



Doctoral Thesis

Forward design of a complex biochemical system

Author(s):

Hold, Christoph

Publication Date:

2015

Permanent Link:

<https://doi.org/10.3929/ethz-a-010432428> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

DISS. ETH NO. 22521

FORWARD DESIGN OF A COMPLEX BIOCHEMICAL SYSTEM

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by
CHRISTOPH HOLD
Dipl.-Ing., Universität Stuttgart

born on 23.06.1980
citizen of Germany

accepted on the recommendation of
Prof. Dr. Sven Panke (ETH Zürich, Switzerland), examiner
Prof. Dr. Jörg Stelling (ETH Zürich, Switzerland), co-examiner
Prof. Dr. Philipp Rudolf von Rohr (ETH Zürich, Switzerland), co-examiner

2015

Abstract

Synthetic Biology aims at establishing engineering's construction process within the field of biology: design, implementation and validation of new biological systems from scratch that have not been established in nature yet. However, this construction process is limited by the availability of robust models that allow the application of forward engineering to the design of novel biological systems.

In this work an enzymatic reaction system for the production of dihydroxyacetone phosphate, a precursor to artificial monosaccharides, is constructed from purified enzymes. The backbone of the reaction system is made up of enzymes from glycolysis that allow the production of DHAP from glucose in upper glycolysis and the regeneration of the required cofactors in lower glycolysis, allowing the operation of the system with only catalytic amount of cofactors driven by the difference in free energies between glucose and the reaction products, either dihydroxyacetone phosphate and lactate or glycerol 3-phosphate and pyruvate.

Due to the size and complexity of the system consisting of 11 enzymes, some of which exhibit additional regulatory properties, a rational construction process has to be model-based and a mechanistic dynamic reaction model is formulated based on available mechanistic information on the enzymes in the literature. For the parameterization of the model the enzymatic system is implemented in a fully identified continuously operated enzyme membrane reactor whose reactor effluent is constantly conditioned for subsequent online mass spectrometry. Omission of additional compound separation steps such as high pressure liquid chromatography allows for an accurate tracking of dynamic concentration changes at a high data density, but generates the possibility of quantification errors through ion suppression and limits in compound identifiability. Therefore, extensive efforts were undertaken to increase the accuracy of measurements, including the implementation of an isotopologue standard for glycolytic substrates and intermediates.

By imprinting different feed- and pulse profiles on the reaction system in the reactor, this setup enabled comprehensive yet accurate access to the enzymatic machinery for manipulation and observation. Standard systems theory input functions like step, ramp, pulse and sine were implemented, allowing for recording of the enzymatic systems responses to defined excitation functions. The system was tested with the identification of a rate model for the first enzyme of the reaction system, glucokinase of *Escherichia coli*. Despite intensive experimental and modelling efforts, it was impossible to fit experimental data well based on the existing information on the enzyme's mechanism. Only once an additional inhibition by one of the reaction products, ADP, was integrated into the model of the reaction rate, satisfactory agreement between parameterized model and experimental data could be obtained by exploiting the experimental freedom enabled by the setup to design system perturbations to specifically address difficult aspects of model identifiability.

Then in a step-wise approach the total system of 11 enzymes with a total of 60 parameters was divided into more manageable subsystems with up to 23 parameters. For these subsystems, parameters were estimated from a limited number of perturbation experiments. This approach enabled the formulation of a fully

parameterized model of the reaction system, which was able to qualitatively and quantitatively accurately reproduce the dynamic concentration changes used for its identification. Remarkably, the vast majority of the parameters was within biologically meaningful ranges, with only a minority of parameters still suffering from issues of identifiability.

The model was also used to generate predictions about optimizing product formation by re-distributing enzyme concentrations or while reducing the concentration of cofactors. Remarkably, the predictions were well confirmed by the corresponding experiments set up according to the model predictions. In terms of optimizing enzyme distributions, the model predictions were even valid outside the range of enzyme and substrate concentrations that had been used to parameterize the model.

Zusammenfassung

Das Ziel der synthetischen Biologie ist es, innerhalb des Forschungsbereichs der Biologie Konstruktionsprozesse analog zu etablierten Ingenieursdisziplinen einzuführen, welche die Formgebung, die Implementierung wie auch die Validierung neuartiger biologischer Systeme beinhalten, die keine natürlichen Vorlagen besitzen. Bisher wird die Etablierung solcher Prozesse jedoch durch die mangelnde Verfügbarkeit hochwertiger und robuster Modelle behindert, welche die gerichtete rationale Konstruktion solcher neuartiger biologischer Systeme erst ermöglichen.

Im Zuge dieser Dissertation wurde ein Reaktionssystem aus aufgereinigten Enzymen für die Produktion von Dihydroxyacetonphosphat, einer Vorstufe künstlicher Monosaccharide, geschaffen. Das Grundgerüst dieses neuartigen Reaktionssystems basiert auf Enzymen welche die Umwandlung von Glukose in DHAP innerhalb der oberen und die Regeneration benötigter Kofaktoren innerhalb der unteren Glykolyse katalysieren. Diese Auslegung erlaubt es das Reaktionssystem mit katalytischen Kofaktorkonzentrationen nur durch die Differenz der Gibbs'schen freien Energie zwischen Glukose und den Reaktionsprodukten, entweder Dihydroxyacetonphosphat und Laktat oder Glycerol 3-phosphat und Pyruvat, zu betreiben.

