MINIREVIEW

Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*

Jens Nielsen & Michael C. Jewett

Center for Microbial Biotechnology, BioCentrum-DTU, Technical University of Denmark, Lyngby, Denmark

Abstract

Industrial biotechnology is a rapidly growing field. With the increasing shift towards a bio-based economy, there is rising demand for developing efficient cell factories that can produce fuels, chemicals, pharmaceuticals, materials, nutraceuticals, and even food ingredients. The yeast *Saccharomyces cerevisiae* is extremely well suited for this objective. As one of the most intensely studied eukaryotic model organisms, a rich density of knowledge detailing its genetics, biochemistry, physiology, and large-scale fermentation performance can be capitalized upon to enable a substantial increase in the industrial application of this yeast. Developments in genomics and high-throughput systems biology tools are enhancing one's ability to rapidly characterize cellular behaviour, which is valuable in the field of metabolic engineering where strain characterization is often the bottleneck in strain development programmes. Here, the impact of systems biology on metabolic engineering is reviewed and perspectives on the role of systems biology in the design of cell factories are given.

Introduction

The yeast *Saccharomyces cerevisiae* serves as a very important model organism for studying the molecular mechanisms underlying complex diseases like cancer, diabetes, and various metabolic disorders. For this reason, genome sequencing was undertaken at an early stage. Chromosome III of *S. cerevisiae* was the first complete chromosome to be sequenced for any organism (Oliver *et al.*, 1992), and the completion of the entire genome sequence in 1996 represented the first available genome for any eukaryote (Goffeau *et al.*, 1996). Because the genomic source code only provides an inventory of parts, functional genomics tools have also been developed using this yeast as a vehicle for observing and quantifying cellular behaviour. These tools include: transcriptome analysis (Lashkari *et al.*, 1997), proteome analysis (Zhu *et al.*, 2001), metabolome analysis (Villas Boas *et al.*, 2005a; Jewett *et al.*, 2006), flux analysis (Sauer, 2006), interactome analysis (Uetz *et al.*, 2000; Lee *et al.*, 2002; Harbison *et al.*, 2004), and locasome analysis (Huh *et al.*, 2003), among others. Complementing today's systems biology tools, extensive compendiums of high-throughput data are available both from specific studies (Hughes *et al.*, 2000)

and in databases (Table 1). The excessive amounts of data available for *S. cerevisiae*, both at the global level and at the molecular level, make this yeast well suited for a coordinated effort in systems biology, where the objective is to obtain a quantitative description of cellular processes, global mapping of all key quantitative interactions within the cell, and ultimately, to predict how and why cells function the way they do (Mustacchi *et al.*, 2006).

As mentioned above, much of the development in the field of genomics and systems biology of yeast is driven by the use of this organism as a model for studying human diseases or human pathogens. *Saccharomyces cerevisiae* is, however, also a very important cell factory. While traditional applications of this yeast include the production of beer, spirits, wine, and bread, the advent of genetic engineering and recombinant DNA (rDNA) technology allowed new opportunities for exploiting *S. cerevisiae* for many compelling bio-based applications. Today, for example, *S. cerevisiae* is used for the commercial production of pharmaceutical protein products like insulin, and several vaccines, including hepatitis and papillomavirus. Looking forward, the use of *S. cerevisiae* as a cell factory for the production of rDNA proteins is likely to become even more important.

Table 1. Internet resources for yeast systems biology data

Database	Website
Saccharomyces Genome Database	http://www.yeastgenome.org/
Stanford MicroArray Database	http://genome-www5.stanford.edu/
Yeast deletion project	http://www-sequence.stanford.edu/group/yeast_deletion_project/data_sets.html
Transcriptional regulatory code of yeast	http://web.wi.mit.edu/young/regulatory_code/
MIPS comprehensive yeast genome database	http://mips.gsf.de/genre/proj/yeast/
Comprehensive systems biology database	http://csbdb.mpimp-golm.mpg.de/
Proteome bioknowledge database	http://www.proteome.com/control/tools/proteome
Yeast GFP fusion localization database	http://yeastgfp.ucsf.edu/
Protein interaction database	http://www.ebi.ac.uk/intact/site/index.jsf
General repository for interaction datasets	http://www.thebiogrid.org/
Yeast search for transcriptional regulators and consensus tracking	http://www.yeastract.com/
Cold Spring Harbor Laboratory	www.reactome.org/

Pioneering breakthroughs in engineering the yeast *Pichia pastoris* to have human glycosylation pathways have enabled the production of homogeneously glycosylated proteins at high levels (Gerngross, 2005; Hamilton *et al.*, 2006; Li *et al.*, 2006). It is expected that transferring this technology to *S. cerevisiae* will enable recombinant glycoprotein production at much lower costs and with higher efficacy, than cell cultures that are conventionally used to express glycoproteins. This promises to expand the antibody market as well as markets for other humanized glycoproteins that today can only be produced at high costs. Another traditional application of *S. cerevisiae* is the production of ethanol (often referred to as bioethanol to distinguish it from ethanol being produced from petrochemicals). The production of bioethanol has experienced a dramatic increase in the last couple of years. With increasing oil prices and for numerous geopolitical reasons, there is financial, social, and political pressure to increase the production of bioethanol to be used as a renewable fuel. Besides the use of *S. cerevisiae* as a cell factory for the production of biofuels, this organism has also been exploited for the production of other chemicals like organic acids, e.g. lactic acid (Porro *et al.*, 1999; Ishida *et al.*, 2006) and pyruvate (van Maris *et al.*, 2003), glycerol (Geertman *et al.*, 2006), and more complex natural products, e.g. isoprenoids (Yamano *et al.*, 1994; Ro *et al.*, 2006; Shiba *et al.*, 2006) and polyketides (Kealey *et al.*, 1998; Mutka *et al.*, 2006; Wattanachaisaereekul *et al.*, 2007). With these developments, the market value of products derived from fermentations with *S. cerevisiae* is expected to increase further in the future, and much above the general market growth (see Fig. 1).

