Learning Modules for Computational Systems Biology

The main features of BioSym™

BioSym™ is an interactive, blended learning bio-modeling training course.

- BioSym™ addresses important questions relevant to the emerging field of Systems Biology
- Real life data are used for model building
- The modular structure allows one to incorporate individual models into science curricula at institutions with different needs
- It is of interest to institutions that do not have the full competence to offer courses in Computational Systems Biology

- The concept is based on tested didactical scenarios
- Experienced teachers, scientists and e-learning experts, are involved in teaching BioSym™ courses
- It incorporates webbased training, team work and e-collaboration
- It promotes time independent active participation (distance learning)
- It offers links to data bases relevant for modeling topics

- BioSym™ is based on widely used mathematical software packages
- It makes use of OLAT for the organization of courses
- Lectures can be streamed with specialized lecture recording software and presented in OLAT

- BioSym™ contains modules which can be used in basic as well as in advanced courses
- Learners acquire skills which make BioSym™ useful for “marketable” professional advancement
- New contents can be added at any time which assures sustainable usability for a long period of time

Examples from BioSym

Poster presentations by members of the BioSym™ group

- BioSym™ - A Systems Biology Learning Network
- A computational modeling approach to systems biology
- Analysis of complex biological systems through computational mathematical modeling
- Bio-Thermodynamics: Understanding glycolysis with quantum chemistry
- Modeling of metabolic networks: A computational approach to functional systems biochemistry and metabolic engineering
- Selection and adaptation in microbial communities: A computational modeling approach to ecosystem complexity
- Eco-genomics of rumen communities: How similar, in an evolutionary sense, are cellulases from different rumen microbes?

Kurt Hanselmann (kurt.hanselmann@hispeed.ch), Christoph Fuchs (christoph.fuchs@mnf.uzh.ch), Stefan Schafroth (stefan_erst@gmx.ch), Maja A. Lazzaretti-Ulmer (maja.lazzaretti@math.uzh.ch), Roman Kälin (rkaelin@amath.uzh.ch), Stefan Brammertz (sbrammertz@gmail.com), Hans Ulrich Suter (suter@physik.uzh.ch)

BioSym™ - Computational Systems Biology
biosym@mnf.uzh.ch or i-research.training@hispeed.ch

BioSym™

A Systems Biology Learning Network produced by the BioSym™ team, http;//www.biosym.uzh.ch**

swiss | i-research & training

A Systems Biology Learning Network

Teaching / Learning Objectives

- Supporting mathematical and quantitative approaches in the life sciences
- incorporating physical and chemical principles into biological understanding
- familiarizing students with the power of modeling biological processes and systems
- promoting conceptual teaching and learning of Systems Biology
- Training the use of MATLAB and its tool boxes

Learning approach

- blended learning
- modular design
- encouraging continuing education
- facilitating distance learning

Information management

Find best means of professional information dissemination

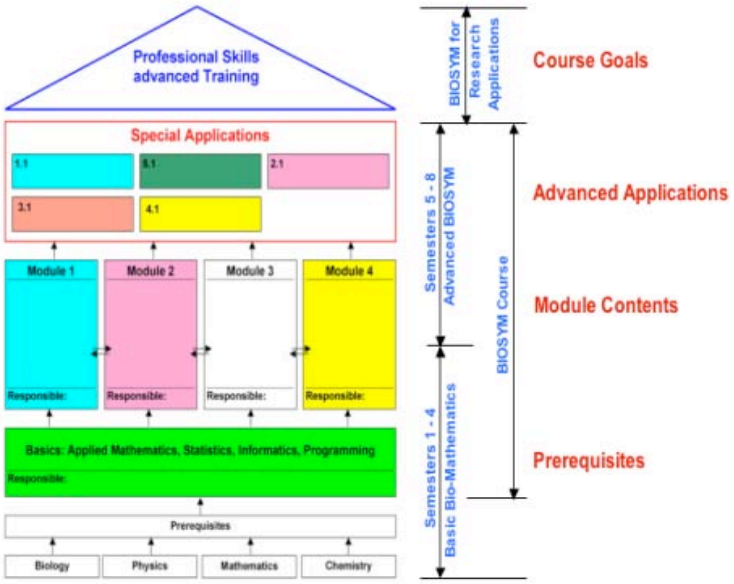
Instruct access to information in libraries and data banks

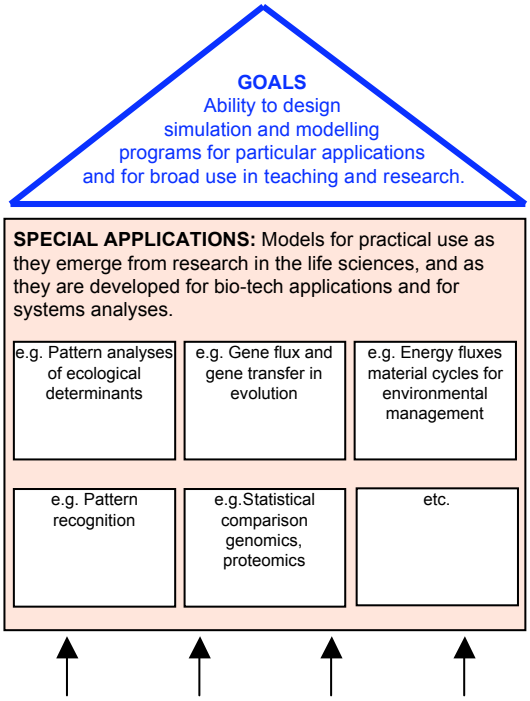
Offer information processing / evaluating techniques

Suggest efficient teaching skills and learning strategies

Validate teaching approaches towards learning success

Curricular Integration



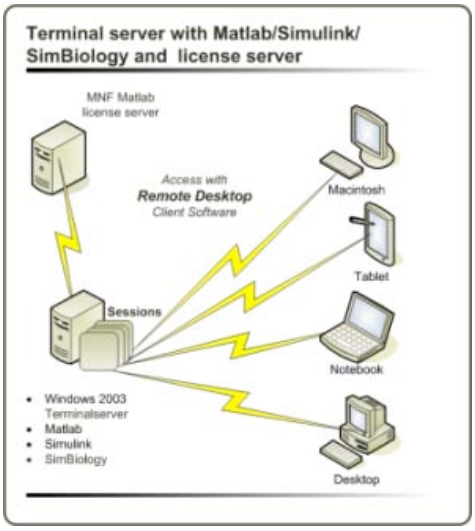


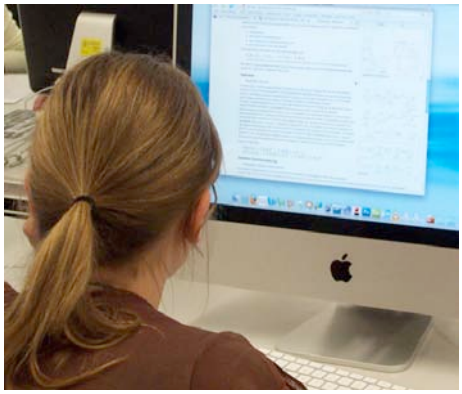
Module design

- BioSym™ introduces key molecular, cellular, organismic and systems concepts
- BioSym™ emphasizes the quantitative and integrative nature of biological problems

Learning Environment

- Interactive modules via OLAT
- Matlab Classroom Licenses
- Microsoft Terminal Server
- Recorded Lessons on Flash Media Server





**** Network Partners**: University Zurich, ETH Zurich, ZH Winterthur, University Fribourg, Ruhr University Bochum, WHO Geneva, Roche Basel, University Hospital Basel. **Collaborators UZH:** Barbour Andrew D. math., Brammertz Stefan biol., Fuchs Christoph[1] inform., Hanselmann Kurt[2] biol., Heinzmann

A computational modeling approach to systems biology

Christoph Fuchs (christoph.fuchs@mnf.uzh.ch) and Kurt Hanselmann (kurt.hanselmann@hispeed.ch), BioSym™ - Computational Systems Biology, Institute of Mathematics, University of Zürich, Winterthurerstr. 190, 8057 Zürich

Today it has become essential to employ mathematical models as research tools at all levels of biology. BioSym is a compilation of interactive models which can be used to study biological systems quantitatively, from the molecular to the ecosystem level. The models are based on biological and physicochemical principles which can be expressed with mathematical algorithms. They are offered under http://www.biosym.unizh.ch/index.php. BioSym contains classical deterministic models and more complex stochastic ones (e.g. epidemics, metabolic networks, gene regulation and metabolic control, physiology, gene/protein evolution etc.). On an advanced level, it introduces models which can assist users in designing quantitative experiments with proper boundary conditions and handling large data sets. Systems biology with BioSym is a logical step towards synthesizing details and fragments of knowledge into a more holistic view of biology, and it can serve as a motivation to deal with the complexity inherent to many biological systems.
Courses which are offered by the BioSym team introduce users to model building, show them how to design mathematical models and train them how to use simulations. The learning modules are primarily based on MATLAB and its toolboxes. Many models contain a Java Applet or a Flash animation to illustrate the details of the background.

