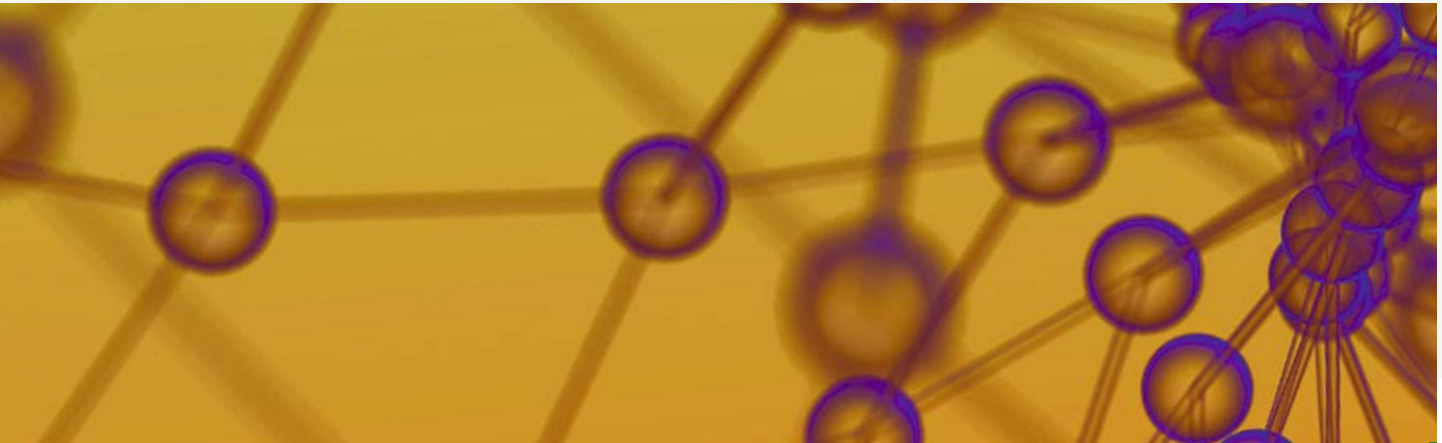


VICTOR SOURJIK

MAX-PLANCK-INSTITUT FÜR TERRESTRICHE MIKROBIOLOGIE, MARBURG

Regeln für die Zusammenarbeit in der Zelle



Proteine agieren in der Zelle niemals allein. Selbst die einfachsten Zellfunktionen, etwa der Transport eines Moleküls durch die Zellmembran oder die Identifikation einer künftigen Zellteilungsstelle, werden normalerweise von Gruppen interagierender Proteine ausgeführt. Diese Gruppen können am besten als Proteinnetzwerke dargestellt werden, wobei die Netzknoten einzelne Proteine und die Verbindungen zwischen den Knoten deren Interaktionen repräsentieren. Je komplexer die Aufgabe ist, umso größer und komplexer ist das zugrunde liegende Netzwerk. Letztlich können sogar alle funktionalen Netze zu einem zellenweiten Netzwerk verbunden werden. Die nächsten großen Herausforderungen in der Biologie bestehen darin, die Regeln zu verstehen, nach denen diese Netzwerke im Kontext einer lebenden Zelle funktionieren, und diese Regeln für ein rationales Design der Netzwerkfunktionen zu nutzen.

VON DER MOLEKULARBIOLOGIE ZUR SYSTEMBIOLOGIE

Eine netzwerkorientierte Sicht auf die Zellbiologie hat in den vergangenen Jahren zunehmend an Akzeptanz gewonnen. Dennoch stellt diese Perspektive eine drastische Paradigmenverschiebung dar, die es noch anzuerkennen gilt. Seit es die Molekularbiologie und die Biochemie gibt, haben sich diese Forschungsbereiche in erster Linie auf die Analyse einzelner Gene oder Proteine konzentriert und dabei einen Ansatz verfolgt, der auf die Zerlegung zellulärer Prozesse in Elementarfunktionen abzielte, die einem einzelnen Gen/Protein zugeschrieben werden können. In den Frühphasen der Molekularbiologie hat sich dieser reduktionistische Ansatz als enorm leistungsstark und einflussreich erwiesen. Er ermöglichte die Entmystifizierung lebender Materie und stellte die gesamte moderne Biologie auf eine solide molekulare

