APPLICATIONS OF BIOCHEMICAL NETWORKS DISCOVERING CONTROL MECHANISMS IN SYSTEMS BIOLOGY

Tatjana Rubina, Latvia University of Agriculture/ Faculty of Information technologies, Liela 2, Latvia,
tatjana.zukova@inbox.lv, +371 25999206
Valters Brusbardis, Latvia University of Agriculture/ Faculty of Information technologies, Liela 2, Latvia,
brusbardisv@inbox.lv
Egils Stalidzāns, Dr.sc.ing., Doc. of Latvia University of Agriculture/ Faculty of Information technologies,
egils.stalidzans@llu.lv

Abstract: Systems biology is a branch of computational biology which studies biological organisms. A goal of systems biology is to understand living organisms at the systems level, combining quantitative information on individual components in order to understand the emergent behaviors that result. Because the mathematic-analytical part of systems biology is not perfect, therefore as new methods of researches come simulations, modeling, computations and other methods. To understand biology at the system level, we must examine the structure and dynamics. Important item studying systems behavior is to find the structure of possible control loops (positive or negative feedback). Prerequisite of a control loop is an oriented cycle in the structure. The approach we used to mapping the structure is graph-based. The network is represented as a directed graph with labeled edges. In this study we examine structure of unique and evolutionary highly conserved process Heat shock response (HSR) with goal to find possible control loops. Thus structural analysis can give insights about control mechanisms, alternatively possible ways of their execution as well as predict possible ways of control alteration.

Keywords systems biology, bioinformatics, mathematical model, biological system, structure of biological system, biochemical network, metabolic networks, control loop, positive (negative) feedback, Heat shock response.

Systems biology and information technologies

Systems biology (SB) is a branch of computational biology which studies biological organisms in their essence and the relationships between the components that make up an organism. SB is not the biology of systems, nor is it the chemistry / physics / molecular genetics of molecules in biological systems (Alberghina and colleagues, 2005). System biology studies biological systems by systematically perturbing them (biologically, genetically, or chemically); monitoring the gene, protein, and informational pathway responses; integrating these data; and ultimately, formulating mathematical models that describe the structure of the system and its response to individual perturbations (Klipp and colleagues, 2005).

A goal of systems biology is to understand living organisms at the systems level, combining quantitative information on individual components in order to understand the emergent behaviors that result (Kholodenkov and colleagues, 2005). System biologists build a system-level

understanding of how the biological world works and solve problems by understanding systems and then applying that knowledge to control them (Paxson and colleagues, 2007).

Because the mathematic-analytical part of systems biology is not perfect, therefore as new methods of researches come simulations, modeling, computations and other methods. While these techniques have great potential in systems biology, biologists have not yet applied them as efficiently (as engineers in traditional disciplines) - both because of the complexity of biological systems and because systems biology research requires contributions from a diverse group of researchers (Paxson and colleagues, 2007) (modelers, mathematics, biologists).

Available data

The last several years' biological research produces increasing volumes of data describing genome sequence of biological organisms, cellular components, their interactions, and states of biological networks for model organisms (Herrgard and colleagues, 2006). This available information about biological entities and their interactions enable us to consider different organisms as molecular systems which control the genetic information

Usually the first available information about a cellular process comes as a prediction of a structure – interconnections between the elements that take part in the process. Numerical characteristics of reactions take much more experimental work and time. Thus extracting of knowledge from structure of networks is the first task that can be completed.

Important item studying systems behavior is to find the structure of possible control loops (positive or negative feedback). Prerequisite of a control loop is an oriented cycle in the structure.

Biochemical networks of the biological system

The identification of most genes encoding the metabolic enzymes of some organism has enable methodologies for the systematic mapping of metabolic networks (Westerhoff and colleagues, 2005). The method first identifies the genes that encode enzymes then identifies the chemical reactions these enzymes catalyze, and then writes for each enzyme which chemical compounds it produces and consumes, and at which stoichiometry (Westerhoff and colleagues, 2005). Component and interaction data including genome sequences, protein complexes and protein – DNA interactions can be used to establish the connectivity of the biochemical networks inside the cell. System-state data types including gene expression, metabolite level, metabolic flux, and high-throughput deletion strain phenotyping data present the states and outputs of these networks (Herrgard and colleagues, 2006). These network models can be used to predict changes in the system states in response to genetic and environmental perturbations.