BioSym™

A Systems Biology Learning Network produced by the BioSym™ team, http;//www.biosym.uzh.ch**

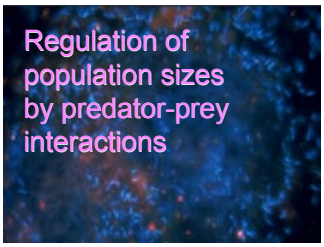
swiss | i-research & training

Computational modeling in systems biololgy

Systems biology of the rumen

Objectives

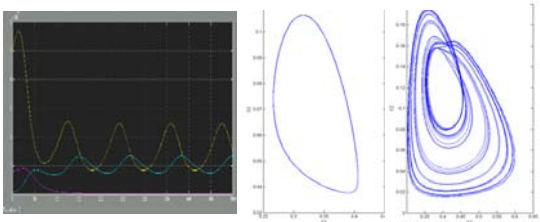
- The BioSym™ **rumen module** consists of interactive mathematical simulation models which allow the student to learn aspects of systems biology from the molecular to the ecosystem level.
- Students get introduced to model building and testing, to using data bases and to analyzing modeling outputs.
- Modeling with BioSym™ employs MATLAB and its tool boxes.
- Training is based on blended-learning scenarios; it incorporates web-based training and interaction, and it supports time independent learning.



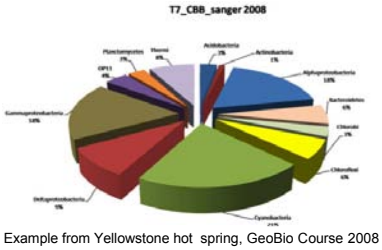
The BioSym™ rumen module contains deterministic and statistical models as well as more complex stochastic ones. On an advanced level, it introduces models which can assist users in designing quantitative experiments with proper boundary conditions and handling large data sets.

$$dX_1/dt = a \cdot X_1 - b_{12} \cdot X_1 \cdot X_2 - b_{13} \cdot X_1 \cdot X_3$$
$$dX_2/dt = b_{21} \cdot X_1 \cdot X_2 - a_2 \cdot X_2$$
$$dX_3/dt = b_{31} \cdot X_1 \cdot X_3 - a_3 \cdot X_3$$

X_1 = prey (yellow), X_2 = predator a (blue), X_3= predator b (purple)

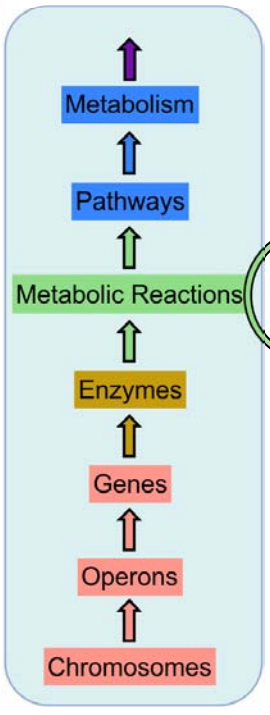


Analyzing metagenomic data from 454 sequencing



Example from Yellowstone hot spring, GeoBio Course 2008

Contents of the rumen module



Interactions
Propagation
Growth

Metabolic networks

Products

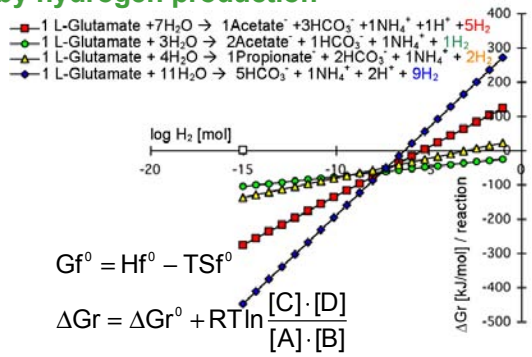
Bio-
Thermodynamics

Educts

Enzyme kinetics

Metagenomics
Genomics
Phylogeny
Gene transfer
Defining
ORFs, regulators
Gene arrangement
Sequence analysis

Regulation of reaction thermodynamics by hydrogen production



$$Gf^0 = Hf^0 - TSf^0$$

$$\Delta Gr = \Delta Gr^0 + RT\ln\frac{[C]\cdot[D]}{[A]\cdot[B]}$$

Phylogenetic relationship of hydrolases within rumen diversity

Regulation of Phosphofructokinase

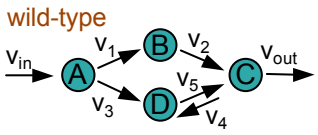
Fructose 6-phosphate + MgATP
→ Fructose 1,6 bisphosphate + MgADP

$$v_{PFK} = \frac{v_{PFK\,max}}{N_{PFK}} \left\{ \frac{\frac{[F6P]}{K_{F6P}}}{1+\frac{[F6P]}{K_{F6P}}} \cdot \frac{\frac{[Mg][ATP]}{K_{MgATP}}}{1+\frac{[Mg][ATP]}{K_{MgATP}}} \right\}$$

Enzymes: Phylogenetic origin

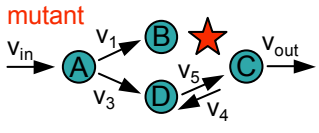
Behavior of knock-out mutants

wild-type



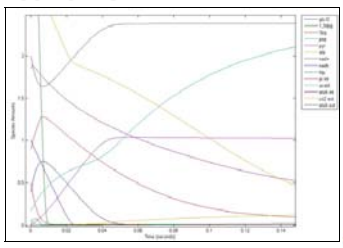
	v_1	v_2	v_3	v_4	v_5	v_{in}	v_{out}
A	-1	0	-1	0	0	1	0
B	1	-1	0	0	0	0	0
C	0	1	0	-1	1	0	-1
D	0	0	1	1	-1	0	0

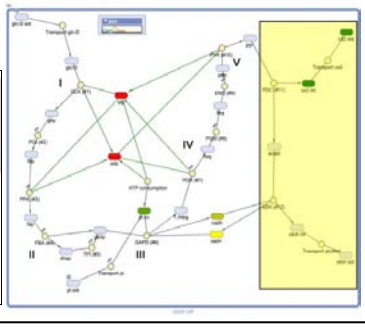
mutant



	v_1	v_2	v_3	v_4	v_5	v_{in}	v_{out}
A	-1	0	-1	0	0	1	0
B	1	0	0	0	0	0	0
C	0	0	0	-1	1	0	-1
D	0	0	1	1	-1	0	0

Simulating metabolic networks





** **Network Partners**: University Zurich, ETH Zurich, ZH Winterthur, University Fribourg, Ruhr University Bochum, WHO Geneva, Roche Basel, University Hospital Basel. **Collaborators UZH:** Barbour Andrew D. math., Brammertz Stefan biol., Fuchs Christoph[1] inform., Hanselmann Kurt[2] biol., Heinzmann Dominik math., Kälin Roman math., Lazzaretti Maja biol., Schafroth Stefan phys., Suter Hans Ulrich chem.