With the extensive fundamental research carried out on *S. cerevisiae* and the substantial industrial interest in this organism as a cell factory, it is obvious to consider the exploitation of the solid knowledge base from genomics and systems biology for future design of improved cell factories. A major hurdle in this exploitation is, however, that many of
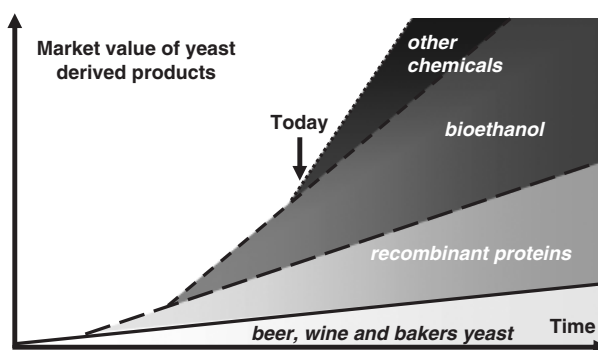

Fig. 1. Illustration of the growing market of yeast biotechnology. The use of *Saccharomyces cerevisiae* for the production of recombinant proteins is expected to grow substantially as more and more products can be produced using yeast as an expression system. The bioethanol market is expected to increase much beyond the current level. Yeast is also expected to be exploited for the production of a wide range of other chemicals in the future.

the high-throughput experimental techniques and bioinformatics algorithms for analysis of these data are not well suited for identification of the rather small adjustments that might occur in metabolism during an industrial fermentation process, e.g. during a fed-batch process used for the production of a recombinant protein. In this review, it will be briefly discussed as to how some of the high-through experimental techniques reported for analysis of yeast can be used in the field of industrial biotechnology. First, however, a definition of systems biology and metabolic engineering will be given.

Systems biology

There are many definitions of systems biology, but most of these contain elements such as mathematical modelling, global analysis (or ome analysis), the whole system is more
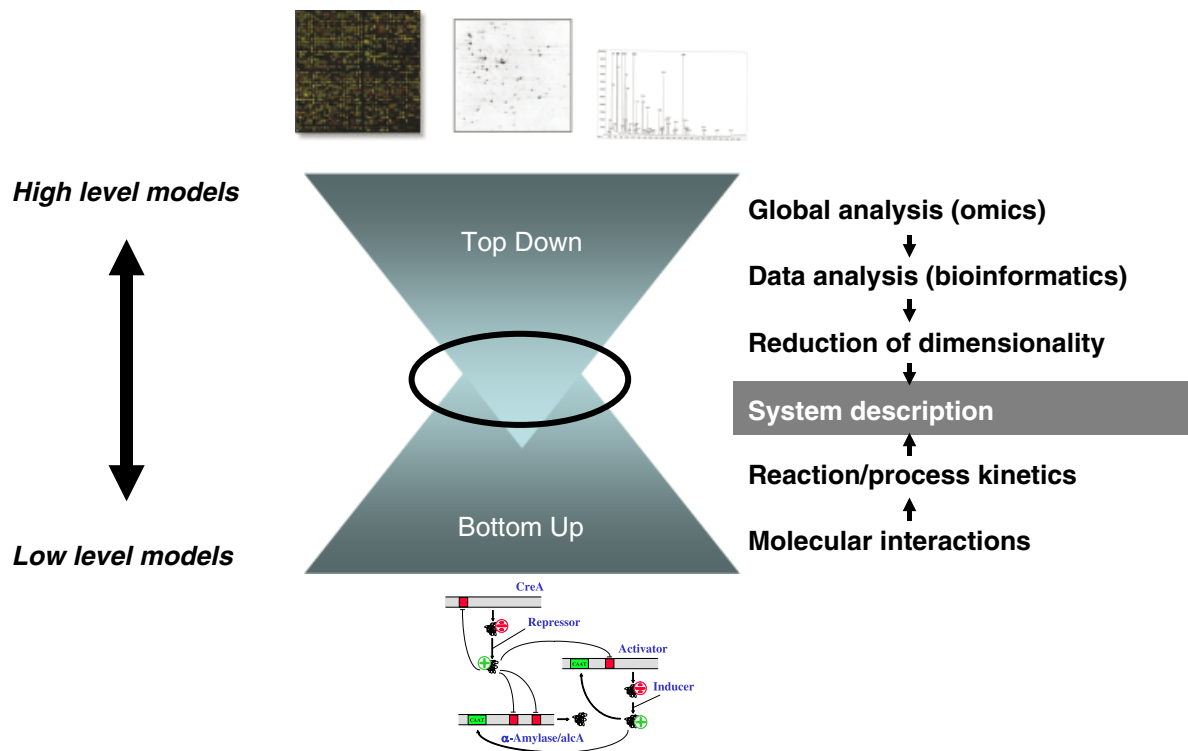
Fig. 2. How top-down and bottom-up systems biology meet in terms of providing a quantitative description of a biological system. In the top-down approach, high-throughput data are applied for identification of structures, connectivity, and possible information on the quantitative interaction between different components. In the bottom-up approach, the system is reconstructed based on biological knowledge, e.g. on molecular interactions.

than the sum of its parts, mapping of interactions between cellular components, and quantification of dynamic responses in living cells (Ideker *et al.*, 2001; Kitano, 2002; Brent, 2004; Stephanopoulos *et al.*, 2004; Kirschner, 2005; Barrett *et al.*, 2006; Bruggeman & Westerhoff, 2007). In most cases, the objective of systems biology is to obtain a quantitative description of the biological system under study, and this quantitative description may be in the form of a mathematical model. In some cases, the model may be the final result of the study, i.e. the model captures key features of the biological system and can hence be used to predict the behaviour of the system under conditions different from those used to derive the model. In other cases, mathematical modelling rather serves as a tool to extract information of the biological system, i.e. to enrich the information content in the data. There is not necessarily a conflict between the two, and generally, mathematical modelling goes hand in hand with experimental work. This partnership exemplifies the view of the essence of systems biology: *to obtain new insight into the molecular mechanisms occurring in living cells or sub-systems of living cells for predicting the function of biological systems through the combination of mathematical modelling and experimental biology.* This does not say anything about the use of global

data, e.g. transcriptome or proteome data, and clearly there are many systems biology studies that do not rely on global data. Mathematical models have, however, been shown to be particularly useful for analysis of global data, as the complexity and integrative nature of biological systems makes it difficult to extract information on molecular processes from global data without the use of models as either scaffolds for the analysis or for hypothesis driven analysis of the data.