To understand biology at the system level, we must examine the structure and dynamics of cellular and organismal function (Kitano) (Alberghina and colleagues, 2005). At the first we construe the structure of biological system that we have chosen.

The relevant literature and information for metabolic and transcription factor gene (for Saccharomyces cerevisiae) is available in the SGD, YPD and MIPS databases. The Metabolic network model consists of the stoichiometric matrix describing all the interconversions between metabolites in the network, maximum flux capacity constraints on all reactions, and a cellular objective function (Herrgard and colleagues, 2006).

The approach we used to mapping the structure is graph-based. The network is represented as a directed graph with labeled edges. The nodes of the graph represent the enzymes, transcriptional factors or elements and edges – reactions or processes.

An example of biological system

Heat shock response (HSR) is a unique and evolutionary highly conserved process found through living beings from prokaryotes to eukaryotes (El-Samad and colleagues, 2005; Kurata and colleagues, 2001; Pockley and colleagues, 2003). Basically HSR is a protection system of the cells against different harmful conditions, including chemical stress (ethanol, heavy metals, etc.), biological stress (inflammation, virus infection, mutant proteins, etc.) and physical stress (elevated temperature, radiation) [El-Samad and colleagues, 2005; Jolly and colleagues, 2000; Pockley and colleagues, 2003; Rieger and colleagues, 2005). A huge effort has been attended to HSR by research area because of high importance in medicine - especially oncology (Jolly and colleagues, 2000). Recently computational biologists have paid attention to HSR. Nevertheless only few researchers have used control theory of engineering to analyze HSR (El-Samad and colleagues, 2005; Kurata and colleagues, 2001).

Different harmful conditions mentioned above in the cells can initiate misfolding and uncoiling of the proteins (Kurata and colleagues, 2001). That will lead to malfunction of the proteins and even more such misfolded proteins are used to form huge aggregates which are harmful to the cells (Petre and colleagues, 2008). The HSR will be turned on by the cells to escape malfunctioning of the system and to fix misfolded and uncoiled proteins (see Fig.1.).

The main elements of the HSR are heat shock proteins (HSP). HSP are categorized in several families according to molecular weight (small HSPs, HSP40, HSP60, HSP70, HSP90, HSP110). However functionality of HSP can occur by two different pathways. Some HSP function as molecular chaperons some as proteases. Chaperons take part in the assembly, stabilization and folding of oligomeric proteins whereas proteases mediate the degradation of damaged proteins. (Kurata and colleagues, 2001; Pockley and colleagues, 2003)

At the stress conditions expression amount and intensity of HSP is smoothly regulated by transcription factors (also known as heat shock factor - HSF). Family of HSF consists of HSF1, HSF2, HSF3 (unique avian transcription factor) and HSF4 (Pirkkala and colleagues, 2001; Lindquist and colleagues, 1988). Nevertheless HSF is able to activate and inactivate expression of HSP by interactions with heat shock elements (HSE) in the HSP gene promoter regions (Pockley and colleagues, 2003).
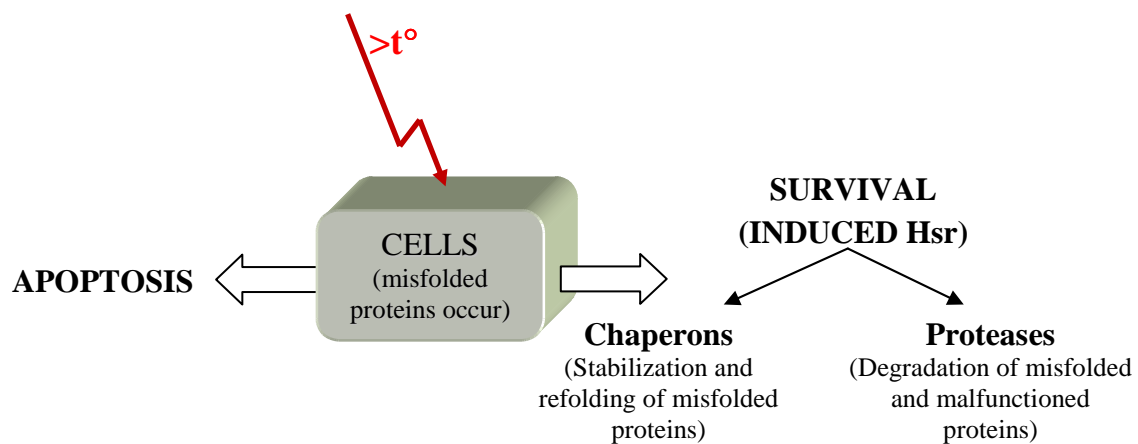