und mechanistische Grundlage. Mit zunehmendem Wissen über zelluläre Systeme wurde jedoch immer deutlicher, dass dieser Blick die Komplexität der zellulären Mechanismen nicht angemessen beschreiben konnte. In der Tat lassen sich viele zelluläre Funktionen nicht einfach auf die Funktionen einzelner Proteine reduzieren, sondern ergeben sich erst aus dem Zusammenspiel mehrerer Komponenten. Das Gesamtsystem ist also mehr als die Summe seiner einzelnen Teile. In der Folge hat sich der Fokus in der biologischen Forschung allmählich von einzelnen Proteinen auf Proteingruppen oder -komplexe verschoben, die als Funktionseinheiten betrachtet werden.

Ein weiterer Grund für diese Verschiebung des Fokus' war das zunehmende Wissen über das hohe Ausmaß an Verbindungen zwischen Proteinen in einer Zelle, welche durch molekulare Untersuchungen zellulärer Prozesse entdeckt wurden. Insbesondere Hochdurchsatz-Techniken („-omik“-Technologien), die hochgradig parallelisierte experimentelle Analysen von Hunderten oder Tausenden von Interaktionen zwischen verschiedenen Proteinen ermöglichen, haben hochkomplexe Interaktionsnetze generiert, die oft Proteine bzw. Einheiten sehr verschiedener Funktionen miteinander verbinden. Die Interpretation der Bedeutung dieser Netzwerke über einzelne Interaktionen hinaus stellte eine enorme Herausforderung dar, die Ansätze erforderlich machte, die sich sehr von denen der konventionellen Molekularbiologie unterscheiden. All diese Entwicklungen führten zur Entstehung der Systembiologie – die man auch als Netzwerkbiologie bezeichnen kann – als einem unabhängigen Zweig der modernen biologischen Forschung.

MATHEMATISCHE MODELLIERUNG IN DER SYSTEMBIOLOGIE

Wie jede neue Disziplin ist die Systembiologie bislang nicht streng definiert worden. In der einfachsten Definition kann jede Analyse, welche alle (oder zumindest mehrere) Gene oder Proteine berücksichtigt, die eine Rolle in einem zellulären Prozess spielen, als Systembiologie bezeichnet werden. In diesem Sinne ist die Systembiologie inzwischen ein üblicher Bestandteil der biologischen Forschung, die sich zunehmend auf neuartige „-omik“-Tools stützt. Mit diesen Techniken können jedoch nur deskriptive Darstellungen zellulärer Netzwerke erzeugt werden, die begrenzte Einsichten in deren Funktionen vermitteln. Beispielsweise erfolgt die Verbreitung eines Signals über ein Netzwerk, das Rezeptoren auf der Oberfläche der Zelle (Eingabe/ Input) mit der Genregulation oder dem Zellverhalten (Ausgabe/ Output) verbindet, häufig über mehrere Schritte und Verzweigungen. Zusätzlich kann diese Signalübertragung durch mehrere positive und negative Rückkopplungs (Feedback)-Schleifen moduliert werden. Obwohl man die Bedeutung einiger dieser regulierenden Verbindungen intuitiv ableiten kann, ist eine Vorhersage darüber, wie sich ein solches Netzwerk bei einer Stimulierung verhält, durch einfache Betrachtung der Netzwerkstruktur nicht oder nur begrenzt möglich.

