Structural analysis of expressed metabolic subnetworks

Oliver Ebenhöh[1] **Wolfram Liebermeister**[2]
oliver.ebenhoeh@rz.hu-berlin.de lieberme@molgen.mpg.de

[1] Institute for Biology, Humboldt University Berlin, 10115 Berlin, Germany
[2] Kinetic Modelling Group, MPI for Molecular Genetics, 14195 Berlin, Germany

Abstract

Expression of metabolic enzymes increases the metabolic capabilities of the cell, but it also consumes resources, and the gene regulatory systems of cells have to handle this tradeoff. To study whether gene expression patterns reflect the varying metabolic needs of the cell, we translated gene expression profiles into sets of active biochemical reactions, which constitute the expressed metabolic subnetworks. The metabolic capacity of a carbon source denotes the number of metabolites that can be produced from this carbon source and some inorganic nutrients. We studied the metabolic subnetworks that are expressed during diauxic shift in yeast and found that the capacities of different carbon sources tend to decrease during the diauxic shift. The subnetwork expressed in the initial glucose environment shows a high capacity of glucose, much higher than networks of the same size arising in later stages of the diauxic shift or from a random selection of reactions. The same holds for the essential capacity, the number of important, constitutive compounds that can be produced from glucose. These findings indicate that gene regulation increases the range of essential compounds that can be obtained from the available nutrients, while minimising the number of expressed enzymes and therefore the burden of protein synthesis.

Keywords: Metabolic network, network expansion, gene expression, optimality principle

1 Introduction

Cells use enzymes to catalyse and to control a huge number of biochemical reactions, which together form a complex and highly connected cellular reaction network. In the last decades, many enzymes from a wide spectrum of organisms have been classified and attributed to biochemical reactions. While most specific enzyme kinetics are still unknown, exhaustive structural information about metabolic networks has become accessible with the emergence of biochemical databases such as KEGG [1, 2] or Reactome [3].

With the topological information of large-scale networks available, several approaches for the structural analysis of metabolism have been developed, such as flux balance analysis [4], elementary flux modes [5], extreme pathways [6], and graph theoretical analyses [7, 8]. Recently, we have proposed the method of network expansion [9, 10], which is particularly useful to assess biosynthetic capacities of metabolic networks when they are provided with a choice of external nutrients. Like the other structural methods, network expansion does not rely on the kinetic properties of enzymes.

During its lifetime, a cell experiences a wide range of environmental conditions, each of which requires specific responses. Regulatory mechanisms such as transcriptional control respond to changes in the environment and activate specific parts of cellular metabolism, while other parts are deactivated. In this work, we provide a conceptional framework to translate gene expression data retrieved from microarray experiments into networks of active enzymatic reactions. Expressed enzymes, and accordingly active reactions, are identified by thresholding the expression data with a tunable threshold parameter. In this paper, we focus on an experiment which measured the mRNA levels during diauxic shift in the yeast *Saccharomyces cerevisiae* [11]. In this well studied example, yeast obtains its

energy from anaerobic fermentation as long as glucose is available and switches to aerobic oxidation after glucose has been depleted.

We analyse the resulting expressed subnetworks by network expansion to study which metabolites can be produced from single specific carbon sources. The resulting synthesis capacities provide functional characteristics of the expressed subnetworks. By tuning the threshold parameter, we can change the size and the functional characteristics of the resulting networks.

Since the production and degradation of enzymes requires energy and material, it seems plausible that the regulatory mechanisms shut down enzymes that do not contribute to the maintainance of cellular functions. This economy of protein production should result in specific expressed subnetworks for different environmental conditions. To test this hypothesis, we compare the essential capacity of the network – the capacity to produce constitutive compounds – to essential capacities of random networks of comparable size. In line with our hypothesis, we find that the expressed subnetworks show a significantly better performance.

2 Translating expression profiles into expressed metabolic subnetworks

The cellular metabolic network comprises all biochemical reactions that are catalysed by enzymes coded in the genome. The specific amount of enzyme molecules is actively controlled by the cell: enzymes that are not needed can be switched down, e. g. by downregulation of the corresponding genes and initiation of their degradation. If some of the enzymes are completely absent, then the remaining active reactions form a subnetwork of the entire metabolic network. In this section we present a method to construct such subnetworks from expression data measured with microarrays.

