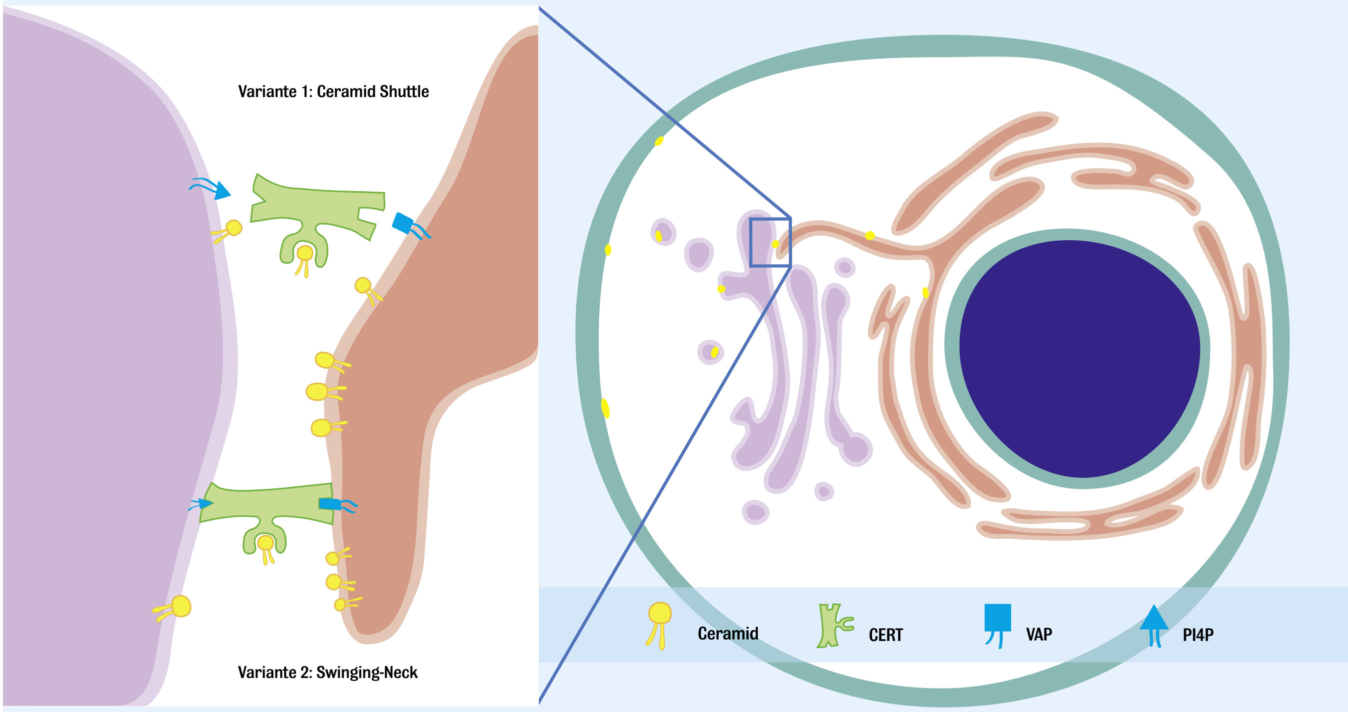
# Mit Simulationstechnik zu neuen Erkenntnissen in der Systembiologie



Lipidtransport
*Mechanismen des Lipidtransportes durch CERT an Membrankontaktstellen zwischen ER und TGN*

Die Systembiologie ist ein aufstrebendes, vergleichsweise junges Forschungsgebiet, deren Ziel ein besseres ganzheitliches Verständnis der Mechanismen biologischer Systeme ist. Biologische Systeme (z.B. Populationen, der Organismen, ein Organ oder auch einzelne Zellen und ihre intrinsischen Signalwege) werden als komplexe dynamische Interaktionsnetzwerke aufgefasst, deren charakteristisches Verhalten eine Folge ihrer Struktur sowie der Dynamik innerhalb des Netzwerkes ist. Um biologische Systeme zu verstehen, ist es daher wichtig, dass die Untersuchung einzelner Bestandteile des Systems im Kontext des gesamten Systems durchgeführt wird. Für ein holistisches Verständnis eines gewählten biologischen Systems lassen sich systemtheoretische und mathematische Modellierungs- und Analysemethoden sowie experimentelle Techniken kombinieren.

## SUMMARY

*Systems biology is a relatively young research field, which was developed in the 20th century. It combines biology, systems theory and simulation technology. Facilitated by enormous developments in experimental techniques as well as by rapidly increasing computing power, we are now able to reach a holistic understanding of biological systems.*

*The simulation of quantitative models of those systems provide interesting new hypotheses, and mathematical analyses help to address recent questions in all biological research areas – ranging for example from the development of new drugs to process optimization in biotechnology.*

*This article focuses on the role of simulation especially for statistical sampling-based approaches for model calibration and the generation of hypotheses. Simulation technology plays a central role here between data acquisition and model analysis. We will exemplify the power of these sampling-based approaches with a cooperation project between the Institute of Systems Theory and Automatic Control (IST) and the Institute of Cell Biology and Immunology (IZI), in which we investigate key regulation processes of protein secretion in mammalian cells.*

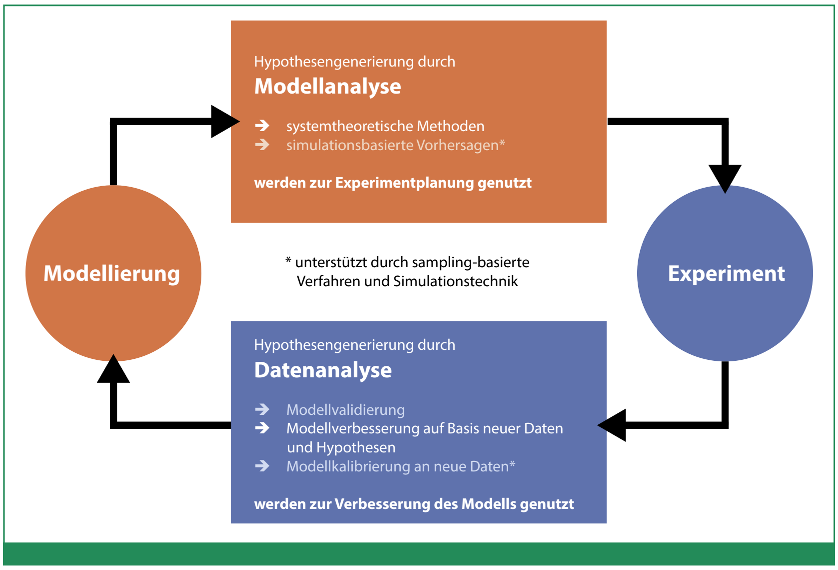
## 1. Einleitung

Die Etablierung der Systembiologie als eigenständiges Forschungsgebiet ist eine logische Konsequenz der enormen Entwicklungen in mehreren Disziplinen in den letzten Jahren. Dies sind zum einen faszinierende neue experimentelle Techniken in der Zell- und Molekularbiologie, insbesondere auf der Genom-, Transkriptom- und Proteomebene sowie im Bereich des sogenannten Imaging, welches es erlaubt, Prozesse und Wechselwirkungen auf molekularer Ebene sichtbar zu machen. Zum anderen führen leistungsfähigere Computersysteme zu großen Fortschritten in der Simulationstechnik und der Entwicklung computerbasierter Analysewerkzeuge für experimentelle Daten und Modelle. Die sich daraus ergebenden Möglichkeiten sind ebenso vielfältig wie die damit verbundenen Herausforderungen. Beispielsweise reichen herkömmliche Methoden zur Datenauswertung schon lange nicht mehr aus, um die rasant wachsende Menge an experimentellen Informationen so zu verarbeiten, dass sie interpretiert und geeignet visualisiert werden können.

Hier sind neue mathematische Methoden gefragt, welche die in den Datensätzen verborgenen Hinweise auf die Strukturen sowie Dynamik innerhalb des untersuchten Systems entschlüsseln helfen.

Das so gewonnene Wissen über Systemstruktur und -dynamik ermöglicht nun die Entwicklung von mathematischen Modellen zu den jeweiligen biologischen Systemen. Diese Modelle können mit Methoden der mathematischen Systemtheorie untersucht werden, um charakteristische Eigenschaften des Systems zu analysieren. Zusätzlich liefern sie Vorhersagen über das Systemverhalten und tragen somit genauso zum Verständnis der biologischen Systeme bei wie Experimente am System selbst.

Diese modellbasierten Simulationen des Systemverhaltens am Computer, sogenannte *in silico* Experimente, haben gegenüber den entsprechenden *in vivo* bzw. *in vitro* Experimenten im Labor viele Vorteile. Sie sind beispielsweise oft günstiger und schneller als Experimente im Labor, Simulationsszenarien können präzise formuliert und damit nachvollziehbar gemacht werden, und es können ethisch bedenkliche Eingriffe an lebenden Pflanzen oder Tieren ersetzt werden.



