

Analyse thermokinetischer Potential-Strom-Kennlinien an einem Modell von *Escherichia coli*

Studienarbeit

im Studiengang
Technische Biologie

vorgelegt von
Waldemar Reusch

Prüfer: Prof. Dr.-Ing. Oliver Sawodny
Betreuer: Dr.-Ing. Michael Ederer

Institut für Systemdynamik, Universität Stuttgart
Prof. Dr.-Ing. O. Sawodny

28. August 2012

Kurzfassung

In den letzten Jahren ist die Menge und Qualität an erhobenen Meßdaten aus den verschiedensten Reaktionsnetzwerken dank der verbesserten Meßtechnik stark angestiegen. Gleichzeitig müssen immer detailliertere mathematische Modelle erstellt werden, um die untersuchten Systeme zu beschreiben und Hypothesen zu evaluieren.

Die kinetische Modellierung gehört zu den gängigsten Modellierungsmethoden. Diese Modelle beruhen auf gewöhnlichen Differentialgleichungen. Ein Nachteil dieser Methode ist allerdings, dass auf die Einhaltung von thermodynamischen Grundsätzen gesondert geachtet werden muss. Diese Schranken werden Wegscheiderbedingungen genannt, und ihre Zahl wächst mit der Größe des untersuchten Systemes.

Mit der Einführung der *thermokinetischen Modellierung* eröffnet sich eine Modellierungsmethode, in welcher die Einhaltung thermodynamischer Grundsätze durch die Struktur des Modelles sicher gestellt wird.

Statt Konzentrationen und Ratenkonstanten sind in der *thermokinetischen Modellierung* Potentiale, Ströme und Widerstände das Hauptaugenmerk. So weist diese Methode interessante Parallelen zu elektrotechnischen Modellsystemen auf. Chemische Spezies können als Kapazitäten betrachtet werden, Reaktionen als Widerstände. In der Elektrotechnik erlauben Spannungs-Strom-Kennlinien die Beurteilung einer Spannungsquelle durch die Darstellung der Quellenspannung in Abhängigkeit vom entnommenen Strom. Eine ähnliche Potential-Strom-Kennlinie kann nun auch für chemische Spezies in einem *thermokinetischen Modell* angelegt werden.

Ziel dieser Arbeit war es, solche Potential-Strom-Kennlinien für die *zentralen Vorläufermoleküle* aus dem *Zentralstoffwechsel* und einige *niedermolekulare Bausteine* zu erstellen und zu beschreiben. Grundlage hierfür war ein von Michael Ederer erstelltes Modell eines *E. coli*-Chemostaten. Dieses Modell musste um *niedermolekulare Bausteine* erweitert werden. Als nächstes galt es, ein Analyseskript zu erstellen, welches automatisiert Simulationen am Modell ausführt, die Daten bewertet und Kennlinien plottet. Schließlich wurde eine Hypothese überprüft, die es ermöglicht, Kennlinien von komplexen Molekülen, anhand vorhandener Kennlinien ihrer *Vorläufermoleküle* abzuschätzen.

Bis auf wenige Ausnahmen zeigten alle erstellten Kennlinien einen bemerkenswert ähnlichen Verlauf, welcher einfach zu erklären ist. Die geschätzten Kennlinien von komplexeren Molekülen lagen in den meisten Fällen sehr nah an den simulierten Kennlinien. Das Analyseskript vereinfachte hier die Arbeit, da es im Batch-Verfahren automatisiert viele Simulationen ausführt und auswertet.

Inhaltsverzeichnis

1 Einleitung und Zielsetzung	1
2 Theoretische Hintergründe	3
2.1 Thermokinetische Modellierung	4
2.2 Das System in der grafische Darstellung	9
2.3 Potential-Strom-Kennlinien	11
2.4 Niedermolekulare Bausteine	13
2.5 Coenzyme	14
2.6 Zielsetzung	15
3 Das Modell in TKMOD	17
3.1 Das <i>TKMOD</i> -Modul	17
3.2 Ein Modell in <i>TKMOD</i>	18
3.3 Der <i>Escherichia coli</i> -Chemostat	20
4 Umsetzung	23
4.1 Ziel Eins: Implementierung von niedermolekularen Bausteinen	23
4.2 Ziel Zwei: Berechnung von Potential-Strom-Kennlinien	24
4.3 Ziel Drei: Schätzung von Kennlinien komplexer Bausteine	27
5 Analyse und Diskussion A: Die Potential-Strom-Kennlinien	29
5.1 Kennlinien der Vorläufermoleküle	30
5.2 Kennlinien der Aminosäuren	33
5.3 Kennlinien der Nucleotide	35
5.4 Kennlinien der Lipopolysaccharidmonomere	36
5.5 Kennlinien der Peptidoglycanmonomere	37
5.6 Diskussion	37

Inhaltsverzeichnis

6 Analyse und Diskussion B: Theoretische Kennlinien im Vergleich mit dem Simulationsergebnis	39
6.1 Vergleich bei Aminosäuren	40
6.2 Vergleich bei Nucleotiden	42
6.3 Vergleich bei Lipopolysaccharidmonomeren	43
6.4 Vergleich bei Peptidoglycanmonomere	44
6.5 Diskussion	44
7 Zusammenfassung und Ausblick	47
Abbildungsverzeichnis	49
Tabellenverzeichnis	50
Quelltextverzeichnis	50
Literaturverzeichnis	51

Kapitel 1

Einleitung und Zielsetzung

Die Komplexität der in der Biologie untersuchten Reaktionsnetzwerke ist in den letzten Jahren mit der steten Verbesserung von Meßtechniken immer weiter gestiegen. Ebenso stieg die Menge und Qualität der erhobenen Meßdaten. Im gleichen Zuge müssen mathematische Modelle und Computersimulationen immer höheren Anforderungen entsprechen, um detaillierte und umfangreiche Hypothesen zu bestätigen oder zu widerlegen [7]. Ein tieferes Verständins von metabolischen Netzwerken ist sowohl für medizinische als auch industrielle Zwecke von Interesse.

Weit verbreitet und erfolgreich ist die Arbeit mit kinetischen Modellen [4], die auf gewöhnlichen Differentialgleichungen beruhen. Bei diesem Lösungsansatz ist jedoch gesondert darauf zu achten, die Gesetze der Thermodynamik nicht zu verletzen. Dies nötigen Schranken wurden 1902 von Wegscheider für einfache chemische Systeme formuliert und 1931 von Onsager weiter ausgebaut [6]. Schuster und Schuster haben die verallgemeinerten Wegscheiderbedingungen formuliert, die für Reaktionsnetzwerke mit verallgemeinerten Massenwirkungskinetiken gelten [10]. Ederer und Gilles haben die *thermokinetische Modellierung* eingeführt, welche die Einhaltung der verallgemeinerten Wegscheiderbedingungen allein durch die Struktur des Modelles garantiert [5].

Der großen Vielfalt an Zellbestandteilen wie Proteinen, Lipiden, Kohlehydraten und Nucleinsäuren liegt eine überraschend kleine Zahl an Vorläufermolekülen zu Grunde. Nur ein Dutzend chemischer Stoffe zählen zu den zentralen Vorläufermolekülen und alle sind Intermediate des zentralen Stoffwechselweges, welcher der Mehrheit an Pro- und Eukaryoten gemeinsam ist. Zu diesem zählt der Citratzyklus, die Glykolyse/Gluconeogenese und der Pentosephosphatweg. Diese Stoffwechselwege liefern nicht nur zentrale Vorläufermoleküle sondern sind auch zuständig für die Energieversorgung [8].

Vorläufermoleküle aus dem zentralen Metabolismus reagieren in Biosynthesewegen zu niedermolekularen Bausteinen und über weitere Pfade zu komplexen Stoffwechselprodukten. Somit gibt es einen steten Abfluss an zentralen Metaboliten aus dem Zentralstoffwechselweg der diesen belastet. In Wildtypen sind einzelne Biosynthesewege in der Regel so reguliert, dass sie eine niedrigere Priorität haben als der Zentralstoffwechsel, so dass die Energieversorgung oder andere Biosynthesewege der Zelle nicht behindert werden. In Computermodellen jedoch kann man diese Abflüsse frei einstellen und die Auswirkungen auf die Konzentrationen der chemischen Komponenten untersuchen.

Kapitel 1 Einleitung und Zielsetzung

Man kann hier einen Vergleich mit einem elektrotechnischen Netzwerk wagen. Die Abflussrate entspricht dem elektrischen Strom, während die Konzentration mit der Spannung verglichen werden kann. Wird der Spannungsquelle ein großer Strom abverlangt, so sinkt die Spannung, wie die Konzentration bei einem großen Abfluss sinkt.

In dieser Arbeit sollen die Konzentrationen von verschiedenen Vorläufermolekülen und niedermolekularen Bausteinen unter eben diesen Belastungen betrachtet werden. Grundlage wird hierbei ein *thermokinetisches Modell* von *Escherichia coli* sein.

In der *thermokinetischen Modellierung* werden Konzentrationen mit thermokinetischen Potentialen beschrieben. Zunächst wird eine Methode entwickelt, um das Potential einer chemischen Spezies in Abhängigkeit von deren Abflussrate darzustellen. So werden die Darstellungen, in Anlehnung an die Elektrotechnik, Potential-Strom-Kennlinien genannt werden.

Als nächstes werden Kennlinien erstellt werden, sowohl für Vorläufermoleküle aus dem Zentralstoffwechsel als auch für niedermolekulare Bausteine aus Biosynthesewegen. Dazu werden Interpretationsansätze erörtert.

Zuletzt wird ein Ansatz bewertet, welcher es erlaubt, Kennlinien von komplexen Syntheseprodukten aus den Kennlinien ihrer Vorläufermoleküle abzuschätzen.

Kapitel 2

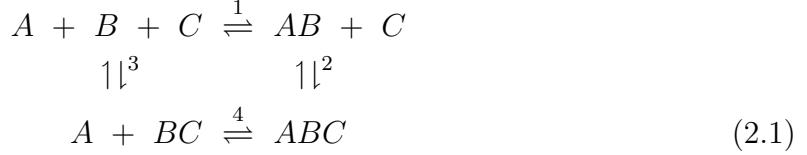
Theoretische Hintergründe

Das Verhalten von biochemischer Reaktionsnetzwerke ist stark abhängig von den Einschränkungen, die durch die Massen- und Energiebilanzen auferlegt werden. So muss einerseits die Masse der Edukte der Masse der Produkte entsprechen. Damit andererseits eine isobare und isotherme Reaktion in die Richtung der Produkte verläuft, muss die aufgewendete Gibb'sche Energie größer sein als die erzeugte Gibb'sche Energie. In einer vorhergehenden Arbeit wurde gezeigt, dass grundlegende Funktions- und Regulationseigenschaften metabolischer Netzwerke allein durch diese Begrenzungen der Netzwerkstruktur beschrieben werden können [11].

Da die stöchiometrischen Verhältnisse in chemischen Netzwerken weitaus besser bekannt sind als die kinetischen Parameter, ist bei der **kinetischen Modellierung** die Beachtung dieser Schranken für die Wahl der kinetischen Parameter essentiell. Die Schranken aus der Massenbilanz können relativ einfach in Massenerhaltungssätzen fest gehalten werden. Bei der korrekten Modellierung von Energieflüssen müssen indes Schranken aus zwei Ansätzen beachtet werden. Zum einen die Schranken, die aus dem zweiten Hauptsatz der Thermodynamik resultieren, zum zweiten solche, die aus dem Prinzip des detaillierten Gleichgewichtes stammen. Diese wurden 1902 von Wegscheider formuliert und werden als *Wegscheider Bedingungen* bezeichnet [6]. Ist das Modellsystem isoliert von äußerlichen thermodynamischen Einflüssen, so postuliert der zweite Hauptsatz, wird es einen Zustand des thermodynamischen Gleichgewichtes erreichen. In diesem thermodynamischen Gleichgewicht, so verlangt das Prinzip des detaillierten Gleichgewichtes, ist die Wahrscheinlichkeit einer Reaktion gleich der Wahrscheinlichkeit der Rückreaktion, somit sind die Reaktionsraten gleich und der Gesamtfluss dieser Reaktion gleich Null.

2.1 Thermokinetische Modellierung

Betrachten wir ein Beispielsystem aus [5]. Die Komponenten A , B und C komplexieren in beliebiger Reihenfolge. Hierdurch entsteht ein Kreislauf mit den Komponenten A , B , C , AB , BC und ABC :



Nehmen wir an, dass es sich bei dem System um eine homogene Phase konstanten Volumens handelt. Des Weiteren handelt es sich bei allen Reaktionen j um Massenwirkungskinetiken mit den kinetischen Parametern $k_{\pm j}$. Sind c_i [mol m^{-3}] die Konzentrationen der chemischen Spezies A_i , so berechnen sich die Molbilanzen \dot{c}_i [$\text{mol m}^{-3} \text{ s}^{-1}$] und Stoffflüsse J_j [$\text{mol m}^{-3} \text{ s}^{-1}$] über die Reaktionen j wie folgt.

$$\begin{aligned} \dot{c}_A &= -J_1 - J_4 & \dot{c}_B &= -J_1 - J_3 & \dot{c}_C &= -J_2 - J_3 \\ \dot{c}_{AB} &= +J_1 - J_2 & \dot{c}_{BC} &= +J_3 - J_4 & \dot{c}_{ABC} &= +J_2 + J_4 \end{aligned} \quad (2.2)$$

$$\begin{aligned} J_1 &= k_{+1} \cdot c_A \cdot c_B - k_{-1} \cdot c_{AB} \\ J_2 &= k_{+2} \cdot c_{AB} \cdot c_C - k_{-2} \cdot c_{ABC} \\ J_3 &= k_{+3} \cdot c_B \cdot c_C - k_{-3} \cdot c_{BC} \\ J_4 &= k_{+4} \cdot c_A \cdot c_{BC} - k_{-4} \cdot c_{ABC} \end{aligned} \quad (2.3)$$

Der zweite Satz der Thermodynamik und das detaillierte Gleichgewicht verlangen nun, dass für das geschlossene System ein Zustand des thermodynamischen Gleichgewichtes existiert, in dem die Gesamtflüsse $J_{j,eq}$ über alle Reaktionen j verschwinden: $J_{1,eq} = J_{2,eq} = J_{3,eq} = J_{4,eq} = 0$. Nun lässt sich das Verhältnis der Reaktionskonstanten über die Gleichgewichtskonzentrationen berechnen: $c_{AB,eq}/(c_{A,eq} \cdot c_{B,eq}) = k_{+1}/k_{-1}$. In diesem Beispiel erkennt man folgende Beziehung:

$$\frac{c_{AB,eq}}{c_{A,eq} \cdot c_{B,eq}} \cdot \frac{c_{ABC,eq}}{c_{AB,eq} \cdot c_{C,eq}} \cdot \frac{c_{BC,eq} \cdot c_{A,eq}}{c_{ABC,eq}} \cdot \frac{c_{B,eq} \cdot c_{C,eq}}{c_{BC,eq}} = 1 \quad (2.4)$$

Dies führt zu dem Schluss, dass das Produkt der Gleichgewichtskonstanten über dem Kreislauf ebenfalls gleich eins sein muss:

$$\frac{k_{+1}}{k_{-1}} \cdot \frac{k_{+2}}{k_{-2}} \cdot \frac{k_{-3}}{k_{+3}} \cdot \frac{k_{-4}}{k_{+4}} = 1 \quad (2.5)$$

Gleichung 2.5 ist die so genannte detailliert Gleichgewichtsbeziehung oder auch *Wegscheider Bedingung*. Ausserhalb dieser Grenze beschreibt das Modell ein System, dessen Flüsse im Gleichgewicht nicht verschwinden, und so die Konzentrationen stets von den Gleichgewichtskonzentrationen abweichen. Dies bedeutet zugleich, dass das Modell einem unmöglichen chemischen *Perpetuum mobile* entspricht [5].

Die *Wegscheider-Bedingungen* stellen einen direkten Zusammenhang zwischen den kinetischen Parametern k_j her. Werden diese verletzt, ist das Modell physikalisch unzulässig.

Betrachten wir im Weiteren die allgemeine Formulierung kinetischer Modelle:

$$\dot{c} = N \cdot J(c, k) \quad (2.6)$$

mit den Konzentrationen $c \in \mathbb{R}^n$, Flüssen $J \in \mathbb{R}^m$, kinetischen Parametern $k \in \mathbb{R}^l$ und der stöchiometrischen Matrix $N \in \mathbb{R}^{n \times m}$, welche die stöchiometrischen Parameter $\nu_{i,j}$ aller Komponenten A_i in allen Reaktionen j enthält.

Biologische Systeme zeichnen sich durch ein hohes Maß an Flexibilität und Robustheit aus. Dieses wird unter anderem durch teilweise redundante metabolische Pfade und Zyklen erreicht. Hieraus folgt jedoch, dass die Zahl n an Reaktanden X_i geringer ist als die Zahl m an Reaktionen j . Die Zahl der *Wegscheider-Bedingungen* aber wird von den Dimensionen und dem Rang der stöchiometrischen Matrix N bestimmt. Ist $r = \text{Rang}(N)$, so ist Zahl an Erhaltungsbeziehungen gegeben als $d_n = n - r$ und die Zahl an metabolischen Zyklen als $d_m = m - r = m - n + d_n$. Dies ist gleichzeitig die Zahl an unabhängigen *Wegscheider-Bedingungen*, die nötig sind, um das Modell thermodynamisch konsistent zu beschreiben. Die Differenz zwischen Reaktionen und Reaktanden schlägt sich hier positiv nieder, und die Zahl an *Wegscheider-Bedingungen* ist bei großen Modellen in der Regel sehr hoch [5].

Die vorangegangene Argumentation setzt die Möglichkeit eines thermodynamischen Gleichgewichtes in einem geschlossenen System voraus. Die *Wegscheider-Bedingungen* gelten aber auch gleichermaßen für Modelle offener Systeme, die das thermodynamische Gleichgewicht nicht erreichen. In diesem Fall stellen sie sicher, dass das Modell das thermodynamische Gleichgewicht erreichen würde, wenn alle Flüsse über die Systemgrenze unterbunden werden, d.h. das System geschlossen wird.

Versucht man die Gleichung 2.5 auf ein großes Netzwerk zu übertragen, so kann der Fall auftreten, dass Gleichgewichtskonstanten in Relation gebracht werden, welche im biochemischen Reaktionsnetzwerk unterschiedlichen Funktionseinheiten oder Kompartimenten angehören. Die Anwendung der *Wegscheider-Bedingungen* gestaltet sich somit sehr unintuitiv. Zusätzlich behindern diese Relationen eine Sensitivitätsanalyse, d.h. die Untersuchung von Änderungen des Systemverhaltens durch Variationen in jedem einzelnen Parameter. Solche Analysen können zu irreführenden Ergebnissen führen, da die Parameter nicht frei wählbar sind, und Parameterkombinationen untersucht werden, die physikalisch nicht möglich sind.

