# Questions and answers on the developments and potential of Systems Biology in 2005; results of the Gosau plenary discussions[1]

## *1. Discussion on Principles of Systems Biology (Sunday 7 pm plenary discussion)*

1. *Biological principles versus engineering principles: Is it risky to take engineering as a background when addressing biological questions? One might miss the variability aspect, so typical of biology and get the wrong answers. Also: in the reverse direction; can engineering learn from biology?* Doyle: Yes, it is risky. Most technology is not quite good. Be selective! But it is easy to misinterpret what is important (not feathers and flapping for flight of airplanes; control was important). Yes, learn from each other; be critical!

2. *Will detailed kinetic information be required for Systems Biology; or, what could it be replaced by?* Hans Westerhoff: "Yes, for ultimate testing of proposed mechanisms, kinetic information will be needed". Reinhart Heinrich: "Yes, but how can such information be obtained." John Doyle: "The answer is: No (for the sake of the argument), because most of the details do not matter (cf. this laptop; the details of the microprocessor and capacitors in it do not matter for its performance; the latter is determined by its programmer and its user). If the parameter does not matter for function, it will be very hard to measure that parameter. However, not measuring is dangerous, because the answer might be wrong. As a consequence: Yes AND No." Igor Goryanin supports John Doyle's position, and then wonders what the success measure is of Systems Biology: "To improve the yield is an example of a success measure. If SB helps it is OK; it then does not matter whether one then has used/measured all the kinetic details." Benjamin Hall: "The answer might depend on the questions one asks; like in Molecular Dynamics, the coarser grained methods may give us new insights, or reduce computational intensity." Mattias Reuss: "Parameter values become less and less important as the network size increases." Reinhart Heinrich argues against this view: "One needs to know some parameters; not all are determined by the network stoichiometries." ??: (Please e-mail the name) "If one is realistic; how much information can one determine? Where should one put the cut off?" Stefan Schuster: "Models should be made with a degree of kinetic detail that depends on the question asked."

---

[1] Report of Questions/answers sessions of the Gosau First FEBS Advanced Lecture Course on Systems Biology;version 4/18/2005 10:58:03 AM
PLEASE CHECK THIS! Please e-mail corrections to hweste@bio.vu.nl
[Notes made by Hans Westerhoff (1,2,3) and Jacky Snoep (4) , edited by Hans Westerhoff, with apologies for possible mistakes]

3. *Why something is happening is often discussed in Systems Biology. However it seems to have either of two meanings:* ***Mechanism versus purpose****. Shouldn't it always be specified which of the two types of question one addresses?* Albert Goldbeter: "*E.g.* the question might be posed 'why are there so many rhythms?' and this could have either of the two meanings. Pulsatile cAMP in Dd has a function and the heart has been optimized for function. The functional types of question is relevant therefore." John Doyle: "The mechanism question is a is How? Question. There are in fact different 'Why' questions. Why do we have eyes?: to see (*i.e.* function right now). Why do they provide functionality?: evolution pressure. In addition, in evolution, things can get a functionality, all of sudden, that they had not evolved for. Also, there are different whys for different contexts. One should be careful with 'why' arguments."

4. *Continue on discussion by Douglas Kell: importance of inductive reasoning vs deductive reasoning; bottom up versus top down, analytic versus synthetic? Network structure is inductive or deductive?* Answer: "Both"

5. *Fundamental differences between metabolic analyses and signal transduction/genetic-network analyses. How to integrate metabolic, signal transduction, gene expression, spatial networks*? Stefan Schuster: "There are 2 main differences, i.e., whether there is mass flow or not [balance equations] and the fact that signal transduction systems are not at steady state." Albert Golbeter: "Yet the two can still be modeled by the same methods." Reinhart Heinrich: "The design is different; signal transduction pathways are subject to fewer constrained interactions." Ursula Kummer: "There is mass flow in signal transduction; still there is always also some sort of enzyme catalyzed reaction." John Doyle: A major difference is in the balance (conservation) laws, which apply to metabolic networks. Signal networks have to obey other laws, such as those related to robustness and fragility. More of these remain to be discovered; they are also softer. A metabolic network exhibits a net flow of mass. There is an additional fundamental difference in that a metabolic pathway is cell autonomous. Signaling pathways are very different between cell types." Todor Vujasinovic: "Metabolic pathways usually maintain homeostasis of the cell. They provide the possibility to live life at steady state. Signal transduction pathways have a different functionality, i.e. the one to adapt to changes, *i.e*. they have different logics." Igor Goryanin commented that metabolic pathways are not in steady state, whereas signal transduction pathways can be.