From the above, it is clear that different mathematical models play a central role in systems biology. The type of model that one will use in a systems biology study will, however, depend completely on the objective of the study. Often, one distinguishes between top-down systems biology and bottom-up systems biology (see Fig. 2). Top-down systems biology is basically a data-driven process, where new biological information is extracted from large data sets. The models used in this kind of study can be soft models like neural networks, graphs, or even statistical models. In many cases, there is not a specific hypothesis and the analysis may be rather inductive (Kell & Oliver, 2004), but often the initial analysis leads to some kind of hypothesis that then leads to establishment of a course model that is then evaluated against the data. An excellent example of this kind of modelling is a study on the yeast cell cycle, where de Lichtenberg *et al.*

(2005) found from analysis of various ome data that key protein complexes are assembled during the cell cycle and not all proteins within these complexes have a cyclic transcription pattern. Bottom-up systems biology is, on the other hand, based on the presence of very detailed knowledge, and this knowledge is then translated into a mathematical formulation that is then used to simulate the behaviour of the system. Generally there is not enough knowledge available to build detailed mechanistic models, and an important element of bottom-up systems biology is therefore an evaluation of different model structures. Very good examples of bottom-up systems biology include detailed modelling of metabolic pathways in yeast (Mauch *et al.*, 2000; Müller, 2006), the High Osmolarity Glycerol (HOG) signal transduction pathway (Klipp *et al.*, 2005), and the α project (Lok & Brent, 2005).

It is difficult to classify mathematical models applied in either top-down or bottom-up systems biology as many different types of models may be used, e.g. models based on ordinary differential equations, stochastic models, stoichiometric models, and graph models. One approach to classify models has, however, been given by Ideker & Lauffenburger (2003), who classified mathematical models used in systems biology as:

(1) high-level models that describe the components and their interactions, and

(2) low-level models that describe the molecular mechanisms underlying interactions of the system components.

Clearly, these classifications are positioned at two extremes, but basically low-level models refer to a bottom-up approach where the system is reconstructed from quantifying all the interactions and the high-level models refer to a top-down approach where structures, interactions, and their strengths are extracted from global data (see Fig. 2). Most bottom-up driven models only describe a subset of the complete biological system, as there is simply not enough quantitative information available to include interactions between all the components within the cell. There is, however, one type of bottom-up model that is fairly global in its approach: a metabolic network model. Metabolic network models are based on collecting the stoichiometry for all metabolic reactions into a stoichiometric matrix. Through the use of flux balance analysis, where the fluxes are constrained such that all intracellular metabolites balance, and linear programming, it is possible to use these stoichiometric models for simulation of growth and product formation (Famili *et al.*, 2003; Forster *et al.*, 2003; Price *et al.*, 2004). As metabolic pathways and architecture are well established, it is possible to expand this modelling concept to cover practically all parts of the metabolism, and it may even be possible to expand these models to cover regulation (Barrett & Palsson, 2006; Herrgard *et al.*, 2006). Thus, even though these models are bottom-up driven, they actually provide considerable information about the connectivity between the different enzymes participating in the metabolic network (Barabasi & Albert, 1999). Therefore, metabolic models are unique as they fulfill the criteria of both high- and low-level models. Genome-scale metabolic models provide a framework for organizing and integrating x-ome data.

Metabolic engineering

Metabolic engineering is an applied science focusing on developing new cell factories or improving existing cell factories (Bailey, 1991; Stephanopoulos & Vallino, 1991; Nielsen, 2001; Tyo *et al.*, 2007). There are several definitions, but most of these are consistent with: *the use of genetic engineering to perform directed genetic modifications of cell factories with the objective to improve their properties for industrial application.* In this definition the word improve is to be interpreted in its broadest sense, i.e. it also encompasses the insertion of completely new pathways with the objective to produce a heterologous product in a given host cell factory. Metabolic engineering is an enabling science, and distinguishes itself from applied genetic engineering by the use of advanced analytical tools for identification of appropriate targets for genetic modifications and possibly even the use of mathematical models to perform *in silico* design of optimized cell factories. Metabolic engineering is therefore often seen as a cyclic process (Nielsen, 2001), where the cell factory is analysed and based on this an appropriate target is identified (the design phase). This target is then experimentally implemented and the resulting stain is analysed again. Thus, similar to systems biology, metabolic engineering involves a continuous iteration between design and experimental work. In recent years, there has been increasing focus on using mathematical models for design (Burgard *et al.*, 2003; Pharkya *et al.*, 2004; Patil *et al.*, 2005b). Hereby, it is expected that metabolic engineering will become faster and more efficient through the development of robust and reliable mathematical models describing the function of cell factories. To be fair, however, it should be noted that a large percentage of the successes in using microorganisms as cell factories have thus far occurred without detailed modelling.

One obvious avenue in metabolic engineering is the heterologous expression of complete biosynthetic pathways leading towards interesting and valuable products. By introducing entire pathways, it is possible to either produce known compounds more efficiently or, even through combinatorial biosynthesis, produce completely new chemical entities that may serve as possible new products, such as food ingredients, nutraceuticals, or pharmaceuticals. There are many examples of exploiting yeast as a cell factory for the production of different chemical entities (Table 2).