In our analysis, the logarithmic expression of a gene (relative to a certain reference value) is accepted as a measure for the catalytic acitivity of the corresponding enzyme. We assume that a chemical reaction is only active if the corresponding expression value exceeds a certain threshold. The method to translate the measured expression values into binary activity profiles is explained in the methods section and schematically depicted in Figure 1.

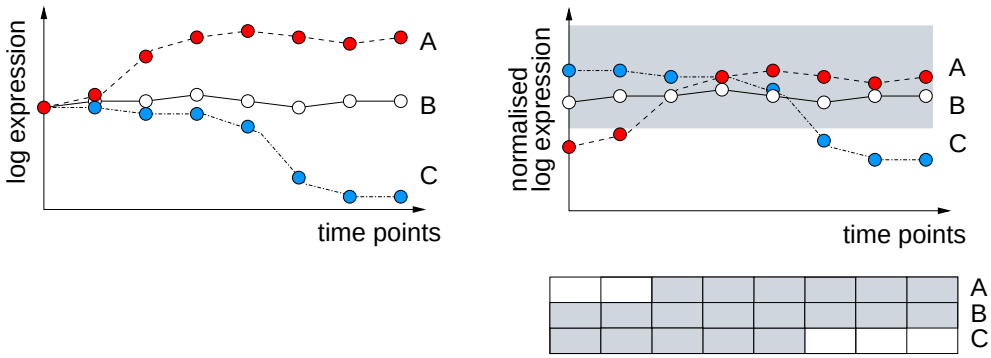

Figure 1: Translating microarray data into binary activity profiles. Top left: Differential expression data for three hypothetical genes A, B, and C. The curves start with a logarithmic value of zero because the sample at the first point in time is used as the reference sample. Top right: Subtracting the gene means shifts the individual curves. Data points above a certain threshold – defined by a $(1 - q)$-quantile of the distribution of all data – are considered expressed (grey), leading to binary activity profiles shown below.

The decision whether a reaction is regarded as active or inactive depends partly on the data preprocessing. The absolute measured values in a microarray experiment (usually channel intensities

for red and green light fluorescence markers) can strongly depend on the individual properties of the microarray chip. Therefore, we shift the mean logarithmic value for each microarray to zero. In addition, the mean logarithmic expression is subtracted for each gene as shown in Figure 1. The resulting value compares the expression of a gene at a certain point to the expression in the other points rather than to a specific (somewhat arbitrarily chosen) reference sample.

We then choose a threshold parameter $0 \leq q \leq 1$ that determines the fraction of genes to be considered expressed. For example, a threshold value of $q = 0.8$ means that the top 80% of all gene expression values will represent expressed genes, while the remaining lowest 20% represent inactive genes. In this way, microarray data are translated into binary expression profiles. We subsequently identify those expressed genes which code for metabolic enzymes and, for reasons of simplicity, we consider all enzymes to be expressed for which at least one expressed gene has been identified. These sets of enzymes define the active reactions which compose the expressed subnetwork for each microarray experiment considered.