Oft werden ubiquitäre Metabolite wie Adenosintriphosphat (ATP), Adenosiniphosphat (ADP) und anorganisches Phosphat (P_i) bei der Modellierung vernachlässigt, ihre Konzentrationen als konstant angenommen und deren Einfluss in die kinetischen Parameter einbezogen. Durch diese Vereinfachung ist aber die Herleitung von *Wegscheider-Bedingungen* nicht mehr möglich. Betrachten wir das Beispiel aus 2.1, und nehmen wir an, dass die Reaktion 1 tatsächlich unter Spaltung von energiereichem ATP zu energiearmen ADP und P_i verläuft: $A + B + ATP \rightleftharpoons AB + ADP + P_i$. Die

kinetischen Ratengleichungen in 2.3 müssen nicht verändert werden, da die Konzentrationen von ATP, ADP und P_i als konstant angenommen werden. Durch das Fixieren der Konzentrationen wirkt jedoch eine äußere thermodynamische Kraft auf das System und verhindert dessen thermodynamisches Gleichgewicht. Solch ein Kreislauf wird dann *futile cycle* genannt, da er durch einen permanenten Einstrom von energiereichen Verbindungen getrieben wird. Da das System nun keinen thermodynamischen Gleichgewichtszustand erreichen kann, gilt die an Gleichung 2.4 angewendete Argumentation nicht mehr, und es kann keine Gleichgewichtsbeziehung hergestellt werden. Dies gilt für alle vereinfachten Stöchiometrien, da nicht entschieden werden kann ob es sich um einen regulären Kreislauf oder einen *futile cycle* handelt.

Ganz im Gegensatz zu elektrotechnischen oder mechanischen Modellen, enthält das in 2.3 dargestellte Modell nicht die treibenden Kräfte der Reaktionen. In den erstgenannten Systemen haben Flüsse und ihre treibenden Kräfte stets die selbe Richtung, und ermöglichen so dem geschlossenen System das Erreichen eines Gleichgewichtszustandes mit verschwindenden Kräften und Strömen. Die Einhaltung des zweiten Satzes und des detaillierten Gleichgewichtes ist somit durch die Modellstruktur gewährleistet.

Im Gegensatz zu den vorhergehenden Modellierungsmethoden, basiert die **Thermokinetische Modellierung** auf dem thermokinetischen Potential ξ_i der chemischen Spezies X_i und der thermokinetischen Kraft F_j als treibende Kraft in Richtung des Flusses über den Reaktionen j [3]. So wird, wie bei elektrotechnischen oder mechanischen Modellen, die physikalische Konsistenz schon durch die Struktur des Modelles sicher gestellt.

Kapazitäten. Jede Komponente X_i besitzt eine Kapazität C_i [$mol\ L^{-1}$]. Für diese gilt: $c_i = C_i \cdot \xi_i$ mit der Konzentration c_i [$mol\ L^{-1}$] und dem thermokinetischen Potential ξ_i von Komponente i . In einer idealen Lösung ist die Kapazität C_i konstant.

Widerstände. Jede Reaktion j besitzt einen Widerstand R_j [$mol^{-1}L\ s$]. Für diesen gilt $R_j \cdot J_j = F_j$ mit der Reaktionsrate J_j [$mol\ L^{-1}$] und der thermokinetischen Kraft F_j entlang der Reaktion j . Die thermokinetische Kraft besteht aus den antagonistischen Kräften der Edukte und Produkten: $F_j = F_{E,j} - F_{P,j}$ mit $F_{E,j} = \prod_{i \in E_j} \xi_i^{|\nu_{E,i,j}|}$ und $F_{P,j} = \prod_{i \in P_j} \xi_i^{|\nu_{E,i,j}|}$. E_j und P_j ist die hier die Menge der Edukte bzw. Produkte, $\nu_{E/P,i,j}$ die Stöchiometrischen Koeffizienten der Edukte und Produkte X_i in der Reaktion j . Reaktionen mit einer idealen Massenwirkungskinetik besitzen einen konstanten thermokinetischen Widerstand R_j .

Thermodynamisches Gleichgewicht. In diesem Zustand versiegen alle Stoffflüsse J_j und mit ihnen alle treibenden Kräfte F_j . Mit $c_i = C_i \cdot \xi_i$ und $\prod_{i \in E_j} \xi_i^{|\nu_{E,i,j}|} = \prod_{i \in P_j} \xi_i^{|\nu_{E,i,j}|}$ lassen sich die Gleichgewichtskonstanten $K_{eq,j}$ für kinetische Modelle aus den Kapazitäten C_i der *thermokinetischen Modellierung* errechnen.

$$K_{eq,j} = \frac{k_{+j}}{k_{-j}} = \frac{\prod_{i \in P_j} c_{i,eq}^{|\nu_{E,i,j}|}}{\prod_{i \in E_j} c_{i,eq}^{|\nu_{E,i,j}|}} = \frac{\prod_{i \in P_j} C_{i,eq}^{|\nu_{E,i,j}|}}{\prod_{i \in E_j} C_{i,eq}^{|\nu_{E,i,j}|}} \quad (2.7)$$

Somit wird die Gleichgewichtskonstante $K_{eq,j}$ alleinig durch die Kapazitäten bestimmt. Die Einführung des Zusammenhangs zwischen Kräften und Flüssen garantiert im Besonderen die thermodynamische Konsistenz der Modelle.

Konstante Widerstände R_j und Kapazitäten C_i , und somit die linearen Zusammenhänge zwischen Kräften und Strömen bzw. zwischen Potentialen und Konzentrationen, gelten jedoch nur in idealen Lösungen mit idealen Massenwirkungskinetiken. Nicht-ideale Lösungen führen zu nicht konstanten, zustandsabhängigen Widerständen und Kapazitäten [4].

Betrachten wir die Reaktion



Die Berechnung des Flusses J erfolgt über die die Kraft F und den Widerstand R :

$$J = \frac{F}{R} \quad \text{mit } R \geq 0, \text{ hergeleitet in [4].} \quad (2.9)$$

Die treibende Kraft F ihrerseits errechnet sich aus den wirkenden Kräften der Edukte F_E und Produkte F_P , und diese wiederum aus den thermodynamischen Potentialen ξ_i :

$$F = F_E - F_P \quad F_E = \xi_A \cdot \xi_B \quad F_P = \xi_C. \quad (2.10)$$

Die thermodynamischen Potentiale ξ_i errechnen sich aus den Konzentrationen c_i und den Kapazitäten C_i :

$$\xi_i = \frac{c_i}{C_i} \quad \text{für } i = \{A, B, C\}. \quad (2.11)$$

Zusammengefasst ergibt sich folgende Systembeschreibung für das Beispiel in Gleichung 2.8:

$$J = \frac{c_A \cdot c_B}{R \cdot C_A \cdot C_B} - \frac{c_C}{R \cdot C_C}. \quad (2.12)$$

$$J = -\dot{c}_A = -\dot{c}_B = \dot{c}_C \quad (2.13)$$

Allgemein können wir nun ein mathematisches Modell eines Reaktionssystems mit n Komponenten A_i , m Reaktionen j und den Edukten E_j und Produkten P_j wie folgt beschreiben:

$$\begin{aligned} \dot{c}_i &= \sum_j (\nu_{E,i,j} + \nu_{P,i,j} \cdot J_j) \\ C_i \cdot \xi_i &= c_i \\ R_j \cdot J_j &= \prod_{i \in E_j} \xi_i^{|\nu_{E,i,j}|} - \prod_{i \in P_j} \xi_i^{|\nu_{E,i,j}|} \end{aligned} \quad (2.14)$$

Die drei Teile des Modelles in 2.14 stellen die Beschränkungen dar, die das Verhalten des Systems bestimmen. Der erste Teil spiegelt die Massenerhaltung wieder mit den stöchiometrischen Parametern $\nu_{i,j}$. Der zweite Teil repräsentiert das Verhalten des Systems im Gleichgewichtszustand durch die Kapazitäten C . Dieser Teil legt auch die Richtung des Reaktionsflusses fest. Im dritten Teil wird das Verhalten des Modells in Zuständen außerhalb des Gleichgewichtes mit den Widerständen R festgelegt. Nach weiteren Ersetzungen erhalten wir Molbilanzen, die nur noch von der Kapazität C_i , der aktuellen Konzentration c_i und den Widerständen R_j der Reaktionen abhängig sind. Übertragen auf das Beispiel aus 2.1 müssen die Molbilanzen in Gleichung 2.2 ergänzt werden mit den thermokinetischen Potentialen

$$\begin{aligned} C_A \cdot \xi_A &= c_A & C_B \cdot \xi_B &= c_B & C_C \cdot \xi_C &= c_C \\ C_{AB} \cdot \xi_{AB} &= c_{AB} & C_{BC} \cdot \xi_{BC} &= c_{BC} & C_{ABC} \cdot \xi_{ABC} &= c_{ABC} \end{aligned} \quad (2.15)$$

und den kinetischen Ratengleichungen

$$\begin{aligned} R_1 \cdot J_1 &= \xi_A \cdot \xi_B - \xi_{AB} & R_2 \cdot J_2 &= \xi_{AB} \cdot \xi_C - \xi_{ABC} \\ R_3 \cdot J_3 &= \xi_B \cdot \xi_C - \xi_{BC} & R_4 \cdot J_4 &= \xi_A \cdot \xi_{BC} - \xi_{ABC} \end{aligned} \quad (2.16)$$

und führen zu den Simulationsgleichungen

$$\begin{aligned} \dot{\xi}_A \cdot C_A &= -\frac{\xi_A \cdot \xi_B - \xi_{AB}}{R_1} - \frac{\xi_A \cdot \xi_{BC} - \xi_{ABC}}{R_4} \\ \dot{\xi}_B \cdot C_B &= -\frac{\xi_A \cdot \xi_B - \xi_{AB}}{R_1} - \frac{\xi_B \cdot \xi_C - \xi_{BC}}{R_3} \\ \dot{\xi}_C \cdot C_C &= -\frac{\xi_B \cdot \xi_C - \xi_{BC}}{R_3} - \frac{\xi_{AB} \cdot \xi_C - \xi_{ABC}}{R_2} \\ \dot{\xi}_{AB} \cdot C_{AB} &= +\frac{\xi_A \cdot \xi_B - \xi_{AB}}{R_1} - \frac{\xi_{AB} \cdot \xi_C - \xi_{ABC}}{R_2} \\ \dot{\xi}_{BC} \cdot C_{BC} &= +\frac{\xi_B \cdot \xi_C - \xi_{BC}}{R_3} - \frac{\xi_A \cdot \xi_{BC} - \xi_{ABC}}{R_4} \\ \dot{\xi}_{ABC} \cdot C_{ABC} &= +\frac{\xi_{AB} \cdot \xi_C - \xi_{ABC}}{R_2} + \frac{\xi_A \cdot \xi_{BC} - \xi_{ABC}}{R_4} \end{aligned} \quad (2.17)$$

mit dem Zustandsvektor $(\xi_A, \xi_B, \xi_C, \xi_{AB}, \xi_{BC}, \xi_{ABC})$ und den Systemparametern $C_A, C_B, C_C, C_{AB}, C_{BC}, C_{ABC}, R_1, R_2, R_3$ und R_4 . Thermokinetische Modelle (TKM) können mit einem Tupel von Matrizen und Funktionen dargestellt werden. Im Beispiel aus 2.1 würde das das Modell mit dem Tupel in Gleichung 2.18 beschrieben werden. So können sie, ähnlich linearen Gleichungssystemen, durch Äquivalenzumformungen manipuliert werden. Erfüllen die Matrizen und Gleichungen, die das TKM darstellen, gewisse Randbedingungen, kann die Modellgröße mit geeigneten Transformations- und Reduktionsschritten verringert werden. Im besonderen kann das Modell reduziert werden, wenn es Erhaltungsbeziehungen enthält oder ein Widerstand oder eine chemische

Kapazität gleich Null ist. $R_j = 0$ entspricht einem schnellen Reaktionsgleichgewicht, $C_i = 0$ quasi-stationären Reaktanden [3].

$$M = [N_E, N_P, c(\xi), R, \xi_0]$$

mit

$$N_E = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad N_P = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

$$c(\xi) = (C_A \xi_A, C_B \xi_B, C_C \xi_C, C_{AB} \xi_{AB}, C_{BC} \xi_{BC}, C_{ABC} \xi_{ABC})$$

$$R = (R_1, R_2, R_3, R_4)$$

$$\xi_0 = (\xi_{A,0}, \xi_{B,0}, \xi_{C,0}, \xi_{AB,0}, \xi_{BC,0}, \xi_{ABC,0}) \quad (2.18)$$

Während bei der kinetischen Modellierung irreversible Prozesse explizit mit stets nicht-negativen Kinetiken (wie der Michaelis-Menten-Kinetik $J = J_{max} \cdot c_S / (k_m + c_S)$ mit der Substratkonzentration c_S und den Parametern J_{max} und k_m) modelliert werden können, kennt die *thermokinetische Modellierung* keine irreversiblen Reaktionen. Diese stellen den Grenzfall einer reversiblen Reaktion dar, für die $F_E \gg F_P$ gilt und damit auch $\prod_{i \in E_j} \xi_i^{|\nu_{i,j}|} \gg \prod_{i \in P_j} \xi_i^{|\nu_{i,j}|}$.

2.2 Das System in der grafische Darstellung

Bei der Definition der thermokinetischen Modellierung sind die Begriffe Potential, Widerstand und Kapazität schon gefallen, deswegen liegt der Vergleich mit elektrotechnischen Modellen nicht fern. So kann man auch den Fluss J mit dem elektrischen Strom I vergleichen. Diese Analogien verleiten dazu, bei der grafischen Darstellung von TKM auf die etablierten Darstellung von Kapazitäten und Widerständen aus der Elektrotechnik zurückzugreifen. In Abbildung 2.1 wird eine Komponente gleich einer Kapazität dargestellt, eine Reaktion wird als Widerstand illustriert, jeweils mit den Gleichungen welche die jeweiligen Eigenschaften beeinflussen. Der Stöchior ist entfernt verwandt mit einem idealen Transformator. [4] Im Weiteren werden Ähnlichkeiten und Unterschiede betrachtet.

Erhaltungsbeziehungen. Während in elektronischen Schaltkreisen nur ein Ladungsträger beachtet werden muss, muss in chemischen Reaktionsnetzwerken je eine Erhaltungsbeziehung für jeden konservierten Rest (z.b. chemische Elemente) aufgestellt werden. In biochemischen Netzwerken gibt es in der Regel mehr Erhaltungsbeziehungen, da die beteiligten Komponenten konservierte Anteile besitzen.

Nicht-Linearität. Elektrotechnische Modelle können mit linearen Modellgleichungen beschrieben werden, wenn man von einem System mit idealen Leitern, Widerständen

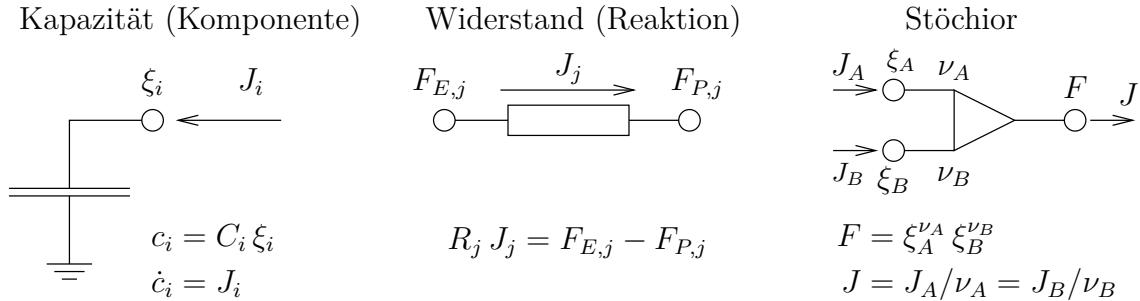


Abbildung 2.1: Grundlegende Elemente der Thermokinetischen Modellierung. Folgende Symbole wurden benutzt: Konzentration c_i , Thermokinetisches Potential ξ_i , Kapazität C_i , Fluss J_j , thermokinetische Kraft F_j , thermokinetische Kräfte der Edukte/Produkte $F_{E/P,j}$, Widerstand R_j , stöchiometrische Koeffizienten $|\nu_i|$. Die Grafik ist entnommen aus Thermokinetic Modeling and Model Reduction of Reaction Networks [3].

und Kapazitäten ausgeht. Durch die Einführung des Stöchiors und der Einwirkung stöchiometrischer Koeffizienten auf die treibende Kraft F können bei der *Thermokinetischen Modellierung* trotz idealer Rahmenbedingungen nicht-lineare Gleichungssysteme entstehen.

Knotenpunktsatz - erstes Kirchhoff'sches Gesetz. Kirchhoff postulierte, dass an einem Ladungsfreien Knotenpunkt in einem System die Summe der eingehenden und ausgehenden Flüsse Null sein muss. Dies gilt auch für thermokinetische Modelle, da dies direkt aus dem Satz der Massenerhaltung hervorgeht.

Maschensatz - zweites Kirchhoff'sches Gesetz. Des Weiteren postulierte Kirchhoff, dass die Summe der Spannungsdifferenzen zwischen den Knotenpunkten einer Masche gleich Null ist. Dasselbe gilt für die chemischen Potentialdifferenzen in einem chemischen Reaktionskreislauf. Definieren wir die chemische Potentialdifferenz als $\Delta\mu_j = R \cdot T \cdot \log(F_{E,j}/F_{P,j})$ und übersetzen Kirchhoff's Maschensatz auf Beispiel 2.1, können wir folgendes sagen: $\Delta\mu_1 + \Delta\mu_2 - \Delta\mu_3 - \Delta\mu_4 = 0$. Nach einer Umformung mithilfe der Logarithmusgesetze schließlich ergibt sich: $F_{E,1}/F_{P,1} \cdot F_{E,2}/F_{P,2} \cdot F_{E,3}/F_{P,3} \cdot F_{E,4}/F_{P,4} = 1$ [4].

Ersatzwiderstand. In der Elektrotechnik vereinfacht die Verwendung von Ersatzwiderständen die Handhabung von komplexen Netzwerken. Die grundlegenden Regeln gelten auch für die *Thermokinetische Modellierung*. (Abbildung 2.2 A) Angenommen die Stöchiometrien von Reaktion 1 und 2 seien gleich. So gilt $F = F_1 = F_2$. Der Gesamtfluss ist die Summe der Einzelflüsse $J = J_1 + J_2 = (1/R_1 + 1/R_2) \cdot F$. Der Gesamtwiderstand ist somit $1/R = 1/R_1 + 1/R_2$. (Abbildung 2.2 B) Angenommen die Produkte von Reaktion 1 seien die Edukte von Reaktion 2, sodass gilt $F_{P,1} = F_{E,2}$. Im Quasi-Gleichgewicht gilt außerdem $J = J_1 = J_2 = F/(R_1 + R_2)$. So ist der Gesamtwiderstand $R = R_1 + R_2$.

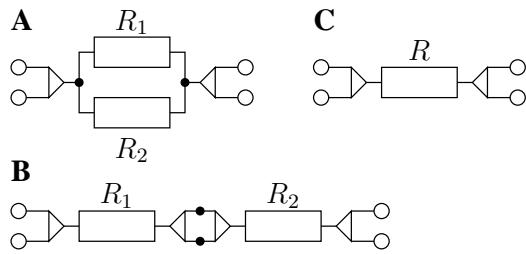


Abbildung 2.2: Parallel- (A) und Serienschaltung (B) von Widerständen. Die Schaltkreise können ersetzt werden durch je einen einzelnen Ersatzwiderstand (C) mit $1/R = 1/R_1 + 1/R_2$ bzw. $R = R_1 + R_2$. Die Grafik ist entnommen aus Thermokinetic Modeling and Model Reduction of Reaktion Networks [3].