The insertion of heterologous pathways for the production of valuable products, in general, does not by itself result

Table 2. Examples of production of heterologous products in yeast

Type of product	Specific application	References
Hormones	Production of insulin and insulin precursors. Through engineering of leader sequences, the productivity of protein production has been increased substantially	Kjeldsen (2000)
Vaccines	Production of hepatitis vaccines. Through expression of a virus surface protein in yeast, an efficient vaccine has been developed	Ishida *et al.* (2006)
Organic acids	Production of lactic acid. Through expression of a heterologous lactic acid dehydrogenase in yeast, lactic acid production was achieved	Porro *et al.* (1999)
Sesquiterpenes	Through heterologous expression of plant genes in yeast, many different sesquiterpenes have been produced. This includes the anti-malarial drug precursor artemisinic acid	Ro *et al.* (2006)
Carotenoids	Through expression of bacterial genes in yeast, β-carotene and lycopene were produced	Yamano *et al.* (1994)
Diterpenoids	Through expression of 10 plant genes in yeast, a major part of the biosynthetic route towards taxol was reconstructed	DeJong *et al.* (2006)
Polyketides	Through combined expression of a polyketide synthetase encoding gene together with an activating enzyme, 6-MSA could be produced in high titers in yeast	Kealey *et al.* (1998) Mutka *et al.* (2006) Wattanachaisaereekul *et al.* (2007)

Table 3. Frequency of precursor metabolites and cofactors in a *Saccharomyces cerevisiae* genome scale model*

Precursor metabolite	No of reactions	Cofactor	No of reactions
Glucose-6P	16	ATP	188
Fructose-6P	18	ADP	146
Ribose-5P	20	NADH	65
Erythrose-4P	6	NAD^+	78
Glyceraldehyde-3P	13	NADPH	78
3-Phosphoglycerate	6	$NADP^+$	86
Phosphoenolpyruvate	12		
Pyruvate	27		
Acetyl-CoA	32		
2-Oxoglutarate	38		
Succinyl-CoA	3		
Oxaloacetate	12		

*The data are taken from the metabolic model developed by Forster *et al.* (2003).

in high-level production of the desired product. In order to improve the yield or productivity, it is generally required to improve the supply of the precursor metabolites and the cofactors required for biosynthesis of the product. All macromolecules and smaller metabolites in nature are derived from only 12 precursor metabolites, i.e. glucose-6P, fructose-6P, ribose-5P, erythrose-4P, glyceraldehyde-3P, 3P-glycerate, phosphoenolpyruvate, pyruvate, acetyl-CoA, 2-oxoglutarate, succinyl-CoA, and oxaloacetate. Besides these 12 precursor metabolites the biosynthesis of metabolites and proteins requires the use of cofactors like NADPH, NADH, and ATP. The frequent use of the 12 precursor metabolites and the cofactors in cellular reactions is illustrated in Table 3. As shown, about 16% of the almost 1200 reactions in a genome-scale metabolic model of *S. cerevisiae* involve ATP (Forster *et al.*, 2003). Not only do cofactors knit

different parts of the metabolism together but also the 12 precursor metabolites participate in a large number of reactions. In fact, the metabolic network of *S. cerevisiae* forms a very dense metabolic graph of enzymes and metabolites, with an average diameter of about 5. This means that it is possible to jump from any enzyme to any other enzyme in the network in only five steps (connecting through any other enzyme or metabolite) (Patil & Nielsen, 2005a) (Fig. 3).

Tight coupling of many different biochemical pathways imposes a major constraint when the objective is to increase the flux towards a specific precursor metabolite. As a result, redirection of fluxes requires a fundamental understanding of the complete network operation and not only on how the fluxes distribute over a few branch points. For this purpose, methods for flux quantification are extremely useful. Metabolic fluxes can either be estimated through the use of flux balance analysis (Nissen *et al.*, 1997; Price *et al.*, 2003) or through the use of C13-labelled substrate feeding followed by analysis of the labelling patterns in intracellular metabolites (Gombert *et al.*, 2001; Sauer, 2006). Owing to their abundance and stability, C13-based methods have conventionally used proteinogenic amino acids to detect labelling patterns. Recently, however, methods for direct analysis in the free pool of metabolites have been developed (van Winden *et al.*, 2005; Wiechert & Noh, 2005; Noh & Wiechert, 2006). Besides providing general information on how the metabolic network is operating under different growth conditions, metabolic flux analysis is very well suited for analysis of the effects of growth on different media (dos Santos *et al.*, 2003a), specific mutations (dos Santos *et al.*, 2003b), and screening of different mutants (Raghevendran *et al.*, 2004; Blank *et al.*, 2005). Beyond applications in yeast, this technology has also been demonstrated to be very useful for analysis of a large collection of *Bacillus subtilis* mutants
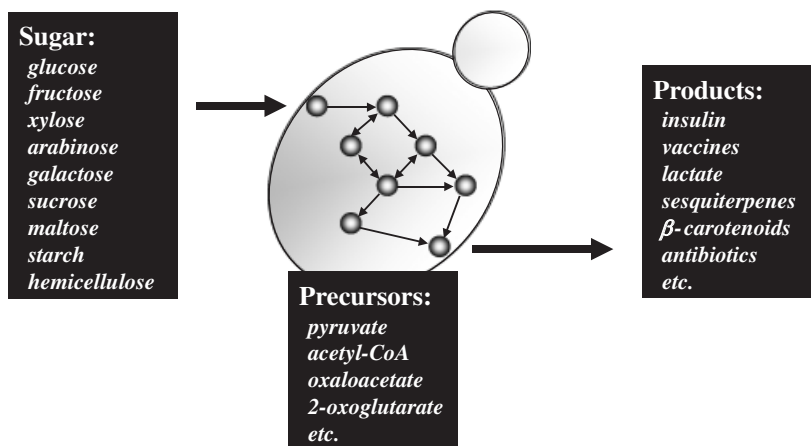
Sugar:
glucose
fructose
xylose
arabinose
galactose
sucrose
maltose
starch
hemicellulose

Precursors:
pyruvate
acetyl-CoA
oxaloacetate
2-oxoglutarate
etc.