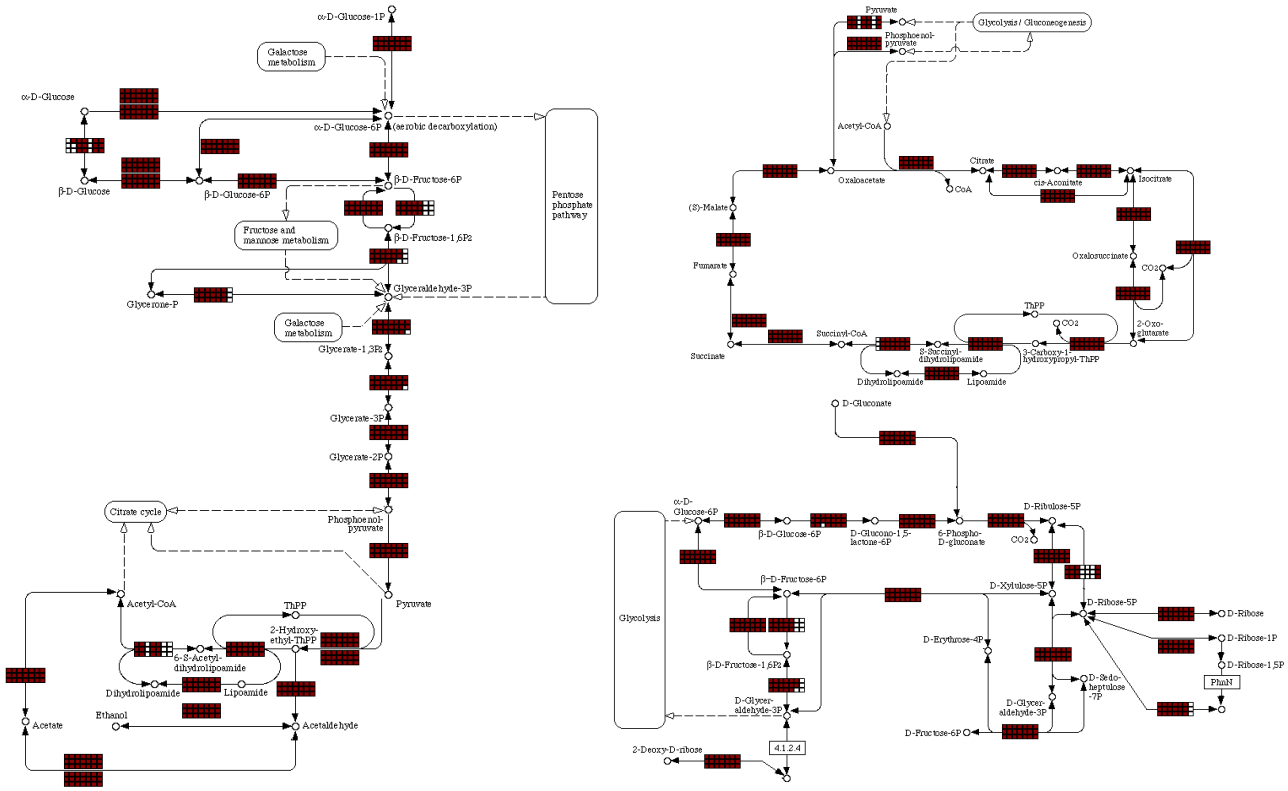

Figure 2: Discrete activity profiles during the diauxic shift in yeast. The panels show the three central metabolic pathways glycolysis, citric acid cycle, and the pentose phosphate pathway. Each reaction is represented by a box showing whether it is active in different phases of the experiment [11] (from left, 9 hours, to right, 21 hours) and for different threshold values (top to bottom: $q = 0.85, 0.825, 0.8$).

Figure 2 shows the results for gene expression in yeast, measured during diauxic shift [11]. In the experiments, yeast cells in batch culture run out of the nutrient glucose and switch their metabolism from fermentation of sugar to respiration based on ethanol. Samples have been taken at seven time points between 9 and 21 hours after application of glucose. In the figure, little boxes at the reactions indicate the binary expression profile during time (from left to right).

While most reactions remain active during the whole experiment, the *fructose bisphosphatase* reaction (converting fructose 6-phosphate into fructose 1,6-bisphosphate), an important step in glycolysis, is shut down in the last two points in time. Increasing the threshold value q (rows within the little boxes) also increases the size of the expressed subnetwork, that is, the number of active reactions.

3 Structural analysis of the expressed metabolic networks

In the previous section, we have shown how gene expression profiles can be translated into expressed metabolic subnetworks. For the diauxic shift data from [11], we have calculated the expressed subnetworks for each time point and for several threshold values. Now we shall investigate structural and functional properties of these subnetworks using the method of network expansion [9, 10]. This method allows a characterisation of metabolic networks in terms of their synthesising capacities. In particular, we determine which biochemical compounds can be produced when inorganic material is abundant while exclusively one carbon containing compound is available. The number of the compounds that can be produced from a particular carbon source will be termed the synthesising capacity, or simply the capacity, of the expressed subnetwork for this particular carbon source. The capacities are calculated using network expansion as described in the methods section.