Um diese Probleme zu lösen, sind mathematische Modelle und Computersimulationen biologischer Netzwerke besonders hilfreich. Während die mathematische Beschreibung natürlicher Prozesse ein grundlegender Bestandteil der Physik ist, ist die Anwendung mathematischer Modelle in der Biologie nach wie vor begrenzt und zum Teil umstritten. Ein Hauptproblem ist die offensichtliche Einzigartigkeit jedes einzelnen biologischen Netzwerkes, was impliziert, dass das aus einem Netzwerk erhaltene Wissen nicht ohne Weiteres auf ein anderes Netzwerk übertragen werden kann. Dies steht in starkem Kontrast zur Physik, wo Systeme gleicher Art mit demselben mathematischen Modell beschrieben werden können. Obwohl die eigentliche Modellierung eines zellulären Netzwerkes häufig relativ einfach mit einer Reihe gekoppelter gewöhnlicher Differenzialgleichungen erfolgen kann, sind in der Regel umfassende Kenntnisse über einzelne Reaktionsparameter innerhalb des Netzwerkes erforderlich, die über langwierige biochemische Experimente erworben werden müssen. Deshalb ist es unrealistisch, solche Informationsmengen für jedes einzelne zelluläre Netzwerk zu erhalten. Dies bedeutet, dass eine detaillierte Modellierung nur für eine relativ kleine Gruppe ausgewählter Netzwerke durchgeführt werden kann. Wenn auch alternative Ansätze der Modellierung möglich sind,

die nur die Struktur der Netzwerke verwenden und kein Wissen über detaillierte Parameter erfordern, so ist die Präzision dieser Modelltypen und ihr Vermögen, die Netzwerkreaktion auf bestimmte Stimuli oder Störungen zuverlässig vorauszusagen, sehr begrenzt.

Angesichts dieser Schwierigkeiten ist nicht zu erwarten, dass in naher Zukunft Computersimulationen biologische Experimente in dem Maße ersetzen können, wie dies in der Physik häufig der Fall ist. Dennoch sind mathematische Modelle zu zellulären Prozessen aus zwei Gründen sehr sinnvoll: Erstens können sie zeigen, ob das beobachtete Verhalten eines Netzwerkes mit dem aktuellen Wissen über die Struktur dieses Netzwerkes prinzipiell erklärbar ist. Überraschenderweise ist dies meistens nicht der Fall, wodurch deutlich wird, dass wesentliche Komponenten der Netzwerkstruktur in den meisten Fällen noch unbekannt sind. Häufig können mathematische Modelle auch auf eine wahrscheinlich fehlende Verbindung hinweisen, wie beispielsweise ein negatives Feedback innerhalb eines Netzwerkes. Der zweite Aspekt mathematischer Modelle ist, dass sie zur Entdeckung allgemeiner Eigenschaften in zellulären Netzwerken ein unerlässliches Werkzeug darstellen.

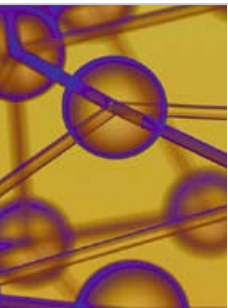
HÄUFIG KÖNNEN MATHEMATISCHE MODELLE AUCH AUF EINE WAHRSCHEINLICH FEHLLENDE VERBINDUNG HINWEISEN, WIE BEISPIELSGEWISSE EIN NEGATIVES FEEDBACK INNERHALB EINES NETZWERKES.

EIGENSCHAFTEN VON ZELLULÄREN NETZWERKEN

Gibt es angesichts der Individualität jedes einzelnen zellulären Netzwerkes eindeutige Hinweise auf die Existenz von Eigenschaften, die gleichermaßen in verschiedenen Netzwerken zu finden sind? Diese Frage ist noch nicht abschließend beantwortet, aber aus der Untersuchung unterschiedlicher Systeme haben sich mehrere Eigenschaften herauskristallisiert, die allgemeiner Natur sein könnten.