Products:
insulin
vaccines
lactate
sesquiterpenes
β-carotenoids
antibiotics
etc.

Fig. 3. Illustration of how a yeast cell factory is used to convert different raw materials (sugars) into a wide range of different products. In all cases the carbon passes through a set of 12 precursor metabolites, which form the building blocks for all organic chemicals found in nature.

(Fischer & Sauer, 2005). Thus, flux analysis today represents a standard technique for rapid phenotypic characterization of metabolically engineered strains, and this tool is likely to gain even wider use in the future.

Another important tool for metabolic characterization is the analysis of the complete set of intracellular and extracellular metabolites associated with a cell, or metabolome analysis (Jewett *et al.*, 2006). Having already shown utility in drug discovery, strain classification (Allen *et al.*, 2003), and functional genomics (Raamsdonk *et al.*, 2001), metabolome analysis is emerging as powerful tool in systems biology research. One of the major challenges currently being addressed is ensuring robust and unbiased quantification of a large number of metabolites. It is inherently difficult to measure intracellular metabolites quantitatively as the very low time constants for turnover of these metabolites require rapid quenching of metabolism (Villas-Boas *et al.*, 2005b). A requirement for the quenching process is that the metabolites do not leak out of the cells, and it is difficult to find a method that can be generally applied to measure different types of metabolites (Villas-Boas *et al.*, 2005c). However, in recent years several robust methods have been developed for analysis of specific groups of metabolites, e.g. sugar phosphates (Gonzalez *et al.*, 1997; Smits *et al.*, 1998; Mashego *et al.*, 2006) and amino and nonamino organic acids (Villas-Boas *et al.*, 2005a). In addition to dynamic developments in refined analytical techniques, advances in internal standardization, another main challenge in quantitative metabolome analysis, are also paving the way for more robust measurements. Heijnen and colleagues have developed an approach that uses extracts from ^{13}C-saturated microbial cultivations to provide an internal standard for all intracellular metabolites to be quantified (Mashego *et al.*, 2004; Wu *et al.*, 2005). This work has created a platform that is independent of ion-suppression effects, of metabolite modifications during extraction, and of variations in instrument response.

Examples of genomics and systems biology studies of relevance for metabolic engineering

Owing to the high connectivity of the different metabolic reactions within the metabolic network, there has been considerable interest in exploiting tools from functional genomics for mapping of global regulatory structures or even using high-throughput experimental techniques provided by the various omics for dissecting how fluxes through different branches of the metabolic network are controlled. This can only be done through the combination of experimental data and mathematical models of one kind or the other. Westerhoff and colleagues have extended the concept of metabolic control analysis for distributing flux control at the hierarchical and metabolic levels (ter Kuile & Westerhoff, 2001; Rossell *et al.*, 2006). Flux control at the hierarchical level means that the flux through a given reaction is controlled by transcription, translation or posttranslational modifications, i.e. modification of the active enzyme concentration, whereas flux control at the metabolic level indicates that the flux is controlled through interaction between the enzyme and the metabolites. To identify coregulated subnetworks within the metabolic network, along with so-called reporter metabolites, the network structure provided by a genome-scale metabolic network can be combined with transcriptome data (Patil & Nielsen, 2005a). In this analysis, reporter metabolites represent hotspots in the metabolic network where there is the most statistically significant transcriptional change between conditions or strains. The concept of reporter metabolites has been extended further to use metabolome data for identification of reporter reactions (Cakir *et al.*, 2006). By mapping reporter reactions with reporter metabolites, it was possible to categorize reactions into metabolically or hierarchically regulated categories (Cakir *et al.*, 2006). Another type of multilevel analysis for capturing how information stored at the genetic level is translated into phenotypic landscapes has

been pursued by the group of Pronk by analyzing the response of yeast to various environmental perturbations. Vertical genomics strategies, integrating molecular measurements from multiple layers of the cellular hierarchy for a particular functional pathway (e.g. mRNA, proteins, and metabolites), have clarified the sequential response of glycolytic reactions in *S. cerevisiae* to a sudden relief of glucose limitation (Kresnowati *et al.*, 2006), and from a broader perspective, provided an insight into transcriptional control using proteomics (Kolkman *et al.*, 2006). While these new approaches for identifying key regulatory structures controlling cellular behaviour hold significant promise to impact metabolic engineering, they have yet to be exploited.

With transcriptome analysis being the most mature and well-implemented omics technique, there has been much focus on whether this can be used to provide information on how the metabolic network is operating (Jewett *et al.*, 2005). One successful application of transcriptome analysis for identification of metabolic engineering targets was the improvement of galactose uptake in *S. cerevisiae*. Through genome-wide transcription analysis of several different mutants with improved galactose uptake, Bro *et al.* (2005) identified that there was up-regulation of PGM2 encoding phosphoglucomutase. By overexpressing the PGM2 gene, galactose uptake could be increased by 80%. Owing to the presence of regulation at the level of translation and at the metabolic level, there is no direct correlation between transcripts and metabolic fluxes (Moxley *et al.*, submitted). However, Stelling *et al.* (2002) introduced so-called control effective fluxes, which are functions of the different elementary flux modes (Schuster *et al.*, 2000) in the metabolic network, and showed that the control effective fluxes correlate quite well with transcription in *Escherichia coli*. In a later study, this was also shown to be the case for *S. cerevisiae* when there was a shift on growth at different carbon sources (Cakir *et al.*, 2004). In order to further look into the possible correlation between metabolic fluxes and transcript levels, Regenberg *et al.* (2006) performed transcriptome analysis at different specific growth rates in chemostat cultures (glucose limited). They identified which genes are decreasing and increasing for increasing specific growth rates. Besides mapping all genes related to the Crabtree effect, i.e. the onset of fermentative metabolism under aerobic growth conditions, they found that genes responsible for catabolism of C2 carbon sources, e.g. ethanol, are transcribed at low specific growth rates. This was further confirmed in a study by Vemuri *et al.* (2007), who, through heterologous expression of oxidases in both the cytosol and in the mitochondria, showed that there is indeed excess capacity of the TCA cycle, but that the onset of the Crabtree effect is caused by lack of capacity for oxidation of NADH in the mitochondria.