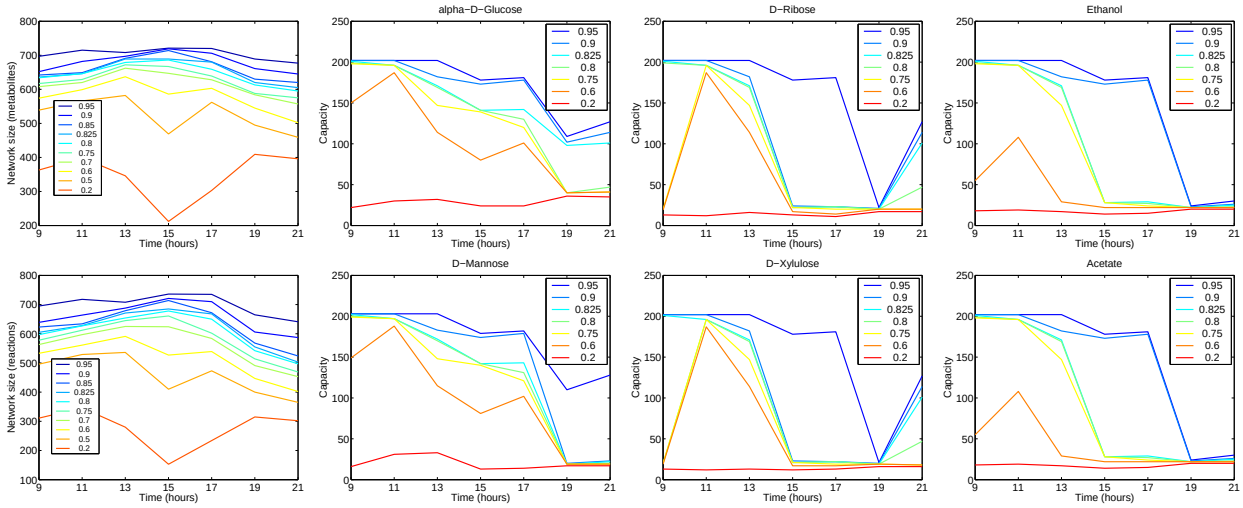

Figure 3: Expressed subnetworks during diauxic shift in yeast. The two left panels show the network size (number of metabolites, top, and number of reactions, bottom) for several threshold values (different lines, the values decrease from top to bottom). The remaining panels show synthesising capacities for selected carbon sources.

As shown in Figure 3, the synthesising capacities for selected carbon sources change during the diauxic shift. The panels in the figure, representing the results for chemically related saubstances, have been arranged in pairs: The panels in the second column show the result for two hexoses, α-D-Glucose and D-Mannose, those in the third column for two pentoses, D-Ribose and D-Xylulose, and the right two panels for the two-carbon metabolites ethanol and acetate. Different lines correspond to different values of the threshold parameter.

For most threshold values and for most carbon sources, the synthesising capacity tends to decrease as glucose becomes limited. At the first two time points (9 and 11 hours), all expressed subnetworks computed with high threshold values ($q \geq 0.8$) display a large synthesising capacity for all examined carbon sources. In fact, the expressed subnetworks can synthesise over 200 metabolites provided that exclusively one of the selected carbon sources and inorganic material is available. Apparently, the metabolic network of yeast expressed under glucose-rich conditions ensures a high capacity for a large number of carbon sources.

For α-D-Glucose, we observe a dramatic drop in synthesising capacity between 17 and 19 hours after glucose application for all reasonably high threshold values ($q \geq 0.6$). Interestingly, this is the point in time when cells run out of glucose [11]. If glucose is no longer available, the ability to synthesise many molecules from glucose is apparently no longer an important cellular function. As

a consequence of glucose deprivation, cells react by shutting off some parts of their metabolism and switching on others. This change in metabolic activity is reflected in the change of the synthesising capacities.

The synthesising capacities for chemically related substances display a similar behaviour during diauxic shift. For example, the capacity for D-Mannose exhibits a similar behaviour as for α-D-Glucose. However, the determined expressed subnetworks for threshold values $q \geq 0.825$ at the late times (19 and 21 hours) still display a rather large glucose capacity (over 100 compounds), while a high mannose capacity requires an extremely high threshold value of $q = 0.95$. This suggests that for the critical threshold values of $q = 0.9$ and $q = 0.825$, the expressed subnetworks still allow for a flexible chemical conversion of glucose but miss exactly those reactions which connect mannose to the remaining glucose metabolism. A closer inspection revealed that this is indeed the case and that the critical reaction is mediated by the enzyme *mannose-6-phosphate ketol-isomerase*, converting mannose-6-phosphate into fructose-6-phosphate. In fact, if the network determined with threshold $q = 0.825$, containing 511 reactions, is structurally modified by adding this single reaction, the synthesising capacity for mannose increases from 22 to over 100. It is biologically reasonable for yeast to maintain certain conversion capabilities for glucose even if external glucose is no longer available, since glucose can internally be produced and act as an intermediate metabolite. However, it is not economic for the cell to invest energy in the production of enzymes that convert mannose into glucose if no mannose is externally available.