2.3 Potential-Strom-Kennlinien

Eines der Ziele, die das Forschungsgebiet der Biotechnologie verfolgt, ist die Optimierung von Biosynthesewegen. Alle anabolischen Pfade beziehen ihre Ressourcen in Form von Vorläufermolekülen aus dem Zentralstoffwechsel. Ein hoher Biosynthesestrom wirkt sich allerdings negativ auf das Potential der beteiligten Vorläufermoleküle aus. Zum einen erschwert dies die Aufrechterhaltung eines hohen Stoffflusses durch den untersuchten und andere Biosynthesewege, zum anderen kann auch die Effizienz oder sogar die Funktionsfähigkeit des Zentralstoffwechsels dadurch beeinträchtigt werden. Vor allem bei zentralen Vorläufermolekülen ist es deswegen interessant, das thermodynamische Potential einer chemischen Spezies bei unterschiedlichen Abflusströmen zu beobachten.

Hier lassen sich thermodynamische Potentiale wieder mit elektrotechnischen Spannungsquellen vergleichen. Auch bei diesen sinkt die verfügbare Spannung an der Quelle mit der Erhöhung des abgeleiteten Stroms. Die Charakterisierung einer Spannungsquelle erfolgt über Spannung-Strom-Kennlinien (siehe Abbildung 2.3 (b)). Eine gute Spannungsquelle kennzeichnet sich durch eine konstante Spannung über weite Bereiche des entnommenen Stroms.

Der Verlauf der Potential-Strom-Kennlinie einer chemischen Spezies in einem thermodynamischen Modell folgt aus dem thermodynamischen Potential, das sich stationär einstellt, wenn die chemische Komponente mit einem bestimmten Strom entnommen wird. Somit stellen Potential-Strom-Kennlinien keinen zeitlichen Verlauf der Simulation dar. Vielmehr entspricht jede Koordinate auf der Kennlinie einem Gleichgewichtszustand bei einem eingestellten Abflusstrom.

Da komplexere biologische Systeme sich in der Regel nicht linear verhalten, können in dem mathematischen Modell Grenzzyklen, Multistabilität oder andere Attraktoren auftreten. Die dann durch das oben beschriebene Verfahren erstellte Kennlinie beschreibt dann den Verlauf der stationären Punkte, die sich quasi-stationär, aus dem ursprünglichen System ohne Abfluss des Vorläufermoleküls, ergeben. So können auch Kennlinien

aufreten die sich in sich zurück biegen. Treten im Modell Bifurkationen auf, kann die so berechnete Kennlinie auch mehrere Äste aufweisen.

Im Rahmen dieser Arbeit wurde bei der Erstellung von Potential-Strom-Kennlinien der Abflussstrom stets normiert auf einen Liter Reaktorvolumen, da dies produktionstechnisch am interessantesten ist. Sofern nicht anders vermerkt, ist dieser gemeint, wenn von einem Strom die Rede ist. Jedoch kann auch die Normierung auf die Zelltrockenmasse von Interesse und unter Umständen bei der Diskussion bestimmter Systemverhalten behilflich sein.

Die Simulationssoftware liefert nicht das thermokinetische Potential der beteiligten chemischen Spezies, sondern nur die Konzentration dieser. Da im Modell die Kapazität C_i einer chemischen Spezies A_i konstant ist, gilt laut Gleichung 2.14 $\xi_{i,t}/\xi_{i,t=0} = c_{i,t}/c_{i,t=0}$. In den Potential-Strom-Kennlinien werden somit Konzentrationen genormt auf die Anfangskonzentration dargestellt, die den thermokinetischen Potentialen genormt auf das Anfangspotential entsprechen.

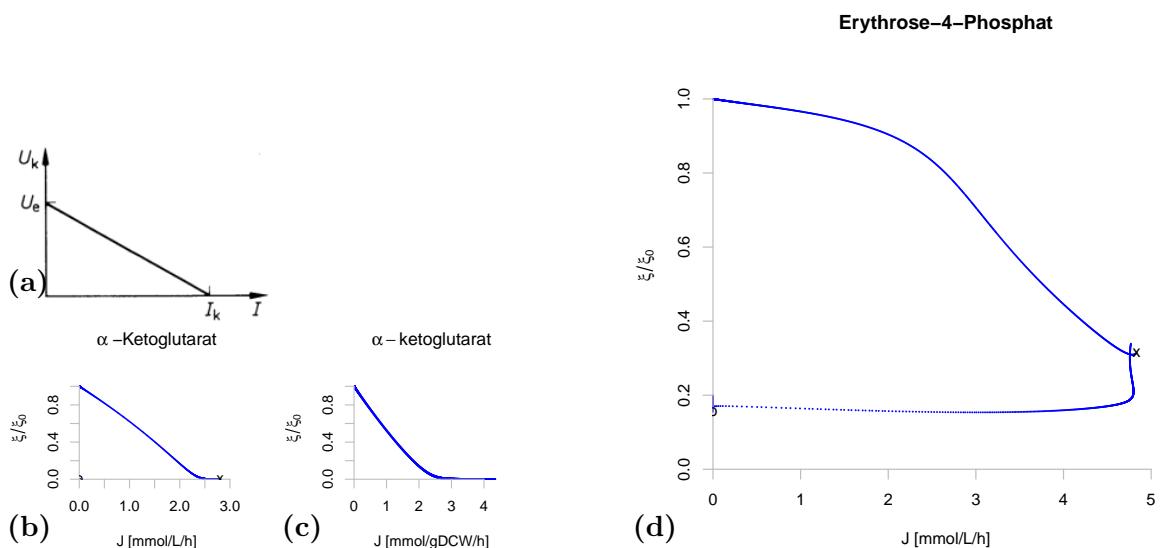


Abbildung 2.3: (a) Spannung-Strom-Kennlinie einer linearen Stromquelle aus der Elektrotechnik. Aufgetragen ist die Klemmenspannung U_k über dem Strom I , mit der Leerlaufspannung U_e und dem Kurzschlusstrom I_k . [12] (b) Potential-Strom-Kennlinie von α -Ketoglutarat akg aus dem thermokinetischen Modell von $E. coli$. Aufgetragen ist das normierte thermokinetische Potential $\xi/\xi_0 = \xi_i/\xi_{i,t=0}$ über den auf einen Liter Reaktorvolumen normierten Fluss J in mmol/L/h. (c) Potential-Strom-Kennline aus dem gleichen System wie (b). Der Fluss ist hier normiert auf die Zelltrockenmasse (DCW). (d) Potential-Strom-Kennlinie von Erythrose-4-Phosphat $e4p$. Der Strom ist hier wieder genormt auf einen Liter Reaktorvolumen. Im Gegensatz zur Kennlinie in (b), besitzt diese Kennlinie mehrere stationäre thermodynamische Potentiale ξ/ξ_0 zu einem gewissen normierten Abflusstrom J . Die Kennlinie beginnt bei $\xi/\xi_0 = 1$. Die Marker x bzw. \circ bezeichnen den Punkt maximalen Abflusstromes bzw. maximaler Ratenkonstante des Abflusstromes (Details in Abschnitt 4.2).

2.4 Niedermolekulare Bausteine

Eines der Ziele dieser Arbeit ist es, Potential-Strom-Kennlinien von niedermolekularen Bausteinen zu berechnen und zu analysieren. Hierzu musste eine zusätzliche Annahme eingeführt werden, welche am Beispiel von Isoleucin erläutert werden soll.

Die vereinfachte Reaktionsgleichung für die Isoleucinsynthese lautet wie folgt:



Die korrekte Berechnung des Potentials ξ_{Ile} des Biosyntheseproduktes Isoleucin *ile* erfordert ein detailliertes Modell des entsprechenden Biosyntheseweges. Da dieses nicht vorliegt, wird angenommen, dass die Reaktionen von den Vorläufermolekülen zum Biosyntheseprodukt im schnellen Gleichgewicht vorliegen. Das thermodynamische Potential von Isoleucin berechnet sich demnach als Produkt der Eduktpotentiale, potenziert mit ihren stöchiometrischen Koeffizienten:

$$\xi_{\text{Ile}} = \frac{\xi_{\text{ATP}}^2 \cdot \xi_{\text{NADPH}}^5 \cdot \xi_{\text{NH}} \cdot \xi_{\text{OAA}} \cdot \xi_{\text{PYR}}}{\xi_{\text{ADP}}^2 \cdot \xi_{\text{NADP}}^5} \quad (2.20)$$

Somit lässt sich jedoch nur das maximal mögliche Potential des Bausteines errechnen. Da die Annahme des schnellen Gleichgewichtes für die meisten Biosyntheseprodukte nicht erfüllt ist, liegt das reale Potential unterhalb dieses Wertes. Das berechnete Potential ist nichtsdestotrotz ein Maß dafür, wie sich die Verfügbarkeit des Bausteines mit dem Biosynthesefluss ändert.

Einzelne Bausteine wurden jedoch in der Berechnung vernachlässigt. In den vereinfachten Reaktionsgleichungen aus Neidhardt [9] kommt zum Beispiel auch atomarer Kohlenstoff *c* vor. Dieser kommt in biologischen Systemen nicht atomar vor sondern wird von verschiedenen Transportmolekülen, wie dem Tetrahydrofolat, im *C1-Stoffwechsel* zur Verfügung gestellt, welcher jedoch nicht modelliert ist. Ebenso ist der Sulfid *s* und Ammonium *nh* Transport nicht modelliert.

Es wird angenommen, dass die drei erwähnten Spezies ein konstantes Potential aufweisen. Bei der Darstellung der Potential-Strom-Kennlinien wird das Verhältnis $\xi_{\text{Ile}}/\xi_{\text{Ile},0}$ aufgetragen, so dass der unbekannte Proportionalitätsfaktor nicht entscheidend ist.

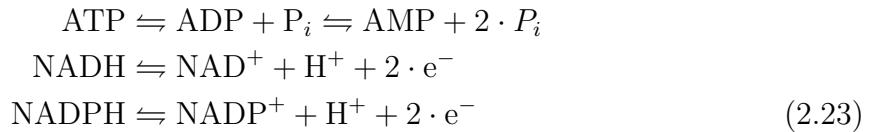
$$\xi_{\text{Ile}} \sim \frac{\xi_{\text{ATP}}^2 \cdot \xi_{\text{NADPH}}^5 \cdot \xi_{\text{OAA}} \cdot \xi_{\text{PYR}}}{\xi_{\text{ADP}}^2 \cdot \xi_{\text{NADP}}^5} \quad (2.21)$$

Verallgemeinert lässt sich das maximale Potential ξ_{bb} eines niedermolekularen Bausteines *A_{bb}* mit den Edukten *E_{bb}* und deren stöchiometrischen Koeffizienten $\nu_{i,bb}$ in folgender Form abschätzen:

$$\xi_{bb} \sim \prod_{i \in E_{bb}} \xi_i^{\nu_{i,bb}} \quad (2.22)$$

2.5 Coenzyme

Bei der Biosynthese werden für den Aufbau von niedermolekularen Bausteinen einerseits *Vorläufermoleküle* aus den zentralen Stoffwechselwegen verwendet, andererseits werden für viele der Umformungs- und Vernüpfungsreaktionen auch Coenzyme benötigt. Coenzyme sind Edukte der Biosynthesereaktionen welche nicht im Reaktionsprodukt aufgehen sondern eine funktionelle Gruppe abgeben oder aufnehmen. Ein Coenzym kommt dementsprechend in mindestens zwei Formen vor: mit und ohne funktioneller Gruppe. Beispiele sind hierfür ATP/ADP/AMP sowie die Coenzympaare NADH/NAD und NADPH/NADP:



Für die Erstellung einer Potential-Strom-Kennlinie für Coenzyme ist deswegen zu beachten, dass bei einem Abstrom von z.B. ATP ein äquivalenter Gegenstrom von ADP und inorganischem Phosphat existiert. Im Modell ist der ATP/ADP-Gradient ξ_{ATP}/ξ_{ADP} und der ADP/AMP-Gradient ξ_{ADP}/ξ_{AMP} aufgrund eines schnellen Gleichgewichtes identisch. Aus diesem Grund ist es ausreichend, den ATP/ADP-Gradienten mit ATP als Edukt und ADP als Produkt zu betrachten. Bei NAD und NADP wird sowohl die oxidierte als auch die reduzierte Form verwendet.

Für die Betrachtung des Potential-Gradienten der Coenzyme wird das Potentialverhältnis ζ eingeführt:

$$\begin{aligned} \zeta_{ATP} &= \frac{\xi_{ATP} \cdot \xi_{ADP,t=0}}{\xi_{ADP} \cdot \xi_{ATP,t=0}} \\ \zeta_{NADH} &= \frac{\xi_{NADH} \cdot \xi_{NAD^+,t=0}}{\xi_{NAD^+} \cdot \xi_{NADH,t=0}} \\ \zeta_{NADPH} &= \frac{\xi_{NADPH} \cdot \xi_{NADP^+,t=0}}{\xi_{NADP^+} \cdot \xi_{NADPH,t=0}} \end{aligned} \quad (2.24)$$

Bei der Erstellung der Potential-Strom-Kennlinien für die betreffenden Moleküle wird das Potentialverhältnis ζ statt des normierten Potentials ξ/ξ_0 gegen den Produktstrom J_{PROD} geplottet.

Bei Potential-Strom-Kennlinien für *niedermolekulare Bausteine* muss unterschieden werden, ob auf einem Biosyntheseweg, der eines der oben genannten Moleküle benötigt, eine funktionelle Gruppe freigesetzt wird oder das Coenzym in das komplexe Molekül als Bestandteil eingeht. Dem entsprechend wird dann in Gleichung 2.21 ξ_{ATP}/ξ_{ADP} durch ζ_{ATP} ersetzt. Ebenso wird bei NADP⁺/NADPH verfahren:

$$\xi_{Ile} \sim \zeta_{ATP}^2 \cdot \zeta_{NADPH}^5 \cdot \xi_{OAA} \cdot \xi_{PYR} \quad (2.25)$$

2.6 Zielsetzung

Ziel dieser Arbeit ist es, das Verhalten von Potentialen von chemischen Komponenten im *thermodynamischen Modell* von *Escherichia coli* bei verschiedenen Abflussströmen zu untersuchen.

Es soll ein Programm entwickelt werden, welches automatisiert eine Reihe von Simulationen startet, und die Ergebnisse in grafischer Form präsentiert. Bei der Simulation können numerische Probleme auftreten, da die Größenordnung der auftretenden Flüsse und Potentiale nicht im Vorfeld abgeschätzt werden können. Diese Probleme sollen von dem Programm erkannt und umgangen werden. Zu der Auswertung soll vor allem die Erstellung von Potential-Strom-Kennlinien für alle zentrale Vorläufermoleküle sowie für alle proteinogene Aminosäuren gehören. Ob die Kennlinien einen aussagekräftigen Verlauf zeigen, ist abhängig von einigen Parametern, die das Programm nach Möglichkeit auch automatisiert optimieren soll.

Die Vorläufermoleküle und Aminosäuren sollen nach ihrem Kennlinienverlauf eingruppiert, die Verläufe diskutiert und Erklärungsansätze geliefert werden.

Schließlich soll ein Ansatz untersucht werden, welcher es erlaubt, die Potential-Strom-Kennlinien von komplexen Syntheseprodukten, wie Aminosäuren und Ribonucleotiden, anhand ihrer Zusammensetzung aus Vorläufermolekülen abzuschätzen.

Kapitel 3

Das Modell in TKMOD

Die von Michael Ederer entwickelte *Mathematica-Paket TKMOD* begann als ein Modul, um das Erstellen und die Simulation von Modellgleichungen komplexer Reaktionssysteme im thermokinetischen Modellformalismus zu vereinfachen. Inzwischen besteht die Erweiterung aus einer Vielzahl an miteinander interagierenden Modulen, die hier anhand dem dieser Arbeit zu Grunde liegenden Modell näher vorgestellt werden sollen.

Untersucht wurde ein von Michael Ederer entwickeltes Modell, welches eine *E. coli* Kultur in einem Chemostat beschreibt. Das Modell wurde im Rahmen dieser Arbeit erweitert.

Zunächst kann man ein *TKMOD* Projekt in drei Teile gliedern: das *TKMOD*-Modul, das Modell und ein Start- und Analyseskript.

Das *TKMOD*-Modul, welches das *Mathematica* Modell anhand von Benutzereinstellungen in den Konfigurationsdateien `rules.mk` und `settings.mk` in eine ausführbare `*.exe` Datei überführt wird im Kapitel 3.1 erläutert. Das Modell `model.mma` an sich, formuliert in der Sprache von *Mathematica* wird im allgemeinen angesprochen in Kapitel 3.2 und im speziellen in Kapitel 3.3. Schließlich wurde in dieser Arbeit noch ein auführendes und analysierendes Modul `resmeas.r` in der Programmiersprache *R* erstellt, welches die Aufführbare Datei `model.exe` mit verschiedenen Parametern aus Parameterdateien `*.par` ausführen kann und die vom Modell generierten Daten `model.dat` verarbeitet. Die Funktionen und die Entwicklung dieses Modules wird in Kapitel 4 näher betrachtet. Die Interaktionen zwischen den einzelnen Bestandteilen des Paketes sind in Abb. 3.1 zusammengefasst.

3.1 Das *TKMOD*-Modul

Das Modul *TKMOD* ist eine Sammlung aus Skripten in den Sprachen *Fortran*, *Mathematica* und *R*, gebündelt unter einem zentralen *MAKEFILE* Skript. Beim Aufrufen dieser werden Benutzereinstellungen aus `settings.mk` und Regeln aus `rules.mk` geladen. Erstere kann angepasste Systempfade beinhalten, während letztere Benutzerspezifische Dateierstellungsregeln enthält.

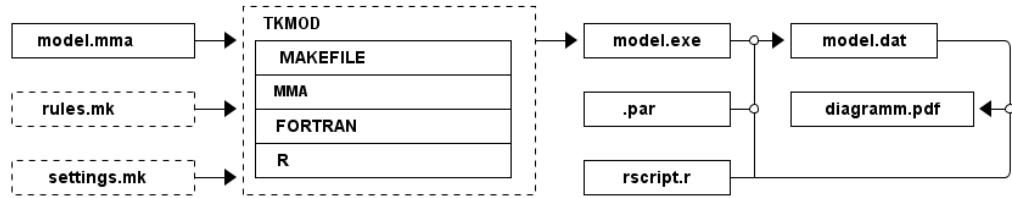


Abbildung 3.1: Die Interaktionen zwischen den *TKMOD* Modulen. Das *TKMOD*-Hauptmodul besteht aus einer Kombination aus *MAKEFILE*, *MMA*, *R* und *FORTRAN* Scripten, welche das Modell aus *model.mma* auslesen, und mit Hilfe von Benutzereinstellungen in *rules.mk* und *settings.mk* in die ausführbare Modelldatei *model.exe* überführen. Diese wird über die Kommandozeile oder über *resmeas.r* gestartet. Durch die Verwendung von Parameterdateien **.par*, welche beim Starten der Simulation eingelesen werden, bleibt das Modell auch nach dem Kompilieren flexibel. Das Analysemodul *resmeas.r* verwertet auch die in *ecoli.dat* enthaltenen Simulationsergebnisse zu Diagrammen in *diagramm.pdf*. Das Modell lag für diese Arbeit komplett vor. Ergänzungen wurden in *ecoli.mma* und *resmeas.r* vorgenommen, Details hierzu sind in Kapitel 4.1 aufgeführt.

Durch die Arbeit mit einem *MAKEFILE* Skript war die Entwicklung in einer Linuxumgebung notwendig. Es war jedoch ausreichend auf die Windows-basierte Linuxumgebung *CYGWIN* auszuweichen.

