System 	Fall 4 <ul style="list-style-type: none"> alle Raten in r_s eindeutig bestimmt System kann inkonsistent sein = „überbestimmtes“ System

Tab. 5.1: Klassifikation eines mit Gleichungssystem (5.2) gegebenen Flusszenarios

5.1.2 Klassifikation von Raten: Spezifische Berechenbarkeit und Redundanz

Redundanz und Bestimmtheit sind *globale* Eigenschaften eines Szenarios (5.2). In unterbestimmten bzw. redundanten Systemen bleibt die *spezifische* Bestimmtheit einer *einzelnen* Rate in r_s bzw. die *spezifische* Redundanz einer *einzelnen* Rate in r_s aber zunächst offen.

Spezifische Berechenbarkeit/Bestimmtheit einer unbekanntten Rate: Ein Element aus r_s ist

- *berechenbar (bestimmt, determiniert)*, wenn es *eindeutig* mit (5.4) bestimmt ist, also jede KQL denselben Wert für diese Komponente von r_s aufweist.
- *nichtberechenbar (unbestimmt)*, wenn es *nicht eindeutig* aus (5.4) berechnet werden kann. Unterbestimmte Systeme haben mind. eine nichtberechenbare Rate.

Spezifische Redundanz („Ausgleichbarkeit“) einer bekannten Rate: Ein Element aus r_s ist

- *nichtredundant*, wenn die Konsistenz des Systems (5.2) unabhängig von dieser Rate ist. In nichtredundanten Systemen gibt es nur nichtredundante Raten.
- *redundant*, wenn die Konsistenz des Systems (5.2) von dieser Rate mitbestimmt wird. Redundante Raten treten nur in redundanten Systemen auf und werden auch als „ausgleichbar“ (engl.: *balanceable*) bezeichnet, da sie in inkonsistenten Systemen angepaßt werden sollten.

Feststellen der Berechenbarkeit einer unbekanntten Rate

Gl. (5.5) liefert uns mit der Pseudoinversen für alle Komponenten von r_s eine KQL, der wir die tatsächlich eindeutig berechenbaren Raten von r_s entnehmen können. Bleibt die Frage, wie wir die berechenbaren Raten finden. In [159] wurde eine Methode vorgeschlagen, bei der die Singulärwertzerlegung von N_s untersucht wird. Allerdings wurde der Ansatz weder dort, noch in der angegebenen Referenz klar motiviert oder bewiesen. Einen leichter nachvollziehbaren Ansatz, aus dem wir später auch eine *verallgemeinerte Berechenbarkeitsanalyse* ableiten können (Abschnitt 5.2), konstruieren wir wie folgt:

In einem bestimmten System ($\text{Rang}(\mathbf{N}_0) = \mu$; Fälle 3 und 4 in Tab. 5.1) sind alle unbekanntes Raten eindeutig berechenbar, d.h. mit (5.5) bzw. (5.3) gegeben. Wir können uns daher auf unterbestimmte Systeme (Fall 1 oder 2) beschränken, wo $\text{Rang}(\mathbf{N}_0) < \mu$ ist. Dann besitzt \mathbf{N}_0 einen Nullraum mit Dimension $\mu - \text{Rang}(\mathbf{N}_0)$ und somit gibt es eine (nicht eindeutige) Kernmatrix \mathbf{K}_0 mit $\mu - \text{Rang}(\mathbf{N}_0)$ vielen linear unabhängigen Spalten, so daß $\mathbf{N}_0 \mathbf{K}_0 = \mathbf{0}$ (Nullmatrix). Dann gibt es unendlich viele KQLen \mathbf{r}_0 , wobei (5.5) mit der Pseudoinversen gerade die (partikuläre) Lösung $\mathbf{r}_{0,part}$ mit minimaler Länge liefert [99]. Der gesamte Lösungsraum läßt sich nun allgemeiner darstellen als

$$\mathbf{r}_0 = -\mathbf{N}_0^+ \mathbf{N}_0 \mathbf{r}_0 + \mathbf{K}_0 \mathbf{a} = \mathbf{r}_{0,part} + \mathbf{K}_0 \mathbf{a} \quad (5.6)$$

Der Vektor \mathbf{a} drückt die Unbestimmtheit des Systems aus und ist ein beliebiger Vektor mit $\mu - \text{Rang}(\mathbf{N}_0)$ vielen Komponenten (Freiheitsgraden). Es läßt sich leicht zeigen, daß jedes \mathbf{r}_0 nach (5.6) ebenso eine KQL für (5.2) ist, wie $\mathbf{r}_{0,part}$:

$$\mathbf{N}_0 \mathbf{r}_0 = \mathbf{N}_0 \mathbf{r}_{0,part} + \mathbf{N}_0 \mathbf{K}_0 \mathbf{a} = \mathbf{N}_0 \mathbf{r}_{0,part} \quad (5.7)$$

Mit $\mathbf{r}_{0,part}$ und der Kernmatrix \mathbf{K}_0 können also alle KQLen von (5.2) generiert werden. Allerdings müssen in (5.5) *nicht alle* Komponenten von \mathbf{r}_0 variabel sein: Eine Komponente j in \mathbf{r}_0 ist fixiert, wenn das Skalarprodukt aus j -ter Zeile von \mathbf{K}_0 mit *jedem* beliebigen Vektor \mathbf{a} immer 0 ergibt. Das ist genau dann der Fall, wenn die j -te Zeile von \mathbf{K}_0 eine Nullzeile ist. Damit haben wir das Kriterium für die Berechenbarkeit einer Rate gefunden:

Lemma 5.1: Das j -te Element von \mathbf{r}_0 ist im Szenario (5.2) genau dann eindeutig bestimmt (berechenbar), wenn die j -te Zeile in der Kernmatrix \mathbf{K}_0 von \mathbf{N}_0 eine Nullzeile ist.

Lemma 5.1 besagt letztlich, daß der Lösungsraum (5.6) orthogonal zu einer berechenbaren Komponente j steht. Die Werte berechenbarer Raten können der partikulären Lösung \mathbf{r}_0 in (5.5) entnommen werden. Damit können also Raten (oder allgemeiner: Variablen) in unterbestimmten Systemen determiniert sein – ein Fakt, der selbst in der Literatur zur linearen Algebra kaum erwähnt wird. In der MFA haben diese „frühzeitig“ berechenbaren Flüsse eine besondere physiologische Bedeutung, kennzeichnen sie doch Strukturzwänge im Stoffwechsel, die entsprechende Konsequenzen für das Verhalten des Systems haben (siehe Abschnitt 5.5).

Überprüfen der Redundanz einer (bekannten) Rate

Eine KQL minimiert die $\|\cdot\|_2$ -Norm des Residuenvektors $\mathbf{res} = \mathbf{N}_0 \mathbf{r}_0 + \mathbf{N}_s \mathbf{r}_s$ (siehe (5.4)). Die Residuen lassen sich durch Einsetzen von (5.5) in (5.2) berechnen als:

$$\mathbf{res} = \mathbf{N}_0 \mathbf{r}_0 + \mathbf{N}_s \mathbf{r}_s = \mathbf{N}_0 \mathbf{r}_0 - \mathbf{N}_0 \mathbf{N}_0^+ \mathbf{N}_0 \mathbf{r}_0 = (\mathbf{N}_0 - \mathbf{N}_0 \mathbf{N}_0^+ \mathbf{N}_0) \mathbf{r}_0 = \mathbf{R} \mathbf{r}_0 \quad (5.8)$$

\mathbf{R} ist die Redundanzmatrix [149,159]. In konsistenten Systemen, gib es mindestens eine Lösung, bei der die Residuen den Nullvektor ergeben, d.h.:

$$\mathbf{0} = \text{res} = \mathbf{R}\mathbf{r}_b \quad (5.9)$$

In *nichtredundanten* Systemen ist (5.9) automatisch erfüllt, da dann \mathbf{R} eine Nullmatrix ist. In redundanten Systemen existieren dagegen fast immer von Null verschiedene Einträge in \mathbf{R} . Gemessene Raten in \mathbf{r}_b , deren korrespondierende Spalten in \mathbf{R} nur Nullen enthalten, haben jedoch definitiv keinen Einfluß auf den Fehler (Konsistenz) des Systems, d.h. solche Raten sind *nichtredundant* [159]. Die j -te Komponente in \mathbf{r}_b ist dagegen *redundant*, d.h. mitentscheidend für die Konsistenz des Szenarios, wenn die entsprechende j -te Spalte in \mathbf{R} *keine Nullspalte* ist. In den meisten realen Fällen verursachen redundante Raten in \mathbf{r}_b tatsächlich Inkonsistenz, d.h. einen nicht verschwindenden Residuenvektor. Redundante Raten sollten deshalb vor der Berechnung der berechenbaren Raten angepaßt werden, um ein konsistentes System zu erhalten. Dabei berücksichtigt man gewöhnlich auch die Varianz-Covarianz-Matrix der normalverteilten Meßfehler (d.h. die Genauigkeit) der gemessenen Raten \mathbf{r}_b . Dies führt zu einer gewichteten Kleinste-Quadrate-Schätzung $\hat{\mathbf{r}}_b$, die dann konsistent ist und statt \mathbf{r}_b in (5.5) für die Berechnung der unbekanntenen Raten verwendet wird. Aus der Differenz zwischen $\hat{\mathbf{r}}_b$ und \mathbf{r}_b lassen sich u.U. große Meß- oder Modellfehler bzgl. eines vorgegebenen Konfidenzintervalls detektieren. Die Vorgehensweise ist ausführlich beschrieben in [149,160]. Uns begegnen redundante Systeme nur noch in den Beispielen.

5.1.3 Umsetzung im FluxAnalyzer

Hat man im *FluxAnalyzer* die Werte von (gemessenen) Raten in den interaktiven Flußkarten eingegeben, kann die Flußanalyse gestartet werden („Flux analysis ...“). Zunächst wird dann intern die Redundanz des Systems überprüft. In nichtredundanten Szenarien werden die berechenbaren Raten sofort ermittelt und ausgegeben. Bei redundanten System kann der Benutzer zwischen verschiedenen Schätzverfahren auswählen, die in den (evtl.) korrigierten Meßwerten ein konsistentes System liefern (Varianzen der Messungen können einbezogen werden: Tab. 4.1). Erst dann werden die berechenbaren Raten ermittelt. In redundanten Systemen überprüft ein Konsistenztest schließlich noch die Wahrscheinlichkeit von groben Meß- oder Modellfehlern. Mit einer Sensitivitätsanalyse kann der (lineare) Einfluß von vorgegebenen Raten auf die bestimmten Werte von berechenbaren Raten festgestellt werden. Wegen (5.5) entnimmt man diese Werte folgender Matrix:

$$\frac{d\mathbf{r}_a}{d\mathbf{r}_b} = -\mathbf{N}_a^* \mathbf{N}_b \quad (5.10)$$

5.1.4 Beispiele

Um die oben skizzierte Vorgehensweise zu illustrieren, sind in Abb. 5.1 drei Flußszenarien im Netz NB2 (Abb. 2.3) inkl. der dazugehörigen Matrizen \mathbf{K}_s und \mathbf{N}_s dargestellt (die Werte berechenbarer Raten wurden gemäß (5.5) bestimmt und sind leicht überprüfbar). Szenario 1 zeigt, daß unterbestimmte Systeme berechenbare Raten aufweisen können. Szenario 2 ist exakt determiniert, wo der Nullraum nur noch den Nullvektor enthält. Szenario 3 demonstriert, daß ein unterbestimmtes System redundant sein kann. Der Redundanzmatrix von Szenario 3 (nicht gezeigt) kann man entnehmen, daß die Raten R4 und R7 redundant sind, während R3 nichtredundant ist. R4 und R7 (und damit das Szenario) sind inkonsistent, da sie im konsistenten Fall beide identisch sein müssen. Sind die Messungen beider Raten normalverteilt mit gleicher Varianz, so ergibt die Schätzung die korrigierten Werte $\hat{R}4 = \hat{R}7 = 2$, mit denen die berechenbaren Raten R9 und R10 bestimmt werden.

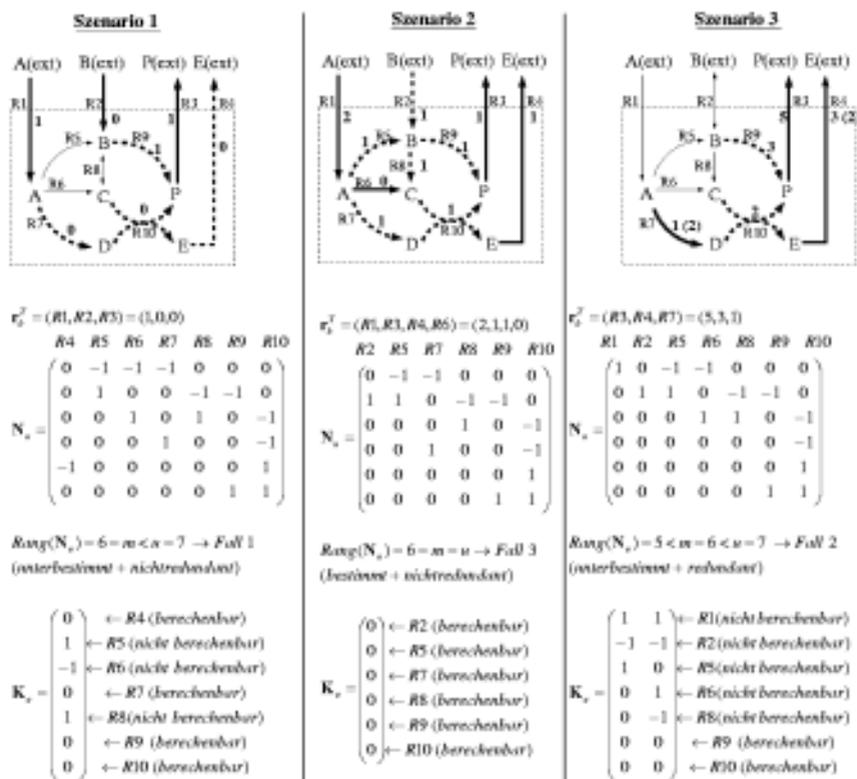


Abb. 5.1: Drei Flußszenarien aus dem Netz NB2 (Abb. 2.4). Die gemessenen Raten wurden mit dicken, die berechneten mit dicken gestrichelten Pfeilen gekennzeichnet.

5.2 Verallgemeinerte Berechenbarkeitsanalyse in stationären Reaktionsnetzwerken

Der letzte Abschnitt hat gezeigt, wie wir in einem konkreten Szenario die Berechenbarkeit einer Rate feststellen können. Allerdings können auch allgemeinere Fragen auftreten, wie

- (i) Welche Kombination von Messungen erlaubt die Berechnung einer bestimmten Rate? Oder welche zusätzlich gemessenen Raten in einem gegebenen Szenario (5.2) determinieren das gesamte System oder eine bestimmte Rate?
- (ii) Welche Raten können niemals durch Messungen berechenbar werden?

Zur Beantwortung dieser Fragen benötigen wir eine detailliertere Analyse des Nullraums von N_s , da dessen Gestalt nach Lemma 5.1 die spezifische Berechenbarkeit der unbekannt Raten festlegt. Beziehen wir uns im Folgenden auf N_s (stöchiometrische Matrix der unbekannt Raten eines Szenarios (5.2)), so schließt das den Spezialfall (5.1) ohne bekannte Raten mit ein, wo $N=N_s$ ist. Wir erinnern uns, daß der Nullraum von N_s alle Vektoren \mathbf{k} mit

$$\mathbf{0} = N_s \mathbf{k} \quad (5.11)$$

enthält. Eine, i.a. nicht eindeutige, Darstellung des Nullraums liefert die Kernmatrix \mathbf{K}_s , die eine Basis des Nullraums enthält. Zunächst wollen wir untersuchen, wie sich Nullraum und Kernmatrix verändern, wenn wir eine weitere Rate in einem Szenario (5.2) vorgeben. Die neue Matrix \tilde{N}_s ist schnell abgeleitet: wir brauchen nur die entsprechende Spalte in N_s zu streichen (und in N_f anhängen). Können wir auch den neuen Nullraum von \tilde{N}_s ableiten? Das würde uns helfen, solche Auswahl von Messungen zu treffen, mit der das Kriterium von Lemma 5.1 für eine gesuchte Rate erfüllt wird. Zunächst einmal gilt:

Lemma 5.2: Man betrachte das Szenario (5.2) und nehme an, ein Vektor \mathbf{k} liege im Nullraum der Matrix N_s . Eine unbekannt Rate, die z -te Komponente von \mathbf{r}_s , wird nun zusätzlich gemessen. Der Vektor $\tilde{\mathbf{k}}$, der durch das Streichen der z -ten Komponente von \mathbf{k} entsteht, liegt im Nullraum der neuen Matrix \tilde{N}_s , falls die gestrichene z -te Komponente Null ist.

Diese Aussage folgt daraus, daß solche Linearkombinationen von Spalten in N_s noch immer mit \tilde{N}_s realisierbar sind, bei denen nicht auf die z -te Spalte von N_s zugegriffen wurde. Das gilt natürlich auch für jene, die den Nullvektor ergeben (5.11). Wir illustrieren Lemma 5.2 am simplen Netz NB1 (Abb. 2.2). Graphik, N und eine Kernmatrix \mathbf{K} seien noch einmal gegeben:

The diagram shows a reaction network with two species, A and B, enclosed in a dashed box. Reaction R1 is an input from the left to species A. Reaction R2 is a cycle from A to B and back to A. Reaction R3 is a cycle from B to A and back to B. Reaction R4 is an output from species B to the right.

$$N = \begin{pmatrix} R1 & R2 & R3 & R4 \\ 1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix}$$

$$K = \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{matrix} R1 \\ R2 \\ R3 \\ R4 \end{matrix} \quad (5.12)$$

Wir betrachten nun den Fall, daß R1 gemessen wird. Die erste Spalte in \mathbf{N} verschwindet folglich und mit ihr eine Dimension des Nullraums. Die einzig verbleibende Dimension des neuen Nullraums (d.h. Spalte in $\tilde{\mathbf{K}}_s$) können wir wegen Lemma 5.2 aus der zweiten Spalte von \mathbf{K} ableiten, da diese an der ersten, zu streichenden Komponente R1 eine Null aufweist :

$$\tilde{\mathbf{N}}_s = \begin{pmatrix} R2 & R3 & R4 \\ -1 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}, \quad \tilde{\mathbf{K}}_s = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \begin{matrix} R2 \\ R3 \\ R4 \end{matrix} \quad (5.13)$$

Leider funktioniert diese Vorgehensweise nur selten. Einfaches Gegenbeispiel: Man betrachte in (5.12) nun Rate R2 als gemessen. Beide Spalten in \mathbf{K} würden nicht „überleben“, da sie R2 involvieren. Dennoch hat natürlich auch in diesem Fall der neue Nullraum die Dimension 1. Die Kernmatrix scheint keine geeignete Darstellung des Nullraums zu sein, mit der wir den neuen Nullraum nach Entfernen einer Spalte von \mathbf{N}_s konstruieren könnten. Geeigneter wäre eine Menge von Vektoren, angeordnet in einer Matrix \mathbf{D}_s , mit folgenden Eigenschaften:

- (E1) $\mathbf{0} = \mathbf{N}_s \mathbf{D}_s$, d.h. alle Spalten von \mathbf{D}_s liegen im Nullraum von \mathbf{N}_s
 - (E2) Für alle \mathbf{k} , die im Nullraum von \mathbf{N}_s liegen (5.11), gibt es ein \mathbf{v} mit $\mathbf{k} = \mathbf{D}_s \mathbf{v}$.
 - (E3) Nach Streichen einer Spalte z in \mathbf{N}_s läßt sich das neue $\tilde{\mathbf{D}}_s$ aus \mathbf{D}_s ableiten, indem Lemma 5.2 auf jede Spalte \mathbf{d} von \mathbf{D}_s angewendet wird: Ist das z -te Element in \mathbf{d} Null so wird \mathbf{d} nach Streichen dieses z -ten Elements in $\tilde{\mathbf{D}}_s$ berücksichtigt und sonst nicht.
 - (E4) \mathbf{D}_s ist minimal, d.h. keine Teilmenge der Spalten von \mathbf{D}_s erfüllt (E1)-(E3).
- (5.14)

Die Menge \mathbf{E}_s der Elementarmoden, ermittelt bzgl. \mathbf{N}_s und angeordnet als Spalten in der Matrix \mathbf{D}_s , weisen diese Eigenschaften auf, sofern alle Reaktionen (Spalten in \mathbf{N}_s) als reversibel betrachtet werden! (E1) entspricht (B1) in (2.13); (E2) gilt wegen (2.23); (E3) ergibt sich aus (2.16) und (E4) wird gerade durch die Unzerlegbarkeit (B3) in (2.13) sichergestellt.

Für \mathbf{D}_s wird (B2) von (2.13) nicht gefordert, d.h. die Reversibilität der Reaktionen wird hier nicht berücksichtigt und darf es auch nicht (siehe auch Abschnitt 7.3.4). Wir tun so, als wären alle Reaktionen im System reversibel, berechnen die EMen und bilden daraus die Spalten von \mathbf{D}_s . Damit wird sichergestellt, daß der gesamte Nullraum von \mathbf{N}_s abgedeckt wird und somit (E2) für alle Vektoren \mathbf{k} des Nullraums gilt. Da alle Reaktionen reversibel sind, gilt das auch für die berechneten EMen. Der Einfachheit halber betrachten wir deshalb – im Unterschied zu (2.14) – jeden dieser EMen in \mathbf{D}_s nur einmal. Das ist kein Problem, da wir in (E2) beliebige Vektoren \mathbf{v} mit u.U. negativen Komponenten zulassen.

Mit \mathbf{D} kennzeichnen wir wieder den Spezialfall (5.1) ohne gemessene Raten. Wir müssen dazu also die EMen auf Basis der reversiblen Stöchiometrie in \mathbf{N} berechnen. In NB1 ergibt sich:

$$\mathbf{D} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad (5.15)$$

Die ersten beiden Vektoren in \mathbf{D} entsprechen den beiden EMen von NB1, wo alle Reaktionen irreversibel sind (vgl. Abb. 2.8). Der dritte Vektor (dritte Flußverteilung in Abb. 2.8) ist nur im vollständig reversiblen Fall ein weiterer EM und ist hier in \mathbf{D} zu berücksichtigen.

In [77] wurde \mathbf{D}_0 als *elementare Repräsentation des Nullraums* von \mathbf{N}_0 bezeichnet. *Kombinatorische Repräsentation* wäre vielleicht noch passender, da die Spalten in \mathbf{D}_0 letztlich alle Kombinationen *minimal* linear abhängiger Spalten in \mathbf{N}_0 repräsentieren. Wie die EMen sind die Spalten in \mathbf{D}_0 bis auf Skalierung eindeutig bestimmt. \mathbf{D}_0 enthält mind. so viele Elemente wie eine Basis des Nullraums von \mathbf{N}_0 , meistens aber deutlich mehr. Sehr nützlich ist die Tatsache, daß Lemma 5.1 wegen (E1) und (E2) in (5.14) auch für \mathbf{D}_0 gilt: Eine Komponente j aus \mathbf{r}_0 ist berechenbar, wenn die j -te Zeile in \mathbf{D}_0 eine Nullzeile ist. Mit diesen und obigen Eigenschaften von \mathbf{D}_0 können wir die zu Beginn des Abschnitts gestellten Fragen beantworten:

Anwendung (i): Welche Messungen machen eine Rate berechenbar?

Die z -te Komponente in \mathbf{r}_0 sei bisher nicht eindeutig berechenbar (Zelle j in \mathbf{D}_0 keine Nullzeile). Um sie berechenbar zu machen, bauen wir eine Menge \mathbf{M} (zusätzlich) zu messender Raten wie folgt auf: Alle solche Spalten \mathbf{d} in \mathbf{D}_0 (oder \mathbf{D}) müssen „entfernt“ werden, in denen die z -te Komponente nicht Null ist. Das erreichen wir, wenn in jeder dieser Spalten mindestens eine zu messende Rate $y \in \mathbf{M}$ ebenfalls ungleich Null (d.h. in \mathbf{d} involviert) ist. Als Beispiel solle R4 in NB1 berechenbar gemacht werden. (5.15) verrät uns, daß R4 im ersten und zweiten Vektor von \mathbf{D} auftaucht. In beiden Vektoren muß nun mindestens eine zu messende Rate enthalten sein. Eine Möglichkeit wäre, R2 und R3 zu messen. Naheliegender aber ist, nur R1 zu messen, denn diese Rate ist ebenfalls ungleich Null in beiden Vektoren. Wir erhalten in letzterem Fall für die neuen Matrizen:

$$\tilde{\mathbf{N}}_0 = \begin{pmatrix} R2 & R3 & R4 \\ -1 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}, \quad \tilde{\mathbf{K}}_0 = \begin{pmatrix} (-1) & R2 \\ 1 & R3 \\ 0 & R4 \end{pmatrix}, \quad \tilde{\mathbf{D}}_0 = \begin{pmatrix} | & | & | & | & | \\ \hline 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} (-1) & R2 \\ 1 & R3 \\ 0 & R4 \end{pmatrix} \quad (5.16)$$

$\tilde{\mathbf{D}}_0$, die kombinatorische Darstellung des neuen Nullraums) leitet sich nach Lemma 5.2 ab: Wir müssen die erste Zeile streichen (da R1 nun gemessen wird), sowie alle Spalten, in denen die R1-Komponente ungleich Null ist. $\tilde{\mathbf{D}}_0$ zeigt uns nun, daß die Zeile für R4 wie gewünscht Null und damit berechenbar ist (man kann leicht sehen, daß im steady-state $R4=R1$ gelten

muß). In diesem speziellen Fall gilt sogar $\tilde{\mathbf{K}}_s = \tilde{\mathbf{D}}_s$, da nur noch eine Dimension im Nullraum vorliegt. Würden wir statt R1 die Rate R2 messen, ergibt sich analog:

$$\tilde{\mathbf{N}}_s = \begin{pmatrix} R1 & R3 & R4 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \quad \tilde{\mathbf{K}}_s = \begin{pmatrix} 1 & R1 \\ 1 & R3 \\ 1 & R4 \end{pmatrix}, \quad \tilde{\mathbf{D}}_s = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & R2 \\ 1 & R3 \\ 1 & R4 \end{pmatrix} \quad (5.17)$$

Wir hätten demzufolge keine weitere Rate bestimmt. Wir können nun aber mit $\tilde{\mathbf{D}}_s$ ganz analog fortfahren und den Nullraum sukzessive weiter einschränken. In (5.17) würde die Messung einer beliebigen weiteren Rate alle anderen Raten und damit das gesamte System berechenbar machen. In (5.16) würde dagegen nur die zusätzliche Messung von R2 oder R3 das System vollständig determinieren, während die zusätzliche Messung der ohnehin berechenbaren Rate von R4 ein redundantes System zur Folge hätte.

Zunächst bleibt hier die Frage offen, wie man zu einer vorgegebenen Menge zu bestimmender Raten, systematisch *alle möglichen* Kombinationen von Messungen ermitteln kann, die diese Anforderung erfüllen. Die geeignetste Meßkombination ließe sich dann identifizieren. Ein solches Verfahren wird in Kap. 7 mit Hilfe der Minimalen Schnittmengen konstruiert.

Anwendung (ii): Identifikation niemals berechenbarer Raten

Wir können nun auch leicht Reaktionen identifizieren, deren Rate unter keinen Umständen mit der MFA (über Massenbilanzierung) bestimmbar sind. Solche Raten können weder gemessen, noch durch beliebige Kombination meßbarer Raten berechenbar gemacht werden. Letzteres tritt ein, wenn in mind. einer Spalte in \mathbf{D} (oder \mathbf{D}_s) nur solche Reaktionen von Null verschiedene Einträge haben, deren Raten nicht meßbar sind. Trotz Messungen können wir diese Spalte dann nämlich niemals aus \mathbf{D}_s und somit auch nicht aus dem Nullraum eliminieren, d.h. sie verbleibt immer als Freiheitsgrad im System. Ein typisches Beispiel liefert uns NB1: Da R1 und R4 über die Systemgrenzen gehen, würden sie im realen Fall wohl meßbaren Raten entsprechen. Damit könnten wir aber R2 und R3 nicht berechenbar machen, denn der dritte Vektor in \mathbf{D} involviert nämlich gerade ausschließlich diese (nicht meßbaren) Raten R2 und R3. Das führt zu der Feststellung, daß R2 und R3 niemals bestimmbar wären.

In realen Situationen können solche intrazellulären Raten nur noch über ^{13}C -Markierungsexperimente mit darauffolgender Bestimmung von Isotopomerenverteilungen mittels NMR oder GC-MS bestimmt werden [16,38,154,169]. Diese Methoden werden für die Quantifizierung metaboler Flußverteilungen immer wichtiger, da sie der natürlich gegebenen Unterbestimmtheit metaboler Netze einen hohen Grad an Meßinformation entgegensetzen. Die hier vorgestellte Methode könnte auch bei der Planung solcher Experimente hilfreich sein.

Umsetzung im FluxAnalyzer

Die Vektoren von \mathbf{D} bzw. \mathbf{D}_s können im *FluxAnalyzer* als „spezielle“ EMEN berechnet werden. Man muß dabei lediglich die Option „Check reversibility“ ausschalten (Abb. 4.6 links). Mit dem Selektions-Tool lassen sich oben beschriebene Anwendungen leicht umsetzen.

5.3 Parallelen mit der Beobachtbarkeitsanalyse in der Regelungstechnik

Berechenbarkeitsanalysen nach Abschnitten 5.1 und 5.2 sind letztlich auf beliebige lineare Gleichungssysteme anwendbar. Einen zur Berechenbarkeit ähnlichen Begriff, die *Beobachtbarkeit*, kennt man aus der Regelungstechnik. Die Beobachtbarkeit beschreibt eine globale Struktureigenschaft eines untersuchten Systems oder Regelkreisgliedes [39]. Häufig werden dabei lineare (oder linearisierte) zeitinvariante Systeme betrachtet, die folgende Gestalt haben:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}\end{aligned}\quad (5.18)$$

Dabei ist \mathbf{x} der $(q,1)$ -Zustandsvektor des Systems, \mathbf{y} der $(m,1)$ -Ausgangsvektor und \mathbf{u} ist der $(p,1)$ -Eingangsvektor. \mathbf{A} ist die (q,q) -Systemmatrix, die das dynamische Verhalten von \mathbf{x} diktiert, \mathbf{B} die (q,p) -Eingangsmatrix, \mathbf{C} die (m,q) -Ausgangsmatrix und \mathbf{D} die (m,p) -Durchgangsmatrix. *Beobachtbarkeit* in einem solchen System liegt vor, wenn man bei bekanntem \mathbf{y} und \mathbf{u} den Vektor \mathbf{x} eindeutig ermitteln kann. Als hinreichende und notwendige Bedingung für Beobachtbarkeit muß das *Kalman-Kriterium* gelten [39]. Danach ist (5.18) beobachtbar, wenn die $(q \cdot m, q)$ Beobachtbarkeitsmatrix \mathbf{Q} mit

$$\mathbf{Q} = \begin{pmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \dots \\ \mathbf{C}\mathbf{A}^{r-1} \end{pmatrix}\quad (5.19)$$

vollen Rang q besitzt. Nach Tab. 5.1 wird hier letztlich gefordert, daß ein lineares Gleichungssystem der Gestalt $\mathbf{a} = \mathbf{Q}\mathbf{x}$ bestimmt (vollständig berechenbar) sein muß.

Auch in (5.18) kann man die *spezifische* Beobachtbarkeit einzelner Komponenten von \mathbf{x} hinterfragen, worauf in Standard-Lehrbüchern relativ selten eingegangen wird. Lemma 5.1 liefert dafür einen einfachen Test: die i -te Komponente von \mathbf{x} ist beobachtbar, sofern in einer Kernmatrix von \mathbf{Q} die i -te Zeile eine Nullzeile ist. Auch die verallgemeinerte Berechenbarkeitsanalyse nach Abschnitt 5.2 ist analog für (5.18) anwendbar, d.h. über eine Analyse der Elementarmoden bzgl. \mathbf{Q} kann man nach Kombinationen von zu messenden Größen suchen, die bestimmte (oder alle) Komponenten in \mathbf{x} beobachtbar machen. (Allerdings würde man hier davon ausgehen, daß jede gemessene Größe ein Vielfaches einer Zustandsgröße x_i liefert).

5.4 Realisierbarkeit eines Flußszenarios

Im Abschnitt 5.1 haben wir ein Flußszenario (5.2) als konsistent bezeichnet, wenn die vorgegebenen Raten r_b tatsächlich eine ausgeglichene Flußverteilung zulassen. Mit anderen Worten: der Term $-N_b r_b$ muß im Bildbereich von N_a liegen, d.h. die Residuen in (5.9) sind Null. Allerdings kann ein Flußszenario konsistent in diesem algebraischen Sinne sein, ohne daß jedoch tatsächlich eine realisierbare Flußverteilung mit den vorgegebenen Raten im Reaktionsnetz existiert. Ein einfaches Beispiel: im Netz NB2 (Abb. 2.3) seien die Reaktionsraten $R1=0$ und $R4=1$ bekannt. Dieses Szenario wäre unterbestimmt und nichtredundant und damit konsistent im Sinne von (5.9). Berechenbar wären die Raten von $R7$ und $R10$, die mit der von $R4$ identisch sind. Dennoch kann man sich leicht überzeugen, daß keine *realisierbare* Flußverteilung mit diesen Vorgaben existiert. Die notwendige Bereitstellung von A (Edukt in $R7$) wäre dann nämlich nicht möglich, da $R5$ und $R6$ irreversibel sind. (Produkt E kann nur produziert werden, wenn A aufgenommen wird, d.h. $R1>0$ ist). In einem größeren Netzwerk würde der Betrachter vermutlich arglos den berechneten Werten vertrauen.

Die Ursache für diese Art der Inkonsistenz liegt in einer Verletzung der Vorzeichenbedingungen (Reversibilitäten) bzgl. der Reaktionsraten. Eine Erweiterung des Konsistenzbegriffs, der diese Nebenbedingungen an die Reaktionsraten ebenfalls berücksichtigt, ist daher sinnvoll:

(5.20) Realisierbarkeit: Ein Flußszenario ist *realisierbar*, wenn es zum vorgegebenen Vektor r_b der bekannten Raten mind. eine Lösung r_v gibt, so daß

- (i) $N_a r_a = -N_b r_b$ (Konsistenz nach (5.9))
- (ii) für jedes $r_{a,i}$ ($= i$ -te Komponente in r_a) werden evtl. vorgegebene Schranken für minimale (α_i) und maximale (β_i) Reaktionsrate eingehalten, d.h. $\alpha_i \leq r_{a,i} \leq \beta_i$
(Bedingung (ii) muß natürlich ebenso für die Komponenten der vorgegebenen Raten in r_b gelten)

In den meisten Fällen kennt man nur Vorzeichenbedingungen für die Reaktionsraten ($\alpha_i = 0$ für irreversible Reaktionen) und sonst keine weiteren Schranken (und setzt sie deshalb auf $-\infty$ bzw. $+\infty$). Realisierbarkeit bedeutet dann letztlich, daß es mind. eine zulässige, stationäre Flußverteilung $r = (r_a, r_b)$ gibt, die im Flußkegel F liegt (2.21), wobei der Anteil r_b gerade vorgegeben ist. Werden auch von Null und Unendlich verschiedene Schranken vorgegeben, wird der Kegel nach oben begrenzt.

Kriterium (ii) ist für die vorgegebenen Raten r_b sofort überprüfbar. Mit (5.9) läßt sich auch leicht testen, ob es überhaupt ein r_a gibt, mit dem (i) erfüllt ist. Wesentlich schwieriger dagegen ist es zu überprüfen, ob es ein r_a gibt das (i) *und* (ii) erfüllt. Sofern ein System voll bestimmt ist, brauchen wir nur das einzig mögliche r_a nach (5.5) zu berechnen und dann (i) und (ii) zu überprüfen. Doch wie geht man im unterbestimmten Fall vor (wie im Beispiel oben)?

Hier greifen wir auf einen Trick aus der linearen Optimierung zurück. Dieser liefert in der ersten Phase des Simplexalgorithmus eine *zulässige Lösung* (hier äquivalent mit *realisierbare*

Lösung) für vorgegebene Nebenbedingungen bzw. weist nach, daß eine zulässige Lösung gar nicht existiert [40]. Dazu definieren wir hier folgendes lineares Optimierungsproblem

$$\begin{aligned} \text{Nebenbedingungen: } \mathbf{N}_s \mathbf{r}_s + \mathbf{z} &= [\mathbf{N}_s \mathbf{I}] \begin{pmatrix} \mathbf{r}_s \\ \mathbf{z} \end{pmatrix} = \mathbf{b} = -\mathbf{N}_s \mathbf{r}_b \\ \alpha_i &\leq r_{s,i} \leq \beta_i \\ z_i &\geq 0 \end{aligned} \quad (5.21)$$

$$\text{Zielfunktion: } \sum_{i=1}^m z_i = z_1 + z_2 + \dots + z_m = \min!$$

\mathbf{z} ist zunächst ein beliebiger $(m,1)$ -Vektor, \mathbf{I} die (m,m) -Einheitsmatrix. Damit dieses Verfahren funktioniert, müssen alle Elemente im Vektor $\mathbf{b} = -\mathbf{N}_s \mathbf{r}_b$ nichtnegativ sein. Für alle $b_i < 0$ multipliziert man sonst b_i und die i -te Zeile in \mathbf{N}_s mit -1 und setzt dann diese in (5.21) ein, was keinen Einfluß auf die Lösung \mathbf{r}_s in (5.2) hat.

Zunächst einmal existiert immer eine zulässige Lösung $\begin{pmatrix} \mathbf{r}_s \\ \mathbf{z} \end{pmatrix}$ in (5.21), nämlich mit $\mathbf{r}_s = \mathbf{0}$ und $\mathbf{z} = \mathbf{b}$. Davon ausgehend wird nun die Zielfunktion minimiert. Ihr theoretisch minimalster Wert überhaupt ist 0 und tritt genau dann ein, wenn $\mathbf{z} = \mathbf{0}$ ist. Das wäre gleichbedeutend damit, daß das Flußszenario realisierbar ist, denn dann erfüllt das gleichzeitig ermittelte \mathbf{r}_s sowohl (i) als auch (ii) in (5.20). Ergibt sich also im Ergebnis dieser linearen Optimierung der Nullvektor für \mathbf{z} , ist das Szenario realisierbar und sonst nicht.

Dieser Test auf Realisierbarkeit kann sehr nützlich sein und ermöglicht z.B. inkonsistente Daten oder Netzmodelle zu identifizieren bzw. das Vorhandensein bestimmter Funktionalitäten im Netz zu überprüfen (siehe Beispiel in Abschnitt 5.3). Im *FluxAnalyzer* wurde dazu die Funktion „Check feasibility“ implementiert (Abb. 4.5), die ein vorgegebenes Szenario in einem Netz (unter Berücksichtigung definierter Ratenminima und -maxima) mit oben beschriebener Methode auf Realisierbarkeit überprüft.

5.5 Anwendungsbeispiel: Redoxbalance im Stoffwechsel der *Rhodospirillaceae*

Bisher gibt es in der Literatur kaum Beispiele, wo unterbestimmte Stoffwechselnetze untersucht worden sind. Eine Ausnahme stellt [100] dar, wo allerdings nur kumulierte Stöchiometrien verwendet wurden. Am realen Beispiel des phototrophen Stoffwechsels der schwefelfreien Purpurbakterien (*Rhodospirillaceae*; abgek.: SPB) soll hier gezeigt werden, daß u.U. auch hochgradig unterbestimmte Flußszenarien wichtige Schlußfolgerungen zulassen. In unserem Anwendungsfall werden wir – ohne eine gemessene Reaktionsrate – aufzeigen, wie essentiell das Problem der Redoxbalance in den SPB ist.

Bereits im Abschnitt 3.3 wurden wesentliche Aspekte des Stoffwechsels der SPB beschrieben und das stöchiometrische Modell vorgestellt. Letzteres wurde im *FluxAnalyzer* implementiert

und für die folgenden Berechnungen herangezogen (Abb. 4.7 und 4.8). Wir fokussieren hier zunächst auf photoheterotrophes Wachstum unter anaeroben Bedingungen im Licht mit Succinat als Substrat. Die entsprechende Situation im Stoffwechsel ist als Szenario S0 in Tab. 5.2 beschrieben. Wir können die Aufnahmeraten der anderen Substrate und die von Sauerstoff auf Null setzen. Wir gehen außerdem davon aus, daß außer CO₂ keine Nebenprodukte gebildet werden, was unter optimalen Bedingungen auch beobachtet wird [109]. Alle anderen Reaktionen sind im Netz freigegeben. In diesem Szenario sind zwar vier Reaktionsraten berechenbar, allerdings sind diese trivial (z.B. ist der Fluß durch die Oxidase Null, da kein Sauerstoff aufgenommen werden kann). Es verwundert nicht, daß keine weiteren Reaktionsraten berechnet werden konnten, da keine Angaben über die Aufnahmemenge des Substrats vorlagen.

Reaktion	S0	S1	S2	S2: Sensitivitäten bzgl. SuccCoA:olKG	S2: Sensitivitäten bzgl. G6P::PGluc	S3
<i>Frac_up</i>	0	0	0			0
<i>Mol_up</i>	0	0	0			0
<i>Succ_up</i>		1	1			1
<i>Prop_up</i>	0	0	0			0
<i>O2_up</i>	0	0	0			0
<i>N_up</i>		0.689	0.689	0	0	
<i>S_up</i>		0.016	0.016	0	0	
<i>Ac_up</i>	0	0	0			0
<i>CO2_exchange</i>		-0.805	-0.805	0	0	
<i>Succ_ex</i>	0	0	0			0
<i>Ac_ex</i>	0	0	0			0
<i>Form_ex</i>	0	0	0			0
<i>Hydrogenase</i>	0	0	0			0
<i>Nitrogenase</i>	0	0	0			0
μ		0.076	0.076	0	0	
<i>F1P::F16P</i>	0	0	0	0	0	0
<i>F6P::G6P</i>			0.013	0	1	
<i>F16P::F6P</i>						
<i>F6P::F16P</i>						
<i>F16P::T3P</i>						
<i>DHAP::G3P</i>			-0.557	2	-1	
<i>G3P::DPG</i>			-1.369	5	-2	
<i>DPG::3PG</i>			-1.369	5	-2	
<i>3PG::2PG</i>			-0.068	-1	0	
<i>2PG::PEP</i>			-0.068	-1	0	
<i>PEP::PYR</i>			0.694	-1	0	
<i>Pyruvate</i>			0			
<i>PYR::AcCoA</i>			0.509	-1	0	
<i>AcCoA::Cit</i>			0.071	-1	0	
<i>Cit::ICit</i>			0.071	-1	0	
<i>ICit::olKG</i>			0.071	-1	0	

Reaktion	S0	S1	S2	S2: Sensitivitäten bzgl. SuccCoA: aKG	S2: Sensitivitäten bzgl. G6P: PGIac	S3
aKG::SuccCoA			0			
SuccCoA::Succ			-0.042	-1	0	
Succ::Fam			0.993	-1	0	
Fam::Succ			0			
Fam::Mal			1.064	-1	0	
Mal::OxAc			1.064	-1	0	
CitMalSynth			0			
CitMal::Mes			0	0	0	
Mes::PropCoA		0	0	0	0	
Glyox::Mal			0	0	0	
Prop::PropCoA	0	0	0	0	0	0
PropCoA::MMalCoA		0	0	0	0	
MMalCoA::SuccCoA			0	0	0	
G6P::PGIac			0		[Veränderliche: +1]	
PGIac::PGIac			0	0	1	
PGIac::R15P			0	0	1	
R15P::X5P			-0.522	2	0	
R15P::R5P			-0.192	1	0	
R15P::R115P			0.714	-3	1	0
Rubisco			0.714	-3	1	0
S17PSynth						0
S17P::S7P						0
Transket1			-0.249	1	0	
Transaldol						
Transket2			-0.274	1	0	
OxAc::PEP			0.815	0	0	
Pyr::OxAc			0			
AcCoA::AcP			-0.036	0	0	
AcP::Ac			-0.036	0	0	
Pyr::Fom	0	0	0	0	0	0
Mal::Pyr			0			
SuccCoA::aKG			0	[Veränderliche: +1]		
AcCoA::Pyr			0			
CitLyase			0			
NADHDehydro			-0.993	1	0	
BCI						
Oxidore	0	0	0	0	0	
Photo						0
Transhydro			-1.346	1	-2	
NADH::Fd			0	1	0	
ATPSynth						
ATPdrain						

Tab. 5.2: Szenarien für Wachstum auf Succinat. Vorgegebene Raten: weiße Felder; berechnete Raten: graue Felder, nichtberechenbare Raten: leere Felder (Raten in [mmol/(gTM h)]).

5.5.1 Reduktionsgrad des Substrats, Biomasseausbeute und Netto-CO₂-Freisetzung

Aus diesem Grunde fixieren wir in S1 die Aufnahme rate von Succinat (r_{Succ_up} ; Reaktionsnamen nach Anhang A.3) zunächst auf 1 mmol/(gTM·h). Obwohl noch immer 13 (!) Freiheitsgrade im System vorliegen, sind nun zusätzlich sowohl die Netto-CO₂-Bilanz als auch die Wachstumsrate μ und mit ihr die Aufnahme rates von Sulfat und Ammonium, sowie die anabolen Flüsse zur Biomassesynthese berechenbar. Anabole Rates sind aus Platzmangel in Tab. 5.2 nicht aufgeführt, ergeben sich aber – ebenso wie für Sulfat- und Ammoniumaufnahme – immer proportional aus μ . Die anabolen Flüsse in S1 sind denen in Abb. 4.8 identisch.

Die Berechenbarkeit von μ und $CO_2_exchange$ ist schon etwas überraschend, gibt es doch in diesem stark unterbestimmten Netz viele Reaktionen, die CO₂ freisetzen oder fixieren (Calvin-Zyklus, reduktiver/oxidativer TCA, anaplerotische Reaktionen, oxidativer Pentosephosphat-Weg). Diese gefundene Struktureigenschaft des Netzes steht im engen Zusammenhang mit einer ausgeglichenen Kohlenstoff- und Redoxbalance und hat weitreichende Konsequenzen für den phototrophen Stoffwechsel der *Rhodospirillaceae*. Die berechneten Rates sind nämlich die einzig möglichen, bei denen sowohl die Pools der Kohlenstoff-Intermediate als auch die der Reduktionsäquivalente NAD(P)H simultan im Gleichgewicht gehalten und damit auf Dauer ohne Leer- oder Überlauf balanciert werden können. Die Freisetzung von CO₂ (negatives Vorzeichen von $CO_2_exchange$) deutet darauf hin, daß Succinat – gemessen am mittleren Redoxstatus der Kohlenstoffatome – oxidiert als die Biomasse ist. In [109,155] wurde bereits auf diesen experimentell beobachteten Zusammenhang hingewiesen.

Wir haben in S1 die Aufnahme von Succinat auf Eins fixiert, d.h. S1 stellt den normalisierten Fall dar. Ein veränderter Eingangsstrom von Succinat führt lediglich zu einer multiplikativen Anpassung der berechneten Rates, z.B.:

$$r_{CO_2_exchange} = -0.805 r_{Succ_up} \quad (5.22)$$

Die tatsächliche Aufnahme rate r_{Succ_up} hängt im wesentlichen von der Rate der Photosynthese ab, da diese im phototrophen Stoffwechsel den Hauptanteil der benötigten Energie (ATP) für Synthese- und Erhaltungsprozesse generieren muß. Die ATP-Bilanz selbst hat in S1 *keinen* Einfluß auf die Bestimmtheit der berechneten Rates: setzt man ATP in S1 auf *extern* (d.h. ohne Zeile in N und damit nicht bilanziert), ergibt sich das gleiche Ergebnis. Setzt man dagegen NADH oder NADPH auf *extern*, wird die Berechenbarkeit der ermittelten Rates aufgehoben. Der erzeugte metabolische Strukturzwang ist also eng mit der Redoxbalance verknüpft. Die in den katabolen Prozessen vom Substrat freigesetzten und in den Reduktionsäquivalenten NAD(P)H zwischengespeicherten Elektronen müssen unter anaeroben Bedingungen letztlich vollständig in die Biomasse abfließen (sofern keine weiteren Produkte ausgeschieden werden). Unter aeroben Bedingungen ändert sich das, denn dort – im Gegensatz zur Photosynthese – fließen bei der ATP-produzierenden oxidativen Phosphorylierung Elektronen (auf

Sauerstoff) ab. Bei der Atmung ist somit die Redoxbilanz eng mit der Energiebilanz gekoppelt (Abschnitt 5.5.6), was bei der Photosynthese nicht gilt.

Wir können Szenario S1 analog für andere Substrate durchführen, jeweils mit normalisierter Aufnahme- und Freisetzung des betrachteten Substrates. An der Berechenbarkeit der Raten ändert sich dabei nichts, allerdings ergeben sich veränderte Werte für CO_2 -Freisetzung und μ (Tab. 5.3). Der oben angedeutete Zusammenhang zwischen CO_2 -Freisetzung und Oxidationsstatus des Substrates ist wiederum klar erkennbar. Die größte Freisetzung an CO_2 und auch der größte relative Verlust an Kohlenstoff ist für das am stärksten oxidierte Substrat, nämlich Malat, festzustellen. Beim stark reduzierten Propionat dagegen wird in der Nettobilanz CO_2 sogar fixiert, d.h. Wachstum auf Propionat kann nur erfolgen, wenn zusätzlich CO_2 (bzw. Bicarbonat) angeboten wird. Diese Aussage wurde in Experimenten klar bestätigt [108,125]. Die stärkste Fixierung von Kohlendioxid ist – wie zu erwarten – bei autotrophen Wachstum auf einem H_2/CO_2 -Gemisch notwendig, da hier der gesamte Kohlenstoff für die Biomasse vom CO_2 bereitgestellt werden muß (H_2 -Aufnahme über *Hydrogenase*).

	Succinat	Malat	Acetat	Fructose	Propionat	Wasserstoff (autotroph)
<i>CO₂-Nettobilanz</i>						
berechnet [mmol CO_2 / mmol Subst.]	-0.805	-1.262	-0.175	-0.524	0.195	0.456
gemessen [mmol CO_2 / mmol Subst.]	-0.74 ⁽¹⁾	-1.15 ⁽¹⁾	-0.25 ⁽¹⁾		0.23 ⁽¹⁾	
berechnet [100% - mmol CO_2 / mmol C im Subst.]	-20.1 %	-31.5%	-8.8%	-8.7%	46.5%	∞
<i>Biomasseertrag Y_{X/S}</i>						
berechnet [gTM / mmol Subst.]	0.076	0.066	0.044	0.131	0.076	0.011
berechnet [gTM / g Subst.]	0.644	0.493	0.733	0.727	1.027	5.5
gemessen [gTM / g Subst.]	0.655 ⁽¹⁾	0.476 ⁽¹⁾	0.660 ⁽¹⁾	0.717 ⁽¹⁾		

Tab. 5.3: Berechnete (grau unterlegt) und experimentell gemessene Werte der CO_2 -Nettobilanz (negativ: Freisetzung, positiv: Fixierung) und des Biomasseertrages bei phototrophen Wachstum der *Rhodospirillaceae* auf unterschiedlichen Substraten. Referenzen für gemessene Werte: ⁽¹⁾ nach [108] (Spezies: *R. rubrum*); ⁽²⁾ nach [19] (*R. sphaeroides*); ⁽³⁾ nach [35] (*R. rubrum*); ⁽⁴⁾ nach eigenen Experimenten mit *R. rubrum* [52].

In Tab. 5.3 sind auch Meßdaten aus der Literatur bzw. aus eigenen Experimenten angegeben. Ein Vergleich mit den berechneten CO_2 -Bilanzen zeigt für alle Substrate eine gute Übereinstimmung. Eine leichte Tendenz zur überschätzten Freisetzung von CO_2 ist erkennbar, die auf eine vom Modell abweichende Biomassezusammensetzung rückschließen läßt, d.h. in den Experimenten in [108] war die Biomasse vermutlich etwas oxidiert. Setzt man im *FluxAnalyzer* bspw. den Anteil der Lipide etwas herab und dafür andere (mehr oxidierte) entsprechend herauf, bekommt man eine noch bessere Übereinstimmung bzgl. der CO_2 -Bilanz.

Da auch die Wachstumsrate berechenbar ist, können wir sofort den Biomasseertrag $Y_{X/S}$ in [gTM / mmol Substrat] aus dem Quotienten von μ und der Substrataufnahmerate berechnen.

Teilen wir noch durch das Molekulargewicht des Substrates erhalten wir die Einheit [gTM / g Substrat] (Tab. 5.3). Bei Wachstum auf Propionat ist durch die Nettofixierung von Kohlenstoff die Ausbeute pro Gramm Substrat größer Eins ist. Diese Ausbeute wird nur noch beim autotrophen Wachstum mit molekularem Wasserstoff übertroffen. Die theoretisch bestimmten Werte für Y_{XS} entsprechen den maximal möglichen Ausbeuten die nur dann erreicht werden können, wenn keine weiteren Produkte ausgeschieden werden. Tatsächlich scheint das in der Realität der Fall zu sein, stimmen doch auch hier berechnete Werte gut mit Meßdaten überein.

5.5.2 Bedeutung des Calvin-Zyklus für die Redoxbalance

Im Szenario S2 sollen die Freiheitsgrade für (normiertes) photoheterotrophes Wachstum auf Succinat weiter eingeschränkt werden. Wir beziehen uns dabei zunächst konkret auf *Rhodospirillum rubrum*, wo folgende zusätzlichen Annahmen zu S1 laut angegebenen Literaturquellen sinnvoll erscheinen:

- kein oxidativer Pentosephosphat-Weg in *R. rubrum* wegen fehlender Glucose-6-phosphat-Dehydrogenase [53,137] $\rightarrow G6P::PGlac = 0$
- geringe oder keine Enzymaktivitäten wurden für *R. rubrum* bei Wachstum auf Malat für folgende Enzyme festgestellt [67]: Malat-Enzym, Pyruvat-Carboxylase, PEP-Synthase $\rightarrow Mal::Pyr = Pyr::Ox = Pyr::PEP = 0$
- die drei charakteristischen Enzyme des reduktiven TCAs zeigen relativ geringe Aktivitäten und sind vermutlich evolutionäre „Überreste“ [67,155]; wir nehmen deshalb zunächst an, daß sich die Flüsse durch die (oxidative und unter anaeroben Bedingungen reprimierte) α -Ketoglutarat-Dehydrogenase und die (reduktive) α -Ketoglutarat-Synthetase neutralisieren und die beiden anderen Reaktionen des reduktiven TCA eine vernachlässigbare Rate aufweisen $\rightarrow SuccCoA::aKG = aKG::SuccCoA = CitLyase = AcCoA::Pyr = 0$
- Citramalat-Zyklus nicht aktiv (keine Acetat-Verwertung) $\rightarrow CitMalSynth = 0$
- da ein Nettofluß von Succinat zu Fumarat unter diesen Bedingungen stattfinden muß setzen wir den Fluß durch die Fumarat-Reductase auf Null $\rightarrow Fum::Succ = 0$

Mit diesen Annahmen werden nun viele weitere Raten berechenbar, obwohl das System noch immer drei Freiheitsgrade hat (Tab. 5.2; das Szenario S2 ist außerdem in Abb. 4.7 und 4.8 dargestellt). Wie zu erwarten ergeben sich für die in S1 determinierten Raten (z.B. für μ und $CO_2_exchange$) exakt die gleichen Werte, schließlich haben wir S1 in S2 nur weiter eingeschränkt. Einige andere interessante Aspekte fallen auf. So ist ein beträchtlicher Fluß durch die vom Enzym Rubisco katalysierte Schlüsselreaktion (*Rubisco*) des Calvin-Zyklus festzustellen, bei der CO_2 fixiert wird. Das ist zunächst erstaunlich, hatten wir doch bereits in Szenario S1 eine wesentliche Nettofreisetzung von CO_2 berechnet. Hier können nun frühere Spekulationen [125,155] quantitativ belegt werden: Die Reflexierung eines Teils des freigesetzten CO_2 (z.B. aus den Reaktionen $Ox::PEP$ und $Pyr::AcCoA$) dient nicht nur der Minderung des

Verlusts von Kohlenstoff, sondern ist als Senke für überschüssige Reduktionsäquivalente sogar *notwendig*. Das wird am (negativen) Fluß durch die Reaktion *G3P::DPG* deutlich, die in dieser Richtung Bestandteil des Calvin-Zyklus ist und NADH konsumiert. Auch *Transket1* und *Transket2* zeigen mit ihren Raten in Richtung des Calvin-Zyklus.