The two pentoses show almost identical synthesising capacities. Interestingly, for intermediate threshold values ($0.1 \leq q \leq 0.9$), the capacity drops significantly between 13 and 15 hours (a good time before complete glucose exhaustion) and partly recovers between 19 and 21 hours (after complete exhaustion). This indicates that there exists some connection between the regulation of pentose metabolism and the external availability of hexoses. Obviously, investigation of the three central metabolic pathways glycolysis, citric acid cycle and pentose phosphate pathway (Figure 2) is not sufficient to explain this behaviour in detail. However, their examination hints at an important role of *ribose-5-phosphate isomerase*, transforming D-ribose-5-phosphate into D-ribulose-5-phosphate: the presence / absence of this reaction strongly correlates with high / low synthesising capacity for pentoses – this reaction is inactive for the three points 15, 17, and 19 hours for a wide range of threshold values.

Surprising, and far more difficult to explain, is the result for ethanol (and acetate, displaying almost exactly the same behaviour). It is known that during diauxic shift, yeast changes from anaerobic to aerobic metabolism. After glucose deprivation, ethanol is used as both energy and carbon source. Therefore, one would intuitively expect that the synthesising capacity for ethanol increases during diauxic shift. However, this is apparently not the case. To find explanations for this behaviour on the network level, we have tried to identify critical reactions that are responsible for the decreased synthesising capacity (see Methods). When comparing the expressed subnetworks at 17 hours for the threshold values $q = 0.9$ (network size of 677 reactions, capacity for ethanol of 178 compounds) and $q = 0.825$ (641 reactions, 29 compounds), we could identify a reaction that is lost in the smaller network, which seems to inflict a dramatically reduced synthesising capacity. However, this reaction, catalysed by the enzyme *aminoadipate semialdehyde oxidoreductase*, is involved in lysine degradation and no direct connection to the metabolism of ethanol is apparent.

While the capacity found for glucose seems to reflect the external availability of glucose, the capacity for ethanol is apparently not a good indicator for the absence of glucose. Further investigations will be necessary to understand the relation between the synthesising capacity of ethanol and the shift to aerobic oxidation during glucose depletion [11].

4 Performance of the expressed subnetworks

Enzymes are necessary to catalyse the chemical reactions that transform external nutrients into biomass. However, the production, maintainance, and degradation of enzymes requires energy and

material, and the presence of enzymes increases the osmotic pressure. All this puts a burden on the cell, so the amount of enzyme molecules should be kept small.

Different optimality-based approaches [12, 13, 14, 15, 16] have been suggested to model the compromise between metabolic performance and the cost of enzyme production. Under certain circumstances, it seems to be optimal for the cell to switch off parts of the metabolic network. An analysis of the Lac operon [16] has shown that the operon should be shut off completely if lactose is only available in small amounts.

Here we characterise networks by their capability to produce metabolites from a selected carbon source. In this framework, enzyme expression enables the cell to produce biomass precursors and other essential metabolites (which we together term *constitutive metabolites*) from the external resources. This benefit is counteracted by the cost of enzyme production, quantified by the number of active reactions. Thus ideally, gene expression should lead to expressed subnetworks that maximise the synthesising capacity for the available carbon sources, while minimising the network size.

Our distinction between active and inactive reactions depends on an arbitrary threshold q, and it may be that the computed capacities just reflect our specific way of data preprocessing, a specific choice of the threshold value q, or simply the network size. To rule out these possibilities, we show that for a range of threshold values the expressed networks found in reality yield higher benefits than reference networks of the same size and thus the same cost of enzyme production. As reference networks, we consider random subnetworks constructed from the complete yeast network by randomly chosing a given number of reactions. The underlying null hypothesis is that the gene expression values do not at all reflect the metabolic function of the genes.