Eine solche Eigenschaft, die den meisten Netzwerken gemein zu sein scheint, ist die Robustheit. Die Bedeutung der Robustheit – der Unempfindlichkeit gegenüber Störungen – für zelluläre Netzwerke ist nicht überraschend, sind diese Netzwerke doch extensiven internen und externen

Störungen ausgesetzt. Interne Störungen sind vielen biochemischen Reaktionen innerhalb der Zelle innewohnend, weil die meisten Reaktionen nur unter Beteiligung einer relativ kleinen Kopienzahl von Komponenten erfolgen. Obwohl sich die Durchschnittswahrscheinlichkeit einer bestimmten Reaktion (z. B. Protein-Modifizierung) feststellen lässt, kann die Anzahl von Proteinen, die zu einem bestimmten Zeitpunkt und in einer speziellen Region der Zelle tatsächlich kollidieren und miteinander reagieren, erheblich fluktuieren – typischerweise in der Größenordnung einer Quadratwurzel der Kopienzahlen der beteiligten Komponenten. Die resultierenden Fluktuationen im Output einer einzigen speziellen Reaktion können sich über die Netzwerke verbreiten und grundlegend das Netzwerkverhalten beeinträchtigen. In vielen Fällen können solche schnellen stochastischen Fluktuationen in der Netzwerkaktivität durch das geringere Tempo der nachgelagerten Output-Reaktionen „herausgefiltert“ werden. Das ist aber nicht immer der Fall. Insbesondere die Fluktuationen in der Aktivität der Gentranskription können eine beträchtliche Variabilität in den Kopienzahlen der Netzwerkproteine über die Zeit oder zwischen den Zellen einer Population untereinander erzeugen. Eine solche Variation, die als Genexpressionsrauschen bezeichnet wird, hat normalerweise direkte und lang anhaltende Effekte auf die Netzwerkleistung.



BIOLOGISCHE NETZWERKE UNTERSCHIEDEN SICH NICHT WESENTLICH VON KOMPLEXEN, VON MENSCHEN GESCHAFFENEN SYSTEMEN, WIE STROMNETZEN ODER TRANSPORTNETZWERKEN.

Neben einem solchen inhärenten stochastischen Rauschen wird die Netzwerkleistung grundsätzlich auch durch Änderungen im zellulären oder extrazellulären Umfeld beeinflusst. So variieren die Netzwerkkomponentenmengen nicht nur stochastisch, sondern auch deterministisch aufgrund von Änderungen in der Genregulation. Darüber hinaus wirken sich externe Faktoren wie Temperaturveränderungen auf das Tempo der meisten biochemischen Reaktionen aus, wobei sich die Aktivität eines „durchschnittlichen“ Enzyms bei zehn Grad Temperaturerhöhung verdoppelt. Das bedeutet, dass jedes zelluläre Netzwerk in einem Organismus ohne eine Körpertemperaturkontrolle durch einen täglichen oder saisonalen Temperaturzyklus stark beeinflusst wird.

Aus diesen Gründen ist es sehr wahrscheinlich, dass die Evolution Netzwerke nicht nur hinsichtlich der Durchführung ihrer speziellen Funktion selektiert hat (z. B. Signalübertragung), sondern auch bezüglich einer robusten Durchführung dieser Funktion, unabhängig von internen und externen Störungen. In dieser Hinsicht unterscheiden sich biologische Netzwerke nicht wesentlich von komplexen, von Menschen geschaffenen Systemen, wie Stromnetzen oder Transportnetzwerken, wo die Robustheit des Systems ein wesentlicher Bestandteil jeder technischen Konstruktion ist. Während die Grundsätze einer robusten Konstruktion technischer Systeme bekannt sind, bleiben die Mechanismen, die der Robustheit biologischer Netzwerke zugrunde liegen, noch aufzuklären.