Warum werden überhaupt überschüssige Reduktionsäquivalente erzeugt? Zur Einspeisung des Succinats in den Zentralstoffwechsel wird ein Großteil des aufgenommenen Substrates zunächst mittels Succinat-Dehydrogenase (*Succ::Fum*) in Fumarat überführt (nur ein kleiner Anteil wird zum benachbarten Präkursor *SuccCoA* geleitet). Dabei werden auf der Stufe des Ubichinons Elektronen in die membranständige Elektronentransportkette eingebracht. Netto muß genau die gleiche Menge Elektronen wieder herausgeführt werden, da die zyklische Photosynthese keine Elektronen verbraucht. Dies kann mit dem energiegetriebenen, umgekehrten Transport der Elektronen (Redoxskala bergauf, Kap. 9) vom Ubichinon auf NAD erfolgen. In der Tat wird für *NADHDehydro* ein negativer (umgekehrter) Fluß berechnet. Zwei weitere Elektronen werden von NAD bei der Malat-Dehydrogenase aufgefangen. Erst dann kann ein Großteil des Kohlenstoffs über die PEP-Carboxykinase (*Oxa::PEP*) aus dem TCA abfließen, wobei CO₂ freigesetzt wird. Bereits an dieser Stelle sind 2 mol NADH produziert worden, die in der Gesamtbilanz zu einem Überschuß beitragen.

Damit werden also pro mmol aufgenommenes Succinat insgesamt 1.519 mmol CO₂ freigesetzt und davon 0.714 mmol im Calvin-Zyklus erneut fixiert (was zur negativen Nettobilanz von 0.805 mmol führt, Tab. 5.3). Der berechnete Wert für den refixierten Anteil des CO₂ stimmt relativ gut mit der Angabe von 0.5 mmol in [109] überein.

	Succinat	Malat	Acetat	Fructose	Propionat	Wasserstoff (autotroph)
Rate <i>Rubisco</i> [mmol / mmol Subst.]	0.714	0.183	0.186	0.367	0.714	0.531
Rate <i>Rubisco</i> [mmol / mmol C im Subst.]	0.179	0.046	0.093	0.062	0.238	se

Tab. 5.4: Berechnete Raten für die Schlüsselreaktion des Calvin-Zyklus („*Rubisco*“), normiert auf die Aufnahme rate des Substrates bzw. auf die Aufnahme rate von (molarem) Kohlenstoff im Substrat.

Szenario S2 können wir wie bei S1 analog auch wieder für die anderen Substrate durchführen (bei Acetat wird *Pyr::AcCoA*=0 gesetzt und dafür der vermutliche Aufnahmeweg über die Citramalat-Synthase (*CitMalSynth*) „geöffnet“ [7,64]). Bei allen betrachteten Substraten ergibt sich ein Fluß durch den Calvin-Zyklus (Tab. 5.4 zeigt die relativen Flüsse durch die Reaktion *Rubisco*). Dieser ist bei Malat am geringsten, nicht zuletzt weil es das am stärksten oxidierte Substrat ist (vgl. mit Tab. 5.3). Auf der anderen Seite ist bei Succinat, obwohl oxidiertes als Fructose, ein größerer Fluß durch den Calvin-Zyklus notwendig, erst recht wenn man auf die mit dem Substrat aufgenommenen Kohlenstoffatome (Tab. 5.4) normiert. Zu bemerken wäre außerdem noch, daß bei Propionat, wo eine positive CO₂-Bilanz zu verzeichnen war, ein Großteil der Fixierung über das Enzym Propionyl-CoA-Carboxylase läuft (Reaktion *Prop-CoA::MMalCoA*). Bei lithoautotrophen Wachstum mit Wasserstoff übernimmt der Calvin-

Zyklus nicht nur die Rolle der Elektronensenke, sondern stellt gleichzeitig die einzige Quelle organischer Kohlenstoffverbindungen dar. Der in Tab. 5.4 angegebene Fluß für autotrophes Wachstum erscheint niedrig, ist allerdings auf verwerteten Wasserstoff normiert. Um die gleiche Menge *Biomasse* wie z.B. auf Succinat zu produzieren, muß der Fluß durch *Rubisco* fünfmal so hoch sein wie bei Succinat.

5.5.3 Zusammenhang zwischen Calvin-Zyklus und reduktivem/oxidativem TCA

Wir hatten in S2 den Fluß durch $aKG::SuccCoA$, eine Reaktion des reduktiven TCAs, auf Null gesetzt, da die katalysierende α -Ketoglutarat-Synthetase (α -KGS) in *R. rubrum* nur mit relativ niedrigen Aktivitäten nachgewiesen worden ist. Wir wollen nun dennoch hinterfragen, welche Auswirkungen ein Fluß durch diese Reaktion, bei der CO_2 fixiert und eine Reduktionäquivalente konsumiert wird, auf die anderen Flüsse hätte. Als geeigneten Ansatz können wir dazu die Sensitivitäten der berechneten Raten in S2 bzgl. $SuccCoA::aKG$ nach Gl. (5.10) bestimmen. Dabei stellt sich heraus, daß nur einige Raten überhaupt sensitiv auf diesen Fluß reagieren (Tab. 5.2). Da die Rate von $SuccCoA::aKG$ in S1 noch „frei“ wählbar war, sind insbesondere die in S1 berechenbaren Raten insensitiv bzgl. dieser Rate. Sensitiv reagieren dagegen die Flüsse im TCA, die nun stärker in die reduktive Richtung, d.h. entgegen dem Uhrzeigersinn getrieben werden (bzw. oxidative Flüsse wie durch $Mal::Ox$ nehmen ab). Das ist intuitiv klar, dient doch $SuccCoA::aKG$ gerade einem reduktive(re)n TCA. Eine zweite wesentliche Umverteilung erfolgt in Reaktionen des Calvin-Zyklus. Dessen Bedeutung zur Beseitigung überschüssiger Redoxkraft nimmt nun nämlich relativ drastisch ab: ein $mmol$ Umsatz in $SuccCoA::aKG$ spart bspw. die dreifache Menge in der Reaktion *Rubisco* ein (Tab. 5.2). Der Fluß durch $G3P::DPG$ in NADH-verbrauchender Richtung wird sogar um die fünffache Menge reduziert. Das zeigt, wie stark der reduktive TCA den Calvin-Zyklus entlasten kann, auch wenn vermutlich nur geringe Raten in *R. rubrum* möglich sind.

Diese Erkenntnis läßt sogar eine Hypothese für ein beobachtetes Phänomen zu: Während eine *Rubisco*-Mutante von *R. rubrum* noch in der Lage ist, photoheterotroph auf Malat zu wachsen, gilt das für die *Rubisco*-Doppelmutante von *R. sphaeroides* nicht (Doppelmutante, da diese Spezies zwei Formen des Enzyms *Rubisco* hat) [33,90,168]. Für Wachstum auf Malat hatten wir in S2 eine relativ geringe Rate für *Rubisco* berechnet (≈ 0.18). Auch bei Wachstum

auf Malat ist die Sensitivität $\frac{dr_{rubisco}}{dr_{SuccCoA::aKG}} = -3$, d.h. mit $r_{SuccCoA::aKG} = 0.06$ könnte *R. rubrum*

den fehlenden Fluß durch *Rubisco* ausgleichen. In *R. sphaeroides* dagegen ist das nicht möglich, da das Enzym α -KGS noch nicht einmal mit geringen Aktivitäten und auch nicht im Genom nachgewiesen werden konnte (Abschnitt 3.3).

Mit einer Realisierbarkeitsanalyse nach Abschnitt 5.3 läßt sich sogar zeigen, daß bei gleichzeitigem Ausfall der CO_2 -fixierenden Enzyme *Rubisco* und α -KGS Wachstum auf Succinat

gar nicht möglich wäre. Selbst im viel weniger eingeschränkten Szenario S1 würde keine realisierbare Flußverteilung existieren, setzte man beide Raten auf Null. Zu diesem Ergebnis käme man durch „einfache“ MFA nicht, da in diesem modifizierten Szenario $\bar{S}1$ dieselben Raten wie in S1, sowie nur eine weitere Rate ($R15P::R15P$) berechenbar sind. In deren berechneten Werten wird die vorhandene Inkonsistenz des System nicht sichtbar.

Die für $SuccCoA::aKG$ berechneten Sensitivitäten stellen mit negativem Vorzeichen die Sensitivitäten für die entgegengesetzt laufende Reaktion $aKG::SuccCoA$ der α -Ketoglutarat-Dehydrogenase dar (einzige Ausnahme: Sensitivität bzgl. $NADH:Fd$ ist Null, da statt reduziertem Ferredoxin hier $NADH$ als Reduktionsäquivalente genutzt wird). Wird der TCA also stärker in oxidativer Richtung betrieben, dann muß der Calvin-Zyklus mit rapide steigendem Fluß die dabei freigesetzten Redoxäquivalente auffangen. Das ist vermutlich der Grund, warum der TCA an dieser Stelle unterbrochen bzw. sogar in reduktiver Richtung betrieben wird. Die Reprimierung der α -Ketoglutarat-Dehydrogenase unter anaeroben Verhältnissen ist auch bei anderen Bakteriengruppen eindeutig nachgewiesen [51].

Mit der unter anaeroben Bedingungen exprimierbaren Nitrogenase erkannte man eine weitere Möglichkeit, wie die SPB unter stark reduzierten Bedingungen den Calvin-Zyklus von seiner Rolle als Redoxsenke „entlasten“ können. Die Nitrogenase benötigt zur Fixierung von Stickstoff nämlich Redoxäquivalente und setzt Wasserstoff als Nebenprodukt frei [66]. Berechnet man die Sensitivitäten der Raten in S2 bzgl. *Nitrogenase* so stellen wir z.B. für *Rubisco* in der Tat einen negativen Wert fest (nicht gezeigt). Allerdings ist die Nitrogenase sehr energieaufwendig und führt durch den Verlust von Elektronen auch zu einer reduzierten Biomasseausbeute. Außerdem wirkt Ammonium reprimierend auf ihre Expression [66].

5.5.4 Calvin-Zyklus, Transhydrogenase und oxidativer Pentosephosphat-Weg

In Szenario S2 hatten wir den Fluß im oxidativen Pentosephosphat-Weg (oxPPW) wegen der offensichtlich nicht vorhandenen Glucose-6-phosphat-Dehydrogenase in *R. rubrum* auf Null gesetzt. Da der oxPPW bekanntermaßen in vielen Organismen als $NADPH$ -Quelle dient, stellt sich hier die Frage, wie die für die anabolen Reaktionen benötigte Elektronenquelle generiert wird. Als eine Möglichkeit kommt die Isocitrat-Dehydrogenase ($ICit::aKG$) in Frage. Allerdings kann unter anaeroben Bedingungen kaum eine Nettoproduktion erfolgen, da wie oben angedeutet der oxidative TCA nur schwach betrieben wird bzw. beim rückwärtslaufenden reduktiven TCA sogar $NADPH$ verbraucht wird. Tatsächlich bleibt in S2 nur noch die Transhydrogenase übrig, die aus $NADH$ mit geringem Energieeinsatz $NADPH$ produziert. In S2 wurde für *Transhydro* folglich auch ein relativ hoher Fluß berechnet.

In *R. sphaeroides* und *R. capsulatus* wurde dagegen die Existenz des oxPPW eindeutig nachgewiesen [21]. Wie ändert sich die Lage in S2, wenn ein Fluß durch den oxPPP führt? Die Sensitivitätsanalyse bzgl. $G6P::PGlac$ gibt Aufschluß darüber (Tab. 5.2): In der Tat wird der

Fluß durch *Transhydro* reduziert und zwar mit einem Proportionalitätsfaktor von -2 (α PPW generiert zwei NADPH). Allerdings können wir auch feststellen, daß der α PPW einen höheren Fluß durch den Calvin-Zyklus verursacht (Faktor 1 für *Rubisco*). Wird also der α PPW zur Produktion von NADPH herangezogen, wird das dabei freigesetzte CO_2 ebenfalls über den Calvin-Zyklus aufwendig refixiert. Letztlich kann man sich die gemeinsame Operation von α PPW und Calvin-Zyklus als eine „virtuelle Transhydrogenase“ vorstellen (Abb. 5.2). Allerdings müssen dabei – im Vergleich zur Transhydrogenase – erhebliche Flüsse umgesetzt werden und die Nettoreaktion $2 \text{ NADH} + 3 \text{ ATP} \rightarrow 2 \text{ NADPH}$ ist energetisch ungünstig. Aus diesen Gründen können wir hier die Vermutung anstellen, daß der α PPW unter photoheterotrophen Bedingungen, ähnlich wie in Pflanzen [117], weitgehend reprimiert ist.



Abb. 5.2: „Transhydrogenase“-Funktion beim Zusammenspiel von α PPW und Calvin-Zyklus

Auf jeden Fall können wir hier festhalten, daß die Transhydrogenase eine wichtige Rolle bei der Redoxbalance unter phototrophen Bedingungen spielt. Bereits in [90] wurde darauf hingewiesen und Messungen in *R. rubrum* zeigen, daß unter phototrophen Bedingungen eine deutlich höhere Transhydrogenase-Aktivität zu verzeichnen ist, als unter aeroben [53]. Unter aeroben Bedingungen wird der TCA klar oxidativ betrieben, so daß NADPH in großen Mengen auch über die Isocitrat-Dehydrogenase bereitgestellt werden kann.

Eine Alternative für die NADPH-Synthese stellt generell noch das Malat-Enzym dar, dessen Reaktionsrate in S2 auf Null gesetzt war. Die Sensitivität von *Transhydro* bzgl. *Mal::Pyr* in S2 ist -1 . Auf *Rubisco* hätte eine Verschiebung dieser Flüsse keinen Einfluß.

5.5.5 Verallgemeinerte Berechenbarkeitsanalyse für phototrophen Wachstum

An zwei Beispielen soll die in Abschnitt 5.2 eingeführte verallgemeinerte Berechenbarkeitsanalyse illustriert werden. Die elementare (kombinatorische) Repräsentation des Nullraums \mathbf{D}_s in Szenario S1 besitzt 132198 Elemente (Spalten). Sie werden als Elementarmoden der stöchiometrischen Matrix \mathbf{N}_s der Reaktionen mit unbekanntem Raten von S1 bestimmt, wobei man alle Reaktionen als reversibel betrachtet. Die hohe Anzahl drückt die starke Unterbestimmtheit von S1, d.h. die große Dimension des Nullraums von \mathbf{N}_s aus. Ein Element von \mathbf{D}_s ist (es werden nur Reaktionen mit stöchiometrischen Koeffizienten ungleich Null angegeben):

$$\{4 \text{ PEP}::\text{Pyr}, 4 \text{ Pyr}::\text{OxA}, 4 \text{ OxA}::\text{PEP}, 3 \text{ Photo}, 3 \text{ BCI}, 4 \text{ ATPSynth}\} \quad (5.23)$$

Hierbei handelt es sich um einen bekannten *futile cycle* aus anaplerotischen Reaktionen um Pyruvat und PEP, der seinen Energieverbrauch hier durch die Photosynthese speist. Da in (5.23) keine Reaktion die Systemgrenze überschreitet, wäre wohl keine der zugehörigen Raten (ohne ^{13}C Experimente) meßbar. Nach Abschnitt 5.2 könnten wir damit eigentlich keine Rate der in (5.23) enthaltenen Reaktionen mittels MFA bestimmen, selbst wenn alle anderen Raten im System meßbar wären. Erst durch die (begründete) Annahme $\text{Pyr}::\text{OxA} = 0$ in S2 fällt (5.23) aus dem Nullraum heraus, was die anderen Reaktionsraten in S2 überhaupt berechenbar macht.

Ein anderes Element von \mathbf{D}_0 in S1, das auch durch die zusätzlichen Annahmen in S2 nicht betroffen ist und deshalb als eines von 6 Elementen in \mathbf{D}_0 bzgl. S2 erhalten bleibt, ist

$$\{3 \text{ Photo}, 3 \text{ BCI}, 4 \text{ ATPSynth}, 4 \text{ ATPdrain}\} \quad (5.24)$$

Dieser Elementarmodus kennzeichnet die Bereitstellung von ATP für wachstumsunabhängige Erhaltungsprozesse, wobei die ATP-Synthese über zyklische Photosynthese erfolgt. Bereits oben hatten wir erwähnt, daß die ATP-generierende Photosynthese, im Gegensatz zur Respiration, keine Reduktionsäquivalente benötigt oder verbraucht, was hier nochmals deutlich wird. Nur wenn man den ATP-Bedarf für Erhaltungsprozesse quantitativ abschätzen kann, sind die Raten der in (5.24) involvierten Reaktionen für phototrophe Bedingungen mit MFA bestimmbar. Unter aeroben Bedingungen im Dunkeln können wir allerdings $\text{Photo}=0$ setzen, wodurch (5.24) aus dem resultierenden Nullraum herausfällt. Damit wäre die Berechenbarkeit der anderen drei Reaktionen in (5.24) zumindest erst einmal theoretisch wieder möglich. Dieses Beispiel zeigt, daß die Berechenbarkeit einer Rate auch von den jeweiligen Umweltbedingungen abhängen kann.

5.5.6 Chemotrophes Wachstum unter aeroben Bedingungen

Die *Rhodospirillaceae* wachsen sehr gut auch respirativ unter aeroben Bedingungen im Dunkeln. Können wir für eine normalisierte Substrataufnahme, ähnlich wie in Szenario S1, bereits Strukturzwänge ableiten? Dazu „öffnen“ wir im Szenario S3 die Sauerstoffzufuhr (O_2_{up}) und setzen – physiologisch untermauert – spezifische Reaktionen des aerob reprimierten Calvin-Zyklus (S17PSynth, Rubisco), sowie der Photosynthese *Photo* auf Null. Obwohl sogar weniger Freiheitsgrade (11) verbleiben als in S1 (13) kann uns S3 praktisch keine weiteren Erkenntnisse liefern. Selbst wenn wir die in S2 zusätzlich fixierten Raten (mit Ausnahme der unter aeroben Bedingungen für den oxidativen TCA verwendeten Reaktion $\text{aiKG}::\text{SuccCoA}$) auch hier mit Null vorbelegen würden, wäre, außer einigen offensichtlichen Nullraten, nichts bestimmbar.

Die Ursache der stark eingeschränkten Berechenbarkeit in diesem Szenario liegt, wie bereits in Abschnitt 5.5.1 erläutert, in der nun durch die Respiration gegebenen Kopplung zwischen Redox- und Energiehaushalt. In diesem Szenario bleibt unbestimmt, wieviel des aufgenommenen Substrates zur Bereitstellung von NADH für die Respiration verwendet wird, und wieviel für die Synthese von Biomasse verbleibt.

6 Kombinatorische Elementarmodenanalyse

In Kap. 2 wurde die grundlegende Bedeutung der Elementarmoden für die Metabole Pathwayanalyse erläutert. Eine typische Anwendung ist das Erkennen qualitativ unterschiedlicher Verwertungswege im Netz. Jedem EM können quantitative Eigenschaften zugeordnet werden (wie Ausbeutekoeffizient oder Anzahl involvierter Reaktionen/Metabolite), die einen Vergleich von EMen bzgl. verschiedener Kriterien ermöglichen.

Hier wollen wir einen Schritt weitergehen und die gesamte Menge der Elementarmoden hinsichtlich verschiedener kombinatorischer und statistischer Beziehungen untersuchen. Damit werden wir z.B. die Relevanz (Bedeutung) von Reaktionen für verschiedene Umweltbedingungen abschätzen können. Zunächst benötigen wir einige formale Vereinbarungen.

6.1 Spezifische Selektion einer Teilmenge von Elementarmoden

Umweltbedingungen oder bestimmte Zustände eines Stoffwechselnetzes werden wir im folgenden jeweils mit einer Menge S von einfachen Booleschen Regeln spezifizieren. Diese können bzgl. einer Reaktion i zwei mögliche Fälle kennzeichnen:

(a) die Reaktion i wird genutzt (Reaktionsrate $r_i \neq 0$): $S = \{i\}$.

(b) eine Reaktion i wird nicht genutzt ($r_i = 0$), wofür wir $S = \{-i\}$ schreiben.

Wird bzgl. einer Reaktion i keine Spezifikation gegeben, so kann diese genutzt werden (man beachte den wesentlichen Unterschied zu (a)). Wie im Kap. 5 nennen wir S auch ein Szenario, wobei im Unterschied zu Kap. 5 keine absoluten Werte angegeben werden. In NB2 (Abb. 2.3) repräsentiert z.B. $S1 = \{-R1\}$ die Bedingung, daß R1 nicht genutzt wird. Je nach Kontext charakterisieren wir damit eine Umweltbedingung (Substrat A liegt nicht vor) oder einen Stoffwechselzustand (Reaktion R1 wird nicht genutzt oder ist reprimiert). Man kann auch Kombinationen solcher Regeln konstruieren, die dann alle gleichzeitig gelten sollen (UND-Verknüpfung). So spezifiziert $S2 = \{-R2, R4\}$ in NB2 den Zustand ohne Aufnahme/Ausscheidung von B und mit Ausscheidung von E (über R4). Schließlich lassen wir im allgemeinsten Fall auch logische Kombinationen solcher Szenarien zu, wie $S1$ ODER $S2$.

Bezeichnen wir mit $E = \{e^1, e^2, \dots\}$ die Menge der EMen eines Netzes, so repräsentiert E (nicht-kursiv) eine Matrix mit den EMen als Spalten. Matrix E hat folglich q Zeilen (Reaktionen) und $|E|$ viele Spalten. Mit der Menge E^S bzw. Matrix E^S identifizieren wir die Teilmenge all derjenigen EMen, die eine vorgegebene Menge S an Spezifikationen erfüllt. So wäre $E^{S2} = E^{\{-R2, R4\}} = \{EM3, EM4\}$ (vgl. Abb. 2.9). E^S bzw. E^S lassen sich unter Ausnutzung der Eigenschaft (2.16) leicht aus dem Muster, d.h. aus der Binärdarstellung der EMen ermitteln, die als Matrix E_{EM} abgelegt werden kann. Dabei ist $E_{EM,ik} = 1$, falls Reaktion i im EM k genutzt wird ($e_i^k = E_{ik} \neq 0$) und sonst $E_{EM,ik} = 0$.

6.2 Flexibilität des Netzes und Relevanz einer Reaktion für vorgegebene Umweltbedingungen

6.2.1 Relative Auftrittshäufigkeiten von Reaktionen und Flexibilität eines Szenarios

In diesem Abschnitt betrachten wir alle EMen eines Netzes mit der gleichen Gewichtung, d.h. keinem EM wird eine (physiologisch) höhere oder niedrigere Bedeutung zugemessen. Bereits mit Gl. (2.23) hatten wir erwähnt, daß jede realisierbare stationäre Flußverteilung \mathbf{r} im Stoffwechsellnetz durch eine nichtnegative Linearkombination aller EMen darstellbar ist:

$$\mathbf{r} = \sum_{j=1}^{|\mathcal{E}|} \gamma_j \mathbf{e}^j = \mathbf{E}\boldsymbol{\gamma}; \quad \gamma_j \geq 0 \text{ für alle } j \in \{1, \dots, |\mathcal{E}|\} \quad (6.1)$$

Liegt mindestens ein EM vor, können wir mit Gl. (6.1) keine eindeutige Flußverteilung \mathbf{r} ableiten (schon ein einzelner EM wäre skalierbar und damit \mathbf{r} nicht mehr eindeutig). Auch wenn \mathbf{r} bekannt ist, muß $\boldsymbol{\gamma}$ in (6.1) nicht eindeutig bestimmt sein, d.h. mehrere Zerlegungen von \mathbf{r} in EMen sind möglich. Zwei wichtige Aspekte ergeben sich dennoch aus (6.1):

Bedeutung (Relevanz) einzelner Reaktionen

Gl. (6.1) legt folgendes Verfahren zur Abschätzung der Relevanz einer Reaktion i nahe: Je häufiger Reaktion i in den EMen auftritt (d.h. wenn $e_i^k = \mathbf{E}_{ik} \neq 0$ für viele k), desto wahrscheinlicher ist es, daß Reaktion i mit einer entsprechenden Rate $r_i \neq 0$ in einer stationären Flußverteilung \mathbf{r} involviert ist. Um die Relevanz der Reaktionen so abzuschätzen, berechnen wir also ihre *relativen Auftrittshäufigkeiten* \mathbf{h} in den EMen, die leicht aus der Binärdarstellung der EMen ermittelt werden kann:

$$\mathbf{h} = \frac{1}{|\mathcal{E}|} \mathbf{E}_{\text{bin}} \mathbf{1} = \frac{1}{|\mathcal{E}|} \mathbf{E}_{\text{bin}} \begin{pmatrix} 1 \\ \dots \\ 1 \end{pmatrix} \quad (6.2)$$

Dabei ist $\mathbf{1}$ ein Vektor mit $|\mathcal{E}|$ vielen Einsen. h_i kann auch als *Auftrittswahrscheinlichkeit* der Reaktion i in den EMen betrachtet werden: $h_i = P(e_i^k \neq 0)$.

Will man spezifische Umwelt- oder Randbedingungen analysieren, definiert man ein Szenario S , das die zulässige Menge der EMen auf \mathcal{E}^S einschränkt, d.h. für (6.1) erhält man:

$$\mathbf{r} = \mathbf{E}^S \boldsymbol{\gamma}^S; \quad \gamma_j^S \geq 0 \text{ für alle } j \in \{1 \dots |\mathcal{E}^S|\} \quad (6.3)$$

Mit oben eingeführter Notation ermitteln sich dann die relativen Auftrittshäufigkeiten \mathbf{h}^S bzgl. S analog aus der Binärdarstellung $\mathbf{E}_{\text{bin}}^S$ der selektierten EMen:

$$\mathbf{h}^S = \frac{1}{|\mathcal{E}^S|} \mathbf{E}_{\text{kin}}^S \mathbf{1} = \frac{1}{|\mathcal{E}^S|} \mathbf{E}_{\text{kin}}^S \begin{pmatrix} 1 \\ \dots \\ 1 \end{pmatrix} \quad (6.4)$$

Vektor $\mathbf{1}$ enthält hier $|\mathcal{E}^S|$ viele Einsen. Das Element h_i^S kann auch als *bedingte* Auftretswahrscheinlichkeit der Reaktion i verstanden werden: $h_i^S = \text{Pr}(e_i^k \neq 0 \mid S)$.

Diskutieren wir im folgenden über Szenarien, so ist der Spezialfall der leeren Spezifikation $S = \{\}$, wo alle EMen des Netzes angesprochen werden, mit eingeschlossen.

Ein wichtiger Extremfall sind *essentielle* Reaktionen, für die $h_i^S = 1$ gilt. Eine essentielle Reaktion ist in allen selektierten EMen involviert und deshalb auch an jeder nichttrivialen Flußverteilung \mathbf{r} (die S genügt) beteiligt. Entfernen einer essentiellen Reaktion i ($S' = \{S, -i\}$) „zerstört“ wegen (2.16) alle EMen und damit alle möglichen Flußverteilungen \mathbf{r} bzgl. des Szenarios S in Gl. (6.3). Sofern einer essentiellen Reaktion genau ein Enzym und dessen Gen(e) zugeordnet sind, können letztere ebenfalls als essentiell eingestuft werden. Kann eine essentielle Reaktion hingegen von mehreren vorhandenen (Iso-)Enzymen katalysiert werden, sind letztere einzeln betrachtet nicht essentiell. Im umgekehrten Fall können multifunktionale Enzyme wie die Transketolase (bei deren Entfernen mehrere Reaktionen gleichzeitig eliminiert werden) essentiell sein, ihre einzelnen katalysierten Reaktionen dagegen nicht.

Den essentiellen Reaktionen völlig entgegengesetzt sind *ausgeschlossene* Reaktionen, die in keinem EM involviert sind und für die folglich $h_i^S = 0$ ist. Solche Reaktionen haben für Realisierungen eines Szenarios praktisch keine Relevanz, da ihre resultierenden Reaktionsraten r_i in Gl. (6.3) immer Null sind. Die Zelle wird eine ausgeschlossene Reaktion unter Umweltbedingungen, die dem Szenario entsprechen, vermutlich reprimieren. Die in Abschnitt 2.4.1 beschriebenen *blockierten Reaktionen*, die über die Kernmatrix identifiziert wurden, erfüllen das Kriterium $h_i^S = 0$ für jedes S . Im Gegensatz zur Nullraumanalyse können wir hier auch die Vorzeichenbedingungen für die Reaktionsraten und zusätzliche Einschränkungen durch ein spezifiziertes Szenario S berücksichtigen.

In den meisten Fällen sind nur wenige Reaktionen ausgeschlossen oder essentiell, d.h. die relativen Auftretshäufigkeiten (abgekürzt relAHen) liegen zwischen 0 und 1. Dennoch liefert uns h_i^S grobe Information über die Relevanz einer Reaktion in einem gegebenen Szenario.

Flexibilität und (strukturelle) Robustheit

Da in (6.1) bzw. (6.3) über alle zur Verfügung stehenden EMen summiert wird, ist die Anzahl der EMen $|\mathcal{E}|$ bzw. $|\mathcal{E}^S|$ ein Maß für die Flexibilität des Netzes bzw. des Szenarios. Eine große Anzahl von EMen eröffnet viele, qualitativ unterschiedliche Kombinationsmöglichkeiten, um eine Flußverteilung zu realisieren – das System ist flexibel. Flexibilität kann hier als Re-

dundanz verstanden werden, denn mit der Anzahl der EMen wird das Vorhandensein paralleler und genetisch unabhängiger Routen (Abschnitt 2.4.3) quantitativ ausgedrückt. Im Extremfall $|E^S| = 0$ gibt es überhaupt keine Möglichkeit, nichttriviale stationäre Flußverteilungen zu konstruieren, die den Bedingungen eines spezifizierten Szenarios gehorchen.

Betrachtet man die Flexibilität spezifischer auf der Stufe der einzelnen Reaktionen, so kann das Netz bzgl. einer Reaktion mit hoher $relAH$ h_i^S nur wenig flexibel reagieren. Solche Reaktionen ermöglichen zwar viele Realisierungen, können aber gerade deshalb selbst kaum ersetzt werden, da nur ein kleiner Anteil der EMen $(1-h_i^S)$ ohne sie operieren kann. Für essentielle Reaktionen gibt es überhaupt keine Alternativen (Abb. 6.1).

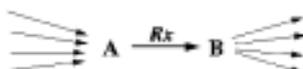


Abb. 6.1: Essentielle Reaktionen (hier: R_x) ermöglichen dem Netz flexible Realisierungen (hier: $4 \cdot 4 = 16$ EMen), sind aber selbst nicht zu ersetzen.

Wie wir unten zeigen werden, sind Flexibilität und Redundanz eng mit (struktureller) Robustheit verknüpft. Robustheit eines Systems ist definiert als relative Insensitivität gegenüber Änderungen in dessen Parametern [72,93,166]. Hier betrachten wir Änderungen in der Struktur des Reaktionsnetzes. Sollten Reaktionen (zufällig) ausfallen, z.B. durch Mutation oder Inhibition, so bedeutet eine hohe Flexibilität eine höhere Wahrscheinlichkeit dafür, daß dennoch Routen im Netz vorhanden sind, die ohne diese Reaktion(en) operieren können. Je größer also die Flexibilität, desto robuster sollte die Struktur gegen externe oder interne Störungen sein.

Zur Illustration oben eingeführter Begriffe betrachten wir drei Szenarien im Netz NB2. In Tab. 6.1 sind die jeweils selektierten EMen (vgl. mit Abb. 2.9) und die relativen Häufigkeiten der Reaktionen angegeben. $S1$ entspricht „Wachstum“ auf A ohne Aufnahme/Ausscheidung von B. Vier EMen erfüllen diese Spezifikation. Da $h_{P_2}^S = 1$ ist, können wir schlußfolgern, daß bei diesem Szenario Metabolit P immer ausgeschieden wird, d.h. R3 ist hier (neben der geforderten Reaktion R1) essentiell. Die anderen Reaktionen haben eine Auftretswahrscheinlichkeit von 0.5, sind damit also ersetzbar. In $S2$ betrachten wir Wachstum auf B (ohne Aufnahme von A). Nur ein EM erfüllt diese Spezifikation, Wachstum auf B ist also weniger flexibel als auf A. Die beiden im einzigen EM involvierten Reaktionen R3 und R9 sind damit (neben der geforderten Reaktion R2) automatisch essentiell. Alle anderen Reaktionen sind dagegen ausgeschlossen und damit ohne Bedeutung für die Realisierung dieses Szenarios, ihre Raten müssen hier sogar Null sein. In $S3$ betrachten wir wie in $S1$ Wachstum auf A ohne Produktion/Konsumption von B und fordern zusätzlich, daß E produziert wird. Natürlich muß die selektierte Menge an EMen in $S3$ eine Teilmenge von der in $S1$ sein. Hier sind nun R7 und R10 (neben den geforderten R1 und R4) essentiell, während R9 überhaupt keine Relevanz mehr hat. R5, R6 und R8 sind an Realisierungsmöglichkeiten beteiligt, aber strukturell ersetzbar.

Szenario	R ⁵	EM-Nr. (Abb. 2.9)	relative Auftrittshäufigkeiten \hat{h}_i^6									
			R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
S1 = {R1, ¬R2}	4	1,2,3,4	1	0	1	0,5	0,5	0,5	0,5	0,5	0,5	0,5
S2 = {¬R1, R2}	1	8	0	1	1	0	0	0	0	0	1	0
S3 = {R1, ¬R2, R4}	2	3,4	1	0	1	0	0,5	0,5	1	0,5	0	1

Tab 6.1: Verschiedene Szenarien in NB2: Selektierte EMen und relative Auftrittshäufigkeiten.

6.2.2 Anwendung auf *E. coli*: Wachstum auf verschiedenen Substraten

Im folgenden analysieren wir die EMen des Zentralstoffwechsels von *E. coli* (Abschnitt 3.2, Anhang A.2, Abb. 4.9) bzgl. des Wachstums auf jeweils einem der vier Substrate Glucose, Succinat, Acetat und Glycerin. Die EMen wurden so berechnet, daß jeweils nur ein Substrat aufgenommen werden konnte⁷. Die 5 Produkte konnten dagegen (auch parallel) ausgeschieden werden. Tab. 6.2 zeigt eine Übersicht der mit dem FluxAnalyzer berechneten EMen, einschließlich ihrer Aufteilung nach verschiedenen Szenarien.

	Glucose	Acetat	Glycerin	Succinat
Insgesamt für jedes Substrat	27100 (100%)	599 (100%)	11333 (100%)	4250 (100%)
davon (Szenarien):				
(1) mit Wachstum: [wsw]	21592 (80%)	363 (61%)	9479 (84%)	3421 (80%)
(1.1) rein aerobes Wachstum [wsw, O2_ap]	19801 (73%)	363 (61%)	9479 (84%)	3421 (80%)
(1.2) anaerobes Wachstum [wsw, ¬O2_ap]	1791 (7%)	0 (0%)	0 (0%)	0 (0%)
(2) mit ATP-Nettosynthese [ATPwsw]	2660 (10%)	42 (7%)	851 (8%)	280 (7%)
(2.1) unabh. v. Wachstum [¬wsw, ATPwsw]	879 (3%)	30 (5%)	274 (2%)	103 (2%)
(3) Roten ohne Nutzen [¬wsw, ¬ATPwsw]	4629 (17%)	206 (34%)	1580 (14%)	726 (17%)

Tab. 6.2: EMen für *E. coli* für Wachstum auf verschiedenen Substraten: Gesamtanzahl und Verteilung bzgl. unterschiedlicher Kriterien (Szenarien).

Betrachten wir zunächst die Gesamtanzahl der zur Verfügung stehenden EMen, so bestätigt sich die intuitive Erwartung, daß die Verwertung von Glucose am flexibelsten ist. Mit nur 599 EMen (2,2% im Vergleich zu Glucose) hat Acetat die kleinste Anzahl genetisch unabhängiger Verwertungswege. Die Analyse des Szenarios „anaerobes Wachstum“ demonstriert, wie mit EMen Funktionalität präzisiert werden kann: nur für Glucose gibt es EMen, die Wachstum ohne Sauerstoff (und ohne externe Elektronenakzeptoren) ermöglichen, was in Übereinstimmung mit experimentellen Beobachtungen steht. Ein anderer interessanter Aspekt in Tab. 6.2 ist, daß jeweils nur ein kleiner Anteil der EMen netto ATP produzieren kann, insbesondere wenn das unabhängig von Wachstum geschehen soll. Letzteres ist aber von eminenter Bedeutung, da die Zelle einen wachstumsunabhängigen ATP-Bedarf für Erhaltungsprozesse (*Maintenance*) hat [83,149]. Beträchtlich dagegen ist der Anteil solcher EMen, die Substrat

⁷ Jedem Substrat ist ein Szenario Substrat zugeordnet, das die Aufnahme der anderen Substrate ausschließt. Für Wachstum auf Glycerin gilt z.B. das Szenario: Glycerin = {¬Glc_PTS_ap, ¬Glc_ATP_ap, ¬Ac_ap, ¬Suc_ap}.

verbrauchen, aber dabei weder Biomasse noch ATP generieren. Eine optimale Regulation des zellulären Stoffwechsels wird einen großen Fluß entlang dieser „nutzlosen“ Routen (*fruitless routes*) vermeiden. In den meisten Fällen sind diese EMen mit Ausscheidung von Produkten wie Lactat oder Formiat verbunden, nur bei einigen entsteht ausschließlich CO_2 .

Erwähnt soll an dieser Stelle noch sein, daß die Gesamtanzahlen der EMen in Tab. 6.2 (erste Zeile) jeweils einen EM ohne Substrataufnahme einschließen. Dieser wird gebildet aus einem zyklischen Fluß zwischen Succinat-Dehydrogenase und Fumarat-Reductase. Generell läßt sich sagen, daß EMen ohne Verbrauch von extern bereitgestellter Energie (hier: ohne Aufnahme eines Substrates) thermodynamische Inkonsistenzen offenbaren [119]. In der Tat wurde die Fumarat-Reductase-Reaktion hier vereinfacht modelliert, da das an dieser Reaktion beteiligte Menachinon [83] in einem gemeinsamen Pool mit Ubichinon bilanziert wurde (Anhang A.1). Da Menachinon ein negativeres Redoxpotential hat, liefert es bei der Aufnahme von Elektronen vermutlich weniger Energie, da die Redoxpotentialdifferenz zu potentiellen Elektronenlieferanten (wie NADH) kleiner ist als bei Ubichinon. Dann würde auch dieser EM entfallen, da er ohne Energieeintrag nicht mehr operieren könnte. Die Tatsache, daß kein weiterer, thermodynamisch unmöglicher EM gefunden wurde, ist ein Fingerzeig für die Konsistenz des stöchiometrischen Modells und demonstriert eine weitere nützliche Anwendung von EMen.

Abb. 6.2 zeigt die Relevanzen einiger ausgewählter Reaktionen anhand ihrer relAH $h_{\text{Substrat},i}$, jeweils für Wachstum auf einem der vier Substrate. Die anaplerotische PEP-Carboxylase-Reaktion ($\text{PEP}::\text{OxA}$) dient bekanntermaßen zum Auffüllen der Präkursoren-Pools im TCA, insbesondere bei glykolytisch verwerteten Substraten. In der Tat ist die relAH von $\text{PEP}::\text{OxA}$ in EMen von Glycerin und Glucose recht hoch. Dagegen hat $\text{PEP}::\text{OxA}$ bei Wachstum auf Succinat eine geringere Bedeutung, da Succinat direkt in den TCA eintritt und deshalb TCA-Komponenten *entgegen* der Richtung von $\text{PEP}::\text{OxA}$ aus dem TCA *heraus*-transportiert werden müssen (was sich tatsächlich in einer hohen relAH der dafür relevanten Reaktionen $\text{Mal}::\text{OxA}$ und $\text{OxA}::\text{PEP}$ widerspiegelt; nicht gezeigt). Bei Wachstum auf Acetat ist es ähnlich. Dieses Substrat gelangt über die Zwischenstufe Acetyl-CoA und dann mit Hilfe des Glyoxylat-Zyklus sofort in den TCA. Der Glyoxylat-Zyklus ist für Wachstum auf Acetat sogar essentiell. Dieses bei entsprechenden Mutanten experimentell beobachtbare Ergebnis wird auch durch die 100%-ige Auftretswahrscheinlichkeit der Reaktion der Isocitrat-Lyase, einem der beiden Enzyme des Glyoxylatzyklus, in den EMen von Acetat bestätigt (Abb. 6.2). Eine nicht zu unterschätzende Relevanz dieses Zyklus ergibt sich auch für Glucose und Glycerin, da er dort ebenfalls als möglicher anaplerotischer Weg in Frage kommt, wenn auch mit deutlich geringerer Relevanz als die PEP-Carboxylase. Für Succinat ist der Glyoxylat-Weg aus den gleichen Gründen wie bei der PEP-Carboxylase von geringer Bedeutung. Das Beispiel der Isocitrat-Lyase zeigt, daß die Klassifikation einer Reaktion oder eines Gens nach essentiell/nicht essentiell i.a. von den Umweltbedingungen – hier vom Substrat – abhängen.

Die Pyruvat-Kinase schließlich, die PEP irreversibel zu Pyruvat konvertiert, wird wieder vornehmlich von den beiden glykolytisch abzubauenenden Substraten genutzt. Dabei ist $h_{Pyr,PEP}^{Glycolic, (a)}$ signifikant größer als $h_{Pyr,PEP}^{Gluco, (a)}$. Das ist darin begründet, daß bei der Aufnahme von Glucose über das Phosphotransferase-System (PTS) bereits eine beachtliche Menge an PEP zu Pyruvat umgesetzt wird, was sich in einer geringeren Relevanz der Pyruvat-Kinase manifestiert. Bei Succinat und Acetat wird Pyruvat vornehmlich über das Malat-Enzym synthetisiert.

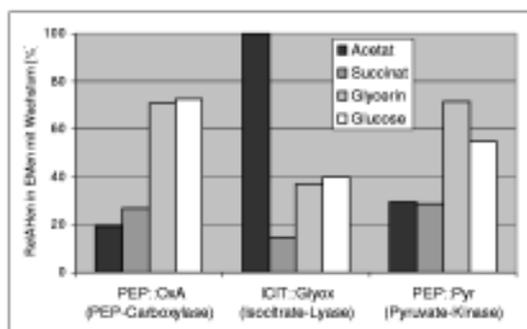


Abb. 6.2: Relative Auftretshäufigkeiten $h_i^{Substrat, (a)}$ einzelner Reaktionen in EMen mit Wachstum (jeweils für jedes Substrat).

Schon diese Beispiele bestätigen, daß die relAHen tatsächlich Rückschlüsse auf die Relevanz einzelner Reaktionen ermöglichen. Über die gewählte Selektion von EMen kann dabei jeweils auf bestimmte Umweltbedingungen (Substrat, aerob/anaerob) oder/und Zellfunktionen (z.B. Wachstum, Synthese eines speziellen Produktes) fokussiert werden.

In biologischen Studien wird die Relevanz einzelner Reaktionen oft über „knock-out-Mutanten“ untersucht, wo Gene metabolischer Enzyme und damit Reaktionen im Netz deletiert sind. In Abb. 6.3a untersuchen wir den Zusammenhang zwischen dem experimentell bestimmten Phänotyp (lebensfähig/nicht lebensfähig) solcher Deletionsmutanten von *E. coli* mit der entsprechenden relAH der stillgelegten Reaktion. (Im Anhang B findet sich eine Einzelaufzählung der Mutanten und relAHen.) Jeder der 90 Punkte in Graphik 6.3a repräsentiert eine Mutante, die einem der vier Substrate ausgesetzt wurde (x-Achse). Auf der y-Achse ist die entsprechende relAH $h_i^{Substrat, (a)}$ der deletierten Reaktion i aufgetragen. Ein klarer Zusammenhang zwischen dem experimentellen Phänotyp und der relAH der entfernten Reaktion kann festgestellt werden: Bei 34 der insgesamt 40 nicht lebensfähigen Mutanten wurden essentielle Reaktionen entfernt, d.h. jeder EM mit Wachstum auf dem entsprechenden Substrat involviert die gelöschte Reaktion. Auch in den anderen 6 nicht lebensfähigen Mutanten hat die gelöschte Reaktion zumindest eine hohe Relevanz. Vermutlich reichen dort die wenigen verbleibenden alternativen Pathways aus kinetischen (oder regulatorischen) Gründen nicht aus. 48 der 50

experimentell lebensfähigen Mutanten verfügen tatsächlich über alternative EMen nach der Deletion, da in diesen Fällen die relAH der gelöschten Reaktion unter 100% lag. In nur zwei Fällen konnte eine Mutante auf dem dargebotenen Substrat wachsen, obwohl die deletierte Reaktion als essentiell eingestuft wird. In beiden Fällen (einmal für Substrat Succinat, einmal Glycerin) handelte es sich um die reversible Reaktion $G6P \rightarrow F6P$, die vom im Gen pgi kodierten Enzym Glucosephosphat-Isomerase katalysiert wird. Im Modell wird diese Reaktion bei Wachstum auf Succinat oder Glycerin als letzter gluconeogenetischer Schritt zur Synthese von Glucose-6-phosphat (G6P) benötigt, das in relativ geringen Mengen in die Biomasse eingebaut wird. Generell sind solche *falsch negativen* Prädiktionen ein Beweis für ein inakkurates Netzmodell. Und tatsächlich scheint G6P in der Realität austauschbar zu sein [13]. Klassifizierten wir die Überlebensfähigkeit der untersuchten Mutanten ausschließlich danach, ob eine Reaktion essentiell ist oder nicht, so wären hier 81 von 90 Prädiktionen (90%) korrekt. Diese 0/1-Prädiktionen sind immer mit denen der *Flux Balance Analysis* (Abschnitt 2.4.2) identisch, man vergleiche z.B. die Ergebnisse hier mit denen in [29]. Allerdings lassen die relAHen differenziertere Aussagen über die Relevanz der gelöschten Reaktion zu.

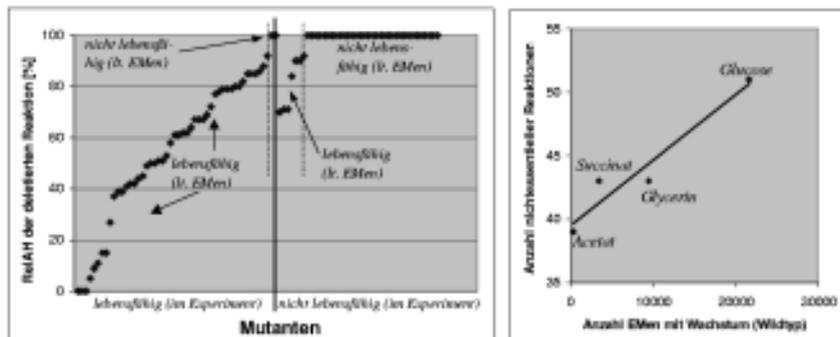


Abb. 6.3: (a) (links) Gegenüberstellung von Mutanten-Phänotypen im Experiment und der relativen Auftretenshäufigkeit $h_i^{relAH(i)}$ der jeweils deletierten Reaktion i . Eine genaue Aufschlüsselung ist in Anhang B gegeben. (b) (rechts) Korrelation zwischen der Anzahl EMen und der Anzahl nichtessentieller Reaktionen (jeweils bzgl. Wachstum).

Der Zusammenhang zwischen Flexibilität und struktureller Robustheit wurde oben bereits angesprochen. In Abb. 6.3b wird getestet, ob tatsächlich eine Korrelation besteht zwischen der Anzahl der EMen und der Fähigkeit, strukturelle Änderungen – hier bei Ausfall einzelner Reaktionsschritte – zu tolerieren. Dazu tragen wir für jedes Substrat die Anzahl der EMen mit Wachstum gegen die Anzahl nichtessentieller Reaktionen auf. Tatsächlich gibt es bei Wachstum auf Glucose die größte und bei Acetat die kleinste Menge an Reaktionen, deren Ausfall über alternative Wege abgefangen werden kann. Das Beispiel Glycerin und Succinat zeigt aber, daß nicht immer eine größere Anzahl an EMen eine geringere Ausfallwahrscheinlichkeit bei einer zufälligen Deletion impliziert.

Dennoch scheint die Menge genetisch unabhängiger Pathways, zumindest solange man unterschiedliche Szenarien (Umweltbedingungen) in ein und demselben Netz vergleicht, als grobe Kenngröße für die strukturelle Robustheit geeignet zu sein. Problematisch wird ein quantitativer Vergleich bei Netzen mit unterschiedlicher Größe. So kann ein großes Netz mit sehr vielen EMen dennoch einige essentielle Reaktionen aufweisen, während man leicht ein kleines Netz mit wenigen EMen konstruieren kann, das keine essentielle Reaktion hat. Die Beziehung zwischen Netzgröße und EMen ist zudem nicht linear (siehe Kap. 8). Einen anderen Vorschlag zur quantitativen Bewertung der strukturellen Robustheit (oder *Fragilität*) beschreibt Kap. 7. Dort werden wir mit den *Minimalen Schnittmengen* gezielt nach solchen (u.U. Mehrfach-)Mutanten suchen, bei denen eine Funktion wie Wachstum ausgeschlossen ist.

Ein anderer, hier nicht weiter diskutierter Aspekt ist die Frage, inwieweit Mutanten, die als lebensfähig eingestuft wurden, tatsächlich noch die ursprüngliche (stöchiometrische) Performance aufweisen (z.B. bzgl. maximalen Biomassertrag). Der Leser sei hier auf [148] verwiesen, in der einige der hier dargestellten Ergebnisse komprimiert publiziert wurden.

6.2.3 Anwendung für *Rhodospirillaceae*: Phototropher und respirativer Stoffwechsel auf verschiedenen Substraten

In ähnlicher Weise wie für *E. coli* sollen hier auch die EMen der schwefelfreien Purpurbakterien (SPB) diskutiert werden. Die EMen wurden wiederum *separat* für die drei Substrate Succinat, Acetat und Fructose berechnet (Tab. 6.3). Als Grundlage diente das in Abschnitt 3.3 definierte Netz (Abb. 4.7 und 4.8) mit den möglichen Nebenprodukten Acetat, Formiat und Succinat. Die Nitrogenase-Reaktion wurde nicht berücksichtigt, außerdem wurden von vornherein die drei Reaktionen des reduktiven TCA (*SuccCoA::alKG*, *AcCoA::Pyr*, *CitLyase*) ausgeschlossen. Bereits in Abschnitt 3.3 wurde darauf hingewiesen, daß die Enzyme weder genetisch noch biochemisch in *R. sphaeroides* oder *R. capsulatus* nachgewiesen werden konnten. Die Bedeutung einer möglicherweise geringen Aktivität des reduktiven TCA in *R. rubrum* wurde bereits in Abschnitt 5.5.3 diskutiert. Die Glucose-6-phosphat-Dehydrogenase wurde berücksichtigt, obwohl dieses Enzym in *R. rubrum* vermutlich nicht vorhanden ist (vgl. Abschnitt 5.5.4). Insofern sind die berechneten EMen insbesondere für *R. sphaeroides* und *R. capsulatus* relevant. Dennoch können viele Schlußfolgerungen zumindest qualitativ auch auf *R. rubrum* übertragen werden: alle diejenigen der hier betrachteten EMen, bei denen die zusätzliche Forderung $\{-G6P::PGlac\}$ erfüllt ist, sind auch für *R. rubrum* gültig.

Obwohl ein direkter Vergleich mit dem *E. coli* - Netz mit Vorsicht zu bewerten ist (z.B. konnten im *E. coli* Netz andere Produkte ausgeschieden werden), sind signifikant höhere Anzahlen an EMen für die SPB festzustellen (man vgl. z.B. Substrate Succinat und Acetat in Tab. 6.2 und 6.3). Die größere Flexibilität ergibt sich insbesondere durch die Photosynthese, die als zusätzliche Energiequelle zur Verfügung steht. Aussagen zur relativen Aufteilung der EMen bzgl. Wachstum, ATP-Nettosynthese und *futile routes* sind ähnlich wie bei *E. coli*.

Für die *Rhodospirillaceae* macht es hier Sinn, die EMen bzgl. des Einsatzes von Sauerstoff und/oder Licht in vier unterschiedliche Szenarien aufzuteilen (Punkte (4)-(7) in Tab. 6.3). Wir fokussieren hier vor allem auf die Unterschiede zwischen aerobem (respirativem) Stoffwechsel im Dunkeln ($O2_{up}, \neg Photo$) und anaerobem Metabolismus mit Photosynthese ($\neg O2_{up}, Photo$). Dem respirativen Stoffwechsel stehen bei allen drei Substraten, insbesondere bei Succinat, deutlich mehr Routen zur Verfügung als anaerob mit Photosynthese. Besonders auffällig ist die Tatsache, daß für Succinat und Acetat unter rein phototrophen Bedingungen nur ein einziger EM existiert, der eine Nettoproduktion von ATP ermöglicht und damit Energie für den Erhaltungsstoffwechsel zur Verfügung stellen kann. In beiden Fällen wird ATP aus der zyklischen Photosynthese ohne Substratverbrauch generiert (er entspricht exakt (5.24)). Auch rein fermentative Routen, die theoretisch parallel zur Respiration oder Photosynthese möglich wären, können bei Succinat und Acetat – ähnlich wie bei *E. coli* – weder zur ATP-Produktion, geschweige denn zum Wachstum herangezogen werden (Tab 6.3, Punkt (7)). Lediglich bei Fructose gibt es phototroph einige zusätzliche Routen für die ATP-Produktion, außerdem kann Fructose auch fermentativ verwertet werden.

	Succinat	Acetat	Fructose
insgesamt für jedes Substrat	11156 (100%)	2267 (100%)	59267 (100%)
davon (Szenarios):			
(1) mit Wachstum (μ)	10764 (96%)	1903 (84%)	55704 (94%)
(2) mit ATP-Nettosynthese (ATP_{net})	1074 (10%)	50 (2%)	4475 (8%)
(3) Routen ohne Nutzen ($\neg \mu, \neg ATP_{net}$)	359 (3%)	341 (15%)	2018 (5%)
(4) mit Respiration UND Photosynthese ($O2_{up}, Photo$)	2857 (26%)	736 (32%)	12880 (22%)
(5) mit Respiration, ohne Photosynthese ($O2_{up}, \neg Photo$)	6667	993	30985
	(61%/100%)	(44%/100%)	(52%/100%)
(5.1) mit Wachstum ($O2_{up}, \neg Photo, \mu$)	6572 (59%/96%)	768 (34%/77%)	28406 (48%/92%)
(5.2) mit ATP-Nettosynthese ($O2_{up}, \neg Photo, ATP_{net}$)	1073 (10%/16%)	49 (2%/5%)	4275 (7%/14%)
(6) mit Photosynthese / anaerob ($Photo, \neg O2_{up}$)	1430	536	11869
	(13%/100%)	(24%/100%)	(20%/100%)
(6.1) mit Wachstum ($Photo, \neg O2_{up}, \mu$)	1377 (12%/96%)	468 (21%/87%)	11678 (20%/98%)
(6.2) mit ATP-Nettosynthese ($Photo, \neg O2_{up}, ATP_{net}$)	1 ($<0.01\%/0.07\%$)	1 ($0.04\%/0.1\%$)	21 ($0.04\%/0.18\%$)
(7) fermentativ (ohne Resp./Photos.) ($\neg Photo, \neg O2_{up}$)	2	2	3533
	($<0.01\%/100\%$)	(0.1%/100%)	(6%/100%)
(7.1) mit Wachstum ($\neg Photo, \neg O2_{up}, \mu$)	0 (0%/0%)	0 (0%/0%)	2805 (5%/79%)
(7.2) mit ATP-Nettosynthese ($\neg O2_{up}, \neg Photo, ATP_{net}$)	0 (0%/0%)	0 (0%/0%)	179 (0.3%/5%)

Tab 6.3: EMen im Stoffwechsel der schwefelfreien Purpurbakterien für Wachstum auf verschiedenen Substraten: Gesamtanzahl und Verteilung bzgl. unterschiedlicher Kriterien (Szenarien).