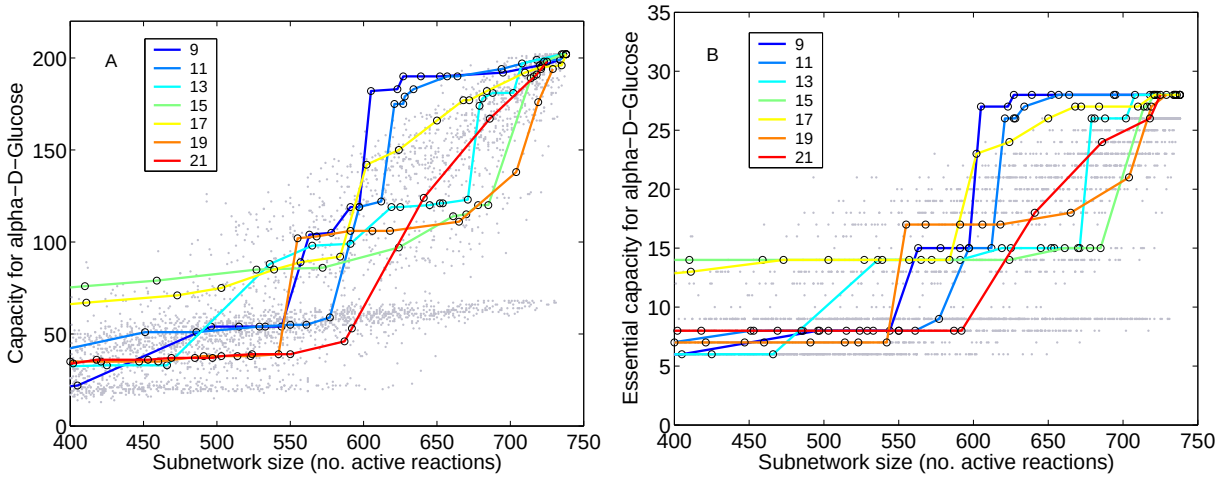

Figure 4: Performance of the expressed subnetworks during diauxic shift. Left: Synthesising capacity for α-D-Glucose is plotted versus subnetwork size. Each light grey point refers to a random subnetwork. The lines correspond to subnetworks expressed during diauxic shift [11], each line represents an experimental point (evaluated with different threshold values q). Right: Essential capacities for α-D-Glucose denoting the number of constitutive compounds obtainable from glucose. In both cases, the expressed network under glucose conditions performs better than expressed networks at later points and random networks of the same size.

To test our hypothesis of increased capacity for the nutrient α-D-Glucose, we plotted in Figure 4a the synthesising capacity against the network size for different choices of the threshold parameter q. Grey points correspond to random networks, while each line depicts an expressed subnetwork for different choices of the threshold q. As expected, larger random networks (x-axis) tend to yield larger capacities for glucose (y-axis). The same holds for the subnetworks expressed during the diauxic shift: at a low threshold, both the network size and the capacity for glucose are small. When the network

becomes larger, also the capacity for glucose increases; however, when compared to random networks, high capacities are already reached for networks of moderate size. We found that the network expressed under glucose conditions, thresholded to a size of 600 reactions, has a higher capacity than any of 10000 random networks of the same size. At a size of 650 reactions, two of the random networks have higher capacity (p-value $\approx 2 \cdot 10^{-4}$), while at 700 reactions, the p-value raises to 0.084.

Altogether, the subnetwork expressed at the beginning of the experiment performs better on glucose than the subnetworks at intermediate or late times where no glucose is available, and significantly better than the randomly chosen subnetworks.

Since not all metabolites play an important role in metabolism, we have defined an ad-hoc list of 76 constitutive metabolites which are essential for growth (see Table 1). We repeated the above analysis determining the essential capacity, under which we understand the number of constitutive metabolites that can be produced from glucose as the sole carbon source. The results are plotted in Fig. 4b. As for Fig. 4a, it can be observed that expressed subnetworks perform significantly better than random subnetworks. Interestingly, even for later times and low threshold values, the essential capacity remains very high as compared to randomly selected subnetworks.

Table 1: Constitutive metabolites. This ad-hoc list contains metabolites which are considered to be essential for cell growth. It comprises most amino acids and important cofactors. The list has been compiled by determining those metabolites which appear in at least 90% of the metabolic networks of all organisms contained in the KEGG database.

H_2O	ATP	NAD^+	NADH
NADPH	$NADP^+$	ADP	Orthophosphate
CoA	CO_2	Pyrophosphate	NH3
UDP	S-Adenosyl-L-methionine	AMP	S-Adenosyl-L-homocysteine
Pyruvate	Acetyl-CoA	L-Glutamate	GDP
Glycine	L-Alanine	GTP	L-Lysine
L-Aspartate	CMP	L-Arginine	CTP
L-Glutamine	L-Serine	L-Methionine	Phosphoenolpyruvate
UTP	L-Tryptophan	L-Phenylalanine	H^+
ITP	L-Tyrosine	L-Cysteine	Tetrahydrofolate
IDP	UMP	Glyceronephosphate	CDP
D-Ribose5-phosphate	L-Leucine	dATP	L-Histidine
5,10-Methylenetetrahydrofolate	GMP	L-Proline	L-Valine
L-Threonine	dADP	dGTP	dAMP
dGDP	dTDP	dTMP	dUMP
L-Isoleucine	dCTP	dTTP	dUTP
dCDP	dUDP	Thiamintriphosphate	Zymosterol
Heme	dGMP	L-1-Pyrroline-3-hydroxy-5-carboxylate	FAD
3',5'-CyclicAMP	L-Asparagine	dCMP	Glutathione