Wird beim Start des *MAKEFILE* Skriptes eine ausführbare Datei angefordert, startet ein *Mathematica* Skript, welches die Modelldatei *model.mma* einliest und in die Programmiersprache *Fortran* überführt. Im nächsten Zug wird das Ergebnis mit Hilfe von *DASKR* - einem Fortran-Paket zur Simulation von Differential-Algebra-Systemen [2] - zu der ausführbaren Simulationsdatei *model.exe* kompiliert. Diese Datei kann schließlich über die Kommandozeile oder über ein Skript, z.B. auch über das *MAKEFILE* Skript, gestartet werden. Der Verlauf der Simulation kann durch verschiedene Kommandozeilenparameter von *model.exe* auch nach der Kompilierung weitgehend beeinflusst werden. Die Ausgabedatei *model.dat* beinhaltet den zeitlichen Verlauf sämtlicher Konzentrationen und Flüsse, sowie Volumen, Temperatur und pH-Wert aller Kompartimente.

3.2 Ein Modell in *TKMOD*

Das thermokinetisches Modell wird als ein Objekt in *Mathematica* erstellt. Es setzt sich aus 4 Typen von Objekten zusammen: Parameter *P*, Zellkompartimente *Compartment*, Reaktanden *Compound* und Flüsse *Flux*. Diese wiederum besitzen unterschiedliche Eigenschaften, die am folgenden Beispiel, Quelltext 3.1 , erläutert werden.

Quelltext 3.1: Ein einfaches Beispiel für eine *TKMOD* Datei

```

1 NetworkDescription [ "simple circle" ,
{
5   Compartiment [ "c" ,
    "Size"          -> 1 ,
    "Scale"         -> 100 ,
    "ElectricalPotential" -> 0 ,
    "Temperature"   -> 298 ,
    "SizeDot"       -> 0 ,
    "Dilution"      -> 0 ] ,
10  Compound [ "A" ,
    "Capacity"      -> 0.8 ,
    "Potential"     -> P[ "potential_a" , "Value" ->3] ,
    "Compartiment"  -> "c" ,
    "Scale"         -> 1 ] ,
15  Compound [ "B" , ... ]
Compound [ "C" , ... ]
Flux [ "AB" ,
    "Educts"        -> "A" ,
    "Products"       -> "B" ,
    "Compartiment"  -> "c" ,
    "Resistance"    -> 1 ] ,
20  Flux [ "BC" , ... ]
Flux [ "CA" , ... ]
}]
```

Die Definition des Modelles beginnt mit dem Objekttyp **NetworkDescription**. Erstes Element des Objektes ist der Netzwerkname, im Falle des Beispieles “simple circle”. Es folgt eine Liste der Komponenten des Systems, welche wiederum Objekte sind und mit der Typezeichung anfangen - z.B. **Compartiment** - gefolgt vom Objektnamen - z.B. “c”. Nach diesem folgt eine Reihe von Zuweisungen von Werten zu Parametern. Verschiedene Elementtypen benötigen unterschiedliche Parameter. So Besitzt ein Element vom Typ **Compound** die Eigenschaft **Capacity**, während einem **Flux** eine **Resistance** zu eigen ist.

Ein Beispiel für die Benutzung von Parametern **P** ist auf der rechten Seite der Zuweisung von Zeile 12 in Quelltext 3.1 zu finden. Diese nehmen eine besondere Stellung im Modell ein, da sie auch nach dem Kompilieren des Modelles **model.mma** zu einer ausführbaren Simulationsdatei **model.exe** über Kommandozeilenparameter beeinflusst werden können. Die Zuweisung der Eigenschaft **Value** übergibt dem Parameter einen Standartwert.

3.3 Der *Escherichia coli*-Chemostat

Insgesamt bestand das ursprüngliche *thermokinetische Modell* des Reaktionsnetzwerkes von *E. coli* aus 6 Kompartimenten, 152 Reaktanden, 153 Stoffflüssen und 202 Parametern. Als zentrale Vorläufermoleküle wurden 15 Reaktanden näher untersucht (siehe Tabelle 3.1, Abschnitt Vorläufermoleküle). Im Rahmen dieser Studienarbeit wurde das Modell um einige Bausteine erweitert (siehe Kapitel 4.1).

Das Modell beschränkt sich auf folgende Stoffwechselwege und Reaktorflüsse:

- Transportwege am Reaktor, an der periplasmatischen Membran und an der Zellmembran
- Glykolyse/ Gluconeogenese
- Pentosephosphatweg
- Krebszyklus (Citratzyklus/TCA)
- Pyruvat Metabolismus
- Oxidative Phosphorylierung

Eine grafische Darstellung des Modellsystems ist in Abb. 3.2 zu sehen. In Tabelle 3.1 sind die wichtigsten chemischen Spezies mit ihrer Abkürzung und vollen Bezeichnung zu finden. Im Folgenden werden chemische Spezies vorwiegend mit ihren Kürzeln bezeichnet. Die vollständigen Bezeichnungen der Enzyme sowie die EC-Nummer und eine Verweis zu weiteren Informationen sind in der Modellbeschreibung `ecoli.mma` zu finden. In diesem Modell nehmen zwei Stoffflüsse eine gesonderte Rolle ein.

Der dem **Wachstum** entsprechende Fluss J_{GROWTH} stellt den Zuwachs an Zellmasse im Reaktor dar. Als Edukte E_{GROWTH} gehen hier alle 15 Vorläufermoleküle ein (NAD deprotoniert, NADP protoniert). Produkte P_{GROWTH} der Reaktion sind unter anderem Glucose-3-Phosphat, Acetat, Succinat, Fumarat und CO_2 . Die empirischen Parameter k_{GROWTH_A} , k_{GROWTH_B} und die stöchiometrischen Parameter $\nu_{i,GROWTH}$ fließen ebenfalls in die Berechnung des Flusses ein. Der Fluss J_{GROWTH} beeinflusst die Volumenänderung der verschiedenen Zellkompartimente in dem Modell.

$$J_{GROWTH} = k_{GROWTH_A} \cdot \left(k_{GROWTH_B} + \sum_{i \in E_{GROWTH}} \nu_{i,GROWTH} * \log(c_i) \right) \quad (3.1)$$

Im Modell ist die Möglichkeit eines **Produktionsstromes** implementiert, der dem System eine bestimmte chemische Spezies oder simulaten mehrere Spezies in einem bestimmten stöchiometrischen Verhältnis entzieht. Die Größe des Stromes J_{PROD} ist definiert als Produkt aus der Konzentration c_i der chemischen Komponente A_i , potenziert mit dem stöchiometrischen Parameter $\nu_{i,PROD}$, und der Ratenkonstante k_{PROD} . Sollen

dem System mehrere Stoffe mit den stöchiometrischen Parametern $\nu_{i,PROD}$ entzogen werden, so gilt folgende, verallgemeinerte Beschreibung für den Produktionsfluss:

$$J_{PROD} = k_{PROD} \cdot \prod_i c_i^{\nu_{i,PROD}} \quad (3.2)$$

Tabelle 3.1: Liste relevanter chemischer Spezies und ihrer Abkürzungen. Die vollen Namen aller im Modell auftauchenden Spezies sind in der Modelldatei `ecoli.mma` zu finden.

Vorläufermoleküle		Ribonucleotide		Weitere Bausteine	
3pg	3-Phosphoglycerat	met	Methionin	udpglc	UDP-Glucose
accoa	Acetyl-CoA	phe	Phenylalanin	etham	Ethanolamin
akg	α -Ketoglutarat	pro	Prolin	ohma	Hydroxy-Myristinsäure
atp	Adenosintriphosphat	ser	Serin	cmpkdo	Cytosinmonophosphat-3-deoxy-D-manno-octulosonate
dhap	Dihydroxyacetophosphat	thr	Threonin	ndphep	NDP-Heptose
e4p	Erythrose-4-Phosphat	trp	Tryptophan	tdpglca	TDP-Glucuronsäure
f6p	Fructose-6-Phosphat	tyr	Tyrosin	udpnacga	UDP-N-Acetylglucosamin
g6p	Glucose-6-Phosphat	val	Valin	udpnacma	UDP-N-Acetylmuraminsäure
nad(h)	Nicotinamidadenindinucleotid (protoniert)	Ribonucleotide		dap	Diaminopimelic acid
nadp(h)	Nicotinamidadenindinucleotidphosphat (protoniert)	atpdn	Adenosintriphosphat (de novo Synthese)	glc _m	Glucose(als Glycogen Monomer)
oaa	Oxalacetat	gtp	Guanosintriphosphat		
pep	Phosphoenolpyruvat	ctp	Cytosintriphosphat		
pyr	Pyruvat	utp	Uraciltriphosphat		
r5p	Ribulose-5-Phosphat	datp	Deoxyadenosintriphosphat		
succoa	Succinyl-CoA	dgtp	Deoxyguanosintriphosphat		
		dctp	Deoxycytosintriphosphat		
		dntp	Deoxytyrosintriphosphat		
Aminosäuren		Weitere Bausteine			
ala	Alanin	udpglc	UDP-Glucose		
arg	Arginin	etham	Ethanolamin		
asn	Asparagin	ohma	Hydroxy-Myristinsäure		
asp	Aspartat	cmpkdo	Cytosinmonophosphat-3-deoxy-D-manno-octulosonate		
cys	Cystein	ndphep	NDP-Heptose		
gln	Glutamin	tdpglca	TDP-Glucuronsäure		
glu	Glutamat	udpnacga	UDP-N-Acetylglucosamin		
gly	Glycin	udpnacma	UDP-N-Acetylmuraminsäure		
his	Histidin	dap	Diaminopimelic acid		
ile	Isoleucin	glc _m	Glucose(als Glycogen Monomer)		
leu	Leucin				
lys	Lysin				

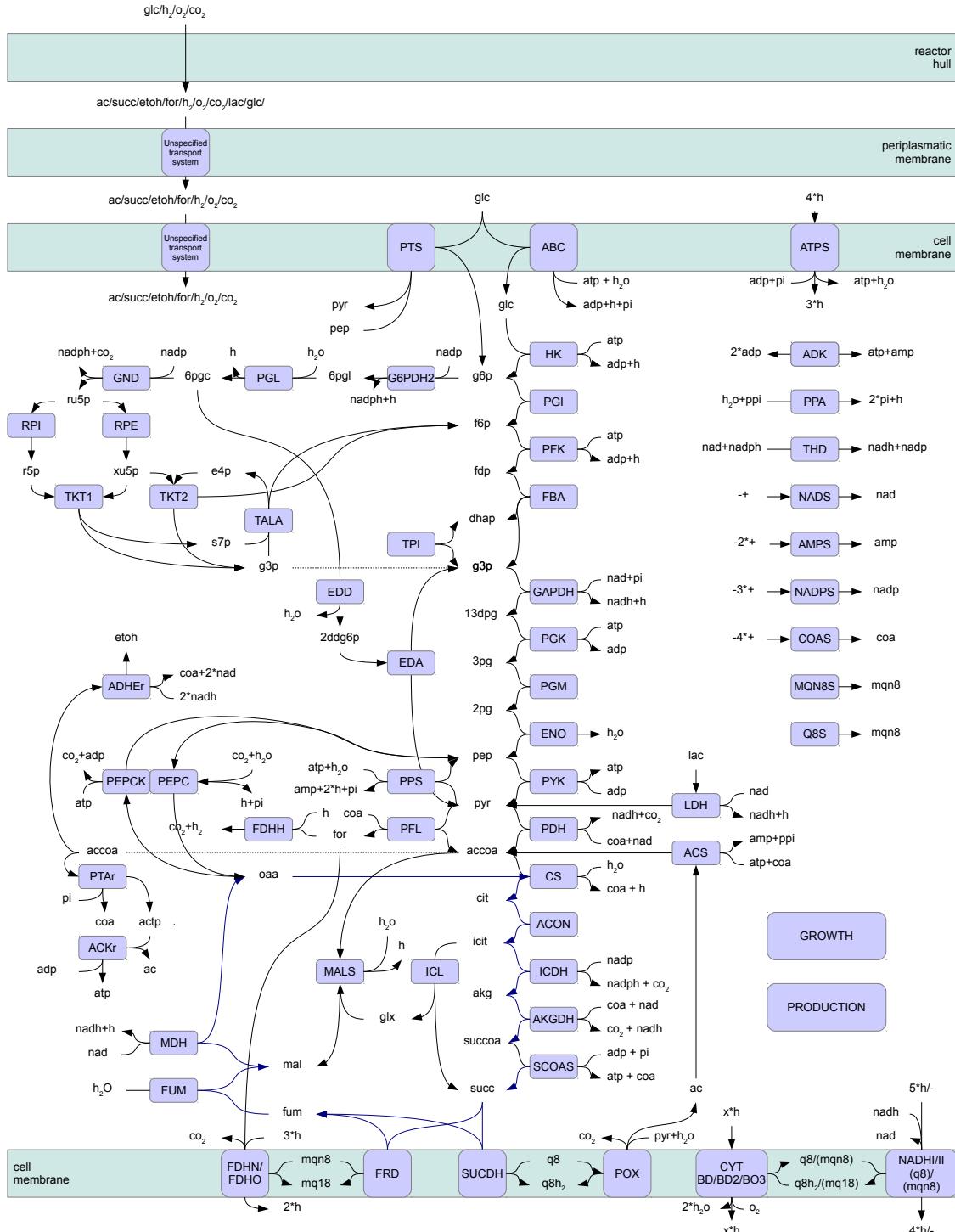


Abbildung 3.2: Das *E. coli* Modellsystem. Dargestellt sind alle chemischen Spezies (Compounds) ohne spezielle Markierung und die beteiligten Enzyme in Kästchen. Diese entsprechen, mit den zugehörigen Richtungspfeilen, den Flüssen aus dem Modell. Der TCA Zyklus ist mit blauen Flusspfeilen markiert. Die Compounds $\text{g}3\text{p}$ und accoa kommen, der Übersichtlichkeit halber, je zwei mal vor und sind mit einer gepunkteten Linie verbunden. Die Compounds h_2o , co_2 , h , pyr , pep , atp , adp , pi , nad , nadph und coa kommen ubiquitär vor und sind deshalb mehrfach dargestellt.

Kapitel 4

Umsetzung

Im nächsten Kapitel wird nun genauer darauf eingegangen, wie die gesetzten Ziele umgesetzt wurden, welche Probleme sich dabei in den Weg gestellt haben und wie diese gelöst wurden.

4.1 Ziel Eins: Implementierung von niedermolekularen Bausteinen

Das erste Ziel, das in dieser Studienarbeit verfolgt wurde, war die Erweiterung des Modells von *E. coli* um einige *niedermolekulare Bausteine*. Diese mussten einerseits als Objekt in der in Kapitel 3.2 eingeführten Form in das Modell implementiert werden. Zum anderen muss der in Gleichung 3.2 aufgeführte Produktionsfluss J_{PROD} um die niedermolekularen Bausteine erweitert werden.

Im bearbeiteten Modell von *E. coli* wurde die Erweiterung um niedermolekulare Bausteine durch ein zweites *Mathematica*-Skript realisiert. Ein Befehl innerhalb der Modelldatei `ecoli.mma` lädt die vereinfachten Reaktionsgleichungen für die niedermolekularen Bausteine aus der Datei `bb.mma`, und verwendet diese, um sowohl neue Komponenten für das Modell, als auch den Produkfluss J_{PROD} zu erstellen. Ein Beispiel für die Aminosäuren Alanin *ala* und Phenylalanin *phe* ist in Quelltext 4.1, Zeilen 1-2, zu sehen.

Die niedermolekularen Bausteine, die dem Modell hinzugefügt wurden, sind in Tabelle 3.1 unter den Überschriften **Aminsäuren**, **Ribonucleotide** und **Weitere Bausteine** aufgezählt.

In Kapitel 2.4 wurde auch erwähnt, das gewisse chemische Spezies bei der Berechnung vernachlässigt werden. Welche das im Detail sind, ist ebenfalls in der Datei `bb.mma` notiert, und zu sehen in Quelltext 4.1 Zeile 3.

Quelltext 4.1: Ein Beispiel für eine Zusammensetzungsgesetz in `bb.mma`

```
1 BB = { "ala" -> "nadph" - "nadp" + "nh" + "pyr" ,  
         "phe" -> "atp" - "adp" + 2 "nadph" - 2 "nadp" + "nh" \\ + "e4p" + 2 "pep" }  
disregard = {"s" -> 0, "c" -> 0, "nh" -> 0 };
```

Die Berechnung des **Potentials eines niedermolekularen Bausteines** erfolgt über die in Gleichung 2.22 angegebene Formel. Über eine Schleife im *Mathematica*-Skript wurde für jeden betrachteten Baustein ein Objekt in *Mathematica* angelegt, welches der in Kapitel 3.2 vorgestellten Form entspricht.

Der in Gleichung 3.2 beschriebene **Produktionsstrom** musst erweitert werden, um die Produktion von Bausteinen zu ermöglichen. In Kapitel 2.4 wurde die Annahme eingeführt, dass die Produktionswege von den Vorläufermolekülen zum niedermolekularen Baustein im schnellen Gleichgewicht vorliegen. Dieser Annahme folgend, entspricht die Entnahme eines niedermolekularen Bausteines aus dem System der Entnahme seiner Vorläufermoleküle. Somit kann die in Gleichung 3.2 dargestellte Formel verwendet werden, um den Produktionsstrom zu berechnen, wobei die stöchiometrischen Produktionskoeffizienten $\nu_{i,PROD}$ durch die stöchiometrischen Biosynthesekoeffizienten $\nu_{i,bb}$ ausgetauscht werden. Werden mehrere niedermolekulare Bausteine A_{bb} gleichzeitig abgezogen, müssen die stöchiometrischen Parameter $\nu_{i,bb}$ der Vorläufermoleküle A_i über alle produzierten niedermolekularen Bausteine A_{bb} aufsummiert werden. Demnach berechnet sich der Produktionsstrom nach folgender verallgemeinerten Gleichung:

$$J_{PROD} = k_{PROD} \cdot \prod_i c_i^{\sum \nu_{i,bb}} \quad (4.1)$$

4.2 Ziel Zwei: Berechnung von Potential-Strom-Kennlinien

Das Hauptziel der Arbeit war es, Potential-Strom-Kennlinien für verschiedene *Vorläufermoleküle* und *niedermolekulare Bausteine* zu erstellen und zu analysieren.

Um diese Kennlinien zu erstellen, wurde das Modellsystem simuliert. Während der Simulation wird der Abflussstrom einer chemischen Spezies über den Simulationsverlauf erhöht und die zugehörige Konzentration gemessen. Damit gewährleistet wird, dass das Modell zu Simulationsbeginn in einem Gleichgewichtszustand vorliegt, wird der Abfluss erst nach einer gewissen Simulationszeit eingestellt. Eine hohe Gesamtzeit der Simulation erlaubt einen langsamem Anstieg des Abflussstromes und stellt sicher, dass sich während des Anstiegs ein Quasi-Gleichgewichtszustand einstellt.