Auch bei den SPB findet sich jeweils nur ein EM ohne Nettoenergieverbrauch, d.h. ohne Substrataufnahme oder Photosynthese. Es handelt sich um den gleichen Minizyklus aus Succinat-Dehydrogenase und Fumarat-Reductase wie bei *E. coli* (siehe Erläuterung oben). Deshalb können wir auch hier von einem thermodynamisch konsistenten Netzmodell ausgehen.

In Abb. 6.4 sehen wir die relativen zweier ausgewählter Reaktionen für rein phototrophes ($\hat{h}_{Substrat, (\neg O2_{up}, Photo, \mu)}$) bzw. rein respiratives ($\hat{h}_{Substrat, (O2_{up}, \neg Photo, \mu)}$) Wachstum der SPB, je-

weils sortiert nach Substraten. Für die Schlüsselreaktion des Calvin-Zyklus (*Rubisco*) ergeben sich erwartungsgemäß die jeweils höchsten Level in EMen mit phototrophem Wachstum. Bereits im Abschnitt 5.5.2 hatten wir die eminente Bedeutung des Calvin-Zyklus – auch bei einer Nettofreisetzung von CO_2 – für photoheterotrophes Wachstum diskutiert. Aufgrund der notwendigen Balance der Redoxäquivalente ist diese Reaktion bei phototrophem Wachstum auf Acetat und Succinat sogar essentiell. (Hinweis: unter Einbeziehung der Nitrogenase-Reaktion gäbe es einige EMen ohne Involvierung von *Rubisco*, allerdings mit herabgesetzter Biomasseausbeute; siehe Abschnitt 5.5.3.) Lediglich mit Fructose als Substrat existieren einige phototrophe Routen, die *Rubisco* nicht involvieren. Allerdings sind diese grundsätzlich an eine Ausscheidung von Produkten geknüpft, was zu Einbußen in der Effizienz des Stoffwechsels führt. Das Beispiel *Rubisco* zeigt erneut, daß die relAH tatsächlich Funktion und Relevanz einer Reaktion widerspiegelt, was insbesondere in funktional weniger gut charakterisierten Stoffwechsel(teil)netzen hilfreich ist. Die wichtige Bedeutung des Calvin-Zyklus als Redoxsenke – auch für photoheterotrophes Wachstum der SPB – konnte hier bestätigt werden.

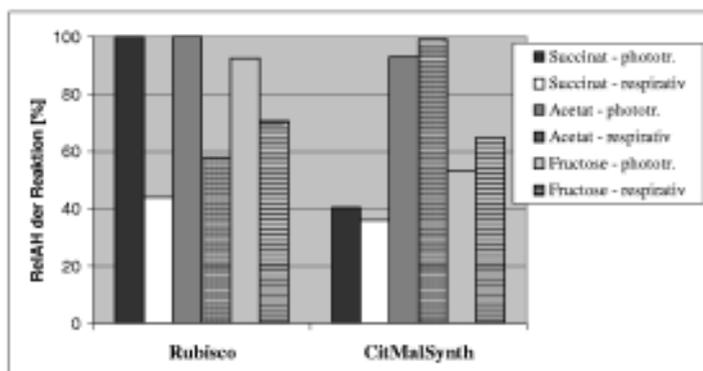


Abb. 6.4: Relative Auftretshäufigkeiten zweier ausgewählter Reaktionen in EMen für rein phototrophes bzw. rein respiratives Wachstum der SPB auf einem der drei Substrate.

Der praktische Nutzen der kombinatorischen Pathwayanalyse kann auch am Beispiel der Citramalat-Lyase-Reaktion (*CitMalSynth*) illustriert werden, die den Startpunkt des erst jüngst vorgeschlagenen Citramalat-Zyklus darstellt ([7,64], Abschnitt 3.3). Dieser gilt als eine Erklärung dafür, wie die SPB, die über keinen Glyoxylat-Zyklus als anaplerotischen Weg verfügen, auf Acetat wachsen können. Tatsächlich ergeben sich für *CitMalSynth* bei Wachstum auf Acetat sowohl im photo- als auch im chemotrophen Fall die jeweils höchsten relAHen, während die Reaktion für die anderen Substrate einen deutlich geringeren Anteil an der Flexibilität hat (Abb. 6.4). Dennoch kommt man hier zu der etwas überraschenden Feststellung, daß Wachstum auf Acetat auch ohne Citramalat-Zyklus zumindest strukturell möglich wäre, da *CitMalSynth* nicht essentiell ist. Eine genauere Analyse aller EMen ohne Beteiligung von *CitMalSynth* zeigt, daß der TCA stark in oxidativer Richtung gefahren wird. Das dabei entste-

hende NADH wird über den Calvin-Zyklus zur Refixierung eines Teils des freigesetzten Kohlendioxids verwendet. Das Primärprodukt, 3-Phosphoglycerat, steht dann als Kohlenstoffquelle zur Auffüllung der 12 Präkursoren-Pools zur Verfügung, jene im TCA können nun über die anaplerotische Reaktion der Pyruvat-Carboxylase erreicht werden. Anhang C zeigt so einen (in diesem Fall phototrophen) EM im *FluxAnalyzer*. In allen zum Citramalat-Zyklus alternativen Routen ist *Rubisco* involviert, d.h. bzgl. der Spezifikation $[Ac_up, \mu, -CitMalSynth]$ ist diese Reaktion essentiell. Da das Enzym *Rubisco* aber aerob reprimiert ist [155], können diese alternativen Routen für aerobe Verhältnisse praktisch ausgeschlossen werden. Außerdem sind diese EMen energetisch sehr aufwendig und implizieren große Flüsse, die vermutlich auch anaerob/phototroph kaum realisierbar sind. Darum werden die Verwertungswege, die den Citramalat-Zyklus umgeben, vermutlich keine Rolle spielen. Dieses Beispiel legt nahe, die EMen differenzierter nach ihrer Effizienz (Nutzen/Aufwand) zu bewerten.

6.3 Gewichtete Relevanzen

Im letzten Abschnitt haben wir jeden EM mit der gleichen Gewichtung betrachtet und die Relevanz einer Reaktion über ihre relative Auftrittshäufigkeit bewertet. Diese Vorgehensweise berücksichtigte alle Realisierungsmöglichkeiten des Systems, also die gesamte Flexibilität, läßt dabei aber zwei wichtige Kriterien außer acht:

- (i) *Effizienz*: Oft treten große Unterschiede in der *Performance* der EMen auf. So können einige EMen eine sehr hohe Ausbeute bzgl. Biomasse oder ATP aufweisen und/oder benötigen für gleiche Syntheseleistungen geringere Flüsse als andere. Da zumindest Bakterien bzgl. des Wachstums optimale Phänotypen aufweisen können [27,61], sollte dieser Aspekt hier entsprechend berücksichtigt werden.
- (ii) *Reaktionsraten in den EMen*: In (6.2) bzw. (6.4) haben wir nur die binäre Darstellung der EMen verwendet. Nun kann es durchaus passieren, daß eine Reaktion bzgl. zweier Szenarien jeweils hohe relAHen zeigt, aber in nur einem der beiden Szenarien auch mit jeweils großen Umsätzen involviert ist. Für eine stärker quantitativ orientierte Bestimmung der Relevanz einer Reaktion sollten deshalb auch deren Koeffizienten in den EMen mit einfließen, was eine vorhergehende Normierung der EMen voraussetzt.

Der folgende allgemeine Ansatz einer *gewichteten Elementarmodenanalyse* versucht nun, diese beiden Kriterien mit einzubinden:

$$\sigma^S = \mathbf{E}^S \mathbf{w}^S \quad (6.5)$$

Die *Flexibilität* des betrachteten Szenarios S wird weiterhin über die Gesamtheit der selektierten EMen in \mathbf{E}^S berücksichtigt. Der $(\mathbb{R}^S, 1)$ -Vektor \mathbf{w}^S soll hier ausschließlich nichtnegative Einträge haben und soll sowohl die *Normierung* der EMen, als auch ihre Gewichtung bzgl. eines zu definierenden *Effizienzkriteriums* übernehmen. Auch die *Reaktionsraten* in den EMen

fließen nun über E^S mit in die Berechnung von σ^S ein. Dieser $(q,1)$ -Vektor σ^S enthält die gewichteten Relevanzen einer jeden Reaktion bzgl. des Szenarios S . Letztlich repräsentiert σ^S wegen (6.3) eine zulässige stationäre Flußverteilung. Wir wollen uns an dieser Stelle aber offenhalten, die Matrix E^S in (6.5) gegen andere Darstellungen auszutauschen, z.B. durch die Matrix E_{Abs}^S , welche die Absolutwerte von E^S enthält. Tauscht man E^S gegen E_{Abs}^S aus und setzt $w_i = \frac{1}{|E^S|}$ für alle i dann entspricht σ^S den relAHen, d.h. dann ist $\sigma^S = h^S$ (Gl. (6.4)).

Sind wir an den gewichteten Relevanzen der Reaktionen bzgl. der Biomassesynthese auf einem bestimmten Substrat Sub interessiert (Symbol: $\sigma^{S(Sub)}$), so könnte eine konkrete Ausprägung für (6.5) wie folgt gestaltet werden: Als Gewichte $w_j^{S(Sub)}$ wählen wir

$$w_j^{S(Sub)} = \frac{e_j(\mu)}{\sum_i e_i(\mu)} \cdot \frac{1}{e_{Sub, Aufnahme}^j \cdot Y_{X/Sub}^{max}} \quad (6.6)$$

Der erste Faktor in (6.6) ist ein (normiertes) Effizienzkriterium für den j -ten EM. Dabei wird die spezifische Effizienz $e_j(\mu)$ des j -ten EM hier so definiert, daß der Nutzen (erzeugte Biomasse) ins Verhältnis zum Aufwand (Summe der absoluten Flüsse) gesetzt wird:

$$e_j(\mu) = \frac{e_{\mu}^j}{\sum_i |e_i^j|} \quad (6.7)$$

Indirekt wird durch (6.7) auch ein hoher Substratverbrauch negativ bewertet. Die zu realisierenden Flüsse können tatsächlich als Aufwand angesehen werden, da die entsprechenden Enzyme synthetisiert werden müssen. Die (zusätzliche) Einbeziehung thermodynamischer oder enzymespezifischer Parameter für jede Reaktion wäre sinnvoll, sofern diese Daten zur Verfügung stehen. Durch die Division der spezifischen Effizienz durch die Summe aller Effizienzen in (6.6) erfolgt eine Normierung, so daß die Summe aller dieser Quotienten 1 ergibt.

Der zweite Faktor in (6.6) normiert den zu gewichtenden EM j . Zunächst wird durch die Substrataufnahme $e_{Sub, Aufnahme}^j$ geteilt, was voraussetzt, daß nur EMen mit Verwertung des jeweiligen Substrates betrachtet werden. Sollen später Relevanzen für verschiedene Substrate miteinander verglichen werden (s.u.), muß auch der Ausbeutekoeffizient des betrachteten Substrates berücksichtigt werden. In (6.6) wird die maximale Biomasseausbeute $Y_{X/Sub}^{max} = \max_j (e_{\mu}^j / e_{Sub, Aufnahme}^j)$ verwendet, alternativ könnten auch experimentell bestimmte Ausbeuten eingesetzt werden. Mit (6.5)-(6.7) ermittelt sich $\sigma^{S(Sub)}$ damit nun wie folgt:

$$\sigma^{S(Sub)} = E_{Abs}^{S(Sub)} w^{S(Sub)} \quad (6.8)$$

Die Verwendung von Absolutwerten in den EMen verhindert, daß sich positive und negative Einträge einer reversiblen Reaktion neutralisieren und damit die Bedeutung der Reaktion für die Flexibilität „verwischen“ würden. EMen ohne Wachstum werden entsprechend mit 0 bewertet. Im Gegensatz zu den relAHen liegt $\sigma^{\mu(\text{Sub})}$ nicht mehr zwischen 0 und 1, ist aber auf jeden Fall nichtnegativ. Anhang D illustriert am Beispiel von NB2 die Berechnung und Interpretation von gewichteten Relevanzen, die nach obigem Schema berechnet werden.

Bisher haben wir in $\sigma^{\mu(\text{Sub})}$ nur die erzeugte Biomasse als Effizienzkriterium berücksichtigt. Schon mehrfach wurde betont, daß die Zelle ATP für wachstumsunabhängige Erhaltungsprozesse generieren muß. Deshalb sollte als zweites funktionales Kriterium für die Zelle auch die Produktion von ATP (summativ) in die Gewichtung der EMen einbezogen werden:

$$w_j^{\mu+\text{ATP}(\text{Sub})} = \frac{\varepsilon_j(\mu)}{\sum_k \varepsilon_k(\mu)} \cdot \frac{1}{e_{\text{Sub_aufnahme}}^j \cdot Y_{X/\text{Sub}}^{\text{max}}} + \frac{\varepsilon_j(\text{ATP})}{\sum_k \varepsilon_k(\text{ATP})} \cdot \frac{1}{e_{\text{Sub_aufnahme}}^j \cdot Y_{\text{ATP}/\text{Sub}}^{\text{max}}} \quad (6.9)$$

Die spezifischen Effizienzen $\varepsilon_j(\text{ATP})$ bzgl. Netto-ATP-Synthese ergeben sich analog zu (6.7)

$$\varepsilon_j(\text{ATP}) = \frac{e_{\text{ATP}}^j}{\sum_i |e_i^j|}, \quad (6.10)$$

wobei e_{ATP}^j die Nettosynthese an ATP beschreibt (im stöchiometrischen Modell von *E. coli* bzw. SPB entspricht das der Reaktionsrate von *ATPdrain*). Damit erhalten wir die gewichteten Relevanzen der Reaktionen bzgl. Biomasse- und ATP-Synthese bei Wachstum auf *Sub*:

$$\sigma^{\mu+\text{ATP}(\text{Sub})} = E_{\text{Sub}}^{\text{[Sub]}} w^{\mu+\text{ATP}(\text{Sub})} \quad (6.11)$$

$\sigma^{\mu+\text{ATP}(\text{Sub})}$ wurde in [148] als *control-effective fluxes* (CEFs) eingeführt, da mit ihnen Rückschlüsse auf die Regulation des Stoffwechsels möglich sind (s.u.). Die CEFs liefern quantitative Abschätzungen für die physiologischen Relevanzen der Reaktionen bei Wachstum auf dem entsprechenden Substrat. Der Begriff „Substrat“ kann hier weit gefaßt werden, so könnte man z.B. auch die Photosynthese als eine extern zur Verfügung gestellte (Energie-) Ressource betrachten. Werden mehrere Substrate zur Verwertung angeboten, kann man diese kumulativ (mit entsprechender Gewichtung) in $e_{\text{Sub_aufnahme}}^j$ berücksichtigen.

Da die Koeffizienten in den EMen in die Berechnung von gewichteten Relevanzen eingehen, kann es durchaus passieren, daß eine essentielle Reaktion nur einen kleinen CEF-Wert aufweist (z.B. wenn sie für die Produktion einer essentiellen Komponente zuständig ist, die nur einen sehr kleinen Anteil an der Biomasse hat). Ähnlich wie bei mRNA-Transkriptmessungen

liegt die besondere Anwendung der CEFs im differentiellen Vergleich ihrer Levels bzgl. verschiedener Bedingungen; hier bei Wachstum auf zwei verschiedenen Substraten *Sub1* und *Sub2*. Für jede Reaktion *i* bilden wir dazu das Verhältnis θ_i der CEF-Werte:

$$\theta_i(\text{Sub1}, \text{Sub2}) = \frac{\sigma_i^{\text{P}^+ \text{ATP}}(\text{Sub1})}{\sigma_i^{\text{P}^+ \text{ATP}}(\text{Sub2})} \quad (6.12)$$

Mit θ setzen wir also die gewichteten Relevanzen der Reaktion *i* auf *Sub1* und *Sub2* in Beziehung. Aufgrund der zusätzlichen Einbeziehung von Effizienz und Umsatz zeigen sich differentielle Unterschiede in θ signifikanter als im Verhältnis von relAHen. So ergibt sich z.B. im *E. coli* - Netz: $\theta_{\text{C}_{60}::\text{G}_{60}}(\text{Acetat}, \text{Glucose})=20.54$ während das entsprechende Verhältnis aus den relAHen nur 2.51 ist (vgl. Abb. 6.2). Die besondere Bedeutung der Isocitrat-Lyase bei Wachstum auf Acetat wird also über $\theta_{\text{C}_{60}::\text{G}_{60}}$ quantitativ viel stärker ausgedrückt.

Es erscheint attraktiv, die Verhältnisse der CEFs mit denen von mRNA-Transkriptmessungen von stationären Kulturen zu vergleichen und damit zu testen, ob sich θ als theoretisches (prädiktives) Transkriptverhältnis eignet. Die mRNA-Levels spiegeln die Ergebnisse genetischer Regulationsvorgänge in der Zelle wider. Die Expression eines Gens, das ein Stoffwechsellenzym kodiert, sollte sich dabei an der Relevanz der entsprechenden Reaktion – nach ähnlicher Maßgabe wie die CEFs – für die jeweils vorliegenden Bedingungen orientieren. Für *E. coli* sind bereits einige Expressionsprofile von Kulturen im exponentiellen Wachstum mittels DNA-Microrarrays bestimmt worden. Für nachfolgende Analysen wurden publizierte Daten aus [106] (Substratpaar Acetat/Glucose) bzw. [105] (Glycerin/Glucose) herangezogen. Anhang E gibt eine Einzelauflistung.

Abb. 6.5a zeigt die 50 vergleichbaren Datenpaare aus gemessenen und theoretischen Transkriptverhältnissen für Acetat/Glucose. Den Berechnungen der $\theta(\text{Acetat}, \text{Glucose})$ lagen die EMen aus Abschnitt 6.2.2 zugrunde. Eine gute Übereinstimmung experimenteller und theoretischer Werte ist erkennbar. Drei besonders grobe Ausreißer können detektiert werden: Für die Gene *pta* und *pckA* wurde jeweils ein relativ großer Wert $\theta(\text{Acetat}, \text{Glucose})$ ermittelt, da sie strukturell die Aufnahme von Acetat gewährleisten können und somit bei Wachstum auf Acetat eine größere Relevanz haben sollten. Die beobachtete Herabregulation dieser Gene kann nur damit erklärt werden, daß die parallele Acetat-Synthase (Gen *acs*) die Acetataufnahme allein übernimmt, während *pta* und *pckA* vermutlich für die Acetat-Ausscheidung zuständig sind und dann in der Tat auf Glucose höhere Relevanz haben [106]. Die signifikant erhöhte Expression des Gens *aspA* (Aspartase) bei Wachstum auf Acetat wurde von θ klar unterschätzt. Eine mögliche Erklärung, die im Zusammenhang mit dem Redoxhaushalt steht, wird im Abschnitt 6.4 gegeben. Läßt man die drei Ausreißer unbeachtet, so ergibt sich für die verbleibenden 47 Punkte in Abb. 6.5a eine sehr gute lineare Korrelation mit $r^2=0.60$.

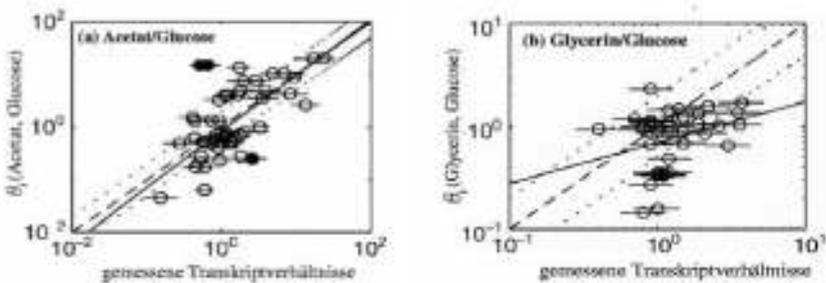


Abb. 6.5 (a): Gegenüberstellung gemessener und mittels $\beta(\text{Acetat, Glucose})$ prädizierter differentieller Expression metabolischer Gene bei exponentiellem Wachstum auf Acetat bzw. Glucose (logarithmische Darstellung). Ausgefüllte Kreise entsprechen den groben Ausreißern, horizontale Linien in den Kreisen deuten das 95%-Konfidenzintervall der Messung an (Anhang E). Die durchgezogene Linie erhält man durch lineare Regression. Die gestrichelte Linie entspricht der Identität. Die beiden gepunkteten Linien zeigen zweifache Abweichung nach oben bzw. unten.
(b) Gemessene und prädizierte Genexpression bei Glycerin/Glucose (analog nach (a)).

Die experimentellen und theoretischen Transkriptverhältnisse für Glycerin/Glucose korrelieren nicht so gut (Abb. 6.5b). Mehrere Gründe kommen dafür in Betracht: (i) Die Verwertungswege von Glucose und Glycerin sind stärker überlappend als bei Acetat, da beide Substrate letztlich glykolytisch abgebaut werden. Insbesondere unter dem Blickwinkel der relativ geringen Präzision der DNA-Microarray-Chiptechnologie [105,106] sind die resultierenden Veränderungen in den CEFs bzw. in der Genexpression für eine Korrelationsanalyse vermutlich nicht signifikant genug (vgl. Skalen in Abb. 6.5a und 6.5b). (ii) Es existieren bisher unbekannte Routen. In diesem Falle kann der Vergleich von experimentellen und theoretischen Transkriptdaten Hypothesen liefern, wo das Netzmodell unvollständig ist. (iii) Die Expression erfolgt bei Glycerin nicht im Sinne der strukturellen Relevanzen (CEF's), z.B. aufgrund spezifischer regulatorischer Zwänge.

Zusammenfassend können wir festhalten, daß die aus den CEF's ermittelten, theoretischen Transkriptquotienten β den Einfluß der Netzstruktur auf die Genregulation quantitativ wiedergeben können, zumindest in einem groben Maßstab. Man beachte, daß die CEF's bzw. β -Werte ebenso wie mRNA-Transkriptdaten nicht direkt mit den tatsächlich auftretenden Flüßen korrelieren müssen, obwohl ein indirekter Zusammenhang allemal bestehen sollte. Die hierarchisch untergeordnete Ebene der metabolischen Regulation (z.B. allosterische Kontrolle der Enzymaktivität) kann die Feinjustierung der Flüße (auch kurzfristig) vornehmen und so für stöchiometrisch optimale Flußverteilungen sorgen. Eine rein nach optimaler Ausbeute ausgerichtete Genexpression im Sinne der FBA-Methode (Abschnitt 2.4.2) ist aber unwahrscheinlich: Verwendet man für die oben durchgeführte Analyse nur die jeweils optimalen EMen ergibt sich z.B. für Acetat/Glucose eine viel schlechtere Übereinstimmung in prädizierten und gemessenen Transkriptverhältnissen [148]. Flexibilität scheint also ein wichtiges

Kriterium für die (langfristige) zelluläre Regulation zu sein, um auf kurzfristige Veränderungen – z.B. bei der Akkumulation eines Metaboliten – reagieren zu können.

Auf zwei generelle Probleme der theoretischen Transkriptverhältnisse θ sei noch hingewiesen: Nur einige Umweltgrößen, wie die „Substratverfügbarkeit“, lassen sich stöchiometrisch beschreiben und auf eine zugeordnete Menge von EMen abbilden. Die Substratkonzentration kann bspw. nicht berücksichtigt werden, obwohl diese vermutlich einen Einfluß auf die Genexpression haben wird. Ein anderes Problem sind Isoenzyme, da hier die Expressionsdaten zweier Gene der Relevanz *einer* Reaktion gegenüberstehen (bei multifunktionalen Enzymen ist es genau umgekehrt). Als grobe Annäherung kann entweder der Durchschnitt der Expressionsverhältnisse der Isoenzyme verwendet werden oder man bezieht sich nur auf das Gen bzw. Enzym, das nach bisherigen Kenntnissen quantitativ dominierend über das andere ist.

Bei den gewichteten Relevanzen σ^{P+ATP} haben wir bisher ein Effizienzkriterium mit Bezug auf Biomasse- und ATP-Ausbeute verwendet, das aus Sicht der Zelle physiologisch sinnvoll erscheint. Prinzipiell kann man beliebige Zielfunktionen in die Definition von Effizienzen bzw. Gewichten einbauen. Eine Herangehensweise im Sinne des Metabolic Engineering wäre folgende: Man wählt als Effizienzkriterium die maximale Synthese eines gewünschten Produktes und berechnet dafür die entsprechenden gewichteten Relevanzen (vgl. Anhang D). Ein Vergleich mit den „natürlichen“ Relevanzen σ^{P+ATP} könnte dann Hinweise geben, an welchen Stellen über genetische Modifikationen eingegriffen werden sollte.

6.4 Strukturelle Kopplungen zwischen Reaktionen

Bisher haben wir die Relevanz einzelner Reaktionen im Bezug auf spezifische Umweltbedingungen oder Funktionalitäten untersucht. Wir wollen nun unseren Blickwinkel dahingehend verschieben, daß wir die Relevanzen von Reaktionen für die Operation *einer beliebig* vorgegebenen Reaktion R_x untersuchen. Dabei betrachten wir zunächst keine Einschränkung bzgl. der Umweltbedingungen. Praktisch haben wir es also mit dem Szenario $S=[R_x]$ zu tun. Wie in Abschnitt 6.1 kennzeichnet E^{R_x} die Menge der EMen, in denen R_x involviert ist, und E^{R_x} die entsprechende Matrix dieser EMen. $E_{bin}^{R_x}$ ist die Binärdarstellung von E^{R_x} .

Analog zu Gl. (6.4) können wir die relAH jeder Reaktion in den EMen in E^{R_x} bestimmen:

$$h^{R_x} = \frac{1}{|E^{R_x}|} E_{bin}^{R_x} \mathbf{1} = \frac{1}{|E_{bin}^{R_x}|} E_{bin}^{R_x} \begin{pmatrix} 1 \\ \dots \\ 1 \end{pmatrix} \quad (6.13)$$

In Abschnitt 6.2.1 hatten wir h_i^S als (ungewichtete) Relevanz der i -ten Reaktion für das Szenario S betrachtet. Dementsprechend ist $h_i^{R_x}$ die Relevanz von i für die Operation von R_x . Mit dem besonderen Fokus auf R_x bezeichnen wir $h_i^{R_x}$ auch als (ungewichtete) *strukturelle*

Kopplung von i zur Reaktion R_x . Man beachte, daß im allgemeinen Fall $h_{R_x}^{R_i} \neq h_{R_i}^{R_x}$ gilt. So ist z.B. in NB1 (Abb. 2.8): $h_{R_2}^{R_1} = 1 \neq h_{R_1}^{R_2} = 0.5$. Sämtliche Kopplungen können übersichtlich als numerische oder graphische *Kopplungsmatrix* dargestellt werden. Abb. 6.6 zeigt beide Darstellungen, die sich aus den EMen von NB2 (Abb. 2.9) ergeben.

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
R1	1	0.43	0.71	0.43	0.43	0.43	0.43	0.57	0.29	0.43
R2	0.75	1	0.5	0.25	0.25	0.25	0.25	0.5	0.25	0.25
R3	0.83	0.33	1	0.5	0.33	0.33	0.5	0.5	0.5	0.5
R4	1	0.33	1	1	0.33	0.33	1	0.67	0	1
R5	1	0.33	0.67	0.33	1	0	0.33	0.33	0.33	0.33
R6	1	0.33	0.67	0.33	0	1	0.33	0.67	0.33	0.33
R7	1	0.33	1	1	0.33	0.33	1	0.67	0	1
R8	1	0.5	0.75	0.5	0.25	0.5	0.5	1	0.25	0.5
R9	0.67	0.33	1	0	0.33	0.33	0	0.33	1	0
R10	1	0.33	1	1	0.33	0.33	1	0.67	0	1

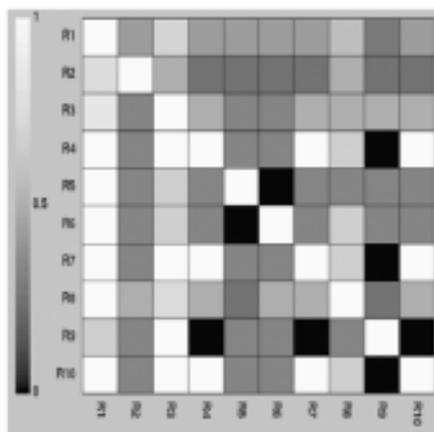


Abb. 6.6: Kopplungsmatrix für NB2: numerische und graphische Darstellung (M_{R_i} repräsentiert $h_{R_i}^*$).

Anhand der *strukturellen Kopplungen* können wir untersuchen, wie stark einzelne Reaktionen operativ miteinander verknüpft sind. Als eine strenge Kopplung wurden bereits in Abschnitt 2.4.1 die *Enzyme Subsets* (ESs) erwähnt. Das ist eine Menge von Reaktionen, die gemeinsam operieren müssen. In der Notation der strukturellen Kopplung bedeutet das, daß für jedes Paar $\{R_x, R_y\}$ aus dem ES gelten muß: $h_{R_x}^{R_y} = h_{R_y}^{R_x} = 1$, d.h. jede Reaktion aus dem ES benötigt für ihre Funktion zwingend alle anderen aus dem ES, wie z.B. $\{R4, R7, R10\}$ in NB2 (Abb. 6.6). Nicht selten können „einseitig“ strenge Kopplungen zwischen Reaktionspaaren identifiziert werden. In einem solchen Fall ist die erste Reaktion für die zweite essentiell ($h_{R_x}^{R_y} = 1$), was umgekehrt aber nicht gilt ($h_{R_y}^{R_x} < 1$). In NB2 ergibt sich z.B. für $\{R3, R9\}$ so eine hierarchische Konstellation. Dort ließe sich also präzisieren, daß, sofern R9 aktiv ist, auch R3 operieren muß. R3 dagegen benötigt R9 nur in 50% seiner EMen.

Der entgegengesetzte Fall von strengen Kopplungen sind solche Paare von Reaktionen, die in keinem EM gemeinsam auftreten, d.h. für die $h_{R_x}^{R_y} = h_{R_y}^{R_x} = 0$ gilt. $\{R5, R6\}$ und $\{R9, R10\}$ seien exemplarisch für NB2 genannt. Wir bezeichnen solch ein Paar als „sich (in den EMen) gegenseitig ausschließende Reaktionen“. Sie sind entweder funktional losgelöst voneinander oder in zueinander parallelen Routen lokalisiert. Entfernt man eine der beiden Reaktionen, bleiben alle EMen mit Beteiligung der anderen erhalten. Zwei parallele irreversible Reaktionen (von Isoenzymen) sind ein triviales Beispiel sich gegenseitig ausschließender Reak-

tionen [75]. Im allgemeinen folgt aus der Beteiligung der einen Reaktion an einer konkreten Flußverteilung aber nicht, daß die jeweils andere unbeteiligt ist, da sich die Flußverteilung aus der Summe aller EMen ergibt (Gl. (6.1)).

Die meisten strukturellen Kopplungen sind gewöhnlich nicht streng klassifizierbar, sondern schwanken zwischen stark (nahe bei 1) und schwach gekoppelt (nahe bei 0). Starke Kopplungen weisen auf eine gemeinsame Regulation hin, wie es bei ESs häufig der Fall ist [75,143]. Schwache Kopplungen lassen dagegen auf eine unabhängige Regulation schließen, wie z.B. bei den drei Aspartokinase-Isoenzymen in *E. coli* [75,83]. Isoenzyme dienen wohl prinzipiell als zusätzliche Freiheitsgrade für die genetische Regulation eines Flusses.

Eine Aufgabe für die Zukunft könnte sein, alle Reaktionen des Netzes anhand ihrer strukturellen Kopplungen mit einem Cluster-Algorithmus ähnlich wie in [42,57] hierarchisch zu gruppieren. Ein Problem besteht darin, daß die Kopplungen wie oben erwähnt nicht symmetrisch sind, was von Cluster-Algorithmen aber i.a. vorausgesetzt wird. Man müßte deshalb abgeleitete oder alternative (symmetrische) Kopplungsmaße verwenden, wie z.B. den Korrelationskoeffizienten für jeweils zwei Reaktionen. Letzterer wird bzgl. aller EMen ermittelt, erwies sich aber als weniger gut geeignet, Kopplungen quantitativ zu erfassen. Generell könnte man auch, analog zu den gewichteten Relevanzen, gewichtete Kopplungsmaße definieren, wo die EMen bzgl. eines vorzugebenen Kriteriums noch gewichtet werden.

Bisher haben wir strukturelle Kopplungen bzgl. *aller* EMen bestimmt, d.h. jede denkbare Konstellation (z.B. bzgl. Substratverfügbarkeit) wurde in (6.13) berücksichtigt. Wollen wir die generelle strukturelle Abhängigkeit zweier Reaktionen untersuchen, ist das auch notwendig. Für bestimmte Anwendungen ist es aber sinnvoll, Kopplungen, analog wie für die Relevanzen in Abschnitt 6.2, bzgl. eines eingeschränkten Szenario S zu betrachten. Wir kennzeichnen diese *partiellen Kopplungen* zu Rx als $h^{Rx,S}$, die sich wie in (6.4) ergeben. (Hinweis: Superskript „ Rx, S “ steht letztlich für das zusammengesetzte Szenario $\{Rx \text{ UND } S\}$).

Mit den partiellen Kopplungen können wir nun spezielle Fragen angehen, wie: „Angenommen, der Organismus wächst auf einem bestimmten Substrat, und es sei zusätzlich bekannt, daß die Reaktion Rx involviert ist. Welche Konsequenzen hat das für die Relevanzen der anderen Reaktionen?“ Genau diese Frage kann sich bspw. bei der Interpretation von Transkriptionsdaten stellen. So hatten wir in Abschnitt 6.3 erwähnt, daß für *E. coli* bei Wachstum auf Acetat unerwartet eine erhöhte Expression für die Aspartase auftrat, die durch die gewichteten Relevanzen nicht erklärt werden konnten. Um die Ursachen oder/und Konsequenzen dieser Beobachtung zu untersuchen, wurden die strukturellen Kopplungen $h^{Asp:Fum, \{Ac, Asp\}}$ zur Aspartase-Reaktion $Asp::Fum$ im *E. coli* - Netz berechnet. Das bei weitem auffälligste Ergebnis war dabei, daß sich für die Transhydrogenase $h^{Transhydro, \{Ac, Asp\}} = 0$ ergab, während dieselbe Reaktion ohne Einschränkung auf $Asp::Fum$ eine nicht geringe Auftretshäufigkeit von $h^{Ac, Asp, Asp} = 0.53$ aufwies. Bzgl. Wachstum auf Acetat sind $Asp::Fum$ und $Transhydro$ also sich gegenseitig ausschließende Reaktionen. Damit können wir vermuten, daß die Transhy-

drogenase-Reaktion im durchgeführten Experiment wahrscheinlich eine geringe Relevanz hatte (siehe auch Anhang E). Tatsächlich läßt sich mit der Aspartase ein relativ kleiner Zyklus verwirklichen, der eine zur Transhydrogenase äquivalente Funktion realisieren kann (Abb. 6.7). Dieser Zyklus besteht aus 5 Reaktionen und leistet im Nettoumsatz die Konvertierung von NADPH zu NADH. Alle beteiligten Gene wiesen bei Acetat im Vergleich zur Glucose auch eine mehr oder weniger stark erhöhte Expression auf (Anhang E).

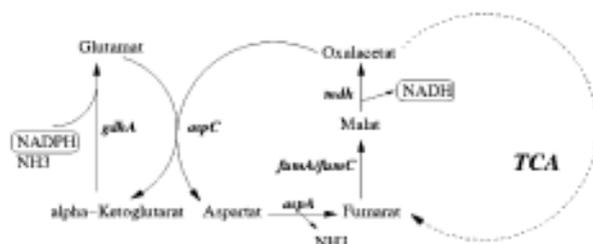


Abb. 6.7: Transhydrogenase-Funktion eines Mini-Zyklus im Stoffwechsel von *E. coli* unter Einbeziehung der Aspartase. In einem Umlauf wird netto ein NADPH zu ein NADH transformiert. Beteiligte Reaktionen: Aspartat-Transaminase (Gen *aspC*), Glutamat-Dehydrogenase (*gdhA*), Aspartase (*aspA*), Malat-Dehydrogenase (*mdh*), Fumarase (*fumA* / *fumC*).

Die mögliche Involvierung dieses Zyklus läßt zwei Fragen aufkommen. Zunächst, wofür wird die Transhydrogenase-Funktion des Zyklus bei Acetat als Substrat benötigt? Eine mögliche Antwort könnte folgende sein: Bei Wachstum auf Acetat wird der TCA intensiver genutzt, als bei Glucose. Dadurch erhöht sich der Fluß durch die Isocitrat-Dehydrogenase, die NADPH produziert. Sofern nun diese zwangsläufige Produktion von NADPH die benötigten Menge für anabolische Reaktionen überschreitet, muß ein Ausgleich erfolgen, der über eine Transhydrogenase-Funktion in Richtung NADH gewährleistet wird. NADH selbst wird in reichlichen Mengen für die Atmung benötigt. Unterstützt wird diese hypothetische NADPH-Überschuss-Situation durch die Untersuchungen in [106], bei der für Acetat eine verstärkte Expression eines Gens festgestellt wurde, das möglicherweise ein NADP-abhängiges Malat-Enzym kodiert (im *E.coli* Netzmodell wurde nur das NAD-abhängige Malat-Enzym berücksichtigt).

Eine zweite Frage, die sich stellt, ist, warum für die Konvertierung von NADPH zu NADH nicht die Transhydrogenase selbst verwendet wird. In der Arbeit [13] wurde gezeigt, daß die membranständige Transhydrogenase (Gen *putAB*) wohl vornehmlich für die andere Richtung zuständig ist, also für die Synthese von NADPH aus NADH. Es gibt zwar eine zweite (cytoplasmatische) Transhydrogenase (Gen *UdhA*), die in der gleichen Richtung wie der Aspartase-Zyklus arbeitet [13], aber vermutlich ist letzterer dennoch zumindest unterstützend aktiv.

Die skizzierten Anwendungen der strukturellen Kopplungen demonstrieren nochmals, welches Potential die kombinatorische Elementarmodenanalyse auch für die Interpretation genomischer Daten (*functional genomics*) hat.

6.5 Kombinatorische Elementarmodenanalyse im *FluxAnalyzer*

Nach der Berechnung von EMen kann man im *FluxAnalyzer* mit Hilfe des Selektions-Tools (Abschnitt 4.4) ein beliebiges Szenario S über 8 Kriterien spezifizieren. Daraufhin wird die entsprechende Menge E^S der zulässigen EMen selektiert, die dann den in diesem Kapitel beschriebenen kombinatorischen/statistischen Analysen zugänglich sind (Abb. 4.6 rechts):

- Berechnung relativer Auftrittshäufigkeiten; essentielle und ausgeschlossene Reaktionen können so identifiziert werden
- Berechnung gewichteter Relevanzen (CEF's; mit Vorgabe des Effizienzkriteriums und des Substrates) und theoretischer Transkriptverhältnisse (θ)
- Berechnung struktureller Kopplungen und Ausgabe der Kopplungsmatrix (Abb. 6.6 wurde mit dem *FluxAnalyzer* erstellt)

7 Minimale Schnittmengen: Elementare Fehlermoden und effiziente Eingriffe in metabolischen Reaktionsnetzwerken

EMen sind *kleinste (unzerlegbare) funktionale Einheiten* eines Stoffwechselnetzes. In diesem Kapitel werden wir unsere Perspektive dahingehend verändern, daß wir nach kleinsten *Eingriffen (Schnitten)* im Reaktionsnetz suchen, die eine bestimmte Funktionalität ausschalten. Zunächst werden wir das Konzept und einen Algorithmus zur Berechnung von solchen *minimalen Schnitten* in Reaktionsnetzen erarbeiten, ehe dann weitreichende Anwendungen vorgestellt werden.

7.1 Definition von minimalen Schnitten

Zur Illustration der einzuführenden Begriffe betrachten wir das in Tab. 7.1 gezeigte fiktive Reaktionsnetz, welches im folgenden als NB4 bezeichnet wird. NB4 besteht aus 6 internen Metaboliten (externe Metabolite außerhalb der Systemgrenzen wurden nicht mit eingezeichnet) und 9 Reaktionen, von denen lediglich R4 reversibel ist. Die Stöchiometrie der Reaktionen kann leicht der Graphik entnommen werden. Wir werden hier insbesondere an der Reaktion ZR interessiert sein, die wir daher als *Zielreaktion* bezeichnen. Die Rate dieser Reaktion drückt aus, welche Menge des Metaboliten X synthetisiert und mit ZR über die Systemgrenzen hinaus exportiert wird.

Wollen wir die qualitativ unterschiedlichen Pathways in NB4 identifizieren, die eine Synthese von X ermöglichen, so bietet sich wieder die Berechnung der EMen an. In NB4 ergeben sich insgesamt vier EMen, von denen drei die Reaktion ZR involvieren, d.h. wo X vom System produziert und exportiert wird (Tab. 7.1).

Man nehme nun an, die Produktion von X soll durch gezielte Manipulationen verhindert werden. Als Ziel dieser Manipulationen würde man fordern, daß nach dem Eingriff keine ausgeglichene (steady-state) Flußverteilung mehr realisierbar sein soll, in der ZR operiert. Denn dann wäre in der Tat eine kontinuierliche Produktion von X unmöglich. Als struktureller Eingriff käme die Inaktivierung (sinnbildlich: das *Herausschneiden*) einer oder mehrerer Reaktionen in Frage, was biologisch dem Deletieren von Genen der katalysierenden Enzyme oder anderen Manipulationen zur Inhibierung der Enzymaktivität entsprechen würde. Das führt uns zur Definition einer *Schnittmenge*:

Definition (7.1): In einem Reaktionsnetz sei eine Zielreaktion vorgegeben. Eine Menge C von Reaktionen wird bzgl. der Zielreaktion *Schnittmenge* genannt, wenn nach dem Entfernen aller in C enthaltenen Reaktionen keine realisierbare steady-state Flußverteilung im Netz mehr existiert, welche die Zielreaktion involviert.

(Hinweis: Der Begriff *Schnittmenge* (abgekürzt SM) sollte hier nicht mit dem mengentheoretischen Schnitt von Mengen verwechselt werden. Statt dessen ist hier eine Menge herauszuschneidender Reaktionen gemeint. Eindeutiger ist da der Begriff *cut set* aus dem Englischen).

Reaktionsnetz NB4		R1	R2	R3	R4	R5	R6	R7	R8/ZR
Elementarmoden (hervorgehoben: mit Beteiligung von ZR)									
EM1	1	1	1	-1	0	0	0	0	0
EM2	1	0	0	0	0	1	1	1	1
EM3	2	1	1	0	1	0	0	0	1
EM4	1	0	0	1	1	0	0	0	1
Minimale Schnittmengen (Zielreaktion: ZR)									
MSM0									×
MSM1	×								
MSM2					×	×			
MSM3					×		×		
MSM4					×			×	
MSM5		×		×		×			
MSM6			×	×		×			
MSM7		×		×			×		
MSM8			×	×			×		
MSM9		×		×				×	
MSM10				×	×			×	

Tab. 7.1: Netz-Beispiel NB4: Reaktionsnetzwerk, Elementarmoden und minimale Schnittmengen.

Eine *triviale* SM ist natürlich immer die Zielreaktion selbst: [ZR]. Es gibt aber einige Gründe, warum wir an anderen potentiellen SMn interessiert sein könnten: (1) Möglicherweise möchte man Reaktionen am Anfang eines entsprechenden Pathways schneiden, um unnötige Reaktionsschritte im Netz zu vermeiden (in NB4 wäre ZR am Ende des Pathways). (2) Manchmal können „Zielreaktionen“ nur indirekt inaktiviert werden. Ein Beispiel ist die (kumulierte) Biomassessynthese, die in stöchiometrischen Modellen oft als Reaktion behandelt wird, aber keiner realen Reaktion und somit keinem Gen oder Enzym zugeordnet ist. (3) Eventuell möchten wir die strukturellen „Geführungspotenziale“ bzgl. einer vorgegebenen Funktionalität erkennen. (4) Man kann obige Definition leicht auf *mehrere* Zielreaktionen erweitern (s.u.), deren gleichzeitiges Versagen oft effizienter durch das Schneiden anderer Reaktionen erreicht werden kann, als durch das Deletieren der Zielreaktionen selbst.

Eine nicht sehr intelligente Wahl für eine SM wäre das Herausschneiden aller Reaktionen außer ZR selbst. So ist z.B. $CI = \{R5, R8\}$ in NB4 völlig ausreichend, um einen Fluß entlang ZR und damit die Produktion von X zu unterbinden. CI ist sogar sehr effizient, da das Entfernen von nur R5 oder nur R8 eine Funktion der ZR nicht ausschließen kann, d.h. {R5} und {R8} sind keine SMn. Wir haben somit eine *minimale* SM gefunden, weil keine echte Teilmenge von CI eine SM ist:

Definition 7.2: Eine Schnittmenge C wird *minimale Schnittmenge* (MSM; engl.: *minimal cut set*) genannt, wenn keine echte Teilmenge von C eine Schnittmenge ist.

Eine weitere SM in NB4 wäre $C2=\{R2,R4,R6\}$, die ebenfalls minimal ist. Im Vergleich mit $C1$ zeigt sich, daß sich die Minimalität einer MSM also nicht auf die Anzahl ihrer Elemente bezieht, sondern darauf, daß kein Element ausgeschlossen werden können, ohne daß die Eigenschaft einer SM verloren geht. Diese Charakteristik ist der Unzerlegbarkeit der EMEN ((B3) in (2.13)) ganz analog. Eine dritte SM bzgl. ZR wäre $C3=\{R2,R5,R7\}$. Diese SM wäre allerdings nicht minimal, da bereits $\{R5,R7\}$ eine SM ist und somit R2 eine redundante Deletion darstellt. Damit wäre $C3$ im Sinne unseres Anliegens weniger effizient. $C4=\{R1\}$ ist in NB4 die einzige MSM, die nur eine Reaktion enthält (außer der trivialen SM). Einzelne Reaktionen, die für sich bereits eine MSM darstellen, sind die *essentiellen* Reaktionen bzgl. der Zielreaktion (vgl. mit Kap. 6.2) und somit besonders geeignete Schnittkandidaten.

In NB4 gibt es insgesamt 11 MSMn, einschließlich der trivialen (Tab. 7.1). Ein Algorithmus zur Berechnung der MSMn wird später gegeben. Zunächst noch einige Bemerkungen:

Bemerkung (1): Das Entfernen aller Reaktionen einer MSM *garantiert* eine Fehlfunktion der Zielreaktion, solange die angenommene Netzstruktur korrekt ist. Das gilt nicht für echte Teilmengen einer MSM. Allerdings kann es passieren, daß unter physiologischen Bedingungen doch bereits eine Teilmenge einer MSM zum Versagen der ZR ausreicht. Ursache können regulatorische und kapazitive Restriktionen des Netzes sein, die hier nicht berücksichtigt werden. Führt z.B. das Entfernen einer Reaktion einer MSM durch regulatorische Zwänge zum physiologisch bedingten Abschalten einer anderen Reaktion einer MSM, dann wäre diese MSM eigentlich nicht mehr minimal. Sprechen wir im folgenden also über MSMn und Netzwerkfehlfunktionen, so ist das grundsätzlich von einer rein strukturellen Perspektive aus zu sehen. Sollten allerdings regulatorische Zwänge bekannt sein, so lassen sich diese bei der Berechnung der MSMn berücksichtigen (s.u.). Fragestellungen, die auf eine graduelle Änderung der Performance abzielen (z.B. Reduktion der maximalen Produktausbeute auf einen Wert größer Null) betrachten wir hier nicht.

Bemerkung (2): Die Inaktivierung aller Reaktionen einer MSM unterdrückt die Zielreaktion garantiert. Allerdings können andere Pathways aktiv bleiben. Entfernt man z.B. in NB4 $MSM2=\{R5,R6\}$, so können die Metabolite B und C trotzdem noch aus A produziert werden. In einigen Applikationen wird es gerade das Ziel sein, solche MSMn zu finden, welche die Zielreaktion unterdrücken, dabei aber andere Pathways „verschonen“.

Bemerkung (3): Betrachten wir $MSM4=\{R5,R8\}$, so ist leicht einzusehen, daß X nicht mehr produziert werden kann, da alle Pathways zu X unterbrochen sind. Für $MSM6=\{R3,R4,R6\}$ ist das dagegen nicht mehr so offensichtlich, da eine Synthese von X über R1, R2 und R5 noch realisierbar erscheint. MSMn garantieren aber immer, daß die Funktion nur noch mit einer unausgeglichenen Flußverteilung möglich ist. Eine kontinuierliche Produktion von X nach

Entfernen von **MSM6** würde zur Akkumulation von Metabolit B führen und kann somit aus physiologischen Gründen nicht dauerhaft aufrechterhalten werden.

Bemerkung (4): Bisher sind wir von nur *einer* Zielreaktion ausgegangen. Man kann die Definitionen 7.1 und 7.2 für (M)SMn aber leicht so erweitern, daß allgemeiner eine *Menge* an Zielreaktionen berücksichtigt werden kann. In diesem Fall fordert man dann, daß nach dem Entfernen einer (M)SM keine steady-state Flußverteilung mehr realisierbar ist, die *irgendeine* der Zielreaktionen involviert. Möchte man z.B. *jede* ausgeglichene Flußverteilung im Netz ausschließen, so sind *alle* Reaktionen des Netzes Zielreaktionen. Sämtliche hier durchgeführte Betrachtungen ergeben sich daher völlig analog für den Fall mehrerer Zielreaktionen.

7.2 Algorithmus zur Berechnung minimaler Schnittmengen

Die MSMn bzgl. einer definierten Zielreaktion gehören der Potenzmenge der Menge der Reaktionen an und sind eindeutig bestimmt. In NB4 kann man die MSMn noch relativ leicht finden, in größeren Netzwerken benötigen wir jedoch ein systematisches Berechnungsschema, das hier abgeleitet werden soll. Dieses Schema muß garantieren,

- (a) daß die berechneten MSMn tatsächlich Schnittmengen sind, d.h. jede MSM „zerstört“ alle stationären Flußverteilungen, bei denen die Zielreaktion beteiligt ist.
- (b) daß jede MSM tatsächlich *minimal* ist.
- (c) daß *alle* MSMn gefunden werden.

Im hier entwickelten Algorithmus wird für Bedingung (a) wiederum die Tatsache ausgenutzt, daß sich jede realisierbare steady-state Flußverteilung \mathbf{r} als nichtnegative Linearkombination der Elementarmoden \mathbf{e}^j des Netzes darstellen läßt (2.23):

$$\mathbf{r} = \sum_j \gamma_j \mathbf{e}^j \quad ; \quad \gamma_j \geq 0 \quad (7.1)$$

Wollen wir also garantieren, daß die Rate r_k der Zielreaktion in allen Flußverteilungen \mathbf{r} Null ist, dann muß jeder EM an der k -ten Stelle eine Null haben. Damit gilt:

Lemma 7.1: Eine (minimale wie nichtminimale) Schnittmenge C erfüllt folgende Bedingung: Für jeden EM, der die Zielreaktion (mit einer Rate ungleich Null) involviert, gibt es mindestens eine Reaktion in C , die ebenfalls in diesem EM involviert ist.

Lemma 7.1 garantiert, daß alle EMen mit Beteiligung der Zielreaktion verschwinden, wenn die Reaktionen einer SM aus dem Netz entfernt werden. Wir sagen auch, daß eine SM alle EMen der Zielreaktion „abdeckt“. Lemma 7.1 nutzt explizit Eigenschaft (2.16) aus: Nach dem Entfernen von Reaktionen konstituiert sich die neue Menge der EMen aus all denjenigen des ursprünglichen Netzes, bei denen diese Reaktionen nicht involviert sind. Eine Neuberechnung

der EMen nach dem Entfernen von Reaktionen bleibt somit erspart [140]. Da die *extremen Pathways* diese Eigenschaft nicht besitzen (Abschnitt 2.4.4), können sie i.a. nicht für die Berechnung der MSMn mit dem unten angegebenen Algorithmus herangezogen werden.

Lemma 7.1 legt ein sukzessives Verfahren zur Konstruktion einer SM nahe: man startet mit einer Reaktion und fügt zu dieser *vorläufigen* SM solange weitere Reaktionen hinzu, bis alle EMen mit Beteiligung der Zielreaktion durch die dann entstandene SM abgedeckt sind. Unter zusätzlicher Berücksichtigung der geforderten Minimalität (b) und Vollständigkeit (c) ergibt sich folgender Pseudocode für den Algorithmus zur Bestimmung der MSMn:

-
- 1) Berechne alle EMen im gegebenen Netz
 - 2) Definiere die Zielreaktion ZR
 - 3) Wähle alle EMen aus, in denen die Rate von ZR nicht Null ist und speichere diese EMen als binäres Array em_zr ($em_zr[j][i]=1$ falls die Reaktion i im EM j involviert ist)
 - 4) Initialisiere die Arrays msm und vor_msm wie folgt (msm enthält die MSMn; vor_msm die vorläufigen MSMn; beide Arrays enthalten Mengen von Reaktionsindizes): hänge $\{i\}$ an msm an, falls Reaktion i essentiell für ZR ist (d.h. wenn $em_zr[j][i]=1$ für alle EMen j); sonst hänge $\{i\}$ an vor_msm an
 - 5) For $j=2$ TO MAX_MSM_GRÖSSE // MAX_SM_GRÖSSE = maximale Kardinalität der MSMn
 - 5.1) $neu_vor_msm=[]$ // für neue vorläufige MSMn
 - 5.2) FOR $i=1$ TO q // q = Anzahl Reaktionen im Netz
 - 5.2.1) Entferne alle (Index-)Mengen von vor_msm , die Reaktion i enthalten
 - 5.2.2) Finde alle Mengen von Reaktionen in vor_msm , die mindestens einen EM von em_zr nicht abdecken, an dem Reaktion i beteiligt ist; füge in diese Mengen die Reaktion i mit ein und speichere diese neuen vorläufigen MSMn im Array $temp_vor_msm$
 - 5.2.3) Verwerfe alle $temp_vor_msm$, die eine Obermenge der bereits gefundenen (und in msm) gespeicherten MSMn sind
 - 5.2.4) Finde von den beibehaltenen Mengen in $temp_vor_msm$ diejenigen, die jetzt alle EMen in em_zr abdecken und hänge sie an msm an; hänge alle anderen an neu_vor_msm an
 - ENDFOR
 - 5.3) IF neu_vor_msm leer
 - 5.3.1) Break
 - ELSE
 - 5.3.2) $vor_msm=neu_vor_msm$
 - ENDIF
- ENDFOR
- 6) Resultat: das Array msm enthält die MSMn.
-

Erläuterungen/Anmerkungen zum Algorithmus:

- Alle verwendeten Arrays speichern Mengen von Reaktionsindizes (1 ... q), die effektiv in Bit-Arrays abgelegt werden können und somit schnelle Bitoperationen zulassen.
- Reaktionen eines Enzyme Subsets (ESs) treten in EMen nur gemeinsam auf (Abschnitt 6.4). Bei den MSMn dreht sich diese Beziehung um: in keiner MSM ist mehr als eine Reaktion eines ESs enthalten. Die minimale Eigenschaft würde sonst sofort verlorengehen, da nach Löschen einer Reaktion eines ESs schon per Definition folgt, daß die anderen Reaktionen dieses ESs nicht mehr operativ sein können. Man kann das leicht anhand der MSMn in NB4 überprüfen, wo mit {R2,R3} und {R6,R7,R8} zwei ESs enthalten sind. So

tritt in keiner MSM, die R6 enthält, R7 oder R8 auf. Allerdings gibt es für jede MSM, die R6 enthält, äquivalente MSMn, wo R6 durch R7 bzw. R8 ausgetauscht ist. Diese generelle Eigenschaft kann zur Steigerung der Effizienz in einem Vorverarbeitungsschritt (vor der Hauptschleife) ausgenutzt werden. Dabei löscht man alle Reaktionen (Spalten in *em_zr*) eines ESs mit Ausnahme eines „Vertreters“. Nach der Berechnung können leicht die äquivalenten MSMn gebildet und an *msm* angehängt werden. Enthält also ein ES e Reaktionen und ist der Vertreter an z MSMn beteiligt, so werden $z(e-1)$ äquivalente MSMn an *msm* angehängt. Diese Vorgehensweise kann u.U. beträchtliche Einsparungen an Speicher und Rechenzeit bringen.