Eine andere grundsätzliche Eigenschaft zellulärer Netzwerke ist ihre Modularität. Modularität bedeutet, dass Netzwerke, die bestimmte spezifische zelluläre Aufgaben durchführen (z. B. Teilung von Chromosomen während der Zellteilung), relativ isoliert sind und autonom voneinander fungieren. Netzwerkmodularität schafft im Wesentlichen eine Ebene an Funktionalität innerhalb der Zelle (Module), welche zwischen dem Proteinnetzwerk der Gesamtzelle und dem einzelner Proteine liegt. Auch hier finden sich Parallelen zur Technik: Die meisten komplexen Maschinen, etwa PKWs oder Flugzeuge, werden aus funktionellen Modulen zusammengebaut, die unabhängig vom Rest der Maschine optimiert, produziert, kontrolliert und ausgetauscht werden können. Die Vorteile eines modularen Designs biologischer Netzwerke sind wohl vergleichbar. Die Evolution optimiert die Funktionalität einzelner Module und verschaltet sie ohne große Auswirkungen auf den Rest des Netzwerkes neu oder tauscht sie aus. Modularität ermöglicht zudem die Feinabstimmung einzelner Funktionen und verbessert die Netzwerk-Robustheit, weil Störungen in einem Modul nicht automatisch auf den Rest des zellulären Netzwerkes übertragen werden.

VON DER SYSTEMBIOLOGIE ZUR SYNTHETISCHEN BIOLOGIE


Angesichts der Parallelen zwischen technischen Systemen und zellulären Netzwerken lag der natürliche nächste Schritt der Biotechnologie nahe, das rationale Design neuartiger zellulärer Netzwerke mit vorgegebenen Eigenschaften in Angriff zu nehmen. In einem gewissen Umfang erfolgen solche Arbeiten bereits seit Jahrzehnten in den Forschungsfeldern der Molekularbiologie, der Gentechnik und der Biotechnologie. Allerdings waren derartige Netzwerkmodifikationen typischerweise relativ einfach und betrafen lediglich den

Austausch oder die Einbringung eines oder sehr weniger Gene in einem Organismus. Die bisher erreichten Fortschritte in der Systembiologie versprechen die Biotechnologie zu revolutionieren, indem sie das Netzwerkdesign in Form der synthetischen Biologie auf ein qualitativ neues Niveau bringen. In ihrer Reinform, die zur Jahrhundertwende von ihren frühen Vertretern befürwortet wurde, sollte die synthetische Biologie eine echte technische Disziplin sein, in der das Netzwerkdesign unter Verwendung vollstandardisierter Teile mit Hilfe einer speziellen Software am Computer (*in silico*) erfolgt. Der Computer verifiziert dann die Funktionalität der konstruierten Netzwerke durch Verhaltenssimulationen – überwiegend vergleichbar mit den Vorgängen, die heute bei der Konstruktion eines Autos oder eines Flugzeuges zum Einsatz kommen. Bemerkenswerterweise hätte ein solches Design nicht notwendigerweise einen Bezug zu natürlich existierenden zellulären Netzwerken ähnlicher Funktion, sondern würde ein rein künstliches Netzwerk kreieren.

Schnell wurde jedoch deutlich, dass diese anfängliche Aussicht der synthetischen Biologie übermäßig optimistisch war. Ein solches rationales Design neuartiger Netzwerke muss eine ganze Reihe großer Herausforderungen überwinden. Es stellte sich heraus, dass die biologischen Bauteile extrem schwer zu standardisieren und von unerwünschten Einflüssen innerhalb der Zelle zu isolieren waren, die ihre Aktivitäten in unvorhergesehener Weise beeinträchtigten. Die *in silico* entwickelten Netzwerke waren deshalb in den meisten Fällen innerhalb der Zelle nicht funktional, weil die Proteinmengen oder -aktivitäten zu stark von den gewünschten Werten abwichen. Diese Probleme wurden mit einer zunehmend größeren Zahl von Netzwerkkomponenten immer offensichtlicher, da selbst kleinste Abweichungen vom gewünschten Verhalten innerhalb des Netzwerkes verstärkt wurden. Und selbst synthetische Netzwerke, die unter bestimmten Bedingungen funktionierten, wurden unter anderen Bedingungen funktionsunfähig, weil ihnen offensichtlich die benötigte Robustheit gegenüber Störungen fehlte. Folglich ist das rationale Design funktionaler synthetischer Netzwerke aktuell auf relativ einfache und kleine Systeme beschränkt, die nicht annähernd die Komplexität der meisten natürlichen zellulären Netzwerke aufweisen.

