**This is a postprint of an article published in
eckwer, W.-D., Jahn, D., Zeng, A.-P., Hempel, D.C.
System biotechnological approaches to process development
[Systembiotechnologische ansätze zur prozessentwicklung]
(2006) Chemie-Ingenieur-Technik, 78 (3), pp. 193-208.**

Systembiotechnologische Ansätze zur Prozessentwicklung*

Wolf-Dieter Deckwer¹, Dieter Jahn², An-Ping Zeng³, Dietmar C. Hempel⁴

Abstract

Mit Hilfe der Bioinformatik kann aus dem Genom eines Mikroorganismus das potenzielle metabolische Netzwerk rekonstruiert werden. Netzwerke stellen die eigentliche Basis der Systembiologie bzw. Systembiotechnologie dar. Für eine Systembiotechnologie, die auf Prozesse und Produkte fokussiert ist, müssen die potenziellen Netzwerke in geeigneter Weise funktionalisiert, d.h. der Phänotyp ermittelt werden, der optimale Prozesse bzw. Produkte und Produktivitäten liefert. Die Funktionalität der genregulatorischen und metabolischen Netzwerke wird durch die „Omics“-Techniken abgefragt. Da die Funktionalität vom Environom abhängt, ist es erforderlich, diese Techniken in koordinierter Weise möglichst simultan anzuwenden. Das Fluxom (Flussverteilung der zentralen Stoffwechselwege) erfordert derzeit besonderes Interesse, das sowohl unter Variation genetischer Eigenschaften (Metabolic Engineering) als auch Änderungen des Environoms zu ermitteln ist. Im Rahmen einer im SFB 578 „Vom Gen zum Produkt“ angestrebten prozessorientierten Systembiotechnologie stellen sich für den Verfahrenstechniker vielfache Aufgaben, die nur in Kooperation mit Molekularbiologen und Bioinformatikern zu lösen sind. Dazu gehört neben der Bereitstellung definierter, reproduzierbarer Kultivierungsverfahren in hochinstrumentierten Bioreaktoren die Erfassung kinetischer Zusammenhänge in Abhängigkeit vom Environom sowie von

* Auszugsweise vorgetragen von W.-D. Deckwer auf der GVC-DECHEMA-Vortrags- und Diskussionstagung „Systembiotechnologie für industrielle Prozesse“, 1.-4. Mai 2005, Braunschweig

¹ Prof. Dr. W.-D. Deckwer, TU-BCE, GBF, Mascheroder Weg 1, 38124 Braunschweig, email: wdd@gbf.de

² Prof. Dr. D. Jahn, TU-BS, Institut für Mikrobiologie, Biozentrum, Spielmannstraße 7, 38106 Braunschweig, email: d.jahn@tu-bs.de

³ Prof. Dr. A.-P. Zeng, AG Systembiologie, GBF, Mascheroder Weg 1, 38124 Braunschweig, email: aze@gbf.de

⁴ Prof. Dr. D.C. Hempel, TU-BS, Institut für Bioverfahrenstechnik, Gaußstraße 17, 38106 Braunschweig, email: d.hempel@tu-bs.de

Netzwerksimulationen. Die Erkennung des Effekts ingenieurtypischer Parameter auf die Funktionalität der Netzwerke steht jedoch erst am Anfang.

Einleitung

Unter Systembiologie wird ein systematischer und holostischer Ansatz verstanden, um Systeme des Lebens ganzheitlich bzw. zumindest möglichst umfassend zu begreifen und quantitativ zu beschreiben. Systembiologie ist z.Zt. außerordentlich aktuell, denn nachdem auch die vollständige genetische Information (Genom) vieler Organismen bekannt ist, stellt sich die Frage nach der Vorausberechenbarkeit des phänotypischen Erscheinungsbildes oder: Kann man Leben berechnen? Jeder seriöse Wissenschaftler wird diese Frage derzeit mit einem klaren Nein beantworten. Aber nach einigem Nachdenken muss man doch erkennen, dass die Biowissenschaften auf dem besten Wege zu diesem anspruchsvollen Ziel sind. Tatsächlich ist in den Biowissenschaften in außerordentlich verkürztem Zeitrahmen das zu beobachten, was Physik, Chemie und Ingenieurwissenschaften in vergangenen Jahrhunderten bzw. Jahrzehnten erfuhren, nämlich den Einzug der Mathematik, die die zahlreichen Beobachtungen und Messungen ordnen und quantitativ beschreiben kann und exakte Voraussagen ermöglicht. Die konsequente Anwendung und Umsetzung der präzisen mathematisch erfassten Gesetzmäßigkeiten von Physik (und Chemie) hat zu einem rasanten technischen Fortschritt geführt.

Gegenüber physikalischen und chemischen Prozessen gelten biologische Vorgänge als ungemein vielschichtig und komplex. Um hier Zusammenhänge und Gesetzmäßigkeiten zu erkennen und zu quantifizieren, benötigt man zunächst Methoden zur Erhebung von ausreichend charakteristischen Kenngrößen auf verschiedenen Ebenen. In den vergangenen 15 Jahren wurden Hochdurchsatzmethoden zur Analyse zentraler biologischer Prozesse etabliert. Sie sind momentan wesentliche Werkzeuge der modernen Lebenswissenschaften, die sich weiterhin in stürmischer Entwicklung befinden. Mit solchen Hochdurchsatzverfahren wurde eine Datenflut erzeugt, deren Aufarbeitung zu einem neuen Fachgebiet, der Bioinformatik, führte. Diese ist ein unentbehrliches Werkzeug zur Klassifizierung und Bewertung der anfallenden Datenmengen, zur Erkennung von Gesetzmäßigkeiten und deren Umsetzung in mathematische bzw.

systembiologische Modelle. Dies erfordert von Biologen nicht nur die interdisziplinäre Zusammenarbeit mit Mathematikern, sondern auch mit Chemikern, Mediziner und Ingenieuren, denn neben dem grundsätzlichen Erkenntnisgewinn stehen gleichzeitig Anwendungen besonders im Bereich der Medizin und Biotechnologie im Fokus des Interesses. Wohin der Weg mit welchen Mitteln führen soll und wozu das Ganze überhaupt gebraucht wird, soll im Folgenden kurz skizziert werden (siehe auch den Beitrag von D.C.Hempel in diesem Heft). Zum Verständnis des bisherigen Entwicklungsstands und möglicher zukünftiger Vorgehensweisen ist ein kurzer Rückblick sinnvoll.

Grundprinzipien biologischer Prozesse

In ihrer Grundstruktur sind biologische Prozesse erkannt und lassen sich weitgehend auf physikalische und chemische Gesetzmäßigkeiten zurückführen. Abb. 1 zeigt, wie die im Erbgut in Form von Genen gespeicherten Informationen durch geeignete Stimuli in Botschaften (mRNA) umgeschrieben (Transkription) und an den Ribosomen daraus Proteine synthetisiert werden (Translation). Der Zusammenhang zwischen der Gensequenz, die aus nur vier Nukleinsäuren zusammengesetzt ist, und der Aminosäuresequenz des Proteins ist dabei eindeutig. Das gilt allerdings nicht für dessen räumliche Struktur. Die Proteine können als „Genprodukte“ aufgefasst werden, nämlich als Ergebnis des vom Gen ausgehenden Informationsflusses, obwohl die (20) Aminosäuren die eigentlichen Systembausteine der Proteinbildung darstellen. Die Proteine ihrerseits wirken überwiegend als Enzyme, d.h. sie sind Katalysatoren für die Umsetzung von Substraten und Stoffwechselintermediaten (Metabolite) und letztlich verantwortlich für alle katabolischen und anabolischen Vorgänge in der Zelle einschließlich möglicherweise involvierter Transportprozesse über die Zellwand und zwischen Zellkompartimenten.

Häufig sind Gene, die für eine bestimmte Funktion zuständig sind, in sogenannten Operons organisiert, die einer gemeinsamen Kontrolle unterliegen. Ein besonders gut untersuchtes Beispiel ist das *lac*-Operon, das breite biotechnologische Anwendungen gefunden hat. Die Regulation der Gene bzw. Operons, d.h. ob der Informationsfluss initiiert wird und die biokatalytischen Proteine in der Zelle produziert werden, kann relativ einfach sein, wie am Beispiel des für die mikrobielle

Quecksilberresistenz kodierenden *mer*-Operons in Abb. 2 gezeigt ist [1,2]. In Abwesenheit von Hg(II) werden die Strukturgene *mer TPAB* nicht abgelesen, da der Promoter P durch das Regularprotein R blockiert ist. Durch Hg(II) wird die Repression aufgehoben, indem dieses an R bindet und dadurch der Promoter frei ist für die Andockung der RNA-Polymerase. So kann die Ablesung der Strukturgene erfolgen und die Proteine synthetisiert werden, die die Transformation der für die Zelle toxischen Hg(II)-Verbindungen in flüchtiges metallisches Quecksilber Hg(0) bewirken. Sinkt die Hg(II)-Konzentration unter einen bestimmten Schwellenwert wird die Proteinbiosynthese wieder ausgeschaltet durch Bindung von R an P.

