

Methoden zur integrierten Analyse metabolischer Netzwerke unter stationären und instationären Bedingungen

Sebastian Aljoscha Wahl

Forschungszentrum Jülich GmbH
Institut für Biotechnologie-2 (IBT-2)

Methoden zur integrierten Analyse metabolischer Netzwerke unter stationären und instationären Bedingungen

Sebastian Aljoscha Wahl

Schriften des Forschungszentrums Jülich
Reihe Gesundheit / Health

Band/Volume 1

ISSN 1866-1785

ISBN 978-3-89336-506-7

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte Bibliografische Daten
sind im Internet über <<http://dnb.d-nb.de>> abrufbar.

Herausgeber und Vertrieb: Forschungszentrum Jülich GmbH
Zentralbibliothek, Verlag
D-52425 Jülich
Telefon (02461) 61-5368 · Telefax (02461) 61-6103
e-mail: zb-publikation@fz-juelich.de
Internet: <http://www.fz-juelich.de/zb>

Umschlaggestaltung: Grafische Medien, Forschungszentrum Jülich GmbH

Druck: Grafische Medien, Forschungszentrum Jülich GmbH

Copyright: Forschungszentrum Jülich 2008

Schriften des Forschungszentrums Jülich
Reihe Gesundheit / Health Band / Volume 1

D 467 (Diss., Siegen, Univ., 2007)

ISSN 1866-1785

ISBN 978-3-89336-506-7

Vollständig frei verfügbar im Internet auf dem Jülicher Open Access Server (JUWEL)
unter <http://www.fz-juelich.de/zb/juwel>

Alle Rechte vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (Druck, Fotokopie oder in einem anderen Verfahren) ohne schriftliche Genehmigung des Verlages reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Methoden zur integrierten Analyse metabolischer Netzwerke unter stationären und instationären Bedingungen

Abstract

Das Überleben biologischer Organismen beruht zu einem großen Teil auf deren Stoffwechselleistungen. In den Zellen laufen dabei unzählige durch Enzyme katalysierte Stoffumwandlungen (biochemischen Reaktionen) ab, die sowohl Energie als auch Material zur Synthese von Biomasse und Produkten bereitstellen. Diese Reaktionen werden in Modellen in Form von metabolischen Netzwerken dargestellt. Deren Modellierung, Simulation und Analyse gewinnt vor allem im Metabolic Engineering stetig an Bedeutung und die metabolischen Netzwerke werden immer detaillierter und umfangreicher. Sie sind deshalb nur noch rechnerunterstützt handhabbar.

Metabolische Untersuchungen werden je nach Fragestellung unter verschiedenen experimentellen Bedingungen durchgeführt. Eine erste Unterteilung wird aufgrund des Zustands vorgenommen: 1. Stationäre Experimente werden zur Quantifizierung einer Stoffflusslage unter den gegebenen Bedingungen (z.B. eines Produktionsprozesses) eingesetzt. 2. Instationäre Experimente zielen auf eine kinetische Untersuchung der enzymatischen Reaktionen ab und können so auch für Vorhersagen unter anderen Umgebungsbedingungen genutzt werden. Je nach Fragestellung und den verfügbaren Messdaten unterscheiden sich auch die erstellten Modelle. In der Literatur ist bislang keine auf alle Systeme anwendbare Systematik der Modellierung dargestellt. Dabei gibt es zwischen den verschiedenen metabolischen Modellen (stationär und kinetisch) einige Überschneidungen. Sie basieren zu einem großen Teil auf vergleichbaren Modellierungsansätzen. Auch ähneln sich die im Verlauf der Analyse benötigten Hilfsmittel. In dieser Arbeit wurde deshalb eine einheitliche Systematik eingeführt, die auf alle hier untersuchten Systeme anwendbar ist.

Bei der stationären Betrachtung metabolischer Netzwerke ergeben sich meist relativ einfache Gleichungssysteme. Für deren Analyse sind bereits eine Vielzahl von Methoden etabliert. Allerdings ist festzustellen, dass entwickelte theoretische Konzepte selten zur Interpretation experimenteller Beobachtungen herangezogen werden. In Kapitel 11 wird deshalb eine erweiterte Methode vorgestellt, die theoretische Ansätze zur Interpretation gemessener Flusslagen nutzt. Die Erstellung und Analyse kinetischer Modelle ist bereits wesentlich aufwendiger. Insbesondere sind für die Modellierung notwendige Informationen über Enzymmechanismen und metabolische Regulationen nur teilweise vorhanden oder sogar widersprüchlich. Im Rahmen der Arbeit ist deshalb ein Vorgehen entwickelt worden, dass schon vor der eigentlichen Modellierung einsetzt und ermöglicht aus den vorhandenen Messdatenreihen Hypothesen zu extrahieren. Anhand der Hypothesen wird dann eine Vielzahl möglicher Modelle erstellt und an die Messdaten angepasst. Hierbei zeigt sich, dass unterschiedliche Modelle die experimentellen Daten vergleichbar gut wiedergeben können. Glücklicherweise kann beobachtet werden, dass diejenigen Modelle welche die Messdaten gut reproduzieren, einige gemeinsame Eigenschaften und Mechanismen aufweisen. Diese Überschneidungen geben deutliche Hinweise auf wichtige Modellbestandteile. Für einige Anwendungen wie z.B. die gezielte Verbesserung von Produktionsstämmen anhand von Modellvorhersagen werden allerdings genauere Modelle, vor allem auch die Werte der kinetischen Parameter benötigt.

In den letzten Kapiteln der Arbeit werden verschiedene Ansätze diskutiert, wie detailliertere Modelle identifiziert werden können. Zum Einen wird vorgeschlagen, mehrere Experimente mit gewissen genetischen Veränderungen durchzuführen und anhand eines gekoppelten Modells auszuwerten. Mit diesem Ansatz kann eine höhere Vorhersagegenauigkeit der kinetischen Parameter erzielt werden. Zum Anderen wird ein neuer experimenteller Ansatz vorgeschlagen, in dem der Stoffwechsel nicht nur metabolisch sondern auch isotopisch durch die Zugabe von markiertem Kohlenstoff (^{13}C) ausgelenkt wird. Anhand einer Simulationsstudie mit einem kleinen Beispielsystem wird dieser neue, bislang noch nicht untersuchte Zustand analysiert und diskutiert.

Methods for the Integrated Evaluation of Metabolic Networks under Stationary and Non-Stationary Conditions

Abstract

In biological organisms a variety of enzyme catalyzed reactions are taking place that deliver energy as well as precursors for biomass and products. These reactions are modeled by metabolic networks. The simulation and analysis of these networks is getting more important especially for „Metabolic Engineering“ approaches. Current models are very detailed and nowadays can only be handled with computer techniques. Metabolic analyses are performed under different experimental conditions depending on the aim of the studies. Firstly metabolic stationary conditions are applied to quantify metabolic fluxes under e.g. production conditions. Secondly non-stationary conditions that aim at a kinetic characterization of the metabolic interactions. Kinetic models can in contrast to stationary fluxes be used for prediction of further states.

The modeling approaches differ depending on the available measurements. In literature up to now there is no systematic approach that could be applied to the different conditions. Nevertheless there are several overlaps and the approaches are based on the same chemical and physical laws. In case of stationary metabolic analyses relatively simple linear equation systems are used. There are a variety of methods for analysis available. Nevertheless, these theoretical concepts are often not applied for the interpretation of measured data. In the presented work a method is developed that enables to use a theoretic method for the interpretation of a measured flux distribution.

Kinetic models are much more evolved. The information needed for modeling is only partly available and sometimes even contradictorily. Therefore in this work a concept was developed that enables to extract hypotheses on reaction mechanisms from measured time-series data. Based on these hypotheses a variety of possible models are set up and fitted to the data. It was observed that several models are able to reproduce the measured data. Fortunately these models share common characteristics that seem to be essential for the correct reproduction of the data.

Several approaches are discussed how the identification of kinetic parameters can be improved. It is shown that the simultaneous evaluation of experiments with two genetically different strains does improve the statistical accuracy. From a simulative study it is derived that an additional labeling pulse will further increase the statistical accuracy up to a factor of ten.

Inhaltsverzeichnis

| | |
|--|-----------|
| Verwendete Symbole | 5 |
| I Einleitung & Methoden | 11 |
| 1 Metabolic Engineering | 13 |
| 1.1 Entwicklung der weißen Biotechnologie | 13 |
| 1.2 Motivation und Methoden des Metabolic Engineering | 14 |
| 1.3 Modellierung im Metabolic Engineering | 16 |
| 2 Ziele und Gliederung | 19 |
| 2.1 Schwerpunkte der vorliegenden Arbeit | 19 |
| 2.2 Gliederung der Arbeit | 20 |
| 3 Untersuchte biologische Systeme | 25 |
| 3.1 Bakterieller Stoffwechsel und Regulationsmechanismen | 26 |
| 3.2 L-Phenylalanin Produzenten | 29 |
| 4 Experimentelles Vorgehen | 33 |
| 4.1 Überblick der experimentellen Bedingungen | 33 |
| 4.2 Stationäre Stoffflussanalyse (SFA) | 35 |
| 4.3 Stationäre Markierungsexperimente (^{13}C SFA) | 36 |
| 4.4 Stimulus-Response Experimente | 37 |
| 4.5 LC-MS/MS Messtechnik | 39 |
| 5 Modellbildung | 43 |
| 6 Metabolische Netzwerk-Modelle | 49 |
| 6.1 Beispielsystem zur Verdeutlichung der Modellerstellung | 49 |
| 6.2 Massenbilanzen | 49 |
| 6.3 Enzym-Kinetiken | 51 |
| 6.4 Vereinfachte Kinetische Ansätze | 53 |
| 6.5 Bilanzen für ^{13}C Markierungszustände | 56 |

| | | |
|-----------|--|------------|
| 7 | Simulation, Systemanalyse und Statistische Methoden | 61 |
| 7.1 | Schema der Simulation und Systemanalyse | 61 |
| 7.2 | Fehlerübertragung | 65 |
| 8 | Anwendung der verschiedenen Modellansätze | 71 |
| 8.1 | Metabolisch stationäre Modelle | 71 |
| 8.2 | Stöchiometrische Stoffflussanalyse | 71 |
| 8.3 | Metabolische Netzwerk Analyse | 74 |
| 8.4 | ^{13}C Stoffflussanalyse | 74 |
| 8.5 | Instationäre ^{13}C Stoffflussanalyse | 79 |
| 8.6 | Metabolisch instationärer Zustand | 79 |
| 9 | Auswertung mit Visuellen Werkzeugen | 85 |
| 9.1 | Visualisierung von Sensitivitäten | 85 |
| 9.2 | Kovarianz-Matrizen | 86 |
| 9.3 | Multidimensionale Skalierung | 88 |
| II | Stationäre Modellierung und Methoden zur Auswertung | 91 |
| 10 | Auswertung von ^{13}C Markierungsexperimenten | 93 |
| 10.1 | Sensor-Reaktor Konzept | 93 |
| 10.2 | Schätzung der extrazellulären Flüsse | 94 |
| 10.3 | Korrektur der ^{13}C Markierungsmessung | 97 |
| 10.4 | Netzwerk zur Auswertung | 98 |
| 10.5 | Flussschätzungen und statistische Auswertung | 100 |
| 10.6 | Ergebnisse und Diskussion | 103 |
| 10.7 | Möglichkeiten zur Steigerung der Bestimmtheit | 105 |
| 11 | Interpretation stationärer Stoffflusslagen | 107 |
| 11.1 | Sensitivitätsanalyse stationärer Netzwerke | 108 |
| 11.2 | Limitierungs-Potenzial: Reduzierung des theoretischen Maximums | 110 |
| 11.3 | Sensitivität des beschränkenden Flusses | 112 |
| 11.4 | Auswertung der L-Phenylalanin Experimente | 114 |
| 11.5 | Ergebnisse und Diskussion | 118 |

| | |
|--|------------|
| III Kinetische Modellierung und Auswertung | 121 |
| 12 SRE Auswertungsstrategie | 123 |
| 12.1 Auswertungsstrategie | 123 |
| 12.2 Messdaten Aufarbeitung | 126 |
| 12.3 Eingesetzte Spline-Glättung | 129 |
| 12.4 Fokussierte Modellierung | 130 |
| 12.5 Modellfamilien | 131 |
| 12.6 Parameteranpassung | 131 |
| 12.7 Statistische Untersuchung | 132 |
| 12.8 Modell-Ranking | 132 |
| 13 Dynamische Modelle | 133 |
| 13.1 Zeitlich versetzte Korrelationsanalyse der Messdaten | 133 |
| 13.2 Untersuchte Modell Varianten | 136 |
| 13.3 Auswertung der mechanistischen Modellfamilie | 138 |
| 13.4 Modelle mit generischen Michaelis-Menten Termen | 139 |
| 13.5 LinLog Modellfamilie | 139 |
| 14 Statistische Untersuchung | 147 |
| 14.1 Parameterbestimmtheit im mechanistischen Modellierungsansatz | 147 |
| 14.2 Bestimmtheit im LinLog Modellansatz | 149 |
| 14.3 Effektor-Einflüsse | 151 |
| 14.4 Einfluss der Eingangsgößen (Splines) in fokussierten Modellen | 152 |
| IV Erweiterte Auswertung | 157 |
| 15 Steigerung der Parameterbestimmtheit | 159 |
| 15.1 Messung des Steady-State Flusses | 160 |
| 15.2 Gekoppelte Anpassung mehrerer Datensätze | 161 |
| 16 Modell-Analyse und -Konsistenz | 167 |
| 16.1 Abschätzung maximaler Flussraten | 167 |
| 16.2 Vergleich simulierter und notwendiger Flüsse | 168 |
| 16.3 Optimierung der Enzymexpression für den Steady State | 171 |
| 16.4 Einfluss der extrazellulären Konzentrationen | 173 |
| 16.5 Thermodynamische Betrachtung | 174 |

| | | |
|-----------|--|------------|
| V | Metabolisch und Isotopisch Instationäre Modellierung | 179 |
| 17 | Simulationsstudie: Instationäre ¹³C Experimente | 181 |
| 17.1 | Modellaufbau | 181 |
| 17.2 | Untersuchte Szenarien | 183 |
| 17.3 | Ergebnisse aus dem Statistischen Vergleich | 186 |
| 17.4 | Versuchsplanung | 187 |
| 17.5 | Vereinfachte ¹³ C Modelle | 190 |
| 18 | MS/MS Messmodell | 197 |
| 18.1 | Keine Fragmentierung | 198 |
| 18.2 | Mit Fragmentierung in Q2 | 198 |
| 18.3 | Kompakte Darstellung des Messmodells | 200 |
| 18.4 | Massenisotopen-Korrektur für MS/MS Daten | 200 |
| 18.5 | Zusammenfassung | 204 |
| 19 | Zusammenfassung, Diskussion und Ausblick | 205 |
| A | Auswertung der ¹³C Markierungsexperimente | 211 |
| A.1 | Reaktionen und Ergebnisse der Flusschätzungen | 211 |
| B | Anhang zu den Dynamischen Modellen | 217 |
| B.1 | | 222 |
| C | Simulation von dynamischen Markierungsexperimenten | 225 |
| C.1 | Stöchiometrie, Kinetik und C-Atom Transitionen - Spirallus | 225 |

Verwendete Symbole

Abkürzungen - Allgemein

| Abkürzung | Bedeutung |
|------------------|---|
| BFGS | Broyden-Fletcher-Goldfarb-Shanno (Optimierungsverfahren) |
| BTM | Biotrockenmasse |
| CLS | Sensitivität des reduzierten Optimums (Constraint Limitation Sensitivity) |
| <i>E. coli</i> | <i>Escherichia coli</i> |
| GC | Gaschromatographie |
| ESI | Elektrosprayionisierung |
| fbr | Feed-back resistent |
| FQS | Fehlerquadratsumme |
| HPLC | Hochdruck Flüssigkeitschromatographie (High Pressure Liquid Chromatography) |
| IBT | Institut für Biotechnologie der Forschungszentrum Jülich GmbH |
| <i>in silico</i> | (Lat:im Silizium) Am Rechner durchgeführte Experimente (Simulation) |
| <i>in vitro</i> | (Lat: im Reagenzglas) Experimente unter künstlichen Bedingungen |
| <i>in vivo</i> | (Lat: im Leben) Experimente im lebenden System |
| IPTG | Isopropyl- β -D-thiogalactopyrosid |
| LC | Liquid Chromatography |
| LC-MS | Kopplung von Flüssigkeitschromatographie (Liquid Chromatography) und Massenspektroskopie (Mass Spectrometry) |
| LC-MS/MS | LC in Kopplung mit einem Massenspektrometer mit 2 Massenfiltern (z.B. Triplequadrupole MS) |
| LP | Limitierungs Potenzial (Limitation Potential) |
| LSODA | Methode zur numerischen Lösung von Differentialgleichungen (Livermore Solver for Ordinary Differential Equations) |

| Abkürzung | Bedeutung |
|------------|--|
| M3L | XML Format zur Beschreibung metabolischer Netzwerke (Metabolic Modeling Markup Language) |
| MatVis | Visualisierungswerkzeug (Matrix Visualizer) |
| MetVis | Visualisierungswerkzeug (Metabolic Visualizer) |
| MM | Michaelis-Menten Kinetik |
| MMT | Simulationswerkzeug (Metabolic Modeling Tool) |
| <i>m/z</i> | Masse zu Ladungsverhältniss |
| OD | Optische Dichte |
| PPP | Pentosephosphat-Weg (Pentose-Phosphate-Pathway) |
| SBML | XML Format zur Beschreibung metabolischer Netzwerke (Systems Biology Markup Language) |
| SRE | Stimulus-Response-Experiment |
| SSQ | Fehlerquadratsumme (Sum of Squares) |
| TCA | Trikarbonsäure-Zyklus (Tri-Carbon Acid Cycle) |

Abkürzungen - Metabolite

| Abkürzung | Bedeutung | Abkürzung | Bedeutung |
|-----------|---------------------------------------|------------------------------|---|
| 2-PG | 1-Phosphoglycerat | Mal | Malat |
| 3-PG | 3-Phosphoglycerat | OAA | Oxaloacetat |
| 6-PG | 6-Phosphogluconat | NAD | Diphosphopyridinnucleotid, oxidiert |
| AcCoA | Acetyl-Coenzym A | NADH | Diphosphopyridinnucleotid, reduziert |
| ADP | Adenosindiphosphat | NADP | Diphosphopyridinnucleotid-Phosphat, oxidiert |
| aKG | 2-Oxo-Ketoglutarat | NADPH | Diphosphopyridinnucleotid-Phosphat, reduziert |
| AMP | Adenosinmonophosphat | NH ₄ ⁺ | Ammonium |
| ATP | Adenosintriphosphat | P _i | anorganischer Phosphat |
| CHO | Chorismat | PPA | Prephenat |
| Cit | Citrat | PPY | Phenylpyruvat |
| DAHP | 3-Desoxy-Arabino-Heptulose-7-Phosphat | Rib5P | Ribose-5-Phosphat |
| DHS | 3-Dehydroshikimat | Rul5P | Ribulose-5-Phosphat |
| DHQ | 3-Dehydroquinat | RNA | Ribonuclein Säure |
| DNA | Desoxyribonuclein Säure | S3P | 5-Enolpyruvylshikimat-3-Phosphat |
| DPG | Diphosphoglycerat | Shik | Shikimat |
| E4P | Erythrose-4-Phosphat | S7P | Septulose-7-Phosphat |
| EPSP | 5-Enolpyruvylshikimate-3-Phosphat | Succ | Succinat |
| F6P | Fructose-6-Phosphat | Tyr | Tyrosin |
| FBP | Fructose-1,6-Bis-Phosphat | Xyl5P | Xylose-5-Phosphat |
| G6P | Glucose-6-Phosphat | X5P | Summenpool aus Xyl5P, Rib5P, Rul5P |
| GAP | Glyceraldehyd-3-Phosphat | | |
| Glc | Glucose | | |
| Gly | Glycin | | |
| ICit | Isocitrat | | |

Abkürzungen - Enzyme

| Abkürzung | Bedeutung | EC |
|-----------|--|----------|
| AATA | Transaminase für aromatische Aminosäuren | 2.6.1.57 |
| CS | Chorismat-Synthase | 4.6.1.4 |
| CM | Chorismat-Muthase | 5.4.99.5 |
| DAHPS | DAHP-Synthase | 4.1.2.15 |
| DHQS | Dehydroquinat Synthase | 4.6.1.3 |
| DHQDH | Dehydroquinat Dehydratase | 4.2.1.10 |
| EPSPS | EPSP Synthase | 2.5.1.19 |
| GDH | Glucose-6-Dehydrogenase | 1.1.1.49 |
| SDH | Shikimat Dehydrogenase | 1.1.1.25 |
| SK | Shikimat Kinase | 2.7.1.71 |
| PDH | Prephenat Dehydratase | 4.2.1.51 |
| PTS | Phosphotransferase System | - |
| TA | Transaldolase | 2.2.1.2 |
| TK | Transketolase | 2.2.1.1 |

Abkürzungen - Gene

| Abkürzung | Gen |
|-------------|---|
| <i>wt</i> | Wildtyp-Variante des Gens |
| <i>fbr</i> | Feedback-resistente Variante des Gens (gezielt modifiziert) |
| <i>aroA</i> | EPSP Synthase |
| <i>aroB</i> | DHQ Synthase |
| <i>aroC</i> | Chorismat Synthase |
| <i>aroD</i> | DHQ Dehydratase |
| <i>aroE</i> | Shikimat Dehydrogenase |
| <i>aroF</i> | DAHP Synthase (L-Tyrosin sensitiv) |
| <i>aroG</i> | DAHP Synthase (L-Phenylalanin sensitiv) |
| <i>aroH</i> | DAHP Synthase (L-Tryptophan sensitiv) |
| <i>pheA</i> | Chorismat Muthase/Prephenat Dehydratase (L-Phe sensitiv) |
| <i>pps</i> | PEP Synthase |
| <i>tal</i> | Transaldolase |
| <i>tktA</i> | Transketolase |
| <i>tyrA</i> | Chorismat Muthase/Prephenat Dehydratase (L-Tyr sensitiv) |
| <i>tyrB</i> | Transaminase der aromatischen Aminosäuren |

Mathematische Symbole

| Symbol | Einheit | Bezeichnung |
|--------------|---------|--|
| $\mathbf{0}$ | | Nullvektor der Länge n , $\mathbf{0} = (0, 0, \dots, 0)^T$ |
| $\mathbf{1}$ | | Vektor der Länge n mit $\mathbf{1} = (1, 1, \dots, 1)^T$ |
| \mathbf{I} | | Einheitsmatrix, $\mathbf{I} = \text{diag}\{1, \dots, 1\}$ |
| α | | Vektor der Parameter |

| Symbol | Einheit | Bezeichnung |
|--|--|---|
| A | | Metabolit A bzw. Metabolit-Pool A |
| A#0100 | | Isotopomer #0100 des Metaboliten A bzw. des Metabolit-Pools A |
| a_{0100} | | Isotopomeren-Anteil A#0100 |
| \mathbf{c} | [mM] | Vektor der (intrazellulären) Konzentration |
| \mathbf{x}_c | [mM] | Vektor des Isotopomeren-Konzentrationen |
| \mathbf{x} | | Isotopomerenverteilungs-Vektor (eines oder mehrerer Metaboliten) |
| v1 | | Reaktionsschritt v1 |
| v_1 | [mM _{intra} s ⁻¹] | Reaktionsgeschwindigkeit (Fluss) der Reaktion v1 |
| \mathbf{v} | | Vektor der Flüsse, z.B. $\mathbf{v} = (v_{\text{upt}}, v_1, v_2, \dots)^T$ |
| \mathbf{v}_f | | Vektor der <i>freien</i> Flüsse, z.B. $\mathbf{v}_f = (v_{\text{upt}}, v_6, v_7)^T$ |
| $\hat{\mathbf{v}}$ | | Aus der Optimierung geschätzter Flussvektor |
| μ | [h ⁻¹] | Wachstumsrate |
| r_P | [mmol h ⁻¹ g _{BTM} ⁻¹] | Produktbildung |
| r_S | [mmol h ⁻¹ g _{BTM} ⁻¹] | Substrat-Aufnahmerate |
| \mathbf{P} | | Hilfsmatrix zur Zuordnung von Vektorelementen, Matrix Beschneidung, etc. |
| \mathbf{m}_X | | Massenisotopomerenverteilung des Metaboliten X |
| \mathbf{y} | | Messung (Vektor) |
| $\hat{\mathbf{y}}$ | | Erwartete Messung (Vektor) |
| \mathbf{M} | | Messmatrix |
| ϕ | | Verdünnungsfaktor (z.B. durch Extraktion) |
| ω | | Skalierungsfaktor |
| γ | | Verhältnis von Peakfläche zu Konzentration (MS-Messung) |
| ε | | (Mess-)Fehler |
| $\text{Cov}(\boldsymbol{\alpha})$ | | Kovarianzmatrix der Parameter |
| $\text{Cov}(\mathbf{y})$ | | Kovarianzmatrix der Messung |
| $\text{Sens}_{\boldsymbol{\alpha}}^{\mathbf{x}}$ | | Sensitivität der Variablen \mathbf{x} bezüglich der Parameter $\boldsymbol{\alpha}$ |
| σ | | Standardabweichung |
| γ | | Natürliche Isotopen-Wahrscheinlichkeits-Verteilung |
| $\boldsymbol{\xi}$ | | Vektor der Probenname-Zeitpunkte |
| $\Delta_f G^0$ | | Gibbs-Energie eines Moleküls unter Standardbedingungen |
| $\Delta_r G^0$ | | Gibbs-Energie einer Reaktion unter Standardbedingungen |
| χ^2 | | Summe der Fehlerquadrate. |
| ι | | Einfluss eines Effektors |
| G^0 | [J K ⁻¹ mol ⁻¹] | Gibbs-Energie der Verbindung unter Standardbedingungen |
| $\Delta_r G$ | [J K ⁻¹ mol ⁻¹] | Freiwerdende Reaktionsenergie (reaction) |
| $\Delta_f G$ | [J K ⁻¹ mol ⁻¹] | Energie die zur Bildung benötigt wird (formation) |

Indizes

| Index | Beschreibung |
|----------------|---|
| upt | Index für den Fluss in die Zelle (Substrat-Aufnahme) |
| reactor | Index für den Bezug zum Reaktor |
| intra | Index für den Bezug zur intrazellulären Umgebung |
| ex | Index für den Bezug zur extrazellulären Umgebung |
| feed | Index für den Bezug zur (Substrat)Zugabe |
| inp | Index für Eingang-Substrate (Substrat-Zusammensetzung) |
| P | Bezug zum Produkt (z.B. Produktbildungsrate) |
| aE | Alle Effektoren |
| nE | Keine Effektoren |
| min | Minimal |
| max | Maximal |
| opt | Optimal(e) (z.B. Flussverteilung, Probennahme-Zeitpunkte, etc.) |

Teil I
Einleitung & Methoden

Kapitel 1

Metabolic Engineering

1.1 Entwicklung der weißen Biotechnologie

Mikroorganismen, vor allem Bakterien und Pilze, sind für den Unbefangenen zunächst für die Schäden bekannt, die sie bei Mensch, Tier und Pflanze verursachen. Ihr Auftreten als "Schädlinge" beschäftigt die human-, veterinär- und phytomedizinische Mikrobiologie. Mikroorganismen treten jedoch überwiegend in der Rolle der "Nützlingle" auf. Sowohl im Naturkreislauf (z.B. Kohlenstoff- und Stickstoff-Kreislauf) als auch in der Industrie sind Mikroorganismen wertvolle Helfer.

Der Mensch nutzt seit Jahrtausenden die Stoffwechselleistung von Mikroorganismen. Klassische Produkte der Biotechnologie sind Bier, Wein, Essig, Brot und Milchprodukte. Der Einsatz von biotechnologischen Verfahren beruhte dabei auf eher empirischen Forschungen. Dass "kleine Tierchen", wie der Entdecker der Bakterien Antoine van Leeuwenhoek sie nannte, Grund der alkoholischen Gärung waren, wusste man lange Zeit nicht. Robert Koch zählte im vorangegangenen Jahrhundert zu den ersten, die Mikroorganismen im Labor gezüchtet und untersucht haben. Erst die Entdeckung der DNS durch WATSON UND CRICK (1953) führte zur heutigen Gentechnik, die neue Möglichkeiten für die Biotechnologie eröffnete.

Seitdem hat sich das Spektrum biotechnologischer Verfahren enorm vergrößert. Neben der Veredelung von landwirtschaftlichen Produkten (Lebensmitteltechnologie) werden Mikroorganismen in der Umwelttechnologie, im Pharmabereich und zur Herstellung von Feinchemikalien eingesetzt. Neben kleinen Molekülen, wie z.B. Aminosäuren, organischen Säuren und Antibiotika, werden inzwischen auch Impfstoffe und andere komplexe Substanzen, wie z.B. Humaninsulin mit Mikroorganismen hergestellt. Die vielfältigen Anwendungen werden unter dem Begriff der "weißen Biotechnologie" oder auch "industriellen Biotechnologie" zusammengefasst.

Trotz der Fortschritte gibt es weiterhin vielfältige ungeklärte Zusammenhänge im mikrobiellen Stoffwechsel, er bildet den Gegenstand aktueller Forschungsarbeiten. Das Spektrum der Untersuchungen umfasst sowohl genetische, molekularbiologische, mikrobiologische als auch prozesstechnische Arbeiten. Zur Durchführung und Auswertung der Experimente wird Wissen aus vielen verschiedenen Disziplinen benötigt.

1.2 Motivation und Methoden des Metabolic Engineering

Seit längerem werden mit den Methoden der Gentechnik gezielte Eingriffe zur Steigerung der natürlichen Stoffwechsellleistungen vorgenommen. Oftmals muss jedoch festgestellt werden, dass die Produktbildung nicht das gewünschte, maximal mögliche Potential erreicht. Aus dem Substrat (meist Glucose oder Stärke) werden neben dem Zielprodukt vor allem CO_2 , Acetat und weitere Zwischenprodukte aus dem Stoffwechsel gebildet. Wieso es trotz der Modifikationen zur Bildung von Nebenprodukten kommt und nicht die volle Produktivität erreicht wird, kann bislang meist nur vermutet werden.

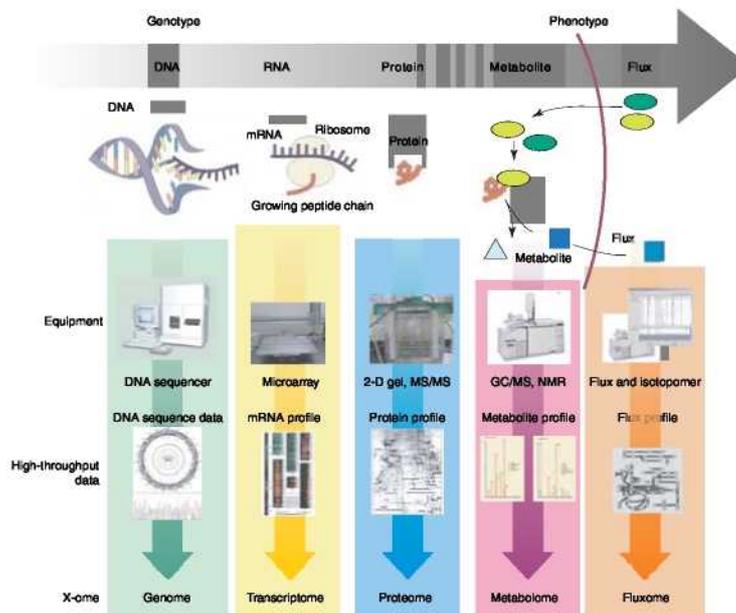


Abbildung 1.1: In der Systembiologie untersuchte Ebenen der Regulation von Mikroorganismen (aus (LEE ET AL., 2005)).

Antworten kann nur eine detaillierte und quantitative Analyse der Stoffwechsellvorgänge in der Zelle geben. Das hoch komplexe und extrem vernetzte Zusammenspiel genetischer, transkriptorischer, (post-)translatorischer und metabolischer Prozesse zu verstehen, erfordert die integrierte Anwendung unterschiedlicher Methoden und Disziplinen. Sie werden unter den Schlagwörtern “Metabolic Engineering” (BAILEY, 1991) und “Systembiologie” (DOUTHEIL, 2002; KITANO, 2002) entwickelt und angewendet. Abb. 1.1 gibt einen Überblick über die untersuchten Eigenschaften der Zellen und den Methoden, die unter den Schlagwörtern “Genomics”, “Transcriptomics”, “Proteomics” und “Metabolomics” geführt werden. Ziel ist die systematische Erforschung des (mikrobiellen) Stoffwechsels, um

mögliche Stoffwechsellängspässe auf dem Weg zum Produkt zu identifizieren (LEE ET AL., 2005; MAHADEVAN ET AL., 2005; VEMURI UND ARISTIDOU, 2005).

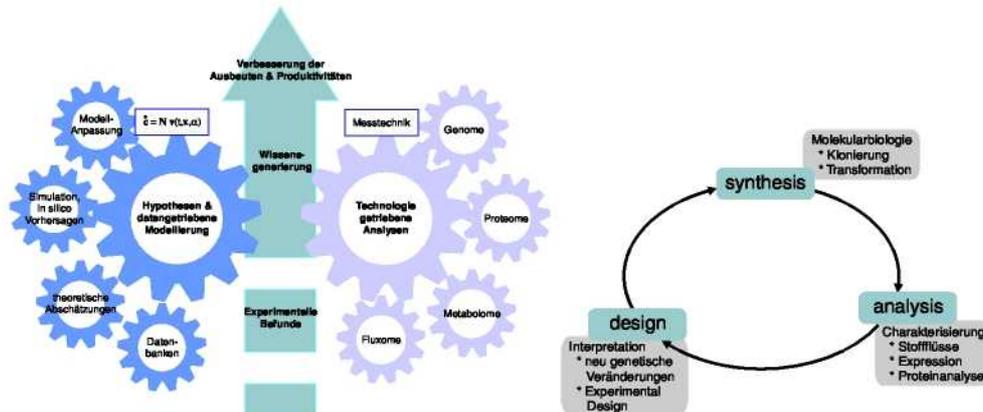


Abbildung 1.2: a) Modellgestützte Auswertung von Experimenten. Zur Verfügung stehen verschiedene mathematische Werkzeuge und Datenbanken. Im Vordergrund dieser Arbeit steht die datengetriebene Auswertung von experimentellen Daten (nach (LEE ET AL., 2005)). b) "Metabolic Engineering Cycle": Ausgehend vom aktuellen Wissenstand wird ein Stamm vorgeschlagen ("synthesis") und ein Experiment durchgeführt ("analysis"). Die modellgestützte Auswertung wird zu weiteren Hypothesen führen, die in neuen Experimenten untersucht werden ("design") (nach (NIELSEN, 2001))

Der Modellbildung und Simulation des Stoffwechsels kommt hierbei immer mehr Bedeutung zu. Ziel der Modellierung ist nicht nur ein qualitatives, sondern vor allem das quantitative Verständnis der metabolischen und zunehmend auch genetischen bzw. regulatorischen Zusammenhänge in der Zelle (Abb. 1.2a). Um aussagekräftige Modelle zu erhalten, ist die Kopplung an experimentelle Daten von entscheidender Bedeutung. Nur durch den kombinierten Einsatz von experimentellen und mathematischen Methoden können die komplexen Zusammenhänge in der Zelle verstanden werden.

NIELSEN (2001) schlägt ein iteratives Vorgehen für die Wissensgenerierung vor. Er nennt es den "metabolic engineering cycle" (Abb. 1.2b). Die Untersuchung beginnt mit einer genetischen Veränderung ("synthesis"). Die durch die Modifikation eintretenden Veränderungen werden auf den verschiedenen Ebenen (Abb. 3.2) untersucht ("analysis"). Aus der modellgestützten Auswertung gehen verschiedene Hypothesen hervor, die in weiteren Experimenten mit anderen genetischen Veränderungen genauer überprüft werden müssen ("design"). Um aussagekräftige Experimente, die zur Diskriminierung zwischen den verschiedenen Hypothesen herangezogen werden können werden modellgestützt Versuchsplanungen durchgeführt.

Die im Rahmen dieser Arbeit vorgestellten Methoden unterstützen die Schritte des “metabolic engineering cycles”. Es werden Methoden zur gezielten Auswertung und Interpretation entwickelt. Daraus lassen sich leicht Hypothesen zur Steigerung der Produktivität ableiten. Zudem wird die Planung von weiteren Experimenten durch Methoden der Versuchsplanung unterstützt. Versuchsplanungsstrategien werden auch zur *a priori* Beurteilung von neuen, bislang nicht durchgeführten experimentellen Bedingungen eingesetzt.

1.3 Modellierung im Metabolic Engineering

Bereits im vorangegangenen Abschnitt wurde die modellgestützte Auswertung angesprochen. Ist das Ziel die quantitative Beschreibung von Zusammenhängen, ist die Modellierung und Simulation eines der wichtigsten Hilfsmittel. In den Ingenieurwissenschaften können fast alle untersuchten Systeme mit Hilfe von Modellen untersucht, analysiert oder gar entwickelt werden (beispielsweise elektrische Schaltkreise, chemische Reaktionen, Strömungen, u.v.a.).

Die Anwendung von Modellierung und Simulation nimmt bei biotechnologischen Untersuchungen stetig zu. Aufgrund der hohen Komplexität der Stoffwechselnetzwerke ist eine modellgestützte Auswertung dringend notwendig. Vor allem können die für das “metabolic engineering” relevanten Größen, wie z.B. die Stoffflüsse, Enzymaktivitäten und -Affinitäten nicht direkt aus den gemessenen Daten abgeleitet werden. Zur Bestimmung müssen quantitative Modelle eingesetzt werden.

Besonders bekannt sind stöchiometrische Modellierungsansätze, insbesondere die Stoffflussanalyse (STEPHANOPOULOS ET AL., 1998) und Netzwerkanalyse (SCHUSTER ET AL., 2000). Stöchiometrische Modelle beruhen auf der Massenerhaltung. Mit der Stoffflussanalyse werden aus gemessenen extrazellulären Stoffflüssen die intrazellulären Stoffflüsse geschätzt (z.B. (VALLINO UND STEPHANOPOULOS, 1993; BONARIUS ET AL., 1996; DELGADO UND LIAO, 1997)). Anhand der stöchiometrischen Modelle können unter Hinzunahme thermodynamischer Grenzen theoretische Ausbeuten berechnet werden und Phenotypen vorhergesagt werden (z.B. (SCHUSTER ET AL., 1999; EDWARDS UND PALSSON, 2000; BURGARD ET AL., 2001; MAHADEVAN UND SCHILLING, 2003)).

Stöchiometrische Analysen geben eine Beschreibung der Stoffflussverteilung eines bestimmten physiologischen Zustands. Die Modelle können nicht zur Vorhersage von Stoffflüssen unter veränderten Bedingungen oder einer veränderten genetischen Ausstattung eingesetzt werden. Zur Vorhersage wird neben der Stöchiometrie auch eine quantitative Beschreibung der kinetischen Interaktionen benötigt. So entstehen reaktionskinetische Modelle, die neben der Stöchiometrie Enzymkinetiken enthalten. Zur Bestimmung der *in*

in vivo Enzymkinetiken werden Experimente die mehrere Zustände durchlaufen benötigt. PETERSEN ET AL. (2001) führte mehrere Stoffflussanalysen unter verschiedenen Bedingungen durch und nutzt die daraus gewonnenen Daten für ein reaktionskinetisches Modell.

Um die aufwendige Durchführung mehrerer Experimente zu umgehen, werden Stimulus-Response Experimente eingesetzt. Der zelluläre Metabolismus wird in diesen Experimenten stark ausgelenkt, so dass die Zelle verschiedene metabolische Zustände durchläuft (z.B. (CHASSAGNOLE ET AL., 2002; THEOBALD ET AL., 1997; VISSER ET AL., 2002; OLDIGES ET AL., 2004)). Ziel des reaktionskinetischen Ansatzes ist, konkrete Vorschläge zur Verbesserung der Stoffwechselleistung zu ermöglichen (VISSER ET AL., 2004a).

Kapitel 2

Ziele und Gliederung

2.1 Schwerpunkte der vorliegenden Arbeit

Die vorliegende Arbeit versteht sich als Beitrag zum Metabolic Engineering, insbesondere im Bereich der Modellierung und Simulation. Betrachtet man die Modellbildung innerhalb des Metabolic Engineering, fällt auf, dass die Modellbildung für jede Fragestellung eigene Methoden vorschlägt. Damit geht die Systematik der Modellbildung und Simulation etwas verloren.

In dieser Arbeit wird deshalb ein auf die unterschiedlichen Experimente allgemein anwendbares Konzept entwickelt. Die Problemstellung gliedert sich in die zwei großen Bereiche Modellbildung und Simulation. Die erste Aufgabe der Modellierung ist, das betrachtete System so zu abstrahieren, dass Gleichungen zur Beschreibung des Verhaltens aufgestellt werden können. Mit der Lösung der Gleichungen beschäftigt sich die Simulationstechnik. Aus den langjährigen Anwendungen in anderen Ingenieurwissenschaften steht hier ein breites Spektrum an Methoden zur Verfügung, das nun auf biologische Systeme übertragen werden muss.

Diese Verbindung von Modell und der experimentellen Umgebung bildet einen Schwerpunkt der vorliegenden Arbeit. Es werden breit anwendbare Konzepte zur integrierten Modellierung von Experimenten ausgearbeitet. Das Experiment wird durch folgende Eigenschaften charakterisiert:

- **Untersuchter Organismus:** Er ist durch eine bestimmte genetische Ausstattung charakterisiert. Die Genetik gibt vor, welche Reaktionen in der Zelle ablaufen können.
- **Experimentelle Bedingungen:** Durch die eingestellten experimentellen Bedingungen wird vorgegeben, welche Zustände beobachtet werden (physiologische Bedingungen, stationär oder dynamisch),
- **Probenaufarbeitung:** Die Probenaufarbeitung hängt von der eingesetzten Messtechnik und den beobachteten Zustandsgrößen ab (z.B. Zellaufschluss),
- **Messtechnik:** Sie gibt vor, welche Größen des Systems beobachtet werden können.

Diese Bestandteile werden in der Modellierung durch Bausteine abgebildet (Abb. 7.1). Das Modell besteht aus

- Netzwerkmodell: In das Netzwerkmodell gehen die genetische Ausstattung sowie der beobachtete Zustand ein.
- Messmodell: Zur Abbildung der Probenaufarbeitung und der Messung wird ein Messmodell eingesetzt.
- Messdatenaufarbeitung: Teilweise müssen die Messdaten aufgrund weiterer Einflüsse korrigiert werden, die mit dem Messmodell nur umständlich abgebildet werden können (z.B. Verzerrung von Markierungsmessungen durch natürliche Massenisotope).

Durch diesen modularen Modellaufbau lassen sich Modelle einfach an die gegebenen experimentellen Bedingungen anpassen. Steht das Modell, können die verschiedenen Verfahren aus der Simulationstechnik, insbesondere Simulation, Parameteranpassung, statistische Analyse und Versuchsplanung angewendet werden.

Einen weiteren Schwerpunkt der Arbeit bilden Strategien zur Interpretation der Modelle. Es werden Konzepte zur Auswertung von stationären Stoffflusslagen vorgestellt, die die gemessenen Flüsse in Bezug zum Ziel einer maximalen Ausbeute setzt. In die Analyse fließen unter anderem Beobachtungen aus der Thermodynamik ein.

2.2 Gliederung der Arbeit

Die Gliederung der vorliegenden Arbeit orientiert sich an der zuvor geschilderten Vorgehensweise. Zunächst werden die Grundlagen der Modellierung (**Kapitel 5**) und Simulation (**Kapitel 7**) allgemein dargestellt. Zentrales Element ist dabei immer wieder die Frage, wie sich Messfehler auf die Ergebnisse der Parameterschätzung auswirken. Die eingesetzten Konzepte der Fehlerübertragung werden in Abschnitt 7.2 ausführlich dargestellt. Die Anwendung der Methoden und die Lösungswege für die verschiedenen Fragestellungen wird in **Kapitel 8** vorgestellt.

Die Methoden werden in den folgenden Kapiteln zur Auswertung verschiedener Experimente eingesetzt. Es wird gezeigt, wie die Schritte der Probenaufarbeitung integriert werden und wie sich anhand der modularen Modelle neben den geschätzten Parametern auch die Genauigkeit der Schätzung bestimmen lässt. Ausgehend von den ausgewerteten Experimenten können Vorhersagen zur Auswirkung genetischer Veränderungen getroffen werden. Vor allem können auf der Grundlage dieses Experimente neue Experimente geplant werden. Angestrebt wird, ein Experiment so durchzuführen, dass genauere Ergebnisse erreicht werden. Dazu werden Methoden des Experimental Design eingesetzt und die erreichbaren Genauigkeiten abgeschätzt.

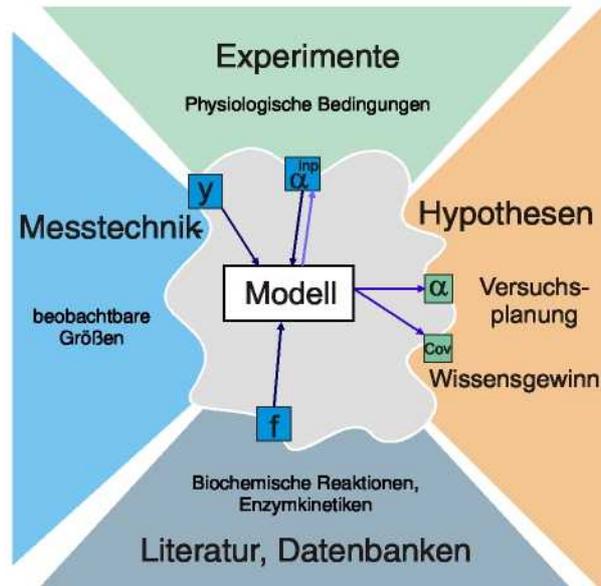


Abbildung 2.1: Um die Methoden aus der Simulationstechnik effektiv zur Auswertung biologischer Experimente einzusetzen, folgt der Modellaufbau dem experimentellen Aufbau. Die Kopplung zwischen Experiment und Modell bilden einen Schwerpunkt dieser Arbeit. Das mit der "Umgebung" gekoppelte Modell wird als Black Box mit Ein- und Ausgangsgrößen betrachtet. Für den Umgang mit dem Modell an sich stehen viele Methoden aus der Simulationstechnik zur Verfügung. Insbesondere die Parameteranpassung, statistische Untersuchungen und Versuchsplanung können nach allgemeinen Schemata durchgeführt werden.

Alle in die Anpassung des Modells einfließenden Messungen werden daher mit möglichst genauen Fehlerabschätzungen versehen. Da die Messdaten meist in vielen Schritten generiert werden und folglich entsprechend umgerechnet werden müssen, ist es entscheidend, für jeden der vorgenommenen Schritte auch die Fehler zu betrachten.

Die Auswertung stationärer ^{13}C Experimente gehört im Metabolic Engineering bereits zum Standard (WIECHERT, 2001). In der Literatur wird meist wenig über die Aufarbeitung der Messdaten, vor allem der extrazellulärer Stoffflüsse als auch der Markierungsmessungen berichtet. Üblicherweise werden eher empirische Abschätzungen getroffen. In **Kapitel 10** wird daher ein makroskopisches Modell für den quasi-stationären Zustand während der Markierungsphase eingesetzt. Anhand des Modells können alle Konzentrationsmessungen über die Zeit zur Schätzung der Flussmessung einbezogen werden. Zudem kann auch eine detaillierte Fehlerrechnung, d.h. die Übertragung von Fehlern bei der Konzentrationsmessung auf die Flussschätzung durchgeführt werden.

Anhand der berechneten Stofffluss-Verteilungen sollte festgestellt werden, welche Stoffflüsse die Produktivität beeinflussen. In der Literatur sind zur Interpretation stationärer

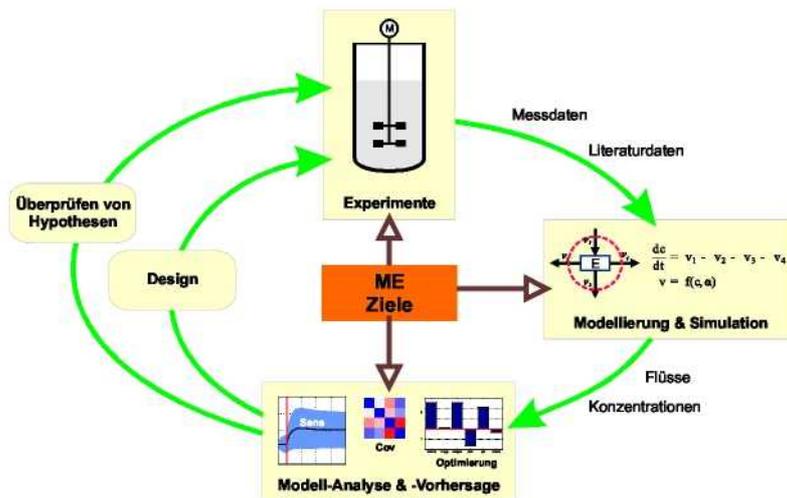


Abbildung 2.2: Integriertes Vorgehen im Metabolic Engineering: Ausgehend von einer Zielsetzung, z.B. der Steigerung der Produktausbeute, werden Experimente geplant und durchgeführt. Aus den Messdaten werden, mit Hilfe der Modellierung und Simulation, die für das Metabolic Engineering relevanten Größen, die intrazellulären Flüsse oder Enzymaktivitäten bestimmt. Mit dem Modell können Hypothesen über die metabolischen Zusammenhänge erstellt werden und die Auswirkungen von Veränderungen am biologischen System vorhergesagt werden. Um die Hypothesen bzw. Vorhersagen zu überprüfen werden Experimente, die möglichst viel Informationen für die gegebene Fragestellung liefern, gezielt geplant (Experimental Design).

Stofffluss-Verteilungen kaum geeignete Methoden zu finden. Die Metabolic Control Analysis (MCA) kann nur für Modelle mit reaktionskinetischen Termen eingesetzt werden, stationäre Modelle lassen sich mit diesen Methoden nicht untersuchen. In **Kapitel 11** werden zwei neue Methoden vorgeschlagen, die Konzepte aus der Metabolic Flux Analysis und Metabolic Network Analysis aufgreifen. Anhand der Auswertungen und Analysen in Kapitel 10 und 11 wird sodann ein "idealer" Stamm entworfen und Vorschläge zur Verbesserung des experimentellen Vorgehens abgeleitet.

Etwas komplexer ist das Vorgehen beim Erstellen reaktionskinetischer Modelle auf Basis von Stimulus-Response Experimenten. Zurzeit ist nur wenig über das *in vivo* Enzymverhalten und ihre kinetischen Parameter bekannt. Das Vorgehen der Modellerstellung muss diesem Unwissen Rechnung tragen und es muss versucht werden, möglichst alle denkbaren kinetischen und stöchiometrischen Zusammenhänge in Betracht ziehen. Für die Modellierung bedeutet dies, dass statt einem einzigen eine Vielzahl von Modellen aufgestellt und an die Daten angepasst werden muss. In **Kapitel 12** wird eine Strategie zur Auswertung von Stimulus-Response Experimenten vorgeschlagen.

Zunächst wird dabei versucht, Modelle nicht nur anhand des in Datenbanken gesammelten Wissens zu erstellen, sondern möglichst “unvoreingenommen” auf die Daten zuzugehen. Statistische Verfahren werden eingesetzt um aus den Messdaten Hypothesen über die Zusammenhänge abzuleiten. Nimmt man Literaturwissen und die Hypothesen aus den Messdaten zusammen, entstehen für die untersuchten Reaktionsschritte eine Vielzahl von möglichen Modellen, die in Modellvarianten ausgedrückt werden (**Kapitel 12**). Die Hypothesen werden mit verschiedenen kinetischen Ansätzen, d.h. mechanistischen Reaktionskinetiken (**Kapitel 13**), sowie phänomenologischen Ansätzen beschrieben. Das in dieser Arbeit eingesetzte Modellierungswerkzeug MMT2 (HAUNSCHILD, 2006) unterstützt den Modellierer durch die Möglichkeit, Modellfamilien zu definieren. Die Umgebung ermöglicht es mit einer Vielzahl von Modellen zu arbeiten, diese an die Messdaten anzupassen und zu analysieren.

Modell-Varianten, die die Daten gut wiedergeben können, werden einer genaueren Untersuchung unterzogen. Zur Analyse der Parameterbestimmtheit werden statistische Methoden (**Kapitel 14**) eingesetzt. Im Gegensatz zu stationären Experimenten sind die Sensitivitäten der Konzentrationen gegenüber den Modell-Parametern zeitabhängige, hochdimensionale Matrizen. Zur Interpretation der großen Datenmengen werden daher visuelle und statistische Hilfsmittel benötigt. In **Kapitel 9** werden die im Visualisierungswerkzeug MatVis (QELI, 2007) implementierten Verfahren kurz geschildert.

Die Visualisierung der Sensitivitäten und Kovarianzen wird hauptsächlich für die Modell-Analyse eingesetzt. Sie helfen dem Modellierer, einen schnellen Überblick zu erhalten und zu erkennen, wo Besonderheiten auftreten. Bei der Beurteilung des regulatorischen Einflusses von Inhibitoren und Aktivatoren in metabolischen Netzwerken kann die Sensitivitätsanalyse allerdings nur sehr bedingt eingesetzt werden. Die Sensitivitäten der Flüsse bezüglich der Konzentrationen können z.B. im Fall einer Sättigung zu Fehlinterpretationen führen. Obwohl fast vollständige Inhibierung vorliegt, wird daher eine sehr geringe Sensitivität berechnet. In **Kapitel 9** wird eine neue Methode zur Quantifizierung und Visualisierung von Effektor-Einflüssen beschrieben (NOACK, 2005) und zur Interpretation des Modellverhaltens in **Kapitel 16** eingesetzt.

Aus der Modell-Analyse wird auch deutlich, dass reaktionskinetische Ansätze durch eine breitere Datenbasis gestützt werden müssen. Ohne zusätzliche Informationen ist nur ein sehr geringer Teil der Modell-Parameter bestimmt. Mögliche Wege, zu besser bestimmten Modellen zu kommen, werden in dieser Arbeit diskutiert:

Gekoppelte Auswertung mehrerer Experimente: Die Stimulation des metabolischen Netzwerks *in vivo* ist experimentell stark eingeschränkt. Durch das Zucker-Aufnahmesystem (PTS) und Katabolitrepession können nur wenige Substrate zur (star-

ken) Auslenkung des Metabolismus genutzt werden. Durch genetische Eingriffe insbesondere Knock-Outs und Überexpressionen können an bestimmten Stellen gezielte Veränderungen der Stoffwechselwege erreicht werden. Experimente mit definiert unterschiedlichen Stämmen unter definierten, vergleichbaren experimentellen Bedingungen können gekoppelt ausgewertet werden. Aufgrund der genetischen Veränderungen ergeben sich andere Konzentrationsbereiche, der durchflossene Konzentrationsbereich erweitert sich. Ein Modell, das gleichzeitig an verschiedene Experimente angepasst wird, ist wesentlich besser bestimmt (**Abschnitt 15.2**).

Zusätzliche Flussmessungen: Bislang gelingt es noch nicht, den Verlauf der extrazellulären Konzentrationen während des Puls-Experiments so genau zu messen, dass daraus Flüsse berechnet werden können. Eine statistische Untersuchung soll klären, ob die extrazellulären Flüsse (z.B. im Steady State vor dem Puls) zu einer besseren *in vivo* Parameterschätzung führen (**Abschnitt 15.1**).

Puls mit markiertem Substrat: In der stationären Stoffflussanalyse wird seit langem erfolgreich ^{13}C markiertes Substrat zur Steigerung der Genauigkeit eingesetzt (WIECHERT, 2001). In einer Simulationsstudie wird untersucht, ob der Einsatz von Markierung auch für reaktionskinetische Modelle sinnvoll sein kann (**Kapitel 17**).

Kapitel 3

Untersuchte biologische Systeme

Dieses Kapitel dient dem Leser als kurze Einführung in den bakteriellen Stoffwechsel und die zellulären Regulationsmechanismen. Danach werden die in dieser Arbeit untersuchten Stämme und der Aromatenbiosyntheseweg geschildert.

Für das Metabolic Engineering sind vor allem gut charakterisierte Organismen interessant, deren genetische Ausstattung bekannt ist. Zur Entwicklung von Methoden werden deshalb vor allem *Saccharomyces cerevisiae* (VAN DAM ET AL., 2002; RAAMSDONK ET AL., 2001; VISSER ET AL., 2004b), *Escherichia coli* (ALMAAS ET AL., 2004; BURGARD ET AL., 2001; EDWARDS ET AL., 2001), *Bacillus subtilis* (DAUNER ET AL., 1999; FISCHER UND SAUER, 2005; GOEL ET AL., 1993) und *Corynebacterium glutamicum* (MARX ET AL., 1996; GRAAF, 2000; PETERSEN ET AL., 2000; VALLINO UND STEPHANOPOULOS, 1993) eingesetzt. Die Stöchiometrie der Reaktionen des Zentralstoffwechsels und vieler Biosynthesewege sind bekannt. Abb. 3.1 zeigt den Zentralstoffwechsel und die Abzweigungen zu den insgesamt zwanzig Aminosäuren.

Das im Metabolic Engineering formulierte Ziel einer gerichteten Verbesserung der Stoffwechselaktivität zur Steigerung der Produktivität lässt sich am besten am Beispiel eines industriell relevanten Prozesses untersuchen. Hier ergeben sich konkrete Fragestellungen, die mit den entwickelten Methoden beantwortet werden sollen und in einem weiteren Experiment überprüft werden können.

In dieser Arbeit werden Modelle für einen auf *E.coli* basierenden Produzenten der aromatischen Aminosäure L-Phenylalanin vorgestellt. Das Bakterium *Escherichia coli* gehört zur Gattung der Prokaryonten (SCHLEGEL, 1992), besitzt also insbesondere keine abgetrennten Reaktionsräume (Kompartimente). Prokaryonten besitzen zudem im Gegensatz zu Eukaryonten keinen Zellkern. *E. coli* ist eines der am besten charakterisierten Mikroorganismen, die gesamte DNA ist sequenziert und den meisten Genen kann eine Funktion zugeordnet werden (EISENBERG ET AL., 2000).

3.1 Bakterieller Stoffwechsel und Regulationsmechanismen

Bakterien sind in ihrer Umwelt ständig wechselnden Bedingungen ausgesetzt. In ihrer natürlichen Umgebung herrschen meist Mangelbedingungen, die Zellen müssen versuchen, mit wenigen Nährstoffen auszukommen und in der Lage sein, ihren Stoffwechsel schnell auf die verfügbaren Substrate umzustellen. Die für die biochemischen Reaktionen notwendigen Katalysatoren sind die Enzyme. Enzyme sind Makromoleküle, die aus Aminosäuren

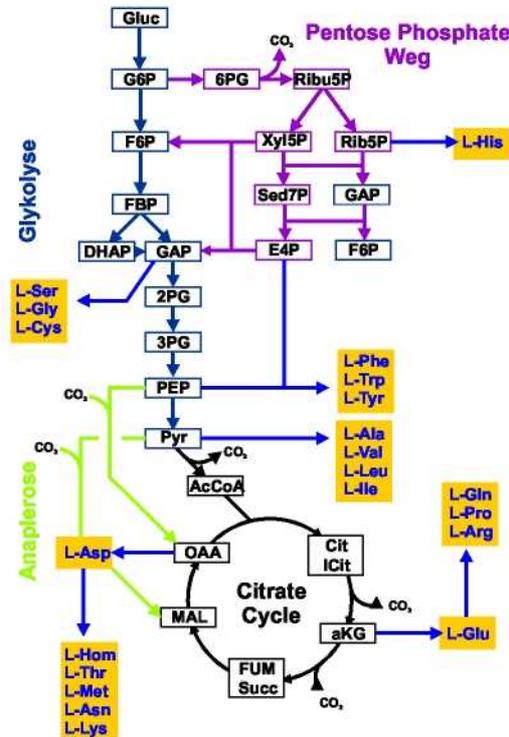


Abbildung 3.1: Zentralstoffwechsel (Glykolyse, Pentose Phosphat Weg, Tricarbonsäure Zyklus und Anaplerose). L-Phenylalanin zweigt bei E4P und PEP ab.

sowie teilweise aus Zuckern und Metall-Ionen bestehen. Deren “Bauplan”, insbesondere die Abfolge der Aminosäuren ist auf der DNA codiert. Auf ihrem Genom (DNA) tragen Bakterien bis zu 6000 Gene (In *E.coli* K-12 wurden 4288 Gene gefunden (BLATTNER ET AL., 1997)). Unter Laborbedingungen wird davon nur ein Teil wirklich exprimiert, d.h. abgelesen und das entsprechende Protein synthetisiert. Die Regulation der Proteinsynthese ist hoch komplex und verfügt über eine Vielzahl unterschiedlicher Mechanismen (Abb. 3.2),

Transkription: Bei der Enzymsynthese wird die genetische Information nicht direkt von der DNA, sondern über einen Zwischenträger, die mRNA (messenger Ribonucleinsäure) weitergegeben. Das Kopieren einer DNA-Sequenz für ein bestimmtes Enzym auf die mRNA wird Transkription genannt. Die für die Transkription zuständigen Enzyme, die RNA-Polymerasen besitzt spezielle Untereinheiten (σ -Faktoren), die den Beginn (Promotor) der abzulesenden DNA-Sequenzen erkennen. Der Zugang zum Promotor wird über verschiedene Regulatorproteine (RP) gesteuert (Abb. 3.2, mit '1' gekennzeichnet). Bislang kennt man vermutlich nur einen kleinen Teil der

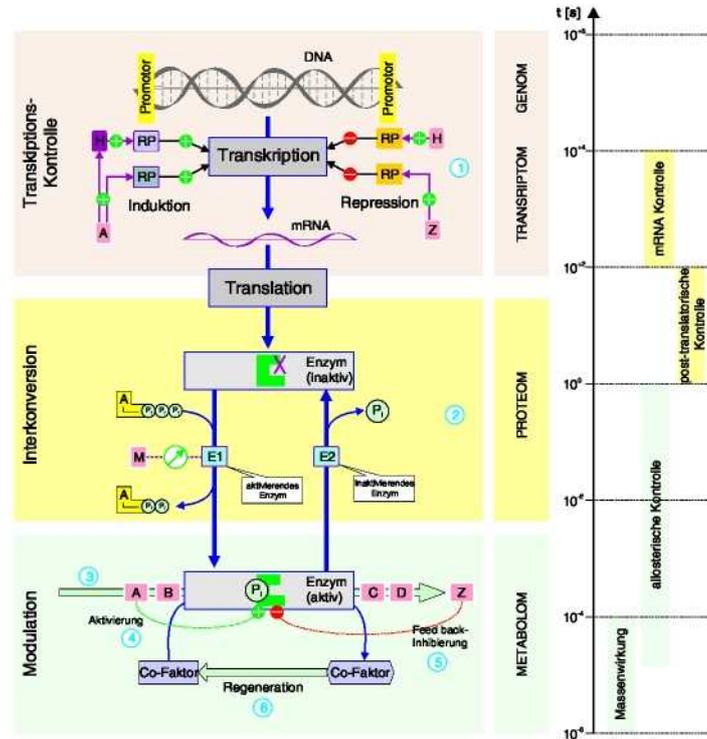


Abbildung 3.2: Schematische Darstellung der verschiedenen zellulären Regulationsebenen. Die Enzymsynthese beginnt mit der Transkription, die über Repressoren und Induktoren gesteuert wird. Die Ribosomen synthetisieren anhand der entstandenen mRNA das Enzym. Die Aktivität des Enzyms wird durch die post-translatorische Regulation (Interaktion) und allosterische Effektoren kontrolliert (aus KOOLMAN UND RÖHM (1998)).

Regulatorproteine, die auch als Transkriptionsfaktoren (TF) bezeichnet werden. Ein RP, das die Transkription steigert, wird "Induktor" genannt, ein RP, das die Transkription hemmt, bezeichnet man als "Repressor". Die Aktivität der RPs hängt meist von weiteren, kleinen Molekülen, die an das RP binden ab. Diese Art die Aktivität zu beeinflussen wird als Ligand-abhängige Regulation bezeichnet.

Translation: In der Transkription wurde eine bestimmte DNA-Sequenz auf mRNA kopiert. Für die Herstellung des entsprechenden Enzyms muss diese mRNA nun an die Ribosomen, "Enzymfabriken" der Zelle binden. Der Aufbau des Proteins (Biosynthese) durch das Ribosom anhand der mRNA Sequenz wird Translation genannt. Ähnlich wie bei der Transkription können bestimmte Moleküle an die mRNA binden und die Biosynthese beschleunigen oder verhindern (WILLIAMS ET AL., 2002). Die Regulation der Translation wird Translationskontrolle genannt. Wichtig für die

Menge an hergestelltem Protein ist auch die Halbwertszeit von mRNA. Sie wird auf ungefähr zwei bis zwanzig Minuten abgeschätzt (BERG ET AL., 2005). Es gibt Hinweise, dass die Bindung bestimmter Moleküle die Stabilität und damit die Halbwertszeit verändern können (CAPONIGRO UND PARKER, 1996; SACHS, 1993).

Interkonversion: Die genetische und translatorische Regulation reicht jedoch nicht aus, den Stoffwechsel unter sich ändernden Bedingungen schnell genug anzupassen. Der Auf- und Abbau von Enzymen ist dafür zu langsam. Die so genannte post-translatorische Regulation verändert das Enzym und steuert so dessen Aktivität. Dies geschieht zum einen durch weitere Enzyme (Interkonversion) wie z.B. Protein-Kinasen oder Dephosphorylierung (Abb. 3.2, mit '2' gekennzeichnet), zum anderen durch kleine Moleküle die an das Enzym binden (Modulation, s.u.). Wie auch bei der Transkriptionskontrolle sind die an der Interkonversion beteiligten Enzyme von Liganden abhängig. Diese Ebene der Regulation ist bereits sehr schnell, da kein Protein *de novo* synthetisiert werden muss, sondern die Enzymaktivität durch kovalente Modifikationen (Phosphorylierung, Methylierung, etc.) verändert wird (BENDT ET AL., 2003).

Modulation: Die am schnellsten agierende Regulation ist die allosterische Modulation der Enzyme durch Liganden. Neben den Bindungsstellen für Substrate und Produkte besitzen einige Enzyme weitere Bindungsstellen (Abb. 3.2, Punkte '4' und '5'). Binden an diesen Stellen die passende Moleküle verändert sich die sterische Konformation des Enzyms, d.h. dessen räumliche Faltung. Dadurch kann die Aktivität des Enzyms gehemmt (inhibiert) oder auch verstärkt (aktiviert) werden. Die allosterischen Effektoren (Inhibitoren oder Aktivatoren) binden nur an das Enzym, werden aber nicht umgesetzt. Häufig zu beobachten ist die Inhibierung von Eintrittsreaktionen in bestimmte Synthesewege, z.B. der Aminosäuresynthese, durch die Endprodukte des Synthesewegs. Dieser Inhibierungstyp wird Feedback Inhibierung genannt. Neben der allosterischen Modulation können auch nicht-allosterische Inhibitionen auftreten. Der Effektor bindet hierbei an das aktive Zentrum des Enzyms und verdrängt das Substrat (kompetitive Inhibierung). Diese Moleküle reagieren meist nur sehr langsam, oder werden gar nicht umgesetzt. Teilweise binden Effektoren auch an den Enzym-Substrat-Komplex und verlangsamen so die Reaktion (unkompetitive Inhibierung).

Viele der in der Zelle ablaufenden Reaktionen sind Cofaktor abhängig. Für die Reaktion wird als zweites Substrat z.B. der universelle Energieträger ATP oder die Reduktionsäquivalente NAD(P)H bzw. NAD(P)⁺ benötigt. Die Regeneration von ATP findet hauptsächlich

lich in der oxidativen Phosphorylierung (Atmungskette) statt. NAD^+ wird vorwiegend im Zitratzyklus reduziert, NADP^+ dagegen im oxidativen Pentose-Phosphat-Weg. Die Regulation dieser Stoffwechselwege nimmt über die Cofaktoren somit Einfluss auf eine Vielzahl weiterer Reaktionen.

3.2 L-Phenylalanin Produzenten

L-Phenylalanin ist ein Endprodukt des aromatischen Aminosäure Synthesewegs. Der erste Teil des vom Zentralstoffwechsel abzweigenden Biosynthesewegs wird Shikimat-Biosyntheseweg genannt (Abb. 3.3). In Pflanzen ist dieser Syntheseweg stark ausgeprägt. Ungefähr 20% des aufgenommenen Kohlenstoffs fließt in den Shikimat-Weg (HASLAM, 1993). Aus Shikimat werden u.a. Pigmente und Stoffen zur Abwehr von Insekten synthetisiert. In Bakterien wird Shikimat hingegen hauptsächlich zur Biosynthese der drei aromatischen Aminosäuren L-Phenylalanin, L-Tyrosin und L-Tryptophan benötigt.

Die biotechnologische Produktion von aromatischen Aminosäuren mit Bakterien ist seit langem Gegenstand der Forschung (DELL UND FROST, 1993; PATNAIK ET AL., 1995; PITTARD, 1996; FLORES ET AL., 1996). Für Tiere und Menschen ist L-Phenylalanin eine der essenziellen Aminosäuren und muss über die Nahrung aufgenommen werden. L-Phenylalanin wird als Futtermittelzusatz in relativ großen Mengen benötigt. Daneben ist L-Phenylalanin als Ausgangsstoff zur Synthese von L-Aspartam (Zuckerersatzstoff) industriell relevant (BONGAERTS ET AL., 2001).

Die Synthese von L-Phenylalanin aus zwei Molekülen Phosphoenolpyruvat (PEP) und einem Molekül Erythrose-4-Phosphat (E4P) erfolgt in zehn enzymatischen Reaktionsschritten (Abb. 3.3). Der Reaktionsweg ist eng mit dem Zentralstoffwechsel verknüpft. Neben den Ausgangsstoffen PEP und E4P werden ein ATP, ein NADPH und ein Molekül Glutamat (Glu) für die Synthese benötigt.

Bekannt ist, dass der Shikimat-Syntheseweg durch eine Vielzahl von Feedback-Inhibitionen reguliert ist (Abb. 3.4). Die erste Regulation erfolgt bei der Eintrittsreaktion, die durch die Desoxy-Heptulosonat-7-Phosphat Synthase (DAHPS) katalysiert wird. Die Endprodukte des Aromaten-Biosynthesewegs sind L-Phenylalanin (Phe), L-Tyrosin (Tyr) und L-Tryptophan (Trp), welche die Zelle zum Wachstum (Proteinsynthese) benötigt. Um eine ausreichende Versorgung sicherzustellen, verfügt *E. coli* über drei DAHPS Iso-Enzyme (Enzyme für ein und dieselbe enzymatische Reaktion), die jeweils von einem der Endprodukte inhibiert werden. Die entsprechenden Gene sind *aroH*, *aroF* und *aroG*. Um den Stofffluss zum jeweils benötigten Endprodukt zu leiten, sind am Verzweigungspunkt Chorismat weitere Feedback-inhibierte Enzyme zu finden. Die Reaktion zu Prephenat (Chorismat Mutase / Prephenat Dehydratase) wird sowohl für Tyr als auch Phe benötigt, *E. coli*

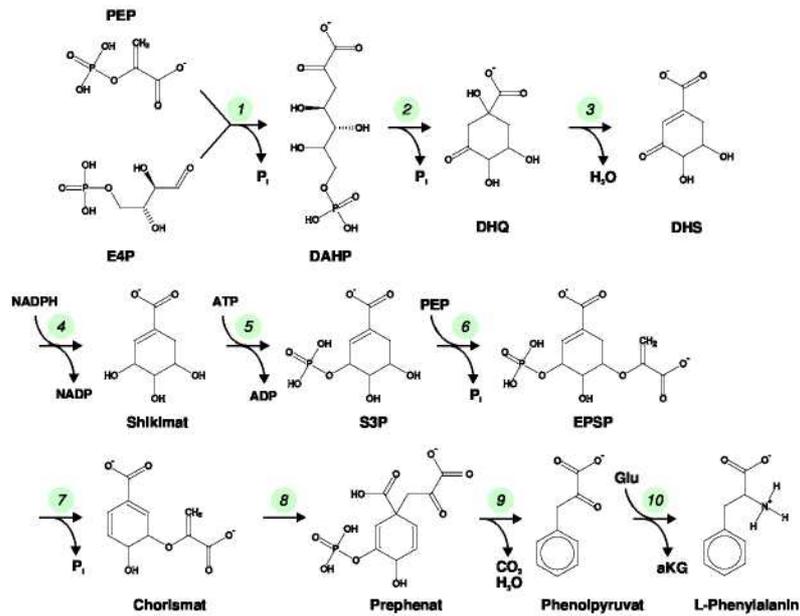


Abbildung 3.3: Die Synthese von L-Phenylalanin erfolgt in zehn Schritten. (1) DAHP Synthase, (2) DHQ Synthase, (3) DHQ Dehydratase, (4) Shikimat Dehydrogenase, (5) Shikimat Kinase, (6) EPSP Synthase, (7) Chorismat Synthase, (8) Chorismate Mutase, (9) Prephenat Dehydratase und (10) Transaminase.

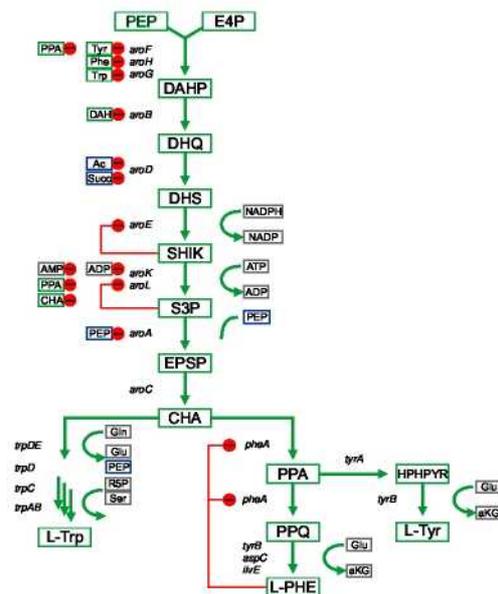


Abbildung 3.4: In der Literatur beschriebene metabolische Modulationen im Shikimat- und Aromatenbiosynthese-Weg.

besitzt für die Reaktion zwei Iso-Enzyme (Gene *pheA*, *tyrA*) mit Feedback-Inhibierung (Phe bzw. Tyr inhibiert, Abb. 3.4).

Um einen L-Phenylalanin Überproduzenten zu erhalten, werden diese Feedback-Inhibierungen entfernt. Am IBT-1 (Institut für Biotechnologie 1 des Forschungszentrum Jülich) wurden insbesondere folgende Stämme konstruiert. Aufgelistet ist lediglich eine Auswahl der in dieser Arbeit vorkommenden Stämme:

E. coli 4pF20 mit dem Genotyp $\Delta(pheA\ tyrA\ aroF)$ / pJF119EH *aroF^{fbr}pheA^{fbr}amp*.

E. coli 4pF78 mit dem Genotyp $\Delta(pheA\ tyrA\ aroF)$ / pJF119EH *aroF^{fbr}pheA^{fbr}aroB amp*.

E. coli 4pF81 mit dem Genotyp $\Delta(pheA\ tyrA\ aroF)$ / pJF119EH *aroF^{wild}pheA^{fbr}aroB aroL amp*.

Die chromosomale Deletion von *tyrA* verhindert einen Abfluss von Prephenat in die Tyrosin-Synthese. Gleichzeitig wird der Stamm durch diese Deletion Tyrosin auxotroph, d.h. nur unter Zugabe von Tyrosin im Medium ist Wachstum möglich. Die Gene *pheA* (Chorismat Mutase/Prephenat Dehydratase) und *aroF* (DAHP Synthase) wurden durch die Feedback-resistenten Varianten *aroF^{fbr}* (JOSSEK ET AL., 2001) und *pheA^{fbr}* auf dem Plasmid ersetzt. Das Plasmid pJF119EH (FÜRST ET AL., 1986) wird durch Zugabe von IPTG induziert (*tac*-Promotor).

Auf dem Plasmid des Stamms *E. coli* 4pF78 ist zusätzlich das Gen für die DHQ Synthase integriert. *E. coli* 4pF81 trägt statt der Feedback-resistenten DAHP Synthase das Gen des Wildtyps. Zusätzlich ist auf dem Plasmid auch das Gen *aroL* (Shikimat Kinase II).

Kapitel 4

Experimentelles Vorgehen

4.1 Überblick der experimentellen Bedingungen

Dieses Kapitel stellt die im Metabolic Engineering am häufigsten verwendeten experimentellen Bedingungen dar. Gezeigt wird, welche Methoden zur Auswertung verwendet werden können und welche Größen beobachtet bzw. bestimmt werden können. Experimente mit kultivierten Zellen unter definierten Bedingungen bilden die Grundlage für die datengestützte Modellierung (Kapitel 10 und 12).

Das experimentelle Vorgehen hängt dabei stark von der Zielsetzung der Experimente ab. Zur schnellen Charakterisierung von klonierten Stämmen werden meist einfache Schüttelkolben-Experimente durchgeführt. Ist das Ziel eine genauere Untersuchung der Stoffflüsse, werden Bioreaktoren eingesetzt. In den Reaktoren können die Umgebungsbedingungen geregelt und so zum einen vergleichbare und reproduzierbare Bedingungen eingestellt werden. Zum anderen kann die Steuerung der Umgebung gezielt eingesetzt werden, um die Zellen in bestimmten Zuständen zu beobachten.

Abbildung 4.1 gibt einen Überblick über mögliche experimentelle Bedingungen und den daraus gewonnenen Informationen. Unterschieden wird zunächst zwischen Experimenten unter

- **metabolisch stationären** Bedingungen: Unter diesen Bedingungen bleiben die Konzentrationen und Flüsse über den Beobachtungszeitraum konstant. Gemessen werden die extrazellulären Flüsse. Diese Art von Analyse wird Stoffflussanalyse (SFA) genannt.
- **metabolisch instationären** Bedingungen: Hier verändern sich die Konzentrationen und Flüsse über den Beobachtungszeitraum. Gemessen werden die intrazellulären Konzentrationen (Reaktionskinetische Experimente).

Mit stationären Experimenten lassen sich Aussagen über die intrazelluläre Stoffflussverteilung unter den gegebenen Umgebungsbedingungen treffen. Da mit den extrazellulären Flussmessungen nicht genügend Messdaten zur Bestimmung aller Flüsse zur Verfügung stehen, wird oftmals mit ^{13}C markierten Substraten gearbeitet. Neben Konzentrationen und Flüssen werden nun auch Markierungsanreicherungen beobachtet. Der Markierungszustand wird durch die Isotopomere (Verteilung der möglichen Markierungszustände),

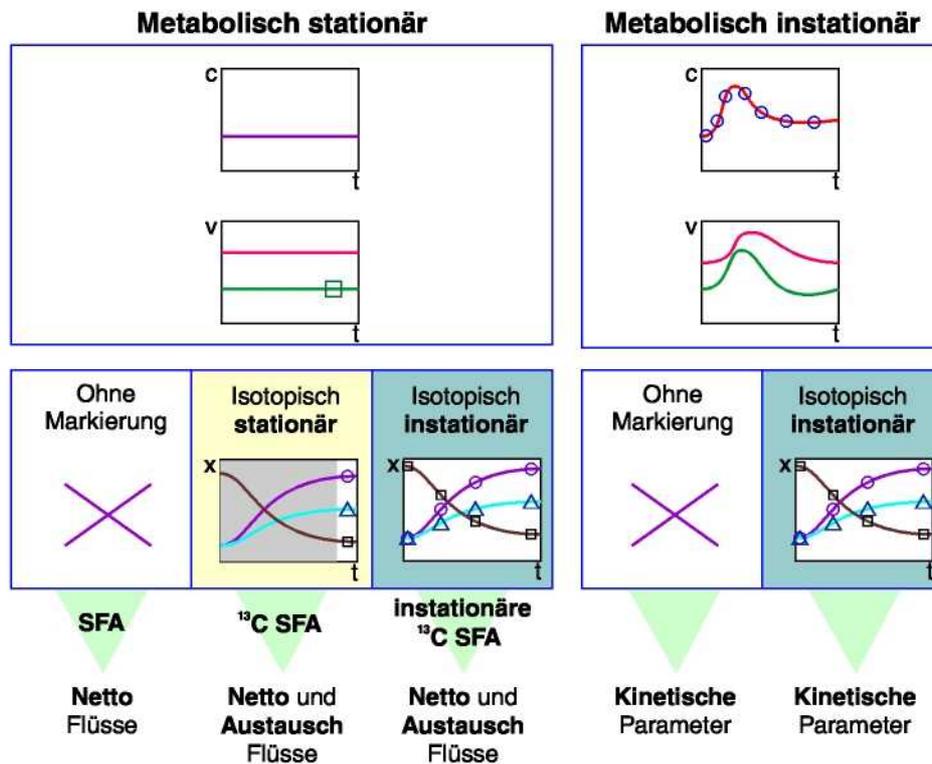


Abbildung 4.1: Je nach Zielsetzung der Analyse werden verschiedene experimentelle Bedingungen eingestellt. Soll die Stoffflusslage bestimmt werden, wird ein metabolisch stationärer Zustand benötigt. Wird keine besondere Genauigkeit benötigt, wird ein Experiment ohne Markierung durchgeführt. Gemessen werden die extrazellulären Stoffflüsse. Höhere Genauigkeiten lassen sich durch ¹³C Markierungsexperimente erreichen. Beobachtet werden kann die zeitliche Anreicherung der Markierung (isotopisch instationär) oder nach ausreichend langer Zeit der isotopisch stationäre Zustand. Sollen in vivo Enzymkinetiken bestimmt werden, wird ein metabolisch instationäres Experiment durchgeführt.

beschrieben. Wird das Experiment unter metabolisch stationären Bedingungen durchgeführt, reichert sich Markierung in den Metaboliten an und erreicht schließlich einen stationären Zustand. Beobachtet werden kann der

- **isotopisch stationäre** Zustand, der sich nach einer ausreichend langen Zeit einstellt (¹³C SFA), oder der
- **isotopisch instationäre** Zustand, bei dem die Markierungsanreicherung über die Zeit verfolgt wird (instationäre ¹³C SFA).