6. *Does Systems Biology have a role in rationalizing research, such as to specifty that phosphatases are more interesting targets than kinases? And are they?* Hans Westerhoff: "Yes, but it depends on which aspect of the dynamics are relevant." Igor Goryanin: "Drug companies prove that both are important." Ursula Klingműller: "Yes but this was already known. Systems Biology must move forward, *i.e.* further than this. Such further movement will depend on high quality experimental data."

7. *Is it justified to assume maximality/optimality. Is evolution complete enough? Is there, perhaps in this aspect an essential difference between the evolution of genetic networks in eukaryotes and those in prokaryotes?* Douglas Kell: "No, but one should ask this question only when one knows the objective function (knows what the relevant biological function is). Mostly evolution will not be able to keep up." Uri Alon: "I agree that in many cases one does not know the objective function, but in many laboratory conditions one can put an objective function in place. There is a trade off between being fully optimal and being able to evolve. Also: we do not have an optimum for a single thing; many objective functions may be relevant at the same time." Dennis Vitkup: "What is known is that the production of offspring matters. The problem is to map 'function' onto this fitness; this also depends on the type of conditions for evolution (test-tube versus real world conditions)."

### Issues left for Monday's 7 pm discussion:

8. *When studying a protein/role. Could Systems Biology help identifying the role if not everything is known?*
9. *How could/should Molecular Dynamics and single molecule biochemistry contribute to Systems Biology?*

## *2. Discussion on Tools for Systems Biology (Monday 7 pm plenary discussion)*

1. *How is Icat performing with membrane proteins quantitatively and qualitatively? And: wat is the price of proteomics?* Rüdi Äbersold: "This has nothing to do with Icat specifically. It has to do with solubility. The solution is to digest, dissolve and then go ahead. The price of proteomics is 200 k€ for an ion trap mass spectrometer and 40 k€ for chromatography."

2. *To what extent are we able to predict biological functions from structural network information? Do we need kinetic models? "*Metabolic stoichiometries may be used as a start; adding regulatory constraints would help." Kell: "It all depends on what one means with function." Matthias Reuss: "It depends on what you model, e.g. whether you are interested in dynamics." Hans Westerhoff: "The network stoichiometry method will often not work for drug target design because of homologies between host and parasite. The difference then should be sought in the kinetics. If gene A activates, and gene B inhibits according to the literature, only a kinetic model may figure out what the total effect is."

3. *How could quantitative proteomics help with determining activation / deactivation rates and (other) kinetic parameters* . Uri Alon: The problem with kinetic parameters will be solved soon because of technology development, *e.g.* RNA half lives will be determined through addition of inhibitors of mRNA

synthesis and then hybridization arraying as function of time. 10 years ahead we will have lots of parameter values, so better be prepared." Rüdi Äbersold: "Suicide inhibitors for various enzyme classes will be important. So will be isotope tagged covalent adduct formation only when the enzyme is active. The covalent adduct can then be isolated and the amount of *active* enzymes can be determined as function of time. In neurobiology proteins involved in the signaling are largely unknown. Therefore, so far, one cannot take into account his information. Therefore it is highly important to develop this field further."

4. *What could be the role of molecular modeling for understanding consequences of mutations, and ultimately for drug design*? "Molecular modeling could link local changes to global changes; e.g. effects of methyl groups on binding constants, of which thorugh systems biology methods then the effects on the funtioing of the system can be established." Ursula Kummer: "Yes, an example has been the. predicting of a kinetic parameter that was unknown, i.e. an association rate constant with superoxide radicals. Mattias Reuss: "An example of such a role is in detoxification. There is a difficulty to detoxify certain drugs, which is due to their structure. Molecular modeling helps to understand this." Benjamin Hall: "Normal mode analysis can predict the frequency of motions. Many mutations not themselves affecting binding still did affect this indirectly through affecting the dynamic modes of the protein."

5. *What could be the contribution of Systems Biology to the integration of knowledge up from molecular level to the various cellular, organism and patient levels?* Igor Goryanin: "The company 'Physiomics' claims that they can do this."