Deutlich komplexer als der in Abb. 2 gezeigte Induktionsmechanismus durch das Substrat Hg(II) ist die Regulation der Tryptophanbiosynthese, die vereinfacht in Abb. 3 wiedergegeben ist. Das *trp*-Operon umfasst die Gene, die für die 5 Syntheseschritte vom Chorismat, dem gemeinsamen Vorläufer der aromatischen Aminosäuren, zu Tryptophan kodieren und zusätzlich den Promoter und Operator sowie ein Regulationsgen enthalten. Hier liegt einerseits Inhibition der Aktivität der Syntheseproteine, andererseits eine effektive Feedback-Repression auf Transkriptionsebene durch das Produkt Tryptophan vor. In letzterem Fall wirkt Tryptophan als Korepressor, indem es an das inaktive Repressorprotein (Aporepressor) bindet und den Holo-repressorkomplex bildet, der spezifisch an das *trp*-Operon bindet und dadurch eine Ablesung der Strukturgene durch die RNA-Polymerase solange verhindert, bis die Tryptophankonzentration unter einen Grenzwert gesunken ist. Für die Überproduktion von Tryptophan ist daher eine effektive Exkretion aus dem Zellinneren von erheblicher Bedeutung [3]. Wie in Abb. 4 angedeutet, lassen sich die maßgeblichen Gleichungen auf Transkriptions- und Translationsebene formulieren und die daraus resultierenden Enzymkonzentrationen in die Metabolit (Tryptophan)-Massenbilanz einführen [4,6]. Bei kontinuierlichem Wachstum ($\mu = \text{konst.}$) führt die Lösung des Gleichungssystems als Folge der Regulation des *trp*-Operon durch Repression und Inhibition zu ausgeprägten phasenverschobenen intrazellulären Oszillationen für mRNA, die Synthesenzyme und das Produkt, die auch experimentell beobachtet werden (Abb. 5).

Das Beispiel der Genregulation und Produktbildung bei der Tryptophan-Biosynthese verdeutlicht deren komplexe Natur, die aber bei isolierter Betrachtung noch einer

mathematischen Beschreibung zugänglich ist, die die realen Verhältnisse zumindest qualitativ widerspiegelt. Die Regulation der Gene und die physiologische Antwort der Zellen auf zellinterne und -externe Reize gestaltet sich außerordentlich komplex, da Gene, Proteine und Metabolite auf vielfache Art interagieren. Das ist unmittelbar einsichtig angesichts der Größe von Genom, Transkriptom, Proteom und Metabolom (s. Abb. 1). In Einzelfällen sind Dutzende von Proteinen in die Regulation einzelner Gene bzw. Operons involviert. Die Erkennung und das Verständnis der komplexen Signalnetzwerke steht noch am Anfang, doch lassen sich schon jetzt zumindest einige wichtige mathematische Grundstrukturen der Genregulation erkennen [7,8]. Eine Behandlung auf Systemebene erfordert die Berücksichtigung aller zellulären Informations- und Stoffwechselflüsse, was zu integrierten Netzwerken führt, die bisher nur ansatzweise zur Verfügung stehen.

Im Omics-Zeitalter

Momentan ist alles „omics“. Diese Endung beschreibt methodische Ansätze; deren Ziel ist die Erfassung möglichst aller in der Zelle vorkommenden Moleküle eines speziellen Typs (DNA, mRNA, Proteine, Metabolite). Die Natur und Zusammensetzung dieser Moleküle, die nur eine begrenzte Stabilität aufweisen, hilft, den momentanen Zustand der Zelle zu verstehen. Die Information für alle zellulären Funktionen ist in Form von DNA gespeichert. Deshalb nennt man die Erforschung der gesamten DNA (Genom) eines Organismus „Genomics“. Ziel dieser Untersuchung ist daher die Zusammensetzung des Genoms eines Organismus mit 10^7 bis 10^9 Bausteinen aufzuklären. Dazu sind gerade in den letzten Jahren erhebliche methodische Fortschritte in der automatisierten DNA-Sequenzbestimmung erzielt worden, so dass wir heute auf eine Sammlung von mehr als 200 komplett bestimmten Genomen zurückgreifen können.

Die Identifizierung der funktionellen Abschnitte des Genoms, d. h. der Gene, ist das nächste Ziel der Genomics. Gleichzeitig sollen die Abschnitte (Promotoren) in der Nähe der Gene gefunden werden, die die Umschreibung des Gens in eine Botschaft (mRNA, Transkript) steuern. Die Zusammenfassung aller Botschaften nennt man Transkriptom, es gibt Auskunft über die momentane Arbeitsweise der Zelle. Die Technik zu dessen Analyse nennt man entsprechend Transkriptomics. Sie basiert auf

der sogenannten DNA-Array/Chip-Technologie. Dabei werden für jedes Gen eines Genoms Fängermoleküle auf einen Chip immobilisiert. Die mRNAs einer Zelle werden dann so markiert, dass sie leuchten können. Je mehr leuchtende mRNA nun ein Fänger auf dem Chip gebunden hat, desto stärker leuchtet er, d.h. desto größer ist die Menge des intrazellulär produzierten Botschaftermoleküls. Die Analyse des Transkriptoms unter Variation gezielt eingestellter externer Parameter, wie z.B. Wachstumstemperatur, Nährstoffzusammensetzung, Sauerstoffversorgung etc., gibt Auskunft darüber, welche Gene in der Zelle eingeschaltet sind. Hier entsteht ein großes Datenaufkommen.

Im letzten Schritt des zellulären Datentransfers werden nun die Botschaften (mRNAs) in die eigentlichen zellulären Funktionsträger, die Proteine, übersetzt. Diese erfüllen als Enzyme primär Funktionen in der Katalyse des zellulären Stoffwechsels. Auch das Skelett der Zelle, Signaltransfer, Regulations- und Differenzierungsprozesse zu unterschiedlichen Zelltypen werden meist über Proteine vermittelt. Man kann sich vorstellen, wie wichtig es für das Verständnis der zellulären Leistungen ist, diese „Arbeiter“ direkt nachzuweisen. Die Gesamtheit aller Proteine, die sich unter bestimmten Bedingungen in der Zelle aufhalten, nennt man Proteom, entsprechend die Methodik zu dessen Ermittlung Proteomics. Sie basiert auf der Isolation der Proteine aus Zellen und deren Auftrennung nach Ladung und Größe über zweidimensionale Gelelektrophorese. Man kann so bis zu 1000 Proteine auftrennen, die nach Anfärbung als Spots auf dem 2D-Gel sichtbar werden. Ein Beispiel zeigt Abb. 6. Mittels physikalischer Methoden (Massenspektrometrie) und bioinformatischer Werkzeuge (Datenbanken) gelingt es, die Identität der einzelnen Spots aufzuklären, d.h. welches Protein sie genau darstellen und welche Reaktionen sie katalysieren, falls es sich um Enzyme handelt. Auch kann über die Intensität der Proteinspots auf deren Menge zurückgeschlossen werden, allerdings nicht auf deren katalytische Aktivität in der Zelle.

Dazu bedarf es eines weiteren methodischen Ansatzes, der als Metabolomics bezeichnet wird und zum Ziel hat, alle Metabolite, also das Metabolom, simultan zu erfassen. Bei den Metaboliten handelt es sich um Stoffwechselzwischenprodukte, sie sind also die Substrate und Produkte der enzymatischen, in der Zelle ablaufenden Reaktionen. Je nach Zelltyp sind 500 bis 2000 Metabolite in der Zelle anzutreffen.

Sie können mittels chromatographischer und spektroskopischer Methoden identifiziert und quantifiziert werden.

Rekonstruktion von Netzwerken

Wie kann man nun die Datenfülle, die aus den Omics-Techniken erhältlich ist, nutzen? Das soll im Folgenden u.a. an Beispielen demonstriert werden, die an der Technischen Universität Braunschweig im Rahmen des DFG-Sonderforschungsbereichs 578 „Vom Gen zum Produkt“ bearbeitet werden. Dabei geht es um die Produktion von rekombinanten Proteinen in *Bacillus megaterium* und *Aspergillus niger*. Ziel ist zumindest die ansatzweise Aufstellung systembiotechnologischer Modelle für diese Organismen und das möglichst quantitative Verständnis der Expression und Sekretion heterologer Proteine (Biokatalysatoren, Pharmaproteine, speziell Antikörper) in Abhängigkeit von der Ernährungsphysiologie und ingenieurtypischen Betriebsparametern (Environom). Nach einer vom BMBF ausgegebenen Definition ist die Systembiologie ein neues interdisziplinäres Forschungskonzept, das die Entschlüsselung metabolischer und regulatorischer Netzwerke im Organismus zum Ziel hat, um physiologische Leistungen quantitativ und systematisch zu erfassen, zu modellieren und zu simulieren. So ist das metabolische Netzwerk eines Organismus eine notwendige Voraussetzung für das Verständnis seiner Physiologie und seines phänotypischen Verhaltens. Die Entschlüsselung des Netzwerks auf Systemebene ist von erheblicher Bedeutung sowohl für biotechnologische Anwendungen als auch für die biomedizinische Forschung wie die Erkennung von Pathogenitätsmechanismen und die Entwicklung neuer Wirkstoffe.