Fig.1. Stress induced HSR.

At the normal conditions HSF is presented at monomer stage. As soon as stress occurs monomers of HSF dimerize and trimerize (see molecular reactions 1 and 2 mentioned below). Only trimers of HSF can form conjunction between HSE (see molecular reaction 3 mentioned below). Newly formed complex will activate transcription of the HSP gene which results in newly synthesized HSP. Destruction of trimers and complex between HSF and HSE will inactivate the system (Petre and colleagues, 2008).

Biochemical network

Several researcher groups (Ion Petre and co-workers (Petre and colleagues, 2008), H. Kurata and co-workers (Kurata and colleagues, 2001), Theodore Rieger and co-workers (Rieger and colleagues, 2005)) have presented mathematical model of HSR. Mathematical model of Eukaryotic HSR presented by Ion Petre and co-workers was used as basic to analyze control loops of HSR. Originally model presented by Ion Petre and co-workers consists of 18 molecular reactions (see below) (Petre and colleagues, 2008).

1. $2 * hsf \leftrightarrow hsf2$
2. $hsf + hsf2 \leftrightarrow hsf3$
3. $hsf3 + hse \leftrightarrow hsf3{:}hse$
4. $hsf3{:}hse \rightarrow hsf3{:}hse + mhsp$

Explanation of abbreviations:

hsf – monomer of heat shock factor;

hsf2 – dimer of heat shock factor;

hsf3 – trimer o f heat shock factor;

5. hsp + hsf ↔ hsp:hsf

6. hsp + hsf2 → hsp:hsf + hsf

7. hsp + hsf3 → hsp:hsf + 2 * hsf

8. hsp + hsf3:hse → hsp:hsf + 2 * hsf + hse

9. hsp → 0

10. prot → mfp

11. hsp + mfp ↔ hsp:mfp

12. hsp:mfp → hsp + prot

13. hsf → mhsf

14. hsp → mhsp

15. hsp + mhsf ↔ hsp:mhsf

16. hsp:mhsf → hsp + hsf

17. hsp + mhsp ↔ hsp:mhsp

18. hsp:mhsp → 2 * hsp

hse – heat shock element;

hsp – heat shock protein;

prot – protein

mfp – misfolded protein;