Die Höhe des Abflusstromes wird durch den Parameter k_{PROD} eingestellt (siehe Gleichung 4.1). Dessen Maximalwert \tilde{k}_{PROD} wird bei Simulationsbeginn mit dem Kommandozeilenparameter $P["PROD_k"]$ und das Modell übergeben. Das Modell benutzt eine Hilfsvariable k_{sim} , um den Abflusstrom im Simulationsverlauf zu erhöhen und wieder zu verringern. Diese Veränderung kann linear über der Simulationszeit erfolgen, oder aber mit einer höheren Potenz. Eingestellt wird dies mit dem Parameter $P["ramp"]$. Die Berechnung des Stromes erfolgt also wie folgt:

$$0 \leq k_{sim}(t, P["ramp"]) \leq 1$$

$$J_{PROD} = k_{sim} \cdot \tilde{k}_{PROD} \cdot \prod_i c_i^{\sum \nu_{i,bb}} \quad (4.2)$$

In Gleichung 4.2 sieht man, dass der Produktstrom J_{PROD} ein Produkt ist aus der Variable \tilde{k}_{PROD} und mehreren Konzentrationen c_i , welche mit ihren stöchiometrischen Parametern $\nu_{i,bb}$ potenziert werden. Wenn die Konzentration einer Spezies unter einen bestimmten Minimalwert fällt, kann es zu Folge haben, dass das der Strom J_{PROD} trotz steigender Ratenkonstante k_{PROD} sinkt.

Während der Wert von k_{sim} in der ganzen Simulation zwischen 0 und 1 liegt, und die stöchiometrischen Parameter $\nu_{i,bb}$ in der Regel im Zahlenraum von 1 bis 10 liegen, bewegen sich die Konzentrationen c_i in einem Bereich von mehreren Zehnerpotenzen. Der Strom J_{PROD} schwankt, als Produkt, in noch größeren Bereichen, was zu unerwünschten Konsequenzen führen kann, die im folgenden erläutert werden.

Die Simulation kann vor dem Erreichen der angezielten Produktion bei $k_{sim} = 1$ abbrechen. Zumeist werden hier die numerischen Grenzen des Lösungsalgorithmus überschritten. Dies wiederum ist oft ein Indiz für unphysiologische Konzentrationen von chemischen Spezies.

Ist der maximale Strom sehr groß, d.h. \tilde{k}_{PROD} zu groß gewählt, können zwei verschiedene Probleme auftreten. Einerseits kann es zu numerischen Fehlern kommen, was zu einem vorzeitigen Abbruch der Simulation führt. Andererseits resultiert die Wahl eines hohen \tilde{k}_{PROD} -Wertes bei gegebenen Ausgabeschrittweite in einem großen Abstand zwischen den Quasi-Gleichgewichtszuständen. Die Potential-Strom-Kennlinie ist so möglicherweise nicht ausreichend aufgelöst und ein interessantes Verhalten des Systems wird übersehen.

Ist der maximale Strom zu gering, so läuft die Simulation ohne numerische Fehler ab, doch erreicht das Modell dann auch nicht seine maximal mögliche Produktion, und die untersuchten thermokinetischen Potentiale zeigen keine oder nur wenig Schwankungen.

Es ist nun wichtig, einen Wert für k_{PROD} zu finden, welcher den maximalen Produktionsstrom bei $k_{sim} = 1$ in einen experimentell interessanten Bereich rückt.

Da Konzentrationen von chemischen Komponenten in einem komplexen Reaktionsnetzwerk über viele Wege miteinander verknüpft sind, und sich gegenseitig beeinflussen, lässt sich der Parameter \tilde{k}_{PROD} nicht a priori abschätzen. Um einen geeigneten Skalierungsfaktor zu finden, müssen die Simulationsdaten bewertet werden und die Simulationen, falls nötig, mit einem veränderten Parameter \tilde{k}_{PROD} gestartet werden. In dieser Arbeit wurde eine Routine entwickelt, die automatische einen geeigneten \tilde{k}_{PROD} -Wert bestimmt.

Die Bewertung der Simulation erfolgt in dem analysierenden R-Skript `resmeas.r`. Zum einen wird der Verlauf der Variable k_{sim} aus einer abgeschlossenen Simulation zur Bewertung herangezogen. Erreicht diese nicht den Wert 1, so ist die Simulation vor dem

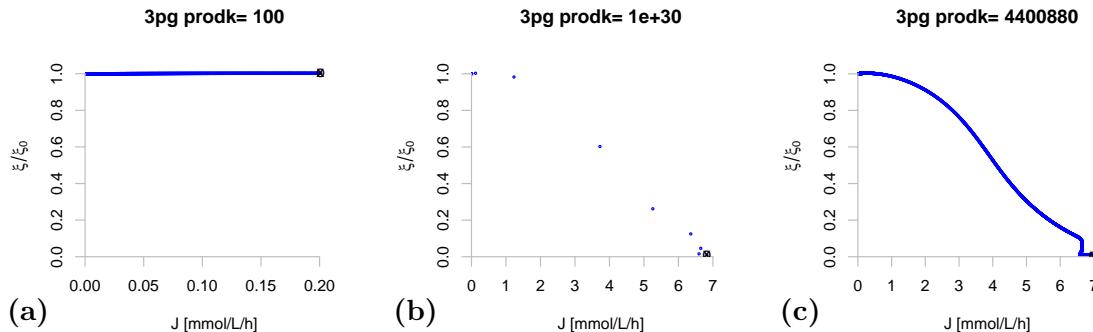


Abbildung 4.1: Beispiele für Potential-Strom-Kennlinien bei unterschiedlichen Geschwindigkeitsparametern \tilde{k}_{PROD} . Bei (a) ist \tilde{k}_{PROD} klein gewählt. Der resultierende Strom J hat so eine kaum erkennbare Auswirkung auf das normierte Potential ξ/ξ_0 . Bei (b) ist \tilde{k}_{PROD} hingegen sehr groß. Die Simulation erreicht in nur wenigen Schritten ihr Produktionslimit (x), und beschreibt das Verhalten des Modellsystems nur lückenhaft. Bei (c) ist der Parameter \tilde{k}_{PROD} geeignet gewählt, um die Potential-Strom-Kennlinie darzustellen.

Erreichen des angestrebten Produktionsflusses, in den allermeisten Fällen aufgrund numerischer Fehler, abgebrochen, und es kann davon ausgegangen werden, dass die der Produktionsfluss eine zu große Belastung für das System dargestellt hat. Für die nächste Simulation sollte demnach ein geringerer Wert für k_{PROD} gewählt werden. Erreicht k_{sim} den Wert 1, und ist auch der Produktstrom J_{PROD} zum gleichen Zeitpunkt maximal, ist die Simulation erfolgreich verlaufen, das Produktionslimit des Modellsystems jedoch nicht erreicht.

Das maximal mögliche k_{PROD} , mit dem die Simulation gerade noch vollständig durchläuft, ohne aufgrund numerischer Fehler abzubrechen, exakt zu erreichen, ist aufgrund der Modell- und Simulationsstruktur nicht möglich. Daher wird ein geeigneter Wert für k_{PROD} wie folgt definiert:

- k_{sim} erreichte mindestens den Wert 0.99
- der Zeitpunkt des höchsten Produktflusses liegt nicht bei $k_{sim} = 1$

Erweist sich k_{PROD} als zu hoch, so berechnet sich dessen Wert für die nächste Simulation als Produkt des maximal erreichten k_{sim} und dem letzten k_{PROD} . War k_{PROD} zu niedrig gewählt, wird der Wert mit einem vorher festgelegten Multiplikator vervielfacht.

Das Analyseskript in `resmeas.r` führt die Schritte Simulation - Analyse - Modifikation mehrmals hintereinander aus, und findet im Idealfall einen Wert für k_{PROD} , welcher der obigen Definition für einen geeigneten Wert entspricht.

Ist ein geeigneter Wert gefunden, so gibt das Skript die Simulationsdaten in einem Potential-Strom-Diagramm wider. In diesem Schritt wird auch das Potential ξ gegen das Startpotential ξ_0 und der Fluss J_{PROD} gegen das Reaktorvolumen normiert. Es erfolgt zudem eine Markierung der Meßwerte: (x) am Punkt des maximalen Produktionsflusses und (o) am Umkehrpunkt der Simulation, an dem k_{sim} den Wert 1 erreicht.

4.3 Ziel Drei: Schätzung von Kennlinien komplexer Bausteine

Viele Potentiale ξ_i in einem biochemischen Netzwerk sind voneinander abhängig, da sie sich biochemische Pfade teilen oder aufeinander aufbauen. Dies gilt auch für die *zentralen Vorläufermoleküle*, die alle dem zentralen Stoffwechsel entspringen. Wird dem Citratzyklus beispielsweise SuccinylCoA *succoa* entzogen, so sinkt auch das Potential von Oxalacetat *oaa*.

Wird dieser Sachverhalt vernachlässigt, so lässt sich eine Schätzung für das Potential eines Produktes, welches dem Modell mehrere *Vorläufermoleküle* gleichzeitig entzieht, berechnen. Für diese Schätzung werden Daten aus den Potential-Strom-Kennlinien der *Vorläufermoleküle* benutzt, das Modell muss hierfür nicht erweitert werden und es muss keine neue Simulation durchgeführt werden.

Als Grundlage für die Schätzung dient die Formel zur Berechnung des Potentials eines *Baustein* *bb* aus Gleichung 2.22. Um das Potential eines Produktes aus seinen Edukten zu berechnen, wird das Potential ξ als Funktion des zugehörigen Abstromes J_{PROD} betrachtet. Da der Strom $J_{i,PROD}$ der *Vorläufermoleküle* das jeweils um den stöchiometrischen Koeffizienten $\nu_{i,bb}$ vervielfachte des Stromes $J_{bb,PROD}$ des *Baustein* ist, muss auch das entsprechende Potential der Vorläufermoleküle zu einem erhöhten Strom betrachtet werden.

$$J_{i,PROD} = J_{bb,PROD} \cdot \nu_{i,bb} \quad (4.3)$$

Nehmen wir uns das Beispiel der Isoleucinsynthese aus Gleichung 2.19 zur Hilfe, so kann man wie folgt das Potential von Isoleucin *Ile* schätzen:

$$\xi_{Ile}(J_{PROD}) \sim \frac{\xi_{ATP}^2(2 \cdot J_{PROD}) \cdot \xi_{NADPH}^5(5 \cdot J_{PROD}) \cdot \xi_{OAA}(J_{PROD}) \cdot \xi_{PYR}(J_{PROD})}{\xi_{ADP}^2(2 \cdot J_{PROD}) \cdot \xi_{NADP}^5(5 \cdot J_{PROD})} \quad (4.4)$$

Allgemein lässt sich nun schreiben:

$$\xi_{bb}(J_{bb,PROD}) = \prod_{i \in E_{bb}} \xi_i(\nu_{i,bb} \cdot J_{bb,PROD})^{\nu_{i,bb}} \quad (4.5)$$

Weiter muss beachtet werden, dass einige Kennlinien von *Vorläufermolekülen* keinen stets abfallenden Verlauf aufweisen, und sich somit für eine Flussrate mehrere Potentiawerte vorhanden sein können. Deswegen werden für die Betrachtung nur J_{PROD} und ξ -Werte bis zum ersten Flussmaximum einbezogen.

Kapitel 5

Analyse und Diskussion A: Die Potential-Strom-Kennlinien

In diesem Kapitel werden nun die erstellten Potential-Kennlinien von *Vorläufermolekülen* und *niedermolekularen Bausteinen* vorgestellt und analysiert. Die Kennlinien werden, wenn möglich, gruppiert werden, und es werden Ansätze erörtert, die das unterschiedliche Verhalten der *Vorläufermolekülpotentiale* erklären. Hierzu werden zum Teil weitere Darstellungen des Simulationsverlaufes zur Hilfe genommen. Ebenfalls werden die Kennlinien teilweise auf den Verlauf bis zum Punkt des maximalen Produktionsflusses beschränkt, da darüber hinaus numerische/systematische Fehler auftreten, die das Potential unphysiologisch werden lassen.

5.1 Kennlinien der Vorläufermoleküle

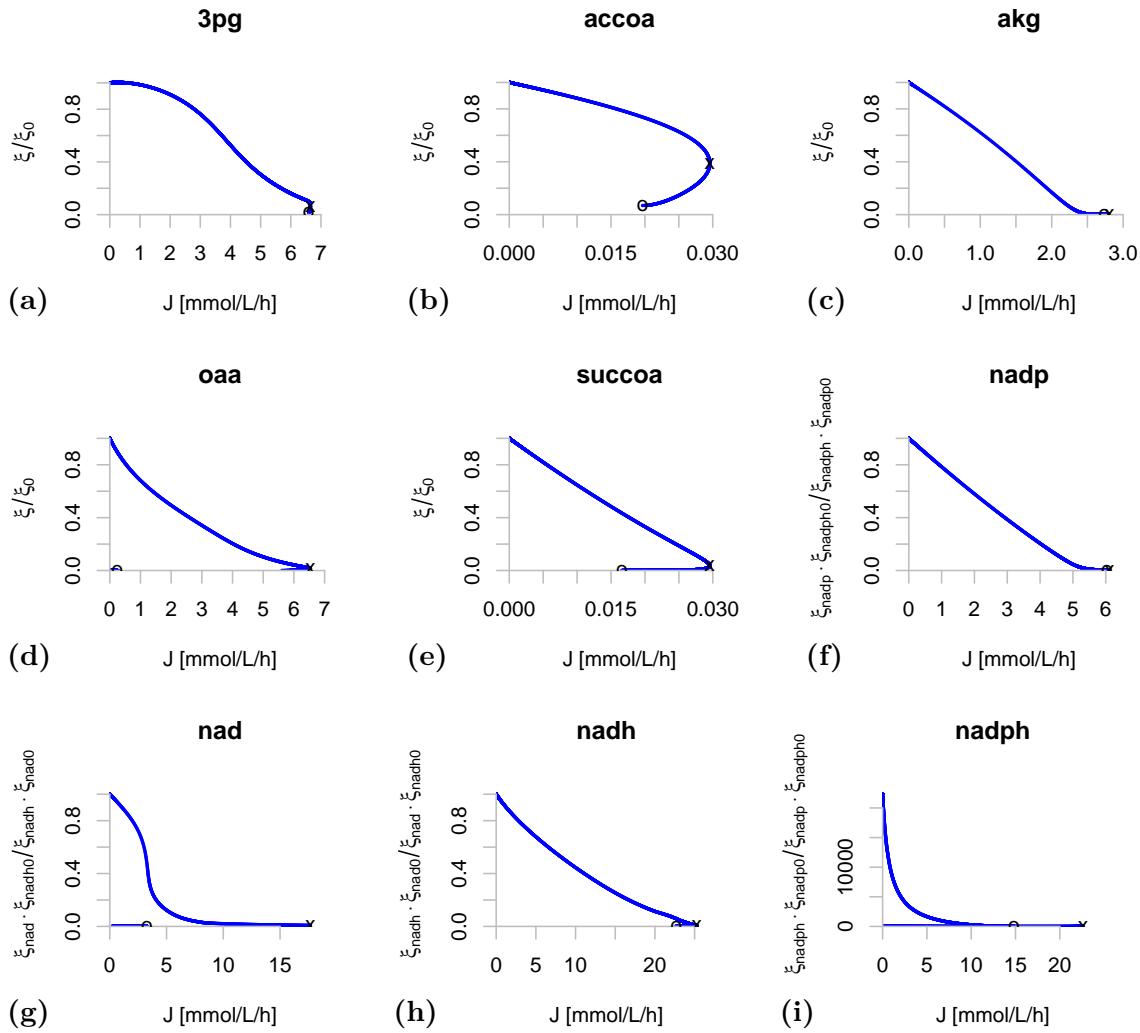


Abbildung 5.1: Potential-Strom-Kennlinien von Vorläufermolekülen mit regulärem Verlauf. x Punkte mit maximalem Produktionsstrom J_{PROD} , o Simulationsswendepunkt, an dem die Ratenkonstante k_{PROD} wieder verringert wird. **a** 3pg, **b** accoa, **c** akg, **d** oaa, **e** succoa, **f** nadp, **g** nad, **h** nadh, **i** nadph. Zu erkennen ist, dass die Kennlinie bis zum maximalen Produktionsstrom x stets fällt. Danach fällt sowohl das Potential als auch der Strom bis zum Umkehrpunkt o. Hier kehrt die Kennlinie um und verläuft deckungsgleich zurück zu ihrem Ursprung bei $J_{PROD} = 0, \xi/\xi_0 = 1$. Bei nad, nadp, nadh, nadph ist statt des normierten thermodynamischen Potentials ξ/ξ_0 das Potentialverhältnis ζ zum jeweils analogen Molekül gegen den Produktstrom aufgetragen. Die Potential-Strom-Kennlinien in dieser Gruppe zeichnen sich dadurch aus, dass das Potential ξ_i/ξ_0 mit der steigender Ratenkonstante k_{PROD} stets fällt. Dies führt zu einer Kennlinie, in der das Potential mit steigendem Strom J_{PROD} bis zum maximalen Produktionsstrom bei x abfällt. Danach fällt auch der Strom, jedoch ohne dass das Potential sich erholt, bis zum Umkehrpunkt der Simulation bei o, ab dem die Geschwindigkeitskontante k_{PROD} wieder verringert wird.

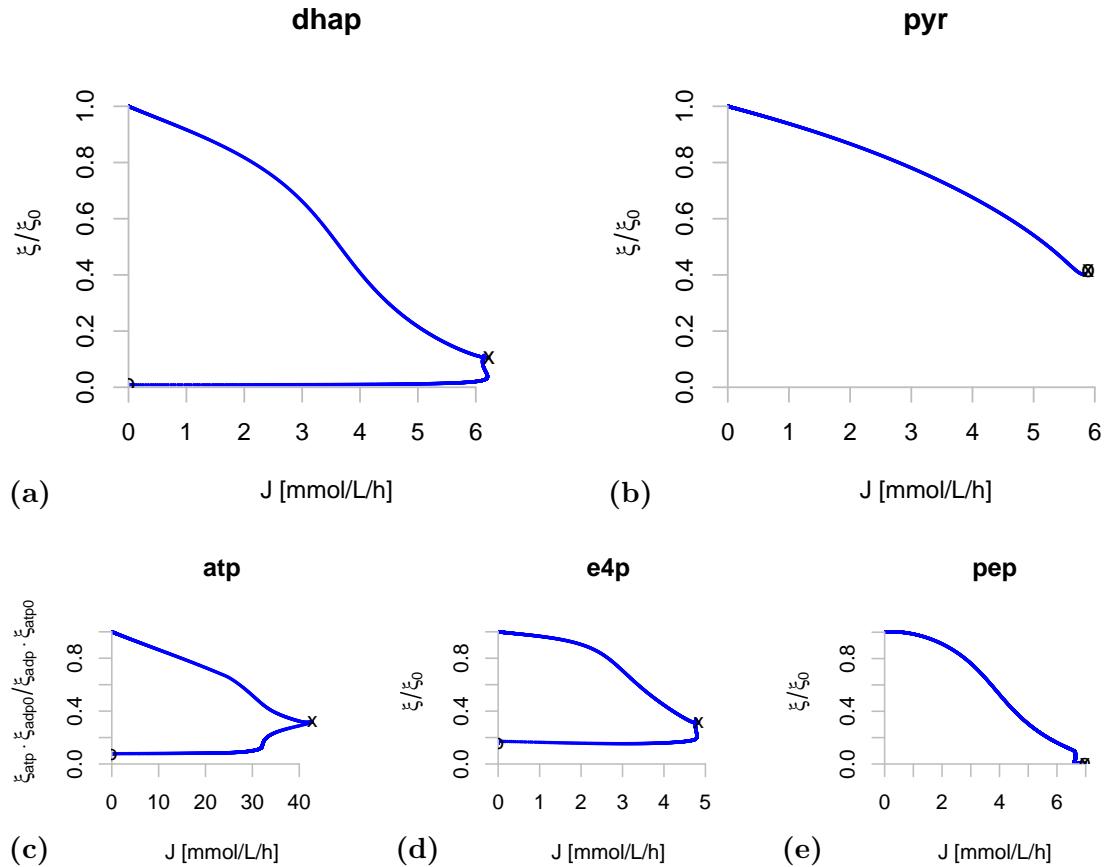


Abbildung 5.2: Potential-Strom-Kennlinien von Vorläufermolekülen mit *fast* regulärem Verlauf. \times Punkte mit maximalem Produktionsfluss J_{PROD} , \circ Simulationswendepunkt, an dem die Ratenkonstante k_{PROD} wieder verringert wird. Die Kennlinien in dieser Gruppe sind ähnlich zu denen aus der vorhergehenden, jedoch weisen sie kleine Besonderheiten auf. **a** *atp*, **b** *dhap*, **c** *e4p*: Diese Kennlinien weisen einen „Sprung“ am Punkt des maximalen Stromes auf. Das Potential steigt, während der Strom fällt, danach fällt das Potential wieder während der Strom wiederum steigt um kurz vor dem maximalen Strom wieder zu fallen. **d** *pep*: Die Kennlinie fällt bis zu einem Wendepunkt, an dem der Strom etwas zurückgeht. Wenn das Potential beinahe zusammengebrochen ist, steigt der Strom wieder bis zum maximalen Strom. Hier ist zudem k_{PROD} so gewählt, dass es sich der Wendepunkt \circ mit dem Punkt maximalen Stromes \times überlagert. Bei **e** *pyr* macht das Potential kurz vor erreichen des maximalen Stromes einen kurzen Sprung nach oben. Dies könnte allerdings ein Hinweis darauf sein, dass die ersten numerischen Fehler auftreten.