- Deletionen von multifunktionalen Enzymen (wie der Transketolase) lassen sich leicht adäquat berücksichtigen. Dabei muß beachtet werden, daß bei der Deletion eines solchen Enzyms (bzw. dessen Gens) mehrere Reaktionen gleichzeitig im Netz inaktiviert werden. Deshalb faßt man alle Reaktionen eines multifunktionalen Enzyms in *einer* Spalte in *em_zr* zusammen, die für solche EMen eine 1 enthält, an denen mind. eine dieser Reaktionen beteiligt ist.
- Regulatorische Zwänge (siehe Bemerkung (1) oben) lassen sich berücksichtigen, indem solche EMen, die aufgrund bekannter Regelkreise nicht auftreten können, von vornherein aus *em_zr* ausgeschlossen werden (analog wie in [22]). Aufgrund der bekannten Katabolitenrepression lassen sich z.B. für *E. coli* solche EMen sofort ausschließen, in denen Glucose- und Lactoseaufnahme gleichzeitig auftreten. Damit wird die Menge der EMen, die von jeder MSM überdeckt werden muß, reduziert.
- Sollen mehrere Zielreaktionen gleichzeitig berücksichtigt werden (Bemerkung (4)), selektiert man im Schritt (3) alle EMen, in denen mind. eine der Zielreaktionen involviert ist.
- Schritt (5.2.1) vermeidet redundante Kombinationen.
- Die Menge der aktuell abgedeckten EMen kann für jede vorläufige MSM auch separat gespeichert und nach jedem Schritt aktualisiert werden. Das kann Schritt (5.2.2) beträchtlich beschleunigen, allerdings erhöht sich auch der Speicherbedarf immens.
- Beim Test auf Minimalität im Schritt (5.2.3) wird genau wie bei der Berechnung von EMen überprüft, ob eine Menge in einer anderen enthalten ist (Gl. (4.1)). Deshalb können hier praktisch die gleichen Subroutinen wie beim Berechnen der EMen verwendet werden (Abschnitt 4.4), was nochmals die konzeptionelle Analogie von EMen und MSMn unterstreicht.
- Wie die Berechnung der EMen ist auch die Bestimmung der MSMn ein kombinatorisch komplexes Problem, da die Anzahl der MSMn zusammen mit der Anzahl der EMen bei steigender Netzgröße exponentiell wächst (vgl. Kap. 8). Die Anzahl der in Abschnitt 7.3 diskutierten MSMn von *E. coli* (Tab. 7.2) ist jeweils kleiner als die der zugrundeliegenden EMen. Allerdings ist die Anzahl der *vorläufigen* MSMn während der Berechnung bis zu 100-mal höher. Im Gegensatz zu den EMen läßt sich obiger Algorithmus für die MSMn aber ausschließlich mit effizienten Bit-Arrays und Bit-Operationen programmieren. Deshalb wird letztlich in den meisten Fällen der Speicher limitierender Faktor sein.

- Da die benötigte Rechenzeit dennoch schnell ansteigen kann (in Tab. 7.2 bis zu 30 Stunden), ist die Option, die maximale Größe der zu berechnenden MSMn zu beschränken (MAX_MSM_GROESSE, Schritt (5)), ein großer Vorteil. Der Algorithmus garantiert, daß nach der Beendigung der i -ten (äußeren) Schleife alle MSMn mit i oder weniger Elementen gefunden und in *msm* abgelegt worden sind.

	Acetat	Succinat	Glycerin	Glucose
# EMen mit Wachstum	363	3421	9479	21592
# MSMn (Zielreaktion: Wachstum)	245	1255	2970	4225
maximale Anzahl vorläufiger MSMn (während Berechnung)	3563	69628	344196	902769
benötigte Rechenzeit (Intel Pentium, 1MHz; 4 GB RAM)	7 sec	20 min	5.42 h	29.67 h
<i>Fragilitätskoeffizienten F_i einzelner Reaktionen bzgl. Wachstum</i> (in Klammern: Größe der kleinsten MSM, in der die jeweilige Reaktion auftritt)				
<i>F16P::F6P</i> (Fructosebiphosphatase)	1 (1)	1 (1)	1 (1)	0.102 (6)
<i>ATPSynth</i> (ATP-Synthase)	1 (1)	0.325 (3)	0.141 (3)	0.149 (3)
<i>SuccCoA::Succ</i> (Succinyl-CoA-Synthetase)	0.207 (2)	0.145 (2)	0.125 (2)	0.131 (2)
<i>PEP::OxAc</i> (PEP-Carboxylase)	0.128 (2)	0.117 (2)	0.120 (2)	0.143 (2)
<i>Mal::Pyr</i> (Malat-Enzym)	0.5 (2)	0.5 (2)	0.114 (2)	0.123 (2)
<i>R15P::X5P</i> (Ribulose-Phosphat-Epimerase)	0.198 (2)	0.135 (2)	0.128 (2)	0.148 (2)
F (Netzwerk-Fragilität bzgl. Wachstum)	0.783	0.718	0.699	0.643

Tab. 7.2: Überblick über Anzahl und Eigenschaften von MSMn im Zentralstoffwechsel von *E. coli*, jeweils bzgl. Wachstum auf einem der vier verschiedenen Substrate. Die angegebenen Rechenzeiten schließen die Zeit zum Berechnen der EMen nicht mit ein. Weitere Erläuterungen in Abschnitt 7.3.

Umsetzung im FluxAnalyzer

Der oben angegebene Algorithmus zur Berechnung von MSMn wurde im *FluxAnalyzer* implementiert. Beliebige Zielreaktionen können dadurch definiert werden, daß vor der Berechnung die abzudeckende (zu löschende) Menge an EMen durch den Benutzer selektiert wird (Abschnitt 4.4). Vor der Berechnung können außerdem multifunktionale Enzyme deklariert, bestimmte Reaktionen ausgeschlossen und MAX_MSM_GROESSE, also die maximale Kardinalität der MSMn, angegeben werden. Der Algorithmus wurde unter den oben angegebenen Aspekten so effizient wie möglich programmiert. Nach der Berechnung können die MSMn über ein Control Panel (ähnlich dem in Abb. 4.6) in den Flußkarten angezeigt und spezifische Selektionen von bestimmten MSMn ausgeführt werden (z.B. alle MSMn, die eine bestimmte Reaktion (nicht) enthalten). Für solche Selektionen von MSMn können wiederum statistische Eigenschaften bestimmt werden, einschl. des unten eingeführten Fragilitätskoeffizienten. Auf Wunsch können alle diejenigen EMen selektiert werden, die NICHT durch eine gegebene MSM abgedeckt werden, also nach deren Entfernen funktional bleiben.

7.3 Anwendungen

Die konzeptuellen Eigenschaften der MSMn eröffnen eine ganze Reihe nützlicher Anwendungen, die bei der gezielten Suche nach möglichen Eingriffspunkten im Netz oder allgemeiner bei der Analyse der strukturellen Fragilität eines (metabolen) Reaktionsnetzes hilfreich sind. Zur Illustration nutzen wir die MSMn, die sich im Stoffwechselmodell von *E. coli* ergeben (siehe Übersicht in Tab. 7.2). Sie wurden bzgl. der Zielfunktion „Wachstum“ ermittelt – jeweils separat für die vier betrachteten Substrate. Bei der Berechnung der MSMn konnte demnach auf diejenigen in Abschnitt 6.2.2 berechneten EMen zurückgegriffen werden, die *muc* involvierten (zweite Zeile in Tab. 6.2).

7.3.1 Identifikation von Targets zur Unterdrückung zellulärer Funktionen

Die wohl naheliegendste Anwendung der MSMn ist die Suche nach Angriffspunkten, um gezielt bestimmte Funktionen (wie Transport, Synthese, Umwandlung von Metaboliten) des metabolen Netzes unterdrücken zu können. Die Identifizierung von optimalen Targets spielt sowohl bei der Arzneimittelforschung [24] als auch beim Metabolic Engineering [149,170] eine entscheidende Rolle. Die MSMn sind ideal für diesen Zweck geeignet, weil sie die effizienten (da minimalen) Manipulationsmengen darstellen, die eine Zielfunktion garantiert ausschalten. Ein Screening aller MSMn erlaubt dann gezielt nach der geeignetsten Manipulationsmenge zu suchen, die in der Praxis verschiedene Anforderungen erfüllen muß:

- In den meisten Fällen wird es wünschenswert sein, mit möglichst wenigen Eingriffsstellen auszukommen, d.h. man sucht MSMn mit wenigen Elementen.
- Andere Pathways im Netz sollen möglichst gering von den Eingriffen betroffen sein, d.h. man sucht solche MSMn, wo diese Pfade – zumindest strukturell – noch intakt sind (Bemerkung (2)). Welche Pathways nach dem Löschen einer MSM funktional bleiben, erkennt man an den EMen, die nicht durch diese MSM abgedeckt sind. In NB4 enthalten z.B. ~~MSM012134~~ keine Reaktion, die in EM1 involviert ist. Entfernt man eine dieser MSMn, so ist zwar die Zielreaktion außer Kraft gesetzt, EM1 bleibt aber intakt und damit die (prinzipielle) Fähigkeit des Netzes, Metabolite B und C zu produzieren.
- Einige Reaktionen können u.U. nur schwer oder nur mit größerem Aufwand eliminiert oder inhibiert werden, z.B. wenn mehrere Isoenzyme existieren. Dann wird man sich auf jene MSMn konzentrieren, die diese Reaktionen nicht enthalten.

7.3.2 Falsifikation der Netzstruktur und Mutanten-Phänotyp-Prädiktionen

Das Wegschneiden einer kompletten MSM muß bei korrekter Netzstruktur zum Ausfall der Zielfunktion führen. Da es sich hier um eine aus der Struktur abgeleitete Prädiktion handelt, können MSMn auch zur systematischen Verifikation von hypothetischen oder „frisch“ rekon-

struierten Netzen herangezogen werden. Bezüglich der Zielfunktion „Wachstum“ ist die Prozedur klar vorgezeichnet: Enthält eine Menge deletierter Reaktionen (z.B. in Einfach-, Doppel-, Tripel-, ... Mutanten) eine MSM vollständig, dann (Fall 1) sollte dies zu einem nicht lebensfähigen Phänotypen des entsprechenden Organismus führen. Ist dagegen keine MSM in der Deletionsmenge enthalten, dann (Fall 2) wäre Wachstum zumindest strukturell weiterhin möglich. Eine falsche Prädiktion für Fall (1) wäre eine falsch negative Prädiktion und ein Beweis für eine inkorrekte Netzstruktur. Eine inkorrekte Prädiktion für Fall (2) wäre eine falsch positive Prädiktion und damit wegen Bemerkung (1), Abschnitt 7.1, zwar ein Hinweis, aber kein Beweis für eine inkonsistente Netzstruktur. Insbesondere „marginal tolerierbare“ Deletionen, bei denen von einer großen MSM alle Reaktionen bis auf eine entfernt werden, könnten zu falsch positiven Prädiktionen führen. So gibt es für *E. coli* bei Wachstum auf Glucose einige MSMn mit 12 Elementen (Abb. 7.1), und es erscheint unwahrscheinlich, daß die Deletion von 11 Reaktionen einer solchen MSM toleriert würden (eine Überprüfung wäre dennoch reizvoll). Zu beachten ist dabei, daß in oben genannten MSMn tatsächlich alle 12 Deletionen zur Unterdrückung der Zielreaktion beitragen. Bei einem linearen, nichtessentiellen Pathway mit 11 Reaktionen wäre zwar ebenfalls eine Deletion dieser 11 Reaktionen tolerierbar, jedoch wären 10 dieser Deletionen redundant.

Die Verwendung nichtredundanter und damit minimaler Deletionsmengen ist für die Falsifikation von Netzwerken wichtig. Man nehme z.B. an, in NB4 (Tab. 7.1) gäbe es zusätzlich folgende, noch nicht erkannte Reaktion: $Rx: B + C + D = X$. Würde man im Extremfall tatsächlich alle Reaktionen deletieren, außer der unbekanntes Rx und der Zielreaktion ZR, so würde man neben dem hohen experimentellen Aufwand auch indirekt den Pfad von Rx zu ZR zerstören und damit die Netzstruktur nicht falsifizieren können. Dagegen könnte die Deletion von MSM4 tatsächlich zeigen, daß eine bisher unbekannte Reaktion im Netz existiert.

Phänotyp-Prädiktionen von Einfachmutanten von *E. coli* sind bereits mit FBA [29] und Elementarmodenanalyse ([114] und Abschnitt 6.2.2) durchgeführt worden und ergaben sehr gute Übereinstimmungen mit der Lebensfähigkeit realer Mutanten. Solche Analysen sind äquivalent mit der Suche nach MSMn mit nur einem Element (= essentielle Reaktionen). Allerdings ermöglichen die MSMn darüber hinaus die systematische Suche nach sinnvollen Mehrfachmutanten. Die hier berechneten MSMn könnten als eine Basis für eine systematische Falsifizierung des zentralen Katabolismus von *E. coli* dienen, indem man entsprechende Mutanten auf den jeweiligen Substraten wachsen läßt. So wäre z.B. $D1 = \{R15P::X5P, TransHydro, SuccCoA::Succ\}$ eine MSM für alle vier Substrate, d.h. nur Teilmengen dieser drei Reaktionen wären als Deletionsmenge tolerierbar. Dagegen wäre $D2 = \{PEP::OxAc, ICit::Glyox\}$

- keine SM für Wachstum auf Succinat,
- eine MSM für Wachstum auf Glucose und Glycerin und
- eine SM, jedoch keine MSM bei Wachstum auf Acetat (da bereits Deletion von *ICit::Glyox* Wachstum auf Acetat ausschließt).

Folglich kann *E. coli* die Doppeldeletion *D2* prinzipiell nur auf Succinat tolerieren.

7.3.3 Strukturelle Fragilität und Robustheit

MSMn beschreiben minimale Voraussetzungen, die zu einer Fehlfunktion in der Zelle führen. Deshalb eignen sich MSMn auch für eine „Risikoanalyse“, bei der man die Fragilität (oder invers: Robustheit) einer Netzstruktur bzgl. der Zielfunktion studiert. Im folgenden nehmen wir dazu an, daß jede Reaktion des Netzes mit der gleichen (i.a. geringen) Wahrscheinlichkeit ausfallen kann (eine Berücksichtigung unterschiedlicher Wahrscheinlichkeiten wäre denkbar). Daraus folgt direkt, daß kleine MSMn mit höherer Wahrscheinlichkeit Ursache des Ausfalls der Zielreaktion sein können und sie daher die kritischsten *Fehlermoden* darstellen.

Abb. 7.1 zeigt die Größenverteilungen der MSMn von *E. coli*, jeweils bzgl. Wachstum auf einem der vier Substrate. Es ist leicht abzulesen, daß sich die größten (und meisten) MSMn für Glucose ergeben, während z.B. für Acetat weniger und im Mittel kleinere MSMn auftreten. Für Glucose gibt es MSMn mit bis zu 12 Reaktionen, während bei Acetat nur sehr wenige mehr als 5 Elemente aufweisen. Die Anzahl essentieller Reaktionen, also von MSMn mit einem Element, ist bei Acetat um 12 höher als für Glucose. Generell zeigt sich, daß Anzahl und Größe der MSMn mit der Anzahl der EMen (Tab. 7.2) korrelieren. Bereits in Abb. 6.3b wurde der Zusammenhang zwischen Anzahl der EMen und struktureller Robustheit (dort gemessen an der Anzahl nichtessentieller Reaktionen) dargelegt, der sich in den MSMn differenzierter widerspiegelt: je mehr EMen für die Zielreaktion zur Verfügung stehen, d.h. desto flexibler das System ist, desto mehr Reaktionen müssen im Mittel für ein Versagen der Zielreaktion (hier Wachstum) *gleichzeitig* ausfallen. Dieser Zusammenhang leitet sich auch aus dem oben gegebenen Algorithmus zum Berechnen der MSMn ab, wo jeder EM durch die MSMn abgedeckt werden muß. Da sich mit steigender Kardinalität der MSMn mehr Kombinationsmöglichkeiten auszuschneidender Reaktionen ergeben (zumindest so lange die Kardinalität kleiner als $q/2$ ist), steigt auch die Anzahl der MSMn.

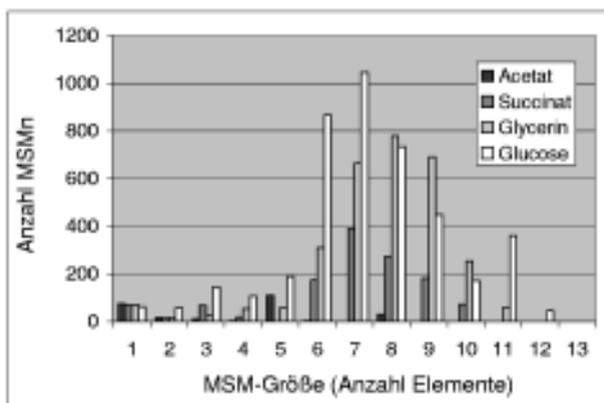


Abb. 7.1: Größenverteilungen der MSMn im Zentralstoffwechsel von *E. coli* bei Wachstum (= Zielreaktion) auf verschiedenen Substraten.

Allerdings sind Flexibilität und strukturelle Robustheit (bzw. Fragilität) nicht immer eindeutig positiv (bzw. negativ) korreliert. Abb. 7.2 zeigt so ein Gegenbeispiel: Netze (a) und (b) besitzen die gleiche Anzahl an Reaktionen, (internen + externen) Metaboliten und EMen. Dennoch ist Netz (a) offensichtlich fragiler als (b), da es drei essentielle Reaktionen und nur eine MSM mit zwei Elementen aufweist, während Netz (b) nur eine essentielle Reaktion und vier MSMn mit zwei Elementen enthält. Fällt also eine der fünf Reaktionen zufällig aus, wäre Netz (a) mit größerer Wahrscheinlichkeit nicht mehr funktionstüchtig.



Abb. 7.2: Zwei Netze mit gleicher Netzdimension und Anzahl EMen, aber unterschiedlicher Fragilität.

Die MSMn ermöglichen also eine subtilere Fragilitäts- oder Robustheitsanalyse als die bloße Anzahl an EMen. Mit den *kleinen* MSMn lassen sich strukturelle „Engpässe“, die für die Zielfunktion am „gefährlichsten“ sind, direkt identifizieren. Zusätzlich erlauben die MSMn nicht nur die *globale* Fragilität des Netzes zu beurteilen, sondern auch den *spezifischen* Beitrag jeder einzelnen Reaktion: Tritt eine Reaktion vorwiegend in größeren MSMn auf, ist ihr Ausfall für die Zielfunktion weniger kritisch, da zum Versagen letzterer noch viele weitere Elemente ausfallen müßten. Als ein quantitatives Maß definieren wir für jede Reaktion i den (spezifischen) *Fragilitätskoeffizienten* F_i , der sich aus dem Reziproken der mittleren Größe aller derjenigen MSMn ergibt, in denen die Reaktion i auftritt:

$$F_i = \frac{1}{\text{Mittelwert}(\{|MCS_j| : i \in MCS_j\})} \quad (7.2)$$

Je kleiner F_i , desto unkritischer ist diese Reaktion für die Operation der Zielfunktion. Per Definition ist $F_i = 0$, falls Reaktion i in überhaupt keiner MSM auftritt. Der maximale Wert 1 tritt für die kritischsten Netzelemente überhaupt auf, nämlich für essentielle Reaktionen. Den Fragilitätskoeffizienten von NB4 (Tab. 7.3) kann man z.B. entnehmen, daß neben der essentiellen Reaktion R1 die Reaktion R5 am kritischsten für die Operation von ZR ist ($F_i = 0.5$). Fällt R5 aus, dann sind R2, R3 und R4 automatisch bedeutungslos für die Funktion von ZR.

Tab. 7.3 zeigt, daß auch F_i und die relative Auftretshäufigkeit h_i^{ZR} der Reaktion i in den EMen bzgl. ZR korrelieren, aber erneut nicht immer streng monoton. So treten R6, R7 und R8 in genauso vielen EMen wie R2, R3 und R4 auf, zeigen aber einen größeren Fragilitätskoeffizienten als die zweite Gruppe. Der Grund dafür ist, daß nach Ausfall einer Reaktion aus der

ersten Gruppe, bereits ein weiterer Schnitt bei R5 die Zielreaktion unterdrückt. Dagegen kann nach Löschen von R2 oder R3 oder R4 noch eine beliebige weitere nichtessentielle Reaktion ausfallen – ZR bliebe funktional.

	R1	R2	R3	R4	R5	R6	R7	R8	ZR
F_i	1	1/3	1/3	1/3	1/2	3/8	3/8	3/8	1
h_i^{20}	100%	33%	33%	33%	67%	33%	33%	33%	100%

Tab. 7.3: Fragilitätskoeffizienten und relative Auftretshäufigkeiten bzgl. ZR in NB4.

Tab. 7.2 zeigt Fragilitätskoeffizienten einiger ausgewählter Reaktionen im *E.coli*-Netz. Sie hängen streng vom jeweiligen Substrat ab. Es ist intuitiv klar, daß die Fructosebisphosphatase völlig unkritisch für Wachstum auf Glucose ($F_i = 0.102$), aber dafür essentiell als gluconeogenetischer Schritt für die anderen drei Substrate ist. Der relativ hohe F_i -Wert für das Malat-Enzym bei Wachstum auf Acetat und Succinat kann durch den notwendigen Abfluß von Kohlenstoff aus dem TCA erklärt werden, der durch die Substrate dort hineingelangt. Dafür gibt es nur noch eine zweite Alternative, nämlich die PEP-Carboxykinase.

Alternative Fragilitätskoeffizienten wären denkbar. So ist die Größe der kleinsten MSM, in der eine Reaktion i auftritt, ebenfalls von wichtiger Bedeutung. In *E. coli* ist z.B. F_i für die Reaktion des Malat-Enzyms auf Glucose relativ klein (Tab. 7.2). Allerdings gibt es eine kritische Konstellation (MSM mit nur zwei Elementen): fallen Malat-Enzym und Malat-Dehydrogenase gleichzeitig aus, ist Wachstum auf Glucose nicht mehr möglich.

Zur Quantifizierung der Fragilität des gesamten Netzes wird hier schließlich der *Netzfragilitätskoeffizient* F eingeführt, der sich als mittlerer spezifischer Fragilitätskoeffizient aller q Reaktionen ergibt:

$$F = \frac{1}{q} \sum_i F_i \quad (7.3)$$

Im Unterschied zur Flexibilität (Anzahl EMen) diskriminiert diese Größe nun tatsächlich die beiden Netze in Abb. 7.2: $F_{(a)} = 0.8 > F_{(b)} = 0.6$. Für NB4 ergibt sich ein Wert von 0.514, die F -Werte für Wachstum von *E. coli* auf den verschiedenen Substraten finden sich in Tab. 7.2.

7.3.4 Minimale Meßkombinationen

MSMn lassen sich auch so interpretieren, daß durch ihre Deletion alle Freiheitsgrade im Netz (bzgl. der Zielreaktion) beseitigt werden und kein stationärer Fluß mehr möglich ist. Um die Reduktion von Freiheitsgraden geht es auch in der metabolen Flußanalyse (MFA, Kap. 5), wo einige Reaktionsraten mit dem Ziel gemessen werden, die (algebraischen) Freiheitsgrade für unbekannte Raten zu eliminieren. Mit der verallgemeinerten Berechenbarkeitsanalyse für (sta-

tionäre) Reaktionsnetze wurde in Abschnitt 5.2 eine Methode eingeführt, mit der man gezielt nach solchen Meßkombinationen suchen kann, die eine unbekannte Rate r_u einer Reaktion u berechenbar machen. Die Vorgehensweise beinhaltet zwei Schritte: (1) Berechne die EMen, wobei *alle* Reaktionen als reversibel zu betrachten sind. (2) Konstruiere die Menge \mathbf{M} der zu messenden Raten so, daß in allen EMen von (1), in denen die Reaktion u involviert ist, mindestens eine zu messende Reaktionsrate aus \mathbf{M} involviert ist. Beim Vergleich dieses Schemas mit Lemma 7.1 stellen wir fest, daß die Menge \mathbf{M} eine Schnittmenge bzgl. der Zielreaktion u ist. Das bedeutet, wenn wir die MSMn des *komplett reversiblen* Netzes bzgl. der Zielreaktion u bestimmen, finden wir alle minimalen und nichtredundanten Meßkombinationen, die r_u berechenbar machen.

Die Forderung, das Netz als vollständig reversibel zu betrachten, mag wenig intuitiv erscheinen. Wir illustrieren das in Abb. 7.3, wo R2 als Zielreaktion im Netz NB1 betrachtet werden soll. Die linke Seite von Abb. 7.3 demonstriert, wie die MSMn bzgl. der Funktion von R2 berechnet werden. Es gibt zwei EMen in NB1, von denen einer R2 involviert. Eine MSM muß diesen minimal abdecken, wofür es, neben der trivialen $MSM0=\{R2\}$, genau zwei Möglichkeiten gibt: $MSM1=\{R1\}$ und $MSM2=\{R4\}$. Löschen wir also R1 oder R4, so kann R2 nicht operieren. Für die Berechenbarkeit reichen diese MSMn aber nicht aus, denn messen wir z.B. nur R1, kann die Reaktionsrate von R2 nicht eindeutig abgeleitet werden (außer im Spezialfall, wo $r_{R1}=0$ ist und damit $r_{R2}=0$ direkt folgt; deshalb ist R1 gerade eine MSM). Die rechte Seite von Abb. 7.3 zeigt, wie wir die minimalen Meßkombinationen zur eindeutigen Bestimmung von r_{R2} finden: Hierzu müssen wir nun auch einen dritten EM in Betracht ziehen, der sich zusätzlich ergibt, wenn alle Reaktionen als reversibel betrachtet werden. Damit involvieren nun zwei EMen die Reaktion R2. Als MSMn ergeben sich daraus nun $\{R1,R3\}$ und $\{R3,R4\}$, die in der Tat die einzigen nichtredundanten Meßkombinationen sind, die R2 berechenbar machen. Hieraus läßt sich schlußfolgern, daß R3 auf jeden Fall gemessen werden muß, wie bereits in Kap. 5.2 abgeleitet wurde.

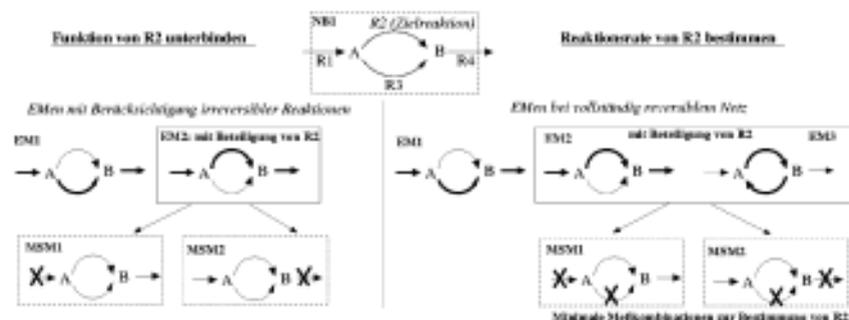


Abb. 7.3: MSMn in NB1 (Zielreaktion R2): Unterbinden der Funktion von R2 (links) und minimale Meßkombinationen zur eindeutigen Bestimmung der stationären Rate von R2 (rechts). Die triviale $MSM0=\{R2\}$ wurde jeweils nicht mit eingezeichnet.

Minimale Meßkombinationen aus der Netzgraphik abzuleiten, wird in NB4 bereits deutlich schwerer. Betrachtet man NB4 als vollständig reversibel, so ergeben sich 9 EMen und daraus 16 MSMn (= minimale Meßkombinationen) bzgl. Zielreaktion ZR. Die kleinste MSM enthält nun 2 Elemente, d.h. mind. 2 Messungen sind notwendig. So ist R1 zwar essentiell für ZR, ihre gemessene Rate r_{R1} reicht allein i.a. aber nicht zur Bestimmung von r_{ZR} aus. Dagegen bleibt z.B. [R5,R8] auch im vollständig reversiblen Netz als MSM erhalten. Aus praktischen Gründen kann man oft nur Raten von solchen Reaktionen messen, welche die Systemgrenze überschreiten, in NB4 also R1, R3, R4 und R6. In diesem Fall stellt sich [R1,R3] als optimale Meßkombination heraus, da sie als einzige der dann in Frage kommenden MSMn nur 2 Elemente aufweist. Selbst in diesem einfachen Beispiel ist es nicht leicht zu erkennen, daß die Rate von ZR aus jener von R1 und R3 berechnet werden kann ($r_{ZR}=r_{R1}-r_{R3}$). Alle anderen MSMn mit nur extrazellulären Reaktionen, haben drei Elemente (NB4 hat $9-6=3$ Freiheitsgrade!). Die Tatsache, daß bereits 16 minimale Meßkombinationen in einem kleinen Netz wie NB4 auftreten, unterstreicht nochmals die kombinatorische Komplexität solcher Untersuchungen. Möchte man *alle* Raten im Netz bestimmen (d.h. alle Reaktionen sind Zielreaktionen), so enthalten zwangsläufig alle MSMn des reversiblen Netzes so viele Elemente, wie es Freiheitsgrade ($q-\text{Rang}(\mathbf{N})$) im Netz gibt. Natürlich läßt sich die hier beschriebene Methode prinzipiell für jedes lineare Gleichungssystem einsetzen.

7.3.5 Elimination von Erhaltungsrelationen

Bereits in Abschnitt 2.5 wurde darauf hingewiesen, daß man die Erhaltungsrelationen (ERen) eines Reaktionsnetzes in stöchiometrischen und oft auch in dynamischen Modellen eliminieren möchte, um dann mit einem reduzierten Modell fortzufahren. Dazu müssen Zeilen (Metabolite) aus der stöchiometrischen Matrix \mathbf{N} entfernt werden, so daß danach keine lineare Abhängigkeiten mehr zwischen den verbleibenden Zeilen in der reduzierten Matrix \mathbf{N}_r existieren, also $\text{Rang}(\mathbf{N}_r) = \text{Anzahl Metabolite (Zeilen) in } \mathbf{N}_r$ ist. Dazu betrachte man nun die MSMn bzgl. der *elementaren* Erhaltungsrelationen (EERen). Letztere berechnen sich ganz analog wie die Elementarmoden, sie werden allerdings aus der Transponierten von \mathbf{N} ermittelt (Abschnitt 2.5). Berechnet man dann die MSMn bzgl. der EERen mit allen Metaboliten als „Zielmetaboliten“, so stellen diese MSM dann alle möglichen Kombinationen von aus \mathbf{N} zu entfernenden Metaboliten (Zeilen) dar, die zur Elimination der ERen in \mathbf{N}_r führen.

7.4 Minimale Schnittmengen in der Risikoanalyse und Graphentheorie

Während EMen minimale funktionale Kombinationen von Netzelementen sind, stellen MSMn minimale (irreduzible, unzerlegbare, elementare) Fehlermoden dar. Minimale Fehlermoden sind auch zentraler Gegenstand bei Sicherheits- und Zuverlässigkeitsanalysen, die für indu-

strielle Anlagen durchgeführt werden [34,147]. Dort betrachtet man Fehlerbäume, bei denen *Ereignisse (events)* über Boolesche Operationen miteinander verknüpft werden können, d.h. Fehlerbäume sind nichtrekursive Boolesche Netzwerke. Sogenannte *basic events* stellen die Eingänge (Blätter) des Baumes dar (wie z.B. „Strom fällt aus“), *intermediate events* erhält man durch logische Verknüpfungen von *events* (wie „Strom fällt aus“ ODER „Maschine steht“). Auf der obersten Stufe steht das *top event*, das gewöhnlich einen ersten Systemfehler (Havarie) darstellt, den es zu vermeiden gilt. Eine Schnittmenge in so einem Fehlerbaum ist eine Menge von *basic events*, die zum *top event* führt, und eine minimale Schnittmenge besitzt keine echte Teilmenge, die das *top event* auslösen kann. Dieser Ansatz ist ganz analog zum hier verfolgten: unser *top event* ist der Ausfall der Zielreaktion. Die Deletion einer Reaktion ist ein *basic event* und die Beseitigung mehrerer Reaktion ist ein *intermediate event*. Leider können die optimierten Algorithmen zur Suche nach MSMn in Fehlerbäumen hier aber nicht eingesetzt werden, da wir den Fehlerbaum gerade nicht kennen – wir wissen im Vorhinein nicht, welche Kombinationen von Deletionen zum Ausfall der Zielreaktion führen. Allerdings, will man das Konzept der MSMn auch auf zelluläre Signaltransduktionsnetze erweitern – was sehr attraktiv erscheint – dann könnten diese Algorithmen möglicherweise hilfreich sein. Ansätze zur Beschreibung von Signaltransduktionen mittels Boolescher Netzwerke gibt es bereits [44]. Sofern letztere nicht rekursiv sind, könnten Methoden aus der Risikoanalyse zum Einsatz kommen.

Den Begriff der MSMn gibt es auch in der Graphentheorie [54]. Dort stellen MSMn solche minimalen Mengen aus Kanten dar, deren Deletion zum Zerfall des Graphen in zwei oder mehrere Komponenten führt, d.h. für einige Paare von Knoten gibt es keinen verbindenden Pfad mehr. Obwohl eine gewisse Ähnlichkeit mit der hier verwendeten Definition von MSMn offensichtlich ist, können graphentheoretische Methoden zur Bestimmung von MSMn in Reaktionsnetzwerken i.a. nicht verwendet werden. In Abschnitt 2.3 wurde bereits auf die hypergraphische Natur von stöchiometrischen Netzwerken hingewiesen, deren Eigenschaften bei einer Transformation in Graphendarstellung teilweise verloren gehen. Abb. 2.6 illustriert den Unterschied zwischen MSMn in Graphen und Hypergraphen: Entfernt man Reaktion R2 im Netz NB3, dann kann Metabolit E nicht mehr aus A produziert werden, d.h. {R2} ist eine MSM. Im bipartiten bzw. Substrat-Graphen von NB3 bleibt aber weiterhin eine Verbindung (Pfad) zwischen A und E bestehen und somit wäre {R2} keine MSM im graphentheoretischen Sinne.

8 Kombinatorische Komplexität der Elementarmodenbestimmung

Die vorangegangenen Kapitel demonstrierten die tiefgreifende Bedeutung der Elementarmoden für funktionale Analysen in Stoffwechselnetzen. In größeren Netzen sind solche Studien nur eingeschränkt durchführbar, da die Anzahl möglicher Pathways rasant ansteigt. In diesem Kapitel gehen wir dieser kombinatorischen Komplexität auf den Grund. Wir beginnen mit einer oberen Schranke für die maximale Anzahl an EMen bei einer vorgegebene Netzgröße.

8.1 Maximal mögliche Anzahl von Elementarmoden

Jeder EM e ist durch die involvierten Reaktionen $R(e)$ eindeutig definiert (2.12). Wegen (B1) in (2.13) repräsentiert $R(e)$ eine Teilmenge der Spalten (Reaktionen) von N , die linear abhängig sind. Diese Spalten sind sogar *minimal* linear abhängig, denn wegen (B3) in (2.13) repräsentiert jede echte Teilmenge von $R(e)$ eine Menge von Spalten aus N , die linear *unabhängig* sind.

Eine obere Schranke für die maximal mögliche Anzahl minimal linear abhängiger Mengen an Spaltenvektoren einer Matrix wurde in [153] gegeben. Diese Spaltenmengen wurden dort als *Simplexe* bezeichnet, in der Theorie der Matroide nennt man sie auch *Circuits* [111]. Simplexe und Circuits erfüllen demnach (B1) und (B3) von (2.13), berücksichtigen im allgemeinen aber keine Vorzeichenbedingungen nach (B2). Übertragen auf unsere Anwendung werden damit die EMen bzgl. des vollständig reversiblen Netzes betrachtet, was genau der kombinatorischen Darstellung des Nullraums nach Abschnitt 5.2 entspricht. Da diese Menge an EMen wegen (2.17) immer eine Obermenge der tatsächlich (thermodynamisch) realisierbaren EMen ist, können wir die obere Schranke an Simplexen erst recht als obere Schranke S_{\max} für die Anzahl S an EMen übernehmen (q = Anzahl Reaktionen):

$$S \leq S_{\max} = \binom{q}{\text{Rang}(N)+1} = \frac{q!}{(q - \text{Rang}(N) - 1)! (\text{Rang}(N) + 1)!} \quad (8.1)$$

Zu beachten ist, daß reversible EMen in S bzw. S_{\max} nur einmal gezählt werden (ansonsten müßte (8.1) noch mit 2 multipliziert werden). Die höchstmögliche Anzahl S_{\max} an EMen tritt dann auf, wenn alle Reaktionen reversibel sind und jede Teilmenge aus $\text{Rang}(N)$ vielen Spaltenvektoren linear unabhängig ist, also eine Basis aller q Spaltenvektoren in N darstellt. In diesem Fall bilden folglich alle Kombinationen aus $\text{Rang}(N)+1$ vielen Spalten einen EM, deren Anzahl durch den Binomialkoeffizienten in (8.1) ausgedrückt wird. Sofern keine Erhaltungsrelationen im Netz auftreten bzw. diese entfernt wurden und außerdem die Anzahl der Metabolite m kleiner als die der Reaktionen ist (sonst $S_{\max} = 0$), so ist $\text{Rang}(N)=m$ und damit:

$$S \leq S_{\max} = \binom{q}{m+1} = \frac{q!}{(q-m-1)! (m+1)!} = \frac{q \cdot (q-1) \cdot (q-2) \cdot \dots \cdot (q-m)}{(m+1)!} \quad (8.2)$$

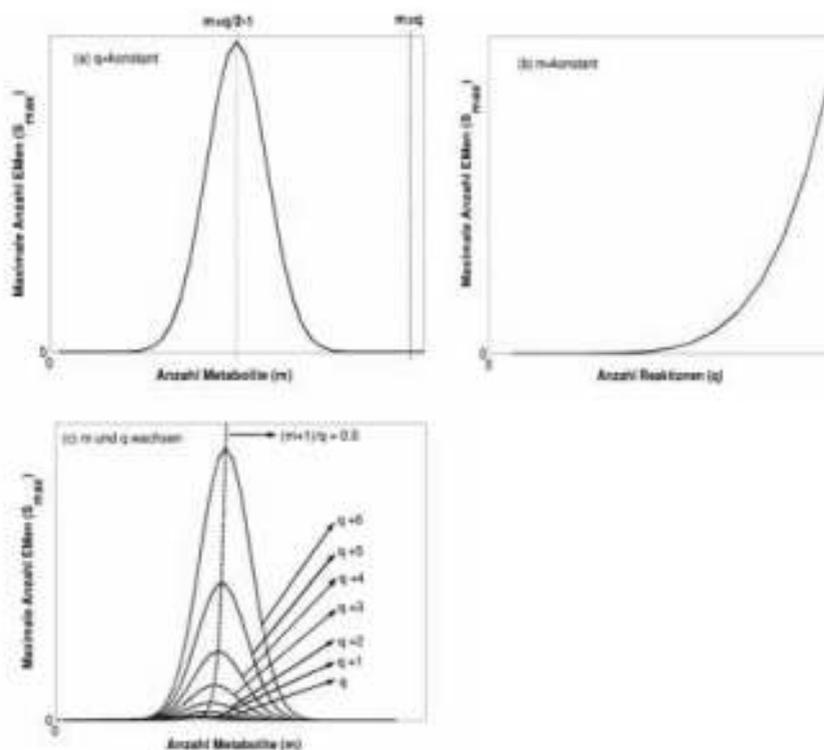


Abb. 8.1: Abhängigkeit der Schranke S_{max} : (a) bei steigender Anzahl an Metaboliten; (b) bei steigender Anzahl an Reaktionen und (c) wenn Reaktionen **und** Metabolite zahlenmäßig zunehmen.

Zur vereinfachten Illustration geben wir im folgenden von Fall (8.2) aus. Abb. 8.1 zeigt, wie sich eine steigende Netzgröße auf S_{max} auswirkt. In Abb. 8.1a wird zunächst q festgehalten und m variiert (z.B. indem ursprünglich externe Metabolite nun als intern betrachtet werden). Sofern q gerade ist ergibt sich ein Maximum bei $m = q/2 - 1$, andernfalls gibt es zwei benachbarte Maxima bei $m = (q-3)/2$ bzw. bei $m = (q-1)/2$. Hat das Netz also etwa halb so viele Metabolite wie Reaktionen ($m/q = 1/2$), so liegt die „ungünstigste“ Konstellation vor, da sich dann kombinatorisch die meisten EMen ergeben können. In realistischen Netzen liegt dieses Verhältnis aber viel näher bei 1 (also rechts vom Maximum). Daher wird sich bei Berücksichtigung weiterer Metabolite im Netz die Schranke und in den meisten Fällen auch die tatsächliche Anzahl an EMen reduzieren. Allerdings gibt es auch Fälle, wo nach Einfügen eines Metaboliten S_{max} fällt, aber S steigt (natürlich nicht über S_{max} hinaus). Dies gilt insbesondere für stark verknüpfte Metabolite, da diese auf vielfache Weise, nämlich durch viele verschiedene Reaktionen, im Gleichgewicht gehalten werden können. Betrachtet man daher z.B.

ATP in einem Stoffwechselmodell zunächst als extern (unbalanciert) und dann als intern (m erhöht sich um 1), so steigt häufig die Anzahl an EMen, obwohl S_{max} sinkt. Bei Metaboliten, die an nur wenigen Reaktionen beteiligt sind, ist das selten der Fall.

Abb. 8.1b zeigt die Abhängigkeit von S_{max} , wenn die Anzahl der Reaktionen bei gleichbleibender Anzahl berücksichtigter Metabolite erhöht wird. Fügt man eine neue Reaktion in ein bestehendes Netz ein, bleiben als Umkehrung von (2.16) die EMen des ursprünglichen Netzes immer erhalten, daher kann S_{max} nur streng monoton steigen. In der Tat steigt S_{max} in (8.2) um den Faktor $(q+1)(q-m)$ wenn man q um Eins erhöht. Die resultierende Steigung bei fortlaufender Erhöhung ist aber nicht exponentiell! Im rechten Term in Gl. (8.2) treten nämlich immer nur $(m+1)$ Faktoren im Zähler auf, während der Nenner unverändert bleibt, d.h. die Kurve in Abb. 8.2b wird durch ein Polynom $(m+1)$ -ten Grades beschrieben.

Im realen Fall bedeutet steigende Netzgröße allerdings, daß sowohl die Anzahl der Metabolite, als auch die der Reaktionen wächst. Nimmt man ein fixes Verhältnis $m/q < 1$ an, so ergibt sich nun ein exponentieller Anstieg von S_{max} wenn q und, im entsprechenden Verhältnis, m vergrößert werden. Die gestrichelte Linie in Abb. 8.1c zeigt die explosiv wachsende Schranke S_{max} für den Fall, daß m und q im (ungünstigsten) Verhältnis von $(m+1)/q=1/2$ wachsen.

Als Zwischenresultat können wir also festhalten, daß die tatsächliche Anzahl an EMen, auch wenn die obere Schranke S_{max} nicht erreicht wird (s.u.), maßgeblich von der Netzgröße (q und m), sowie vom Verhältnis zwischen m und q abhängt.

8.2 Realistische Netze: $S \ll S_{max}$

Tab. 8.1 (letzte Zeile) zeigt die Anzahlen an EMen, die sich für das *E.coli*-Netz bei verschiedenen Substraten ergaben (vgl. mit Tab. 6.2). Zusätzlich wurde hier auch der Fall betrachtet, wo alle Substrate gleichzeitig zur Verfügung stehen (letzte Spalte). Die jedem Szenario zugrunde liegende Anzahl an Metaboliten bzw. Reaktionen ist in Tab. 8.1 angegeben (erste Zeile) und ermöglicht die Bestimmung von S_{max} nach Gl. (8.2). Vergleicht man die oberen Schranken mit der jeweils tatsächlich auftretenden Menge an EMen, so ist eine enorme Überschätzung festzustellen, bei Glucose z.B. mit einem Faktor von ca. 10^{16} .

	Glucose	Acetat	Glycerin	Succinat	alle 4 Substrate gleichzeitig
# Metabolite/Reaktionen [# irreversible Reaktionen]	89106 [78]	89105 [77]	89106 [78]	89105 [77]	89110 [82]
obere Schranke (S_{max})	$3.69 \cdot 10^{16}$	$5.57 \cdot 10^{17}$	$3.69 \cdot 10^{16}$	$5.57 \cdot 10^{17}$	$4.39 \cdot 10^{21}$
kompriertes Netz: # Metabolite/Reaktionen	3046 [28]	3045 [27]	3147 [29]	3045 [27]	3151 [33]
kompriertes Netz: obere Schranke (S_{max})	$5.12 \cdot 10^{11}$	$1.67 \cdot 10^{11}$	$7.52 \cdot 10^{11}$	$1.67 \cdot 10^{11}$	$4.85 \cdot 10^{15}$
benötigte Rechenzeit	1027 sec	3.8 sec	53.6 sec	16.4 sec	ca. 50 Stunden
Anzahl EMen (S)	27100	599	11333	4250	507633

Tab. 8.1: EMen-Berechnungen und Abschätzungen für verschiedene Substratszenarien im *E.coli*-Netz.

Die Tatsache, daß die Anzahl an EMen in (komplexeren) metabolischen Netzen oft zwar sehr hoch, aber dennoch weit unter dem theoretischen Maximum liegt, hat folgende Gründe:

(1) *Kurze Routen:* Die Länge $L=R(\epsilon)$ ="Anzahl involvierter Reaktionen" der EMen ist in Stoffwechselnetzen meistens viel kürzer als theoretisch möglich ($L_{max}=Rang(N)+1=m+1$).

Findet man einen EM mit $L < L_{max}$, so reduziert sich S_{max} um den Wert $\binom{q-L}{m+1-L}$. Dieser

Binomialkoeffizient gibt die Anzahl derjenigen $(m+1)$ -elementigen Teilmengen der q Reaktionen an, welche die L Reaktionen des EM enthalten und damit nicht mehr als EM in Frage kommen. (Hinweis: Im Extremfall $S = S_{max}$ haben alle EMen die Länge L_{max} , s.o.)

Abb. 8.2 zeigt die Längenverteilung der EMen bei Glucose als Substrat. Alle EMen haben eine geringere Länge als $L_{max}=89+1=90$. Die bimodale Verteilung ergibt sich durch EMen mit Wachstum (größere Längen) bzw. ohne Wachstum.

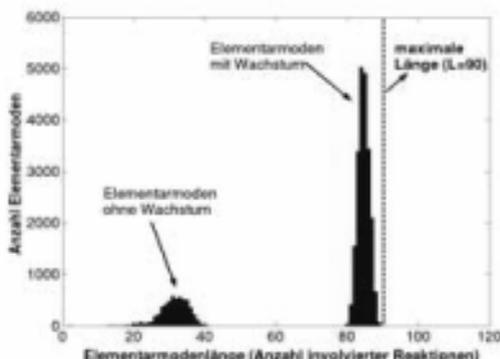


Abb. 8.2: Längenverteilung der Elementarmoden im *E.coli*-Netz bei Glucose als Substrat.

Kurze Weglängen sind physiologisch sicher bedeutsam (z.B. geringerer Installationsaufwand, weniger fehleranfällig), ergeben sich aber auch aus einem anderen Grund. Man betrachte zwei Vektoren $v_1 = (v_{11}, v_{12})^T$ und $v_2 = (v_{21}, v_{22})^T$ aus dem \mathbb{R}^2 . Die je zwei reellen Komponenten beider Vektoren sollen zufällig aus dem Intervall $[-3,3]$ gewählt werden. Wie wahrscheinlich ist es, daß die beiden Vektoren linear abhängig sind? Praktisch Null, da $v_2 = v_1 \cdot v_{21}/v_{11}$ (falls $v_{11} \neq 0$) bzw. $v_2 = 0$ (falls $v_{11} = 0$) *exakt* gelten muß, die Komponenten aber zufällig aus einem reellen Intervall gewählt werden. Deshalb weisen Matrizen, deren reelle Einträge zufällig gewählt werden, so gut wie immer die maximale Anzahl minimal linear abhängiger Spalten nach Gl. (8.1) auf.

Anders verhält es sich dagegen, wenn nur ganze Zahlen aus einem endlichen Intervall als Einträge in der Matrix erscheinen, wie es bei stöchiometrischen Matrizen (mit Ausnahme der

Stöchiometrie der Biomassesynthese) praktisch immer der Fall ist. Nimmt man obiges Beispiel und wählt die Komponenten zufällig aus $\{-3, -2, -1, 0, 1, 2, 3\}$, so ist die Wahrscheinlichkeit, zwei linear abhängige Vektoren zu erhalten, nun größer Null. Kleine EMen (und daraus folgend $S \ll S_{\max}$) sind also zumindest teilweise auch ein Ergebnis der ganzzahligen Stöchiometrie von Reaktionsnetzwerken.

(2) Enzyme Subsets (ESs) treten oft in metabolen Reaktionsnetzen auf, häufig als lineare Reaktionssequenzen. Für viele Untersuchungen ist es völlig ausreichend, die Reaktionen eines ESs in einer (kumulierten) Reaktion zusammenzufassen (Abschnitt 2.4.1). Auch bei der Berechnung von EMen verfährt man auf diese Weise ([141] und Abschnitt 4.4), was letztlich zu einer kleineren Anzahl an Reaktionen und meistens auch Metaboliten führt. Die Anzahl der EMen in diesem komprimierten Netz ist der im originalen identisch, allerdings folgt aus der reduzierten Netzgröße gewöhnlich eine kleinere obere Schranke. Im Falle von *E. coli* verbleiben im komprimierten Netz jeweils nur etwa die Hälfte der Reaktionen bzw. ein Drittel der Metabolite, was die obere Schranke bereits drastisch senkt. Dennoch überschätzt letztere noch immer sehr stark den wahren Wert (Tab. 8.1). ESs reduzieren also die mögliche Anzahl an EMen.

(3) Irreversible Reaktionen: Viele Stoffwechselreaktionen können als praktisch irreversibel betrachtet werden. Der Ausschluß der Rückrichtung einiger Reaktionen führt wegen (2.17) fast immer zu einer geringeren oder selten gleichbleibenden Anzahl an EMen. In NB2 ist z.B. $S_{\max}=120$ und $S=12$, wären alle Reaktionen reversibel, so ergäben sich 21 EMen.

Die Reduktion der oberen Schranke S_{\max} durch irreversible Reaktionen ist i.a. schwer abzuschätzen. Der Anteil irreversibler Reaktionen in Stoffwechselnetzen ist nicht gering (vgl. Tab. 8.1) und sie sind ein wesentlicher Grund dafür, warum die Menge an (gültigen) EMen stark eingeschränkt wird. Man erinnere sich z.B. an Abschnitt 5.5.5: Dort hatten wir für ein Szenario die kombinatorische Repräsentation D_n des Nullraums – also die Menge der EMen ohne Einschränkungen der Reaktionsrichtungen – berechnet. Es ergaben sich 132198 EMen. Berücksichtigt man hingegen die Restriktionen in der Reversibilität, so bleiben gerade einmal 43 EMen übrig, also nur 0.03%. Irreversible Reaktionen sind daher auch physiologisch von großer Bedeutung: Sie machen den Dschungel aus verschlungenen Stoffwechselwegen übersichtlicher und setzen markante Stellen, an denen Ströme und Prozesse unumkehrbar werden. Die Koordination und Regulation des Stoffwechsels wird sich dadurch vermutlich viel einfacher gestalten lassen.

(4) Schwache Konnektivität der Metabolite: Stoffwechselnetze sind i.a. schwach verknüpft, das heißt die meisten Metabolite sind nur an wenigen Reaktionen beteiligt (siehe Abschnitt 2.3). Für solche Spezies gibt es folglich nur wenige Kombinationsmöglichkeiten von Reaktionen, die den Metaboliten im Gleichgewicht halten können. Soll ein EM z.B. über einen Metaboliten führen, der nur an zwei Reaktionen beteiligt ist, gibt es nur eine Konstellation, nämlich die, wo beide Reaktionen verwendet werden. Ist die Spezies an drei Reaktionen

beteiligt, gibt es dagegen schon maximal vier (dreimal je zwei Reaktionen oder alle zusammen). Der unmittelbare Zusammenhang zwischen Konnektivität und Anzahl an EMen läßt sich auch am Algorithmus zur Berechnung der EMen erkennen, da dort die EMen aus gerade solchen Kombination in einem Tableau ermittelt werden [141].

Da die Konnektivität in metabolen Netzen einem Potenzgesetz folgt ([65], Abschnitt 2.3), gibt es einige wenige Metabolite wie ATP oder NAD(P)H, die sehr stark verknüpft sind. Diese sind, wie in Abschnitt 8.1 erwähnt, durch ihre vielfach kombinierbaren Ein- und Ausgänge für größere Anzahlen an EMen „verantwortlich“. Deshalb wurde in [145] auch vorgeschlagen, Metabolite mit hoher Konnektivität als extern zu betrachten, um die Anzahl an EMen gering zu halten. Ob dadurch tatsächlich eine Reduktion in S erzielt wird hängt, wie oben gezeigt, u.a. aber auch vom Verhältnis m/q ab.

Wie schwer sich die tatsächliche Anzahl an EMen i.a. abschätzen läßt, zeigt sich z.B. in *E. coli* bzgl. der Substrate Succinat bzw. Acetat: Obwohl sich beide Szenarien nur in der Substrataufnahmereaktion unterscheiden, ergibt sich für Succinat die siebenfache Anzahl an EMen. Dies bestätigt aber auch, wie sensitiv eine (in diesem Falle wesentliche) Veränderung in der Netzstruktur durch die EMen erfaßt wird.

8.3 Rechenaufwand

Wie oben gesehen haben, daß die Anzahl an EMen im Stoffwechsel aus verschiedenen Gründen viel niedriger als theoretisch möglich ist. Dennoch erfordert ihre Berechnung in größeren Netzen beträchtliche Rechen- und Speicherressourcen und ist, je nach Netzgröße und Ausstattung des Computers, zum Teil gar nicht realisierbar. Tab. 8.1 zeigt die benötigte Rechenzeit für die einzelnen Szenarios für *E. coli*. Im Extremfall, wo alle Substrate gleichzeitig angeboten wurden und sich ca. eine halbe Million EMen ergaben, benötigte ein Intel Pentium IV (1 GHz) mehr als zwei Tage. (Man beachte, daß die Rechenzeit u.a. auch von der Sortierung des Tableaus abhängig ist.) Trägt man die finale Anzahl an EMen gegen die jeweils benötigte Rechenzeit auf, ergibt sich eine quadratische Abhängigkeit (Abb. 8.3). Diese Tatsache läßt sich darauf zurückführen, daß im Algorithmus die vorläufigen EMen miteinander kombiniert und außerdem gegeneinander auf die Elementareigenschaft getestet werden müssen (Gl. (4.1)), was bei steigender Anzahl an EMen einen quadratisch steigenden Aufwand $o(S^2)$ impliziert. Wenn wir davon ausgehen, daß S wie S_{max} bei zunehmender Netzgröße (mit m/q konstant) exponentiell ansteigt – wenn auch gedämpft im Vergleich zu S_{max} – so ist die Berechnung der EMen folglich nicht in polynomialer Zeit möglich.