01

Der systembiologische Zirkel

*Der Modellierer, typischerweise ein Mathematiker, Physiker oder Ingenieur, steht in ständigem Austausch mit dem experimentellen Biologen. Der Biologe formuliert eine Fragestellung, welche dann in die Modellsprache übersetzt werden muss. Daraufhin überlegen sich beide eine Modellvorhersage, die zunächst simuliert wird. Nun muss diese Vorhersage mit Daten aus einem geeigneten neuen Experiment untermauert oder widerlegt werden. Im ersten Fall führt dies zu einer Modellverfeinerung, im zweiten Fall zu einer Modelländerung oder Erweiterung, und das Spiel beginnt erneut.*

Doch wie genau gelangen Systembiologen nun zur ihren Ergebnissen? Der Weg von der Idee zur Erkenntnis lässt sich in diesem Bereich gut mit dem „systembiologischen Zirkel“, wie in **(01)** dargestellt, erklären: Ausgehend von einer biologischen Fragestellung oder Hypothese wird, meist basierend auf aktueller Literatur und existierenden Daten, ein erstes Modell erstellt. Systemtheoretische Modellanalysen und simulationsbasierte Vorhersagen generieren daraufhin Hypothesen, welche in weiteren Experimenten getestet werden können. Die Interpretation der neu erhobenen Daten führt wiederum zu Hypothesen, die zur strukturellen Verbesserung und Erweiterung des Modells genutzt werden. Die erhobenen Datensätze werden zudem zur Kalibrierung des erweiterten Modells verwendet. Modellbildung und Simulation spielen hierbei an mehreren Stellen eine entscheidende Rolle. Wie genau simulationsbasierte Vorhersagen, die Kalibrierung und Validierung der hierfür nötigen Modelle funktioniert, soll im Folgenden näher beleuchtet werden.

## 2. Der Weg von der Modellierung bis zur Vorhersage

### 2.1 Datengetriebene Modellierung

In unseren systembiologischen Projekten verwenden wir häufig *datengetriebene Modellierung*. Hierbei werden bei der Erstellung des Modells neben bereits vorhandenem Vorwissen über das System auch die experimentellen Rahmenbedingungen miteinbezogen. Die daraus entstehenden Modelle sind spezifisch für die mit den vorhandenen Daten zu beantwortende Fragestellung optimiert und erfordern von der ersten Modellvariante bis zum finalen Modell eine enge Zusammenarbeit von Modellierer und Experimentator.

Bei einem datengetriebenen Modellierungsansatz wählt man die Anzahl der zu schätzenden Parameter in dem Modell so klein wie möglich, während die Modellstruktur so komplex wie nötig gehalten wird. Diese Vorgabe führt dazu, dass Modelle aus einem datengetriebenen Modellierungsansatz speziell bei der Modellkalibrierung sowie bei der Erstellung von Hypothesen aus Modellsimulationen ihre Stärken haben. Bei zu einfachen Modellen können die in den Daten enthaltenen Informationen indes nicht optimal ausgenutzt werden. Außerdem werden vielleicht nicht alle Eigenschaften des biologischen Systems abgebildet. Für ein sehr komplexes Modell dagegen, das grundsätzlich viele verschiedene Verhaltensweisen zeigen kann, ist es oft schwierig die Modellparameter zu bestimmen. Zum einen wird das Lösen des Schätzproblems mit wachsender Anzahl von Parametern selbst schon schwieriger, zum anderen besteht die Gefahr des *Overfitting*. Das heißt, dass man bei der Kalibrierung das Modell nicht nur an die relevanten Informationen über das System, sondern auch an die zufälligen Rauscheigenschaften des vorliegenden Datensatzes anpasst. Diese Gefahr besteht besonders bei Datensätzen mit großen Unsicherheiten – 20 Prozent Messfehler sind keine Seltenheit in der Biologie – und wenigen Wiederholungsmessungen. Das Modell kann in Gefahr laufen Ausreißer in den Daten als Systemverhalten zu interpretieren. Als Folge erhält man ein Modell mit kleinem Trainingsfehler, da die Anpassung an die zur Schätzung verwendeten Daten sehr gut ist. Allerdings wird das Modell nicht sehr gut generalisieren und große Fehler bei der Vorhersage neuer Szenarien machen. Durch die Abstimmung der Modellstruktur an die experimentellen Rahmenbedingungen wird also eine maximale Übertragung der in den Datensätzen enthaltenen Informationen auf das Modell ermöglicht. Damit können die zur Kalibrierung verwendeten Daten ausreichend gut beschrieben werden, und das Modell liefert gleichzeitig gute Vorhersagen.

### 2.2. Punktschätzer und sampling-basierte Parameterschätzmethoden

Hat man sich für ein Modell entschieden und die entsprechenden Gleichungen aufgestellt, so ist der nächste Schritt die *Modellkalibrierung*, also das Anpassen des Modells an die tatsächlichen experimentellen Daten. Bei parametrisierten Modellen heißt das, Werte für alle Modellparameter so zu finden, dass die Modellsimulationen mit den gemessenen Daten möglichst gut übereinstimmen. Ein üblicher Ansatz besteht darin, die Parameterschätzung als Optimierungsproblem zu formulieren. Ein kleiner Abstand zwischen Simulation und Daten wäre hier etwa ein Optimierungsziel, und wir benötigen ein geeignetes Maß für diesen Abstand, unsere *Zielfunktion*.

Oft werden beispielsweise *Kleinste-Quadrate-Schätzer*, welche die Summe der quadratischen Fehler zwischen Modellsimulation und Messdaten minimieren, oder *Maximum-Likelihood-Schätzer* verwendet. Letztere gehen davon aus, dass die Messdaten aus einem stochastischen Prozess erzeugt wurden. Mögliche Messausgänge werden somit als Zufallsvariablen interpretiert, und die vorliegenden Messdaten stellen eine Stichprobe aus deren Wahrscheinlichkeitsverteilung dar. Die zu optimierende Zielfunktion ist hier die *Likelihood-Funktion*, die für jeden Parametersatz die Wahrscheinlichkeit wiedergibt, die vorliegenden Daten zu beobachten. Der Maximum Likelihood Schätzer liefert somit Parameter, welche diese Wahrscheinlichkeit maximieren.

Solche Schätzer werden auch als *Punktschätzer* bezeichnet, da sie für jeden Parameter einen einzigen optimalen Punkt bezogen auf das Abstandsmaß liefern. Für den Fall, dass man ausreichend große Datensätze zur eindeutigen Identifizierung der Parameter sowie ein Modell vorliegen hat, welches die Daten gut erklären kann, sind diese Punktschätzer geeignet und können auch für weitere Analysen verwendet werden. Dies ist in der Systembiologie besonders für das Kalibrieren quantitativer Modelle allerdings selten der Fall. Oft sind Messungen auf molekularer Ebene schwierig, teuer und aufwendig, so dass die Datensätze klein im Vergleich zur Modellkomplexität sind und nicht genügend Informationen enthalten, um alle Parameterwerte eindeutig zu bestimmen. Man kann zum Beispiel nicht alle wichtigen Moleküle eines Signalwegs messen, so dass man nicht beobachtbare Variablen im Modell hat, oder es stehen für Zeitreihen nur wenige Messzeitpunkte zur Verfügung. Die Daten werden in diesem Fall auch als *sparse* bezeichnet, und es sind nicht alle Parameter identifizierbar. Das Optimierungsproblem ist in diesem Fall *schlecht-gestellt*, da es keine eindeutige Lösung hat. Es ist allgemein nicht immer offensichtlich, ob man es mit einem schlecht-gestellten Problem zu tun hat. In der Praxis deuten unterschiedliche Ergebnisse bei einer Optimierung mit unterschiedlichen Startparametern, die aber ähnliche Zielfunktionswerte liefern, auf ein solches schlecht-gestelltes Problem hin. In diesem Fall kann man ad hoc nicht entscheiden, welcher dieser Parametersätze nun der Beste ist.

Für schlecht-gestellte Probleme reichen Punktschätzer zur weiteren Analyse also nicht aus. Es gibt sehr unterschiedliche Ansätze, um mit diesem Problem umzugehen. Die Theorie schlecht-gestellter inverser Probleme stellt zum Beispiel *Regularisierungsverfahren* bereit, bei denen die Zielfunktion für die Optimierung neben einem Term zur Anpassung an die Messdaten auch einen datenunabhängigen *Regularisierungsterm* enthält, welcher zu komplexe Modelle bestraft. Hierdurch soll zum einen Overfitting vermieden werden, und zum anderen soll das Optimierungsproblem in ein gut-gestelltes Problem mit eindeutiger Lösung umgewandelt werden.