Für die Auswertung von metabolisch und isotopisch stationären Experimenten sind Methoden vorhanden. Flussgeschätzungen auf der Grundlage isotopisch instationärer Experi-

mente sind sowohl experimentell als auch rechnerisch extrem aufwendig und werden zum heutigen Zeitpunkt erstmals durchgeführt. Es liegt auf der Hand, dass die Genauigkeit der Flusserschätzung anhand mehrerer über die Zeit genommener Proben wesentlich gesteigert werden kann.

Ist das Ziel der Analysen, nicht nur die intrazellulären Flüsse, sondern auch die Eigenschaften und Regulation der Enzyme *in vivo* zu bestimmen, werden metabolisch instationäre Experimente benötigt. Der notwendige experimentelle und rechnerische Aufwand ist auch bei diesen Experimenten beträchtlich.

Ein Experiment, das unter metabolisch und isotopisch instationären Bedingungen durchgeführt wird, ist vom experimentellen Aufwand (fast) vergleichbar. Im Ergebnis erhält man wiederum die kinetischen Parameter der enzymatischen Reaktionsschritte. Es ist zu erwarten, dass deren Genauigkeit im Vergleich zu Experimenten ohne Markierungszugabe wesentlich gesteigert wird, da mehr Messungen zur Verfügung stehen.

Die einzelnen experimentellen Aufbauten und Messtechniken werden in den folgenden Abschnitten detaillierter beschrieben.

4.2 Stationäre Stoffflussanalyse (SFA)

Zur Bestimmung der intrazellulären Stoffflüsse in einem bestimmten physiologischen Zustand, werden Stoffflussanalyse Experimente durchgeführt. Die geforderten stationären Bedingungen lassen sich am besten in einem kontinuierlich betriebenen Reaktor einstellen (Abb. 4.2C). Bei dieser Betriebsweise wird kontinuierlich Medium mit Substrat zugegeben (Volumenfluss \dot{V}_{feed}) und das Volumen V im Reaktor über den Füllstand konstant gehalten. Bei dieser Betriebsweise stellt sich nach einer gewissen Zeit ein stationärer Zustand ein. Insbesondere kann die Wachstumsrate μ der Kultur durch die Verdünnungs-Rate $D = \dot{V}_{\text{feed}}/V_{\text{reactor}}$ eingestellt werden (maximal bis zum Erreichen der kritischen Durchflussrate $D = \mu_{\text{max}}$). Im Gleichgewichtszustand bleiben die Konzentrationen \mathbf{c} und Flüsse \mathbf{v} innerhalb und außerhalb der Zelle konstant. Im Batch oder Fed-Batch Betrieb können über einen gewissen Zeitraum stationäre Flüsse beobachtet werden (z.B. bei Substratsättigung mit maximaler, konstanter Wachstumsrate). Hier spricht man von quasi-stationären Bedingungen.

Über eine Bilanzierung der Massen werden die extrazellulären Flüsse \mathbf{v}_{ext} (z.B. Substrat-, Sauerstoffaufnahme und Produktbildung) ermittelt. Mit diesen Flüsse und einigen Annahmen über den Energiehaushalt, insbesondere der Bilanzierung aller intrazellulären Metabolite (inklusive der Energie- und Reduktions-Äquivalente), können anhand eines Netzwerkmodells intrazelluläre Flüsse berechnet werden (STEPHANOPOULOS ET AL.,

1998; VALLINO UND STEPHANOPOULOS, 1993). Bei einigen Annahmen ist jedoch Vorsicht geboten. Insbesondere folgende Punkte sollen hier genannt werden:

- Bei der Bilanzierung der Energie wird angenommen, dass die Zelle keine Energie “verschwendet”. In verschiedenen Studien wurde inzwischen belegt, dass in der Zelle “futile cycles” existieren (metabolische Reaktionen die im Kreis laufen und dabei Energie verbrauchen ohne ein Produkt herzustellen) (PETERSEN ET AL., 2000; MARX ET AL., 1996). Mit rein stöchiometrischen Ansätzen können “futile cycles” nicht entdeckt werden.
- Die Regeneration von ATP durch die oxidative Phosphorylierung wird aus der Sauerstoffaufnahme und dem P/O-Verhältnis (Phosphate to Oxygen) bestimmt. Das P/O-Verhältnis wird meist aus der Literatur entnommen, verschiedene Arbeiten zeigen jedoch, dass der Wert von den experimentellen Bedingungen abhängt (SAUER UND BAILEY, 1997; BRAND, 1994).

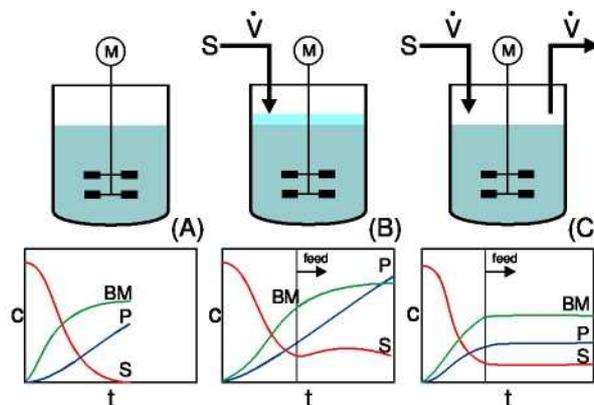


Abbildung 4.2: Betriebsweisen von Bioreaktor-Prozessen. (A) Batch-Prozess (Satzbetrieb), (B) Fed-Batch (Satzbetrieb, mit zusätzlicher Fütterung), (C) Chemostat (kontinuierlich betriebener Reaktor).

4.3 Stationäre Markierungsexperimente (^{13}C SFA)

Wie bei der stöchiometrischen SFA wird die Zelle im metabolisch (quasi-) stationären Zustand beobachtet. Die Bedingungen der Experimente sind somit (fast) identisch (Abb. 4.2). Zugeführt wird jetzt jedoch ein spezifisch ^{13}C markiertem Substrat. Um den metabolisch und isotopisch stationären Zustand zu erreichen, müssen die stationären Bedingungen allerdings vergleichsweise länger aufrecht erhalten werden. Gemessen werden nun neben den

extrazellulären Flüssen auch Markierungsanreicherungen. Durch diese zusätzlichen Informationen können die intrazellulären Flüsse genauer und auf weniger Annahmen basierend bestimmt werden. Im Vergleich zur stöchiometrischen SFA können vor allem parallele und bidirektionale Flüsse bestimmt werden.

- Parallele Reaktionswege: Einige Metabolite können durch unterschiedliche, parallele Reaktionswege synthetisiert werden (z.B. Lysin-Synthese in *C. glutamicum*). Mit einer rein stöchiometrischen Analyse können die Flüsse der parallelen Wege nicht getrennt bestimmt werden.
- Bidirektionale Flüsse: Die Zellen verfügen zum Wachstum auf verschiedenen Nährstoffen über eine Vielzahl von Enzymen. Vor allem muss sichergestellt werden, dass die Zellen alle notwendigen Precursor (Vorstufen) der einzelnen Biosynthesewege bereitgestellt werden. Viele Reaktionsschritte sind deshalb in beiden Richtungen möglich. Ist Glucose das Substrat, ist die Glycolyse aktiv. Wird Citrat das einzige Substrat vorgegeben, muss die Zelle Intermediate der Glycolyse und des PPP, die zum Wachstum notwendig sind (z.B. PEP, GAP, DAHP, R5P, E4P) bereitstellen. Die Umkehr der Glycolyse, die Gluconeogenese, erfolgt teilweise über die gleichen Enzyme, meist sind andere Enzyme wie z.B. die PEP Synthase für die "Umkehr"-Schritte zuständig (BERG ET AL., 2005). Es ist anzunehmen, dass die Reaktionen teilweise auch gleichzeitig ablaufen. Mit einem rein stöchiometrischen Ansatz können diese bidirektionalen Flüsse (Reaktionen mit gleichzeitiger Hin- und Rückreaktion) nicht getrennt bestimmt werden.

Die Flussberechnung wird zudem verlässlicher, da man auf Annahmen verzichten kann. Insbesondere wird bei der ^{13}C SFA die Energie nicht bilanziert. Die Annahmen einer geschlossenen Energiebilanz sind kritisch (s.o.).

Die Auswertung (quasi-) stationärer Markierungsexperimente und die ^{13}C Modellierung werden in Abschnitt 6.5 detailliert besprochen.

4.4 Stimulus-Response Experimente

Ziel von metabolisch instationären Experimenten ist die *in vivo* Bestimmung von Enzymkinetiken. Anhand von stationäre Experimenten können Flusslagen eines bestimmten Zustands der Zelle ermittelt werden. Wie diese Flusslage zustande kommt wie sich Veränderungen der genetische Ausstattung auswirken kann nur anhand reaktionskinetischer Beschreibungen erfolgen.

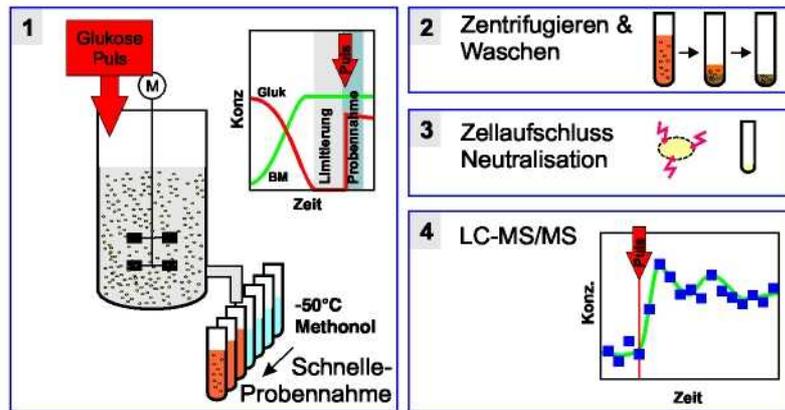


Abbildung 4.3: Durchführung von Stimulus-Response-Experimenten: Um eine möglichst große Auslenkung des Stoffwechsels zu erreichen, werden die Zellen vor dem Puls Substrat-limitiert, durch die Zugabe eines konzentrierten Substrat-Pulses wird der Metabolismus stark ausgelenkt. Die Antwort der Zelle wird mit der schnellen Probennahme-Einheit mit 3-4 Proben pro Sekunde beobachtet. In den Probennahme-Röhrchen befindet sich -50°C kaltes Methanol um den Metabolismus zum Probennahmezeitpunkt anzuhalten. Zur Messung der intrazellulären Konzentrationen wird die Probe abzentrifugiert, gewaschen und aufgeschlossen.

In den Ingenieur-Wissenschaften wird zur Analyse des Systemverhaltens häufig die Sprungantwort aufgenommen und charakterisiert. Dieser Ansatz wird im Metabolic Engineering meist "Stimulus-Response" Experiment (SRE) genannt (Abb. 4.3) (OLDIGES ET AL., 2004). Um eine möglichst große Auslenkung des metabolischen Netzwerks zu erreichen, wird die Aktivität des Metabolismus zunächst stark reduziert, indem das Substrat limitiert wird. Die limitierte Kultur wird dann ruckartig durch einem hochkonzentrierten Substrat-Puls aus dem Gleichgewichtszustand ausgelenkt. Die Antwort auf den Puls wird durch die Messung der intrazellulären Metabolit Konzentrationen beobachtet (THEOBALD ET AL., 1997). Da die Reaktion des Metabolismus auf den Puls extrem schnell ist, muss der experimentelle Aufbau einige Anforderungen erfüllen:

Pulsaufgabe: Der Substrat-Puls muss im gesamten Reaktor schnell und gleichmäßig erfolgen um ein Verwischungseffekt bzw. inhomogene Bedingungen im Reaktor zu vermeiden. Eine sehr schnelle Durchmischung ist deshalb unabdingbar, im verwendeten Bioreaktor (20L) wurde eine Mischzeit von $\Theta = 650\text{ ms}$ gemessen (BUCHHOLZ ET AL., 2001).

Probennahme: Um die Antwort des Metabolismus zeitlich detailliert aufzunehmen, müssen mehrere Proben pro Sekunde genommen werden. Ansätze schneller Probenahmen sind in (LANGE ET AL., 2001; BUCHHOLZ ET AL., 2001; SCHÄFER

ET AL., 1999) beschrieben. Mit der in Jülich eingesetzten schnellen Probennahme können 160 Proben in **40 s**, d.h. einer Frequenz von **4 – 5 Hz** entnommen werden.

Stoppen der Enzymaktivität: Um ein Abbild des Systemzustands zum Zeitpunkt der Probennahme sicherzustellen, muss der Metabolismus bzw. die Enzymaktivität sofort gestoppt werden. Das Anhalten des Stoffwechsels wird durch Quenchen in **–50 C** kaltem Methanol (DE KONIG UND VAN DAM, 1999; SCHÄFER ET AL., 1999) erreicht. Zusammen mit der Probe aus dem Reaktor stellt sich eine Mischtemperatur von ca. **–20 C** ein.¹ Man geht davon aus, dass bei dieser niedrigen Temperatur die Enzyme inaktiv sind (DE KONIG UND VAN DAM, 1999).

Intrazelluläre Messung: Es muss versucht werden, eine Methode zu finden, bei der keine extrazellulären Metabolite die Messung beeinflussen und gleichzeitig keine intrazellulären Metabolite verloren gehen. Zudem muss verhindert werden, dass es zu weiteren enzymatischen Umsetzungen während der Probenaufarbeitung kommt. Die Zellen werden zunächst durch Zentrifugieren vom Fermentationsmedium getrennt. Es werden gekühlte Zentrifugen (**–20 C**) verwendet. Der Überstand wird verworfen und das Zellpellet resuspendiert. Die resuspendierten Zellen werden dann chemisch mit 43%iger Perchlorsäure (ebenfalls bei **–20°C**, (BUCHHOLZ ET AL., 2001)) aufgeschlossen². Die saure Lösung wird vor der Messung mit Kaliumcarbonat neutralisiert. Zum Entfernen von Zellwandbestandteilen sowie gebildetem Salz (KClO_4) wird nochmals zentrifugiert.

4.5 LC-MS/MS Messtechnik

Nach dem Zellaufschluss liegen die intrazellulären Metabolite in Lösung vor. Die Anzahl intrazellulärer Metabolite in Bakterien schwankt je nach Quelle sehr stark, z.B. EcoCyc: 1048 (KESELER ET AL., 2005), (KARP ET AL., 2000; KANEHISA UND GOTO, 2000; SELKOV ET AL., 1996). Mit hochmoderne FTMS (Fourier Transformierte MS) AHARONI ET AL. (2002) kann die Masse der Moleküle so genau gemessen werden, dass deren chemische Zusammensetzung bestimmt werden kann. Bei Bakterien geht man von 500 bis zu 3000 niedermolekularen Substanzen (Metaboliten) aus (*S. cerevisiae* \approx **600**, (RAAMSDONK ET AL., 2001), Pflanzen schätzungweise 3000-20000 (PICHERSKY UND GANG, 2000)). Die Messtechnik muss in der Lage sein, diese Vielzahl von Metaboliten getrennt zu quantifizieren. Eine weitere Schwierigkeit bei der intrazellulären Messung sind die extrem

¹In der Literatur werden noch weitere Verfahren vorgeschlagen. HAJJAJ ET AL. (1998) verwenden flüssigen Stickstoff, GONZALES ET AL. (1997) schlagen kochendes Ethanol zum Quenchen vor.

²Alternative Methoden sind der Aufschluss mit heißem Ethanol (GONZALES ET AL., 1997) oder Trichloracetat (KOPPENSCHLÄGER UND AUGUSTIN, 1967)

niedrigen Konzentrationen in der Zelle. Durch den Zellaufschluss und die anschließende Neutralisation wird die Lösung zudem stark verdünnt (ca. 25-35fach). Die eingesetzte Messtechnik muss daher extrem sensitiv sein.

Mit einer Kopplung von chromatographischer Trennung (mittels HPLC, High Performance Liquid Chromatography) und Massenspektroskopie (MS) kann eine Trennung aufgrund physikalischer bzw. chemischer Eigenschaften und der Masse erreicht werden (Abb. 4.4) (VAN DAM ET AL., 2002; BUCHHOLZ ET AL., 2001). Die Koppelung dieser beiden Messtechniken ist erst seit kurzem überhaupt möglich. Die HPLC arbeitet bei sehr hohen Drücken und in wässriger Lösung, im MS-Gerät hingegen werden gasförmige Ionen im Hochvakuum getrennt (ARPINO UND GUIOCHON, 1982). Bei der Entwicklung der HPLC-Methode muss auf eine für die MS verträgliche Wahl von Eluent und Puffer geachtet werden. Eluent und Puffer müssen flüchtig sein um bei der Ionisierung (ca. 300 °C) zu verdampfen. Bei der in dieser Arbeit vorgestellten Messungen von OLDIGES (2004) wurden zwei in Serie geschaltete HPLC-Säulen (Nucleodex beta-OH 250, Macherey-Nagel) verwendet. Eluiert wurde mit einem Gradient (Eluenten: 12 mM NH₄Ac und 20% A + 80% Methanol). Die chromatographische Trennung ist dabei so ausgelegt, dass vor allem der Einfluss von Matrix-Molekülen, wie z.B. Überresten der Zellwand oder der Proteine, verringert wird.

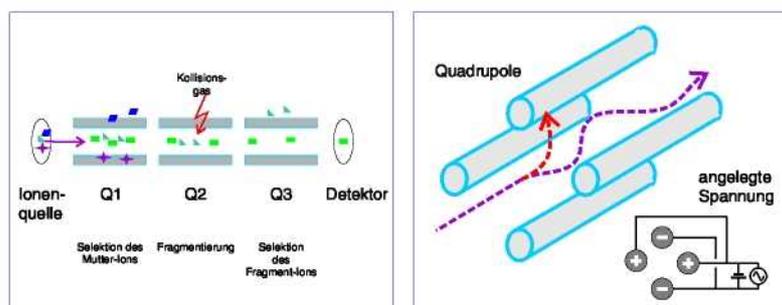


Abbildung 4.4: (a) Prinzip der MS/MS Messtechnik. Im ersten Quadrupol (Q1) wird eine bestimmte Masse "durchgelassen", der zweite Quadrupol (Q2) dient zur Fragmentierung. Im dritten Quadrupol wird wiederum eine bestimmte (Fragment-)Masse gewählt, Ionen mit diesem m/z Verhältnis treffen auf den Detektor. (b) Schematischer Aufbau eines Quadrupols. An den vier Stäben wird eine Gleichspannung und eine Wechsellspannung angelegt. In dem aufgebauten elektrischen Feld werden die Ionen von ihrer Flugbahn ausgelenkt. Je nach angelegter (Wechsel-)Spannung passieren Ionen mit einem bestimmten Masse zu Ladungsverhältnis m/z den Quadrupol. Zu leichte bzw. zu schwere Ionen fliegen seitlich aus dem Feld.

Der "Detektor", das MS-Gerät, besteht aus einer Serie von "Massenfiltern". Die Moleküle werden beim Eintritt in das MS zunächst ionisiert. Durchgesetzt hat sich hier die Elektro-

sprayionisation (ESI) (HOPFGARTNER ET AL., 2004). Die jetzt geladenen Analyten, werden durch ein starkes elektrisches Feld beschleunigt und treten in den ersten von drei Quadrupolen (Q1). Der Quadrupol besteht aus vier Zylindern, die ein elektromagnetisches Wechselfeld aufbauen. Die durchfliegenden Ionen werden in diesem Feld ausgelenkt, wie stark die Auslenkung ist, hängt von ihrem Masse- zu Ladungsverhältnis (m/z) ab. Bei einer bestimmten angelegten Frequenz können Ionen mit einem Masse- zu Ladungsverhältnis in Resonanz gebracht werden und so den Quadrupol passieren – Ionen mit einem anderen m/z -Verhältnis fliegen aus dem Feld. Der zweite Quadrupol (Q2) wird als Kollisionsraum genutzt. Die Moleküle stoßen mit einem Kollisionsgas (meist Stickstoff) zusammen und zerfallen in Fragmente (Abb. 4.4). Das Fragmentierungsmuster ist von der chemischen Struktur und den kovalenten Bindungsstärken abhängig. Durch diesen Schritt können Analyten mit gleichem Masse- zu Ladungsverhältnis noch aufgrund einer unterschiedlichen Fragmentierung getrennt werden. Im dritten Quadrupol (Q3) kann das gewünschte Fragment selektiert werden und trifft auf den Detektor.

Die Auswertung der Spektren sowie eine Fehlerabschätzung der Messungen wird in Abschnitt 12.2 beschrieben.

Kapitel 5

Modellbildung für das Metabolic Engineering

Die für das Metabolic Engineering besonders interessanten Größen, die *in vivo* Stoffflüsse, die *in vivo* Enzym-Konzentrationen, Enzym-Aktivitäten und metabolischen Regulationen können nicht direkt gemessen werden. Der Zugang zu diesen Größen kann nur über quantitative Modelle hergestellt werden. Das Modell beschreibt, wie die nicht beobachtbaren Größen, die *in vivo* Flüsse, bzw. Enzym-Konzentrationen, -Aktivitäten und -Affinitäten, auf das beobachtbare Verhalten wirken.

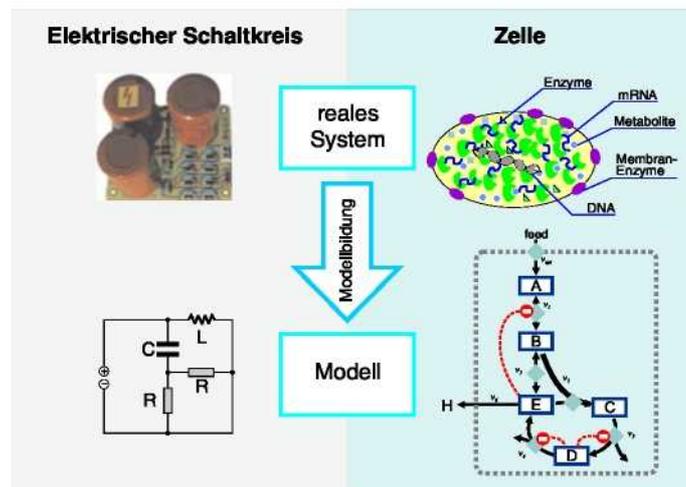


Abbildung 5.1: Analogien beim Vorgehen der Modellierung. Die Bestandteile des elektrischen Schaltkreises besitzen bestimmte Funktionen (z.B. Speicherung von Ladung in Kondensatoren) und sind über Leiterbahnen miteinander verbunden. Im elektrischen Netzwerk gilt die Erhaltung der Ladung. Im Metabolischen System bilden enzymatische Reaktionen die Verbindungen zwischen den einzelnen Metabolit-Pools (Speichern). Hier gilt die Erhaltung der Masse.

Die Herausforderung besteht nun darin, die Techniken der Modellierung aus den Ingenieurwissenschaften auf biologische Systeme zu übertragen. Die Modellierung metabolischer Netzwerke wird oft mit der Modellierung elektrischer Schaltkreise verglichen (WIECHERT, 2002) (Abb. 5.1). Grundlage der Modellerstellung bei elektrischen Schaltkreisen ist die Erhaltung der Ladung, die vergleichbar mit der Erhaltung der Masse metabolischer Reaktionsnetzwerke ist. Die Funktion der einzelnen Bauteile elektrischer Schaltkreise (Transistoren, Widerstände, etc.) kann mit einigen Grundgesetzen beschrieben

werden. Dabei werden einige vereinfachende Annahmen getroffen, z.B. wird der Einfluss der Temperatur vernachlässigt.

Die Modelle in diesen Disziplinen haben eine erstaunliche Reife erlangt. Beispielsweise können Schaltkreise oder ganze Prozessoren am Computer entwickelt werden und die Modelle geben das Verhalten des realen Schaltkreises sehr exakt wieder.

Die Modellierung biologischer Systeme ist in gewisser Weise komplexer, und erfordert eine andere Herangehensweise. Vor allem müssen folgende Punkte in Betracht gezogen werden:

Unbekannte Wechselwirkungen]: Im Gegensatz zu z.B. elektrischen Schaltkreisen, bei denen die einzelnen Bauteile durch klar definierte elektrische Leiterbahnen verbunden sind, befinden sich bei biologischen Systemen Metabolite, Enzyme und andere Zellbestandteile gemeinsam im Cytoplasma (Prokaryonten) bzw. in Kompartimenten (Eukaryonten und höhere Zellen). Prinzipiell kann jede Substanz mit jeder anderen interagieren (PAZOS ET AL., 2005). Einige Enzyme sind in der Lage, mehrere Reaktionen zu katalysieren. Teilweise muss aus diesem Grund auch die Stöchiometrie metabolischer Netzwerke hinterfragt und erweitert werden (VAN WINDEN ET AL., 2003).

Kritischer ist die Situation beim Aufstellen der möglichen regulatorischen Einflüsse von Metaboliten auf Enzyme. Die regulatorische Struktur ist nicht immer bekannt (HOFMANN UND KOPPERSCHLAGER, 1982; PAZOS ET AL., 2005), da bei der *in vitro* Analyse von Enzymen meist nur ein bestimmtes Substanzspektrum untersucht wird.

In vitro und in vivo Untersuchungen: Im Gegensatz zu rein technischen Systemen können die einzelnen Komponenten, die Gene, Transkription, Translation, Enzyme und Metabolite, nicht einzeln, isoliert im Reagenzglas (*in vitro*) untersucht und beschrieben werden. Enzyme sind komplexe Makromoleküle, deren Funktion stark von den Umgebungsbedingungen, wie z.B. dem pH-Wert, der Temperatur, der Salzkonzentrationen oder dem Reduktionspotential abhängen. Bislang ist es nicht möglich, die *in vivo* intrazellulären Bedingungen *in vitro* zu reproduzieren. Eine weitere Schwierigkeit besteht in der *in vivo* Situation "viel Enzym, wenig Substrat", ca. 40% des Zellvolumens wird von Makromolekülen eingenommen (CHEBOTAREVA ET AL., 2004). Studien mit makromolekülähnlichen Substanzen (z.B. PEG) zeigen, dass Crowding die thermodynamische Aktivität der Metabolite erhöht, gleichzeitig aber die Diffusionsgeschwindigkeit und damit die Reaktionswahrscheinlichkeit erniedrigt (MATHEWS, 1993; ROHWER ET AL., 1998). Das Ausmaß dieser beiden

Effekte hängt von der Art der jeweiligen Reaktion ab (CHEBOTAREVA ET AL., 2004).

Metabolic Channeling, die direkte Übergabe eines Reaktanten an das nächste Enzym, wird sowohl für die Glykolyse (SHEARER ET AL., 2005; CLEGG UND JACKSON, 1990), den TCA (SUMEGI ET AL., 1993), den Urea-Zyklus (COHEN UND KUDA, 1996) und verschiedenen Biosynthesewegen (SRERE, 1987) vermutet. ATKINSON (1977) beschreibt gar die Notwendigkeit von Channeling. Aufgrund der (*in vitro*) gemessenen K_M -Werte der Enzyme wären die intrazellulären Konzentrationen zu niedrig um eine ausreichende Aktivität hervorzurufen.

Eine weitere Schwierigkeit besteht in der *in vitro* Charakterisierung membrangebundener Enzyme, wie z.B. dem Zuckeraufnahmesystem (PTS) (HOFMANN UND KOPPERSCHLAGER, 1982). Diese Enzyme müssen zur *in vitro* Analyse zunächst aus der Membran gelöst werden. Ob das beobachtete Verhalten auf das *in vivo* Verhalten übertragen werden kann, ist zumindest fraglich.¹

Enzym-Modifikation: Zellen verfügen über eine post-translatorische Regulationsebene, die Enzymaktivitäten gezielt verändert (Interkonversion, Abb. 3.2). Enzyme können durch kovalente Reaktionen, wie z.B. Phosphorylierung, Methylierung, aktiviert oder inhibiert werden (BENDT ET AL., 2003; ARANDA ET AL., 2006). Die reaktionskinetische Beschreibung einer enzymatischen Reaktion hängt somit auch vom Zustand des Enzyms ab. Ergebnisse aus *in vitro* Untersuchungen, deren Fermentationsbedingungen (zur Herstellung des Enzyms) nicht vergleichbar sind, sollten deshalb mit Vorsicht übertragen werden.

Organisation des Stoffwechsels: Wissenschaftler versuchen auf verschiedenen Wegen, die grundlegende Konzepte bei der Organisation von metabolischen Netzwerken und deren Regulationsstrategien ausfindig zu machen. Allerdings muss man bei der Anwendung "menschlicher" Design Kriterien auf den bakteriellen Stoffwechsel darauf achten, dass die Organisation des Stoffwechsels nicht durch den Menschen und seine Herangehensweise entstanden ist. Werden also bestimmte zelluläre Ziele, wie z.B. bestmögliches Wachstum oder gute Anpassungsfähigkeit, als Kriterium unterstellt, wird unter Umständen eine falsche Hypothese zugrunde gelegt. Der Modellierer sollte deshalb versuchen, dem Stoffwechsel möglichst unvoreingenommen gegenüberzutreten und die Informationen aus den gemessenen Daten abzuleiten.

¹Die Untersuchung metabolischer Netzwerke *in vitro* lebt augenblicklich unter dem Schlagwort "Synthetic Biology" wieder auf (TORRES ET AL., 1986).

Aus diesen Gründen ergibt sich die Forderung, die Zelle als Gesamtsystem *in vivo* zu untersuchen. Nur auf diesem Weg wird es gelingen, die Zusammenhänge im lebenden Organismus zu entschlüsseln. Die experimentellen Zugänge dazu wurden in Kapitel 4 dargestellt. Zur Auswertung der aus den Experimenten gewonnen Messdaten, müssen einige vereinfachende Annahmen getroffen werden. Die wichtigsten, allgemein anerkannten sind:

Keine Konzentrations-Gradienten im Bioreaktor : Jede Zelle im Bioreaktor befindet sich unter den selben Umgebungsbedingungen. Es gibt im Fermentor räumlich keine Unterschiede in der Temperatur, der Nährstoffkonzentrationen, dem Druck und dem pH-Wert. Diese Annahme kann v.a. für kleine Reaktoren mit hoher Durchmischung getroffen werden.

Keine Konzentrations-Gradienten innerhalb der Zelle: Auch innerhalb der Zelle muss eine sofortige Durchmischung und eine homogene Stoffverteilung angenommen werden. Bei einem Durchmesser von ca. $2\ \mu\text{m}$ (*E. coli*) und der Annahme ungehinderter Diffusion stellen sich homogene Bedingungen extrem schnell ein (GERSHON ET AL., 1985). Die Gültigkeit dieser Abschätzung für das intrazelluläre Milieu mit Barrieren durch die hohe Dichte an Makromolekülen ist sicherlich fraglich (s. oben), muss jedoch getroffen werden um handhabbare Modelle zu erstellen.

Kein Metabolic Channeling: Eng verknüpft mit dem Konzentrations-Gradienten ist das "Metabolic Channeling". Channeling hätte enormen Einfluss auf die Modellbildung, da die Konzentrationen innerhalb eines Channels sich von denjenigen außerhalb wesentlich unterscheidet (SAVAGEAU, 1991). Trotz der o.g. Hinweise auf die Existenz von Channeling, ist die Annahme homogen verteilter Metabolite zwingend notwendig.

Durchschnittszelle: Es wird angenommen, dass sich die im Reaktor befindlichen Zellen alle gleich verhalten. Sie unterscheiden sich weder hinsichtlich ihres Zellzyklus, noch in der genetischen oder enzymatischen Ausstattung. Streng genommen kann diese Annahme nur für synchrone Kulturen getroffen werden. Allgemein wird bei "einfachen" Zellen wie Prokaryonten die Annahme einer Durchschnittszelle akzeptiert. Z.B. BROCK ET AL. (2003) beschreibt, dass sich die DNA von schnell wachsende *E.coli* Kulturen in ununterbrochener Teilung befinden. Dies spricht gegen die These ausgeprägter Zellzyklen mit unterschiedlichen Stoffflusslagen.

Keine Unterschiede in der Morphologie: Dieser Punkt ist eng an die Forderung der Durchschnittszelle geknüpft. Z.B. beim Wachstum von Pilzen und Hefen kann es zu unterschiedlich spezialisierten Zellen kommen²

Auf der Grundlage dieser Annahmen ergibt sich fast zwangsläufig eine Modellstruktur, die auf Massenbilanzen und Termen zur quantitativen Beschreibung der Stoffflüsse in Abhängigkeit der Stoffkonzentrationen aufbaut (Abschnitt 6.2–6.3). Die einzelnen Bausteine zur Modellerstellung, die Massenbilanzen (¹³C Massenbilanzen) sowie Kinetiken werden in den folgenden Abschnitten detailliert ausgearbeitet. Welche Gleichungen in das Modell einfließen, hängt von der Zielsetzung der Modellierung und dem durchgeführten Experimenten und verfügbaren Daten ab.

Im Metabolic Engineering werden vor allem folgende Modelle eingesetzt:

Makroskopische Modelle: Die Zelle wird als Black-Box aufgefasst, die eine Stoffumwandlung vornimmt (FREDRICKSON ET AL., 1970). Bilanziert werden die extrazellulären Konzentrationen im Reaktor sowie die Biomasse selbst. Diese Modelle werden meist zur Optimierung von Produktionsbedingungen eingesetzt (BRIK TERNBACH, 2005).

Metabolische Netzwerk-Modelle Bilanziert werden die intrazellulären Konzentrationen in der Zelle. Sie sind durch eine Vielzahl von Reaktionen miteinander verbunden, die das Reaktionsnetzwerk darstellen. Metabolische Netzwerke werden stationär und zunehmend auch dynamisch modelliert (RIZZI ET AL., 1997; LANGE ET AL., 2001; CHASSAGNOLE ET AL., 2002).

Ganz-Zell Modelle In die Modellbildung wird neben den Metaboliten auch die Transkription und Translation aufgenommen (PALSSON, 2004; DE JONG, 2002). Auf genetischer Ebene sind bereits eine Vielzahl von regulatorischen Mechanismen bekannt und in Datenbanken, wie z.B. KEGG (KANEHISA UND GOTO, 2000) abrufbar. Meist fehlt jedoch eine quantitative Beschreibung der regulatorischen Zusammenhänge.

Das Schema der Modellerstellung ist für die verschiedenen Typen vergleichbar (Abb. 7.1), sie unterscheiden sich jedoch wesentlich im Grad der Detaillierung und der Zielsetzung. Mit dem Modell werden zunächst die bekannten Zusammenhänge in Form von Kinetiken und Massenbilanzen formuliert. Im Modell sind somit zum einen bekannte (bzw.

²Bei der Modellierung der Penicilin-Produktion werden segregierte Modelle, die auch die Morphologie beschreiben eingesetzt (ZANGIROLAMI ET AL., 1997)

allgemein angenommene) Zusammenhänge, wie z.B. eindeutige Zuordnungen aufgrund der Stöchiometrie (Abschnitt 8.2) oder C-Atom-Transitionen in der ^{13}C Stoffflussanalyse (Abschnitt 6.5) enthalten. Zum anderen sind die für das Metabolic Engineering besonders interessanten, nicht direkt beobachtbaren Größen (intrazelluläre Flüsse, Aktivitäten und Affinitäten) als Unbekannte, aus den experimentellen Daten zu bestimmende Größen enthalten (Abb. 7.1).

Kapitel 6

Metabolische Netzwerk-Modelle

Akzeptiert man die zuvor getroffenen Annahmen und Vereinfachungen (Kapitel 5) können Modelle metabolischer Netzwerke auf Basis von Stoffflüssen und Poolgrößen aufgestellt werden. Dazu werden im Folgenden die verwendeten Modellgleichungen erläutert. Sie bilden die Grundlage für die Simulationen und Analysen der vorliegenden Arbeit. Die Lösung, Analyse und Interpretation der Gleichungen für die einzelnen Modelltypen wird im nächsten Kapitel beschrieben.

6.1 Beispielsystem zur Verdeutlichung der Modellerstellung

Zur Erläuterung wird das in Abb. 6.1 dargestellte kleine Beispielnetzwerk eingesetzt. Für die Notation werden folgende Symbole und Bezeichner verwendet:

Metabolite: Zur besseren Unterscheidung im Text werden für Metabolite Großbuchstaben in serifenloser Schrift verwendet, z.B. **A**. Die Konzentrationen werden in den Formeln mit den selben Symbolen dargestellt.

Flüsse: Flüsse werden mit Kleinbuchstaben ebenfalls in serifenloser Schrift dargestellt (z.B. v_1). In den Formeln hingegen wird der Fluss in Kursivschrift (v_1) angegeben.

Das Beispiel-Netzwerk ist eine Abstraktion des Zentralstoffwechsels - **A_{ex}** ist das extrazelluläre Substrat, die Metabolite **A** und **B** fassen die Glykolyse zusammen. Die Metabolite **C**, **D** und **E** stellen eine Abstraktion des Tricarbonsäure Zyklus (TCA) dar. Metabolit **F** steht für CO_2 , also Kohlenstoff, der das System verlässt. Ein weiterer Abfluss geht aus dem TCA in die Biomasse (hier zu **E_{ex}**). Auch bei den enzymatischen Reaktionen wurde versucht, ein vereinfachtes Abbild der im Stoffwechsel auftretenden Mechanismen wiederzugeben (Tab. C.1). Einige Reaktionsschritte sind reversibel, können also in beide Richtungen ablaufen (Tab. in Abb. 6.1). Die Reaktionsschritte v_{upt} , v_3 und v_4 werden durch ihre Produkte bzw. allosterisch reguliert. Die Substrat-Aufnahme v_{upt} wird durch das Produkt der Reaktion, **A**, inhibiert. Die Reaktion v_3 ist ebenfalls produktinhibiert, der Reaktionsschritt v_4 wird allosterisch durch Metabolit **C** inhibiert.

6.2 Massenbilanzen

Ein metabolisches Netzwerk ist eine Abstraktion der in der Zelle ablaufenden biochemischen Prozesse. Die in der Zelle frei verteilten Metabolite werden im metabolischen Netz-

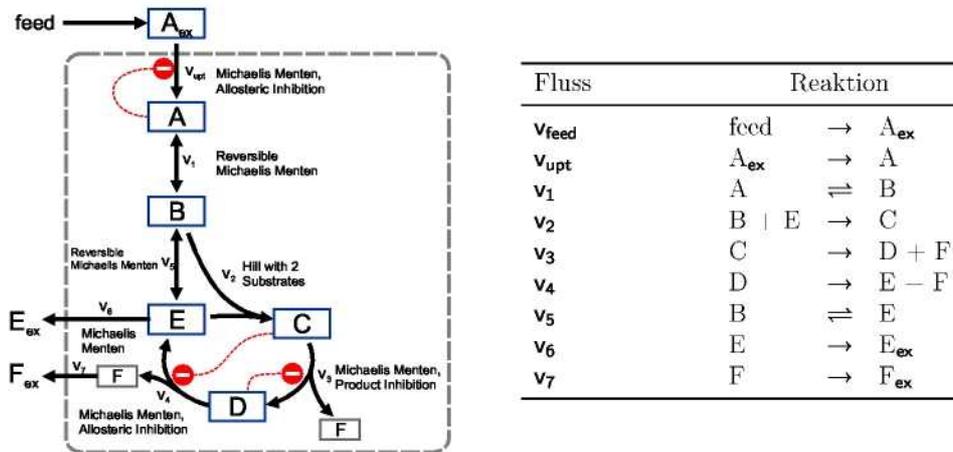


Abbildung 6.1: Beispielnetzwerk Spirallus mit den Reaktionsgleichungen.

werk als Speicher (Pool) aufgefasst, die Stoffumwandlungen können dann als Austausch zwischen den Pools dargestellt werden.

Bilanziert wird die Masse. Alle chemischen Substanzen, die von der Zelle aufgenommen werden befinden sich im System, werden umgewandelt oder verlassen das System wieder. Die Masse bleibt dabei erhalten. Damit kann für jeden Pool des metabolischen Netzwerks eine Bilanz aufgestellt werden. Sie beschreibt die Veränderung der Konzentration durch die Zu- und Abflüsse aus dem jeweiligen Pool. Zur Verdeutlichung wird die Konzentration des Pools B im Beispielnetzwerk (Abb. 6.1) betrachtet:

$$\dot{B} = v_1 - v_5 - v_2 \tag{6.1}$$

Die Massenbilanz wird für alle intrazellulären Pools aufgestellt. Besitzt eine Zelle Kompartimente, müssen die Volumina der Kompartimente beachtet werden. Mit dem Übergang in ein anderes Kompartiment wechselt der Bilanzraum. Im laufenden Beispiel werden auch die Konzentrationen im Reaktor bilanziert. Werden die Flüsse auf den intrazellulären Raum bezogen, ist für den Verbrauch des extrazellulären Substrats A_{ex} das Volumen der beiden Bilanzräume zu beachten:

$$v_{A_{\text{ex}} \rightarrow A} = \frac{V_{\text{intra}}}{V_{\text{reactor}}} v_{\text{upt}} \tag{6.2}$$

Die Massenbilanzen für das betrachtete Beispielsystem mit den extrazellulären Metaboliten \mathbf{A}_{ex} , \mathbf{E}_{ex} und \mathbf{F}_{ex} lauten somit (\dot{V}_{feed} ist der Volumenstrom des Feeds, c_{feed} die Konzentration des Feeds):

$$\begin{aligned}
 \dot{\mathbf{A}} &= v_{\text{upt}} - v_1 \\
 \dot{\mathbf{B}} &= v_1 - v_2 - v_5 \\
 \dot{\mathbf{C}} &= v_2 - v_3 \\
 \dot{\mathbf{D}} &= v_3 - v_4 \\
 \dot{\mathbf{E}} &= v_4 + v_5 - v_2 - v_6 \\
 \dot{\mathbf{F}} &= v_3 + v_4 - v_7
 \end{aligned}
 \quad
 \begin{aligned}
 \dot{\mathbf{A}}_{\text{ex}} &= v_{\text{feed}} - \beta v_{\text{upt}} \\
 \dot{\mathbf{E}}_{\text{ex}} &= \beta v_6 \\
 \dot{\mathbf{F}}_{\text{ex}} &= \beta v_7
 \end{aligned}
 \quad
 \begin{aligned}
 v_{\text{feed}} &= \frac{\dot{V}_{\text{feed}} c_{\text{feed}}}{V_{\text{reactor}}} \\
 \beta &= \frac{V_{\text{intra}}}{V_{\text{reactor}}}
 \end{aligned}
 \quad (6.3)$$

Alle aufgestellten Gleichungen können unter Verwendung einer Matrix-Notation dargestellt werden. Dazu werden die Systemvariablen (Metabolitkonzentrationen bzw. Poolkonzentrationen) sowie die Flüsse (Skalare oder Funktionen) in Vektoren zusammengefasst:

$$\begin{aligned}
 \mathbf{c} &= [\mathbf{A}_{\text{ex}}, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}, \mathbf{F}, \mathbf{E}_{\text{ex}}, \mathbf{F}_{\text{ex}}]^T \\
 \mathbf{v} &= [v_{\text{feed}}, v_{\text{upt}}, v_1, v_2, v_3, v_4, v_5, v_6, v_7]^T
 \end{aligned}$$

Zur Verknüpfung der Variablen zum Gleichungssystem (6.3) wird die stöchiometrische Matrix \mathbf{N} aufgestellt, die sich direkt aus den Massenbilanzen ableiten lässt:

$$\dot{\mathbf{c}} = \mathbf{N} \mathbf{v} \quad \text{mit: } \mathbf{N} = ()$$

6.3 Enzym-Kinetiken

Enzyme sind die Katalysatoren biochemischer Reaktionen. Sie binden meist sehr spezifisch ihre Substrate, wandeln sie um und geben das Produkt ab. Allosterisch regulierte Enzyme besitzen neben der Bindungsstelle für die Substrate weitere spezifische Bindungsstellen für Aktivatoren und Inhibitoren (Abb. 3.2). Bindet ein Molekül an dieser Stelle verändert das Enzym seine räumliche Gestalt und gewinnt oder verliert dadurch einen Teil seiner Aktivität (allosterische Inhibierung/Aktivierung). Die Umsetzungsgeschwindigkeit der enzymatisch katalysierten Reaktion kann je nach Mechanismus und Anzahl von Substraten, Inhibitoren und Aktivatoren von einer Vielzahl von Konzentrationen abhängen. In der Literatur werden verschiedene Ansätze zur Formulierung von Enzymkinetiken dargestellt. Am bekanntesten ist die mechanistische Herleitung, zurzeit werden aber auch verstärkt phänomenologische Ansätze diskutiert (z.B. LinLog-Kinetiken).

Mechanistische Enzym-Kinetik

Eine mechanistische Enzymkinetik leitet sich aus einem Modell des Enzymmechanismus ab. Modelliert werden die auftretenden Zustände, wie z.B.:

- freies Enzym E
- Enzym-Substrat-Komplex ES
- Enzym-Produkt-Komplex EP
- weitere Komplexe, z.B. Enzym-Inhibitor-Komplex EI

Meist wird nun das Massenwirkungsgesetz zur Beschreibung der Reaktionen zwischen den Komplexen angenommen. Ist die Substratkonzentration wesentlich höher als die Menge Katalysator E lassen sich die Gleichungen vereinfachen, da $d/dt = 0$ als gute Näherung eingesetzt werden kann. Mit dieser Näherung können die Konzentrationen von E, S, P und I analytisch bestimmt werden. Als Beispiel werden die Mechanismen der reversible Michaelis-Menten Kinetik sowie zwei Mechanismen mit Inhibitoren besprochen (Abb. 6.2).

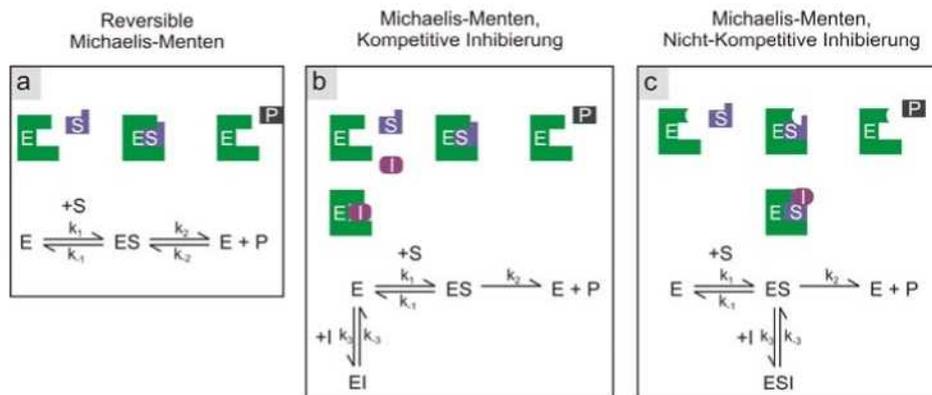


Abbildung 6.2: Schematische Darstellung verschiedener Enzym-Mechanismen. Aus diesen Modellen lassen sich die kinetischen Gleichungen ableiten. (a) Reversibler Michaelis-Menten Mechanismus, (b) Kompetitive Inhibition, (c) Nicht-Kompetitive Inhibition. Daneben gibt es viele weitere Mechanismen, u.a. auch gemischte Inhibition, die beide Inhibitor-Komplexe (EI und ESI in (b) und (c)) enthalten.

Die Herleitung der kinetischen Terme basiert wiederum auf den Massenbilanzen. Zumindest für einfache Mechanismen, wie die hier dargestellten, ist die Herleitung fast trivial und in Standardwerken wie z.B. CORNISH-BOWDEN (1995); BISSWANGER (2002) übersichtlich dargestellt.

Zur Lösung wird angenommen, dass sich für die Enzym-Substrat- bzw. Enzym-Produkt-Komplexe der stationäre Zustand sehr schnell einstellt und die Bildung der Komplexe keinen Einfluss auf die Substrat-, Produkt und Inhibitorkonzentration hat. Im Gleichgewicht sind die Konzentrationsänderungen der einzelnen Enzymkomplexe Null. Die Gleichungen werden nach der Produktbildungsrate umgestellt und die Parameter der Mas-

senwirkungskinetiken (meist mit k bezeichnet) mit Verhältnissen zusammengefasst. Die reversible Michaelis-Menten Kinetik lautet nach Umstellen und Ersetzen wie folgt:

$$v = v_{max} \frac{\left(S - \frac{P}{K_{eq}}\right)}{K_{M,S} + S + K_{M,S} \frac{P}{K_{M,P}}} \quad (6.4)$$

Die Mechanismen in Abb. 6.2a und b werden in analoger Weise behandelt. In den Massenbilanzen erscheint jeweils ein neuer Komplex (**EI** bzw. **ESI**). Nach Umstellen und Zusammenfassen der Parameter, ergeben sich die folgenden Kinetiken BISSWANGER (2002):

$$\begin{aligned} \text{b: } v &= v_{max} \frac{S}{K_{M,S} \left(1 + \frac{1}{K_I}\right) + S} \\ \text{c: } v &= v_{max} \frac{S}{K_{M,S} + \left(1 + \frac{1}{K_I}\right) S} \end{aligned}$$

Die in den mechanistischen Kinetiken enthaltenen Parameter lassen sich direkt im biochemischen Kontext interpretieren. Der v_{max} -Wert gibt die maximale katalytische Aktivität wieder, und ist proportional zur Enzymkonzentration. Die K_M und K_I -Werte sind Michaelis-Menten Konstanten und geben die Affinität des Enzyms gegenüber dem Substrat, Produkt oder Inhibitor wieder. Im Fall der einfachen, nicht inhibierten Michaelis-Menten Kinetik wird bei $S = K_{M,S}$ die halbe maximale Reaktionsgeschwindigkeit $v = 0.5 v_{max}$ erreicht.

6.4 Vereinfachte Kinetische Ansätze

Die Herleitung einer mechanistischen Kinetik aus den einzelnen Komplexen wird für mehrere Substrate, Produkte bzw. Modifikatoren sehr aufwendig (z.B. Verfahren nach King-Altman (BISSWANGER, 2002)). Zudem werden zur Beschreibung eine Vielzahl von Parametern (Gleichgewichtskonstanten, etc.) benötigt. All diese Parameter aus den experimentell verfügbaren Daten zu bestimmen gelingt nur noch in wenigen Fällen.

Deshalb wird versucht, vereinfachte Kinetiken einzusetzen, die das Verhalten des wahren Mechanismus approximieren. Die Komplexität der mechanistischen Kinetik entsteht vor allem durch die gegenseitige Beeinflussung der beteiligten Substrate, Produkte und Modifikatoren. Wird angenommen, die gegenseitigen Beeinflussungen sind gering, können die einzelnen Substanzen als unabhängig angesehen werden. Ihr Einfluss auf die Reaktionsgeschwindigkeit kann dann durch wesentlich weniger Parameter beschrieben werden.

In der Literatur werden vor allem drei approximative Ansätze diskutiert, so genannte LinLog Kinetiken, generische Michaelis-Menten Ansätze und PowerLaw Terme.

6.4.1 Approximation mit generischen Michaelis-Menten Termen

Der generische Michaelis-Menten Ansatz beruht auf der einfachen Michaelis-Menten Kinetik. Nehmen mehrere Substanzen Einfluss auf die enzymatische Umsetzungsgeschwindigkeit, werden sie als unabhängig wirkend angenommen. Ausgehend von der einfachen irreversiblen (oder auch reversiblen) Michaelis-Menten Kinetik werden weitere Substrate \mathbf{S} , Produkte \mathbf{P} , Inhibitoren \mathbf{I} und Aktivatoren \mathbf{A} multiplikativ hinzugenommen. Die approximierte Kinetik kann somit wie folgt dargestellt werden:

$$v = v_{max} \frac{S_1}{K_{M,S_1} + S_1} \frac{S_2}{K_{M,S_2} + S_2} \cdots \left(1 - \frac{I_1}{I_1 + K_{I_1}}\right) \left(1 - \frac{I_2}{I_2 + K_{I_2}}\right) \cdots \left(1 + \frac{a_1 A_1}{A_1 + K_{A_1}}\right) \left(1 + \frac{a_2 A_2}{A_2 + K_{A_2}}\right) \cdots \quad (6.5)$$

Dieser Ansatz bildet wichtige Eigenschaften der Michaelis-Menten-Kinetik, wie z.B. Sättigung, ab. Durch die multiplikative Verknüpfung der Terme entstehen allerdings schnell Artefakte, beispielsweise wird eine kompetitive Inhibierung bei hoher Substratkonzentration $S \gg K_{M,S}$ wenig inhibieren, während durch den generischen Produktterm $\left(1 - \frac{I}{I+K_I}\right)$ schnell eine vollständige Inhibierung eintritt.

6.4.2 LinLog Approximation

In den letzten Jahren haben sich vor allem logarithmische Ansätze (WESTERHOFF UND VAN DAM, 1987; NIELSEN, 1997; VISSER UND HEIJNEN, 2003; VISSER ET AL., 2004a; HEIJNEN, 2005) zur Approximierung mechanistischer Kinetiken durchgesetzt. Meist berufen sich die Autoren auf Beobachtungen von ONSAGER (1937) zurück. Die genaue Herleitung ist ausführlich in WESTERHOFF UND VAN DAM (1987) dargestellt. In dieser Arbeit sei lediglich auf den Bezug zur Thermodynamik verwiesen. Die thermodynamische Triebkraft ($\Delta_r G$) einer Reaktion mit den Substrat- und Produktkonzentrationen (S_i , $i = 1 \dots n_S$) und P_k , $k = 1 \dots n_P$ ist laut ALBERTY (2003):

$$\begin{aligned} \Delta_r G &= G_{P_1}^0 + G_{P_2}^0 + \dots - G_{S_1}^0 - G_{S_2}^0 - \dots + RT \ln \frac{P_1 P_2 \dots}{S_1 S_2 \dots} \\ &= \Delta G^0 + RT \ln \frac{P_1 P_2 \dots}{S_1 S_2 \dots} \end{aligned} \quad (6.6)$$

Die thermodynamische Triebkraft wird von den Stoffeigenschaften (G^0) und den Konzentrationen bestimmt. ONSAGER (1937) beschreibt, dass die Geschwindigkeit der Reaktion $S_1 + \dots + S_{n_S} \rightleftharpoons P_1 + \dots + P_{n_P}$ nahe des Gleichgewichts proportional zur Affinität ($A = RT \ln(K_{eq} P S^{-1})$).

Die Geschwindigkeit enzymatisch katalysierte Reaktionen hängt neben der thermodynamischen Triebkraft auch von regulatorischen Einflüssen ab. Um regulatorische Einflüsse von Inhibitoren und Aktivatoren in eine approximierete Kinetik aufzunehmen, schlagen WESTERHOFF UND VAN DAM (1987) einen zu Substraten bzw. Produkten analogen Term mit empirischen Koeffizienten vor (z.B. für einen Inhibitor $i_1 \ln I_1$ in Gl. 6.7). Der allgemeine Ansatz lautet dann (HEIJNEN, 2005):

$$v = e (k + s_1 \ln S_1 + \dots s_{n_S} \ln S_{n_S} - p_1 \ln P_1 - \dots - p_{n_P} \ln P_{n_P} + a_1 \ln A_1 + \dots a_{n_A} \ln A_{n_A} - i_1 \ln I_1 - \dots i_{n_I} \ln I_{n_I}) \quad (6.7)$$

Dieser Ansatz wurde von WU ET AL. (2005) aufgenommen und durch Einführen eines Referenz-Steady-States in ein leichter zu interpretierendes Format überführt. Durch die Normierung der Flüsse \mathbf{v} mit \mathbf{v}^0 (steady-state Fluss), der Enzymkonzentrationen \mathbf{e} mit \mathbf{e}^0 (Referenz Enzymkonzentration) sowie der Metabolitkonzentrationen \mathbf{c} mit \mathbf{c}^0 ergibt sich eine Verbindung zur Metabolic Control Analysis (MCA, Abschnitt 8.6.2) (HEINRICH ET AL., 1977; KACSER UND BURNS, 1973; HEIJNEN, 2005). Die in Gl. (6.8) auftauchenden Matrizen \mathbf{E}^0 enthalten dann die Elastizitäten ϵ^0 auf den Diagonalen:

$$\frac{\mathbf{v}}{\mathbf{v}^0} = \begin{bmatrix} \mathbf{e} \\ \mathbf{e}^0 \end{bmatrix} \left(\mathbf{1}^T + \mathbf{E}_x^0 \ln \left(\frac{\mathbf{c}}{\mathbf{c}^0} \right) + \mathbf{E}_{c^0} \ln \left(\frac{\mathbf{c}_{\text{ex}}}{\mathbf{c}_{\text{ex}}^0} \right) \right) \quad (6.8)$$

PowerLaw Approximation

Etwas älter als der LinLog Ansatz ist der von SAVAGEAU (1970) geprägte “Power-Law” Ansatz. In seiner “Biological System Theory” wird die Reaktionsgeschwindigkeit v phänomenologisch aus Produkten von Potenztermen der beteiligten Metabolite dargestellt:

$$v = s_1 S_1^{\alpha_1} s_2 S_2^{\alpha_2} \dots p_1 P_1^{-\beta_1} p_2 P_2^{-\beta_2} \dots \quad (6.9)$$

Ein Vorteil dieser Darstellung ist die Möglichkeit, durch Logarithmieren lineare Gleichungen zu erhalten. Da dieser Ansatz bei der Reproduktion der Messdaten keine zufriedenstellenden Ergebnisse lieferte, wird auf eine weitere Darstellung verzichtet.

6.4.3 Zusammenfassen der stöchiometrischen und kinetischen Gleichungen

Mit den bisher beschriebenen Formeln kann nun ein vollständiges Modell aufgestellt werden, dass sowohl die stöchiometrischen als auch kinetischen Gleichungen enthält. Es entsteht ein nicht-lineares Differentialgleichungssystem, dass das dynamische Verhalten des metabolischen Netzwerks abbilden kann. Mit den in Abschnitt 6.2 eingeführten Vektoren

\mathbf{c} und \mathbf{v} , den im Parametervektor $\boldsymbol{\alpha}$ zusammengefassten kinetischen Parametern sowie den gegebenen experimentellen Parametern $\boldsymbol{\alpha}^{\text{inp}}$

$$\begin{aligned}\mathbf{c} &= [\mathbf{A}_{\text{ex}}, \mathbf{A}, \mathbf{B}, \dots, \mathbf{F}_{\text{ex}}]^T \\ \mathbf{v} &= [v_{\text{upt}}, \dots, v_7]^T \\ \boldsymbol{\alpha} &= [{}^{\text{upt}}v_{\text{max}}, {}^{\text{upt}}K_{M,A}, {}^{\text{upt}}K_I, \dots, {}^{v7}v_{\text{max}}]^T \\ \boldsymbol{\alpha}^{\text{inp}} &= [V_{\text{intra}}, V_{\text{reactor}}, {}^{\text{feed}}v(t)]^T\end{aligned}$$

kann das allgemeine Modell kompakt aufgestellt werden:

$$\dot{\mathbf{c}} = \mathbf{N} \mathbf{v}(\mathbf{c}, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}) \quad (6.10)$$

Die Lösung dieses nicht-linearen Differentialgleichungssystems ist mit numerischen Verfahren möglich (Abschnitt 8.6).

6.5 Bilanzen für ^{13}C Markierungszustände

Zur Beschreibung eines ^{13}C -Markierungsexperiments werden neben den Konzentrationen auch die Markierungszustände der einzelnen Metabolite bilanziert (Abb. 4.1 in Anhang C). Diese Bilanzen geben den Zusammenhang zwischen Flüssen (bzw. Kinetiken), Konzentrationen und Markierungszuständen quantitativ wieder.

Ein Metabolit X mit n C-Atomen kann 2^n verschiedene Markierungsmuster aus ^{13}C und ^{12}C Atomen annehmen (Abb. 6.3b). Die Markierungsmuster bilden die Isotopomere, die in dieser Arbeit in der Notation von WIECHERT UND GRAAF (1997) dargestellt werden. Als Beispiel wird der Metabolit D betrachtet – er besitzt 2^3 Isotopomeren-Zustände, die Bezeichnung D#000 wird für D mit nur unmarkierten ^{12}C Atomen verwendet. D#001 bedeutet, dass C₃ ein markiertes ^{13}C Atom ist, C₁ und C₂ sind unmarkiert. Der vollständig markierte Zustand wird mit D#111 dargestellt.

Zur Bilanzierung der Markierungszustände werden alle Isotopomere eines Pools betrachtet und die Menge aller in diesen Pool ein- und austretenden Isotopomere bilanziert. Bei der Bilanzierung werden folgende Notationen verwendet:

B#ij : Bezeichnet das Isotopomer mit der Markierung #ij.

B_{ij}: enthält die Konzentration des Isotomers B#ij.

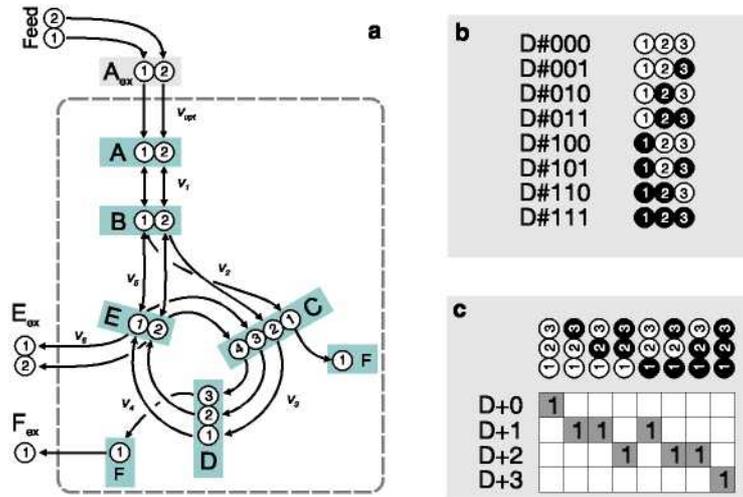


Abbildung 6.3: a) C-Atom Transitionsnetzwerk: Dieses Netzwerk gibt die Transitions (Übergänge) der C-Atome in den einzelnen Reaktionsschritten wieder. b) Isotopomeren-Zustände für den C3-Körper D. c) Aufbau einer Messmatrix für Massenisotopomere: Zuordnung der Isotopomeren-Zustände zu den Massenisotopomeren.

B: enthält die Konzentration des Pools B, d.h. der Summe aller Isotopomeren-Pools $B_{\#ij}$.

$$B = \sum_{i=0}^1 \sum_{j=0}^1 B_{ij}$$

b_{ij} : beschreibt den Anteil des Isotopomers $B_{\#ij}$ am Gesamtpool B:

$$b_{ij} = \frac{B_{ij}}{B}$$

Grundlage der Bilanzierung ist das C-Atom Transitionsnetzwerk (Abb. 6.3a), das die Transitions der einzelnen C-Atome in den einzelnen Reaktionsschritten wiedergibt. Im Gegensatz zur rein stöchiometrischen Stoffflussanalyse können auch bidirektionale Flüsse bestimmt werden, d.h. Flüsse von Reaktionen, die zur selben Zeit vorwärts und rückwärts ablaufen (v^{\rightarrow} , v^{\leftarrow}).

Zum besseren Verständnis der Bilanzierung werden die Gleichungen für die Pools $B_{\#ij}$ und $C_{\#ijkl}$ aufgestellt. Die Veränderung der Konzentration B_{ij} der Isotopomeren-Pools ergibt sich aus den jeweiligen Zu- und Abflüssen:

- Der Zufluss $v_{B_{\#ij}}^{\rightarrow}$ in den Isotopomeren-Pool $B_{\#ij}$ über v_1^{\rightarrow} entspricht genau dem Anteil des Isotopomeren-Pools $A_{\#ij}$ am Gesamtpool A multipliziert mit dem v_1^{\rightarrow} Fluss sowie dem Zufluss aus $E_{\#ij}$ über den Fluss v_5^{\leftarrow} :

$$\begin{aligned} \text{Zufluss in } B\#ij : v_{B\#ij}^{\vec{}} &= \frac{A_{ij}}{A} v_1^{\vec{}} + \frac{E_{ij}}{E} v_5^{\leftarrow} \\ &= a_{ij} v_1^{\vec{}} + e_{ij} v_5^{\leftarrow} \end{aligned}$$

- Der Abfluss aus $B\#ij$ über $v_2^{\vec{}}$ ergibt sich aus dem Anteil b_{ij} multipliziert mit dem Fluss $v_2^{\vec{}}$. Analog dazu wird der Abfluss über $v_5^{\vec{}}$ und den Rückfluss v_1^{\leftarrow} aufgestellt:

$$\text{Abfluss aus } B\#ij : v_{B\#ij}^{\leftarrow} = b_{ij} (v_2^{\vec{}} + v_5^{\vec{}} + v_1^{\leftarrow})$$

Die Differentialgleichungen für die Isotopomeren $B\#ij$ lauten folglich:

$$\dot{B}_{ij} = a_{ij} v_1^{\vec{}} + e_{ij} v_5^{\leftarrow} - b_{ij} (v_2^{\vec{}} + v_5^{\vec{}} + v_1^{\leftarrow})$$

Etwas komplexer sind die Gleichungen für bimolekulare Reaktionsschritte. Hier müssen alle möglichen Kombinationen von Isotopomeren-Substraten in Betracht gezogen werden. Beispielhaft wird der Isotopomeren-Pool $C\#0110$ bilanziert. Die Wahrscheinlichkeit, dass die beiden notwendigen Isotopomeren-Substrate $B\#01$ und $E\#10$ miteinander reagieren und über v_2 $C\#0110$ bilden, ist über das Produkt ihrer Anteile, $b_{01} e_{10}$ gegeben. Folglich können die Zuflüsse in den Isotopomeren-Pools $C\#ijkl$ mit den jeweiligen Produkten der Anteile b_{ij} und e_{kl} aufgestellt werden:

$$\text{Zufluss in } C\#ijkl : v_{C\#ijkl}^{\vec{}} = (b_{ij} e_{kl}) v_2^{\vec{}}$$

Gehen aus einer Reaktion zwei oder mehr Produkte hervor, treten Summen von Isotopomerenanteilen auf. Beispielhaft wird die Reaktion v_3 betrachtet. v_3 führt von C zu D und F . Das Isotopomer $D\#001$ kann aus $C\#0010$ und $C\#0011$ hervorgehen. Im Zufluss summieren sich somit die Anteile $c_{0010} + c_{0011}$. Analog dazu wird der Zufluss nach $F\#$ aufgestellt. Nach $F\#0$ fließt die Summe der Anteile $c_{0000} + c_{0001} + c_{0010} + c_{0011} + c_{0100} + c_{0101} + c_{0110} + c_{0111}$:

$$\text{Zufluss in } F\#i \text{ aus } C\#ijkl : v_{F\#i}^{\vec{}} = \sum_{j=0}^1 \sum_{k=0}^1 \sum_{l=0}^1 c_{ijkl} v_3^{\vec{}}$$

$$\text{Zufluss in } D\#jkl : v_{D\#jkl}^{\vec{}} = \sum_{i=0}^1 c_{ijkl} v_3^{\vec{}}$$

Wie bereits bei den stöchiometrischen Gleichungen bietet sich die Einführung einer Matrix-Notation der Isotopomeren-Bilanzgleichungen an. Folgende Vektoren werden eingeführt:

Konzentrationen der einzelnen Isotopomere, \mathbf{x}_c : An Stelle der Poolgrößen \mathbf{c} in den stöchiometrischen Modellen, tritt der Vektor mit den Konzentrationen der einzelnen Isotopomere:

$$\mathbf{x}_c = [\mathbf{A}_{\text{ex},00}, \mathbf{A}_{\text{ex},01}, \mathbf{A}_{\text{ex},10}, \mathbf{A}_{\text{ex},11}, \mathbf{A}_{00}, \mathbf{A}_{01}, \dots, \mathbf{F}_{\text{ex}10}, \mathbf{F}_{\text{ex}11}]^T$$

Flüsse: \mathbf{v} Der Vektor mit den Flüssen \mathbf{v} erweitert sich um die Rückreaktionen bzw. Austauschflüsse:

$$\mathbf{v} = [v_{\text{upt}}^{\rightarrow}, v_1^{\rightarrow}, v_1^{\leftarrow}, v_2^{\rightarrow}, \dots, v_7^{\rightarrow}]^T$$

Parameter α : Der Vektor der Parameter enthält alle kinetischen Parameter:

$$\alpha = [v_{\text{max}}^{v1}, K_{mA}^{v1}, K_1^{v1}, \dots, v_{\text{max}}^{v7}]^T$$

Input-Parameter α^{inp} : Dieser Vektor enthält Parameter die durch das Experiment vorgegeben, d.h. in der Parameteranpassung nicht angepasst werden, hier z.B. das Input Markierungsmuster, den Feedstrom, etc.:

$$\alpha^{\text{inp}} = [v_{\text{feed}}^{\rightarrow}, \text{feed}_{00}, \text{feed}_{01}, \text{feed}_{10}, \text{feed}_{11}]^T$$

Das Modell zur Simulation der Konzentrationsverläufe der Isotopomeren, kann somit wie folgt dargestellt werden:

$$\dot{\mathbf{x}}_c = f(\alpha, \alpha^{\text{inp}})$$

In die Funktion f gehen die kinetischen Terme als auch die C-Atom Transitionen ein. Eine kompakte Darstellung der Gleichungen, vergleichbar zu Systemen ohne Markierung gelingt nicht mehr so leicht. Aufgrund der bei bimolekularen Reaktionen auftretenden bilinearen Terme reicht zur vollständigen Darstellung der Gleichungen eine zweidimensionale Matrix nicht mehr aus. WIECHERT (1995) führte daher eine dreidimensionale Matrix ein, deren Aufbau im Rahmen dieser Arbeit nicht notwendig ist, da das System in anderen Koordinaten wieder mit zweidimensionalen Matrizen beschreibbar wird (Abschnitt 8.4).

Kapitel 7

Simulation, Systemanalyse und Statistische Methoden

7.1 Schema der Simulation und Systemanalyse

Im vorangegangenen Kapitel wurden die Grundlagen der Modellbildung dargestellt. Dieses Kapitel befasst sich nun mit der Simulation, Systemanalyse und statistischen Auswertung. Das Vorgehen wird allgemeingültig, d.h. auf alle in dieser Arbeit auftretenden Modelle formuliert (Abb. 7.1).

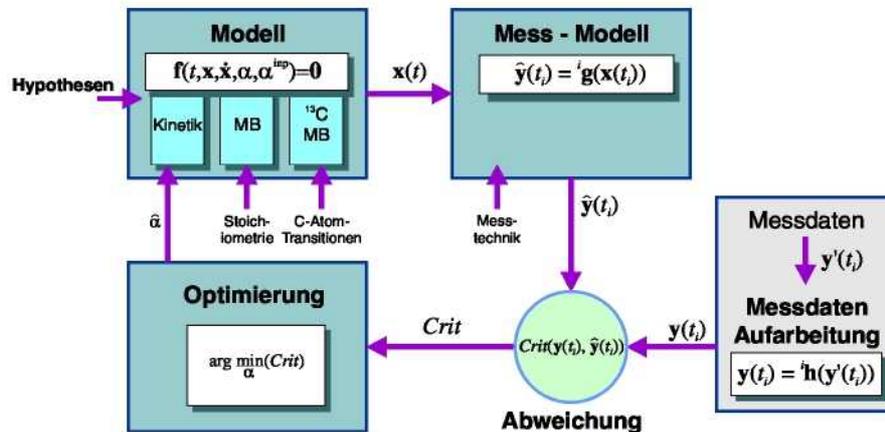


Abbildung 7.1: Allgemeines Schema des Modellaufbaus: In die Funktion \mathbf{f} gehen die Stöchiometrie (Massenbilanzen MB), die angenommenen Kinetiken sowie die C-Atom-Transitionen ein. Der Zusammenhang zwischen Modell und Messdaten wird über Messmodelle ${}^i\mathbf{g}$ hergestellt. Da die Proben teilweise in mehreren Schritten aufbereitet werden müssen, oder verzerrt vorliegen, ist unter Umständen ein Hilfsmodell \mathbf{h} zur Abbildung der Probenaufarbeitung und Korrektur notwendig. Aus den Abweichungen von beobachteten und simulierten Verhalten wird eine Maßzahl Crit bestimmt, die zur Anpassung der Parameter $\hat{\alpha}$ herangezogen wird.

Die aufgestellten Gleichungen lassen sich als implizite Funktionen \mathbf{f} in Abhängigkeit der Zeit t , den Parametern α , den Eingangsgrößen α^{inp} und den Zustandsvariablen \mathbf{x} , bzw. deren Ableitungen $\dot{\mathbf{x}}$ darstellen:

$$\mathbf{f}(t, \alpha, \alpha^{\text{inp}}, \mathbf{x}, \dot{\mathbf{x}}) = \mathbf{0} \quad (7.1)$$

Aus der implizit gegeben Funktion werden die Systemzustände $\mathbf{x}(t)$ bestimmt (s. Simulation),

$$\mathbf{x} = \mathbf{x}(t, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0), \quad \text{mit: } \mathbf{x}(t=0) = \mathbf{x}_0 \quad (7.2)$$

Meist können die Zustände \mathbf{x} , bzw. $\dot{\mathbf{x}}$, nicht direkt beobachtet werden. Um das Systemmodell mit den messbaren Größen \mathbf{y} zu verbinden, wird ein Modell für die Messung benötigt. Die Funktionen ${}^i\mathbf{g}$ berechnen für den Systemzustand $\mathbf{x}(t)$ zum Zeitpunkt t_i die erwarteten Messungen $\hat{\mathbf{y}}(t_i)$:

$$\hat{\mathbf{y}}(t_i) = {}^i\mathbf{g}(\mathbf{x}(t_i, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0)) \quad (7.3)$$

Die Funktionen \mathbf{f} und ${}^i\mathbf{g}$ bilden das Modell. Für die Anwendung müssen folgende Problemstellungen gelöst werden:

Simulation: Aufgabe der Simulation ist bei gegebener Funktion \mathbf{f} (Gl. (7.1)) die Lösung für \mathbf{x} und $\hat{\mathbf{y}}$ (Gl. 7.2) zu berechnen.

| | |
|---|--|
| Gegeben: $\boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0$ | Gesucht: $\mathbf{x} = \mathbf{x}(t, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0)$ $\mathbf{y} = {}^i\mathbf{g}(\mathbf{x}(t_i, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0))$ |
|---|--|

Zum Einsatz kommen je nach Problemstellung verschiedene Methoden:

- Lineare Systeme können z.B. mit dem Gauß-Verfahren gelöst werden.
- Für die Lösung nichtlinearer Systeme werden z.B. das Newton-Verfahren oder andere iterative Lösungsverfahren verwendet.
- Liegt ein Differentialgleichungssystem vor, werden Verfahren zur numerischen Integration, z.B. das Euler-Verfahren, Runge-Kutta oder andere eingesetzt.

Die Berechnung der (erwarteten) Messungen $\hat{\mathbf{y}}(t_i)$ ergibt sich durch Einsetzen des Systemzustands \mathbf{x} (Gl. (7.3)).

Parameteranpassung: Aufgabe der Parameteranpassung ist es, aus N gegebenen Messreihen $\mathbf{y}(t_i)$ (zu mehreren Zeitpunkten $\boldsymbol{\xi} = (t_1, \dots, t_N)^T$) und gegebenen Eingangsgößen $\boldsymbol{\alpha}^{\text{inp}}$ die Parameter bestmöglich zu schätzen.

| | |
|---|---|
| Gegeben: $\mathbf{y}(t_i), \boldsymbol{\alpha}^{\text{inp}}$ | Gesucht: $\hat{\boldsymbol{\alpha}}, \hat{\mathbf{x}}_0$ |
|---|---|

Diese Problemstellung wird als eine Minimierung des Kriteriums *Crit* formuliert. Meist wird die gewichtete Fehlerquadratsumme (FQS) als Kriterium der Abweichung verwendet. Hierbei wird die Abweichung zwischen der Modellberechnung und

dem Messwert anhand der Varianz gewichtet, d.h. Messwerte mit kleinen Messfehlern erhalten einen höheren Einfluss auf die FQS als ungenaue Messwerte. Die Gewichtung wird in Gl. 7.4 durch die Doppelstriche und die Kovarianzmatrix $\Sigma_{\mathbf{y}(t_i)}$ dargestellt. Ein sehr übliches Vorgehen ist, die Fehlerquadratsumme aus den Quadraten der Residuen geteilt durch die Varianzen zu berechnen. Bei der Minimierung der FQS müssen die Systemgleichungen $\mathbf{f}(t, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}, \dot{\mathbf{x}}) = \mathbf{0}$ erfüllt sein. Sie werden als Nebenbedingung gesetzt.

$$\begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\mathbf{x}}_0 \end{pmatrix} = \arg \min_{\substack{\boldsymbol{\alpha}, \mathbf{x}_0 \\ \mathbf{f}(t, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}, \dot{\mathbf{x}}) = \mathbf{0}}} \sum_{i=1}^N \|\mathbf{y}(t_i) - {}^i\mathbf{g}(\mathbf{x}(t_i, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0))\|_{\Sigma_{\mathbf{y}(t_i)}}^2 \quad (7.4)$$

Zum Einsatz kommen hierzu

- Lineare Regression,
- Ableitungsfreie iterative Optimierungsverfahren, wie z.B. das Simplex-Verfahren (MICHALEWICZ UND FOGEL, 2002) oder evolutionäre bzw. genetische Optimierungsverfahren (BEYER, 2001).
- Ableitungsgestützte iterative Optimierungsverfahren, z.B. das Gradienten- oder Newton-Verfahren (DEUFLHARD, 1995).

χ^2 -Test: Der χ^2 -Test wird zur Überprüfung und Beurteilung der Parameteranpassung eingesetzt. Der χ^2 -Wert ist die gewichtete Fehlerquadratsumme mit den geschätzten Parameterwerten $\hat{\boldsymbol{\alpha}}$ und $\hat{\mathbf{x}}_0$:

$$\chi^2 = \sum_{i=1}^N \|\mathbf{y}(t_i) - {}^i\mathbf{g}(\mathbf{x}(t_i, \hat{\boldsymbol{\alpha}}, \boldsymbol{\alpha}^{\text{inp}}, \hat{\mathbf{x}}_0))\|_{\Sigma_{\mathbf{y}(t_i)}}^2 \quad (7.5)$$

Sensitivitätsanalyse: Die Sensitivität beschreibt den Einfluss der Parameter $\boldsymbol{\alpha}$ oder der Anfangswerte \mathbf{x}_0 auf das Systemverhalten $\mathbf{x}(t)$ bzw. die erwarteten Messungen $\hat{\mathbf{y}}$.

| | |
|---|--|
| Gegeben: $\boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0$ | Gesucht: $\text{Sens}_{\boldsymbol{\alpha}}^{\mathbf{x}}(t), \text{Sens}_{\mathbf{x}_0}^{\mathbf{x}}(t)$ $\text{Sens}_{\boldsymbol{\alpha}}^{\mathbf{y}}(t), \text{Sens}_{\mathbf{x}_0}^{\mathbf{y}}(t)$ |
|---|--|

Die Sensitivitäten berechnen sich aus der Ableitungen des Systemzustands $\mathbf{x}(t)$ nach den Parametern, bzw. Anfangswerten:

$$\text{Sens}_{\boldsymbol{\alpha}}^{\mathbf{x}}(t) = \frac{d\mathbf{x}}{d\boldsymbol{\alpha}} \quad \text{Sens}_{\mathbf{x}_0}^{\mathbf{x}}(t) = \frac{d\mathbf{x}}{d\mathbf{x}_0} \quad (7.6)$$

Die benötigten Ableitungen werden aus den implizit gegebenen Systemgleichungen (7.1) nach der Kettenregel gebildet:

$$\frac{df}{d\alpha} + \frac{\partial f}{\partial t} \frac{dt}{d\alpha} + \frac{\partial f}{\partial \alpha^{\text{inp}}} \frac{d\alpha^{\text{inp}}}{d\alpha} + \frac{\partial f}{\partial \mathbf{x}} \frac{d\mathbf{x}}{d\alpha} + \frac{\partial f}{\partial \mathbf{x}_0} \frac{d\mathbf{x}_0}{d\alpha} + \frac{\partial f}{\partial \dot{\mathbf{x}}} \frac{d\dot{\mathbf{x}}}{d\alpha} = 0 \quad (7.7)$$

Viele der auftretenden Ableitungen sind Null. Keine Rolle spielen in den hier untersuchten Modellen die Ableitungen $d\alpha^{\text{inp}}/d\alpha$, $d\mathbf{x}_0/d\alpha$ und $dt/d\alpha$. Vereinfacht lautet das Gleichungssystem dann:

$$\frac{df}{d\alpha} + \frac{\partial f}{\partial \mathbf{x}} \frac{d\mathbf{x}}{d\alpha} + \frac{\partial f}{\partial \dot{\mathbf{x}}} \frac{d\dot{\mathbf{x}}}{d\alpha} = \mathbf{0} \quad (7.8)$$

Die Ableitungen $d\dot{\mathbf{x}}/d\alpha$ beinhalten die Ableitungen nach der Zeit t und den Parameter α . Die Reihenfolge der Ableitungen ist vertauschbar:

$$\frac{d\dot{\mathbf{x}}}{d\alpha} = \frac{d}{dt} \left(\frac{d\mathbf{x}}{d\alpha} \right) = \left(\frac{d\dot{\mathbf{x}}}{d\alpha} \right)$$

Die Gleichungen werden umgestellt und es ergibt sich ein Differentialgleichungssystem, in das die Lösung $\mathbf{x}(t)$ eingeht:

$$\frac{d\dot{\mathbf{x}}}{d\alpha} = -\frac{\partial f}{\partial \dot{\mathbf{x}}}^{-1} \left(\frac{\partial f}{\partial \mathbf{x}} \frac{d\mathbf{x}}{d\alpha} + \frac{df}{d\alpha} \right)$$

Diese Differentialgleichung wird wiederum in impliziter Darstellung wiedergegeben:

$$\mathbf{f}'(t, \alpha, \alpha^{\text{inp}}, \mathbf{x}, \frac{d\mathbf{x}}{d\alpha}, \left(\frac{\dot{\mathbf{x}}}{\alpha} \right)) = \mathbf{0}$$

Meist wird die Sensitivität der Messungen $\mathbf{y}(t_i)$ gegenüber den Parametern benötigt. Durch Anwenden der Kettenregel erhält man:

$$\text{Sens}_{\alpha}^{\mathbf{y}}(t) = \frac{d\mathbf{y}}{d\alpha} = \frac{d\mathbf{y}}{d\mathbf{x}} \frac{d\mathbf{x}}{d\alpha}$$

Fehlerübertragung: Zusammen mit der Sensitivitätsanalyse bildet die Fehlerübertragung ein wesentliches Werkzeug im Rahmen dieser Arbeit. Allgemein lässt sich mit der Fehlerübertragung bestimmen, wie sich Fehler in Eingangsgrößen auf die berechneten Größen auswirken. Anhand der Fehlerrechnung können auch Aussagen über die Identifizierbarkeit und die Abhängigkeiten der Ausgangsgrößen getroffen werden. Das Vorgehen wird in Abschnitt 7.2.2 verdeutlicht.