Es kommt also zunächst darauf an, aus den “Omics“-Daten die Netzwerke zu etablieren. Die Information darüber liegt im Genom, dessen Entschlüsselung mit bioinformatischen Werkzeugen die Mechanismen der Regulation einzelner Gene erkennen lässt und unabhängig davon das gesamte potentielle metabolische Netzwerk enthält, indem es für alle vom betreffenden Organismus ausführbaren Leistungen (Reaktionen) die Katalysatoren bereitstellt. Die tatsächlich vom Organismus unter den gewählten Kultivierungsbedingungen hergestellten Proteine lassen sich andererseits aus dem Proteom ohne Kenntnis des Genoms bestimmen.

So können durch massenspektrometrische Analyse der Spots des in Abb. 6 gezeigten 2 D-Gels die momentan unter den speziellen physiologischen Bedingungen in der Zelle vorhandenen Proteine identifiziert werden und eine Proteindatenbasis für diesen Organismus (*B.megaterium*) erstellt werden [9]. Daraus lässt sich mit Kenntnissen der Biochemie ein partielles, funktionelles metabolisches Netzwerk etablieren, wie Abb. 7 für den zentralen C-Metabolismus und die Biosynthese der Aminosäuren zeigt.

Allerdings ist die Auswertung des Proteoms ohne Kenntnis des Genoms und des daraus rekonstruierbaren metabolischen potentiellen Netzwerks mühsam. Der Königsweg ist daher, zunächst aus dem Genom *in silico* das metabolische Netzwerk aufzustellen. Für die Rekonstruktion, Visualisierung und Analyse metabolischer Netzwerke sind in den letzten Jahren verschiedene bioinformatische Methoden und Werkzeuge entwickelt worden, wie u. a. KEGG [10], WIT [11], EcoCyc, MetaCyc [12] und die metabolischen Karten von Boehringer Mannheim [13]. Grundsätzlich basiert die Rekonstruktion metabolischer Netzwerke auf der Identifikation metabolischer Enzyme (EC-Nummer) und der dazugehörigen biochemischen Reaktionen. Der erste Schritt besteht darin, aus den Genomdaten einzelne Gene bzw. kodierende Sequenzen (CDS) zu finden. Im nächsten Schritt werden diese genutzt, um Datenbanken nach deren möglicher Funktion abzufragen. Sind die Funktionen den einzelnen Genen zugeordnet, so wird daraus im letzten Schritt das metabolische Netzwerk rekonstruiert. Diese Vorgehensweise erfordert genomische Daten hoher Qualität, d.h., geringe Fehler in den kodierenden Sequenzen, welche erfahrungsgemäß eine 8 bis 10fache Abdeckung erfordert, insbesondere auch, um die Annotierung der einzelnen Gene zu erkennen. Die Vollständigkeit des genombasierten metabolischen Netzwerkes hängt andererseits von der zur Funktionszuordnung der Gene verwendeten Enzym- und Reaktionsdatenbank ab. Eine erweiterte und revidierte Reaktionsdatenbank wurde kürzlich von Ma und Zeng [14] vorgeschlagen. Die allgemeine vereinfachte Vorgehensweise bei der Rekonstruktion metabolischer Netzwerke ist in Abb. 8 schematisch wiedergegeben. Aus dem annotierten Genom werden die Enzyme und Reaktionen identifiziert. Unter Berücksichtigung nichtenzymatischer Reaktionen und Reversibilitätsbeziehungen lässt sich dann die Metabolit-Verbindungsmatrix erstellen und daraus, wie in Abb. 8 angedeutet, das metabolische Netzwerk rekonstruieren. Die Punkte des Netzwerkes bedeuten Enzymnummern bzw. Reaktionen, während die

Verbindungslinien die Metabolite darstellen. Mit der Größe des Genoms des Organismus steigt auch die Anzahl möglicher Reaktionen und Interaktionen. Die potentiellen metabolischen Netzwerke sind daher außerordentlich komplex und kaum zu „durchschauen“. Zur Analyse ihrer globalen Struktur und Eigenschaften werden Bioinformatik-Methoden eingesetzt, um z. B. topologische Kenngrößen zu ermitteln. Das gezeigte metabolische Netzwerk (Abb. 8) kann weiter modular strukturiert werden.

Seit kurzem stehen genomische Daten für *A.niger* und *B.megaterium* zur Verfügung. Allerdings ist die Annotation nicht vollständig und die Daten weisen zum Teil eine weniger als 5-fache Abdeckung der Genomsequenz auf und sind daher noch fehlerbehaftet. Die Aufstellung der metabolischen Netzwerke aus solchen unvollständigen Genomdaten ist mit Hilfe des Programms IdentiCS (Identification of Coding Sequences) [15] gelungen. Dabei handelt es sich um eine retrospektive Technik, mit der ausgehend von der Reaktionsdatenbank die kodierenden Gensequenzen identifiziert werden und daraus das metabolische Netzwerk rekonstruiert wird. Diese Vorgehensweise weist einige Nachteile auf. So lassen sich die Genstrukturen und Proteinsequenzen nicht exakt ermitteln, auch das Auffinden neuer Gene ist nicht möglich. Doch hat sich IdentiCS zur Aufstellung metabolischer Netzwerke grundsätzlich außerordentlich bewährt. Die bei IdentiCS verwendete umgekehrte Abfragetechnik ermöglicht selbst bei geringer (4fach) Genomabdeckung eine Identifizierung der kodierenden Sequenzen zu über 95%. Abb. 9 zeigt beispielhaft das metabolische Netzwerk von *A.niger* bereits in modularer Struktur. Es enthält 434 Enzyme und 1172 Metabolite. Man erkennt dicht gepackte Bereiche mit umfangreichen Verzweigungen und Verknüpfungen, die dem zentralen C- und N-Metabolismus (insbesondere Glycolyse, Pyruvat-Stoffwechsel, Pentosephosphatweg, Citronensäurecyklus) entsprechen und Ausläufer ohne nennenswerte Verzweigungen, die zur Biosynthese der Lipide, Aminosäuren und Purine führen. Natürlich kann die in Abb. 9 gezeigte Netzwerkstruktur unmittelbar in die üblichen Diagramme der Stoffwechselwege umgesetzt werden, z.B. unter Verwendung der KEGG-Struktur.

Direkte Information darüber, ob ein Gen eingeschaltet ist und mRNA produziert, gibt das Transkriptom (s. Abb. 1 und 3). Das sogenannte transkriptionelle regulatorische

Netzwerk (TRN) spielt bei allen zellulären Funktionen eine zentrale Rolle, da es die Expression der Gene reguliert und vielfach den Endpunkt von Signaltransduktionsketten darstellt. Aus der Analyse von Transkriptom-Daten ist es gelungen, zumindest für die gut untersuchten Organismen wie *Escherichia coli* und *Saccharomyces cerevisiae* genomweite, systemumfassende TRNs aufzustellen [16-19]. Dabei zeigte sich, dass bestimmte Gene nur dann exprimieren, wenn gleichzeitig andere aktiv sind, was zu einer insgesamt hierarchischen Struktur der TRNs führt. Als Beispiel zeigt Abb. 10 das kürzlich von Ma et al. [19] vorgeschlagene TRN von *E.coli*, das sich über 9 hierarchische Ebenen erstreckt. An der Spitze der Pyramide stehen die globalen Regulatorgene wie *crp*, das cAMP-Rezeptorprotein, welches insgesamt 112 weitere Gene reguliert. Weitere globale Regulatoren sind u.a. *rpos* (Stationärphasen-Sigmafaktor), *cspA,E* (cold shock-Proteine), *soxR* (Antwort auf oxidativen Stress) und *rpoN* (RNA-Polymerase Sigma-54-Untereinheit). Die aus Abb. 10 erkennbaren fächerförmigen hierarchischen Muster zeigen, dass die Befehlsübertragung fast ausnahmslos über Feed-Forward-Loops erfolgt und innerhalb der TRN keine Rückkopplung auftritt. Trotz vieler Überlappungen der fächerförmigen Muster ist das rein transkriptorische Netzwerk daher im Vergleich zum metabolischen Netzwerk relativ übersichtlich strukturiert. Die vielfältigen Wechselwirkungen der DNA und mRNA zum metabolischen Netzwerk, d.h. Interaktionen mit der Proteom- und Metabolomebene, sind allerdings noch unberücksichtigt. Die Zusammenführung der metabolischen und transkriptorisch regulatorischen Netzwerke zu integrierten Netzwerken auf Systemebene ist eine Herausforderung für die Zukunft [18].