Aus der Statistik haben sich sogenannte *Bayes'sche Lernverfahren* entwickelt, die eine statistisch konsistente Beschreibung der in den Daten enthaltenen Informationen über alle Modellparameter in Form von Wahrscheinlichkeitsverteilungen liefern. Bei solchen Ansätzen werden also sowohl die Daten als auch die Parameter als Zufallsvariablen mit zugrunde liegenden Wahrscheinlichkeitsverteilungen interpretiert. Neben der Likelihood-Funktion zur Beschreibung der Datengenerierung repräsentiert eine *a-priori-Wahrscheinlichkeitsverteilung* den aktuellen Wissensstand über die noch nicht angepassten Parameter. Die Leistungssteigerung von Bayes'schen Verfahren im Vergleich zu Punktschätzern wurde am IST an konkreten Anwendungsbeispielen demonstriert.

Man interessiert sich nun bei der Bayes'schen Parameterschätzung für die *a-posteriori-Verteilung*, eine Wahrscheinlichkeitsverteilung über den Parametern, die den aktuellen Wissensstand nach Einbezug der vorliegenden Messdaten wiedergibt. Diese ist nach dem Satz von Bayes proportional zum Produkt aus a-priori-Verteilung und Likelihood-Funktion. Für detailliertere Informationen über Bayes'sche Lernverfahren und den Satz von Bayes sei hier auf **(b01)** verwiesen. Die a-priori Verteilung in Bayes'schen Lernverfahren kann in bestimmten Fällen auch die Rolle des Regularisierungsterms einnehmen, was man als *Bayes'sche Regularisierung* bezeichnet. In diesem Fall haben wir es mit einem gut-gestellten Problem zu tun, und es reichen Punktschätzer wie beispielsweise der *Maximum-a-posteriori Schätzer*, welcher die a-posteriori Wahrscheinlichkeit maximiert, für weitere Analysen aus.

Da dies für unsere Modelle im allgemeinen nicht der Fall ist, werden globale Informationen über die a-posteriori Verteilung benötigt. Diese Verteilung ist meist nicht analytisch zugänglich und wird üblicherweise durch *sampling-basierte Ansätze* genähert. Hierbei werden repräsentative Stichproben aus der a-posteriori Verteilung erzeugt (*Sampling*), mit deren Hilfe man Informationen über die Verteilung selbst ableiten kann. Sampling-basierte Verfahren sind sehr mächtig, da sie Informationen über Modellparameter inklusive Unsicherheiten enthalten, die auch zur Schätzung von Modellvorhersagen mit Unsicherheiten verwendet werden können.

Die Erzeugung solcher Stichproben ist je nach Modell und Datenlage allerdings nicht immer einfach. In unseren Anwen-

### DER SATZ VON BAYES

Der Satz von Bayes geht auf den Mathematiker Thomas Bayes (1701 - 1761) zurück. Er ermöglicht es, bedingte Wahrscheinlichkeiten umzukehren. Für eine Beobachtung $A$ gibt es eine mögliche Erklärung $B$, und die Wahrscheinlichkeit, dass $A$ und $B$ gemeinsam eintreten, ist gegeben durch:

$$P(A,B) := P(A|B)P(B) = P(B|A)P(A),$$

Die bedingte Wahrscheinlichkeit, dass $B$ die Beobachtung $A$ verursacht hat, ergibt sich nach Umstellen zu:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}.$$

Dabei ist $P(B)$ die von $A$ unabhängige Wahrscheinlichkeit, dass $B$ zutrifft. Sie wird auch als Vorwissen bzw. *a-priori Wahrscheinlichkeit* bezeichnet. $P(A)$ bezeichnet die Wahrscheinlichkeit für Beobachtung $A$, und $P(A|B)$ ist ein Maß dafür, wie wahrscheinlich es ist $A$ bei gegebenem $B$ zu beobachten.

Dieser Satz gilt auch für Wahrscheinlichkeitsdichten für reelle Zufallsvariablen, z.B. von Beobachtungen wie Proteinkonzentrationen (Daten $D$) und Modellparametern $\theta$. In einem Bayes'schen Ansatz zur Parameterschätzung wird die *a-posteriori Verteilung* $p(\theta|D)$ untersucht. Diese beschreibt die Wahrscheinlichkeit für Parameterwerte $\theta$ bei gegebenen Daten $D$, und ist nach dem Satz von Bayes gegeben durch:

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)},$$

wobei $P(D|\theta)$ die *Likelihood Funktion* bezeichnet.

### DER SATZ VON BAYES - EIN BEISPIEL

Für ein Würfelexperiment darf Spieler A zwischen drei Würfeln mit 6, 12 oder 20 Seiten wählen, und er kann sich entscheiden ein oder zweimal zu würfeln. Spieler B soll anhand der gewürfelten Summe entscheiden, welche Variante Spieler A wohl gewählt hat. Spieler A teilt ihm die Zahl 12 mit. Es gibt sechs mögliche Erklärungen $B = [n;s]$ ($n$ Würfe à $s$ Seiten) für den Wurf mit den folgenden Wahrscheinlichkeiten:

$P(A = 12|B = [n,s])$

| | 1 Wurf | 2 Würfe |
|---|---|---|
| 6 Seiten | 0 | 1/36 |
| 12 Seiten | 1/12 | 11/144 |
| 20 Seiten | 1/20 | 11/400 |

Die gesuchte Umkehr erhält man mit der Annahme, dass a priori alle Varianten gleich wahrscheinlich sind ($P(B_i) = 1/6$), und $P(A = 12) = 0.044$ durch den Satz von Bayes:

$P(B = [n,s]|A = 12)$

| | 1 Würfel | 2 Würfel |
|---|---|---|
| 6 Seiten | 0 | 0.105 |
| 12 Seiten | 0.314 | 0.288 |
| 20 Seiten | 0.189 | 0.104 |

Nach dem Satz von Bayes bekommt somit die einfachste Erklärungs-Variante (ein 12-Seitiger Würfel) die höchste Wahrscheinlichkeit $P(12|[1,12]) = 0.314$ zugeordnet. Einfach bedeutet in dem Fall: die Variante mit den wenigsten kombinatorischen Möglichkeiten, jedoch nicht so einfach, dass sie die Beobachtung gar nicht erklärt $[1,6]$ oder sehr unwahrscheinlich ist $[2,6]$. Nach dem Satz von Bayes sollte sich Spieler B also für diese Variante entscheiden.

**b01**

dungen kommen hierfür *Markov Chain Monte Carlo* (*MCMC*) Verfahren zum Einsatz, die in **(b02)** näher erläutert werden. Diese Sampling Verfahren sind im allgemeinen sehr rechenintensiv, da sie viele Auswertungen der Zielverteilung erfordern. Bei Systemen von nichtlinearen Differenzialgleichungen, wie wir sie verwenden, muss hierfür zur Auswertung der Likelihood-Funktion das Modell viele Male numerisch integriert werden. Dies macht die Rechenzeit zum limitierenden Faktor, so dass eine Anwendung dieser Methoden in der Praxis auf Modelle mittlerer Größe beschränkt ist. Um dies zu verbessern hat das IST in den letzten Jahren sehr effektive sampling-basierte Methoden zur Parameterschätzung und Experimentplanung speziell für Differentialgleichungsmodelle entwickelt. Hierbei spielt die Simulationstechnik eine zentrale Rolle, da effiziente numerische Simulationsverfahren der Schlüssel zur Laufzeitoptimierung sind.

### 2.3 Die Vorhersage – Mehr als nur der Kaffeesatz

In den meisten Fällen interessiert man sich nicht so sehr für die Parameterwerte direkt, sondern vielmehr für Vorhersagen, welche mit dem kalibrierten Modell getroffen werden können. So könnte man beispielsweise ein neues, experimentell noch nicht getestetes Szenario simulieren und damit eine Modellvorhersage für dieses Szenario erhalten. Verwendet man für solche Vorhersagen Parameterwerte, die aus Punktschätzern erhalten wurden, so bekommt man eine einzelne Lösung. In einem Bayes'schen Kontext sind zusätzlich darüber hinausgehende Informationen über Unsicherheiten in Form von Wahrscheinlichkeitsverteilungen enthalten, so dass sich aus der a-posteriori-Verteilung im Parameterraum prinzipiell eine entsprechende Verteilung für die Vorhersage ableiten lässt. Für ein Differenzialgleichungsmodell wäre das beispielsweise eine sich zeitlich ändernde Wahrscheinlichkeitsdichte im Zustandsraum. Diese beinhaltet natürlich erheblich mehr Information als eine einzelne Modellvorhersage. So würde man zum Beispiel einer großen Abweichung zwischen Modellvorhersage und Messung an einem Zeitpunkt weniger Bedeutung beimessen, wenn die Modellvorhersage für diesen Punkt eine große Varianz aufweist.