[1] biosym@mnf.uzh.ch, [2] i-research.training@hispeed.ch

Analysis of complex biological systems through computational mathematical modeling

Stefan Schafroth (stefan_erst@gmx.ch) and Kurt Hanselmann (kurt.hanselmann@hispeed.ch), BioSym - Computational Systems Biology, Institute of Mathematics, University of Zürich, Winterthurerstr. 190, 8057 Zürich

Studies dealing with the regulation of metabolic or hereditary processes in a cell, or with the mode of action of a drug in an organ or the behavior of organisms in communities and their responses to ecosystem determinants are often very complex processes. Mathematical approaches allow one to reduce the complexity of biological systems to understandable models and to describe processes and interactions quantitatively. However, every model is an idealization of the real world; models describe only those mechanisms that contribute essentially to observed or postulated phenomena. Mathematical models require that either well defined data sets are available from the literature or that unknown model parameters can be estimated from experience or expert knowledge. Another reason for applying computational modeling in biology is the generation and validation of hypotheses. A well constructed model can lead to predictions, which can then be tested experimentally. Deviations between the predictions and the actual observation can lead the investigator to improve the model and to design new experiments.
The poster presents an overview of the modeling workflow, it summarizes mathematical approaches for statistical significance tests, time series analysis as well as deterministic and stochastic kinetic models. They are illustrated with examples from different fields taken from BioSym, a Systems Biology Modeling Network.

BioSym™

A Systems Biology Learning Network produced by the BioSym™ team, http;//www.biosym.uzh.ch**

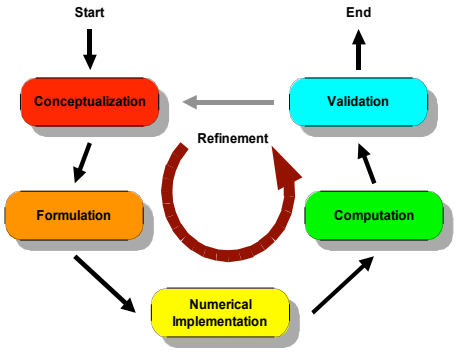
swiss | i-research & training

Analysis of Complex Biological Systems

Mathematical modeling in biology: 3 good reasons

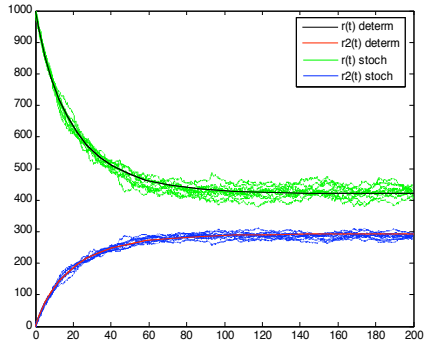
- Managing complexity and handling uncertainty: A model is always an idealization of the real world using only well defined input data.
- Modeling requires abstraction: The model describes only those underlying mechanisms that contribute most strongly to the observed phenomenon.This results in a reduction of complexity.
- Generation and validation of hypotheses: A good model can produce observable predictions. Deviations of predictions from actual observations can lead to model improvement.

A modeling workflow consists of five stages



In the refinement process these stages are repeatedly executed in a virtually never-ending process which generates models of increasing generality and validity.

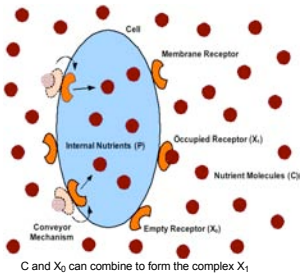
The use of deterministic and stochastic algorithms



LDE: Stable cycles with period k. The red line represents the trajectory (time course) of the system in the phase plane.

Models can illustrate simple relationships

Example: How bacteria consume substrates. Application of a rate flow model.

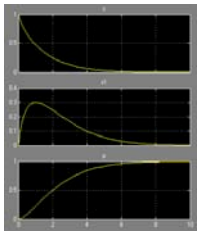


C and X_0 can combine to form the complex X_1

$$C + X_0 \xrightarrow{k_1} X_1 \quad ; \quad X_1 \xrightarrow{k_{-1}} C + X_0 \quad ; \quad X_1 \xrightarrow{k_2} P + X_0$$

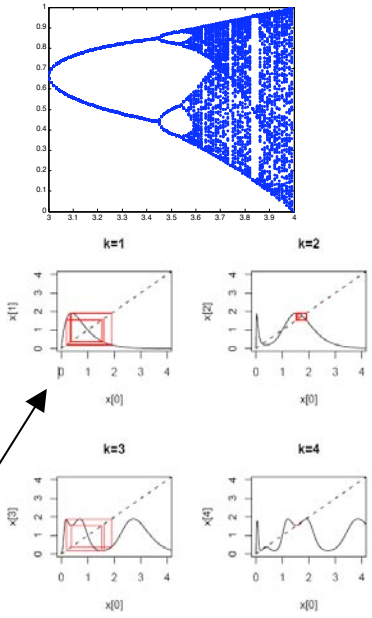
$dC/dt = -k_1C^*X_0 + k_{-1}X_1$
$dX_0/dt = -k_1C^*X_0 + k_{-1}X_1 + k_2X_1$
$dX_1/dt = k_1C^*X_0 - k_{-1}X_1 - k_2X_1$
$dP/dt = k_2X_1$

C external nutrients
X_0 unoccupied receptor
X_1 occupied receptor
P internal nutrients
X_0+X_1 = constant



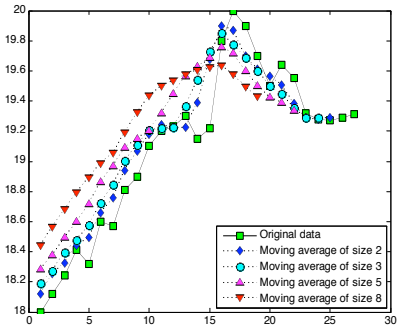
Simple models can show complex behaviour

In 1976 the Australian theoretical ecologist Robert May showed that simple first order difference equations can have very complicated or even unpredictable dynamics. The Logistic Difference Equation (**LDE**) is a model to explore the route into chaotic behaviour. The route to chaos starts with period doublings.



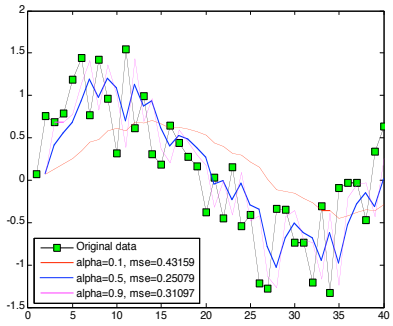
Basic techniques for time series analysis

Time series data often arise when monitoring physical processes. Time series analysis accounts for the fact that data points taken over time may have an internal structure (such as auto-correlation, trend or seasonal variation) that should be accounted for.

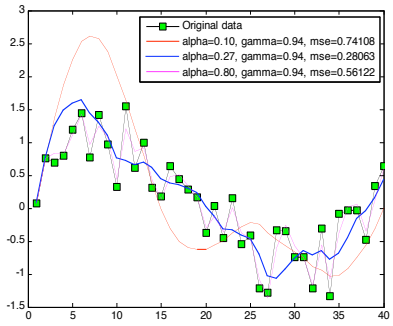


Exponential smoothing

Exponential smoothing assigns exponentially decreasing weights as the observations get older. This is in contrast to single moving averages where past observations are weighed equally. Exponential smoothing is a very popular scheme to produce a smoothed time series.



Double exponential smoothing uses two constants and is better at handling trends.



** **Network Partners**: University Zurich, ETH Zurich, ZH Winterthur, University Fribourg, Ruhr University Bochum, WHO Geneva, Roche Basel, University Hospital Basel. **Collaborators UZH:** Barbour Andrew D. math., Brammertz Stefan biol., Fuchs Christoph[1] inform., Hanselmann Kurt[2] biol., Heinzmann Dominik math., Kälin Roman math., Lazzaretti Maja biol., Schafroth Stefan phys., Suter Hans Ulrich chem.