Netzwerke stellen die eigentliche Wissensbasis der Systembiologie dar. Sie erlauben Organismenvergleiche (comparative genomics) sowie das Erkennen hierarchischer Strukturen und funktionaler Abhängigkeiten. Im Mittelpunkt des Interesses steht insbesondere die Funktionalität der metabolischen Netzwerke, die in Abhängigkeit von Änderungen auf physiologischer und genetischer Ebene (Einsatz von Mutanten, speziell Knock-out-Mutanten) sowie verfahrenstechnischer Einflussgrößen zu untersuchen ist. Aus dem Netzwerk lässt sich die Stoßrichtung für genetische Änderungen zur Erreichung bestimmter Ziele (Produktoptimierung) festlegen, aber auch die Wirksamkeit bereits erfolgter Modifikationen auf genetischer Ebene überprüfen. Die Einflüsse der Ernährungsphysiologie und der bei der

Organismenkultivierung vorliegenden verfahrenstechnischen Parameter (s. Abb. 11) werden oft unter dem Begriff Environom zusammengefasst. Es kommt also darauf an, zu prüfen, welche Teile des aus dem Genom rekonstruierten metabolischen Netzwerk unter den durch das Environom festgelegten einstellbaren Bedingungen aktiv sind. Dazu dienen die „Omics“-Techniken, nämlich die Analyse des Proteoms, Metaboloms und Secretoms. Unter dem letzten versteht man die aus der Zelle in das Kultivierungsmedium insgesamt sekretierten Verbindungen. Informationsreich im Hinblick auf die Funktionalität des metabolischen Netzwerks und somit die phänotypische Ausprägung des betrachteten Organismus, ist speziell die Analyse der metabolischen Flüsse (MFA = Metabolic Flux Analysis) in Abhängigkeit von genetischen Variationen und Änderungen/Störungen des Environoms. Das metabolische Fluxom kann aufgrund experimenteller Einschränkungen natürlich nicht so detailreich sein wie das Metabolom, liefert aber quantitative Aussagen über die Beteiligung einzelner Stoffwechselwege. Solche Kenntnisse stellen die rationale Grundlage dar, einerseits für gezielte Experimente aufgrund von bestimmten Hypothesen und Netzwerksimulationen, sowie andererseits für das gezielte Metabolic Engineering und spezielle verfahrenstechnische Maßnahmen.

Metabolische Flussanalyse

Grundlage der Analyse aktiver metabolischer Netzwerke ist die Massenbilanzierung. Dabei müssen eine Vielzahl von Reaktionen mit Verzweigungen und Verknüpfungen berücksichtigt werden, wobei die stöchiometrischen Zusammenhänge zwischen den Metaboliten abgesehen von wenigen Ausnahmen aus der Biochemie bekannt sind. Üblicherweise laufen biologische Prozesse unter instationären Bedingungen ab, unter denen eine metabolische Bilanzierung außerordentlich schwierig ist, da entsprechende Daten nicht mit ausreichender Genauigkeit und Fülle zur Verfügung stehen. Wichtigste Voraussetzung für eine erfolgreiche Bilanzierung ist daher das Vorliegen stationärer Zustände. Das erfordert Untersuchungen in stationären kontinuierlichen Kulturen, damit die Organismen im statistischen Mittel in gleichen Zuständen vorliegen. Durch Messung der extrazellulären Flüsse für Substratverbrauch, O₂-Aufnahme, Biomasse-Erzeugung, CO₂-Entwicklung sowie Produkt- und Nebenproduktbildungen stehen diese dann für eine Bilanzierung zur Verfügung. Das Bilanzgleichungssystem ist allerdings meist unterbestimmt und lässt

sich nur durch zusätzliche Annahmen und Einschränkungen lösen. Die erhaltene Lösung ist in Bezug auf die intrazellulären Metabolitflüsse jedoch wegen des Auftretens paralleler und reversibler Stoffwechselwege und metabolischer Zyklen nicht eindeutig (Futile Cycles = scheinbar nutzlose Kreisläufe). Die über das sogenannte P/O-Verhältnis gekoppelten Redox- und Energiebilanzen für die Kofaktoren NADH und NADPH und den Energietransporter ATP stellen einen besonderen Unsicherheitsfaktor bei der Bilanzierung dar. Hier könnte die quantitative Vermessung intrazellulärer Kenngrößen z.B. durch NMR-Messungen oder die Techniken zur Metabolom-Erfassung weiterhelfen.

Stattdessen wird eine andere leistungsfähige Methode derzeit überwiegend eingesetzt, nämlich die Verwendung von markierten Substraten mit stabilen Isotopen. Bevorzugt erfolgt die Markierung bestimmter C-Atome des Substrats mit ^{13}C , z. B. mit $[1-^{13}\text{C}]$ - oder $[\text{U-}^{13}\text{C}]$ -Glucose. Da die Stöchiometrie und Mechanismen der katabolischen und anabolischen Stoffwechselwege detailliert bekannt sind, lässt sich die Verteilung der Isotopomere (gebildet aus Isotop und Isomer) eines Metabolits für jedes markierte Substrat vorausberechnen bzw. lassen sich aus der gemessenen Isotopomerenverteilung eines bestimmten Metaboliten oder einer Metabolitgruppe quantitative Schlussfolgerungen über die relativen Anteile der Flüsse der zentralen Stoffwechselwege ziehen. Da zumindest einige der extrazellulären Metabolitflüsse bekannt sind, lassen sich aus den Isotopomerenverteilungen auch die Absolutwerte der intrazellulären Metabolitflüsse berechnen. Zur Ermittlung der Isotopomerverteilung können nach geeigneter Probenvorbereitung NMR-Verfahren und massenspektrometrische Analysemethoden herangezogen werden. Beide Techniken sind, besonders was die Datenauswertung betrifft, aufwendig, erfordern umfangreiche Korrekturen sowie vertiefte mathematische Kenntnisse und weisen spezifische Vor- und Nachteile auf [20-25]. Eine prägnante Zusammenfassung sowohl der experimentellen als auch der mathematischen Probleme, die bei der metabolischen Flussanalyse mittels ^{13}C -markierter Substrate mit NMR- als auch MS-Techniken auftreten, findet sich bei Wiechert [25].

Fluxom-Messungen für *Bacillus megaterium*

Die Vorgehensweise der Ermittlung der metabolischen Flüsse soll am Beispiel des im Rahmen des SFB 578 untersuchten *B. megaterium*s aufgezeigt werden. Abb. 12 zeigt die experimentelle Grundordnung einer kontinuierlichen Kultivierung, die neben anderen Nährstoffen mit Glucose als C-Quelle, die einen bestimmten Anteil ^{13}C -Markierung aufweist, versorgt wird. Bei der aeroben Kultivierung werden Glucoseverbrauch, Acetat- und Biomassebildung, sowie die CO_2 -Entwicklung und das $^{12}\text{CO}_2/^{13}\text{CO}_2$ -Verhältnis ermittelt. Biomasseproben werden nach Lyophilisation total hydrolysiert und die gebildeten Aminosäuren nach Derivatisierung gaschromatographisch getrennt und massenspektrometrisch detektiert. Da die Biosynthese der Aminosäuren, wie aus Abb. 7 (und 14) zu ersehen ist, ausschließlich aus Molekülen erfolgt, die Bestandteile des zentralen C-Metabolismus (Glycolyse, Pentosephosphatweg, TCA-Zyklus) sind, kann aus dem vermessenen Isotopomeren-Muster der proteinogenen Aminosäuren die Isotopomerenverteilung der zentralen C-Metabolite berechnet werden. Auf Basis eines Isotopomerenmodells des metabolischen Netzwerks lassen sich daraus die Metabolitflüsse bestimmen. Es sei an dieser Stelle darauf hingewiesen, dass die metabolische Flussanalyse hinsichtlich der Durchführung der Experimente und der involvierten Analytik außerordentlich zeit- und kostenaufwendig ist und erfahrenes Personal erfordert. Ebenso ist die Auswertung der vielen anfallenden Daten und ihre statistische Bewertung komplex und erfordert mathematische Kenntnisse und Erfahrung in der Behandlung umfangreicher Gleichungssysteme. Das wird deutlich, wenn man bedenkt, dass ein Metabolit mit nur 3 C-Atomen $2^3 = 8$ Isotopomere aufweist und deren Anzahl bei 5 C-Atome der derivatisierenden Aminosäure und ihrer bei der Ionisierung auftretenden Bruchstücke bereits auf $2^6 = 64$ ansteigt. Bei der massenspektrometrischen Analyse der Isotopomeren sind umfangreiche Korrekturen zur Berücksichtigung der natürlichen Isotopenverteilung notwendig. Andererseits sind die Methoden zur Fluxomermittlung heute soweit fortgeschritten, dass ähnlich den anderen „Omics“-Techniken Hochdurchsatzmethoden für das Fluxom in Sicht sind. Sauer [26] hat kürzlich darauf hingewiesen, dass die Genauigkeit der Massenspektrometrie ausreichend ist, um die relativen Flussverhältnisse auch in miniaturisierten Reaktoren ohne nennenswerte Einbußen an Informationsgehalt zu bestimmen. Natürlich muss dabei die Bedingung des stationären Fließgleichgewichts fallen

gelassen werden. Jedoch stellt sich das Problem, wie klein Bioreaktoren sein dürfen, um reproduzierbare Ergebnisse zu erhalten. Das ist bis in den 1 ml Maßstab von Tiefkammern von Mikrotiterplatten gelungen [26], womit für viele Fragestellungen eine passable Hochdurchsatztechnologie zur relativen Fluxomerfassung zur Verfügung steht.