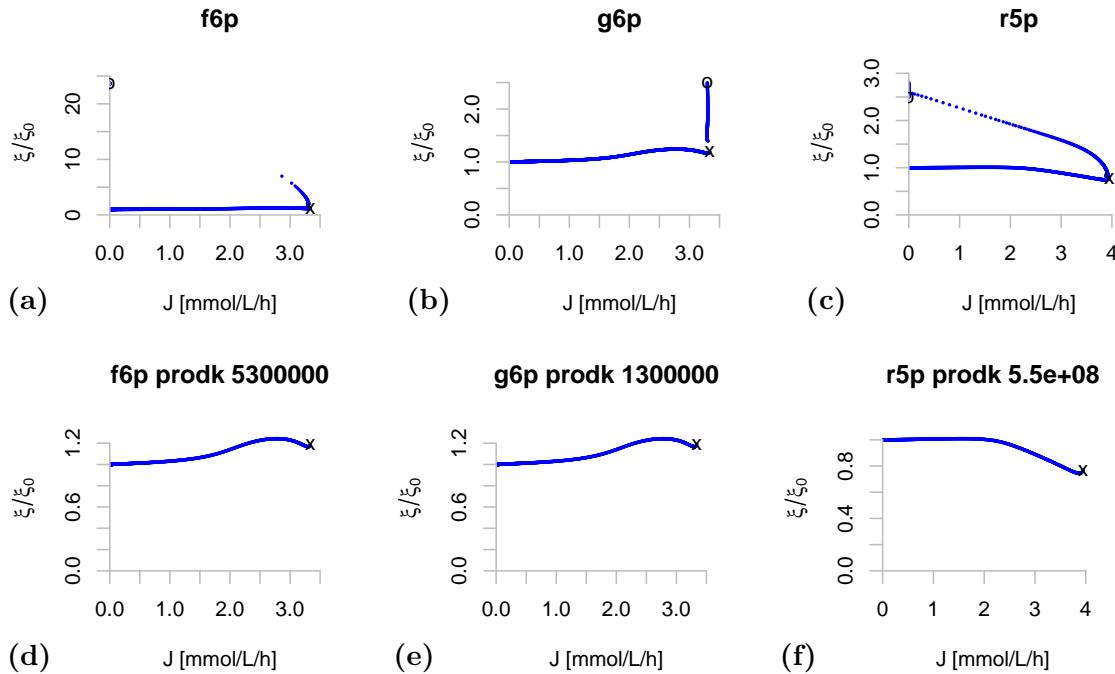


Abbildung 5.3: Potential-Strom-Kennlinien von Vorläufermolekülen mit nach oben gekehrtem Verlauf. x Punkte mit maximalem Produktionsfluss J_{PROD} , o Simulationswendepunkt, an dem die Ratenkonstante k_{PROD} wieder verringert wird. **a,b,c:** vollständiger Verlauf **d,e,f:** gekürzter Verlauf. Vor Erreichen des maximalen Stromes steigt und fällt das Potential von $f6p$ und $g6p$, das von $r5p$ steigt währenddessen nur minimal, und fällt danach. Die Potential-Strom-Kennlinien von $g6p$ und $f6p$ sind theoretisch gleich, weil zwischen diesen Spezies ein schnelles Gleichgewicht vorliegt. Die Unterschiede in der Darstellung sind auf die Auswahl von unterschiedlichen maximalen Ratenkonstanten \tilde{k}_{PROD} zurückzuführen. Das Potential fällt in allen drei Fällen. Bei den Zuckern $f6p$, $g6p$ und $r5p$ steigt das Potential mit steigendem k_{PROD} nach erreichen des Maximalstromes.

5.2 Kennlinien der Aminosäuren

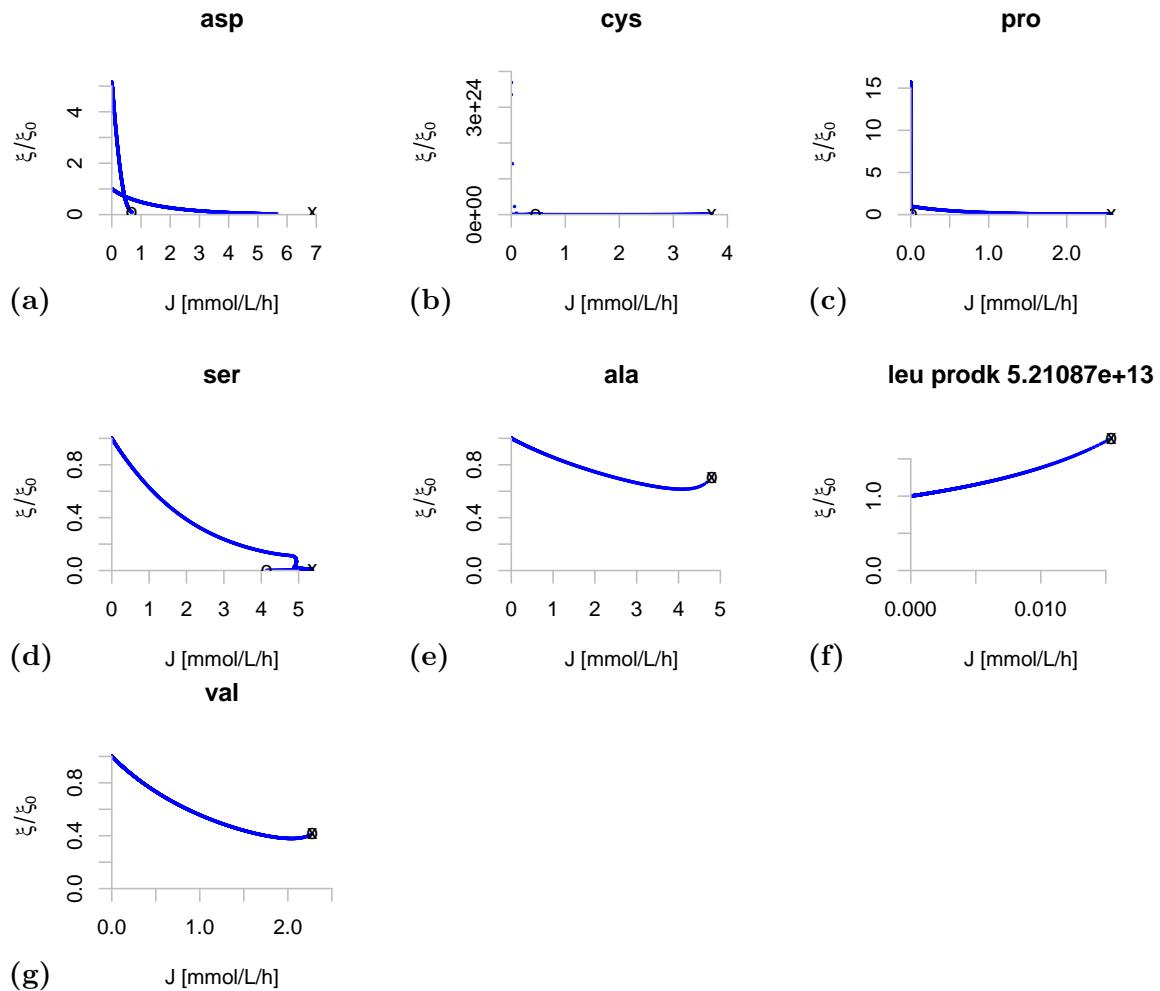


Abbildung 5.4: Potential-Strom-Kennlinien von Aminosäuren mit nicht regulärem Verlauf. x Punkte mit maximalem Produktionsfluss J_{PROD} , o Simulationswendepunkt, an dem die Ratenkonstante k_{PROD} wieder verringert wird. **a asp:** Bei erreichen des maximalen Stromes bricht derselbe ein. Auch das Potential ist zu diesem Zeitpunkt nahe 0. Nachdem die Ratenkonstante wieder erhöht wird, fällt der Strom noch weiter, während das Potential bis auf das fünffache des Anfangspotentials steigt. **b cys, c pro:** Die Darstellung der Kennlinien ist hier auf den Verlauf bis zum Punkt des maximalen Produktstromes x begrenzt, da das Potential danach auf ein Vielfaches des Anfangspotentials steigt. **d ser:** Vor dem Erreichen des Maximalstromes fällt der Produktstrom etwas ab, während das Potential in dem Bereich stets fällt. **e ala, g val:** Vor Erreichen des maximalen Stromes steigt das Potential. **f leu:** Das Potential steigt mit steigender Ratenkonstante, und die Simulation bricht schon bei einem geringen Strom aufgrund numerischer Fehler ab.

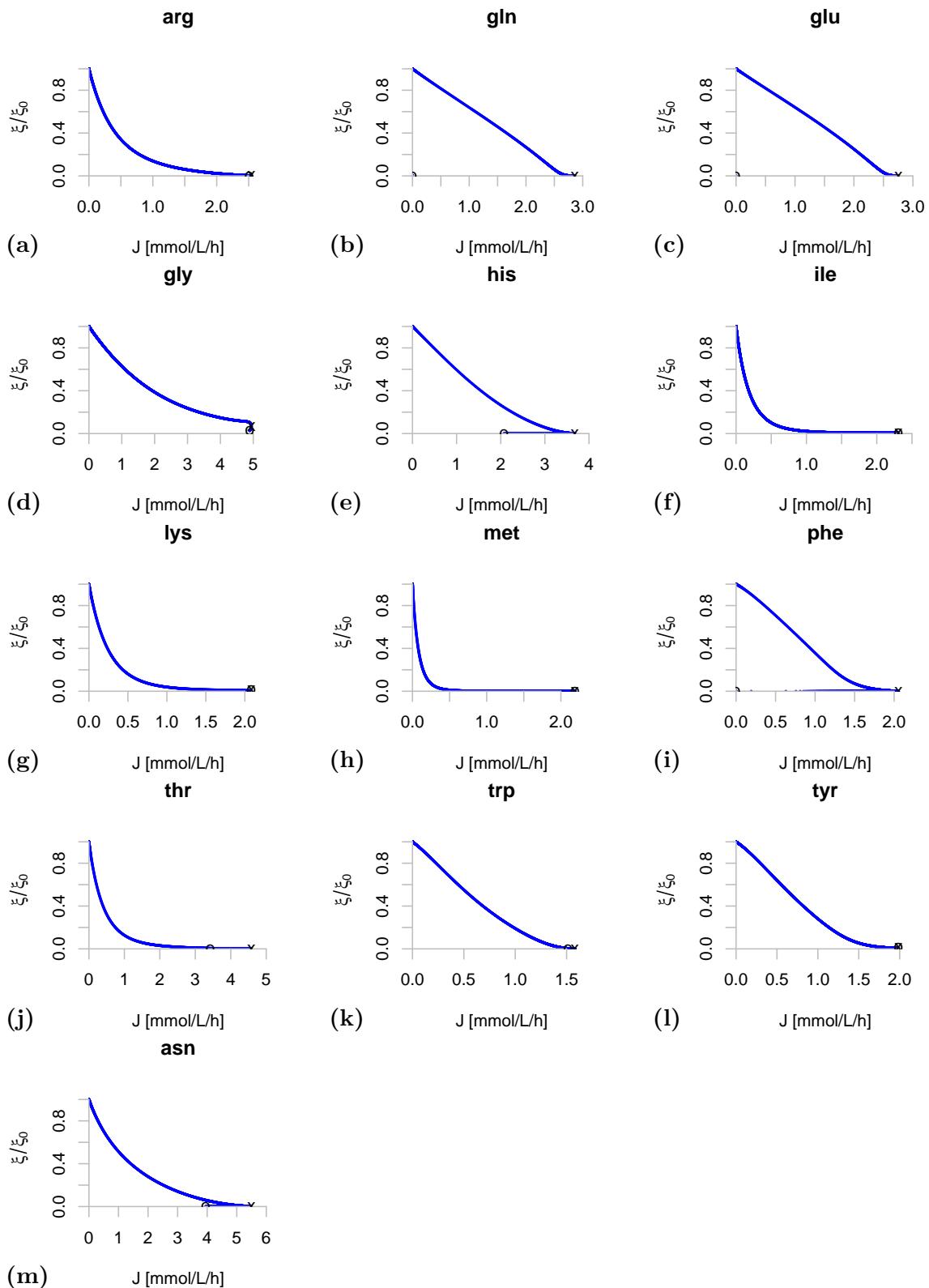


Abbildung 5.5: Potential-Strom-Kennlinien von Aminosäuren mit regulärem Verlauf. x Punkte mit maximalem Produktionsfluss J_{PROD} , o Simulationswendepunkt, an dem die Ratenkonstante k_{PROD} wieder verringert wird. Während bei **a** *arg*, **d** *gly*, **f** *ile*, **g** *lys*, **h** *met*, **k** *trp*, **l** *tyr*, **m** *asn* und **n** *val* das System kurz nach dem erreichen des Maximalstromes an ihre numerischen Grenzen stößt, bleibt es bei **b** *gln*, **c** *glu*, **e** *his*, **i** *phe* und **j** *thr* auch für höhere k_{PROD} -Werte funktionsfähig.

5.3 Kennlinien der Nucleotide

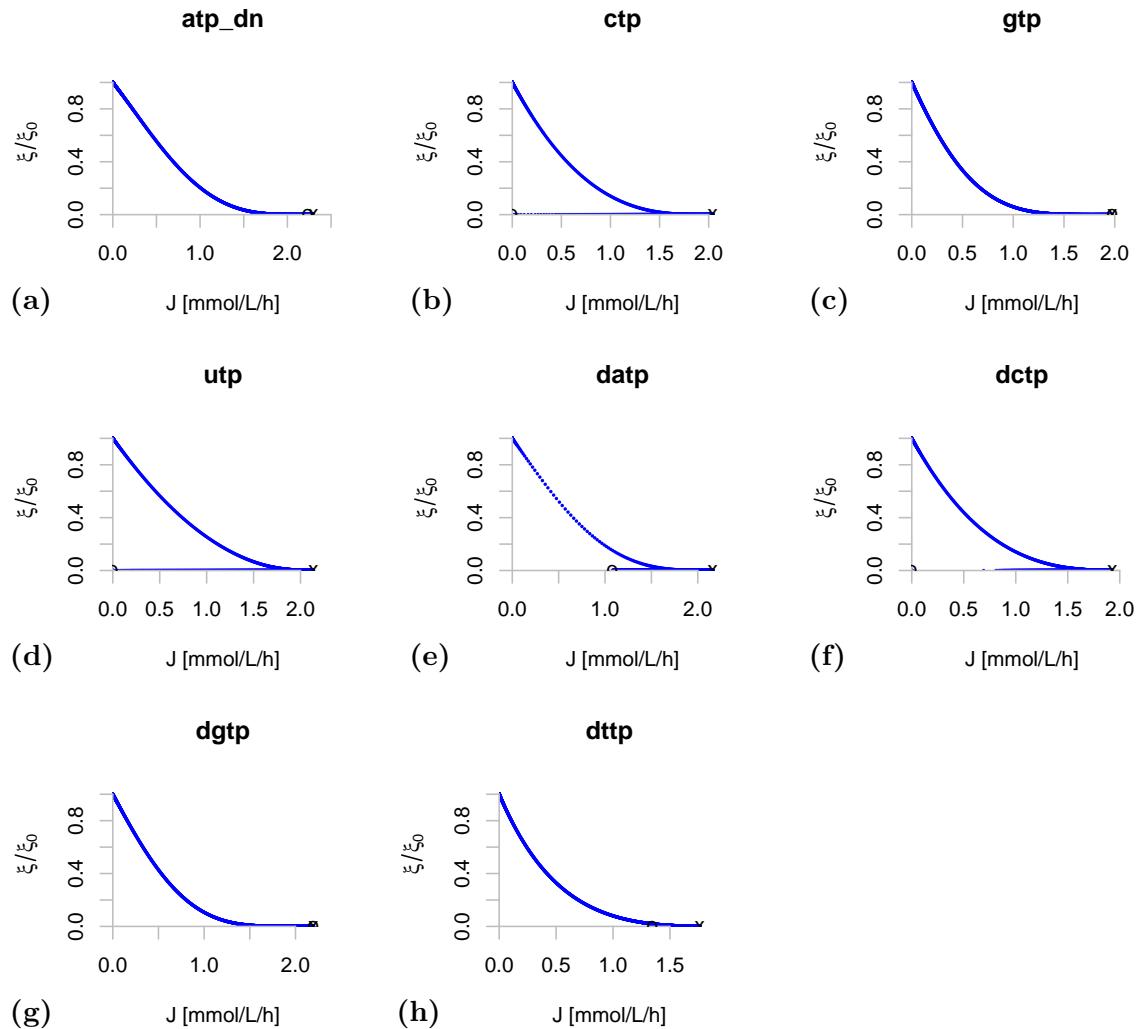


Abbildung 5.6: Potential-Strom-Kennlinien von Nucleotiden. x Punkte mit maximalem Produktionsfluss J_{PROD} , \circ Simulationswendepunkt, an dem die Ratenkonstante k_{PROD} wieder verringert wird. *De-novo*-Synthese von **a** atp, **b** ctp, **c** gtp, **d** utp, **e** datp, **f** dctp, **g** dgtp und **h** dtpp. Im Unterschied zu dem ATP, welches in Abbildung 5.2 betrachtet wird, wird hier die *de novo*-Synthese von ATP und den anderen Nucleotiden betrachtet. Die Kennlinien weisen alle einen regulären Verlauf auf, in dem das Potential stets fällt und der Strom bis zum Erreichen eines Grenzpotentials x steigt und dann fällt bis zum Wendepunkt der Simulation \circ . Der absteigende Verlauf ist nicht immer zu erkennen, da die Simulation oft kurz nach Erreichen des Maximalstromes abbricht.