[1] biosym@mnf.uzh.ch, [2] i-research.training@hispeed.ch

Bio-Thermodynamics: Understanding glycolysis with quantum chemistry

Hans Ulrich Suter (suter@physik.uzh.ch) and Kurt Hanselmann (kurt.hanselmann@hispeed.ch), BioSym - Computational Systems Biology, Institute of Mathematics, University of Zürich, Winterthurerstr. 190, 8057 Zürich

With the software GAMESS-US it has become possible to calculate the thermodynamic properties, such as Enthalpy and Free Energy of formation with an accuracy of about 4 kJ/mole for any molecular species. Calculations are based on the geometrical structure of molecules, which are available in PDB-databases. As an example, we calculated the Enthalpy of alpha-D-Glucose as 1215 kJ/mole and the free energy as -907 kJ/mole, which agrees well with experimental value of -917.2 kJ/mole for 25°C. Unfortunately, quantification of the interaction of molecules with the aqueous cytoplasmic matrix of a living cell (solvatation effect) is not yet possible, and the calculations for large molecules requires long calculation times. Using standardized quantum chemistry methods we calculated thermodynamic values for a number of biomolecules and designed bio-thermodynamic models for intermediary reactions of the glycolysis pathway. Values for most glycolysis intermediates have not been determined experimentally and can only be obtained through calculation. Special care needs to be taken to calculate the correct protonated state of the carboxylic acid intermediates for cytoplasmic pH-conditions. The poster will outline the calculation procedure and illustrate the usefulness of the approach in systems bio-thermodynamics with a few examples.

BioSym™

A Systems Biology Learning Network produced by the BioSym™ team, http;//www.biosym.uzh.ch**

swiss | i-research & training

Understanding glycolysis with quantum chemistry

Bio-thermodynamic models in metabolism

Standardized quantum chemistry methods were employed to calculate thermodynamic values for a number of molecules, which then served to design bio-thermodynamic models for reactions of the Emden-Meyerhof-Parnas pathway (glycolysis).

Glycolysis

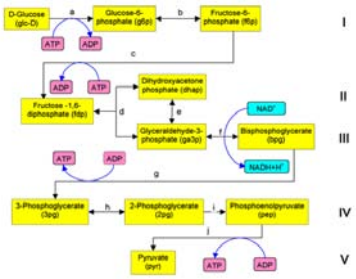


Fig.1 The 5 steps in glycolysis: **I**:Phosphorylation of glucose associated with ATP "investments"; **II**: Splitting of a C6 sugar into two C3 compounds; **III**: Oxidation and first ATP gain; **IV**: Phosphoglyceraldehyde to phosphoenolpyruvate transformation and **V**: Second ATP gain.

Calculation method illustrated for α–D–Glucose

Assumption: Molecules are synthesized from the elements:

$6C + 6H_2 + 3O_2 \longrightarrow C_6H_{12}O_6$

Enthalpy, entropy, free energy (Gibbs energy) of formation are defined as:

$Hf° = E_{el}+E_{vib}(T)+E_{trans}(T)+E_{rot}(T)+E_{sol}(T)+RT$

$Sf° = S_{vib}(T)+S_{trans}(T)+S_{rot}(T)$

$Gf° = Hf° - TSf°$

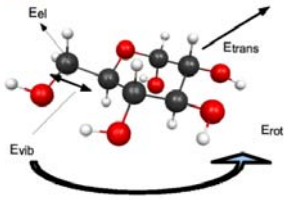


Fig.2 Calculations based on approx. geometrical structure and number of electrons. Data from PDB-databases. The solvent is water.

Table 1 Geometry and electronic energies are optimized employing GAMESS.

	b3lyp* Hf	b3lyp* Gf	g3mp2* Hf	g3mp2* Gf
C	-38.1320	-38.1228	-38.0561	-38.0579
H_2	-1.1650	-1.1799	-1.1668	-1.1816
O_2	-150.2503	-150.2726	-150.1610	-150.1610
glc	-686.9737	-687.0232	-686.2639	-686.3151

Values in Hartree; 1 Hartree = 627.51 kcal/mol = 2626 kJ/mol

* Method of calculation in GAMESS. glc = $C_6H_{12}O_6$ = glucose

$Hf°_{glc} = Hf_{glc} - 6Hf_C - 6Hf_{H2} - 3Hf_{O2} = -1163$ kJ/mol

$Gf°_{glc} = Gf_{glc} - 6Gf_C - 6Gf_{H2} - 3Gf_{O2} = -854.7$ kJ/mol

Add solvatation enthalpy: -52.2 kJ/mol

$Hf°_{glc, solv} = -1215$ kJ/mol

$Sf°_{glc} = -1.035$ kJ/mol

$Gf°_{glc, solv} = -1215 - 298.15\ (-1.035) = -907$ kJ/mol

(experimental value = –917.2 kJ/mol)

Calculations for glycolysis intermediates

Table 2 Standard enthalpies and energies of formation

Molecule in H_2O**	Hf° [kJ/mol]	Gf° [kJ/mol]	Method of calculation
glc-D	-1215	-907	g3mp2
g6p2-	-2816	-2504	b3lyp
f6p2-	-2761	-2312	b3lyp
fdp4-	-4434	-4699	b3lyp
dhap2-	-2194	-2088	g3mp2
ga3p2-	-2210	-2101	g3mp2
bpg4-	-4598	-4485	g3mp2
3pg3-	-2886	-2775	g3mp2
2pg3-	-2526	-2395	g3mp2
pep3-	-1679	-1604	g3mp2
pyr1-	-584	-501	g3mp2
ADP3-	-5545	-5154	b3lyp
ATP4-	-7594	-7223	b3lyp
mNAD1+	-812	-953	g3mp2
mNADH	-835	-674	g3mp2

** For abbreviations see figure 1

Calculations with the g3mp2 method give maximum useful precision (error of about 6 kJ/mol) within an optimal calculation time. Values for larger molecules (ATP, ADP) were calculated with the b3lyp/6–31G** method, whose error is up to 10x larger. For NAD^+ and NADH we only calculated the nicotinamide fragment, which changes its structure during the redox process. For dissolution of molecules in water we applied the PCM calculation method.

Energetic analysis of glycolysis

$\Delta Gr^* = Gf_{product} - Gf_{educt}$

C6 and C3 metabolites only

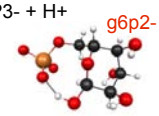
$\Delta Gr^* > 0$ = energy loss

$\Delta Gr^* < 0$ = energy gain

glc-D + ATP4- => g6p2- + ADP3- + H+

Gluco-kinase

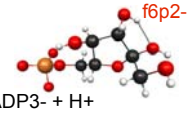
$\Delta Gr^* = -1597$ [kJ/mol]



g6p2- <=> f6p2-

Phosphoglucose-isomerase

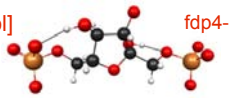
$\Delta Gr^* = +183$ [kJ/mol]



f6p2- + ATP4- => fdp4- + ADP3- + H+

Phosphofructo-kinase

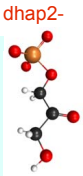
$\Delta Gr^* = -2387$ [kJ/mol]



fdp4- <=> dhap2- + ga3p2-

Fructosebisphosphate-aldolase

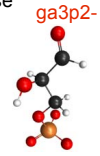
$\Delta Gr^* = +510$ [kJ/mol]



dhap2- <=> ga3p2-

Triosephosphate-isomerase

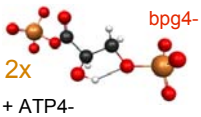
$\Delta Gr^* = -13$ [kJ/mol]



ga3p2- +NAD1+ +Pi2- <=> bpg4- +NADH +H+

Glyceraldehydephosphate-dehydrogenase

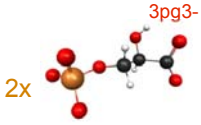
$\Delta Gr^* = -2384$ [kJ/mol]



bpg4- + ADP3- <=> 3pg3- + ATP4-

Phosphoglycerate-kinase

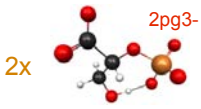
$\Delta Gr^* = +1710$ [kJ/mol]