5 Discussion

Differential expression leads to a temporal up- and downregulation of enzymes. Complete downregulation of enzymes switches off the corresponding biochemical reactions, reducing the metabolic network to a current expressed subnetwork. We presented a heuristic method to construct expressed metabolic subnetworks from expression data.

Our method of thresholding the expression data is based on the following argument: we assume that a low expression value of an enzyme leads to a small activity. For saturable kinetics, this implies a small maximal velocity and hence a small reaction rate. Even if no reaction is completely switched off, we assume that the downregulated reactions are so slow that their contribution to metabolism can be neglected, and we compute the synthesising capacities arising from the remaining, sufficiently fast reactions. We expect that using enzyme concentration instead of mRNA concentrations will lead to more reliable results.

The size of the determined subnetworks depends critically on an arbitrary selected threshold parameter q. However, Figure 2 shows that the results for the three central metabolic pathways (glycolysis,

the citric acid cycle and the pentose phosphate pathway) are stable within a certain parameter range ($0.8 \leq q \leq 0.85$), and also the binary activity profiles for most important enzymes remain unchanged.

We characterised the calculated subnetworks at each point in time and for different threshold parameters by the synthesising capacities for various carbon sources. By comparing the composition of the calculated subnetworks with their capacities, we could identify reactions whose absence or presence critically influences the capacities for several carbon sources. For some cases, such as the mannose capacity, a critical reaction within the metabolic subnetwork could be identified. However, for other examples, like the ethanol capacity, it is not obvious why the presence or absence of the critical reactions leads to such extreme changes in capacity. Considering the high connectedness and complicated wiring of a metabolic network, it should not be surprising that many phenomena cannot be related to single reactions.

While the synthesising capacities for hexoses show plausible time courses – the capacity for glucose should decrease once glucose is no longer available – the time course for ethanol is not directly and intuitively related to the textbook picture of the diauxic shift. Intuituvely, we would expect the ethanol capacity to increase during diauxic shift but we find that this is not the case. A possible explanation for this discrepancy is as follows: It is conceivable that production of new compounds from ethanol requires additional cofactors which we have not considered in the calculations based on network expansion. In such a case the determined capacity of ethanol will be low even if there is no net consumption of the cofactor in reality.

One important function of metabolism is to provide precursors for biomass production and other metabolites which are indispensable for a wide class of reactions. Continuously growing cells need to produce them: thus we scored metabolic networks by the number of (1) all compounds and (2) all constitutive compounds they can produce from glucose. We compared the performance of expressed subnetworks to a large number of random networks and observed that for a range of network sizes, the expressed networks lead to a significantly high yield even for small subnetwork size. This result supports the hypothesis that the expression of enzymes follows an optimality principle: the cell seems to ensure that a maximal number of essential metabolites can be produced by a minimal set of active reactions. Studies to corroborate and extend this finding may comprise a more detailed modelling that accounts for the exact composition of the growth medium and compares the results for different experimental conditions.

6 Methods

6.1 Construction of expressed subnetworks from gene expression data

Expression data were downloaded from [11]. The expressed subnetworks for the different time points were constructed as follows: (1) Microarray data were transformed to log10 ratios (intensity in sample divided by intensity in the reference sample). (2) Multiple values for the same open reading frame (ORF) were averaged (arithmetic mean of log values). Missing values were replaced by values guessed from a linear regression model that is also used in analysis of variance: column, row, and total means were computed from the non-missing values. For each missing value, we imputed the sum of the corresponding row and column means minus the total mean. (3) The data were normalised by subtraction of array mean values and subsequent subtraction of the gene mean values. (4) Expressed ORFs were determined by thresholding the expression values. We consider the distribution of all expression values in a data set. Given the quantile threshold q, an ORF is considered expressed if the corresponding expression value is above the $(1-q)$-quantile. (5) We downloaded the metabolic network of yeast from the KEGG database [1, 2]. The data set contains a mapping between chemical reactions and open reading frames. An enzyme is considered expressed if at least one of the corresponding ORFs (as indicated by KEGG) is expressed. We also tested other variants of data preprocessing: in step (3), we considered (i) unnormalised data and (ii) data that were only normalised for the array

mean. In step (4), we also tried thresholding per array. These different preprocessed data lead to similar general conclusions (not shown).