6. *Reliability of data in Systems Biology. How does gene dosage influence the Systems Biology approach?*
Douglas Kell: :The biggest problem here is that those data were usually taken under nonphysiological conditions (*e.g.* pH 10). One should now measure again, rather than to go back to early literature." Hans Westerhoff: "We need to be highly critical. We are often being seduced into not being this, using parameter values we are in critical need of when modeling and that we then dig up out of the literature. Standardization is necessary, also of quality control."

7. *Should we use single cell techniques, or should average data on cell populations be enough?* Uri Alon: "Average data on cell populations tell us a lot. But some things cannot be seen when averaging over cell populations. Individual cells do different things, even though genetically identical. A certain fraction of cells will not respond. This has been designed (through evolution). Another phenomenon where one needs to look at single cells is with oscillations. Oscillations out of phase would not be seen in populations. Also sharp transitions will not be seen in cell populations. Methods are being developed, *e.g*. image analysis, arrays of cells. Roland Eils: "One important reason for engaging in single cell analysis is that compartmentation matters. This would be missed on the basis of a population wide analysis. A 2$^{nd}$ reason is that many methods are now scalable, e.g. cell-array

technologies. Many of these arguments also apply to single molecules' technologies."

8. *How to deal with cell/individual diversity for personalised drug design?*
   See above.

9. *What is the sensitivity of systems behavior to parameter values versus its sensitivity to network structure? Which of the two is most important for dynamic behavior?* Roland Eils: "The simple answer is that both are important. If one has a dynamic model one can address both issues. *In silico* methodologies can be easier than experimental ones." Stefan Schuster: "When establishing a model one should first establish a structure." Reinhart Heinrich: "One may have a large model, but may still not be able to fit anything: it is not so that a large model will fit all behavior." Sune Danø: "It is possible to infer from dynamics a motif of moving variables, i.e. 'behavior'". ??: " When you know the structure of the network, you can already deduce that some things are not important for network function."

10. *How can we use existing databases effectively? What additional experiments are needed? Hypothesis versus data driven approaches, which is needed most?*
    See above.

    *How to design experiments to determine kinetic parameters for models?*
    Roland: this is computationally addressable; optimal experimental design methods exist. This is something to be moved forward. Kell: however we need many more parameters than known to the experimental design field. Sune Danø: main problem with building model is to get the entire behaviour. One would like to see in experiments MANY properties many variables, many phenomena in which the system as a whole responds. Reuss: Experimental design's Fischer info matrix requires good estimates of parameters, *i.e.* circular. Grosu: may be used to clarify the limits of parameter estimation methods. It is easier to develop the method of parameter estimation than to develop a new experimental technique.

## REMAINING QUESTIONS MONDAY

11. Uncertainty in parameters. How can we deal with this?
12. How SB could help to analyze microarray data?

## *3. Discussion on the Systems Biology of Unicellular Organisms (Wednesday 7 pm plenary discussion)*

1. *Bas Teusink: There have been many of talks about E. coli and yeast. Suppose you are from an industry and want a different organism (e.g. L. lactis ). Can one transfer the problem & knowledge? Could one identify classes of problems, which are shared by classes of organisms, like eukaryotic microorganisms, versus prokaryotic microorganisms? Or unicellular versus multicellular?* Mattias Reuss: "Tools can be transferred. The influence of structure on dynamic behavior can be transferred." Hans Westerhoff: "One should be able to compare, because the different organisms sometimes have the same optimization function. But , ???: "Even different strains of *E. coli* do not behave similarly." Stefan Hohmann: There are amazing differences even between strains of the same organism. Even pathways organized in a similar way show differences in architecture between closely related species." Bas Teusink: "Uri Alon said there are motifs in regulatory structures. This suggests that one can transfer knowledge."

2. *Uwe Sauers work showed that it takes time before precursors are translated; therefore the method cannot be used dynamically. What newer methods could be used?* Uwe Sauer: "The observation is correct. The choice was to use amino acid in biomass. However, there are lots of flux analysis methods used in mammalian cells; these are less robust (as a tool) however. Van Winden (Delft) develops a method based on free metabolites. When cells are not always in exponential phase, then that method may not work; because it is an average over heterogeneous behavior." Mattias Reuss: "We are working on a new technique that looks at the dynamics of isotopomers."