Was die Deutung der massenspektrometrischen Analyse der proteinogenen Aminosäuren betrifft, so können häufig vereinfachte Beziehungen anstatt des vollständigen Isotopomerenmodells herangezogen werden, da nur das Verhältnis der Flüsse interessiert. Leicht interpretierbar ist z.B. die Beteiligung des Pentosephosphatwegs (PPP) aus der Isotopomerenverteilung der aromatischen Aminosäuren bei Einsatz von [$1\text{-}^{13}\text{C}$]-Glukose (s. Abb.14). Da im oxidativen Teil des PPP dann gerade das markierte C-Atom der Glukose als CO_2 freigesetzt wird und Histidin nur aus Pentose-5-phosphat sowie Phenylalanin und Tyrosin nur aus Erythrose-4-phosphat und Phosphoenolpyruvat (PEP) entstanden sein können, deuten die Isotopomerengehalte der aromatischen Aminosäuren über den natürlichen Wert hinaus auf die Reversibilität der Transaldolase- und Transketolase-Reaktionen vom Fruktose-6-phosphat und Triose-3-phosphat hin zum Pentose-5-phosphat und Erythrose-4-phosphat. Ein anderes einleuchtendes Beispiel zeigt Abb.13. Hier ist wiederum für *B.megaterium* MS 941 die Isotopomerenverteilung der C3-Komponenten 3-Phosphoglycerat (3PG), PEP und Pyruvat (Pyr) dargestellt. Es zeigt sich klar, dass für 3PG und PEP kein signifikanter Unterschied in den m+0- und m+1-Massen auftritt, speziell die m+1-Massen sind mit 6-7% fast identisch, während die für m+2 praktisch vernachlässigbar sind. Dagegen ist der Unterschied gegenüber Pyr ausgeprägt, insbesondere beträgt der Anteil für m+2 hier 13%. Für Pyr muss daher angenommen werden, dass ein zusätzlicher Stoffwechselweg auftritt, der neben der Glykolyse, die markierte C-Atome in Pyr einführt, speziell zur Erhöhung der m+2-Fraktion seiner Isotopomeren beiträgt. Die Intermediate PEP und Pyc stellen wichtige Schlüsselverbindungen zwischen Glykolyse und TCA-Zyklus dar und geben Auskunft über den Ablauf von anaplerotischen und glykoneogenetischen Reaktionen, die energetisch von erheblicher Bedeutung sind. Die aus den Messdaten berechneten Massenflussverhältnisse sind in Abb.14 ersichtlich, in der die absoluten Flüsse für zwei Wachstumsraten von *B.megaterium* MS 941 wiedergegeben sind.

Bei Verwendung von $[1-^{13}\text{C}]$ -Glukose als C-Quelle spielt der TCA-Zyklus mit einer Serie von Kreislaufreaktionen eine wichtige Rolle in der Erzeugung höher markierter Intermediate. Einerseits bewirkt das Enzym PyrDH den Einbau von Acetyl-CoA und andererseits die Aktivität der Pyrcarboxylase die Einschleusung von markiertem CO_2 in den TCA-Zyklus. Sind markierte Verbindungen im TCA-Zyklus vorhanden, so können diese mit Hilfe des Malatenzyms in Pyr überführt werden, was, wie beobachtet, zum Anstieg des m+2-Anteils des Pyr führt (s. Abb.13). Andererseits zeigt der untersuchte Organismus keinerlei Aktivität für die PEPcarboxykinase, die die Reaktion vom Oxalacetat (OAA) zu PEP katalysiert. Das erklärt die Stabilität der Isotopomerenverteilung des PEP im Vergleich zu 3PG und die Änderungen gegenüber Pyr, zumal die Reaktion von PEP zu Pyr irreversibel verläuft.

Tabelle 1 gibt einen Vergleich der untersuchten *B. megaterium*-Stämme mit anderen *Bacilli*-Arten. Was die *B. megaterium*-Stämme betrifft, so fällt auf, dass für WH 320 im Vergleich zu MS 941 1. der oxidative Teil des PPP stärker ausgeprägt und damit gleichzeitig seine Reversibilität deutlich heruntergefahren ist, 2. der Fluss vom Pyr zum Ac-CoA zugenommen hat und 3. generell erniedrigte Flüsse um den Pyr-Knoten auftreten. Im Vergleich zu den Literaturergebnissen [27,28] mit anderen *Bacilli*-Arten treten vergleichbare Flüsse in der Glykolyse auf. Auffallend ist dagegen, dass von den Rückreaktionen vom TCA-Zyklus zu den C3-Körpern der Glykolyse entweder die PEPcarboxykinase katalysierte Umsetzung von OAA zu PEP oder die vom Malatenzym katalysierte Reaktion von Malat zu Pyr aktiv ist. Für keine der bisher betrachteten *Bacilli*-Arten wurde beobachtet, dass beide Reaktionen auftreten.

Ergebnisse zum Fluxom

Die in Tabelle1 angegebenen eigenen Daten verdeutlichen, dass die über Massenspektrometrie der proteinogenen Aminosäuren erhaltenen Fluxomdaten der zentralen Stoffwechselwege relativ grosse Fehler aufweisen. Das ist zum einen bedingt durch Unsicherheiten, die mit der exakten Bestimmung der Zusammensetzung der Biomasse zusammenhängen, zum anderen mit der Methodik selbst [25]. Die beobachteten Änderungen sind i.a. jedoch so signifikant, dass sie ausserhalb des Messfehlers liegen und die getroffenen Schlussfolgerungen tendenziell richtig sind. Wenn man daher das Fluxom des zentralen Metabolismus als

prägnanteste und phänotypisch relevanteste Ausprägung eines Organismus in Betracht zieht, so stellt sich die Frage, inwieweit sowohl Änderungen im Genom als auch im Environom phänotypisch sichtbar werden, wobei besonders das letztere für den Verfahrenstechniker von Interesse ist. Zunächst soll das Genom betrachtet werden. Bei Untersuchungen mit den *B. megaterium*-Stämmen MS 941 und WH 320, wobei es sich bei WH 320 um eine genetisch nicht definierte, chemische Mutante handelt, treten charakteristische Änderungen in der Reversibilität des PPP, dem Trioseanteil aus der Glykolyse und der Bildung von OAA aus Pyr sowie von Pyr aus Malat auf, jedoch bleibt das generelle Bild erhalten. Das ist in Übereinstimmung mit einer umfassenden Studie von Fischer und Sauer [29], die Fluxome von 137 Mutanten von *B. subtilis* untersuchten und keine signifikanten Änderungen feststellen konnten. Die Ursache wird darin vermutet, dass auch unter suboptimalen Stoffwechselbedingungen durch regulatorische Effekte auf Transkriptomebene eine „Standby“-Operationsfähigkeit der Organismen beibehalten wird, die sich auf Kosten des Wachstums an wechselnde Umweltbedingungen anpassen kann. Offenbar sind Rigidität und Robustheit des integralen funktionellen Netzwerks allgemeine evolutionäre Auslegungsprinzipien von Organismen.

Es muss daher die Frage gestellt werden, ob das Fluxom der zentralen Stoffwechselwege allein geeignet sein kann, als indikatorisches Instrument für die heterologe Proteinproduktion, die im SFB 578 angestrebt wird, zu dienen. Man kann abschätzen, dass im Vergleich zur Biomassebildung der Stoff- und Energiebedarf für die Herstellung von Fremdproteinen nur wenige Prozent ausmachen kann. Daher scheinen andere „Omics“-Methoden eher geeignet zu sein, in Erfahrung zu bringen, durch welche Limitationen die Produktion und Sekretion heterologer Proteine in den extrazellulären Raum in aktiver Form begrenzt ist. Vermutlich ist das Ausschleusungssystem überlastet und die Bereitstellung von Helfer- und Faltungsproteinen mangelhaft.

Ganz anders sind natürlich solche Prozesse hinsichtlich der Bedeutung des Fluxoms einzuschätzen, wenn es um die massive Überproduktion eines Primärmetaboliten geht, der entweder direkt als Zielprodukt in den zentralen Stoffwechselwegen auftritt oder unmittelbar ein Vorläufer eines solchen darstellt. Dazu sind einige Beispiele, bei

denen „Omics“- und speziell Fluxomics-Daten zur Verbesserung der Produktivität beigetragen haben, in Tabelle 2 zusammengetragen.

Prozessorientierte Systembiotechnologie und Verfahrenstechnik

In Abb.11 sind sehr global die extrazellulären Signale, die auf eine Zelle einwirken können, dargestellt und die auf Basis des vorhandenen Genoms den Phänotyp der Zelle prägen bzw., wie es auch formuliert werden kann, die Funktionalität des Genoms und der daraus resultierenden Netzwerke festlegen. Es versteht sich, dass das Nährstoffangebot die meisten Stimuli für das metabolische Netzwerk (Physiologie) zur Verfügung stellt. Mit Kenntnis des Fluxoms und Metaboloms lassen sich nun gezielt Änderungen im Genom vornehmen, um z.B. das Substratspektrum zu erweitern oder einzuengen bzw. substratverbrauchende störende Nebenreaktionen zu unterbinden. Dem Verfahrenstechniker obliegt es dann, im Reaktor Bedingungen einzustellen, damit die Substratversorgung oder die Produktabführung über das Medium optimal ist. Das erfordert u.a. die Kenntnis der Substrataufnahmerate, wobei überwiegend der Monod-Ansatz angenommen wird. Da der Prozess wegen $K \ll S$ sich vorwiegend im Bereich der Substratsättigung befindet, stellt die Einhaltung der Mischzeit in der flüssigen Phase bei dem Scale-up biotechnologischer Prozesse keine echte Herausforderung dar. Die Situation ändert sich jedoch, wenn eine Produktinhibierung oder Substratüberschusshemmung auftritt. Im ersteren Fall ist eine Produktseparation anzustreben, im letzteren sind Zufütterungsstrategien anzuwenden und für eine möglichst schnelle Gleichverteilung des Substrats über den Reaktionsraum zu sorgen, damit die Reaktionsrate auf ihren Maximalwert eingestellt werden kann.