3pg3- <=> 2pg3-

Phosphoglycerate-mutase

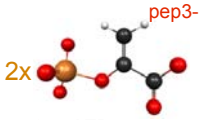
$\Delta Gr^* = +380$ [kJ/mol]



2pg3- <=> pep3- + H_2O

Enolase

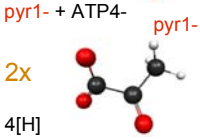
$\Delta Gr^* = +791$ [kJ/mol]



pep3- + ADP3- + H+ <=> pyr1- + ATP4-

Pyruvate kinase

$\Delta Gr^* = +1103$ [kJ/mol]



glc-D => 2 pyr1- + 2H+ + 4[H]

$\Delta Gr^* = -95$ [kJ/mol]

ADP3- +Pi2- + H+ <=> ATP4- + H_2O

ADP kinase

$\Delta Gr^* = -2069$ [kJ/mol]

References

GAMESS:http://www.msg.ameslab.gov/GAMESS/gamess.html

PCM: http://www.cup.uni-muenchen.de/oc/zipse/compchem/solv/pcm.html

G3MP2: http://www.cup.uni-muenchen.de/oc/zipse/compchem/thermo/G3MP2.html

**** Network Partners**: University Zurich, ETH Zurich, ZH Winterthur, University Fribourg, Ruhr University Bochum, WHO Geneva, Roche Basel, University Hospital Basel. **Collaborators UZH:** Barbour Andrew D. math., Brammertz Stefan biol., Fuchs Christoph[1] inform., Hanselmann Kurt[2] biol., Heinzmann Dominik math., Kälin Roman math., Lazzaretti Maja biol., Schafroth Stefan phys., Suter Hans Ulrich chem.

[1] biosym@mnf.uzh.ch, [2] i-research.training@hispeed.ch

Modeling of metabolic networks: A computational approach to functional systems biochemistry and metabolic engineering

Stefan Brammertz (sbrammertz@gmail.com) and Kurt Hanselmann (kurt.hanselmann@hispeed.ch), BioSym - Computational Systems Biology, Institute of Mathematics, University of Zürich, Winterthurerstr. 190, 8057 Zürich

The biochemistry of individual reactions in the Embden-Meyerhof-Parnas pathway (glycolysis), the Krebs cycle (citric acid cycle) and the Calvin-Benson cycle (pentose phosphate pathway) are well established. These three pathways and a number of related ones play key roles in cellular processes of many aerobic and anaerobic, prokaryotic and eukaryotic organisms. We made an attempt to design mathematical models for the quantitative analysis and dynamic simulation of these pathways. The models are based on Michaelis-Menten rate equations and mass transfer concepts; the software Simbiology (The Mathworks) is employed for model design. The models allow one to study interactions between different processes with linked biochemical reactions, the regulation of enzymes and process optimization. Enzyme parameters (K_m, K_i, v_{max}, etc.) and concentrations of metabolites are compiled from different databases available on the www (BRENDA, KEGG, ExPASy, etc.) and from scientific publications. The values are then screened for reliability and missing values are chosen based on expert knowledge.
Dynamic models are excellent learning and research tools because they allow one to study the role of individual enzymes within complex cellular metabolic networks which may lead to new hypotheses. Numerous options can be tested *in silico* before one designs and carries out experiments *in vivo* or *in vitro*.

BioSym™

A Systems Biology Learning Network produced by the BioSym™ team, http;//www.biosym.uzh.ch**

swiss | i-research & training

Modeling metabolic networks

A computational approach to functional biochemistry and metabolic engineering

Computational Models in Enzyme Kinetics

Dynamic models are excellent learning and research tools because they allow one to study the role of individual enzymes within complex cellular networks. We developed mathematical models based on SimBiology (The Mathworks) for the quantitative analysis and dynamic simulation of metabolic pathways like EMP (glycolysis), oCAC (Krebs cycle) and rCBB (Calvin cycle).

1. Glycolysis in yeast (3)

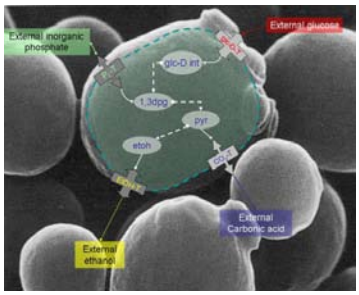


Fig.1 Glycolysis in yeast, a prerequisit for ethanol production

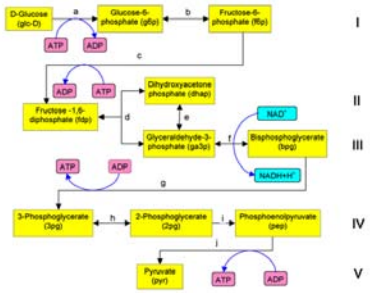


Fig.2 Glycolysis (EMP) from glucose to pyruvate can be divided into 5 steps. **I**: Phosphorylation of glucose associated with ATP "investment"; **II**: Splitting of a C6 sugar into two C3 compounds; **III**: Oxidation and first ATP gain; **IV**: Glycerate-phosphoenolpyruvate transformation, **V**: Second ATP gain.

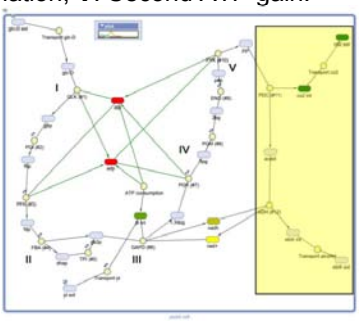


Fig.3 Glycolysis model in SimBiology

Results of a simulation

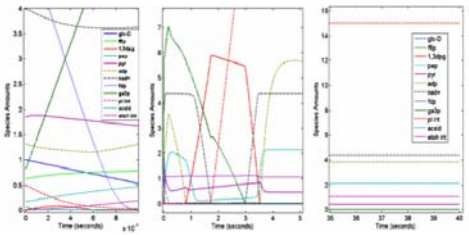
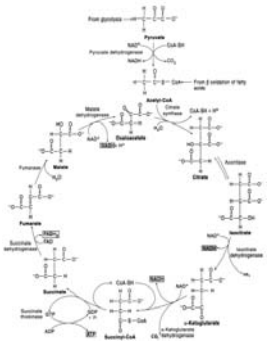


Fig.4 Adjustment of metabolite concentrations with different time resolutions.
Most intermediates quickly reach ± constant intracellular steady state concentrations.

2. oCAC in Mitochondria (1)

Fig.5 The oxidative citric acid cycle (oCAC) oxidizes catabolites to CO_2 and produces anabolic intermediates. It is located in mitochondria, but it must be linked with processes that take place in other cell compartments.



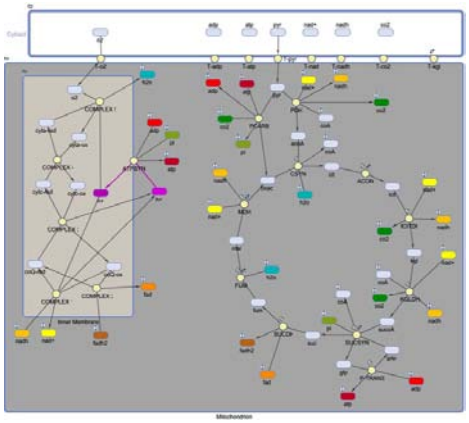
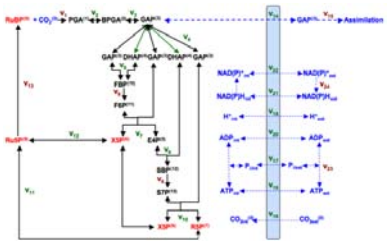


Fig.6 oCAC model in SimBiology

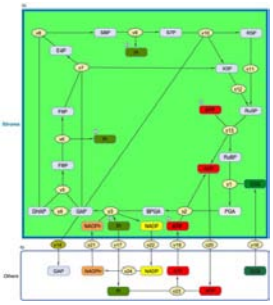
3. rCBB in Chloroplasts (2)

The reductive Calvin-Benson-Bassham cycle (rCBB) accounts for CO_2 fixation in the stroma of chloroplasts and in many autotrophic bacteria and a few archaea. It is linked to other cell compartments for the supply of ATP and NAD(P)H needed for the regeneration of the CO_2 acceptor, ribulose-1,5 bisphosphate (Fig.7).



