

Mathematische Zellmodelle – Fiktion oder greifbare Wirklichkeit ?

Autor:

Prof. Dr. Hermann-Georg Holzhütter
Humboldt-Universität Berlin
Medizinische Fakultät (Charité)
Institut für Biochemie
Leiter der Arbeitsgruppe „Mathematische Systembiologie“

Monbijoustr. 2
D-10117 Berlin
Tel.: (030) 450 528 166
Fax: (030) 450 528 942
mail: hergo@charite.de

„Keinerlei Glaubwürdigkeit ist in jenen Wissenschaften, die sich der mathematischen Wissenschaften nicht bedienen oder keine Verbindung zu ihnen haben“ (Leonardo da Vinci).

Systembiologie: Quid novi?

Systembiologie lautet das allerorts zu vernehmende neue Schlagwort, mit dem eine grundlegend neue Vorgehensweise in der biochemischen und molekularbiologischen Forschung apostrophiert werden soll. Das Ziel besteht darin, komplexe Lebensvorgänge wie etwa die Differenzierung von Stammzellen zu spezialisierten organotypischen Zellen nicht allein zu beschreiben, sondern durch Zurückführung auf die ihnen zugrunde liegenden Prozesse auf molekularer Ebene kausal zu verstehen und – als höchste Stufe dieses Verstehens – durch Computersimulationen nachzubilden. Diese Zielstellung ist nicht neu. Die Erklärung der seit Jahrtausenden bekannten und genutzten Bildung von Ethanol durch Hefekulturen als letzter Reaktionsschritt eines vielstufigen Abbauweges (Glykolyse) von Glukose, der bereits vor etwa 70 Jahren durch Forschungsarbeiten von *Emden* und *Meyerhoff* vollständig rekonstruiert wurde, ist nicht anderes als Systembiologie, odergenauer gesat: Systembiochemie. Und die mathematische Modellierung dieses Stoffwechselweges durch grundlegende Arbeiten von Heinrich und Rapoport [1] im Zeitraum 1970 – 1985 spiegelt sehr genau den Inhalt mathematischer Systembiologie wider, nämlich die Computersimulation eines komplexen biochemischen Reaktionssystems, die es ermöglicht, den Einfluss eines einzelnen Elementarprozesses auf das Verhalten des gesamten Systems quantitativ abzuschätzen. Neu ist gegenwärtig allerdings die Dimension der Zielstellungen in der mathematischen Systembiologie. Es geht nicht mehr allein um die Erforschung der molekularen Grundlagen einzelner

Reaktionswege einschließlich ihrer mathematischen Modellierung, sondern um die mathematische Modellierung kompletter Zellen und – als ultimatives Fernziel – kompletter Organismen. Entscheidender Impetus für diese ehrgeizige Zielstellung ist die Tatsache, daß es mit Hilfe moderner biochemischer Hochdurchsatzmethoden (RNA-, DNA- und Protein-Microarrays; 2-D Gel-Elektrophorese von Proteingemischen; Metabolitanalytik mittels Kapillar-HPLC und GC/MS) möglich geworden ist, eine fast vollständige Fotografie des molekularen Zustandes einer Zelle anzufertigen, d.h. gleichzeitig die Konzentration von Stoffwechselprodukten, Proteinen und mRNA-Molekülen zu bestimmen. Durch zeitliches Aneinanderreihen dieser „molekularen Fotografien“ entsteht ein Film von der zellulären Dynamik auf molekularer Ebene. Die große Herausforderung besteht nun darin, diese Dynamik auf molekularer Ebene vorhersagbar zu machen. Um beispielsweise abschätzen zu können, wie sich die pharmakologisch herbeigeführte Hemmung einer bestimmten Cyclin-abhängigen Kinase auf die Länge des G1-Phase des Zellzyklus auswirkt, ohne die entsprechend aufwendigen Experiment durchzuführen zu müssen, bedarf es eines mathematischen Modells, das sowohl gen-regulatorische, signalübertragende und metabolische Prozesse umfasst.