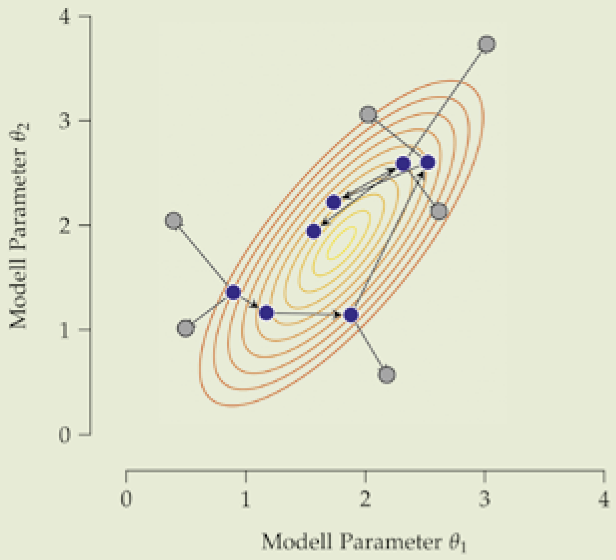
In der Praxis wird eine solche Verteilung der Vorhersage meist ebenfalls durch eine empirische Schätzung mit Hilfe der a-posteriori Stichprobe genähert. Auch hierzu ist die Simulation des Modells mit den Parameterwerten aus der Stichprobe nötig. Eine solche sampling-basierte Generierung von Vorhersagen mit anschließender Dichteschätzung ist daher ebenfalls sehr rechenintensiv. Methoden der Simulationstechnik sind hier also sehr gefragt, steigern sie doch die Effizienz der Berechnungen.

Soviel zur Theorie – ganz konkret im Einsatz helfen solch sampling-basierten Ansätze in der systembiologischen Forschungspraxis z. B. bei der Untersuchung molekularer Regulationsmechanismen. Dies wollen wir in den nächsten Abschnitten mit einem Kooperationsprojekt des IST und des IZI an der Universität Stuttgart erläutern. Dazu beginnen wir vorerst mit einem kleinen Exkurs in die Biologie des Golgi-Komplex und seiner Bedeutung.

## MCMC IN DER PARAMETERSCHÄTZUNG

*Markov Chain Monte Carlo (MCMC)* Sampling ist ein Verfahren zur effizienten Generierung von Stichproben aus hochdimensionalen Verteilungen $p(\theta)$. Mit Hilfe einer Markovkette mit Übergangswahrscheinlichkeit $p(\theta'|\theta)$ wird abhängig vom aktuellen Parameter $\theta$ ein neues $\theta'$ vorgeschlagen, welches mit einer bestimmten Akzeptanzwahrscheinlichkeit $\alpha$ angenommen wird. Dabei ist $\alpha$ so gewählt, dass die akzeptierten $\theta$, welche in untenstehender Abbildung blau markiert sind, eine Stichprobe aus der Zielverteilung darstellen. Diese Zielverteilung ist in der Abbildung durch Höhenlinien repräsentiert. Ein bekannter und oft verwendeter MCMC Sampling-Algorithmus ist der *Metropolis-Hastings-Algorithmus*:

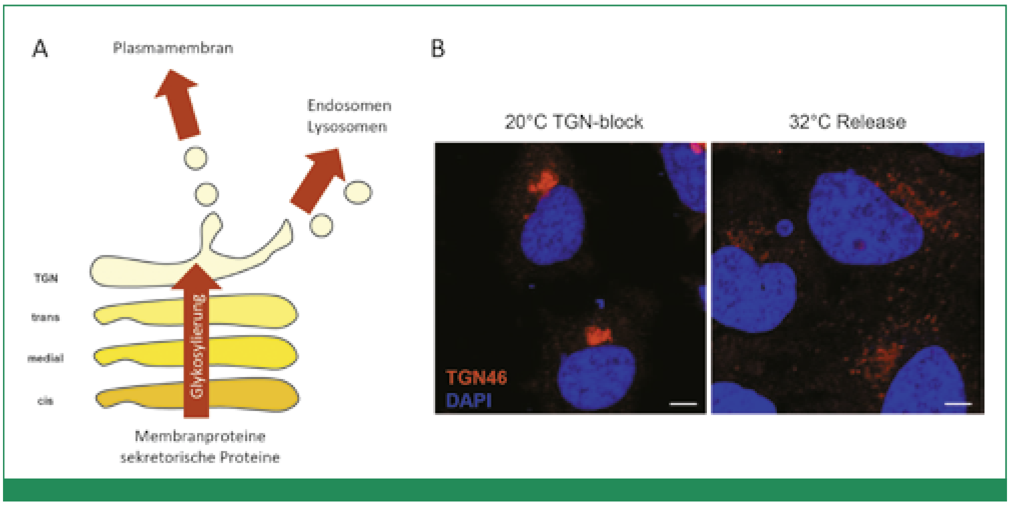
1. Initialisiere die Markovkette mit $\theta_0$ und $p(\theta'|\theta)$ und setze $i = 0$
2. Ziehe $\theta'$ aus $p(\theta'|\theta_i)$
3. Setze $\alpha := \min\left(1, \frac{p(\theta')}{p(\theta_i)} \frac{p(\theta_i|\theta')}{p(\theta'|\theta_i)}\right)$
   und
   $$\theta_{i+1} = \begin{cases} \theta' & \text{mit Wahrscheinlichkeit } \alpha \\ \theta_i & \text{sonst} \end{cases}$$
4. Setze $i = i + 1$ und gehe zu 2



b02

## 3. Molekulare Regulationsmechanismen am Golgi Apparat – Ein kleiner Einblick in die Welt der Zellbiologie

Der von dem Wissenschaftler und Nobelpreisträger Camillo Golgi 1898 entdeckte Golgi-Komplex zählt zu den Organellen eukaryotischer Zellen und stellt eine charakteristische polar aufgebaute Membranstruktur dar, die bei der Sekretbildung eine wichtige Rolle spielt. Membranproteine und Proteine, die sezerniert werden, gelangen nach ihrer Synthese am Endoplasmatischen Retikulum (ER) zunächst in den cis-Golgi, wo sie schrittweise während der Passage zum medial- und trans-Golgi durch Glykosylierung modifiziert, um schließlich am trans-Golgi-Netzwerk (TGN) in unterschiedliche Transportvesikel verpackt zu werden. Der Zielort dieser Transportvesikel können andere interne membranumhüllte Organellen wie z.B. Lysosomen, Endosomen oder die Plasmamembran sein, wo lösliches Cargo-Protein wie beispielsweise Antikörper oder Hormone in die extrazelluläre Umgebung freigesetzt wird **(02)**. Darüber hinaus spielt der Golgi-Komplex eine wichtige Rolle in der Zellpolarität, die für den Aufbau von epithelialen Zellverbänden und Organstrukturen, aber auch für die gerichtete Zellbewegung während der Wundheilung essentiell ist.



02

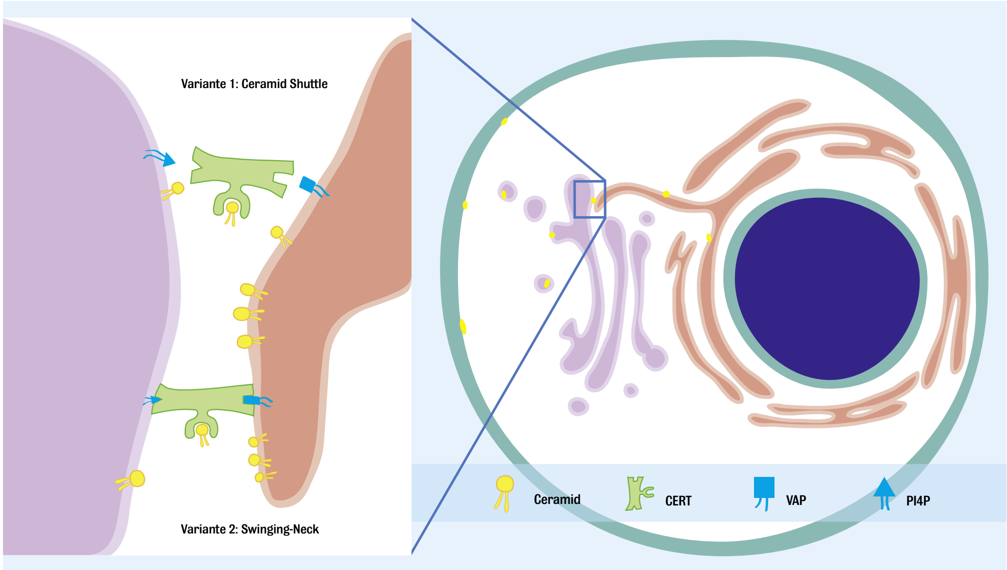
Proteintransport durch den Golgi-Komplex *(A) Membranproteine und sekretorische Proteine werden durch den Golgi-Komplex in cis/medial/trans Richtung geschleust und durch Glykosylierung modifiziert. Am TGN erfolgt die Sortierung und Verpackung in Vesikel, die zur Plasmamembran oder zu Endosomen und Lysosomen transportiert werden. (B) Ein in HeLa-Zellen eingebrachtes rot-markiertes Membranprotein (TGN46) akkumuliert durch eine Senkung der Temperatur auf 20°C am TGN. Nach Erwärmen der Zellen auf 32°C (Release) verlässt das Membranprotein das TGN in Vesikeln, die zur Plasmamembran transportiert werden. Der Zellkern ist in blauer Farbe dargestellt. Größenmaßstab 5 µM.*