Conceptual models (Fig.7) allow one to define reactions, enzymes and compartments for designing mathematical models in SimBiology, e.g. rCBB (Fig.8)

Fig.8 rCBB-model CO_2 fixation and RuBP regeneration in the stroma. ATP and NADPH production takes place in other cell compartments



Results of an rCBB simulation

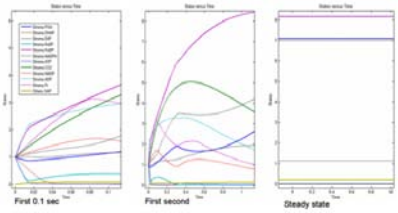


Fig.9 Adjustment phases to steady state in the stroma; normalized starting conditions

Discussion

The most critical steps in metabolic modeling are compiling experimental numerical values for:

- the concentration of metabolites and coenzymes under steady-state conditions,
- the characteristic kinetics and regulatory sensitivities of enzymes (K_m, K_i, v_{max}, etc.).

Most commonly used databases are: BRENDA, KEGG and SWISSPROT. Expert knowledge is required to calculate reasonable modeling values from a variety of experimental data obtained under different conditions.

References

1. KELLY, Patrick J. et al. (1979). The tricarboxylic acid cycle in *Dictyostelium discoideum* a model of the cycle at preculmination and aggregation. Biochem. J. 184, 589-597
1. PETTERSSON, Gosta and Ulf Ryde-Pettersson (1988). A mathematical model of the Calvin photosynthesis cycle. Eur. J. Biochem. 175, 661 -672
2. TEUSINK, Bas et al. (2000). Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. Eur. J. Biochem. 267, 5313-5329

**** Network Partners**: University Zurich, ETH Zurich, ZH Winterthur, University Fribourg, Ruhr University Bochum, WHO Geneva, Roche Basel, University Hospital Basel. **Collaborators UZH:** Barbour Andrew D. math., Brammertz Stefan biol., Fuchs Christoph[1] inform., Hanselmann Kurt[2] biol., Heinzmann Dominik math., Kälin Roman math., Lazzaretti Maja biol., Schafroth Stefan phys., Suter Hans Ulrich chem.

[1] biosym@mnf.uzh.ch, [2] i-research.training@hispeed.ch

Selection and adaptation in microbial communities: A computational modeling approach to ecosystem complexity

Roman Kälin (rkaelin@amath.uzh.ch), Munti Yuhana (myhn@gmx.ch) and Kurt Hanselmann (kurt.hanselmann@hispeed.ch), BioSym - Computational Systems Biology, Institute of Mathematics, University of Zürich, Winterthurerstr. 190, 8057 Zürich

Stability and dynamics of an ecosystem depends on the ability of its organisms to interact with each other and to quickly respond to perturbations. We have studied changes in microbial community compositions in a remote high mountain lake that seasonally passes through extremes of environmental changes. The ecosystem was analyzed applying molecular techniques which are based on biomolecular indicators and combined with measurements of physicochemical ecosystem determinants. The diversity of organisms is overwhelming and, due to the variability of parameter combinations under natural conditions, one can seldom observe similar population compositions under seemingly similar environmental settings. Instead, numerous community patterns emerge from the lake's population pool which allow one to create hypotheses and concepts about the role of selection and adaptation in community regulation.

We have developed a computational "selection-adaptation model" based on extended Lotka-Volterra algorithms that allows one to simulate population development and disappearance with predetermined parameter assignments. The investigator can define stabilizing and destabilizing mechanisms and follow population diversity changes.

An understanding of ecosystem complexity cannot be reached by observation and experimentation alone. Good theoretical models help one to carry out numerous simulations *in silico* and to define those environmental determinants and organismic characteristics that might play essential regulatory roles.

BioSym™

A Systems Biology Learning Network produced by the BioSym™ team, http;//www.biosym.uzh.ch**

swiss | i-research & training

Selection and adaptation in microbial communities

How can diversity be altered ?

Microbial ecosystems that contain diverse populations respond to variable conditions by adjusting community homeostasis rapidly.

This flexibility requires ...

- minimal population sizes
- energy and resources for reproduction
- responses to environmental signals
- gene exchange mechanisms
- fitness of organisms for change
- selection of positive mutations
- interacting populations
- many more

Here, we introduce computational approaches to diversity modeling applying Matlab and Simulink (The Mathworks).

Concepts of a selection / adaptation model

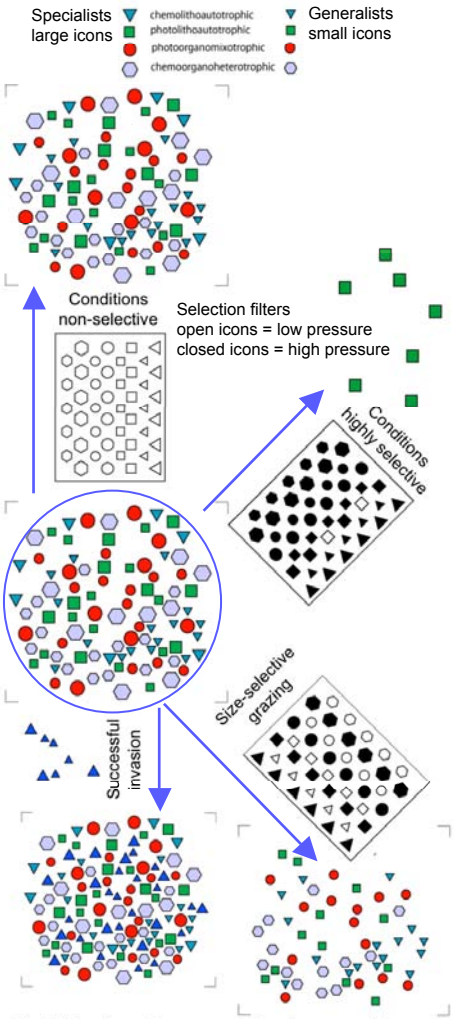


Fig.1 Modes of community changes through selective „filtering“ (examples)

4-population competition model with 4 regulatory variables[3]

$$dX_1/dt = a_1 \cdot X_1 - a_{12} \cdot X_1 \cdot X_2 - a_{13} \cdot X_1 \cdot X_3 - a_{14} \cdot X_1 \cdot X_4 - a_{11} \cdot X_1^2$$
$$dX_2/dt = a_2 \cdot X_2 - a_{21} \cdot X_1 \cdot X_2 - a_{23} \cdot X_1 \cdot X_3 - a_{24} \cdot X_1 \cdot X_4 - a_{22} \cdot X_2^2$$
$$dX_3/dt = a_3 \cdot X_3 - a_{31} \cdot X_1 \cdot X_3 - a_{32} \cdot X_2 \cdot X_3 - a_{34} \cdot X_3 \cdot X_4 - a_{33} \cdot X_3^2$$
$$dX_4/dt = a_4 \cdot X_4 - a_{41} \cdot X_1 \cdot X_4 - a_{42} \cdot X_2 \cdot X_4 - a_{43} \cdot X_3 \cdot X_4 - a_{44} \cdot X_4^2$$

$$a_1 = a_{1A} \cdot A^*_1 + a_{1B} \cdot B^*_1 + a_{1C} \cdot C^*_1 + a_{1D} \cdot D^*_1$$
$$a_2 = a_{2A} \cdot A^*_2 + a_{2B} \cdot B^*_2 + a_{2C} \cdot C^*_2 + a_{2D} \cdot D^*_2$$
$$a_3 = a_{3A} \cdot A^*_3 + a_{3B} \cdot B^*_3 + a_{3C} \cdot C^*_3 + a_{3D} \cdot D^*_3$$
$$a_4 = a_{4A} \cdot A^*_4 + a_{4B} \cdot B^*_4 + a_{4C} \cdot C^*_4 + a_{4D} \cdot D^*_4$$