Die Hoffnung, in Bälde über derartige mathematische Ganzzell-Modelle verfügen zu können, gründet sich auf zwei Tatsachen. Erstens sind auf Grund der erwähnten Fortschritte auf dem Gebiet der Molekülanalytik gute Voraussetzungen gegeben, den notwendigen Zyklus der Modellvalidierung (Modell) → (Experiment) → (verändertes Modell) rascher und vollständiger zu durchlaufen, als etwa noch vor 10 Jahren. Zweitens ist zu konstatieren, dass die technischen und personellen Voraussetzungen für die Aufstellung und Simulation von komplexen mathematischen Modellen in den letzten Jahren sehr viel besser geworden sind. Es sind weltweit beträchtliche Anstrengungen unternommen worden, um im akademischen und im kommerziellen Sektor neue Unternehmen und Verbände auf dem Gebiet der mathematischen Systembiologie zu gründen. Dazu gehören beispielsweise das von der International Union of Physiological Sciences getragene „Physiome“-Projekt (<http://www.physiome.org/>), das an der Keio-Universität Japan mit mehr als 100 Millionen Euro geförderte „E-Cell“-Projekt (<http://www.e-cell.org/>), das in Seattle (USA) ansässige „Institute for Systems Biology“ (<http://www.systemsbiology.org/>) oder das niederländische

„Yeast Silicon Cell“ – Projekt (<http://www.bio.vu.nl/hwconf/Silicon/>). Auch in Deutschland gibt es seit Beginn dieses Jahres ein vom BMBF gefördertes Verbundprojekt „Systembiologie“, das auf die Untersuchung der molekularen Grundlagen ausgewählter Funktionen von Leberzellen ausgerichtet ist. Eine wesentliche Aktivität der neu geschaffenen Einrichtungen war und ist die Schaffung von Programmierungs-Plattformen für die Modellierung. Davon gibt es inzwischen mehr als 30 verschiedene mehr oder weniger komfortable Varianten. Neuartige, auf die Bedürfnisse der Systembiologie zugeschnittene Sprachen wie z.B. SMBL (Systems Biology Markup Language) wurden geschaffen, um die Struktur und Funktion biologischer Systeme in formaler Weise zu repräsentieren und damit den Transfer von mathematischen Modellen zwischen verschiedenen Software-Paketen wie *BioSpice*, *DBSolve*, *E-Cell*, *Gepasi*, *Jarnac*, *StochSim* und *Virtual Cell* zu erleichtern. Was nun eigentlich nur noch fehlt, sind die oftmals vollmundig angekündigten Modelle kompletter Zellen oder Organismen. Die Situation erinnert etwas an einen Reiter, der sich mit Sattel, Stiefeln und Reitanzug bestens ausgerüstet hat, dem aber noch das Pferd fehlt.

In den zurückliegenden drei Jahrzehnten sind zahlreiche mathematische Modelle kleinerer biochemischer Reaktionssysteme, vornehmlich auf dem Gebiet des Zwischenstoffwechsels, entstanden, von denen einige im Modell-Repositoryum JWS-Online (<http://jjj.vbi.vt.edu/>) zur eigenen Benutzung verfügbar sind. Die in der Literatur bereits vorhandenen mathematischen Modelle decken aber nur einen geringen Teil des gesamten zellulären Reaktionsnetzwerkes ab. Der größte Teil des zellulären Netzwerkes höherer eukaryotischer Zellen wird nach dem jetzigen Kenntnisstand von signalübertragenden und gen-regulatorischen Wegen gebildet, zu denen bisher kaum Modelle (siehe z.B. [2,3]) existieren. Das einzige bislang vorzeigbare Ganzzellmodell existiert für die reife rote Blutzelle, den Erythrozyten [4]. Diese Zelle besitzt keine Organellen mehr und betreibt dementsprechend nur noch einen sehr eingeschränkten Metabolismus. Für eine „richtige“ Zelle ist ein vergleichbares kinetisches Modell nicht verfügbar. Wird sich das in den kommenden Jahren ändern? Hier ist aus verschiedenen Gründen Realismus und von wissenschaftlichem Ernst getragene Bescheidenheit angesagt.

Mathematische Zellmodelle: Was ist zur Zeit wirklich möglich?

Die Zelle ist eine räumliche Struktur und alle in der Zelle stattfindenden molekularen Prozesse können streng genommen nur im Kontext der subzellulären Kompartimente, in und an denen sie ablaufen, verstanden werden. Subzelluläre makromolekulare Strukturen wie Membranen oder Filamente sind einem ständigen Umbau und räumlichen Bewegungen unterworfen. Die Mechanismen ihrer Assemblierung und die Dynamik ihrer räumlichen Verteilung - Prozesse die praktisch nur an einzelnen Zellen untersucht werden können – sind bisher experimentell schlecht charakterisiert. Dementsprechend sind mathematische Ansätze, die die Selbstorganisation räumlicher Molekülstrukturen beschreiben, gegenwärtig äußerst rudimentär. Mathematische Zellmodelle im strengen Sinne, mit deren Hilfe sich die Menge und räumliche Position eines beliebigen Moleküls bei bekannten äußeren Bedingungen berechnen lässt, sind schon aus diesem Grunde für die nächsten Jahrzehnte kein realistisches Ziel.