An diesem einfachen Beispiel wird bereits deutlich, dass der Kinetik der Reaktionen, die in das metabolische Netzwerk involviert sind, eine wichtige Rolle zukommt, ein Punkt, der bisher nur vereinfacht betrachtet wurde (s. z. B. Abb. 4), aber für eine effiziente Verfahrensführung essentiell ist. Es ist klar, dass die makroskopischen Gesetze für Wachstum, Substratverbrauch und Produktbildung angesichts der komplexen Netzwerkstruktur der Zelle nur als Grenzfälle von Relevanz sein können, nicht aber das dynamische Verhalten erfassen können, dem biotechnologische Prozessen aufgrund ihres Batch-Charakters zwangsläufig unterworfen sind. Hinzuweisen ist auch darauf, dass aus „Omics“-Methoden keine unmittelbar

verwertbaren kinetischen Informationen zugänglich sind. Allenfalls durch simultane Ermittlung von Proteom (liefert Schätzwerte für die Enzymmasse, jedoch nicht über dessen Aktivität), Metabolom und Fluxom, unterstützt durch *in vitro*-Aktivitätsmessungen und Sensitivitätsuntersuchungen (Metabolic Control Analysis), lassen sich kinetische Daten entwickeln. Hier wird eine wichtige Aufgabe des Bioverfahrenstechnikers gesehen, wobei in vielen Fällen auf die umfangreiche biochemische Literatur zurückgegriffen werden kann. Erforderlich sind auch die Durchführung von Netzwerksimulationen, eine Aufgabe, die dem Verfahrenstechniker entgegen kommt, da das dazu verwendete Simulationswerkzeug für die Auslegung komplexer Chemie- und Trennanlagen, sowie Regel- und Steuerkreisen zumindest partiell zur Verfügung steht.

Das Problem der Sauerstoffversorgung aerober Kulturen in Grossreaktoren und bei hohen Zelldichten wird weiterhin von Aktualität für die Verfahrenstechnik bleiben, die angesichts der Verfügbarkeit von Netzwerken auf dieser Ebene durch sorgfältig geplante Experimente und Simulationen angegangen werden muss. Fragen die in diesem Zusammenhang von Bedeutung sind, sind die Fluxomänderungen beim Übergang von aeroben zu anaeroben Bedingungen, beim Auftreten von oxidativem Stress sowie zu dem eigentlichen Charakter mikroaerober Prozesse (z.B. 2,3-Butandiol).

Auch Scherstress, z.B. durch intensives Rühren, ist von Einfluss auf die Ausbildung bestimmter Phänotypen. Bei Prokaryonten konnten solche Effekte bisher kaum eindeutig nachgewiesen werden, im Gegensatz zu Eukaryonten, speziell tierischen Zellen, die wegen möglicher schädigender Schereffekte stets mit besonderer Vorsicht behandelt werden. Kürzlich konnten jedoch auch für einen typischen Prokaryonten (*Bacillus subtilis*), der unter drastisch unterschiedlichen Scherbedingungen kultiviert wurde, deutliche Effekte nachgewiesen werden. Sahoo et al. [46] beobachteten das Wachstum von *Bacillus subtilis* in einem Schüttelkolben bei vernachlässigbarer Scherrate und in einem Bioreaktor mit Couette-Strömung bei einer Scherrate (γ) von bis zu 1482 s^{-1} . Abb. 15 zeigt, dass bei der höheren Scherrate kleinere Zellen auftreten, zusätzlich sich aber höhere Zelldichten (190%) mit deutlich erhöhten metabolischen Aktivitäten einstellen. Unter sonst gleichen Bedingungen stieg bei erhöhter Scherung die Wachstumsrate auf 245% und die

Katalaseaktivität auf 335%. Sahoo et al. [46] konnten klar zeigen, dass bei erhöhter Scherrate reaktive Sauerstoffspezies (Radikale) auftreten und mit der Expression des allgemeinen Stressregulators σ_B auf transkriptionsebene korrelieren. Auslöser für die Veränderung des Phänotyps ist vermutlich, dass bei hohen Scherraten eine in der Zellmembran lokalisierte NADH-Oxidase derart stimuliert wird, dass die nachgewiesenen reaktiven Sauerstoffspezies entstehen und daraus die Stressantwort resultiert.

Anders als bei Prokaryonten reagieren Pilze wie der im SFB 578 untersuchte *Aspergillus niger* außerordentlich empfindlich auf Scherstress [47-49]. Bei diesem filamentös wachsenden Pilz liegt bereits am Anfang der Kultivierung eine hohe Sensitivität gegenüber Scherung vor. Die phänomenologischen Vorgänge der Sporenaggregation und Auskeimung sind in Abb.16 dargestellt [47,48]. Die Konidien unterliegen zunächst einer ersten spontanen Aggregationsphase; nach etwa 6 h beginnt eine deutliche Auskeimung und damit verbunden eine zweite Aggregationsphase, die zu fadenartigen Gebilden und schließlich zur Ausbildung von Pellets führt. Das Wachstum der Pellets hinsichtlich Größe und Dichte ebenso wie die Produktbildung (Glukoamylase) zeigen, wie in Abb. 17 dargestellt, ausgeprägte Abhängigkeiten vom Leistungseintrag [49]. Alle phänomenologischen Vorgänge lassen sich mit makroskopischen Ansätzen unter Einführung von empirischen Koeffizienten hervorragend beschreiben [47-51], doch ist bisher der Brückenschlag zu den biologischen Parametern z.B. charakteristische Veränderungen in den „Omics“- Ebenen und Netzwerkstrukturen nicht bzw. erst ansatzweise methodisch gelungen [52].

Schlussfolgerung

Die derzeitige Situation der auf die Entwicklung und Verbesserung von Bioprozessen und Bioprodukten ausgerichteten Systembiotechnologie und die Stellung der Verfahrenstechnik darin ist in Abb. 18 dargestellt. Im Prinzip geht es darum, den Organismus auf Basis des vorhandenen Genoms bzw. gezielt hergestellter Mutanten mit Hilfe des Environoms derart zu funktionalisieren, dass sich der Phänotyp mit den gewünschten Leistungen ergibt.

Ausgehend von den Gensequenzen ergeben sich mit Hilfe der Bioinformatik in deterministischer Weise die potentiellen Transkriptome und Proteome, sowie die dazugehörigen Netzwerke, deren Funktionalität für ein bestimmtes, einstellbares Environom durch die „Omics“ Methoden ermittelt wird. Dabei ist es wichtig, dass möglichst viele „Omics“ Daten (Transkriptom, Proteom, Metabolom, Fluxom, Sekretom) unter exakt gleichen Bedingungen ermittelt werden. Das ist derzeit auch deshalb von Bedeutung, weil die eingesetzten Methoden häufig fehlerbehaftet sind und man deshalb auf Aussagen über relative Änderungen angewiesen ist. Es ist daher ebenfalls wichtig, die „Omics“ Daten unter gezielter Variation von Eigenschaften des Environoms, z.B. pH, Substratangebot, hydrodynamische Parameter wie Leistungseintrag oder Scherrate zu erfassen, um relevante Zusammenhänge zu erkennen. Je nach Wissensstand lassen sich dann neue Experimente anlegen, die sowohl neue Ergebnisse zum Einfluss des Environoms berücksichtigen als auch neue Mutanten erfordern können (Metabolic Engineering). Dieses iterative systembiotechnologische Vorgehen erfordert von der Aufgabenstellung her zwingend die koordinierte Zusammenarbeit von Molekularbiologen, Bioinformatikern und Verfahrenstechnikern. Für den Verfahrenstechniker steht zunächst im Vordergrund streng reproduzierbare Kultivierungs- und Produktionsbedingungen unter systematischer Variation von Kenngrößen des Environoms zu realisieren und dabei die makroskopischen Kenngrößen (Substratverbrauch und -eintragsrate, Biomassebildung, etc.) und das Sekretom (Produktbildung) simultan während des Kultivierungsverlaufs zu erfassen. Eine wichtige Aufgabe ist weiterhin die Ermittlung kinetischer Ansätze und Daten für Netzwerkreaktionen aus „Omics“-Methoden, um daraus zusammen mit Bioinformatikern Netzwerksimulationen durchzuführen sowie mit Molekularbiologen neue gezielte und verifizierende Experimente als Ansatzpunkte für ein erfolgreiches Metabolic Engineering zu etablieren. Dies macht deutlich, dass in einer auf Bioprozesse fokussierten Systembiotechnologie für den Verfahrenstechniker wichtige Aufgaben anstehen, die von ihm allerdings erfordern, seinen bisherigen Kenntnisstand zumindest auf metabolische Netzwerke und deren Handhabbarkeit zu erweitern.