In der vorliegenden Arbeit werden die Lösungen für

1. stationäre stöchiometrische Modelle (Abschnitt 8.2),
2. stationäre Markierungsexperimente (Abschnitt 8.4),
3. metabolisch dynamische Modelle (Abschnitt 8.6) sowie
4. metabolisch und isotopisch dynamische Modelle (Abschnitt 17)

dargestellt.

7.2 Fehlerübertragung

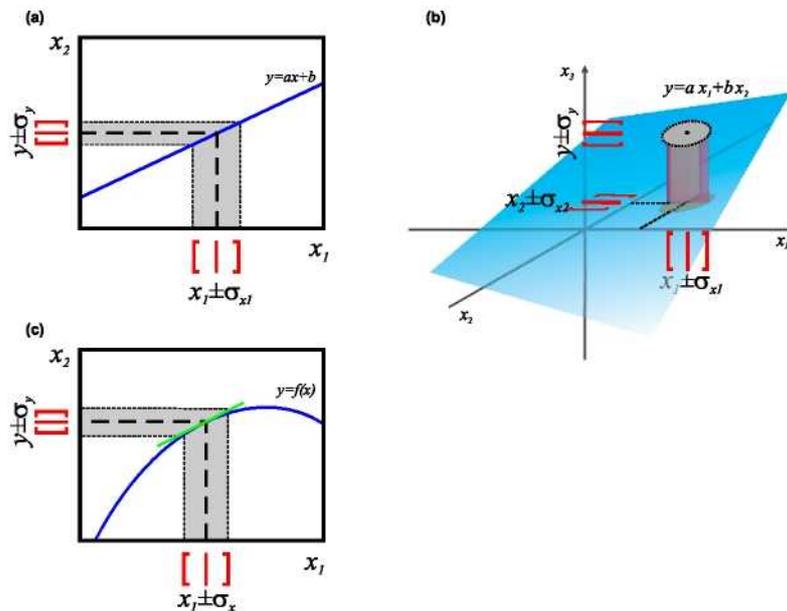


Abbildung 7.2: Fehlerübertragung: (a) für eine lineare Funktion $y = a x$, (b) eine lineare Funktion $y = a x_1 + b x_2$. (c) Liegt eine nichtlineare Funktion vor, wird die Fehlerübertragung durch Linearisieren der Funktion in x_1 ermittelt.

Besonders beim Arbeiten mit biologischen, stark verrauschten Daten, sollte für die Parameterschätzung $\hat{\alpha}$ eine Fehlerabschätzung getroffen werden. Grundlage bildet die Fehlerfortpflanzung, wie sie für einfache Funktionen in Abb. 7.2 angedeutet wird. Betrachten wir zunächst den einfachen Fall einer eindimensionalen linearen Funktion (Abb. 7.2a). Wird für x_1 ein normalverteilter Fehler $\sigma(x_1)$ angenommen, ergibt sich für y folgender Fehler:

$$\sigma^2(y) = \frac{\partial y}{\partial x_1} \sigma^2(x_1) \frac{\partial y}{\partial x_1} \quad \sigma(y) = \sqrt{a^2 \sigma^2(x_1)} = a \sigma(x_1) \quad (7.9)$$

Im Fall mehrdimensionaler Funktionen mit mehreren fehlerbehafteten Größen \mathbf{x} , müssen in der Fehlerübertragung alle Fehler $\sigma(\mathbf{x})$ in Betracht gezogen werden. Hierbei müssen die Fehler der gemessenen Größen weiter charakterisiert werden. Entscheidend ist, wie stark die Eingangsgrößen \mathbf{x} (x_1 und x_2 in Abb. 7.2b) voneinander abhängen. Abhängigkeiten werden in der multivariaten Statistik durch die Kovarianzmatrix $\mathbf{Cov}(\mathbf{x})$ beschrieben. Diese Matrix kann graphisch interpretiert werden.

$$(\mathbf{x} - \mu)^T \mathbf{Cov}(\mathbf{x} - \mu) = 1 \quad (7.10)$$

Die Lösung der Gleichung (7.10) bildet einen Ellipsoiden innerhalb derer mit einer Wahrscheinlichkeit von 68% (2σ -Intervall) der wahre Wert liegt.

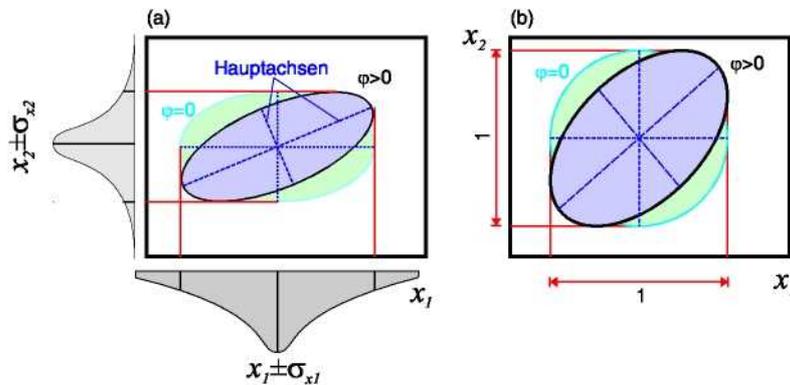


Abbildung 7.3: Graphische Darstellung der Kovarianzmatrix (a) und der Korrelationsmatrix (b). Die blauen Ellipsen geben die Gestalt für nichtkorrelierte Variablen x_1 und x_2 wieder. Die in lila gezeichneten Ellipsen deuten die Gestalt für positiv korrelierte ($\varphi > 0$) Variablen an.

In der Kovarianzmatrix befinden sich die Varianzen $\sigma^2(x_i)$ entlang der Diagonalen. Sind die Nebendiagonalen nicht besetzt, liegen die Achsen des Ellipsoiden parallel zu den Koordinatenachsen (Abb. 7.3a, $\text{Corr}_{i,j}(\mathbf{x}) = \varphi_{i,j} = 0$). Die Projektion des Ellipsoiden auf die jeweiligen Achsen gibt die zweifache Standardabweichung wieder.

Bestehen Abhängigkeiten der Variablen \mathbf{x} (z.B. x_1 und x_2), enthält die Kovarianz Einträge in Nebendiagonal-Elementen (in Abb. 7.3a $\text{Corr}_{1,2}(\mathbf{x}) = \varphi > 0$). Durch den Eintrag verändert sich die Lage des Ellipsoiden. Die Projektion auf die Achsen ergibt wieder die zweifache Standardabweichung.

Die Abhängigkeiten $\text{Cov}_{i,j}(\mathbf{x})$ lassen sich in der Korrelationsmatrix ($\text{Corr}(\mathbf{x})$) besser interpretieren (Abb. 7.3b). Die Korrelationsmatrix wird durch normieren der Kovarianzmatrix gebildet:

$$\text{Corr}_{i,j}(\mathbf{x}) = \frac{\text{Cov}_{i,j}(\mathbf{x})}{\sqrt{\text{Cov}_{i,i}(\mathbf{x}) \text{Cov}_{j,j}(\mathbf{x})}} \quad (7.11)$$

Auf der Hauptdiagonalen befinden sich jetzt Einsen, auf den Nebendiagonalen finden sich Einträge zwischen $-1 \leq \varphi_{i,j} \leq 1$. Der Betrag $|\varphi_{i,j}|$ des Eintrags gibt die Höhe der Abhängigkeit wieder, das Vorzeichen die Richtung der Abhängigkeit. Positive Korrelation der Variablen x_i und x_j bedeutet, dass bei einer Abweichung um Δ in x_i , x_j ebenfalls um einen Betrag $x_j + k \Delta$, mit: $k > 0$ abweicht. Bei negativer Korrelation wird x_j entgegengesetzt reagieren, d.h. $x_j - k \Delta$.

In Abb. 7.2b ist die Fehlerübertragung einer zweidimensionalen Funktion $y = a x_1 + b x_2$ angedeutet. Angenommen werden unabhängige Fehler $\sigma(x_1)$ und $\sigma(x_2)$. Die Hauptachsen des Ellipsoids (in der x_1, x_2 -Ebene) liegen somit parallel zu den Koordinaten x_1 und x_2 . Wie im eindimensionalen Fall werden auch die Fehler übertragen. Aus dem Schnitt eines Zylinders mit der Grundfläche des Ellipsoids mit der Ebene $y = a x_1 + b x_2$ lässt sich die Kovarianz der abhängigen Größe y ($\text{Cov}(y)$) bestimmen. Sie ergibt sich aus den Fehlern in \mathbf{x} ($\text{Cov}(\mathbf{x})$) durch:

$$\text{Cov}(y) = \mathbf{J} \text{Cov}(\mathbf{x}) \mathbf{J}^T \quad \text{mit: } \mathbf{J} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}} \quad (7.12)$$

Für das Beispiel in Abb. 7.2b ergibt sich also:

$$\begin{aligned} \text{Cov}(y) &= \mathbf{J} \text{Cov}(\mathbf{x}) \mathbf{J}^T = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} \sigma^2(x_1) & \\ & \sigma^2(x_2) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = a^2 \sigma^2(x_1) + b^2 \sigma^2(x_2) \\ \sigma(y) &= \sqrt{a^2 \sigma^2(x_1) + b^2 \sigma^2(x_2)} \end{aligned}$$

In Abb. 7.2c wird die linearisierte Fehlerübertragung nichtlinearer Funktionen dargestellt. Die Näherung durch Linearisierung ist nur in einem infinitesimal kleinen Bereich um die Entwicklungsstelle gültig. Für viele Systeme zeigt sich jedoch eine akzeptable Näherung auch für größere Bereiche.

7.2.1 Fehlerübertragung in der Regression

Die beschriebene Fehlerübertragung lässt sich auf alle behandelten Systeme anwenden. Die Funktionen müssen dazu allerdings entsprechend aufgestellt werden. Meist wird die Fehlerübertragung der Messwerte $\mathbf{y}(t_i)$ auf die unbekannt Parameter $\boldsymbol{\alpha}$ benötigt. Die dazu benötigte Inverse wird aus einem Regressionsansatz gewonnen. Die Schätzung $\hat{\boldsymbol{\alpha}}$ wurde mittels der Kleinsten-Quadrate Methode bestimmt. Für lineare Systeme der Gestalt

$$\mathbf{y} = \mathbf{A} \boldsymbol{\alpha} \quad (7.13)$$

ist die Lösung durch

$$\hat{\boldsymbol{\alpha}} = (\mathbf{A}^T \text{Cov}^{-1}(\mathbf{y}) \mathbf{A})^{-1} \mathbf{A}^T \text{Cov}(\mathbf{y}) \mathbf{y} \quad (7.14)$$

gegeben. Die Übertragung von Messfehlern ($\text{Cov}(\mathbf{y})$) auf die geschätzten Parameter $\text{Cov}(\hat{\boldsymbol{\alpha}})$ durch die Regression ergibt sich zu:

$$\text{Cov}(\hat{\boldsymbol{\alpha}}) = (\mathbf{A}^T \text{Cov}^{-1}(\mathbf{y}) \mathbf{A})^{-1} \quad (7.15)$$

Diese Fehlerübertragung lässt sich durch Linearisieren der verketteten Modellfunktion(en) ${}^i\mathbf{g}(\mathbf{x}(t_i), \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0)$ anwenden (Abb. 7.4).

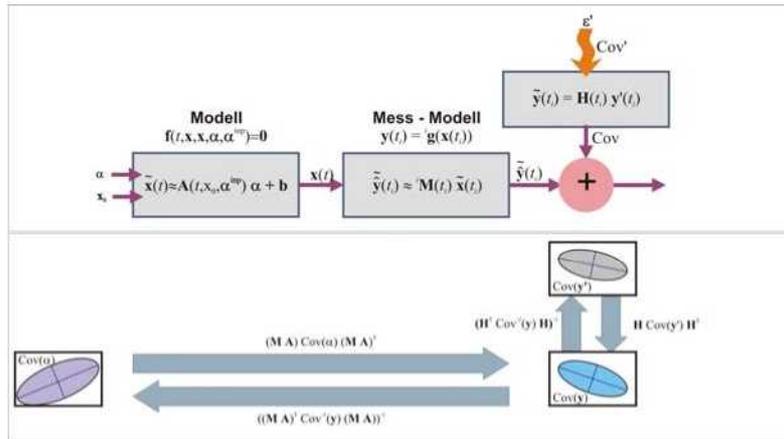


Abbildung 7.4: Anwendung der Fehlerübertragung auf die allgemeine Modellstruktur.

Betrachten wir zunächst den einfachen Fall einer Messreihe $\mathbf{y}(t_1)$ zum Zeitpunkt $t = t_1$. Die Taylor-Näherung der verketteten Funktion ${}^1\mathbf{g}(\mathbf{x}((t, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0)))$ in $t = t_1$ sowie in $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}$ für $\boldsymbol{\alpha}$ lautet dann:

$${}^1\tilde{\mathbf{y}}(t_1) = {}^1\mathbf{g}(\mathbf{x}(t_1, \hat{\boldsymbol{\alpha}}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{x}_0)) + \text{Sens}_{\boldsymbol{\alpha}}^{\tilde{\mathbf{y}}}(t = t_1, \boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}) (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}})$$

Die Fehlerübertragung nach Gl. 7.15 ergibt somit:

$$\text{Cov}(\boldsymbol{\alpha}) = ((\text{Sens}_{\boldsymbol{\alpha}}^{\tilde{\mathbf{y}}})^T(t_1) \text{Cov}^{-1}(\mathbf{y}(t_1)) \text{Sens}_{\boldsymbol{\alpha}}^{\tilde{\mathbf{y}}}(t_1))^{-1}$$

Liegen Messungen $\mathbf{y}(t_i)$ der N Messzeitpunkte $\boldsymbol{\xi} = [t_1, t_2, t_3, \dots, t_N]^T$ vor, wird der Gesamtmessvektor aus den einzelnen Messreihen $\mathbf{y}(\boldsymbol{\xi}) = [\mathbf{y}(t_1), \mathbf{y}(t_2), \dots, \mathbf{y}(t_N)]^T$ gebildet. Auf der rechten Seite stehen die in $t = t_i$ und $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}$ linearisierten Funktionen ${}^i\tilde{\mathbf{y}}(t_i, \boldsymbol{\alpha})$:

$$\underbrace{\begin{pmatrix} \mathbf{y}(t_1) \\ \mathbf{y}(t_2) \\ \vdots \\ \mathbf{y}(t_N) \end{pmatrix}}_{\mathbf{y}(\boldsymbol{\xi})} = \underbrace{\begin{pmatrix} {}^1\tilde{\mathbf{y}}(t_1, \boldsymbol{\alpha}) \\ {}^2\tilde{\mathbf{y}}(t_2, \boldsymbol{\alpha}) \\ \vdots \\ {}^N\tilde{\mathbf{y}}(t_N, \boldsymbol{\alpha}) \end{pmatrix}}_{\tilde{\mathbf{y}}(\boldsymbol{\xi}, \boldsymbol{\alpha})} + \underbrace{\begin{pmatrix} \boldsymbol{\varepsilon}(t_1) \\ \boldsymbol{\varepsilon}(t_2) \\ \vdots \\ \boldsymbol{\varepsilon}(t_N) \end{pmatrix}}_{\boldsymbol{\varepsilon}(\boldsymbol{\xi})} \quad (7.16)$$

Die Fehlerfortpflanzung auf die geschätzten Parameter lautet dann folglich:

$$\mathbf{Cov}(\boldsymbol{\alpha}) = \left(\mathbf{Sens}_{\boldsymbol{\alpha}}^{\hat{y}T}(\boldsymbol{\xi}) \mathbf{Cov}^{-1}(\mathbf{y}(\boldsymbol{\xi})) \mathbf{Sens}_{\boldsymbol{\alpha}}^{\hat{y}}(\boldsymbol{\xi}) \right)^{-1} \quad (7.17)$$

mit: $\mathbf{Sens}_{\boldsymbol{\alpha}}^{\hat{y}} = \begin{pmatrix} \mathbf{Sens}_{\boldsymbol{\alpha}}^{\hat{y}}(t_1) \\ \mathbf{Sens}_{\boldsymbol{\alpha}}^{\hat{y}}(t_2) \\ \vdots \\ \mathbf{Sens}_{\boldsymbol{\alpha}}^{\hat{y}}(t_N) \end{pmatrix}^T$, $\mathbf{Cov}(\mathbf{y}(\boldsymbol{\xi})) = \begin{pmatrix} \mathbf{Cov}(\mathbf{y}(t_1)) & & & \\ & \mathbf{Cov}(\mathbf{y}(t_2)) & & \\ & & \ddots & \\ & & & \mathbf{Cov}(\mathbf{y}(t_N)) \end{pmatrix}$

7.2.2 Untersuchungen anhand der Fehlerübertragung

Anhand der aus der Fehlerübertragung bestimmten Kovarianzmatrix $\mathbf{Cov}(\hat{\boldsymbol{\alpha}})$ können für die Modellinterpretation wichtige Größen abgeleitet werden:

1. **Vertrauensbereich:** Der Vertrauensbereich wird aus den Diagonalelementen, den Varianzen $\mathbf{Var}(\hat{\alpha}_i) = \mathbf{Cov}_{i,i}(\hat{\boldsymbol{\alpha}})$ entnommen. Die Standardabweichung des Parameters α_i ist

$$\sigma(\hat{\alpha}_i) = \sqrt{\mathbf{Var}(\hat{\alpha}_i)}$$

2. **Korrelation:** Durch normieren der Kovarianzmatrix $\mathbf{Cov}(\hat{\boldsymbol{\alpha}})$ nach Gl. (7.11) erhält man die Korrelationsmatrix $\mathbf{Corr}(\hat{\boldsymbol{\alpha}})$. Sie gibt Aufschluss über die Abhängigkeiten der Parameter $\hat{\boldsymbol{\alpha}}$. Hohe Korrelationen finden sich vor allem für Parameter, die ähnlich oder aber genau entgegengesetzt auf die Messungen wirken. Hoch korrelierte Parameter lassen sich deshalb nicht getrennt identifizieren, da jeweils der andere dieselbe Wirkung auf die Messung haben kann. Ein Beispiel ist in Abb. 9.2 dargestellt.
3. **Hauptkomponenten-Analyse:** Hinter der Hauptkomponenten-Analyse verbirgt sich eine Zerlegung der Kovarianzmatrix in ihre Eigenvektoren \mathbf{U} und die n Eigenwerte $\boldsymbol{\lambda}$,

$$\mathbf{Cov}(\hat{\boldsymbol{\alpha}}) = \mathbf{U} \mathbf{D} \mathbf{U}^T \quad \mathbf{D} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \quad (7.18)$$

Die Eigenwerte werden dabei absteigend sortiert.

$$\lambda_1 > \lambda_2 > \dots > \lambda_n$$

In der graphischen Interpretation geben die Eigenwerte die Länge der Achsen des Ellipsoiden (Abb. 7.3a) wieder. Die Eigenvektoren beschreiben die Richtung der Achsen. Der größte Eigenwert gehört zur längsten Achse des Ellipsoiden und enthält

somit die Richtung (bzw. Parameter) durch die ein Großteil der Varianz entsteht. Der Anteil a_1 der erklärten Varianz durch die längste Achse beträgt

$$a_1 = \frac{\lambda_1}{\sum_{i=1}^n \lambda_i}$$

Anhand der Hauptkomponentenanalyse können so schnell die im Modell schlecht bestimmten Parameter gefunden werden. Liegen auf den Achsen mehrere Parameter, deutet dies zudem auf hohe Korrelationen dieser Parameter hin. Abbildung 9.3 zeigt eine Hauptkomponentenzerlegung für die Kovarianzmatrix in Abb. 9.2.

4. **D-Kriterium:** Das D -Kriterium ist ein in der Versuchplanung häufig eingesetztes Maß (PAZMAN, 1986). Es berechnet sich aus der Determinante der Kovarianzmatrix:

$$D = \det(\text{Cov}(\boldsymbol{\alpha})) \quad (7.19)$$

Graphisch kann man sich dies als das Volumen des Ellipsoiden (Abb. 7.3a) vorstellen.

Um die statistische Güte verschiedener Experimente zu vergleichen, definiert man ein Referenz-Experiment mit dem D -Kriterium D_{ref} . Zum Vergleich der Parameterbestimmtheit wird der Informations-Index I verwendet (MÖLLNEY ET AL., 1999):

$$\begin{aligned} \bar{D}(\boldsymbol{\alpha}) &= \sqrt[2 \dim(\boldsymbol{\alpha})]{D} \\ I &= \frac{D_{ref}}{\bar{D}(\boldsymbol{\alpha})} \end{aligned} \quad (7.20)$$

Zwei Schritte sind hier von Bedeutung. Wie erwähnt beschreibt das D -Kriterium das Volumen des Ellipsoiden, die $2 \dim(\boldsymbol{\alpha})$ te Wurzel gibt (bis auf einen konstanten Faktor) die durchschnittliche Länge der Achsen des Ellipsoiden wieder (Abb. 7.3a). Mit dem Informationsindex I wird diese Länge mit anderen Experimenten verglichen. Da \bar{D} proportional zur Länge der Achsen ist, kann I als durchschnittliche Verbesserung bzw. Verschlechterung der Standardabweichungen interpretiert werden.

In Abschnitt 9 wird die visuelle Darstellung von Kovarianz-, Korrelations- und der in die Hauptkomponenten zerlegten Matrix anhand eines Beispiels beschrieben.

Kapitel 8

Anwendung der verschiedenen Modellsätze

8.1 Metabolisch stationäre Modelle

Der stationäre Zustand ist ein Sonderfall des dynamischen Verhaltens. Konzentrationen und Flüsse haben hier einen Gleichgewichtszustand erreicht und verändern sich über die Zeit nicht. Dies vereinfacht die Modellgleichungen wesentlich, da die rechte Seite der Differentialgleichungen (6.10) im stationären Zustand Null ist und aus dem ODE-System so ein Gleichungssystem wird. Für die stationäre Stoffflussanalyse ergibt sich zudem ein lineares System.

8.2 Stöchiometrische Stoffflussanalyse

Das experimentelle Vorgehen zur Bestimmung intrazellulärer Stoffflüsse wurde in Abschnitt 4.2 besprochen. Die Beschreibung der Auswertung orientiert sich an Abb. 7.1. Das im ersten "Block" zu lösende Gleichungssystem (Gl. (7.1)) des Netzwerks besteht nur aus den stoichiometrischen Gleichungen (Abb. 4.1). Parameter der Gleichungen sind die Flüsse \mathbf{v} und das zu lösende System vereinfacht sich somit zu:

$$f(\boldsymbol{\alpha}) = \mathbf{0} \quad \boldsymbol{\alpha} = \mathbf{v} \quad (8.1)$$

Die zu erfüllenden stöchiometrischen Gleichungen lauten:

$$\mathbf{N} \mathbf{v} = \mathbf{0} \quad (8.2)$$

Die Simulation, d.h. aus gegebenen Flüssen \mathbf{v} einen Systemzustand \mathbf{x} zu bestimmen ist für die SFA nicht relevant. Gültige Flusslagen \mathbf{v} erfüllen die Stöchiometrie und alle Systemvariablen sind bereits bekannt.

8.2.1 Messmodell

Für die Verknüpfung des Modells mit den Messdaten (Funktion \mathbf{g} in Abb. 7.1) ergeben sich für die n gemessene Flüsse \mathbf{y}_i die einfachen Messgleichungen:

$$\begin{aligned} \hat{y}_i &= {}^i \mathbf{g}(v_j) = v_j, \quad \text{mit: } i = 1, \dots, n \\ \hat{\mathbf{y}} &= \mathbf{M} \mathbf{v} \end{aligned} \quad (8.3)$$

Da die Flüsse \mathbf{v} linear abgebildet werden, ist auch das Gesamtmodell linear.

8.2.2 Parameteranpassung

Damit vereinfacht sich die Parameteranpassung, d.h. der Lösung nach den Parametern \mathbf{v} . Der in Kapitel 7 angedeutete Weg kann linear gelöst werden. Die Minimierung unter Nebenbedingungen Gl. (7.4) vereinfacht sich zu ($\boldsymbol{\alpha}$ wird mit \mathbf{v} ersetzt):

$$(\hat{\mathbf{v}}) = \arg \min_{\mathbf{N} \mathbf{v} = \mathbf{0}} \sum_{i=1}^n \|y_i - \hat{y}_i(\mathbf{v})\|_{\Sigma_y}^2, \quad \text{mit: } \hat{\mathbf{y}} = \mathbf{M} \mathbf{v} \quad (8.4)$$

Die Einhaltung der Nebenbedingung kann wie folgt sichergestellt werden.

1. Die stöchiometrischen Gleichungen $\mathbf{N} \mathbf{v} = \mathbf{0}$ bilden ein Gleichungssystem mit $n_f = \dim(\mathbf{N}) - \text{rank}(\mathbf{N})$ Freiheitsgraden (STRANG, 1976).
2. Wählt man n_f linear unabhängige Flüsse \mathbf{v}_f aus \mathbf{v} , können alle restlichen Flüsse \mathbf{v}_d berechnet werden.
3. Das Gleichungssystem wird dazu wie folgt dargestellt:

$$\mathbf{0} = [\mathbf{N}_f \ \mathbf{N}_d] \begin{pmatrix} \mathbf{v}_f \\ \mathbf{v}_d \end{pmatrix} \quad (8.5)$$

4. Die abhängigen Flüsse \mathbf{v}_d sind somit:

$$\mathbf{v}_d = -\mathbf{N}_d^{-1} \mathbf{N}_f \mathbf{v}_f \quad (8.6)$$

Wählt man statt \mathbf{v} die Flüsse \mathbf{v}_f als Parameter, kann die Minimierung (Gl. (7.4)) ohne Nebenbedingungen formuliert werden:

$$(\hat{\mathbf{v}}) = \arg \min_{\mathbf{v}_f} \sum_{i=1}^n \|y_i - \hat{y}_i^*(\mathbf{v}_f)\|_{\Sigma_y}^2, \quad \text{mit: } \hat{\mathbf{y}}_i^*(\mathbf{v}_f) = \mathbf{M} \mathbf{P} \begin{pmatrix} \mathbf{I} \\ -\mathbf{N}_d^{-1} \mathbf{N}_f \end{pmatrix} \mathbf{v}_f \quad (8.7)$$

Die erwarteten Messungen $y_i(\mathbf{v})$ werden zu $y_i^*(\mathbf{v}_f)$, da sich durch den Übergang von \mathbf{v} zu \mathbf{v}_f und \mathbf{v}_d die Funktion verändert (Gl. (8.7)). Die Einheitsmatrix \mathbf{I} erhält die Dimension entsprechend der Dimension des Flussvektors \mathbf{v}_f , mit der Hilfsmatrix \mathbf{P} wird aus dem Vektor der freien Flüsse der Gesamtvektor gebildet:

$$\mathbf{v} = \mathbf{P} \begin{pmatrix} \mathbf{v}_f \\ \mathbf{v}_d \end{pmatrix}$$

Die Lösung der Minimierung in Gl. (8.7) kann mit der linearen Regression bestimmt werden (Gl. (7.14)):

$$\hat{\mathbf{v}}_f = (\mathbf{A}^T \text{Cov}^{-1}(\mathbf{y}) \mathbf{A})^{-1} \mathbf{A}^T \text{Cov}(\mathbf{y}) \mathbf{y} \quad \text{mit: } \mathbf{A} = \mathbf{M} \mathbf{P} \begin{pmatrix} \mathbf{I} \\ -\mathbf{N}_d^{-1} \mathbf{N}_f \end{pmatrix} \quad (8.8)$$

8.2.3 Sensitivitätsanalyse und statistische Berechnungen

Die Sensitivitätsanalyse stationärer Netzwerke spielt in Kapitel 11 eine wichtige Rolle zur Interpretation gemessener Stoffflusslagen. In diesem Abschnitt wird die Sensitivitätsanalyse vor allem im Hinblick auf statistische Untersuchungen, insbesondere der Fehlerfortpflanzung besprochen.

Für den Fall eines stationären Stoffwechselnetzwerks ergibt sich eine leicht zu lösende Problemstellung. Die Sensitivität der *freien* Flüsse \mathbf{v}_f auf die abhängigen Größen \mathbf{v}_d wird durch Ableiten der Gleichung 8.6 nach \mathbf{v}_f bestimmt:

$$\mathbf{Sens}_{\mathbf{v}_f}^{\mathbf{v}_d} = \frac{\partial \mathbf{v}_d}{\partial \mathbf{v}_f} = -\mathbf{N}_d^{-1} \mathbf{N}_f \quad (8.9)$$

Genauso kann die Sensitivität der Messungen $\hat{\mathbf{y}}$ gegenüber \mathbf{v}_f bestimmt werden. Abgeleitet wird Gleichung 8.3:

$$\mathbf{Sens}_{\mathbf{v}_f}^{\hat{\mathbf{y}}} = \frac{\partial \hat{\mathbf{y}}}{\partial \mathbf{v}_f} = \mathbf{M} \mathbf{P} \begin{pmatrix} \mathbf{I} \\ -\mathbf{N}_d^{-1} \mathbf{N}_f \end{pmatrix} \quad (8.10)$$

Mit der Sensitivität kann nun auch die Fehlerfortpflanzung berechnet werden. Nach Gl. (7.15) wird bei bekannter Kovarianz der Messungen ($\mathbf{Cov}(\mathbf{y})$) die Kovarianz der Parameter, hier also der *freien* Flüsse zu

$$\mathbf{Cov}(\mathbf{v}_f) = \left(\mathbf{Sens}_{\mathbf{v}_f}^{\hat{\mathbf{y}}} \mathbf{Cov}^{-1}(\hat{\mathbf{y}}) \mathbf{Sens}_{\mathbf{v}_f}^{\hat{\mathbf{y}T}} \right)^{-1} \quad (8.11)$$

Anhand der bekannten Kovarianz der *freien* Flüsse können nun auch die Kovarianzen der abhängigen Flüsse mit der Fehlerfortpflanzung und den Sensitivitäten in Gl. (8.9) bestimmt werden.

8.2.4 Werkzeuge

Der mühseligste Schritt bei der Berechnung von Stoffflüssen ist das Aufstellen der teilweise sehr großen stoichiometrischen Matrix. Diese Aufgabe kann von verschiedenen Werkzeugen übernommen werden. Aufgrund der guten Integrierbarkeit in MatLab eignet sich MetaTool (Version 5.0). Alle weiteren Schritte der Berechnung lassen sich leicht in MatLab umsetzen.

KLAMT ET AL. (2003) stellen ein in MatLab umgesetztes Tool zur Berechnung stationärer Stoffflüsse mit grafischer Oberfläche relativ günstig zur Verfügung. Die Firma InSilico bietet ein wesentlich umfangreicheres, ebenfalls mit grafischer Oberfläche ausgestattetes Softwaretool (InSilico Discovery) an (MAUCH, 2005).

8.3 Metabolische Netzwerk Analyse

Die Methoden der Metabolischen Netzwerk Analyse (Metabolic Network Analysis, MNA) zielen nicht auf die Bestimmung intrazellulärer Flüsse. Im Vordergrund stehen vor allem folgende Fragestellungen:

1. Welcher maximale Fluss bzw. welche Ausbeute kann für ein bestimmtes Produkt erreicht werden (MAHADEVAN UND SCHILLING, 2003; PHARKYA ET AL., 2004),
2. Welche Stoffflusslagen können in einem gegebenen metabolischen Netzwerk auftreten (SCHUSTER ET AL., 2000; RAMAKRISHNA ET AL., 2001; SCHILLING ET AL., 2000),
3. Kann das Netzwerk in Module, bzw. funktionelle Einheiten getrennt werden (HEINRICH UND SCHUSTER, 1998),
4. Gibt es "essenzielle" Reaktionsschritte bzw. Funktionen im Netzwerk (EDWARDS ET AL., 2001).

Die ersten beiden Fragestellungen sind für das Metabolic Engineering mit konkreten Stämmen und Produktionszielen wichtig. Die Fragen 3 und 4 sind von der Grundlagenforschung motiviert und helfen bei der Suche nach den metabolischen Netzwerken zugrunde liegenden Design-Kriterien wie z.B. Stabilität, Flexibilität oder Modularität (EDWARDS UND PALSSON, 2000). Die Methoden zur Berechnung maximaler Ausbeuten und möglicher Stoffflusslagen werden in Kapitel 11 ausführlich anhand eines einfachen Beispiels und des L-Phenylalanin-Netzwerks vorgestellt.

8.3.1 Werkzeuge

Genutzt wurde vor allem das von PFEIFFER ET AL. (1999) bereitgestellte Werkzeug MetaTool in der Version 5.0. Mit diesem Tool können die Elementarmoden des Netzwerks bestimmt werden. Zur Berechnung der optimalen Stoffflusslage wurde die Funktion `lprsm` (nach FANG UND PUTHENPURA (1993)) der freien MatLab-Toolbox `matlog` eingesetzt.

8.4 ^{13}C Stoffflussanalyse

Das Gleichungssystem der Markierungsbilanzen (Abschnitt 6.5) ist im Vergleich zu den stöchiometrischen Gleichungen wesentlich komplexer. Zum einen besitzen die Modelle ein

Vielfaches an Gleichungen, zum anderen treten nichtlineare Terme auf (für bimolekulare Reaktionen, Abschnitt 6.5). Für die Simulation wird die Lösung der Gleichungen bezüglich der Verteilungen \mathbf{x} (bzw. den Konzentrationen $\mathbf{x}_c = \mathbf{c P x}$) der Isotopomere benötigt:

$$\mathbf{x} = \mathbf{x}(\mathbf{v}, \boldsymbol{\alpha}^{\text{inp}}) \quad (8.12)$$

Erstaunlicherweise gibt es für die Lösung eine lineare Transformation, mit der das nichtlineare Gleichungssystem in ein stufenweise lineares System mit nichtlinearen inhomogenen Termen überführt werden kann (WIECHERT UND GRAAF, 1997). Dazu betrachtet man anstatt der Isotopomere die so genannten Cumomere (Kunstwort aus “**C**umulated Isotopomers”). In dieser Darstellung wird die Markierungs-Verteilung in markierten (Massen-)Fragmenten dargestellt. Statt den Markierungszuständen 1 (¹³C) und 0 (¹²C) hat das Cumomer die Zustände 1 (¹³C) und x (¹³C oder ¹²C). So besitzt der Metabolit D (8 Isotopomerenzustände) wieder 8 Cumomerenzustände. Die Verteilung der Cumomerenanteile ${}^c\mathbf{d}$ kann aus der Verteilung der Isotopomere \mathbf{d} durch eine lineare Transformation ermittelt werden :

$$\begin{aligned} d_{xxx} &= d_{000} + d_{001} + d_{010} + d_{011} + d_{100} + d_{101} + d_{110} + d_{111} \\ d_{xx1} &= d_{001} + d_{011} + d_{101} + d_{111} \\ d_{x1x} &= d_{010} + d_{011} + d_{110} + d_{111} \\ d_{1xx} &= d_{100} + d_{101} + d_{110} + d_{111} \\ d_{x11} &= d_{011} + d_{111} \\ d_{1x1} &= d_{101} + d_{111} \\ d_{11x} &= d_{110} + d_{111} \\ d_{111} &= d_{111} \end{aligned} \quad \mathbf{d} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{d}$$

Hervorzuheben sind,

1. D_{xxx} , bildet sich aus Summe aller Isotopomeren-Anteile. Für die Konzentration des Cumomers D_{xxx} gilt somit:

$${}^c d_{xxx} = \sum_{i=0}^1 \sum_{j=0}^1 \sum_{k=0}^1 d_{ijk} = 1 \quad {}^c D_{xxx} = {}^c d_{xxx} D = D$$

Die Bilanzen der beliebig markierten Cumomere geben folglich die stöchiometrischen Gleichungen wieder.

2. D_{111} , ist in Cumomeren- und der Isotopomeren Darstellung identisch $D_{\#111} = D_{111}$.

WIECHERT UND GRAAF (1997) sortiert die Cumomere anhand ihres “Gewichtes”, d.h. der Anzahl der 1'en. Stufe 0 enthält die beliebig markierten Cumomere (z.B. D_{xxx})

und gibt somit die Stöchiometrie wieder. Stufe 1 enthält alle Cumomere mit einer 1, Stufe 2 alle Cumomere mit zwei markierten C-Atomen, usw. bis Stufe m mit den Cumomeren der vollmarkierten längsten C-Ketten. Für jede Gewichts-Stufe können die Cumomeren-Gleichungen aufgestellt werden. Die Gleichungen der einzelnen Stufen sind jeweils nur von der Lösung leichterer Stufen abhängig. Die Gleichungen für das Beispielsystem lauten dann:

$$\begin{aligned}
 \text{Stufe 0: } & \mathbf{A}_{xx} : 0 = \text{feed}_{xx} v_{\text{upt}}^{\rightarrow} - \mathbf{a}_{xx} v_1^{\rightarrow} \\
 & \mathbf{B}_{xx} : 0 = \mathbf{a}_{xx} v_1^{\rightarrow} + \mathbf{b}_{xx} (-v_1^{\leftarrow} - v_2^{\rightarrow} - v_5^{\rightarrow}) \\
 & \mathbf{C}_{xx} : 0 = \mathbf{b}_{xx} \mathbf{e}_{xx} v_2^{\rightarrow} - \mathbf{c}_{xxx} v_3^{\rightarrow} \\
 & \mathbf{D}_{xx} : 0 = \mathbf{c}_{xxx} v_3^{\rightarrow} - \mathbf{d}_{xxx} v_4^{\rightarrow} \\
 & \mathbf{E}_{xx} : 0 = \mathbf{d}_{xxx} v_4^{\rightarrow} + \mathbf{b}_{xx} v_5^{\rightarrow} + \mathbf{e}_{xx} (-v_2^{\rightarrow} - v_6^{\rightarrow}) \\
 & \mathbf{F}_x : 0 = \mathbf{c}_{xxx} v_3^{\rightarrow} + \mathbf{d}_{xxx} v_4^{\rightarrow} - \mathbf{f}_x v_7^{\rightarrow} \\
 \text{Stufe 1: } & \mathbf{A}_{x1} : 0 = \text{feed}_{x1} v_{\text{upt}}^{\rightarrow} - \mathbf{a}_{x1} v_1^{\rightarrow} \\
 & \mathbf{A}_{1x} : 0 = \text{feed}_{1x} v_{\text{upt}}^{\rightarrow} - \mathbf{a}_{1x} v_1^{\rightarrow} \\
 & \mathbf{B}_{x1} : 0 = \mathbf{a}_{x1} v_1^{\rightarrow} + \mathbf{e}_{x1} v_5^{\leftarrow} + \mathbf{b}_{x1} (-v_1^{\leftarrow} - v_2^{\rightarrow} - v_5^{\rightarrow}) \\
 & \mathbf{B}_{1x} : 0 = \mathbf{a}_{1x} v_1^{\rightarrow} + \mathbf{e}_{1x} v_5^{\leftarrow} + \mathbf{b}_{1x} (-v_1^{\leftarrow} - v_2^{\rightarrow} - v_5^{\rightarrow}) \\
 & \mathbf{C}_{xxx1} : 0 = \mathbf{b}_{xx} \mathbf{e}_{x1} v_2^{\rightarrow} - \mathbf{c}_{xxx1} v_3^{\rightarrow} \\
 & \mathbf{C}_{xx1x} : 0 = \mathbf{b}_{xx} \mathbf{e}_{1x} v_2^{\rightarrow} - \mathbf{c}_{xx1x} v_3^{\rightarrow} \\
 & \mathbf{C}_{x1xx} : 0 = \mathbf{b}_{x1} \mathbf{e}_{xx} v_2^{\rightarrow} - \mathbf{c}_{x1xx} v_3^{\rightarrow} \\
 & \mathbf{C}_{1xxx} : 0 = \mathbf{b}_{1x} \mathbf{e}_{11} v_2^{\rightarrow} - \mathbf{c}_{1xxx} v_3^{\rightarrow} \\
 & \mathbf{D}_{xx1} : 0 = \mathbf{c}_{xxx1} v_3^{\rightarrow} - \mathbf{d}_{xx1} v_4^{\rightarrow} \\
 & \mathbf{D}_{x1x} : 0 = \mathbf{c}_{xx1x} v_3^{\rightarrow} - \mathbf{d}_{x1x} v_4^{\rightarrow} \\
 & \mathbf{D}_{1xx} : 0 = \mathbf{c}_{x1xx} v_3^{\rightarrow} - \mathbf{d}_{1xx} v_4^{\rightarrow} \\
 & \mathbf{E}_{x1} : 0 = \mathbf{d}_{x1x} v_4^{\rightarrow} + \mathbf{b}_{x1} v_5^{\rightarrow} + \mathbf{e}_{x1} (-v_5^{\leftarrow} - v_2^{\rightarrow} - v_6^{\rightarrow}) \\
 & \mathbf{E}_{1x} : 0 = \mathbf{d}_{1xx} v_4^{\rightarrow} + \mathbf{b}_{1x} v_5^{\rightarrow} + \mathbf{e}_{1x} (-v_5^{\leftarrow} - v_2^{\rightarrow} - v_6^{\rightarrow}) \\
 & \mathbf{F}_1 : 0 = \mathbf{c}_{1xxx} v_3^{\rightarrow} + \mathbf{d}_{xx1} v_4^{\rightarrow} - \mathbf{f}_1 v_7^{\rightarrow} \\
 \text{Stufe 2: } & \mathbf{A}_{11} : 0 = \text{feed}_{11} v_{\text{upt}}^{\rightarrow} - \mathbf{a}_{11} v_1^{\rightarrow} \\
 & \mathbf{B}_{11} : 0 = \mathbf{a}_{11} v_1^{\rightarrow} + \mathbf{e}_{11} v_5^{\leftarrow} + \mathbf{b}_{11} (-v_1^{\leftarrow} - v_2^{\rightarrow} - v_5^{\rightarrow}) \\
 & \mathbf{C}_{xxx11} : 0 = \mathbf{b}_{xx} \mathbf{e}_{11} v_2^{\rightarrow} - \mathbf{c}_{xxx11} v_3^{\rightarrow} \\
 & \mathbf{C}_{x1x1} : 0 = \mathbf{b}_{x1} \mathbf{e}_{x1} v_2^{\rightarrow} - \mathbf{c}_{x1x1} v_3^{\rightarrow} \\
 & \mathbf{C}_{x11x} : 0 = \mathbf{b}_{x1} \mathbf{e}_{1x} v_2^{\rightarrow} - \mathbf{c}_{x11x} v_3^{\rightarrow} \\
 & \mathbf{C}_{1xx1} : 0 = \mathbf{b}_{1x} \mathbf{e}_{x1} v_2^{\rightarrow} - \mathbf{c}_{1xx1} v_3^{\rightarrow} \\
 & \mathbf{C}_{1x1x} : 0 = \mathbf{b}_{1x} \mathbf{e}_{1x} v_2^{\rightarrow} - \mathbf{c}_{1x1x} v_3^{\rightarrow} \\
 & \mathbf{C}_{11xx} : 0 = \mathbf{b}_{11} \mathbf{e}_{xx} v_2^{\rightarrow} - \mathbf{c}_{11xx} v_3^{\rightarrow} \\
 & \mathbf{D}_{x11} : 0 = \mathbf{c}_{xxx11} v_3^{\rightarrow} - \mathbf{d}_{x11} v_4^{\rightarrow} \\
 & \mathbf{D}_{1x1} : 0 = \mathbf{c}_{x1x1} v_3^{\rightarrow} - \mathbf{d}_{1x1} v_4^{\rightarrow} \\
 & \mathbf{D}_{11x} : 0 = \mathbf{c}_{x11x} v_3^{\rightarrow} - \mathbf{d}_{11x} v_4^{\rightarrow} \\
 & \mathbf{E}_{11} : 0 = \mathbf{d}_{x11} v_4^{\rightarrow} + \mathbf{b}_{11} v_5^{\rightarrow} + \mathbf{e}_{11} (-v_5^{\leftarrow} - v_2^{\rightarrow} - v_6^{\rightarrow}) \\
 & \vdots \\
 \text{Stufe 4: } & \mathbf{C}_{111} : 0 = \mathbf{b}_{11} \mathbf{e}_{11} v_2^{\rightarrow} - \mathbf{c}_{1111} v_3^{\rightarrow}
 \end{aligned}$$

Diese ausführliche Darstellung zeigt die für die effiziente Lösung wichtigen Eigenschaften der Cumomeren-Darstellung. Das System vereinfacht sich wesentlich, da

1. alle vollständig beliebig markierten Cumomeren-Anteile durch Einsen ersetzt werden können.
2. die in der vorherigen Stufe ($i - 1$) berechneten Größen in Stufe i eingesetzt werden können. Dadurch entfallen die nichtlinearen Terme. So ist für die Berechnung der Bilanz von $C_{x_1 x_1}$ (Stufe 2) die Lösung für B_{x_1} und E_{x_1} notwendig. Beide Cumomere wurden bereits in Stufe 1 berechnet. Mit dem Einsetzen der Lösung wird aus der Multiplikation von zwei Unbekannten ein konstanter Term. In WIECHERT ET AL. (1999) wird gezeigt, dass alle nichtlinearen Terme ersetzt werden können und die Lösung der i ten Stufe nur von Lösungen der Stufen $0 < i$ abhängen.

Für gegebene Flüsse \mathbf{v} werden nun die Verteilung der Cumomeren \mathbf{x} berechnet. Wie das Gleichungssystem wird auch der Vektor der Cumomeren in Gewichte unterteilt. Der Vektor ${}^0\mathbf{x}$ enthält alle Cumomere mit Gewicht 0, ${}^1\mathbf{x}$ alle Cumomere mit Gewicht 1, usw. Die Lösung der Stufe Null ist trivial, liegt ein die Stöchiometrie erfüllender Vektor \mathbf{v} vor, sind die Gleichungen erfüllt:

$$\mathbf{1} = {}^0\mathbf{x}$$

Für Stufe 1 wird das lineare Gleichungssystem

$$\mathbf{0} = {}^1\mathbf{A}(\mathbf{v}) {}^1\mathbf{x} + {}^1\mathbf{b}(\mathbf{v})$$

gelöst. Die Matrix ${}^1\mathbf{A}(\mathbf{v})$ ist verwandt mit der stöchiometrische Matrix. Multipliziert mit dem Vektor der Cumomeren (mit Gewicht 1 ${}^1\mathbf{x}$) bildet sie die rechte Seite der Cumomerenbilanzen ab. Statt der stöchiometrischen Koeffizienten stehen jedoch die jeweils speisenden bzw. verbrauchenden Flüsse. Der Vektor ${}^1\mathbf{b}(\mathbf{v})$ enthält die zu ersetzenden Terme, z.B. $\text{feed}_{x_1} v_{\text{opt}}^{\rightarrow}$ (feed_{x_1} ist durch die Substratzufuhr vorgegeben). In Stufe 2 wird die Abhängigkeit deutlicher:

$$\mathbf{0} = {}^2\mathbf{A}(\mathbf{v}) {}^2\mathbf{x} + {}^2\mathbf{b}(\mathbf{v}, {}^1\mathbf{x})$$

Jetzt fließen auch Ergebnisse aus ${}^1\mathbf{x}$ ein. Eingesetzt werden z.B. $\mathbf{b}_{x_1} \mathbf{e}_{x_1} v_2^{\rightarrow}$ in der Bilanz für $C_{x_1 x_1}$. Allgemein lauten die Gleichungen der i ten Stufe:

$$\mathbf{0} = {}^i\mathbf{A}(\mathbf{v}) {}^i\mathbf{x} + {}^i\mathbf{b}(\mathbf{v}, {}^1\mathbf{x}, {}^2\mathbf{x}, \dots, {}^{i-1}\mathbf{x}), \quad i = 1, \dots, m$$

Die Verteilung der Cumomeren ${}^i\mathbf{x}$, $i = 1 \dots m$ ergibt sich durch Anwenden von Verfahren der linearen Algebra (z.B. Gauss-Eliminierung) (STRANG, 1976).¹

8.4.1 Messmodell

Mit der Lösung der Systemgleichungen liegen die Verteilungen \mathbf{x} und die Flüsse \mathbf{v} vor. Gemessen werden können extrazelluläre Flüsse \mathbf{w} und bestimmte Eigenschaften der Markierungszustände \mathbf{x} (z.B. die Verteilung der Massenisotopomeren). MÖLLNEY ET AL. (1999) schlägt dazu ein allgemeines, skalierbares Messmodell \mathbf{g} vor. Dieses Modell konstruiert ein Messsignal aus Linearkombinationen der Markierungszustände \mathbf{x} . Die Messsignale werden dabei in Messgruppen unterteilt, die mit Hilfe von Skalierungsfaktoren $\boldsymbol{\omega}$ auf die Skala des Messgeräts gebracht werden. Mit diesem Ansatz können Roh-Messdaten verwendet werden, insbesondere ist keine Normierung der Messungen notwendig (die statistische Nachteile mit sich zieht (WAHL, 2002)). Das Messmodell kann auf ${}^1\text{H-NMR}$, ${}^{13}\text{C-NMR}$, 2D-NMR sowie MS-Messdaten (Abschnitt 18) angepasst werden:

$$\hat{\mathbf{y}} = \mathbf{g}(\mathbf{x}(\mathbf{v}, \boldsymbol{\alpha}^{\text{inp}}), \boldsymbol{\omega}) \quad (8.13)$$

8.4.2 Parameteranpassung

Dem Schema folgend muss nun wieder die Minimierung unter Nebenbedingungen (Gl. (7.4)) durchgeführt werden. Für die metabolisch und isotopisch stationäre SFA mit n gemessenen Flüssen und m gemessenen Markierungsverteilungen $\mathbf{y}_\mathbf{x}$ zu:

$$\hat{\mathbf{v}} = \arg \min_{\substack{\mathbf{v}, \boldsymbol{\omega} \\ \mathbf{f}(\mathbf{v}, \mathbf{x}, \boldsymbol{\alpha}^{\text{inp}}) = \mathbf{0}}} \|\mathbf{y} - \hat{\mathbf{y}}\|_{\Sigma_\mathbf{y}}^2 \quad \text{mit: } \mathbf{y} = \begin{pmatrix} \mathbf{w} \\ \mathbf{y}_\mathbf{x} \end{pmatrix} \quad \hat{\mathbf{y}} = \mathbf{g}(\mathbf{x}(\mathbf{v}, \boldsymbol{\alpha}^{\text{inp}}), \boldsymbol{\omega}) = \begin{pmatrix} \mathbf{M}_\mathbf{w} \mathbf{v} \\ \boldsymbol{\omega} \mathbf{M}_\mathbf{x} \mathbf{x} \end{pmatrix} \quad (8.14)$$

Wie im Fall der stöchiometrischen SFA kann die Nebenbedingung $\mathbf{f}(\mathbf{v}, \mathbf{x}, \boldsymbol{\alpha}^{\text{inp}}) = \mathbf{0}$ eliminiert werden, indem die Flüssen \mathbf{v} in *freie* und *abhängige* Flüsse (\mathbf{v}_f und \mathbf{v}_d) geteilt werden. Aufgrund der nun enthaltenen Vor- und Rückflüsse vergrößert sich allerdings die Anzahl *freier* Flüsse.

Die Minimierung kann allerdings nicht mehr explizit (vgl. Gl. (8.8)) gelöst werden. Sie wird mit nichtlinearen, iterativen Optimierungsalgorithmen (z.B. Simplex, BFGS oder evolutionären Verfahren) durchgeführt (WIECHERT ET AL., 2001).

¹(ISERMANN ET AL., 2004) setzen Methoden der Graphentheorie zur Erkennung von Strukturen in den Gleichungen ein. Durch geschicktes Umsortieren und Diagonalisieren der Matrizen sind wesentlich schnellere Lösung möglich.

8.4.3 Sensitivitäten und Parameterbestimmtheit

Mit der in Kapitel 7 beschriebenen Vorgehensweise können nun auch die Sensitivitäten der Markierungsverteilungen \mathbf{x} und der abhängigen Flüsse \mathbf{v}_d berechnet werden. Da es sich trotz der Kaskadierung weiterhin (als Gesamtsystem) um ein nichtlineares System handelt wird in $\hat{\mathbf{v}}, \hat{\omega}$ linearisiert. Die Sensitivitäten und Parameterbestimmtheiten für stationäre Markierungssysteme ist in (WIECHERT UND GRAAF, 1997) ausführlich dargestellt.

8.4.4 Werkzeuge

In dieser Arbeit wurde ausschließlich das Software-Werkzeug ^{13}C FLUX eingesetzt. Es ist eine sehr flexible Simulationsumgebung. Das ^{13}C Modell wird in einem Textfile (FTBL-Format) definiert, wobei sich der Modellaufbau an der Notation von WIECHERT (1995) orientiert. Im ersten Abschnitt werden das Netzwerk mit den jeweiligen C-Atom Transitionen aufgeführt, danach werden die Netto- und Austauschflüsse gesetzt. Der Benutzer wählt hier die *freien* und abhängigen Variablen und kann Grenzen für bestimmte Flüsse setzen. Im letzten Abschnitt werden die Messdaten für die extrazellulären Flüsse und gemessene Markierungen eingetragen.

8.5 Instationäre ^{13}C Stoffflussanalyse

In dieser Arbeit werden keine Modelle im metabolisch stationären, isotopisch instationären Zustand erstellt. Interessierte Leser können in den ausführlich Arbeiten von NÖH ET AL. (2006) Simulationsstudien und erste Auswertungen (NÖH ET AL., 2007) nachlesen.

8.6 Metabolisch instationärer Zustand

Mit den in den Abschnitten 6.2 und 6.3 beschriebenen Gleichungen kann das dynamische Verhalten eines metabolischen Netzwerks simuliert werden. Der Konzentrationsverlauf $\mathbf{c}(t)$ der Metabolit Pools sowie der extrazellulären Substrate und Produkte wird durch Differentialgleichungen dargestellt. Die rechte Seite wird durch die stöchiometrische Matrix \mathbf{N} und den Fluss-Vektor $\mathbf{v}(\mathbf{c}, \boldsymbol{\alpha})$ mit den Enzym-Kinetiken aufgebaut (explizite Darstellung des allgemeinen, impliziten Modells \mathbf{f}):

$$\dot{\mathbf{c}} = \mathbf{N} \mathbf{v}(\mathbf{c}, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}), \quad \text{mit: } \mathbf{c}(t=0) = \mathbf{c}_0 \quad (8.15)$$

Die Lösung dieser Differentialgleichungen kann mit verschiedenen Werkzeugen berechnet werden. Eine weitere Aufbearbeitung ist nicht notwendig. Die Differentialgleichung (8.15) ist nichtlinear und muss numerisch gelöst werden (z.B. Euler, Runge-Kutta, Rosenbrock).

Beispielhaft wird das Modell im Anhang betrachtet (metabolisches Netzwerk in Abb. C.1, Kinetiken in Tab. C.1 und Parameterwerte in Tab. C.1). In der Simulation wird zum Zeitpunkt $t = 0\text{ s}$ der Feedstrom \dot{V}_{feed} von $0.21/\text{h}$ kurzzeitig (für 0.25 s) auf $6001/\text{h}$ erhöht (Puls-Zugabe). Die extrazelluläre Substratkonzentration steigt durch die Puls-Zugabe extrem schnell auf 1.5 mM an (Abb. 8.1). Mit dem Anstieg des extrazellulären Substrats steigt sofort die Substrat-Aufnahmerate v_{upt} und die intrazellulären Konzentrationen (A, B, C, D und E).

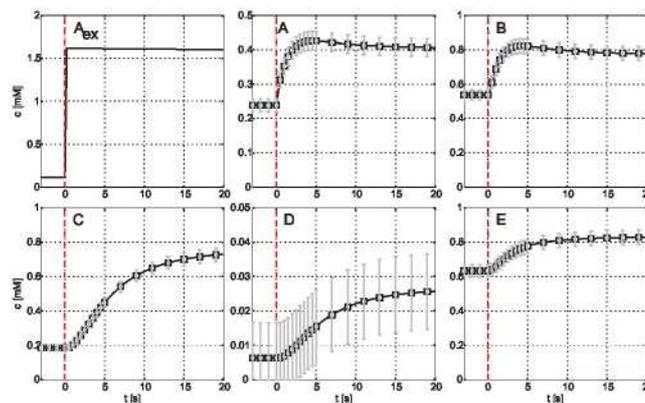


Abbildung 8.1: Simulationsergebnisse für die Poolgrößen. Die Pools A und B zeigen ein leichtes Überschwingen durch die Pulsaufgabe. Die Fehlerbalken berechnen sich aus den in Kap. 17 getroffenen Annahmen zur Messgenauigkeit.

8.6.1 Messmodell

Im ersten Moment scheint es nicht sinnvoll, auch für dynamische Netzwerkmodelle eine Messmodell aufzustellen, da die Abbildung der simulierten Konzentrationen direkt auf die gemessenen Konzentrationen \mathbf{y} abgebildet werden. Es treten jedoch Sonderfälle auf, die mit Hilfe eines Messmodell gehandhabt werden können:

- **Relative** Messung: Ist für einen Analyten kein Standard verfügbar, kann die absolute Konzentration nicht bestimmt werden.
- **Isobare** Messung: Hier fallen zwei oder mehr Metabolite auf den selben Peak, d.h. sowohl die chromatische Retentionszeit als auch das m/z -Verhältnis sind identisch, z.B. F6P und G6P. Ausgewertet werden kann dann nur das Summensignal.

Dem Schema folgend wird das Messmodell zu:

$$\begin{aligned}\hat{y}(t_i) &= {}^i\mathbf{g}(\mathbf{c}(t_i, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{c}_0)) \\ \hat{y}(t_i) &= {}^i\mathbf{M} \mathbf{c}(t_i, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{c}_0)\end{aligned}\quad (8.16)$$

8.6.2 Sensitivitätsberechnung

Im Gegensatz zu den zuvor behandelten stationären Systemen spielt die Nicht-Linearität unter Umständen eine gewichtigere Rolle. Es ist anzunehmen, dass die Parameter je nach Systemzustand und Parametrisierung unterschiedlichen Einfluss auf das Systemverhalten nehmen können². In der Systemanalyse werden deshalb verschiedene Methoden der Sensitivitätsberechnung vorgeschlagen (SALTELLI ET AL., 2000), um “global” gültige Sensitivitäten zu berechnen.

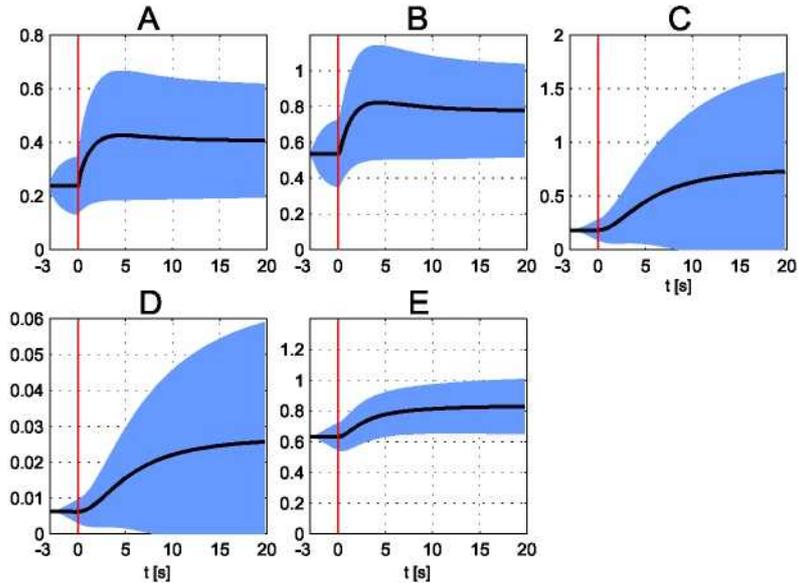


Abbildung 8.2: Sensitivität der Metabolit Pools gegenüber dem Parameter $v_{\text{max}}^{\text{upt}}$. Die Sensitivität wird unskaliert als Schlauch um den simulierten Konzentrationsverlauf dargestellt.

In dieser Arbeit wird ausschließlich die “lokale” Sensitivität betrachtet. “Lokal” bedeutet in diesem Zusammenhang, dass die Sensitivität für den untersuchten Parameterzustand $\hat{\boldsymbol{\alpha}}$ bestimmt wird. Diese auf den augenblicklichen Systemzustand und die augenblickliche Parametrisierung eingeschränkte Sensitivität ist für den in dieser Arbeit benötigten Zusammenhang maßgeblich, da der Parametereinfluss im Bereich einer gefundenen Schätzung $\hat{\boldsymbol{\alpha}}$

²Für stationäre ^{13}C Systeme konnte gezeigt werden, dass die Näherung durch Linearisierung das wahre Systemverhalten für relativ große Parametervariationen noch gut wiedergeben kann.

beurteilt werden soll. Auch in der folgenden Kovarianzberechnung ist die augenblickliche Parametrisierung maßgeblich (die Funktion f wird in $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}$ abgeleitet). Aufgrund der Nichtlinearität ist diese Sensitivität nur für infinitesimal kleine Änderungen gültig³. Die partielle Ableitung der Differentialgleichung (8.15) nach den Parametern $\boldsymbol{\alpha}$ lautet dann (totales Differential, (DEUFELHARD UND BORNEMANN, 2002)):

$$\mathbf{Sens}_{\boldsymbol{\alpha}}^{\mathbf{c}} = \frac{d}{dt} \frac{d\mathbf{c}}{d\boldsymbol{\alpha}} = \mathbf{N} \left(\frac{\partial \mathbf{v}}{\partial \mathbf{c}} \frac{d\mathbf{c}}{d\boldsymbol{\alpha}} + \frac{d\mathbf{v}}{d\boldsymbol{\alpha}} \right) \quad (8.17)$$

Abbildung 8.2 zeigt die Sensitivität (nicht normiert) der Metabolite bezüglich des Parameters $^{upt}v_{max}$.

Oftmals werden die Sensitivitäten als Konzentrationskoeffizienten wiedergegeben (MAUCH ET AL., 1997). Sie geben wieder, wie sich eine prozentuale Änderung eines Parameters auf die Zustandsgröße auswirkt (ebenfalls prozentual). D.h., die Sensitivität wird auf Parameter- und Zustandsgrößen normiert.

$$\mathbf{C}^{\mathbf{c}} = \text{diag}\{\mathbf{x}\}^{-1} \mathbf{Sens}_{\boldsymbol{\alpha}}^{\mathbf{c}} \text{diag}\{\boldsymbol{\alpha}\} \quad (8.18)$$

Neben der Sensitivität der Konzentrationen \mathbf{c} wird die Sensitivität der Flüsse \mathbf{v} benötigt. Die Flüsse sind abhängig der Konzentrationen und Parameter, deren Sensitivität lässt sich wieder nach der Kettenregel bestimmen:

$$\mathbf{v} = \mathbf{v}(\mathbf{c}, \boldsymbol{\alpha}) \quad \mathbf{Sens}_{\boldsymbol{\alpha}}^{\mathbf{v}} = \frac{d\mathbf{v}}{d\mathbf{c}} \frac{d\mathbf{c}}{d\boldsymbol{\alpha}} + \frac{d\mathbf{v}}{d\boldsymbol{\alpha}} \quad (8.19)$$

Auch diese Sensitivitäten werden meist normiert als Flusskontrollkoeffizienten wiedergegeben. Sie lauten dann:

$$\mathbf{C}^{\mathbf{v}} = \text{diag}\{\mathbf{v}\}^{-1} \mathbf{Sens}_{\boldsymbol{\alpha}}^{\mathbf{v}} \text{diag}\{\boldsymbol{\alpha}\} \quad (8.20)$$

Werden die Sensitivitäten der Messungen $\hat{\mathbf{y}}$ benötigt, wird das Messmodell nach $\boldsymbol{\alpha}$ abgeleitet:

$$\frac{\partial \hat{\mathbf{y}}}{\partial \boldsymbol{\alpha}}(t_i) = {}^i\mathbf{M} \frac{\partial \mathbf{c}}{\partial \boldsymbol{\alpha}}(t_i) \quad (8.21)$$

³Stichproben haben gezeigt, dass für Parametervariationen im Bereich von 10% immer noch eine sehr gute Näherung erreicht wird.

8.6.3 Parameteranpassung

In die Parameteranpassung fließen im Gegensatz zu den zuvor besprochenen stationären Modellen nun die zeitabhängigen Messungen $\mathbf{y}(t_i)$ ein. Für einen Zeitpunkt t_k lautet die Messgleichung:

$$\hat{\mathbf{y}}(t_k) = {}^k g(\mathbf{c}(t_k, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{c}_0)) \quad (8.22)$$

Die Messgleichungen aller Zeitpunkte t_k , $k = 1 \dots N$ werden in vektorieller Form zusammengefasst:

$$\underbrace{\begin{pmatrix} \mathbf{y}(t_1) \\ \mathbf{y}(t_2) \\ \vdots \\ \mathbf{y}(t_N) \end{pmatrix}}_{\hat{\mathbf{y}}(\boldsymbol{\xi})} = \underbrace{\begin{pmatrix} {}^1 g(\mathbf{c}(t_1, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{c}_0)) \\ {}^2 g(\mathbf{c}(t_2, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{c}_0)) \\ \vdots \\ {}^N g(\mathbf{c}(t_N, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{c}_0)) \end{pmatrix}}_{\boldsymbol{\xi} g(\mathbf{c}(\boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\alpha}^{\text{inp}}, \mathbf{c}_0))} \quad (8.23)$$

Die (beste) Schätzung der Parameter $\hat{\boldsymbol{\alpha}}$ wird wieder durch Minimieren der Fehlerquadratsumme bestimmt. Die Minimierung kann mit verschiedenen Algorithmen (Suplex (ROWAN, 1990), BFGS (ZHU ET AL., 1997) oder einem evolutionären Verfahren (BEYER, 2001)) durchgeführt werden.

Zum Ausschluss physiologisch oder thermodynamisch nicht zulässiger Lösungen, können die Parameter-, Fluss- und Konzentrationsbereiche beschränkt werden. Diese können als Nebenbedingungen der Optimierung eingeführt werden.

8.6.4 Statistische Auswertung

Zur statistischen Beurteilung der gefundenen Parameterschätzung $\hat{\boldsymbol{\alpha}}$ wird die Parameter-Kovarianzmatrix bestimmt. Benötigt werden wieder die Sensitivitäten aus Gl. (8.17) und die Messgleichungen in Gl. (8.23). Die Kovarianz der Parameter kann nach Gleichung (7.17) bestimmt werden.

8.6.5 Werkzeuge

Zur Lösung der dargestellten Gleichungen stehen verschiedene Werkzeuge zur Verfügung. Zu nennen sind v.a. Gepasi (MENDES, 1997), JWS Online (OLIVIER UND SNOEP, 2004) und E-Cell (TAKAHASHI ET AL., 2003). In der vorliegenden Arbeit wird ausschließlich das in Siegen entwickelte "Metabolit Modelling Tool 2" (MMT2) eingesetzt (HAUNSCHILD, 2006). Die Modellbeschreibung wird in einem XML-File (HAUNSCHILD,

2005) im von HAUNSCCHILD (2006) genannten M3L-Format (ein Dialekt von “System Biology Markup Language”, SBML, (HUCKA ET AL., 2003)) an den Generator übergeben (Abb. 8.3).

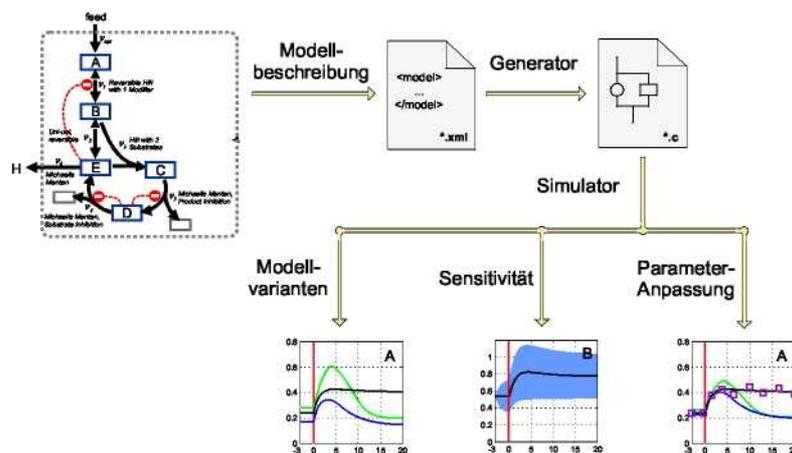


Abbildung 8.3: Schema der Simulation mit MMT2 - Das Modell wird als M3L-Datei vom Generator `mnt2_mk3im` eingelesen und ein Source-Code mit dem Modell (Simulator) generiert. Der Simulator wird kompiliert und kann über die Kommandozeile oder über MatLab gesteuert werden. Es können Simulationsläufe, Sensitivitäts-Berechnungen oder Parameter-Anpassungen durchgeführt werden.

Um eine schnelle Lösung zu erhalten, muss die Berechnung der rechten Seite der Differentialgleichung beschleunigt werden. In der Simulationstechnik bewährt hat sich die automatische Generierung von Source-Code. D.h., das in M3L-codierte Modell wird eingelesen und in C-Code umgesetzt. Zwar ergibt sich dadurch zunächst ein längerer Compilierungslauf, die Berechnungen werden aber wesentlich beschleunigt. Zur Lösung der Differentialgleichungen wird die Funktion `lsoda` (HINDMARSH, 1983) eingesetzt.

Mit MMT2 können auch die Sensitivitäten (Gl. (8.17)) berechnet werden. Für die Aufgabe der Parameteranpassung kann die Software auf einem Cluster-Computer ausgeführt werden. Nebenbedingungen werden in MMT2 über “penalty” Funktionen implementiert, die je nach Größe der Überschreitung des vorgegebenen Bereichs, der FQS χ^2 eine zusätzliche Abweichung P addiert. Um die Überschreitungen P in das gewünschte Verhältnis zu den Abweichungen χ^2 zu setzen, wird die Abweichung mit einem (wählbaren) Faktor p multipliziert⁴. Sowohl bei den Flüssen als auch den Konzentrationen wird die Abweichung der Simulation unter- oder oberhalb des zulässigen Bereichs über die Zeit t_{out} der Überschreitung integriert:

⁴Der Faktor kann im generierten C-Code verändert werden

$$P = p_v P_v + p_c P_c + p_\alpha P_\alpha \quad (8.24)$$

$$P_\alpha = \begin{cases} 0, & \alpha_{min} \leq \alpha \leq \alpha_{max} \\ \alpha_{max} - \alpha, & \alpha > \alpha_{max} \\ \alpha - \alpha_{min}, & \alpha < \alpha_{min} \end{cases} \quad (8.25)$$

$$P_c = \begin{cases} 0, & \mathbf{c}_{min} \leq \mathbf{c}(t) \leq \mathbf{c}_{max} \\ \int_{t_{out}} \mathbf{c}_{max} - \mathbf{c}(t) dt, & \mathbf{c}(t) > \mathbf{c}_{max} \\ \int_{t_{out}} \mathbf{c}(t) - \mathbf{c}_{min} dt, & \mathbf{c}(t) < \mathbf{c}_{min} \end{cases} \quad (8.26)$$

$$P_v = \begin{cases} 0, & \mathbf{v}_{min} \leq \mathbf{v}(t) \leq \mathbf{v}_{max} \\ \int_{t_{out}} \mathbf{v}_{max} - \mathbf{v}(t) dt, & \mathbf{v}(t) > \mathbf{v}_{max} \\ \int_{t_{out}} \mathbf{v}(t) - \mathbf{v}_{min} dt, & \mathbf{v}(t) < \mathbf{v}_{min} \end{cases} \quad (8.27)$$

Die Simulationsergebnisse können in csv-Dateien oder über ein MatLab-Interface abgerufen werden. MatLab ermöglicht dann einen effizienten Umgang mit den Simulationsdaten. In MatLab kann dann z.B. die gesamte statistische Auswertung vorgenommen werden.

Kapitel 9

Unterstützung der Auswertung mit visuellen Werkzeugen

In der Simulation sowie den statistischen Untersuchungen entstehen große Datenmengen. Im folgenden werden unterstützende visuelle Werkzeuge vorgestellt, insbesondere die von QELI (2007) entwickelte Visualisierung von metabolischen Netzwerken MetVis und zur Visualisierung großer Matrizen MatVis.

9.1 Visualisierung von Sensitivitäten

Zur Visualisierung wird die in Zahlen enthaltene Information in eine visuelle Darstellung in Farben transformiert. Die Matrix-Struktur bleibt dabei erhalten. Da die Einträge je nach Parameter unterschiedliche Skalen aufweisen, müssen die Daten in eine skaleninvariante Größe transformiert werden. Die naheliegende Skalierung sind die Fluss- bzw. Konzentrationskontrollkoeffizienten, die in Abschnitt 8.6.2 bereits eingeführt wurden. Die normierten Sensitivitäten müssen für die Visualisierung nun noch auf den Wertebereich $[-1; 1]$ projiziert werden. Dabei soll sichergestellt werden, dass die Skalierung über die Zeit gleich bleibt. So können Sensitivitäten verschiedener Zeitpunkte verglichen werden.

$$\mathbf{S}_{\alpha}^{vis} = \frac{1}{\max_k \max_i \max_j |\mathbf{C}_{\alpha_j}^{c_i}(t_k)|} \text{diag}\{\mathbf{c}\} \mathbf{C}_{\alpha}^c(t)$$

mit: $\mathbf{C}_{\alpha}^c(t) = \text{diag}\{\mathbf{x}(t)\}^{-1} \mathbf{Sens}_{\alpha}^c(t) \text{diag}\{\alpha\}$ (9.1)

Studien anhand einiger Beispiele zeigen, dass mit dieser Skalierung nur wenig Information dargestellt wird (Abb. 9.1 oben). Die Sensitivität einiger Pools bzw. Flüsse ist im Vergleich zu anderen sehr ausgeprägt. Als alternative Skalierung wird deshalb nun jede Zeile (d.h. für jeden Pool bzw. Fluss) getrennt skaliert

$$\mathbf{S}_{\alpha}^{vis, c_i} = \frac{1}{\max_k \max_j |\mathbf{C}_{\alpha_j}^{c_i}(t_k)|} \mathbf{C}_{\alpha}^{c_i}(t)$$

(9.2)

Zum Vergleich der Zeilen wird an der rechten Seite die Zeilennorm dargestellt. Mit dieser Skalierung werden Zeilen mit kleiner Zeilennorm vergleichsweise größer dargestellt (Abb. 9.1 unten).

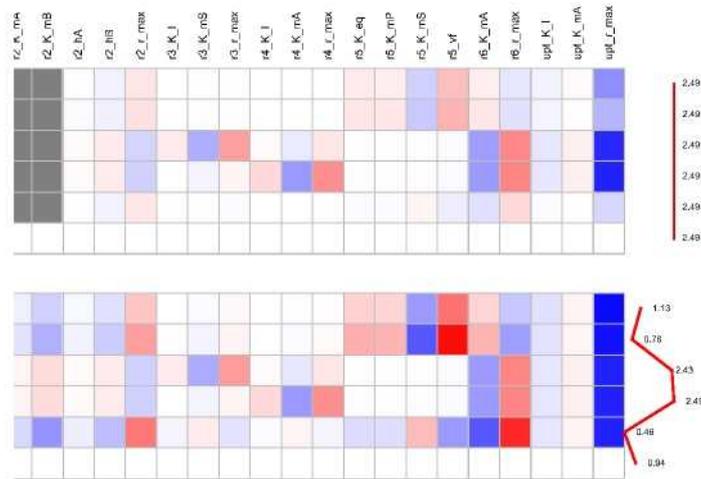


Abbildung 9.1: Darstellung der Sensitivitätsmatrix für das laufende Beispiel zum Zeitpunkt $t = 10s$. Oben wurde die globale Skalierung verwendet, im unteren Plot die Zeilennorm.

Eine weitere Skalierung ergibt die so genannte Output-Matrix. Sie bezieht die statistische Qualität der vorliegenden Messungen mit ein. Die Sensitivität der Messungen wird mit der Kovarianz der Messungen gewichtet:

$$\text{Out}(t_i) = {}^i \text{Sens}_{\alpha}^c(t_i) \text{Cov}^{-1}(\mathbf{y}(t_i))$$

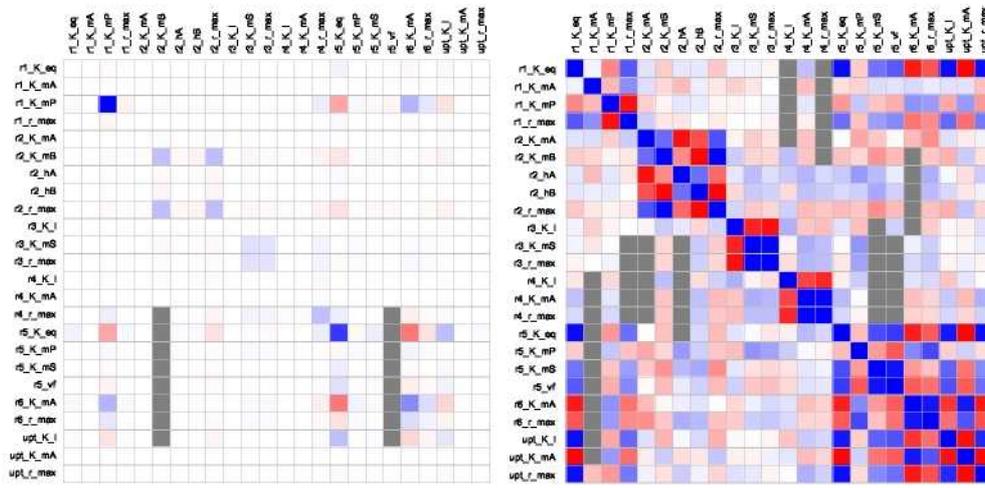
Für die Visualisierung muss dann eine weitere Skalierung, z.B. nach Gl. (9.1) vorgenommen werden.

9.2 Kovarianz-Matrizen

MatVis (QELI, 2007) wurde bewusst allgemein gehalten um beliebige Matrizen visuell darzustellen. So können auch die Kovarianzmatrizen eingelesen und visualisiert werden. Dazu wird eine Datei übergeben, die sowohl die Rohdaten als auch für die Visualisierung skalierte Daten (Wertebereich $[-1;1]$) enthält. Die Informationen der Kovarianzmatrix können wie folgt dargestellt werden:

Global skaliert: Die Kovarianz-Matrix wird auf den größten Eintrag (liegt auf der Diagonalen) skaliert.

Zeilennormiert: Jede Zeile der Kovarianz-Matrix wird auf ihren (absolut) größten Eintrag normiert. Die jeweilige Zeilennorm wird rechts der Matrix wiedergegeben.



(a) global skaliert

(b) Korrelations-Matrix

Abbildung 9.2: Darstellung der Kovarianzmatrix. Links: Global skaliert, Rechts als Korrelationsmatrix.

Korrelations-Matrix: Sie enthält die Daten bereits im Bereich $[-1; 1]$ (s. Abschnitt 7.2.2). Auf der rechten Seite werden die Varianzen der Kovarianzmatrix dargestellt.

Hauptkomponenten: Die Matrix mit den Eigenvektoren wird in der Matrixdarstellung visualisiert. Die Eigenwerte werden rechts neben der Matrix dargestellt.

In Abb. 9.2 sind die verschiedenen Skalierungsmethoden für das Beispielnetzwerk (Abb. 6.1) dargestellt.

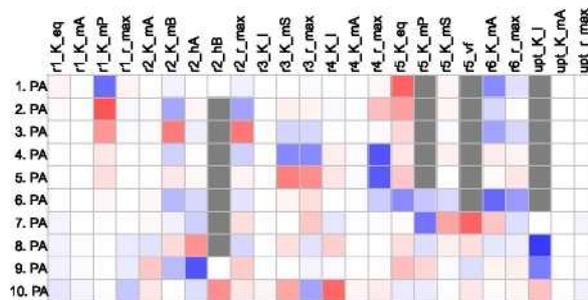


Abbildung 9.3: Hauptkomponentenanalyse der Kovarianzmatrix in Abb. 9.2.