Zellen und Organismen sind dynamische Systeme. Während Daten zur Identifizierung und Quantifizierung ihrer Variablen, den Molekülen, die einschlägigen Datenbanken zum Überlaufen bringen, wächst unser Wissen über die zugrunde liegenden Prozesse ungleich langsamer. Die an der Signalübertragung und der Genexpression beteiligten enzymatischen Prozesse sind bisher kaum charakterisiert. Kritisch anzumerken ist hier die rückläufige Zahl von Publikationen, die sich mit der Isolierung und Charakterisierung von Enzymen befassen (Abb. 1).

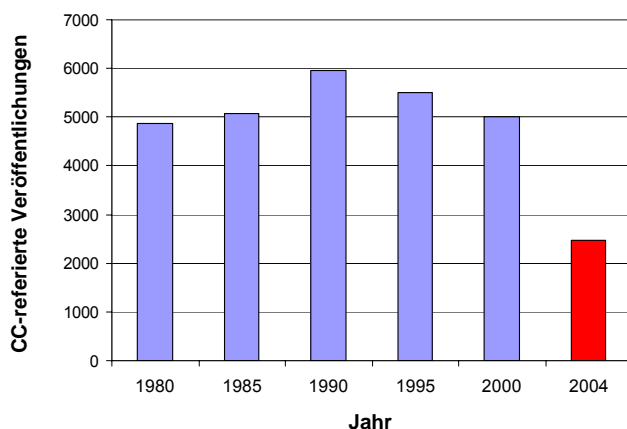


Abb. 1 Zahl von wissenschaftlichen Veröffentlichungen, die in den C(urrent) C(ontents) referiert sind und als Schlüsselworte „Enzym(e)“ und „Kinetik“ enthalten

Solche Arbeiten versprechen im Augenblick wenig Ruhm und Ehr'. Daraus resultiert eine große Diskrepanz zwischen der Vielzahl von gemessenen Molekülmustern und wenig zuverlässige Information über die zugrunde liegenden Prozesse auf signal-übertragender und gen-regulatorischer Ebene.

Also sind alternative mathematische Ansätze gefragt, mit denen sich vorhandene Informationen integrieren, zumindest empirische Regeln ableiten und qualitativen Vorhersagen des Systemverhaltens machen lassen. Hier sind gegenwärtig drei unterschiedliche mathematische Ansätze zu nennen.

Zum einen erfreut sich die Statistik, die jetzt im Gewand der Bioinformatik erscheint, einer großen Renaissance bei der Aufdeckung von Zusammenhängen in Hochdurchsatzdaten. Statistische Modelle ersetzen die tatsächlichen, durch biochemische Prozesse definierten Relationen zwischen den Molekülen durch einfache, meistens lineare Pseudo-Relationen. In diesem Sinne lassen sich abstrakte „genetische Netzwerke“ [5] oder „Protein-Netzwerke“ definieren, zwischen deren Elementen nach statistisch signifikanten Zusammenhängen gesucht wird. Eine Feststellung wie „Gen X und Gen Y werden gegenläufig exprimiert“ lässt natürlich völlig offen, ob das immer so sein muß und ob dieser Zusammenhang zufälliger Natur ist oder durch einen Regelkreis zustande kommt, dem die Genprodukte von Gen X und Gen Y angehören. Immerhin aber lassen sich gewisse Muster in den ‚molekularen Fotografien‘ erkennen und empirische Regeln für ihr Auftreten ableiten.

Ein zweiter mathematischer Ansatz, der oft als „strukturelle Modellierung“ bezeichnet wird [6,7], untersucht in systematischer Weise alle Wege, die ein gegebener Stoff durch aufeinander folgende chemische Umwandlungen durchlaufen kann. Für die Anwendung dieser Methode müssen lediglich die Stöchiometrien bekannt sein, in denen sich die beteiligten Moleküle ineinander umwandeln. Die Aufgabenstellung entspricht etwa der Bestimmung aller Routen zwischen zwei Orten in einer Straßenkarte. Auf diese Weise kann geprüft werden, ob ein gegebenes Netzwerk überhaupt funktionsfähig ist und welche Pfade vom Ausfall eines Prozesses, z.B. im Falle eines Gendefekts, betroffen sind [8].