3. *Frank van Enckevort*: *Some information is already available in the literature, through genome sequence analysis. One now wishes to link this up to experimental/biochemical data about interactions/associations. However, who determines how good the latter data is? How about the validation?* Hans: There are different qualities/meanings in these data. For instance, with respect to interaction data there are three methods, which give highly different results, as they should; hexokinase and phosphofructokinase should score as interactive in text mining but not in a 2-hybrid assay. Bas Teusink: "Peer Boork in Heidelberg has a website where associations are inferred from the literature ( the associations suggested from different methods are compared on that website)."

4. *Benjamin Hall: There is a limited number of experimental methods, which often yield contradictory information. How to deal with this? And how to deal with the flood of data coming in? Isn't there a possibility that small/low scale data is being underestimated in terms of its importance? Should we not go for low throughput-high quality-strong focus data? An example is the issue of transporters. Should one not rather do a transport assay than a genome wide*

*array study?* Bas Teusink: "One should take high throughput data as a suggestion starting point."

5. *How can small labs (and how can they be motivated to) produce data that are comparable with data coming from large scale approaches.* Thomas Eißing was surprised that after Stelling's talk which said that high-thoughput data were of little help for producing the model, the PI's did not jump up and attack him for saying this, and defend high throughput data. Stefan Hohmann: "It depends a lot on the type of data. Of transcriptome data, the quality may be good enough. Protein-interaction data are only suggestive if at all useful." Uwe Sauer: "High throughput data's usefulness depends on the issue addressed. YSBN (i.e. the Yeast Systems Biology Network) groups the data that are most comparable. … One should separate 'connecting/correlating the data' from 'analyzing the data'. High throughput should be very useful for the former. High throughput data are also highly useful for probabilistic models." Marta Cascante: "To motivate the small labs: we have seen many examples of very small models or very limited experiments bringing highly important advances. Its is important to begin from a good question (*cf*. the posters)." Hans Westerhoff: "High throughput experiments are highly important for establishing weak correlations, hence weak mechanisms (because the large numbers produce statistical significance), not for simple strong mechanisms." Benjamin Hall: "High throughput approaches tend to be biased, e.g. by not looking at membrane proteins. One should not (as high throughput does) look only at the easy targets, e.g. membrane proteins." Here the target is not easy and one needs the low throughput analyses."

6. *Could we not extract more out of the floods of data? How do we get the hypotheses out? What to look for? Should we start an initiative to produce high throughput data in one central place, i.e. a data warehouse?*
   See above.

7. *How to fit parameters? Should we not produce common /standard methods to do this? Can one establish a method for parameter fitting for each type of question/model?* Sune Danø: "You cannot have just one easy way to fit parameters." Jacky Snoep: "From the Silicon cell point of view we fit on the individual components only. Systems Biology behavior should not be fitted, but predicted and then validated. This is important also for putting models together." Mattias Reuss: "What standard methods, numerical optimization techniques should one implement? One should have many. To comment on Jacky's statement; it is dangerous to have a strategy that focuses on individual enzymes; in the three dimensional freedom of a test tube one obtains different parameter values than the ones that pertain to the *in vivo* situation. There is a 3D versus 2 D case with different parameter values." Hans Westerhoff: "We cannot continue to have it that *in vivo* is different from *in vitro*. A molecule in a cell sits in an aqueous environment which may be crowded; so we do an *in vitro* experiment in an aqueous environment with macromolecular crowders." ???: "A relevant example is phosphofructokinase where different methods should be used in

parallel." Jacky Snoep: "One should remain precise about construction versus validation."

8. *With respect to Jörg Stelling's budding yeast model: How about translational regulation? How to approach this within the context of the existing models?* Stefan Hohmann: "One can do analyses on polysomal RNA and see which mRNA's are actually being translated." Hans Westerhoff: "This (i.e. whether there is transcriptional versus translational regulation, and how there is of each) is now all addressable by 'vertical genomics'-Systems Biology, and should be addressed and solved in the next 5 years for a number of organisms."