9.3 Multidimensionale Skalierung

Die Multidimensionale Skalierung ist stark verwandt mit der Hauptkomponenten-Analyse, jedoch allgemeiner einsetzbar. Betrachtet wird der ‘‘Abstand’’ $d_{i,j}$ der Beobachtungen \mathbf{y}_i und \mathbf{y}_j mit jeweils n Merkmalen. Übliche Distanzmaße sind:

Manhattan-Distanz:

$$d(\mathbf{y}_i, \mathbf{y}_j) = L_1(\mathbf{y}_i, \mathbf{y}_j) = \sum_{k=1}^n |y_i - y_j|$$

Euklidischer Abstand

$$d(\mathbf{y}_i, \mathbf{y}_j) = L_2(\mathbf{y}_i, \mathbf{y}_j) = \sum_{k=1}^n \sqrt{|y_i - y_j|^2}$$

Minkowski-Metrik:

$$d(\mathbf{y}_i, \mathbf{y}_j) = L_p(\mathbf{y}_i, \mathbf{y}_j) = \sum_{k=1}^n \sqrt[p]{|y_i - y_j|^p}$$

Der Abstand der n Beobachtungen \mathbf{y} kann in einem n -dimensionalen Raum durch Punkte mit den entsprechenden Abständen dargestellt werden. Mehr als drei Dimensionen können visuell nicht mehr wiedergegeben werden. Mit dem Ansatz der Multidimensionalen Skalierung wird nun versucht, die Abstände des n dimensionalen Raumes so gut wie möglich in zwei oder drei Dimensionen wiederzugeben, die Position der n Punkte im Raum kann dabei auf verschiedenen Wegen bestimmt werden. In MatVis implementiert ist die klassische Multidimensionale Skalierung (COX UND COX, 2001) und das Sammon-Mapping (SAMMON, 1969).

Diese Verfahren werden nun eingesetzt, um die Sensitivitätsmatrix dynamischer Modelle zu untersuchen. Zu einem Zeitpunkt t_k wird der Abstand der Parameter-Sensitivitäten in zwei bzw. drei Dimensionen visualisiert (Abb. 9.4). Aus dieser Abbildung lassen sich Aussagen über die Ähnlichkeit von Sensitivitäten in Bezug auf das (beobachtbare) Systemverhalten treffen. Deutlich zeigt sich im Inneren eine ‘‘Parameterwolke’’, die ähnlichen, sehr geringen Einfluss auf das Modellverhalten nimmt. Einige Parameter stetzen sich von dieser Wolke ab (z.B. $^{upt}v_{max}$, $^{v2}v_{max}$, $^{v3}v_{max}$, $^{v4}v_{max}$, $^{v3}K_{mA}$, $^{v4}K_{mA}$). Diese Parameter bestimmen zu diesem Zeitpunkt hauptsächlich das Modellverhalten. Aus Abb. 9.1 erkennt man den auf alle Metabolite positiven Einfluss von $^{upt}v_{max}$, $^{v3}v_{max}$ und $^{v4}v_{max}$ erniedrigen die Konzentrationen von C und D. Die K_m -Werte der Reaktionen v3 und v4 wirken genau entgegengesetzt, eine höhere Michaelis-Menten Konstante wird die Geschwindigkeit dieser beiden Reaktionen verlangsamen.

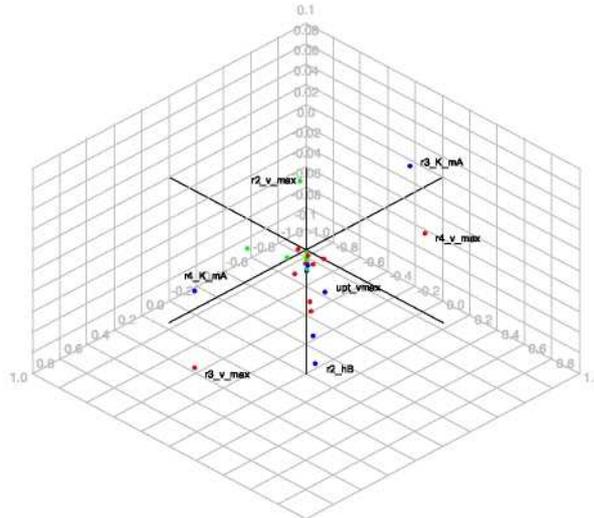


Abbildung 9.4: Klassische MDS der Sensitivitätsmatrix in Bezug auf die Parameter zum Zeitpunkt $t = 10s$.

Die Methode der Multidimensionalen Skalierung erkennt das unterschiedliche Verhalten der Parameter und gibt so einen schnellen Überblick. In MatVis wurde zusätzlich ein Clustering-Algorithmus (k -means, (JAIN ET AL., 1999)) implementiert (QELI, 2007).

Teil II

Stationäre Modellierung und Methoden zur Auswertung

Kapitel 10

Auswertung von ^{13}C Markierungsexperimenten im metabolisch stationären Zustand

10.1 Sensor-Reaktor Konzept

Industrielle Fermentationen laufen in großen Reaktoren mit einem Volumen von 50 L bis mehreren Kubikmetern. Ein ^{13}C Markierungsexperiment in diesem Maßstab durchzuführen ist extrem kostenintensiv (^{13}C markierte Glucose kostet derzeit ca. 70-100 € pro Gramm). Um Stoffflussanalysen auch unter Produktionsbedingungen anzuwenden, wurde das Sensor-Reaktor Konzept entwickelt (TAKORS ET AL., 2004; EL MASSAOUDI ET AL., 2003).

Neben dem Produktions-Reaktor wird ein kleiner "Sensor"-Reaktor betrieben. Um die Stoffflussanalyse durchzuführen, wird der Sensor-Reaktor aus dem Produktions-Reaktor befüllt und eine Parallel-Fermentation gestartet. Um die Produktionsbedingungen im Sensor-Reaktor abzubilden, wird im "Master-Slave-Modus" fermentiert. D.h. die Istwerte (z.B. pH, pO_2 , Feed, etc.) des Produktions-Reaktors sind die Sollwerte des Sensor-Reaktors (Abb. 10.1). In Versuchen wurde gezeigt, dass die Fermentationen bis zu 3 Stunden parallel geführt werden können (EL MASSAOUDI ET AL., 2003).

^{13}C Markierungsexperimente werden unter metabolisch stationären Bedingungen durchgeführt (WIECHERT, 2001). Insbesondere sollten während des Experiments die Flüsse konstant bleiben und ein über die Zeit konstantes ^{13}C Markierungsmuster des Substrats vorliegen.

Direkt nach der Inokulierung wird ein Markierungspuls aufgegeben. Zusammen mit dem noch im Fermentor enthaltenen unmarkierten Substrat stellt sich ein bestimmter Markierungsanteil ein. Um diesen Anteil konstant zu halten, wird nun ein Glucose-Feed mit dem selben Markierungsanteil eingeschaltet. In den hier durchgeführten Experimenten wurde ein Anteil von 25% ^{13}C markierter Glucose angestrebt.

Untersucht wurde ein L-Phenylalanin produzierender *E.coli* Stamm (*E.coli* 4pF81) auf der Basis von *E.coli* K12 LJ110 (ZEPPENFELD ET AL., 2000) (Abschnitt 3.2). Chromosomal deletiert wurden die Gene *pheA*, *tyrA*, *aroF*, die für die Enzyme Chorismat Mutase/Prephenat Dehydratase, Chorismat Mutase/Prephenat Dehydrogenase sowie die Tyrosin sensitive DAHP Synthase stehen. Das Plasmid (pJF119EH) trägt die Gene *aroF pheA^{fb}aroB aroL* für die Enzyme DAHP Synthase (Tyrosin feedback-inhibiert), Choris-

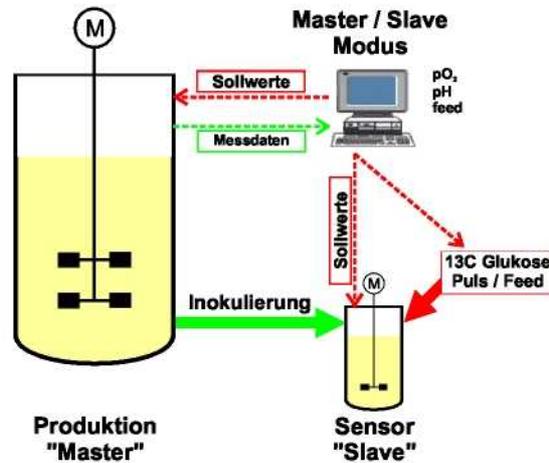


Abbildung 10.1: Funktionsschema des Sensor-Reaktor-Systems: Zu einem bestimmten Zeitpunkt wird der Sensor-Reaktor aus dem großen Produktions-Reaktor befüllt. Die Fermentation wird im Master/Slave Modus durchgeführt. D.h. die Istwerte des Produktions-Reaktors sind die Sollwerte des Sensor-Reaktors. Die Markierung wird durch ein Puls (^{13}C Glukose) zugegeben. Wird mit einem Feed gearbeitet, enthält dieser ebenfalls ^{13}C Glukose.

mat Mutase/Prephenat Dehydratase (eine feedback resistente Variante), 3-Dehydroquinate Synthase und Shikimat Kinase II (s. Abb. 10.2). Diese Gene stehen unter der Kontrolle des IPTG-induzierbaren Promotors P_{tac} . Das Plasmid trägt zudem *ampR* (Ampicillin-Resistenz) zur Selektion der plasmidtragenden Stämme. Medium sowie Fermentationsbedingungen werden in (WAHL ET AL., 2004b) beschrieben.

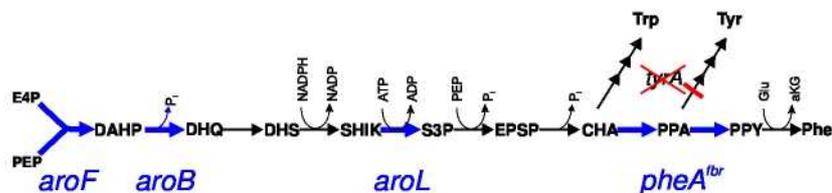


Abbildung 10.2: Der Stamm *E. coli* 4pF81 hat den Genotyp *E. coli* LJ110 $\Delta(pheA\ tyrA\ aroF)$ pJF119EH *aroF pheA^{fbr} aroB aroL ampR*. Durch die Deletion von *tyrA* ist der Stamm Tyrosin auxothroph.

10.2 Schätzung der extrazellulären Flüsse

Zur Schätzung der extrazellulären Flüsse wird ein Modell zur Beschreibung des makroskopischen Fed-Batch Betriebs aufgestellt. In der hier eingesetzten ^{13}C Modellierung geht man für den Beobachtungszeitraum von stationären Flüssen aus. Das bedeutet, dass sich die extrazellulären und intrazellulären spezifischen Substrat-Aufnahmerate, die Wach-

tumsrate sowie die Produktbildungsraten \mathbf{r} über die Zeit nicht ändern $d\mathbf{r}/dt = \mathbf{0}$. Nicht stationär sind hingegen die Konzentrationen. Im Fed-Batch Betrieb gibt es keinen Abfluss, d.h. Biomasse und Produkte akkumulieren über die Zeit im Reaktor.

Im Fed-Batch Betrieb verändert sich durch den Feed und die Probenahmen ständig das Fermentationsvolumen V im Sensor-Reaktor. Der Zufluss $\dot{V}_{\text{feed}}(t)$ mit der Substratkonzentration S_{feed} wird als Profil vom Master-Reaktor vorgegeben und entspricht so den Zugaben im großen Produktionsreaktor.

Zur Beschreibung des quasi-stationären Fed-Batch Experiments wird ein Differentialgleichungssystem mit den Zustandsvariablen V (Reaktor-Volumen), X (Biomassekonzentration [g/l]), S (Substratkonzentration [mM]), P_1 (Produktkonzentration [mM]) und die Nebenproduktkonzentrationen P_2, \dots, P_n aufgestellt. Das Volumen ändert sich durch den Zufluss $\dot{V}_{\text{feed}}(t)$:

$$\dot{V} = \dot{V}_{\text{feed}}(t)$$

Die Biomassekonzentration X ergibt sich aus dem Biomassewachstum μX und der Verdünnung durch den Feed \dot{V}_{feed} :

$$\dot{X} = \mu X - \frac{\dot{V}_{\text{feed}}(t)}{V} X$$

Die Substratkonzentration S ergibt sich aus dem Zufluss $\dot{V}_{\text{feed}}(t)$ mit der Konzentration S_{feed} , der spezifischen Substrat-Aufnahmerate r_S sowie der Verdünnung durch den Zufluss:

$$\dot{S} = \frac{\dot{V}_{\text{feed}}(t)}{V} S_{\text{feed}} - r_S X - \frac{\dot{V}_{\text{feed}}(t)}{V} S \quad (10.1)$$

Die Produktkonzentration ergibt sich aus der Produktbildungsrate $r_{P_i} X$ und der Verdünnung zu:

$$\dot{P}_i = r_{P_i} X - \frac{\dot{V}_{\text{feed}}}{V} P_i \quad \text{mit: } i = 1 \dots n \quad (10.2)$$

Zur besseren Handhabung wird das Differentialgleichungssystem in Vektornotation zusammengefasst:

$$\dot{\mathbf{c}} = \mathbf{f}(\mathbf{c}, \mathbf{r}, \boldsymbol{\alpha}^{\text{inp}}) \quad (10.3)$$

$$\begin{aligned} \mathbf{c} &= (V, X, S, P_1, \dots, P_n)^T \\ \mathbf{r} &= (\mu, r_S, r_{P_1}, \dots, r_{P_n})^T \\ \boldsymbol{\alpha}^{\text{inp}} &= (\dot{V}_{\text{feed}}(t), S_{\text{feed}})^T \\ \mathbf{c}_0 &= (X_0, S_0, P_{1,0}, \dots, P_{n,0})^T \end{aligned}$$

Damit liegt dieses Modell wieder in der allgemeinen Darstellung vor und die in Abschnitt 8.6 beschriebenen Anwendungen lassen sich direkt übertragen.

Um auch die Probenahmen in die Betrachtung einzubeziehen, wird die Lösung des Systems zu Probenahmezeiten unterbrochen und das Volumen um die Entnahme $V_{\text{sample}}(t_k)$

verringert. Danach wird die Simulation mit dem verringerten Volumen fortgesetzt. Zusätzliche Unterbrechungen der Lösung ergeben sich bei einer Veränderung des Feed-Stroms.

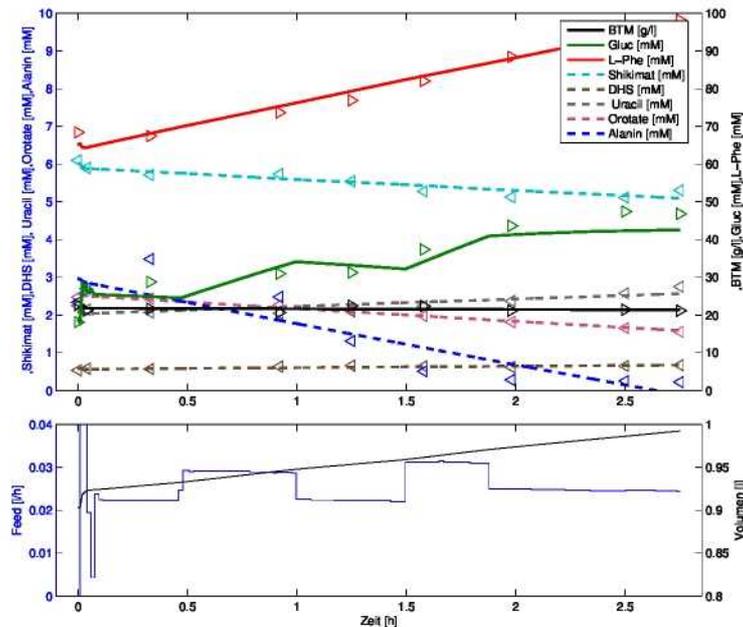


Abbildung 10.3: Oberer Plot: Anpassung der extrazellulären Konzentrationen unter der Annahme stationärer Raten (erstes Sensor-Reaktor Experiment). Im unteren Plot sind die Glucose-Zufuhr und der Verlauf des Reaktorvolumens V dargestellt.

Die Ergebnisse für die drei aufeinander folgenden Sensor-Reaktor Experimente sind in Tab. 10.1 zusammengefasst, die Anpassung des ersten Experiments (Phase A) ist in Abb. 10.3 dargestellt. Mit angegeben sind auch die jeweiligen Standardabweichungen der Schätzungen. Sie wurden anhand der Fehlerübertragung aus den Standardabweichungen der 9 Messzeitpunkte bestimmt. Es fällt auf, dass die Flüsse relativ genau bestimmt werden. Die Korrelationen sind in den hier betrachteten Auswertungen sehr gering, da die die Zellen nicht mehr wachsen ($\mu = 0$).

10.3 Korrektur der ^{13}C Markierungsmessung

Im Gegensatz zu Markierungsexperimenten im Chemostat wird sich für den Fed-Batch Betrieb kein stationärer Markierungszustand für die Produkte bzw. die Biomasse einstellen. Im Chemostaten kommt es durch den Durchfluss zu einem Auswaschen der zu Beginn des Experiments unmarkierten Produkte (bzw. Biomasse). Im Fed-Batch Betrieb bleibt der unmarkierte Anteil jedoch im Reaktor.

Tabelle 10.1: Geschätzte extrazelluläre Flüsse für die drei Sensor-Reaktor-Experimente.

| Rate | | Phase | | |
|----------------|--------------------------|------------------|------------------|------------------|
| | | A (14 – 16.8 h) | B(17.2 – 20 h) | C(20.5 – 23.3 h) |
| μ [1/h] | | 0.01 ± 0.01 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| Substr. | Aufnahme | 3.42 ± 0.02 | 2.70 ± 0.02 | 2.40 ± 0.02 |
| | [mM/h/g _{BTM}] | | | |
| L-Phenylalanin | | 0.66 ± 0.03 | 0.45 ± 0.02 | 0.25 ± 0.02 |
| | [mM/h/g _{BTM}] | | | |
| Shik. | [mM/h/g _{BTM}] | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| DHS | [mM/h/g _{BTM}] | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| Uracil | [mM/h/g _{BTM}] | 0.01 ± 0.00 | 0.02 ± 0.00 | 0.01 ± 0.00 |
| Orotat | [mM/h/g _{BTM}] | -0.01 ± 0.00 | -0.01 ± 0.00 | -0.01 ± 0.00 |
| Alanin | [mM/h/g _{BTM}] | -0.05 ± 0.00 | -0.03 ± 0.00 | -0.02 ± 0.02 |

Somit kann die für Chemostat Experimente übliche Wash-Out Korrektur zur Extrapolation des stationären Zustands (MARX ET AL., 1996; VAN WINDEN ET AL., 2001; WIECHERT UND NOH, 2005) nicht durchgeführt werden. Die stationäre Markierungsanreicherung wird stattdessen über eine Verdünnungskorrektur geschätzt. Dazu wird der Anteil des vor dem Markierungsexperiment vorhandenen Produkts aus der Markierungsmessung “herausgerechnet”.

Tabelle 10.2: ^{13}C Markierungsanreicherungen (aufgrund der Verdünnung korrigierte Werte) für die ^{13}C Modellierung.

| | Messung im Überstand | | Anteil neu synthetisierten L-Phe | stationäre Markierung | |
|----------|----------------------|-------|----------------------------------|-----------------------|--------|
| | C3 | C5+C9 | p | C2 | C5+C9 |
| A | 4.50% | 3.90% | 39.26% | 9.76% | 8.23% |
| B | 3.80% | 3.10% | 23.66% | 12.51% | 9.55% |
| C | 3.20% | 2.60% | 11.83% | 18.86% | 13.78% |

Um diese Korrektur durchführen, wird angenommen, dass im Zentralstoffwechsel sofort nach dem Markierungspuls der isotopisch stationäre Zustand erreicht wird. D.h. von Beginn des Markierungsexperiments an wird markiertes Produkt hergestellt. Diese Annahme ist gültig, wenn die Zeitkonstante zum Erreichen des stationären Zustands (Zentralstoffwechsel) im Vergleich zur Dauer des Experiments kurz ist¹.

Während der Markierungsphase wird neues, markiertes Produkt gebildet. Am Ende liegt es mit einem Anteil p des gesamten Produkts vor. Die Markierungsmessung am Ende des Experiments (y') kann als Mischung des Messsignals y_x des Anteils p und dem Messsignal y_n des natürlich markiert vorliegenden Anteils $1 - p$ dargestellt werden:

¹Erste Experimente und Simulations-Studien zeigen, dass der Zentralstoffwechsel zumindest für die Glykolyse bereits nach ca. 10 Minuten einen stationären Zustand erreicht (GRÖNKE, 2007)

$$y' = p y_x + (1 - p) y_n$$

Bei bekannten Anteilen p und bekannter Signalstärke für unmarkiertes Produkt kann nach der unbekanntem Messung y_x umgestellt werden. Gemessen wurde hier mittels ^1H NMR, die Höhe des Singulets ergibt sich aus der Markierungsanreicherung an des untersuchten C-Atoms. Mit den Produktmengen n_1 am Anfang und n_2 am Ende der Markierungsphase² und einer natürlichen Wahrscheinlichkeit $^{13}\text{C}\gamma$ ergibt sich y_x zu

$$\begin{aligned} y' &= p y_x + (1 - p) ^{13}\text{C}\gamma & \text{mit: } p &= \frac{n_2 - n_1}{n_2} & ^{13}\text{C}\gamma &= 0.011 \\ y_x &= \frac{y' n_2 - n_1 ^{13}\text{C}\gamma}{n_2 - n_1} \end{aligned} \quad (10.4)$$

Die so korrigierten Markierungsmessungen der drei Experimente sind in Tabelle 10.3 dargestellt. Neben der Markierungsanreicherung wird auch der Messfehler für die stationäre Markierungsanreicherung benötigt. Wird die Messung y' mit einem Fehler $\sigma(y')$ (für die Konzentrationsmessungen wurde ein Fehler von 10% angenommen) und die Mengen n_1 , n_2 mit den Fehlern $\sigma(n_1)$, $\sigma(n_2)$ bestimmt (berechnen sich aus der Simulation des Konzentrationsverlaufs), ergeben sich für die korrigierte Messung y_x nach der Fehlerfortpflanzung die in Tabelle 10.3 angegebenen Fehler.

10.4 Netzwerk zur Auswertung

Das ^{13}C Modell beschreibt die Stöchiometrie sowie die C-Atom-Transitionen im Netzwerk. Das aufgestellte metabolische Netzwerk (Abb. 10.4 und Tab. A.2) enthält die Reaktionen des Zentralstoffwechsels (Glykolyse, Pentose-Phosphat-Weg und Zitronensäure-Zyklus) sowie des Aromatenbiosynthesewegs. Das Netzwerk musste allerdings im Vergleich zu Untersuchungen in der Wachstumsphase mit Messungen in der Biomasse (z.B. (MARX ET AL., 1996)) wesentlich vereinfacht werden. Bei der hier durchgeführten Untersuchung der Produktionsphase reichert sich die Markierung nur im Zentralstoffwechsel und dem Produkt L-Phenylalanin an. Zum Zeitpunkt der Untersuchungen waren intrazelluläre Markierungsmessungen noch nicht möglich, zur Verfügung stehen lediglich die Markierungsmessungen des extrazellulär angereicherten Produkts. Das Modell beinhaltet folgende Reaktionen und Vereinfachungen:

²Die Mengen n_1 und n_2 werden aus der Anpassung des makroskopischen Modells entnommen und bezüglich der Probenahmen korrigiert. Eine Probe zu Beginn des Markierungsexperiments enthält hauptsächlich unmarkiertes Produkt, während gegen Ende das markiert synthetisierte Produkt überwiegt.

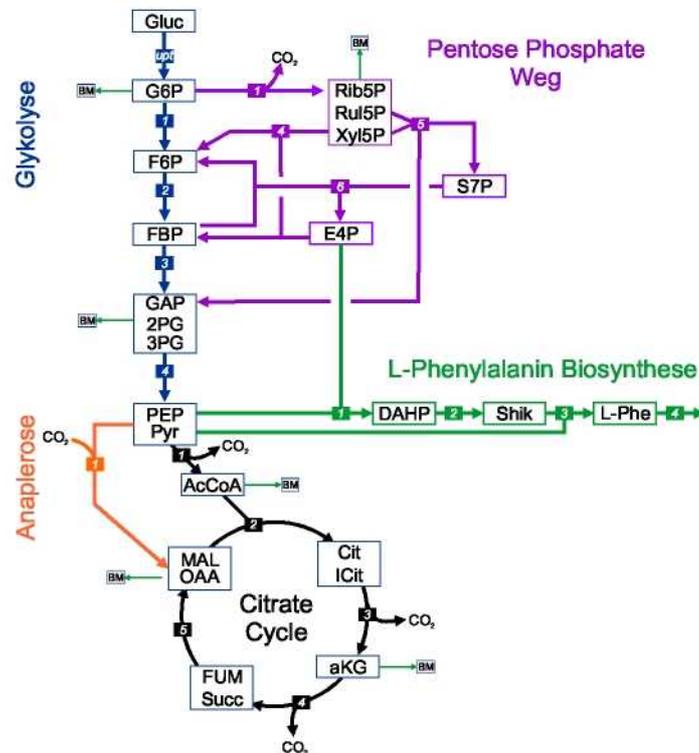


Abbildung 10.4: Metabolisches Netzwerk des *L*-Phenylalanin produzierenden Stamm *E. coli* 4pF81. Einige Metabolit-Pools wurden zur Erzwingung der Identifizierbarkeit zusammengefasst. Die zum Modell gehörigen C-Atom-Transitionen sind in Tabelle A.2 aufgeführt.

Glykolyse: Die Glucose-Aufnahme durch das PTS-System (Phosphotransferase System) wird im Modell durch zwei Flüsse (1-¹³C markiert und natürlich markiert) wiedergegeben um das extrazelluläre Markierungsverhältnis abbilden zu können. Die Metabolite DAHP, GAP, 2-PG und 3-PG wurden zu einem Pool zusammengefasst. Genauso mussten PEP und Pyr in einem Pool zusammengefasst werden, da die Flüsse zwischen PEP und Pyr mit den vorliegenden Messungen nicht identifizierbar sind. Des weiteren wurde der Austauschfluss zwischen F6P und FBP aus dem Modell entfernt.

Pentose-Phosphat Weg: Die Pools Rib5P, Rul5P und Xyl5P stehen im schnellen Gleichgewicht und wurden in einem Pool zusammengefasst. Die Austauschreaktionen ppp4, ppp5 und ppp6 mussten aus dem Modell genommen werden, da sie mit den vorhandenen Messungen nicht identifizierbar sind.

Zitronensäure-Zyklus: Cit und ICit wurden in einem Pool zusammengefasst, genauso OAA und Mal. Zwischen Succ und Mal wird eine Scrambling-Reaktion³ eingefügt, da Succ symmetrisch ist.

Anaplerotische Reaktionen: In *E. coli* sind anaplerotische Reaktionen zwischen PEP und OAA (PEP Carboxylase / PEP Carboxykinase) sowie Pyr und Mal (Malate-Enzym) beschrieben. Im Modell mussten die verschiedenen Reaktionen in einer Reaktion zwischen den gelumpten Pools PEP/Pyr und OAA/Mal zusammengefasst werden – zudem kann nur der Nettofluss bestimmt werden. Aus dem Modell genommen wurden auch die anaplerotischen Reaktionen im TCA – unter Glucose gesättigten Bedingungen wird üblicherweise davon ausgegangen, dass der Glyoxylat-Shunt inaktiv ist (FISCHER UND SAUER, 2003; PEEKHAUS UND CONWAY, 1998).

Nicht in das Modell aufgenommen wurden der Entner-Doudorhoff Weg (ED), da angenommen wird, dass dieser Reaktionsweg unter aeroben Bedingungen inaktiv ist (FISCHER UND SAUER, 2003).

10.5 Flussschätzungen und statistische Auswertung

Die Schätzung der Flüsse wurde mit der Software 13C-FLUX (WIECHERT ET AL., 2001) durchgeführt. Die Anpassung an die Messdaten wurde auf dem Siegener “Simtec“-Cluster durchgeführt. Zunächst wurde mit einem evolutionären Verfahren gearbeitet um möglichst ein globales Minimum im Flussraum zu finden. Diese Fluss-Schätzung wurde dann mit einem lokalen Optimierungsverfahren (BFGS) verfeinert. Die geschätzten intrazellulären Flüsse sind in Tab. A.3 dargestellt.

Zur Beurteilung der Anpassungsgüte wird zunächst die Fehlerquadratsumme mit den tabellierten χ^2 -Werten verglichen. Die beobachtete Abweichung liegt für das Experiment C über dem nach der χ^2 -Tabelle zulässigen Wert (Tab. A.1). Die FQS sollte für einem Vertrauensbereich $\beta = 0.995$ zwischen 10.59 und 0.01 liegen. Besonders auffällig ist dabei der Beitrag der CO_2 Ausscheidungsrate. Sowohl im ersten als auch dritten Markierungsexperiment (Phase A bzw. C) liegt der geschätzte Wert über der gemessenen Ausscheidungsrate (im ersten Experiment wurde die Detektorgrenze von 5% CO_2 erreicht).

Für die Beurteilung der Flusslagen ist die statistische Untersuchung mit entscheidend. Zunächst erkennt man einen deutlichen Abfall der statistischen Qualität je später das Experiment durchgeführt wird, d.h. von Phase A nach C. Dies ist darin begründet, dass die kor-

³Scrambling bedeutet, dass die Zuordnung der C-Atome durch diesen Reaktionsschritt nicht eindeutig ist. Die Reaktion wird durch zwei Reaktionen a und b mit unterschiedlichen Zuordnungen (meist einer Spiegelung) dargestellt (MARX ET AL., 1996)

Tabelle 10.3: Mit ^{13}C -FLUX geschätzte Flüsse mit Standardabweichung.

| | | A | | B | | C | |
|----------------------|-----------------|-------|-------|-------|-------|-------|-------|
| | | Fluss | Abw. | Fluss | Abw. | Fluss | Abw. |
| Glykolyse | upt1 | 0.84 | 0.023 | 0.67 | 0.018 | 0.56 | 0.016 |
| | upt2 | 2.53 | 0.023 | 2.02 | 0.018 | 1.71 | 0.016 |
| | emp1 | 1.86 | 0.442 | 1.69 | 0.585 | 1.99 | 1.066 |
| | emp2 | 2.66 | 0.149 | 2.21 | 0.196 | 2.09 | 0.356 |
| | emp3 | 2.66 | 0.149 | 2.21 | 0.196 | 2.09 | 0.356 |
| | emp4 | 5.43 | 0.161 | 4.45 | 0.202 | 4.09 | 0.358 |
| Pentose-Phosphat-Weg | ppp1 | 1.50 | 0.443 | 1.01 | 0.585 | 0.28 | 1.066 |
| | ppp2 | 0.80 | 0.295 | 0.52 | 0.390 | 0.10 | 0.711 |
| | ppp3 | 0.70 | 0.148 | 0.49 | 0.195 | 0.18 | 0.356 |
| | ppp4 | 0.10 | 0.148 | 0.03 | 0.195 | -0.09 | 0.355 |
| | ppp5 | 0.70 | 0.148 | 0.49 | 0.195 | 0.18 | 0.356 |
| | ppp6 | 0.70 | 0.148 | 0.49 | 0.195 | 0.18 | 0.356 |
| TCA | tcc1 | 4.21 | 0.178 | 3.53 | 0.213 | 3.52 | 0.364 |
| | tcc2 | 4.21 | 0.178 | 3.53 | 0.213 | 3.52 | 0.364 |
| | tcc3 | 4.21 | 0.178 | 3.53 | 0.213 | 3.52 | 0.364 |
| | tcc4 | 4.21 | 0.178 | 3.53 | 0.213 | 3.52 | 0.364 |
| | tcc5 | 4.21 | 0.178 | 3.53 | 0.213 | 3.52 | 0.364 |
| | ana1 | 0.01 | 0.020 | 0.02 | 0.002 | 0.03 | 0.020 |
| Shikimat-Weg | shik1 | 0.60 | 0.024 | 0.45 | 0.022 | 0.27 | 0.020 |
| | shik2 | 0.60 | 0.024 | 0.45 | 0.022 | 0.27 | 0.020 |
| | shik3 | 0.60 | 0.024 | 0.45 | 0.022 | 0.27 | 0.020 |
| | phe | 0.60 | 0.024 | 0.45 | 0.022 | 0.27 | 0.020 |
| Neben-Prod. | CO ₂ | 14.73 | 0.295 | 12.01 | 0.246 | 11.08 | 0.224 |
| | uracil | 0.01 | 0.020 | 0.02 | 0.002 | 0.03 | 0.020 |

rigierten Markierungsmessungen mit immer höheren Ungenauigkeiten behaftet sind. Im letzten Experiment ist die Produktivität bereits stark abgesunken und während der Markierungsphase wurden nur noch ca. 12% des Produkts neu gebildet. Die ^{13}C Markierungsanreicherung wird dadurch stark verdünnt und das ^{13}C Signal wird sehr schwach. Zudem ergibt sich aufgrund der starken Verdünnung ein sehr hoher Korrekturfaktor (Tab. 10.3). Die Messfehler der extrazellulären Raten bleiben hingegen in etwa konstant (Tab. 10.1). Sie sind unabhängig von der Verdünnung, sie hängen lediglich von der Genauigkeit der Konzentrationsmessungen ab.

In Abb. 10.5 sind die wichtigsten statistischen Größen visuell dargestellt (für Experiment A). Folgendes lässt sich beobachten:

Korrelation (Abb. 10.5a): Zu erkennen ist eine geringe Korrelation der Flüsse **ppp1** und **upt1**, ansonsten sind keine gegenseitigen Beeinflussungen der *freien* Flüsse zu erkennen. Dies ist auf die starken Vereinfachungen des metabolischen Netzwerks zurückzuführen (insbesondere keine Austauschflüsse).

Parameter Sensitivität (Abb. 10.5b): Die Sensitivitäten $\partial\mathbf{y}/\partial\mathbf{v}_f$ werden gewichtet nach den Messfehlern $\mathbf{Cov}(\mathbf{y})$ dargestellt. Es zeigt sich:

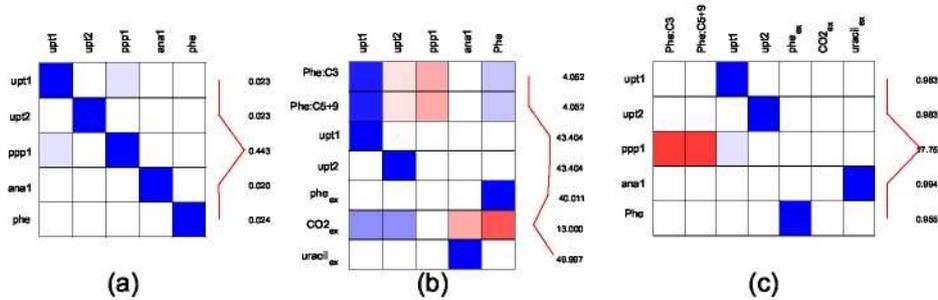


Abbildung 10.5: Visuelle Darstellung der Parameter Korrelationsmatrix, d.h. der Korrelation der freien Flüsse $\text{Corr}(\mathbf{v}_f)$ (links). Die Sensitivität $(\partial \mathbf{y} / \partial \mathbf{v}_f)$ (mitte) gibt wieder, wie stark die Messsignale $\hat{\mathbf{y}}$ von den freien Flüssen beeinflusst wird. Die Output Sensitivität $(\partial \mathbf{v}_f / \partial \mathbf{y})$ (rechts) zeigt die Sensitivität der freien Flüsse gegenüber den Messungen.

- Die Markierungsmessungen *Phe-C₃* und *Phe-C₅₊₉* zeigen eine positive Sensitivität gegenüber den Flüssen *upt1* ($1\text{-}^{13}\text{C}$ Glucose-Aufnahmerate) und *phe* (L-Phenyl^l-alanin Bildung). Eine negative Sensitivität erkennt man für die Flüsse *ppp1* und *upt2*. Durch *ppp1* wird das markierte $^{13}\text{C}_1$ abgespalten (CO_2), der Markierungsanteil in L-Phe sinkt. *upt2* ist die Aufnahmerate der unmarkierten Glucose, erhöht sich diese Rate, wird der Markierungsanteil ebenfalls geringer. *ana1* nimmt keinen Einfluss auf die Markierungsmessungen⁴
- *ana1* hängt direkt mit dem Abfluss des C4-Körpers Mal/OAA aus dem TCA zusammen. *ana1* beeinflusst zudem die CO_2 Messung.
- Die CO_2 Rate wird von allen Flüssen (außer *ppp1*) beeinflusst⁵.
- Die drei freien Flüsse *upt1*, *upt2* und *phe* werden gemessen, diese Messungen sind folglich direkt sensitiv.

Output-Sensitivitäten Aus den Sensitivitäten (Abb. 10.5c) lässt sich schnell erkennen, dass die drei Flüsse *upt1*, *upt2* und *phe* mit der Genauigkeit der entsprechenden Messungen bestimmt werden (Sensitivität beträgt ca. 0.95-0.98). *ana1* wird über die Messung von *ura* und CO_2 festgelegt. Als einziger Fluss ist *ppp1* hauptsächlich von den Markierungsmessungen abhängig. Die Anreicherungs-Messung wird unter anderem durch den Fluss *upt1* beeinflusst, somit entsteht die beobachtete Korrelation (Abb. 10.5a) *ppp1* und *upt1*.

⁴Einfluss nehmen würde ein Austauschfluss *ana1_{xch}*, der in diesem Netzwerk nicht enthalten ist.

⁵Eigentlich ist hier eine direkte Beeinflussung zu erwarten. Durch die Wahl der freien Flüsse wird bewirkt, dass eine Erhöhung von *ppp1* zu einem niedrigeren Abfluss in den TCA führt. Somit wird zwar mehr CO_2 in *ppp1* gebildet, die CO_2 -Bildung im TCA nimmt aber ab.

10.6 Ergebnisse und Diskussion

Während der Produktionsphase sinkt die Ausbeute an L-Phenylalanin deutlich ab. Zu Beginn (Phase A) erreicht der Stamm noch eine Ausbeute von ca. 18% $\text{mol/mol}_{\text{upt}}$, am Ende (Phase C) sind es nur noch knapp 12% $\text{mol/mol}_{\text{upt}}$ (Abb. 10.6). Mit der Abnahme der Produktivität beobachtet man eine Zunahme der TCA-Aktivität und eine Abnahme des Flusses in den oxidativen Pentose-Phosphat-Weg. Über die Gründe dieses Verhaltens kann bei stationären Stoffflussdaten lediglich spekuliert werden. Die TCA-Aktivität hängt wesentlich von dem Fluss durch die Pyruvat-Kinase (PykA/PykF) und der entgegengerichteten PEP-Synthase (Pps) ab. Der erhöhte Abfluss kann somit durch ein Nachlassen der Pps-Aktivität als auch einem Anstieg der PykA/PykF Aktivität erklärt werden. In Betracht ziehen kann man folgende Hypothesen:

- Es ist bekannt, dass PykA/PykF durch FBP aktiviert werden (VALLE ET AL., 1996). Nimmt man an, dass bei einem erhöhten Fluss durch die Glykolyse auch die FBP Konzentration ansteigt, würde PykA/PykF stärker aktiviert.
- In der Literatur wird eine Inhibierung der PEP-Synthase (Pps) durch Metabolite des TCA (OAA und aKG) beschrieben (COOPER UND KORNBERG, 1967; CHULAVATNATOL UND ATKINSON, 1973). Eine erhöhte aKG-Konzentration könnte u.a. durch ein erhöhten Glutamatbedarf für die Transaminierung von PPY bei hohen L-Phenylalanin Konzentrationen hervorgerufen werden. Die Transaminierung ist eine Gleichgewichtsreaktion ($\Delta_R G^0 \sim 0 \text{ kJ}/(\text{K M})$). Damit die Reaktion in Richtung L-Phe abläuft, muss eine hohe Eduktkonzentration (PPY, Glu) vorliegen. Glu wird aus aKG regeneriert. Je nach extrazellulärer Ammoniumkonzentration wird Glu über die Glutamat-Dehydrogenase (hohe NH_4^+ Konzentration) oder den Urea-Zyklus (sehr hohe Affinität) bereitgestellt ((LENGELETER ET AL., 1999)).

Bei den hier ausgewerteten Experimenten musste das Netzwerk stark vereinfacht werden. Aufgrund der eingesetzten Messtechnik lagen zudem nur wenige Markierungsmessungen vor. Zusätzlich werden die Markierungsanreicherungen durch unmarkierte Substanzen vor dem Markierungs-Experiment verdünnt. Zu bedenken ist zusätzlich, dass die Annahme eines sehr schnellen Einstellens des isotopisch stationären Zustands im Zentralstoffwechsel getroffen werden muss. Fraglich ist der Einfluss von möglichen Markierungssenken, wie z.B. den unmarkierten Proteinen. Neuere Experimente zeigen (NÖH ET AL., 2007), dass Proteinabbau möglicherweise die Markierung der Zentralstoffwechsel-Metabolite beeinflusst. Eine Abschätzung der Auf- und Abbauraten der Proteine im wachstumslimitierten

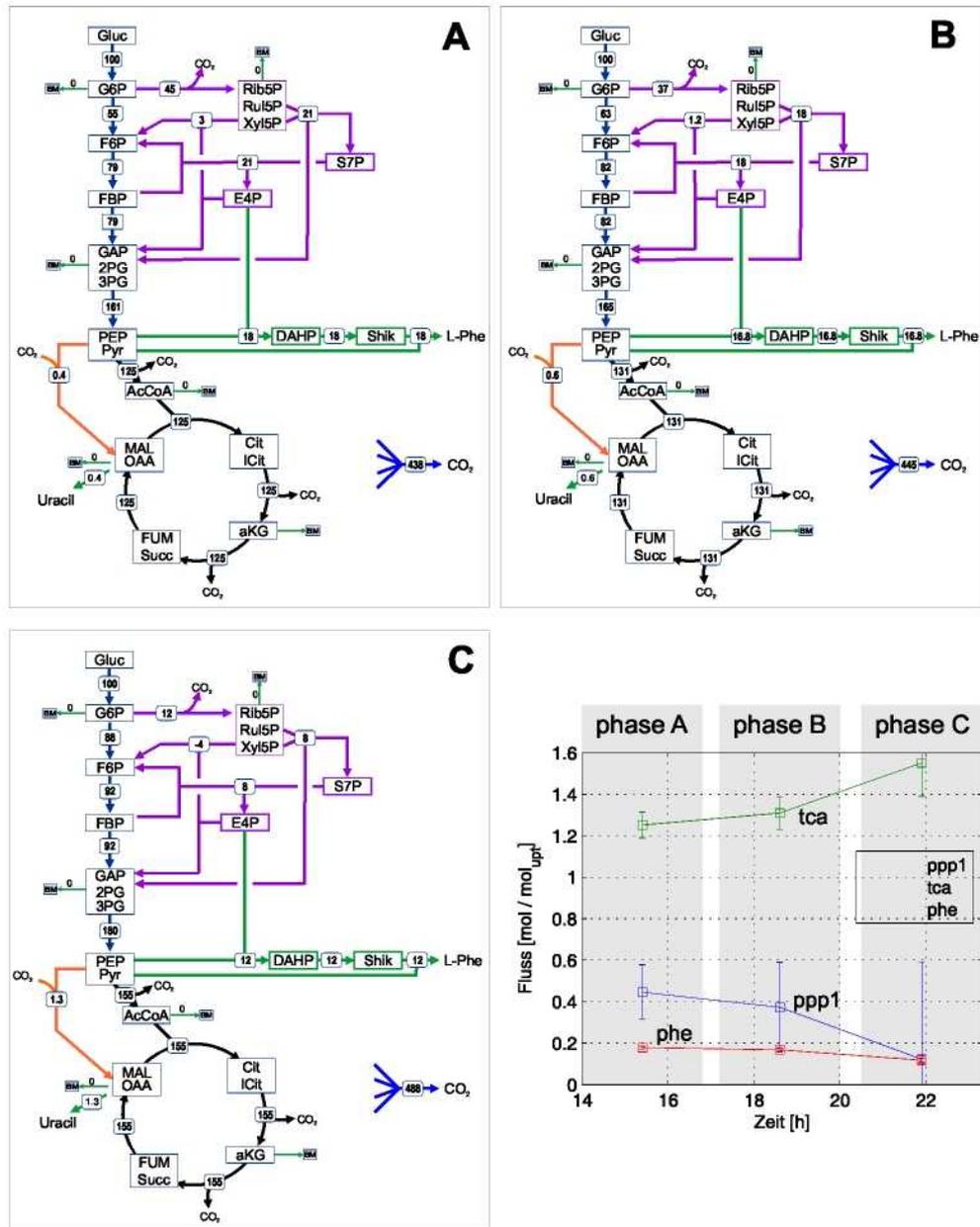


Abbildung 10.6: Darstellung der relativen Flüsse für die drei Experimente.

Zustand liegt zur Zeit noch nicht vor. Aufschluss über die Abbauraten könnten gezielte isotopisch instationäre Markierungsexperimente (Methodik in (NÖH ET AL., 2006)) liefern.

Ein deutlicher Vorteil der Sensor-Reaktor Experimente ist, dass jetzt zeitliche Veränderungen der Stoffflüsse unter Produktionsbedingungen beobachtet werden können. Trotz der hohen Standardabweichungen lassen sich auch bei dem hier untersuchten Stamm klare Tendenzen zu erkennen.

10.7 Möglichkeiten zur Steigerung der Bestimmtheit

Die Möglichkeiten, das experimentelle Vorgehen zu verändern sind bei Sensor-Reaktor Experimenten leider gering. Folgende Parameter können angepasst werden, um eine höhere Genauigkeit zu erreichen:

Input-Substrat: Statt nur 1-¹³C markierter Glucose kann ein Mix aus z.B. 1-¹³C und vollmarkierter U-¹³C Glucose zugegeben werden. Einzige Einschränkung bei der Wahl der Substrat-Zusammensetzung ist die Vorgabe, vor dem ¹³C Puls in der Substrat-Sättigung zu bleiben, sowie durch den Puls eine Konzentration von 10 g/l Glucose nicht zu überschreiten (um eine erhöhte Acetat-Bildung zu vermeiden). Zudem müssen im Master- und Sensorreaktor vergleichbaren Bedingungen herrschen, d.h. die Zugabe sollte nicht zu veränderten Umgebungsbedingungen führen. Die Substrat-Zusammensetzung wird in der Optimierung deshalb auf einen Anteil $p \geq 50\%$ nicht-markierter Glucose beschränkt.

Markierungsdauer: Um den Anteil des markierten Produkts zu steigern, kann die Dauer der Markierungsphase verlängert werden. Bedingung sind stationäre Flüsse sowie ein paralleler Fermentationsverlauf im Master- und Sensor-Reaktor.

Markierungsmessungen: Neue hoch sensitive Messgeräte, insbesondere Massenspektrometer ermöglichen die Messung der intrazellulären Markierungsanreicherung. Insbesondere können die Markierungen einiger Metabolite des Zentralstoffwechsel gemessen werden. Der Zentralstoffwechsel ist ein durchflossenes System, der Anteil unmarkierter Substanzen wird ausgewaschen. Somit kommt es nicht zu der negativen Auswirkung der Markierungsverdünnung, vor allem zu späten Zeitpunkten der Fermentation.

Unter diesen Vorgaben wurde eine Versuchsplanung für Experimente mit einem dem L-Phenylalanin-Produzenten ähnlichen Stamm durchgeführt. Untersucht werden sollen wiederum drei Phasen: Späte Wachstumsphase, frühe und späte Produktionsphase. Ziel ist

eine bestmögliche Identifizierung der anaplerotischen Flüsse `ana1` sowie des Abflusses in den oxidativen Pentose-Phosphat-Weg `ppp1`.

Die Optimierung sucht nach dem bestmöglichen Substrat-Mix aus natürlicher, $1\text{-}^{13}\text{C}$ und $\text{U-}^{13}\text{C}_6$ (vollmarkierter) Glucose⁶. Da mit der Puls-Zugabe maximal 50% markierte Glucose zugegeben werden soll (s.o.), wird in der Versuchsplanung die Nebenbindungen $p \geq 50\%$ (p ist der Anteil unmarkierter Glucose) gesetzt⁷.

Bei der Untersuchung der Wachstumsphase kann auf NMR-Messungen der proteinogenen Aminosäuren zurückgegriffen werden. In den nachfolgenden Experimenten in der Produktionsphase kann nur noch die Anreicherung im Produkt `3,4-CHD` gemessen werden. Im Gegensatz zu dem beschriebenen Experiment soll die Anreicherung zusätzlich intrazellulär gemessen werden.

Daraus ergeben sich folgende Substratzusammensetzungen:

- Für das erste Experiment wird eine Substratzusammensetzung 50:17:33 (unmarkiert: $1\text{-}^{13}\text{C}:\text{U-}^{13}\text{C}_6$) vorgeschlagen. Im Vergleich zu den ausgewerteten Experimenten können mit dieser Zusammensetzung und den zusätzlichen Markierungsmessungen die Austauschflüsse der anaplerotischen Reaktionen zumindest in ihrer Größenordnung bestimmt werden. Auch die Austauschflüsse im Pentose-Phosphat-Weg können in ihren Größenordnungen abgeschätzt werden. Die restlichen Flüsse des Netzwerks sind im Vergleich besser bestimmbar.
- Der Substrat-Mix für die Phasen B und C wird mit (unmarkiert: $1\text{-}^{13}\text{C}:\text{U-}^{13}\text{C}_6$)=50:21:29 berechnet. Durch die Erhöhung des Markierungsanteils wird eine bessere Fluss-schätzung ermöglicht.

Inzwischen liegen die Messdaten der drei Experimente vor und Karsten Grönke (Fermentationsgruppe, IBT-2) wertet die Experimente aus. Neben den NMR-Daten sind zudem mittels LC-MS/MS einige intrazelluläre Markierungsanreicherungen gemessen worden (GRÖNKE, 2007).

⁶Untersucht wurden weitere verfügbare Markierungsmuster, aufgrund der extrem hohen Kosten beschränkt sich die Darstellung hier auf $1\text{-}^{13}\text{C}$ und $\text{U-}^{13}\text{C}_6$ Glucose

⁷Durch diese Vorgabe werden gleichzeitig auch die Kosten des Markierungsexperiments niedrig gehalten.

Kapitel 11

Methoden zur Interpretation stationärer Stoffflusslagen

Aus den Stoffflusskarten in Abb. 10.6 lassen sich Tendenzen ablesen. Im Verlauf der Produktionsphase nimmt z.B. der Fluss in den oxidativen Pentosephosphat-Weg ab, während der Fluss in den Zitratzyklus zunimmt und die Produktion immer weiter zurückgeht. Welche Flüsse des Netzwerks wirken am stärksten auf die Absenkung der Produktausbeute? Diese klassische Frage lässt sich für stationäre Netzwerkmodelle bislang nur unzureichend beantworten, da nur wenige Methoden zur Interpretation gemessener Stoffflusslagen vorliegen.

DELGADO UND LIAO (1997) stellen eine sensitivitätenbasierte Methode vor, die “Inverse Flux Analysis” genannt wird, vor. Dabei werden interessante intrazelluläre Flüsse des Netzwerks gewählt und deren Einfluss auf die Wachstums- oder Produktionsrate bestimmt. Die Gleichungen für diesen Ansatz wurden in Kapitel 8.2, Seite 73 in etwas anderem Zusammenhang dargestellt. Durch diesen Ansatz ergeben sich allerdings Nachteile, die mit der Wahl der Flüsse zusammenhängen. Im folgenden Beispiel wird erläutert, wieso die Wahl der freien Flüsse Einfluss auf die Sensitivitäten nimmt.

Einen anderen Weg beschreiten SCHILLING ET AL. (2000). Sie betrachten die stöchiometrischen Gleichungen im Zusammenhang mit einem Optimierungskriterium. Sie stellen sich die Frage, inwieweit eine Metabolitkonzentration das Optimierungsziel beeinflusst, indem sie die Sensitivität der rechten Seite der Bilanzgleichung auf die Lösung des Optimierungsproblems bestimmen. Diesen Ansatz kann man als den Einfluss eines zusätzlichen, “virtuellen” Flusses in den betrachteten Pool verstehen (SCHILLING ET AL., 2000). Mit dieser Methode wird der Einfluss einer Bilanzgleichung quantifiziert. Zur Interpretation einer gemessenen Stoffflusslage kann sie jedoch nicht herangezogen werden.

Um die Lücke bei der Interpretation zu schließen, werden in diesem Kapitel zwei neue Konzepte vorgestellt. Sie greifen Methoden der beiden beschriebenen Ansätze auf und erweitern so die Methoden der stationären Stoffflussanalyse und metabolischen Netzwerkanalyse (WAHL ET AL., 2006b). Sie verwenden sowohl Optimierungskriterien als auch Sensitivitäten um die gemessene Stoffflusslage in den möglichen Lösungsraum des Systems einzuordnen und in diesem Kontext zu interpretieren. Vorgeschlagen wird die Berechnung von “Limitierungs-Potenzialen” (LP) und der “Sensitivitäten der beschränkenden Flüsse” (CLS - Constraint Limitation Sensitivities) (WAHL ET AL., 2006b), die anhand eines ein-

fachen Beispiels erläutert und dann zur Untersuchung der drei gemessenen Stoffflusskarten aus Kapitel 10 (Abb. 10.6) eingesetzt werden.

11.1 Sensitivitätsanalyse stationärer Netzwerke

Zur Berechnung stationärer Netzwerksensitivitäten schlagen DELGADO UND LIAO (1997) eine auf partiellen Ableitungen basierte Methode vor. Sie leitet sich direkt aus den Grundlagen in Kapitel 8.2 ab. Um das Verhalten dieser Sensitivitäten zu erläutern, muss ein Netzwerk mit mindestens drei Freiheitsgraden betrachtet werden. Aus diesem Grund wird das Beispielnetzwerk (Abb. 6.1) um einen Abfluss aus B erweitert (Abb. 11.1). Betrachtet werden soll die Sensitivität des Produktflusses v_6 gegenüber dem intrazellulären Fluss v_2 . Das Gleichungssystem wird nach *freien* und *abhängigen* Flüssen umgestellt. Um die Sensitivität des Produktflusses v_p (hier v_6) gegenüber dem intrazellulären Fluss v_i (hier v_2) zu erhalten, muss v_i im Vektor der *freien*, v_p im Vektor der *abhängigen* enthalten sein. Im laufenden Beispiel werden drei *freie* Flüsse zum vollbestimmten Gleichungssystem benötigt.

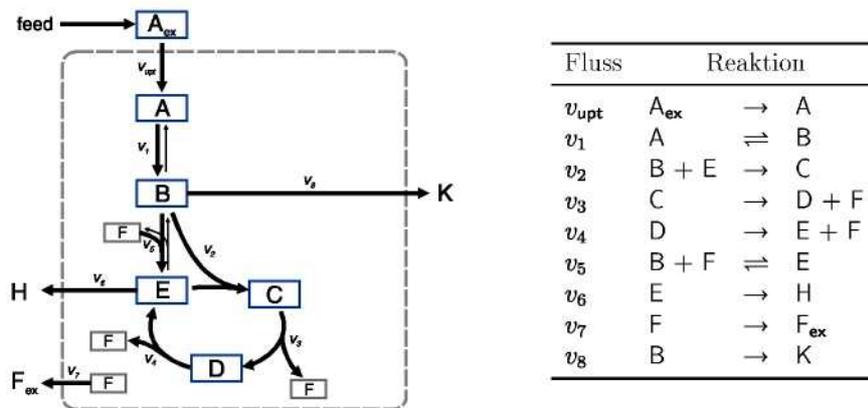


Abbildung 11.1: Beispielnetzwerk zur Sensitivitäts-Analyse. Das gewünschte Produkt ist H, das aus dem Vorläufer E gebildet wird. Metabolit F ist mit CO_2 eines realen Netzwerks vergleichbar. Er ist intern verfügbar, überflüssiges F kann über die Reaktion v_7 in das Medium abgegeben werden. Die einzelnen Reaktionen sind in der Tabelle rechts dargestellt, reversible Reaktionen sind mit Doppelpfeilen gekennzeichnet.

Die Wahl, welche *freien* Flüsse gewählt werden, ist dem Modellierer überlassen. Meist wird die Substrat-Aufnahmerate normiert ($v_{\text{upt}} = 1$), womit ein Freiheitsgrad festgelegt wird. Im Beispiel bleiben dann noch zwei Freiheitsgrade. Betrachtet werden hier die beiden

Möglichkeiten $[v_{\text{upt}}, v_2, v_7]^T$ und $[v_{\text{upt}}, v_2, v_8]^T$. Die Gleichung des abhängigen Flusses v_6 (eine Zeile in $\mathbf{N}_d^{-1} \mathbf{N}_f \mathbf{v}_f$, Gl. (8.6)) lautet dann:

$$v_6 = 2 v_2 - v_7 \quad \text{bzw.} \quad v_6 = v_{\text{upt}} - v_2 - v_8 \quad (11.1)$$

Die Sensitivität lautet somit:

$$\begin{aligned} \mathbf{v}_f &= [v_{\text{upt}}, v_2, v_7]^T & \frac{\partial v_6}{\partial v_2} &= 2 & \frac{\partial v_6}{\partial v_7} &= -1 \\ \mathbf{v}_f &= [v_{\text{upt}}, v_2, v_8]^T & \frac{\partial v_6}{\partial v_2} &= -1 & \frac{\partial v_6}{\partial v_8} &= -1 \end{aligned} \quad (11.2)$$

Diese Sensitivitäten zeigen etwas zunächst Unerwartetes. Je nach Wahl der freien Flüsse \mathbf{v}_f treten unterschiedliche Sensitivitäten (hier für v_2) auf. Dies verdeutlicht, was hinter der Wahl der freien Flüsse steht. Sie legt die Koordinaten für die Lösung des Gleichungssystems fest. Die abhängigen Flüsse ergeben sich aus bestimmten Kombinationen der (gewählten) freien Flüsse, der Einfluss eines bestimmten freien Flusses (z.B. v_2) in der Linearkombination kann sich in einem anderen Koordinatensystem verändern.

Die partielle Ableitung $\partial \mathbf{v}_{d,i} / \partial \mathbf{v}_{f,j}$ gibt die Sensitivitäten des Flusses $\mathbf{v}_{d,i}$ bezüglich $\mathbf{v}_{f,j}$ unter der Annahme, dass die restlichen freien Flüsse $v_{f,k}$, $k \neq j$ konstant bleiben, an. Wird das Koordinatensystem gewechselt, ändern sich die Koordinaten $v_{f,k}$, $k \neq j$. Damit ergibt sich für die Berechnung des Flusses $\mathbf{v}_{d,i}$ eine andere Linearkombination und somit können sich auch die Ableitung verändern. Zusammenfassend zeigt sich, dass diese Methode zur Interpretation der gemessenen Stoffflusslage nur bedingt geeignet ist, da die Abhängigkeit von der Wahl der freien Flüsse zu einigen Nachteilen führt. Insbesondere gilt:

- Es können nur bestimmte freie Flüsse \mathbf{v}_f gewählt werden, da nicht jede Wahl aus \mathbf{v} das Unabhängigkeitspostulat erfüllt.
- Die Sensitivität ist unter der Annahme definiert, dass alle Flüsse $(\mathbf{v}_f)_k$, $k \neq j$ konstant bleiben.
- Gleichzeitig kann nur die Sensitivität von maximal der Anzahl der *freien* Flüsse verglichen werden.
- Die Sensitivitäten aus Analysen mit unterschiedlich gewählten freien Flüssen können nicht verglichen werden, da sie aus unterschiedlichen Koordinatensystemen stammen.

Aus diesen Beobachtungen leitet sich die Forderungen nach einer von den Koordinaten unabhängigen Methode zur Interpretation der Stoffflusslage ab. Aus der Metabolic Network Analysis (vgl. Kapitel 8.3) sind verschiedene Methoden zur theoretischen Analyse von Netzwerkmodellen bekannt. Sie basieren zum Großteil auf Optimierungskriterien, die

ohne die Wahl freier Flüsse auskommen. Diese Methoden werden im Folgenden zur Interpretation der gemessenen Stoffflusslage eingesetzt. Dabei wird die Idee verfolgt, den untersuchten Fluss im Zusammenhang des möglichen Lösungsraum zu analysieren.

11.2 Limitierungs-Potenzial: Reduzierung des theoretischen Maximums

Die Berechnung der theoretisch maximalen Ausbeute kann als lineares Optimierungsproblem unter Nebenbedingungen formuliert werden (MAHADEVAN UND SCHILLING, 2003). Zum einen müssen die Massenbilanzen (Gl. (8.2)) erfüllt sein, zum anderen müssen thermodynamische Grenzen von Flüssen eingehalten werden (Abschnitt 8.3). Bei einigen Reaktionsschritten kann aufgrund sehr hoher ΔG^0 -Werte davon ausgegangen werden, dass sie (fast) ausschliesslich in einer Richtung ablaufen. Diese irreversible Reaktionsschritte fließen als Ungleichungen $\mathbf{v}_{irr} \geq \mathbf{0}$ in die Optimierung ein. Zur leichteren Interpretation der Ergebnisse, wird die Aufnahmerate v_{upt} auf 100% skaliert.

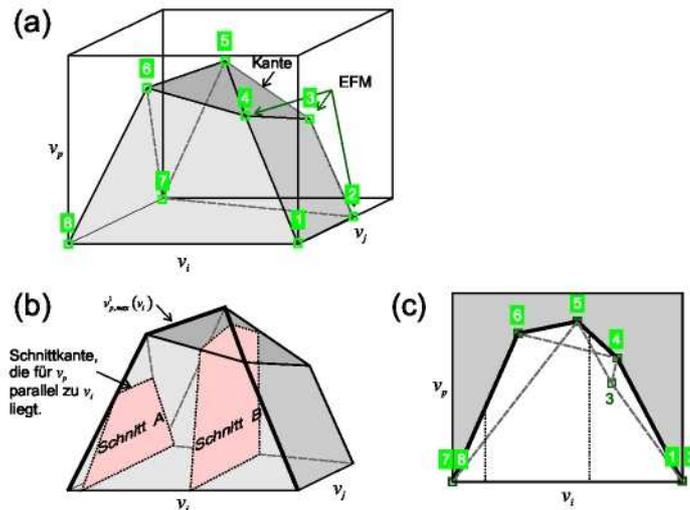


Abbildung 11.2: (a) Skizze des Lösungsraums unter Nebenbedingungen $\mathbf{v}_{irr} \geq \mathbf{0}$ für drei Flüsse v_i , v_j und v_p . Die Kanten und vor allem die Eckpunkte haben besondere Eigenschaften die zur Berechnung ausgenutzt werden können. (b) Die Einschränkung durch die zusätzliche Nebenbedingung $v_i = w_i$ kann als Schnitt durch den Polyeder bei $v_i = w_i$ (Schnitt A und B) veranschaulicht werden. (c) Durch "Aufzeichnen" des jeweils höchsten Punktes entlang des Schnitts entsteht die Funktion ${}^1v_{p,max}$. Sie kann als Projektion der oberen Grenze des Lösungsraumes auf die v_i - v_p Ebene verstanden werden.

Den unter diesen Bedingungen möglichen (skalierten) Lösungsraum der Massenbilanzen kann man sich als Polyeder im Flussraum vorstellen (Abb. 11.2). Ein Fluss v_i kann in-

nerhalb der Grenzen des Lösungsraums variiert werden. Ausserhalb des Raumes können die Massenbilanzen nicht mehr erfüllt werden, ohne gegen eine oder mehrere der Irreversibilitäts-Bedingungen zu verstossen. Die in diesem Raum maximal erreichbare Wachstumsrate oder Produktausbeute v_p kann man als Lösung eines Optimierungsproblems gewinnen:

$$v_{p,max} = \max v_p, \text{ so dass: } \begin{cases} \mathbf{N} \mathbf{v} = \mathbf{0} \\ v_{\text{upt}} = 1 \\ \mathbf{v}_{\text{irr}} \geq \mathbf{0} \end{cases} \quad (11.3)$$

Das Problem kann mit Hilfe der Linearen Programmierung (LP) gelöst werden. Man erhält so den maximal möglichen Fluss $v_{p,max}$ und die zu diesem Maximum gehörige Stoffflussverteilung $\mathbf{v}_{p,max}$.

Die berechnete optimale Stoffflusslage kann nun mit einer gemessenen Stoffflusslage verglichen werden. Um die beobachteten Abweichungen im Zusammenhang zu interpretieren, wird ein bestimmter, gemessener Fluss $v_i = w_i$ betrachtet und in die Optimierung integriert. Grundlage bildet die Frage, wie das Netzwerk sich unter der eingeschränkten Optimalitätsbedingung ausrichtet und inwieweit das Optimum von dem beobachteten Fluss v_i abhängt. Die Lineare Programmieraufgabe erhält dann eine weitere Bedingung, die den Fluss v_i auf den Wert w_i setzt:

$${}^1v_{p,max}(w_i) = \max v_p, \text{ so dass: } \begin{cases} \mathbf{N} \mathbf{v} = \mathbf{0} \\ v_{\text{upt}} = 1 \\ \mathbf{v}_{\text{irr}} \geq \mathbf{0} \\ v_i = w_i \end{cases} \quad (11.4)$$

Die eingeschränkte maximale Produktausbeute ${}^1v_{p,max}(w_i)$ wird nun mit dem theoretischen Maximum ohne die Einschränkung des gemessenen Flusses ${}^0v_{p,max}$ verglichen. Der Quotient $\Phi_p(w_i)$ gibt an, wie stark der gemessene Fluss w_i das erreichbare Maximum verringert:

$$\Phi_p(w_i) = \frac{{}^1v_{p,max}(w_i)}{{}^0v_{p,max}} \quad (11.5)$$

Als Beispiel wird wiederum das Netzwerk aus Abb. 11.1 gewählt. In Abbildung 11.3 ist der maximale Fluss (${}^1v_{p,max}$) über v_i (hier v_2 bzw. v_8) aufgetragen. Für den Abfluss v_8 zeigt sich ein leicht zu interpretierendes Ergebnis. Je höher der Abfluss aus **B** wird, umso niedriger fällt das erreichbare Maximum ${}^1v_{p,max}(v_8)$ aus. Für den Fluss v_2 zeigt sich ein etwas anderes Bild, dass auch die zuvor beobachtete Abhängigkeit von der Wahl der freien Flüsse erklärt (vgl. Abschnitt 11.3). Ist der Fluss v_2 Null, ist der im Kreislauf ($C \rightarrow D \rightarrow E$) inaktiv und das zweite Substrat F der Reaktion v_5 wird nicht mehr gebildet. Somit ist auch $v_5 = 0$ und der Vorläufer E des Produktes H kann weder über v_4 noch v_5 gebildet werden. Steigt der Fluss durch den Kreislauf, erhöht sich die erreichbare Produktausbeute

bis zu einem Wert von ${}^1v_{p,max} = 0.66$ bei $v_2 = 0.33$. Wird v_2 größer 0.33 , sinkt die Produktausbeute wieder ab. Bei $v_2 = 1$ wird der Vorläufer E vollständig im Kreis verbraucht und lediglich in F umgesetzt, es entsteht kein Produkt mehr.

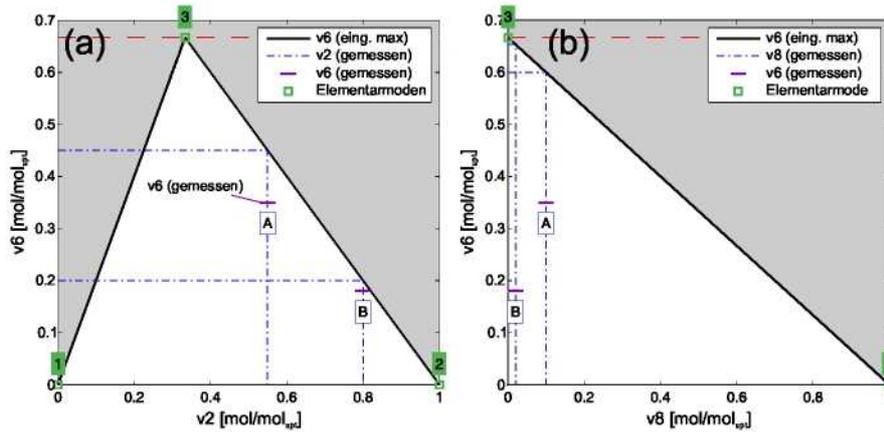


Abbildung 11.3: Darstellung des maximal erreichbaren Produktausbeuten über den Flüssen v_1 (a) und v_6 (b). Die angenommenen gemessenen Flüsse sind mit A und B gekennzeichnet. Die Elementarmoden sind mit den Zahlen 1 – 3 dargestellt.

11.3 Sensitivität des beschränkenden Flusses

Aus Abbildung 11.3 lässt sich neben der Reduzierung des Maximums auch die Sensitivität des erreichbaren Flusses ${}^1v_{p,max}$ ablesen. Sie spiegelt sich in der lokalen Steigung der Geraden ${}^1v_{p,max}$ wieder. Eine Veränderung des einschränkenden Flusses v_i führt zu einer Veränderung der erreichbaren Maximums. Zur Berechnung der Sensitivität

$$\frac{d^1v_{p,max}}{dw}(w_i) \quad (11.6)$$

wird zunächst noch einmal die allgemeine Lösung der stöchiometrischen Massenbilanzen und der Ungleichungen betrachtet (Abb. 11.2). Die Flächen bzw. Kanten des Polyeders geben die Grenzen des Lösungsraumes an. Die Kanten, vor allem aber die Eckpunkte haben dabei besondere Eigenschaften – hier stoßen eine oder mehrere (Eckpunkt) der gegebenen Ungleichungen gleichzeitig an den Grenzwert $\mathbf{v}_{irr} \geq \mathbf{0}$.

Es kann gezeigt werden, dass diese Eckpunkte den möglichen Lösungsraum aufspannen, d.h. jeder andere Punkt des Lösungsraumes als Linearkombination der Eckpunkte dargestellt werden kann (SCHUSTER ET AL., 2000). Zur Bestimmung der Eckpunkte kann die konvexe Hülle des linearen Systems bestimmt werden, ein stärker biologisch motivierter Zugang ergibt sich über die Elementarmodenanalyse (Elementary Flux Modes, EFM)

(SCHUSTER UND HILGETAG, 1994; SCHUSTER ET AL., 1999). Ein Elementarmodus beschreibt eine Stoffflusslage, bei der ein Maximum an Flüssen “ausgeschaltet”, d.h. Null ist (Abschnitt 8.3).

Daraus ergibt sich der direkte Zusammenhang zu den Kanten und Eckpunkten in Abb. 11.2. Sie werden durch die Teilmenge der Elementarmoden mit irreversiblen Flüssen $\mathbf{v}_{irr} = \mathbf{0}$ gebildet (d.h. alle Elementarmoden, bei denen reversible Reaktionen ausgeschaltet werden und nicht gleichzeitig an eine der Grenzen $\mathbf{v}_{irr} = \mathbf{0}$ erreicht wird, werden ausgeblendet). Die Berechnungsalgorithmen für Elementarmoden sind ausgereift und auch für große Netzwerke anwendbar (SCHUSTER UND SCHUSTER, 1993).

Für das Beispielnetzwerk werden 3 Elementarmoden gefunden (Tab. 11.3). Diese drei Moden beschreiben den Weg zu den drei Produkten K (EFM 1), F_{ex} (EFM 2) und H (EFM 3). Damit wird nochmals deutlich, dass jede erreichbare Flusslage durch eine Linearkombination dieser Moden beschrieben werden kann.

Tabelle 11.1: Elementarmoden für das Beispielnetzwerk in Abb. 11.1. Die Aufnahme rate wurde auf $v_{upt} = 1$ skaliert.

| EM | v_{upt} | v_1 | v_2 | v_3 | v_4 | v_5 | v_6 | v_7 | v_8 |
|----|-----------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 2 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 2 | 0 |
| 3 | 1 | 1 | 0.333 | 0.333 | 0.333 | 0.667 | 0.667 | 0 | 0 |

Mit Hilfe der Elementarmoden kann auch die Ableitung $\frac{d^1 v_{p,max}}{d w}(w_i)$ berechnet werden. Das Vorgehen wird anhand von Abbildung 11.2 verdeutlicht. Die Berechnung des erreichbaren Maximums unter der Nebenbedingung $v_i = w_i$ bildet einen Schnitt durch den Polyeder entlang $v_i = w_i$ (Schnitt A und B zeigen dies für unterschiedliche w_i). Das erreichbare Maximum ist der höchste Punkt in v_p entlang dieses Schnitts. Die Funktion $^1 v_{p,max}(v_i)$ kann als Projektion der oberen Grenze des Lösungsraumes auf eine Ebene $v_i - v_p$ verstanden werden.

Aus Abb. 11.2 lässt sich eine weitere Eigenschaft klar erkennen. Eine Kante im Lösungsraum ist eine Linearkombination von zwei Elementarmoden. Die Kanten (und damit auch die Funktion $^1 v_{p,max}(v_i)$) sind zwischen zwei Elementarmoden stetig und linear. Die in Abb. 11.2c dargestellte Linie ist also eine Projektion der jeweils höchsten Kante des Polyeders. Die Elementarmoden (EFM 4,5,6) bilden Knickstellen in der Funktion $^1 v_{p,max}(v_i)$. Aufgrund der Linearität und Stetigkeit zwischen zwei Moden, kann die Sensitivität über die Steigung zwischen zwei Elementarmoden bestimmt werden. Zusammengefasst wird die Sensitivität in folgenden Schritten berechnet:

1. Berechnung aller q Elementarmoden (mit Normierung der Aufnahme rate auf $v_{\text{upt}} = 1$):

$$\text{EFM}_{\mathbf{k}} = [v_{\mathbf{k},1}, \dots, v_{\mathbf{k},i}, \dots, v_{\mathbf{k},p}, \dots, v_{\mathbf{k},r}], \mathbf{k} = 1 \dots q$$

2. Der mögliche Fluss-Raum wird durch die konvexe Hülle aller Elementarmoden $\text{EFM}_{\mathbf{k}}$ umschlossen.
3. Die Elementarmoden werden auf die $v_i - v_p$ Ebene projiziert, alle anderen Flüsse werden ‘ausgeblendet’.
4. Die Elementarmoden mit den höchsten Flüssen v_p werden durch Linien verbunden. Sie entsprechen der Projektion der höchsten Kante, der konvexen Hülle.
5. Um die Sensitivität zu berechnen, werden die $v_i = w_i$ einschliessenden Elementarmoden bestimmt und die Steigung ausgerechnet. Diese Steigung entspricht der Sensitivität des erreichbaren Maximum gegenüber dem Fluss v_i .

In Abbildung 11.2 lässt sich erkennen, dass häufig zwei oder mehr Moden in der Projektion auf einen Punkt fallen. Dies tritt dann auf, wenn der erreichbare maximale Fluss durch mehrere Stoffflusslagen dargestellt werden kann (MAHADEVAN UND SCHILLING, 2003). Mehrdimensional betrachtet entspricht dies einem Schnitt durch den Polyeder, der eine in v_p parallele Kante in einer Richtung $v_{\mathbf{k}}, \mathbf{k} \neq i$ aufweist (Schnitt A in Abb. 11.2. Die Lösung für $d^1 v_{p,\text{max}}/d w(w_i)$ ist jedoch weiterhin eindeutig, da die Steigung zwischen den Elementarmoden paralleler Kanten zwangsläufig gleich ist. Die Berechnung wurde in MatLab unter Nutzung von METATOOL umgesetzt.

11.4 Auswertung der L-Phenylalanin Experimente

Die Methoden werden im folgenden zur Interpretation der drei gemessenen Stoffflusslagen eingesetzt. Im Gegensatz zu dem in den ^{13}C Analysen eingesetzten Modell wird ein stöchiometrisches Netzwerk mit den Co-Metaboliten ATP, ADP, AMP, NADH/NAD⁺ sowie NADPH/NADP⁺ aufgebaut (Abb. 11.4). Die Bilanzierung der Co-Metabolite NAD⁺ und NADP⁺ ist aufgrund der Massenerhaltungen $\text{NADH} + \text{NAD}^+ = \text{konst.}$, sowie $\text{NADPH} + \text{NADP}^+ = \text{konst.}$ redundant und kann somit ausgelassen werden¹.

Zur Regeneration von ATP aus $\text{ADP} + \text{P}_i$ über die oxidative Phosphorylierung (NOGUCHI ET AL., 2004) wird eine Reaktion resp definiert, die pro Proton 1.8 ATP liefert (P/O ratio, Reaktionen sind in Tab. A.1 dargestellt). Um einen möglichen Energieverlust durch ‘futile

¹So genannte Erhaltungsgleichungen (conserved moieties) führen zu stöchiometrischen Matrizen die nicht vollen Rang besitzen.

cycling“ abzubilden wird die Reaktion *fut* eingeführt, ATP wird hier zu ADP verbraucht. Zur Regeneration von AMP wird die Reaktion *e1* ($\text{AMP} + \text{ATP} = 2 \text{ADP}$) in das Modell aufgenommen. In vielen stöchiometrischen Modellen wird die Reaktion zwischen PEP und Pyr reversibel aufgestellt. Die durch die Pyruvat-Kinase und PEP Synthase katalysierten Reaktionen unterscheiden sich jedoch hinsichtlich der benötigten Co-Metabolite. Um PEP aus Pyr zu gewinnen, wird ein ATP zu AMP verbraucht (KESELER ET AL., 2005). Die beiden Richtungen müssen deshalb als getrennte Reaktionen dargestellt werden.

Im letzten Reaktionsschritt zur Synthese von L-Phenylalanin, die Transaminase-Reaktion wird Glutamat zu 2-Oxo-Glutarat umgesetzt. Um das verbrauchte Glutamat wieder zu regenerieren, wird NADPH benötigt. Vereinfacht wird der Verbrauch in den Reaktionsschritt *phe* aufgenommen.

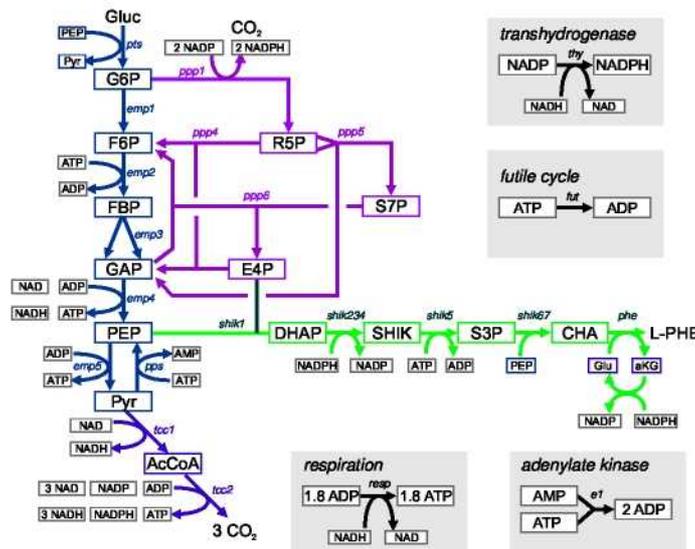


Abbildung 11.4: Stöchiometrisches Modell für die interpretation der Ergebnisse aus den ^{13}C Stoffflussanalysen. Zusätzlich werden auch die Co-Metabolite ATP, ADP, AMP sowie NADH und NADPH modelliert. Die zugehörigen Reaktionsgleichungen sind in Tabelle A.1 dargestellt.

Für dieses stöchiometrische Modell ergeben sich 18 Elementarmoden (Tab. A.1).

11.4.1 Limitierungs-Potentiale intrazellulärer Stoffflüsse

Mit der Methode werden vier interessante Reaktionsschritte untersucht: die Reaktionen *tcc1* (Pdh, Nettofluss in den TCA), *ppp1* (Glucose-6-Phosphat-Dehydrogenase), *ppp4* (Transketolase) sowie *ppp6* (Transaldolase).

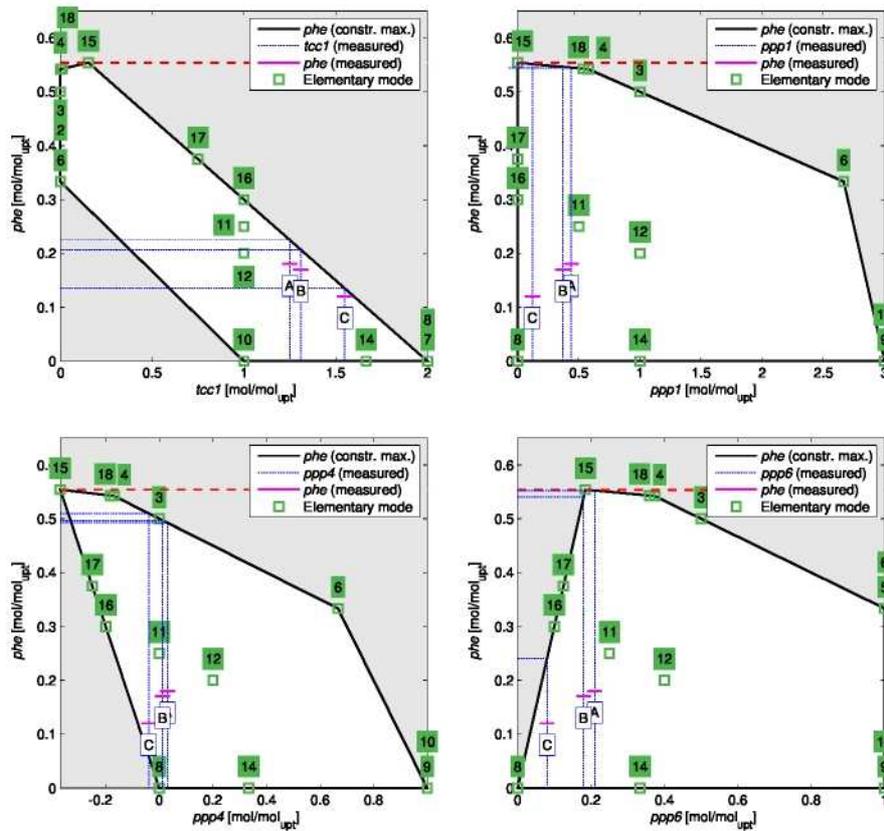


Abbildung 11.5: Projektionen der Funktion $v_{phe,max}^1$ für die Flüsse $tcc1$, $ppp1$, $ppp4$ und $ppp6$. Die gepunkteten Linien zeigen die gemessenen Flussraten (z.B. $tcc1$), die Linie über dem Kasten zeigt die gemessene Produktausbeute phe . Das bei diesem Fluss noch erreichbare Maximum ist mit der horizontal gepunkteten Linie angedeutet.