Effiziente Algorithmen sind daher absolut notwendig, um die EMen zumindest in Netzen mit mittlerer Komplexität berechnen zu können. In Abschnitt 4.4. wurden einige Maßnahmen beschrieben, mit denen der Algorithmus im *FluxAnalyzer* optimiert wurde. Erste Ansätze für eine Parallelisierung der Berechnung wurden bereits vorgeschlagen [128].

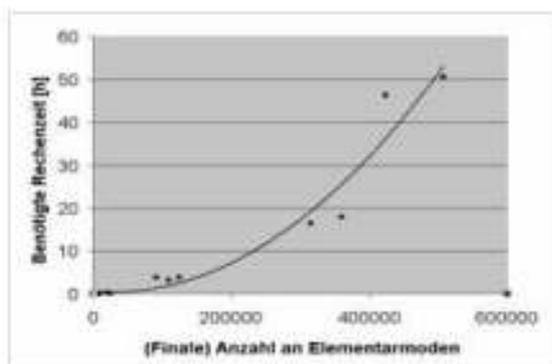


Abb. 8.3: Abhängigkeit der benötigten Rechenzeit von der (finalen) Anzahl an EMen. Neben den in Tab. 8.1 angegebenen Rechenzeiten wurden einige weitere Szenarios mit einbezogen. Die Berechnung erfolgte jeweils mit dem FluxAnalyzer auf einem Rechner mit Intel Pentium, 1 GHz und 4 MB RAM. Die eingezeichnete quadratische Kurve ergab sich aus dem Fitten der Datenpunkte.

Dennoch, in genom-skaligen Stoffwechselnetzen mit mehreren hunden Reaktionen und Metaboliten (vgl. [29]: $q=720$, $m=436$) erscheint die Bestimmung aller EMen kaum möglich. Dann können nur noch Approximationen betrachtet werden. Als eine Möglichkeit kommt die Bestimmung „lokaler EMen“ in kleineren Teilnetzen in Frage, die man aus einer (meistens heuristischen) Zerlegung des Gesamtnetzes erhält [130,145]. Alternativ könnte man Reaktionen von milderem Interesse im Originalnetz weglassen. Die dann berechneten EMen wären zwar eine u.U. stark reduzierte, aber zumindest *gültige* Teilmenge der EMen des gesamten Netzes, was für EMen von lokalen Netzen i.a. nicht gilt.

9 Von der Struktur zur Dynamik: Modellierung der Elektronentransportkette und der Redoxregulation in photosynthetischen Bakterien

In den Abschnitten 3.3, 5.5 und 6.2.3 wurden zentrale Fragestellungen des Redoxstoffwechsels in schwefelfreien Purpurbakterien (SPB, *Rhodospirillaceae*) untersucht. Basierten die Analysen bisher auf rein stöchiometrischen Informationen, soll hier nun der Bogen zu einem dynamischen Modell geschlagen werden. Bezüglich des Stoffwechsels wird dieses zwar auf die membranständige Elektronentransportkette (ETK), einem zentralen Umschlagplatz von Redox- und Energieäquivalenten, beschränkt bleiben, bezieht dafür aber auch die redoxgesteuerte Regulation der Expression photosynthetischer Apparate und anderer Komponenten der ETK mit ein. Diese Studien sollen zu einem tieferen Verständnis der noch unvollständig aufgeklärten Signaltransduktionswege beitragen, mit denen die SPB an verschiedenen Stellen in der ETK ihren Redoxzustand erfassen und ein Umschalten zwischen aerobem und photosynthetischem Wachstum vornehmen können. Obwohl wir weitgehend an den stationären Lösungen des Modells interessiert sein werden, wird sich zeigen, daß eine Elementarmodenanalyse zwar die möglichen Operationsweisen der ETK offenlegt, jedoch obige Fragestellungen ohne dynamisches Modell nicht untersucht werden können.

Wie bisher stützen wir uns auf publizierte Erkenntnisse der folgenden drei wichtigen Vertreter der SPB: *R. rubrum*, *R. capsulatus* und *R. sphaeroides*. Da Ergebnisse aus eigenen Experimenten mit *R. rubrum* in das Modell einfließen sollen, wird, so weit möglich, auf *R. rubrum* fokussiert. Literatur zur redoxgesteuerten Regulation findet sich aber zumeist nur über *R. capsulatus* und *R. sphaeroides*, auf die dann zurückgegriffen wird.

9.1 Physiologie und Modellierung der Elektronentransportkette

9.1.1 Die Elektronentransportkette der *Rhodospirillaceae*

Die zwei wichtigsten Methoden der SPB zur Energiegewinnung sind, unter anaeroben Bedingungen in Licht, zyklische Photosynthese und, im aeroben Milieu, Respiration (Abschnitt 3.3). In beiden Fällen finden Elektronentransportprozesse in der Membran statt, die in Abb. 9.1 dargestellt sind und uns bereits im Stoffwechselmodell der SPB begegnet waren (siehe „Membranteil“ in Abb. 4.7). Respirative und photosynthetische ETK überlappen sich stark und haben als gemeinsames Ziel die Generierung protonenmotorischer Kraft, die (u.a.) zur Synthese der universellen Energieeinheit ATP genutzt werden kann.

Der Elektronentransport in der ETK folgt einem Redoxpotentialgefälle und wird räumlich über *Redox-Carrier* vermittelt. Diese können Elektronen temporär aufnehmen, diffusiv transportieren und über verschiedene *Enzymkomplexe* in der Membran an andere Redox-Carrier oder Elektronenakzeptoren wieder abgeben. Die wichtigsten Redoxträger der SPB in der Membran sind Ubichinon und Cytochrom c_2 .

Ubichinon (oxidierte Form: Q; reduzierte Form: QH₂) ist lipophil und befindet sich daher direkt in der Membran. Es kann 2 Elektronen aufnehmen und bindet dabei auch 2 Protonen:



R. rubrum besitzt ein weiteres Chinon, nämlich *Rhodochinon* (in *R. sphaeroides* und *R. capsulatus* nicht nachgewiesen [62]), andere Vertreter der SPB auch *Menachinon*. Diese Chinone besitzen ein niedrigeres Standard-Redoxpotential als Ubichinon und sind daher für bestimmte Elektronenübergänge (z.B. bei der Fumarat-Reduktion) thermodynamisch besser geeignet, als Ubichinon. Da die Rolle der alternativen Chinone in den SPB noch nicht verstanden ist, werden sie im Modell nicht berücksichtigt.

Cytochrome sind kleine Proteine mit einer prosthetischen Häm-Gruppe. Das wichtigste Cytochrom der SPB ist Cytochrom c₂, dessen Eisenatom in der Häm-Gruppe wie bei allen Cytochromen ein Elektron aufnehmen kann:



Cytochrom c₂ ist wasserlöslich und wandert deshalb an der Außenseite der Membran entlang. Andere Cytochrome, wie c₇ in *R. capsulatus* und *R. sphaeroides*, sind nachgewiesen worden, die offensichtlich für die Respiration eingesetzt werden [173]. Im Modell betrachten wir nur einen (zusammengefalteten) Pool an Cytochromen, deren Repräsentant c₂ sein wird, da es insbesondere für die Photosynthese dominierende Bedeutung hat.

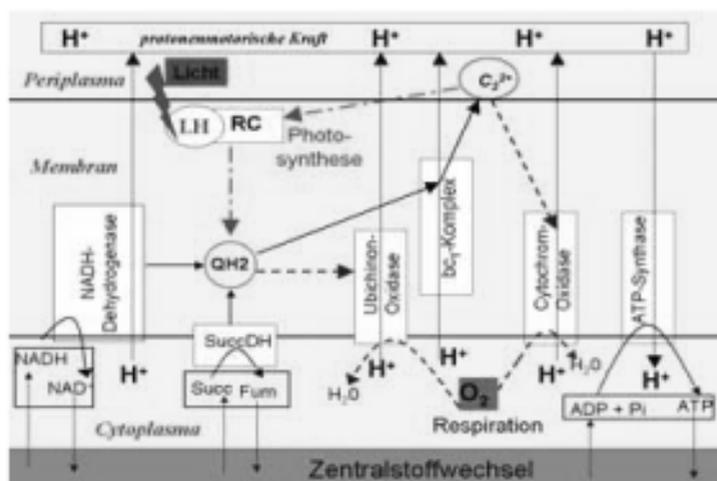


Abb. 9.1: Elektronentransportkette der Rhodospirillaceae. Bis auf Protonentransport (gepunktete Linien) stehen alle Pfeile innerhalb der Membran für Elektronentransport. Sämtliche Pfeile sind als prinzipiell reversibel anzusehen, mit Ausnahme rein respirativer (Strichellinie) und rein photosynthetischer Prozesse (Strichpunktlinie).

NADH bzw. *NADPH* sind die wichtigsten Elektronenträger im Cytoplasma und gehen aus *NAD*⁺ bzw. *NADP*⁺ durch Übergang zweier Elektronen und Binden eines Protons hervor:



Die zentrale Rolle von NAD(P)H als Reduktionsäquivalente im Zentralstoffwechsel wurde bereits in vorhergehenden Abschnitten studiert. Die besondere Bedeutung von NADH für die Respiration liegt darin, daß es Elektronen in die ETK einschleust. Umgekehrt fließen unter bestimmten Bedingungen auch Elektronen aus der ETK auf NAD ab (s.u.).

Verschiedene Enzymkomplexe in der Membran katalysieren den Übergang von Elektronen von einem Redox-Carrier auf einen anderen oder den Ein- und Austritt von Elektronen in die bzw. aus der ETK (Abb. 9.1). Thermodynamisch wandern die Elektronen dabei energetisch bergab, d.h. bzgl. des Redoxpotentials bergauf (Abb. 9.2). Einige Komplexe koppeln den Übergang der Elektronen von einem höheren zu einem niedrigeren Energieniveau an das Pumpen von Protonen von der Innenseite der Membran in das Periplasma (Protonenpumpen). Das daraus entstehende Gefälle in der Protonenkonzentration zwischen Innen- und Außenseite bildet die protonenmotorische Kraft und kann von der ATP-Synthase zur Produktion von ATP genutzt werden. Die wichtigsten Enzymkomplexe in der Membran der SPB, sind folgende (detaillierte Beschreibungen in [96]):

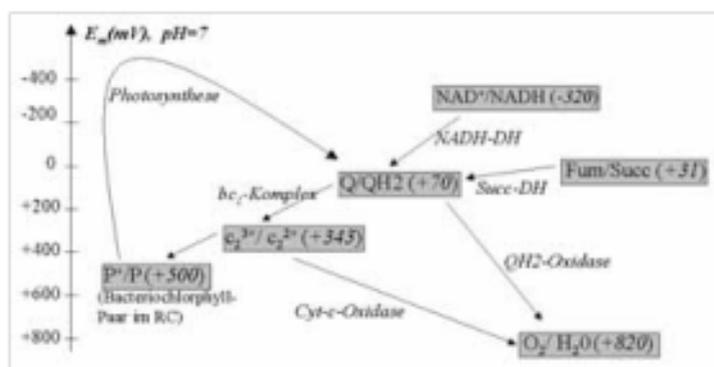


Abb. 9.2: Standard-Redoxpotentiale von Redoxpaaren, die in der Elektronentransportkette involviert sind. Die Pfeile deuten Transport bzw. Übergabe von Elektronen an.

NADH-Dehydrogenase (Komplex I; abgek. NADH-DH): Überträgt die zwei Elektronen von NADH auf Ubichinon, wobei letzteres gemäß (9.1) zwei Protonen bindet. Zusätzlich werden cytoplasmatische Protonen (H_i^+) ins Periplasma gepumpt (H_p^+), wobei hier eine Stöchiometrie von vier angenommen wird. Netto ergibt sich:



In der umgekehrten Richtung ermöglicht die NADH-Dehydrogenase auch den energetisch aktiven Transport von Elektronen aus der ETK heraus (*reverse electron flow*).

Succinat-Dehydrogenase/Fumarat-Reductase (Komplex II, abgek. Succ-DH): Die Succ-DH ist Bestandteil des TCA und oxidiert Succinat zu Fumarat, wobei Ubichinon reduziert wird.

Protonen werden dabei nicht über die Membran gepumpt, was angesichts des geringen Unterschiedes in den Standard-Redoxpotentialen von Succinat und QH₂ nicht wundert (Abb. 9.2):



Der Pfeil in (9.5) ist reversibel, da die Reaktion in Abhängigkeit von den Konzentrationen der beteiligten Reaktionspartner thermodynamisch auch in die andere Richtung realisierbar ist. In diesem Fall übernimmt das Enzym die Funktion einer Fumarat-Reductase, die, wie oben bereits erwähnt, auch als eigenständiges Enzym in *R. rubrum* existiert.

Andere Enzymkomplexe, die Elektronen direkt von einem Metaboliten – ohne Umweg über NADH – auf Q übertragen (wie die Glycerin-Dehydrogenase), werden hier nicht betrachtet.

bc₁-Komplex (Komplex III, Ubichinon-Cytochrom-c-Oxidoreductase): Dieser Komplex transferiert netto zwei Elektronen von QH₂ auf zwei Cytochrome. Der eigentliche Übertragungsprozess ist als *Q-cycle* bekannt [96] und gestaltet sich recht kompliziert. Dabei geben zwei QH₂ jeweils ihre beiden Elektronen an den Komplex ab und entlassen dabei ihre Protonen ins Periplasma. Zwei der Elektronen werden zur Reduktion zweier c₂³⁺-Moleküle verwendet, die anderen beiden werden zur erneuten Reduktion eines (oxidierten) Ubichinons genutzt, d.h.:



NADH-DH, Succ-DH und bc₁-Komplex können sowohl in der phototrophen, als auch in der respirativen ETK involviert sein. Die folgenden beiden benötigen dagegen Sauerstoff als finalen Elektronenakzeptor und sind somit spezifisch für die Atmungskette:

Cytochrom-c-Oxidase (Komplex IV): Dieser Komplex überträgt bei der Respiration Elektronen vom Cytochrom auf den finalen Akzeptor, nämlich Sauerstoff. Dabei werden wiederum Protonen in den Außenraum gepumpt:



Die hier betrachteten Spezies besitzen eine cbb₃-Oxidase, wobei *R. sphaeroides* zusätzlich über eine bzgl. Sauerstoff niedrigaffine aa₃-Oxidase verfügt [102]. Letztere wird im Modell nicht separat berücksichtigt.

Ubichinon-Oxidase: Mit der Ubichinon-Oxidase verfügen die SPB über einen alternativen Weg, Elektronen direkt von QH₂ auf Sauerstoff zu übertragen [163]:



Diese Oxidase ist weniger relevant bei hohen Sauerstoffkonzentrationen, greift aber unterstützend bei reduzierten Bedingungen ein, z.B. bei semiaerobem Wachstum [151,163].

Die Reaktionen der beiden Oxidasen sind irreversible Schritte. Neben Sauerstoff können die SPB viele andere finale Elektronenakzeptoren verwenden (wie DMSO), die für uns hier aber keine Rolle spielen. Die Photosynthese benötigt nun folgende Komponenten:

Photosynthetischer Apparat, bestehend aus Reaktionszentrum (RC) und Lichtsammel- oder Antennenkomplex (LH). Die LH sind Proteinkomplexe, in denen Carotinoide (die für die typische Purpurfärbung der SPB sorgen) und Bacteriochlorophyll (Bchl) inseriert sind. Diese Photosynthesepigmente ermöglichen das Absorbieren von Photonen und die Weiterleitung der Energie zum RC. Man unterscheidet Lichtsammelkomplexe 1 und 2 (LH1 und LH2). LH1, bestehend aus 16 Monomeren, ummantelt das RC und liegt deshalb in 1:1-Stöchiometrie mit dem RC vor [96]. Proteine des LH1 und RC sind folgerichtig auch in einem Operon angeordnet (*puf*, [60]). LH2-Komplexe bestehen aus 9 Monomeren, sind weiter entfernt vom RC lokalisiert und dienen dem zusätzlichen Auffangen von Photonen. Sie werden insbesondere unter Schwachlichtbedingungen ausgebildet, um genügend Anregungen für das RC zur Verfügung zu stellen. *R. rubrum* bildet dabei innerhalb der SPB eine Ausnahme, da es über kein LH2 verfügt [60]. Der Vorteil für die Modellierung besteht darin, daß aus der gemessenen Menge an Bchl sofort die Menge an RC und LH1 abgeleitet werden kann: RC enthält 4 Bchl-Moleküle und LH1 32, d.h. 36 Bchl-Moleküle ergeben ein RC und ein LH1 [96]. Die vom LH zum RC geleitete Anregung wird dort für den eigentlichen photosynthetischen Schritt genutzt, dessen Mechanismus komplex, aber relativ gut verstanden ist [96,107]. Egentliches Ziel der Reaktionssequenzen im RC ist die Reduktion von Q, wobei c_2 als Elektronenspender fungiert. Da letzteres ein höheres Redoxpotential aufweist (Abb. 9.2), ist gerade der Einsatz von (Licht-)Energie notwendig. Abb. 9.3 zeigt ein vereinfachtes Modell, das von 8 Zuständen des RC ausgeht und schnelle Zwischenschritte vernachlässigt [96]. Im ersten Schritt muß ein freies Q an das RC binden (ergibt RCs1). Im nächsten Schritt wird die Energie eines Photons genutzt, um ein Bchl eines sogenannten „speziellen Bchl-Paares“ (P) im RC zu oxidieren (P^+). Das Elektron gelangt dabei auf eine Zwischenstufe (Z, RCs2). Nachfolgend liefert ein reduziertes Cytochrom ein Elektron, um das oxidierte Pigment wieder in den Originalzustand zurückzuführen (RCs3). Thermodynamisch ist das realisierbar, da P ein höheres Standard-Redoxpotential hat als c_2 (Abb. 9.2). Von Z wandert das Elektron weiter zu Q, was zum Semiquinon $Q^{\cdot-}$ reduziert wird (RCs4). Die genannten Schritte wiederholen sich, so daß ein voll reduziertes Q^{2-} entsteht (RCs7). Nachfolgende Bindung von zwei Protonen und Loslösen des entstandenen QH_2 schließt den Zyklus mit folgender Nettostöchiometrie:

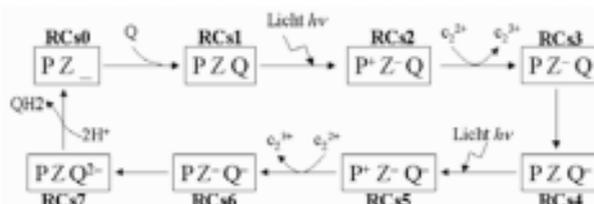


Abb. 9.3: Vereinfachtes Modell der Reduktion von Ubichinon (Q) während der Photosynthese im Reaktionszentrum. P=Bacteriochlorophyll-Paar, Z=Zwischenstufe (Q_{\cdot}). Nähere Erläuterungen im Text.

ATP-Synthase

Die ATP-Synthase kann den in der ETK erzeugten Protonengradienten an eine ATP-Synthase koppeln (vgl. Abschnitt 9.2), wobei Protonen vom Periplasma zurück ins Cytoplasma fließen:



9.1.2 Operationsweisen der Elektronentransportkette: Elementarmoden

An dieser Stelle bietet es sich an, die oben beschriebenen Reaktionen und Komponenten der ETK als Reaktionsnetzwerk aufzufassen und dieses mittels Elementarmoden zu untersuchen. Tatsächlich erhält man so die physiologisch relevanten Operationsweisen der ETK.

Das Netz besteht aus den 7 Reaktionen der Enzymkomplexe: (9.4), (9.5) und (9.7)-(9.11). Bis auf die Ubichinon-Oxidase sind diese Reaktionen auch im Stoffwechselmodell der *Rhodospirillaceae* (Abschnitt 3.3) enthalten. Der Unterschied besteht hier darin, daß wir die Metabolite außerhalb der ETK als externe (gepufferte) Metabolite ansehen und damit die ETK gegen den „Rest“ des Stoffwechsels kapseln. Folglich sind NAD, NADH, ATP, ADP, Succinat, Fumarat, Pi, H_i⁺, H₂O, O₂ und Photonen externe Spezies, während QH₂, c₂ (reduziert) und H_p⁺ als interne (bilanzierte) Metabolite behandelt werden. c₂ (oxidiert) und Q stehen in Erhaltungskorrelation mit c₂ (reduziert) bzw. QH₂.

Tab. 9.1 zeigt die (vereinfachten) Stöchiometrien der Reaktionsgleichungen und die resultierenden 9 EMen in der ETK. Die Bilanzen von NADH, Succinat, ATP, Photonen und O₂ wurden mit angegeben, um Verbrauch bzw. Produktion externer Metabolite verfolgen zu können.

		Photosynthese				Respiration				Fumarat-
		ATP-Synthase	Umgekehrter Elektronentransport			ATP-Synthase				reduktion
		EM1	EM2	EM3	EM4	EM5	EM6	EM7	EM8	EM9
NADH-DH:	$\text{NADH} \leftrightarrow 4 \text{H}_p + \text{QH}_2$		-1	-1	-3			2	6	3
Succ-DH:	$\text{Succinat} \leftrightarrow \text{QH}_2$		1	3	5	6	2			-3
bcl:	$\text{QH}_2 \leftrightarrow 2 \text{c}_2 + 4 \text{H}_p$	3	1				2		6	
Q-Oxidase:	$\text{QH}_2 + 0.5 \text{O}_2 \rightarrow 2 \text{H}_p$			2		6		2		
c-Oxidase:	$2 \text{c}_2 + 0.5 \text{O}_2 \rightarrow 2 \text{H}_p$				2		2		6	
Photosynthese:	$2 \text{c}_2 + 2 \text{Photo} \rightarrow \text{QH}_2$	3	1							
ATP_Synthase:	$3 \text{H}_p \rightarrow \text{ATP}$	4				4	4	4	20	4
<i>Bilanzen externer Metabolite</i>										
Photonen		-6	-2							
NADH (negativ zu NAD)			+1	+1	+3			-2	-6	-3
Succinat (negativ zu Fumarat)			-1	-3	-5	-6	-2			+3
O ₂				-1	-1	-3	-1	-1	-3	
ATP (negativ zu ADP und P _i)		+4				+4	+4	+4	+20	4

Tab. 9.1: Die Elementarmoden der Elektronentransportkette der *Rhodospirillaceae* (Abb. 9.1).

Die EMen in Tab. 9.1 sind bereits nach Photosynthese (EM1, EM2), Respiration (EM3-EM8) und Fumarat-Reduktion (EM9) sortiert. Letztgenannter ermöglicht die Gewinnung von ATP

durch Reduktion von Fumarat zu Succinat, ohne dabei die Photosynthese oder Sauerstoff für die Respiration zu benötigen. Diese Operationsweise ist daher eine wichtige Methode der Energiegewinnung unter anaeroben Bedingungen in Dunkelheit [53].

Photosynthetisch gibt es zwei (u.U. parallele) funktionale Moden: EM1 repräsentiert gerade die *zyklische Photosynthese*. Die Bilanz externer Metabolite zeigt, daß dabei Lichtenergie in chemische Energie des ATP umgewandelt wird und dabei keine Reduktionsäquivalente netto konsumiert oder produziert werden. EM2 zeigt den *reverse electron flow*: Der von der Photosynthese erzeugte Protonengradient wird hier nicht zur ATP-Synthese, sondern zum aktiven Transport von Elektronen aus der ETK heraus zum NADH genutzt. Wenn Elektronen über die Succ-DH in die ETK gelangen (z.B. bei Wachstum auf Succinat), ist dieser Modus essentiell, da es sonst zu einem Elektronenstau in der ETK kommen würde.

Der rückwärtige Elektronenfluß kann auch respirativ auftreten (mit Cytochrom- oder Ubichinon-Oxidase; EM3 bzw. EM4). Allerdings wird man diesen Prozeß, zumindest in der Nettobilanz, wohl nur selten beobachten, da die eingebrachten Elektronen normalerweise zum Sauerstoff wandern werden. Nur in relativ kurzzeitigen Überschusssituationen (z.B. bei schnellem Absinken der Sauerstoffkonzentration) mag dieses Phänomen auch unter aeroben Bedingungen auftreten.

Die anderen respirativen EMen (EM4-EM8) dienen der ATP-Synthese. Bezüglich des Elektronenspenders (NADH oder Succinat) und der Oxidase (Ubichinon- oder Cytochrom-Oxidase) stehen jeweils zwei Optionen zur Verfügung. Optimale ATP-Ausbeute wird mit NADH als Donor und der Cytochrom-Oxidase erzielt (EM8, P/O-Quotient=3.33). Succinat als Elektronenspender und Beteiligung der Ubichinon-Oxidase reduzieren die Ausbeute; sind beide beteiligt, ergibt sich als ungünstigster P/O-Quotient 0.66 (EM5). Wohl aus diesem Grund wird die Ubichinon-Oxidase vorwiegend unter semiaeroben Bedingungen exprimiert, wo sie überschüssige Elektronen in der ETK durch eine hohe Sauerstoffaffinität abführen kann [152].

Obwohl die EMen unser Verständnis für die *Funktionen und Operationsweisen* der ETK festigen, können wir für das tatsächliche *Verhalten* der ETK unter verschiedenen Bedingungen nur bedingt Vorhersagen treffen. So bestimmen z.B. die *Konzentrationen* von Fumarat und Succinat, sowie der Redoxzustand des Ubichinonpools, ob die Succ-DH thermodynamisch als Fumarat-Reductase agieren kann oder nicht. Diese Größen werden in den EMen nicht erfaßt. Aus diesem Grund wenden wir uns dem dynamischen Modell der ETK zu.

9.1.3 *Dynamisches Modell der Elektronentransportkette*

In diesem Abschnitt soll der grundlegende Aufbau des *kinetischen Modells* der ETK erläutert und die wichtigsten Modellgleichungen vorgestellt werden. Das Modell für die ETK, das die in Abb. 9.1 skizzierten Prozesse und Komponenten enthält, ist z.T. in Anlehnung an die Arbeit von Korzeniewski [78] entwickelt worden. Dort wurde ein Modell für die oxidative Phosphorylierung in Mitochondrien aufgestellt, die der Atmungskette der SPB ausgesprochen

ähnlich ist. Zusätzlich wurde die Photosynthese, sowie die Regulation (Abschnitt 9.2) mit integriert. Der komplette Satz an Modellgleichungen findet sich in Anhang F.

Bereits mehrfach wurde erwähnt, daß die protonenpumpende Tätigkeit von Membrankomplexen der Aufrechterhaltung eines elektrochemischen Protonengradienten dient, aus dem sich die protonenmotorische Kraft (*protonmotive force*; abgek. *pmf* oder Δp) konstituiert. Diese setzt sich zusammen aus dem Konzentrationsgradienten (pH-Gradient: $\Delta \text{pH} = \text{pH}_p - \text{pH}_c$) und dem Membranpotential ($\Delta \Psi$) [96]:

$$pmf \text{ (mV)} = \Delta \Psi - 2.3 \cdot R \cdot T \cdot \Delta \text{pH} / F \quad (9.12)$$

(R =Gaskonstante; F =Faraday-Konstante, $T=298\text{K}$ =Temperatur). *pmf* steht in direkter Verbindung mit $\Delta \bar{\mu}_p$, das die Änderung ΔG der (Gibbsschen) freien Enthalpie beschreibt, wenn ein Mol Protonen vom Periplasma (außen) über die Membran zurück ins Cytoplasma fließt:

$$\Delta \bar{\mu}_p \text{ (kJ/mol)} = \Delta G = -F \cdot \Delta \Psi + R \cdot T \cdot \ln \frac{[H^+]_c}{[H^+]_p} = -F \cdot \Delta \Psi + 2.3 \cdot R \cdot T \cdot \Delta \text{pH} = -F \cdot pmf \quad (9.13)$$

Im Gegensatz zu ΔpH ist das Membranpotential $\Delta \Psi$ nicht allein über innere und äußere Protonenkonzentration definiert, sondern stellt einen integralen Wert dar, der die Verteilung der Ionen auf beiden Seiten der Membran widerspiegelt. Obwohl nur näherungsweise und nur unter bestimmten Bedingungen korrekt [144], nehmen wir im Modell einen linearen Zusammenhang zwischen ΔpH und $\Delta \Psi$ an, d.h. aus ΔpH können wir $\Delta \Psi$ über einen vorgegebenen Faktor berechnen. Dieser wurde aus [172] abgeleitet und besagt, daß sich ca. 30% des *pmf* aus dem pH-Gradienten ergeben ($\Delta \text{pH} \approx 1$). Der pH-Wert im Innern der Zelle (pH_c) wird als konstant 7 angenommen. Die Bilanzierung der Protonen im Periplasma [H_p^+] erfolgt über die berechneten Flußraten in den Enzymkomplexen (s.u.) und über einen Parameter, der spezifiziert, wieviel Mol H_p^+ je Liter Zellvolumen ins Periplasma gepumpt werden müssen, damit sich die Konzentration [H_p^+] um ein mol/l erhöht. Dieser (geschätzte) Parameter hängt von der Pufferkapazität der Protonen im Periplasma und vom Volumen des Periplasmas ab. Zusätzlich wird im Modell berücksichtigt, daß bei steigendem *pmf* Protonen verstärkt über die Membran zurückfließen, ohne dabei Arbeit zu verrichten (*proton leak*).

Mit Hilfe von *pmf* kann die ATP-Synthase die Phosphorylierung von ADP zu ATP betreiben (9.11). Als kinetischen Ansatz setzen wir die Reaktionsrate der ATP-Synthase negativ proportional zur Änderung der (Gibbsschen) freien Energie in dieser Reaktion:

$$r_{ATP\text{Synthase}} = -k_{ATP\text{Synth}} \Delta G_{ATP\text{Synth}} = k_{ATP\text{Synth}} (\Delta G_{0,\text{pH}=7}^{ATP\text{Synth}} - RT \ln \frac{[ATP]}{[ADP] \cdot [P_i]} - h \cdot \Delta \bar{\mu}_p) \quad (9.14)$$

Der Parameter h gibt die Menge an Protonen H_p^+ an, die bei der Synthese von 1 ATP über die ATP-Synthase ins Cytoplasma zurückfließen. Gewöhnlich ist $h = 3$. $\Delta G_{\text{ATP-Hydrolyse}}^{\text{standard}}$ ist die freigesetzte Gibbsche Energie der ATP-Hydrolyse ($ATP^+ + H_2O \leftrightarrow ADP^3 + Pi^2 + H_p^+$) unter biologischen Standardbedingungen (mit $pH=7$) und beträgt -30.5 kJ/mol . Die Phosphatkonzentration $[P_i]$ in der Zelle wird als Parameter des Modells festgelegt.

Schließlich muß in (9.14) noch das Verhältnis von ATP/ADP bekannt sein. Da diese Metabolite in vielen anderen Reaktionen des Stoffwechsels als Cofaktoren auftreten, ist es praktisch unmöglich, eine vollständige Bilanz von ATP und ADP auf Basis eines kinetischen Modells zu geben. Auf der anderen Seite können wir nicht davon ausgehen, daß z.B. unter aeroben, semiaeroben und phototropen Bedingungen immer gleiche Konzentrationsverhältnisse vorliegen, d.h. wir können in (9.14) i.a. keinen konstanten Wert annehmen. Deshalb nutzen wir folgende Strategie: Die Konzentrationen von ATP und ADP werden im Modell für die jeweils simulierte Phase vorgegeben (bspw. von Messungen in steady-state Experimenten), d.h. ATP und ADP sind (variable) Parameter und keine Zustandsgrößen. Nachteil dieser Methode ist, daß im realen Fall beim Wechsel der Umweltbedingungen (z.B. aerob zu phototroph) die Konzentrationsverhältnisse nicht sofort einen neuen Level annehmen werden, während im Modell sofort der neue Wert festgelegt wird. Da uns hier aber ohnehin die Gleichgewichtszustände und weniger die transienten Phasen interessieren werden, sollte diese Vorgehensweise unkritisch sein. Dafür müssen wir uns um den ATP-Stoffwechsel abseits der ETK nicht mehr kümmern und „kapseln“ so das Modell gegen den Rest des Stoffwechsels ab (Abb. 9.1). Die gleiche Vorgehensweise wird auch bei NADH/NAD und Succinat/Fumarat angewandt.

Entscheidend für die Dynamik der ETK sind die Geschwindigkeiten der Elektronentransportprozesse, die von den oben beschriebenen Enzymkomplexen katalysiert werden. Für die meisten Komplexe ist die Funktion klar definiert, wie aber die Elektronen in den molekularen Maschinen von einem Redox-Carrier auf einen anderen gelangen, ist oft noch unverständlich. So besteht die NADH-DH bspw. aus 14 Untereinheiten mit mehreren Eisen-Schwefel-Zentren und der genaue Weg der Elektronen und Protonen ist noch nicht aufgeklärt [96].

Deshalb betrachten wir die Übertragung der Elektronen vom Donator zum Akzeptor in den meisten Fällen approximativ als einen Schritt, d.h. wir gehen von einer schnellen Weiterleitung der Elektronen innerhalb des Enzymkomplexes aus. Damit können wir die Nettostöchiometrien (9.4)-(9.9) als Redoxreaktionen auffassen. Treibende Kraft von Redoxreaktionen ist ein Gradient im Redoxpotential der beteiligten Redoxpaare. Ein Redoxpaar kennzeichnet oxidierte und reduzierte Form einer Spezies [96]. So sind Q/QH_2 , c_2^{2+}/c_2^{2+} und $NAD^+/NADH$ die Redoxpaare der Redox-Carrier und Gl. (9.1)-(9.3) zeigen, wie sie durch Elektronenaufnahme bzw. -abgabe auseinander hervorgehen. Das Standard-Redoxpotential $E_{m,7}$ eines Redoxpaares (ermittelt unter Standardbedingungen in 1-molarer Lösung; hier für biologisch relevanten pH-Wert von 7) deutet die Neigung an, Elektronen zu binden (positiveres $E_{m,7}$) oder abzugeben (negativeres $E_{m,7}$). Abb. 9.2 zeigt die Standard-Redoxpotentiale der beteiligten Redoxpaare in der ETK. Für reale Bedingungen müssen aber aktuelle Konzentrationen der oxidier-

ten/reduzierten Form der Spezies S und die Temperatur mit in das aktuelle Redoxpotential E einbezogen werden (*Nernstsche Gleichung*):

$$E = E_{m,7} + \frac{RT}{zF} \cdot \ln \frac{[S^{ox}]}{[S^{red}]} \quad (9.15)$$

z ist die Anzahl transportierter Elektronen. Je positiver (9.15) ist, desto stärker zieht Spezies S Elektronen an. Da Elektronen nie frei auftreten, sind an einer Redoxreaktion immer mind. zwei Redoxpaare beteiligt: Donor-Redoxpaar (Spezies D) und Akzeptor-Redoxpaar (Spezies A). Für die Gesamtbilanz des Redoxpotentials im System ergibt sich dann:

$$\Delta E = \Delta E_{m,7} + \frac{RT}{zF} \cdot \ln \frac{[A^{ox}][D^{red}]}{[A^{red}][D^{ox}]} \quad (9.16)$$

Das Redoxpotential steht direkt mit der Änderung der freien Energie in Beziehung, die sich während der Redoxreaktion ergibt:

$$\Delta G = -z F \Delta E \quad (9.17)$$

Wie bei der ATP-Synthase setzen wir die Reaktionsrate der Redoxreaktionen proportional zu $-\Delta G$ (und damit proportional zu ΔE). Für einige Komplexe müssen wir in (9.17) aber noch berücksichtigen, daß gleichzeitig Protonen gegen die pmf über die Membran transportiert werden, was die Triebkraft reduziert.

Wir illustrieren das am Beispiel der NADH-DH, die zwei Elektronen von NADH auf Ubichinon (Q) überträgt und dabei vier Protonen nach außen pumpt (Gl. (9.4)). In (9.16) ergibt sich:

$$\Delta E_{NADH-DH} = \Delta E_{m,7}^{NADH/QH_2} + \frac{RT}{2F} \cdot \ln \frac{[Q][NADH]}{[QH_2][NAD]} \quad (9.18)$$

Es ist $\Delta E_{m,7}^{NADH/QH_2} = 390\text{mV}$, da $E_{m,7}^{QH_2} = +70\text{mV}$ und $E_{m,7}^{NAD/NADH} = -320\text{mV}$. Die Konzentrationen von Q und QH_2 sind Zustandsgrößen im Modell, während das Verhältnis $[NADH]/[NAD]$ aus oben genannten Gründen vorgegeben wird.

Für die Reaktionsrate $r_{NADH-DH}$ nutzen wir Gl. (9.17), wobei also $z=2$ ist. Allerdings muß vom treibenden Redoxpotentialunterschied noch der Aufwand abgezogen werden, der zum Pumpen der vier Protonen gegen die Richtung der pmf nötig ist:

$$r_{NADH-DH} = -k_{NADH-DH} \Delta G_{NADH-DH} = k_{NADH-DH} F (2\Delta E_{NADH-DH} - 4 pmf) \quad (9.19)$$

Die Multiplikation von pmf mit vier resultiert aus den vier gepumpten Protonen. Zwar ist $\Delta E_{m,7}^{NADH/QH_2}$ recht groß (siehe Gefälle in Abb. 9.2), dennoch kann die Reaktionsrate in

(9.19) auch negativ werden – ganz in Abhängigkeit von aktuellen Konzentrationen und pmf . Dieser Fall entspricht dem oben diskutierten *umgekehrten* Elektronenfluß, wo Elektronen mit Hilfe des pmf aus der ETK (von QH_2 zu NAD) herausströmen (EM2-EM4; Tab. 9.1).

Eine andere Besonderheit tritt bei Enzymkomplexen auf, wo Elektronen quer zur Membran transportiert werden. Das ist z.B. beim bc_1 -Komplex der Fall, wo die Elektronen vom lipophilen QH_2 (in der Membran) zum c_2 transportiert werden, das an der äußeren Membranseite sitzt. Dieser Quertransport wird begünstigt durch das *Membranpotential* und wird mit dem letzten Term in folgender Gleichung für r_{bc_1} berücksichtigt.

$$r_{bc_1} = -k_{bc_1}[bc_1] \Delta G_{bc_1} = k_{bc_1}[bc_1] F(2\Delta E_{bc_1} - 4pmf + 2\psi) \quad (9.20)$$

Da die Konzentration des bc_1 -Komplexes im Modell eine explizite Zustandsgröße ist (während für $[NADH-DH]=\text{konstant}$ gilt), muß diese explizit in (9.20) berücksichtigt werden.

Die Rate der Succ-DH kann analog zu (9.19) beschrieben werden, wobei keine Protonen oder Elektronen vertikal zur Membran transportiert werden. Für die Ubichinon- und c_2 -Oxidase wurde wie in [78] jeweils eine Zwischenstufe innerhalb des Komplexes berücksichtigt, die per Annahme im Redoxgleichgewicht mit QH_2 bzw. c_2^{2+} sind. Von dort fließen die Elektronen mit Michaelis-Menten-Kinetik auf Sauerstoff ab. Die Photosynthese im RC wurde nach Abb. 9.3 über Massenwirkungsgesetz beschrieben.

9.2 Physiologie und Modellierung der Redoxregulation

Die Synthese photosynthetischer Apparate und ihr Einbau in die intracytoplasmatische Membranen (Chromatophoren), die sich bei phototrophem Wachstum ausbilden, sind aufwendige Prozesse. Deshalb ist es nicht überraschend, daß die SPB diese Synthese regulieren und entsprechend den Umgebungsbedingungen anpassen. Wenn die Photosynthese als Energielieferant dienen soll, werden die in einem Cluster angeordneten *Photosynthesegene* (PS-Gene, [2]) exprimiert und die Synthese der photosynthetischen Komplexe gestartet. Zu den wichtigsten PS-Genen gehören das *puf-Operon* (kodiert Proteine des RC und LH1), *pwhA* (große Unter-einheit des RC), die *bch*-Gene (Bacteriochlorophyll-Synthese) und die *crt*-Gene (Carotinoid-Synthese). Das *puc*-Operon in *R. sphaeroides* und *R. capsulatus* kodiert Proteine des LH2. Neben diesen PS-Genen gibt es eine Reihe anderer Gene, deren Expression die Zusammensetzung der ETK beeinflusst und daher ebenfalls beim Wechsel vom respirativen zum photosynthetischen Wachstum angepaßt werden. Dazu gehören Gene für die Tetrapyrrol-Synthese (*hem*-Gene), Cytochrom c_2 (*cycA*) und cbb_2 -Oxidase (*cco*-Operon).

Drei Signale aus der Umwelt sind für die Regulation der PS-Gene entscheidend (Abb. 9.4):

- *Sauerstoff* ist das dominierende Signal. Unabhängig davon, ob eine Lichtquelle zur Verfügung steht oder nicht, wachsen SPB unter aeroben Bedingungen respirativ und die Bil-

dung photosynthetischer Komplexe ist reprimiert. Unter semiaeroben Bedingungen, ($pO_2 < 1\%$) wird die Membran- und Pigmentsynthese (auch im Dunkeln) induziert.

- Die *Lichtintensität* ist zu einem deutlich geringeren Anteil als Sauerstoff an der Regulation der PS-Gene beteiligt. Während photosynthetischem Wachstum wird bei geringer Lichtintensität die Produktion von Pigmenten erhöht, um letztlich mehr Lichtsammelkomplexe (bei *R. sphaeroides* und *R. capsulatus* insbesondere vom Typ LH2) zur Verfügung zu stellen und somit die Ausbeute an absorbierten Lichtquanten zu steigern.
- Das angebotene *Substrat* scheint ebenfalls einen Einfluß auf die Expression der PS-Gene zu haben, wie sich aus Studien mit *R. rubrum* schlußfolgern ließ [46,53]: Bei semiaerobem Wachstum (im Dunkeln) auf einem Substratgemisch mit Succinat und Fructose ergab sich eine signifikant höhere Pigmentierung der Zellen, als ausschließlich auf Succinat oder Fructose. In [53] wurde die Hypothese aufgestellt, daß die gewählte Substratkombination *indirekt* die Expression der PS-Gene beeinflußt, weil Signale, die eigentlich oben genannte Größen (Sauerstoffkonzentration, Lichtintensität) erfassen sollen, moduliert werden.

Die zentrale Frage bei der Regulation der PS-Gene ist, wie O_2 -Konzentration und Lichtintensität gemessen werden. Nach bisherigem Kenntnisstand weiß man, daß zumindest einige Signalwege diese Größen nicht direkt erfassen. Bereits in den 50-er Jahren wurde die Hypothese aufgestellt, daß der *Redoxzustand* der ETK eine integrierende Größe für O_2 -Konzentration und Lichtintensität ist und deshalb ein geeignetes Signal für die Regulation der PS-Gene sein könnte [19]. In den 90-er Jahren konnten Signaltransduktionswege in *R. sphaeroides* und *R. capsulatus* identifiziert werden, die für die Regulation der PS-Gene maßgeblich und in beiden Spezies weitgehend homolog sind. Die drei wichtigsten Sensoren bzw. Regulatoren in *R. sphaeroides* sind folgende (Abb. 9.4; nach Übersichtsartikel [32,102,103]):

PrrB/PrrA (*R. capsulatus*: RegA/RegB): Hierbei handelt es sich um ein typisches Zweikomponenten-System [83]. PrrB ist eine membranständige Sensor-Histidin-Kinase. Unter bestimmten Bedingungen (s.u.) führt PrrB eine Autophosphorylierung durch und kann daraufhin den löslichen Response Regulator PrrA phosphorylieren. PrrA-P kann dann an Promotor-Sequenzen binden und maßgeblich die Expression vieler PS-Gene aktivieren. Zusätzlich aktiviert PrrA-P auch die Expression von c_2 und *inhibiert* die Transkription der *cbb_3*-Oxidase [32,131]. Die globale Rolle des PrrBA-Systems spiegelt sich schließlich darin wider, daß es auch die Fixierung von Kohlendioxid und Stickstoff reguliert [66].

Bleibe die Frage, wann PrrB autophosphoryliert und die Kaskade in Gang setzt. Eine zentrale Rolle spielt hier die *cbb_3*-Oxidase: Man nimmt an, daß unter aeroben Bedingungen ein inhibitorisches Signal von der die *cbb_3*-Oxidase zu PrrB geleitet wird und dessen Kinase-Aktivität hemmt (bzw. Phosphatase-Aktivität erhöht). Zunächst war unklar, welcher Mechanismus das inhibierende Signal hervorruft. Unterstützt durch die Analyse verschiedener Mutanten stellten Oh und Kaplan [102] die Hypothese auf, daß ein hoher *Fluß* an Elektronen durch die Oxidase (von c_2 auf Sauerstoff) zu einem inhibitorischen Signal führt, das vermutlich über einen drit-

ten Mitspieler, PrrC, zu PrrB geleitet wird [31]. Folglich ist die Kinase-Aktivität von PrrB unter anaeroben Bedingungen am größten.

Im Modell existieren entsprechend zwei Zustandsgrößen, *PrrB-P* und *PrrA-P*, die für den phosphorylierten Anteil der Kinase bzw. des Regulators stehen. Die Gesamtkonzentrationen von Kinase und Regulator sind auf 1 normiert, d.h. $PrrA=1-PrrA-P$ und analog für PrrB. Die Autophosphorylierung von PrrB ist kinetisch in Abhängigkeit des Umfanges des Elektronenflusses durch die Oxidase inhibiert. *PrrB-P* kann PrrA phosphorylieren und dephosphoryliertes PrrB kann *PrrA-P* dephosphorylieren. *PrrA-P* beeinflusst dann maßgeblich die Expressionsraten von PS-Genen und von anderen ETK-Komponenten (Abb. 9.4 und Anhang F).

PpsR (*R. capsulatus*: CrtJ): PpsR ist ein Repressor von PS-Genen. Seine voll inhibierende Funktion entfaltet er unter aeroben Bedingungen, wo PpsR als Tetramer an seine Zielpromotoren auf der DNA bindet und so das Ablesen der Gene verhindert [88]. Welches Signal die Aktivität von PpsR tatsächlich bestimmt, ist noch nicht völlig geklärt. Vieles deutet nun aber darauf hin, daß PpsR den Redoxzustand des Ubichinonpools erfäßt: je oxidiert dieser ist, desto stärker ist die Repressoraktivität von PpsR. Allerdings ist bisher nicht eindeutig klargestellt oder nachgewiesen worden, wie z.B. die Lichtintensität das Q/QH_2 -Verhältnis beeinflusst. Hier wird uns die Modellierung wichtige Ergebnisse liefern (s.u.).

PpsR steht nicht im direkten Kontakt mit der ETK [88,103]. In *R. sphaeroïdes* konnte mit AppA ein Gegenspieler von PpsR identifiziert werden. Dieser Antirepressor besitzt ein redoxaktives Zentrum (FAD) und scheint darüber den Q/QH_2 -Redoxzustand zu erfassen. Ist AppA oxidiert, kann PpsR als Repressor aktiv werden. Ist der Ubichinonpool dagegen stärker reduziert (z.B. unter anaeroben Bedingungen) geht auch AppA verstärkt in den reduzierten Zustand über und kann dann den Repressor PpsR blockieren [88]. Daneben scheint AppA auch als Blaulichtsensor fungieren zu können, wodurch die Antirepressor-Aktivität moduliert wird [88]. Darauf gehen wir hier aber nicht weiter ein.

Im Modell existieren zwei – wiederum normierte – Zustandsgrößen zum Abbilden des PpsR-Redoxsignalweges: *AppA_red* (reduzierter Anteil von AppA) und *PpsR_blocked* (blockierter Anteil von PpsR = inaktive Form). AppA tauscht mit dem Q/QH_2 -Pool Elektronen aus (kinetisch in Analogie zu Gl. (9.16) und (9.17)) und erfäßt so den Redoxzustand des Ubichinons. *AppA_red* kann PpsR blocken. Der verbleibende Anteil von PpsR ($= 1 - PpsR_blocked$) wirkt als Repressor für die Synthese von Bchl, RC und LH (Abb. 9.4).

FnrL: Dieser dritte Regulator ist ein Homolog des globalen anaeroben Regulators Fnr von *E. coli* [83]. FnrL geht unter anaeroben Bedingungen in seine aktive (Dimer-)Form über, in der es an die DNA binden und dort als Aktivator für die Gene der ebb_2 -Oxidase und für Gene der Tetrapyrrol-, Bchl- und Cytochrom c_2 -Synthese wirkt [32,103]. FnrL repräsentiert auch andere Gene, die für uns hier aber nicht von Bedeutung sind.

Nicht genau verstanden ist bisher, wie unter aeroben Bedingungen die aktive Form des FnrL (Dimer) in eine inaktive (Monomer) überführt wird, allerdings scheint Sauerstoff wie bei Fnr in *E. coli* direkt für die Konformationsänderung von FnrL verantwortlich zu sein [102].

Dementsprechend bezeichnet *FnrL_dim* im Modell den differenzierten Anteil von *FnrL*, der direkt über einen Inhibitionsterm aus der aktuellen Sauerstoffkonzentration berechnet wird.

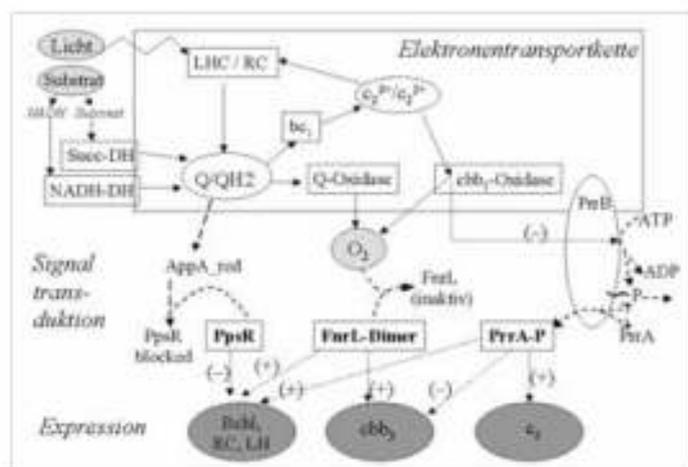


Abb. 9.4: Redoxabhängige Signaltransduktionswege der Rhodospirillaceae zur Regulation der Expression von Komponenten der ETK, insbesondere photosynthetischer Komplexe. Umweltreize (Sauerstoff, Substrat, Licht) sind hellgrün unterlegt, die Gene der ETK-Komponenten in dunkelgrün. Signalwege sind gestrichelt gezeichnet.

In der Literatur finden sich noch keine Studien, die sich mit obigen Signalwegen in *R. rubrum* beschäftigen. Allerdings wurden bereits ein *PpsR*-Homolog [1], sowie *PrtA*-Konsensus-Sequenzen auf der DNA [52] identifiziert. Insofern erscheint es gerechtfertigt, in *R. rubrum* zumindest approximativ von einer analogen Regulation auszugehen, wie bei den Verwandten *R. sphaeroides* und *R. capsulatus*. Generell ist das Modell nur an wenigen Stellen spezifisch für *R. rubrum*, so daß die Resultate auch für die anderen beiden Vertreter relevant sein sollten.

Expressionsraten von Komponenten der ETK

PrtA-P, *PpsR* und *FnrL* wirken oft in Kombination auf die Expression verschiedener Gene ein und konkurrieren oder kooperieren um Bindestellen auf der DNA [102,152]. Im Modell wurden dazu Inhibierungs- bzw. Aktivierungs-Terme multiplikativ verknüpft. Weitere Annahmen sind folgende:

- Die Expression von Ubichinon und *bc1*-Komplex erfolgen approximativ im konstanten Verhältnis zur Expressionsrate von *c2* (siehe auch [43]).
- Die Regulation der Ubichinon-Oxidase erfolgt ähnlich der *cbb1*-Oxidase (allerdings mit geringerem konstitutiven Expressionslevel, da sie aerob kaum exprimiert ist [152]).
- Entsprechend den Gegebenheiten in *R. rubrum* erfolgt die *Bchl*-Synthese im konstanten Verhältnis zu *RC* und *LH*, d.h. auch *LH* und *RC* sind im Verhältnis konstant.

9.3 Modellstudien

Das dynamische Modell der ETK (einschl. der Regulation) wurde in MATLAB implementiert und besteht aus 21 Zustandsgrößen, sowie 32 algebraisch abgeleiteten Größen (Anhang F). 58 Parameter sind im Modell enthalten, die größtenteils geschätzt wurden, und zwar so, daß vorhandene (steady-state) Meßpunkte wiedergegeben werden konnten (Abb. 9.5).

In diesem Abschnitt werden wir Modellsimulationen für verschiedene Sauerstoffkonzentrationen und Lichtintensitäten besprechen. Dabei soll untersucht werden, welche Redoxzustände sich in der ETK unter den jeweiligen Bedingungen einstellen und ob und wie die drei Regulationssysteme PrtBA, AppA/PpsR und FnrL die Expression der PS-Gene geeignet koordinieren können. Insbesondere soll die Frage geklärt werden, wie die Lichtintensität und das Substrat den Redoxzustand des Ubichinonpools und damit über PpsR die Expression photosynthetischer Komponenten beeinflussen.

9.3.1 Simulationen für verschiedene Sauerstoffkonzentrationen und Lichtintensitäten

Abb. 9.5 zeigt Simulationsergebnisse für vier verschiedene (aufeinanderfolgende) Umweltkonstellationen:

- (i) [0 – 200 h] : aerobes Wachstum im Dunkeln (Gelöstsauerstoff= $pO_2=10\%$)
- (ii) [200 – 400 h] : semiaerobes Wachstum im Dunkeln ($pO_2=0.1\%$)
- (iii) [400 – 600 h] : anaerobes Wachstum im Starklicht ($400 \mu E/(m^2 \cdot sec)$)
- (iv) [600 – 800 h] : anaerobes Wachstum im Schwachlicht ($25 \mu E/(m^2 \cdot sec)$)

Wie in Abschnitt 9.2 erläutert, werden die Verhältnisse von NADH/NAD, ATP/ADP und Succinat/ Fumarat für jede der vier Phasen (separat) vorgegeben. Für die ersten beiden wurden eigene Meßwerte und Angaben aus der Literatur herangezogen, die sich bei Wachstum auf Malat bzw. Succinat ergaben [52,135,136]. Das Succinat/Fumarat-Verhältnis wurde mangels verfügbarer Daten in allen vier Phasen auf einen gemeinsamen Schätzwert gesetzt.

Aerobe Wachstumsphase

Bei aeroben Verhältnissen liegt die ETK in einem stark oxidierten Zustand vor, was sich insbesondere am Redoxzustand des Ubichinons (*Ubi_charge*) zeigt: nur etwa 8% liegt reduziert als QH_2 vor, in Übereinstimmung mit ermittelten Meßdaten [52]. Elektronen werden zwar mit hohem Durchsatz von NADH, sowie auch von Succinat in die ETK eingeschleust (EM6 und EM8 von Tab. 9.1 dominieren), Sauerstoff als finaler Akzeptor zieht die Elektronen aber ebenso zügig ab. Die Limitierung liegt hier also in der Anlieferung von Elektronen, was sich auch im niedrigen NADH/NAD-Verhältnis zeigt. Im steady-state pegelt sich pmf bei 186 mV ein und liegt damit, wie in den anderen drei Phasen auch, im physiologisch nachgewiesenen Bereich, was die Validität des gewählten Modellansatzes unterstreicht.

Die dominant vorliegenden Formen der drei Regulatoren PrtA-P (hauptsächlich unphosphoryliert und damit inaktiv), PpsR (weitgehend frei als Repressor) und FnrL (hauptsächlich als

inaktiver Monomer) lassen in dieser Phase eine Expression von photosynthetischen Komplexen kaum zu. Dementsprechend sind die Kulturen im Experiment weitgehend farblos.



Abb. 9.5: Ausgewählte Modellgrößen in Simulationen der ETK (vier verschiedene Wachstumsphasen; Dauer jeweils 200 Stunden). Mittlere Lichtintensität (*mean_irradiance*) und Sauerstoffkonzentration (pO_2) sind für jede der vier Phasen vorgegeben, ebenso die Verhältnisse von NADH/NAD, ATP/ADP und Succinat/Fumarat (erste Zeile). Kreise in den Plots deuten Meßwerte in *steady-state* Experimenten an (c_2 - und Ubichinon-Meßpunkte, sowie Bchl-Levels für aerobe und semiaerobe Bedingungen nach [52]; Bchl-Meßwerte anaerob im Stark- und Schwachlicht nach [49]). Diskussion im Text.