DIE REGELN ZELLULÄRER NETZWERKE

Diese offensichtliche Diskrepanz zwischen den natürlichen Netzwerken und unseren derzeitigen Möglichkeiten im künstlichen Netzwerkdesign zeigt uns, dass wir noch immer nicht verstehen, was ein funktionales Netzwerk ausmacht. Es ist zwar unstrittig, dass die Biologie nicht auf eigenen



ZELLULÄRE SYSTEME KANN MAN SICH ALS EINE ART SPRACHE VORSTELLEN – MIT EINZELNEN AMINOSÄUREN ALS BUCHSTABEN, PROTEINEN ALS WÖRTERN UND DER GESAMTEN ZELLE ALS TEXT.

– über die Gesetze der Physik hinausgehenden – Sondergesetzen beruht. Dennoch könnte es bestimmte Regeln geben, welche dem evolvierten Design zellulärer Netzwerke zugrunde liegen. Im Gegensatz zu den von Menschen hergestellten Maschinen, deren Konstruktionsregeln rational festgelegt werden, könnten die Regeln für das zelluläre Netzwerkdesign natürlicherweise als Produkt des evolutionären Selektionsprozesses entstehen. Der Prozess der natürlichen Netzwerk-Modifizierung ist zwar nicht gerichtet, sondern geschieht durch zufällige Modifikation bestehender Teile oder Zugabe neuer Teile und Module, aber die Evolution selektiert letztlich solche Netzwerke aus, die ihre Funktion unter physiologischen Beschränkungen, die die Netzwerkfunktion in der Zelle grundsätzlich beeinflussen, optimal erfüllen. So können beispielsweise Netzwerke danach ausgewählt werden, ob sie die Kriterien der Robustheit und der Modularität, wie oben erwähnt, erfüllen. Trotz der scheinbaren Zufälligkeit des Evolutionsprozesses kann die Auswahl dennoch zu bestimmten gemeinsamen Regeln darüber führen, wie ein evolutionär erfolgreiches Netzwerk strukturiert ist und funktioniert.

Um diese Thematik zu veranschaulichen kann man gewisse Parallelen zur menschlichen Sprache ziehen. Zelluläre Systeme kann man sich als eine Art Sprache vorstellen – mit einzelnen Aminosäuren als Buchstaben, Proteinen als Wörtern und der gesamten Zelle als Text. In dieser Struktur repräsentieren zelluläre Netzwerke oder Module, die bestimmte Funktionen ausführen, die Sätze. Die Sätze einer Sprache werden nach bestimmten Regeln konstruiert – wenn auch kein Naturgesetz zwingend dafür sorgt, dass sprachliche Sätze eine bestimmte grammatikalische Struktur annehmen. Dennoch können Grammatikregeln infolge unserer kognitiven Prozesse auf natürlichem Wege entstehen, wie es in dem Konzept der Universalgrammatik von Noam Chomsky vorgeschlagen wurde. In vergleichbarer Weise könnten infolge der Art des Evolutionsprozesses und genereller Einschränkungen der Netzwerkfunktionen universelle Regeln entstehen, die für die Organisation zellulärer Netzwerke gelten.



MIKROORGANISMEN SIND DIE EINZIGEN ORGANISMEN, DIE DIE ERDE WÄHREND DES GROSSTEILS IHRER GESCHICHTE BEVÖLKERT HABEN UND DIE DESHALB EVOLUTIONÄR WESENTLICH ÄLTER ALS TIERE UND PFLANZEN SIND.