In Abb. 11.5A ist die Funktion $v_{phe,max}^1(tcc1)$ dargestellt. Der weiße Bereich deutet den Lösungsraum innerhalb der konvexen Hülle des projizierten Polyeders an. Interessanterweise wird für die maximale L-Phenylalanin Ausbeute eine geringe TCA Aktivität ($0.15 \text{ mol/mol}_{upt}$) benötigt um genügend Energie in Form von ATP und NADPH für die Rückreaktion von Pyr zu PEP (PEP Synthase) und den Shikimat-Weg (Shikimat Kinase, Shikimat Dehydratase und die Transaminierung) zu liefern. Steigt der Fluss über $0.15 \text{ mol/mol}_{upt}$, sinkt die Produktausbeute wieder konstant ab. Für die geschätzten Flüsse der Phasen A, B und C ist bereits eine starke Absenkung der erreichbaren Ausbeute zu beobachten, sie wird durch den Abfluss in den TCA von $0.55 \text{ mol/mol}_{upt}$ auf 0.22 (bzw. 0.21 , 0.13) abgesenkt. Damit ergibt sich für $tcc1$ bei den gemessenen Flüssen eine Limitierung von 59% bzw. 63% , 76% für die Phasen A bzw. B und C (Tab. 11.2). Die gemessene phe

Ausbeute liegt mit 0.18 bzw. 0.17, 0.12 bereits nah an der erreichbaren eingeschränkten Ausbeute ${}^1v_{p,max}(tcc1)$. Durch eine optimale Ausrichtung der restlichen Netzwerkflüsse könnte in Phase A lediglich eine Steigerung um ca. 17% erreicht werden.

Tabelle 11.2: Limitierungs-Potenzial der Flüsse *tcc1*, *ppp1*, *ppp4* und *ppp6* auf die Produktausbeute *phe*.

| | tcc1 | ppp1 | ppp4 | ppp6 |
|----------|------|-------|------|-------|
| A | 0.59 | 0.02 | 0.11 | 0.003 |
| B | 0.63 | 0.01 | 0.10 | 0.03 |
| C | 0.76 | 0.005 | 0.08 | 0.57 |

Ein wesentlich anderes Bild ergibt sich für den Fluss *ppp1*. Die gemessenen Flüsse senken das erreichbare Maximum nur sehr geringfügig ab (Tab. 11.2). Im Verlauf der Produktionsphase nähert der Fluss sich dem für die maximale Ausbeute berechneten Fluss $ppp1_{phe,max} = 0$ an und in der letzten Phase des Experiments liegt die Reduzierung durch *ppp1* bei nur noch 0.5%. Die gemessene *phe*-Rate ist weit von dem noch erreichbaren Maximum entfernt. D.h. durch ein optimales Ausrichten anderer Flüsse kann bei jeder der geschätzten *ppp1*-Raten fast die maximale Ausbeute erreicht werden.

Ähnliches zeigt sich für die Flüsse des nicht-oxidativen Pentose-Phosphat-Wegs, *ppp4* und *ppp6*. Bei den gemessenen Flüssen ergibt sich eine geringe Absenkung des erreichbaren Maximums. Eine Ausnahme ergibt sich für *ppp6* in der letzten Phase, hier kommt es zu einer Reduzierung um 57%. Die Ergebnisse für den Pentose-Phosphat-Weg müssen jedoch mit Vorsicht interpretiert werden, da die Standardabweichungen der geschätzten Flüsse sehr hoch ausfallen (Tab. A.3).

In der Literatur finden sich ähnliche Hinweise. In einigen Publikationen wird auf die Bedeutung der Transketolase (*ppp4*) und Transaldolase (*ppp6*) hingewiesen. DRATHS ET AL. (1992) und LI ET AL. (1999) zeigten, dass eine Überexpression von *tktA* (Transketolase A) die Syntheserate aromatischer Aminosäuren erhöht. Die Feststellung, dass vor allem der hohe Abfluss in den TCA zu einer Reduzierung der Ausbeute führt, wird durch Beobachtungen von PATNAIK UND LIAO (1994); PATNAIK ET AL. (1995); BERRY (1996); GOSSET ET AL. (1996); YI ET AL. (2002) und CHANDRAM ET AL. (2004) gestützt. Sie zeigten, dass die gleichzeitige Überexpression von *tktA* und *pps* den Fluss in den Shikimat-Biosyntheseweg steigern.

Sensitivitäten

Die Sensitivität der Reduzierung wird über die in Abschnitt 11.3 beschriebene Methode berechnet. In Tabelle 11.3 sind die Sensitivitäten der untersuchten Flüsse *tcc1*, *ppp1*, *ppp4*

und ppp6 dargestellt. Bei der Variation des Flusses tcc1 über den möglichen Lösungsraum ($0 \leq tcc1 \leq 2$) durchläuft $v_{phe,max}^1(tcc1)$ einige Elementarmoden (4, 18, 15, 17, 16 und 7,8). Für die gemessenen Flüsse tcc1 befindet sich das reduzierte Maximum immer zwischen den Elementarmoden 16 und 7 bzw. 8, die Sensitivität ist -0.3 . Bei ppp6 ergibt sich ein Wechsel von -0.064 zu 3 in Phase C des Experiments.

Tabelle 11.3: Einige Sensitivitäten $\partial v_{p,max}^1 / \partial v_i(w_i)$. Die Werte in den Klammern geben die umschließenden Elementarmoden an.

| | tcc1 | ppp1 | ppp4 | ppp6 |
|---|--|---|--|---|
| A | $-0.3 \begin{bmatrix} 16 \\ 7,8 \end{bmatrix}$ | $-0.021 \begin{bmatrix} 4 \\ 18 \end{bmatrix}$ | $-0.25 \begin{bmatrix} 3,2 \\ 5,6 \end{bmatrix}$ | $-0.064 \begin{bmatrix} 15 \\ 18 \end{bmatrix}$ |
| B | $-0.3 \begin{bmatrix} 16 \\ 7,8 \end{bmatrix}$ | $-0.021 \begin{bmatrix} 15 \\ 18 \end{bmatrix}$ | $-0.25 \begin{bmatrix} 4 \\ 3,2 \end{bmatrix}$ | $-0.064 \begin{bmatrix} 15 \\ 17 \end{bmatrix}$ |
| C | $-0.3 \begin{bmatrix} 16 \\ 7,8 \end{bmatrix}$ | $-0.021 \begin{bmatrix} 15 \\ 18 \end{bmatrix}$ | $-0.25 \begin{bmatrix} 4 \\ 3,2 \end{bmatrix}$ | $3.0 \begin{bmatrix} 7,8 \\ 16 \end{bmatrix}$ |

11.5 Ergebnisse und Diskussion

Mit den beiden vorgestellten Methoden kann der Einfluss der verschiedenen Netzwerkflüsse quantitativ erfasst und im Zusammenhang, z.B. in Bezug auf die Produktausbeute interpretiert werden. Die Methoden geben Aufschluss über die mögliche Limitierung der Produktausbeute durch einen bestimmten (gemessenen) Fluss. Mit der Sensitivität des erreichbaren Maximums bezüglich des untersuchten Flusses kann zudem abgeschätzt werden wie sich eine Flussänderung auf die Produktausbeute ausübt.

Im Gegensatz zu der von DELGADO UND LIAO (1997) definierten Sensitivität wählt nicht der Experimentator die *freien* Flüsse. Sie sind zur Berechnung der Sensitivität natürlich weiterhin notwendig, ergeben sich jedoch implizit aus der Berechnung der beiden Elementarmoden zwischen denen der gemessene Fluss liegt. Aus dieser Argumentation lässt sich ein weiterer Vorteil der vorgestellten Methode ableiten: Neben der Sensitivität, ist auch deren Gültigkeitsbereich durch die einschließenden Elementarmoden festgelegt.

Die Sensitivität des erreichbaren Maximums wird bei Flüssen relevant, die zu einer hohen Limitierung der Produktausbeute beitragen. Liegt die gemessene Produktausbeute nahe des noch möglichen (relativen) Flusses, so sind die stöchiometrische Grenzen des Netzwerks fast erreicht. Nur wenn es gelingt, diesen Fluss zu erhöhen (bzw. erniedrigen), kann die Produktausbeute gesteigert werden. Die Sensitivität der Grenzlinie ist in diesem Fall eine gute Näherung der "wirklichen" Sensitivität (vorausgesetzt, das angenommene Netzwerk spiegelt die Realität ausreichend genau wieder).

Die Methoden wurden in dieser Arbeit auf ein relativ einfaches metabolisches Netzwerk angewendet. Werden zusätzliche Reaktionen eingefügt und bidirektionale Flüsse angenom-

men, so wird der Nutzen der vorgestellten Methoden vermutlich noch deutlicher hervortreten. Denkbar ist es zudem, zwei gemessene Flüsse gekoppelt zu betrachten und deren gemeinsame Limitierung zu bestimmen. Definiert wird hierzu das Maß ${}^2v_{p,max}(w_i, w_j)$. Der Lösungsraum wird dann in zwei Richtungen "fixiert" und das Maximum in diesem Raum berechnet.

Die Limitierungs-Potenziale zeigen, wie die Stoffflusslage verändert werden muss. Sie geben somit auch Hinweise auf die Stammentwicklung: Es zeigt sich, dass eine TCA-Aktivität von **15% mol/mol_{upt}** benötigt wird um den Biosyntheseweg ausreichend mit Cometaboliten und Reduktions-Äquivalenten zu versorgen. Der oxidative Pentosephosphatweg sollte ausgeschaltet sein.

Dieses zunächst irritierende Ergebnis lässt sich aufgrund der bessern Ausbeuten der Regenerierung von Reduktionsäquivalenten und Energieträgern erklären. Viel Energie wird für die PEP Synthase ($2 \times 0.85 \text{ ATP mol/mol}_{\text{upt}}$) sowie die Shikimat-Kinase ($0.55 \text{ ATP mol/mol}_{\text{upt}}$) Reaktion. Zudem werden $2 \times 0.55 \text{ NADPH mol/mol}_{\text{upt}}$ für die Shikimat-Dehydrogenase sowie die Reaminierung von **aKG** zu **Glu** benötigt. In der Bilanz kommen auf den oxidativen PPP $2 \times \text{NADPH mol/mol}_{\text{CO}_2}$. Im TCA-Zyklus (ab Pyruvat) werden $4 \times \text{NADH}$, 1 NADPH sowie 1 ATP unter dem Verlust von 3 CO_2 regeneriert. Unter der Annahme einer ausreichenden Trans"-hy"-dro"-gen"-ase-Aktivität stehen somit **5 NADPH** und **1 ATP** zur Verfügung. Die erreichte Effektivität von **5.9 NADPH/3CO₂** (unter Annahme eines P/O-Verhältnisses von 1.8, d.h. **1.8 ATP** durch **1 NADH** regeneriert werden können) liegt zunächst unter der des oxidativen PPP. Durch den Abfluss in den oxidativen PPP nimmt allerdings der Fluss durch die Glykolyse, u.a. auch der **ATP** und **NADH** generierenden Reaktion **emp₄** von **1.26** (EM 15) auf **1.08** ab. Diese Differenz von **0.18** entspricht für die o.g. Ausbeute eine Steigerung um $0.18 + 0.18 \cdot 0.9 = 0.342$, d.h. über diese Mode werden **6.242 NADPH/3CO₂** gebildet. Die Effektivität, Cometabolite über den TCA (sowie untere Glykolyse) zu regenerieren liegt somit etwas über der Effektivität des oxidativen PPP.

Um einen zu hohen Abfluss in den TCA zu verhindern könnte die Pyruvat Dehydrogenase chromosomal deletiert werden. Abflüsse über die anaplerotischen Reaktionen werden durch Deletion der PEP Carboxykinase (*pck*) verhindert. Die weiteren anaplerotischen Enzyme (*ppc*, *sfcA* und *maeB*) katalysieren Reaktionen, die unter normalen Bedingungen in Richtung PEP bzw. Pyr ablaufen. Um die notwendige Aktivität in der Produktionsphase sicherzustellen, wird Acetat zugeführt (Abb. 11.6). Die TCA-Aktivität erhöht sich von 15% auf 31%, da etwas mehr L-Phe gebildet wird und die PEP Synthaseaktivität verstärkt wird (von 85% auf 100%).

Während der Wachstumsphase muss der TCA zudem mit C4-Körpern aufgefüllt werden (OAA und aKG werden als Vorstufen für einige Aminosäuren benötigt). Hier bietet sich

Glutamat an. Es wird in der Gegenwart von Glucose aufgenommen und ist zudem Aminogruppen Donor der Transaminase aromatischer Aminosäuren (AAT). In der Produktionsphase könnte aus 1 mol Glc und 0.31 mol Ac 0.6 mol L-Phenylalanin gewonnen werden (Abb. 11.6).

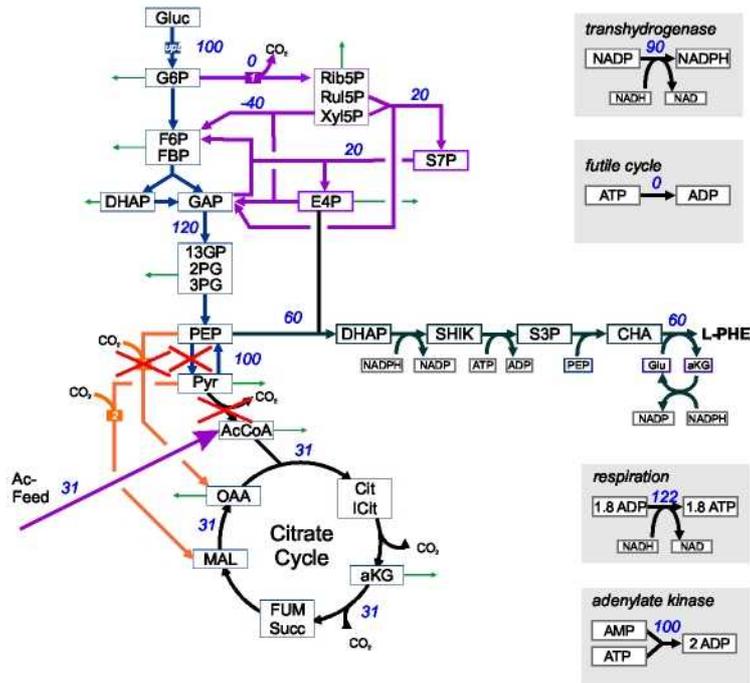


Abbildung 11.6: Stoffflussverteilung eines idealen L-Phenylalanin Produzenten. Die TCA Aktivität wird durch die Zugabe von Acetat "gesteuert", während der Wachstumsphase wird Glutamat zum Auffüllen zugegeben.

Teil III

Kinetische Modellierung und
Auswertung

Kapitel 12

Auswertungsstrategie für Stimulus-Response Experimente

12.1 Auswertungsstrategie

Im Kapitel zu den Grundlagen der Modellierung (Kapitel 5) wurde gezeigt, dass der einzige Zugang zu *in vivo* Enzymkinetiken in Experimenten mit intakten Zellen liegt. Geschildert wurde ebenfalls, dass die zur Modellerstellung notwendige Informationen über mögliche Wechselwirkungen, insbesondere allosterische Inhibierung und Aktivierung von Enzymen *in vitro*, teilweise fehlen. Desweiteren können, aufgrund der unterschiedlichen experimentellen Bedingungen *in vivo* und *in vitro*, *in vitro* gefundene Enzym-Mechanismen nicht ohne weiteres auf die lebende Zelle übertragen werden.

Um diese Hürden zu nehmen, wird eine Strategie zur Auswertung von Stimulus-Response-Experimenten (SRE) in mehreren Schritten vorgeschlagen (Abb. 12.1) (WAHL ET AL., 2006a).

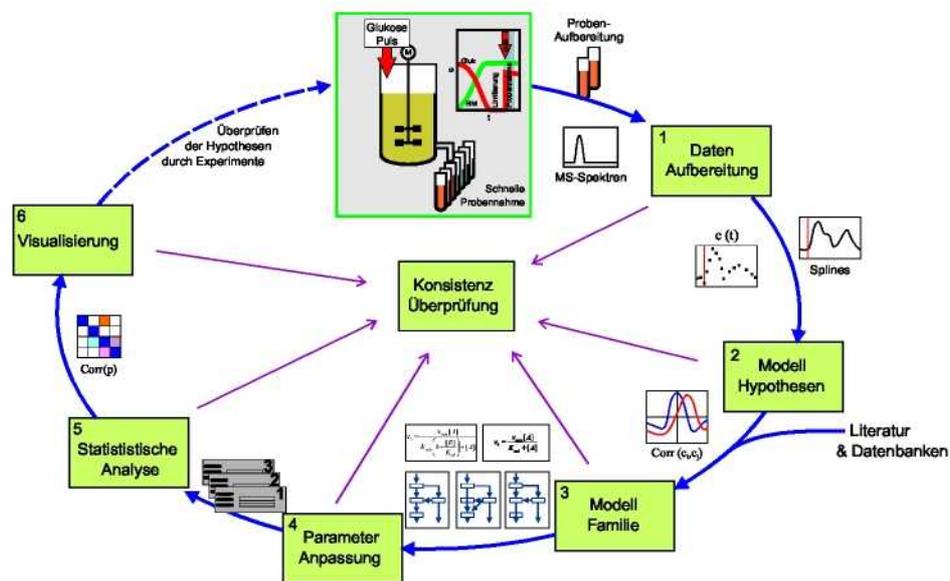


Abbildung 12.1: Strategie zur Auswertung von Stimulus-Response-Experimenten.

- 1. Aufarbeiten der Messdaten:** Um aus den gemessenen MS-Spektren $A_i(t)$ die intrazellulären Konzentrationen $c_i(t)$ zu erhalten, müssen die Aufarbeitungsschritte

(BUCHHOLZ ET AL., 2001) und die Besonderheiten der MS/MS-Messung beachtet werden. Neben den Konzentrationen sind für die Modellierung auch die Standardabweichungen $\sigma_i(t)$ der Messungen von Bedeutung.

Das fokussierte Modell ist ein “offenes” System, das mit dem Rest des Stoffwechsels verbunden werden muss. Aufgrund des starken Messrauschens können die gemessenen Konzentrationen nicht direkt zur Beschreibung des Zustands an den Systemgrenzen eingesetzt werden. Das Rauschen würde zu starken Störungen der modellierten Metabolit-Pools führen. Glättende Splines (DIERCKX, 1975) können flexibel an die Daten angepasst werden. Je nach Rauschanteil wird manuell eine geeignete Glättung gewählt ¹. Für die vorliegenden Messdaten und die gegebene Modellstruktur (Differentialgleichungssystem) haben sich stückweise definierte, stetige Splines der allgemeinen Gestalt

$$\bar{c}(t) = \sum_{i=0}^N \alpha_{j,i} t^i \text{ für } t_j \leq t < t_{j+1} \quad (12.1)$$

bewährt. Neben der Glättung liefern sie auch die zur Simulation notwendige Interpolation. Der eingesetzte Differentialgleichungslöser (Isoda) kommt mit den aus den Splines berechneten Interpolationen gut zurecht.

2. **Modell-Hypothesen:** Auch für kleinere Modelle (wie die im folgenden diskutierten fokussierten Modelle, Abschnitt 12.4) können nicht alle theoretisch möglichen Wechselwirkungen von Enzymen mit Metaboliten berücksichtigt werden. Zunächst werden alle in der Literatur beschriebenen *in vitro* bekannten Enzym Mechanismen gesammelt. Weitere Hypothesen über mögliche Interaktionen können aus einer Korrelations-Analyse der Messdaten abgeleitet werden. Die Idee, Modellhypothesen aus Korrelationen abzuleiten, begründet sich aus der Gestalt der Kinetiken. Steigt beispielsweise die Substratkonzentration, erhöht sich die Reaktionsgeschwindigkeit der enzymatischen Umsetzung und die Produktkonzentration wird steigen. In diesem Fall wird man eine positive Korrelation von Substrat und Produkt beobachten. Ähnlich verhält es sich für Inhibitoren oder Aktivatoren. Steigt eine Inhibitorkonzentration, wird das Enzym stärker gehemmt und es wird weniger Produkt gebildet. Es ergibt sich eine negative Korrelation zwischen Inhibitor und Produkt. Bei Aktivatoren wird man wiederum eine positive Korrelation beobachten. Steigt die aktivierende Wirkung durch eine gestiegene Aktivatorkonzentration, wird mehr Produkt gebildet. Aktivator und Produkt verhalten sich also mit gleichem Vorzeichen.

¹Die Spline Glättung kann auch automatisch, z.B. mit der “Risk estimation” Methode oder dem Verfahren der Cross-Validation (LEE, 2003) parametrisiert werden.

Die Reaktion auf eine Veränderung einer Substrat- oder Inhibitorkonzentration tritt unter Umständen zeitlich versetzt auf, da eine Veränderung des Flusses nicht zu einer sofort merklichen Konzentrationsänderung führt (v.a. für größere Pools). Die Messdaten werden deshalb mit Hilfe der “Time Lagged Correlation-Analysis” auch auf zeitlich versetzte Korrelationen untersucht (ARKIN ET AL., 1997). Auch hierbei bietet es sich an, die Messdaten mit Splines zu glätten. Sie sind zum einen stark verrauscht und zudem liegen sie nicht nicht-äquidistant vor. Jeder geglättete Konzentrationsverlauf $\tilde{c}_i(t)$ wird mit einem zeitlich versetzten Verlauf $\tilde{c}_j(t + \tau)$ korreliert:

$$\rho_{i,j}(\tau) = \frac{s_{i,j}(\tau)}{\sqrt{s_{i,i}(\tau) s_{j,j}(\tau)}} \quad s_{i,j}(\tau) = \langle [\tilde{c}_i(t) - \bar{c}_i] [\tilde{c}_j(t + \tau) - \bar{c}_j] \rangle \quad (12.2)$$

Auffallend hohe Korrelationen deuten auf Wechselwirkungen im Reaktionsnetzwerk (Abschnitt 13.1) hin.

3. **Erstellen einer Modellfamilie:** Aus den Hypothesen der einzelnen Reaktionsschritte ergibt sich eine Vielfalt von möglichen Modell-Kombinationen. Sind für die q Reaktionsschritte je n_k Varianten möglich, ergeben sich $n = n_1 n_2 \dots n_q$ mögliche Kombinationen. Bei größeren Netzwerken kann es daher leicht zu einer enormen kombinatorischen Vielfalt kommen. Einige Kombinationen können zum Glück ausgeschlossen werden. Z.B. können stöchiometrisch nicht sinnvolle Modelle anhand einer Elementarmodenanalyse ausgeschlossen werden (HAUNSCHILD ET AL., 2004). Die Modellierungsumgebung MMT2 unterstützt den Aufbau der Modellfamilie durch die Möglichkeit, für jeden Reaktionsschritt kinetische Varianten zu definieren (Erweiterung des SBML Dialekts (HUCKA ET AL., 2003)). Zusätzlich können vom Benutzer nicht sinnvolle Kombinationen ausgeschlossen werden.
4. **Parameteranpassung:** Das Anpassen einer Modellfamilie stellt ein diskret kontinuierliches Optimierungs-Problem dar. Zum einen muss die richtige Modellstruktur (diskreter Teil), zum anderen die jeweiligen Modell-Parameter bestimmt werden (kontinuierlicher Teil). Liegt eine sehr große Modellfamilie vor, sollten Beziehungen zwischen den Mitgliedern einer Modellfamilie definiert werden. In vielen Fällen, können Kinetiken durch Hinzufügen bzw. Eliminieren einzelner Parametern ineinander überführt werden. Modelle, die sich nur geringfügig unterscheiden, sollten im Suchraum dicht zusammenliegen. In MMT2 wurde für die Anordnung ein gerichteter Abhängigkeitsbaum implementiert (HAUNSCHILD ET AL., 2006). Die Optimierung im diskret-kontinuierlichen Suchraum kann mit dieser Methode gezielt gesteuert werden. Wird nach einer gegebenen Zeit ein hoffnungsvoller Modellkandidat gefunden, sollte die kontinuierliche Parameteranpassung verstärkt an Mo-

dellen in der Nachbarschaft durchgeführt werden. Zur Parameteranpassung stand der Rubenscluster der Universität Siegen mit 256 Opteron Prozessoren, sowie der Hochleistungs-Rechencluster JUMP zur Verfügung.

5. **Modell-Ranking:** Aus der Parameteranpassung werden üblicherweise mehrere Modellvarianten mit einer akzeptablen Anpassung hervorgehen. Dies Kandidaten werden nun anhand eines Kriteriums *Crit*, z.B. dem Akaike-Informationskriterium (BURNHAM UND ANDERSON, 2002), sortiert. Diese, die Daten gut reproduzierenden Varianten der Modellfamilie, sollten bestimmte charakteristische Eigenschaften teilen, die für die Reproduktion der Daten notwendig sind. Die Überschneidungen geben somit Hinweise auf durch die Daten gut bestimmte Mechanismen.
6. **Statistische Untersuchung:** MMT2 liefert neben den simulierten Konzentrationsverläufen und den Flüssen auch alle Sensitivitäten $dc/d\alpha$ (MAUCH ET AL., 1997) der Konzentration bezüglich der Parameter (HAUNSCHILD, 2006) (vgl. auch Abschnitt 8.6.2).

Anhand der Sensitivitäten können statistische Untersuchungen durchgeführt werden. Vor allem kann die Bestimmtheit der Parameterschätzung berechnet werden (Gl. (7.17)). Die Parameterkovarianz gibt Aufschluss über die Zusammenhänge der Parameter (Korrelationen) und den erreichten Vertrauensbereich durch die gegebenen Messdaten (Kapitel 7).

Im folgenden werden die einzelnen Schritte von der Messdatenaufbereitung bis zur statistischen Untersuchung genauer beleuchtet.

12.2 Messdaten Aufarbeitung

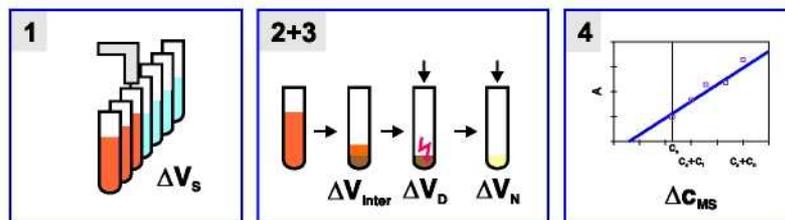


Abbildung 12.2: Schritte der Probenbearbeitung und mögliche Fehlerquellen. 1: Schwankungen des Probennahme-Volumens ΔV_S , 2+3: Fehler beim: Interzell-Volumen V_{inter} , Resuspensions-Volumen V_R , dem zugegebenen Volumen zum Zellaufschluss V_D und dem Volumen zur Neutralisation V_N . 4: Fehler bei der MS/MS Messung.

Um quantitative intrazelluläre Konzentrationen aus den MS/MS-Messungen zu erhalten, müssen alle Schritte der Probenaufbearbeitung und die Besonderheiten der MS-Messtechnik beachtet werden (Abb. 12.2). In die Berechnung der intrazellulären Konzentrationen gehen folgende Größen ein:

Probennahme-Volumen V_S : Das Probevolumen V_S ist mit zufälligen Fehlern durch Druckschwankungen und durch unterschiedliche Verweilzeiten der Probennahmefläße unter dem Ventil behaftet.

Interzell-Volumen V_{inter} : Nach dem Zentrifugieren verbleibt eine kleine Menge des Überstands zwischen den Zellen (Interzell-Volumen).

Resuspensions- V_R , Zellaufschluss- V_D und Neutralisations-Volumen V_N : Hier treten zufällige Pipettierfehler auf ($\Delta V_R, \Delta V_D, \Delta V_N$).

LC-MS/MS Messung c_P : Im Gegensatz zu HPLC-Messungen muss der Einfluss der Probenmatrix, d.h. aller in der Probe vorkommenden Substanzen, beachtet werden. Die Ionisierung (ESI, Abschnitt 4.5) der Analyten vor dem Eintritt in das Vakuum hängt nicht nur von der Konzentration des zu messenden Analyten sondern auch den zum selben Zeitpunkt eluierenden Substanzen ab. Dieser Effekt wird "Ion suppression" genannt (BONFIGLIO ET AL., 1999). Zur Quantifizierung wird deshalb die Methode der Standard-Additionsreihe eingesetzt (ARDREY, 2003). Der Probe mit der Konzentration c_P werden bekannte Konzentrationen $c_{add,k}$, $k = 1 \dots n$ des Analyten zugegeben. Da die jeweiligen Additionen $c_{add,1}, \dots, c_{add,n}$ bekannt sind, kann nun auch die Konzentration der Probe c_P bestimmt werden. Die Berechnung kann aus einer Darstellung im Koordinatensystem Signalstärke A über der Konzentration c erfolgen (Abb. 12.3). Zu beachten ist, dass durch die Standard-Addition die Konzentration c_P der Probe verdünnt wird (in den hier durchgeführten Experimenten ergibt sich eine Verdünnung auf $2/3 c_P$). Da jeder Probe das selbe Volumen $V_{add,i}$ zugegeben wird, ergeben sich auf der x-Achse die Konzentrationen $2/3 c_P + c_{add,i}$. In diesem Koordinatensystem wird deutlich, dass angenommen werden muss, dass die Regression durch $0, 0$ geht ².

Die Steigung γ der Regressionsgeraden gibt das Verhältnis der Peakfläche A zur Konzentration c wieder.

²Dieser Annahme steht entgegen, dass bei der MS-Messungen eine gewisse Konzentration notwendig ist um ein Signal zu erhalten. Im extrem niedrigen Konzentrationsbereich kommt es zudem zu nicht-Linearitäten.

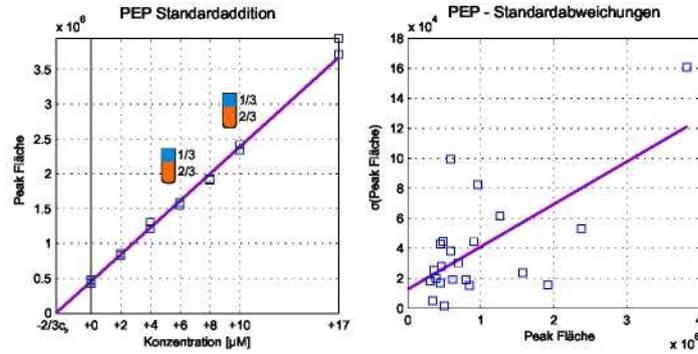


Abbildung 12.3: Rechts: Darstellung der Regression für die Standard-Additionsreihe am Beispiel PEP. Auf der x -Achse wird die Konzentration der Probe aufgetragen, auf der y -Achse die gemessenen Peakflächen. Links: MS/MS Messfehler über der mittleren Peakfläche. Dargestellt ist die Standardabweichung σ_A über der mittleren Peakfläche \bar{A} von Mehrfachmessungen. Die Regressionsgerade gibt ein Fehlermodell unter der Annahme eines Rauschens aus Grundrauschen und signalabhängigem Rauschen wieder.

Um die Fehler $\sigma(c)$ der MS-Messung abzuschätzen, wird ein lineares Fehlermodell angenommen. Der Fehler setzt sich aus einem Grundrauschen σ_g und einem von der Signalintensität A abhängigen relativen Fehler $\sigma_r = a A$ zusammen (Abb. 12.3):

$$\sigma_{MS/MS} = \sigma_g + a A \quad (12.3)$$

Der Datensatz enthält Doppel-Messungen der Standard-Additionsreihe und Doppelbestimmungen der ersten 16 Proben nach dem Puls. Auf Basis dieser Daten wurde das in Gl. (12.3) dargestellte Fehlermodell parametrisiert. Aufgrund der wenigen Messdaten ist eine deutliche Streuung vorhanden (Abb. 12.3b). Der Trend, dass der absolute Messfehler (σ_A) mit der Signalstärke A zunimmt, ist dennoch zu erkennen und findet sich auch für alle anderen gemessenen Analyten wieder.

Mit diesen Größen kann die intrazelluläre Konzentration aus der Peakfläche A , dem spezifischen intrazellulären Volumen (ν), dem Probenvolumen (V_S) und den Zugaben (V_N Neutralisation, V_{inter} Interzellvolumen, V_E Extraktionsvolumen) berechnet werden (BUCHHÖLZ ET AL., 2001):

$$c_{\text{intra}} = \frac{c_{\text{BTM}} V_S \nu + V_E + V_N + V_X}{c_{\text{BTM}} V_S \nu} \frac{A}{\gamma} \quad (12.4)$$

Bis auf die Messung der Peakfläche A können die Schritte der Probenaufarbeitung für die Probe nur einmal durchgeführt werden. Nach der linearen Fehlerfortpflanzung ergibt sich aus den einzelnen Fehlern σ der Größen V_E , V_N , V_{inter} , V_S , ν , γ und A (DEGENRING, 2004):

$$\text{Var}(c_{\text{intra}}) = \mathbf{J} \text{Cov}(V_E, V_N, V_{\text{inter}}, V_S, \nu, \gamma, A) \mathbf{J}^T \quad \mathbf{J} = \frac{\partial c_{\text{intra}}}{\partial (V_E, V_N, V_{\text{inter}}, V_S, \nu, \gamma, A)^T} \quad (12.5)$$

Mit diesen Schritten kann jeder Messung der entsprechende Messfehler zugeordnet werden.

12.3 Eingesetzte Spline-Glättung

Aus der Messdaten Aufarbeitung sind die Konzentrationen $c(t_i)$ und Fehler $\sigma_c(t_i)$ der Zeitpunkte t_i , $i = 1 \dots N$ bekannt. Für die Splineglättung werden in der ersten und zweiten Ableitung stetige kubische B-Splines³

$$\tilde{c}(t) = \sum_{j=1}^{n-4} \alpha_j B_j(t) \quad (12.6)$$

eingesetzt ($B_j(t)$ bezeichnet den normierten kubischen B-Spline mit den Stützstellen $\lambda_j, \lambda_{j+1}, \dots, \lambda_{j+4}$). Der Spline soll so durch die Daten gelegt werden, dass keine schnellen Veränderungen stattfinden und gleichzeitig die Messwerte so gut wie möglich genähert werden. Die Koeffizienten $\alpha_1, \dots, \alpha_{n-4}$ der inneren Knotenpunkte werden aus einer Minimierung der Veränderungsgeschwindigkeiten (3. Ableitung)

$$\nu = \sum_{j=5}^{n-4} \delta_j^2 \quad (12.7)$$

unter der Nebenbedingung einer maximalen Abweichung von den N Messungen

$$\theta = \sum_{i=1}^N \varepsilon_i^2 \leq S, \quad \text{mit: } \varepsilon_i = \left(\frac{\tilde{c}(t_i) - c(t_i)}{\sigma_c(t_i)} \right)^2 \quad (12.8)$$

gewonnen (DIERCKX, 1975). Der Wert ν bildet die Summe der Sprünge δ_j in der dritten Ableitung von $\tilde{c}(t)$ an der Stützstelle λ_j . Der Wert θ gibt die Summe der gewichteten Abweichungen wieder.

Wählt man für S einen großen Wert, wird eine große Abweichung zwischen den gemessenen und den geglätteten Konzentrationen zugelassen, und daher die Minimierung von ν kaum eingeschränkt wird. Wählt man eine geringe Glättung S , ergibt sich eine "harte" Nebenbedingung. Um den Messdaten folgen zu können, müssen stärkere Sprünge in der dritten Ableitung hingenommen werden.

Zur Berechnung der B-Spline-Koeffizienten wurde die Funktion `e02bec` aus der NAG-C Library (The Numerical Algorithms Group Ltd, Oxford, UK) verwendet. In MatLab wurde eine Benutzeroberfläche und die Umrechnung in stückweise Polynome hinzugefügt.

³B-Splines sind stückweise definierte Polynome (Gl. (12.1)) in einer parameter-reduzierten Form.

12.4 Fokussierte Modellierung

Mit der fokussierten Modellbildung sollen die Zusammenhänge in einem speziellen Biosyntheseweg detailliert untersucht werden, während der Rest des Netzwerks so weit wie möglich ausgeblendet wird. Der Biosyntheseweg wird in diesem Ansatz aus dem Netzwerk "ausgeschnitten", an den Systemgrenzen dieses Sub-Netzwerks werden die Verbindungen zum Zentralstoffwechsel durch den Einsatz von Splines hergestellt (Abb. 12.4). Die Splines geben den Verlauf der außerhalb des Biosynthesewegs liegenden Metabolite wieder, und das Modell beschreibt den Konzentrationsverlauf in Abhängigkeit der Parameter- und Splinewerte.

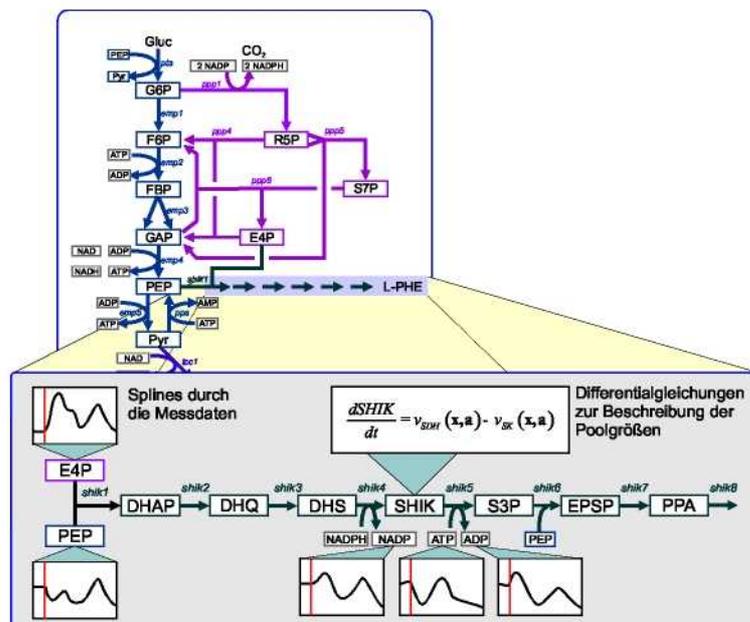


Abbildung 12.4: Fokussierte Modellbildung: Der Biosyntheseweg wird aus dem Netzwerk "ausgeschnitten". Der Konzentrationsverlauf von Metaboliten außerhalb des Biosynthesewegs wird durch Splines beschrieben und wird in den Reaktionskinetiken berücksichtigt.

Das Sub-Netzwerk stellt ein offenes System dar, vor allem sind die Flüsse über die Systemgrenzen nicht messbar. Der Fluss in und aus dem System ergibt sich aus dem kinetischen Ansatz, der Eingangs- und der Ausgangsreaktion (shik1 und shik7).

12.5 Modellfamilien

Unter einer Modellfamilie versteht man im hier gebrauchten Zusammenhang die m kinetischen Modelle, die sich in bestimmten Merkmalen, teilweise in der Stöchiometrie (\mathbf{N}^i), vor allem aber in der metabolischen Regulation, insbesondere den verwendeten Enzymmechanismen (\mathbf{v}^i) unterscheiden:

$$\dot{\mathbf{c}}^i = \mathbf{N}^i \mathbf{v}^i(\mathbf{c}^i, \boldsymbol{\alpha}^i) \quad i = 0 \dots m \quad (12.9)$$

Mit dem Modellierungswerkzeug MMT2 werden für jeden Reaktionsschritt mögliche Varianten formuliert, die Software generiert daraus alle möglichen Kombinationen und fasst sie in einer Modell-Familie zusammen. Stöchiometrische Varianten lassen sich als Spezialfall einer kinetischen Variante, der Null-Kinetik formulieren. Mit der Null-Kinetik wird eine Reaktion ausgeschaltet; So können in einer Modellfamilie Reaktionswege detailliert durch mehrere Reaktionsschritte oder vereinfacht als ein Reaktionsschritt aufgestellt werden.

12.6 Parameteranpassung

Meist kann ein Zusammenhang zwischen den verschiedenen Varianten einer Reaktionskinetik hergestellt werden. So kann z.B. die irreversible Michaelis-Menten-Kinetik als Vereinfachung der reversiblen Variante angesehen werden ($K_{eq} \rightarrow \text{inf}$). Diese Zusammenhänge ermöglichen, systematische Abhängigkeiten der Modelle in einer Familie herzustellen. Insbesondere kann eine gerichtete Beziehung von zwei Modellen i und j hergestellt werden ($i \prec j$). Diese Beziehungen kann für die Parameteranpassung großer Modellfamilien genutzt werden. Wurde für das Modell i bereits eine gute Anpassung gefunden, können einige Parameter des Modells i in j übernommen werden. Durch die gleichzeitige Anpassung verschiedener Modelle (diskret) und der Parameter (kontinuierlich) entsteht ein diskret kontinuierliches Optimierungsproblem.

MMT2 verfügt über einen an diese Aufgabe angepassten Algorithmus, der auf einem Cluster-Computer ausgeführt werden kann. Der Algorithmus steuert die Optimierung der verschiedenen Modelle (diskret). In der Nachbarschaft guter Modellkandidaten wird die Rechenleistung verstärkt auf die Parameteranpassung dieser Modelle gerichtet. Durch diesen Algorithmus wird die Parameteranpassung großer Modellfamilien wesentlich beschleunigt, offensichtlich nicht relevante Modelle werden nicht weiter angepasst (HAUNSCHILD, 2006; HAUNSCHILD ET AL., 2006).

12.7 Statistische Untersuchung

Die statistische Untersuchung der Modellparameter basiert auf den in Kapitel 7 beschriebenen Methoden der nicht-linearen Regressionsanalyse. Für die gemessenen Konzentrationen $\mathbf{c}(t)$ wird für jeden Zeitpunkt t_i eine Regressionsgleichung aufgestellt (Gl. (8.23)) und die Fehlerfortpflanzung anhand der Sensitivitäten (Gl. (8.21)) bestimmt. Auf diesem Weg können nun die Parameterkovarianz $\mathbf{Cov}(\hat{\boldsymbol{\alpha}})$ (Gl. (7.17)) und Parameterkorrelation $\mathbf{Corr}(\hat{\boldsymbol{\alpha}})$ bestimmt werden.

Zur genaueren Analyse der Zusammenhänge können Standardmethoden der multivariaten Statistik eingesetzt werden. Besonders hilfreich ist dabei die Hauptkomponentenanalyse (PEARSON, 1901).

12.8 Modell-Ranking

Aus der Parameteranpassung geht eine Vielzahl von möglichen Modellkandidaten hervor. Es ist nun Aufgabe des Modellierers, ein Sortierkriterium \mathbf{Crit} für die Anpassungsgüte aufzustellen. In der statistischen Analyse hat sich das Akaike Informationskriterium als ein brauchbares Maß erwiesen (BURNHAM UND ANDERSON, 2002):

$$AIC = 2k + n \frac{SSQ}{n} \quad (12.10)$$

Aus allen in der Modellfamilie enthaltenen Modellen ($\mathbf{Mod}^{\mathbf{F}}$), werden nun nur noch diejenigen betrachtet, die ein gesetztes Maß $\mathbf{Crit} \leq T$ erfüllen. Die Modelle $\mathbf{Mod}^{\mathbf{S}}$ sind eine Teilmenge aller Modelle:

$$\mathbf{Mod}^{\mathbf{S}} \subset \mathbf{Mod}^{\mathbf{F}} \forall \mathbf{Crit}(\mathbf{Mod}) \leq T \quad (12.11)$$

Jetzt werden gemeinsame Eigenschaften dieser Modelle gesucht, zunächst mit dem Schwerpunkt, strukturelle bzw. mechanistische Überschneidungen zu finden. Bei der Strukturellen Analyse muß die Parameterbestimmtheit nicht in Betracht gezogen werden. In späteren Analysen, z.B. wenn Aussagen über Substratsättigung, Gleichgewichtslagen, Inhibitoraffinitäten, etc. abgeleitet werden sollen, wird die Parameterbestimmtheit in die Analyse einbezogen.

Kapitel 13

Dynamische Modelle zur Auswertung von Stimulus-Response Experimenten

In diesem Kapitel werden fokussierte Modelle für den Shikimat-Biosyntheseweg anhand der Daten aus dem Experiment mit *E. coli* 4pF78 vorgestellt. Die Modellerstellung und Auswertung folgt der im vorangegangenen Kapitel 12 diskutierten Herangehensweise. Die Modelle basieren auf:

1. Mechanistischen Kinetiken, die in der Literatur beschrieben werden (Tabelle B.1)
2. Phänomenologischen Ansätzen (LinLog und generische Michaelis Menten Kinetiken), die aus den beobachteten Korrelationen erstellt werden.

13.1 Zeitlich versetzte Korrelationsanalyse der Messdaten

Die zeitlich versetzten Korrelationen geben Hinweise auf die Zusammenhänge und regulatorischen Wechselwirkungen des metabolischen Netzwerks. Sie dienen somit zum Einen zur Überprüfung der Messdaten, zum Anderen helfen sie, Modellhypothesen abzuleiten. Um die zeitlich versetzten Korrelationen zu berechnen, müssen die Konzentrations-Zeitverläufe zunächst geglättet werden. Hier wird mit gering geglätteten Splines gearbeitet um die Dynamik für die Korrelationsanalyse auszunutzen. In Abb. 13.1 sind die Korrelationen der Metabolite aus dem Shikimat-Biosyntheseweg dargestellt. Gezeigt werden die Korrelationen mit den beiden Eingangssubstraten PEP und E4P (bzw. dem gemessenen Konzentrationsverlauf des Summensignals X5P aus Xyl5P, Rib5P und Rub5P), dem Produkt der Eingangssubstrate, den Energiemetaboliten ATP, ADP und AMP sowie dem Reduktions-Äquivalent NADP⁺.

Die ersten Metabolite DAH(P), DHS und DHQ des Reaktionsweges korrelieren mit dem Verlauf von X5P (0.73-0.69). Die Korrelation dieser Metabolite mit dem Profil von PEP ist noch niedriger (<0.5). Konsistent mit dem erwarteten Verhalten ist die höhere Korrelation mit dem Produkt X5P*PEP (DAHP: 0.78, DHQ: 0.75 und DHS: 0.67) die entlang der Reaktionskette abnimmt.

Dieses Verhalten kann auch bei den Korrelationen der Folgemetabolite mit dem ersten Metaboliten DAH(P) beobachtet werden. Für DHQ beträgt die Korrelation noch 0.85, danach 0.70, 0.69 und 0.63 für S3P.

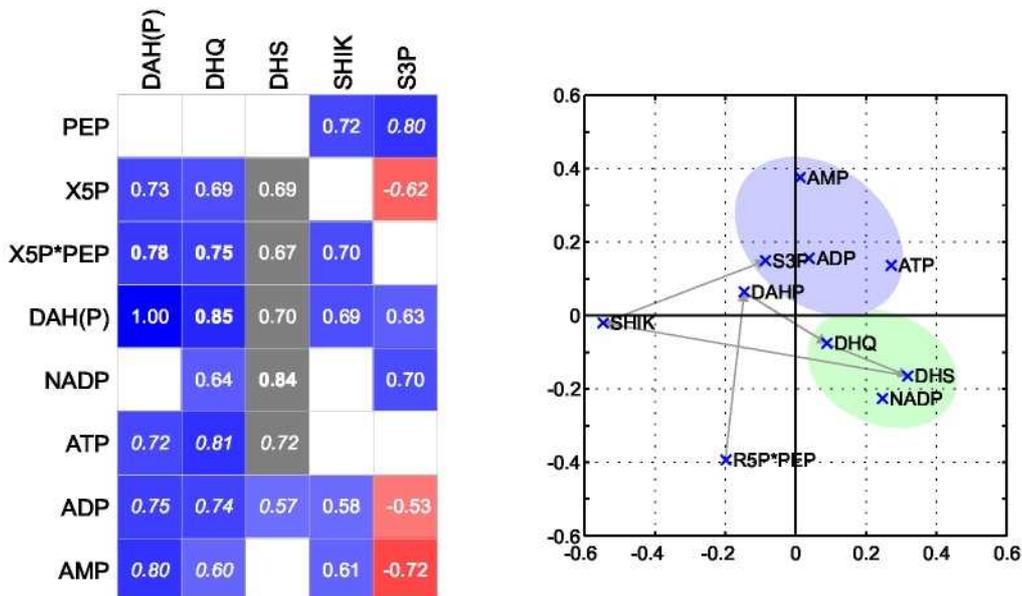


Abbildung 13.1: Links: Maximale Korrelationen der Metabolite des Shikimat Biosynthesewegs mit den metabolischen Vorläufern und Co-Metaboliten. Positive Korrelationen sind blau, negative rot dargestellt. Die Farbsättigung gibt die Größe wieder, Korrelationen $\text{Corr} \leq 0.5$ werden nicht dargestellt. Rechts: Dimensions-reduzierte Darstellung der maximalen Korrelationen nach der Methode von ARKIN ET AL. (1997).

Für SHIK und S3P zeigen sich schwer interpretierbare Korrelationen. SHIK korreliert stark mit dem Produkt X5P*PEP (0.70), stärker als der "näher" gelegenen Metabolite DHS. Auffällig ist auch die Korrelation mit dem PEP Verlauf (0.72). Der Verlauf von S3P korreliert positiv mit PEP (0.80), die Korrelation mit X5P ist dagegen negativ (-0.62). Mit dem Produkt der Eingangssubstrate (X5P*PEP) ist keine Korrelation mehr zu beobachten < 0.5 . Folgende Erklärungsmuster können abgeleitet werden:

DAHP Synthase: DAH(P) zeigt hohe Korrelationen mit dem Produkt der beiden Substrate, sowie mit NADP^+ . Angenommen wird eine irreversible Eingangsreaktion.

DHQ Synthase: DHQ zeigt eine hohe Korrelation mit DAHP und eine geringere mit NAD^+ (0.67, nicht dargestellt). Die Korrelation könnte auf die in der Literatur beschriebene Abhängigkeit von katalytischen Mengen NAD^+ (MOORE ET AL., 1994) zurückgeführt werden. Es werden ein irreversibler, reversibler und NAD^+ abhängiger (LinLog) Mechanismus geprüft.

DHQ Dehydratase: Auch bei dieser Reaktion ist eine hohe Korrelation mit NAD^+ und NADP^+ zu beobachten. Eine Aktivierung durch NAD(P)^+ ist nicht bekannt. NADPH

ist an der folgenden, DHS verbrauchenden Reaktion beteiligt. Der Verlauf von NADPH konnte nicht gemessen werden. Es wird allgemein angenommen, dass die Summen-Konzentration von $\text{NADP}^+ + \text{NADPH}$ in dem relativ kurzen Beobachtungszeitraum konstant bleibt. Die positive Korrelation mit NADP^+ kann somit auch als negative Korrelation mit NADPH angesehen werden. NADPH scheint den Verbrauch von DHS anzutreiben. Für die DHQ Dehydratase werden aus diesem Grund nur eine irreversible und reversible Reaktionsrate geprüft.

Shikimat Dehydrogenase: Das Produkt der Shikimat Dehydrogenase zeigt keine eindeutigen Korrelationen, das Substrat DHS eine relativ hohe, positive Korrelation mit NADP^+ . Das Produkt zeigt zudem eine höhere Korrelation zu PEP als die dichter gelegenen Metabolite DAHP, DHQ und DHS. Unter Umständen deutet dies auf eine Aktivierung der Shikimat Dehydrogenase durch PEP hin. Den Modellen wird ein reversibler sowie ein durch PEP aktivierter Reaktionsmechanismus zugrunde gelegt.

Shikimat Kinase: Das Produkt S3P zeigt kleine, negative Korrelationen mit ATP, ADP und AMP, sowie eine positive Korrelation mit PEP. An der Reaktion ist ATP als Substrat beteiligt, die beobachtete negative Korrelation kann deshalb nicht erklärt werden (Eine höhere Substratkonzentration sollte S3P verstärkt bilden). Die negative Korrelation mit ADP und AMP deutet auf eine Inhibition der Kinase Reaktion durch ADP und AMP hin. Aufgrund der positiven Korrelation mit PEP könnte eine Aktivierung der SK erwägt werden. Möglich ist auch eine Inhibition der S3P verbrauchenden EPSP Synthase. In der Literatur wird eine Substratinhibition durch sehr hohe PEP-Konzentrationen ($K_I \approx 8 \text{ mM}$) (GRUYS ET AL., 1992) beschrieben. Denkbar ist zudem, dass PEP eine indirekte Wirkung über Folgemetabolite (z.B. PPA) ausübt, die inhibierend auf die EPSP Synthase wirken.

EPSP Synthase: Für diesen Reaktionsschritt konnte nur der Verlauf des Substrats gemessen werden. Es wird ein Mechanismus mit und ohne Inhibition durch PEP aufgenommen. In der Literatur wird eine mögliche Inhibition durch ADP (GAERTNER, 1972) genannt. In der Modellfamilie werden Varianten mit und ohne Inhibition durch ADP und PEP untersucht.

Dephosphorylierung von S3P: Eine durch eine Dephosphatase katalysierte Reaktion von S3P zu SHIK wird aufgenommen.

Einige der abgeleiteten Hypothesen finden sich in der Literatur wieder (Kinetiken in Tabelle B.1). Schwierigkeiten ergeben sich vor allem für die beobachteten Korrelationen von SHIK und S3P. Die abgeleiteten Hypothesen für die Shikimat Kinase und EPSP

Synthase können nur teilweise durch Literaturangaben bestätigt werden. Die Hypothese einer Inhibierung der Shikimat Kinase durch Folgemetabolite, z.B. PPA oder PPY, wird in einem "verlängerten" LinLog-Modell umgesetzt (Tab. 13.3) ¹.

13.2 Untersuchte Modell Varianten

Tabelle 13.1: Zusammengefasste Darstellung der mechanistischen Modellfamilie. Reversible Mechanismen sind durch einen Doppelpfeil, irreversible durch eine durchgestrichene Rückreaktion. Inhibitoren sind mit roten Kreisen mit weißem Minuszeichen angedeutet.

| DAHP S PEP + E4P ⇌ DAHP + P _i | DHQ S DAH(P) ⇌ DHQ+P _i | DHQ DH DHQ ⇌ DHS | SHIK DH DHS+NADPH ⇌ SHIK+NADP | SK SHIK+ATP → S3P+ADP | EPSP S S3P+PEP → EPSP+P _i |
|---|---|----------------------------|--|------------------------------------|---|
| 2 Substrate Michaelis Menten ↔ | Michaelis Menten ↔ | Michaelis Menten ↔ | Michaelis Menten ↔ | ⊖ AMP ⊖ AMP ⊖ ADP | Ordered Bi-Bi (DEC E-PEP-PEP) ↔ |
| Ordered Bi-Bi (DEC EP) ↔ | ⊖ DAH(P) Michaelis Menten ↔ | Michaelis Menten ↔ | ⊖ SHIK Michaelis Menten ↔ | ↔ ↔ | Ordered Bi-Bi schnelles GG ↔ |
| 2 | 2 | 2 | 2 | 3 | 2 |

=96 Varianten

In Tabelle 13.1 sind die untersuchten mechanistischen Modellvarianten kurz zusammengefasst. Die Kinetiken sind in Tabelle B.1 aufgeführt. Nicht für jede Reaktion konnte eine für den untersuchten Stamm passende Kinetik gefunden werden – meist sind Enzyme aus anderen Stämmen oder Organismen untersucht worden. Teilweise wird in der Literatur nur ein Mechanismus vorgeschlagen ohne eine kinetische Gleichung anzugeben. Bei genauer untersuchten Enzymen, findet man nicht nur Unterschiede in den Parameterwerten, sondern auch bei den angenommenen Mechanismen (Tab. 13.1). So ergeben sich bis zu drei verschiedene Mechanismen pro Reaktionsschritt, die aufgestellte Modellfamilie enthält insgesamt 96 Varianten. Sind Parameter aus den *in vitro* Experimenten gegeben, werden diese als Startwerte für die Modellanpassung verwendet.

Diesen Kinetiken teilweise sehr ähnlich sind Kinetiken auf Basis von generischen Michaelis-Menten Termen. Für diesen Ansatz wurde ebenfalls eine Modellfamilie aufgestellt (Tab. 13.2). Einige der beobachteten Korrelationen benötigen kinetische Terme mit mehreren Substraten und mehreren Inhibitoren. Diese können zwar mechanistisch aufgestellt werden, die Anzahl an Parametern steigt jedoch stark an, so dass die komplexeren Mechanismen nur

¹Da für die Metabolite EPSP, CHO, PPA und PPY keine Messungen vorliegen, ist die Aussicht auf Erfolg sehr gering. Die besonderen Schwierigkeiten werden in Abschnitt 13.5 diskutiert.

Tabelle 13.2: Reaktionskinetische Varianten für die Modellierung mit dem generischen Michaelis-Menten Ansatz (Tab. B.4). Reversible Mechanismen sind durch einen Doppelpfeil, irreversible durch eine durchgestrichene Rückreaktion. Inhibitoren sind mit roten Kreisen mit weißem Minuszeichen angedeutet.

| DAHP S PEP + E4P \rightleftharpoons DAHP + P _i | DHQ S DAHQ \rightleftharpoons DHQ+P _i | DHQ DH DHQ \rightleftharpoons DHS | SHIK DH DHS+NADPH \rightleftharpoons SHIK+NADP | SK SHIK+ATP \rightarrow S3P+ADP | EPSP S S3P+PEP \rightarrow EPSP+P _i |
|--|--|---|--|--|---|
| ↔ | \rightleftharpoons | \rightleftharpoons | \rightleftharpoons | <div style="display: flex; align-items: center; justify-content: center;"> - AMP ↔ </div> | ↔ |
| <div style="display: flex; align-items: center; justify-content: center;"> - DAH(P) ↔ </div> | <div style="display: flex; align-items: center; justify-content: center;"> - DAH(P) ↔ </div> | ↔ | <div style="display: flex; align-items: center; justify-content: center;"> - SHIK ↔ </div> | <div style="display: flex; align-items: center; justify-content: center;"> - AMP - ADP ↔ </div> | <div style="display: flex; align-items: center; justify-content: center;"> - PEP ↔ </div> |
| 2 | 2 | 2 | 2 | 2 | 2 |

=64 Varianten

in die LinLog-Familie aufgenommen wurden (z.B. SK mit 2 Substraten, 2 Inhibitoren und einem Aktivator).

Zudem wurde das LinLog-Modell um eine Dephosphorylierung von S3P und eine zusammengefasste Reaktion für die Chorismat Muthase, Chorismat Synthase und die Prephenat Dehydratase erweitert. Diese Reaktion wird zum einen NADPH-abhängig und NADPH-unabhängig modelliert (Aktivierung). Eine Zusammenfassung der aufgestellten LinLog-Modellfamilie ist in Tabelle 13.3 zu finden.

Tabelle 13.3: Zusammengefasste Darstellung der LinLog Modellfamilie (die Kinetiken sind in Tab. B.3 aufgeführt). Reversible Mechanismen sind durch einen Doppelpfeil, irreversible durch eine durchgestrichene Rückreaktion. Inhibitoren sind mit roten Kreisen mit weißem Minuszeichen angedeutet. Aktivatoren sind anhand eines grünen Kreises mit Pluszeichen zu erkennen.

| DAHP S PEP + E4P \rightleftharpoons DAHP + P _i | DHQ S DAHQ \rightleftharpoons DHQ+P _i | DHQ DH DHQ \rightleftharpoons DHS | SHIK DH DHS+NADPH \rightleftharpoons SHIK+NADP | SK SHIK+ATP \rightarrow S3P+ADP | EPSP S S3P+PEP \rightarrow EPSP+P _i | CM / PPAS EPSP \rightleftharpoons PPA+P _i | S3P dephos. S3P \rightleftharpoons SHIK+P _i |
|---|--|---|--|---|--|--|---|
| \rightleftharpoons | \rightleftharpoons | \rightleftharpoons | \rightleftharpoons | <div style="display: flex; align-items: center; justify-content: center;"> ↔ - PPx </div> | \rightleftharpoons | \rightleftharpoons | <div style="display: flex; align-items: center; justify-content: center;"> - Nicht vorhanden </div> |
| | <div style="display: flex; align-items: center; justify-content: center;"> \rightleftharpoons + NAD⁺ </div> | | <div style="display: flex; align-items: center; justify-content: center;"> \rightleftharpoons + PEP </div> | <div style="display: flex; align-items: center; justify-content: center;"> - AMP - PPx ↔ - AMP, PPx + PEP </div> | <div style="display: flex; align-items: center; justify-content: center;"> \rightleftharpoons - ADP </div> | <div style="display: flex; align-items: center; justify-content: center;"> \rightleftharpoons + NADPH </div> | <div style="display: flex; align-items: center; justify-content: center;"> ↔ </div> |
| | 2 | | 2 | 4 | 2 | 2 | 2 |

=128 Varianten

Untersucht wurde auch, ob eine Repräsentation im PowerLaw Format gelingt. Die ersten Studien zeigten jedoch sehr deutlich, dass eine Anpassung in diesem Format nicht gelingt. Aufgrund der Struktur von Produkten scheint offensichtlich keine ausreichende Näherung

der benötigten Reaktionsgeschwindigkeiten im auftretenden Konzentrationsbereich möglich. Auf eine Darstellung dieser Modelle wird deshalb verzichtet.

13.3 Auswertung der mechanistischen Modellfamilie

Im Verlauf der Untersuchungen wurden sehr viele Parameteranpassungen auf den Cluster-Computern “Rubens” (Universität Siegen) und JUMP (ZAM, Forschungszentrum Jülich) durchgeführt. Diese Parameteranpassungen bilden die Datenbasis zur Auswertung der Modellfamilie. Jede der möglichen 96 Varianten wurde mindestens 10 mal von unterschiedlichen Startwerten ausgehend optimiert. Die Optimierungsläufe wurden nach 2 Stunden abgebrochen. Bei der verwendeten Modellgröße und -komplexität verbessert sich die Anpassung nach ca. 1 Stunde kaum noch.

Die Anpassungen der besten vier Varianten unterscheidet sich hauptsächlich im Verlauf von SHIK und S3P. Die Mechanismen der Variante 61 (beste Anpassung ist mit * gekennzeichnet), sind in Tab. 13.1) grün hinterlegt. Variante 12 unterscheidet sich in einigen Mechanismen:

| Modell | Reaktionsschritte | | | | | |
|--------|-------------------|------------|------------|------------|------------------|---------------|
| | DAHPS | DHQS | DHSDH | SDH | SK | EPSPS |
| 61 | Bi-Bi DEC | MM, 1S, 1P | MM, 1S, 1P | MM, 2S, 2P | ADP und AMP inh. | Rapid Eq. |
| 12 | Hill | MM, 1S, 1P | MM, 1S, 1P | MM, 2S, 2P | AMP inh. | Ordered Bi-Bi |

Die Häufigkeiten einzelner Reaktions-Varianten sind in Abb. 13.3 auf Grundlage verschiedener Modellmengen dargestellt. Die erste Säule (blau) gibt die Häufigkeiten in der Menge aller Modelle mit einer $FQS \leq FQS_{\min} + 2\%$ wieder. Eine Modell-Variante kann in dieser Menge häufiger auftreten (alle Modell-Varianten werden mindestens 10 mal von verschiedenen Startwerten aus angepasst). Den grünen und roten Säulen liegen die Modell-Varianten mit $FQS \leq FQS_{\min} + 5\%$ bzw. $+10\%$ zugrunde.

Am deutlichsten ausgeprägt sind die Häufigkeiten der SDH. Selbst in der Menge mit schlechteren Modellkandidaten liegt hauptsächlich ein Gleichgewichts-Mechanismus zugrunde (Abb. 13.3). Auch bei der DHQS bleibt eine gewisse Häufigkeit auch in den schlechteren Anpassungen erhalten. Bei der Shikimat-Kinase-Reaktion ist zu beachten, dass sich die beiden Inhibierungs-Mechanismen sehr ähneln (Abb. 14.6), da der Verlauf von ADP und AMP stark korreliert. Die Tendenz zu einem regulierten Mechanismus ist deshalb wesentlich höher, in der Summe ergeben sich 81% (blau). Bei den restlichen Reaktionen gleicht sich die Verteilung aus.

An dieser Stelle wird auf eine weitere Diskussion der Häufigkeiten verzichtet. In den anschließenden Kapiteln wird gezeigt, dass die Modellanpassung an die Daten des Daten-

satzes *E.coli* 4pF78 zwar gelingt, die Parameterbestimmtheit jedoch sehr gering ausfällt. Deshalb wird eine Methode zur Erweiterung der Datenbasis vorgestellt und ein Verfahren zur parallelen Anpassung mehrerer Datensätze eingesetzt. Die Bestimmtheit des Modells verbessert sich durch dieses Vorgehen, die Unterscheidung zwischen den einzelnen Varianten wird dadurch ebenfalls klarer.

13.4 Modelle mit generischen Michaelis-Menten Termen

Abb. 13.4 zeigt die Anpassung der besten fünf Modellvarianten für die generischen Michaelis-Menten basierten Modelle. Eine Darstellung der gefundenen Häufigkeiten ist in Abb. 13.5 zu finden. Hier sind deutliche Parallelen zu den gefundenen Mechanismen der mechanistischen Modellfamilie (Abb. 13.5) zu erkennen.

Im Vergleich zur Modellfamilie auf Basis komplexer mechanistischer Kinetiken verschieben sich die Tendenzen etwas. Da in der mechanistischen Beschreibung die Varianten mit irreversibler DHQ Synthase kaum auftraten, wurde sie nicht in die generische Michaelis-Menten Familie aufgenommen und ein durch ein Substrat inhibierten Mechanismus ersetzt. Dieser setzt sich anscheinend durch, alle in die Analyse einbezogenen Anpassungen besitzen diesen Mechanismus. Deutlich sind auch die Tendenzen für die DHSDH und SDH. Wie in der vorherigen Modellfamilie setzen sich Gleichgewichts-Mechanismen durch. Für die EPSPS wurde zusätzlich eine durch ADP inhibierte Kinetik aufgestellt. Er kann sich allerdings nicht wirklich durchsetzen, lediglich eine kleine Mehrheit (61%) der besseren Anpassungen enthalten diesen Mechanismus.

Für diese Modell-Familie wurde ebenfalls eine Anpassung an die zwei vorhandenen Datensätze durchgeführt (nicht dargestellt).

13.5 LinLog Modellfamilie

Die erweiterte LinLog Modellfamilie enthält 128 verschiedene Varianten für den Shikimat-Biosyntheseweg. Bei der Anpassung (Abb. 13.6) zeigt sich ein zur mechanistischen Modellfamilie ähnliches Bild, die besten Anpassungen unterscheiden sich in ihrer Güte kaum. Erstellt man das Histogramm für die Anpassungen mit maximal 2%, 5% und 10% höherer Fehlerquadratsumme lassen sich wiederum einige Tendenzen erkennen.

Die Mechanismen können nicht direkt miteinander verglichen werden, da einige in der mechanistischen Modellfamilie enthaltene Mechanismen im LinLog-Format nicht dargestellt werden können, oder aber in der gleichen Darstellung münden (z.B. ergibt sich für eine reversible und irreversible Michaelis-Menten Kinetik derselbe LinLog-Ausdruck).

Neu hinzu gekommen sind die Aktivierung der Shikimat Dehydrogenase durch PEP, sowie die Reaktionen der Chorismat-Mutase/Prephenat Dehydratase und eine mögliche Dephosphorylierung von S3P zu Shikimat.

Die Aktivierung durch PEP kommt in guten Modellanpassungen gehäuft vor (80%). Anscheinend können die Messdaten mit diesem Mechanismus besser wiedergegeben werden. In den etwas schlechteren Anpassungen überwiegt der Mechanismus ohne Aktivierung (87%). Diese deutliche Verschiebung der Tendenz kann durch die sehr niedrige Anzahl (5) "sehr guter" Anpassungen ($FQS \leq FQS_{\min} + 2\%$) aus insgesamt über 1000 Parameteranpassungen erklärt werden. Offensichtlich wurde bei diesen Anpassungen ein gutes (lokales) Minimum gefunden. Im "Mittelfeld" befinden sich wesentlich mehr Modelle (102), die den Mechanismus der Aktivierung nicht enthalten.

Auch in der LinLog-Anpassung zeigt sich eine Inhibierung der Shikimat Kinase Reaktion. Eine Inhibierung durch ADP ist implizit durch den reversiblen Mechanismus gegeben, da ADP Produkt der Reaktion ist. Hinzu kommt eine mögliche Inhibierung durch den nachfolgenden Metabolit PPA und eine mögliche Aktivierung durch PEP. In den "sehr guten" Anpassungen überwiegt der alle Regulationen enthaltende Mechanismus (80% mit PEP Aktivierung, Inhibierung durch ADP und AMP). Auch hier zeigt sich im "Mittelfeld" eine andere Häufigkeit. Sie verteilt sich auf die Mechanismen mit Inhibierung durch PPA, Inhibierung durch PPA und AMP und den alle möglichen Regulationen enthaltenden Mechanismus. Diese Beobachtung lässt wieder den Schluss zu, dass für die fünf besten Anpassungen eine besonders gute Anpassung geglückt ist, und das Optimierungsgelände bei der Vielzahl möglicher Einflüsse viele lokale Minima aufweist. In diesem Gelände eine gute Anpassung zu finden gelang in diesem Fall nur in wenigen Optimierungsläufen. Offensichtlich stabiler läuft die Optimierung für die etwas einfacheren Mechanismen. Allerdings wurde für einen Mechanismus ohne Regulation keine einzige Anpassung mit einer Fehlerquadratsumme $FQS \leq FQS_{\min} + 10\%$ gefunden. Diese Beobachtung deutet stark auf eine Regulation der Shikimat Kinase. Welcher Mechanismus vorliegt kann mit den jetztigen Daten nicht geklärt werden.

Obwohl sich für die CM / PPADH deutliche Häufungen zeigen, muss bei der Ableitung von Hypothesen Vorsicht bewahrt werden. Da weder für EPSP noch PPA Messdaten vorliegen, ist der Einfluss der Parameter auf die FQS extrem niedrig und nur indirekt durch die Feedback-Inhibierung auf die Shikimat-Kinase Reaktion und dem Abfluss aus dem Syntheseweg gegeben.

Auch bei der möglichen Dephosphorylierung von S3P zu SHIK zeigen sich Tendenzen, man muss jedoch beachten, dass mit dieser Reaktion zwei parallele Wege zwischen SHIK und S3P entstehen. Sie unterscheiden sich lediglich durch die Co-Metabolite (bei der De-

phosphorylierung entsteht kein ATP). Es gelingt kaum, den Fluss durch den Syntheseweg zu bestimmen. Parallele Reaktionswege aufzulösen, ist ohne weitere Anhaltspunkte, d.h. lediglich aus den Konzentrationsverläufen, nicht möglich. Wieder gilt, dass sich in solch einem Fall offensichtlich der "einfachere" Mechanismus durchsetzt, da das Optimierungsgelände mit weniger Freiheitsgraden sicherer zu einer guten Anpassung führt.

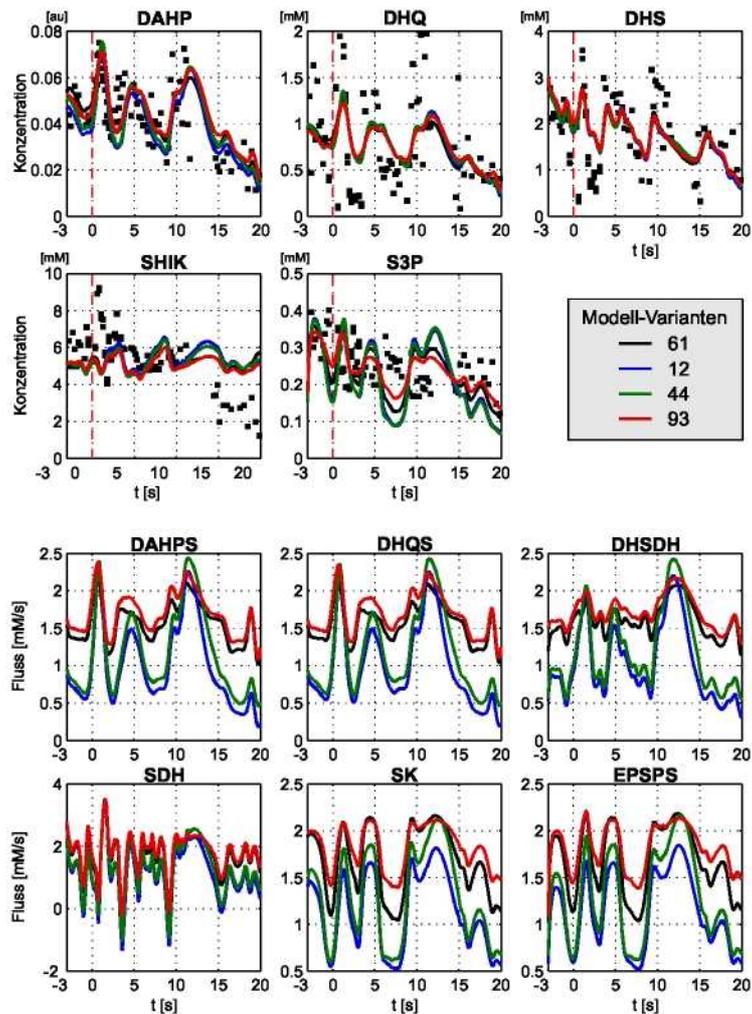


Abbildung 13.2: Plot der Konzentrationen (oben) und Flüsse (unten) der besten Modellanpassungen aus der mechanistischen Modellfamilie (Tab. 13.1). Die FQS der besten vier Anpassungen unterscheidet sich kaum ($\leq 1\%$). Der Fluss durch den Shikimat-Biosyntheseweg wurde in der Optimierung auf $v \leq 2.5$ mM/s beschränkt.

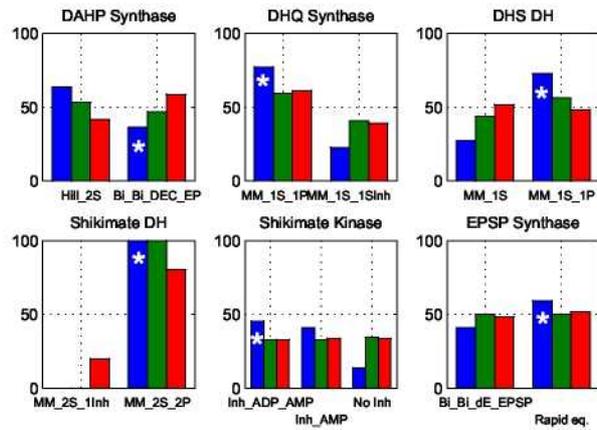


Abbildung 13.3: Häufigkeiten der kinetischen Varianten in der mechanistischen Modellfamilie. In die Betrachtung einbezogen wurde Modelle mit maximal 2% (blau), 5% (grün) und 10% (rot) höherer FQS (im Vergleich zur besten Anpassung *).

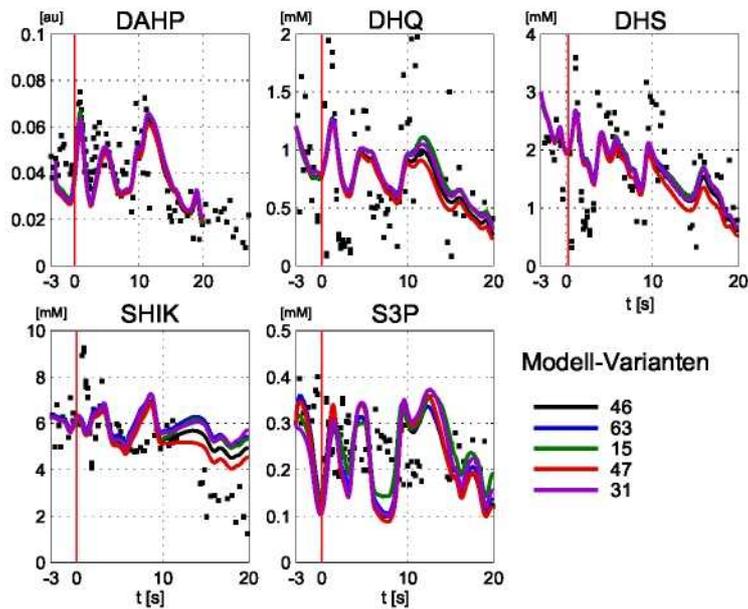


Abbildung 13.4: Anpassung der besten fünf Modellkandidaten mit generischen Michaelis-Menten Kinetiken.

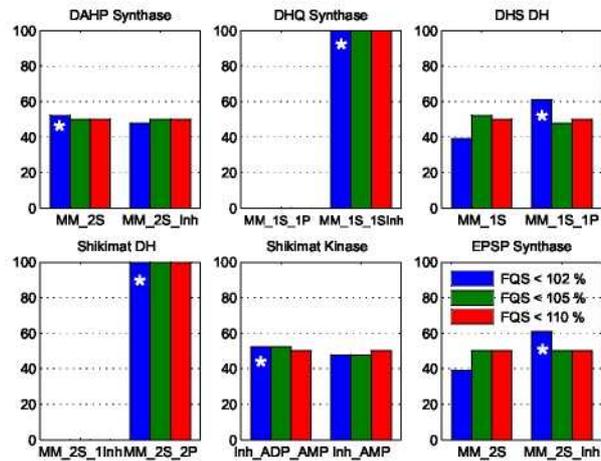


Abbildung 13.5: Histogramm der Häufigkeit bestimmter Mechanismen für die Modellfamilie basierend auf generischen Michaelis-Menten Kinetiken. Der Mechanismus der besten Anpassung ist mit einem Stern gekennzeichnet.

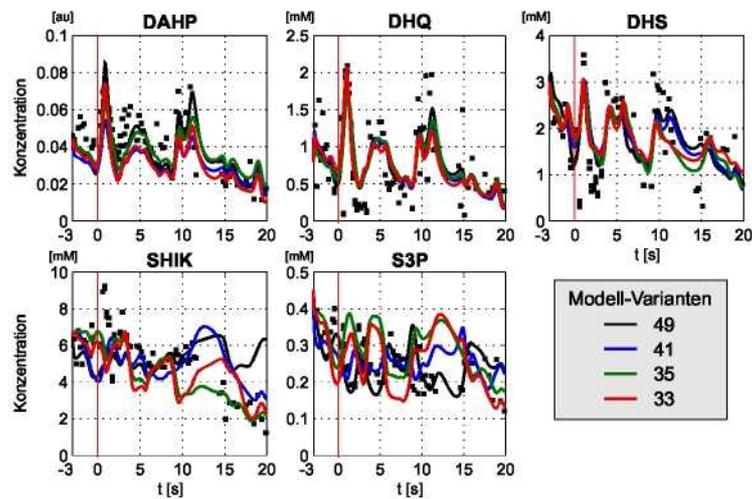


Abbildung 13.6: Anpassung der besten vier Varianten auf Basis von LinLog-Kinetiken.

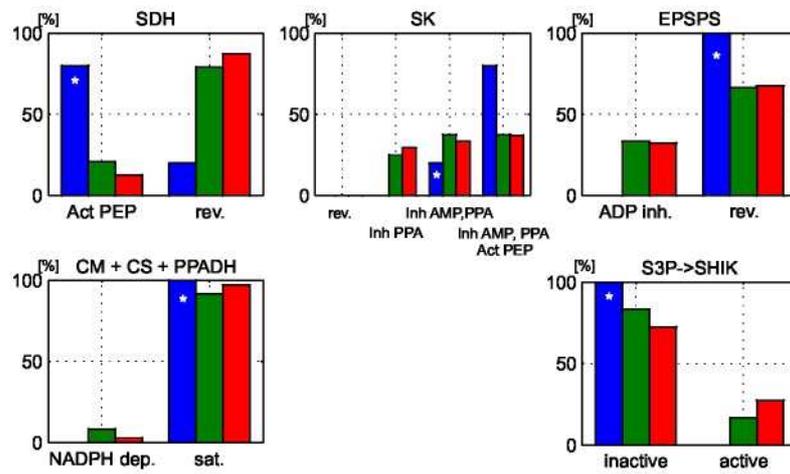


Abbildung 13.7: Häufigkeiten der möglichen Mechanismen der LinLog-Modellfamilie. In die Betrachtung einbezogen werde Modelle mit 2% (blau), 5% (grün) und 10% (rot) höherer FQS.

Kapitel 14

Statistische Untersuchung fokussierter dynamischer Modelle

Statistisch genauer untersucht wird zunächst das mechanistische Modell (beste Anpassung). Daran anschließend wird das LinLog Modell etwas genauer analysiert und auf eine Besonderheit aufgrund des kinetischen Aufbaus aufmerksam gemacht.

Am Ende des Kapitels wird der Einfluss der Eingangssignale, der Splines betrachtet. Da sie aus gemessenen Größen abgeleitet werden, müssen auch sie in die Fehlerbetrachtung einbezogen werden. Exemplarisch wird eine Monte-Carlo Studie eines vereinfachten Eingangssignals untersucht.

Die Ergebnisse der statistischen Auswertung zeigen deutlich, dass die Parameterbestimmtheit mit den zur Zeit vorliegenden Daten auch mit vereinfachten Modellansätzen unzureichend ist. In den anschließenden Kapiteln werden deshalb Möglichkeiten zur Steigerung der Bestimmtheiten diskutiert.