- X_i = size of population (pop.) i, i = 1, 2, 3, 4
- A, B, C, D = regulatory settings: radiation, nutrients, fitness, gene exchange, etc.
- a_i = growth rate of pop.i altered by A,B,C,D
- a_{ij} = influence of pop.j on growth of pop.i

- a_{iJ} = effect of regulatory setting J on growth rate a_i of population i, i = 1,2,3,4, J = A,B,C,D; $a_{iJ} > 0$: growth stimulated, $a_{iJ} < 0$: growth hindered
- J^*_i = normalized impact $0 \leq J^*_i \leq 1$

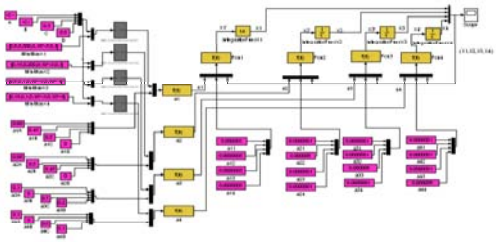


Fig.2 Simulink / Matlab diagram implementing the eight equations above

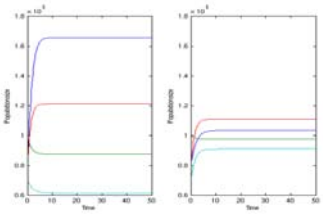


Fig.3 Example: Development of 4 populations under 2 different regulatory settings
left: A = 18.7, B = 458.5, C = 0.0004, D = 0.89
right: A = 1.2, B = 176.4, C = 0.0008, D = 0.01

Outcome of simulations with Applet

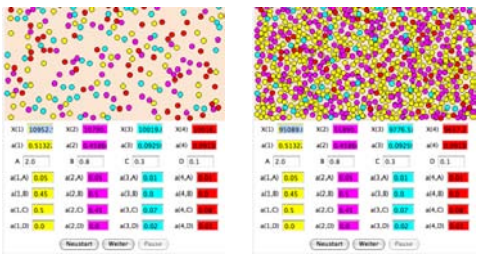


Fig.4 Start and end simulations for a 4 populations /4 regulatory settings model

The Lotka-Volterra-like model[1,2]

$$x_i(t+1) = x_i(t)^* \exp(r - \Sigma_j b_{ij} * x_j(t))$$
$$i = 1,\ldots,n$$

in matrix notation:

$$X(t+1) = X(t) * \exp(R - B * X(t))$$

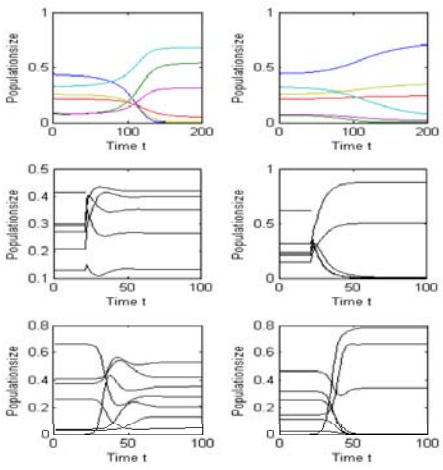


Fig.5 Responses of communities to change

Top row: What will happen if the initial population size changes only slightly?

Middle row: What will happen if, in a stable community, the most common population goes extinct? **Left**: All remaining populations survive, but steady state composition changes; **right**: some other populations go extinct as well.

Bottom row: What will happen if a new population can establish itself in a steady state community? **Left**: All populations survive, but steady state composition changes; **right**: some populations go extinct

Conclusions

Computational models may help to understand complex community changes that can or cannot be analyzed experimentally (Fig.6)

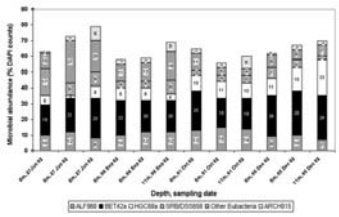


Fig.6 Enumeration of microbial community composition in Lake Jöri by FISH analysis. % of hybridized cells in relation to total detected DAPI counts.

References

[1] Anthony R. Ives, Stephen R. Carpenter (2007) *Stability and Diversity of Ecosystems.* Science 317, 58

[2] BioSym: M-14-03 System dynamics of a multi-species ecosystem model

[3] SIMOLIFE: Microbiology / Selection-Adaptation

** **Network Partners**: University Zurich, ETH Zurich, ZH Winterthur, University Fribourg, Ruhr University Bochum, WHO Geneva, Roche Basel, University Hospital Basel. **Collaborators UZH:** Barbour Andrew D. math., Brammertz Stefan biol., Fuchs Christoph[1] inform., Hanselmann Kurt[2] biol., Heinzmann Dominik math., Kälin Roman math., Lazzaretti Maja biol., Schafroth Stefan phys., Suter Hans Ulrich chem.

[1] biosym@mnf.uzh.ch, [2] i-research.training@hispeed.ch

Eco-genomics of rumen communities: How similar, in an evolutionary sense, are cellulases from different rumen microbes?

Maja A. Lazzaretti-Ulmer (maja.lazzaretti@math.uzh.ch) and Kurt Hanselmann (kurt.hanselmann@hispeed.ch), BioSym - Computational Systems Biology, Institute of Mathematics, University of Zürich, Winterthurerstr. 190, 8057 Zürich

The rumen is a complex ecosystem. Its microbiota comprises mostly anaerobic bacteria and archaea, anaerobic, ciliated protozoa and anaerobic fungi. Cellulose $(C_6H_{10}O_5)_n$ is enzymatically hydrolized in a first step by cellulases produced by some members of the microbiota. The resulting di- and monosaccharides are then further utilized by the same and by other microbes of the community, which produce volatile fatty acids, CO_2, CH_4 and a number of other metabolites.
We retrieved amino acid sequence information for cellulase proteins for a number of rumen microorganisms (*Butyrivibrio fibrisolvens, Clostridium longisporum, Fibrobacter succinogenes, Prevotella ruminicola, Ruminococcus albus* and *Ruminococcus flavefaciens*) from different data bases as well as of *Pyrococcus abyssi*, an Archaeon, which is not a member of any rumen community, and compared them employing the Pfam Protein Families Database tools and the softwares ClustalX and PHYLIP. The resulting phylogenetic tree was then compared with the phylogenetic tree made for the same microorganisms based on their 16S rRNA data. The two trees revealed interesting differences, which suggest that cellulase genes were in some cases obtained by horizontal gene transfer. It is surprising that this should have been happened between microorganisms of different domains and the transfer path between mesophilic bacteria and thermophilic archaea remains to be further investigated.

BioSym™

A Systems Biology Learning Network produced by the BioSym™ team, http;//www.biosym.uzh.ch**

swiss | i-research & training

Evolutionary eco-genomics of rumen cellulases

Enzymes for biotechnology from rumen organisms

- In herbivorous ruminants, such as cattle, dairy cows, goats and sheep, fibrous plant polymers (cellulose, hemi-celluloses etc.) are hydrolized and fermented in a series of complex catabolic reactions, which are carried out in the rumen by many different anaerobic microorganisms (Fig.1, Tab.1)
- Fibrolytic enzymes are produced exclusively by the rumen microbiota
- In this study we focus on the evolutionary relationships among cellulases (Fig.2, step 2; Fig.3)
- How much orthology, how much xenology can we detect among homologous cellulases?