Durch intensive Forschung in den letzten Jahren unter anderem auch am IZI ist das Wissen um die zentralen Moleküle, die für die Sortierung und Verpackung von Proteinen am TGN verantwortlich sind, zum Teil aufgeklärt worden. Ein komplexes Zusammenspiel aus Lipiden und Proteinen ist für die Ausbildung von Transportvesikeln am TGN notwendig. Ein besonders wichtiges Lipid in Golgi-Membranen ist die monophosphorylierte Form von Phosphatidylinositol (PI4P), welches als Signallipid Proteine mit PH-Domäne an Membranen rekrutiert. Eine PH-Domäne ist dabei eine Sequenz in diesem Protein, welche es ihm ermöglicht an bestimmte Membranen anzudocken. Ein solches PH-Domänen Protein ist das Lipidtransferprotein CERT, dessen Aufgabe darin besteht, das Lipid Ceramid von ER Membranen aufzunehmen und zum TGN zu transportieren. Die am TGN lokalisierte Sphingomyelinsynthase wandelt Ceramid in die beiden Lipide Sphingomyelin und Diacylglycerol um, die wiederum für die Ausbildung von Transportvesikeln am TGN unabdingbar sind. Ein funktionsunfähiges CERT-Protein führt damit zu Störungen im zellulären Lipidhaushalt und damit verbundenen Defekten im Membran- und Proteintransport, so dass das Zellüberleben gefährdet ist.

Es ist daher leicht nachvollziehbar, dass eine erhöhte oder erniedrigte Konzentration von CERT mit pathophysiologischen Veränderungen wie sie in Krebszellen zu finden sind im Zusammenhang steht. Des weiteren ist die Abhängigkeit intrazellulärer pathogener Viren und Bakterien von der CERT-Funktion der Wirtszelle bekannt. Studien des IZI in Zusammenarbeit mit der Prozessentwicklung von Boehringer Ingelheim Pharma GmbH lieferten den Beweis dafür, dass sich die einzigartige Funktion des CERT-Proteins im Lipidtransfer und der Golgi-Funktion auch für biotechnische Anwendungen ausnutzen läßt. So werden komplexe therapeutische Proteine wie zum Beispiel Antikörper heutzutage standardmäßig mit Hilfe von Säugerzellen produziert. Da Säugerzellen im Gegensatz zu Bakterien in der Kultivierung sehr anspruchsvoll sind, ist der Bioprozess zur Produktion solcher Proteine mit enormen Kosten verbunden. Ein Ansatz der Bioprozessoptimierung ist die Steigerung der Sekretionsleistung der verwendeten Produktionszellen. In der Tat konnten wir in der Vergangenheit zeigen, dass die genetische Modifikation der Produktionszellen durch stabile Erhöhung der CERT Konzentration zu einer signifikant gesteigerten Produktivität in sogenannten fed-batch Kultivierungen führt. Diese Ergebnisse eröffnen neue Wege für verbesserte Produktionsprozesse der Zukunft.

CERT selbst wird durch ein komplexes, nur im Teil verstandenes Proteinnetzwerk, das komplexe Rückkopplungsmechanismen enthält, reguliert. Zu den beteiligten Molekülen gehören auch die Lipidkinase Phosphatidylinositol 4-kinase III beta (PI4KIIIß), die am Golgi Komplex PI4P produziert und die Proteinkinase D (PKD), die sowohl CERT als auch PI4KIIIß Funktionalität durch direkte Phosphorylierung steuert. Der genaue Lipidtransport-Mechanismus durch CERT an sogenannten membrane contact sites (MCS), an denen sich ER- und TGN-Membranen in unmittelbarer räumlicher Nähe befinden, ist ebenfalls noch ungeklärt. Zwei verschiedene Modelle werden derzeit diskutiert: CERT könnte gleichzeitig über seine PH-Domäne mit Golgimembranen und einem zweiten, spezifischen Bindemotif mit dem ER verbunden sein, so dass nur die hydrophobe Ceramidbindetasche zwischen den beiden Organell-Membranen hin und her schwingt (Swinging Neck Modell). Alternativ könnte CERT sequenziell an diese unterschiedlichen Membranen binden und die kurze Distanz zwischen den Membranen per Diffusion zurücklegen (Shuttle Modell). Beide Modellvarianten sind in **(03)** dargestellt.

Die effektive Nutzung dieses CERT-Netzwerkes, beispielsweise im Rahmen der Optimierung von Produktionszellen, erfordert ein tiefgehendes Verständnis über die Interaktionen zwischen den beteiligten Molekülen. Um die molekulare Regulation und Wirkungsweise des komplexen CERT-Netzwerkes verstehen zu können, ist deshalb ein mathematischer Modellierungsansatz, welcher Rückkopplungs- und Transportmechanismen erklären kann, essentiell.



**03**

Mechanismen des Lipidtransportes durch CERT an Membrankontaktstellen zwischen ER und TGN

*CERT bindet an TGN Membranen über die Interaktion seiner PH-Domäne mit PI4P. Die Bindung an das ER erfolgt über das ER Protein VAP. Der Ceramidtransport erfolgt entlang eines Konzentrationsgradienten vom ER zum TGN. Variante 1 stellt die Hypothese auf, daß CERT sequenziell an ER und TGN Membranen bindet und den kurzen Weg zwischen den Membranen per Diffusion zurücklegt (Shuttle). Variante 2 postuliert, dass CERT gleichzeitig über seine PH-Domäne mit Golgi-Membranen und über das spezifische Bindemotif mit dem ER verbunden ist, so dass nur die hydrophobe Ceramidbindetasche zwischen den beiden Organell-Membranen hin und her schwingt (Swinging Neck).*

Wie aber gehen wir bei der Erstellung eines solchen Modells konkret vor? Und insbesondere: Können wir mit Hilfe unserer Modelle tatsächlich etwas lernen über die Transportmechanismen des CERT Proteins? Wie wir sehen werden, eignen sich auch hier die zuvor erklärten sampling-basierten Ansätze sehr gut! Beispielhaft wollen wir im Folgenden erläutern, wie unsere Modelle und Analysemethoden die Erforschung der genauen Mechanismen des CERT Transports unterstützen können.

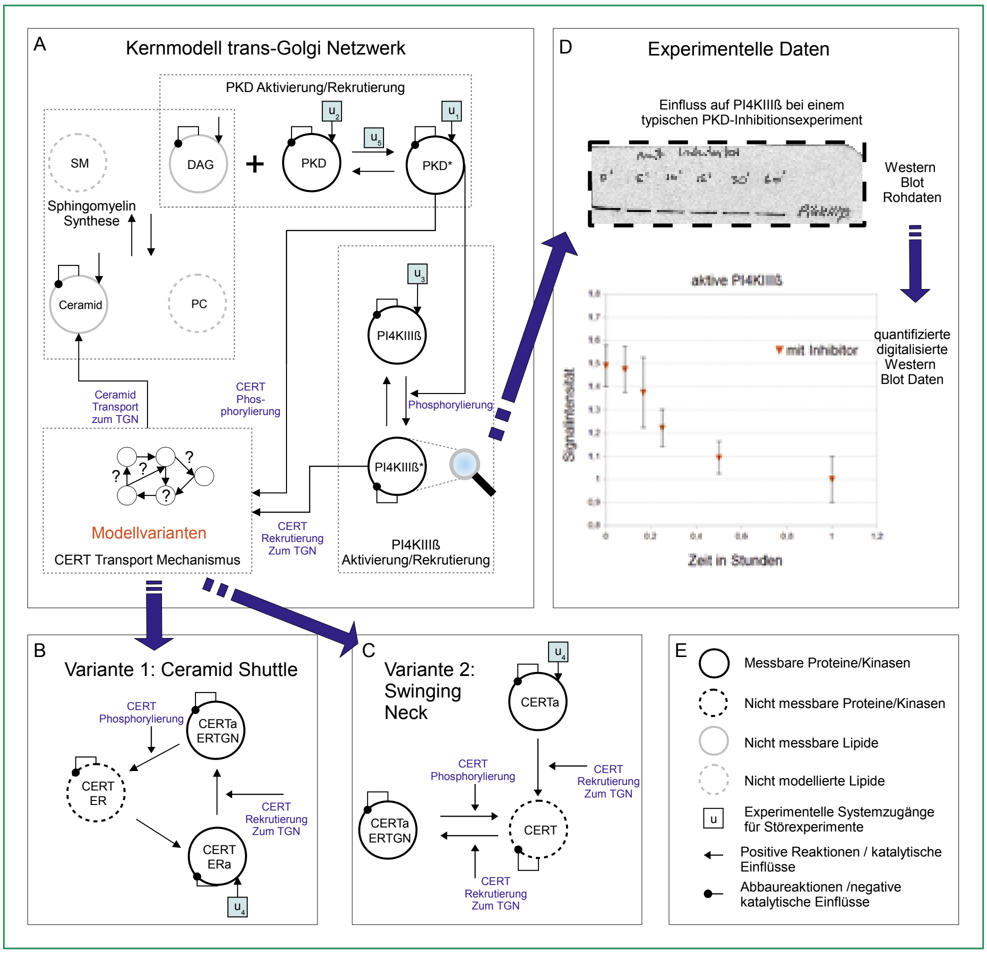
## 4. Den molekularen Mechanismen des Regulationsnetzwerks von CERT auf der Spur