mhsf – misfolded heat shock factor;
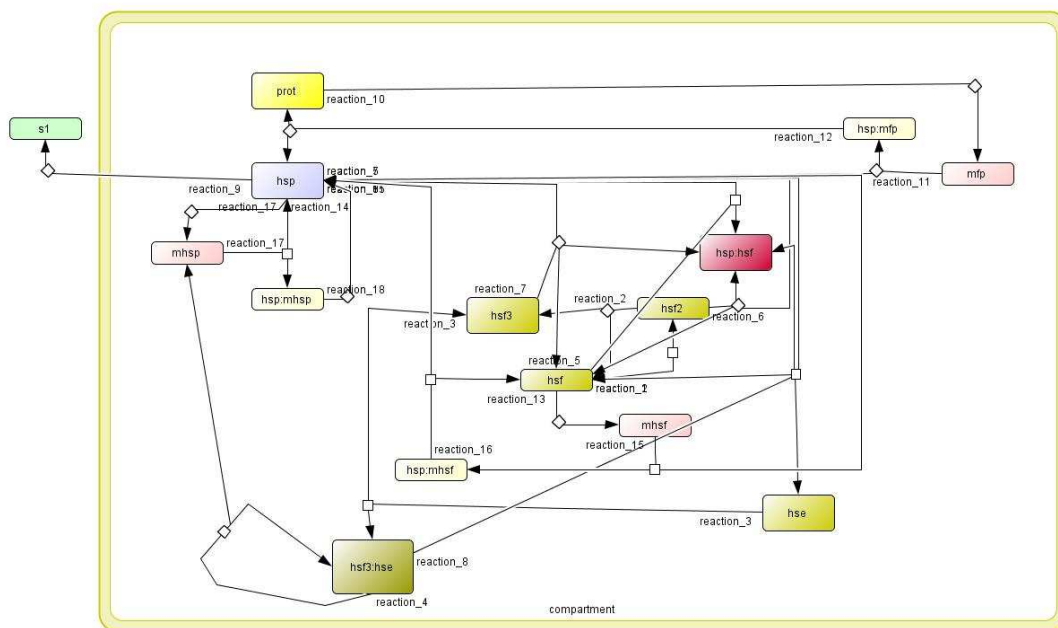
mhsp – misfolded heat shock protein

Fig.2. Topological structure of biochemical network of HSR.

Ion Petre and co-workers used Copasi 4.4 software to analyze mathematical model of HSR. For better understanding of the structure of HSR Copasi file of molecular reactions converted to SBML file to visualize by computer software CellDesigner (see Fig.2.)

Search of control loops

CellDesigner represents HSR as a graph enabling diagnoses structure and conjunction among elements of HSR. Analyses of structure gives confirmation to known control loops and gives

opportunity to search for new candidate loops. Even more structure of known control loops can lead to alternative realizations of control loops.

	reaction 01.	reaction 02.	reaction 03.	reaction 04.	reaction 05.	reaction 06.	reaction 07.	reaction 08.	reaction 09.	reaction 10.	reaction 11.	reaction 12.	reaction 13.	reaction 14.	reaction 15.	reaction 16.	reaction 17.	reaction 18.
hsf	-2	-1	0	0	-1	1	2	2	0	0	0	0	-1	0	0	1	0	0
hsp	0	0	0	0	-1	-1	-1	-1	-1	0	-1	1	0	-1	-1	1	-1	2
mhsp	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	-1	0
hsf2	1	-1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0
hsp:hsf	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0
mfp	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0
mhsf	0	0	0	0	0	0	0	0	0	0	0	0	1	0	-1	0	0	0
hsf3	0	1	-1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0
prot	0	0	0	0	0	0	0	0	0	-1	0	1	0	0	0	0	0	0
hsp:mhsf	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0
hsp:mhsp	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1
hse	0	0	-1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
hsp:mfp	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0
hsf3:hse	0	0	1	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0

Fig.3. Stoichiometry matrix of HSR.

Manual search of oriented cycles in the structure can be easy in case of a small network (up to 10 nodes). In case of bigger networks mathematical algorithms should be applied using stoichiometric matrix of a control network that contains the same information as in Fig.2. An example of stoichiometry matrix is shown in Fig.3.

Further targets

Cellular behavior is determined not only by available biological entities, but mainly by their dynamic interactions and individual properties. Activities of most if not all of the enzymes involved in cellular metabolism are regulated by end products and intermediates of corresponding pathways. This complex network with positive and negative feedback as well as genetic regulation of expression level provide flexible adaption of metabolic network to fast and low changes in external environment correspondingly (Demin and colleagues, 2005).

In the next step of research we plan to analyze the features of evolutionary development of structure and it's dynamic.