5.4 Kennlinien der Lipopolysaccharidmonomere

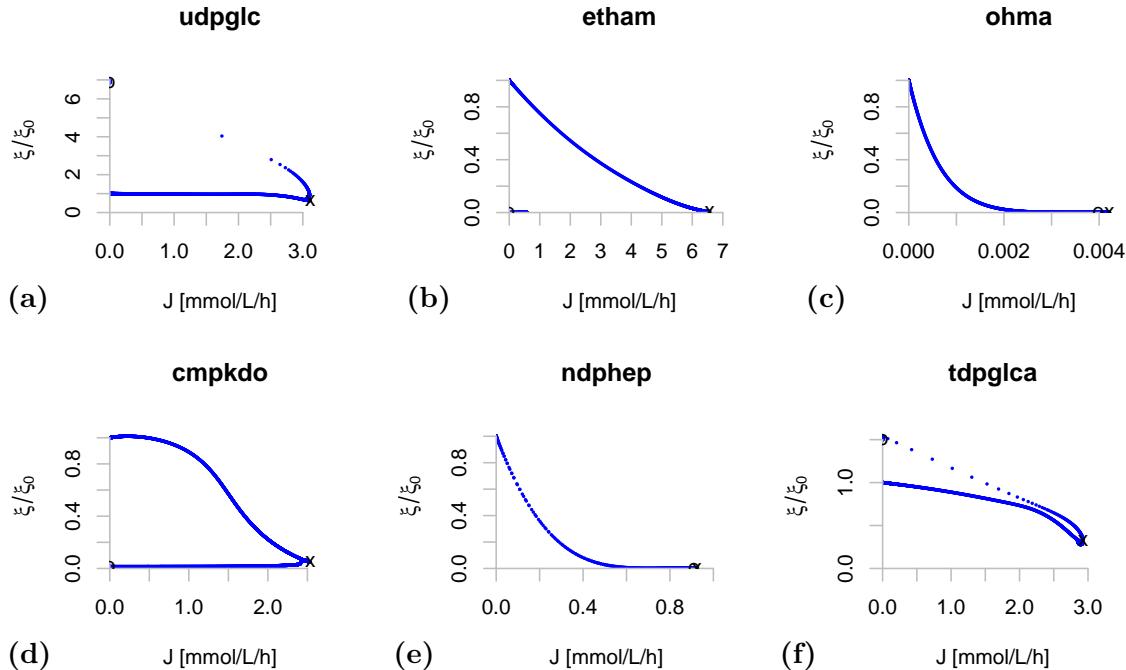


Abbildung 5.7: Potential-Strom-Kennlinien von Lipidbausteinen. x Punkte mit maximalem Produktionsfluss J_{PROD} , o Simulationswendepunkt, an dem die Ratenkonstante k_{PROD} wieder verringert wird. a *udpglc*: Bis zum Erreichen des maximalen Produktflusses bleibt das Potential mit steigender Ratenkonstante stabil, bzw fällt leicht. Nach dem der maximale Fluss erreicht wurde, steigt das Potential mit weiter steigender Ratenkonstante, während der Fluss stark sinkt. Ähnlich sieht das Schaubild bei *tdpglca* aus, wobei hier das Potential etwas weiter fällt vor dem Maximalstrom, und nicht so stark steigt danach. b *etham* c *ohma* e *ndphep* Diese Kennlinien zeigen einen regulären Verlauf. Sie unterscheiden sich nur in der Höhe des Maximalstromes, und in der Steigung der Kennlinie. d *cmpkdo* Der Verlauf dieser Kennlinie ist fast regulär, doch es fällt eine geringe Steigung im Potential bei niedrigen k_{PROD} -Werten auf.

5.5 Kennlinien der Peptidoglycanmonomere

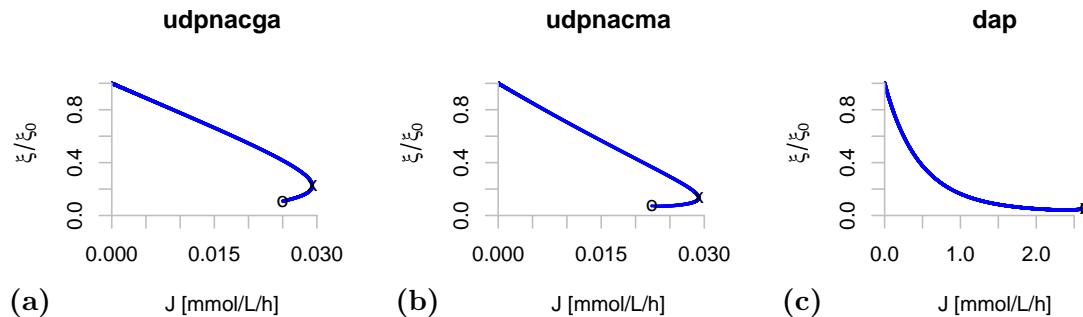


Abbildung 5.8: Potential-Strom-Kennlinien von Peptidoglycanmonomeren. \times Punkte mit maximalem Produktionsfluss J_{PROD} , \circ Simulationswendepunkt, an dem die Ratenkonstante k_{PROD} wieder verringert wird. a *udpnacga* b *udpnacma* und c *dap* welche alle einen regulären Kennlinienverlauf aufweisen.

5.6 Diskussion

Als regulär wird in dieser Arbeit eine Kennlinie bezeichnet, in der das Potential mit steigender Ratenkonstante k_{PROD} fällt und der Abstrom bis zu einem Maximalstrom steigt, und danach wieder fällt. Zwar mögen viele Kennlinien sehr unterschiedlich aussehen, doch treffen diese Kriterien bei genauerer Betrachtung zu. Solche Kennlinien sind zu finden bei Vorläufermolekülen (Abbildung 5.1), Aminosäuren (Abbildung 5.5), Nucleotiden (5.6), Lipopolysaccharidmonomeren (Abbildung 5.7 (b), (c), (d), (e)) und Peptidoglycanmonomeren (Abbildung 5.8). Die Unterschiede die sich innerhalb dieser Kategorisierung finden entfallen hauptsächlich auf die Art der Krümmung bis zum Maximalstrom - konkav (Steigung von ξ/J mit steigendem k_{PROD} fallend), konvex (Steigung von ξ/J mit steigendem k_{PROD} steigend), oder sigmoidal. Des Weiteren kann man unterscheiden, dass das Modellsystem bei einer über den Maximalstrom erhöhten Ratenkonstante bei manchen Produkten stabil bleibt, bei anderen aber numerische Fehler erfährt und die Simulation abbricht.

Einige Moleküle weisen Kennlinien auf, die geringfügig von dem oben beschriebenen abweichen. ATP, DHAP und Erythrose-4-Phosphat (Abbildung 5.2 (a), (b), (c)) zum Beispiel weisen einen “Sprung” am Punkt des maximalen Stromes auf, an dem das Potential etwas steigt, während der Strom fällt. Nach diesem Sprung jedoch fällt das Potential auch wieder. Ein Erklärungsansatz hierfür ist das Auftreten von Bifurkationen bei einem extremen Abstrom und der damit verbundenen Belastung auf das Modellsystem.

Bei DHAP, Erythrose-4-Phosphat, Phosphoenolpyruvat (Abbildung 5.2 (b), (c), (d)) und Serin (Abbildung 5.4 (d)) weist der Strom zwei Maxima auf. Während das lokale Maximum bei DHAP und Erythrose-4-Phosphat bei einem höheren k_{PROD} -Wert

liegt als das Hauptmaximum, liegt es bei Phosphoenolpyruvat und Serin bei einem niedrigeren k_{PROD} -Wert.

Die Saccharide f6p, g6p, r5p (Abbildung 5.3 und einige ihrer Polymere wie UDP-Galactose und TDP-Glucosamin (Abbildung 5.7 (a), (f)), weisen einen nach oben gekehrten Kennlinienverlauf auf. Diese Verläufe zeichnen sich dadurch aus, dass das Potential nach dem erreichen des Maximalen Abstromes mit steigender Ratenkonstante extrem steigt. Auch vor dem Erreichen des Maximalstromes weisen diese Moleküle ein überdurchschnittlich stabiles Potential auf, welches teilweise über dem Anfangspotential liegt. Ein Erklärungsansatz ist das Auftreten einer Pyruvatlimitierung. Die eingeführten Saccharide können nicht weiterverarbeitet werden und akkumulieren somit.

Der Abstrom der Lipopolysaccharidmonomere überträgt sich auf den Abstrom an Sacchariden, sodass hier der selbe Effekt entsteht.

Bei Aspartat, Cystein und Prolin (Abbildung 5.4 (a), (b), (c)) verlaufen die Kennlinien bis zum Erreichen des Maximalstromes regulär. Nach dem Überschreiten dieses Punktes steigt das Potential jedoch in immense Höhen, während der Strom zusammenbricht. Dieses Verhalten kann man einem ähnlichen Effekt wie bei den Sacchariden zuschreiben. Da das Zellvolumen nach dem Erreichen des Maximalstromes sehr stark fällt, treten numerische Fehler bei der Berechnung von Fluss und Potential auf.

Die Kennlinien von Alanin und Leucin (Abbildung 5.4 (e), (f)) entsprechen nicht dem regulären Verlauf, und sind schwer zu deuten. In beiden Fällen steigt das Potential mit steigender Ratenkonstante bevor die Simulation abbricht.

Kapitel 6

Analyse und Diskussion B: Theoretische Kennlinien im Vergleich mit dem Simulationsergebnis

In diesem Kapitel werden die Kennlinien aus der Simulation (Abschnitt 4.2) verglichen mit der Schätzung (Abschnitt 4.3). Die schwarzen Kennlinien entsprechen den Potential-Strom-Verhältnissen in der Simulation. Im Gegensatz zu den Kennlinien aus Kapitel 5 sind die Kennlinien in diesem Kapitel alle auf den Verlauf vor dem Erreichen des Produktstrommaximums gekürzt wurden.

Die roten Kennlinien gehen aus der in Kapitel 4.3 vorgestellten Berechnung hervor. In demselben Kapitel wurde erklärt, dass die Schätzung davon ausgeht, die Potentiale der an einer Biosynthese beteiligten *Vorläufermoleküle* beeinflussen sich nicht gegenseitig.

6.1 Vergleich bei Aminosäuren

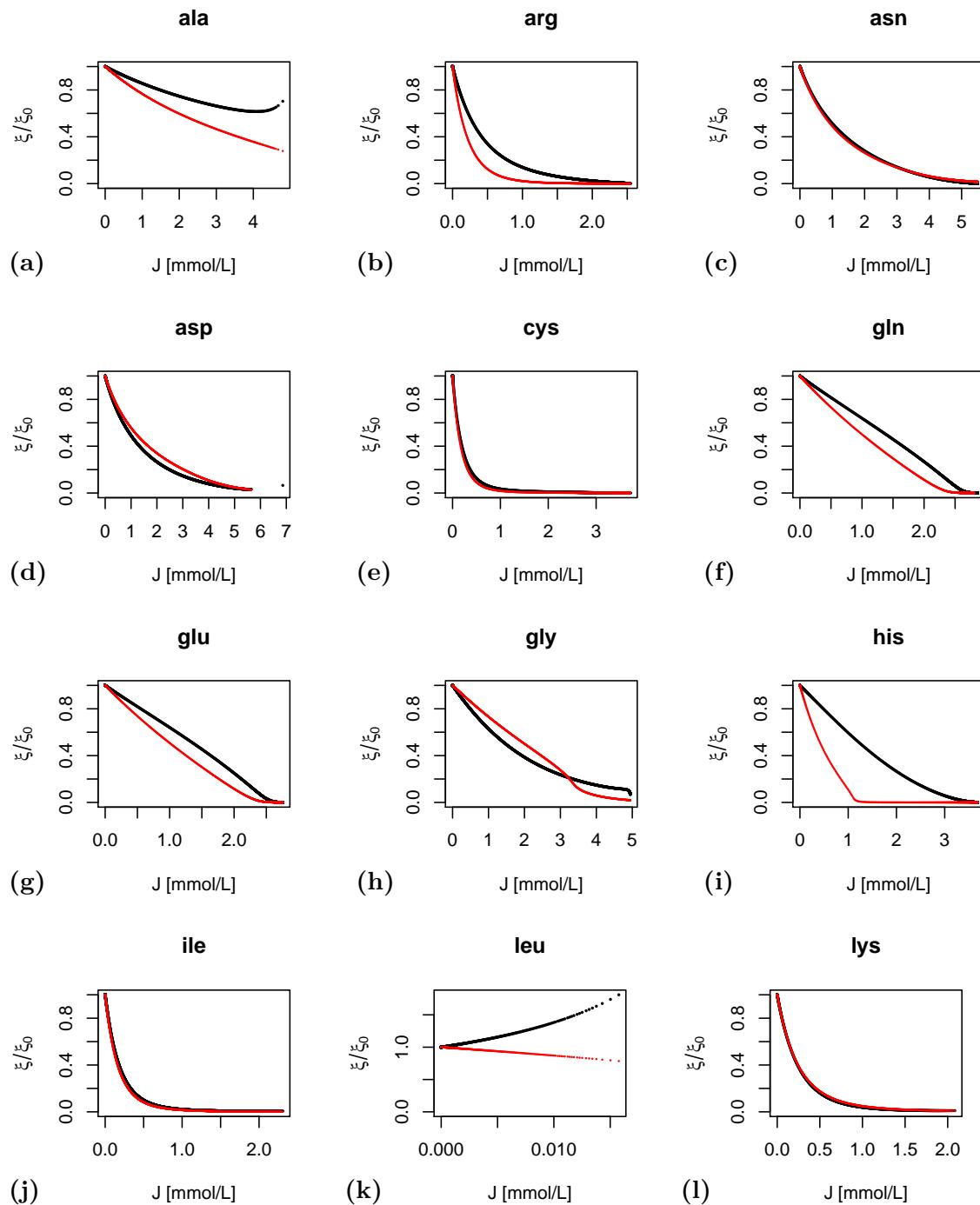


Abbildung 6.1: (Abbildung wird fortgesetzt)

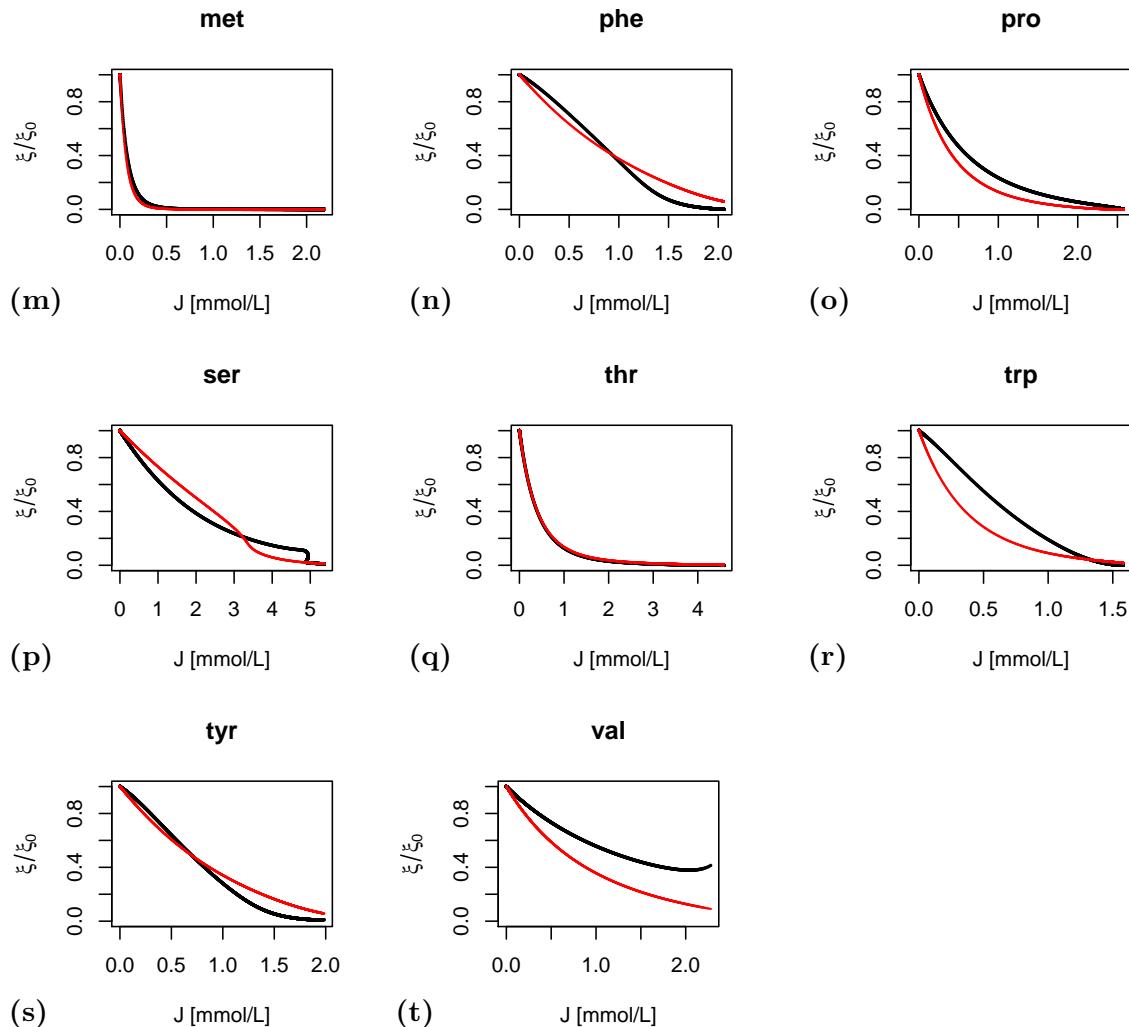


Abbildung 6.1: Vergleich der Kennlinien aus Simulation und Berechnung bei Aminosäuren. rot: berechnete Kennlinie, schwarz: Kennlinie aus der Simulation. Die Kennlinien aus Simulation und Schätzung verlaufen bei bemerkenswert vielen Aminosäuren nahezu deckungsgleich (*asn, cys, ile, lys, met*, und *thr*) oder zumindest äußerst ähnlich (*arg, asp, gln, glu, gly, phe, pro, ser, trp* und *ser*). Lediglich bei *ala, his, leu* und *val* weichen die beiden Kurven deutlich voneinander ab.