Ein dritter alternativer mathematischer Ansatz basiert auf der Anwendung von Optimierungsmethoden. Die Grundannahme besteht hier darin, daß alle molekularen Interaktionen vermöge einer Vielzahl von Regulationsmechanismen so eingerichtet sind, dass Zellen und Organismen ihre aktuellen Funktionen in optimaler Weise ausführen können. Mit einer solchen Prämisse genügt bereits die Kenntnis der Stöchiometrien, mit denen sich die beteiligten Moleküle ineinander umwandeln, um die für das Erfüllen einer gegebenen zellulären Leistung „optimalen“ chemische Umsatzraten zu berechnen [9]. Auf diese Weise konnte z.B. der Effekt von Enzymdeletionen auf die Wachstumsrate von *E. coli*

quantitativ gut vorhergesagt werden [10]. Um dieses recht aussichtsreiche mathematischen Konzept auch auf komplexe eukaryotische Zellen anwenden zu können, ist das Prinzip der Flußminimierung vorgeschlagen worden [11], wonach der Stoffwechsel so reguliert wird, daß die gegebenen Leistungen der Zelle mit einem minimalen Aufwand an externen Ressourcen erfüllt werden. Solche Optimierungsansätze tragen natürlich stark hypothetischen Charakter und sind deshalb in erster Linie dazu gedacht, Heuristiken für weitere zielgerichtete Experimente zu liefern.

Zusammenfassend kann eingeschätzt werden, daß ein mechanistisches, die Kinetik der zugrunde liegenden molekularen Prozesse berücksichtigendes mathematisches Modell von Zellen oder gar Organismen in den nächsten Jahren nicht realisierbar erscheint. Analog dem Vorgehen bei der mathematischen Modellierung von Stoffwechselsystemen wird man sich mit der mathematischen Modellierung kleiner, relativ autonomer Teilnetzwerke, z.B. einzelner Signalwege, begnügen müssen. Solche Arbeiten, die eine enge Zusammenarbeit von Modellierern und experimentellen Forschern erfordern, sind allerdings von großem Wert, um Hypothesen über die molekulare Architektur von Reaktionswegen zu testen. Daneben gibt es eine Reihe von alternativen mathematischen Ansätzen, die grundlegende Zusammenhänge in komplexen Netzwerken sichtbar machen und damit wichtige Impulse für weitere zielgerichtete experimentelle Untersuchungen liefern können.

Literatur

- [1] Rapoport, T.A. & Heinrich, R. (1975) Mathematical analysis of multienzyme systems. I. Modelling of the glycolysis of human erythrocytes. *Biosystems* 7(1), 120-9
- [2] Swameye, I., Müller, T.G., Timmer, J., Sandra, O. & Klingmüller (2003) Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modelling. *Proc Natl Acad Sci U S A* 100(3), 1028-33
- [3] Lee, E., Salic, A., Krüger, R., Heinrich, R. & Kirschner, M.W. (2003) The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt Pathway. *PLoS Biol* 1(1), E10
- [4] Jamshidi, N., Edwards, J.S., Fahland, T., Church, G.M. & Palsson, B.O. (2001) Dynamic simulation of the human red blood cell metabolic network. *Bioinformatics* 17(3), 286-7
- [5] Lee, I., Date, S.V., Adai, A.T. & Marcotte, E.M. (2004) A probabilistic functional network of yeast genes. *Science* 306, 1555-1558

- [6] Schuster, S., Dandekar, T. & Fell, D. A. (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering: *Trends Biotechnol* 17, 53-60
- [7] Schilling, C. H., Letscher, D. & Palsson, B. O. (2000) Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective, *J Theor Biol* 203, 229-48
- [8] Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S. & Gilles, E. D. (2002) Metabolic network structure determines key aspects of functionality and regulation, *Nature* 420, 190-193
- [9] Cornish-Bowden, A. & Cardenas, M. L. (2002) Metabolic balance sheets, *Nature*. 420, 129-30
- [10] Fong, S. & Palsson, B.O. (2004) Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. *Nature Genetics* 36, 1056 - 1058
- [11] Holzhütter, H.G. (2004) The Principle of Flux Minimization and its Application to Estimate Stationary Fluxes in Metabolic Networks. *Eur. J. Biochem.* 271, 2905–2922