Wie können diese Regeln verstanden werden? Ähnlich wie in der Sprachanalyse müsste auch hier der erste Schritt darin bestehen, eine ausreichende Anzahl von „Sätzen“ – Beispielen zellulärer Netzwerke – mit Hilfe einer Kombination aus experimentellen und rechnergestützten Methoden zu analysieren. Experimentell würde dies erfordern, die Struktur eines Netzwerks (d. h. Protein-Protein-Interaktionen und Protein-Reaktionen) abzubilden und zu erforschen, wie diese Interaktionen durch Stimulierung oder Störung des Netzwerkes, beispielsweise durch Änderungen in der Menge oder in den Aktivitäten einzelner Proteine, beeinträchtigt werden. Mit Hilfe einer computergestützten Analyse müsste man dann aus der Struktur und Funktion einzelner Netzwerke – trotz ihrer Verschiedenheit – allgemeine Regeln extrahieren. Solche Analysen können dazu beitragen aufzuklären, was für die Robustheit oder Modularität von Netzwerken verantwortlich ist. Die computergestützte Analyse kann zudem dazu benutzt werden, den gesamten Raum möglicher Netzwerke zu erforschen, die eine spezielle Funktion erfüllen können - durch eine zufällige Neuerschaltung eines Satzes von Proteinen und nachfolgende *in silico*-Überprüfung ihrer Funktionen in einem Prozess, der der natürlichen Evolution ähnelt. Netzwerke, die die gewünschte Funktionalität aufweisen, könnten weiter auf ihre Sekundäreigenschaften getestet werden, etwa Robustheit, und dann mit den natürlichen Netzwerken derselben Funktion verglichen werden. Idealerweise würden sich diese beiden Gruppen überlappen und bestätigen, dass wir verstehen, wie diese natürlichen Netzwerke durch die Evolution ausgewählt werden. Das sollte uns in die Lage versetzen, weitere grundsätzliche Prinzipien evolutionärer Netzwerkdesigns zu entdecken – was nicht nur für die synthetische Biologie, sondern auch für die Technik nützlich wäre.

MIKROORGANISMEN ALS MODELLSYSTEME

Die Aufklärung der gemeinsamen Prinzipien hinter natürlichen intrazellulären Netzwerken und der Regeln ihrer evolutionären Entwicklung würde einen riesigen Fortschritt für die Erforschung aller zellulären Organismen bedeuten. In dieser Hinsicht sind Mikroorganismen besonders attraktive Modelle für die System- und die synthetische Biologie. Insbesondere sind ihre zellulären Netzwerke – speziell in Bakterien – bedeutend einfacher als die in Tier- oder Pflanzenzellen. In Kombination mit dem Vorteil einer wesentlich einfacheren genetischen Manipulation in mikrobiellen Modellen hat dies zur Folge, dass die Funktion dieser Netzwerke tatsächlich in sehr hoher Detailgenauigkeit analysiert werden kann, und zwar sowohl experimentell als auch computergestützt. Darüber hinaus sind die Funktionen von Netzwerken in Mikroorganismen typischerweise einfacher zu definieren, wobei die meisten von ihnen direkt an der Wahrnehmung der Zellumgebung oder des Zustandes innerhalb der Zelle beteiligt sind. Letztlich sind Mikroorganismen die einzigen Organismen, die die Erde während des Großteils ihrer Geschichte bevölkert haben und die deshalb evolutionär wesentlich älter als Tiere und Pflanzen sind. Die evolutionäre Optimierung mikrobieller Netzwerke hat also vermutlich bereits zu einer bestmöglichen Lösung geführt. Diese definierten Funktionen und die angenommene Optimalität erleichtern theoretische Analysen der Netzwerke, da wir wissen (oder zumindest annehmen können), welche Eigenschaften und Funktionen des Netzwerkes evolutionär ausgewählt wurden und wie die Evolution diese Fragen gelöst hat.