Der derzeitige Stand der Systembiotechnologie ist gekennzeichnet durch die schnell fortschreitende, von der Bioinformatik forcierte Konstruktion von biologischen

Netzwerken, die sich hierarchisch und modular strukturieren lassen und so für eine Anwendung nutzbar werden. Die anstehenden Aufgaben der prozessorientierten Systembiotechnologie sind zunächst die Funktionalisierung der Netzwerke, was mit Hilfe koordinierter Experimente erfolgen muss. Der derzeitige Kenntnisstand erfordert für jeden Organismus noch gesonderte Überlegungen. Vor allem steht die Forschung hinsichtlich der Erkennung des Einflusses ingenieurtypischer Kenngrößen auf die Funktionalisierung der Netzwerke erst am Anfang.

Danksagung

Der Deutschen Forschungsgemeinschaft danken wir für die Unterstützung der Arbeiten im Rahmen des SFB 578 „Vom Gen zum Produkt“.

Literatur

- [1] J.M. Horn, M. Brunke, W.-D. Deckwer, K.N. Timmis, Appl. Environ. Microbiol., **1994**,60,357/362
- [2] T. Barkay, S.M. Miller, A.O. Summers, FEMS Microbiol. Rev., **2003**,27,355/38
- [3] R. Krämer, FEMS Microbiol. Rev., **1994**, 13,75/94
- [4] S. Sinha, Biotechn. Bioeng.,**1988**,31,117/124
- [5] B.T. Koh, M.G.S. Yap. Biotechn. Bioeng.,**1993**,41,707/714
- [6] Z.-L. Xiu, A.-P. Zeng, W.-D: Deckwer, J. Biotechn., **1997**,58,125/140
- [7] J.J. Tyson, K.C. Chen, B. Novak, Curr. Opinion Cell Biol., **2003**,15,221/231
- [8] H.M. Sauro, B.N. Kholodenko, Prog. Biophys. Molec. Biol., **2004**,86,5/43
- [9] W. Wang, R. Hollmann, T. Fürch, M. Nimtz, M. Malten, D. Jahn, W.-D. Deckwer, Proteome Sci. **2005**,3,4 (<http://www.proteomesci.com/content/3/1/4>)
- [10] M. Kaneshi, S. Goto, Nucl. Acids Res., **2000**,28,27/30 (<http://www.genome.ad.jp/kegg>)
- [11] R. Overbeek et al., Nucl. Acids Res., **2000**,28,123-125 (<http://www.wit.mcs.anl.gov/WIT>)
- [12] P.D. Karp, M. Riley, M. Saier, I.T. Paulsen, S.M. Paley, A. Pellegrini-Toole, Nucl. Acids Res., **2000**,28, 56/59 (<http://www.biocyc.org/ecocyc/>)
- [13] G. Michael, Biochemical Pathways, **1999**, Spektrum Akademischer Verlag, Heidelberg, Berlin
- [14] H. Ma, A.-P. Zeng, Bioinformatics, **2003**, 19, 270/277
- [15] J. Sun, A.-P. Zeng, BMC Bioinformatics, **2004**, 5, 112 (<http://www.biomedcentral.com/1471-2105/5/112>)
- [16] T.I. Lee et al., Science, **2002**, 298, 799/804
- [17] S.S. Shen-Orr, R. Milo, S. Mangan, U. Alon, Nat. Genet.,**2002**, 31, 64-68
- [18] M.J. Herrgard, M.W. Covert, B.O. Palsson, Curr. Opinion Biotechnol., **2004**, 15,70/77
- [19] H. Ma, J. Buer, A.-P. Zeng, BMC Bioinformatics, **2004**, 5,199 (<http://www.biomedcentral.com/1471-2105/5/199>)
- [20] B. Christensen, J. Nielsen, Adv. Biochem. Eng. Biotechnol., **1999**, 66,209/231
- [21] C. Wittmann, E. Heinzle, Biotechnol. Bioeng., **1999**, 62, 739/750

- [22] C. Wittmann, M. Hans, E. Heinzle, Anal. Biochem., **2002**, 307, 379/382
- [23] E. Fischer, U. Sauer, Eur. J. Biochem., **2003**, 270, 880/891
- [24] E. Fischer, N. Zamboni, U. Sauer, Anal. Biochem., **2004**, 325, 308/316
- [25] W. Wiechert, Metabol. Eng., **2001**, 3, 195/206
- [26] U. Sauer, Curr. Opinion Biotechnol., **2004**, 15, 58/63
- [27] M. Dauner, S. Tazio, U. Sauer, J. Bacteriol., **2001**, 183,7308/7317
- [28] T. Christiansen, B. Christensen, J. Nielsen, Metab. Eng., **2002**,4, 159/169
- [29] E. Fischer, U. Sauer, Nat. Genet., publ. online 8 May, **2005**;
doi:10.1038/ng1555
- [30] K. Menzel, K. Ahrends, A.-P. Zeng, W.-D. Deckwer, Biotechn. Bioeng.,
1998, 60, 617/626
- [31] W. Wang, J. Sun, M. Hartlep, W.-D. Deckwer, A.-P. Zeng, Biotechn. Bioeng.;
2003, 83, 525/536
- [32] W. Wang, J. Sun, M. Nimtz, A.-P. Zeng, W.-D. Deckwer, Proteome
science,**2003**, 1, 6 (<http://www.proteomesci.com/content/>)
- [33] J. Sun, J.van den Heuvel, P. Soucaille, Y. Qu, A.-P. Zeng, Biotechn.
Progress, **2003**, 19, 263/273
- [34] J. Flores, Metabolic Flux Distribution, Modelling and Process Optmization
of Xanthan Production, VDI Fortschritt-Berichte Reihe 17,Nr. 172, VDI
Verlag Düsseldorf GmbH, **1998**
- [35] B.J. Eickmanns, D. Rittmann, H. Sahm, J. Bacteriol., **1995**, 177, 774/782
- [36] A. Marx, A.A. de Graaf, W. Wiechert, L. Eggeling, H.Sahm, Biotechnol.
Bioeng., **1996**, 49, 111/129
- [37] S. Petersen, A. A. De Graaf, L. Eggeling, M. Möllney, W. Wiechert, H.
Sahm, J. Biol. Chem., **2000**,275,35932/35941
- [38] H. Sahm, L. Eggeling, A. A. De Graaf, Biol. Chem., **2000**,381,899/910
- [39] C. Wittmann, E. Heinzle, Metab. Eng., **2001**, 3,173/191
- [40] C. Wittmann, E. Heinzle, Appl. Env. Microbiol., **2002**,68,5843/5859
- [41] M. Dauner, T. Storni, u. Sauer, J. Bacteriol., **2001**, 183, 7308/7317
- [42] M. Dauner, U. Sauer, Biotechn. Bioeng.,**2001**, 76, 132/143
- [43] N. Zamboni, E. Fischer, A. Muffler, M. Wyss, H.-P. Hohmann,U. Sauer,
Biotechn. Bioeng., **2004**,89,219232
- [44] N. Zamboni, E. Fischer, D. Laudert, S. Aymerich, H.-P. Hohmann,
U. Sauer, J. Bacteriol., **2004**, 186,4528/4534
- [45] C. Wittmann, M. Hans, W. Bluemke, Yeast, **2002**, 19, 1351/1363
- [46] S. Sahoo, R.K. Verma,A.K. Suresh, K.K. Roa, J.Bellare, G.K. Suraishkumar,
Biotechn. Prog., **2003**, 19, 1689/1696
- [47] L.H. Grimm, S. Kelly, J. Hengstler, A. Göbel, R. Krull, D.C. Hempel, Biotechn.
Bioeng., **2004**, 87, 213/218
- [48] L.H. Grimm, S.Kelly, R.Krull, D.C. Hempel, Biotechn. Bioeng. (**2005**), im Druck
- [49] S. Kelly, L.H. Grimm, J. Hengstler, E. Schultheis, R. Krull, D.C. Hempel,
Bioprocess Biosyst. Eng., **2004**, 26, 315/323
- [50] U. Rinas, H. El-Enshasy, M. Emmeler, A. Hille, D.C. Hempel, H. Horn, Chem.
Eng. Sci., **2005**, 60,2729/2739
- [51] A. Hille, T.R. Neu, H.Horn, D.C.Hempel, Biotechn. Bioeng. (**2005**), im Druck
- [52] A.Jungebloud, K.Bohle, Y.Goecke, M.Haesner, C.Corde, H.Horn, D.C.Hempel,
Analytical Chemistry (**2005**), eingereicht

Tabelle 1: Ursprung der metabolischen Intermediate des zentralen C-Metabolismus berechnet aus der Isotopomerenverteilung der proteinogenen Aminosäuren.