14.1 Parameterbestimmtheit im mechanistischen Modellierungsansatz

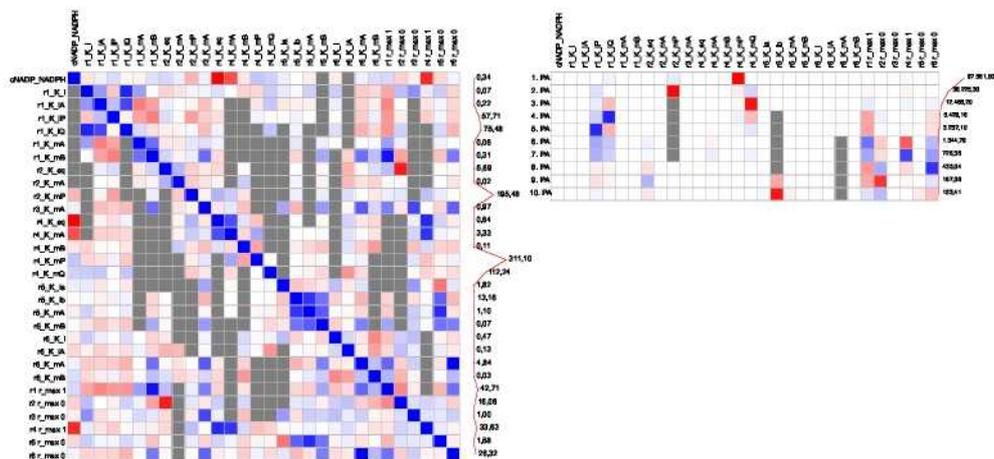


Abbildung 14.1: Parameter-Korrelationsmatrix für die beste mechanistische Anpassung. Rechts sind die ersten Hauptachsen der Kovarianz dargestellt. Der Parameter $SDH_{K_{m,p}}$ erklärt bereits 60% der beobachteten Kovarianz.

Die Parameterbestimmtheit wird nach den Methoden aus Kapitel 7 bestimmt. Die statistische Untersuchung zeigt deutlich, dass nur wenige Parameter bestimmt werden können. Konkret sind lediglich 26% der Parameter mit einer Standardabweichung kleiner 100% bestimmbar.

Um ein klareres Bild der Bestimmtheiten zu bekommen, wird die Kovarianzmatrix mittels Hauptkomponentenanalyse (PCA, Abb. 14.1) untersucht. Auf der ersten Hauptachse fällt der Parameter $^{SDH}K_{m,P}$ auf. Allein dieser Parameter erklärt 60% der Kovarianz. Auf der 2. Hauptachse ($\sim 20\%$) finden sich drei Parameter, die Affinitäten $^{DHQS}K_{m,P}$, $^{DAHPS}K_{i,P}$ und die Aktivität $^{DAHPS}v_{max}$. Ein Teil dieser Parameter beschreibt auch die Richtung der dritten Hauptachse ($\sim 14\%$). Zudem sind sehr hohe Korrelationen zu beobachten. Besonders auffällig sind die Korrelationen $\text{Corr}(c_{\text{NADP+NADPH}}, ^{SDH}K_{eq}) = 0.99$, sowie $\text{Corr}(^{DHQDH}K_{m,A}, ^{DHQDH}v_{max}) = 0.97$.

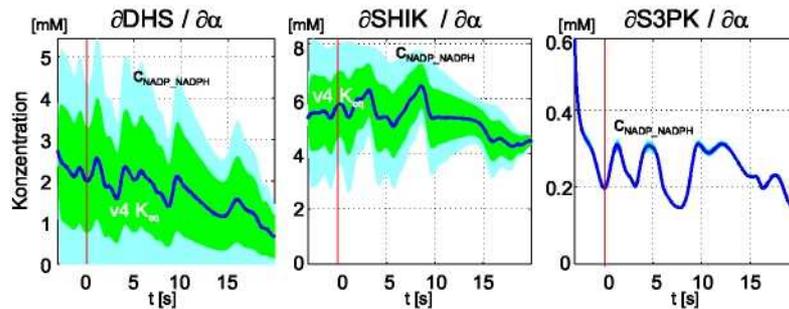


Abbildung 14.2: Sensitivität der Konzentrationen bezüglich der Parameter $c_{\text{NADP+NADPH}}$ und $^{SDH}K_{eq}$ als Schlauch um den simulierten Verlauf. Die Sensitivität der Parameter gleicht sich sehr, die Gestalt des Schlauches ist fast identisch.

Die Korrelation $\text{Corr}(c_{\text{NADP+NADPH}}, ^{SDH}K_{eq})$ kann durch die Betrachtung der SDH Reaktion erklärt werden. Die Gleichgewichtsreaktion $\text{DHS} + \text{NADPH} \rightleftharpoons \text{SHIK} + \text{NADP}^+$ wird durch die Konzentration der Substrate getrieben, hohe Produktkonzentrationen verlangsamen die Reaktion. $^{SDH}K_{eq}$ beschreibt die Lage des Gleichgewichts, $c_{\text{NADP+NADPH}}$ die Gesamtmenge von $\text{NADPH} + \text{NADP}^+$. Wird der Parameter $c_{\text{NADP+NADPH}}$ erhöht, steigt die NADPH-Konzentration ($\text{NADPH} - c_{\text{NADP+NADPH}} - \text{NADP}$, NADP ist durch den Spline gegeben). Damit ist mehr Substrat vorhanden, die Reaktion läuft verstärkt in Richtung der Produkte. Der gleiche Effekt kann auch durch Verändern der Gleichgewichtskonstanten $^{SDH}K_{eq}$ erreicht werden. Wird das Gleichgewicht in Richtung der Produkte verschoben, erhöht sich die Reaktionsgeschwindigkeit. Dieser Effekt kann anhand der Sensitivität der Konzentrationen erkannt werden (Abb. 14.2).

Genauso kann die Korrelation der anderen Parameter erklärt werden. Der Zusammenhang ist auf einer auf den Mittelwert und Varianz normierten Darstellung schnell erkennbar

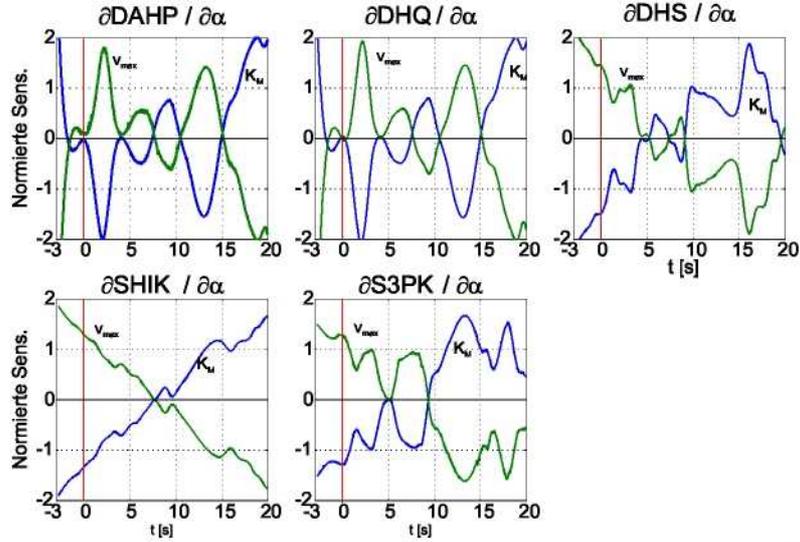


Abbildung 14.3: Auf den Mittelwert und die Varianz normierte Darstellung (Parameter $\text{DHQDH } K_{m,A}$ und $\text{DHQDH } v_{max}$). Aus dieser Darstellung lässt sich schnell die entgegengesetzte Wirkung erkennen. In beiden Fällen entsteht eine hohe Parameter-Korrelation.

($\mu(\text{Sens}_{\alpha_i}^{c_j})$ ist der Mittelwert der Sensitivität der Konzentration j bezüglich α_i über die Zeit. $\text{VarSens}_{\alpha_i}^{c_j}$ die Varianz über die Zeit):

$$\bar{S}_{\alpha_i}^{c_j} = \frac{\text{Sens}_{\alpha_i}^{c_j}(t) - \mu(\text{Sens}_{\alpha_i}^{c_j})}{\text{Var}(\text{Sens}_{\alpha_i}^{c_j})}$$

In Abb. 14.3 zeigt sich deutlich, dass der v_{max} und K_m -Wert der DHQDH unter den untersuchten Bedingungen einen genau entgegengesetzten Einfluss auf alle Messdaten ausüben. Aufgrund der Messungen wird es folglich nicht möglich sein, diese Parameter getrennt zu identifizieren.

14.2 Bestimmtheit im LinLog Modellansatz

Der LinLog-Ansatz wird aufgrund seiner relativ geringen Anzahl von Parametern und der einfachen Struktur, sowie der Eigenschaft, einige mechanistische Kinetiken in weiten Bereichen gut nähern zu können, bevorzugt eingesetzt (HEIJNEN, 2005). Im Fall der fokussierten Modellierung und ohne Steady-State Flussdaten zeigt sich allerdings ein unerwartetes Bild. Zum einen sinkt die Parameteranzahl nicht wesentlich. Sind im mechanistischen Modellansatz je nach Variante 25 bis 34 Parameter zu schätzen, sind es für den LinLog-Ansatz 25 bis 30. Im Gegensatz zum mechanistischen Ansatz wird im Lin-

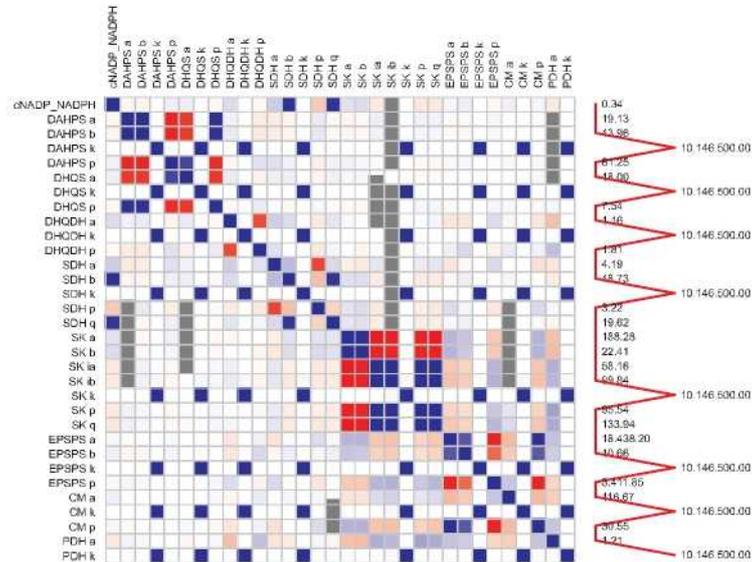


Abbildung 14.4: Parameter-Korrelation im LinLog Modellansatz. Sehr schnell erkennt man, dass die k -Parameter linear abhängig sind und anhand von Konzentrationsmessungen nicht bestimmt werden können.

Log Ansatz immer ein reversibler Mechanismus angenommen¹. Vorteile ergeben sich also erst, wenn im mechanistischen Ansatz kompliziertere Mechanismen auftreten. Denn eine einfache, irreversible Michaelis-Menten Kinetik besitzt weniger Parameter (2: v_{max} , K_M) als der äquivalente LinLog-Ansatz (3: k , a , p). Für eine Reaktion mit 2 Substraten und 2 Produkten müssen insgesamt 5 Parameter geschätzt werden. Handelt es sich um eine irreversible Reaktion kommt man mit einem vereinfachten mechanistischen Ansatz bereits mit 3 Parametern (v_{max} , K_{mA} , K_{mB}) aus.

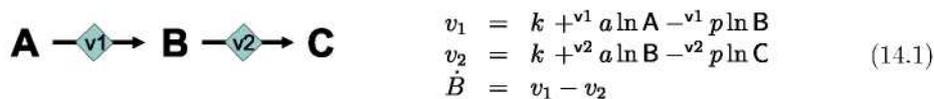


Abbildung 14.5: Ausschnitt einer linearen Reaktionskette in LinLog-Darstellung.

Betrachtet man die Korrelationen im LinLog-Ansatz (Abb. 14.4) im Vergleich zur zuvor bestimmten Korrelation der mechanistischen Parameter, erkennt man eine höhere Abhängigkeit (Korrelation) der Parameter innerhalb einer Kinetik. Teilweise wird auch eine sehr hohe Korrelation von Parameter aufeinander folgender Reaktionen beobachtet (z.B. $DAHPS_p$, $DHQ a$). Diese Korrelationen begründen sich in der Tatsache, dass aufeinander folgende Reaktionen jeweils einen Parameter mit demselben \ln -Term aufweisen. Als Beispiel

¹Irreversible Reaktionen werden durch Beschränken auf nicht negative Flüsse erreicht, der Parameter für das Produkt geht jedoch trotzdem in die Kinetik ein.

ist in Abb. 14.5 der Ausschnitt einer linearen Reaktions-Sequenz dargestellt, die Parameter $v^1 p$ und $v^2 a$ stehen jeweils vor $\ln \mathbf{B}$ (Gl. (14.1)). Hier sollte man sich wieder ins Gedächtnis rufen, dass Konzentrationen gemessen werden. Eine Steigerung der Konzentration \mathbf{B} kann durch Absenken von $v^1 p$, oder Erhöhen von $v^2 a$ erreicht werden. Der Einfluss dieser beiden Parameter ist so gut wie identisch, da jeweils mit demselben $\ln \mathbf{B}$ -Term multipliziert wird. Es entsteht eine hohe Korrelation.

Im fokussierten Modell ergeben sich die größten Schwierigkeiten bei der Bestimmung der k -Parameter. Nimmt man an, die Konzentration von \mathbf{A} wird durch einen Spline aus den Messdaten vorgegeben und \mathbf{B} ist wieder eine gemessene Konzentration, so kann der k -Parameter nicht bestimmt werden. $v^1 k$ und $v^2 k$ sind linear abhängig: werden beide Parameter in gleichem Maße um Δk erhöht oder erniedrigt, verändert sich zwar der Fluss durch die Reaktionen $v_1 = v_1 + \Delta k$ und $v_2 = v_2 + \Delta k$, die Konzentration von \mathbf{B} bleibt allerdings unverändert, da sich sowohl der Zu- als auch der Abfluss um Δk ändert.

Vergleicht man den allgemeinen LinLog-Ansatz mit dem normierten Ansatz (HEIJNEN, 2005), fällt auf, dass die k -Parameter nicht mehr in den Gleichungen auftreten. Stehen für den Steady-State sowohl Fluss- als auch Konzentrations-Messungen zur Verfügung, können die k -Parameter "eliminiert" werden. Nimmt man für das Beispielsystem (Abb. 14.5) einen gemessenen Referenz Steady-State ($\mathbf{v}^0, \mathbf{c}^0$) vor dem Puls an:

$$\begin{aligned} v_1^0 &= k + a \ln \mathbf{A}^0 - p \ln \mathbf{B}^0 \\ v_2^0 &= k + a \ln \mathbf{B}^0 - p \ln \mathbf{C}^0 \end{aligned} \quad (14.2)$$

und normiert die allgemeine LinLog-Form in Abb. 14.5, ergibt sich folgendes Format:

$$\frac{v_1}{v_1^0} = 1 + \frac{a}{v_1^0} \ln \frac{\mathbf{A}}{\mathbf{A}^0} - \frac{p}{v_1^0} \ln \frac{\mathbf{B}}{\mathbf{B}^0} \quad (14.3)$$

$$\frac{v_2}{v_2^0} = 1 + \frac{a}{v_2^0} \ln \frac{\mathbf{B}}{\mathbf{B}^0} - \frac{p}{v_2^0} \ln \frac{\mathbf{C}}{\mathbf{C}^0} \quad (14.4)$$

Mit dieser Normierung wird allerdings implizit angenommen, der Steady-State sei fehlerfrei gemessen (sowohl für die Konzentrationen als auch die Flüsse). Ein vergleichbares Vorgehen wäre, Gleichung (14.2) nach k umzustellen und als Linearkombination der a und p -Parameter darzustellen².

14.3 Effektor-Einflüsse

Der Einfluss von Effektoren auf die Reaktionsgeschwindigkeit wird nach der in Kapitel 9 vorgestellten Vorschrift beurteilt. Im mechanistischen Modell finden sich für die beste

² $\mathbf{A}^0, \mathbf{B}^0$ und v_1^0 sind bekannt und geben die Faktoren vor a und p

Modellvariante zwei Inhibitoren, die gleichzeitig auf die Shikimat Kinase Reaktion wirken. Alternativ wurde eine Variante aufgestellt, die nur eine Inhibierung durch ADP enthält. Unter den besten Anpassungen finden sich auch einige Modelle mit dem vereinfachten Mechanismus.

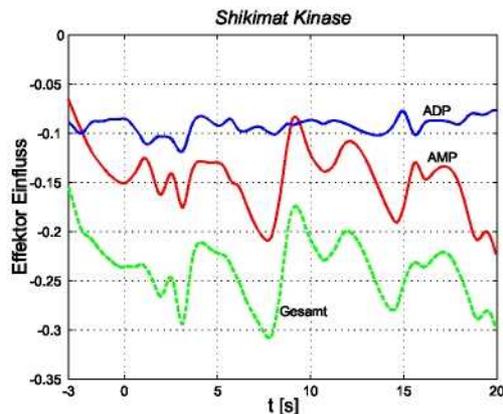


Abbildung 14.6: Darstellung der Effektorwirkungen nach der Methode von NOACK ET AL. (2006). AMP und ADP reduzieren die Reaktionsgeschwindigkeit um maximal 30%. AMP wirkt dabei stärker inhibierend (maximal 22%) als ADP (maximal 12%).

In Abb. 14.6 erkennt man, dass die Reaktion durch den Einfluss von AMP und ADP um bis zu 30% langsamer wird. Die Inhibierung durch AMP ist im Vergleich zu ADP ungefähr doppelt so stark ausgeprägt (Maximum bei ca. 22%, ADP ca. 12%). Zudem kann beobachtet werden, dass sich die Inhibierung durch ADP und AMP über die Zeit sehr ähneln. Grund dafür ist der qualitativ vergleichbare Verlauf von ADP und AMP. Damit erklärt sich auch, warum sich hier keine der beiden Modellvarianten eindeutig absetzt. Es kann lediglich festgestellt werden, dass eine Inhibierung vorliegt, ob sie durch ADP oder AMP oder gar beide hervorgerufen wird, kann anhand der gemessenen Verläufe nicht differenziert werden.

14.4 Einfluss der Eingangsgößen (Splines) in fokussierten Modellen

In diesem Abschnitt soll der Frage nachgegangen werden, inwieweit das Glätten der Eingangsgößen (Splines) Einfluss auf das Modellverhalten nimmt. Es soll untersucht werden, wie sich Fehler bei der Konzentrationsmessung in den Splines zeigen und wie sie auf das fokussierte Modell wirken.

Um diese Effekte genauer zu untersuchen, werden Monte-Carlo Simulationsstudien für den Shikimat Biosyntheseweg durchgeführt. Zur Verdeutlichung des Einflusses werden generische Eingangssignale verwendet, die sich an den gemessenen Signalen orientieren. Die Messdaten werden mit normalverteilten Fehlern behaftet, geglättet, und für die Simulation verwendet. Die Simulationsergebnisse werden dann statistisch untersucht.

14.4.1 Monte-Carlo Studie anhand des Modells für den Shikimat-Biosyntheseweg

Die gemessenen Konzentrations-Zeitverläufe der beiden Pulsexperimente zeigen ein sehr komplexes Verhalten. Für die Untersuchung des Glättens und dem Einfluss auf das Modellverhalten wurden deshalb einfachere Eingangs-Funktionen generiert. Das Modell wurde so parametrisiert, dass vor dem Puls ein Steady-State vorherrscht und der Steady-State Fluss durch den Syntheseweg mit der Abschätzung in Kapitel 16 übereinstimmt. Die Modellgleichungen (Tab. B.5) und Parameterwerte (Tab. B.6) sind im Anhang B.1 zu finden. Die vorgegebenen (“wahren”) Eingangssignale (Splines der Precursor PEP und X5P, sowie der Co-Metabolite ATP, ADP, AMP und NADH/NAD⁺) sind in Abb. 14.7 blau dargestellt. In der Studie werden zu bestimmten Zeitpunkten Konzentrationen gemessen und mit einem Rauschen behaftet.

Das Rauschen setzt sich aus Fehlern bei der Proben-Aufarbeitung σ_ϕ (Verdünnungsfehler) und der LC-MS/MS Messung $\sigma_{MS/MS}$ zusammen. Der Wert ϕ^* ist die wahre Verdünnung, $A^*(t)$ der wahre MS-Messwert, $A^*(t_0)$ ist die Peak-Fläche der Steady-State Konzentration:

$$\begin{aligned}\sigma_\phi &= 0.1 \phi^* \\ \sigma_{MS/MS} &= \sigma_g + a A, \quad \sigma_g = 0.01 A^*(t_0), \quad a = 0.02 A^*(t)\end{aligned}$$

Die so verrauschten Daten werden geglättet (Beispiel einer Glättung ist in Abb. 14.7 rot dargestellt). Verwendet wurden wieder Smoothing-Splines (Gl. (12.6)) mit unterschiedlichen Glättungsfaktoren (im dargestellten Beispiel wurde $smf = 50$ gewählt).

Die so gewonnenen Splines werden in das Modell übertragen (M3L-Datei) und als Eingangssignal verwendet. Die Simulation wird mit den Anfangsparametern durchgeführt und beobachtet, wie sich die simulierten Konzentrationsverläufe verändern.

Die Messung der Eingangssignale, die Glättung und Simulation wird nun mehrmals wiederholt (Monte-Carlo Studie, hier wurden 100 Durchläufe durchgeführt).

Betrachtet man zunächst die Eingangssignale, zeigt sich ein erwartetes Verhalten. Mit einem steigenden Glättungsfaktor geht zunehmend die Dynamik verloren (nicht dargestellt). Daneben zeigen sich auch aber einige andere Effekte. Zum einen erkennt man, dass einzelne Interpolationen im Vergleich zu den Orginaldaten unterschiedliche Charakteristiken

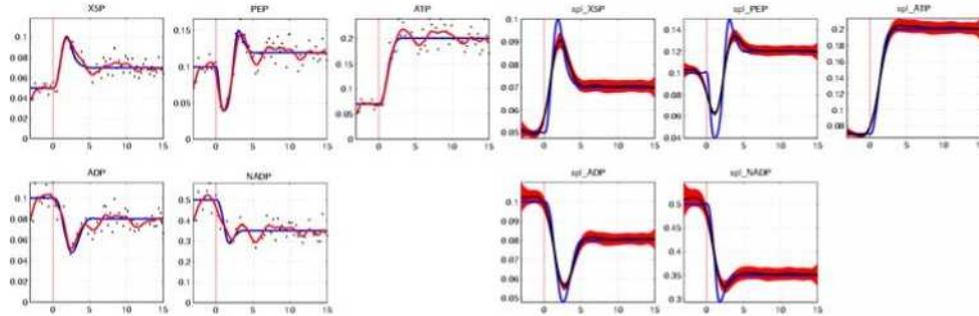


Abbildung 14.7: Links: Originale (unverrauschte) Eingangssignale (blaue Kurven) mit einer beispielhaften verrauschten Messung. Durch diese Messung wird die Glättung gelegt (rot). Rechts: Monte-Carlo Studie der Glättung (100 Läufe). Die Standardabweichungen der geglätteten Konzentrationen sind als Schlauch dargestellt. Die Messdaten wurden hier mit 10% Fehler beim Aufschluss und 3% MS/MS Messfehler behaftet.

aufweisen (z.B. die Oszillationen von NADP^+ vor und nach dem Puls, Abb. 14.7). Sehr deutlich zeigt sich, dass der Spline an den Rändern wesentlich stärker "schwankt" als im mittleren Bereich. Der Variations-Schlauch ist bis zur dritten Messung wesentlich breiter (Abb. 14.7).

Klar erkennbar ist zudem, dass durch die Spline-Glättung auch zeitliche Verschiebungen entstehen. In Bereichen mit hoher Dynamik neigt der Spline zu einer Verzerrung des zeitlichen Maximums. Die Glättung von PEP zeigt diesen Effekt am besten. Der Spline gleitet schon vor dem Puls in das erste Minimum, dieses wird etwas zu früh erreicht, das Maximum bei $t = 3\text{s}$ hingegen zu spät.

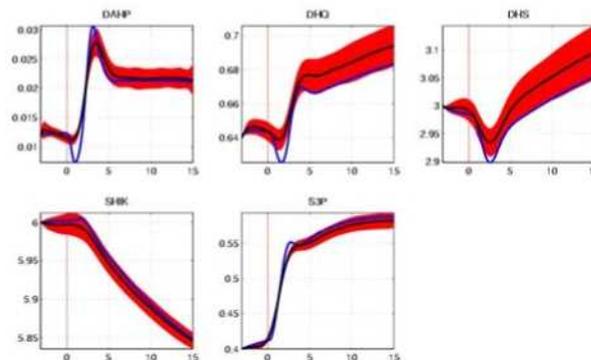


Abbildung 14.8: Simulation mit den verrauschten Eingangs-Splines (100 Monte-Carlo-Wiederholungen). Der Fehler in den Splines pflanzt sich in der Simulation fort und führt zu systematischen Abweichungen.

Vor allem diese Abweichungen wirken sich nun auf die Simulation aus. Im Modell mit dem verrauschten Eingang integriert sich die Abweichung, da der Eingangsspline den Fluss,

d.h. die erste Ableitung der Konzentrationen beeinflusst (Abb. 14.8). Vor allem bei DHQ, DHS und SHIK zeigt sich eine deutliche Abweichung der verrauschten Simulationen im Vergleich zur Simulation mit den vorgegebenen originalen Eingangssignalen (Abb. 14.8). Der Mittelwert aus der Monte-Carlo Studie liegt hier über der originalen Simulation. Zum einen wird für PEP das Minimum nicht vollständig erreicht. Zum anderen steigt die Konzentration von X5P zu früh an. Der Fluss in den Syntheseweg ist somit zunächst zu groß und akkumuliert in DHQ, DHS und SHIK.

Aus diesen Ergebnissen zeigt sich, dass durch die Glättung systematische Abweichungen auftreten können. Im Modell werden Stofftransportprozesse beschrieben. Eine zeitlich auftretende Abweichung wird somit in die Folgemetabolite "transportiert" und kann auch zeitlich länger vorliegen. Es empfiehlt sich deshalb, verschiedene Glättungseinstellungen gegeneinanderzustellen.

Teil IV
Erweiterte Auswertung

Kapitel 15

Steigerung der Parameter -bestimmtheit

Aus der statistischen Untersuchung der verschiedenen Modellansätze zeigt sich deutlich, dass zur Modellidentifizierung weitere Messdaten nötig sind. Denkbar sind folgende Ansätze:

1. **Erweitern der Messdatenbasis:** Neben den Konzentrationsmessungen könnten weitere Informationen, wie z.B. der Steady-State Fluss durch den Biosyntheseweg vor dem Puls in die Betrachtungen einbezogen werden. Dieser Ansatz wird in Abschnitt 15.1 betrachtet.
2. **Weitere Auslenkungen:**
 - Verwenden eines alternativen Substrats, das den Stoffwechsel an anderer Stelle auslenkt (z.B. Glutamat, GAP, etc.).
 - Gezielte Veränderung der Stoffwechselwege durch Knock-Outs oder Überexpression.

Durch diese Experimente entstehen weitere Messdaten, die den durchlaufenen Konzentrationsbereich wesentlich erweitern können. Um eine zusätzliche Auslenkung des Systems in die Parameterschätzung einzubeziehen, müssen zwei oder mehr Experimente gekoppelt ausgewertet werden. Im Rahmen dieser Arbeit liegen zwei Datensätze sehr ähnlicher Stämme vor (*E.coli* 4pF20 und *E.coli* 4pF78), die sich in der Überexpression eines Gens *aroB* unterscheiden. Diese beiden Experimente werden gekoppelt ausgewertet (Abschnitt 15.2).

3. **^{13}C markierte Glucose:** Vor allem aus den Analysen mit LinLog Modellen zeigt sich, dass bei der Betrachtung von Konzentrationsänderungen wenig über den Durchfluss durch einen Pool ausgesagt werden kann. Der Einsatz von ^{13}C markierter Glucose könnte hier wesentlich mehr Informationen bereitstellen. Im Gegensatz zum Puls mit natürlicher Glucose kann beobachtet werden, wie die vor dem Puls unmarkierten Pools mit markiertem Material durchflossen werden. So kann anhand der Markierungsanreicherung beispielsweise unterschieden werden, ob ein Anstieg einer Metabolit-Konzentration durch einen gehemmten Abfluss (langsame Markierungsanreicherung) oder einen stärkeren Zufluss bei gleichem Abfluss (schnelle Mar-

kierungsanreicherung) bedingt wird. Dieser Ansatz wird in Kapitel 17 ausführlich behandelt.

15.1 Messung des Steady-State Flusses

Um Fluss-Informationen auch in der Parameteranpassung ausnutzen zu können, müsste die Software MMT2 angepasst werden – dies war im Rahmen der Dissertation von HAUNSCHILD (2006) nicht mehr möglich. Der Einfluss auf die Parameterbestimmtheit kann deshalb nur *a posteriori* bei der Berechnung der Parameter-Kovarianz untersucht werden. Die Informationen werden durch zusätzliche Messung \mathbf{w} mit den Genauigkeiten $\sigma(\mathbf{w})$ in die Berechnung der Parameter-Kovarianz aufgenommen. Die Berechnung der Sensitivitäten der Flüsse gegenüber den Parametern wurden in Kapitel 7, Gl. (8.19) dargestellt.

In der folgenden Untersuchung wird angenommen, dass der steady-state Fluss durch den Shikimat Biosyntheseweg mit einer Genauigkeit $\sigma(\mathbf{w}) = 0.001\text{mM}/(\text{s})$ bestimmt werden kann. Durch diese zusätzliche Messung werden vor allem die v_{\max} -Werte besser bestimmt. Zudem verringern sich die Korrelationen. So ist z.B. $\text{Corr}(\text{DAHPS } v_{\max}, \text{DHQDH } K_{mA}) = 0.42$ um einen Faktor 2.83 niedriger. Ähnlich deutlich ist der Abfall der Korrelationen von $\text{Corr}(\text{DHQS } v_{\max}, \text{DHQDH } K_{mA}) = -0.36$.

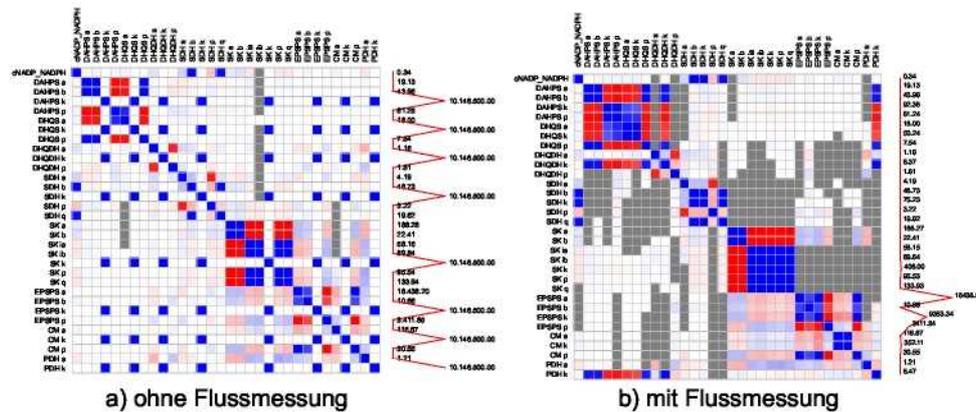


Abbildung 15.1: Darstellung der Korrelationsmatrizen für das LinLog-Modell (Variante 49) unter der Annahme, dass keine Steady-State Flussmessung (a) vorliegt und (b) eine Flussmessung $v_{SS} = w_{SS} \pm 0.001\text{mM}/\text{s}$.

Abb. 15.1 zeigt die Korrelationsmatrizen für ein LinLog-Modell (bestes Modell). Der Einfluss von Fluss-Messungen auf die Parameterbestimmtheit für Modelle im LinLog-Format ist im Vergleich zu mechanistischen Modellen deutlicher, der Informations-Index beträgt hier $I = 2.1$. Im vorangegangenen Kapitel wurde bereits auf die Besonderheit der k -Parameter im LinLog-Format aufmerksam gemacht. Werden Steady-State Fluss-

messungen nicht zur Normierung, sondern als zusätzliche (fehlerbehaftete) Information in die Parameterschätzung einbezogen, können die k -Parameter wesentlich besser bestimmt werden. Die mit Faktoren > 1000 bestimmte Steigerung ist mit Vorsicht zu betrachten, da die Sensitivitäten der Konzentrationen bezüglich der k -Parameter zu linear abhängigen Zeilen in der Parameter-Kovarianzmatrix führen und sie somit singulär ist. In der Darstellung der Korrelationsmatrix (Abb. 15.1) ist der Effekt der Flussmessung deutlich zu erkennen. Die für die Konzentrationen linear abhängigen k -Parameter bilden Punkte hoher Korrelationen (a), mit Flussmessung werden diese Abhängigkeiten entkoppelt.

15.2 Gekoppelte Anpassung mehrerer Datensätze

Durch die gekoppelte Auswertung mehrerer Experimente kann die Datenbasis wesentlich erweitert werden. Vor allem können verschiedene Auslenkungen den durchlaufenen Konzentrationsbereich wesentlich erweitern und somit Kinetiken besser bestimmt werden. Denkbar sind Experimente mit verschiedenen Substraten oder auch mit genetisch veränderten Stämmen. Beispielhaft wird die Möglichkeit eines "genetischen" Pulses untersucht. Es wurden Pulsexperimente mit dem Stamm *E.coli* 4pF20 und 4pF78 durchgeführt. Die beiden Stämme unterscheiden sich lediglich durch ein Gen auf dem Plasmid (4pF78 trägt zusätzlich *aroB*).

Um das Modell zu erstellen, müssen die Kopplungsbedingungen festgelegt werden (Abb. 15.2):

- Es werden für jedes Experiment die jeweiligen Splines als Eingangsdaten verwendet.
- Die Enzymeigenschaften (z.B. K_m und K_i -Werte) sind in beiden Stämmen gleich, die Parameter werden in beiden Modellen gleichgesetzt.
- Die Expression bzw. Enzymaktivität die in den v_{max} -Werten wiedergegeben wird, ist in den beiden Stämmen unterschiedlich. Jedes Modell hat seinen eigenen v_{max} -Parametervektor.

Die Kopplung der Enzymeigenschaften kann in MMT2 durch den Einsatz *globaler* Parameter erreicht werden. Diese Parameter sind im Gegensatz zu den in den Kinetiken definierten *lokalen* Parametern im gesamten Modell verfügbar (HAUNSCHILD, 2005).

Wie schon im ersten Modellierungsansatz werden auch im Doppelmodell mehrere kinetische Varianten zugelassen (Tab. 13.1). Um die Kopplung sicherzustellen, müssen jetzt gleichzeitig in beiden Modellen die selben Kinetiken gewählt werden. Dies wurde manuell durch eine Auswahl der Modellvarianten erreicht.

Die Anpassung gelingt nicht ganz so gut wie für jeden Datensatz getrennt, die wesentlichen Tendenzen werden aber weithin gut wiedergegeben. Die Varianten-Anpassung (Abb. 15.3)

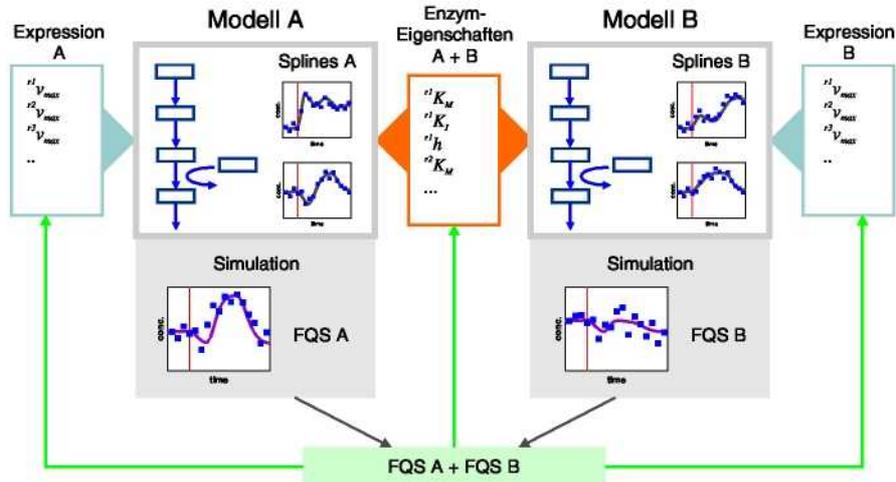


Abbildung 15.2: Kopplung der Modelle für die beiden Experimente 4pF78 und 4pF20. Die Enzym-eigenschaften werden gleichgesetzt, die Expression (v_{max}) ist unterschiedlich. Die Abweichungen werden addiert und bilden eine gemeinsame FQS der beiden Modelle.

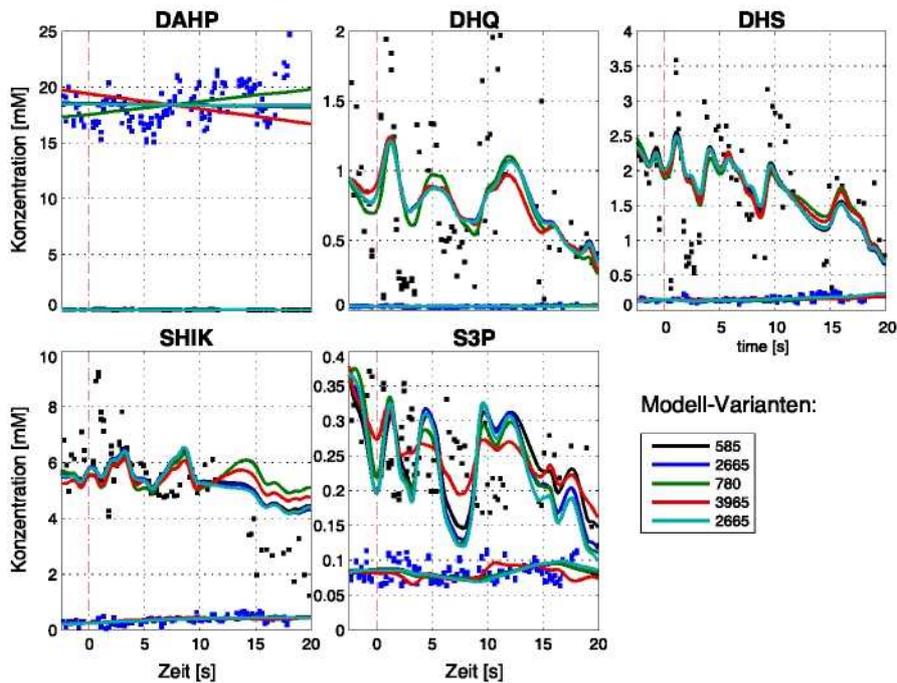


Abbildung 15.3: Die vier besten Anpassungen (Varianten) an die Meßdaten für die beiden Experimente mit *E.coli* 4pF78 und 4pF20.

zeigt ein zum Ein-Datensatz Modell vergleichbares Bild. Wiederum können einige Varianten das gemessene Verhalten vergleichbar reproduzieren. Die Auswertung der besten Modell-Varianten ist in Abb. 15.4 zusammengefasst. Untersucht wird die Häufigkeit der kinetischen Varianten der einzelnen Reaktionsschritte für Modell-Varianten mit bis zu 2% (5%, 10%) höherer FQS.

Zu erkennen sind zur Anpassung an den Datensatz 4pF78 ähnliche Tendenzen. Für die Eingangsreaktion wird offensichtlich der Ordered-Bi-Bi-Mechanismus bevorzugt. Ungefähr 80% der besten Modellanpassungen (bei $FQS \leq 102\%$) enthalten diesen Mechanismus. Für die DHQ Synthase zeigt sich ein eher unerwartetes Bild. Während für die Einzel-Anpassung ein inhibierter Mechanismus gefunden wurde, wird in der Doppelmodell-Anpassung der Gleichgewichts-Mechanismus favorisiert (66%). Die DAHP Konzentration im Datensatz 4pF20 liegt um Größenordnungen über der bei 4pF78 gemessenen. Traut man dem Ergebnis, war die Aktivität des Enzyms in 4pF20 sehr gering und nicht durch die extrem hohe Substratkonzentration inhibiert.

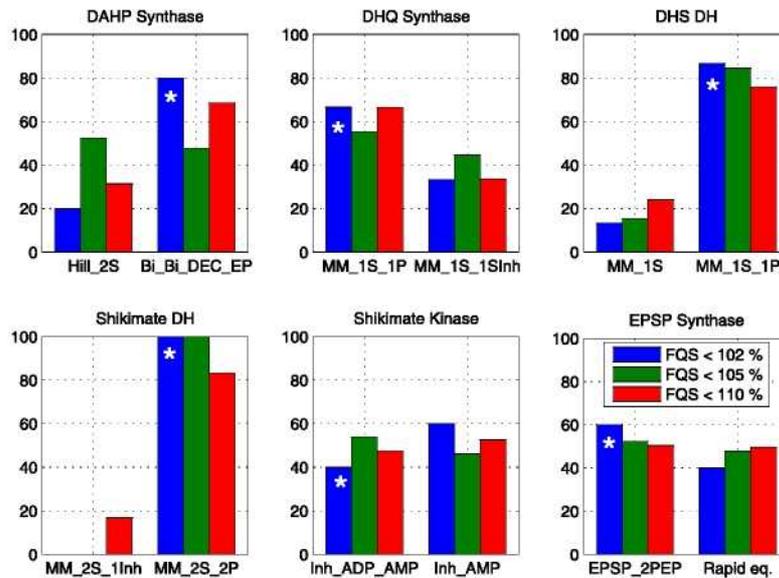


Abbildung 15.4: Häufigkeiten der in den besten Modellen auftretenden Enzymmechanismen – Bei der Berechnung wurden Modell-Anpassungen berücksichtigt, deren FQS maximal 2% (5%, 10%) über der FQS der besten Anpassung liegen.

Wie im Einzelmodell wird sowohl bei der DHS Dehydratase als auch der Shikimat Dehydrogenase Reaktion der Gleichgewichts-Mechanismus vorgezogen. Im Fall der SDH enthalten alle Anpassungen die Gleichgewichts-Variante ($FQS \leq 102\%$). Selbst etwas schlechtere Kandidaten ($FQS \leq 105\%$) enthalten fast ausschließlich Varianten mit Gleichgewichts-

reaktionen. Das eindeutige Ergebnis für die Shikimat Dehydrogenase widerspricht der von DELL UND FROST (1993) *in vitro* beschriebenen Inhibierung durch das Produkt Shikimat. Der aus den *in vitro* Daten bestimmte K_I -Wert von **0.16mM** (wurde aus dem Diagramm in (DELL UND FROST, 1993) abgelesen) scheint allerdings in jedem Fall *in vivo* nicht vorzuliegen. Bei Shikimatkonzentrationen im Bereich von mehreren mM würde die Reaktion vollständig inhibiert.

Für die Shikimat Kinase zeigt sich keine so deutliche Tendenz¹. Auch für den letzten Reaktionsschritt sind keine ausgeprägten Häufigkeiten zu beobachten. Modelle mit Ordered-Bi-Bi-Mechanismus schneiden in der Anpassung geringfügig besser ab, 60% der Modelle enthalten ihn.

15.2.1 Vergleich der Enzym-Aktivitäten

Aus der Kopplung der beiden Modelle (Abb. 15.2) ergibt sich eine weitere interessante Analyse. Die beiden Stämme *E.coli* 4pF20 und 4pF78 können hinsichtlich der Enzymaktivitäten verglichen werden. Die unabhängigen v_{max} Vektoren werden gegenübergestellt (Tab. 15.1).

Klar zu erkennen ist der wesentlich höhere v_{max} -Wert (Aktivität) der DHQ Synthase Reaktion im Stamm 4pF78. In den vier besten Anpassungen liegt der v_{max} -Wert durchschnittlich mehr als 100fach über dem des Stamms 4pF20. Die Aktivität der DAHP Synthase, sowie der DHQ Dehydratase scheinen dagegen in etwa konstant zu bleiben. Diese Beobachtung stimmen mit den Genotypen überein. Der Stamm 4pF78 trägt *aroB* (DHQ Dehydratase) auf dem Plasmid, 4pF20 nicht. Die Plasmide beider Stämme enthalten *aroF^{br}* (DAHP Synthase).

Für die restlichen Reaktionen sind für den Stamm 4pF78 höhere Aktivitäten zu beobachten, deren Ausmaß jedoch sehr stark zwischen den Varianten schwankt. Zudem sind die berechneten Standardabweichungen sehr hoch.

15.2.2 Parameterbestimmtheit im Doppelmodell

Zur Beurteilung des Doppel-Modell Ansatzes wird die Parameterbestimmtheit mit der der Einzel-Modelle verglichen. Untersucht werden zwei Anpassungen, die Modell-Varianten 585 und 2665. Um vergleichbare Bedingungen herzustellen, werden die Parameterwerte der Doppel-Modell Anpassung auf die beiden Einzel-Modelle übertragen. Zum Vergleich der statistischen Güte wird ein in der Versuchsplanung übliches Maß, das *D*-Kriterium herangezogen (Gl. (7.19)). Da im Doppelmodell mehr Parameter enthalten sind (jedes

¹Mechanismen ohne Regulation wurden in diesen Anpassungen nicht in die Modellfamilie aufgenommen. In den Studien mit den jeweiligen Einzelmodellen erschien diese Variante nicht relevant.

Tabelle 15.1: Vergleich der Enzym-Aktivität (v_{max} -Parameter) der Modelle für den Stamm 4pF78 und 4pF20. Dargestellt sind die Quotienten $^{4pF78}v_{max}/^{4pF20}v_{max}$ sowie die dazugehörigen Standardabweichungen. Die teilweise sehr großen Standardabweichungen ergeben sich im Fall sehr kleiner v_{max} -Werte im Nenner mit relativ großen Standardabweichungen der Parameter (vgl. Abschnitt 15.2.2)

| Enzym | Vergleich der Enzym-Aktivität - 4pF78 / 4pF20 | | | |
|-----------------|---|---------------|---------------|-------------|
| | 585 | 2665 | 780 | 3965 |
| DAHP Synthase | 1.0 ± 0.05 | 1.0 ± 0.19 | 24.8 ± 1681 | 1.0 ± 0.19 |
| DHQ Synthase | 282.4 ± 486 | 322.6 ± 1588 | 368.5 ± 75878 | 66.3 ± 10.6 |
| DHS DH | 1.9 ± 0.59 | 1.6 ± 89.5 | 0.5 ± 0.08 | 0.2 ± 15.7 |
| Shikimat DH | 80.0 ± 1049 | 80.0 ± 850.28 | 37.4 ± 298 | 33.9 ± 155 |
| Shikimat Kinase | 25.4 ± 1267 | 8.9 ± 39.6 | 8.1 ± 158 | 2.2 ± 16.9 |
| EPSP Synthase | 23.9 ± 8.01 | 30.2 ± 18.9 | 37.0 ± 278 | 3.3 ± 53.9 |

Modell hat seinen \mathbf{v}_{max} -Vektor, sowie je einen Parameter für die Gesamt-Konzentration von NADP + NADPH), wird das D -Kriterium in den Informations-Index I (Gl. (7.20)) umgerechnet.

Als Referenz I_{ref} wird das Einzelmodell des Stamms *E.coli* 4pF78 gewählt, es enthält $\dim(\alpha) = 30$ Parameter und hat ein D -Kriterium von $D_{ref,78} = 1.42 \cdot 10^{-12}$. Das Doppelmodell hat $\dim(\alpha_{78+20}) = 37$ Parameter bei einem D -Kriterium $D = 2.07 \cdot 10^{-33}$. Offensichtlich erfüllt sich damit die Hoffnung auf eine höhere Parameterbestimmtheit. Der Informations-Index beträgt $I = 3.08$, d.h. die Standardabweichungen sind im Durchschnitt dreimal niedriger. Die Verteilung der Standardabweichungen über die Parameter wird in Abb. 15.5 als akkumulierter Plot dargestellt. Aufgrund der logarithmischen Skalierung erkennt man nicht auf den ersten Blick die Steigerung um einen Faktor drei. Sie ist für einzelne Parameter sehr unterschiedlich, wodurch teilweise ein weiter Abstand zwischen den Kurven zu erkennen ist.

In der Tabelle in Abb. 15.5 ist die Bestimmtheit der im Doppelmodell gut bestimmten Parameter im Vergleich zu den beiden Einfach-Modellen dargestellt. Für den Parameter ${}^5K_{Ib}$ (letzte Zeile in der Tabelle für Variante 585) ist die erwartete Akkumulation der Information zu beobachten, die Standardabweichung sinkt von $\pm 200\%$ auf $\pm 50\%$. Erstaunlicherweise wird aber auch die Bestimmtheit von Parametern verbessert, die nur in einem der beiden Modelle vorkommen. So kann beispielsweise der Parameter ${}^{v2(B)}v_{max}$ im Doppelmodell wesentlich besser bestimmt werden (statt $\pm 205\%$, jetzt $\pm 35\%$) obwohl er nur in Modell B vorkommt. Die Erklärung ist in den Korrelationen zu finden. Der Parameter ${}^{v2(B)}v_{max}$ korrelierte stark mit ${}^{v2}K_{eq}$ (der in beiden Modellen vorkommt). Die Standardabweichung von ${}^{v2}K_{eq}$ ist im Einzelmodell 4pF20 (Modell B) extrem groß ($\pm 12225\%$), im

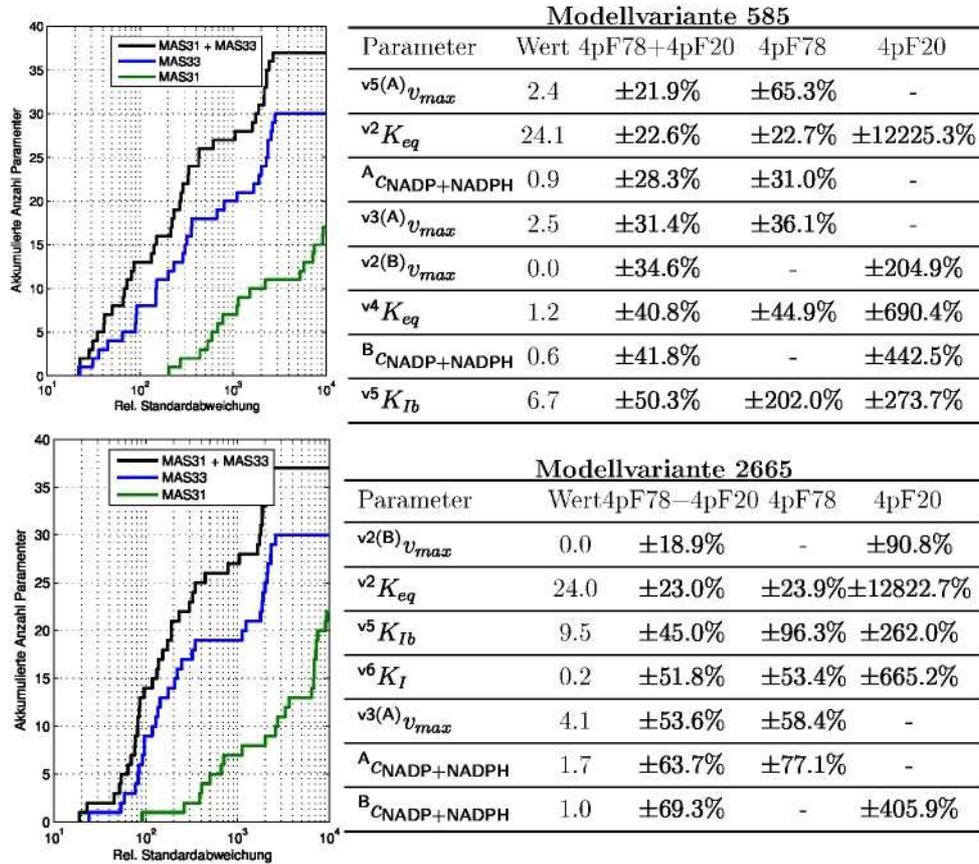


Abbildung 15.5: Darstellung der relativen Fehler der Parameter für die beiden Datensätze (4pF20 und 4pF78) sowie der des Doppelmodells als akkumulierte Verteilung.

Doppel-Modell beträgt sie $\pm 23\%$. In den weiteren Modellvarianten ist ein vergleichbares Verhalten zu beobachten (Abb. 15.5).

Kapitel 16

Modell-Analyse und -Konsistenz

In den folgenden Abschnitten werden weitere Methoden zur Modell-Analyse und Ansätze zur Überprüfung der Modellaussagen diskutiert. Aus den Simulation erhält man neben dem Konzentrationsverlauf auch die Flüsse. Es wird diskutiert, ob die simulierten intrazellulären Flüsse mit den zu erwartenden Werten übereinstimmen (Abschnitt 16.1).

Das Modell kann verwendet werden um gezielte Vorschläge zur Stammverbesserung abzuleiten. Optimierte werden die Enzymexpressionen. Im Vergleich der beiden Stämme wird die Plausibilität der vorgeschlagenen Enzymexpressionen diskutiert (Abschnitt 16.3).

Einige Enzym-Mechanismen lassen Rückschlüsse auf das thermodynamische Gleichgewicht zu. Allgemein wird angenommen, dass Reaktionen mit einem hohen thermodynamischen Gefälle in der Zelle reguliert werden. Reaktionen, die in der Nähe des thermodynamischen Gleichgewichts ablaufen, folgen dagegen meist einen reversiblen Mechanismus. Das thermodynamische Gefälle wird anhand einer Näherung der Gibbschen Energien bestimmt (Abschnitt 16.5).

16.1 Abschätzung maximaler Flussraten

Vor der Puls-Aufgabe befinden sich die Bakterien in einem limitierten stationären Zustand. Die zugeführte Glucose wird sofort von den Zellen aufgenommen. Die spezifische (extrazelluläre) Glucose-Aufnahmerate v_{upt}^0 im Steady-State vor dem Puls ergibt sich somit aus der im Feed zugeführten Menge ($\dot{V}_{\text{feed}} c_{\text{feed}}$):

$$v_{\text{upt}}^0 = \frac{\dot{V}_{\text{feed}} c_{\text{feed}}}{V_{\text{reactor}} c_{\text{BTM}}} = 0.752 \frac{\text{mmol}}{\text{g}_{\text{BTM}} \text{h}} \quad (16.1)$$

Die CO_2 Bildungsrate wurde online gemessen. In den Aromaten-Biosyntheseweg kann während der Limitierung maximal noch der verbleibende Kohlenstoff fließen. Da auf dem Weg zu L-Phenylalanin ein Molekül CO_2 abgespalten wird, berechnet sich der maximal möglich Fluss zu ($n_{\text{C,Glc}}$ = Anzahl der C-Atome in Glc, $n_{\text{C,phe}}$ = Anzahl der C-Atome von L-Phe):

$$v_{\text{phe}}^0 = \frac{n_{\text{C,Glc}} v_{\text{upt}}^0 - v_{\text{CO}_2}^0}{n_{\text{C,phe}}} = 0.132 \frac{\text{mmol}}{\text{g}_{\text{BTM}} \text{h}} = 0.0132 \frac{\text{mM}_{\text{intra}}}{\text{s}} \quad (16.2)$$

Für die Umrechnung wurde eine spezifischen Dichte von $\rho = 2.78 \text{ ml}_{\text{intra}}/\text{g}_{\text{BTM}}$ (l_{intra} bezeichnet intrazelluläres Volumen) angenommen (BUCHHOLZ ET AL., 2001).

Ein Zugang zu einer Abschätzung nach dem Puls ergibt sich aus der Betrachtung der maximal möglichen Ausbeute (Abschnitt 11.2) und der im Experiment beobachteten maximalen Glucose-Aufnahmerate. In der Wachstumsphase erreicht der Stamm eine Substrat-Aufnahmerate $v_{\text{upt}} \approx 10 \text{ mmol}/(\text{g}_{\text{BTM}} \text{ h})$. Es ist durchaus möglich, dass die Glucoseaufnahme nach dem Puls auch über diesen Wert steigt (CHASSAGNOLE ET AL., 2002). Es kann aber angenommen werden, dass im Mittel über den Beobachtungszeitraum dieser Wert zumindest nicht wesentlich überschritten wird. Damit kann auch eine Obergrenze für den Fluss in den Aromaten-Biosyntheseweg abgeschätzt werden. Die maximale Ausbeute wurde mit $Y_{\text{max}} = 0.55 \text{ mM}/\text{mM}$ bestimmt:

$$v_{\text{phe}} = Y_{\text{max}} v_{\text{upt,max}} = 0.55 \frac{\text{mol}}{\text{mol}_{\text{upt}}} 10 \frac{\text{mmol}}{\text{g}_{\text{BTM}} \text{ h}} = 0.55 \frac{\text{mM}_{\text{intra}}}{\text{s}} \quad (16.3)$$

In den gemessenen intrazellulären Konzentrationen werden Änderungen mit bis zu $v_{\text{upt,max}} = 10 \text{ mM}_{\text{intra}}/\text{s}$ beobachtet (*E.coli* 4pF78, je nach Spline-Glättung auch höher). Auch vor dem Puls sind hohe Konzentrationsänderungen zu beobachten, die eindeutig gegen die hier vorgenommenen Abschätzungen verstoßen (Maximum vor dem Puls sollte um $0.0132 \text{ mmol}_{\text{intra}}^{-1} \text{ s}^{-1}$) liegen. Aus dieser Abschätzung wird deutlich, dass die hohe Dynamik der gemessenen Konzentrationen nicht mit der extrazellulären Akkumulation übereinstimmt und es in der Modellierung nicht gelingen kann, intra- und extrazelluläres Verhalten gleichzeitig und übereinstimmend wiederzugeben. Wieso es zu diesen extremen Konzentrationsänderungen kommt, kann an dieser Stelle nicht geklärt werden. Dazu müssen weitere gezielte Experimente, unter Umständen mit Markierung durchgeführt werden. Aus diesen Daten könnte man erkennen, ob stark fluktuierende Austauschflüsse oder Messfehler die hohe Dynamik verursachen.

16.2 Vergleich der simulierten und notwendigen Flüsse zur Wiedergabe der Messdaten

Die Konzentrationsänderungen der intrazellulären Pools entstehen durch die zu- und abfließenden Reaktionen. Die Beziehung von Flüssen und Konzentrationsänderungen ist durch die Stöchiometrie vorgegeben und kann genutzt werden, den notwendigen Fluss zu rekonstruieren. BARDOW UND MARQUARDT (2004) verwenden diesen Zusammenhang zur inkrementellen Modellidentifizierung. Statt eines Differentialgleichungssystems muss in den Flusskoordinaten nur ein Gleichungssystem gelöst und angepasst werden muss.

Betrachtet man den in dieser Arbeit untersuchten linearen Syntheseweg, können die zur Wiedergabe der Messungen notwendigen Flüsse bei bekanntem Zu- und Abfluss in und

aus dem Syntheseweg bestimmt werden. Da die Messdaten mit starkem Rauschen behaftet sind und die Ableitung der Konzentrationen benötigt wird, wird mit geglätteten Konzentrationen (Abb. 16.1) gearbeitet.

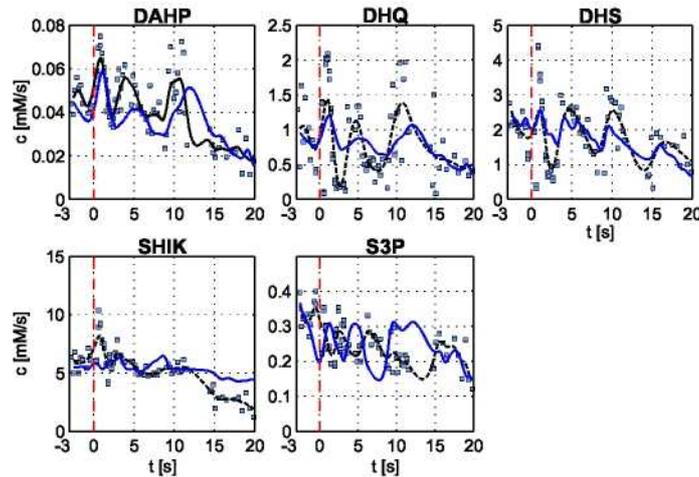


Abbildung 16.1: Konzentrations-Zeitverlauf der gemessenen Konzentrationen mit Spline-Glättung (gestrichelte Linie) und der besten Modellanpassung (Doppelmodell Nr. 585) für den Datensatz *E.coli* 4pF78.

Schwierigkeiten bereitet der Ein- und Ausfluss aus dem Biosyntheseweg. Messtechnisch ist kein Zugang zum (intrazellulären) Fluss über die Systemgrenze möglich.

Hier wurden deshalb die aus der Simulation gewonnenen Zu- und Abflüsse eingesetzt. In Abb. 16.2 sind die aus den Konzentrationen und der Stöchiometrie berechneten Flüsse (gestrichelt) und die Flüsse aus der Simulation (blau) dargestellt. Zu beobachten sind:

- DAHP Synthese: Der Fluss der DAHP Synthese stammt aus der Simulation, simulierter und rekonstruierter Fluss stimmen somit per Definition überein.
- DHQ Synthese: Simulierter und notwendiger Abfluss aus dem Pool (DHQ Synthese) sind sich sehr ähnlich. Da die Konzentrationsänderungen im DAHP Pool extrem niedrig sind, müssen Ein- und Ausfluss fast identisch sein. Obwohl hier der simulierte Zu- und Abfluss den rekonstruierten Flüssen entspricht, erkennt man in Abb. 16.1 bereits Abweichungen zwischen simulierten und geglätteten Konzentrationen. Dies deutet darauf hin, dass der aus der Simulation entnommene Fluss in den Aromaten-Biosyntheseweg fehlerbehaftet ist.
- DHQ Dehydratase: Für diese Reaktion sind deutliche Abweichungen zu beobachten. Der rekonstruierte Fluss erreicht ein Maximum von $\overset{\text{DHQDH}}{\tilde{v}} \approx 2.3 \text{ mM/s}$, die Simulation kommt auf maximal $\overset{\text{DHQDH}}{\tilde{v}} \approx 1.45 \text{ mM/s}$.

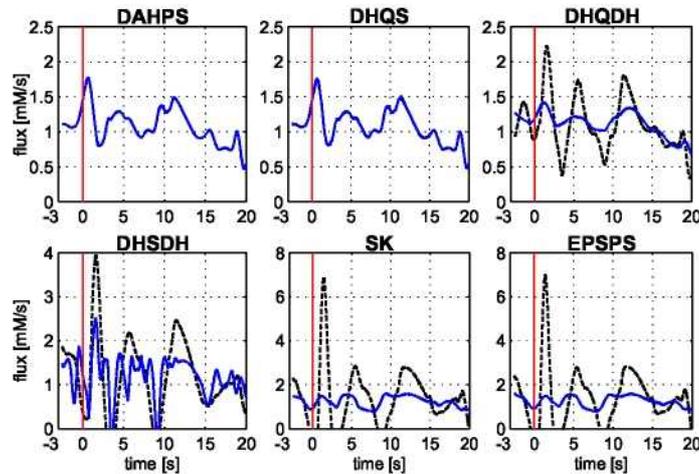


Abbildung 16.2: Vergleich der Flüsse aus der Simulation (blau) und der Berechnung aus den geglätteten Konzentrationsmessungen und der Stöchiometrie des Biosynthesewegs (gestrichelt).

- Shikimat Kinase: Für diesen Reaktionsschritt sind die Unterschiede am deutlichsten. Um die gemessenen Konzentrationsänderungen in SHIK wiederzugeben wird ein maximaler Fluss von 8 mM/s benötigt. Die Simulation ist nicht in der Lage, diese großen Flussänderungen wiederzugeben. Der Fluss wurde auf den Bereich bis maximal 2.5 mM/s beschränkt (Abschnitt 16.1). Zu beobachten sind hier auch leichte zeitliche Verschiebungen. Der zweite Anstieg im Fluss der SK Reaktion ist in der Simulation zu früh (ca. 1 s). Genauso ist der in der Simulation beobachtete Abfall der Konzentration ca. 2 s zu früh.
- EPSP Synthase: Da die Konzentrationen von S3P im Vergleich zu SHIK wieder sehr gering sind, sind Ein- und Ausfluss aus dem Pool fast identisch. Die für die Shikimat Kinase beobachteten Abweichungen sind somit auch für die EPSP Synthase zu beobachten.

Die Zeitverzögerungen sprechen dafür, dass eine Feedback Inhibierung der Shikimat Kinase durch Nachfolgemetabolite vorliegt. Deren Konzentrationsprofil wäre stärker verzögert und könnte zur beobachteten zeitlichen Verschiebung führen. Nicht auszuschließen ist allerdings auch, dass der Effekt mit durch die Glättung hervorgerufen wird (s. auch Abschnitt 14.4)

16.3 Optimierung der Enzymexpression für den Steady State

Wird angenommen, das Modell für den Aromaten-Biosyntheseweg gibt die Realität ausreichend genau wieder, kann es zur Vorhersage optimaler Enzymexpressionen eingesetzt werden.

Die Zelle kann nur eine begrenzte Menge Enzym produzieren (BROWN, 1991). Aufgrund der Homeostase (VISSER UND HEIJNEN, 2003) können auch die intrazellulären Konzentrationen nicht beliebig verändert werden. Diese beiden Einschränkungen müssen bei der Variation der Expression beachtet werden.

Die Enzymmengen \mathbf{e} bzw. die zur Enzymmenge proportionalen \mathbf{v}_{\max} -Werte werden aus einer Optimierung des steady-state Flusses (\mathbf{v}) durch den Biosyntheseweg bestimmt. Die Optimierung unter Nebenbedingungen lautet dann (\mathbf{v}_{\max}^0 sind die Enzymmengen des Ausgangszustands, \mathbf{v}_{\max} die veränderten Enzymmengen, \mathbf{c}^0 sind die Konzentrationen im gemessenen steady-state (Tab. 16.1), \mathbf{c} die sich einstellenden Konzentrationen mit den Enzymmengen \mathbf{v}_{\max} , der Skalar n_v gibt die Anzahl der Enzyme, n_c die Anzahl Pools wieder)¹:

$$\mathbf{v}_{\max}^{\text{opt}} = \arg \max_{\mathbf{v}_{\max}} v_{phe}(\mathbf{c}, \mathbf{c}_{\text{ex}}, \mathbf{v}_{\max}, \boldsymbol{\alpha}) \quad \text{so dass:} \quad \begin{aligned} \mathbf{N} \mathbf{v} &= \mathbf{0} \\ \sum_{i=1}^{n_v} \frac{v_{\max,i}}{v_{\max,i}^0} &\leq 1.25 n_v \\ \frac{1}{n_c} \sum_{j=1}^{n_c} \frac{|c_j - c_j^0|}{c_j^0} &\leq 5 n_c \end{aligned} \quad (16.4)$$

Die außerhalb des Biosyntheseweges liegenden Konzentrationen (\mathbf{c}_{ex}^0 , linke Tabelle) bleiben während der Optimierung unverändert, die intrazellulären Konzentrationen (rechte Tabelle) ergeben sich aus der steady-state Bedingung. Abb. 16.3 zeigt das Schema der Optimierung. In der inneren Schleife werden iterativ die sich einstellenden Konzentrationen und Flüsse berechnet. In der äußeren Schleife läuft die iterative Optimierung.

Betrachtet man zunächst die vorgeschlagenen Aktivitäten für den Stamm *E. coli* 4pF20, fällt die geforderte, verstärkte Aktivität der DHQ Synthase und Shikimat Kinase auf (Abb. 16.4a). Eine Verdreifachung des Steady-State Flusses wird durch die Steigerung der Aktivitäten (DHQS 3x, SK 2.5x) vorhergesagt. Zu beachten ist jedoch, dass diese Optimierung unter der Annahme konstanter Konzentrationen der Precursor durchgeführt wurde. Die Steigerung kann wesentlich niedriger ausfallen, wenn durch den höheren Bedarf die Konzentrationen der Precursor und Co-Metabolite sinken. Die geforderte Überexpression der DHQ Synthase (*aroB*) ist im Stamm *E. coli* 4pF78 umgesetzt und führte zu einer Steigerung der Produktivität.

¹Die in (VISSER UND HEIJNEN, 2003) gesetzten Grenzen beziehen sich auf ein Modell des Zentralstoffwechsels. Da hier nur ein Teil betrachtet wird, werden die Grenzen vor allem der Konzentrationen größer gewählt.

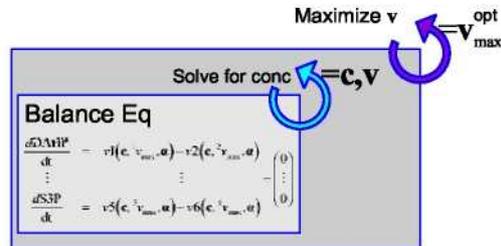


Abbildung 16.3: Optimierung des steady-state Flusses (äußere Schleife). Variiert werden die Enzymexpressionen (v_{max} -Werte). Aus dem Steady-State $Nv(c, c_{ex}, v_{max}, \alpha) = 0$ werden die Konzentrationen und Flüsse bestimmt. Das nicht-lineare Gleichungssystem wird iterativ mit dem Newton-Verfahren gelöst (innere Schleife).

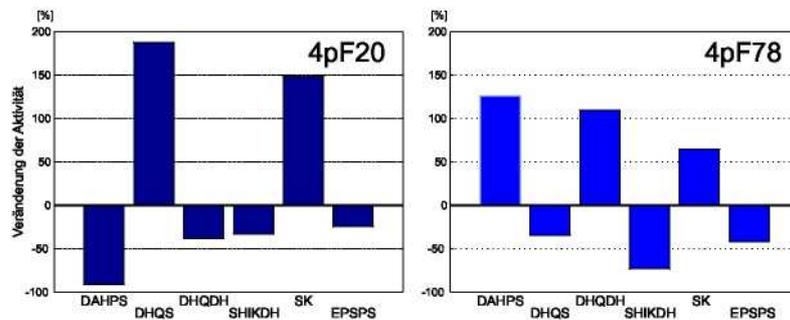


Abbildung 16.4: Vorgeschlagene Veränderung der Enzym-Aktivitäten (v_{max}^{opt}/v_{max}^0) zur Steigerung des Steady-State Flusses. Links für den Stamm 4pF20, rechts für 4pF78. Der Fluss durch den Syntheseweg wird um 200% (4pF20) bzw. 110% (4pF78) gesteigert (unter der Annahme, die Konzentrationen der Precursor und Co-Metabolite bleibt erhalten).

Wird die Optimierung für den Stamm *E.coli* 4pF78 durchgeführt, erkennt man, dass nun drei Enzymaktivitäten erhöht werden sollten (Abb. 16.4b). Der Fluss durch den Aromaten-Biosyntheseweg wird mehr als verdoppelt ($v = 2.1 v_{ref}$). Die bereits überexprimierte Eingangsreaktion (DAHP Synthase) weist offensichtlich eine zu geringe Aktivität auf. Ihre Aktivität sollte 2.2fach erhöht werden. Die Aktivität der schon überexprimierten DHQ Synthase ist anscheinend ausreichend, im optimierten Syntheseweg wird die Aktivität sogar reduziert. Offensichtlich ist durch die Überexpression bereits eine Überkapazität vorhanden. Von allen erniedrigten Aktivitäten wird die der DHQ Synthase allerdings am wenigsten herabgesenkt (An dieser Stelle sei auch auf die gegebene Nebenbedingung einer maximal 1.25fach Erhöhung der Enzymmenge (Gl. (16.4)) hingewiesen. Einige Aktivitäten müssen herabgesenkt werden, um eine verstärkte Aktivität anderer, stärker limitierender Reaktionsschritte zu ermöglichen.)

Als nächste hinderliche Reaktion wird die DHQ Dehydratase ermittelt, ihre Aktivität wird im optimierten Zustand um 100% erhöht. Offensichtlich hoch aktiv ist die Shikimat Dehydrogenase, im optimierten Zustand werden nur ca. 20% der Aktivität benötigt. Die Aktivität der nachfolgenden Reaktion (Shikimat Kinase) muss wiederum erhöht werden (Faktor 1.8).

Tabelle 16.1: Links: Aus den Mittelwerten nach dem Puls bestimmte Konzentrationen \mathbf{c}^0 für die Optimierung des Steady-State Flusses. Rechts: Konzentrationen \mathbf{c} im optimierten Steady-State.

| Metabolit | Konzentration [mM] | | Metabolit | Konzentration [mM] | |
|-----------|--------------------|-------|-----------|--------------------|-----------|
| | 4pF20 | 4pF78 | | 4pF20 | 4pF78 |
| PEP | 0.06 | 0.05 | DAHP | 0.27[au] | 0.039[au] |
| X5P | 0.34 | 0.09 | DHQ | 0.13 | 0.65 |
| ATP | 0.17 | 0.3 | DHS | 2.47 | 10.05 |
| ADP | 0.37 | 0.2 | SHIK | 3.19 | 8.83 |
| AMP | 4.0 | 4.0 | S3P | 0.77 | 0.98 |
| NADP | 0.2 | 0.2 | | | |
| NADPH | 0.4 | 0.71 | | | |

Diese Ergebnisse decken sich weitestgehend mit den experimentellen Beobachtungen. Es wurde festgestellt, dass die Feed-back resistente Variante der DAHP Synthase (*aroF^{br}*) im Vergleich zur nicht modifizierten Variante eine geringere Aktivität aufweist. Der Stamm *E. coli* 4pF81 mit *aroF* und *aroL* (Shikimat Kinase) Überexpression erreicht eine höhere L-Phenylalanin Produktivität. Über den Einfluss der DHQ Dehydratase gibt es kaum experimentelle Befunde. DELL UND FROST (1993) berichten, dass dieser Schritt einen geringen Einfluss auf die Reaktionsgeschwindigkeit der Aromaten-Biosynthesewegs nimmt, da es sich um eine Gleichgewichtsreaktion handelt. Wenn jedoch wirklich zu wenig Katalysator vorhanden ist, stellt sich das Gleichgewicht unter Umständen nur langsam ein, so dass durchaus eine Limitierung denkbar ist.

Die Optimierung zeigt auch, dass zur Steigerung des Flusses die Erhöhung nur einer Enzymaktivität nicht ausreicht. Offensichtlich gibt es in diesen Modellen kein "Bottleneck"-Enzym, im Fall von 4pF20 muss die DHQS als auch die SK überexprimiert werden. Der Stamm 4pF78 trägt die DHQS auf dem Plasmid. Zur weiteren Steigerung müssen drei Enzym-Aktivitäten verstärkt werden, die im vorangegangenen Optimierungsschritt ausreichende DAHPS-, DHQDH-, und die bereits für 4pF20 unzureichende SK-Aktivität.

16.4 Einfluss der extrazellulären Konzentrationen

Untersucht wird die Sensitivität des Flusses durch den Shikimat Biosyntheseweg bezüglich der externen Konzentrationen PEP, X5P, NADP, ATP, ADP und AMP. Die Sensitivität der Konzentrationen c_{ex} auf den Steady-State (Tab. 16.1) ist in Abb. 16.5 wiedergegeben.

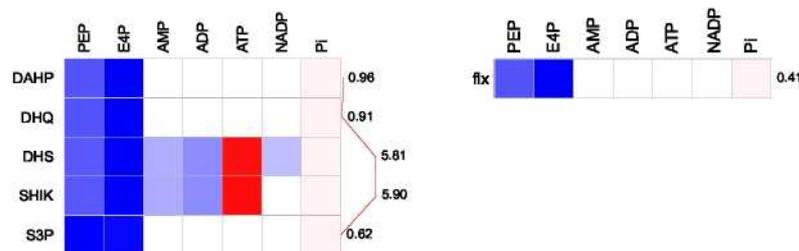


Abbildung 16.5: Sensitivität der Konzentrationen (links) bzw. des Steady-State Flusses (rechts) bezüglich der extrazellulären Konzentrationen c_{ex} . Die Sensitivitäten sind als Konzentrations- bzw. Fluss-Kontroll-Koeffizienten dargestellt (MCA).

Auffällig ist, dass offensichtlich die beiden Precursor PEP und X5P den größten Einfluss auf die Konzentrationen nehmen. Die Sensitivität von PEP ist dabei bis S3P etwas niedriger als der X5P Einfluss. ATP, ADP, AMP und NADP beeinflussen wesentlich die Konzentrationen von DHS und SHIK. Vor allem ATP fällt hier auf. Die Sensitivität auf die DHS und SHIK Konzentration ist vergleichsweise hoch. Genauso die NADP-Konzentration. Steigt sie an, fällt die DHS Konzentration. Schon in den Modellhypothesen wurde eine starke Abhängigkeit von DHS gegenüber NADP bzw. NADPH vermutet, die sich nun im Modell bestätigt.

Betrachtet man die Sensitivität des Flusses durch das System zeigt sich wieder die starke Abhängigkeit von den Precursor-Konzentrationen. ATP, ADP, AMP und NADP spielen eine untergeordnete Rolle. P_i wird in die Kinetiken mit aufgenommen (hier als konstante Konzentration), da ein Einfluss auf die DAHP Synthase beschrieben wurde (SIMPSON UND DAVIDSON, 1976). Eine hohe P_i Konzentration hat einen negativen Einfluss auf den Fluss durch den Aromaten-Biosyntheseweg. Dies lässt sich auch aus der thermodynamischen Betrachtung ableiten, da hohe P_i -Konzentrationen das Gleichgewicht verschieben (P_i wird im Reaktionsweg mehrmals abgespalten und tritt als Nebenprodukt auf).

16.5 Thermodynamische Betrachtung

Betrachtet man die im Syntheseweg ablaufenden Reaktionen, lassen sich aus Parallelen zu anderen Reaktionen folgende Thesen ableiten:

DAHP Synthase: Die beiden Substrate sind energiereich, vor allem der Verbrauch eines energiereichen PEP legt eine thermodynamisch bevorzugte Reaktion zu DAHP nahe.

DHQ Synthase: Die Phosphatabspaltung ist thermodynamisch günstig. Zudem kommt es zum Ringschluss (Aldol-Kondensation).

DHQ Dehydratase: Hydrierungen sind allgemein bevorzugte Reaktionen (z.B. Fumarate mit $\Delta G = -3.8$). Dehydrierungen weisen meist ein leicht zu den Substraten geneigtes thermodynamisches Gleichgewicht auf.

Shikimat Dehydrogenase: Diese Reaktion ist die Reduktion eines Ketons zu einem sekundären Alkohol unter Verbrauch von NADPH. Die umgekehrte Richtung, die Oxidation eines sekundären Alkohols ist nicht bevorzugt (z.B. Malat Dehydrogenase $\Delta G^0 = +29.7$)

Shikimat Kinase: Die Übertragung von anorganischem Phosphat von ATP auf Shikimat kann als bevorzugte Richtung angenommen werden (Phosphatgruppe an S3P hat vermutlich eine niedrigere Energie).

EPSP Synthase: Gruppen-Transfer mit PEP als Donor ist vermutlich die bevorzugte Richtung.

Chorismat Synthase: Durch die Eliminierung von P_4 wird eine Doppelbindung eingeführt. Die Phosphatabspaltung ist normalerweise eine energetisch günstige Reaktion (z.B. Phosphat-Ester Hydrolyse $\Delta G = -15$).

Chorismat Muthase: Die Pyruvyl-Gruppe wird vom Enol in die stabilere Keto-Form umgesetzt. Sie ist somit bevorzugt.

Prephenat Dehydrogenase: Decarboxylierung und Aromatisierung ist bevorzugt.

Transaminase: Die Transaminase-Reaktion ist eine Gleichgewichtsreaktion.

Eine genauere Abschätzung kann durch die Berechnung der thermodynamischen Gleichgewichte erreicht werden. Für die Intermediate des Shikimat Biosynthesewegs sind kaum ΔG^0 -Werte tabelliert. Die für die Berechnung notwendigen Werte werden deshalb nach der "Group contribution" Methode von MAVROVOUNIOTIS (1991) abgeschätzt. Nach MAVROVOUNIOTIS (1991) werden die Moleküle in charakteristische Gruppen zerlegt, denen jeweils ein bestimmter ΔG^0 -Wert zugewiesen wird (s.Abb. 16.6).

| Group | ΔG^0 [kcal/mol] | Group | ΔG^0 [kcal/mol] |
|--------------|-------------------------|------------|-------------------------|
| origin | -23.6 | origin | -23.6 |
| -COOH | -72 | -COOH | -72 |
| >CO | -27.2 | Ring >C< | -12.8 |
| -CH2- | 1.7 | Ring -CH2- | 6.1 |
| >CH- | -4.8 | Ring >CH- | -2.2 |
| >CH- | -4.8 | Ring >CH- | -2.2 |
| >CH- | -4.8 | Ring >CO | -27.4 |
| -CH2- | 1.7 | Ring -CH2- | 6.1 |
| -OH | -32 | -OH(ter) | -30.5 |
| -OH | -32 | -OH(sec) | -32 |
| -OH | -32 | -OH(sec) | -32 |
| -O-PO3(prim) | -29.5 | Total | -225.5 |
| Total | -259.3 | | |

Abbildung 16.6: Zerlegung in charakteristische Gruppen der Intermediate DAHP und DHQ nach der Methode von MAVROVOUNIOTIS (1991) zur Berechnung der ΔG^0 -Werte.

Tabelle 16.2: $\Delta_r G^0$ -Werte der Reaktionsschritte von E4P und PEP zu L-Phenylalanin auf Basis der berechneten ΔG^0 -Werte nach der Methode von MAVROVOUNIOTIS (1991).

| Reaktion | Edukte | Produkte | ΔG^0 [$\frac{kJ}{Kmol}$] |
|-------------------------|------------------------------|--|------------------------------------|
| DAHP Synthase | PEP + E4P + H ₂ O | DAHP + P _i | -72.9 |
| DHQ Synthase | DAHP | DHQ + P _i | -95.5 |
| DHQ Dehydratase | DHQ | DHS + H ₂ O | -8.0 |
| Shikimat Dehydrogenase | DHS + NADPH | SHIK + NADP ⁺ | -48.8 |
| Shikimat Kinase | SHIK + ATP | S3P + ADP | -118.8 |
| EPSP Synthase | S3P + PEP | EPSP + P _i | 9.2 |
| Chorismat Synthase | EPSP | CHO + P _i | -62.0 |
| Chorismat Muthase | CHO | PPA | -147.8 |
| Prephenat Dehydrogenase | PPA | PPY + CO ₂ + H ₂ O | -103.7 |
| Transaminase | PPY + Glu | L-PHE + aKG | 0.0 |

Aus Tabelle 16.2 erkennt man, dass sich die meisten im ersten Abschnitt aufgestellten Hypothesen bestätigen. Thermodynamisch ungünstig ist nach der Abschätzung nach der Group-Contribution-Methode die EPSP Synthase-Reaktion sowie die Transaminierung. Bei diesen Reaktionsschritten müssen sich höhere Substrat-Konzentrationen (im Vergleich zu den Produkten) einstellen, um eine ausreichende Triebkraft für die Reaktion herzustellen. Leider sind für diese Reaktionsschritte bislang keine intrazellulären Konzentrationen gemessen worden.