Fig.1 The rumen community consists of a great diversity of bacteria, archaea and eukarya, e.g. protozoa and fungi.

6 Levels of a rumen food web

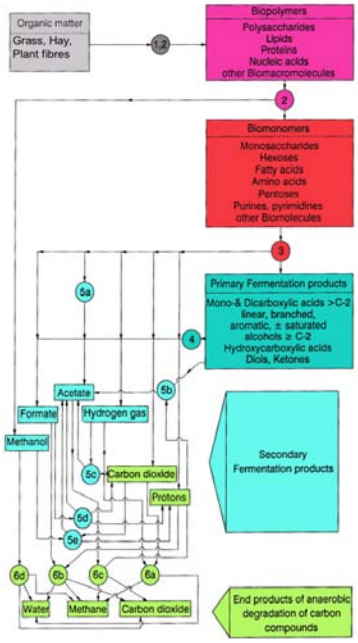


Fig.2 1 Lysis of cells and tissues, 2 Hydrolysis of biopolymers, 3 Primary fermentation, 4 Secondary fermentation, 5 Acetogenesis, 6 Methanogenesis

Functionally homologous fibrolytic enzymes are present in taxonomically very different rumen microorganisms (e.g. bacteria; Tab.1)

ORGANISMS	1 Starch and/or Glycogen degradation	2 Cellulose degradation	3 Pectin degradation	4 Hemicellulose degradation	5 Sugar metabolism, Glycolysis	6 Acetogenesis	7 Butyrogenesis	8 Lactate formation	9 Succinate formation	10 Lactate conversion to propionate	11 Succinate conversion to propionate	12 Formate formation	13 Methanogenesis from H_2, CO_2 and formate
Ruminobacter amylophilum	X				X	X			X			X	
Succinimonas amylolytica	X				X	X			X	X			
Selenomonas ruminantium	X				X	X		X			(X)		
Selenomonas ruminantium subsp. lactilytica	X				X	X			(X)	X			
Streptococcus bovis	X		X		X			X					
Ruminococcus flavefaciens		X		X	X	X			X			X	
Ruminococcus albus		X		X	X	X						X	
Fibrobacter succinogenes		X		X	X	X			X			X	
Butyrivibrio fibrisolvens		X		X	X	X	X	X				X	
Lachnospira multiparus			X	X	X	X		X				X	
Lactobacillus spp.					X			X					
Schwartzia succinivorans											X		
Veillonella parvula						X				X	(X)		
Megasphaera elsdenii					X	X	X			X	(X)		
Methanobrevibacter ruminantium													X
Methanomicrobium mobile													X

Catabolic Abilities

( ) some strains only or only by resting cells

Table 1. Selected Bacteria and Archaea of the rumen microbiota and their catabolic abilities.

Cellulases

Cellulase: EC 3.2.1.4:
Endohydrolysis of 1,4-β-D-glucosidic linkages

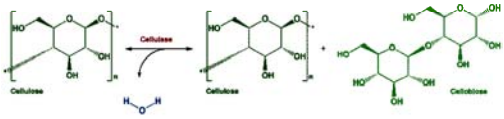


Fig.3 Cellulose + H_2O ⇔ Cellulose + Cellobiose
KEGG reaction R02886

Approach

- Retrieve amino acid sequence information for cellulase proteins from different data bases for *B. fibrisolvens*, *C. longisporum*, *F. succinogenes*, *P. ruminicola*, *R. albus, R. flavefaciens* as well as of *Pyrococcus abyssi*, an Archaeon, that is not a member of any rumen community (Tab.2, Fig.4).
- Analyze the sequences employing Pfam Protein Families Database tools, ClustalX, PHYLIP and Bioinformatics toolbox (Matlab).
- Construct a phylogenetic tree for cellulases and compare to the phylogenetic tree for the same microorganisms based on 16S rRNA gene sequence data (Fig.5)

References

- From BioSym Module „Gene transfer and evolution (Rumen enzymes)“
- Cohen, G.N., et al. (2003). An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*. Molecular Microbiology 47(6): 1495–1512.
- Garcia-Vallvé, S., Romeu, A. and Palau, J. (2000). Horizontal Gene Transfer of Glycosyl Hydrolases of the Rumen Fungi. Molecular Biology and Evolution 17(3): 352-361.

Cellulases, alignment, phylogeny

Table 2. Protein sequence identification and corresponding abbreviation (* needed in the application with ClustalX and PHYLIP)

Pfam and UniProtKB Entry name	UniProtKB Primary accession number	Protein name	Origin of the protein	Abbreviation*
Rumen Bacteria:				
GUN1_BUTFI	P20847	Endoglucanase 1	***Butyrivibrio fibrisolvens***	Butyri_fib
GUNA_CLOLO	P54937	Endoglucanase A precursor	***Clostridium longisporum***	Clostr_lon
Q59445_FIBSU	Q59445	Endoglucanase 3 precursor	***Fibrobacter succinogenes***	Fibrob_suc
Q9ZN63_PRERU	Q9ZN63	Cellulase	***Prevotella ruminicola***	Prevot_rum
GUN1_RUMAL	P16216	Endoglucanase 1 precursor	***Ruminococcus albus***	Rumino_alb
O05143_RUMFL	O05143	Endoglucanase A precursor	***Ruminococcus flavefaciens***	Rumino_fla
Archaeon isolated from a deep-sea hydrothermal vent (as outgroup):				
Q9V052_PYRAB	Q9V052	Major extracellular endo-1,4-betaglucanase	***Pyrococcus abyssi***	Pyroco_aby

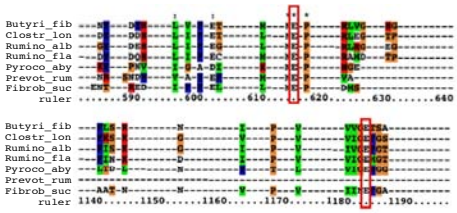


Fig.4 Details of aligned aminoacid sequences of cellulases from the studied microorganisms. The Pfam predicted active sites (2 glutamates, E) are very well conserved. (The sequence from *Prevotella ruminicola* cellulase is shorter).

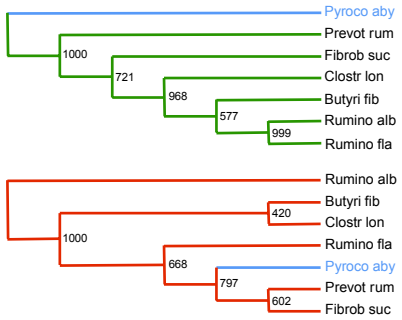


Fig.5 Phylogenetic trees of six rumen bacteria and an Archaeon (*P. abyssi*), based on 1000 bootstrap samples each for the 16S-rRNA (top) and the amino acid sequences of the cellulase (below).

Conclusions

- Differences in phylogenies: cellulase genes were in some cases probably obtained by horizontal gene transfer (HGT).
- HGT of endoglucanase (celA) from the rumen bacterium *F.succinogenes* to the rumen fungi *Orpinomyces joyonii* was postulated by Garcia-Vallvé et al. (2000).
- Cohen et al. (2003) reported on proteins, presumably of bacterial origin, including some from mesophilic bacteria, to be present in the *Pyrococcus abyssi* genome. This implies that HGT could have happened between organisms of different domains.
- The transfer path between mesophilic bacteria and thermophilic archaea remains to be investigated further.

** **Network Partners**: University Zurich, ETH Zurich, ZH Winterthur, University Fribourg, Ruhr University Bochum, WHO Geneva, Roche Basel, University Hospital Basel. **Collaborators UZH:** Barbour Andrew D. math., Brammertz Stefan biol., Fuchs Christoph[1] inform., Hanselmann Kurt[2] biol., Heinzmann Dominik math., Kälin Roman math., Lazzaretti Maja biol., Schafroth Stefan phys., Suter Hans Ulrich chem.

[1] biosym@mnf.uzh.ch, [2] i-research.training@hispeed.ch