Semiaerobe Wachstumsphase

Sinkt die Konzentration von Gelöstsaauerstoff auf einen niedrigen Wert unter 1%, wird Sauerstoff zur limitierenden Ressource. Unmittelbar nach Absenkung stauen sich die Elektronen in der ETK sofort an und sorgen dafür, daß der Ubichinonpool kurzzeitig zu über 90% reduziert ist und der Zustrom von Elektronen fast zum Erliegen kommt. Als Folge des verringerten Flusses von Elektronen zum Sauerstoff in der *ebb₃*-Oxidase (*Flux_{per}Oxidase*) phosphoryliert PrrB zunächst sich selbst und dann PrrA. Gleichzeitig geht FnrL verstärkt in die aktive Dimer-Form über. Außerdem steht AppA in Kontakt mit der ETK und ist nun, wie Ubichinon, stärker reduziert. So kann AppA nun als Antirepressor aktiv werden und PpsR blockieren. Als eine Folge dieser Signaltransduktionen wird die Expression der PS-Gene (*Bchl*, *RC*) gestartet und eine Pigmentierung setzt ein – obwohl im Dunkeln kein physiologischer Nutzen daraus gezogen werden kann. Anders ist es im Fall der verstärkten Synthese von Redox-Carriern und insbesondere der erhöhten Expression von *ebb₃*- und Ubichinon-Oxidase (vgl. [151]). Letztere führt die Elektronen nun auch direkt von QH₂ auf Sauerstoff ab und sorgt somit für eine Entlastung des Ubichinonpools, wenn auch auf Kosten einer geringeren Energieausbeute (EM5 und EM7, Tab. 9.1). Im Gleichgewicht ist die ETK jetzt reduzierter als im voll aeroben Zustand, allerdings konnten die Redoxzustände auf einem niedrigeren Niveau als unmittelbar nach Umschaltung stabilisiert werden. Der Elektronenfluß ist aber verringert, insbesondere relativ zur Menge an *ebb₃* (*Flux_{per}Oxidase*), wodurch weiterhin ein Signal über PrrBA anliegt. Aufgrund des reduzierten Stromes durch die ETK liegt im steady-state auch ein geringerer Level an *pmf* und ein geringerer Strom durch die ATP-Synthase vor, was die Wachstumsrate bei ATP-Limitierung verringert (hängt stark vom pO₂ ab).

Anaerobes Wachstum in Starklicht

Wachsen die SPB nun unter vollständig anaeroben Verhältnissen im Licht, kann die Respiration keinesfalls mehr als Energiequelle genutzt werden und es erfolgt eine vollständige Umschaltung auf Photosynthese. Da nun kein Sauerstoff mehr vorhanden ist, liegen PrrA-P und FnrL fast vollständig phosphoryliert bzw. als Dimer vor, was zu einer weiter erhöhten Expression von Photosynthesekomplexe führt (siehe *Bchl*-Meßwert). Der Ubichinonpool ist im steady-state nur wenig reduzierter (je nach Parameterwerten sogar oxidiert) als bei semiaeroben Verhältnissen, weshalb sich die Repressoraktivität von PpsR nur marginal verschiebt. Da PrrA-P einen reprimierenden Effekt auf die Expression der Oxidasen hat, sinkt deren Konzentration nun wieder ab [151]. Allerdings muß die *ebb₃*-Oxidase weiterhin auf einem niedrigeren Level vorhanden sein, um das PrrBA-Signals generieren zu können.

Die ETK wurde nun also von Respiration auf Photosynthese umgestellt. Die Reaktionszentren liegen dabei mit einem nur geringen Anteil im Zustand RCs1 bzw. RCs4 vor (Abb. 9.3). Letztere beiden benötigen für ihren Übergang in den jeweils nächsten Zustand Lichtquanten, die hier in sättigender Intensität die Lichtsammelkomplexe erreichen.

Die Summe ein- und austretender Elektronen in die ETK muß auch bei der Photosynthese im steady-state Null sein. In der Tat werden die über die Succinat-Dehydrogenase eingetragenen Ladungen (*Succ_{DH}flux*) über die NADH-Dehydrogenase (*NADH_{DH}flux*) mittels *reverse*

electron flow wieder heraustransportiert (vgl. EM2, Tab. 9.2). Die Raten hängen dabei vom NADH/NAD- bzw. Succinat/Fumarat-Quotienten ab.

Anaerobes Wachstum im Schwachlicht

Unter Schwachlichtbedingungen wird die Lichtintensität für die Photosynthese limitierend: nunmehr verharren 80% der Reaktionszentren im Zustand RCs1 oder RCs4 und „warten“ dort auf Anregungen, die von den LHs durch Absorption von Photonen zur Verfügung gestellt werden müssen. Um weiterhin einen hohen photosynthetischen Fluß zu gewährleisten, steigert *R. rubrum* die zellulären Konzentrationen von RC und LH im gleichen Verhältnis. Die benötigte Verstärkung der Expression kann im Modell nur noch über AppA/PpsR signalisiert werden, da sowohl PrrBA, als auch PnrL unter anaeroben Bedingungen bereits automatisch voll aktiviert sind und somit nicht mehr im Sinne der Lichtintensität differenzieren können. Damit AppA die Repressoraktivität von PpsR weiter einschränken kann, muß der Ubichinonpool reduzierter sein, als unter Starklichtbedingungen. Und genau das tritt auch ein (Abb. 9.5). Ein Grund ist der erhöhte NADH-Spiegel in der Zelle, der vermutlich daher rührt, daß nicht genügend ATP für die Redoxsenke Calvin-Zyklus bereitgestellt werden kann. Allerdings kann das nicht die alleinige Ursache sein, denn setzt man das NADH/NAD-Verhältnis auf den gleichen Wert wie unter Starklichtbedingungen, ändern sich die Redoxzustände von c_2 und Ubichinon nur geringfügig. Im Vergleich zu Starklichtbedingungen hätte man vermuten können, daß sich die Elektronen auf Seiten von c_2 anstauen, da sie dort auf die vom RC vermittelte Rückübertragung auf Q warten. Tatsächlich ist der Redoxzustand aber sowohl vom c_2 als auch vom Ubichinon reduzierter. Ursache ist folgende: Aufgrund des verringerten Flusses der zyklischen Photosynthese werden weniger Protonen nach außen gepumpt, woraufhin das *pmf* abfällt. Das hemmt nun den *reverse electron flow* von Succinat über QH_2 zu NAD, da hierbei *pmf* eine entscheidende Rolle spielt (siehe Gl. (9.19), die für den umgekehrten Fluß mit -1 multipliziert werden muß). Selbst wenn keine Elektronen vom Succinat in die ETK fließen, führt die verringerte *pmf* wegen Gl. (9.19) zu einem reduzierteren Zustand des Ubichinonpools, da dann die Elektronen vom NADH stärker in die ETK „drücken“.

Man könnte meinen, der beobachtete Effekt könnte von den Werten der (geschätzten) Modellparametern abhängen. Um dieser Frage nachzugehen, wurde mit einem evolutionären Algorithmus [123] eine Parametersuche durchgeführt. Als Zielkriterium sollten die in Abb. 9.5 dargestellten Meßwerte wiedergeben werden und zusätzlich die Forderung erfüllt sein, daß der Ubichinonpool im Schwachlicht weniger reduziert ist als im Starklicht. Es konnte auch nach längerer Rechenzeit keine Parameterkonstellation gefunden werden, die diese Forderung erfüllt. Somit scheint das Verhalten der ETK im Bezug auf den Redoxzustand des Ubichinonpools unter Schwachlichtbedingungen eine Struktureigenschaft des Modells zu sein.

Aufgrund der verstärkten Inaktivierung des Repressors PpsR im Schwachlicht wird die Synthese photosynthetischer Apparate nun also verstärkt (*PSC_synth*), wodurch der *Gesamtfluß* der Photosynthese erhöht und das *pmf* – allerdings auf einem niedrigeren Level – stabilisiert wird. RCs1 und RCs4 bleiben in *R. rubrum* aber weiterhin die dominierenden Zustände der

RC, da sich die spezifische Kapazität der Reaktionszentren zum Absorbieren von Photonen nicht verändert hat. *R. capsulatus* und *R. sphaeroides* verfolgen hier die alternative Strategie, verstärkt Lichtsammelkomplexe vom Typ LH2 einzubauen und damit das LH/RC-Verhältnis zu erhöhen. Im Modell ergäbe sich dann eine gleichmäßigere Verteilung der acht Zustände des RC und ein höheres *pmf* im steady-state (nicht gezeigt). Da außerdem der Aufwand für die LH2-Synthese (zumindest stöchiometrisch) geringer ist als für RC+LH1, erscheint diese Strategie effektiver.

Ein anderer wichtiger Aspekt der erhöhten Konzentration von Bchl, LH und RC im Schwachlicht soll hier noch erwähnt werden: Die u.U. verringerte Wachstumsrate bei Schwachlicht führt automatisch zu einer geringeren Ausdünnung dieser Komponenten. Da photosynthetische Komplexe scheinbar keinem aktiven Abbau unterliegen, kann die reduzierte Wachstumsrate erheblich zur erhöhten Konzentration dieser Komponenten beitragen. Hat z.B. eine Zellkultur A die doppelte Konzentration an Pigmenten wie eine andere Kultur B, wächst aber nur mit der halben Wachstumsrate, so ist die Syntheserate in beiden Kulturen identisch!

9.3.2 Der Redoxzustand des Ubichinons als integrierendes Signal

Wie wir im letzten Abschnitt gesehen haben, scheint der Redoxzustand des Ubichinonpools ein geeignetes Signal zur Regulation der PS-Gene in den SPB zu sein. Ubichinon hat eine ganz zentrale Position in der ETK. Da der über diesen Redox-Carrier laufende Transport von Elektronen von vielen intra- und extrazellulären Größen beeinflusst wird (Abb. 9.4), ist der Redoxzustand des Ubichinons ein stark integrierendes Signal:

- (i) Eine geringe (oder nicht vorhandene) *Sauerstoffkonzentration* führt zu einer stärkeren Reduktion des Ubichinons, da sich Elektronen zurückstauen.
- (ii) Eine geringere *Lichtintensität* führt, wie oben erläutert, zu einer stärkeren Reduktion des Ubichinons aufgrund des fallenden *pmf*.
- (iii) Das NADH/NAD-Verhältnis beschreibt den Redoxzustand des gesamten Stoffwechsels und bestimmt den Druck von Elektronen in den Ubichinonpool.
- (iv) Auch das Succinat/Fumarat-Verhältnis ist von Bedeutung für den Eintrag von Elektronen über die Succinat-Dehydrogenase.

Da die Konzentrationsquotienten NADH/NAD und Succinat/Fumarat sehr wahrscheinlich vom gewählten Substrat mitgeprägt werden, hat das Substrat indirekt einen Einfluß auf den Redoxzustand des Ubichinonpools und somit auf die App/PpsR-gesteuerte Expression photosynthetischer Komponenten. An dieser Stelle können wir nun der Hypothese für den in Abschnitt 9.2 erwähnten Effekt einer verstärkten Expression von PS-Genen bei semiaeroben Wachstum von *R. rubrum* im Dunkeln auf Succinat/Fructose nachgehen (vgl. [53]): Bei dieser Substratkombination liegt vermutlich sowohl der NADH-Spiegel (durch die relativ stark reduzierte Fructose) als auch die Succinat-Konzentration (durch kontinuierliche Aufnahme des Substrates) auf einem hohen Level vor, was einen relativ starken Elektronendruck erzeugt und den Ubichinonpool stärker reduziert, als es bspw. bei Wachstum auf nur einem der beiden

Substrate der Fall wäre. Die Folge ist dann eine verstärkte Expression der PS-Gene unter semiaeroben Bedingungen durch die verminderte PpsR-Repressoraktivität. Vermutlich handelt es sich hier um einen unvermeidbaren Nebeneffekt des integrierenden Redoxsignals.

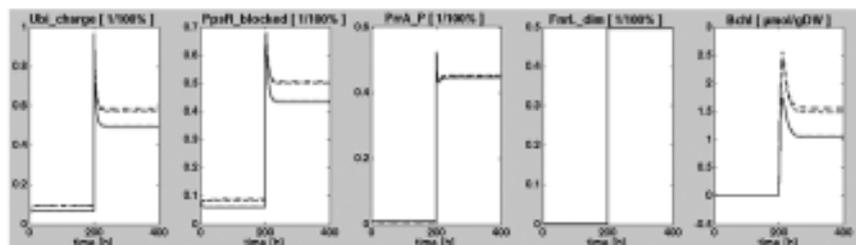


Abb. 9.6: Steigerung der Expression photosynthetischer Gene (als Indikator: Bchl-Konzentration) bei Erhöhung des NADH/NAD- bzw. Succinat/Fumarat-Verhältnisses unter aeroben (0-200h) und semiaeroben (200-400h) Bedingungen. Durchgehende Linie: jeweils gleiche Bedingungen wie in Abb. 9.5; gepunktete Linie: Succinat/Fumarat-Verhältnis um 100% erhöht; gestrichelte Linie: NADH/NAD-Verhältnis um 100% erhöht; Strichpunktlinie: beide Verhältnisse um jeweils 100% erhöht.

Abb. 9.6 zeigt Simulationen dazu. NADH/NAD- und Succinat/Fumarat-Verhältnisse wurden zunächst einzeln und dann zusammen für aerobes und semiaerobes Wachstum um jeweils 100% im Vergleich zur Simulation in Abb. 9.5 erhöht. Der maximale Effekt zeigt sich, wenn beide Verhältnisse gleichzeitig erhöht werden. Unter aeroben Bedingungen steigt der reduzierte Anteil des Ubichinons dabei nur gering von etwa 8 auf 10%, was zu einer leicht erhöhten Menge an blockiertem Repressor PpsR führt. Da die Levels von PprA-P und FnrL fast bzw. völlig unverändert niedrig sind, ist weiterhin, wie im Experiment beobachtet, keine Expression von Bchl festzustellen. Auch unter semiaeroben Bedingungen bleiben die Signale von PprBA und FnrL praktisch unverändert (hier allerdings auf einem mittleren Niveau), jedoch zeigt sich jetzt ein signifikanter Anstieg von Ubi_charge (von 49 auf 59%), was zu einer verstärkten Inaktivierung von PpsR und damit zum verstärkten Ablesen der PS-Gene führt. Infolge dessen steigt die Konzentration photosynthetischer Komponenten weiter an. Das NADH/NAD-Verhältnis hat an diesem Effekt den größeren Anteil.

Neben den vier oben genannten Faktoren gibt es weitere, die den Redoxzustand des Ubichinons mitbestimmen. Werden alternative Elektronenakzeptoren (wie DMSO oder TMAO) angeboten, fließen Elektronen direkt vom QH₂ ab und sorgen somit für einen oxidierteren Zustand, was tatsächlich zu einer geringeren Expression von PS-Genen führt [52]. Auch das ATP/ADP-Verhältnis hat einen zumindest indirekten Einfluß. Dieses Verhältnis bestimmt das Niveau des pmf mit. Je niedriger es ist, desto geringer ist das pmf, da die ATP-Synthase dann schneller ATP produziert (Gl. (9.14)) und somit schneller protonenmotorische Energie verbraucht. Ein geringeres pmf beeinflusst aber wiederum die Geschwindigkeit der ablaufenden Prozesse in der ETK, was sich entsprechend auf den Redoxzustand des Ubichinonpools aus-

wirkt. In einer Simulation stellte sich heraus, daß ein geringeres ATP/ADP-Verhältnis unter allen Bedingungen ebenfalls zu einer stärkeren Reduktion des Ubichinons führt.

Dennoch scheint der Redoxzustand des Ubichinonpools im wesentlichen den *globalen Redoxzustand* der Zelle widerzuspiegeln. Ubichinon ist zentraler Knotenpunkt von Elektronenströmen und sein Redoxzustand, beeinflußt von verschiedensten Umweltreizen, signalisiert kumulativ die globale Verfügbarkeit von Elektronen. Vermutlich deshalb ist diese integrierende Kenngröße ein geeignetes Signal für die Regulation. Auch in anderen Organismen wurde nachgewiesen, daß der Redoxzustand des Ubichinonpools als Signal für die Regulation verschiedener Gene dient [45].

Eine ganz analoges Beispiel, bei der in einem Transportsystem der *Zustand einer Komponente* ein Signal für die Regulation darstellt, ist das Phosphotransferase-System (PTS) zur Zuckeraufnahme in *E. coli* (Abb. 9.7). Im Gegensatz zur ETK, wo Elektronen transportiert werden, wird beim PTS über mehrere Zwischenstufen eine Phosphatgruppe vom Phosphoenolpyruvat (PEP) bis zur finalen Phosphorylierung des Zuckers (z.B. Glucose) übertragen, wobei letzterer gleichzeitig über die Membran in das Zellinnere transportiert wird [83]. Signale werden nun über den Phosphorylierungszustand einzelner Zwischenstufen des PTS abgegriffen. Von besonderer Bedeutung ist z.B. das Enzym EIIA^{Glc} vom Glucose-PTS, da es sowohl am Transport von Glucose beteiligt ist und gleichzeitig als Signalgeber für die Kataboliten-Repression dient. Sinkt die äußere Konzentration von Glucose ab, so steigt der Anteil vom phosphorylierten EIIA^{Glc} (EIIA^{Glc}-P) an, da die Phosphatgruppen nicht mehr auf Glucose übertragen werden können. EIIA^{Glc}-P steigert nun die Aktivität der Adenylatzyklase, einem Enzym, das die Synthese von zyklischem AMP (cAMP) aus ATP katalysiert. Dieses Alarmon der Zelle sorgt nun z.B. dafür, daß die Expression anderer Zucker-PTS über den Regulator CRP aktiviert wird (CRP Modulon). Liegt EIIA^{Glc} dagegen vorwiegend unphosphoryliert vor, so wird die Aufnahme anderer Zucker reprimiert.

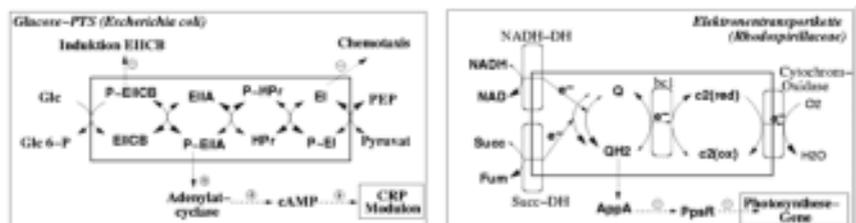


Abb. 9.7: Transportsystem als Signalgeber: Glucose-PTS in *E. coli* (links) und Elektronentransportkette der *Rhodospirillaceae* (rechts). Um eine einfache (und analoge) Darstellung zu bewahren, wurde nur der respirative Arm der ETK über die Cytochrom-Oxidase in der Grafik berücksichtigt.

Der Phosphorylierungszustand von EIIA spielt also eine ganz analoge Rolle wie der Redoxzustand des Ubichinon in der ETK. Während ein hoher Anteil von EIIA^{Glc}-P eine geringe Kon-

zentration von Glucose im Medium signalisiert, deutet ein hoher Anteil von QH_2 auf einen reduzierten Zustand der ETK hin, der durch fehlende Elektronenakzeptoren (Sauerstoff, alternative Akzeptoren, evtl. auch NAD) oder/und auf eine geringe Lichtintensität schließen läßt. Erhöhte Konzentrationen von $EIIA^{Glc}$ -P, als auch von QH_2 werden letztlich durch gebremste Flüsse verursacht. Als Gegenmaßnahme kommt es zur verstärkten Expression alternativer Zuckeraufnahmesysteme (über cAMP und CRP) bzw. von Photosystemen (über PpsR).

Ein wichtiger Aspekt blieb bisher undiskutiert: Sind es die *Verhältnisse* $EIIA^{Glc}$ -P/ $EIIA^{Glc}$ bzw. QH_2 /Q, aus denen die Signale generiert werden, oder sind es die *absoluten Mengen* an $EIIA^{Glc}$ -P bzw. QH_2 ? Nur solange die Gesamtmengen an $EIIA^{Glc}$ bzw. Ubichinon konstant bleiben, spielt dieser Aspekt keine Rolle. Im Modell der ETK der SPB bestimmt das Verhältnis QH_2 /Q das über AppA abgegriffene Signal, das ebenso wie kinetische Prozesse in der ETK mittels Nernstscher Gleichung beschrieben wurde.

9.3.3 Zusammenfassung zum dynamischen Modell der ETK

Die Simulationen im vorangegangenen Abschnitt haben gezeigt, daß ein Zusammenspiel der drei Signalwege über PrrBA, FnrL und AppA/PpsR tatsächlich eine koordinierte Regulation photosynthetischer Gene – entsprechend phänomenologischer Beobachtungen – ermöglichen. Sofern die Modellannahmen korrekt sind, können wir mit den oben gewonnenen Erkenntnissen darauf schließen, daß sowohl FnrL als auch PrrBA vornehmlich das Sauerstoffsignal erfassen und deshalb dominierend für das Umstellen der ETK beim Wechsel vom aeroben zu semiaeroben bzw. vom respirativen zum photosynthetischen Wachstum zuständig sind. Andere Umweltreize werden durch diese beiden Regulatoren nur gering oder, insbesondere unter anaeroben Bedingungen, gar nicht erfaßt. Da FnrL und PrrBA also (fast) „reine“ Sauerstoffsignale sind, regulieren sie auch andere Gene, die nicht unmittelbar mit der ETK in Zusammenhang stehen, deren Expression aber ebenfalls vom Sauerstoff abhängig ist. So reguliert FnrL z.B. anaerobe Respiration auf verschiedenen Elektronenakzeptoren [175] und aktiviert PrrA-P (bzw. RegA-P) die Stickstoff- und Kohlendioxidfixierung [66]. Daß mit FnrL und PrrBA zwei Signalwege zur sauerstoffabhängigen Regulation existieren, mag der Option gewidmet sein, beim Übergang vom aeroben zum anaeroben Stoffwechsel stufenweise auf verschiedene Sauerstoffkonzentrationen anzusprechen. So kann semiaerobes Wachstum als separater Operationsmodus etabliert werden, wo z.B. die Expression der Ubichinon- und cbb_2 -Oxidase am höchsten ist.

Ein stärker integrierendes (Redox-) Signal, das nun u.a. auch auf Lichtintensität reagiert, wird dagegen vom AppA/PpsR-Weg verarbeitet. PpsR reprimiert die PS-Gene unter oxidativen, und damit insbesondere aeroben Bedingungen, kann aber außerdem die Expression photosynthetischer Komponenten auch der Lichtintensität anpassen. Aufgrund der starken Integration des verarbeiteten Signals kann es passieren, daß andere Einflüsse, wie z.B. Reduktionsgrad des Substrates, dieses Signal (evtl. ungewollt) modulieren.

10 Zusammenfassung / Summary

Die beiden ineinandergreifenden thematischen Schwerpunkte der vorliegenden Arbeit sind

- (i) die Ausarbeitung neuer Methoden für die metabolische Netzwerk- und Pathwayanalyse und der Nachweis ihrer praktischen Anwendbarkeit in realen Stoffwechselnetzen.
- (ii) die Modellierung von Redoxprozessen in schwefelfreien Purpurbakterien. Dies umfasst den Redoxstoffwechsel (unter Heranziehung der Methoden von (i)), sowie die von Redoxsignalen gesteuerte Expression photosynthetischer Gene.

Das Ziel metabolischer Strukturanalysen ist es, Zusammenhänge zwischen Struktur, Funktion und Regulation in komplexen Stoffwechselnetzen herzustellen. Mit den Elementarmoden als fundamentales Werkzeug kann ein Großteil der relevanten Struktureigenschaften eines stöchiometrischen Netzwerkes extrahiert werden. Drei neue Analysemethoden wurden hier auf Basis dieser unzerlegbaren Stoffwechselrouten entwickelt. Bei der *kombinatorischen Pathwayanalyse* wird mittels statistischer Untersuchungen der EMen sowohl die Flexibilität des Netzes, als auch die relative Bedeutung einzelner Reaktionen für verschiedene Umweltkonstellationen abgeschätzt. Dabei können nicht nur die für ein Szenario essentiellen Netzelemente detektiert, sondern auch graduelle Bewertungen vorgenommen werden. Mit den *control-effective fluxes* wurde ein plausibles quantitatives Maß vorgeschlagen, das eine differentielle Abschätzung der Bedeutung einer Reaktion bzgl. zweier verschiedener Szenarios ermöglicht und dabei die Kriterien Effizienz und Flexibilität berücksichtigt. Ein Vergleich der rein strukturell ableitbaren CEFs mit differentiellen Expressionsdaten von Enzymen des Zentralstoffwechsels in *E. coli* zeigte eine gute Korrelation, was zumindest semi-quantitative Prädiktionen ermöglicht und den Zusammenhang zwischen der Struktur eines Netzes und seiner Regulation unterstreicht. In ähnlicher Art und Weise kann die Bestimmung *struktureller Kopplungen* bei der Interpretation von Genexpressionsdaten behilflich sein. Eine (strukturelle) Modularisierung eines Reaktionsnetzes auf Basis der strukturellen Kopplungen wäre ebenfalls denkbar, was aber in Zukunft noch gezeigt werden müßte.

Mit den *minimalen Schnittmengen* wurde ein neuer Ansatz eingeführt, der elementare Fehlermoden in einem Stoffwechselnetz beschreibt. Damit bilden sie das Pendant zu den kleinsten funktionalen Einheiten darstellenden Elementarmoden, obgleich sie aus letzteren berechnet werden können. Die minimalen Schnittmengen sind sehr gut geeignet, um die Fragilität (bzw. Robustheit) einer Netzstruktur zu untersuchen. Mit dem hier vorgeschlagenen Fragilitätskoeffizienten können kritische Netzelemente, deren Ausfall besonders schwer kompensierbar ist, identifiziert werden, außerdem ist eine quantitative Bewertung der gesamten Fragilität des Netzes möglich. Dabei stellte sich heraus, daß strukturelle Robustheit zwar häufig, aber nicht immer mit der Flexibilität, also der Anzahl an Realisationsmöglichkeiten, korreliert ist.

Minimale Schnittmengen sind ideale Wegweiser für geplante Struktureingriffe, bei denen gezielt Funktionen im Stoffwechsel unterdrückt werden sollen. Damit ist reichlich Potential für Anwendungen im Metabolic Engineering oder in der Medikamentenforschung gegeben.

Minimale Schnittmengen entfernen auf minimale Weise sämtliche Freiheitsgrade in einem Reaktionsnetzwerk. Aufgrund dieser Eigenschaft lassen sie sich auch für eine ganz andere Applikation einsetzen, nämlich zur Bestimmung minimaler Meßkombinationen, die eine bestimmte stationäre Reaktionsrate berechenbar machen. Das theoretische Fundament dazu wurde im Kap. 5 hergeleitet. Zunächst wurde gezeigt, daß trotz globaler Unterbestimmtheit eines Szenarios einige Raten dennoch (eindeutig) berechenbar sein können, was sich bei den *Rhodospirillaceae* auch am praktischen Beispiel demonstrieren läßt (s.u.). Ein allgemeines Kriterium zur Überprüfung der Berechenbarkeit einer unbekannt Rate wurde hergeleitet. Darauf aufbauend wurde mit der kombinatorischen (elementaren) Repräsentation des Nullraums der stöchiometrischen Matrix, die wiederum auf dem Konzept der Elementarmoden beruht, eine Darstellung gefunden, aus der unmittelbar Meßkombinationen zur Berechnung unbekannter Flüsse abgeleitet werden können. Der Bezug zu den minimalen Schnittmengen, und damit zu einem systematischen Berechnungsschema *minimaler* Meßkombinationen, wurde hergeleitet. Derartige Betrachtungen sind nützlich bei der Planung von Experimenten und wurden in metabolen Flußanalysen in dieser Art bisher nicht durchgeführt.

Bei den oben beschriebenen Anwendungen zeigte sich immer wieder, daß es vor allem die ausgezeichneten *kombinatorischen* Eigenschaften sind, die den Elementarmoden eine Schlüsselstellung in der Netzstrukturanalyse verleihen. Die Elementarmoden in ihrer Gesamtheit weisen einen enormen Gehalt an Information auf: sie repräsentieren sämtliche irreduziblen funktionalen Teilnetze eines Stoffwechselnetzes. Die damit verbundene kombinatorische Komplexität bei der Berechnung der Elementarmoden fordert ihren Preis und läßt es – zumindest derzeit – noch nicht zu, sie in genom-skalierten Stoffwechselnetzen zu bestimmen. Eine weitere Optimierung oder Parallelisierung des Berechnungsalgorithmus bleibt daher eine Aufgabe für die Zukunft.

Mit dem *FluxAnalyzer* wurde eine graphische Benutzeroberfläche geschaffen, in der sowohl die hier neu entwickelten, als auch viele andere Methoden der Netzstrukturanalyse integriert wurden. Damit steht nun eine benutzerfreundliche Software zur Verfügung, die in einer Umgebung vielfältige topologische Studien in metabolen Netzen ermöglicht.

Die hier vorgestellten Konzepte wurden spezifisch für Stoffwechselnetze hergeleitet. Da es sich bei den Elementarmoden aber letztlich um Vektoren handelt, die aus einer (hier stöchiometrischen) Matrix abgeleitet werden, lassen sich viele der Anwendungen verallgemeinern und insbesondere auf beliebige lineare Gleichungssysteme, bei Bedarf auch mit Vorzeichenrestriktionen in den Variablen, übertragen. Die Suche nach geeigneten Meßkombinationen von Variablen, aus denen andere eindeutig abgeleitet werden können, wäre ein Beispiel.

Unter Anwendung der vorgestellten Konzepte konnten eine Reihe wichtiger Aspekte für den Redoxstoffwechsel der *Rhodospirillaceae* abgeleitet werden. Eine Analyse des aufgestellten Stoffwechselnetzes zeigte auf, welche globalen Strukturveränderungen im Zentralstoffwechsel beim Übergang vom aeroben zum (anaeroben) phototrophen Metabolismus notwendig sind, um den Redoxstoffwechsel unter den neuen Bedingungen ausgewogen gestalten zu können. Dabei ergeben sich gewisse Strukturzwänge, die intuitiv nicht mehr faßbar sind. So konnte in Übereinstimmung mit Experimenten nachgewiesen werden, daß der CO₂-fixierende Calvin-Zyklus bei photoheterotrophen Stoffwechsel, trotz einer Nettofreisetzung von CO₂, essentiell für die Balance von Redoxäquivalenten ist. Die Studien zeigten auch, warum das Verhältnis zwischen NAD(H)- und NADP(H)-produzierenden Reaktionsflüssen von besonderer Bedeutung ist. In diesem Zusammenhang konnte die Hypothese abgeleitet werden, daß die Transhydrogenase unter phototrophen Bedingungen möglicherweise eine besonders wichtige Rolle spielt, während dem oxidativen Pentosephosphat-Weg im phototrophen Szenario quantitativ eine geringe Bedeutung zugewiesen wurde.

Eine Auswertung der Elementarmoden der *Rhodospirillaceae* zeigte, daß der hypothetische Citramalat-Zyklus – wie früher prognostiziert – Wachstum auf dem Substrat Acetat gewährleisten kann. Unerwarteter Weise konnten auch alternative Verwertungsmöglichkeiten von Acetat gefunden werden, die aber als relativ unwahrscheinlich eingestuft werden können.

Für den Redoxstoffwechsel der *Rhodospirillaceae* sind die Vorgänge in der Elektronentransportkette (ETK) von besonderem Interesse. Die ETK ist zentraler Kontaktpunkt für Elektronenströme, sie erzeugt protonenmotorische Kraft und mit ihr Energie in Form von ATP. Sie unterliegt beim Wechsel der Umgebungsbedingungen, insbesondere bzgl. Sauerstoffkonzentration und Lichtintensität, fundamentalen Veränderungen in ihrer Zusammensetzung (z.B. Einbau von Photosystemen), die z.T. über direkt in der ETK erfaßte Redoxsignale ausgelöst werden. Wollen wir sowohl die metabolische als auch die regulatorische Funktion der ETK studieren, reichen die auf Quasistationarität basierenden Strukturanalysen nicht mehr aus, da nun Konzentrationen explizit berücksichtigt werden müssen. Deshalb wurde ein dynamisches Modell für die ETK der *Rhodospirillaceae* erstellt, das die Prozesse in kinetischer Abhängigkeit von Redoxpotentialunterschieden beteiligter Komponenten beschreibt. Dieses zunächst rein metabolische Modell kann in der Tat, in Abhängigkeit eingestellter Parameter und Umgebungsbedingungen, die verschiedenen Operationsweisen der ETK nachbilden, die auch über die Elementarmoden identifiziert wurden. Durch die zusätzliche Berücksichtigung der drei wichtigsten Signalwege AppA/PpsR, PrrA/PrrB und FnrL wurde auch die Regulation der Expression von (insbesondere photosynthetischen) ETK-Komponenten mit im Modell integriert. Als ein wesentliches Ergebnis konnte gezeigt werden, daß die z.T. hypothetischen Vorstellungen bzgl. des Abgreifens der Redoxsignale zumindest qualitativ mit experimentellen Messungen von exprimierten ETK-Komponenten vereinbar sind.

Eine besonders wichtige Schlußfolgerung konnte auch für den Redoxzustand des Ubichinons, der vermutlich über den AppA/PpsR-Signalweg erfaßt wird, gezogen werden. Ubichinon liegt

in der Reihenfolge aerob, semiaerob, anaerob im Starklicht und anaerob im Schwachlicht zunehmend reduzierter vor, was bisher vermutet, aber experimentell noch nicht belegt werden konnte. Damit reagiert dieses Signal sowohl auf Sauerstoff als auch auf die Lichtintensität, weshalb es in der Tat ideal für die Regulation photosynthetischer Gene geeignet ist. Da Ubichinon ein zentraler Knotenpunkt in der ETK ist, wird sein Redoxzustand auch von vielen anderen Größen beeinflusst. Daher können Nebeneffekte im erfaßten Signal enthalten sein, welche die eigentlich primär zu erfassende Information u.U. modulieren können (hier z.B. verursacht durch bestimmte Substratkombinationen).

Wie das Beispiel der ETK der *Rhodospirillaceae* zeigt, bleiben dynamische Modelle zum Studium des Verhaltens biochemischer Netzwerke unverzichtbar. Dennoch zeigt sich immer häufiger, daß die Topologie eines zellulären Netzwerks bereits entscheidend dessen Funktion und Verhalten determiniert und nachgewiesene Eigenschaften, wie z.B. Robustheit, bestimmte Anforderungen im Design der Netze implizieren. Deshalb bleibt es auch in der Zukunft eine wichtige Aufgabe, die Zusammenhänge zwischen Struktur, Funktion und Regulation solcher molekularen Netzwerke in der Zelle zu erkennen und sie mathematisch formal zu beschreiben. Für Stoffwechselnetze stehen dafür mittlerweile viele Werkzeuge zur Verfügung. Noch weitgehend in den Anfängen sind dagegen z.B. Strukturanalysen in Signaltransduktionsnetzen. Hier muß geklärt werden, ob stöchiometrische Analysen zumindest teilweise übertragbar sind, oder ob andere Methoden eine – parameterunabhängige – Bestimmung von Schlüsseleigenschaften dieser Netze aus ihrer Struktur gewährleisten können.

Summary

Introduction and Motivation

Structural analysis of biochemical networks yields deep insights into topology, functionality and capabilities of biological networks and has become an important approach for systems biology. Especially metabolic networks have been examined extensively, as their structure is often well-known. One main goal of this thesis is the elaboration of new methods for metabolic network and pathway analysis and to prove their applicability in realistic networks. The central metabolism of *E. coli* as well as the metabolic network of the purple nonsulfur bacteria (PNB; *Rhodospirillaceae*) serve to demonstrate the novel approaches in practice. Compared to *E. coli*, the metabolism of the facultatively photosynthetic PNB is much less understood. Much attention has been drawn to the question, how the PNB manage to balance their redox metabolism under different environmental conditions, especially during photoheterotrophic growth. A stoichiometric analysis of their metabolic network can lead to a deeper understanding of the necessary (re)routing of metabolic fluxes. A related and also not yet fully understood process in the PNB is the reorganization of the electron transport chain when switching from aerobic growth (in the dark) to phototrophic growth (anaerobically) in light. It is assumed that redox signals govern the gene regulation, in particular the synthesis of photosynthetic complexes. To study these metabolic and regulatory processes in the electron transport chain of the PNB, a dynamic model is set-up and analyzed.

Results

The concept of *elementary (flux) modes*, which are smallest functional sub-networks, has become a fundamental framework for studying numerous structural and functional properties in metabolic networks. Here, three new analytic methods relying on elementary modes are introduced: (1) combinatorial pathway analysis, (2) minimal cut sets and (3) generalized observability analysis of stationary metabolic fluxes.

A combinatorial (or statistical) analysis of elementary modes enables to assess the flexibility of the network as well as the relative importance of a single metabolic reaction for a given environmental scenario (e.g. with respect to the availability of substrates, oxygen, light and so on). Essential reactions, which occur in all elementary modes, can be identified allowing phenotype predictions. Phenotypes of *E. coli* predicted by elementary-modes analysis agree well with experimental results of *E. coli* mutants. Even if the relative frequency, by which a reaction is utilized in the set of elementary modes, is not 100%, it provides a simple yet useful measure of the structural importance of a reaction for the respective scenario. This kind of analysis is extended by introducing *control effective fluxes* (CEFs). These quantitative measurements of "reaction importance" consider also the absolute flux values of the reactions as well as the efficiencies of the modes. Transcript ratios of metabolic genes in *E. coli* obtained from growth experiments on two different substrates are compared with CEF ratios yielding a significant correlation. Since CEFs are derived completely from the structure, these results

emphasize that the metabolic network structure has profound implications on its own regulation.

In a similar way, *structural couplings* can be determined by elementary-mode analysis which enable to quantify the dependence of one reaction from the operation of the others. It is demonstrated that structural couplings might give clues to how the network is regulated and that they can be valuable for interpreting gene expression data.

Elementary modes are minimal functional units. For describing also minimal *failure modes* the concept of minimal cut sets (MCSs) is introduced. A minimal cut set (MCS) is a minimal (irreducible) set of reactions in the network whose inactivation will definitely lead to a failure in certain network functions. A (structural) failure means here that a continuous operation of one or several objective reactions is not possible anymore. An algorithm for computing the MCSs with respect to user-defined objective reactions is devised. Interestingly, the MCSs can be computed from their counterparts, the elementary modes.

A number of potential applications of MCSs are outlined. First, the set of MCSs are ideally suited for assessing the fragility (or inverse: robustness) of a metabolic network. *Specific fragility coefficients* can be computed from the MCSs which allow for an identification of single network elements which are crucial for the operation of the objective reaction(s). A (*global*) *network fragility coefficient* is also introduced quantifying the overall fragility of the network. A fragility analysis of the MCSs in the *E. coli* network model reveals that the specific as well as the overall fragility depend strongly on the chosen substrate, hence, on the environmental conditions.

An even more important practical application of MCSs is the search for optimal cuts in the network (reaction or enzyme or gene deletions) which repress certain metabolic functions. Hence, MCSs may be helpful for target identification in drug discovery or for metabolic engineering.

A network failure is guaranteed when an MCS is removed from the network. Such a prediction can be used to falsify a given network structure. MCSs are ideal candidates for network falsification experiments since they contain by definition only a minimal set of necessary cuts.

MCSs remove in a minimal way all degrees of freedom in a reaction network. Therefore, MCSs can also represent *minimal measurement sets* rendering a certain (unknown) reaction rate observable (calculable). This fact can be valuable for metabolic flux analysis.

For this purpose, a generalized framework for studying the observability of stationary fluxes in reaction networks is devised. First, a general criterion for detecting the observable reaction rates is given. This criteria relies on an investigation of the null space of the stoichiometric matrix related to the reactions with unknown rate. An analysis of the metabolic network model of the *Rhodospirillaceae* demonstrates that also underdetermined systems may have observable parts implying important constraints for the network overall behavior (see below).

For a systematic search of suitable measurement sets, the *combinatorial representation of the null space* is introduced. This representation comprises all elementary modes occurring in the (hypothetically) completely reversible network and is of fundamental importance for a generalized observability/calculability analysis of stationary reaction rates. In particular, it allows one to construct those sets of measurable rates that enable a calculation of a certain unknown rate. It can then be shown that the determination of non-redundant (minimal) measurements sets is equivalent to the computation of minimal cut sets with respect to the combinatorial null space representation. Besides, rates which are never calculable by the metabolite balancing equation may be easily detected by this method.

The set of elementary modes comprises all irreducible functional sub-networks and provides therefore a combinatorial representation of a metabolic network. The combinatorial properties of elementary modes are fundamental for many theoretical network studies utilizing elementary modes (including those given here). However, for the same reason, the determination of elementary modes becomes highly complex in larger networks. Here, an upper boundary for the maximal number of elementary modes is provided. In realistic networks it turns out that these upper boundaries are much larger than the real number of modes. The main reasons for this fact are the relatively high proportion of irreversible reactions, the occurrence of short pathways, the weak connectivity of metabolites and certain structural couplings like enzyme subsets. Nevertheless, the number of modes in larger networks is considerably large and efficient algorithms for the computation of pathways are required.

The newly devised algorithms as well as already existing methods for structural network analysis were integrated in a new software tool, the *FluxAnalyzer*. The *FluxAnalyzer* is a package for MATLAB® and facilitates structural studies in metabolic networks within a graphical user interface. Arbitrary metabolic network models can be composed and then linked with network graphics leading to *interactive flux maps*. These maps allow for user input and display of calculation results within a network visualization. A large and powerful collection of tools and algorithms for structural network analysis can be applied interactively. This toolbox enables to perform metabolic flux analysis, flux optimization, computation of basic topological and graph-theoretical features and, in particular, pathway analysis by elementary modes including the new applications introduced in this work.

By using the generalized observability analysis and combinatorial pathway analysis new insights into the redox metabolism of the *Rhodospirillaceae* can be obtained. A challenging task for these bacteria is the switch from respirative (aerobic) to photosynthetic (anaerobic) growth where the synthesis of reducing equivalents in the central metabolism must be greatly diminished. In some cases, electron sinks are required additionally for maintaining redox poise. This leads to inherent constraints posed by the metabolic network structure which are non-intuitive and require therefore stoichiometric network analysis. Here, in agreement with experimental results, the network model predicts that the Calvin cycle, which fixes CO₂ and

consumes NADH, is essential for photoheterotrophic growth on diverse substrates. This is even the case if a considerable CO₂ net release occurs. The model also predicts correctly the CO₂ net release as well as the biomass yield for phototrophic growth on different substrates. Remarkably, these values are independent of the ATP overall requirement in the cell, as could be shown by observability analysis. This is not the case for aerobic growth where redox and energy metabolism are strongly coupled stoichiometrically. The studies performed here also suggest that the ratio between NADH and NADPH production is of great importance. Related to this aspect, stoichiometric network analysis reveals that the transhydrogenase might play a particular role during phototrophic growth (as confirmed in experiments), whereas the oxidative pentose phosphate pathway is presumably of low importance under these conditions.

Elementary-modes analysis also shows that the hypothetical citramalate cycle of the *Rhodospirillaceae* can indeed provide a way for growth on acetate and thus compensate the missing glyoxylate shunt. Unexpectedly, even alternative pathways for metabolizing acetate can be found. However, due to low efficiency, it is unlikely that these routes are utilized in reality. Another surprising result is that during phototrophic growth on some substrates only one elementary mode is available for growth-independent ATP production (needed for maintenance processes).

The processes in the electron transport chain (ETC) are of special importance for the redox metabolism since the ETC is a central market place of reducing equivalents. A fundamental reorganization of the ETC in the membrane takes place when environmental conditions changes, in particular oxygen tension and, to a smaller extent, light intensity. This reorganization includes redistribution of fluxes as well as synthesis and incorporation of new components, like photosynthetic complexes or oxidases. The regulation of the reconfiguration of the ETC is mediated by redox signals, which are at least partially generated in the ETC itself. Since the concentrations (or ratios thereof) of certain ETC components determine these redox signals, the assumption of pseudo steady-state as used for structural network analyses is not sufficient. Therefore, for studying metabolic as well as regulatory functions of the ETC in *Rhodospirillaceae*, a dynamic model was set-up. The kinetic description of the metabolic flows in the ETC, which are predominately electron exchange (i.e. redox) reactions, relies on the driving force built-up by redox potential differences of participating species. The model also accounts for the proton-motive force, which is generated during these processes and required for the synthesis of ATP. The metabolic part of this model can drive the (8) fundamental operational modes of the ETC, as can be seen by the corresponding elementary modes. Furthermore, the three most important regulators controlling the synthesis of ETC components in *Rhodospirillaceae* are taken into account in the model: FnrL, AppA/PpsR and the two-component system PrrB/PrrA. Performing simulations with this model several conclusions can be drawn. First, by incorporating the three main regulatory pathways in the model, experimental measurements of the expression of (in particular photosynthetic) ETC components can be reproduced, at least qualitatively.

Second, several important conclusions concern the redox state of the ubiquinon pool, which can, so far, not be measured experimentally *in vivo*. The model predicts that ubiquinone is oxidized at most under respirative conditions, it gets more reduced under semi-aerobic and anaerobic high-light conditions and is most reduced when exposed anaerobically to low-light. This result suggests that the response curve of the redox state of ubiquinone, which is probably triggered via the AppA/PpsR system, seems to be a suitable signal for controlling photosynthesis gene expression. Whereas PrrBA and FnrL respond predominately to the oxygen tension, AppA/PpsR trigger the oxygen signal *and* the light intensity. Since the electron carrier ubiquinon is a central node in the ETC, its redox state reflects the global redox state. For this reason, such a strongly integrating signal might be modulated up to a certain extent by secondary factors, here, for example, by the available substrate.

Using a transport system at the same time as a signal generator seems to be an efficient and effective strategy and is also known from the glucose phosphotransferase system (PTS) in *E. coli*.

The model of the ETC of the *Rhodospirillaceae* shows that dynamic models of cellular systems are the ultimate goal in systems biology because they provide the complete systems behavior. However, it is more and more accepted that the topologies of cellular networks determine key aspects of their functions and behavior. Methods for structural network studies are therefore gaining in importance, also due to the fact, that the network structure is usually much easier accessible than a full parametrization of its kinetic model. For metabolic network analysis, a large corpus of methods is now available. The future will show whether similar successful approaches can also be developed for signaling or genetic networks.

Anhang

Anhang A: Stoffwechselnetze *Escherichia coli* und *Rhodospirillaceae*

A.1 Symbole und Namen der internen Metabolite

In fet geschrieben: Makromoleküle.

* nur im Netz von *Escherichia coli* betrachtet; † nur im Netz der *Rhodospirillaceae* betrachtet

Abk.	Metabolit / Komponente	Abk.	Metabolit / Komponente
G6P	Glucose-6-phosphat	Ala	Alanin
F6P	Fructose-6-phosphat	Cys	Cystein
F16P	Fructose-1,6-bisphosphat	Asp	Aspartat
DHAP	Dihydroxyacetonphosphat	Glu	Glutamat
Glyc3P	Glycerin-3-phosphat	Phe	Phenylalanin
G3P	Glyceraldehyd-3-phosphat	Gly	Glycin
DPG	1,3-Bisphosphoglycerat	His	Histidin
3PG	3-Phosphoglycerat	Ile	Isoleucin
2PG	2-Phosphoglycerat	Lys	Lysin
PEP	Phosphoenolpyruvat	Leu	Leucin
Pyr	Pyruvat	Met	Methionin
AcCoA	Acetyl-CoA	Asn	Asparagin
Cit	Citrat	Pro	Prolin
ICit	Isocitrat	Gln	Glutamin
αKG	α-Ketoglutarat (= 2-Oxoglutarat)	Arg	Arginin
SuccCoA	Succinyl-CoA	Ser	Serin
Succ	Succinat	Thr	Threonin
Fum	Fumarat	Val	Valin
Mal	Malat	Trp	Tryptophan
OxA	Oxalacetat	Tyr	Tyrosin
Glyox	Glyoxylat	rATP	ATP für RNA-Synthese
R5P	Ribose-5-phosphat	rGTP	GTP für RNA-Synthese
R15P	Ribulose-5-phosphat	rCTP	CTP für RNA-Synthese
E4P	Erythrose-4-phosphat	rUTP	UTP für RNA-Synthese
X5P	Xylulose-5-phosphat	dATP	ATP für DNA-Synthese
S7P	Sedoheptulose-7-phosphat	dGTP	GTP für DNA-Synthese
PGluc	6-Phosphogluconolacton	dCTP	CTP für DNA-Synthese
PGluc	6-Phosphogluconate	dTTP	TTP für DNA-Synthese
ATP	ATP (Adenosintriphosphat)	mit_FS	"Mittlere" Fettsäure (bzgl. Länge und Sättigung)
NADH	NADH	UDPGlc	UDP-Glucose
NADPH	NADPH	CDPEth	CDP-Ethanolamin
QH2	reduziertes Ubichinon; vereinfachend ist diesem Pool auch reduziertes Menachinon (<i>E. coli</i>) bzw. Rhodochinon (<i>R. rubrum</i>) zugeordnet	OH_myr_ac	β-OH-Fettsäuren (in LPS)
Hp	Protonen im Periplasma	C14_0_FS	C(14:0) Fettsäuren (in LPS)
O2	Sauerstoff	CMP_KDO	CMP-2-keto-3-desoxyoctonsäure
CO2	Kohlendioxid (gelöst als Bicarbonat)	NDPHep	NDP-Heptose

Abk.	Metabolit / Komponente	Abk.	Metabolit / Komponente
N	Ammonium	TDPGlc	TDP-Glucosamin
S	Sulfid	UDP_NAG	UDP-N-Acetylglucosamin
Glyc	Glycerin	UDP_NAM	UDP-N-Acetylmuraminsäure
AcP	Acetylphosphat	di_am_pim	Diaminopimelinsäure
Ac	Acetat	ADPGlc	Glycogen-Monomer
Fd_red	reduzierendes Ferredoxin *	CitMal	Citramalat *
S17P	Seduheptulose-1,7-bisphosphat *	Mes	Mesaconat *
R115P	Ribulose-1,5-bisphosphat *	Prop	Propinat *
c2_red	reduzierendes Cytochrom c2 (und c') *	PropCoA	Propionat-CoA *
Form	Formiat		Methyl-Malonyl-CoA (Pool für beide Formen: S (=D) und R (=L)) *
Lac	Lactat *	MMalCoA	
Adh	Acetaldehyd *	Protein	Protein
Eth	Ethanol *	RNA	RNA
Chor	Chorismat	DNA	DNA
PRPP	5-Phosphoribosyl-1-pyrophosphat	PepGly	Peptidoglycan (Murein)
MTHF	5-10-Methylen tetrahydrofolat	LPS	Lipopolysaccharide
AspSAld	Aspartate-4-semialdehyd	Glycogen	Glycogen
HSe	Homoserin	Lip	Lipide
bchl	Bacteriochlorophyll a *	PolyP	Polyphosphat *
hydry_but	β -Hydroxybuttersäure-CoA *	PHB	Polyhydroxybutyrat *
		Bchl	Bacteriochlorophyll *

A.2. *Escherichia coli*: Reaktionsgleichungen (berücksichtigen nur die betrachteten internen Metabolite; z. T. kumulierte Stöchiometrien)

Reaktion	Stöchiometrie / Reversibilität	Gene	Enzym / Reaktionsbeschreibung
Substrataufnahme:			
Glc_PTS_up	1 PEP → 1 G6P + 1 Pyr	psg	Glucoseaufnahme über Glucose-PTS
Glc_ATP_up	1 ATP → 1 G6P	gik	Glucoseaufnahme + Phosphorylierung mit Glucokinase
Succ_up	→ 1 Succ		Succinatenaufnahme
Glyc_up	→ 1 Glyc		Glycerinaufnahme
O2_up	→ 1 O2		Sauerstoffaufnahme
N_up	→ 1 N		Ammoniumaufnahme
S_up	4 ATP + 4 NADPH → 1 S		Sulfataufnahme und assimilarische Reduktion zu Sulfid
Ac_up	→ 1 Ac		Acetatnaufnahme
Zentralstoffwechsel:			
G6P→F6P	1 G6P ↔ 1 F6P	pgi	Glucosephosphat-Isomerase
F16P→F6P	1 F16P → 1 F6P	fbp	Fructosebiphosphatase
F6P→F16P	1 F6P + 1 ATP → 1 F16P	pfkA	6-Phosphofruktokinase
F16P→F3P	1 F16P ↔ 1 DHAP + 1 G3P	fbx	Fructosebiphosphat-Aldolase
DHAP→G3P	1 DHAP ↔ 1 G3P	tpiA	Triosephosphat-Isomerase
G3P→DHG	1 G3P ↔ 1 DHG + 1 NADH	gapA	Glyceraldehyd-3-phosphat-Dehydrogenase
DHG→3PG	1 DHG ↔ 1 3PG + 1 ATP	pgk	Phosphoglycerat-Kinase
3PG→2PG	1 3PG ↔ 1 2PG	gpmA	Phosphoglyceratmutase
2PG→PEP	1 2PG ↔ 1 PEP	eno	Enolase
PEP→PYR	1 PEP → 1 Pyr + 1 ATP	pykA, pykF	Pyruvat-Kinase I / II
Pyr→PEP	1 Pyr + 2 ATP → 1 PEP	ppsA	PEP-Synthetase; (Verbrauch von 2 ATP eingesetzt, da nicht ADP sondern AMP entsteht)

Reaktion	Stöchiometrie / Reversibilität	Gene	Enzym / Reaktionsbeschreibung
Pyr::AcCoA	1 Pyr → 1 AcCoA + 1 NADH + 1 CO ₂	aceEF	Pyruvat-Dehydrogenase
AcCoA::Cit	1 AcCoA + 1 Oxa → 1 Cit	glkA	Citrat-Synthase
Cit::ICit	1 Cit ↔ 1 ICit	acnB	Acetat-Hydratase
ICit::alKG	1 ICit ↔ 1 alKG + 1 NADPH + 1 CO ₂	icdA	Isocitrat-Dehydrogenase (NADP-spezifisch)
alKG::SuccCoA	1 alKG → 1 SuccCoA + 1 NADH + 1 CO ₂	sucAB	Oxoglutarat-Dehydrogenase
SuccCoA::Succ	1 SuccCoA ↔ 1 Succ + 1 ATP	sucCD	Succinyl-CoA-Synthetase
Succ::Fum	1 Succ → 1 Fum + 1 QH ₂	sdhC	Succinat-Dehydrogenase
Fum::Succ	1 Fum + 1 QH ₂ → 1 Succ	frdA	Fumarat-Reduktase; Elektronendonator: Membrin oder (als umgekehrte Succinat-Dehydrogenase) Ubichinon
Fum::Mal	1 Fum ↔ 1 Mal	fumA, fumB, fumC	Fumarase A (aerobes Isoenzym) bzw. B und C (anaerob)
Mal::OxA	1 Mal ↔ 1 OxA + 1 NADH	mdh	Malat-Dehydrogenase
ICit::Glyox	1 ICit → 1 Succ + 1 Glyox	aceA	Isocitrat-Lyase
Glyox::Mal	1 AcCoA + 1 Glyox → 1 Mal	glcB	Malat-Synthase
G6P::PGLac	1 G6P ↔ 1 PGLac + 1 NADPH	zwf	Glucose-6-phosphat-Dehydrogenase
PGLac::PGLac	1 PGLac → 1 PGLac	pgl	6-Phosphogluconatase
PGLac::R5P	1 PGLac → 1 R5P + 1 NADPH + 1 CO ₂	gud	Gluconat-6-phosphat-Dehydrogenase, decarboxylierend
R5P::X5P	1 R5P ↔ 1 X5P	rpe	Ribulose-5-phosphat-3-Epimerase
R5P::R5P	1 R5P ↔ 1 R5P	rpiA, rpiB	Ribosephosphat-Isomerase
Transket1	1 R5P + 1 X5P ↔ 1 G6P + 1 5TP	tkxA, tkxB	Transketolase (Isoenzyme 1/2)
Transaldol	1 G3P + 1 5TP ↔ 1 F6P + 1 E4P	talA, talB	Transaldolase B
Transket2	1 E4P + 1 X5P ↔ 1 F6P + 1 G3P	tkxA, tkxB	Transketolase (Isoenzyme 1/2)
OxA::PEP	1 OxA + 1 ATP → 1 PEP + 1 CO ₂	peckA	Phosphoenolpyruvat-Carboxylase
PEP::OxA	1 PEP + 1 CO ₂ → 1 OxA	ppc	Phosphoenolpyruvat-Carboxylase
AcCoA::Adh	1 AcCoA + 1 NADH ↔ 1 Adh	adhE	Acetaldehyd-Dehydrogenase
Adh::Eth	1 Adh + 1 NADH ↔ 1 Eth	adhE	Alkohol-Dehydrogenase
AcCoA::AcP	1 AcCoA ↔ 1 AcP	pta	Phosphotransacetylase