Im linearen Syntheseweg sollten sich Konzentrationen einstellen, die dem ersten Hauptsatz entsprechend zu einem negativen ΔG -Wert führen:

$$\begin{aligned}\Delta_f G &= \sum_{i=1}^{n_P} \Delta G_{P_i}^0 - \sum_{i=1}^{n_S} \Delta G_{S_i}^0 + R T \ln \frac{\prod_{i=1}^{n_P} a_{P_i} P_i^{\gamma_{P_i}}}{\prod_{i=1}^{n_S} a_{S_i} S_i^{\gamma_{S_i}}} \\ &= \Delta_f G^0 + R T \ln \frac{\prod_{i=1}^{n_P} a_{P_i} P_i^{\gamma_{P_i}}}{\prod_{i=1}^{n_S} a_{S_i} S_i^{\gamma_{S_i}}}\end{aligned}\quad (16.5)$$

Interessanterweise lässt sich aus den Werten auch eine Hypothese zur Regulation enzymatischer Reaktionen bestätigen. Vor allem thermodynamisch bevorzugte Reaktionen werden allosterisch reguliert. Bekannt ist, dass die Eingangsreaktion, sowie die Chorismat Synthase und -Muthase durch das Endprodukt L-Phenylalanin inhibiert werden.

Die hier durchgeführte thermodynamische Analyse basiert auf geschätzten ΔG^0 -Werten. Zudem ist zu beachten, dass in der Zelle nicht Normalbedingungen herrschen. Teilweise sind für einzelne enzymatische Reaktionen des Shikimat Biosynthesewegs auch experimentell bestimmte ΔG^0 verfügbar. Im nächsten Schritt sollten dann die ΔG^0 -Werte auf intrazelluläre Bedingungen umgerechnet werden. An der ETH Zürich wird zurzeit an einem Tool (KÜMMEL ET AL., 2006) gearbeitet, dass eine Umrechnung nach der Methode von ALBERTY (2003) vornimmt.

Teil V

Metabolisch und Isotopisch Instationäre Modellierung

Kapitel 17

Simulationsstudie: Metabolisch instationäre ^{13}C Experimente

Die statistischen Untersuchungen der vorangegangenen Kapitel haben gezeigt, dass zur *in vivo* Parameterbestimmung weitere Informationen bereitgestellt werden müssen. Bei der Durchführung stationärer Stoffflussanalysen ist der Einsatz von ^{13}C markiertem Substrat sehr erfolgreich und ermöglicht eine wesentlich genauere Bestimmung der intrazellulären Stoffflüsse im Vergleich zum rein stöchiometrischen Ansatz. Die Frage, ob die Markierungsinformationen auch bei der Bestimmung von kinetischen Parametern *in vivo* Erfolg versprechend ist, wurde noch nicht detailliert untersucht.

Um die Qualität der Parameterschätzung reaktionskinetischer ^{13}C Modelle abzuschätzen, wird ein kleines Beispielsystem (Abb. 17.1) untersucht. Es werden verschiedene Annahmen über verfügbare Messdaten im Vergleich zu einem Experiment ohne ^{13}C Markierung betrachtet.

Die Experimente der Simulationsstudie werden analog zum Stimulus-Response Experiment ohne Markierung durchgeführt. Das metabolische Netzwerk wird vor dem Puls Substrat-limitiert und zum Zeitpunkt $t = 0\text{ s}$ durch die Zugabe eines hoch konzentrierten markierten Substrat-Pulses aus dem limitierten Gleichgewichtszustand ausgelenkt (Abb. 4.3).

Am Beispiel des in Abb. 17.1 dargestellten kleinen Beispielsystems wird die Komplexität und der Informationsgehalt metabolisch und isotopisch instationärer ^{13}C Experimente diskutiert. Untersucht werden unterschiedliche Szenarien bezüglich der vorhandenen Messdaten und deren Qualität.

Die Experimente der Simulationsstudie werden analog zum Stimulus-Response-Experiment ohne Markierung durchgeführt. Das metabolische Netzwerk wird vor dem Puls Substrat-limitiert und zum Zeitpunkt $t = 0\text{ s}$ durch die Zugabe eines hoch konzentrierten ^{13}C markierten Substrat-Pulses aus dem Gleichgewichtszustand ausgelenkt.

17.1 Modellaufbau

Zur Beschreibung des metabolisch und isotopisch instationären Zustands wird ein Differentialgleichungsmodell benötigt, das sowohl den Konzentrations- als auch den Anreicherungsverlauf darstellen kann (Abb. 4.1). Das Erstellen der Modellgleichungen, insbesondere

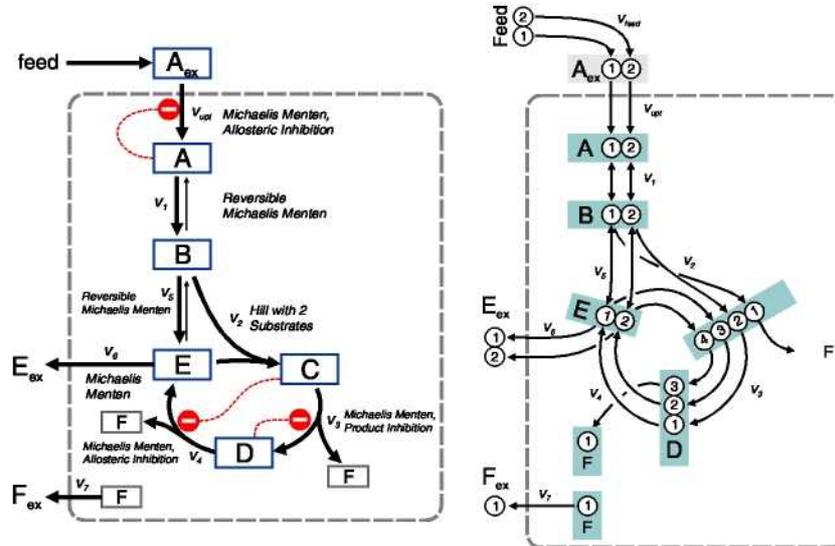


Abbildung 17.1: (a) Stoichiometrie des Beispielnetzwerk für die dynamische Simulationsstudie mit ^{13}C markiertem Substrat. (b) Das zugehörige C-Atom-Transitions-Netzwerk.

re der Massenbilanzen $\dot{\mathbf{c}}$, Kinetiken \mathbf{v} und der ^{13}C Bilanzen $\dot{\mathbf{x}}$ ist in den Abschnitten 6.2, 6.3 und 6.5 erläutert. Zusammenfassend wird das Modell aus diesen Bausteinen erstellt:

Kinetik Der Fluss wird in Abhängigkeit der vorliegenden Konzentrationen durch reaktionskinetische Terme beschrieben.

Metabolit Massenbilanzen Die Konzentrationen verändern sich im metabolisch dynamischen Zustand, die Konzentrationen werden durch Differentialgleichungen beschrieben.

Isotopomeren Massenbilanzen Wie die Konzentrationen sind auch die ^{13}C Anreicherungen instationär und werden durch Differentialgleichungen abgebildet.

Die Summe der Isotopomeren-Konzentrationen eines Pools X ergibt die Gesamtkonzentration des Pools X . Aufgrund dieser Abhängigkeit, kann der Konzentrationsvektor \mathbf{c} aus dem Modell eliminiert und durch Summen des nachfolgend definierten Vektors $\mathbf{x}_{\mathbf{c}}$ ersetzt werden. Das Differentialgleichungs-System kann folglich mit folgende Variablen vollständig beschrieben und parametrisiert werden:

Vektor der Isotopomeren-Konzentrationen $\mathbf{x}_{\mathbf{c}}$: Dieser Vektor fasst alle im Netzwerk vorkommenden Konzentrationen der Isotopomere zusammen:

$$\mathbf{x}_{\mathbf{c}} = [\mathbf{A} \mathbf{a}_{00}, \mathbf{A} \mathbf{a}_{01}, \dots, \mathbf{E} \mathbf{e}_{11}]^T = [\mathbf{A}_{00}, \dots, \mathbf{A}_{11}, \dots, \mathbf{E}_{11}]^T$$

Kinetiken: $\mathbf{v}(\mathbf{c}, \boldsymbol{\alpha})$ Mit diesem Funktionsvektor werden alle im Netzwerk vorkommenden Enzymkinetiken zusammengefasst:

$$\mathbf{v}(\mathbf{c}, \boldsymbol{\alpha}) = [v_{\text{upt}}, v_1, v_2, \dots, v_6]^T$$

Parameter: $\boldsymbol{\alpha}$ Dieser Vektor beinhaltet alle Parameter der Reaktions-Kinetiken:

$$\boldsymbol{\alpha} = [\text{upt}v_{\text{max}}, \text{upt}K_{mA}, \text{upt}K_I, \dots, v_6K_M]^T \quad (17.1)$$

Input Parameter : $\boldsymbol{\alpha}^{\text{inp}}$ Dieser Vektor beinhaltet alle experimentell vorgegebenen Parameter (z.B. Feedstrom, Input Markierungsverteilung)

$$\boldsymbol{\alpha}^{\text{inp}} = [u_{\text{feed}}, c_{\text{feed}}, \text{feed}_{00}, \dots, \text{feed}_{11}]^T$$

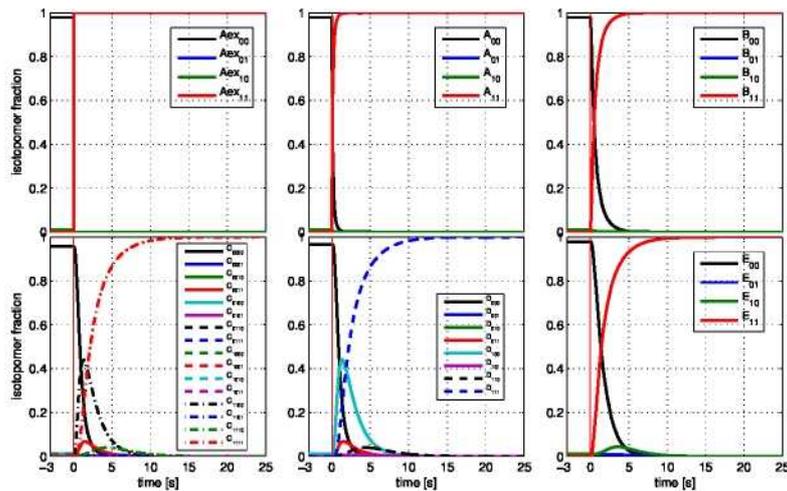


Abbildung 17.2: Simulationsergebnisse für die Isotopomerenverteilung. Extrazellulär (A_{ex}) wird durch den markierten Puls eine Markierungsanreicherung von ca. 90% vorgegeben. In den Pools C und D treten kurzzeitig Mischisotopomere z.B. C#0011 auf. Erreicht der Pool E den stationären Zustand, ist auch C vollständig markiert.

17.2 Untersuchte Szenarien

Die Ergebnisse der Simulationsstudien werden im Vergleich zu einem Referenz-Experiment ohne ^{13}C Markierung untersucht. Besonders interessant ist der Vergleich der Parameterbestimmtheiten und der Korrelationen, die anhand einer Regressionsanalyse berechnet werden können. Sie hängen wesentlich von den Annahmen über die verfügbaren Messungen

Tabelle 17.1: Messbare Informationen der verschiedenen Szenarien.

| Pool | Szenario | | |
|---------|------------------|--------------------|-----------------------|
| | S_{all} | $S_{\text{D-rel}}$ | $S_{\text{C,D-frag}}$ |
| A, B, E | | | |
| C | | | |
| D | | | |

| | |
|--|---|
| | Ohne Fragmentierung |
| | Fragmentierung in C, und C ₂ +C ₂ |
| | Gemessene Massenisotopomere |
| | Messbare Poolgröße |
| | Nicht messbare Poolgröße |

und deren Genauigkeit ab. Für die durchgeführten Studien werden folgende Bedingungen und Szenarien angenommen:

Probennahme: Über einen Zeitraum von $-5 \dots 20 \text{ s}$ können 20 Proben entnommen werden können. Vor dem Puls werden 4 Proben genommen, nach dem Puls die ersten 5 s alle 0.5 s , danach alle zwei Sekunden. (In Abschnitt 17.4 werden die Probennahmezeiten mittels Methoden der Versuchsplanung optimiert.)

Messtechnik: Mit einem LC-MS/MS Messgerät können sowohl der Markierungszustand also auch die intrazelluläre Konzentration gemessen werden. Teilweise wird auch die Messung von Fragmenten, d.h. eines Teils des C-Skeletts möglich sein. Die Anzahl der messbaren Metabolite sowie der verfügbaren Fragmente wird in den Szenarien variiert (Tab. 17.1). In Szenario S_{all} wird angenommen, dass die Poolgrößen aller Metabolit gemessen werden und alle Massenisotopomere zur Verfügung stehen. In $S_{\text{D-rel}}$ ist die Poolgröße **D** nicht bestimmbar, die Massenisotopen-Verteilung ist jedoch zugänglich. Im letzten Szenario $S_{\text{C,D-frag}}$ können alle Poolgrößen bestimmt werden, bei der Messung der Massenisotopen-Verteilung kommt es bei den Metaboliten **C** und **D** zu einer Fragmentierung in Q3. Das erste C-Atom spaltet sich ab und man erhält zusätzliche Informationen über die Markierungsverteilung (s. auch Kapitel 18).

Für die Messfehler werden folgende Annahmen getroffen:

Fehler beim Zell-Aufschluss: Bei der Probenaufbearbeitung (Quenching, Waschen, Zellaufschluss und Neutralisation, Abb. 17.3) kommt es zu zufälligen Fehlern, die in den Verdünnungsfaktor ϕ der Probe eingehen. In den Simulationsstudien wird von einem relativen Fehler von 10% ausgegangen ($\phi^*(t)$ ist die wahre Verdünnung zum Zeitpunkt t):

$$\phi(t) = \phi^*(t) \pm \sigma_\phi \quad \sigma_\phi = 0.1 \phi^*(t)$$

Dieser Fehler wirkt sich auf die Konzentrations-Messungen \mathbf{c} , nicht aber auf die Messung der relativen Markierungsverhältnisse \mathbf{x} aus, da der Verdünnungsfehler

für alle Isotopomeren-Pools gleich ausfällt. Es wird jedoch angenommen, dass der Verdünnungs-Fehler für die verschiedenen Metabolite zu einem Zeitpunkt unterschiedlich ausfallen kann (z.B. durch unterschiedliches Leaking). Zudem wird angenommen, dass zwischen den Probenahmezeitpunkten keine Korrelationen auftreten. Unter diesen Fehlerannahmen gilt dann

$$\text{Corr}(\phi(t)) = \text{diag}\{\sigma_\phi(t)^2\}$$

MS-Messunsicherheit: Der Fehler des MS-Messgerätes besteht aus einem von der Signalstärke (bzw. der Konzentration) $c_P(t)$ unabhängigen Grundrauschen σ_g und einem relativen Anteil $a c_P^*(t)$ ($c_P^*(t)$ ist die wahre Probenkonzentration):

$$c_P(t) = c_P^*(t) \pm \sigma_{MS/MS}(t) \quad \sigma_{MS/MS} = \sigma_g + a c_P^*(t) \quad (17.2)$$

Auch bei dieser Messung wird angenommen, die Fehler seien zufällig und unabhängig. Sowohl zeitlich als auch unter den verschiedenen Metaboliten gibt es keine Korrelationen (Abb. 17.3):

$$\text{Cov}(c_P(t)) = \text{diag}\{\sigma_{MS/MS}^2(t)\}$$

Für die Szenarien (S_{all} , $S_{D\text{-rel}}$ und $S_{C,D\text{-frag}}$) werden Messdaten aus den Simulationen generiert. Der Einfachheit halber wird sowohl die Verdünnung $\phi = 1$ als auch die Skalierung $\gamma = 1$ gesetzt.

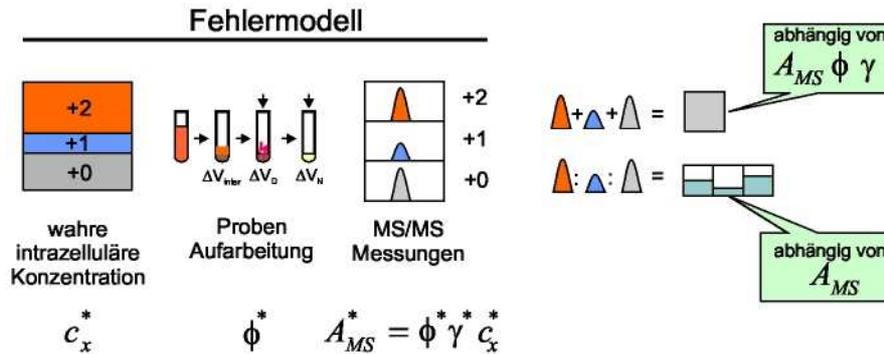


Abbildung 17.3: Fehlermodell für die MS/MS Messungen. Der Fehler der Probenaufarbeitung wirkt sich auf alle Isotopomere eines Pools gleich aus. Der MS/MS Messfehler ist für die einzelnen Isotopomere unabhängig. Damit ergibt sich für die Konzentrationsmessung eine Abhängigkeit von ϕ , γ und der gemessenen Fläche A_{MS} , während die Isotopomerenverteilung nur von A_{MS} , bzw. den Peakflächen-Verhältnissen abhängt.

Der prinzipielle Aufbau des Messmodells wird in Kapitel 18 ausführlich dargestellt. Für das laufende Beispiel ist der Verlauf der Massenisotopomeren-Konzentrationen in Abb. 17.4 für Szenario $\text{S}_{\text{C,D-frag}}$ ($\text{U-}^{13}\text{C}$ Puls) wiedergegeben.

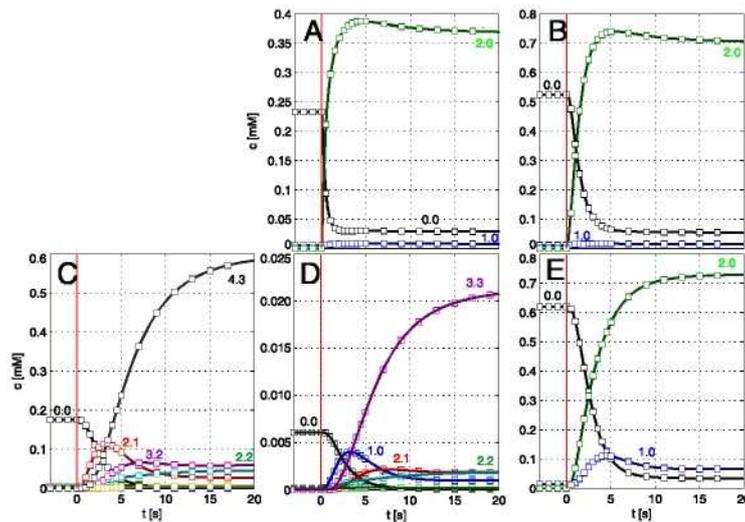


Abbildung 17.4: Aus der Simulation gewonnene Messdaten für Szenario $\text{S}_{\text{C,D-frag}}$ mit $\text{U-}^{13}\text{C}$ markiertem Substrat-Puls. Die Massenspuren sind mit Mutter- und Fragment-Ion-Masse angegeben (0.0 bedeutet $m/z + 0$ im Mutter-Ion, $m/z + 0$ im Fragment, 1.0 $m/z + 1$ im Mutter-Ion, $m/z + 0$ im Fragment).

17.3 Ergebnisse aus dem Statistischen Vergleich

Um die aus den Messdaten gewonnenen Informationen zu vergleichen, wird das D -Kriterium (s. Gl. (7.19)) der Parameterschätzung $\hat{\alpha}$ (ohne $\omega(\xi)$, s. Kap. 18) berechnet. Aus dem D -Kriterium lässt sich zum Vergleich verschiedener Experimente schnell der einfacher zu interpretierende Informations-Index I ableiten (Gl. (7.20)).

In Tab. 17.2 sind die statistischen Güten dargestellt. Zu beobachten ist eine deutliche Verbesserungen der Parameterbestimmtheit durch Zugabe eines ^{13}C markierten Substrats. Das D -Kriterium der Kovarianz-Matrix verkleinert sich von $D = 8.75 \cdot 10^{+11}$ auf $D = 6.10 \cdot 10^{-25}$ (Szenario S_{all}), d.h. im Mittel verkleinern sich die Standardabweichungen um den Faktor $I = 5.67$. Sind zusätzlich Fragmentmessungen der Metabolite C und D verfügbar, steigt der Informationsindex auf $I = 6.00$.

Interessanterweise zeigt sich bei der Analyse der Korrelationen eine Entkoppelung der Parameter verschiedener Reaktionen (Abb. 17.5) im Vergleich zu dem Referenz-Experiment ohne Markierung. Offensichtlich bewirken die zeitlichen Verzögerungen und die zusätzli-

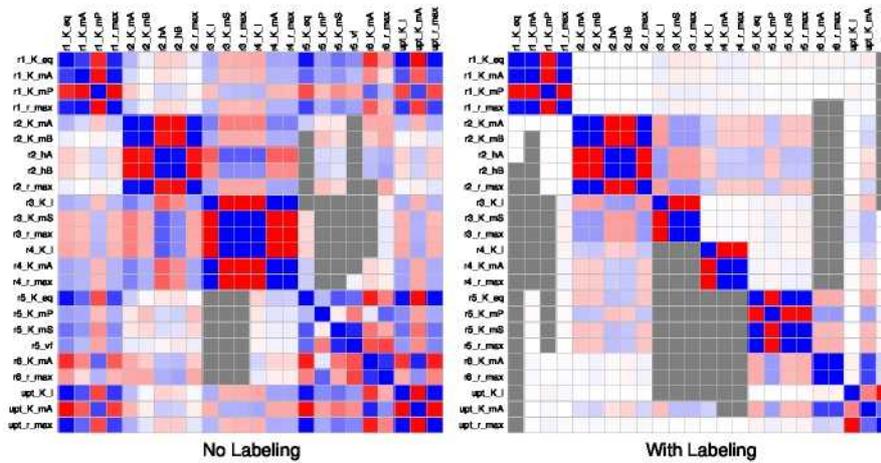


Abbildung 17.5: Parameter-Korrelationen: (a) für das ^{12}C Experiment, (b) für das ^{13}C Experiment $\text{S}_{\text{C,D-frag}}$.

chen Massenspuren, dass die Messungen unterschiedlich sensitiv auf die Parameter reagieren (Abb. 17.6).

Geringer als erwartet ist der Informationsverlust durch nicht messbare Poolgrößen. Wird angenommen, dass die Konzentration D nicht gemessen werden kann (Szenario $\text{S}_{\text{D-rel}}$), fällt der Informationsindex auf $I = 5.35$, d.h. lediglich ca. 6% niedriger als im Fall messbarer Konzentrationen. Der relativ grosse Fehler bei der Bestimmung der Konzentrationen (10%) lässt den Unterschied zu keiner Information offensichtlich sehr gering werden. Die Skalierung ω wird mit einem Fehler von ca. 1 – 18% (je nach Zeitpunkt) geschätzt. Die ω Parameter bei gemessenem \mathbf{y}_ω liegen im Bereich von 0.3 – 8%.

17.4 Versuchsplanung

Mit den Methoden der Versuchsplanung kann auf der Grundlage von angenommenen Parametern, Konzentrationen und Messfehlern das “informativste” Experiment bestimmt werden NÖH UND WIECHERT (2006). Optimierte werden dazu Design-Parameter, die vom Experimentator verändert werden können. Bei metabolisch und isotopisch dynamischen ^{13}C Experimenten können sowohl die Substrat-Zusammensetzung als auch die Probenahme-Zeitpunkte vorgegeben werden. Exemplarisch wird für Szenario $\text{S}_{\text{C,D-frag}}$ eine Optimierung der Probenahmezeitpunkte durchgeführt.

Das in der Versuchsplanung üblichste Optimierungs-Kriterium ist das D-Kriterium. Das D-Kriterium gibt das Volumen der Kovarianzmatrix wieder (Gl. 7.19). Ein niedrigeres D-Kriterium deutet somit auf eine bessere statistische Güte der Parameterschätzung hin. Die Versuchsplanung mit dem Ziel, N optimale Probenahmezeitpunkte im Intervall

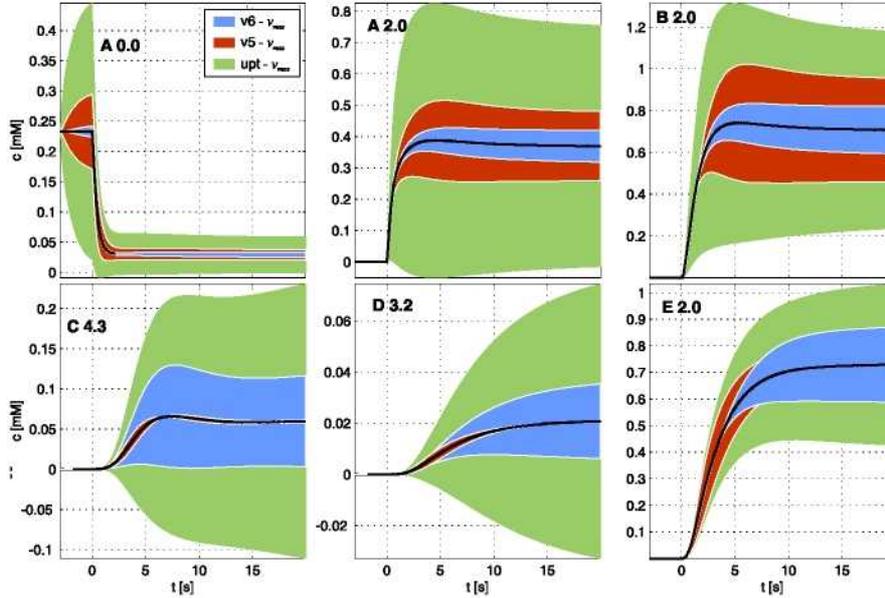


Abbildung 17.6: Vergleich der Sensitivität auf die Messdaten für die v_{max} Parameter des ersten Reaktionsschritts (v_{max}^{upt}), v_{max}^{v5} und der Ausflussreaktion v_{max}^{v6} . Bei einigen Messungen lässt sich ein Vorzeichenwechsel der Sensitivität beobachten (z.B. $\partial y_{C4.3} / \partial v_{max}^{v5}$). Eine Entkoppelung der Parameter ergibt sich v.a. dann, wenn sich der Einfluss der Parameter über die Zeit unterscheidet.

$t_{min} \leq t_1 < t_2 < \dots < t_N \leq t_{max}$ zu bestimmen wird als Minimierung des D -Kriteriums formuliert:

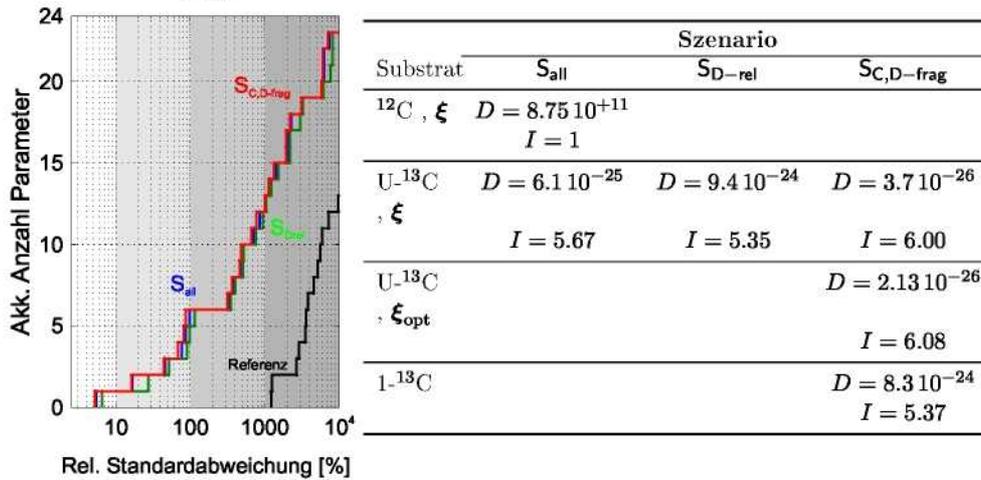
$$\xi_{opt} = \min_{\xi} D(\alpha, \xi) \quad \xi = (t_1, t_2, \dots, t_N)^T \quad (17.3)$$

Diese Minimierung kann auf verschiedenen Wegen angegangen werden. Arbeiten von (NÖH UND WIECHERT, 2006) zeigen, dass die Aufgabe iterativ effektiv gelöst werden kann. Die Schritte bestehen aus:

1. Berechnung der Fisher-Informationsmatrix $\mathbf{FI}(\hat{\alpha}, t)$ für Zeitpunkte im Intervall $[t_{min}; t_{max}]$. Im laufenden Beispiel wurden im Intervall $t = [0; 20]$ alle 0.02 s die $\mathbf{FI}(\hat{\alpha}, t)$ bestimmt.
2. Aus der $\mathbf{FI}(t)$ wird für jeden Zeitpunkt t des Rasters das D -Kriterium berechnet. Da die Kovarianzmatrix $\text{Cov}(\hat{\alpha})$ aus der Inversen der $\mathbf{FI}(\hat{\alpha})$ -Matrix berechnet werden kann, gilt:

$$D(t) = 1 / \det(\mathbf{FI}(\hat{\alpha}, t))$$

Tabelle 17.2: Vergleich der Parameterbestimmtheiten der verschiedenen Szenarien. (a) Die relativen Standardabweichungen der Parameter werden kumuliert dargestellt. Betrachtet man beispielsweise Szenario S_{all} beträgt die geringste Standardabweichung ca. 5% (erste Stufe der roten Kurve). Mit unter 100% Standardabweichung können 6 Parameter bestimmt werden (Anzahl Parameter bei 100% rel. Standardabweichung). (b) Determinanten der Parameter-Kovarianz sowie der jeweilige Informations-Index zum Referenz-Experiment ohne ^{13}C Markierung (erste Zeile).



3. Gewählt wird nun der Zeitpunkt t_1 zu dem das Kriterium minimal ist.

$$t_1 = \arg \min_t D(t), \quad {}^1D = D(t_1)$$

4. Der zweite Probenahme-Zeitpunkt t_2 wird analog bestimmt. Zu beachten ist, dass ein zur Probenahme ausreichender zeitlicher Abstand zur ersten Probe eingehalten wird. Der minimale Abstand zwischen zwei Probenahmen ist t_s (hier 0.2 s). Der zweite Probenahme-Zeitpunkt wird wieder aus der Minimierung der D -Kriteriums bestimmt.

$$t_2 = \arg \min_t D(t_1, t), \quad {}^2D = D(t_1, t_2)$$

5. nach dem selben Schema werden die Probenahmen-Zeitpunkte t_3, \dots, t_N bestimmt.

Dieses Schema wurde für das Szenario $S_{\text{C,D-frag}}$ durchgeführt. Die optimalen Probenahmezeitpunkte ξ_{opt} lauten hier:

$$\xi_{\text{opt}} = (-2.8, -0.9, -0.08, 0.32, 1.28, 1.7, 2.78, 3.56, 4.42, 5.12, 5.54, 5.96, 7.7, 11.5, 11.92, 12.34, 17.52, 18.34, 19.16, 20)^T$$

Durch die Wahl dieser Zeitpunkte verbessert sich das D -Kriterium allerdings lediglich um 1.3% auf $I = 6.08$. Diese Beobachtung lässt vermuten, dass (i) die gewählten Zeitpunkte

bereits nah am Optimum lagen oder (ii) die Identifizierbarkeit bei 20 Messpunkten nicht stark von den genauen Probezeitpunkten abhängt.

17.5 Vereinfachte ^{13}C Modelle

Bereits für das kleine Beispielsystem ergibt sich ein Modell mit 40 Differentialgleichungen und 25 Parametern. Für die Simulation des Zentralstoffwechsel würden mindestens 700 Differentialgleichungen benötigt.

Deshalb wird im Folgenden versucht, ob das Markierungsverhalten vereinfacht durch Massenisotopomere oder die Zustände markiert/unmarkiert wiedergegeben werden kann. Es soll untersucht werden, ob das Verhalten ohne die genauen Positionsinformationen gut genähert werden kann, und ob die aus dem Modell resultierenden Informationen noch ausreichend sind. Angenommen wird ein Pulsexperiment mit vollständig markiertem Substrat und die Verdrängung der vor dem Puls unmarkierten Intermediate betrachtet. Als Messtechnik wird wiederum die LC-MS/MS bevorzugt.

In der Modellierung stellt damit sich die Frage, ob die Simulation der gesamten Isotopomeren-Verteilung notwendig ist, oder das Modell durch Zusammenfassen von Isotopomeren-Pools in eine niedrigere Dimension überführt werden kann. Zwei Varianten der Zusammenfassung sind denkbar:

1. Zusammenfassen der Isotopomeren-Pools zu Massenisotopomeren-Pools,
2. Zusammenfassen in die zwei Zustände markiert und unmarkiert.

Die erste Variante könnte in Bezug auf die eingesetzt MS-Messtechnik sehr interessant sein. Die zweite Variante ist eine weitere Vereinfachung der ersten, die Zusammenfassung der $+1, \dots, +n$ markierten Massenisotope.

Um die verschiedenen Ansätze zu vergleichen und die durch die Vereinfachung hervorgerufenen Fehler werden Simulationen für vollmarkiertes und $1\text{-}^{13}\text{C}$ markiertes Substrat durchgeführt und ausgewertet.

17.5.1 Modellaufbau Massenisotopomere

Die Metabolit-Pools werden in diesem Modell in Massenisotopomeren-Pools ($^{mi}\mathbf{x}$) aufgeteilt. Ein Molekül mit n C-Atomen hat statt 2^n Isotopomeren-Pools dann nur noch $n + 1$ Massenisotopomeren-Pools. Um die Bilanzgleichungen aufzustellen werden wiederum die Zu- und Abflüsse für jeden bilanzierten Metaboliten aufgestellt. Abbildung 17.7 zeigt die Gleichungen für die drei möglichen auftretenden Reaktionstypen (unimolekular, bimolekular und die Spaltungs-Reaktion).

Unimolekulare Reaktionen Betrachtet man die Zu- und Abflüsse in den Pool B des Beispielnetzwerks, ergeben sich für die drei Massenisotopomere $+^0\text{B}$, $+^1\text{B}$ und $+^2\text{B}$ folgende Gleichungen:

$$\begin{aligned} \text{Zufluss in B | i aus A | i: } v_{in} &= \frac{+^i\text{A}}{+^0\text{A} + +^1\text{A} + +^2\text{A}} v_1^- \\ \text{Abfluss aus B | i: } v_{out} &= \frac{+^i\text{B}}{+^0\text{B} + +^1\text{B} + +^2\text{B}} + \frac{+^i\text{B}}{+^0\text{B} + +^1\text{B} + +^2\text{B}} v_5 \\ \text{MB für B+i: } \frac{d +^i\text{B}}{dt} &= v_{in} - v_{out} \end{aligned} \quad (17.4)$$

Bimolekulare Reaktionen Die Massenzuordnung für bimolekulare Reaktionen ist eindeutig (Abb. 17.7). Bei bekannter Substrat-Massenisotopomeren Verteilung kann die Verteilung des Produkts bestimmt werden. Im Beispielnetzwerk ist eine bimolekulare Synthesereaktion (B + E zu C) enthalten. Die Bildung von C lässt sich wie folgt beschreiben:

$$\begin{aligned} v_{B+E \rightarrow C}^{+0} &= \frac{+^0\text{B}}{\sum_{i=0}^{n_B} +^i\text{B}} \frac{+^0\text{E}}{\sum_{i=0}^{n_E} +^i\text{E}} v_3 \\ v_{B+E \rightarrow C}^{+1} &= \frac{+^1\text{B}}{\sum_{i=0}^{n_B} +^i\text{B}} \frac{+^0\text{E}}{\sum_{i=0}^{n_E} +^i\text{E}} v_3 + \frac{+^0\text{B}}{\sum_{i=0}^{n_B} +^i\text{B}} \frac{+^1\text{E}}{\sum_{i=0}^{n_E} +^i\text{E}} v_3 \\ &\vdots \\ v_{B+E \rightarrow C}^{+4} &= \frac{+^2\text{B}}{\sum_{i=0}^{n_B} +^i\text{B}} \frac{+^2\text{E}}{\sum_{i=0}^{n_E} +^i\text{E}} v_3 \end{aligned}$$

Spaltungs-Reaktionen Aus Abb. 17.7 erkennt man schnell, dass es für Spaltungs-Reaktionen keine eindeutige Zuordnung von Massenisotopomeren gibt. Das unmarkierte E#00 entsteht sowohl aus D#000 als auch D#001, d.h. $+^0\text{E}$ kann aus $+^0\text{D}$ und einem Anteil von $+^1\text{D}$ entstehen. Der Anteil aus $+^1\text{D}$ kann nicht bestimmt werden, wenn nur die Summe der Isotopomere D#001+D#010–D#100 betrachtet wird.

Hier kann nur noch genähert werden. Es wird angenommen, die drei Isotopomere die $+^1\text{D}$ bilden seien zu jeweils gleichen Teilen vorhanden. Für die betrachtete Beispielreaktion bedeutet dies, dass $1/3 +^1\text{D}$ zu $+^0\text{E}$ beiträgt:

$$\text{Zufluss in E+0 aus D: } v_{E+0}^- = (+^0\text{d} + 1/3 +^1\text{d}) v_4^- \quad +^k\text{d} = \frac{+^k\text{D}}{+^0\text{D} + +^1\text{D} + +^2\text{D} + +^3\text{D}} \quad (17.5)$$

Betrachtet man das einfach markierte $+^1\text{E}$ (Isotopomere E#01+E#10) ergibt sich ein Zufluss aus den Isotopomeren D#100, D#010, D#101 und D#011. Von den drei $+^1\text{D}$

tragen zwei zu ^{+1}E bei, aus dem Pool der ^{+2}D Isotopomere sind es ebenfalls zwei von drei möglichen Isotopomeren. Mit der Annahme einer Gleichverteilung ergibt sich somit:

$$\text{Zufluss in E+1 aus D: } v_{\text{E+1}}^{\vec{}} = (2/3 \text{ } ^{+1}\text{d} + 2/3 \text{ } ^{+2}\text{d}) v_4^{\vec{}} \quad \text{ } ^{+k}\text{d} = \frac{\text{ } ^{+k}\text{D}}{\text{ } ^{+0}\text{D} \text{ } ^{+1}\text{D} \text{ } ^{+2}\text{D} \text{ } ^{+3}\text{D}} \quad (17.6)$$

Diese Wahrscheinlichkeiten lassen sich leicht verallgemeinern. Für ein Molekül W mit n_W C-Atomen, dass in zwei Metabolite U und V aufgespalten wird, ergeben sich für die Bildung von $\text{U}+i_U$ $i_U = 0 \dots n_U$ die folgenden Flussanteile:

$$\begin{aligned} v_{W \rightarrow V}^{+0} &= 1 \frac{\text{ } ^{+0}\text{W}}{\text{ } ^{+0}\text{W} \text{ } ^{+1}\text{W} + \dots + \text{ } ^{+n_W}\text{W}} v + \frac{\binom{n_V}{0}}{\binom{n_W}{1}} \frac{\text{ } ^{+1}\text{W}}{\sum_{i=0}^{n_W} \text{ } ^{+i}\text{W}} v + \frac{\binom{n_V}{0}}{\binom{n_W}{2}} \frac{\text{ } ^{+2}\text{W}}{\sum_{i=0}^{n_W} \text{ } ^{+i}\text{W}} v \\ &\quad + \dots + \frac{\binom{n_V}{0}}{\binom{n_W}{n_U}} \frac{\text{ } ^{+n_U}\text{W}}{\sum_{i=0}^{n_W} \text{ } ^{+i}\text{W}} v \\ v_{W \rightarrow V}^{+i} &= \frac{\binom{n_V}{i}}{\binom{n_W}{i+0}} \frac{\text{ } ^{+i}\text{W}}{\sum_{i=0}^{n_W} \text{ } ^{+i}\text{W}} v + \frac{\binom{n_V}{i}}{\binom{n_W}{i+1}} \frac{\text{ } ^{+i+1}\text{W}}{\sum_{i=0}^{n_W} \text{ } ^{+i+1}\text{W}} v + \dots \end{aligned}$$

Betrachtet man den Zufluss in ^{+1}E ergibt sich somit:

$$v_{\text{E+1}}^{\vec{}} = \left(\frac{\binom{2}{1}}{\binom{3}{1}} \text{ } ^{+1}\text{d} + \frac{\binom{2}{1}}{\binom{3}{2}} \text{ } ^{+2}\text{d} \right) v_4^{\vec{}}$$

17.5.2 Modellaufbau mit den Zuständen markiert und unmarkiert

Die nächste Vereinfachungsstufe betrachtet nur noch die Zustände markiert oder unmarkiert. Dies ist eine grobe Vereinfachung des detaillierten Isotopomeren-Modells. Wiederum sind es Spaltungs-Reaktionen, die zu Schwierigkeiten bzw. Abweichungen vom Originalmodell führen können. Das Verhältnis beim Übergang lässt sich noch schlechter abschätzen als bei der Modellierung mit Massenisotopomeren.

Bei der Erstellung der Modellgleichungen wird deshalb angenommen, dass bei Spaltungs-Reaktionen aus einem markierten Pool zwei markierte hervorgehen. Bei bimolekularen Synthesereaktionen entsteht ein markierter Pool, wenn mindesten einer der Vorläufer markiert war.

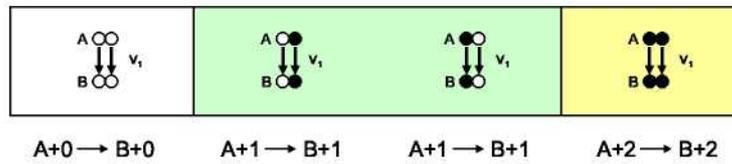
Vergleich der Simulationsergebnisse mit voll markiertem Substrat

Es ist zu erwarten, dass die Simulationsergebnisse aufgrund der notwendigen Annahmen bei Spaltungs-Reaktionen vom Original Isotopomeren-Modell abweichen. Die Positionsinformationen sind vor allem bei spezifisch markierten Substraten z.B. 1- ^{13}C Substrat von

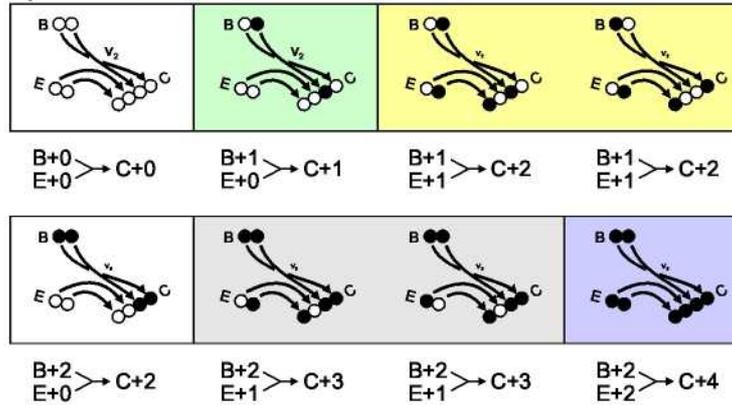
hoher Bedeutung. Wird vollmarkiertes Substrat eingesetzt, kommen spezifische Positionsmarkierungsmuster nur kurzzeitig im Übergang vor. In Abb. 17.8 sind die Simulationsergebnisse für das vollständige Isotopomeren-Modell und das Massenisotopen-Modell dargestellt.

Die nächste Vereinfachung, die Modellierung mit den Zuständen markiert und unmarkiert ist in Abb. 17.9 dargestellt. Erstaunlicherweise sind hier geringere Abweichungen zu beobachten.

a) Unimolekulare Reaktion



b) Bimolekulare Reaktion



c) Spaltungs-Reaktion

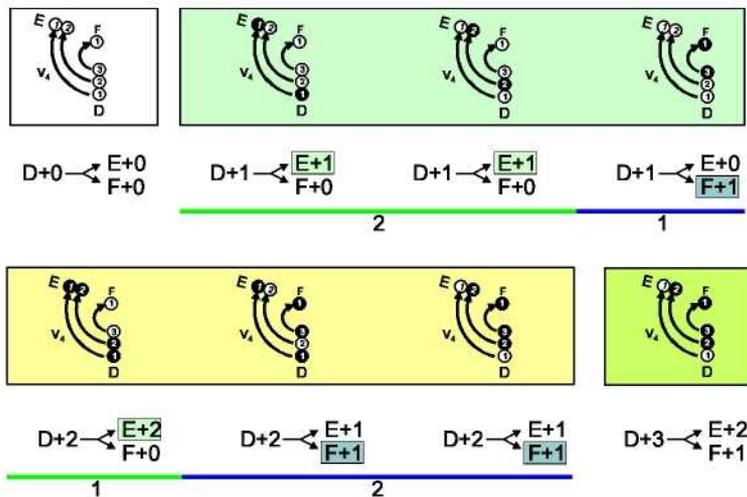


Abbildung 17.7: Betrachtung der Zuordnung für Massenisotopomere. Sowohl für uni- als auch bimolekulare Reaktionen ist die Zuordnung eindeutig. Probleme bereiten Spaltungs-Reaktionen. Hier reicht die Information der Anzahl markierter bzw. unmarkierter C-Atome nicht aus um eine eindeutige Zuordnung der Markierung der beiden entstehenden Moleküle zu treffen. In der Umsetzung wird als Näherung eine Gleichverteilung der markierten Positionen angenommen. Für die Reaktion von D nach E und F ergibt z.B. für D+1 eine Wahrscheinlichkeit von $2/3$ E+1 und $1/3$ F+1.

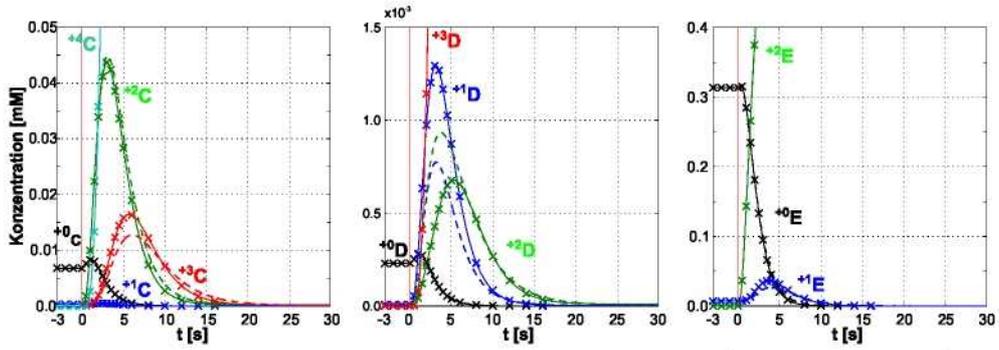


Abbildung 17.8: Abweichungen zwischen dem Isotopomerenmodell (durchgezogene Linie) und dem vereinfachten Massenisotopomeren-Modell (gestrichelte Linie). Die Metabolite A und B werden exakt wiedergegeben, bei den Anreicherungen in den Pools C, D und E sind im Übergang Abweichungen zu beobachten (vergrößerte Darstellung).

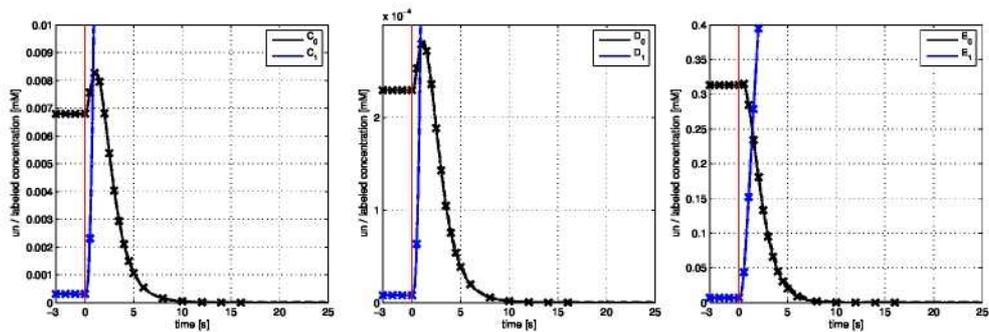


Abbildung 17.9: Das extrem vereinfachte Modell kann den Verlauf der Markierungsanreicherung relativ gut wiedergeben.

Kapitel 18

MS/MS Messmodell

Zentral für die statistischen Abschätzungen des vorangegangenen Kapitels sind die verfügbaren Messdaten. Dieses Kapitel erläutert detailliert, wie die Messgleichungen aufgestellt wurde und wie reale Messdaten, die durch natürliche Massenisotopomere verzerrt vorliegen, korrigiert und integriert werden können.

Das Messmodell von (WIECHERT ET AL., 1999) muss für den metabolisch und isotopisch instationären Zustand erweitert werden. Neben den Isotopomeren-Anteilen werden nun auch die Pool-Konzentrationen betrachtet, zudem sind die Messungen zeitabhängig. Durch diese Erweiterungen erhalten die Skalierungsfaktoren ω des allgemeinen Messmodells eine andere Bedeutung. Der Skalierungsfaktor ω des allgemeinen Messmodells stellt jetzt die Verbindung zwischen gemessener Peakfläche A und der intrazellulären Konzentration c_{intra} her. Die intrazelluläre Konzentration hängt von der Verdünnung ϕ der Probe durch den Aufschluss und dem Verhältnis γ von Peakfläche A zur Probenkonzentration c_P ab ($c_{\text{intra}} = \phi \gamma A$). Quantitative MS-Messungen sind nur möglich, wenn der entsprechende Standard vorliegt. Der Umrechnungsfaktor $c_P = \gamma A$ ist in diesem Fall bekannt. Ist kein quantitativer Standard vorhanden, sind nur die relativen Anreicherungen messbar, der Umrechnungsfaktor γ muss in der Parameter-Anpassung geschätzt werden.

Der Einfachheit halber werden in den Simulationsstudien der Verdünnungsfaktor ϕ und die Skalierung γ zusammengefasst durch ω wiedergegeben¹.

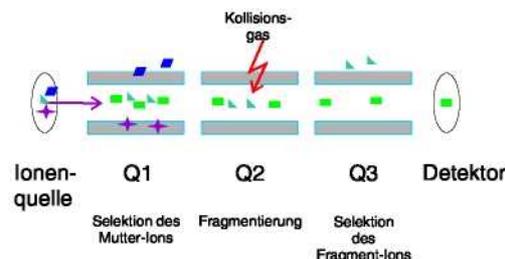


Abbildung 18.1: Schema der Trennung im Triple-Quadrupol. Im ersten Quadrupol (Q1) wird eine bestimmte Masse gefiltert, die in Q2 gelangt. Die Moleküle fallen durch Kollisionen in Fragmente. In Q3 kann eine bestimmte Fragment-Masse selektiert werden, die dann auf den Detektor fällt.

Durch den Einsatz eines Triple-Quadrupol Massenspektrometers (MS/MS, Abb. 18.1) können in Q2 einige Metabolite fragmentiert werden. Neben der Massenisotopomeren-

¹Beide Faktoren sind üblicherweise nur mit relativen Fehlern behaftet, so dass $\omega = \phi \gamma$ auch in der Fehlerrechnung zusammengefasst betrachtet werden können.

Verteilung des Gesamtmoleküls liegen dann zusätzliche Messdaten der Massenisotopomeren-Verteilung vor (Abb. 18.2).

18.1 Keine Fragmentierung

Der Aufbau des Messmodells für ein Gesamtmolekül wird am Beispiel des Metaboliten **A** mit 2 C-Atomen beschrieben. Der Metabolit kommt mit den Massen $m/z+0$ (unmarkiert) $m/z+1$ (ein markiertes C-Atom) und $m/z+2$ (vollständig markiert) vor. Gemessen werden die drei Konzentrationen der Massenisotopomere $\mathbf{y}_A(t_i) = {}^{+0}\mathbf{y}_A, \dots, {}^{+2}\mathbf{y}_A$, die sich aus Linearkombinationen der Isotopomeren-Konzentrationen \mathbf{x}_{cA} ergeben:

$$\begin{aligned} {}^{+0}\mathbf{y}_A &= \omega A_{00} \\ {}^{+1}\mathbf{y}_A &= \omega (A_{01} + A_{10}) \\ {}^{+2}\mathbf{y}_A &= \omega A_{11} \end{aligned}$$

In Matrix-Notation kann die Messung dann wie folgt dargestellt werden:

$$\mathbf{y}_A = \omega \mathbf{M}_A \mathbf{x}_{cA} \quad \mathbf{M}_A = \begin{matrix} & \begin{matrix} \text{A\#00} & \text{A\#01} & \text{A\#10} & \text{A\#11} \end{matrix} \\ \begin{matrix} {}^{+0}\mathbf{y} \\ {}^{+1}\mathbf{y} \\ {}^{+2}\mathbf{y} \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (18.1)$$

18.2 Mit Fragmentierung in Q2

Kommt es in Q2 zu einer Fragmentierung des C-Skeletts, erhält man im Vergleich zur Messung des Gesamtmoleküls zusätzliche Informationen über die Markierungsverteilung (Abb. 18.2). Im Gegensatz zur Messung ohne Fragmentierung, ergibt sich die Intensität aus der Mutter-Ion Masse (Q1) und der Fragment-Ion-Masse, die Q3 passiert. Der Aufbau des Messmodells wird anhand der Fragmentierung von Pyruvat (C3-Körper) erläutert. Im laufenden Beispiel wird das Messmodell auf den Metaboliten **D** angewendet. Bei der Messung von Pyruvat kommt es in Q2 zur Abspaltung von CO_2 , dem C_1 -Atom des C-Skeletts (Abb. 18.2). Um den Messvorgang besser beschreiben zu können, wird das C-Skelett von Pyruvat als Verbindung der beiden Isotopomere **U+V** betrachtet. In Q1 wird die Masse des Mutter-Ion, also das Gesamt-Molekül **U+V** selektiert. Für den C3-Körper Pyruvat ergeben sich die vier Massenspuren $M_{\mathbf{U+V} \ 0}$, $M_{\mathbf{U+V} \ 1}$, $M_{\mathbf{U+V} \ 2}$ und $M_{\mathbf{U+V} \ 3}$. Im zweiten Quadrupol wird die Fragmentierung durchgeführt – abgespalten wird das C_1 -Atom **U**, dass in den Massenisotopen $M_{\mathbf{U}+0}$ und $M_{\mathbf{U}+1}$ auftreten kann. Das geladene

Fragment V kann im dritten Quadrupol selektiert werden. Hier können die Massen M_V-0 , M_V-1 und M_V+2 auftreten. In Abb. 18.2 ist der Weg durch Q1, Q2 und Q3 für die alle acht Isotopomere von Pyr dargestellt. Für den C3 Körper können durch die Fragmentierung fast alle Isotopomere aufgelöst werden. Lediglich die Isotopomere Pyr#010 + Pyr#001 sowie Pyr#101 + Pyr#110 fallen auf die selben Messungen $y_{1:1}$ und $y_{2:1}$. Für längere C-Ketten verschlechtert sich das Verhältnis von auftretenden Isotopomeren zu messbaren Massenisotopomeren allerdings wieder.

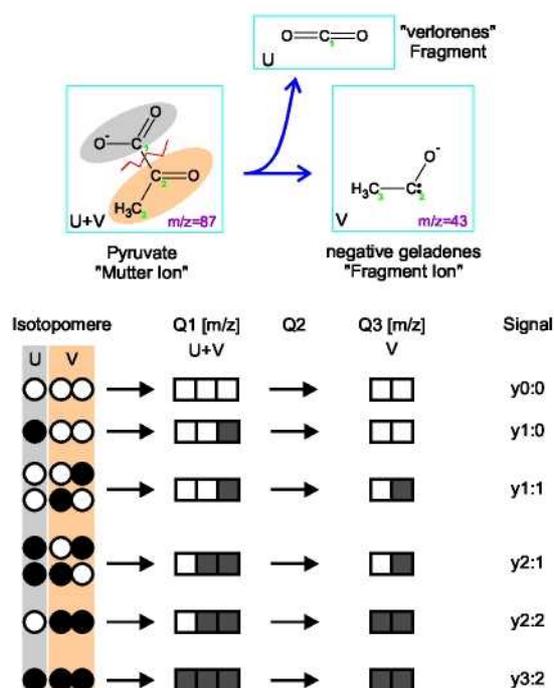


Abbildung 18.2: Fragmentierung des Metaboliten Pyruvat bei der MS/MS Messung. Im Quadrupol Q1 wird die Masse des Mutter-Ions $U+V$ gewählt, in Q2 wird Pyruvat in die Fragmente U (CO_2) und das geladene Fragment V getrennt. In Q3 kann die Masse des Fragments V gewählt werden. Durch die zweistufige Trennung ergibt sich für die Isotopomere (bzw. Summen von Isotopomeren) ein bestimmtes Messsignal (Gl. (18.1)).

Die Messmatrix für Pyruvat lässt sich direkt aus Abb. 18.2 ableiten:

$$\mathbf{y}_D = \omega \mathbf{M}_A \mathbf{x}_c$$

$$\mathbf{M}_A = \begin{matrix} & \begin{matrix} \text{D\#000} & \text{D\#001} & \text{D\#010} & \text{D\#011} & \text{D\#100} & \text{D\#101} & \text{D\#110} & \text{D\#111} \end{matrix} \\ \begin{matrix} y_{0:0} \\ y_{1:0} \\ y_{1:1} \\ y_{2:1} \\ y_{2:2} \\ y_{3:2} \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (18.2)$$

Die für die Modellierung notwendigen Messmodelle können bei bekannter Fragmentierung automatisch mit Hilfe eines MatLab-Skripts aufgestellt werden.

18.3 Kompakte Darstellung des Messmodells

Diese Gleichungen werden für alle messbaren Größen aufgestellt und können durch eine Matrix Notation zusammengefasst werden:

$$\mathbf{y}(t) = \mathbf{P}_{\mathbf{x}\mathbf{c}\mathbf{y}} \begin{pmatrix} \omega_A(t) \\ \vdots \\ \omega_D(t) \\ \omega_E(t) \end{pmatrix} \begin{pmatrix} \mathbf{M}_A & & & \\ & \ddots & & \\ & & \mathbf{M}_D & \\ & & & \mathbf{M}_E \end{pmatrix} \mathbf{x}_c(t) \quad (18.3)$$

Die Matrix $\mathbf{P}_{\mathbf{x}\mathbf{c}\mathbf{y}}$ ist eine Hilfsmatrix, die die Skalierungsfaktoren ω entsprechend der vorliegenden Messungen vervielfältigt. Die Messung wird in Messgruppen, d.h. Messungen mit gleichem Skalierungsfaktor ω zusammengefasst. Der Skalierungsfaktor ω beinhaltet bei instationären Experimenten mit nicht stationären Konzentrationen die Verdünnung durch die Probenaufbearbeitung und das das Verhältnis von Peakfläche zu Konzentration. Die Matrix $\mathbf{P}_{\mathbf{x}\mathbf{c}\mathbf{y}}$ besitzt aufgrund der Messgruppen somit eine Struktur mit "stehenden" Vektoren. Mit dieser Notation kann später das Regressionsmodell kompakt aufgestellt werden.

18.4 Massenisotopen-Korrektur für MS/MS Daten

Die Messung der Massenisotopomere \mathbf{y} des C-Skeletts wird durch das Auftreten von natürlichen Massenisotopen gebundener Atome (z.B. ^1H , ^2H ; ^{16}O , ^{17}O , ^{18}O) linear verzerrt (HELLERSTEIN UND NEESE, 1999). Um die Messdaten in der Stoffflussanalyse

einzusetzen, muss eine Korrektur für die natürlich auftretenden Massenisotope durchgeführt werden. Für MS-Daten wurde von WAHL ET AL. (2004a) eine lineare Matrix-Berechnung vorgeschlagen. Die Herangehensweise, die Verschiebung mit linearen Gleichungen darzustellen und die Korrektur, d.h. den umgekehrten Schritt der Verschiebung über eine Regression zu berechnen, ist auch auf MS/MS Daten anwendbar. Im Gegensatz zur "einfachen" MS-Messung muss die Fragmentierung berücksichtigt werden. In Abb. 18.2 sind alle für das C-Skelett von Pyruvat messbaren Massen dargestellt. Diese Messung wird durch gebundene Atome mit natürlich auftretenden Massenisotopen verschoben.

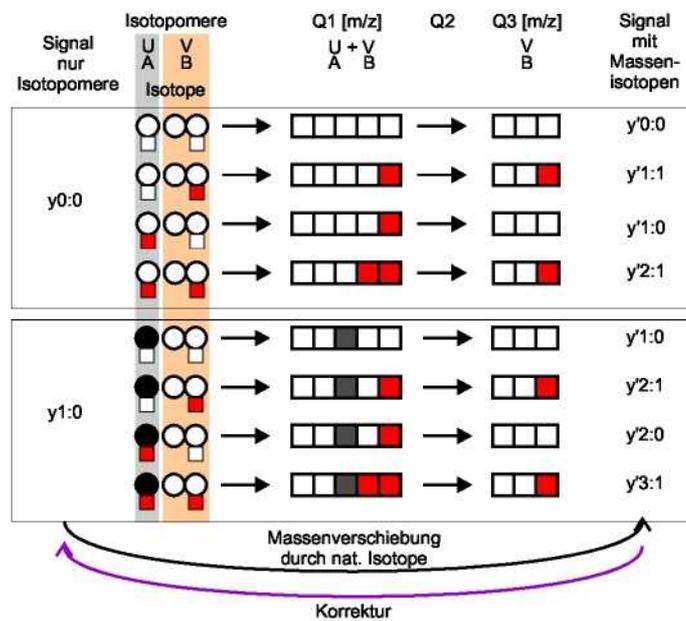


Abbildung 18.3: Schematische Darstellung der Verschiebung der Massenisotopomeren-Messung durch gebundene Atome mit natürlichen Massenisotopen.

In Abb. 18.3 wird gezeigt, wie die messbaren Signale $y_{0:0}$ und $y_{1:0}$ durch zwei gebundene Atome A und B verschoben werden. Die Wahrscheinlichkeit für das Auftreten von Massenisotopen wird mit ($^{+0}\gamma_A$ für Atom A mit der Masse A+0 $^{+1}\gamma_A$ mit der der Masse A+1 usw. angegeben). Mit den Wahrscheinlichkeiten γ erklärt man die Wahrscheinlichkeit der Verschiebung von $y_{0:0}$ auf

- $y'_{0:0}$, wenn beide gebundenen Atome A und B in den Massen A+0 und B+0 vorliegen.

$$y'_{0:0} = {}^{+0}\gamma_A {}^{+0}\gamma_B y_{0:0}$$

- $y'_{1:1}$, wenn A | 0 und B | 1 auftritt.

$$y'_{1:1} = {}^{+0}\gamma_A {}^{+1}\gamma_B y_{0:0}$$

Fragmentierung V die Ladung trägt und detektiert wird. Die Fragmente U und V sind im Prinzip austauschbar. Die zur Korrektur notwendigen Schritte können wie folgt zusammengefasst werden:

1. Zunächst wird der Messvektor \mathbf{y} des C-Skeletts aufgestellt.
2. Die gebundenen Nicht-C-Skelettatome werden den beiden späteren Fragmenten U und V zugeordnet. Nach der Methode von (LEE ET AL., 1991) werden die an U und V gebundenen Atome mit den (zusammengefassten) Isotopenvektoren $\boldsymbol{\gamma}_A$ (für U) und $\boldsymbol{\gamma}_B$ beschrieben.
3. Der Aufbau des verschobenen Messvektors \mathbf{y}' ergibt sich aus den möglichen Massenisotopen des Gesamtmoleküls $U+A+V+B$ und den jeweiligen Fragmenten $U+A$ und $V+B$ (Gl. (18.4)). Das höchst mögliche Fragment in $Q1$ ist die Summe der höchsten Massen für U , A , V und B . Im laufenden Beispiel für den $C3$ Körper $U+V$ mit den beiden Atomen A und $B = +3 + 1 + 1 = +5$. Die höchst mögliche Fragmentmessung ist $V+B$, $+2 + 1 = +3$.
4. Die Verschiebungs-Matrix für die Abbildung $\mathbf{y} \rightarrow \mathbf{y}'$ kann automatisch erstellt werden, lässt sich jedoch nicht mehr allgemein darstellen. Auffallend ist, dass im Vergleich zur einstufigen MS Messung nun Produkte der Isotopenwahrscheinlichkeiten $\boldsymbol{\gamma}_A$ und $\boldsymbol{\gamma}_B$ auftreten.
5. Nicht alle möglichen Massensignale \mathbf{y}' können auch wirklich gemessen werden. Die Zeilen nicht messbarer Signale müssen aus den Gleichungen gestrichen werden. Die Abbildung des Vektors \mathbf{y}' auf den vorliegenden Messvektor \mathbf{z} kann durch eine Matrix \mathbf{P} vorgenommen werden:

$$\mathbf{y}' \rightarrow \mathbf{z} : \quad \mathbf{z} = \mathbf{P} \mathbf{y}' \quad (18.6)$$

6. Der letzte Schritt führt die Eliminierung der natürlichen Massenisotope, die Korrektur der Messdaten durch. Der inverse Schritt von \mathbf{z} zu \mathbf{y} wird mit einer linearen Regression vollzogen:

$$\mathbf{M}\mathbf{M}_z = \mathbf{P} \mathbf{M}\mathbf{M} \quad (18.7)$$

$$\hat{\mathbf{y}} = (\mathbf{M}\mathbf{M}_z)^T \text{Cov}^{-1}(z) \mathbf{M}\mathbf{M}_z)^{-1} \mathbf{M}\mathbf{M}_z^T \text{Cov}^{-1}(z) \mathbf{z} \quad (18.8)$$

18.5 Zusammenfassung

Die Simulationsstudien zeigen, dass eine deutliche Steigerung der statistischen Güte durch die Hinzunahme von ^{13}C Markierungsmessungen erreicht werden kann. Vor allem nimmt

die Korrelation von Parametern unterschiedlicher Reaktionsschritte erheblich ab. Im Fall nicht quantifizierbarer Konzentrationen sinkt die Bestimmtheit erstaunlicherweise nur geringfügig ab. Sind zusätzliche Fragmente messbar, steigen die Parameter Bestimmtheiten an.

Ein vergleichbares Beispielmodell wurde beim Übergang von isotopisch stationären ^{13}C Stoffflussanalysen zu isotopisch instationären Experimenten im metabolisch stationären Zustand untersucht. Die Ergebnisse dieser Studie waren vergleichsweise gut auch auf größere Netzwerke übertragbar.

Die Dimensionen des zu lösenden Differentialgleichungssystems ist im metabolisch stationären und instationären Zustand identisch. NÖH ET AL. (2006) konnten zeigen, dass Systeme dieser Größe im metabolisch stationären Zustand lösbar sind. Im metabolisch instationären Zustand werden die einzelnen Differentialgleichungen komplexer, da statt konstanten Flüssen und konstanten Konzentrationen nun zeitabhängige Größen stehen. Die Anzahl benötigter Gleichungen ist identisch, während die Anzahl der Parameter wesentlich steigt. Statt der (freien) Flüsse sind es hier die kinetischen Parameter der einzelnen Reaktionsschritte.

Diese Art von Systemen lässt sich für größere Netzwerke zwar mit den heute zur Verfügung stehenden Werkzeugen noch nicht umsetzen. Der wesentliche Schritt, zeitlich veränderliche Markierungsanreicherungen zu berechnen, gelingt für den metabolisch stationären Zustand bereits. Die notwendigen Erweiterungen scheinen somit durchaus möglich.

Für das Beispielsystem wurden die Gleichungen manuell in MMT2 eingegeben. Modelle mit längeren C-Ketten und mehr Reaktionen werden sehr unübersichtlich und müssen automatisch erstellt werden.

Eine Dimensionsreduktion erscheint zunächst nicht möglich. Geprüft wurde, ob statt der Isotopomeren (bzw. Cumomeren) eine Beschreibung in Form von Massenisotopomeren möglich ist (Kapitel 17.5). Die Massenisotopomeren Modelle können uni- und bimolekulare Reaktionsschritte richtig wiedergeben, Spaltungs-Reaktionen jedoch nicht. Bei diesen Reaktionen ist die Positionsinformation der Markierungen entscheidend für die Zuordnung in den beiden entstehenden Produkten.

Bereits möglich ist die Aufarbeitung der Messdaten, insbesondere die Korrektur der Massenverteilungen aufgrund natürlich auftretender Isotope. Die Methode ist in Kapitel 18.4 ausgearbeitet und beschrieben.

Kapitel 19

Zusammenfassung, Diskussion und Ausblick

In dieser Arbeit wurden integrative Methoden zur Modellierung biochemischer Reaktionsnetzwerke dargestellt. "Integrativ" versteht sich hier als ein Vorgehen bestehend aus Modell-Bildung, -Interpretation und -Analyse im Lichte der Datenbasis, sowie der Berechnung von Vorhersagen und gezielter Versuchsplanung. Insbesondere wurde versucht, das Vorgehen systematisch darzustellen und auf die verschiedenen untersuchten Systeme anzuwenden. Untersucht wurden:

- metabolisch und isotopisch stationäre,
- metabolisch instationäre, sowie
- metabolisch und isotopisch instationäre Systeme.

In den einzelnen Kapiteln wurden die Ergebnisse bereits zusammengefasst und diskutiert. Im folgenden wird deshalb nicht mehr auf Details eingegangen.

Metabolisch und isotopisch stationäre Auswertung. Im Fall der stationären Experimente wurden makroskopische Modelle zur Auswertung der extrazellulären Konzentrationsmessungen eingesetzt und so konnten erstmalig alle während der quasi-stationären Phasen gemessenen Konzentrationen in die Fluss-schätzung einbezogen werden (Kapitel 10). Zudem konnten so die Fehler der extrazellulären Raten bestimmt werden. Zusammen mit den Markierungsmessungen im Produkt ist ein ^{13}C Modell angepasst worden (WAHL ET AL., 2004b). Zur Interpretation der gemessenen Stoffflusslagen wurden zwei neue Methoden vorgeschlagen (WAHL ET AL., 2006b). Die gemessene Stoffflusslage wird im Raum der möglichen Lösungen dargestellt und im Verhältnis zur theoretisch erreichbaren Ausbeute betrachtet. So lassen sich zwei "Maßzahlen", (Limitation Potential) und dessen Sensitivität (Constrained Limitation Sensitivity) ableiten (Kapitel 11).

Aus den ^{13}C Stoffflussanalysen und deren statistischer Untersuchung wurden Vorschläge zur Steigerung der Genauigkeit der Fluss-schätzung abgeleitet. Die von (GRÖNKE, 2007) durchgeführten stationären Experimente mit einem 3,4-CHD Produzenten zeigen, dass vor allem die Erhöhung des Markierungsanteils zu einer verbesserten Bestimmtheit der Fluss-schätzungen führt.

Metabolisch instationäre Auswertung. Metabolisch dynamische Systeme bilden den Schwerpunkt der vorliegenden Arbeit. Ausgearbeitet wurde ein systematisches Vorgehen in mehreren Stufen WAHL ET AL. (2006a) (Kapitel 12). Zunächst wird eine Korrelationsanalyse durchgeführt um den Raum möglicher Netzwerkstrukturen und regulatorischen Interaktionen zu verkleinern. Aus den Korrelationen lassen sich verschiedene Hypothesen zu dem betrachteten System ableiten (Abschnitt 13.1). Für jeden Reaktionsschritt werden mögliche Varianten aufgestellt und in einer Modellfamilie zusammengefasst. Auf Hochleistungs-Cluster-Computern werden alle Modell-Varianten an die gemessenen Daten angepasst. Statt nur einer Parameteranpassung mit einem Modell erhält der Modellierer eine Vielzahl möglicher Modelle, die die Daten ausreichend genau wiedergeben können. Dies liegt zum einen daran, dass die Messdaten stark verrauscht sind, zum anderen können bestimmte Parameter aus dem im Experiment durchlaufenen Konzentrationsbereich nicht bestimmt werden.

Betrachtet man nun die besten Modell-Varianten, fallen bestimmte, in den Modellen enthaltene Mechanismen auf. Offensichtlich sind bestimmte Reaktionsmechanismen wesentlich, um das beobachtete Verhalten wiedergeben zu können (Abschnitt 13 und 15.2).

In der statistischen Auswertung der Modellanpassungen musste jedoch festgestellt werden, dass die Parameter insgesamt sehr schlecht bestimmt sind (Abschnitt 14).

Dynamisches Modell zur parallelen Anpassung an mehrere Datensätze. Um die Messdatenbasis und die metabolische Auslenkung zu verbreitern, wurden die Modelle parallel an die beiden aufgenommenen Datensätze (*E.coli* 4pF20 und 4pF78) angepasst. Neben dem metabolischen, gibt es jetzt auch einen "genetischen Puls". Diese Modellierung bringt einige Vorteile mit sich:

- Verbesserung der statistischen Bestimmtheit durch Erweitern der Messdatenbasis,
- Möglichkeit eines Vergleichs der Enzym-Aktivitäten der beiden Stämme.

Auch für das Doppelmodell zeigt sich bei der statistischen Betrachtung ein eher ernüchterndes Bild. In der Anpassung des Doppelmodells mit der niedrigsten Fehlerquadratsumme sind nur 8% der Parameter gut (Standardabweichung $\leq 30\%$) bestimmt, 35% schlecht ($\leq 100\%$) und 37% nicht bestimmt ($\leq 1000\%$). Die restlichen 20% der Parameter lassen sich aus den Daten nicht bestimmen. Dies liegt vor allem an dem hohen Messrauschen von teilweise über 20% sowie dem im Verhältnis dazu schwachen biologischen Signal. Die Auslenkung durch den Puls erreicht kaum eine Verdoppelung der Konzentrationen.

Erschwerend kommt hinzu, dass bei dem hier eingeschlagenen Weg der fokussierten Modellbildung der Fluss durch das System kaum aus den intrazellulären Konzentrationsänderungen abgeleitet werden kann. Dies zeigt sich vor allem bei den phänomenologischen LinLog-Modellen.

Interessant sind die Ergebnisse aus dem Vergleich der Enzym-Aktivitäten (Abschnitt 15.2.1). Deutlich erkennbar ist die Übereinstimmung der gefundenen v_{max} -Werte mit dem Genotyp. Die Aktivität der Eintrittsreaktion (DAHP Synthase) ist in beiden Stämmen vergleichbar, die Aktivität der in 4pF78 überexprimierten DHQ Synthase ist im Modell ca. 80fach höher als in 4pF20.

Überprüfung der Modell-Konsistenz. Die gefundenen Modelle wurden anhand verschiedener Abschätzungen auf ihre Konsistenz überprüft. Es wurde gezeigt, dass die im stationären Zustand vor dem Puls beobachteten Konzentrationsänderungen weit über dem theoretisch maximal möglichen Fluss durch den Shikimat-Biosyntheseweg liegen. Auch nach dem Puls ist der benötigte Fluss auch im Mittel über die Zeit höher als der zur Wiedergabe der Messdaten nötige Fluss.

Eine vereinfachte thermodynamische Betrachtung bestätigt die gemessenen Konzentrationsbereiche. Für fast alle Zeitpunkte können die Reaktionen in Richtung des Produkts ablaufen, die $\Delta_f G$ sind weitestgehend < 0 . Thermodynamisch problematisch ist die Reaktion der EPSP Synthase, sie wird nur ablaufen, wenn die EPSP-Konzentrationen (nicht gemessen) extrem niedrig sind.

Optimierung des biosynthetischen Flusses anhand des Modells. Aus einer Optimierung der Enzymaktivitäten für den Steady-State Zustand wurden Vorschläge zur Steigerung des Flusses durch den Biosyntheseweg abgeleitet (Abschnitt 16.3). Bei der Optimierung des Stamms *E. coli* 4pF20 wurden zwei Enzyme vorgeschlagen: DHQ Synthase und Shikimat Kinase. Die Überexpression der DHQ Synthase ist im Stamm 4pF78 vorhanden – sie führte zu einer Steigerung der Produktivität.

Wird die Optimierung für den Stamm 4pF78 durchgeführt, werden drei Überexpressionen vorgeschlagen: DAHP Synthase, DHS Dehydrogenase und Shikimat Kinase. Diese Vorschläge decken sich ebenfalls mit experimentellen Beobachtungen. Der im stationären Zustand untersuchte Stamm 4pF81 hat statt der feedback resistenden DAHP Synthase die Wildtyp-Variante, die eine höhere Aktivität aufweist (vorausgesetzt die Tyrosin-Konzentrationen sind sehr niedrig GERIGK ET AL. (2002)). Zusätzlich wurde die Shikimat Kinase überexprimiert. Dieser Stamm ist im Vergleich zu 4pF81 ein besserer Produzent.

Diskussion und Ausblick

Die vorgestellten dynamischen Modelle sind in der Lage, den prinzipiellen Verlauf der Messdaten wiederzugeben. Bei der Überprüfung der Flüsse durch den Biosyntheseweg ergeben sich jedoch Widersprüche. Im Steady-State vor dem Puls sind sehr starke Konzentrationsänderungen zu beobachten, die auch mit dem theoretischen maximalen Fluss (bei der gemessenen Substrat-Aufnahmerate) nicht erklärt werden können (Abschnitt 16.1). Die dazu benötigten hohen Flüsse lassen sich zur Zeit nicht erklären.

Nimmt die Zelle während der Limitierung unter Umständen weitere Substrate auf? Diese Vermutung konnte anhand der vorliegenden Daten nicht untersucht werden. Dass der Fluss durch den Abbau anderer Substrate ein Vielfaches der Glucose-Aufnahmerate erreicht ist allerdings nicht anzunehmen.

Sind die starken Schwankungen dann Ergebnis der Messunsicherheit? Die Studie zum Einfluss der Spline-Glättung zeigt, dass bereits ein geringes Rauschen zu einer wesentlichen Verzerrung des Verhaltens führen kann (Abschnitt 14.4). Da das biologische Signal im Verhältnis zum Rauschen offensichtlich schwach ist, könnten die beobachteten Konzentrationsänderungen auf das Messrauschen zurückgeführt werden. Setzt man den maximal erreichbaren Fluss als Obergrenze der 1. Ableitung bei der Spline-Glättung, können Modelle mit Flüssen innerhalb der Grenzen gefunden werden. Durch die notwendige starke Glättung verringert sich der durchflossene Konzentrationsbereich jedoch so stark, dass sich die Bestimmtheit der Parameter wesentlich verschlechtert. Eine Unterscheidung verschiedener Mechanismen wird deutlich eingeschränkt, da der Informationsgehalt durch die Glättung verringert wird.

Die Probenaufbereitung von der Probennahme bis zur im MS gemessenen Probe umfasst viele Schritte. Es ist unumgänglich, dass hier zufällige Fehler auftreten. Vermutlich kommt es durch zufällige Schwankungen beim Volumen der Probennahme, sowie beim Trennen von Zellpellet und Überstand insgesamt zu größeren Verdünnungsfehlern, die sich auf alle in der Zelle vorhandenen Konzentrationen gleich auswirkt. Dadurch kommt es zu Korrelationen in den Messdaten, die Konzentrationsmessungen sind nicht mehr als unabhängig zu betrachten. Diese Korrelation muss dann in der Parameteranpassung berücksichtigt werden. Bislang ist dies nicht möglich.

Metabolisch und isotopisch instationäre Systeme. Im letzten Kapitel, der Simulationsstudie zu einem Experiment mit ^{13}C markiertem Substrat, wurde ein Messmodell mit Skalierungsfaktor eingeführt. Die Simulationsergebnisse können mit diesem Messmodell auf die MS-Rohdaten abgebildet werden. Neben dem Fehler der MS-Messung kann der Skalierungsfaktor als fehlerbehaftete (gemessene) Größe betrachtet werden. Er kann

als unabhängig für jede Konzentration oder aber für alle gemessenen Metabolite gleich angenommen werden. Es ist zu überlegen, dieses Messmodell in MMT2 zu implementieren.

Die im letzten Kapitel vorgestellte Simulationsstudie eines metabolisch und isotopisch dynamischen Modells zeigt, dass durch die Markierung der Informationsgehalt drastisch gesteigert werden kann. Die statistischen Ergebnisse lassen sich vermutlich auch auf größere Netzwerke übertragen, NÖH ET AL. (2006) untersuchte ein Netzwerk unter metabolisch stationären unter isotopisch instationären Bedingungen. Die an einem Beispielsystem NÖH UND WIECHERT (2006) vergleichbarer Größe gefundenen Eigenschaften konnten auf ein größeres Netzwerk übertragen werden.

Messtechnisch ist der Zugang zu vielen intrazellulären Metaboliten gegeben (einige Intermediate des Zentralstoffwechsels sowie verschiedener Aminosäure-Biosynthesewege, z.B. Valin und Threonin). Da die einzelnen Massenisotopomere eines Metaboliten in sehr unterschiedlichen Konzentrationen vorkommen (zu Beginn fast vollständig unmarkiert, bei Erreichen des stationären Zustands fast vollkommen markiert), wird eine Linearität des Detektors über mindestens 2-3 Größenordnungen benötigt. Dies erscheint nach den augenblicklichen Messungen möglich.

Um ein metabolisch und isotopisch instationäres Modell des gesamten Zentralstoffwechsels umzusetzen, muss ein entsprechendes automatisiertes Simulationswerkzeug eingesetzt werden. Mit entsprechend angepassten Algorithmen scheint die Simulation und Sensitivitätsanalyse möglich.