Future impact of systems biology on metabolic engineering

Today, only a few examples on how systems biology has impacted metabolic engineering and industrial biotechnology have been seen. However, the introduction of high-throughput experimental techniques has clearly enabled much faster progress in terms of phenotypic characterization of different mutants. In the future, when more advanced mathematical models and bioinformatics algorithms specifically suited for metabolic engineering have been developed, the value of using high-throughput experimental techniques for mapping detailed phenotypes will clearly increase. This is exemplified by the introduction of an algorithm for identification of reporter metabolites (Patil & Nielsen, 2005a), which enables rapid identification of hot-spots in the metabolism based on transcriptome data. It is expected that mathematical models will be used more extensively in the design of metabolic engineering strategies, particularly as recent results have shown that the predictive power of metabolic models is sufficiently good to allow for identification of metabolic engineering targets (Bro *et al.*, 2006). To further capitalize on yeast systems biology in the field of industrial biotechnology, it is, however, important that metabolic models are extended to include regulation, as it is often possible to de-regulate fluxes through engineering of regulatory structures (Ostergaard *et al.*, 2000). For this purpose, detailed kinetic models of signal transduction pathways are expected to be useful, but such detailed models are not necessarily required for identification of metabolic engineering targets, as information about connectivity and type of interaction, i.e. boolean-type information, is often sufficient. At least in the short term, metabolic engineering is likely to benefit more from top-down systems biology than bottom-up systems biology. In the long run, however, it will be desirable to have access to detailed kinetic models as this will enable identification of advanced metabolic engineering strategies that involves fine-tuning activities of specific pathways.

Acknowledgements

References

Allen J, Davey HM, Broadhurst D, Heald JK, Rowland JJ, Oliver SG & Kell DB (2003) High-throughput classification of yeast

mutants for functional genomics using metabolic footprinting. *Nat Biotechnol* **21**: 692–696.

Bailey JE (1991) Toward a science of metabolic engineering. *Science* **252**: 1668–1675.

Barabasi AL & Albert R (1999) Emergence of scaling in random networks. *Science* **286**: 509–512.

Barrett CL & Palsson BO (2006) Iterative reconstruction of transcriptional regulatory networks: an algorithmic approach. *PLoS Comput Biol* **2**: e52.

Barrett CL, Kim TY, Kim HU, Palsson BO & Lee SY (2006) Systems biology as a foundation for genome-scale synthetic biology. *Curr Opin Biotechnol* **17**: 488–492.

Blank LM, Kuepfer L & Sauer U (2005) Large-scale ^{13}C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol* **6**: R49.

Brent R (2004) A partnership between biology and engineering. *Nat Biotechnol* **22**: 1211–1214.

Bro C, Knudsen S, Regenberg B, Olsson L & Nielsen J (2005) Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. *Appl Environ Microbiol* **71**: 6465–6472.

Bro C, Regenberg B, Forster J & Nielsen J (2006) In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab Eng* **8**: 102–111.

Bruggeman FJ & Westerhoff HV (2007) The nature of systems biology. *Trends Microbiol* **15**: 45–50.

Burgard AP, Pharkya P & Maranas CD (2003) OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng* **84**: 647–657.

Cakir T, Kirdar B & Ulgen KO (2004) Metabolic pathway analysis of yeast strengthens the bridge between transcriptomics and metabolic networks. *Biotechnol Bioeng* **86**: 251–260.

Cakir T, Patil KR, Onsan Z, Ulgen KO, Kirdar B & Nielsen J (2006) Integration of metabolome data with metabolic networks reveals reporter reactions. *Mol Syst Biol* **2**: 50.

Dejong JM, Liu Y, Bollon AP, Long RM, Jennewein S, Williams D & Croteau RB (2006) Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* **93**: 212–224.

Famili I, Forster J, Nielsen J & Palsson BO (2003) *Saccharomyces cerevisiae* phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proc Natl Acad Sci USA* **100**: 13134–13139.

Fischer E & Sauer U (2005) Large-scale *in vivo* flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. *Nat Genet* **37**: 636–640.

Forster J, Famili I, Fu P, Palsson BO & Nielsen J (2003) Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res* **13**: 244–253.

Geertman JM, van Maris AJ, van Dijken JP & Pronk JT (2006) Physiological and genetic engineering of cytosolic redox metabolism in *Saccharomyces cerevisiae* for improved glycerol production. *Metab Eng* **8**: 532–542.

Gerngross T (2005) Production of complex human glycoproteins in yeast. *Adv Exp Med Biol* **564**: 139.

Goffeau A, Barrell BG, Bussey H *et al.* (1996) Life with 6000 genes. *Science* **274**: 563–547.

Gombert AK, Moreira dos SM, Christensen B & Nielsen J (2001) Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J Bacteriol* **183**: 1441–1451.

Gonzalez B, Francois J & Renaud M (1997) A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast* **13**: 1347–1356.

Hamilton SR, Davidson RC, Sethuraman N *et al.* (2006) Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* **313**: 1441–1443.

Harbison CT, Gordon DB, Lee TI *et al.* (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**: 99–104.

Herrgard MJ, Lee BS, Portnoy V & Palsson BO (2006) Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*. *Genome Res* **16**: 627–635.

Hughes TR, Marton MJ, Jones AR *et al.* (2000) Functional discovery via a compendium of expression profiles. *Cell* **102**: 109–126.

Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS & O'Shea EK (2003) Global analysis of protein localization in budding yeast. *Nature* **425**: 686–691.

Ideker T & Lauffenburger D (2003) Building with a scaffold: emerging strategies for high- to low-level cellular modeling. *Trends Biotechnol* **21**: 255–262.

Ideker T, Galitski T & Hood L (2001) A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet* **2**: 343–372.

Ishida N, Saitoh S, Ohnishi T, Tokuhiro K, Nagamori E, Kitamoto K & Takahashi H (2006) Metabolic engineering of *Saccharomyces cerevisiae* for efficient production of pure L-(+)-lactic acid. *Appl Biochem Biotechnol* **129–132**: 795–807.

Jewett MC, Oliveira AP, Patil KR & Nielsen J (2005) The role of high-throughput transcriptome analysis in metabolic engineering. *Biotechnol Bioproc Eng* **10**: 385–399.

Jewett MC, Hofmann G & Nielsen J (2006) Fungal metabolite analysis in genomics and phenomics. *Curr Opin Biotechnol* **17**: 191–197.

Kealey JT, Liu L, Santi DV, Betlach MC & Barr PJ (1998) Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc Natl Acad Sci USA* **95**: 505–509.

Kell DB & Oliver SG (2004) Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *Bioessays* **26**: 99–105.

Kirschner MW (2005) The meaning of systems biology. *Cell* **121**: 503–504.

Kitano H (2002) Systems biology: a brief overview. *Science* **295**: 1662–1664.

Kjeldsen T (2000) Yeast secretory expression of insulin precursors. *Appl Microbiol Biotechnol* **54**: 277–286.

Klipp E, Nordlander B, Kruger R, Gennemark P & Hohmann S (2005) Integrative model of the response of yeast to osmotic shock. *Nat Biotechnol* **23**: 975–982.

Kolkman A, Daran-Lapujade P, Fullaondo A, Olsthoorn MM, Pronk JT, Slijper M & Heck AJ (2006) Proteome analysis of yeast response to various nutrient limitations. *Mol Syst Biol* **2**: 2006.26.

Kresnowati MT, van Winden WA, Almering MJ, ten Pierick A, Ras C, Knijnenburg TA, Daran-Lapujade P, Pronk JT, Heijnen JJ & Daran JM (2006) When transcriptome meets metabolome: fast cellular responses of yeast to sudden relief of glucose limitation. *Mol Syst Biol* **2**: 49.

ter Kuile BH & Westerhoff HV (2001) Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Letters* **500**: 169–171.

Lashkari DA, DeRisi JL, McCusker JH, Namath AF, Gentile C, Hwang SY, Brown PO & Davis RW (1997) Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc Natl Acad Sci USA* **94**: 13057–13062.

Lee TI, Rinaldi NJ, Robert F *et al.* (2002) Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**: 799–804.

Li H, Sethuraman N, Stadheim TA *et al.* (2006) Optimization of humanized IgGs in glycoengineered *Pichia pastoris*. *Nat Biotechnol* **24**: 210–215.

de Lichtenberg U, Jensen LJ, Brunak S & Bork P (2005) Dynamic complex formation during the yeast cell cycle. *Science* **307**: 724–727.

Lok L & Brent R (2005) Automatic generation of cellular reaction networks with Moleculizer 1.0. *Nat Biotechnol* **23**: 131–136.

van Maris AJ, Luttik MA, Winkler AA, van Dijken JP & Pronk JT (2003) Overproduction of threonine aldolase circumvents the biosynthetic role of pyruvate decarboxylase in glucose-limited chemostat cultures of *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **69**: 2094–2099.

Mashego MR, Wu L, Van Dam JC, Ras C, Vinke JL, Van Winden WA, Van Gulik WM & Heijnen JJ (2004) MIRACLE: mass isotopomer ratio analysis of U-^{13}C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. *Biotechnol Bioeng* **85**: 620–628.

Mashego MR, van Gulik WM, Vinke JL, Visser D & Heijnen JJ (2006) *In vivo* kinetics with rapid perturbation experiments in *Saccharomyces cerevisiae* using a second-generation BioScope. *Metab Eng* **8**: 370–383.

Mauch K, Vaseghi S & Reuss M (2000) Quantiative analysis of metabolic and signalling pathways in *Saccharomyces cerevisiae*. *Bioreaction Engineering* (Schügerl K & Bellgardt KH, eds), pp. 435–477. Springer, Berlin.

Müller D (2006) Model-assisted analysis of cyclic AMP signal transduction in *Saccharomyces cerevisiae* – cAMP as dynamic coordinator of energy metabolism and cell cycle progression. PhD Thesis, University Stuttgart, Shaker Verlag, Aachen.

Mustacchi R, Hohmann S & Nielsen J (2006) Yeast systems biology to unravel the network of life. *Yeast* **23**: 227–238.

Mutka SC, Bondi SM, Carney JR, Da Silva NA & Kealey JT (2006) Metabolic pathway engineering for complex polyketide biosynthesis in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **6**: 40–47.

Nielsen J (2001) Metabolic engineering. *Appl Microbiol Biotechnol* **55**: 263–283.

Nissen TL, Schulze U, Nielsen J & Villadsen J (1997) Flux distributions in anaerobic, glucose-limited continuous cultures of *Saccharomyces cerevisiae*. *Microbiology* **143**: 203–218.

Noh K & Wiechert W (2006) Experimental design principles for isotopically instationary ^{13}C labeling experiments. *Biotechnol Bioeng* **94**: 234–251.

Oliver SG, van der Aart QJ, Agostoni-Carbone ML *et al.* (1992) The complete DNA sequence of yeast chromosome III. *Nature* **357**: 38–46.

Ostergaard S, Olsson L, Johnston M & Nielsen J (2000) Increasing galactose consumption by *Saccharomyces cerevisiae* through metabolic engineering of the GAL gene regulatory network. *Nat Biotechnol* **18**: 1283–1286.