Reaktion	Stöchiometrie / Reversibilität	Gene	Enzym / Reaktionsbeschreibung
Ac::Ac	1 AcP ↔ 1 ATP + 1 Ac	acsA	Acetat-Kinase
Pyr::form	1 Pyr → 1 AcCoA + 1 Form	pf1B	Pyruvat-Formiat-Lyase
Pyr::Lac	1 Pyr + 1 NADH ↔ 1 Lac	ldhA	Lactat-Dehydrogenase
Glyc::Glyc3P	1 ATP + 1 Glyc → 1 Glyc3P	glpK	Glycerin-Kinase
Mal::Pyr	1 Mal → 1 Pyr + 1 NADH + 1 CO ₂	sdhA	Malat-Enzym (NAD-spezifisch)
Glyc3P::DHAP	1 Glyc3P → 1 DHAP + 1 QH2	glpD	ss-Glycerin-3-phosphat-Dehydrogenase (bei Wachstum auf Glycerin)
DHAP::Glyc3P	1 DHAP + 1 NADH → 1 Glyc3P	gssA	Glycerin-3-phosphat-Dehydrogenase (NAD-spezifisch)
Pyr::Ac	1 Pyr → 1 QH2 + 1 CO ₂ + 1 Ac	poxB	Pyruvat-Oxidase
Ac::AcCoA	2 ATP + 1 Ac → 1 AcCoA	acs	Acetyl-CoA-Synthetase (Verbrauch von 2 ATP, da AMP entsteht)
NADH:Dehydro Oxidase	1 NADH ↔ 1 QH2 + 2 Hp	ruo	NADH-Dehydrogenase I
TransHydro	1 QH2 + 0.5 O ₂ → 2 Hp	cyd	Cytochrom- <i>b</i> -Oxidase
ATP:Synth	1 NADH + 1 Hp ↔ 1 NADPH	ptaA	Transhydrogenase
ATP:atras	3 Hp ↔ 1 ATP	atp	ATP-Synthase (membrangebunden)
	1 ATP →		undefiniert ATP-Verbrauch (malmenase, ...)
Monomere:			
Chor_Synth	2 PEP + 1 E4P + 1 ATP + 1 NADPH → 1 Chor		Synthese von Chorimat
PRPP_Synth	1 R5P + 2 ATP → 1 PRPP		Synthese von 5-Phosphoribosyl-1-pyrophosphat
MTHF_Synth	1 ATP + 1 NADPH ↔ 1 MTHF		Benötigte Co-metabolite für Regeneration von MTHF aus THF (Tetrahydrofolat)
Ala_Synth	1 Pyr + 1 Glu → 1 alKG + 1 Ala		Synthese von Alanin
Val_Synth	2 Pyr + 1 NADPH + 1 Glu → 1 alKG + 1 CO ₂ + 1 Val		Synthese von Valin
Leu_Synth	2 Pyr + 1 AcCoA + 1 NADPH + 1 Gln → 1 alKG + 1 NADH + 2 CO ₂ + 1 Leu		Synthese von Leucin

Reaktion	Stöchiometrie / Reversibilität	Gene	Enzym / Reaktionsbeschreibung
Asa_Synth	2 ATP + 1 N + 1 Asp → 1 Asn	asaA	Asparagin-Synthase A
Asp_Synth	1 OAc + 1 Glu → 1 aKG + 1 Asp	aspC	Aspartat-Transaminase
Asp_Fum	1 Asp → 1 Fum + 1 N	aspA	Aspartase
Asp_AspSAld	1 ATP + 1 NADPH + 1 Asp → 1 AspSAld		Synthese von Aspartat-4-semialdehyd
AspSAld_HSer	1 NADPH + 1 AspSAld → 1 HSer		Synthese von Homoserin
Lys_Synth	1 dl_am_pim → 1 CO2 + 1 Lys		Synthese von Lysin
Met_Synth	1 SuccCoA + 1 MTHF + 1 HSer + 1 Cys → 1 Pyr + 1 Succ + 1 N + 1 Met		Synthese von Methionin
Thr_Synth	1 ATP + 1 HSer → 1 Thr		Synthese von Threonin
Ile_Synth	1 Pyr + 1 NADPH + 1 Glu + 1 Thr → 1 aKG + 1 CO2 + 1 N + 1 Ile		Synthese von Isoleucin
His_Synth	1 ATP + 1 PRPP + 1 Gln → 1 aKG + 2 NADH + 1 His		Synthese von Histidin
Glu_Synth	1 aKG + 1 NADPH + 1 N → 1 Glu	ghlA	Glutarat-Dehydrogenase (NADP-spezifisch)
Glu_Synth	1 ATP + 1 N + 1 Glu → 1 Glu	ghlA	Glutarat-Synthase
Pro_Synth	1 ATP + 2 NADPH + 1 Glu → 1 Pro		Synthese von Prolin
Arg_Synth	1 AcCoA + 4 ATP + 1 NADPH + 1 CO2 + 1 N + 1 Asp + 2 Glu → 1 aKG + 1 Fum + 1 Ac + 1 Arg		Synthese von Arginin
Trp_Synth	1 Chor + 1 PRPP + 1 Gln + 1 Ser → 1 GHP + 1 Pyr + 1 CO2 + 1 Glu + 1 Trp		Synthese von Tryptophan
Tyr_Synth	1 Chor + 1 Glu → 1 aKG + 1 NADH + 1 CO2 + 1 Tyr		Synthese von Tyrosin
Phe_Synth	1 Chor + 1 Glu → 1 aKG + 1 CO2 + 1 Phe		Synthese von Phenylalanin
Ser_Synth	1 3PG + 1 Glu → 1 aKG + 1 NADH + 1 Ser		Synthese von Serin
Gly_Synth	1 Ser → 1 MTHF + 1 Gly		Synthese von Glycin
Cys_Synth	1 AcCoA + 1 S + 1 Ser → 1 Ac + 1 Cys		Synthese von Cystein
rATP_Synth	5 ATP + 1 CO2 + 1 PRPP + 2 MTHF + 2 Asp + 1 Gly + 2 Gln → 2 Fum + 1 NADPH + 2 Glu + 1 rATP		Synthese von ATP für RNA
rGTP_Synth	6 ATP + 1 CO2 + 1 PRPP + 2 MTHF + 1 Asp + 1 Gly + 3 Gln → 2 Fum + 1 NADH + 1 NADPH + 3 Glu + 1 rGTP		Synthese von GTP für RNA

Reaktion	Stöchiometrie / Reversibilität	Gene	Enzym / Reaktionsbeschreibung
dCTP_Synth	1 ATP + 1 Gln + 1 dUTP → 1 Gln + 1 dCTP		Synthese von CTP für RNA
rUTP_Synth	4 ATP + 1 N + 1 PRPP + 1 Asp → 1 NADH + 1 rUTP		Synthese von UTP für RNA
dATP_Synth	1 NADPH + 1 dATP → 1 dATP		Synthese von ATP für DNA
dGTP_Synth	1 NADPH + 1 dGTP → 1 dGTP		Synthese von GTP für DNA
dCTP_Synth	1 NADPH + 1 dCTP → 1 dCTP		Synthese von CTP für DNA
dTTP_Synth	2 NADPH + 1 MTHF + 1 dUTP → 1 dTTP		Synthese von TTP für DNA
sal_FS_Synth	8.24 AcCoA + 7.24 ATP + 13.91 NADPH → 1 mit_FS		Synthese einer "militären" Fettsäure
UDPGLc_Synth	1 GAP + 1 ATP → 1 UDPGLc		Synthese von UDP-Glucose
CDPEib_Synth	1 PEP + 3 ATP + 1 NADPH + 1 N → 1 NADH + 1 CDPEib		Synthese von CDP-Ethanolamin
OH_myr_ac_Synth	7 AcCoA + 6 ATP + 11 NADPH → 1 OH_myr_ac		Synthese von β-OH-Fettsäuren (für LPS)
Cl4_0_FS_Synth	7 AcCoA + 6 ATP + 12 NADPH → 1 Cl4_0_FS		Synthese von C(14:0)-Fettsäuren (für LPS)
CMP_KDO_Synth	1 PEP + 1 R5P + 2 ATP → 1 CMP_KDO		Synthese von CMP-2-Keto-3-desoxyoctonsäure
NDFHep_Synth	1.5 Orp + 1 ATP → 4 NADPH + 1 NDFHep		Synthese von NDP-Heptose
TDPGLc_Synth	1 PEP + 2 ATP + 1 N → 1 TDPGLc		Synthese von TDP-Glucoseamin
UDP_NAG_Synth	1 PEP + 1 AcCoA + 1 ATP + 1 Gln → 1 Gln + 1 UDP_NAG		Synthese von UDP-N-acetylglucosamin
UDP_NAM_Synth	1 PEP + 1 NADPH + 1 UDP_NAG → 1 UDP_NAM		Synthese von UDP-N-acetylmuramin
dl_am_ptm_Synth	1 Pyr + 1 SuccCoA + 1 NADPH + 1 AspSald + 1 Gln → 1 alKG + 1 Succ + 1 dl_am_ptm		Synthese von Diaminoptimelinsäure
ADPGlc_Synth	1 GAP + 1 ATP → 1 ADPGlc		Synthese eines Glycogen-Monomers

Polymer-Synthese (mmol Metabolite / g Polymer) [impliziert in "Biomassensynthese-Reaktion" μ enthalten]:

Proc_Synth	39.9455 ATP + 0.88727 Ala + 0.15818 Cys + 0.41636 Asp + 0.45455 Gln + 0.32 Phe + 1.0582 Gly + 0.16364 His + 0.50182 Ile + 0.39273 Lys + 0.77818 Leu + 0.26546 Met + 0.41636 Asn + 0.38182 Pro + 0.45455 Gln + 0.51091 Arg + 0.37273 Ser + 0.43818 Thr + 0.73091 Val + 0.09818 Trp + 0.23818 Tyr → Proceth
RNA_Synth	1.2488 ATP + 0.80488 rATP + 0.99034 rGTP + 0.61436 rCTP +

Reaktion	Stöchiometrie / Reversibilität	Gen(e)	Enzym / Reaktionsbeschreibung
DNA_Synth	0.66341 rUTP → RNA 4.4129 ATP + 0.7968 dATP + 0.8194 dGTP + 0.8194 dCTP + 0.7968 dTTP → DNA		DNA-Synthase
Lip_Synth	1.4176 Glyc3P + 2.8352 ATP + 1.4176 Ser + 2.8352 mlc_FS → Lip		Lipid-Synthase
LPS_Synth	0.46176 UDPGlc + 0.69118 CDPEth + 0.69118 OH_myr_ac + 0.69118 C14_0_FS + 0.69118 CMP_KDO + 0.69118 NDPHep + 0.46176 TDPGlc → LPS		Lipopolysaccharid-Synthase
PepGly_Synth	5.52 ATP + 2.208 Ala + 1.104 Glu + 1.104 UDP_NAG + 1.104 UDP_NAM + 1.104 di_am_prim → PepGly		Peptidoglycan-Synthase
Glycogen_Synth	6.16 ADPGlc → Glycogen		Glycogen-Synthase
Synthese von Biomasse (g Polymer / g Biomasse):			
μ	0.64 Protein + 0.185 RNA + 0.05 DNA + 0.1 Lip + 0.015 LPS + 0.015 PepGly + 0.015 Glycogen → Biomasse		kumulierte Synthese von einem Gramm Biomasse
Ausscheidungen:			
CO2_ex	1 CO2 ↔		Ausscheidung von Kohlendioxid
Lac_ex	1 Lac →		Ausscheidung von Lactat
Eth_ex	1 Eth →		Ausscheidung von Ethanol
Ac_ex	1 Ac →		Ausscheidung Acetat
Form_ex	1 Form →		Ausscheidung von Formiat (schließt Ausscheidung über Formiat-Dehydrogenase mit ein).

A.3 *Rhodospirillaceae*: Reaktionsgleichungen (berücksichtigen nur die betrachteten internen Metabolite; z.T. kumulierte Stöchiometrien)

Reaktion	Stöchiometrie / Reversibilität	Gen(e)	Enzym / Reaktionsbeschreibung
Substrataufnahme:			
Fruc_up	1 F6P → 1 F1P + 1 Pyr	frsA	Fructoseaufnahme über Fructose-PTS
Mal_up	1 Mal → 1 Mal		Malataufnahme
Succ_up	→ 1 Succ		Succinatenaufnahme
Prop_up	→ 1 Prop		Propionataufnahme
O2_up	→ 1 O2		Sauerstoffaufnahme
N_up	→ 1 N		Ammoniumaufnahme
S_up	4 ATP + 4 NADPH → 1 S		Sulfataufnahme und assimilatrische Reduktion zu Sulfid
Ac_up	→ 1 Ac		Acetataufnahme
Austauschreaktionen (reversibel):			
CO2_exchange	↔ CO2		Aufnahme/Freisetzung von Kohlendioxid
Zentralstoffwechsel:			
F1P::F16P	F1P + ATP ↔ F16P	fruK	1-Phosphofruktokinase
F6P::G6P	1 F6P ↔ 1 G6P	pgl	Glucosephosphat-Isomerase
F16P::F6P	1 F16P → 1 F6P	fbp, cbbF	Fructosebiphosphatase (evtl. auch Pyrophosphatproduzierend, hier nicht berücksichtigt)
F6P::F16P	1 F6P + 1 ATP → 1 F16P	pfkA	6-Phosphofruktokinase
F16P::F3P	1 F16P ↔ 1 DHAP + 1 G3P	fbA, cbbA	Fructosebiphosphat-Aldolase
DHAP::G3P	1 DHAP ↔ 1 G3P	tpiA	Triosephosphat-Isomerase
DHAP::Glyc3P	1 DHAP + 1 NADH → 1 Glyc3P	gpsA	Glycerin-3-phosphat-Dehydrogenase (NAD-spezifisch)

Reaktion	Stöchiometrie / Reversibilität	Gene	Enzym / Reaktionsbeschreibung
GEP::DPG	1 GEP ↔ 1 DPG + 1 NADH	gpaA, cbbG	Glyoxaldehyd-3-phosphat-Dehydrogenase; cbbG liegt im Gencluster für Calvin-Zyklus; auch NAD-spezifisch!
DPG::3PG	1 DPG ↔ 1 3PG + 1 ATP	pgk, cbbK	Phosphoglycerat-Kinase
3PG::2PG	1 3PG ↔ 1 2PG	gpmA	Phosphoglyceromutase
2PG::PEP	1 2PG ↔ 1 PEP	eno	Enolase
PEP::PYR	1 PEP → 1 Pyr + 1 ATP	pykA, pykF	Pyruvat-Kinase I/II
Pyr::PEP	1 Pyr + 2 ATP → 1 PEP	ppsA	PEP-Synthetase; (Verbrauch von 2 ATP eingesetzt, da nicht ADP sondern AMP entsteht)
Pyr::AcCoA	1 Pyr → 1 AcCoA + 1 NADH + 1 CO ₂	aceEF	Pyruvat-Dehydrogenase
AcCoA::Cit	1 AcCoA + 1 OxA → 1 Cit	glfA	Citrat-Synthase
Cit::ICl	1 Cit ↔ 1 ICl	acnA, acnB	Acemalat-Hydrolase
ICl::aIKG	1 ICl ↔ 1 aIKG + 1 NADPH + 1 CO ₂	jeiA	Isocitrat-Dehydrogenase (NADP-spezifisch)
aIKG::SuccCoA	1 aIKG → 1 SuccCoA + 1 NADH + 1 CO ₂	sucAB	Oxoglutarat-Dehydrogenase
SuccCoA::Succ	1 SuccCoA ↔ 1 Succ + 1 ATP	sucCD	Succinyl-CoA-Synthetase
Succ::Fum	1 Succ → 1 Fum + 1 QH ₂	sdhC	Succinat-Dehydrogenase
Fum::Succ	1 Fum + 1 QH ₂ → 1 Succ	fdxA	Fumarat-Reduktase; Elektronendonator: Rhodochinon oder (als umgekehrte Succinat-Dehydrogenase) Ubichinon
Fum::Mal	1 Fum ↔ 1 Mal	fumA	Fumarase
Mal::OxA	1 Mal ↔ 1 OxA + 1 NADH	malH	Malat-Dehydrogenase
CitMalSynth	1 Pyr + 1 Ac ↔ CitMal		Citramalat-Lyase; Start des Citramalat-Zyklus [7,64]
CitMal::Mes	1 CitMal ↔ 1 Mes		Citramalat-Hydrolase (Citramalat-Zyklus)
Mes::PropCoA	1 Mes + 1 AcCoA ↔ 1 PropCoA + 1 Glyox + 1 Ac		zusammengefügte Sequenz aus 3 Reaktionen im Citramalat-Zyklus; z.T. unbekanntes Gen/Enzym (vgl. [7])
Glyox::Mal	1 AcCoA + 1 Glyox → 1 Mal	glbB	Malat-Synthase
Prop::PropCoA	1 Prop + SuccCoA ↔ 1 PropCoA + 1 Succ		Succinat-Propionat-CoA-Transferase
Prop::CoA::MMalCoA	1 PropCoA + 1 CO ₂ + 1 ATP → 1 MMalCoA		Propionyl-CoA-Carboxylase

Reaktion	Stöchiometrie / Reversibilität	Gene	Enzym / Reaktionsbeschreibung
MMAL-CoA: SuccCoA	1 MMALCoA ↔ SuccCoA		Methyl-Malonyl-CoA-Mutase
G6P:PGIac	1 G6P ↔ 1 PGIac + 1 NADPH	zwf	Glucose-6-phosphat-Dehydrogenase (vermutlich nicht in <i>R. rubrum</i> vorhanden)
PGIac:PGIac	1 PGIac → 1 PGIac	pgj	6-Phosphogluconolactonase
PGIac:R15P	1 PGIac → 1 R15P + 1 NADPH + 1 CO ₂	gnd	Gluconat-6-phosphat-Dehydrogenase, decarboxylierend
R15P:X3P	1 R15P ↔ 1 X3P	rpe, cbbE	Ribulose-5-phosphat-3-Epimerase
R15P:R5P	1 R15P ↔ 1 R5P	rpiA	Ribosephosphate-Isomerase
R15P:R115P	1 ATP + 1 R15P ↔ 1 R115P	cbap	Phosphoethalolkinase; Schlüsselreaktion im Calvin-Zyklus
Rubisco	1 R115P + 1 CO ₂ → 2 3PG	cbsM	Rubisco (Ribulosebiphosphat-Carboxylase/Oxygenase); Schlüsselreaktion im Calvin-Zyklus
SUTPSynth	1 E4P + 1 DHAP ↔ 1 S17P	fbA, cbbA	Fructose-1,6-/Sedoheptulose-1,7-bisphosphat-aldolase; Schlüsselreaktion im Calvin-Zyklus
S17P:S7P	1 S17P → 1 S7P	fbp, cbbF	Fructose-1,6-/Sedoheptulose-1,7-bisphosphatase; Schlüsselreaktion im Calvin-Zyklus
Transket1	1 R5P + 1 X5P ↔ 1 G3P + 1 S7P	tkA, tkB, cbbI	Transketolase
Transald1	1 G3P + 1 S7P ↔ 1 F6P + 1 E4P	talB	Transaldolase
Transket2	1 E4P + 1 X5P ↔ 1 F6P + 1 G3P	tkA, tkB, cbbI	Transketolase (isozymne 1/2)
OxAc:PEP	1 OxAc + 1 ATP → 1 PEP + 1 CO ₂	pebA	Phosphoenolpyruvat-Carboxykinase
Pyr:OxAc	1 Pyr + 1 ATP + 1 CO ₂ → 1 OxAc	pta	Pyruvat-Carboxylase
AcCoA:AcP	1 AcCoA ↔ 1 AcP	ackA	Phosphotransacetylase
AcP:Ac	1 AcP ↔ 1 ATP + 1 Ac		Acetat-Kinase; Rückreaktion wird auch über Acetat-Synthase katalysiert (dann mit Verbrauch von 2 ATP)
Pyr:Form	1 Pyr → 1 AcCoA + 1 Form	pfkB	Pyruvat:Formiat-Lyase
Mal: Pyr	1 Mal → 1 Pyr + 1 NADPH + 1 CO ₂	slcA	Malat-Enzym (NADP-spezifisch !)

Reaktion	Stöchiometrie / Reversibilität	Gene(s)	Enzym / Reaktionsbeschreibung
SuccCoA:alKG	1 SuccCoA + 1 CO ₂ + 1 Fd _{ox} → 1 alKG		α-Ketoglutarat-Synthase (reduktiver TCA, vermutlich nur in <i>R. rubrum</i> vorhanden)
AcCoA:Pyr	1 AcCoA + 1 CO ₂ + 1 Fd _{ox} → 1 Pyr		Pyruvat-Synthase (reduktiver TCA, vermutlich nur in <i>R. rubrum</i> vorhanden)
Chl ₂ lyase	1 Cit → 1 OxA + 1 Ac	tauo	Citrat-Lyase (reduktiver TCA)
NADH:hydro	1 NADH ↔ 1 QH ₂ + 4 Hp	pesABC	NADH-Dehydrogenase I
BC1	1 QH ₂ → 2 c ₂ _red + 4 Hp	cro, eta	bc1-Komplex
Oxidase	2 c ₂ _red + 0.5 O ₂ → 2 Hp	puf	Oxidase (cbb3 und aa3)
Photo	[2 Photonen +] 2 c ₂ _red → 1 QH ₂	pmA	Photosynthese im Reaktionszentrum
TransHydro	1 NADH + 1 Hp ↔ 1 NADPH	hup	Transhydrogenase
Hydrogenase	[Wasserstoff H ₂] → 1 QH ₂	hup	Hydrogenase (Aufnahme von H ₂)
Nitrogenase	[Stickstoff N ₂ +] 16 ATP + 8 NADH → 2 N [+ Wasserstoff H ₂]	nif	Nitrogenase (Fixierung von Stickstoff mit H ₂ -Freisetzung)
NADH:Fd	NADH ↔ Fd _{ox}		NADH-Ferredoxin-Oxidoreduktase
ATPSynth	3 Hp ↔ 1 ATP	atp	ATP-Synthase (membrangebunden)
ATPdrain	1 ATP →		undefinierter ATP-Verbrauch (mainenance, ...)
Membrane			
Chc ₂ Synth	2 PEP + 1 EAP + 1 ATP + 1 NADPH → 1 Chc ₂		Synthese von Cheimstat
PRPP ₂ Synth	1 R5P + 2 ATP → 1 PRPP		Synthese von 5-Phosphoribosyl-1-pyrophosphat
MTHF ₂ Synth	1 ATP + 1 NADPH ↔ 1 MTHF		Benötigte Coenzyme für Regeneration von MTHF aus THF (Tetrahydrofolat)
Ala ₂ Synth	1 Pyr + 1 Glu → 1 alKG + 1 Ala		Synthese von Alanin
Val ₂ Synth	2 Pyr + 1 NADPH + 1 Glu → 1 alKG + 1 CO ₂ + 1 Val		Synthese von Valin
Leu ₂ Synth	2 Pyr + 1 AcCoA + 1 NADPH + 1 Glu → 1 alKG + 1 NADH + 2 CO ₂ + 1 Leu		Synthese von Leucin
Asn ₂ Synth	2 ATP + 1 N + 1 Asp → 1 Asn	asnA	Asparagin-Synthase A
Asp ₂ Synth	1 OxA + 1 Glu → 1 alKG + 1 Asp	aspC	Asparat-Transaminase

Reaktion	Stöchiometrie / Reversibilität	Gene(s)	Enzym / Reaktionsbeschreibung
Asp:AspSAld	1 ATP + 1 NADPH + 1 Asp → 1 AspSAld		Synthese von Aspartat-4-semialdehyd
AspSAld:Hiser	1 NADPH + 1 AspSAld → 1 Hiser		Synthese von Homoserin
Lys_Synth	1 di_am_pim → 1 CO ₂ + 1 Lys		Synthese von Lysin
Met_Synth	1 SuccCoA + 1 MTHF + 1 Hiser + 1 Cys → 1 Pyr + 1 Succ + 1 N + 1 Met		Synthese von Methionin
Thr_Synth	1 ATP + 1 Hiser → 1 Thr		Synthese von Threonin
Ile_Synth	1 Pyr + 1 NADPH + 1 Glu + 1 Thr → 1 aIKG + 1 CO ₂ + 1 N + 1 Ile		Synthese von Isoleucin
His_Synth	1 ATP + 1 PRPP + 1 Gln → 1 aIKG + 2 NADH + 1 His		Synthese von Histidin
Glu_Synth	1 aIKG + 1 NADPH + 1 N → 1 Glu	ghlA	Glutamat-Dehydrogenase (NADP-spezifisch)
Gln_Synth	1 ATP + 1 N + 1 Glu → 1 Gln	ghlA	Glutamin-Synthase
Pro_Synth	1 ATP + 2 NADPH + 1 Glu → 1 Pro		Synthese von Prolin
Arg_Synth	1 AcCoA + 4 ATP + 1 NADPH + 1 CO ₂ + 1 N + 1 Asp + 2 Glu → 1 aIKG + 1 Fun + 1 Ac + 1 Arg		Synthese von Arginin
Trp_Synth	1 Chor + 1 PRPP + 1 Gln + 1 Ser → 1 GEP + 1 Pyr + 1 CO ₂ + 1 Glu + 1 Trp		Synthese von Tryptophan
Tyr_Synth	1 Chor + 1 Glu → 1 aIKG + 1 NADH + 1 CO ₂ + 1 Tyr		Synthese von Tyrosin
Phe_Synth	1 Chor + 1 Glu → 1 aIKG + 1 CO ₂ + 1 Phe		Synthese von Phenylalanin
Ser_Synth	1 FPG + 1 Glu → 1 aIKG + 1 NADH + 1 Ser		Synthese von Serin
Gly_Synth	1 Ser → 1 MTHF + 1 Gly		Synthese von Glycin
Cys_Synth	1 AcCoA + 1 S + 1 Ser → 1 Ac + 1 Cys		Synthese von Cystein
cATP_Synth	5 ATP + 1 CO ₂ + 1 PRPP + 2 MTHF + 2 Asp + 1 Gly + 2 Glu → 2 Fun + 1 NADPH + 2 Glu + 1 cATP		Synthese von ATP für RNA
cGTP_Synth	6 ATP + 1 CO ₂ + 1 PRPP + 2 MTHF + 1 Asp + 1 Gly + 3 Glu → 2 Fun + 1 NADH + 1 NADPH + 3 Glu + 1 cGTP		Synthese von GTP für RNA
cCTP_Synth	1 ATP + 1 Gln + 1 cUTP → 1 Glu + 1 cCTP		Synthese von CTP für RNA
cUTP_Synth	4 ATP + 1 N + 1 PRPP + 1 Asp → 1 NADH + 1 cUTP		Synthese von UTP für RNA

Reaktion	Stöchiometrie / Reversibilität	Gene(s)	Enzym / Reaktionsbeschreibung
dATP_Synth	1 NADPH + 1 rATP → 1 dATP		Synthese von ATP für DNA
dGTP_Synth	1 NADPH + 1 rGTP → 1 dGTP		Synthese von GTP für DNA
dCTP_Synth	1 NADPH + 1 rCTP → 1 dCTP		Synthese von CTP für DNA
dTTP_Synth	2 NADPH + 1 MTHF + 1 dUTP → 1 dTTP		Synthese von TTP für DNA
mit_FS_Synth	8.48 AcCoA + 7.48 ATP + (4.2) NADPH → 1 mit_FS		Synthese einer "milderen" Fettsäure in <i>R. rubrum</i> [62]
UDPglc_Synth	1 G6P + 1 ATP → 1 UDPglc		Synthese von UDP-Glucose
CDPEth_Synth	1 3PG + 3 ATP + 1 NADPH + 1 N → 1 NADH + 1 CDPEth		Synthese von CDP-Ethanolamin
OH_myr_ac_Synth	7 AcCoA + 6 ATP + 11 NADPH → 1 OH_myr_ac		Synthese von β-OH-Fettsäuren (für LPS)
C14_0_FS_Synth	2 AcCoA + 6 ATP + 12 NADPH → 1 C14_0_FS		Synthese von C(14:0)-Fettsäuren (für LPS)
hydr_bury_Synth	2 AcCoA + 1 NADH → 1 hydr_bury		Synthese von β-Hydroxybuttersäure-CoA
CMP_KDO_Synth	1 PEP + 1 R5P + 2 ATP → 1 CMP_KDO		Synthese von CMP-2-keto-3-desoxyoctonsäure
NDFHep_Synth	1.5 G6P + 1 ATP → 4 NADPH + 1 NDFHep		Synthese von NDP-Heptose
TDRGlcS_Synth	1 F6P + 2 ATP + 1 N → 1 TDRGlcS		Synthese von TDP-Glucosamin
UDP_NAG_Synth	1 F6P + 1 AcCoA + 1 ATP + 1 Gln → 1 Glu + 1 UDP_NAG		Synthese von UDP-N-acetylglucosamin
UDP_NAM_Synth	1 PEP + 1 NADPH + 1 UDP_NAG → 1 UDP_NAM		Synthese von UDP-N-acetylmuramin
di_am_pim_Synth	1 Pyr + 1 SuccCoA + 1 NADPH + 1 AspSAid + 1 Glu → 1 diKG + 1 Succ + 1 di_am_pim		Synthese von Diaminopimelinsäure
ADPGlc_Synth	1 G6P + 1 ATP → 1 ADPGlc		Synthese eines Glycogen-Monomers
betla_Synth	4 Pyr + 4 GSP + 8 SuccCoA + 3 NADPH + 8 Gly → 1 betla		Synthese von Bacteriochlorophyll a

Polymer-Synthese (mmol Metabolite / g Polymer) [impliziert in "Biomassensynthese-Reaktion" μ enthalten]:

Proc_Synth	39.9455 ATP + 0.88727 Ala + 0.15818 Cys + 0.41636 Asp + 0.45455 Glu + 0.32 Phe + 1.0582 Gly + 0.16364 His + 0.50182 Ile + 0.39273 Lys + 0.77818 Leu + 0.26546 Met + 0.41636 Asn + 0.38182 Pro + 0.45445 Gln + 0.51091 Arg + 0.37273 Ser + 0.43818 Thr + 0.73091 Val + 0.09818 Trp + 0.23818 Tyr → Protein	Protein-Synthese (Aminosäuresammensetzung approximativ von <i>E. coli</i> übernommen)
------------	---	---

Reaktion	Stöchiometrie / Reversibilität	Gene(s)	Enzym / Reaktionsbeschreibung
RNA_Synth	1.2488 ATP + 0.80488 rATP + 0.99024 rGTP + 0.61436 rCTP + 0.66341 rUTP → RNA		RNA-Synthase (Nukleotidzusammensetzung approximativ von <i>E. coli</i> übernommen)
DNA_Synth	4.4129 ATP + 0.59677 dATP + 1.0161 dGTP + 1.0161 dCTP + 0.59677 dTTP → DNA		DNA-Synthase (GTP/CTP-Anteil ca. 63% nach [63])
Lip_Synth	2.8352 ATP + 1.4176 Ser + 1.4176 Glyc3P + 2.8352 mlc_FS → Lip		Lipid-Synthase (Zusammensetzung approximativ von <i>E. coli</i> übernommen)
LPS_Synth	0.46176 UDPGlc + 0.69118 CMPFdh + 0.69118 OHL_myr_ac + 0.69118 C14.0_FS + 0.69118 CMP_KDO + 0.69118 NDPHep + 0.46176 TDPGlc → LPS		Lipopolysaccharid-Synthase (Zusammensetzung approximativ von <i>E. coli</i> übernommen)
PepGly_Synth	5.52 ATP + 2.208 Ala + 1.104 Glu + 1.104 UDP_NAG + 1.104 UDP_NAM + 1.104 di_on_pirn → PepGly		Peptidyl-glycan-Synthase (Zusammensetzung approximativ von <i>E. coli</i> übernommen)
Glycogen_Synth	6.16 ADPGlc → Glycogen		Glycogen-Synthase
PHB_Synth	11.62 hydr_huty → PHB		Synthese von PHB (Polymerisation)
PolyP_Synth	12.66 ATP → PolyP		Synthese von Polyphosphat (Polymerisation)
Behl_Synth	1.145 behla → Behl		Bacteriochlorophyll: Umrechnung mmol → g
Synthese von Biomasse (g Polymer / g Biomasse):			
μ	0.50 Protein + 0.16 RNA + 0.03 DNA + 0.16 Lip + 0.03 LPS + 0.02 PepGly + 0.02 Glycogen + 0.05 PHB + 0.01 Behl + 0.02 PolyP → Biomasse		kumulierte Synthese von einem Gramm Biomasse (Anteil von PHB nach [41], von Lip nach [69], von Behl nach [49]; alle anderen (proportional reduziert) nach [94])
Ausscheidungen:			
Succ_ex	1 Succ →		Ausscheidung von Succinat
Ac_ex	1 Ac →		Ausscheidung Acetat
Form_ex	1 Form →		Ausscheidung von Formiat (schließt: Ausscheidung über Formiat-Dehydrogenase mit ein).

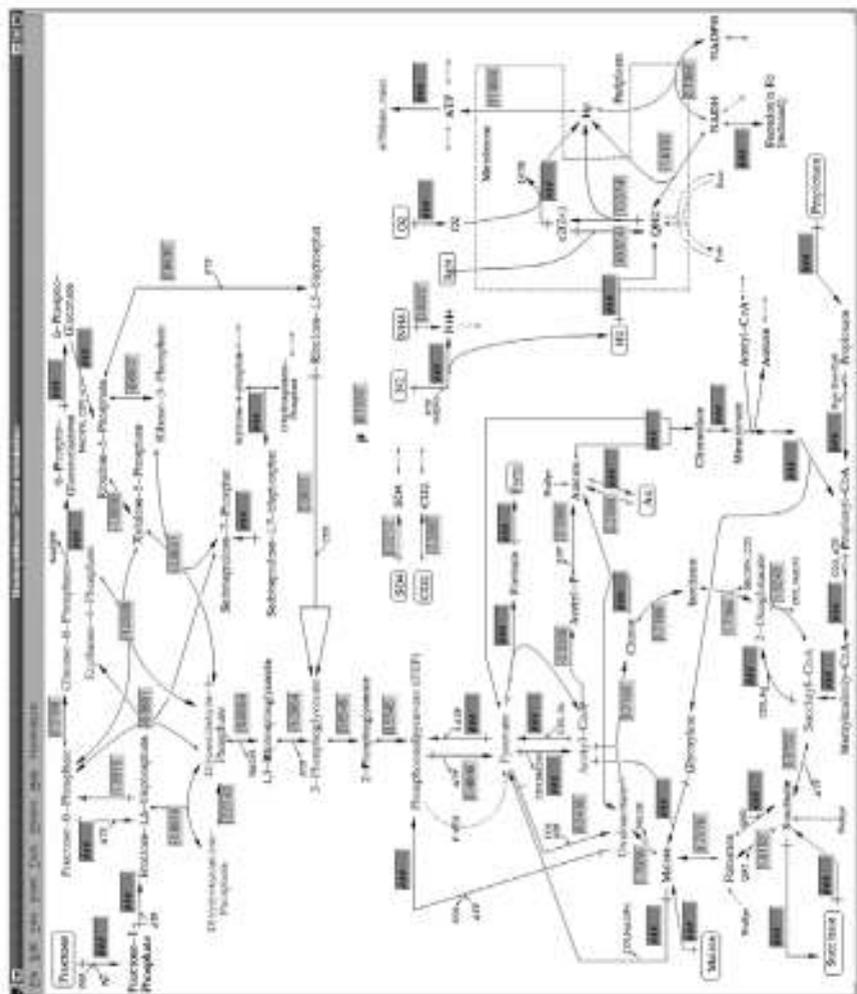
Anhang B: *E. coli* Deletionsstudien

Experimentell beobachtete Phänotypen und relative Auftretshäufigkeiten $k_i^{\text{Substrat}}/k_i^{\text{Referenz}}$ der deletierten Reaktionen (in wachstumsbezogenen EMen für das jeweilige Substrat). Daten für experimentelle Mutanten-Phänotypen wurden hauptsächlich aus [29] entnommen, einige zusätzlich publizierte Daten für die Gene *pta* [15], *pte* [86], *pps*, *sfcA* und *mdh* [161] wurden berücksichtigt.

Substrat	deletiertes Gen	deletierte Reaktion	exp. Phänotyp (1=lebensfähig)	relAH der deletierten Reaktion $k_i^{\text{Substrat}}/k_i^{\text{Referenz}}$
Glucose	<i>fdxA</i>	Pan::Succ	1	0
Succinat	<i>fdxA</i>	Pan::Succ	1	0
Acetat	<i>fdxA</i>	Pan::Succ	1	0
Glycerin	<i>ppsA</i>	Pyr::PEP	1	0.05
Glucose	<i>fbp</i>	Fl6P::F6P	1	0.09
Glucose	<i>ppsA</i>	Pyr::PEP	1	0.11
Succinat	<i>aceA</i>	ICr::Glyox	1	0.15
Succinat	<i>glkB</i>	Glyox::Mal	1	0.15
Succinat	<i>pps</i>	PEP::OxA	1	0.27
Glycerin	<i>glkB</i>	Glyox::Mal	1	0.37
Glucose	<i>aceA</i>	ICr::Glyox	1	0.39
Glucose	<i>glkB</i>	Glyox::Mal	1	0.39
Glycerin	<i>pta</i>	AcCoA::AcP	1	0.41
Glucose	<i>pta</i>	AcCoA::AcP	1	0.42
Acetat	<i>acs</i>	Ac::AcCoA	1	0.42
Succinat	<i>pta</i>	AcCoA::AcP	1	0.44
Acetat	<i>ppsA</i>	Pyr::PEP	1	0.45
Glucose	<i>glk</i>	Gl_ATP_up	1	0.49
Glucose	<i>sfcA</i>	Mal::Pyr	1	0.5
Succinat	<i>ppsA</i>	Pyr::PEP	1	0.5
Glucose	<i>pykAF</i>	PEP::PYR	1	0.51
Glycerin	<i>sfcA</i>	Mal::Pyr	1	0.51
Glycerin	<i>nao</i>	NADHDdehydro	1	0.53
Succinat	<i>pta</i>	TransHydro	1	0.58
Acetat	<i>pta</i>	AcCoA::AcP	1	0.61
Acetat	<i>ackA</i>	AcP::Ac	1	0.61
Glucose	<i>ptsG</i>	Gl_PTS_up	1	0.62
Glycerin	<i>pta</i>	TransHydro	1	0.62
Glucose	<i>sucAB</i>	aKCr::SuccCoA	1	0.64
Glucose	<i>mdh</i>	Mal::OxA	1	0.67
Glucose	<i>pta</i>	TransHydro	1	0.67
Glucose	<i>atp</i>	ATPSynth	1	0.67
Glycerin	<i>mdh</i>	Mal::OxA	1	0.69
Acetat	<i>sfcA</i>	Mal::Pyr	1	0.72
Succinat	<i>mdh</i>	Mal::OxA	1	0.77
Succinat	<i>zwf</i>	GSP::PGluc	1	0.78
Glucose	<i>nao</i>	NADHDdehydro	1	0.79
Glycerin	<i>zwf</i>	GSP::PGluc	1	0.79
Glycerin	<i>rpe</i>	RSP::XSP	1	0.79
Succinat	<i>rpe</i>	RSP::XSP	1	0.8

Substrat	deletiertes Gen	deletierte Reaktion	exp. Phänotyp (Lebensfähig)	relAH der deletierten Reaktion $I_{\text{rel}}^{\text{deletiert}}/I_{\text{rel}}^{\text{wt}}$
Acetat	rpe	R5P: X5P	1	0.8
Glucose	sdhC	Succ: Fum	1	0.82
Glucose	pgl	PGluc: PGIuc	1	0.85
Glucose	god	PGluc: R5P	1	0.85
Glucose	zwf	G6P: PGIuc	1	0.85
Glucose	rpe	R5P: X5P	1	0.86
Glucose	pgi	G6P: F6P	1	0.88
Glucose	cyd	Oxidase	1	0.92
Glycerin	pgi	G6P: F6P	1	1
Succinat	pgl	G6P: F6P	1	1
Succinat	sucAB	aKG: SuccCoA	0	0.7
Glycerin	ppe	PEP: OxA	0	0.71
Acetat	sucAB	aKG: SuccCoA	0	0.71
Acetat	mdh	Mal: OxA	0	0.84
Glucose	fbp	F16P: T3P	0	0.9
Glucose	tpiA	DHAP: G3P	0	0.9
Glucose	pfkA	F6P: F16P	0	0.92
Glucose	glcA	AcCoA: Cit	0	1
Glucose	gapA	G3P: DPG	0	1
Glucose	pgk	DPG: 3PG	0	1
Glucose	eno	2PG: PEP	0	1
Glucose	acnB	Cit: ICit	0	1
Glucose	icdA	ICit: aKG	0	1
Glucose	rpiA	R5P: R5P	0	1
Glycerin	fbp	F16P: F6P	0	1
Glycerin	tpiA	DHAP: G3P	0	1
Glycerin	gapA	G3P: DPG	0	1
Glycerin	pgk	DPG: 3PG	0	1
Glycerin	eno	2PG: PEP	0	1
Glycerin	rpiA	R5P: R5P	0	1
Succinat	fbp	F16P: F6P	0	1
Succinat	sdhC	Succ: Fum	0	1
Succinat	tpiA	DHAP: G3P	0	1
Succinat	gapA	G3P: DPG	0	1
Succinat	pgk	DPG: 3PG	0	1
Succinat	eno	2PG: PEP	0	1
Succinat	rpiA	R5P: R5P	0	1
Acetat	fbp	F16P: F6P	0	1
Acetat	glcA	AcCoA: Cit	0	1
Acetat	sdhC	Succ: Fum	0	1
Acetat	aceA	ICit: Glyox	0	1
Acetat	glcB	Glyox: Mal	0	1
Acetat	tpiA	DHAP: G3P	0	1
Acetat	gapA	G3P: DPG	0	1
Acetat	pgk	DPG: 3PG	0	1
Acetat	eno	2PG: PEP	0	1
Acetat	acnB	Cit: ICit	0	1
Acetat	icdA	ICit: aKG	0	1
Acetat	rpiA	R5P: R5P	0	1
Acetat	atp	ATPSynth	0	1

Anhang C: Elementarmodus mit phototrophem Wachstum der *Rhodospirillaceae* auf Acetat ohne Einbeziehung des Citramalat-Zyklus



Der dargestellte Elementarmodus (nicht involvierte Reaktionen haben dunklen Hintergrund in den Textboxen) zeigt, wie die *Rhodospirillaceae* Acetat phototroph verwenden könnten, ohne dabei den Citramalat-Zyklus zu verwenden.

Anhang D: Beispiel zur Berechnung einer gewichteten Relevanz

Dieser Abschnitt dient der Illustration der in Kap. 6.3 eingeführten gewichteten Relevanzen σ , hier speziell Gleichungen (6.6)-(6.8). Als Beispiel berechnen wir die Relevanzen der Reaktionen im Netz NB2 (Abb. 2.4) bei „Wachstum“ (ausschließlich) auf A ohne Aufnahme von B. Das entspricht dem Szenario $SubA = \{R1, R2 \geq 0\}$, womit hier die ersten 6 EMen von Abb. 2.9 selektiert werden (EM1-EM6; siehe Spalten von $E^{[SubA]}$ in (D.4)). Wir nehmen an, P sei eine Biomassekomponente, so daß wir in Gleichungen (6.6)-(6.8) μ gegen Reaktion R3 (Abfluß von P) ersetzen können. So ergibt sich entsprechend

$$\text{aus (6.6): } w_j^{R3; [SubA]} = \frac{e_j(R3)}{\sum_i e_i(R3)} \cdot \frac{1}{e_{R3}^j \cdot Y_{P;A}^{max}} \quad (D.1)$$

$$\text{aus (6.7): } e_j(R3) = \frac{e_{R3}^j}{\sum_i e_i^j} \quad (D.2)$$

$$\text{aus (6.8): } \sigma^{R3; [SubA]} = E_{Abs}^{[SubA]} w^{R3; [SubA]} \quad (D.3)$$

$$E^{[SubA]} = \begin{pmatrix} 1 & 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix} \begin{matrix} \rightarrow R1 \\ \rightarrow R2 \\ \rightarrow R3 \\ \rightarrow R4 \\ \rightarrow R5 \\ \rightarrow R6 \\ \rightarrow R7 \\ \rightarrow R8 \\ \rightarrow R9 \\ \rightarrow R10 \end{matrix}; \quad E_{Abs}^{[SubA]} = \begin{pmatrix} 1 & 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix} \quad (D.4)$$

Die sechs spezifischen Effizienzen e_j für EM1-EM6 berechnen sich nach (D.2) leicht zu

$$(e_{1M1}, e_{1M2}, e_{1M3}, e_{1M4}, e_{1M5}, e_{1M6}) = (1/4, 1/5, 1/8, 1/7, 0, 0).$$

EM1 hat die höchste spezifische Effizienz, da nur ein kleiner Gesamtfluß zur Produktion von 1 P nötig ist. EM3 und EM4 werden relativ schwach bewertet, da jeweils ein großer Gesamtumsatz zur Synthese von P nötig ist (außerdem ist auch die Ausbeute von P aus A gegenüber EM1 und EM2 halbiert). EM5 und EM6 werden dagegen gar nicht berücksichtigt, da kein P produziert wird (man hätte sie auch von vornherein ausschließen können). Die Summe der spezifischen Effizienzen ist 201/280. Für den maximalen Ertrag von P aus A ergibt sich

$Y_{P,1,4}^{opt} = 1$, der für EM1 und EM2 gleichermaßen gilt. Daraus können wir nun mittels (D.1) die Gewichte für die 6 EMen berechnen:

$$\mathbf{w}^{R3(,SubA)} = (w_{EM1}^{R3(,SubA)}, w_{EM2}^{R3(,SubA)}, w_{EM5}^{R3(,SubA)}, w_{EM4}^{R3(,SubA)}, w_{EM6}^{R3(,SubA)}, w_{EM8}^{R3(,SubA)})^T = \left(\frac{70}{201}, \frac{56}{201}, \frac{17.5}{201}, \frac{20}{201}, 0, 0\right)^T$$

Diesen Gewichtsvektor setzen wir nun schließlich zusammen mit den Absolutwerten der EMen in (D.3) ein, um die Relevanzen $\sigma_j^{R3(,SubA)}$ der 10 Reaktionen für die Produktion von P aus A zu erhalten:

	Reaktion i									
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
$201 \cdot \sigma_j^{R3(,SubA)}$	201	0	163.5	37.5	87.5	76	37.5	73.5	126	37.5

R1 und R3 erhalten die höchsten Relevanzen – sie werden gerade für die Substrataufnahme bzw. Produktausscheidung benötigt. Als nächste folgt dann R9. Diese Reaktion hat eine hohe Relevanz da sie in beiden optimalen EMen (EM1 und EM2) benutzt wird. R5, R6 und R8 folgen mit etwas Abstand, da sie auch an suboptimalen EMen beteiligt sind. Die drei Reaktionen R4, R7, und R10 (die ein *Enzyme Subset* bilden) haben die geringsten Relevanzen, da sie ausschließlich in suboptimalen EMen involviert sind. Dennoch – im Gegensatz zur FBA-Methode – werden sie mit einem Wert größer Null berücksichtigt, da sie zumindest an alternativen Synthesewegen beteiligt sind. R2 dagegen hat Relevanz 0 – die Reaktion ist an keinem EM beteiligt, der die Produktion von P aus A ermöglicht.

Ein Vergleich der hier berechneten gewichteten Relevanzen $\sigma_j^{R3(,SubA)}$ mit den relativen Auftretshäufigkeiten h_i^{S1} im Szenario S1 (Tab. 6.1) zeigt die differenziertere Gewichtung der Reaktionen durch das Effizienzkriterium.

Anhang E: Theoretische und gemessene Transkriptverhältnisse von Genen des Zentralstoffwechsels in *E. coli* für die Substratpaare Acetat/Glucose und Glycerin/Glucose

Die 95%-Konfidenzintervalle der mittels DNA-Microarrays gemessenen Transkriptverhältnisse für das Substratpaar Acetat/Glucose sind der Referenz [106] entnommen. Für Glycerin/Glucose wurden Daten aus [105] verwendet, wobei die 95%-Konfidenzintervalle aus der in [104] angegebenen mittleren Standardabweichung der Messungen berechnet wurden (unter der Annahme normalverteilter Meßfehler). Die theoretischen Transkriptverhältnisse $\theta(\text{Acetat,Glucose})$ bzw. $\theta(\text{Glycerin,Glucose})$ wurden nach Gleichungen (6.6)-(6.12) ermittelt. Die zugrundeliegende Menge an EMen entspricht der in Abschnitt 6.2.2.

Gen	(Reaktion in Anhang A.2)	Messungen: 95%-Konfidenzintervall Acetat/Glucose		$\theta(\text{Acetat,Glucose})$	Messungen: 95%-Konfidenzintervall Glycerin/Glucose		$\theta(\text{Glycerin,Glucose})$
pgi	G6P:F6P	0.57	1.40	0.40	0.75	1.05	0.27
fbp	F16P:F6P	2.50	4.30	1.00	0.75	1.05	1.15
pfkA	F6P:F16P	0.45	0.78	0.06	0.65	0.95	0.15
fbp	F16P:T3P	0.40	0.62	0.22	0.85	1.15	0.16
tpiA	DHAP:G3P	0.65	1.40	0.24	0.75	1.05	2.33
gapA	G3P:DPG	0.30	0.69	0.17	1.05	1.35	1.00
pgk	DPG:3PG	0.46	0.77	0.17	1.05	1.35	1.00
gpmA	3PG:2PG	0.84	4.00	0.28	0.75	1.05	1.00
eno	2PG:PEP	0.34	0.86	0.28	0.65	0.95	1.00
pykA, pykF	PEP:PYR	0.59	1.06	0.59	2.05	2.35	1.58
		(gemittelt)			(pykA)		
ppsA	Pyr:PEP	8.60	21.00	2.68	0.75	1.05	0.67
aceEF	PYR:AcCoA	0.27	0.64	1.53	0.55	0.85	1.19
gltA	AcCoA:Cl	3.50	6.80	10.45	3.35	3.65	1.38
acnA, acnB	Cl:ICl	3.60	13.10	10.45	1.75	2.05	1.38
		(acnB)			(acnB)		
iodA	ICl:alKG	1.30	2.50	5.04	1.05	1.35	1.46
sucAB	alKG:SuccCoA	1.20	3.20	7.49	3.55	3.85	1.72
sucCD	SuccCoA:Succ	1.80	4.70	7.78			1.76
sdhC	Succ:Fum	1.20	2.50	13.63	1.25	1.55	1.49
fumA, fumB, fumC	Fum:Mal	2.20	5.70	3.56	2.45	2.75	1.03
		(fumA)			(fumA)		
mdh	Mal:Ox	2.40	6.20	4.85	1.75	2.05	1.06
aceA	ICl:Glyox	17.4	32.80	20.54	1.35	1.65	1.24
glicB	Glyox:Mal	9.33	31.60	20.54			
zwf	G6P:PGluc	0.41	0.77	0.53	1.05	1.35	0.36
gnd	PGluc:R5P	0.43	0.84	0.53	0.85	1.15	0.36

qpe	R5P: X5P	0.78	1.40	0.52	0.95	1.25	0.33
qpiA, qpiB	R5P: R5P	0.71	1.80	0.68	1.05	1.35	0.49
tkiA, tkiB	Transket1		(qpiA)			(qpiA)	
tkiA, tkiB	Transket1	0.95	1.90	0.51	0.95	1.25	0.34
talA, talB	Transald	0.69	1.30	0.51	0.85	1.15	0.34
tkiA, tkiB	Transket2		(tkiA)			(tkiA)	
tkiA, tkiB	Transket2	0.95	1.90	0.56	0.95	1.25	0.35
pckA	OxA::PEP	4.90	14.40	4.31	2.85	3.15	0.65
ppc	PEP::OxA	0.17	0.45	0.49	0.75	1.05	1.11
adhE	AcCoA::Adh	0.09	0.26	0.05	1.35	1.65	0.69
adhE	Adh::Eth	0.09	0.26	0.05	1.35	1.65	0.69
pta	AcCoA::AcP	0.45	0.96	15.36	0.65	0.95	0.95
ackA	AcP::Ac	0.38	0.71	15.36	0.25	0.55	0.95
pflB	Pyr::Form	1.38	2.24	0.73	1.35	1.65	0.90
ldhA	Pyr::Lac	0.76	1.20	0	0.95	1.25	0.80
stcA	Mal::Pyr	1.20	2.40	4.15			
gpsA	DHAP::Glyc3P	0.91	1.29	1.33			
poxB	Pyr::Ac	0.98	4.27	0.81			
acs	Ac::AcCoA	7.80	11.60	8.80			
nuc	NADHDehydro	0.58	1.45	3.26			
cyd	Oxidase	0.78	1.60	4.05			
prtA	TransHydro	0.28	0.69	0.62	1.05	1.35	1.10
atp	ATPSynth	0.89	1.30	3.86			
asnA	Asn_Synth	0.54	1.05	1.33	3.45	3.75	1.05
aspC	Asp_Synth	1.00	2.00	0.50	0.85	1.15	0.91
aspA	Asp::Fum	1.66	3.98	0.25	1.95	2.25	0.86
gdhA	Glu_Synth	0.76	1.78	0.77	1.55	1.85	0.96
glnA	Gln_Synth	0.34	0.56	1.33	0.55	0.75	1.05

Anhang F: Dynamisches Modell der Elektronentransportkette der *Rhodospirillaceae*

Symbole/Abkürzungen:

- $\text{hill}(S,K,c)$ ist die (normierte) Hill-Kinetik: $\text{hill}(S,K,c) = \frac{S^c}{S^c + K^c}$ ($c = 1 \rightarrow$ Michaelis-Menten-Kinetik; $c > 1 \rightarrow$ sigmoidale Kinetik))
- $\text{inh}(S,K,c)$ ist (normierter) Inhibierungsterm mit: $\text{inh}(S,K,c) = \frac{K^c}{S^c + K^c}$
- $\text{ln}(\dots)$ = natürlicher Logarithmus; $\text{log}(\dots)$ = dekadischer Logarithmus; $\text{exp}(x) = e^x$
- gDW = gram dry weight = gTM = Gramm Trockenmasse

Zustandsgrößen (Z) und algebraische Größen (A) des Modells

(Parameter sind in den Gleichungen *korrekt* geschrieben).