9. *Growth conditions are usually lab-like. Should we not try to establish more natural standard growth conditions, to be used as our standards?* Stefan Hohmann: "This has been discussed in a number of contexts, e.g. in the context of the YSBN." ???: "There are also major disadvantages to standardization; everyone is then doing the same thing. There is less opportunity to find new aspects." Uwe Sauer: "We should make things comparable, *e.g.* 5 students measuring growth rate of *E. coli* in a single lab typically get 5 different results; this we need to get under control through standardization. Chemostats and other conditions should be studied as well however." Bjørn Quistorff: "The keypoint is not to standardize; it is reproducibility that is essential. One should also discuss what the essential issues are that pertain to standardization." Hans Westerhoff: "One absolutely needs standardization. In some present models HXK from yeast at pH 5 sits together with PFK from erythrocytes at pH 7, which is absurd."

### Remaining Wednesday issues (due to lack of time)

10. *Frank van Enckevort*: Do we also want to do ecosystems biology, connected to cell systems bioogy?
11. Solving *E. coli* /yeast. The ultimate goal is to solve the human. Once you understand yourself what then remains to be discovered?
12. In order to solve an organism, we must combine data from different laboratories. How do we organize the definition of the metadata needed? What do we anticipate to be required in order to understand/deduce from the experimental data that what we need for Systems Biology?

## *4. Discussion on the Systems Biology of Multicellular Organisms (Thursday 6:15 pm plenary discussion)*

1. *The quantification of blots is a lot of hard work. How to convince biologists to really produce quantitative data?* Ursula Klingmüller: "It is entirely doable now, immuno-blotting with a special procedure for quantitation now exists. The

procedure can be used by other labs." Jens Timmer: "Many data points in time and small error bars are necessary for Systems Biology. If the modelers produce testable and good models, the biologists will produce the data." Matthias Reuss: "It is important indeed to convince the biologists by the strength of the model."

2. *Why has fluorescent microscopy not been used more, vis-à-vis spatial information in single cells?* Ursula Klingmüller: "Only a limited number of molecules can be visualized. In addition, the GFP that must be overproduced can have side effects, and it is difficult to resolve spatial problems." Matthias Reuss: "Can the resolution at the single cell level help with the problem of inhomogeneous populations? I should give the anser that a single cell is different from a population of cells. Quantitative fluorescent Microscopy and FACS can be used for bacteria and yeast, just as well as immuno-blotting." Ursula Klingmüller: "Sample preparation for microscopy might influence the result. Therefore this method is OK for qualitative but not for quantitative measurements." Response form ??": "GFP fusions might also entail artifacts. A proper protocol can be made for sample preparation in microscopy."

3. *Is it possible to model the interactions between pathogen and host, at a molecular level?* There are projects that try to do this. In the UK Systems Biology projects have received 6 M pound for these types of question.

4. *Can costs and efforts, equipment and protocols be shared?* Ursula Klingmüller: Systems Biology in Germany has its hepatocyte project, in which Standard Operating Procedures must be (and are) developed and shared. Equipment is not the problem." Matthias Reuss: "Will biologist accept such common protocols? Groups should then not further optimize the protocols, which is what they usually do." Ursula Klingmüller: "Protocols can be changed and further developed but this must then again be communicated to al the groups." Matthias Reuss: "A course on wet methods and theory and communication between the groups should be organized. Systems Biology should develop a database with standard methods, such a system exists in Russia."

5. *Immunology as a subject in a next course*

6. *Submission of data to databases. What is the reason that people don't do it?* Lazyness, is there a point in doing it? In the future there will be more curation of data to ensure quality. Is there an advantage for the experimentalist to submit his/her data? Journals should make it obligatory that data are made available. The principle does work for DNA-sequence data. Micro-array data must be submitted to database if published. Not so many databases are available for gene transcription and signal transduction. Schemes for pathway interconnections should be submitted. Too many of the available databases are not structured clearly, a number of main databases should be selected.

7. *Compartmentalization of components in the cell, what models deal with this aspect, is it a problem to model this, or is it absence of data?* No problem for modelers, transport processes must be included. Experimental data has been discussed. Should single cell measurements be used? Not necessary sometimes chemical measurements on populations can be used.

8. *Is it possible to bridge the gap between medicine and molecular biology? Will models be useful to treat patients?* Jens Timmer: "This issue holds a big promise. It might take a while, but eventually we should deliver such models. Currently this is more a hype than a reality. Matthias Reuss: "We must be successful in this aspect, we need these success-stories. Systems Biology models on pharmacological aspects have already shown a usage in this field, have they?"