Patil KR & Nielsen J (2005a) Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc Natl Acad Sci USA* **102**: 2685–2689.

Patil KR, Rocha I, Forster J & Nielsen J (2005b) Evolutionary programming as a platform for *in silico* metabolic engineering. *BMC Bioinform* **6**: 308.

Pharkya P, Burgard AP & Maranas CD (2004) OptStrain: a computational framework for redesign of microbial production systems. *Genome Res* **14**: 2367–2376.

Porro D, Bianchi MM, Brambilla L *et al.* (1999) Replacement of a metabolic pathway for large-scale production of lactic acid from engineered yeasts. *Appl Environ Microbiol* **65**: 4211–4215.

Price ND, Papin JA, Schilling CH & Palsson BO (2003) Genome-scale microbial *in silico* models: the constraints-based approach. *Trends Biotechnol* **21**: 162–169.

Price ND, Reed JL & Palsson BO (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat Rev Microbiol* **2**: 886–897.

Raamsdonk LM, Teusink B, Broadhurst D *et al.* (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* **19**: 45–50.

Raghevendran V, Gombert AK, Christensen B, Kotter P & Nielsen J (2004) Phenotypic characterization of glucose repression mutants of *Saccharomyces cerevisiae* using experiments with ^{13}C-labelled glucose. *Yeast* **21**: 769–779.

Regenberg B, Grotkjaer T, Winther O, Fausboll A, Akesson M, Bro C, Hansen LK, Brunak S & Nielsen J (2006) Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in *Saccharomyces cerevisiae*. *Genome Biol* **7**: R107.

Ro DK, Paradise EM, Ouellet M *et al.* (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940–943.

Rossell S, van der Weijden CC, Lindenbergh A, van Tuijl A, Francke C, Bakker BM & Westerhoff HV (2006) Unraveling the complexity of flux regulation: a new method demonstrated for nutrient starvation in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* **103**: 2166–2171.

dos Santos MM, Gombert AK, Christensen B, Olsson L & Nielsen J (2003a) Identification of *in vivo* enzyme activities in the cometabolism of glucose and acetate by *Saccharomyces cerevisiae* by using ^{13}C-labeled substrates. *Eukaryot Cell* **2**: 599–608.

dos Santos MM, Thygesen G, Kotter P, Olsson L & Nielsen J (2003b) Aerobic physiology of redox-engineered *Saccharomyces cerevisiae* strains modified in the ammonium assimilation for increased NADPH availability. *FEMS Yeast Res* **4**: 59–68.

Sauer U (2006) Metabolic networks in motion: ^{13}C-based flux analysis. *Mol Syst Biol* **2**: 62.

Schuster S, Fell DA & Dandekar T (2000) A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat Biotechnol* **18**: 326–332.

Shiba Y, Paradise EM, Kirby J, Ro DK & Keasling JD (2006) Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* **9**: 160–168.

Smits HP, Cohen A, Buttler T, Nielsen J & Olsson L (1998) Cleanup and analysis of sugar phosphates in biological extracts by using solid-phase extraction and anion-exchange chromatography with pulsed amperometric detection. *Anal Biochem* **261**: 36–42.

Stelling J, Klamt S, Bettenbrock K, Schuster S & Gilles ED (2002) Metabolic network structure determines key aspects of functionality and regulation. *Nature* **420**: 190–193.

Stephanopoulos G & Vallino JJ (1991) Network rigidity and metabolic engineering in metabolite overproduction. *Science* **252**: 1675–1681.

Stephanopoulos G, Alper H & Moxley J (2004) Exploiting biological complexity for strain improvement through systems biology. *Nat Biotechnol* **22**: 1261–1267.

Tyo KE, Alper HS & Stephanopoulos GN (2007) Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends Biotechnol* **25**: 132–137.

Uetz P, Giot L, Cagney G *et al.* (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**: 623–627.

Vemuri GN, Eiteman MA, McEwen JE, Olsson L & Nielsen J (2007) Increasing NADH oxidation reduces overflow metabolism in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* **104**: 2402–2407.

Villas-Boas SG, Moxley JF, Akesson M, Stephanopoulos G & Nielsen J (2005a) High-throughput metabolic state analysis: the missing link in integrated functional genomics of yeasts. *Biochem J* **388**: 669–677.

Villas-Boas SG, Mas S, Akesson M, Smedsgaard J & Nielsen J (2005b) Mass spectrometry in metabolome analysis. *Mass Spectro Rev* **24**: 613–646.

Villas-Boas SG, Hojer-Pedersen J, Akesson M, Smedsgaard J & Nielsen J (2005c) Global metabolite analysis of yeast: evaluation of sample preparation methods. *Yeast* **22**: 1155–1169.

Wattanachaisaereekul S, Lantz AE, Nielsen ML, Andresson OS & Nielsen J (2007) Optimization of heterologous production of the polyketide 6-MSA in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* **97**: 893–900.

Wiechert W & Noh K (2005) From stationary to instationary metabolic flux analysis. *Adv Biochem Eng Biotechnol* **92**: 145–172.

van Winden WA, van Dam JC, Ras C, Kleijn RJ, Vinke JL, van Gulik WM & Heijnen JJ (2005) Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of (13)C-labeled primary metabolites. *FEMS Yeast Res* **5**: 559–568.

Wu L, Mashego MR, van Dam JC, Proell AM, Vinke JL, Ras C, van Winden WA, van Gulik WM & Heijnen JJ (2005) Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ^{13}C-labeled cell extracts as internal standards. *Anal Biochem* **336**: 164–171.

Yamano S, Ishii T, Nakagawa M, Ikenaga H & Misawa N (1994) Metabolic engineering for production of beta-carotene and lycopene in *Saccharomyces cerevisiae*. *Biosci Biotechnol Biochem* **58**: 1112–1114.

Zhu H, Bilgin M, Bangham R *et al.* (2001) Global analysis of protein activities using proteome chips. *Science* **293**: 2101–2105.