Denkbar ist auch, die fokussierte Modellierung auf Experimente mit ^{13}C Puls zu übertragen. Statt des gesamten Netzwerks wird nur der zu untersuchende Biosyntheseweg modelliert. Problematisch sind nun die Eingangssignale. Statt der vollständigen Isotopomerenverteilung sind nur die gemessenen Massenisotopomere verfügbar. Diese Linearkombination aus Isotopomeren kann für die Modellierung genutzt werden, schränkt jedoch die Auswahl der untersuchbaren Biosynthesewege ein. Betrachtet werden können nur Biosynthesewege, die keine Splitt-Reaktionen aufweisen. Für den hier untersuchten Shikimat-Biosyntheseweg könnte ein Modell in Massenisotopomer-Darstellung bis Chorismat eingesetzt werden. Im folgenden Reaktionsschritt wird ein CO_2 abgespalten. Da nur die Massenanteile, nicht jedoch die Positionen der Markierung bekannt sind, kann bei dieser Abspaltung nicht berechnet werden, wie sich die Markierung in die Produkte überträgt.

Bis zum Produkt ohne Abspaltung verlaufen die Biosynthesen von Serin, Cystein (Vorläufer PGA), Threonin, Aspartat (Vorläufer OAA), Glutamat, Glutamin, Prolin (aKG) und Alanin (Pyr). Für die aromatischen Aminosäuren (Tryptophan, Phenylalanin und Tyro-

sin) wäre eine Modellierung bis Chorismat möglich. Bei der Lysin Synthese wird im letzten Schritt CO_2 abgespalten, Großteile des Synthesewegs könnten somit bestimmt werden. Inwieweit diese Modellierung eingesetzt werden kann, wird zur Zeit anhand metabolisch stationärer, isotopisch instationärer Daten des Threonin-Biosynthesewegs ermittelt (STEIN, 2007). Diese Modellierung könnte Aufschluss über den Verlust von Metaboliten durch das Quenching geben. Mit den gemessenen Massenisotopomeren-Verteilungen des Eingangssignals und den gemessenen stationären Konzentrationen und Flüssen sollten die Markierungsanreicherungen wiedergegeben werden können. Daraus ergibt sich eine unabhängige Validierung der Konzentrationsmessungen.

Bestätigen sich die Konzentrationsmessungen, kann der nächste Schritt unternommen werden. Aus stationären Daten lassen sich weder Aussagen über die Mechanismen der Reaktionsschritte, noch über ihre möglicherweise limitierenden Auswirkungen treffen. Das zuvor angesprochene Massen-Isotopomeren-Modell kann auf den instationären Zustand durch die Hinzunahme von Kinetiken erweitert werden. Da in MMT2 noch kein allgemeines Mess-Modell mit Skalierungsfaktor verfügbar ist, müssen die Konzentrationen im richtigen Bereich liegen, die Markierungsanreicherung allein ist im dynamischen Zustand nicht ausreichend. A priori Simulationsstudien könnten zeigen, welche Informationen mit diesem Ansatz über die Parameter der Kinetiken gewonnen werden können.

Anhang A

Auswertung der ^{13}C Markierungsexperimente

A.1 Reaktionen und Ergebnisse der Flussschätzungen

Tabelle A.1: Stöchiometrie des metabolischen Netzwerks (L-Phe) für die Berechnung der optimalen Ausbeuten und "Limitation Potentials".

| Fluss | Edukte | Produkte |
|-------|--------------|--|
| pts | Gluc + PEP | → Gluc6P + Pyr |
| emp1 | Gluc6P | ⇌ Fru6P |
| emp2 | Fru6P + ATP | → FruBP + ADP |
| emp3 | FruBP | ⇌ 2 GA3P |
| emp4 | GA3P + ADP | → PEP + ATP + NADH |
| emp5 | PEP + ADP | → Pyr + ATP |
| pps | Pyr + ATP | → PEP + AMP |
| ppp1 | Gluc6P | → CO ₂ + R5P + 2 NADPH |
| ppp4 | R5P + Ery4P | ⇌ GA3P + Fru6P |
| ppp5 | 2 R5P | ⇌ Sed7P + GA3P |
| ppp6 | GA3P + Sed7P | ⇌ Ery4P + Fru6P |
| tcc1 | Pyr | → AcCoA + NADH |
| tcc2 | AcCoA + ADP | → 2 CO ₂ + ATP + NADPH + 3 NADH |

Anhang A. Auswertung der ^{13}C Markierungsexperimente

| Fluss | Edukte | Produkte |
|-------|-------------------------------|-------------------------|
| shik1 | PEP + Ery4P | → DAHP |
| shik2 | DAHP | → DHQ |
| shik3 | DHQ | ⇌ DHS |
| shik4 | DHS + NADPH | ⇌ Shik |
| shik5 | Shik + ATP | → S3P + ADP |
| shik6 | S3P + PEP | → EPSP |
| shik7 | EPSP | → CHA |
| phe | CHA + NH ₃ + NADPH | → Phe + CO ₂ |
| resp | NADH + 1.8 ADP | → 1.8 ATP |
| fut | ATP | → AMP |
| thy | NADH | → NADPH |
| e1 | AMP + ATP | → 2 ADP |

Tabelle A.2: C-Atom-Transitionen für das ^{13}C Modell. Die Notation wird in (WIECHERT UND GRAAF, 1997) erläutert.

| Reaktion | Edukt | Produkt |
|----------|------------------------------------|--------------------------------------|
| upt_1 | 1- ^{13}C Gluc #abcdef | → G6P #abcdef |
| upt_2 | Gluc #abcdef | → G6P #abcdef |
| emp1 | G6P #abcdef | → F6P #abcdef |
| emp2 | F6P #abcdef | → FBP #abcdef |
| emp3 | FBP #abcdef | → GAP + GAP #cba #def |
| emp4 | GAP #abc | → PEP/Pyr #abc |
| ppp1 | G6P #abcdef | → P5P + CO ₂ #bcdef #a |
| ppp4 | P5P + E4P #abcdef #ABCD | → GAP + F6P #cde #abABCD |
| ppp5 | P5P + P5P #abcde #ABCDE | → S7P + GAP #abABCDE #cde |
| ppp6 | GAP + S7P #abc #ABCDEFG | → E4P + F6P #DEFG #ABCabc |

A.1. Reaktionen und Ergebnisse der Fluss schätzungen

| Reaktion | Edukt | | | → | Produkt | | |
|----------|------------------|---|-----------------|---|---------------------|---|-----------|
| tcc1 | PEP/Pyr #abc | | | → | AcCoA #bc | + | CO2 #a |
| tcc2 | Mal/OAA #abcd | + | AcCoA #AB | → | Cit/ICit #dcbaBA | | |
| tcc3 | ICit #abcdef | | | → | aKG #abcef | + | CO2 #d |
| tcc4 | aKG #abcde | | | → | Succ #bcde | + | CO2 #a |
| tcc5a | Succ #abcd | | | → | Mal/OAA #abcd | | |
| tcc5b | Succ #abcd | | | → | Mal/OAA #dcba | | |
| ana1 | PEP/Pyr #abc | + | CO2 #A | → | Mal/OAA #abcA | | |
| shik1 | PEP/Pyr #abc | + | E4P #ABCD | → | DAHP #abcABCD | | |
| shik2 | DAHP #abcdefg | | | → | Shik #bcdefga | | |
| phe_A | Shik #abcdefg | + | PEP/Pyr #ABC | → | L-Phe #ABCabcdef | + | CO2 #g |
| phe_B | Shik #abcdefg | + | PEP/Pyr #ABC | → | L-Phe #ABCafedcb | + | CO2 #g |
| uracil | Mal/OAA #abcd | + | CO2 #A | → | Uracil #cbaA | + | CO2 #d |

Tabelle A.3: Geschätzte Flüsse aus der Optimierung des ^{13}C Modells für die drei Sensor Reaktor Experimente. Die Standardabweichung der Flussschätzung nimmt mit der Dauer der Fermentation zu, da das ^{13}C Signal immer stärker verdünnt wird.

| | | A | | B | | C | | A | | B | | C | |
|----------------------|-----------------------|-------|-------|-------|-------|-------|-------|--------|--------|--------|--------|--------|--------|
| | | Fluss | Abw. | Fluss | Abw. | Fluss | Abw. | Fluss | Abw. | Fluss | Abw. | Fluss | Abw. |
| Glykolyse | <i>upt1</i> | 0.84 | 0.023 | 0.67 | 0.018 | 0.56 | 0.016 | 100.0% | | 100.0% | | 100.0% | |
| | <i>upt2</i> | 2.53 | 0.023 | 2.02 | 0.018 | 1.71 | 0.016 | | | | | | |
| | <i>emp1</i> | 1.86 | 0.442 | 1.69 | 0.585 | 1.99 | 1.066 | 55.4% | 23.7% | 62.7% | 34.6% | 87.8% | 53.6% |
| | <i>emp2</i> | 2.66 | 0.149 | 2.21 | 0.196 | 2.09 | 0.356 | 79.2% | 5.6% | 82.0% | 8.9% | 92.0% | 17.1% |
| | <i>emp3</i> | 2.66 | 0.149 | 2.21 | 0.196 | 2.09 | 0.356 | 79.2% | 5.6% | 82.0% | 8.9% | 92.0% | 17.1% |
| | <i>emp4</i> | 5.43 | 0.161 | 4.45 | 0.202 | 4.09 | 0.358 | 161.3% | 3.0% | 165.2% | 4.5% | 180.1% | 8.8% |
| Pentose-Phosphat-Weg | <i>ppp1</i> | 1.50 | 0.443 | 1.01 | 0.585 | 0.28 | 1.066 | 44.6% | 29.5% | 37.3% | 58.1% | 12.2% | 384.5% |
| | <i>ppp2</i> | 0.80 | 0.295 | 0.52 | 0.390 | 0.10 | 0.711 | 23.8% | 36.9% | 19.3% | 75.0% | 4.2% | 745.1% |
| | <i>ppp3</i> | 0.70 | 0.148 | 0.49 | 0.195 | 0.18 | 0.356 | 20.8% | 21.2% | 18.0% | 40.2% | 8.0% | 195.5% |
| | <i>ppp4</i> | 0.10 | 0.148 | 0.03 | 0.195 | -0.09 | 0.355 | 3.0% | 147.6% | 1.2% | 584.1% | -3.8% | 410.9% |
| | <i>ppp5</i> | 0.70 | 0.148 | 0.49 | 0.195 | 0.18 | 0.356 | 20.8% | 21.2% | 18.0% | 40.2% | 8.0% | 195.5% |
| | <i>ppp6</i> | 0.70 | 0.148 | 0.49 | 0.195 | 0.18 | 0.356 | 20.8% | 21.2% | 18.0% | 40.2% | 8.0% | 195.5% |
| TCA | <i>tcc1</i> | 4.21 | 0.178 | 3.53 | 0.213 | 3.52 | 0.364 | 125.2% | 4.2% | 131.0% | 6.0% | 155.2% | 10.3% |
| | <i>tcc2</i> | 4.21 | 0.178 | 3.53 | 0.213 | 3.52 | 0.364 | 125.2% | 4.2% | 131.0% | 6.0% | 155.2% | 10.3% |
| | <i>tcc3</i> | 4.21 | 0.178 | 3.53 | 0.213 | 3.52 | 0.364 | 125.2% | 4.2% | 131.0% | 6.0% | 155.2% | 10.3% |
| | <i>tcc4</i> | 4.21 | 0.178 | 3.53 | 0.213 | 3.52 | 0.364 | 125.2% | 4.2% | 131.0% | 6.0% | 155.2% | 10.3% |
| | <i>tcc5a</i> | 2.11 | 0.089 | 1.77 | 0.107 | 1.76 | 0.182 | 62.6% | 4.2% | 65.5% | 6.0% | 77.6% | 10.3% |
| | <i>tcc5b</i> | 2.11 | 0.089 | 1.77 | 0.107 | 1.76 | 0.182 | 62.6% | 4.2% | 65.5% | 6.0% | 77.6% | 10.3% |
| Shikimat-Weg | <i>ana1</i> | 0.01 | 0.020 | 0.02 | 0.002 | 0.03 | 0.020 | 0.4% | 150.9% | 0.6% | 11.7% | 1.3% | 68.8% |
| | <i>shik1</i> | 0.60 | 0.024 | 0.45 | 0.022 | 0.27 | 0.020 | 17.8% | 4.1% | 16.8% | 4.8% | 11.8% | 7.5% |
| | <i>shik2</i> | 0.60 | 0.024 | 0.45 | 0.022 | 0.27 | 0.020 | 17.8% | 4.1% | 16.8% | 4.8% | 11.8% | 7.5% |
| | <i>phe_A</i> | 0.30 | 0.012 | 0.23 | 0.011 | 0.13 | 0.010 | 8.9% | 4.1% | 8.4% | 4.8% | 5.9% | 7.5% |
| | <i>phe_B</i> | 0.30 | 0.012 | 0.23 | 0.011 | 0.13 | 0.010 | 8.9% | 4.1% | 8.4% | 4.8% | 5.9% | 7.5% |
| | <i>phe</i> | 0.60 | 0.02 | 0.45 | 0.022 | 0.27 | 0.020 | 17.8% | 4.1% | 16.8% | 4.8% | 11.8% | 7.5% |
| Neben-Prod. | <i>CO₂</i> | 14.73 | 0.295 | 12.01 | 0.246 | 11.08 | 0.224 | 437.8% | 2.0% | 445.5% | 2.0% | 488.4% | 2.0% |
| | <i>uracil</i> | 0.01 | 0.020 | 0.02 | 0.002 | 0.03 | 0.020 | 0.4% | 150.9% | 0.6% | 11.7% | 1.3% | 68.8% |

A.1. Reaktionen und Ergebnisse der Flusschätzungen

Tabelle A.4: Residuen der Anpassungen, aufgeschlüsselt nach den einzelnen Messungen (die Abweichung ergibt sich aus $\mathbf{x}_{sim} - \mathbf{x}_m$). Das Experiment C weist eine hohe FQS (χ^2) auf, die über dem zulässigen Wert ($\chi_{0.995,2}^2 = 10.59$) liegt. Die Experimente A und B sind statistisch haltbar.

| | | A | | B | | C | |
|---------------------------|--------|----------|------|----------|-------------|----------|-------|
| | | Abw. | FQS | Abw. | FQS | Abw. | FQS |
| Markierungs- Messungen | C5+C9 | 0.013 | 0.28 | 0.020 | 0.26 | 0.006 | 0.01 |
| | C2 | -0.013 | 0.28 | -0.020 | 0.26 | -0.059 | 0.54 |
| Fluss- Messungen | upt1 | -0.008 | 0.11 | 0.000 | 0.00 | -0.015 | 0.84 |
| | upt2 | -0.008 | 0.11 | 0.000 | 0.00 | -0.016 | 0.89 |
| | phe | 0.014 | 0.29 | 0.001 | 0.00 | 0.039 | 3.42 |
| | CO2 | 2.413 | 5.82 | 0.185 | 0.03 | 2.403 | 23.11 |
| | uracil | 0.004 | 0.04 | 0.000 | 0.00 | 0.015 | 0.60 |
| FQS | | | 6.93 | | 0.55 | | 29.40 |

Tabelle A.5: Elementarmoden für das untersuchte L-Phenylalanin-Modell. Die höchste Produktivität stellt sich für die Mode 15 ein.

| EM | emp1 | emp2 | emp3 | emp4 | emp5 | ppp1 | ppp4 | ppp5 | ppp6 | tcc1 | tcc2 | pps | shik1-7 | phe | resp | e1 | thy | fu1 |
|----|-------------|-------------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 1 | 1.00 | 0.88 | 0.88 | 1.50 | 0.00 | 0.00 | -0.25 | 0.13 | 0.13 | 0.75 | 0.75 | 0.25 | 0.38 | 0.38 | 4.50 | 4.55 | 0.00 | 4.30 |
| 2 | 0.00 | 0.50 | 0.50 | 1.00 | 1.60 | 1.00 | 0.00 | 0.50 | 0.50 | 0.00 | 0.00 | 2.60 | 0.50 | 0.50 | 2.00 | 2.60 | -1.00 | 0.00 |
| 3 | 0.00 | 0.50 | 0.50 | 1.00 | 0.00 | 1.00 | 0.00 | 0.50 | 0.50 | 0.00 | 0.00 | 1.00 | 0.50 | 0.50 | 2.00 | 1.80 | -1.00 | 0.80 |
| 4 | 0.42 | 0.63 | 0.63 | 1.08 | 0.00 | 0.58 | -0.17 | 0.37 | 0.37 | 0.00 | 0.00 | 1.00 | 0.54 | 0.54 | 1.16 | 1.00 | -0.07 | 0.00 |
| 5 | -1.67 | 0.00 | 0.00 | 0.67 | 7.93 | 2.67 | 0.67 | 1.00 | 1.00 | 0.00 | 0.00 | 8.93 | 0.33 | 0.33 | 5.33 | 8.93 | -4.67 | 0.00 |
| 6 | -1.67 | 0.00 | 0.00 | 0.67 | 0.00 | 2.67 | 0.67 | 1.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.33 | 0.33 | 5.33 | 4.97 | -4.67 | 3.97 |
| 7 | 1.00 | 1.00 | 1.00 | 2.00 | 26.60 | 0.00 | 0.00 | 0.00 | 0.00 | 2.00 | 2.00 | 25.60 | 0.00 | 0.00 | 12.00 | 25.60 | -2.00 | 0.00 |
| 8 | 1.00 | 1.00 | 1.00 | 2.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 2.00 | 2.00 | 0.00 | 0.00 | 0.00 | 12.00 | 12.80 | -2.00 | 12.80 |
| 9 | -2.00 | 0.00 | 0.00 | 1.00 | 23.60 | 3.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 23.60 | 0.00 | 0.00 | 12.00 | 23.60 | -7.00 | 0.00 |
| 10 | -2.00 | 0.00 | 0.00 | 1.00 | 0.00 | 3.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 12.00 | 11.80 | -7.00 | 11.80 |
| 11 | 0.50 | 0.75 | 0.75 | 1.50 | 0.00 | 0.50 | 0.00 | 0.25 | 0.25 | 1.00 | 1.00 | 0.00 | 0.25 | 0.25 | 7.00 | 7.05 | -1.50 | 7.05 |
| 12 | 0.00 | 0.60 | 0.60 | 1.40 | 0.00 | 1.00 | 0.20 | 0.40 | 0.40 | 1.00 | 1.00 | 0.00 | 0.20 | 0.20 | 8.00 | 8.00 | -2.60 | 8.00 |
| 13 | 0.00 | 0.67 | 0.67 | 1.67 | 25.60 | 1.00 | 0.33 | 0.33 | 0.33 | 1.67 | 1.67 | 24.93 | 0.00 | 0.00 | 12.00 | 24.93 | -3.67 | 0.00 |
| 14 | 0.00 | 0.67 | 0.67 | 1.67 | 0.67 | 1.00 | 0.33 | 0.33 | 0.33 | 1.67 | 1.67 | 0.00 | 0.00 | 0.00 | 12.00 | 12.47 | -3.67 | 12.47 |
| 15 | 1.00 | 0.82 | 0.82 | 1.26 | 0.00 | 0.00 | -0.37 | 0.18 | 0.18 | 0.15 | 0.15 | 0.85 | 0.55 | 0.55 | 0.92 | 0.85 | 0.96 | 0.00 |
| 16 | 1.00 | 0.90 | 0.90 | 1.60 | 0.00 | 0.00 | -0.20 | 0.10 | 0.10 | 1.00 | 1.00 | 0.00 | 0.30 | 0.30 | 6.00 | 6.10 | -0.40 | 6.10 |
| 17 | 1.00 | 0.88 | 0.88 | 1.50 | 8.60 | 0.00 | -0.25 | 0.13 | 0.13 | 0.75 | 0.75 | 8.85 | 0.38 | 0.38 | 4.50 | 8.85 | 0.00 | 0.00 |
| 18 | 0.46 | 0.64 | 0.64 | 1.10 | 0.00 | 0.54 | -0.18 | 0.36 | 0.36 | 0.01 | 0.01 | 0.99 | 0.54 | 0.54 | 1.14 | 0.99 | 0.00 | 0.00 |

Anhang B

Anhang zu den Dynamischen Modellen

Tabelle B.1: In der Literatur beschriebene Mechanismen für den Shikimat Biosyntheseweg.

| Enzym | Vorgeschlagener Mechanismus | Inhibitor | Aktivator |
|-------|--|-----------|------------------|
| DAHPS | Ordered Bi-Bi, Dead-End EP-Komplex ^a | - | - |
| | $\frac{v_{max} AB}{K_{iA} K_B + K_A B + K_B A + AB + \frac{K_A B Q}{K_{iQ}} + \frac{K_i A K_B Q}{K_{iQ}}}$ | | |
| | Hill Kinetik ^b | - | - |
| | $\frac{v_{max} A^{h_A} B^{h_B}}{(A^{h_A} + K_{mA})(K_{mB} + B^{h_B})}$ | | |
| DHQS | Michaelis-Menten | - | - |
| | $\frac{v_{max}}{K_{mA} A + A}$ | | |
| | Michaelis-Menten (NAD ⁺ -abhängig) ^c | - | NAD ⁺ |
| | $\frac{v_{max} AB}{(A + K_{mA})(K_{mB} + B)}$ | | |
| | Michaelis-Menten Substrat inhibiert | DAH(P) | - |
| | $\frac{v_{max} (A/K_{mA})}{1 + \frac{A}{K_{mA}} + \frac{A^2}{K_i}}$ | | |
| DHQDH | Michaelis-Menten ^d | - | - |
| | $\frac{v_{max}}{K_{mA} A + A}$ | | |
| | reversible Michaelis-Menten ^e | - | - |
| | $\frac{v_{max} \left(A - \frac{P}{K_{eq}} \right)}{K_{mA} \left(1 + \frac{P}{K_{mP}} \right) + A}$ | | |

^aDAHPS (Tyr) aus *E.coli* (SCHONER UND HERRMANN, 1976): $K_A = 5.85 \pm 0.15 \mu\text{M}$ $K_B = 96.5 \pm 0.3 \mu\text{M}$ $K_{iA} = 1.96 \pm 0.01 \mu\text{M}$ $K_{iP} = 3.45 \pm 0.27 \text{mM}$ $K_{iQ} = 24.8 \pm 11.8 \mu\text{M}$

^bDAHPS (Trp) aus *E.coli* (AKOWSKI UND BAUERLE, 1997) $h_{A(E4P)} = 2.55 \pm 0.14$; $K_{mA} = 35 \pm 0.7$; $h_{B(PEP)} = 2.15 \pm 0.09 \mu\text{M}$; $K_{mB} = 5.3 \pm 0.1 \mu\text{M}$

^c(ROTENBERG UND SPRINSON, 1978)

^d(CHAUDHURI ET AL., 1986)

^e(DELL UND FROST, 1993)

| Enzym | Vorgeschlagener Mechanismus | Inhibitor | Aktivator |
|-------|---|-------------|-----------|
| SDH | 2 Substrat Michaelis-Menten, Produktinhibierung ^a | - | Shik |
| | $\frac{v_{max} AB}{(A+K_{mA})\left(1+\frac{P}{K_{iP}}\right)\left(K_{mB}\left(1+\frac{P}{K_{iP}}\right)+B\right)}$ | | |
| | Reversible 2 Substrat Michaelis-Menten ^b | - | - |
| | $\frac{v_{max} (AB) - \frac{PQ}{K_{eq}}}{K_{mA}\left(1+\frac{P}{K_{mP}}+A\right)K_{mB}\left(1+\frac{Q}{K_{mQ}}+B\right)}$ | | |
| SK | Michaelis-Menten, 2 Inhibitoren ^c | ADP, AMP | - |
| | $\frac{v_{max} AB}{\left(A+K_{mA}\left(1+\frac{I_a}{K_{iA}}+\frac{I_b}{K_{iB}}\right)\right)\left(K_{mB}\left(1+\frac{I_a}{K_{iA}}+\frac{I_b}{K_{iB}}\right)+B\right)}$ | | |
| | Michaelis-Menten, 1 Inhibitor | AMP | - |
| | $\frac{v_{max} AB}{\left(A+K_{mA}\left(1+\frac{I}{K_i}\right)\right)\left(K_{mB}\left(1+\frac{I}{K_i}\right)+B\right)}$ | | |
| EPSPS | Ordered Bi-Bi, Dead End E*2PEP ^d | PEP (comp.) | - |
| | $\frac{v_{max} AB}{K_{ia}K_{mB}+K_{mA}K_{mB}+K_{mB}A+AB+\frac{B^2K_{mA}}{K_i}}$ | | |
| | Ordered Bi-Bi, rapid equilibrium ^e | | |
| | $\frac{v_{max} AB}{K_{ia}K_{mB}+K_{mB}A+AB}$ | | |

^a(DELL UND FROST, 1993) *E.coli*: $K_{mA}(\text{DHS}) = 72\mu\text{M}$; $K_{mB}(\text{NADPH}) = 30\mu\text{M}$; $K_{iP} = 160\mu\text{M}$

^b(CHAUDHURI ET AL., 1987)

^c(HUANG UND NESTER, 1974) *B.subtilis*, Kein kinetischer Term gegeben, Parameter wurden aus Plots geschätzt: $K_{mA}(\text{SHIK}) = 80\mu\text{M}$; $K_{mB}(\text{ATP}) = 4.4\text{mM}$; $K_{iA}(\text{ADP}) = 3.3\text{mM}$; $K_{iB} = \text{n.a.}$

^d(GRUYS ET AL., 1992) *E.coli*: $K_{mA}(\text{S3P}) = 3.6 \pm 0.2\mu\text{M}$; $K_{mB}(\text{PEP}) = 23 \pm 2\mu\text{M}$; $K_{ia}(\text{S3P}) = 31 \pm 4\mu\text{M}$; $K_{ib}(\text{PEP}) = 200 \pm 20\mu\text{M}$; $K_{i}(\text{PEP}) = 8200 \pm 600\mu\text{M}$

^e(ANDERSON UND JOHNSON, 1990): $K_{ia}(\text{S3P}) = 7\mu\text{M}$; $K_{mB}(\text{PEP}) = 6\mu\text{M}$

Tabelle B.2: Parameteranpassung für das mechanistische Modell (Variante 585, die beste Anpassung für das Doppelmodell).

| Reaktion | Parameter | Wert | Reaktion | Parameter | Wert |
|----------|-------------------|------------|----------|----------------|-----------|
| | $A_{CNADP+NADPH}$ | 0.91mM | SDH | $A : v_{\max}$ | 20mM/s |
| | $B_{CNADP+NADPH}$ | 0.62mM | | $B : v_{\max}$ | 0.25mM/s |
| DAHPS | $A : v_{\max}$ | 12mM/s | | K_{eq} | 1.20 |
| | $B : v_{\max}$ | 12mM/s | | K_{mA} | 0.35mM |
| | K_I | 0.027mM | | K_{mB} | 0.005mM |
| | K_{iA} | 0.177mM | | K_{mP} | 15mM |
| | K_{iP} | 2mM | | K_{mQ} | 5mM |
| | K_{iQ} | 5mM | SK | $A : v_{\max}$ | 2.45mM/s |
| | K_A | 0.005mM | | $B : v_{\max}$ | 0.096mM/s |
| | K_B | 0.0754mM | | K_{IA} | 8.46mM |
| DHQS | $A : v_{\max}$ | 10mM/s | | K_{IB} | 9.27mM |
| | $B : v_{\max}$ | 0.0354mM/s | | K_{mA} | 0.20mM |
| | K_{eq} | 24.10 | | K_{mB} | 0.084mM |
| | K_{mA} | 0.0011mM | EPSPS | $A : v_{\max}$ | 10mM/s |
| | K_{mP} | 15mM | | $B : v_{\max}$ | 0.418mM/s |
| DHQDH | $A : v_{\max}$ | 2.54mM/s | | K_I | 0.495mM |
| | $B : v_{\max}$ | 1.35mM/s | | K_{iA} | 0.005mM |
| | K_{mA} | 0.96mM | | K_{mA} | 1.469mM |
| | | | | K_{mB} | 0.001mM |

Tabelle B.3: Kinetiken für den LinLog Ansatz.

| Enzym | Mechanismus | Inhibitor | Aktivator |
|---|----------------|-----------|------------------|
| DAHPS | rev. | - | - |
| $k + a \ln A + b \ln B - p \ln P - q \ln Q$ | | | |
| DHQS | rev. | - | - |
| | rev. | - | NAD ⁺ |
| $k + a \ln A - p \ln P$ | | | |
| $k + a \ln A - p \ln P + ac \ln AC$ | | | |
| DHQDH | rev. | - | - |
| $k + a \ln A - p \ln P$ | | | |
| SDH | rev. | - | - |
| $k + a \ln A + b \ln B - p \ln P - q \ln Q$ | | | |
| SK | irrev. | - | - |
| | 1 Inhibitor | AMP | - |
| | 1 Aktivator | - | PEP |
| | 1 Akt., 1 Inh. | AMP | PEP |
| $k + a \ln A + b \ln B - p \ln P - q \ln Q - i \ln I$ | | | |
| $k + a \ln A + b \ln B - p \ln P - q \ln Q + ac \ln AC$ | | | |
| $k + a \ln A + b \ln B - p \ln P - q \ln Q + ac \ln AC - i \ln I$ | | | |
| EPSPS | rev. | - | - |
| | 1 Inhibitor | ADP | - |
| | 1 Inhibitor | SHIK | - |
| $k + a \ln A + b \ln B - p \ln P - q \ln Q - i \ln I$ | | | |
| $k + a \ln A + b \ln B - p \ln P - q \ln Q - i \ln I$ | | | |

Tabelle B.4: Kinetiken für den generischen Michaelis-Menten-Ansatz.

| Enzym | Vorgeschlagener Mechanismus | Inhibitor | Aktivator |
|-------|-----------------------------|-----------|-----------|
| DAHPS | irrev. | - | - |
| | Produkt inhibiert | DAHPS | - |
| DHQS | rev. | - | - |
| | Substr. inhibiert | DAH(P) | - |
| DHQDH | rev. | - | - |
| | irrev. | - | - |
| SDH | rev. | - | - |
| | Produkt inhibiert | SHIK | - |
| SK | 2 Inhibitoren | ADP, AMP | - |
| | 1 Inhibitor | ADP | - |
| EPSPS | irrev. | - | - |
| | 1 Inhibitor | PEP | - |

B.1

Tabelle B.5: Kinetiken für das Modell zur Untersuchung der Eingangsgrößen.

| Enzym | Mechanismus | Inhibitor | Aktivator |
|-------|--|-------------|-----------|
| DAHPS | Ordered Bi-Bi, Dead-End EP-Komplex | - | - |
| | $\frac{v_{max} AB}{K_{iA} K_B + K_A B + K_B A + AB + \frac{K_A B Q}{K_{iQ}} + \frac{K_i A K_B Q}{K_{iQ}}}$ | | |
| DHQS | Michaelis-Menten | - | - |
| | $\frac{v_{max}}{K_{mA} A + A}$ | | |
| DHQDH | reversible Michaelis-Menten | - | - |
| | $\frac{v_{max} \left(A - \frac{P}{K_{eq}} \right)}{K_{mA} \left(1 + \frac{P}{K_{mP}} \right) + A}$ | | |
| SDH | 2 Substrat Michaelis-Menten, Produktinhibierung | - | Shik |
| | $\frac{v_{max} AB}{(A + K_{mA}) \left(1 + \frac{P}{K_{iP}} \right) (K_{mB} \left(1 + \frac{P}{K_{iP}} \right) + B)}$ | | |
| SK | Michaelis-Menten, 2 Inhibitoren | ADP, AMP | - |
| | $\frac{v_{max} AB}{\left(A + K_{mA} \left(1 + \frac{I_a}{K_{iA}} + \frac{I_b}{K_{iB}} \right) \right) \left(K_{mB} \left(1 + \frac{I_a}{K_{iA}} + \frac{I_b}{K_{iB}} \right) + B \right)}$ | | |
| EPSPS | Ordered Bi-Bi, Dead End E*2PEP | PEP (comp.) | - |
| | $\frac{v_{max} AB}{K_{iA} K_{mB} + K_{mA} K_{mB} + K_{mB} A + A B + \frac{B^2 K_{mA}}{K_i}}$ | | |

Tabelle B.6: Verwendete Parameterisierung.

| Reaktion | Parameter | Wert | Reaktion | Parameter | Wert | |
|----------|-------------------------|------------------|----------|------------------|------------------|----------|
| DAHPS | $C_{\text{NADP+NADPH}}$ | 0.8mM | SDH | v_{max} | 1.2mM/s | |
| | v_{max} | 1.5mM/s | | K_{mA} | 3mM | |
| | K_I | 1.22mM | | K_{mB} | 0.5mM | |
| | K_{iA} | 0.05mM | | K_I | 7mM | |
| | DHQS | K_{iP} | 0.5mM | SK | v_{max} | 0.35mM/s |
| | | K_{iQ} | 0.1mM | | K_{IA} | 8mM |
| | | K_A | 0.01mM | | K_{IB} | 0.25mM |
| | | K_B | 0.1mM | | K_{mA} | 2.5mM |
| DHQDH | | v_{max} | 0.5mM/s | K_{mB} | 0.03mM | |
| | | K_{mA} | 0.05mM | EPSPS | v_{max} | 1.75mM/s |
| DHQDH | v_{max} | 1.5mM/s | K_I | | 4mM | |
| | K_{eq} | 5 | K_{mA} | | 0.45mM | |
| | K_{mA} | 0.035mM | K_{mB} | | 0.15mM | |
| | K_{mP} | 4mM | | | | |

Anhang C

Simulation von dynamischen Markierungsexperimenten

C.1 Stöchiometrie, Kinetik und C-Atom Transitionen - Spirallus

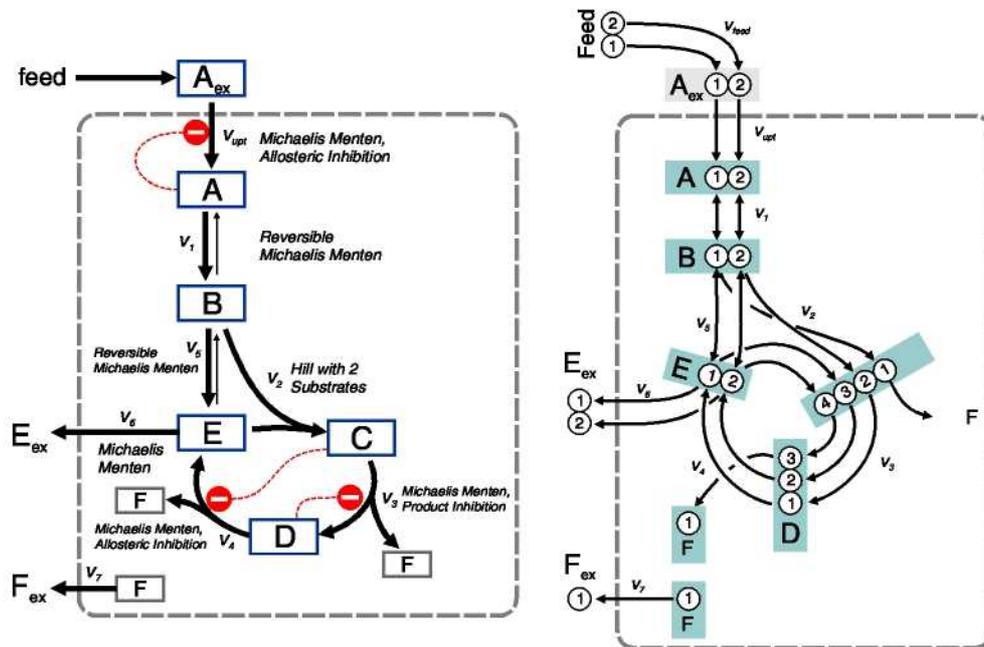


Abbildung C.1: a) Stöchiometrie des Beispielnetzwerk für die dynamische Simulationsstudie mit ^{13}C markiertem Substrat. b) Das zugehörige C-Atom-Transitions-Netzwerk

Tabelle C.1: Für das Modell verwendete Kinetiken. Die eingesetzten Parameterwerte sind in Tabelle C.1 zu finden

| Fluss | Enzym-Mechanismus | Inhibitor | Aktivator | Kinetik |
|-------|--|-----------|-----------|---|
| feed: | Zufluss in den Reaktor | | | $\frac{\dot{V}_{\text{feed}}}{V_{\text{reaktor}}} (C_{\text{feed}} - A_{\text{ex}})$ |
| upt: | Michaelis-Menten, Inhibitor | 1 | B | $\frac{v_{\text{max}} A_{\text{ex}}}{K_{mA} \left(1 + \frac{A}{K_I}\right) + A_{\text{ex}} \left(\frac{A}{K_I}\right)}$ |
| v1: | reversible Michaelis-Menten | | | $\frac{v_{\text{max}} \left(A - \frac{B}{K_{\text{eq}}}\right)}{K_{mA} \left(1 + \frac{B}{K_{mP}}\right) + A}$ |
| v2: | 2 Substrat Hill Kinetik | | | $\frac{v_{\text{max}} B^h A E^{hB}}{(K_{mA} + B^h A) (K_{mB} + E^{hB})}$ |
| v3: | Michaelis-Menten, kompetitiver Inhibitor | 1 | D | $\frac{v_{\text{max}} C}{(K_{mA} \left(1 + \frac{D}{K_I}\right) + C)}$ |
| v4: | Michaelis-Menten, kompetitiver Inhibitor | 1 | C | $\frac{v_{\text{max}} D}{(K_{mA} \left(1 + \frac{C}{K_I}\right) + D)}$ |
| v5: | reversible Michaelis-Menten | | | $\frac{v_{\text{max}} \left(B - \frac{E}{K_{\text{eq}}}\right)}{K_{mA} \left(1 + \frac{E}{K_{mP}}\right) + B}$ |
| v6: | Michaelis Menten | | | $\frac{v_{\text{max}} E}{E + K_{mA}}$ |

Tabelle C.2: C-Atom-Transitionen in der Notation von (WIECHERT UND GRAAF, 1997).

| Fluss | Edukte | | Produkte |
|-------|-------------------|---|-------------------|
| feed: | feed | > | $A_{\mathbf{ex}}$ |
| | #ab | > | #ab |
| upt: | $A_{\mathbf{ex}}$ | > | A |
| | #ab | > | #ab |
| v1: | A | > | B |
| | #ab | > | #ab |
| v2: | B + E | > | C |
| | #ab #cd | > | #abcd |
| v3: | C | > | D F |
| | #abcd | > | #bcd + #a |
| v4: | D | > | E + F |
| | #abc | > | #ab - #c |
| v5: | B | > | E |
| | #ab | > | #ab |
| v6: | E | > | $E_{\mathbf{ex}}$ |
| | #ab | > | #ab |
| v7: | F | > | $F_{\mathbf{ex}}$ |
| | #ab | > | #ab |

Tabelle C.3: Für die Beispiel-Simulation eingesetzte Parameterwerte.

| Fluss | Parameter | Wert |
|-------|---------------------|-------|
| feed | $c_{\mathbf{feed}}$ | 0.01 |
| upt | v_{max} | 1 |
| | K_I | 2 |
| | K_{mA} | 0.001 |
| v1 | v_{max} | 3 |
| | K_{eq} | 3 |
| | K_{mA} | 0.1 |
| | K_{mP} | 3 |
| v2 | v_{max} | 2.5 |
| | K_{mA} | 0.25 |
| | h_A | 2 |
| | K_{mB} | 2 |
| | h_B | 3 |
| v3 | v_{max} | 2 |
| | K_{mS} | 2 |
| | K_I | 0.05 |
| v4 | v_{max} | 3 |
| | K_{mA} | 0.1 |
| | K_I | 1 |
| v5 | v_{max} | 2 |
| | K_{eq} | 4 |
| | K_{mS} | 1 |
| | K_{mP} | 1 |
| v6 | v_{max} | 2 |
| | K_{mA} | 3 |

Literaturverzeichnis

- AHARONI, A., RIC DE VOS, C.H., VERHOEVEN, H.A., MALIEPAARD, C.A., KRUPPA, G., BINO, R., GOODENOWE, D.B. Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry. *Omic*s 2002;**6**(3):217–34.
- AKOWSKI, J.P., BAUERLE, R. Steady-state kinetics and inhibitor binding of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (tryptophan sensitive) from *Escherichia coli*. *Biochem* 1997;**36**(50):15817–15822.
- ALBERTY, R.A. *Thermodynamics of Biochemical Reactions*. John Wiley & Sons, New Jersey, 1 Auflage, 2003.
- ALMAAS, E., KOVACS, B., VICSEK, T., OLTVAI, Z.N., BARABASI, A.L. Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* 2004;**427**(6977):839–843.
- ANDERSON, K.S., JOHNSON, K.A. Kinetic and Structural-Analysis of Enzyme Intermediates - Lessons from EPSP Synthase. *Chem Rev* 1990;**90**(7):1131–1149.
- ARANDA, J.S., SALGADO, E., MUÑOZ-DIOSDADO, A. Multifractality in intracellular enzymatic reactions. *J Theor Biol* 2006;**240**(2):209–217.
- ARDREY, R.E. *Liquid chromatography - mass spectrometry : an introduction*. Analytical techniques in the sciences (AnTS). Wiley, Chichester, 2003.
- ARKIN, A., SHEN, P., ROSS, J. A Test Case of Correlation Metric Construction of a Reaction Pathway from Measurements. *Science* 1997;**277**(5330):1275–1279.
- ARPINO, P.J., GUIOCHON, G. Optimization of the Instrumental Parameters of a Combined Liquid Chromatograph-Mass Spectrometer, Coupled by an Interface for Direct Liquid Introduction. *Journal of Chromatography* 1982;**251**(2):153–164.
- ATKINSON, D. Conservation of solvent capacity. In: *Cellular Energy Metabolism and its Regulation*, 13–31. Academic Press, New York, 1977.
- BAILEY, J.E. Toward a science of metabolic engineering. *Science* 1991;**252**(5013):1668–75.
- BARDOW, A., MARQUARDT, W. Identification of diffusive transport by means of an incremental approach. *Comput Chem Eng* 2004;**28**(5):585–595.
- BENDT, A.K., BURKOVSKI, A., SCHAFFER, S., BOTT, M., FARWICK, M., HERMANN, T. Towards a phosphoproteome map of *Corynebacterium glutamicum*. *Proteomics* 2003;**3**(8):1637–46.
- BERG, J.M., TYMOCZKO, J.L., STRYER, L. *Biochemistry*. Freeman, New York, NY, 5 Auflage, 2005.

- BERRY, A. Improving Production of Aromatic Compounds in *Escherichia coli* by Metabolic Engineering. *TIBTECH* 1996;**14**:250–256.
- BEYER, H. *The Theory of Evolution Strategies*. Springer, Berlin, 2001.
- BISSWANGER, H. *Enzyme Kinetics, Principles and Methods*. WILEY-VCH, 2002.
- BLATTNER, F.R., PLUNKETT, G., R., BLOCH, C.A., PERNA, N.T., BURLAND, V., RILEY, M., COLLADO-VIDES, J., GLASNER, J.D., RODE, C.K., MAYHEW, G.F., GREGOR, J., DAVIS, N.W., KIRKPATRICK, H.A., GOEDEN, M.A., ROSE, D.J., MAU, B., SHAO, Y. The complete genome sequence of *Escherichia coli* K-12. *Science* 1997;**277**(5331):1453–74.
- BONARIUS, H.P.J., HATZIMANIKATIS, V., MEESTERS, K.P.H., GOOIJER, C.D.D., SCHMID, G., TRAMPER, J. Metabolic Flux Analysis of Hybridoma Cells in Different Culture Media Using Mass Balances. *Biotechnol Bioeng* 1996;**50**:299–318.
- BONFIGLIO, R., KING, R.C., OLAH, T.V., MERKLE, K. The effects of sample preparation methods on the variability of the electrospray ionization response for model drug compounds. *Rapid Commun Mass Spectrom* 1999;**13**(12):1175–1185.
- BONGAERTS, J., KRAMER, M., MULLER, U., RAEVEN, L., WUBBOLTS, M. Metabolic engineering for microbial production of aromatic amino acids and derived compounds. *Metab Eng* 2001;**3**(4):289–300.
- BRAND, M. The stoichiometry of proton pumping and ATP synthesis in mitochondria. *The Biochemist* 1994;**16**:20–24.
- BRIK TERNBACH, M. *Modeling Based Process Development of Fed-Batch Bioprocesses: L-Valine Production by Corynebacterium glutamicum*. Dissertation, RWTH Aachen, Aachen, 2005.
- BROCK, T.D., MADIGAN, M.T., MARTINKO, J.M., PARKER, J. *Biology of Microorganisms*. Prentice Hall, 10 Auflage, 2003.
- BROWN, G. Total cell protein concentration as an evolutionary constraint on the metabolic control distribution in cells. *J Theor Biol* 1991;**153**:195–203.
- BUCHHOLZ, A., TAKORS, R., WANDREY, C. Quantification of Intracellular Metabolites in *Escherichia coli* K12 Using Liquid Chromatographic-Electrospray Ionization Tandem Mass Spectrometric Techniques. *Anal Biochem* 2001;**295**:129–137.
- BURGARD, A.P., VAIDYARAMAN, S., MARANAS, C.D. Minimal reaction sets for *Escherichia coli* metabolism under different growth requirements and uptake environments. *Biotechnol Progr* 2001;**17**(5):791–797.
- BURNHAM, K.P., ANDERSON, D.R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer-Verlag, New York, 2002.

- CAPONIGRO, G., PARKER, R. Mechanisms and control of mRNA turnover in *Saccharomyces cerevisiae*. *Microbiol Rev* 1996;**60**(1):233–249.
- CHANDRAM, S.S., YI, J., DRATHS, K.M., VON DAENIKEN, R., WEBER, W., FROST, J.W. Phosphoenolpyruvate Availability and the Biosynthesis of Shikimic Acid. *Biotechnol Prog* 2004:.
- CHASSAGNOLE, C., NOISOMMIT-RIZZI, N., SCHMID, J.W., MAUCH, K., REUSS, M. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnol Bioeng* 2002;**79**(1):53–73.
- CHAUDHURI, S., ANTON, I.A., COGGINS, J.R. Shikimate Dehydrogenase from *Escherichia coli*. *Meth Enzymol* 1987;**142**:315–320.
- CHAUDHURI, S., LAMBERT, J.M., MCCOLL, L.A., COGGINS, J.R. Purification and Characterization of 3-Dehydroquinase from *Escherichia coli*. *Biochem J* 1986;**239**(3):699–704.
- CHEBOTAREVA, N.A., KURGANOV, B.I., LIVANOVA, N.B. Biochemical effects of molecular crowding. *Biochem-Moscow* 2004;**69**(11):1239–1251.
- CHULAVATNATOL, M., ATKINSON, D.E. Phosphoenolpyruvate Synthetase from *Escherichia coli* - Effects of Adenylate Energy Charge and Modifier Concentrations. *J Biol Chem* 1973;**248**(8):2712–2715.
- CLEGG, J., JACKSON, S. Glucose metabolism and the channeling of glycolytic intermediates in permeabilized L-929 cells. *Arch Biochem Biophys* 1990;**278**:452 – 460.
- COHEN, N.S., KUDA, A. Argininosuccinate synthetase and argininosuccinate lyase are localized around mitochondria: An immunocytochemical study. *J Cell Biochem* 1996;**60**:334–340.
- COOPER, R., KORNBERG, H. The mechanism of the phosphoenolpyruvate synthase reaction. *Biophys Acta* 1967;**141**(1):211–213.
- CORNISH-BOWDEN, A. *Fundamentals of Enzyme Kinetics*. Portland Press, London, 1995.
- COX, M., COX, M. *Multidimensional Scaling*. Chapman and Hall, 2001.
- DAUNER, M., BAILEY, J.E., SAUER, U. Metabolic Flux analysis with a Comprehensive Isotopomer Model in *Bacillus subtilis*. *Biotechnol Bioeng* 1999;**76**(2):144–156.
- DE JONG, H. Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. *J Comp Biol* 2002;**9**(1):67–103.
- DE KONIG, W., VAN DAM, K. A method for the determination of changes of glycolytic metabolites in yeast in a subsecond times scale using extraction at neutral pH. *Anal Biochem* 1999;**204**:118–123.

- DEGENRING, D. *Erstellung und Validierung mechanistischer Modelle für den mikrobiellen Stoffwechsel zur Auswertung von Substrat-Puls-Experimenten*. Dissertation, Universität Rostock, Rostock, 2004.
- DELGADO, J., LIAO, J.C. Inverse flux analysis for reduction of acetate excretion in *Escherichia coli*. *Biotechnol Progr* 1997;**13**(4):361–367.
- DELL, K.A., FROST, J.W. Identification and Removal of Impediments to Biocatalytic Synthesis of Aromatics from D-Glucose - Rate-Limiting Enzymes in the Common Pathway of Aromatic Amino-Acid Biosynthesis. *JACS* 1993;**115**(24):11581–11589.
- DEUFELHARD, P., BORNEMANN, F. *Scientific Computing with Ordinary Differential Equations*. Springer-Verlag, New York, 2002.
- DEUFLHARD, P. *Newton Methods for Highly Nonlinear Problems: Theory, Algorithms, Code*. Academic Press, 1995.
- DIERCKX. An algorithm for smoothing, differentiating and integration of experimental data using spline functions. *J Comput Appl Math* 1975;**1**:165–184.
- DOUTHEIL, J. Systems Biology - A great challenge for the european research area. 2002.
- DRATHS, K.M., POMPLIANO, D.L., CONLEY, D.L., FROST, J.W., BERRY, A., DISBROW, G.L., STAVERSKY, R.J., LIEVENSE, J.C. Biocatalytic Synthesis of Aromatics from D-Glucose: The Role of Transketolase. *J Am Chem Soc* 1992;**114**:3956–3962.
- EDWARDS, J.S., IBARRA, R.U., PALSSON, B.O. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat Biotechnol* 2001;**19**(2):125–130.
- EDWARDS, J.S., PALSSON, B.O. Robustness Analysis of the *Escherichia Coli* Metabolic Network. *Biotechnol Progr* 2000;**16**(6):927–939.
- EISENBERG, D., MARCOTTE, E., XENARIOS, I., YATES, T. Protein function in the post-genomic era. *Nature* 2000;**405**:823–826.
- EL MASSAOUDI, M., SPELTHAHN, J., DRYSCH, A., TAKORS, R. Production process monitoring by serial mapping of microbial carbon flux distributions using a novel sensor reactor approach: I - Sensor reactor system. *Metab Eng* 2003;**5**:86–95.
- FANG, S., PUTHENPURA, S. *Linear Optimization and Extensions: Theory and Algorithms*, Kapitel 3.7. Prentice-Hall, Englewood Cliffs, NJ, 1993:.
- FISCHER, E., SAUER, U. Large-scale in vivo fluxes reveal rigidity and suboptimal performance of *B. subtilis* metabolism. *Nature Genetics* 2005;**37**(6):636–640.
- FISCHER, E., SAUER, U. Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS. *Eur J Biochem* 2003;**270**:880–891.

- FLORES, N., XIAO, J., BERRY, A., BOLIVAR, F., VALLE, F. Pathway engineering for the production of aromatic compounds in *Escherichia coli*. *Nat Biotechnol* 1996: **14**(5):620–3.
- FREDRICKSON, A.G., MCGEE, R.D.I., TSCHUCHIYA, A.M. Mathematical Models in Fermentation Processes. *Adv Appl Microbiol* 1970:**23**:419–429.
- FÜRSTE, J.P., PANSEGRAU, W., FRANK, R. AND BLÖCKER, H., SCHOLZ, P., BAGDASARIAN, M., LANKA, E. Molecular Cloning of the Plasmid RP4 Primase Region in a multi-host-range tacP Expression Vector. *Gene* 1986:**48**:119–131.
- GAERTNER, F.H. Purification of two multienzyme complexes in the aromatic-tryptophan pathway of *Neurospora crassa*. *Arch Biochem Biophys* 1972:**151**(1):277–84.
- GERIGK, M., BUJNICKI, R., GANPO-NKWENKWA, E., BONGAERTS, J., SPRENGER, G., TAKORS, R. Process Control for Enhanced L-Phenylalanine Production Using Different Recombinant *Escherichia coli* Strains. *Biotechnol Bioeng* 2002: **80**(7):746–754.
- GERSHON, N., PORTER, K., TRUS, B. The cytoplasmic matrix: Its volume and surface area and the diffusion of molecules through it. *Proc Natl Acad Sci* 1985:**82**:5030–5034.
- GOEL, A., FERRANCE, J., JEONG, J., ATAAL, M. Analysis of Metabolic Fluxes in Batch and Continuous Cultures of *Bacillus subtilis*. *Biotechnol Bioeng* 1993:**42**:686–696.
- GONZALES, B., FRANCOIS, J., RENAUD, M. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast* 1997:**13**:1347–1356.
- GOSSET, G., YONG-XIAO, J., BERRY, A. A Direct Comparison of Approaches for Increasing Carbon Flow to Aromatic Biosynthesis in *Escherichia coli*. *J Indust Microbiol* 1996:**17**(47-52).
- GRAAF, A.A.D. Metabolic Flux Analysis of *Corynebacterium glutamicum*. In: K. SCHÜGERL, K. BELLGARDT, Herausgeber, *Bioreaction Engineering - Modeling and Control*, Chapter 4.3. Springer, 2000:.
- GRUYS, K.J., WALKER, M.C., SIKORSKI, J.A. Substrate Synergism and the Steady-State Kinetic Reaction-Mechanism for EPSP Synthase from *Escherichia coli*. *Biochemistry* 1992:**31**(24):5534–5544.
- GRÖNKE, K. *Further development and optimization of the Sensor Reactor for serial ¹³C-based metabolic flux analysis*. Dissertation, Universität Siegen, 2007.
- HAJJAJ, H., BLANC, P.J., GOMA, G., FRANCOIS, J.M. Sampling techniques and comparative extraction procedures for quantitative determination of intra- and extracellular metabolites in filamentous fungi. *FEMS Microbiol Lett* 1998:**164**:195–200.
- HASLAM, E. *Shikimic acid: metabolism and metabolites*. Wiley, Chichester, 1993.

- HAUNSCHILD, M.D. *MMT2 (v4.0) - Metabolic Modeling Tool Tutorial & Reference Manual*. Department of Simulation, University of Siegen, Siegen, 2005.
- HAUNSCHILD, M.D. *Metabolische Stimulus-Response-Experimente: Modellierung, Simulation und Auswertung*. Dissertation, Universität Siegen, Siegen, 2006.
- HAUNSCHILD, M.D., WAHL, S.A., FREISLEBEN, B., WIECHERT, W. A general framework for large scale model selection. *Optim Method Softw* 2006;**21**:901–917.
- HAUNSCHILD, M.D., WAHL, S.A., VON LIERES, E., QELI, E., FREISLEBEN, B., TAKORS, R., WIECHERT, W. MMT 2: Supporting the Modeling Process for Rapid Sampling Experiments. In: J. LIAO, Herausgeber, *Metabolic Engineering V, Lake Tahoe, California, September 2004*. 2004: .
- HEIJNEN, J.J. Approximative Kinetic Formats used in Metabolic Network Modeling. *Biotechnol Bioeng* 2005;**91**(5):534–545.
- HEINRICH, R., RAPOPORT, S.M., RAPOPORT, T.A. Metabolic regulation and mathematical models. *Prog Biophys Mol Biol* 1977;**32**(1):1–82.
- HEINRICH, R., SCHUSTER, S. The modeling of metabolic systems. Structure, control and optimality. *Biosystems* 1998;**47**(1-2):61–77.
- HELLERSTEIN, M.K., NEESE, R.A. Mass isotopomer distribution analysis at eight years: theoretical, analytic, and experimental considerations. *Am J Physiol* 1999;**276**(6):1146–1170.
- HINDMARSH, A.C. *ODEPACK, A Systematized Collection of ODE Solvers*, Band 1, 55–64. IMACs Transactions on Scientific Computation, Amsterdam, 1983:.
- HOFMANN, E., KOPPERSCHLAGER, G. Phosphofruktokinase from Yeast. *Meth Enzymol* 1982;**90**:49–60.
- HOPFGARTNER, G., VAREGIO, E., TSCHAPPAT, V., GRIVET, C., BOURGOGNE, E., LEUTHOLD, L.A. Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules. *J Mass Spectrom* 2004;**39**:845–855.
- HUANG, L., NESTER, E. Regulation of Aromatic Amino-Acid Biosynthesis in *Bacillus subtilis* 168 - Purification, Characterization, and Subunit Structure of Bifunctional Enzyme 3-Deoxy-D-Arabinoseheptulosonate 7-Phosphate Synthetase-Chorismate Mutase. *J Biol Chem* 1974;**249**(14):4467–4472.
- HUCKA ET AL. The system biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinf* 2003;**19**:524–531.
- ISERMANN, N., WEITZEL, M., WIECHERT, W. Kleene's Theorem and the Solution of Metabolic Carbon Labeling Systems. In: R. GIEGERICH, J. STOYE, Herausgeber, *German Conference on Bioinformatics 2004, Bielefeld*, Band 53. Springer, Bielefeld, 2004: 75–84.

- JAIN, A.K., MURTY, M.N., FLYNN, P.J. Data clustering: a review. *ACM Comput Surv* 1999;**31**(3):264–323.
- JOSSEK, R., BONGAERTS, J., SPRENGER, G.A. Characterization of a new feedback-resistant 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase AroF of *Escherichia coli*. *FEMS Microbiol Letters* 2001;**202**(1):145–148.
- KACSER, H., BURNS, J.A. The control of flux. *Symp Soc Exp Biol* 1973;**27**:65–104.
- KANEHISA, M., GOTO, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucl Acids Res* 2000;**28**(1):27–30.
- KARP, P.D., RILEY, M., SAIER, M., PAULSEN, I.T., PALEY, S.M., PELLEGRINI-TOOLE, A. The EcoCyc and MetaCyc databases. *Nucl Acids Res* 2000;**28**(1):56–59.
- KESELER, I.M., COLLADO-VIDES, J., GAMA-CASTRO, S., INGRAHAM, J., PALEY, S., PAULSEN, I.T., PERALTA-GIL, M., KARP, P.D. EcoCyc: a comprehensive database resource for *Escherichia coli*. *Nucleic Acids Res* 2005;**33**(Database issue):D334–7.
- KITANO, H. Computational systems biology. *Nature* 2002;**420**:206 – 210.
- KLAMT, S., STELLING, J., GINKEL, M., GILLES, E.D. FluxAnalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps. *Bioinformatics* 2003;**19**(2):261–269.
- KOOLMAN, J., RÖHM, K.H. *Taschenatlas der Biochemie*. Georg Thieme Verlag, Stuttgart, 2. Auflage, 1998.
- KOPPENSCHLÄGER, G., AUGUSTIN, H.W. Fehlermöglichkeiten bei der Bestimmung von Metabolitgehalten in Hefezellen. *Experientia* 1967;**23**:623–624.
- KÜMMEL, A., PANKE, S., HEINEMANN, M. Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolome data. *Mol Syst Biol* 2006: Doi:10.1038/msb4100074.
- LANGE, H.C., EMAN, M., VAN ZUIJLEN, G., VISSER, D., VAN DAM, J.C., FRANK, J., DE MATTOS, M.J.T., HEIJNEN, J.J. Improved rapid sampling for in vivo kinetics of intracellular metabolites in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 2001;**75**(4):406–415.
- LEE, S.Y., LEE, D.Y., KIM, T.Y. Systems biotechnology for strain improvement. *Trends in Biotechnology* 2005;**23**(7):349–358.
- LEE, T.C.M. Smoothing parameter selection for smoothing splines: a simulation study. *Computational Statistics and Data Analysis* 2003;**42**:139–148.
- LEE, W.N., BEYERLEY, L., BERGNER, E., EDMOND, J. Mass isotopomer analysis; Theoretical and practical considerations. *Biol Mass Spectrom* 1991;**20**:451–458.
- LENGELER, J.W., DREWS, G., SCHLEGEL, H.G., Herausgeber. *Biology of the Prokaryotes*. Thieme, Stuttgart, 1999.

- LI, K., MIKOLA, M.R., DRATHS, K.M., WORDEN, R.M., FROST, J.W. Fed-Batch Fermentor Synthesis of 3-Dehydroshikimic Acid Using Recombinant *Escherichia coli*. *Biotechnol Bioeng* 1999;**64**(1):61–73.
- MAHADEVAN, R., BURGARD, A.P., FAMILI, I., VAN DIEN, S., SCHILLING, C.H. Applications of metabolic modeling to drive bioprocess development for the production of value-added chemicals. *Biotechnol Bioproc Eng* 2005;**10**(5):408–417.
- MAHADEVAN, R., SCHILLING, C. The Effects of Alternate Optimal Solutions in Constraint-Based Genome-Scale Metabolic Models. *Metab Eng* 2003;**5**(4):264–276.
- MARX, A., DE GRAAF, A.A., WIECHERT, W., EGGELING, L., SAHM, H. Determination of the fluxes in central metabolism of *Corynebacterium glutamicum* by NMR spectrometry combined with metabolic balancing. *Biotechnol Bioeng* 1996;**49**:111–129.
- MATHEWS, C. The cell: Bag of Enzymes or network of channels? *J Bacteriol* 1993;**75**(20):6377–6381.
- MAUCH, K. INSILICO discovery 1.0 - the tool for setup and analysis of in silico cell models. 2005.
- MAUCH, K., ARNOLD, S., REUSS, M. Dynamic sensitivity analysis for metabolic systems. *Chem Eng Sci* 1997;**52**(15):2589–2598.
- MAVROVOUNIOTIS, M.L. Estimation of standard Gibbs energy changes of biotransformations. *J Biol Chem* 1991;**266**(22):14440–14445.
- MENDES, P. Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends Biochem. Sci.* 1997;**22**:361–363.
- MICHALEWICZ, Z., FOGEL, D.B. *How to Solve It: Modern Heuristics*. Springer, Berlin, 1 Auflage, 2002.
- MOORE, J.D., COGGINS, J.R., VIRDEN, R., HAWKINS, A.R. Efficient independent activity of a monomeric, monofunctional dehydroquinase synthase derived from the N-terminus of the pentafunctional AROM protein of *Aspergillus nidulans*. *Biochem J* 1994;**301**:297–304.
- MÖLLNEY, M., WIECHERT, W., KOWNATZKI, D., DE GRAAF, A.A. Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments. *Biotechnol Bioeng* 1999;**66**(2):86–103.
- NIELSEN, J. Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates. *Biochemical Journal* 1997;**321**:133–138.
- NIELSEN, J. Metabolic Engineering. *Appl Microbiol Biotechnol* 2001;**55**:263–283.
- NOACK, S. *Visualisierung und Analyse biochemischer Netzwerkmodelle*. Diplomarbeit, Technische Universität Dresden, Dresden, 2005.

- NOACK, S., WAHL, S., HAUNSCHILD, M., QELI, E., FREISLEBEN, B., WIECHERT, W. Visualizing regulatory interdependencies and parameter sensitivities in biochemical network models. In: I. TROCH, F. BREITENECKER, Herausgeber, *5th International Conference on Mathematical Modelling (MathMod)*, Band Vol. 30,1/30. Argesim Verlag, Vienna, Austria, 2006: 17.
- NOGUCHI, Y., NAKAI, Y., SHIMBA, N., TOYOSAKI, H., KAWAHARA, Y., SUGIMOTO, S., SUZUKI, E. The energetic conversion competence of Escherichia coli during aerobic respiration studied by 31-P NMR using a circulating fermentation system. *J Biochem* 2004;**136**(4):509–15.
- NÖH, K., GRÖNKE, K., LUO, B., TAKORS, R., OLDIGES, M., WIECHERT, W. Metabolic flux analysis at ultra short time scale: isotopically non-stationary ¹³C labeling experiments. *J Biotechnol* 2007;**129**(2):249–267.
- NÖH, K., WAHL, A., WIECHERT, W. Computational tools for isotopically instationary ¹³C labeling experiments under metabolic steady state conditions. *Metab Eng* 2006; **8**(6):554–577.
- NÖH, K., WIECHERT, W. Experimental design principles for isotopically instationary ¹³C labeling experiments. *Biotechnol Bioeng* 2006;**94**(2):234–251.
- OLDIGES, M. *Metabolomanalyse zur Untersuchung der Dynamik im Aromatenbiosyntheseweg in L-Phenylalanin Produzenten von Escherichia coli*. Dissertation, Universität Bonn, Bonn, 2004.
- OLDIGES, M., KUNZE, M., DEGENRING, D., SPRENGER, G.A., TAKORS, R. Stimulation, monitoring, and analysis of pathway dynamics by metabolic profiling in the aromatic amino acid pathway. *Biotechnol Progr* 2004;**20**(6):1623–1633.
- OLIVIER, B.G., SNOEP, J.L. Web-based kinetic modelling using JWS Online. *Bioinformatics* 2004;**20**(13):2143–2144.
- ONSAGER, L. Reciprocal relations in irreversible processes. *Phys Rev* 1937;**37**:405–426.
- PALSSON, B.O. In silico biotechnology - Era of reconstruction and interrogation. *Curr Opin Biotech* 2004;**15**(1):50–51.
- PATNAIK, R., LIAO, J.C. Engineering of Escherichia coli Central Metabolism for Aromatic Metabolite Production with Near Theoretical Yield. *Appl Environ Microbiol* 1994; **10**:3903–3908.
- PATNAIK, R., SPITZER, R.G., LIAO, J.C. Pathway Engineering for Production of Aromatics in Escherichia-Coli - Confirmation of Stoichiometric Analysis by Independent Modulation of AroG, TktA, and Pps Activities. *Biotechnol Bioeng* 1995;**46**(4):361–370.
- PAZMAN, A. *Foundations of Optimum Experimental Design*. Kluwer Academic Publishing, 1986.

- PAZOS, F., RANEA, J.A., JUAN, D., STERNBERG, M.J. Assessing protein co-evolution in the context of the tree of life assists in the prediction of the interactome. *J Mol Biol* 2005;**352**(4):1002–15.
- PEARSON, K. On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine* 1901;**2**(6):559–572.
- PEEKHAUS, N., CONWAY, T. What is for Dinner?: Entner-Doudoroff Metabolism in *Escherichia coli*. *J Bacteriol* 1998;**180**:3495–3502.
- PETERSEN, S., DE GRAAF, A.A., EGGELING, L., MÖLLNEY, M., WIECHERT, W., SAHM, H. In vivo quantification of parallel and bidirectional fluxes in the anaplerosis of *Corynebacterium glutamicum*. *J Biol Chem* 2000;**275**(46):35932–35941.
- PETERSEN, S., MACK, C., GRAAF, A.D., RIEDEL, C., EIKMANN, B., SAHM, H. Metabolic consequences of altered phosphoenolpyruvate carboxykinase activity in *Corynebacterium glutamicum* reveal anaplerotic regulation mechanisms in vivo. *Metab Eng* 2001;**3**(4):344–361.
- PFEIFFER, T., SANCHEZ-VALDENEBRO, I., NUNO, J.C., MONTERO, F., SCHUSTER, S. METATOOL: for studying metabolic networks. *Bioinf* 1999;**15**(3):251–257.
- PHARKYA, P., BURGARD, A.P., MARANAS, C.D. OptStrain: A computational framework for redesign of microbial production systems. *Genome Res* 2004;**14**(11):2367–2376.
- PICHERSKY, E., GANG, D.R. Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plant Sci* 2000;**5**(10):439–45.
- PITTARD, A. Biosynthesis of the Aromatic Amino Acids. In: F. NEIDHARDT, Herausgeber, *Escherichia coli and Salmonella typhimurium - Cellular and Molecular Biology*. ASM press, Washington, DC, 1996:.
- QELI, E. *Visualisation of Biochemical Reaction Networks*. Dissertation, Universität Marburg, Marburg, 2007.
- RAAMSDONK, L.M., TEUSINK, B., BROADHURST, D., ZHANG, N., HAYES, A., WALSH, M.C., BERDEN, J.A., BRINDLE, K.M., KELL, D.B., ROWLAND, J.J., WESTERHOFF, H.V., VAN DAM, K., OLIVER, S.G. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* 2001;**19**(1):45–50.
- RAMAKRISHNA, R., EDWARDS, J.S., MCCULLOCH, A., PALSSON, B.O. Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. *Am J Physiol-Reg I* 2001;**280**(3):695–704.
- RIZZI, M., BALTES, M., U., T., REUSS, M. In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol Bioeng* 1997;**55**:592–608.

- ROHWER, J.M., POSTMA, P.W., KHOLODENKO, B.N., WESTERHOFF, H.V. Implications of macromolecular crowding for signal transduction and metabolite channeling. *Proc Nat Acad Sci* 1998;**95**:10547–10552.
- ROTENBERG, S.L., SPRINSON, D.B. Isotope Effects in 3-Dehydroquinate Synthase and Dehydratase - Mechanistic Implications. *J Biol Chem* 1978;**253**(7):2210–2215.
- ROWAN, T.H. *Functional Stability Analysis of Numerical Algorithms*. Dissertation, Department of Computer Sciences, University of Texas at Austin, 1990.
- SACHS, A.B. Messenger-RNA Degradation in Eukaryotes. *Cell* 1993;**74**(3):413–421.
- SALTELLI, A., CHAN, K., SCOTT, E.M., Herausgeber. *Sensitivity Analysis*. John Wiley & Sons Ltd., 1 Auflage, 2000.
- SAMMON, J. A nonlinear mapping for data structure analysis. *IEEE Trans Comp* 1969;**C-18**(5):401–409.
- SAUER, U., BAILEY, J.E. Estimation of P-to-O ratio in *Bacillus subtilis* and its influence on maximum riboflavin yield. *Biotechnol Bioeng* 1997;**64**:750–754.
- SAVAGEAU, M.A. Biochemical systems analysis. III. Dynamic solutions using a power-law approximation. *J Theor Biol* 1970;**26**(2):215–226.
- SAVAGEAU, M.A. Metabolite Channeling - Implications for Regulation of Metabolism and for Quantitative Description of Reactions *in vivo*. *J Theor Biol* 1991;**152**(1):85–92.
- SCHILLING, C.H., EDWARDS, J., LETSCHER, D., PALSSON, B. Combining Pathway Analysis with Flux Balance Analysis for the Comprehensive Study of Metabolic Systems. *Biotechnol Bioeng* 2000;**71**(4):286–306.
- SCHLEGEL, H.G. *Allgemeine Mikrobiologie*. Georg Thieme Verlag, Stuttgart, 1992.
- SCHONER, R., HERRMANN, K.M. 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase - Purification, Properties, and Kinetics of Tyrosine-Sensitive Isoenzyme from *Escherichia-coli*. *J Biol Chem* 1976;**251**(18):5440–5447.
- SCHUSTER, R., SCHUSTER, S. Refined algorithm and computer program for calculating all non-negative fluxes admissible in steady states of biochemical reaction systems with or without some flux rates fixed. *Comp Appl Biosc* 1993;**9**(1):79–85.
- SCHUSTER, S., DANDEKAR, T., FELL, D.A. Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol* 1999;**17**(2):53–60.
- SCHUSTER, S., FELL, D.A., DANDEKAR, T. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat Biotechnol* 2000;**18**(3):326–332.
- SCHUSTER, S., HILGETAG, C. On elementary flux modes in biochemical reaction systems at steady state. *J Biol Syst* 1994;**2**:165–182.

- SCHÄFER, U., BOOS, W., TAKORS, R., WEUSTER-BOTZ, D. Automated sampling device for monitoring intracellular metabolite dynamics. *Anal Biochem* 1999;**270**:88–96.
- SELKOV, E., BASMANOVA, S., GAASTERLAND, T., GORYANIN, I., GRETCHKIN, Y., MALTSEV, N., NENASHEV, V., OVERBEEK, R., PANYUSHKINA, E., PRONEVITCH, L., E, E.S., YUNUS, I. The metabolic pathway collection from EMP: the enzymes and metabolic pathways database. *Nucl Acids Res* 1996;**24**(1):26–28.
- SHEARER, G., LEE, J.C., KOO, J., KOHL, D.H. Quantitative estimation of channeling from early glycolytic intermediates to CO₂ in intact *Escherichia coli*. *FEBS J* 2005; **272**(13):3260–3269.
- SIMPSON, R.J., DAVIDSON, B.E. Studies on 3-Deoxy-D-Arabinohexulose-7-Phosphate Synthetase(phe) from *Escherichia coli*-K12: 1. Purification and Subunit Structure. *Eur J Biochem* 1976;**70**(2):493–500.
- SRERE, P. Complexes of sequential metabolic enzymes. *Annu Rev Biochem* 1987;**56**:89–124.
- STEIN, N. *¹³C-Metabolomanalyse in Aminosäure-Produzenten am Beispiel von L-Threonin*. Dissertation, Universität Hamburg, 2007.
- STEPHANOPOULOS, G., ARISTIDOU, A., NIELSEN, J. *Metabolic Engineering: Principles and Methodologies*. Academic Press, 1 Auflage, 1998.
- STRANG, G. *Linear Algebra and Its Application*. Academic Press Ltd., London, 1976.
- SUMEGLI, B., SHERRY, A.D., MALLOY, C.R., SRERE, P.A. Evidence for orientation-conserved transfer in the TCA cycle in *Saccharomyces cerevisiae*: carbon-13 NMR studies. *Biochem* 1993;**32**:12725–12729.
- TAKAHASHI, K., ISHIKAWA, N., SADAMOTO, Y., SASAMOTO, H., OHTA, S., SHIOZAWA, A., MIYOSHI, F., NAITO, Y., NAKAYAMA, Y., TOMITA, M. E-cell 2: Multiplatform E-Cell simulation system. *Bioinformatics* 2003;**19**(13):1727–1729.
- TAKORS, R., EL MASSAOUDI, M., WAHL, S.A., DRYSCH, A., DE GRAAF, A.A., WIECHERT, W., WANDREY, C. Der Sensor Reaktor als neuartiges Werkzeug der ¹³C-basierten, seriellen Stoffflussanalyse unter Produktionsprozess-ähnlichen Bedingungen. In: *DECHEMA-GVC Jahrestagung*. Karlsruhe, 2004: .
- THEOBALD, U., MAILINGER, W., BALTES, M., RIZZI, M., REU, M. In vivo analysis of metabolic dynamics in *saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol Bioeng* 1997;**55**(2):305–316.
- TORRES, N.V., MATEO, F., MELENDEZ-HEVIA, E., KACSER, H. Kinetics of metabolic pathways. A system in vitro to study the control of flux. *Biochem J* 1986; **234**(1):169–174.

- VALLE, F., MUNOZ, E., PONCE, E., FLORES, N., BOLIVAR, F. Basic and applied aspects of metabolic diversity: The phosphoenolpyruvate node. *J Ind Microbiol Biotechnol* 1996;**17**(5-6):458–462.
- VALLINO, J.J., STEPHANOPOULOS, G. Metabolic Flux Distributions in *Corynebacterium-Glutamicum* During Growth and Lysine Overproduction. *Biotechnol Bioeng* 1993;**41**(6):633–646.
- VAN DAM, J.C., EMAN, M.R., FRANK, J., LANGE, H.C., DEDEM, VAN GIJS, W.K., HEIJNEN, J. Analysis of glycolytic intermediates in *Saccharomyces cerevisiae* using anion exchange chromatography and electrospray ionization with tandem mass spectrometric detection. *Analytica Chimica Acta* 2002;**460**:209–218.
- VAN WINDEN, W., VERHEIJEN, P., HEIJNEN, S. Possible pitfalls of flux calculations based on C-13-labeling. *Metab Eng* 2001;**3**(2):151–162.
- VAN WINDEN, W.A., VAN GULIK, W.M., SCHIPPER, D., VERHEIJEN, P.J.T., KRABBEN, P., VINKE, J.L., HEIJNEN, J.J. Metabolic flux and metabolic network analysis of *Penicillium chrysogenum* using 2D [C-13, H-1] COSY-NMR measurements and cumulative Bondomer simulation. *Biotechnol Bioeng* 2003;**83**(1):75–92.
- VEMURI, G.N., ARISTIDOU, A.A. Metabolic engineering in the -omics era: Elucidating and modulating regulatory networks. *Microbiol Mol Biol R* 2005;**69**(2):197–217.
- VISSER, D., HEIJNEN, J.J. Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. *Metab Eng* 2003;**5**(3):164–176.
- VISSER, D., SCHMID, J.W., MAUCH, K., REUSS, M., HEIJNEN, J.J. Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics. *Metab Eng* 2004a;**6**(4):378–390.
- VISSER, D., VAN ZUYLEN, G.A., VAN DAM, J.C., EMAN, M.R., PROLL, A., RAS, C., WU, L., VAN GULIK, W.M., HEIJNEN, J.J. Analysis of in vivo kinetics of glycolysis in aerobic *Saccharomyces cerevisiae* by application of glucose and ethanol pulses. *Biotechnol Bioeng* 2004b;**88**(2):157–167.
- VISSER, D., VAN ZUYLEN, G.A., VAN DAM, J.C., OUDSHOORN, A., EMAN, M.R., RAS, C., VAN GULIK, W.M., FRANK, J., VAN DEDEM, G.W.K., HEIJNEN, J.J. Rapid sampling for analysis of in vivo kinetics using the BioScope: A system for continuous-pulse experiments. *Biotechnol Bioeng* 2002;**79**(6):674–681.
- WAHL, S.A. *Statistische Untersuchungen von Messmodellen für die Metabolische Stoffflussanalyse mit ¹³C Markierungsexperimenten*. Diplomarbeit, TU Dresden, Dresden, 2002.
- WAHL, S.A., DAUNER, M., WIECHERT, W. New tools for mass isotopomer data evaluation in C-13 flux analysis: Mass isotope correction, data consistency checking, and precursor relationships. *Biotechnol Bioeng* 2004a;**85**(3):259–268.

- WAHL, S.A., HAUNSCHILD, M.D., OLDIGES, M., WIECHERT, W. Unraveling the Regulatory Structure of Biochemical Networks using Stimulus Response Experiments and Large Scale Model Selection. *IEE Proc Syst Biol* 2006a:275–285.
- WAHL, S.A., MASSAOUDI, M.E., SCHIPPER, D., WIECHERT, W., TAKORS, R. Serial C-13-Based Flux Analysis of an L-Phenylalanine-Producing E-coli Strain Using the Sensor Reactor. *Biotechnol Progr* 2004b:20(3):706–714.
- WAHL, S.A., TAKORS, R., WIECHERT, W. Interpretation of Metabolic Flux Maps by Limitation Potentials and Constraint Limitation Sensitivities. *Biotechnol Bioeng* 2006b:263–272.
- WATSON, J.D., CRICK, F.H. The structure of DNA. *Cold Spring Harb Symp Quant Biol* 1953:18:123–31.
- WESTERHOFF, H.V., VAN DAM, K. *Thermodynamics and control of biological free energy transduction*. Elsevier, Amsterdam, 1987.
- WIECHERT, W. *Metabolische Kohlenstoff-Markierungssysteme: Modellierung, Simulation, Analyse, Datenauswertung*. Habilitationsschrift, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, 1995.
- WIECHERT, W. C-13 metabolic flux analysis. *Metab Eng* 2001:3(3):195–206.
- WIECHERT, W. Modeling and simulation: tools for metabolic engineering. *J Biotechnol* 2002:94(1):37–63.
- WIECHERT, W., GRAAF, A.D. Bidirectional Reaction Steps in Metabolic Networks I. Modeling and Simulation of Carbon Isotope Labeling Experiments. *Biotechnol Bioeng* 1997:55(1):101–117.
- WIECHERT, W., MÖLLNEY, M., ISERMANN, N., WURZEL, W., DE GRAAF, A.A. Bidirectional reaction steps in metabolic networks: III. Explicit solution and analysis of isotopomer labeling systems. *Biotechnol Bioeng* 1999:66(2):69–85.
- WIECHERT, W., MÖLLNEY, M., PETERSEN, S., DE GRAAF, A.A. A universal framework for C-13 metabolic flux analysis. *Metab Eng* 2001:3(3):265–283.
- WIECHERT, W., NOH, K. From stationary to instationary metabolic flux analysis. *Technology Transfer in Biotechnology: From Lab to Industry to Production* 2005:92:145–172.
- WILLIAMS, C., STILLMAN, T., BARYNIN, V., SEDELNIKOVA, S., TANG, Y., GREEN, J., GUEST, J., ARTYMIUK, P. E-coli aconitase B structure reveals a HEAT-like domain with implications for protein-protein recognition. *Nat Struct Biol* 2002:9(6):447–452.
- WU, L., MASHEGO, M.R., VAN DAM, J.C., PROELL, A.M., VINKE, J.L., RAS, C., VAN WINDEN, W.A., VAN GULIK, W.M., HEIJNEN, J.J. Quantitative analysis

- of the microbial metabolome by isotope dilution mass spectrometry using uniformly ^{13}C -labeled cell extracts as internal standards. *Anal Biochem* 2005;**336**(2):164–71.
- YI, J., LI, K., DRATHS, K.M., FROST, J.W. Modulation of Phosphoenolpyruvate Synthase Expression Increases Shikimate Pathway Product Yields in *E. coli*. *Biotechnol Prog* 2002;**18**:1141–1148.
- ZANGIROLAMI, T., JOHANSEN, C., NIELSEN, J., JØRGENSEN, S. Simulation of penicillin production in fed-batch cultivations using a morphologically structured model. *Biotechnol Bioeng* 1997;**56**(6):593–604.
- ZEPPENFELD, T., LARISCH, C., LENGELER, J.W., JAHREIS, K. Glucose Transporter Mutants of *Escherichia coli* K-12 with Changes in Substrate Recognition of IICB(Glc) and Induction Behavior of the ptsG gene. *J Bacteriol* 2000;**182**:4443–4452.
- ZHU, C., BYRD, R.H., NOCEDAL, J. Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization. *ACM Transactions on Mathematical Software* 1997;**23**(4):550 – 560.

1. **Methoden zur integrierten Analyse metabolischer Netzwerke unter stationären und instationären Bedingungen**
von S. A. Wahl (2008), 245 Seiten
ISBN: 978-3-89336-506-7



Band | Volume 1
ISBN 978-3-89336-506-7