Die Notwendigkeit, dieses spezielle Reaktionssystem im Zuge eines modellbasierten rationalen Gestaltungsprozesses auszulegen, war bedingt durch die Größe und Komplexität des Systems, welches in seiner endgültigen Form elf Enzyme beinhaltet, von denen einige zusätzlich zu ihrer katalytischen Aktivität weitere regulatorische Eigenschaften besitzen. Daher wurde zum Zwecke dieser Dissertation ein dynamisches mechanistisches Modell auf Grundlage von in der Fachliteratur verfügbaren enzymatischen Wirkmechanismen erstellt. Um eine vollständige Identifikation der Modellparameter zu ermöglichen, wurde zudem das gesamte enzymatische Reaktionssystem physisch innerhalb eines vollständig charakterisierten Membrandurchflussreaktors realisiert. Zur dynamischen Echtzeitmessung des Ausflusses des Reaktorsystems wurde dieser kontinuierlich aufbereitet und einem Massenspektrometer zugeführt. Durch die so ermöglichte Umgehung zusätzlicher Separationsschritte wie beispielsweise Hochleistungsflüssigkeitschromatographie wurde eine akkurate und zeitlich hochauflösende Bestimmung der sich dynamisch ändernden Stoffkonzentrationen ermöglicht. Um der Gefahr möglicher Messfehler und -ungenauigkeiten, beispielsweise hervorgerufen durch Ionensuppression oder technischen Grenzen in der Stoffidentifizierbarkeit, weitestgehend auszuschließen, wurden intensive Anstrengungen unternommen. Dabei Hervorzuheben ist unter anderem die Etablierung eines Isotopologenstandards für glykolytische Substrate und Zwischenprodukte.

Durch die Einspeisung verschiedener dynamische Zufluss- und Pulsprofile in das Reaktionssystem ermöglichte der Versuchsaufbau einen umfassenden und akkuraten Zugang für die Beeinflussung und die Messung der enzymatischen Maschinerie. Als Eingangsfunktionen wurden etablierte systemtheoretische Profile wie Sprung-, Impuls- und Sinusfunktionen implementiert, welche die Messung der dynamischen Systemantwort auf definierte Anregungen ermöglichten. Der gesamte

Prozess, bestehend aus dem Versuchsaufbau sowie der modellbasierten Interpretation der Messdaten, wurde anhand der Identifikation der Ratenkonstanten für das erste Enzym des Reaktionssystems, der Glucokinase von *Escherichia coli*, getestet. Dabei stellte sich jedoch trotz intensivem experimentellen und Modellierungsaufwands heraus, dass ein ausschliesslich auf bereits vorhandenen Informationen zu den enzymatischen Wirkmechanismen basierender Erklärungsansatz nicht mit experimentellen Messdaten in Einklang zu bringen war. Durch Ausnutzung der durch den Versuchsaufbau gegebenen experimentellen Freiheiten, die die Gestaltung spezifischer dynamischer Systemanregungen zur Analyse komplexer Aspekte im Bereich der Modellidentifizierung ermöglichen, konnte gezeigt werden, dass nur die Integration eines zusätzlichen Produktinhibitionsmechanismus durch ADP in das Modell zu einer zufriedenstellenden Übereinkunft mit den experimentellen Daten führt.

In den weiteren Bemühungen in Richtung einer ganzheitlichen Identifikation wurde das gesamte Reaktionssystem mit elf Enzymen und 60 Parametern schrittweise in mehr handhabbare Teilsysteme mit bis zu 23 Parametern aufgetrennt. Die Parameter dieser Teilsysteme wurden anhand von einer limitierten Zahl von Perturbationsexperimenten bestimmt. Dieser modulare Ansatz ermöglichte es, ein Modell des Gesamtsystems mit einem vollständig identifizierten Parametersatz zu formulieren, welches alle für die Identifikation benutzten Konzentrationsdynamiken sowohl qualitativ wie auch quantitativ akkurat reproduzieren konnte. Beachtenswerter Weise befindet sich die große Mehrzahl der so identifizierten Parameter innerhalb biologisch sinnvollen Wertebereiche, so dass nur einige wenige Parameter noch unter einer mangelnden Identifizierbarkeit leiden.

Des Weiteren wurden anhand des Modells Voraussagen zur möglichen Optimierung der Produktformierung durch eine veränderte Verteilung der Enzymkonzentrationen, auch unter Verringerung der Kofaktorkonzentrationen, getroffen. Experimentelle Validierungen ergaben eine gute Übereinstimmung dieser so getroffenen Voraussagen mit tatsächlich erreichbaren Steigerungen in der Produktformierung. Diese modellbasierten Voraussagen stimmten zudem sogar für Enzym- und Substratkonzentrationen ausserhalb des für die Identifikation genutzten Wertebereichs mit den entsprechenden Messdaten überein.