### 4.1 Von der Modellerstellung ...

Zur Erstellung eines parametrisierten Modells für das Regulationsnetzwerk von CERT wurden zunächst aktuelle Publikationen und Daten aus der Literatur herangezogen. Zusätzlich standen Datensätze aus Experimenten zur Verfügung, welche Informationen über den zeitlichen Verlauf der Aktivität einiger Proteine des Systems enthalten. Da diese Messwerte das Mittel einer gesamten Population von Zellen beschreiben, von dem angenommen werden kann, dass es sich deterministisch verhält, und da keine Information über die Lokalisation der Moleküle innerhalb der Zelle vorliegt, haben wir eine Modellierung mit gewöhnlichen Differenzialgleichungen gewählt. Dies ist mittlerweile ein Standardansatz in der Systembiologie, und es gibt Regeln für das Aufstellen der entsprechenden Gleichungen.

**(04A)** zeigt die Struktur unseres Modells. Variablen in diesem Modell repräsentieren miteinander agierende Lipide und Proteine und deren chemisch modifizierte Formen. Experimentell nicht messbare Variablen sind gestrichelt dargestellt. Die mit „u“ gekennzeichneten Modelleingänge beschreiben mögliche Störexperimente, welche am IZI durchgeführt werden können, wie beispielsweise von außen induzierte Änderungen der Proteinkonzentrationen.

Um unsere Frage nach dem Transportmechanismus von CERT modellbasiert zu untersuchen, wurden zwei Modellvarianten erstellt, die in **(04B)** und **(04C)** sche-



**04**

Modellstruktur des Regulationsnetzwerkes von CERT und experimentelle Daten
*A: Graphische Darstellung des Kerndifferentialgleichungsmodells, welches die Interaktionen der Biomoleküle am TGN beschreibt. Verschiedene Lipide und Proteine beeinflussen sich gegenseitig über chemische Reaktionen. Nicht experimentell messbare Variablen sind gestrichelt dargestellt. Experimentelle Zugänge – mit „u" gekennzeichnet – beschreiben Stellen an denen Störexperimente möglich sind. Einige Modellteile sind bekannt wie z.B. die PI4KIIIβ Aktivierung/Rekrutierung. Andere Modellteile lassen verschiedene Varianten zum Modellvergleich offen, wie z.B. der CERT bedingte Ceramid Transportmechanismus. B & C: Schematische Darstellung der zwei Modellalternativen Ceramid Shuttle (B) und Swinging Neck (C). D: Western Blot Daten aus einem Störexperiment, in dem die PKD Aktivität mit einem Inhibitor geschwächt wurde. Aktive PI4KIIIβ wurde in einer Zeitreihe gemessen. Die Rohdaten werden quantifiziert und digitalisiert, bevor sie ins Modell einfließen. E: Legende.*

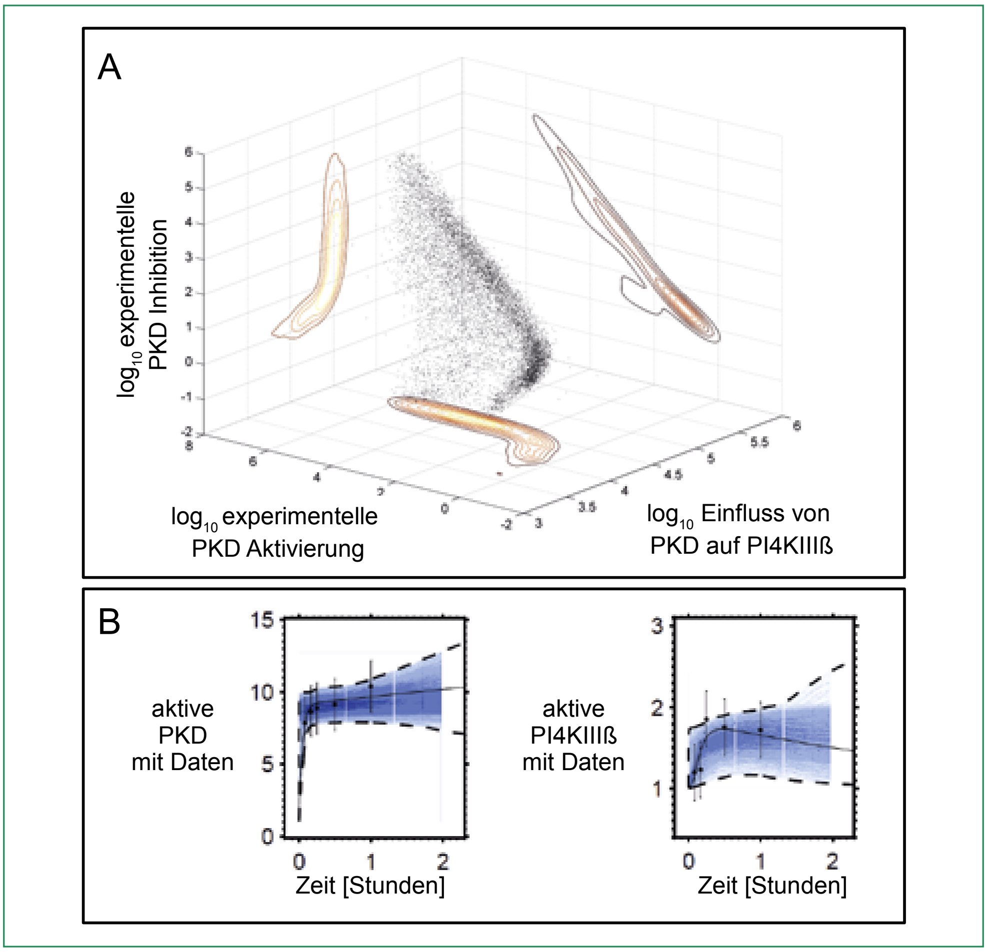
matisch dargestellt sind. Diese beschreiben jeweils die beiden Transporttheorien Ceramid Shuttle und Swinging Neck und können beide jeweils in das Kernmodell eingebettet werden. Die beiden daraus entstandenen Modelle unterscheiden sich nur in den Gleichungen, welche den Ceramid Transport beschreiben. Das Swinging Neck Modell umfasst hierbei 27 Parameter, während die Shuttle Modellvariante durch nur 26 Parameter bestimmt wird. Wir haben es also mit Modellen unterschiedlicher Komplexität und unterschiedlicher Anzahl freier Parameter zu tun, so dass ein reiner Datenfit keine verlässliche Aussage liefert.

**(04D)** zeigt einen exemplarischen Western Blot Datensatz, wie er zur Modellkalibrierung verwendet wurde. Aus den Bilddaten können je nach Färbungsintensität die jeweiligen Proteinkonzentrationen extrahiert werden. Durch die Analyse der Bilddateien mit von uns speziell für dieses Projekt entwickelten Algorithmen werden die Proteinkonzentrationen quantifiziert und so für einen Vergleich mit den entsprechenden Modellvariablen aufbereitet.

Die Relation von messbaren und nicht messbaren Modellvariablen lässt hierbei ein ausbalanciertes Konzept der datengetriebenen Modellierung erkennen: Mit fünf experimentellen Zugängen und sechs messbaren Variablen besitzt das Modell mit seinen insgesamt neun Zustandsvariablen eine gute Beobachtbarkeit.