	Anteil des Gesamtpools [%]				
	<i>B. meg.</i> MS941		<i>B. meg.</i> WH320	<i>B. subtilis</i> ^a	<i>B. clausii</i> ^b
<i>D</i> , h ⁻¹	0.1	0.4	0.1	0.11	0.09
P5P from glucose	56 ± 4	58 ± 6	70 ± 1	<i>n. a.</i>	<i>n. a.</i>
P5P from E4P (TK/TA reactions)	11 ± 5	21 ± 7	6 ± 1	<i>n. a.</i>	<i>n. a.</i>
P5P from trioses (TK reaction)	39 ± 2	29 ± 3	24 ± 1	<i>n. a.</i>	<i>n. a.</i>
E4P from F6P	70 ± 8	58 ± 3	63 ± 3	<i>n. a.</i>	<i>n. a.</i>
Trioses from glycolysis	60 ± 6	78 ± 4	78 ± 2	77	81
PEP from OAA	0	0	0	14	0
PYR from MAL	26 ± 1	23 ± 4	8 ± 7	0	21
OAA from PYR	40 ± 8	62 ± 3	16 ± 1	32	19

^a Dauner et al. [27], ^b Christiansen et al. [28] ; *n. a.* : nicht angegeben

Tabelle 2: Beispiele für Untersuchungen des Fluxoms (und anderer „Omics“-Daten) mit dem Ziel der Produktionssteigerung von extrazellulären Metaboliten.

Organismus	Ziel	Methodik	Referenzen
<i>Klebsiella pneumoniae</i>	1,3-Propandiol	Genomics, Proteomics, Verfahrenstechnik	[30-33]
<i>Xanthomonas campestris</i>	Xanthan	Verfahrenstechnik, stöchiometrische Bilanzierung	[34]
<i>Corynebacterium glutamicum</i>	Lysin	Genomics, Fluxomics, Verfahrenstechnik	[35-40]
<i>Bacillus subtilis</i>	Riboflavin	Genomics, Transkriptomics, Fluxomics	[41-44]
<i>Kluyveromyces marxianus</i>	2-Phenylethanol	Fluxomics	[45]

Liste der Abbildungsunterschriften

- Abb. 1: Molekulare Grundstruktur biologischer Prozesse. Die in der DNA gespeicherte Information wird in Botschaftermoleküle umgesetzt, die an den Ribosomen Proteine herstellen, die als Katalysatoren für einen (oder mehrere) Teilschritte des metabolischen Netzwerks wirken. Die gesamte DNA-Information wird als Genom bezeichnet und umfasst bei Mikroorganismen typischerweise 5000 identifizierte Gene, das Proteom enthält ca. 4000 Proteine und bisher konnten ca. 2000 intrazelluläre Metabolite nachgewiesen werden.
- Abb. 2: Regulation der Quecksilberresistenz (vereinfacht). Die Gene (*T*, *P*, *A*, *B*) des *mer*-Operons stehen unter Kontrolle des Promoters (*P*), der bei Abwesenheit von Hg(II) durch das Regulatorprotein R (Genprodukt des Gens *R*) blockiert ist. Tritt Hg(II) auf, wird R gebunden und die Transkription der Gene *TPAB* setzt ein (Induktion), was zur Umsetzung von toxischen Hg(II)-Verbindungen in metallisches flüchtiges Hg führt [2].
- Abb. 3: Grundstruktur der L-Tryptophan-Biosynthese. Die Gene *EDCBA* des *trp*-Operons werden transkribiert und translatiert. Die resultierenden Enzyme synthetisieren L-Tryptophan aus Chlorisamat. Bei Akkumulation von L-Tryptophan in der Zelle setzt neben Enzyminhibition auch ein Stop der Biosynthese auf Transkriptionsebene durch Feedback-Repression ein, indem der aus Tryptophan und Aporepressor gebildete Repressor den Operator belegt.
- Abb. 4: Metabolitbilanzen über die mikrobielle Zelle. In einfachen Fällen, wie z. B. die L-Tryptophan-Biosynthese, lassen sich die Bilanzen der Metabolite *j*, die an *i* Stoffwechselreaktionen beteiligt sind, unter Berücksichtigung kinetischer Ansätze für Transkription, Translation, Inhibition, Repression, Wachstum und Zerfall (von Enzymen und mRNA) formulieren und eine qualitative Beschreibung erzielen [4-6].
- Abb. 5: Dynamisches Verhalten der L-Tryptophan-Biosynthese. Unter bestimmten Bedingungen treten phasenverschobene andauernde Oszillationen der intrazellulären Konzentrationen von mRNA, der Syntheseeenzyme und des Tryptophans auf [6].
- Abb. 6: Gelelektrophoretische (2-D IEF/SDS-PAGE) Auftrennung der intrazellulären Proteine von *B. megaterium* MS941. Die durch Massenspektrometrie (ESI-QqTOF MS/MS) der Spotextrakte identifizierten Proteine sind gekennzeichnet durch die übliche Nomenklatur.
- Abb. 7: Rekonstruktion des metabolischen Netzwerks von *B. megaterium* MS941 aus Proteomdaten (Abb. 6). Zentraler C-Metabolismus und Biosynthese der Aminosäure (kursiv: noch nicht aus dem Proteom identifizierte Proteine)

- Abb. 8: Rekonstruktion des metabolischen Netzwerks auf Basis der Reaktionsdatenbank. Zur Aufstellung der Metabolitverbindungsmatrix wird die Enzymreaktionsmatrix durch nichtenzymatische Reaktionen und Reservibilitätsbeziehungen ergänzt. Der inverse Abfrageprozess nach Enzymen bzw. Reaktionen ermöglicht mit IdentiCS [15] die schnelle Rekonstruktion potentieller metabolischer Netzwerke bei geringer genomischer Abdeckung
- Abb. 9: Modulare Struktur des potentiellen metabolischen Netzwerks von *A. niger*. Das Netzwerk wurde unter Verwendung von Identics erstellt und modular strukturiert. Insgesamt wurden 1172 Metabolite, die an 434 enzymatischen Reaktionen teilnehmen, berücksichtigt. Man erkennt aus der modularen Darstellung, dass das innere Reaktionsnetz für die katabolischen Prozesse dicht gepackt ist und zahlreiche Verzweigungen bzw. Verknüpfungen aufweist, die bei den aussen angeordneten anabolen Reaktionen für die Biosynthese der Aminosäuren und Purine weitgehend fehlen.
- Abb. 10: Hierarchische und modulare Struktur des transkriptionalen regulatorischen Netzwerks von *E.coli* erweitert von Ma et al. [19] auf Basis von Shen-Orr et al. [17]. Die hierarchische Struktur erstreckt sich über 9 Ebenen nach unten und ein Feedback wird nicht beobachtet.
- Abb. 11: Generell ist die mikrobielle Zelle einer Vielzahl von Einflussgrößen (Environom) ausgesetzt, am besten untersucht ist Nahrungsstress.
- Abb. 12: Experimentelle Anordnung zur Ermittlung der metabolischen Flussverteilung durch partielle Zufütterung mit ^{13}C markierter Glukose im kontinuierlichen Bioreaktor. Durch Analyse des $^{13}\text{C} / ^{12}\text{C}$ -Verhältnis der proteinogenen Aminosäuren der aufgeschlossenen Biomasse kann auf die relative Beteiligung der zentralen Stoffwechselwege rückgeschlossen werden.
- Abb. 13: Massenverteilung der C3-Intermediate an der Reaktionsfläche zwischen Glykolyse und TCA-Zyklus, die teilweise in anaplerotische und glykoneogenetische Reaktionen involviert sind. Die Daten wurden aus kontinuierlichen Kulturen unter identischen Bedingungen ermittelt ($\mu = 0.1 \text{ h}^{-1}$, 10% $[1-^{13}\text{C}]$ -Glukose).
- Abb. 14: Metabolische Flussverteilung eines rekombinanten Stammes von megaterium MS 941 erhalten mit Anpassungsroutine an die experimentellen Daten. Die Daten an den Reaktionspfeilen sindgeschätzte molare Nettoflüsse normalisiert auf die Glukose-aufnahmeraten bei verschiedenen Wachstumsraten von $D = 0,1$ (oben) und $0,4 \text{ h}^{-1}$ (unten). Alle Intermediate, die als Vorläufermoleküle zur Biomassebildung benötigt werden, sind durch einen starken schwarzen Pfeil gekennzeichnet.

- Abb. 15: Wachstum von *B. subtilis* bei verschiedenen Scherraten unter sonst identischen physiologischen Bedingungen [46]. Bei hohen Scherraten bilden sich kleinere Zellen mit erhöhter metabolischer Aktivität.
- Abb. 16: Sporenaggregation und Keimung bei *A. niger*. Bevor Wachstum der fadenförmigen Gebilde oder Pellets einsetzt, treten deutliche Aggregationsphasen auf [47].
- Abb. 17: Der Leistungseintrag beeinflusst entscheidend Pelletdichte und -durchmesser sowie die Produktbildung bezogen auf Biomasse bzw. Pelletvolumen [49].
- Abb.18: Zusammenwirken von Molekularbiologie, Bioinformatik und Verfahrenstechnik für die prozessorientierte Systembiotechnologie.