Symbol	Typ	Einheit	Differentialgleichung / algebraische Gleichung	Bemerkungen
$c2_red$	Z	$\mu\text{mol/gDW}$	$c2_red' = 1000 \cdot (2 \cdot \text{be-l_flux} - 4 \cdot \text{Oxidase_flux} - \text{RC_a2d3_flux} - \text{RC_a4s5_flux}) - \text{mae} \cdot c2_red$	reduziertes Cytochrom c_2 (c_2^{red})
$c2_ox$	A	$\mu\text{mol/gDW}$	$c2_ox = c2 - c2_red$	oxidiertes Cytochrom c_2 (c_2^{ox})
$c2_charge$	A	---	$c2_charge = c2_red/c2$	Anteil des reduzierten c_2 am Gesamtpool
$c2$	Z	$\mu\text{mol/gDW}$	$c2' = c2_synth_mito + c2_synth_max \cdot \text{hill}(\text{PrA_P_Kon_PrA_c2_synth}, 3) - \text{mae} \cdot c2$	Gesamtpool an Cytochrom c_2
Ubi_red	Z	$\mu\text{mol/gDW}$	$\text{Ubi_red}' = 1000 \cdot (\text{NADH_DH_flux} - \text{be-l_flux} - 2 \cdot \text{UbiOxidase_flux} + \text{RC_s7d8_flux} + \text{Succ_DHL_flux}) - \text{mae} \cdot \text{Ubi_red}$	reduziertes Ubichinon (QH2)
Ubi_RC	A	$\mu\text{mol/gDW}$	$\text{Ubi_RC} = \text{RCs1} + \text{RCs2} + \text{RCs3} + \text{RCs4} + \text{RCs5} + \text{RCs6} + \text{RCs7}$	an Reaktionszentren gebundenes Ubichinon
Ubi_ox	A	$\mu\text{mol/gDW}$	$\text{Ubi_ox} = \text{Ubi} - \text{Ubi_RC} - \text{Ubi_red}$	oxidiertes (leeres) Ubichinon (Q)
Ubi_charge	A	---	$\text{Ubi_charge} = \text{Ubi_red}/(\text{Ubi} - \text{Ubi_RC})$	Anteil des reduzierten Ubichinons QH2 am gesamten (freien) Ubichinonpool
Ubi	Z	$\mu\text{mol/gDW}$	$\text{Ubi}' = 10 \cdot (c2_synth_mito + c2_synth_max \cdot \text{hill}(\text{PrA_P_Kon_PrA_c2_synth}, 3)) - \text{mae} \cdot \text{Ubi}$	Gesamtpool an Ubichinon; Expression wie bei c2 im Verhältnis 10:1

Symbol	Typ	Einheit	Differentialgleichung / algebraische Gleichung	Bemerkungen
bc1	Z	$\mu\text{mol/gDW}$	$bc1' = 0.5 \cdot (c2_synth_min + c2_synth_max - hill(PrA_P, Kw_PrA_c2_synth, 3)) - mae \cdot bc1$	bc_1 -Komplex; Expression wie bei c2 im Verhältnis 1:2
Oxidase	Z	$\mu\text{mol/gDW}$	$Oxidase' = Ori_synth_min + Ori_synth_max \cdot inh(PrA_P, Km_Oxysynth_PrA, 4) \cdot hill(PrA_dim, Km_Oxysynth_PrA, 4) - mae \cdot Oxidase$	Cytochrom-Oxidase (cbb); Expression reguliert über PrA_dim und PrA-P
Oxidase_red	A	$\mu\text{mol/gDW}$	$Oxidase_red = Oxidase(1 + (c2_ox)/c2_red) \cdot exp((U2_pmf * 2/2_psl - Ew_A3 + Ew_c2)/ZZZ)$	in Cytochrom b ₅ reduzierte Oxidase ebb; Annahme: (a) b ₅ im Redoxgleichgewicht mit c ₂ (vgl. [78]); (b) eines der beiden bewegten Protonen wird bei Übergabe der Elek. vom c ₂ zum b ₃ ins Peripl. gepumpt
Oxidase_flux	A	nmol/gDW-h	$Oxidase_flux = K_Oxidase_flux \cdot Oxidase_red \cdot c2_red \cdot hill(pO2, Km_Oxidase, 3)$	Fluß durch die Cytochrom-Oxidase
Flux_per_Oxidase	A	nmol/gmol-h	$Flux_per_Oxidase = Oxidase_flux / Oxidase$	Fluß je Cytochrom-Oxidase (PrA-Signal)
UbiOxidase	Z	$\mu\text{mol/gDW}$	$UbiOxidase' = UbiOri_synth_min + UbiOri_synth_max \cdot inh(PrA_P, Kw_Oxysynth_PrA, 4) \cdot hill(PrA_dim, Km_Oxysynth_PrA, 4) - mae \cdot UbiOxidase$	UbiOxidase; Expression reguliert über PrA_dim und PrA-P
UbiOxidase_red	A	$\mu\text{mol/gDW}$	$UbiOxidase_red = UbiOxidase / (1 + Ubi_ox/Ubi_red) \cdot exp((2/2_pmf - Ew_UbiOri + Ew_Ubi)/ZZZ)$	reduzierte UbiOxidase (im Redoxgleichgewicht mit Ubichinonpool; siehe auch Oxidase_red)
UbiOxidase_flux	A	nmol/gmol-h	$UbiOxidase_flux = K_UbiOxidase_flux \cdot UbiOxidase_red \cdot Ubi_red \cdot hill(pO2, Km_UbiOxidase, 1)$	Fluß durch die Ubichinon-Oxidase
RCs0	Z	$\mu\text{mol/gDW}$	$RCs0' = 1000 \cdot (RC_s7s0_flux - RC_s0s1_flux) + PSC_synth - mae \cdot RCs0$	Reaktionszentren im Zustand 0 (Abb. 9.3)
RCs1	Z	$\mu\text{mol/gDW}$	$RCs1' = 1000 \cdot (RC_s0s1_flux - RC_s1s2_flux) - mae \cdot RCs1$	Reaktionszentren im Zustand 1 (Abb. 9.3)
RCs2	Z	$\mu\text{mol/gDW}$	$RCs2' = 1000 \cdot (RC_s1s2_flux - RC_s2s3_flux) - mae \cdot RCs2$	Reaktionszentren im Zustand 2 (Abb. 9.3)
RCs3	Z	$\mu\text{mol/gDW}$	$RCs3' = 1000 \cdot (RC_s2s3_flux - RC_s3s4_flux) - mae \cdot RCs3$	Reaktionszentren im Zustand 3 (Abb. 9.3)
RCs4	Z	$\mu\text{mol/gDW}$	$RCs4' = 1000 \cdot (RC_s3s4_flux - RC_s4s5_flux) - mae \cdot RCs4$	Reaktionszentren im Zustand 4 (Abb. 9.3)
RCs5	Z	$\mu\text{mol/gDW}$	$RCs5' = 1000 \cdot (RC_s4s5_flux - RC_s5s6_flux) - mae \cdot RCs5$	Reaktionszentren im Zustand 5 (Abb. 9.3)
RCs6	Z	$\mu\text{mol/gDW}$	$RCs6' = 1000 \cdot (RC_s5s6_flux - RC_s6s7_flux) - mae \cdot RCs6$	Reaktionszentren im Zustand 6 (Abb. 9.3)
RCs7	Z	$\mu\text{mol/gDW}$	$RCs7' = 1000 \cdot (RC_s6s7_flux - RC_s7s0_flux) - mae \cdot RCs7$	Reaktionszentren im Zustand 7 (Abb. 9.3)
RC	A	$\mu\text{mol/gDW}$	$RC = RCs0 + RCs1 + RCs2 + RCs3 + RCs4 + RCs5 + RCs6 + RCs7$	Konz. totale Menge an Reaktionszentren

Symbol	Typ	Einheit	Differentialgleichung / algebraische Gleichung	Bemerkungen
LH	Z	$\mu\text{mol/gDW}$	$LH = \text{PSC_synth} - \text{mus-LH}$	Leitungsmitterkomplexe; in <i>R. rubrum</i> in fixem Verhältnis zu [RC]
Bchl	A	$\mu\text{mol/gDW}$	$Bchl = 32 \cdot LH + 4 \cdot RC$	
PpsR_blocked	Z	L/100%	$\text{PpsR_blocked} = K_{\text{AppA_PpsR_Block}} \cdot (1 - \text{PpsR_blocked}) \cdot \text{AppA_red} - K_{\text{PpsR_Block_decay}} \cdot \text{PpsR_blocked}$	gesamtes Bacteriochlorophyll (<i>R. rubrum</i>) Blockierter Anteil des Repressors PpsR
AppA_red	Z	L/100%	$\text{AppA_red} = -\text{PpsR_blocked} + K_{\text{AppA_Q}} \cdot (\text{deltaED_AppA_Q} + ZZZZ \cdot \ln(\text{Ubi_red} \cdot (1 - \text{AppA_red})) / (\text{Ubi_ox} \cdot \text{AppA_red}))$	Anteil von reduziertem AppA am Gesamtpool von AppA; hängt vom Redoxzustand des Ubichinons ab
PprB_P	Z	L/100%	$\text{PprB_P} = K_{\text{PrrA_P_dephospho}} \cdot (1 - \text{PprB_P}) \cdot \text{PrrA_P} - K_{\text{PrrA_phospho}} \cdot (1 - \text{PrrA_P}) \cdot \text{PprB_P} - K_{\text{PrrB_P_dephospho}} \cdot \text{PprB_P} + K_{\text{PrrB_phospho}} \cdot \text{inhFlux_per_Oxidase} \cdot K_{\text{PrrB_phospho}} \cdot 3) \cdot (1 - \text{PprB_P})$	Phosphorylierter Anteil der PrrB-Sensorykinase (Autophosphoryl. wird durch Aktivität der Cytochrom-Oxidase inhibiert)
PrrA_P	Z	L/100%	$\text{PrrA_P} = K_{\text{PrrA_phospho}} \cdot (1 - \text{PrrA_P}) \cdot \text{PrrB_P} - K_{\text{PrrA_P_decay}} \cdot \text{PrrA_P} - K_{\text{PrrA_P_dephospho}} \cdot (1 - \text{PrrB_P}) \cdot \text{PrrA_P}$	Phosphorylierter Anteil des Response Regulators PrrA
PrrL_dim	A	L/100%	$\text{PrrL_dim} = \text{inh}(pO_2, K_{\text{pO2_PrrL_dim}}, 4)$	dimensioener Anteil von PrrL
PSC_synth	A	$\mu\text{mol/gDW-h}$	$\text{PSC_synth} = \text{PSC_synth_max} \cdot \text{hill}(\text{PrrA_P}, K_{\text{PrrA_PSC_synth}}, 2) \cdot \text{inh}(1 - \text{PpsR_blocked}, K_{\text{PpsR_PSC_synth}}, 2) - \text{hill}(\text{PrrL_dim}, K_{\text{PrrL_PSC_synth}}, 1)$	Koordinierte Expression der photosynthetischen Gene über PrrA-P, PrrL_dim und PpsR (Aunahme: quasi-stationäres Binden der Regulatoren an die DNA)
light_harvesting	A	---	$\text{light_harvesting} = K_{\text{th}} \cdot LH \cdot \text{hill}(\text{musox_profluance}, K_{\text{th_lightint}}, 4)$	(relative) Menge absorbierten Photons
RC_sbs1_flux	A	nmol/gDW-h	$\text{RC_sbs1_flux} = K_{\text{RC_sbs1_flar}} \cdot \text{RCs0} \cdot \text{Ubi_ox}$	Übergangsrate von RCs0 zu RCs1 (irrev.)
RC_s1s2_flux	A	nmol/gDW-h	$\text{RC_s1s2_flux} = K_{\text{RC_s1s2_flar}} \cdot \text{RCs1} \cdot \text{light_harvestingRC} - K_{\text{RC_s1s2_flux_b}} \cdot \text{RCs2}$	Übergangsrate von RCs1 zu RCs2 (rev.) (Abb. 9.3)
RC_s2s3_flux	A	nmol/gDW-h	$\text{RC_s2s3_flux} = K_{\text{RC_s2s3_flar}} \cdot \text{RCs2} \cdot c2_red - K_{\text{RC_s2s3_flux_b}} \cdot \text{RCs3} \cdot c2_ox$	Übergangsrate von RCs2 zu RCs3 (rev.)
RC_s3s4_flux	A	nmol/gDW-h	$\text{RC_s3s4_flux} = K_{\text{RC_s3s4_flar}} \cdot \text{RCs3} - K_{\text{RC_s3s4_flux_b}} \cdot \text{RCs4}$	Übergangsrate von RCs3 zu RCs4 (rev.)
RC_s4s5_flux	A	nmol/gDW-h	$\text{RC_s4s5_flux} = K_{\text{RC_s1s2_flar}} \cdot \text{RCs4} \cdot \text{light_harvestingRC} - K_{\text{RC_s1s2_flux_b}} \cdot \text{RCs5}$	(analog zu RC_s1s2_flux)

Symbol	Typ	Einheit	Differentialgleichung / algebraische Gleichung	Bemerkungen
RC_s56_flux	A	mmol/(gDW-h)	$RC_s56_flux = K_RC_s23_flux \cdot RCs5_c2_red - K_RC_s23_flux_b \cdot RCs6_c2_ox$	Übergangsrate von RCs5 zu RCs6 (rev.) (analog zu RC_s23_flux)
RC_s6s7_flux	A	mmol/(gDW-h)	$RC_s6s7_flux = K_RC_s34_flux \cdot RCs6 - K_RC_s34_flux_b \cdot RCs7$	Übergangsrate von RCs6 zu RCs7 (rev.) (analog zu RC_s34_flux)
RC_s7s0_flux	A	mmol/(gDW-h)	$RC_s7s0_flux = K_RC_s7s0_flux \cdot RCs7$	Übergangsrate von RCs7 zu RCs0 (irrev.)
NADH_DH_flux	A	mmol/(gDW-h)	$NADH_DH_flux = K_NADHDDehydro_flux \cdot (Em_UbH - 320 + ZZ2 \cdot \ln(NADH_NAD_ratio \cdot Ubi_ox/Ubi_red)) - 4Q \cdot pmf$	Fluß durch NADH-Dehydrogenase
Succ_DH_flux	A	mmol/(gDW-h)	$Succ_DH_flux = K_SuccDehydro_flux \cdot (Em_UbH - 31 + ZZ2 \cdot \ln(Succ_Fwst_ratio \cdot Ubi_ox/Ubi_red))$	Fluß durch Succinat-Dehydrogenase
bc1_flux	A	mmol/(gDW-h)	$bc1_flux = K_bc1_flux \cdot bc1 \cdot (Em_c2 - Em_UbH + ZZ \cdot \ln(c2_ox/c2_red) - ZZ2 \cdot \ln(Ubi_ox/Ubi_red)) - 4Q \cdot pmf + 2Q \cdot psi$	Fluß durch bc1-Komplex
ATPSynthase_flux	A	mmol/(gDW-h)	$ATPSynthase_flux = K_ATPSynthase_flux \cdot (3 \cdot pmf \cdot F - 30.5 - R \cdot T \cdot \ln(ATP_ADP_ratio/Phos))$	Fluß durch die ATP-Synthase
muw	A	l/h	$muw = wsw_swr \cdot hlll(ATPSynthase_flux, Kw_swr_energy, 3)$	Wachstumsrate; Annahme: korreliert mit ATP-Synthase-Fluß
proton_leak	A	mmol/(h)	$proton_leak = K_Hp_leak \cdot pmf$	Rückfluß von Protonen aus dem Periplasma (ohne Energiegewinn)
Hp	Z	mmol/d	$Hp = (4 \cdot NADH_DH_flux + 4 \cdot bc1_flux + 4 \cdot Oxidase_flux + 2 \cdot UbiOxidase_flux - 3 \cdot ATPSynthase_flux) \cdot 410 / pIV - proton_leak$	Konzentration periplasmatischer Protonen, (Zellfläche = 410 gDW/l)
deltaph	A	---	$deltaph = -\log(1/pIV) - pHi$	pH-Differenz: Periplasma-Cytoplasma
psi	A	mV	$psi = -2.303 \cdot ZZ \cdot deltaph \cdot psi_p / (1 - psi_p)$	Membranpotential (fixer Anteil am pmf)
pmf	A	mV	$pmf = psi - 2.303 \cdot ZZ \cdot deltaph$	protonenmotorische Kraft

Parameter

Name	Wert	Bemerkungen
<i>mean_illumance</i>		mittlere Einstrahlungsintensität in $[J/E/(s \cdot m^2)]$; wird für jede Simulation vorgegeben (externe Größe)
<i>pO2</i>		pO_2 (Calciumsenstoff in % Sättigung); wird für jede Simulation vorgegeben (externe Größe)
<i>ATP_ADP_ratio</i>		ATP/ADP-Verhältnis in der Zelle; wird für jede Simulation vorgegeben
<i>NADH_NAD_ratio</i>		NADH/NAD-Verhältnis in der Zelle wird für jede Simulation vorgegeben
<i>Succ_Fum_ratio</i>		Succinat/Fumarat-Verhältnis in der Zelle; wird für jede Simulation vorgegeben
<i>F</i>	0.0965 kJ/(mol·mV)	Faraday-Konstante
<i>R</i>	0.0083 kJ/(mol·K)	Boltzmann-Konstante
<i>T</i>	298 K	Temperatur
<i>ZZ</i>	25.6311 mV	$ZZ=R \cdot T/F$
<i>pH</i>	1	spezifiziert, wieviel $Mol H^+$ je Liter Zellvolumen ins Periplasma gepumpt werden müssen, damit sich die Konzentration $[H^+]$ um ein mol/l erhöht; hängt von der Pufferkapazität der Protonen im Periplasma und vom Volumen des Periplasmas ab (geschätzt)
<i>pH_i</i>	7	pH im Cytoplasma
<i>psi_p</i>	0.7	Anteil des Membranpotentials am pmf (bestimmt aus [172]); $\Delta\psi_{pH} = 1 \rightarrow \Delta\psi_{pH}$ hat rund 30% am pmf
<i>Phos</i>	0.001 M	cytoplasmatische Phosphat-Konzentration
<i>max_rate</i>		
<i>Em_max_energy</i>	0.15 l^{-1}	maximale Wachstumsrate <i>R. rubrum</i> (geschätzt aus [58])
	10	
<i>Em_c2</i>	343 mV	Standard-Redoxpotential ($pH=7$) für Cytochrom c_1 [43]
<i>Em_UIV</i>	70 mV	Standard-Redoxpotential ($pH=7$) für Ubichinon [43]
<i>Em_UIIIa2a1</i>	225 mV	Standard-Redoxpotential ($pH=7$) für Cytochrom b in der Ubichinon-Oxidase (geschätzt)
<i>Em_A3</i>	540 mV	Standard-Redoxpotential ($pH=7$) für Cytochrom b_5 in der cbb ₃ -Oxidase (geschätzt nach [78])

K_ATPSynthase_flux	4.55	
K_NADHDehydro_flux	0.305	
K_SuccDehydro_flux	0.053	
K_Is1_flux	7.03	
K_Hp_leak	0.0002	
K_Oxidase_flux	1.54	
Km_Oxidase	0.4 [%pO2]	
K_UbOxidase_flux	5.5	
Km_UbOxidase	0.2 [%pO2]	
c2_glyox_max	0.12	
c2_glyox_min	0.055	
Km_PyrA_c2_glyox	0.49	
Oxi_glyox_min	0.074	
Oxi_glyox_max	0.08	
Km_Oxizynth_PyrA	0.53	
Km_Oxizynth_Ford	0.33	
Ub/Oxi_glyox_min	0.004	
Ub/Oxi_glyox_max	0.18	
PSC_glyox_max	0.12	
Km_Ford_PSC_glyox	0.85	
Km_PyrA_PSC_glyox	0.32	
Km_PyrB_PSC_glyox	0.32	
deltAE0_AppA_Q	-20	Differenz der Standard-Redoxpotentiale (pH=7) zw. AppA (FAD-Gruppe) und Ubichinon (geschätzt)
K_AppA_Q	5	
K_AppA_PyrB_Block	0.9	
K_PyrB_Block_decay	0.2	

<i>K_PrrB_p_hospho</i>	10	
<i>Km_PrrB_p_hospho</i>	0.7	
<i>K_PrrB_P_diphospho</i>	1	
<i>K_PrrA_p_hospho</i>	10	
<i>K_PrrA_P_diphospho</i>	10	
<i>K_PrrA_P_decay</i>	0.5	
<i>Km_p02_Fwt_dim</i>	0.1	
<i>K_h</i>	3000	
<i>Km_ib_digluur</i>	74 $\mu\text{E}/(\% \text{m}^2)$	nach [3]
<i>K_RC_x03_flux</i>	1400	
<i>K_RC_x12_flux</i>	4	geschätzt nach Angaben in [95]
<i>K_RC_x12_flux_b</i>	1	
<i>K_RC_x23_flux</i>	1800	geschätzt nach Angaben in [95]
<i>K_RC_x23_flux_b</i>	180	
<i>K_RC_x304_flux</i>	2200	geschätzt nach Angaben in [95]
<i>K_RC_x304_flux_b</i>	220	
<i>K_RC_x740_flux</i>	1500	

Parameter ohne Referenzangaben wurden geschätzt.

Literaturverzeichnis

- [1] Addelee, H.A. und Hunter, C.N. (2002) *Rhodospirillum rubrum* possesses a variant of the *bchP* gene, encoding geranylgeranyl-bacteriopheophytin reductase. *Journal of Bacteriology*, **184**, 1578-86.
- [2] Alberti, M., Burke, D.H. und Hearst, J.E. (1995) Structure and sequence of the photosynthesis gene cluster. In: Blankenship, R.E., Madigan, M.T. und Bauer, C.E., editors. *Anoxygenic photosynthetic bacteria*. Kluwer Academic Publishers, Dordrecht, pp. 1083-1106.
- [3] Albiol, J. (1994) Study of the MELiSSA photoheterotrophic compartment. Report ESA-EWP-1808.
- [4] Alexeeva, S., Kort, B., Sawers, G., Hellingwerf, K.J. und de Mattos, M.J.T. (2000) Effects of limited aeration and of the *ArcAB* system on intermediary pyruvate catabolism in *Escherichia coli*. *Journal of Bacteriology*, **182**, 4934-4940.
- [5] Alon, U. (2003) Biological networks: the tinkerer as an engineer. *Science*, **301**, 1866-1867.
- [6] Bailey, J.E. (2001) Complex biology with no parameters. *Nature Biotechnology*, **19**, 503-504.
- [7] Berg, I.A., Krasnikova, E.N. und Ivanovsky, R.N. (2000) Investigation of the dark metabolism of acetate in photoheterotrophically grown cells of *Rhodospirillum rubrum*. *Microbiology*, **69**, 7-12.
- [8] Bertsimas, D. und Tsitsiklis, J.N. (1997) *Introduction to linear optimization*. Athena Scientific, Belmont.
- [9] Bonarius, H.P.J., Schmid, G. und Tramper, J. (1997) Flux analysis of underdetermined metabolic networks: the quest for the missing constraints. *Trends in Biotechnology*, **15**, 308-314.
- [10] Bray, D. (2003) Molecular networks: the top-down view. *Science*, **301**, 1864-1865.
- [11] Buchanan, B.B., Evans, M.C.W. und Arnon, D.I. (1967) Ferredoxin-dependent carbon assimilation in *Rhodospirillum rubrum*. *Arch. Mikrobiol.*, **59**, 32-40.
- [12] Cakir T., Kirdar, B. und Ulgen, K.O. (2004) Metabolic pathway analysis of yeast strengthens the bridge between transcriptomics and metabolic networks. *Biotechnology and Bioengineering*, **86**, 251-60.
- [13] Canonaco F., Hess T.A., Heri S., Wang T., Szyperski T. und Sauer U. (2001) Metabolic flux response to phosphoglucose isomerase knock-out in *Escherichia coli* and impact of overexpression of the soluble transhydrogenase *UdhA*. *FEMS Microbiological Letters*, **204**, 247-252.
- [14] Carlson, R., Fell, D., Sreenc, F. (2002) Metabolic pathway analysis of a recombinant yeast for rational strain development. *Biotechnology and Bioengineering*, **79**, 121-134.
- [15] Chang, D.-E., Shin, S., Rhee, J.-S. und Pann, J.-G. (1999) Acetate metabolism in a *pta* mutant of *Escherichia coli* W3110: importance of maintaining acetyl coenzyme A flux for growth and survival. *Journal of Bacteriology*, **181**, 6656-6663.
- [16] Christensen, B. und Nielsen, J. (1999) Isotopomer analysis using GC-MS. *Metabolic Engineering*, **1**, 282-290.
- [17] Clarke, B.L. (1980) Stability of complex reaction networks. In: Prigogine, I. und Rice, S.A., editors. *Advances in Chemical Physics*, vol. 43. John Wiley, New York, pp.1-215.
- [18] Clarke, B.L. (1988) Stoichiometric network analysis. *Cell Biophysics*, **12**, 237-253.
- [19] Cohen-Bazire, G., Siström, W.R. und Stanier, R.Y. (1957) Kinetic studies of pigment synthesis by nonsulfur purple bacteria. *Journal Cell. Comp. Physiol.*, **49**, 25-68.
- [20] Cohen-Bazire, G., Siström, W.R. und Stanier, R.Y. (1957) Kinetic studies of pigment synthesis by non-sulfur purple bacteria. *J. Cell. Comp. Physiol.*, **49**, 25-68.
- [21] Conrad, R. und Schlegel, H.G. (1974) Different pathways for fructose and glucose utilization in *Rhodospseudomonas capsulata* and demonstration of 1-phosphofructokinase in phototrophic bacteria. *Biochimica et Biophysica Acta*, **358**, 221-225.
- [22] Covert, M. und Palsson, B.O. (2003) Constraints-based models: Regulation of gene expression reduces the steady-state solution space. *Journal of Theoretical Biology*, **221**, 309-325.
- [23] Covert, M.W. und Palsson, B.O. (2002) Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *Journal of Biological Chemistry*, **277**, 28058-28064.
- [24] Dandekar, T. und Sauerborn, R. (2002) Comparative genome analysis and pathway reconstruction. *Pharmacogenomics*, **3**, 245-256.

- [25] Dandekar, T., Schuster, S., Snel, B., Huynen, M. und Bork, P. (1999) Pathway alignment: Application to the comparative analysis of glycolytic enzymes. *Biochemical Journal*, **343**, 115-124.
- [26] Dauner, M. und Sauer, U. (2001) Stoichiometric growth model for riboflavin-producing *Bacillus subtilis*. *Biotechnology and Bioengineering*, **76**, 132-143.
- [27] Edwards J.S., Ibarra, R.U. und Palsson, B.O. (2001) In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature Biotechnology*, **19**, 125-130.
- [28] Edwards, J.S. und Palsson, B.O. (1998) How will bioinformatics influence metabolic engineering? *Biotechnology and Bioengineering* **58**, 162-169.
- [29] Edwards, J.S. und Palsson, B.O. (2000) The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proceedings of the National Academy of Science USA*, **97**, 5528-5533.
- [30] Eiswirth, M., Freund, A. und Ross, J. (1991) Operational procedure toward the classification of chemical oscillations. *Journal of Physical Chemistry*, **95**, 1294-1299.
- [31] Eraso, J.M. und Kaplan, S. (2000) From redox flow to gene regulation: Role of the PrfC protein of *Rhodospirillum rubrum* 2.4.1. *Biochemistry*, **39**, 2052-2062.
- [32] Eraso, J.M. und Kaplan, S. (2002) Redox flow as an instrument of gene regulation. *Methods in Enzymology*, **348**, 216-229.
- [33] Falcone, D.L. und Tabita, F.R. (1993) Complementation analysis and regulation of CO₂ fixation gene expression in a ribulose-1,5-bisphosphate carboxylase-oxygenase deletion strain of *Rhodospirillum rubrum*. *Journal of Bacteriology*, **175**, 5066-5077.
- [34] Fard, N.S. (1997) Determination of minimal cut sets of a complex fault tree. *Computers & Industrial Engineering*, **33**, 59-62.
- [35] Favier-Teodorescu, L., Cornet, J-F. und Dussap, C.G. (2003) Modelling continuous culture of *Rhodospirillum rubrum* in photobioreactor under light limited conditions. *Biotechnology Letters*, **25**, 359-364.
- [36] Feinberg, M. (1987) Chemical reaction network structure and the stability of complex isothermal reactors – I. The deficiency zero and deficiency one theorems. *Chemical Engineering Science*, **43**, 1-25.
- [37] Feinberg, M. (1995) The existence and uniqueness of steady states for a class of chemical reaction networks. *Archive for Rational Mechanics and Analysis*, **132**, 311-370.
- [38] Fischer, E. und Sauer, U. (2003) Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS. *European Journal of Biochemistry*, **270**, 880-991.
- [39] Föllinger, O. (1994) *Regelungstechnik. Einführung in die Methoden und ihre Anwendung*. Hüthig Buch Verlag, Heidelberg.
- [40] Franklin, J. (1980) *Methods of Mathematical Economics*. Springer Verlag, New York.
- [41] Fuller, R.C. (1995) Polyesters and photosynthetic bacteria. In: Blankenship, R.E., Madigan, M.T. und Bauer, C.E., editors. *Anoxygenic photosynthetic bacteria*. Kluwer Academic Publishers, Dordrecht, pp. 1245-1256.
- [42] Gagneur, J., Jackson, D.B. und Casari, G. (2003) Hierarchical analysis of dependency in metabolic networks. *Bioinformatics*, **19**, 1027-1034.
- [43] Garcia, A.P., Venturoli, G., Gad'on, N., Fernandez-Velasco, Melandri, B.A. und Drews, G. (1987) The adaption of the electron transfer chain of *Rhodospseudomonas capsulata* to different light intensities. *Biochimica et Biophysica Acta*, **890**, 335-345.
- [44] Genoud, T. und Metraux, J-P. (1999) Crosstalk in plant cell signaling: structure and function of the genetic network. *Trends in Plant Science*, **4**, 503-507.
- [45] Georgellis, D., Kwon, O. und Lin, E.C.C. (2001) Quinones as the redox signal for the Arc two-component system of bacteria. *Science*, **292**, 2314-2316.
- [46] Ghosh, R., Hardmeyer, A., Thoenen, L. und Bachofen, R. (1994) Optimization of the Sistrom culture medium for large-scale batch cultivation of *Rhodospirillum rubrum* under semiaerobic conditions with maximal yield of photosynthetic membranes. *Applied and Environmental Microbiology*, **60**, 1698-1700.
- [47] Gilles, E.D. (2002) Regelung – Schlüssel zum Verständnis biologischer Systeme. *at - Automatisierungstechnik*, **50**, 7-17.
- [48] Ginkel, M., Kremling, A., Nutsch, T., Rehner, R. und Gilles, E.D. (2003) Modular modeling of cellular systems with ProMoT/Diva. *Bioinformatics*, **19**, 1169-76.

- [49] Gübel, F. (1978) Quantum efficiencies of growth. In: Clayton, R.K. and Sistrom, W.R., editors. *The photosynthetic bacteria*. Plenum Press, New York, pp. 908-925.
- [50] Goto S., Okuno Y., Hattori M., Nishioka T. und Kanehisa M. (2002) LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acids Research*, **30**, 402-404.
- [51] Gottschalk, G. (1986) *Bacterial Metabolism*. Springer-Verlag, New York.
- [52] Grammel, H. (2003) Unveröffentlichte Daten, Max-Planck-Institut für Dynamik komplexer technischer Systeme, Magdeburg.
- [53] Grammel, H., Gilles, E.D und Ghosh, R. (2003) Microaerophilic cooperation of reductive and oxidative pathways allows maximal photosynthetic membrane biosynthesis in *Rhodospirillum rubrum*. *Applied and Environmental Microbiology*, **69**, 6577-6586.
- [54] Gross, J.L. und Yellen, J. (2003) *Handbook of graph theory*. CRC press, New York.
- [55] Happel, J. und Sellers, P.H. (1989) The characterization of complex systems of chemical reactions. *Chemical Engineering Communications*, **83**, 221-240.
- [56] Heinrich, R. und Schuster, S. (1996) *The Regulation of Cellular Systems*. Chapman & Hall, New York.
- [57] Holme, P., Huss, M. und Jeong, H. (2003) Subnetwork hierarchies of biochemical pathways. *Bioinformatics*, **19**, 532-538.
- [58] Holt, S. C. und Marr, A.G. (1965) Effect of light intensity on the formation of intracytoplasmic membrane in *Rhodospirillum rubrum*. *Journal of Bacteriology*, **89**, 1421-1429.
- [59] Hucka, M. et al. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, **19**, 524-31.
- [60] Hunter, C.M. (1995) Genetic manipulation of the antenna complexes of purple bacteria. In: Blankenship, R.E., Madigan, M.T. und Bauer, C.E., editors. *Anoxygenic photosynthetic bacteria*. Kluwer Academic Publishers, Dordrecht, pp. 473-501.
- [61] Ibarra, R.I., Edwards, J.S. und Palsson, B.O. (2002) *Escherichia coli* undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*, **420**, 186-189.
- [62] Imhoff, J.F. und Bias-Imhoff, U. (1995) Lipids, quinones and fatty acids of anoxygenic phototrophic bacteria. In: Blankenship, R.E., Madigan, M.T. und Bauer, C.E., editors. *Anoxygenic photosynthetic bacteria*. Kluwer Academic Publishers, Dordrecht, pp. 179-205.
- [63] Imhoff, J.I. (1995) Taxonomy and physiology of phototrophic purple bacteria and green sulfur bacteria. In: Blankenship, R.E., Madigan, M.T. und Bauer, C.E., editors. *Anoxygenic photosynthetic bacteria*. Kluwer Academic Publishers, Dordrecht, pp. 1-15.
- [64] Ivanovsky, R.N., Krasilnikova, E.N. und Berg, I.A. (1997) The mechanism of acetate assimilation in the purple nonsulfur bacterium *Rhodospirillum rubrum* lacking isocitrate lyase. *Microbiology*, **66**, 621-626.
- [65] Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N. und Barabasi, A.L. (2000) The large-scale organization of metabolic networks. *Nature*, **407**, 651-654.
- [66] Joshi, H.M. und Tabita, F.R. (1996) A global two component signal transduction system that integrates the control of photosynthesis, carbon dioxide assimilation, and nitrogen fixation. *Proceedings of the National Academy of Science USA*, **93**, 14515-14520.
- [67] Joshi, H.M. und Tabita, F.R. (2000) Induction of carbon monoxide dehydrogenase to facilitate redox balancing in a ribulose biphosphate carboxylase/oxygenase-deficient mutant strain of *Rhodospirillum rubrum*. *Archives of Microbiology*, **173**, 193-199.
- [68] Kanehisa, M. und Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, **28**, 27-30.
- [69] Kaplan, S. (1978) Control and kinetics of photosynthetic membrane development. In: Clayton, R.K. und Sistrom, W.R., editors. *The photosynthetic bacteria*. Plenum Press, New York, pp. 808-839.
- [70] Karp, P.D., Riley, M., Paley, S.M. und Pellegrini-Toole, A. (2002) The MetaCyc Database. *Nucleic Acids Research*, **30**, 59-61.
- [71] Kauffman, K.J., Prakash, P. und Edwards, J.S. (2003) Advances in flux balance analysis. *Current Opinions in Biotechnology*, **14**, 491-496.
- [72] Kitano, H. (2002) Systems biology: a brief overview. *Science*, **295**, 1662-1664.
- [73] Klamm, S. und Gilles, E.D. (2004) Minimal cut sets in biochemical reaction networks. *Bioinformatics*, **20**, 226-234.

- [74] Klamt, S. und Stelling, J. (2002) Combinatorial complexity of pathway analysis in metabolic networks. *Molecular Biology Reports*, **29**, 233-236.
- [75] Klamt, S. und Stelling, J. (2003) Two approaches for metabolic pathway analysis? *Trends in Biotechnology*, **21**, 64-69.
- [76] Klamt, S., Schuster, S. und Gilles, E.D. (2002) Calculability analysis in underdetermined metabolic networks illustrated by a model of the central metabolism in purple nonsulfur bacteria. *Biotechnology and Bioengineering*, **77**, 734-751.
- [77] Klamt, S., Stelling, J., Ginkel, M. und Gilles, E.D. (2003) FluxAnalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps. *Bioinformatics*, **19**, 261-269.
- [78] Korzeniewski, B. (1996) Simulation of oxidative phosphorylation in hepatocytes. *Biophysical Chemistry*, **58**, 215-224.
- [79] Kremling, A., Bettenbrock K., Laube, B., Jahreis, K., Lengeler, J.W. und Gilles, E.D. (2001) The organization of metabolic reaction networks. III. Application for diauxic growth on glucose and lactose. *Metabolic Engineering*, **3**, 362-379.
- [80] Kremling, A., Klamt, S., Ginkel, M. und Gilles, E.D. (2004) Werkbench zur Modellbildung, Simulation und Analyse zellulärer Systeme. *it - Information Technology*, **46**, 12-19.
- [81] Le Novère, N. und Shimizu, T.S. (2001) STOCHSIM: modelling of stochastic biomolecular processes. *Bioinformatics*, **17**, 575-576.
- [82] Leiser, J. und Blum, J.J. (1987) On the analysis of substrate cycles in large metabolic systems. *Cell Biophysics*, **11**, 123-138.
- [83] Lengeler, J.W., Drews, G. und Schlegel, H.G. (1999) *Biology of the prokaryotes*. Thieme Verlag, Stuttgart.
- [84] Liao, J.C., Hou, S. und Chao, Y. (1996) Pathway analysis, engineering, and physiological considerations for redirecting central metabolism. *Biotechnology and Bioengineering*, **52**, 129-140.
- [85] Loew, L.M. und Schaff, J.C. (2001) The Virtual Cell: a software environment for computational cell biology. *Trends in Biotechnology*, **19**, 401-406.
- [86] Lyngstadaas, A., Sprenger, G.A. und Boye, E. (1998) Impaired growth of an *Escherichia coli* *rpe* mutant lacking ribulose-5-phosphate epimerase activity. *Biochimica et Biophysica Acta*, **1381**, 319-330.
- [87] Ma, H. und Zeng, A. (2003) Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics*, **19**, 270-277.
- [88] Masuda, S. und Bauer, C.E. (2002) AppA is a blue light photoreceptor that antirepresses photosynthesis gene expression in *Rhodobacter sphaeroides*. *Cell*, **110**, 613-623.
- [89] Mavrouniotis, M. L., Stephanopoulos, G. und Stephanopoulos, G. (1990) Computer-aided synthesis of biochemical pathways. *Biotechnology and Bioengineering*, **36**, 1119-1132.
- [90] McEwan, A.G. (1994) Photosynthetic electron transport and anaerobic metabolism in purple nonsulfur phototrophic bacteria. *Antonie van Leeuwenhoek*, **66**, 151-164.
- [91] Mendes, P. (1997) Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends in Biochemical Sciences*, **22**, 361-363.
- [92] Michal, G. (1999) *Biochemical Pathways*. Spektrum Akademischer Verlag, Heidelberg.
- [93] Morohashi, M., Winn, A.E., Boersik, M.T., Bolouri, H., Doyle, J. und Kitano H. (2002) Robustness as a measure of plausibility in models of biochemical networks. *Journal of Theoretical Biology*, **216**, 19-30.
- [94] Neidhardt, F.C., Ingraham, J.L. und Schaechter, M. (1990) *Physiology of the bacterial cell: A molecular approach*. Sinauer Associates, Sunderland.
- [95] Newman, M.E.J. (2003) The structure and function of complex networks. *SIAM Review*, **45**, 167-256.
- [96] Nicholls, D.G. und Ferguson, S.J. (2002) *Bioenergetics 3*. Academic Press, London.
- [97] Nielsen, J. (1998) Metabolic Engineering: Techniques for analysis of targets for genetic manipulations. *Biotechnology and Bioengineering*, **58**, 125-132.
- [98] Nissen, T.L., Schulze, U., Nielsen, J. und Villadsen, J. (1997) Flux distributions in anaerobic, glucose-limited continuous cultures of *Saccharomyces cerevisiae*. *Microbiology*, **143**, 203-218.
- [99] Noble, B. und Daniel, J.W. (1988) *Applied linear Algebra*. Prentice-Hall, Englewood Cliffs.

- [100] Noorman, H.J., Romeln, B., Luyben, K.Ch.A.M. und Heijnen, J.J. (1996). Classification, error detection, and reconciliation of process information in complex biochemical systems. *Biotechnology and Bioengineering*, **49**, 364-376.
- [101] Nuño, J.C., Sanchez-Valdenbro, I., Perez-Iratxeta, C., Melendez-Hervia, E. und Montero, F. (1997) Network organization of cell metabolism: monosaccharide interconversion. *Biochemical Journal*, **324**, 104-111.
- [102] Oh, J. und Kaplan, S. (2000) Redox signaling: globalization of gene expression. *The EMBO Journal*, **19**, 4237-4247.
- [103] Oh, J. und Kaplan, S. (2001) Generalized approach to the regulation and integration of gene expression. *Molecular Microbiology*, **39**, 1116-1123.
- [104] Oh, M.-K. und Liao, J.C. (2000) DNA microarray detection of metabolic responses to protein overproduction in *Escherichia coli*. *Metabolic Engineering*, **2**, 201-209.
- [105] Oh, M.-K. und Liao, J.C. (2002) Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*. *Biotechnology Progress*, **16**, 278-286.
- [106] Oh, M.-K., Rohlin, L., Kao, K.C. und Liao, J.C. (2002) Global expression profiling of acetate grown *Escherichia coli*. *Journal of Biological Chemistry*, **277**, 13175-13183.
- [107] Okamura, M.Y. und Feher, G. (1995) Proton-coupled electron-transfer reactions of Q_B in reaction centers from photosynthetic bacteria. In: Blankenship, R.E., Madigan, M.T. und Bauer, C.E., editors. *Anoxygenic photosynthetic bacteria*. Kluwer Academic Publishers, Dordrecht, pp. 577-594.
- [108] Ormerod, J.G. (1956). The use of radioactive carbon dioxide in the measurement of carbon dioxide fixation in *Rhodospirillum rubrum*. *Biochemical Journal*, **64**, 373-380.
- [109] Ormerod, J.G. und Sirevag, R. (1983). Essential aspects of carbon metabolism. In: Ormerod, J.G., editor. *The phototrophic bacteria: Anaerobic life in the light*. Blackwell Scientific Publications, Oxford, pp 100-119.
- [110] Overbeek, R., Larsen, N., Pusch, G.D., D'Souza, M., Selkov, E. Jr., Kyrpides, N., Fonstein, M., Maltsev, N. und Selkov, E. (2000) WIT: integrated system for high-throughput genome sequence analysis and metabolic reconstruction. *Nucleic Acids Research*, **28**, 123-125.
- [111] Oxley, J.G. (1992) *Matroid Theory*. Oxford Press, Oxford.
- [112] Palsson, B.O. (2000) The challenges of in silico biology. *Nature Biotechnology*, **18**, 1147-50
- [113] Papin, J.A., Price, N.D. und Palsson, B.O. (2002) Extreme pathway lengths and reaction participation in genome-scale metabolic networks. *Genome Research*, **12**, 1889-1900.
- [114] Papin, J.A., Price, N.D., Edwards, J.E. und Palsson, B.O. (2002) The genome scale metabolic extreme pathway structure in *Haemophilus influenzae* shows significant network redundancy. *Journal of Theoretical Biology*, **215**, 67-82.
- [115] Papin, J.A., Price, N.D., Wiback, S.J., Fell, D.A. und Palsson, B.O. (2003) Metabolic pathways in the post-genome era. *Trends in Biological Sciences*, **28**, 250-258.
- [116] Pfeiffer, T., Sánchez-Valdenbro, I., Nuño, J.C., Montero, F. und Schuster, S. (1999) METATOOL: For studying metabolic networks. *Bioinformatics*, **15**, 251-257.
- [117] Poolman, M.G., Fell, D.A. und Raines, C.A. (2003) Elementary modes analysis of photosynthate metabolism in the chloroplast stroma. *European Journal of Biochemistry*, **270**, 430-439.
- [118] Pramanik, J. und Keasling, J.D. (1997) Stoichiometric model of *Escherichia coli* metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnology and Bioengineering*, **56**, 398-421.
- [119] Price, N.D., Famili, L., Beard, D.A. und Palsson, B.O. (2002) Extreme pathways and Kirchhoff's second law. *Biophysical Journal*, **83**, 2879-2882.
- [120] Price, N.D., Papin, J.A. und Palsson, B.O. (2002) Determination of redundancy and systems properties of *Helicobacter pylori*'s metabolic network using extreme pathway analysis. *Genome Research*, **12**, 760-769.
- [121] Price, N.D., Papin, J.A., Schilling, C.H. und Palsson, B.O. (2003) Genome-scale microbial in silico models: the constraints-based approach. *Trends in Biotechnology*, **21**, 162-169.
- [122] Reed, J.L. und Palsson, B.O. (2003) Thirteen years of building constraint-based in silico models of *Escherichia coli*. *Journal of Bacteriology*, **185**, 2692-2699.
- [123] Rettinger, A. (1995) Parameteroptimierung mittels Evolutionsstrategien. Diplomarbeit am Institut für Systemdynamik und Regelungstechnik, Universität Stuttgart.
- [124] Richard, P., Teusnik, B., Hemker, M.B., Van Dam, K. und Westerhoff, H.V. (1996) Sustained oscillations in free-energy state and hexose phosphates in yeast. *Yeast*, **12**, 731-740.

- [125] Richardson, D.J., King, G.P., Kelly, D.J., McEwan, A.G., Ferguson, S.J. and Jackson, J.B. (1988) The role of auxiliary oxidants in maintaining redox balance during photosynthetic growth of *Rhodospirillum rubrum* on propionate and butyrate. *Archives of Microbiology*, **150**, 131-137.
- [126] Rockafellar, R.T. (1970) *Convex analysis*. Princeton University Press, Princeton.
- [127] Rohwer, J.M. and Botha, P.C. (2001) Analysis of sucrose accumulation in the sugar cane culm on the basis of in vitro kinetic data. *Biochemical Journal*, **358**, 437-445.
- [128] Samatova, N.P., Geist, A., Ostrouchov, G., and Melchko, A.V. (2002) Parallel out-of-core algorithm for genome-scale enumeration of metabolic systematic pathways. *Proceedings of the First IEEE Workshop on High Performance Computational Biology (HCOMB2002)*, Ft. Lauderdale.
- [129] Sauro, H.M. (2000) Jarnac: A system for interactive metabolic analysis. In: Hofmeyr, J., Rohwer, J.M. and Snoep, J.L., editors. *Animating the cellular Map: Proceedings of the 9th International Meeting on BioThermoKinetics*. Stellenbosch University Press, Stellenbosch.
- [130] Schilling, C.H. and Palsson, B.O. (2000) Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. *Journal of Theoretical Biology*, **203**, 249-83.
- [131] Schilling, C.H., Edwards J.S., Letscher, D. and Palsson, B.O. (2001) Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems. *Biotechnology and Bioengineering*, **71**, 286-306.
- [132] Schilling, C.H., Letscher, D. and Palsson, B.O. (2000) Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *Journal of Theoretical Biology*, **203**, 229-248.
- [133] Schilling, C.H., Schuster, S., Palsson, B.O. and Heinrich, R. (1999) Metabolic pathway analysis: Basic concepts and scientific applications in the post-genomic era. *Biotechnology Progress*, **15**, 296-303.
- [134] Schlegel, H.G. (1992) *Allgemeine Mikrobiologie*. Thieme Verlag, Stuttgart.
- [135] Schön, G. (1969) Der Einfluß der Kulturbedingungen auf den ATP-, ADP- und AMP-Spiegel bei *Rhodospirillum rubrum*. *Arch. Mikrobiol.*, **66**, 348-364.
- [136] Schön, G. (1971) Der Einfluß der Kulturbedingungen auf den Nicotin-Adenin-Dinucleotid(phosphat)-Gehalt in Zellen von *Rhodospirillum rubrum*. *Arch. Mikrobiol.*, **68**, 147-163.
- [137] Schön, G. (1972) Substratverwertung und Bacteriochlorophyll-Synthese in anaerober Dunkelkultur von *Rhodospirillum rubrum*: I. Abhängigkeit der Bacteriochlorophyll-Synthese von der Substratkonzentration und dem Elektronenacceptor. *Zbl Bakt Hyg. I. Abt A*, **220**, 380-386.
- [138] Schuster, S. (2003) Metabolic pathway analysis in biotechnology. In: Kholodenko, B.N. and Westerhoff, H.V., editors. *Metabolic engineering in the post-genome era*. Horizon Scientific, Wymondham, pp. 181-208.
- [139] Schuster, S. and Hilgetag, C. (1994) On elementary flux modes in biochemical reaction systems at steady state. *Journal of Biological Systems*, **2**, 165-182.
- [140] Schuster, S., Dandekar, T. and Fell, D.A. (1999) Detection of elementary flux modes in biochemical networks: A promising tool for pathway analysis and metabolic engineering. *Trends in Biotechnology*, **17**, 53-60.
- [141] Schuster, S., Fell, D. and Dandekar, T. (2000) A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnology*, **18**, 326-332.
- [142] Schuster, S., Hilgetag, J.C., Woods, J.H. and Fell, D.A. (2002) Reaction routes in biochemical reaction systems: Algebraic properties, validated calculation procedure and example from nucleotide metabolism. *Journal of Mathematical Biology*, **45**, 153-181.
- [143] Schuster, S., Klamt, S., Weckwerth, W., Moldenhauer, F. and Pfeiffer, T. (2002) Use of network analysis of metabolic systems in bioengineering. *BioProcess and Biosystems Engineering*, **24**, 363-372.
- [144] Schuster, S., Ouhabi, R. Rigoulet, M. and Mazat, J.-P. (1998) Modelling the interrelation between the transmembrane potential and pH difference across membranes with electrogenic proton transport upon build-up of proton-motive force. *Bioelectrochemistry and Bioenergetics*, **43**, 181-192.

- [145] Schuster, S., Pfeiffer, T., Moldenhauer, P., Koch, I. und Dandekar, T. (2003) Exploring the pathway structure of metabolism: decomposition into subnetworks and application to *Mycoplama pneumoniae*. *Bioinformatics*, **18**, 351-61.
- [146] Seressiotis, A. und Bailey, J.E. (1988) MPS: An artificially intelligent software system for the analysis and synthesis of metabolic pathways. *Biotechnology and Bioengineering*, **31**, 587-602.
- [147] Sinnamoni, R.M. und Andrews, J.D. (1997) New approaches to evaluating fault trees. *Reliability Engineering & System Safety*, **58**, 89-96.
- [148] Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S. und Gilles, E.D. (2002) Metabolic network structure determines key aspects of functionality and regulation. *Nature*, **420**, 190-193.
- [149] Stephanopoulos, G.N., Aristidou, A.A. und Nielsen, J. (1998) *Metabolic Engineering*. Academic Press, San Diego.
- [150] Strang, G. (1980) *Linear algebra and its applications*. New York, Academic Press.
- [151] Swem, D.L. und Bauer, C.E. (2002) Coordination of ubiquinol oxidase und cytochrome cbb₃ oxidase expression by multiple regulators in *Rhodobacter capsulatus*. *Journal of Bacteriology*, **184**, 2815-2820.
- [152] Swem, L.R., Elsem, S., Bird, T.H., Swem, D.L., Koch, H.-G., Myllykallio, H., Daldal, F. und Bauer, C.E. (2001) The RegB/RegA two-component regulatory system controls synthesis of photosynthesis and respiratory electron transfer components in *Rhodobacter capsulatus*. *Journal of Molecular Biology*, **309**, 121-138.
- [153] Szalkai, I. (2000) A new general algorithmic method in reaction syntheses using linear algebra. *Journal of Mathematical Chemistry*, **28**, 1-34.
- [154] Szyperski, T. (1998) ¹³C-NMR, MS und metabolic flux balancing in biotechnology research. *Quarterly Reviews of Biophysics*, **31**, 41-106.
- [155] Tabita, F.R. (1995) The biochemistry and metabolic regulation of carbon metabolism and CO₂-fixation in purple bacteria. In: Blankenship, R.E., Madigan, M.T. und Bauer, C.E., editors. *Anoxygenic photosynthetic bacteria*. Kluwer Academic Publishers, Dordrecht, pp. 885-914.
- [156] Theobald, U., Maillinger, W., Baltes, M., Rizzi, M. und Reuss, M. (1997) In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnology and Bioengineering*, **55**, 305-316.
- [157] Tomita, M., Hashimoto, K., Takahashi, K., Shimizu, T.S., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J.C. und Hutchison, C.A.3rd. (1999) E-CELL: software environment for whole-cell simulation. *Bioinformatics*, **15**, 72-84.
- [158] Vallino, J.J. und Stephanopoulos, G. (1993) Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnology and Bioengineering*, **41**, 633-646.
- [159] Van der Heijden, R.T.J.M., Heijnen, J.J., Hellinga, C., Romein, B. und Luyben, K.Ch.A.M. (1994) Linear constraint relations in biochemical reaction systems: I. Classification of the calculability and the balanceability of conversion rates. *Biotechnology and Bioengineering*, **43**, 3-10.
- [160] Van der Heijden, R.T.J.M., Romein, B., Heijnen, J.J., Hellinga, C. und Luyben, K.Ch.A.M. (1994) Linear constraint relations in biochemical reaction systems: II. Diagnosis and estimation of gross errors. *Biotechnology and Bioengineering*, **43**, 11-20.
- [161] Van der Rest, M., Frank, C. und Molebaer, D. (2000) Functions of the membrane-associated and cytoplasmic malate dehydrogenases in the citric acid cycle of *Escherichia coli*. *Journal of Bacteriology*, **182**, 6892-6899.
- [162] Van Dien, S.J. und Lidstrom, M.E. (2002) Stoichiometric model for evaluating the metabolic capabilities of the facultative methylotroph *Methylobacterium extorquens* AM1, with application to reconstruction of C(3) and C(4) metabolism. *Biotechnology and Bioengineering*, **78**, 296-312.
- [163] Varela, J. und Ramirez, J.M. (1990) Oxygen-linked electron transfer and energy conservation in *Rhodospirillum rubrum*. In: Drews, G. und Dawes, E.A., editors. *Molecular biology of membrane-bound complexes in phototrophic bacteria*. Academic Press, New York, pp. 443-452.
- [164] Varma, A., Boesch, B.W. und Palsson, B.O. (1993) Biochemical production capabilities of *Escherichia coli*. *Biotechnology and Bioengineering*, **42**, 59-73.
- [165] Volkman, P. (1994) *Taschenbuch Elektrotechnik + Elektronik. Band 1: Grundlagen*. VDE-Verlag, Berlin.
- [166] von Dassow, G., Meir, E., Munro, E. und Odell, G.M. (2000) The segment polarity network is a robust developmental module. *Nature*, **406**, 188-191.

- [167] Wagner A. and Fell D.A. (2001) The small world inside large metabolic networks. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **268**, 1803-10.
- [168] Wang, X., Falcone, D.L. and Tabita, F.R. (1993) Reductive pentose phosphate-independent CO₂ fixation in *Rhodobacter sphaeroides* and evidence that ribulose biphosphate carboxylase / oxygenase activity serves to maintain the redox balance of the cell. *Journal of Bacteriology*, **175**, 3372-3379.
- [169] Wiechert, W. (2001) ¹³C Metabolic Flux Analysis. *Metabolic Engineering*, **3**, 195-206.
- [170] Wiechert, W. (2002) Modeling and simulation: tools for metabolic engineering. *Journal of Biotechnology*, **94**, 37-64.
- [171] Wolf, J., Passarge, J., Somsen, O.J.G., Snoep, J.L., Heinrich, R. und Westerhoff, H.V. (2000) Transduction of intracellular and intercellular dynamics in yeast glycolytic oscillations. *Biophysical Journal*, **78**, 1145-1153.
- [172] Wraight, C.A., Cogdell, R.J. and Chance, B. (1978) Ion transport and electrochemical gradients in photosynthetic bacteria. In: Clayton, R.K. and Sistrom, W.R., editors. *The photosynthetic bacteria*. Plenum Press, New York, pp. 908-925.
- [173] Zannoni, D. (1995) Aerobic and anaerobic electron transport chains in anoxygenic phototrophic bacteria. In: Blankenship, R.E., Madigan, M.T. und Bauer, C.E., editors. *Anoxygenic photosynthetic bacteria*. Kluwer Academic Publishers, Dordrecht, pp. 949-971.
- [174] Zeigarnik, A.V. (1991) On hypercycles and hypercircuits in hypergraphs. In: Hansen, P., Fowler, P. W. und Zheng, M., editors. *Discrete Mathematical Chemistry, vol. 51 of DIMACS series in discrete mathematics and theoretical computer science*, pp. 377-383.
- [175] Zeilstra-Ryalls, J.H., Gabbert, K., Mouncey, N.J., Kaplan, S. und Kranz, R.G. (1997) Analysis of the *fruL* gene and its function in *Rhodobacter capsulatus*. *Journal of Bacteriology*, **179**, 7264-7273.