6.2 Vergleich bei Nucleotiden

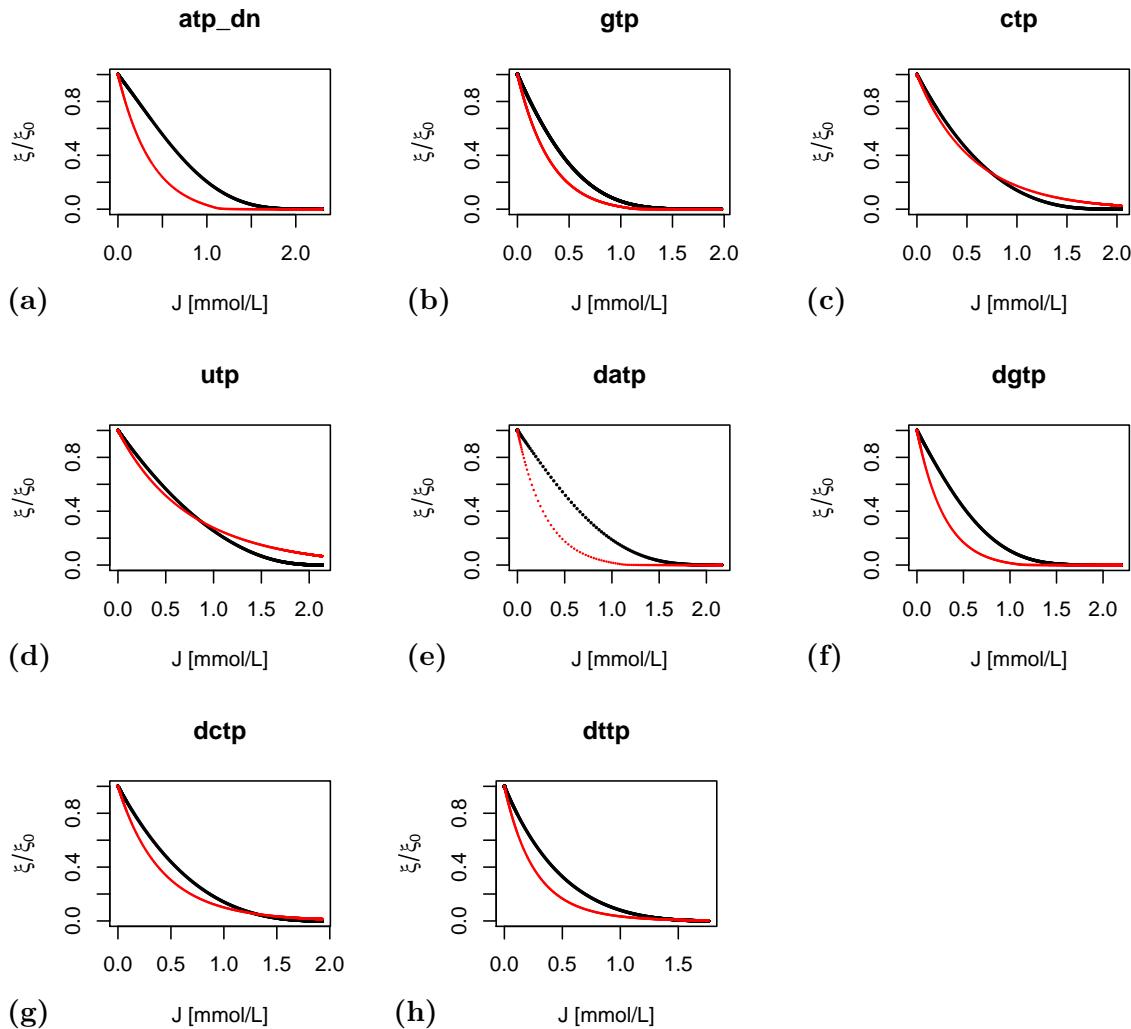


Abbildung 6.2: Vergleich der Kennlinien aus Simulation und Berechnung bei Nucleotiden. rot: berechnete Kennlinie, schwarz: Kennlinie aus der Simulation. Die Kennlinien aus der Berechnung ähneln bei allen Nucleotiden der Kennlinie aus der Simulation. Die größten Unterschiede zeigen sich bei der de-novo Synthese von *a* *atp* und *datp*. Auffällig ist, dass hier wieder die roten kennlinien in der Regel niedriger als die schwarzen liegen.

6.3 Vergleich bei Lipopolysaccharidmonomeren

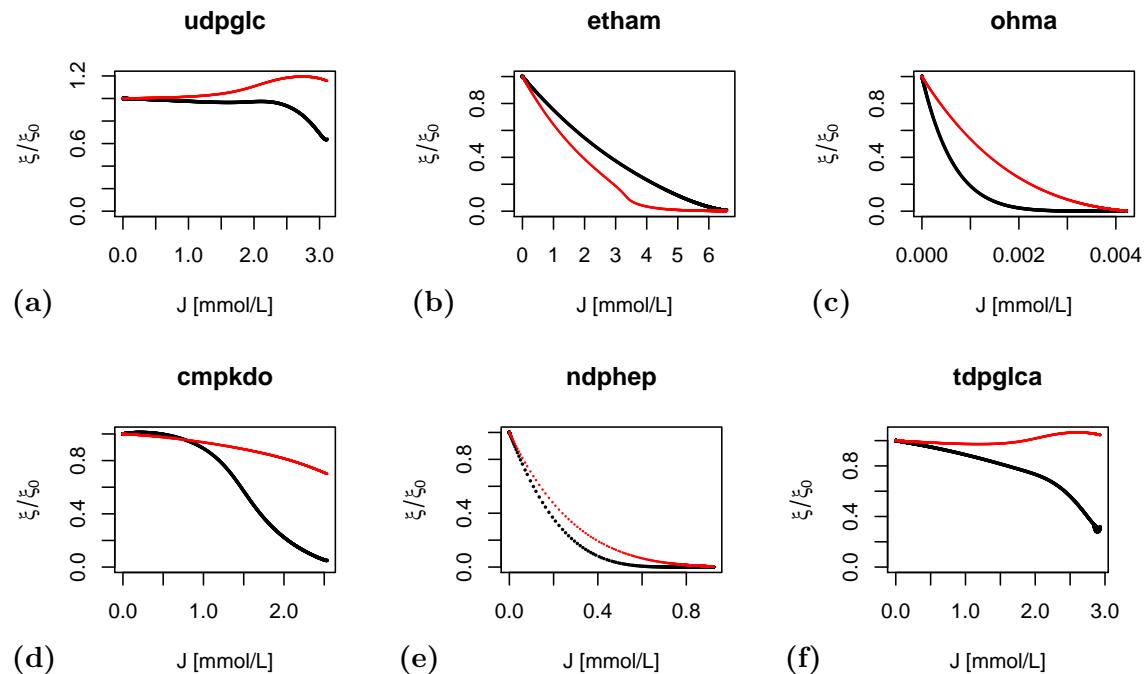


Abbildung 6.3: Vergleich der Kennlinien aus Simulation und Berechnung bei Lipopolysaccharidmonomeren. rot: berechnete Kennlinie, schwarz: Kennlinie aus der Simulation.

6.4 Vergleich bei Peptidoglycanmonomeren

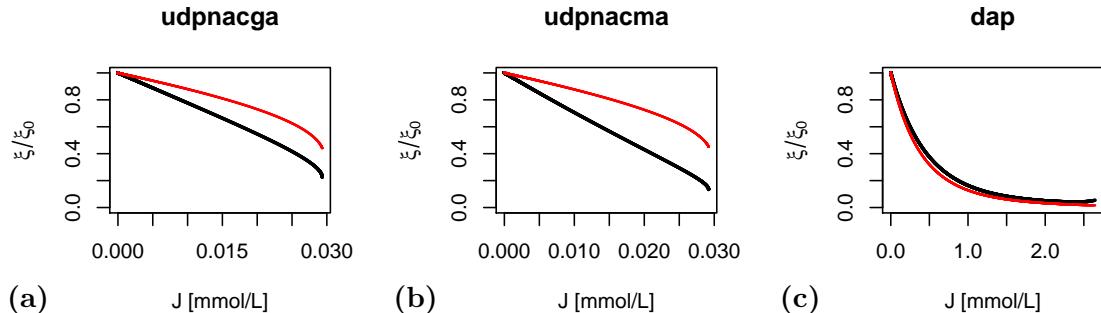


Abbildung 6.4: Vergleich der Kennlinien aus Simulation und Berechnung bei Peptidoglycan- und Glycogenmonomeren. rot: berechnete Kennlinie, schwarz: Kennlinie aus der Simulation.

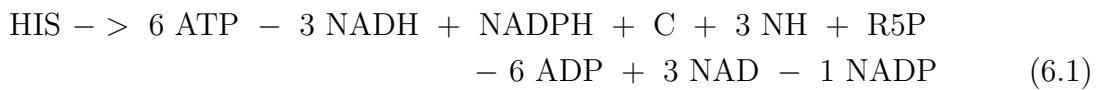
6.5 Diskussion

In Kapitel 4.3 wurde eine Methode erläutert, mit der es möglich ist, die Kennlinie eines Biosyntheseproduktes anhand der Kennlinien seiner Edukte abzuschätzen. Diese Methodik ergibt allerdings eine optimistische Schätzung, da angenommen wird, die Abströme der Vorläufermoleküle beeinflussen sich nicht gegenseitig. Es wird daher erwartet, dass die Potentiale aus der Simulation geringer sind und die Kennlinien daher niedriger verlaufen.

Umso interessanter ist es, dass viele der geschätzten Kennlinien nahezu deckungsgleich oder zumindest sehr ähnlich zu den Kennlinien aus der Simulation verlaufen. In den Fällen, in denen die Kennlinien voneinander abweichen, verläuft die Kennlinie aus der Berechnung zumeist niedriger als die Kennlinie aus der Simulation, das heisst, dass die Potentiale aus der Simulation entgegen der Erwartung höher sind als die Potentiale aus der Berechnung. Trotzdem liegen sie alle in der selben Größenordnung. Die Ausnahmen werden im folgenden besprochen.

Die Kurven von Alanin, Leucin und Valin (Abbildung 6.1 (a), (k), (t)) waren schon bei der Simulation von Potential und Strom schwer zu deuten (Abschnitt 5.2).

Nach der Schätzung hätte das Potential von Histidin (Abbildung 6.1 (i)) schon bei einem viel niedrigeren Strom fallen sollen, als er in der Simulation gefallen ist. Auch bei Valin (Abbildung 6.1 (t)) verläuft die Kennlinie aus der Berechnung weit unter der Kennlinie aus der Simulation. Betrachtet man die Zusammensetzungsregel für Histidin, lässt sich eine mögliche Erklärung formulieren:



Wie man sieht werden bei der Biosynthese Histidin 6 ATP benötigt, im gleichen Zuge jedoch 3 NADH zur Verfügung gestellt. In der aeroben Zellatmung wird NADH oxidiert um ca. 2,5 ATP zu gewinnen ([1]). Vernachlässigt man Ammonium und C₁-Körper, welche nicht modelliert sind, werden bei der Biosynthese von Histidin effektiv ca. 1,5 ATP freigesetzt. In der Schätzung werden diese Wechselwirkungen jedoch vernachlässigt, womit die Schätzung des Potentials niedriger ausfällt als die Simulation desselben.

Beim Vergleichen der Kennlinien aus Simulation und Berechnung bei UDP-Galactose, CMP-KDO und TDP-Glucosamin (Abbildung 6.3 (a), (d), (f)) lässt sich nur unzufriedenstellende Ähnlichkeit feststellen. Dies kann zum größten Teil darauf zurückgeführt werden, dass die Simulierten Kennlinien dieser Moleküle nicht dem regulären Verlauf entsprechen, offensichtlich beeinflusst vom Verhalten ihrer jeweiligen *Vorläufermoleküle* g6p, r5p und f6p.

Die Schätzung für Hydroxy-Myristinsäure und NDP-Heptose (Abbildung 6.3 (c), (e)) verläuft höher als die Simulation. Dies ist auch der Fall bei UDP-N-Acetylglucosamin und UDP-N-Acetylmuraminsäure (Abbildung 6.4 (a), (b)). Es ist möglich dass sich erst hier, bei komplexeren Molekülen mit einer hohen Zahl an *Vorläufermolekülen* der Optimismus der Schätzung niederschlägt, und die Interaktionen zwischen den *Vorläufermolekül*-Strömen ins Gewicht fallen.

Kapitel 7

Zusammenfassung und Ausblick

Der *thermokinetische* Modellformalismus ist ein vielversprechender Ansatz für das Erstellen von thermodynamisch konsistenten mathematischen Modellen von biochemischen Reaktionsnetzwerken. Mit der *TKMOD* Erweiterung für Mathematika wird das Erstellen dieser Modelle zudem stark vereinfacht.

Thermokinetische Potential-Strom-Kennlinien erlauben eine Aussage über die Reaktion eines Modells auf eine Belastung durch die Abfuhr eines oder mehrerer Vorläufermoleküle. Erkenntnisse aus dieser Analyse können Ansätze für eine Optimierung des Produktionsprozesses oder des produzierenden Stammes liefern.

Die Wahl der maximalen Ratenkonstante k_{PROD} hat sich als entscheidend für die Berechnung der Kennlinien erwiesen. In dieser Arbeit wurde ein Algorithmus entwickelt, welcher die Wahl eines geeigneten Wertes übernimmt.

Ein interessantes Ergebnis aus der Analyse der Potential-Strom-Kennlinien ist, dass viele Verläufe regulär sind. Mit steigender Ratenkonstante über der Abflussreaktion steigt die Flussrate, während das Potential sinkt. Fällt das Potential unter ein gewisses Minimum, so sinkt sowohl das Potential als auch die Abflussrate. Diese Relationen passen gut in das Modell eines biochemischen Reaktionsnetzwerkes.

Die Kennlinien der sacchariden Vorläufermoleküle Fructose-6-Phosphat, Glucose-6-Phosphat und Ribulose-5-Phosphat unterscheiden sich jedoch von dem eben beschriebenen Verhalten und es wurde ein Erklärungsansatz angesprochen.

Auch bei den Aminosäuren gab es einige Ausnahmen, von denen im Besonderen Alanin, Leucin und Valin herausstachen. Hier fand sich keine zufriedenstellende Erklärung für den Verlauf der Potential-Strom-Kennlinien.

Die schätzende Berechnung der Potentiale von komplexeren *Bausteinen* lieferte Ergebnisse, die erstaunlich gut zu den Kennlinien aus der Simulation passten. Viele der Kennlinien für Aminosäuren und Nucleotide verlaufen beinahe deckungsgleich oder sehr ähnlich. Bei den Aminosäuren fielen vor allem die schon erwähnten Moleküle Alanin, Leucin und Valin auf.

Bei den komplexeren Lipopolysaccharid- und Peptidoglycanmonomeren traten größere Unterschiede zwischen den Kennlinien der Simulation und der Schätzung auf. Die

Kapitel 7 Zusammenfassung und Ausblick

Qualität der Schätzung scheint demnach mit steigender Zahl an verwendeten *Vorläufermolekülen* zu fallen.

Als weiterführendes Forschungsprojekt würde es sich anbieten, Experimente an einem realen Chemostaten durchzuführen. Der Versuchsaufbau müsste eine Methode beinhalten, welche es ermöglicht gezielt einzelne *Bausteine* oder *Vorläufermoleküle* aus Zelle und Chemostat zu entfernen oder für die Zelle unbrauchbar machen. Aus der Rate, mit der die Inaktivierung oder Abfuhr erfolgt und der intrazellulären Konzentration lässt sich eine Potential-Strom-Kennlinie erstellen und mit den Ergebnissen dieser Arbeit vergleichen.

Abbildungsverzeichnis

2.1	Grundlegenden Elemente der Thermokinetischen Modellierung	10
2.2	Parallel- und Serienschaltung von Widerständen	11
2.3	Beispiele für Strom-Spannungskennlinien	12
3.1	Die <i>TKMOD</i> Module	18
3.2	Das <i>E. coli</i> Modellsystem	22
4.1	Beispiele für Potential-Strom-Kennlinien bei unterschiedlichen Geschwindigkeitsparametern \tilde{k}_{PROD}	26
5.1	Potential-Strom-Kennlinien von Vorläufermolekülen mit regulärem Verlauf	30
5.2	Potential-Strom-Kennlinien von Vorläufermolekülen mit <i>fast</i> regulärem Verlauf	31
5.3	Potential-Strom-Kennlinien von Vorläufermolekülen mit nach oben gekehrtem Verlauf	32
5.4	Potential-Strom-Kennlinien von Aminosäuren mit nicht regulärem Verlauf	33
5.5	Potential-Strom-Kennlinien von Aminosäuren mit regulärem Verlauf . .	34
5.6	Potential-Strom-Kennlinien von Nucleotiden	35
5.7	Potential-Strom-Kennlinien von Lipidbausteinen	36
5.8	Potential-Strom-Kennlinien von Peptidoglycanmonomeren	37
6.1	(Abbildung wird fortgesetzt)	40
6.1	Vergleich der Kennlinien aus Simulation und Berechnung bei Aminosäuren	41
6.2	Vergleich der Kennlinien aus Simulation und Berechnung bei Nucleotiden	42
6.3	Vergleich der Kennlinien aus Simulation und Berechnung bei Lipopolysaccharidmonomeren	43
6.4	Vergleich der Kennlinien aus Simulation und Berechnung bei Peptidoglycan- und Glycogenmonomeren	44

Tabellenverzeichnis

3.1	Liste relevanter chemischer Spezies und ihrer Abkürzungen.	21
-----	--	----

Quelltextverzeichnis

3.1	Ein einfaches Beispiel für eine <i>TKMOD</i> Datei	19
4.1	Ein Beispiel für eine Zusammensetzungssregel in <code>bb.mma</code>	23

Literaturverzeichnis

- [1] Jeremy M. Berg, John L. Tymoczko, and Lubert Stryer. *Biochemistry*. W. H. Freeman, 0007. aufl edition, 2010.
- [2] Peter N Brown, Alan C Hindmarsh, and Linda R Petzold. Using Krylov Methods in the Solution of Large-Scale Differential-Algebraic Systems. *SIAM Journal on Scientific Computing*, 15(6):1467–1488, 1994.
- [3] Michael Ederer. *Thermokinetic Modeling and Model Reduction of Reaction Networks*. PhD thesis, Universität Stuttgart, 2010.
- [4] Michael Ederer and Ernst Dieter Gilles. Thermodynamic-Kinetic Modeling and Electrical Engineering. In *Proceedings of the FOSBE, 2007, Stuttgart, Germany*, pages 135–140, 2007.
- [5] Michael Ederer and Ernst Dieter Gilles. Thermodynamically feasible kinetic models of reaction networks. *Biophysical Journal*, 92(6):1846–57, March 2007.
- [6] Michael Ederer and Ernst Dieter Gilles. Thermodynamic Constraints in Kinetic Modeling: Thermodynamic-Kinetic Modeling in Comparison to Other Approaches. *Engineering in Life Sciences*, 8(5):467–476, October 2008.
- [7] Hiroaki Kitano. Computational systems biology. *Nature*, 420(6912):206–10, November 2002.
- [8] J Lengeler, Gerhart Drews, and Hans G Schlegel. *Biology of the Prokaryotes*. Blackwell Science Ltd, Oxford, UK, December 1998.
- [9] F C Neidhardt, J L Ingraham, and M Schaechter. *Physiology of the bacterial cell: a molecular approach*. Sinauer Associates, 1990.
- [10] Stefan Schuster and Ronny Schuster. A generalization of Wegscheider’s condition. Implications for properties of steady states and for quasi-steady-state approximation. *Journal of Mathematical Chemistry*, 3(1):25–42, January 1989.
- [11] Jörg Stelling, Steffen Klamt, Katja Bettenbrock, Stefan Schuster, and Ernst Dieter Gilles. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420(6912):190–3, 2002.
- [12] Herbert A. Stuart and Gerhard Klages. *Kurzes Lehrbuch Der Physik*. Gabler Wissenschaftsverlage, 2009.

Erklärung des Autors

der Studienarbeit mit dem Titel

Analyse thermokinetischer Potential-Strom-Kennlinien an einem Modell von *Escherichia coli*

Hiermit versichere ich,

1. dass ich meine Arbeit bzw. bei einer Gruppenarbeit den entsprechend gekennzeichneten Anteil der Arbeit selbständig verfasst habe,
2. dass ich keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe,
3. dass die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist,
4. dass ich die Arbeit weder vollständig noch in Teilen bereits veröffentlicht habe und
5. dass das elektronische Exemplar mit den anderen Exemplaren übereinstimmt.

Stuttgart, den 28. August 2012



Waldemar Reusch