Conclusions

1. Holistic approach of systems biology aims to understand and predict cellular processes in quantitative way. However usually quantitative data is not available.

2. Under lack of numerical information about the process of interest structural analysis can be used to study control mechanisms of a process. Graphically presented structure or stoichiometric matrix of biochemical network can be used to identify oriented cycles that are one of prerequisites of control loop existence.

3. Thus structural analysis can give insights about control mechanisms, alternatively possible ways of their execution as well as predict possible ways of control alteration.

Literature

[1] Hans V. Westerhoff and Jan-Hendrik S. Hofmeyr. What is system biology? From genes to functions and back. L.Alberghina, H.V. Westerhoff. Systems Biology: Definitions and Perspectives. Springer (2005). pp.119-141

[2] Boris N. Kholodenko, Frank J.Bruggeman, Herbert M. Sauro. Mechanistic and modular approaches to modeling and inference of cellular regulatory networks, Systems Biology: Definitions and Perspectives. Springer (2005). pp.

[3] Oleg V. Demin, Tatyana Y.Plyusnina, Galina V. Lebedeva, Ekaterina A.Zobova, Eugeniy A. Metelkin, Alex G. Kolupaev, Igor I. Goryanin, Frank Tobin. Kinetic modeling of the E.coli metabolism. L.Alberghina, H.V. Westerhoff. Systems Biology: Definitions and Perspectives. Springer (2005). pp.31-67

[4] Richard Paxson, Kristen Zannella. System biology: Studying the world's most complex dynamic systems// Journal TheMathWorksNews&Notes - June (2007), p.4-7

[5] L.Alberghina, H.V.Westerhoff (Eds.) Systems Biology. Definitions and Perspectives.Springer Verlag Berlin Weidelberg (2005), 408 pp.

[6] Markus J. Herrgard, Baek-Seok Lee, Vasiliy Portnoy, and Bernhard O.Palsson. Integrated analyses of regulatory and metabolic networks reveals novel regulatory mechanisms in Saccharomyces cerevisiae. Genome research (2006).

[7] E. Klipp, R.Herwig, A.Kowald, C.Wierling, H.Lehrach. Systems biology in practice. Concept, Implementation and Application. WILEY-VCH Verlag GmbH & Co, KGaA, Weinheim (2005).

[8] El-Samad H., Kurata H., Doyle C. J., Gross C. A., Khammash M. Surviving heat shock: Control strategies for robustness and performance // PNAS. – v. 102, No. 8 (2005), P. 2736 – 2741.

[9] Jolly C., Morimoto R. I. Role of the heat shock response and molecular chaperones in oncogenesis and cell death // Journal of the National Cancer Institute. – v. 92, No. 19 (2000), P. 1564 – 1572.

[10] Kurata H., El-Samad H., Yi T. M., Khammash M., Doyle J. Feedback regulation of the heat shock response in E. Coli / Proceedings of the 40th IEEE, held in Orlando, Florida USA, December, 2001, p 837 – 842.

[11] Lindquist S., Craig E. A. The heat – heat shock proteins // Annual Review of Genetics. – v. 22, (1988), P. 631 – 677.

[12] Petre I., Mizera A., et al. A new mathematical model for the heat shock response // Proceeding of COMPMOD 2008. – Turku Centre for Computer Science v.47 (2008) pp. 39 – 52.

[13] Pirkkala L., Nykanen P., Sistonen L. Roles of the heat shock transcription factors in regulation of the heat shock response and beyond // The FASEB Journal. – v. 15, No. 7 (2001), P. 1118 – 1131.

[14] Pockley G. Heat shock proteins as regulators of the immune responce // The Lancet. – v. 362, (2003), P. 469 – 476.

[15] Rieger T. R., Morimoto R. I., Hatzimanikatis V. Mathematical modelling of the eukaryotic heat – shock response: dynamics of the hsp70 promoter // Biophysical Journal. – v.88 (2005), P. 1646 – 1658.