## 4.2 ... über die Modellkalibrierung ...

Für die beiden Modellvarianten wurden mittels MCMC Sampling Stichproben aus der a-posteriori Verteilung generiert. **(05A)** zeigt exemplarisch die Stichprobe für drei Parameter des Swinging Neck Modells in logarithmischer Darstellung. Auf den Koordinatenachsen sind die Höhenlinien der aus der Stichprobe empirisch geschätzten Dichtefunktion dargestellt. Aus der Abbildung geht hervor, dass man mit den vor-



Sampling-basierte Parameterschätzung und simulationsbasierte Modellvorhersagen

*A: Darstellung einer Parameterstichprobe, die mit MCMC Sampling ermittelt wurde. Hier wird der Zusammenhang zwischen drei Parametern dargestellt: Die Wirkung von PKD auf PI4KIIIß und zwei experimentelle Zuflüsse ins Modell. Diese Punktwolken können komplexe Formen in hochdimensionalen Räumen annehmen. Um Zusammenhänge zwischen Parametern, wie z.B. Korrelationen, besser zu verstehen, werden sie auf niedriger dimensionale Unterräume projeziert. Die Gesamtstichprobe hatte hier 27 Parameter und daher auch 27 Dimensionen. B: Eine Modellvorhersage des Swinging Neck Modells zusammen mit Trainingsdaten. Die schwarze Linie stellt die Simulation mit dem Maximum a-posteriori Schätzer dar. Die Wahrscheinlichkeitsdichte ist durch die Intensität der Blaufärbung dargestellt. Über 99 Prozent aller Vorhersagen liegen innerhalb der gestrichelten Linie.*

**05**

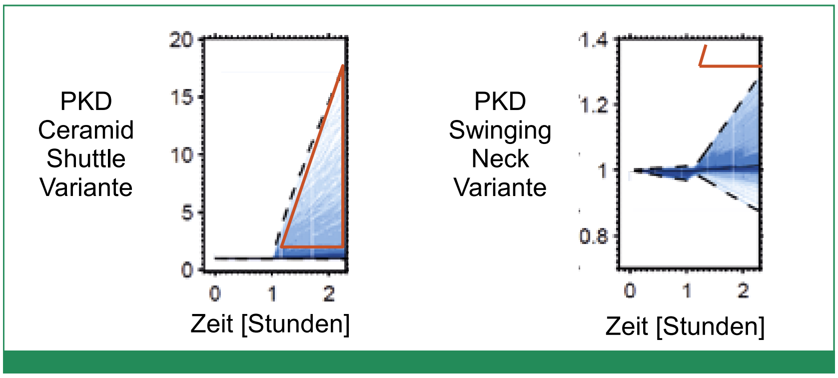
liegenden Daten noch weit davon entfernt ist, alle Parameterwerte genau angeben zu können. Der Bereich mit hoher a-posteriori Verteilung umfasst für zwei der drei Parameter noch mehrere Größenordnungen. Weiterhin erkennt man, dass die Verteilung dieser drei Parameter sehr von einer Normalverteilung abweicht und die Parameter hohe und teilweise nichtlineare Korrelationen aufweisen, wie es für Parameterschätzungen von Differenzialgleichungsmodellen häufig der Fall ist.

Die marginalen a-posteriori Verteilungen für Modellparameter von nicht direkt messbaren Variablen weisen meist eine noch größere Varianz auf. Es existieren also viele unterschiedliche Parametersätze, mit denen das Swinging Neck Modell die Daten etwa gleich gut reproduzieren kann, und zwischen denen man auf Basis der vorliegenden Daten nicht unterscheiden kann. Die Parameterverteilung des Shuttle Modells weist eine ähnlich starke Varianz auf.

Zusammenfassend lässt sich also erkennen, dass wir es hier mit einem sehr schlechtgestellten Optimierungsproblem zu tun haben, und globale sampling-basierte Ansätze für unsere Zwecke generell geeignet sind.

**(05B)** zeigt exemplarisch den Vergleich einer Variablen des kalibrierten Swinging Neck Modells mit einem zur Kalibrierung verwendeten Datensatz auf. Aus diesen Vergleichen lässt sich erkennen, dass die vorliegenden Messdaten von beiden Modellen gut beschrieben werden. Insbesondere lassen diese Daten noch keine signifikante Präferenz für eine der beiden Modellvarianten zu.

Ist das nun der Weisheit letzter Schluss? … Nun, mag vielleicht sein! Allerdings können wir uns an diesem Punkt unsere Modelle zu Nutze machen und modellgestützt nach geeigneten Experimenten suchen, mit denen sich vielleicht zwischen den beiden Hypothesen unterscheiden lässt. Wir werden sehen …



**06**

Vorhersage eines neuen hypothetischen Szenarios

*Das Swinging Neck und das Shuttle Modell treffen jeweils eine Vorhersage über dasselbe Experiment. Die rot umrandete Fläche zeigt, in welchem Bereich sich die Vorhersagen stark unterscheiden.*

Vorab sei aber noch Folgendes zur Laufzeit angemerkt: Die Simulation eines Experimentes mit einem dieser Modelle und einem Parametersatz dauert weniger als eine Zehntelsekunde. Dies mag zunächst schnell erscheinen, die Modellkalibrierung ist jedoch trotzdem sehr zeitintensiv: Um die hier dargestellten Parameterverteilungen zu erzeugen, mussten für jedes Modell vier unterschiedliche experimentelle Szenarien simuliert werden. Je Modell wurden, um eine repräsentative Stichprobe zu erhalten, sechs Millionen Parametersätze aus der a-posteriori Verteilung gezogen und ausgewertet, woraus sich insgesamt 48 Millionen Modellsimulationen ergaben. Dies entspricht 50–60 Stunden Rechenzeit auf einem Kern eines modernen Rechners.

### 4.3 ... zur Modellvorhersage

Unsere sampling-basierten Methoden erlauben es nun, modellbasiert neue Experimente vorzuschlagen, welche geeignet sind um zwischen den beiden Modellvarianten zu unterschieden. Hierzu wurden nun mit beiden Modellen Vorhersagen für neue Szenarien gemacht und dabei untersucht, für welche dieser Szenarien sich die Vorhersagen der beiden Modelle signifikant voneinander unterscheiden. Ein solcher Unterschied zeigte sich beispielsweise beim Vergleich des Anstiegs der PKD Gesamtproteinmenge innerhalb von zwei Stunden nach PKD Aktivierung **(06)**. Das Shuttle Modell schließt hierbei eine starke Erhöhung der PKD Gesamtmenge bis zu einem Faktor 10–12 nicht aus, während das Swinging Neck Modell nur leichte Änderungen in der PKD Gesamtmenge vorhersagt. Der rot eingerahmte Bereich verdeutlicht das Gebiet, in dem sich die Modellvorhersagen stark unterscheiden. Die Chance, eine der beiden Transporttheorien durch Messungen dieser Variablen im biologischen System innerhalb des relevanten Zeitbereichs zu bekräftigen, ist hier demnach sehr hoch. Diese und weitere Experimente sind momentan in Planung, und wir sind gespannt ob sich eine der Hypothesen nach einem weiteren Durchlaufen des systembiologischen Zirkels durchsetzen wird.

## 5. Zu guter Letzt ein paar Schlussworte

In diesem Artikel haben wir die Rolle der Simulation für statistische sampling-basierte Verfahren zur Kalibrierung und Analyse systembiologischer Modelle diskutiert. Bayes'sche Lernverfahren zur Modellkalibrierung, wie sie hier vorgestellt wurden, liefern eine statistisch konsistente Beschreibung des Modells, in dem die Variablen und die Parameter als Zufallsvariablen interpretiert werden. Da man es bei der Parameterschätzung meist mit schlechtgestellten inversen Problemen zu tun hat, sind globale sampling-basierte Verfahren das passende Werkzeug. Sie liefern neben den optimalen Lösungen auch Informationen über Unsicherheiten in den Parametern und auch in Modellvorhersagen, und machen damit eine fundierte Evaluierung des Modells und auch Vergleiche unterschiedlich komplexer Modelle möglich. Hier wurde die Mächtigkeit dieser Methoden anhand eines Beispiels zur modellbasierten Untersuchung des Transportmechanismus des Proteins CERT demonstriert.

Gerade für nichtlineare dynamische Modelle ist der Rechenaufwand für solche globalen Verfahren allerdings extrem hoch, so dass ihre Anwendung, abhängig von der Anzahl der Parameter und den Eigenschaften der Zielfunktion, bisher auf Modelle mittlerer Größe beschränkt ist. Effiziente numerische Simulationsverfahren sowie Methoden zur Modellreduktion spielen hierbei für die Übertragbarkeit auf größere Systeme eine wesentliche Rolle.