6.2 Structural analysis using network expansion

In the method of network expansion (described in detail in [10]), a series of metabolic networks is constructed from a given number of initial substrates (the seed). In each step, the network is expanded by those reactions which utilise only the seed and those metabolites which are products of reactions incorporated in previous steps. By construction, the metabolites within the final network characterise the synthesising capacity of a metabolic network when only the seed substrates are available. For our calculations, we assume that important cofactors, namely ATP/ADP/AMP, NAD(P)H/NAD(P)$^+$ and Coenzyme-A, whose presence is required for many reactions, do not have to be synthesised during the expansion process. We rather assume that they act only in their function as cofactors, i. e. transferring phosphate groups, accepting electrons or transferring acyl groups. The reason for this approach is that these cofactors exist in every functional cell and, when functioning as cofactors, they are neither produced or consumed and consequently their total concentrations remain constant. The cofactor functionality is therefore provided even if they cannot be produced from the seed compounds. The role of cofactors has been extensively discussed in [10].

The synthesising capacities in this work have been calculated as follows: We first identified all non-carbon containing compounds appearing in the yeast metabolism. Then, for each carbon-containing metabolite in the yeast metabolism, we added this metabolite to the previously determined set of non-carbon compounds and used this set as a seed for the network expansion. These calculations were carried out for all carbon-containing compounds and for all determined expressed subnetworks.

We identified critical reactions by the following approach: We investigate the case for two expressed subnetworks corresponding to different threshold values q, for which the larger displays a high and the smaller a low synthesising capacity for some carbon source. We first identify those reactions which appear only in the larger network. For each such reaction, we construct a new subnetwork by removing it from the larger subnetwork and calculate its capacity. A reaction is considered critical if its removal results in a considerably reduced synthesising capacity. Similarly, we construct for each of these reactions a subnetwork by adding the reaction to the smaller of the two subnetworks and again compute the capacity. In this case, a reaction is considered critical if its addition leads to a considerably increased synthesising capacity. For the presented examples, both methods resulted in the same critical reactions.

References

[1] Kanehisa M. A database for post-genome analysis. Trends Genet. 1997;13:375–376.

[2] Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, et al. From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res. 2006;34:D354–357.

[3] Joshi-Tope G, et al. Reactome: a knowledgebase of biological pathways. Nucleic Acids Res. 2005;33 Database Issue:D428–32.

[4] Kauffman KJ, Prakash P, Edwards JS. Advances in flux balance analysis. Current Opinion in Biotechnology. 2003;14:491–496.

[5] Schuster S, Fell DA, Dandekar T. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nature Biotechnology. 2000;18:326–332.

[6] Papin JA, Price ND, Wiback SJ, Fell DA, Palsson BO. Metabolic pathways in the post-genome era. TIBS. 2003;28(5):250–258.

[7] Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási AL. The large-scale organization of metabolic networks. Nature. 2000;407:651–654.

[8] Wagner A, Fell DA. The small world inside large metabolic networks. Proc R Soc Lond B. 2001;268:1803–1810.

[9] Ebenhöh O, Handorf T, Heinrich R. Structural Analysis of Expanding Metabolic Networks. Genome Informatics. 2004;15(1):35–45.

[10] Handorf T, Ebenhöh O, Heinrich R. Expanding Metabolic Networks: Scopes of Compounds, Robustness and Evolution. J Mol Evol. 2005;61:498–512.

[11] DeRisi JL, Iyer VR, Brown PO. Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale. Science. 1997;278:680–686.

[12] Reich JG. Zur Ökonomie im Proteinhaushalt der lebenden Zelle. Biomed Biochim Acta. 1983;42(7/8):839–848.

[13] Klipp E, Heinrich R. Competition for enzymes in metabolic pathways: implications for optimal distributions of enzyme concentrations and for the distribution of flux control. BioSystems. 1999;54:1–14.

[14] Liebermeister W, Klipp E, Schuster S, Heinrich R. A theory of optimal differential gene expression. BioSystems. 2004;76:261–278.

[15] Holzhütter H. The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. Eur J Biochem. 2004;271(14):2905–2922.

[16] Dekel E, Alon U. Optimality and evolutionary tuning of the expression level of a protein. Nature. 2005;436:588–992.