Eine Zukunftsvision im Rahmen des Exzellenzclusters SimTech der Universität Stuttgart ist die Nutzbarmachung systembiologischer und biomechanischer Erkenntnisse in der Medizin, wie beispielsweise die modellgestützte Optimierung der Behandlungen von Patienten mit Medikamenten. Diese Zielsetzung erfordert über die Grundlagenforschung auf Einzelzellebene hinausgehende Multi-Skalen-Ansätze, die mehrere Längen- und auch Zeitskalen umfassen. Man kann sich leicht vorstellen, dass der Simulationstechnik sowohl bei der Erstellung und Simulation solcher Modelle als auch für dessen Kalibrierung eine nicht zu unterschätzende Rolle zukommt. •

*Patrick Weber, Karsten Kuritz,*
*Andrei Kramer, Frank Allgöwer, Monilola Olayioye,*
*Angelika Haußer und Nicole Radde*

## ZUSAMMENFASSUNG

*Die Systembiologie ist ein noch recht junges Forschungsgebiet, welches sich zu Beginn des 20. Jahrhunderts an der Schnittstelle zwischen Biologie, Systemtheorie und Simulationstechnik entwickelt hat. Sowohl enorme Fortschritte im experimentellen Bereich als auch immer leistungsfähigere Computer ermöglichen heute erstmals eine ganzheitliche Betrachtung biologischer Systeme.*

*Die Simulation quantitativer Modelle dieser Systeme liefern interessante neue Hypothesen, und mathematische Analysen helfen aktuelle Fragestellungen in allen Forschungsbereichen der Biologie zu adressieren – von der Entwicklung neuer Medikamente bis hin zur Prozessoptimierung in der Biotechnologie.*

*Dieser Artikel widmet sich der Rolle der Simulation speziell für statistische sampling-basierte Ansätze zur Modellkalibrierung und der Generierung von Hypothesen. Die Simulationstechnologie nimmt bei den hier vorgestellten Methoden eine zentrale Stellung zwischen experimenteller Datenerhebung und theoretischer Systemanalyse ein. Das Potenzial dieser Methoden wird beispielhaft an einem systembiologischen Kooperationsprojekt zwischen dem Institut für Systemtheorie und Regelungstechnik (IST) und dem Institut für Zellbiologie und Immunologie (IZI), in dem wir molekulare Schlüsselprozesse der Proteinsekretion in Säugetierzellen untersuchen, demonstriert.*

## Anmerkung

Die in diesem Artikel dargestellten Analyseergebnisse und deren Interpretation stellen aktuelle Zwischenergebnisse eines laufenden Projektes dar und wurden anhand von vorläufigen, teilweise noch nicht evaluierten Datensätzen erstellt.

## Referenzen

**Publikationen Methoden – Institut für Systemtheorie und Regelungstechnik**

- Weber P, Kramer A, Dingler C, Radde N (2012). Trajectory-oriented Bayesian experiment design versus Fisher A-optimal design: an in depth comparison study. Bioinformatics 28(18), i535–i541.
- Thomaseth C, Weber P, Hamm T, Kashima K, Radde N (2013). Modeling sphingomyelin synthase 1 driven reaction at the Golgi apparatus can explain data by inclusion of a positive feedback mechanism, J Theor Biol 337, 174–180.

**Publikationen Biologie – Institut für Zellbiologie und Immunologie**

- Olayioye MA, Hausser A (2012). Integration of non-vesicular and vesicular transport processes at the Golgi complex by the PKD-CERT network. Biochim Biophys Acta 1821(8), 1096–103.
- Florin L, Pegel A, Becker E, Hausser A, Olayioye MA, Kaufmann H (2009). Heterologous expression of the lipid transfer protein CERT increases therapeutic protein productivity of mammalian cells. J Biotechnol 141(1–2), 84–90.

**Basisliteratur Systembiologie**

- Gelman A, Carlin J, Stern H, Rubin D (2003). Bayesian Data Analysis, Second Edition, Taylor & Francis.
- Klipp E, Liebermeister W, Wierling C, Kowald A, Lehrach H, Herwig R (2011). Systems Biology, Wiley.

## DIE AUTOREN | 1

### FRANK ALLGÖWER

*Frank Allgöwer ist Professor für Systemtheorie und Regelungstechnik und Leiter des gleichnamigen Instituts an der Universität Stuttgart. Er hat in Stuttgart Technische Kybernetik und an der University of California at Los Angeles Angewandte Mathematik studiert und promovierte in der Fakultät Verfahrenstechnik der Universität Stuttgart. Vor seiner Berufung nach Stuttgart im Jahr 1999 hatte er eine Professur für Nichtlineare Systeme im Departement Elektrotechnik der ETH Zürich. Längere Forschungsaufenthalte brachten Frank Allgöwer an das NASA Ames Research Center, das California Institute of Technology, die University of California at Santa Barbara und zur Firma DuPont in Wilmington, Delaware. Sein Hauptarbeitsgebiet ist die Entwicklung und Anwendung systemtheoretischer Methoden zur Analyse und Regelung dynamischer Systeme.*

## DIE AUTOREN | 2

### NICOLE RADDE (3VR)

*Nicole Radde hat an der Technischen Universität Darmstadt Physik studiert, an der Universität zu Köln am Zentrum für Angewandte Informatik promoviert, und war anschließend als Postdoktorandin am Institut für Medizinische Informatik, Statistik und Epidemiologie der Universität Leipzig tätig. Seit Oktober 2008 arbeitet sie als Juniorprofessorin für „Systemtheorie in der Systembiologie“ am Institut für Systemtheorie und Regelungstechnik der Universität Stuttgart. Ihre Forschungsinteressen liegen im Bereich der statistischen sampling-basierten Lernverfahren zur Parameterschätzung und der systemtheoretischen Analyse intrazellulärer Netzwerke.*

### PATRICK WEBER (2VL)

*Patrick Weber hat an der Albert-Ludwigs-Universität Freiburg Mikrosystemtechnik studiert und 2009 sein Diplom erhalten. Im Hauptstudium und in seiner Abschlussarbeit vertiefte er sich in biomedizinischer Mikrofluidik. Seit Anfang 2010 ist er wissenschaftlicher Mitarbeiter am Institut für Systemtheorie und Regelungstechnik an der Universität Stuttgart und bearbeitet in der Systembiologie-Gruppe das Projekt: Modelling of the trans-Golgi network key player interactions.*

### KARSTEN KURITZ (L)

*Nach Studien an der Universität Stuttgart und der University of Linköping erlangte Karsten Kuritz im Jahr 2012 den Abschluss Diplom-Biologe (technisch orientiert). Während des Studiums verbracht er im Jahr 2010 einen 6-monatigen Forschungsaufenthalt bei der Firma Merrimack Pharmaceuticals in Cambridge, Massachusetts. Seit November 2012 ist Karsten Kuritz wissenschaftlicher Mitarbeiter am Institut für Systemtheorie und Regelungstechnik. Sein Forschungsschwerpunkt liegt in der Anwendung von systemtheoretischen Methoden auf Fragestellung der molekularen Zellbiologie.*

### ANDREI KRAMER (2VR)

*Andrei Kramer ist wissenschaftlicher Mitarbeiter in der Systembiologie-Gruppe im Institut für Systemtheorie und Regelungstechnik der Universität Stuttgart. Sein Forschungsgebiet ist die Untersuchung und Weiterentwicklung von Parameter-Sampling Methoden in der Modellierung von biologischen Systemen. Er hat an der Humboldt Universität zu Berlin Physik studiert (Diplom 2008).*

### MONILOLA OLAYIOYE (3VL)

*Prof. Monilola Olayioye absolvierte von 1991 bis 1996 ein Studium der Biotechnologie an der Technischen Universität Braunschweig und promovierte von 1997 bis 2000 am Friedrich Miescher Institut in Basel, Schweiz in der Zellbiologie. Nach einem dreijährigen durch ein EMBO Stipendium finanziertem Postdokaufenthalt am Walter and Eliza Hall Institut in Melbourne, Australien baute Frau Olayioye 2004 im Rahmen des SFB495 „Topologie und Dynamik von Signalprozessen“ eine unabhängige Nachwuchsgruppe an der Universität Stuttgart auf. Die Arbeitsgruppe von Frau Olayioye ist am Institut für Zellbiologie und Immunologie angesiedelt und beschäftigt mit der Aufklärung von molekularen Signalwegen, die der deregulierten Proliferation, Migration und Invasion von Tumorzellen zugrunde liegen. 2011 nahm Frau Olayioye den Ruf auf eine Heisenberg-Professur für Molekulare Tumorzellbiologie an der Universität Stuttgart an.*

### ANGELIKA HAUSSER (R)

*Angelika Hausser ist seit 2003 Arbeitsgruppenleiterin und seit 2010 Akademische Rätin am Institut für Zellbiologie und Immunologie der Universität Stuttgart. Von 1991 bis 1997 studierte sie technische Biologie an der Universität Stuttgart, wo sie 2000 mit einer Arbeit über die Regulation der Proteinkinase D in T-Lymphozyten promoviert wurde. Dem Eliteförderprogramm für Postdoktoranden der Landesstiftung Baden-Württemberg gehörte sie von 2002 bis 2004 an, von 2008 bis 2013 war sie Mitglied im WIN-Kolleg der Heidelberger Akademie der Wissenschaften. Derzeit beschäftigt sie sich mit Signalwegen und molekularen Mechanismen der Proteinsekretion und Zellmigration.*

**Kontakt**
Universität Stuttgart
Institut für Systemtheorie
und Regelungstechnik
Pfaffenwaldring 9
D–70550 Stuttgart
Tel: +49 (0) 711/685-67734
Fax: +49 (0) 711/685-67735
E-Mail:
sekist@ist.uni-stuttgart.de
Internet:
http://www.simtech.
uni-stuttgart.de