Structural Analysis of Expanding Metabolic Networks

Oliver Ebenhöh oliver.ebenhoeh@rz.hu-berlin.de
Thomas Handorf handorf@physik.hu-berlin.de
Reinhart Heinrich reinhart.heinrich@rz.hu-berlin.de
Theoretical Biophysics, Institute of Biology, Humboldt University Berlin, Germany

Abstract

Methods are developed for structural analysis of metabolic networks expanding in size. Expansion proceeds in consecutive generations in which new reactions are attached to the network produced in the previous stage. Different rules are applied resulting in various modes of expansion. Expansion is performed on the set of glycolytic reactions as well as on a very large set of reactions taken from the KEGG database. It is shown that reactions and compounds strongly differ in the generation in which they are attached to the network allowing conclusions for the temporal order of the acquisition during network evolution. The expansion provides efficient tools for detecting new structural characteristics such as substrate-product relationships over long distances.

Keywords: metabolic network, evolution, stoichiometry, optimisation, scope, KEGG database

1 Introduction

Metabolic networks in living cells are characterised by several quantities. Variables such as the concentrations of metabolites or fluxes may change on a short time scale whereas system parameters such as stoichiometric numbers defining the structure of the networks or kinetic constants of enzymatic reactions can be considered to remain constant throughout the life span of a single organism.

Traditionally, mathematical modelling of metabolic networks aims at simulating stationary states or time courses for the system variables with the system parameters used as input quantities. Recently, structural analyses gained much interest since in contrast to enzyme kinetic parameters the stoichiometric properties are often much better known [5]. The development of metabolic databases such as KEGG [3, 4] allows for studying the structure of metabolism on a large scale.

In previous work we performed structural analyses on subnetworks by applying evolutionary optimisation using a limited set of generic reactions. We could show that the stoichiometry of the central energy metabolism, in particular the ordering of ATP and NADH producing and consuming reactions, can be explained under the premise that contemporary networks ensure a high ATP production rate [1, 6]. Furthermore we introduced the concept of stoichiometric robustness characterising the effect of structural changes on the function of metabolic networks [2].

The goal of the present work is to reconstruct extant metabolic systems on a much larger scale by taking into account as many specific reactions as possible. We perform this reconstruction by simulating evolutionary processes starting from a small set of reactions and compounds. During this procedure new reactions are repeatedly attached according to certain rules leading to networks of increasing size. Each individual process we call network expansion. We show that the mutual dependencies of reactions by their substrates and products allow for conclusions on the temporal order of the emergence of enzymatic reactions and subpathways. Such studies will help to elucidate the principles under which network structures evolved towards increasing complexity. In this work we develop methods for this kind of investigation which are tested first on the limited set of reactions of the glycolytic pathway. Application to the large set leads to the definition of new structural characteristics such as the scope of compounds.

2 Basic Assumptions

Any metabolic network depends on substrates which have to be provided by the environment. Through a series of biochemical reactions these substrates get converted into end products which are released into the environment. Such substrates and products are called external compounds of the given network. Under steady-state conditions, the production and consumption of all other (internal) metabolites is balanced resulting in time independent concentrations. Let $\mathbf{C} = (\mathbf{X}, \mathbf{S})^T$ denote the vector of the concentrations of the metabolites such that $\mathbf{X}$ comprises all internal and $\mathbf{S}$ all external compounds.

The dynamic behaviour of the network is governed by a set of differential equations for the metabolite concentrations having the form

$$\frac{d\mathbf{C}}{dt} = \mathbf{N} \cdot \mathbf{V}, \tag{1}$$

where $\mathbf{N}$ is the stoichiometric matrix which can be decomposed into two submatrices $\mathbf{N_X}$ and $\mathbf{N_S}$ for the internal and external compounds, respectively, and $\mathbf{V}$ is the vector of the rates of the biochemical reactions. In principle, all reactions are reversible allowing for positive and negative values of $\mathbf{V}$. Under physiological conditions, however, many reactions proceed only in one direction leading to sign restrictions for the corresponding components of $\mathbf{V}$. In the following, we consider all reactions to be reversible.

The steady-state of the system is determined by the equation

$$\mathbf{N_X} \cdot \mathbf{V} = 0. \tag{2}$$

In the framework of stoichiometric analysis, Eq. (2) is treated as a linear equation system for feasible fluxes $\mathbf{V}$. In general, the solution of the system is not unique with the general solution being a superposition of a set of fundamental solutions. We distinguish between two different types of solutions, one where the fluxes result in a net conversion of external metabolites, $\mathbf{N_S} \cdot \mathbf{V} \neq 0$, and one where the fluxes represent only internal cycles, $\mathbf{N_S} \cdot \mathbf{V} = 0$.

In the present work we investigate expanding network structures leading to sequences of stoichiometric matrices of increasing size. This analysis is performed in view of evolution where metabolic networks have acquired new metabolic capabilities in form of additional enzymatic reactions. On shorter time scales, changes in network size could be mediated by a gene transfer as can be observed between bacteria.

For deriving rules for network expansion, we assume that a newly acquired biochemical reaction may be especially useful to an organism if it makes it less dependent from external resources or provides it with new resources. This can be fulfilled by reactions allowing to treat one or more compounds to be internal which were external before the acquisition. If the acquired reaction depends on metabolites not present in the network before, these compounds must be considered external.

However, for some chemicals which were always highly abundant in the environment there was no need for an organism to balance their production and consumption. In our analysis we consider such compounds to remain external during the whole process of network expansion. We investigate several scenarios with different sets of these highly abundant compounds.

The computational analysis of network expansion is performed by an algorithm which can be summarised as follows:

1. The process starts with a set of one or more biochemical compounds acting as a seed of the expanding network.

2. From the set of all available reactions a set of candidate reactions is determined. Candidate reactions are defined by the property that they can in principle be attached to the current network. Candidate reactions must fulfil one of the two expansion conditions: At least one of its compounds must already be present in the current network (weak expansion condition – C1);

all of its substrates or all of its products must already be present (strong expansion condition – C2).

3. From the set of candidate reactions, one or more are attached to the current network according to specific selection rules. In this work, the selection rules utilise Eqs. 2 and 1, for example for deciding which reactions perform net conversions under steady state conditions (see Section 3.1).

Steps 2 and 3 are repeated continuously until no candidate reaction can be found. Each loop leads to a new network generation.

3 Expansion on the Set of Glycolytic Reactions

In this section we investigate network expansion on a limited set of reactions. Candidate reactions are chosen from the thirteen reactions of the glycolytic pathway (see Figure 1). Since this pathway supplies energy in form of ATP, consumption of ATP by other processes (ATPases) is also taken into account (the reaction placed on top of the first glycolytic reaction at the left end of Figure 1).

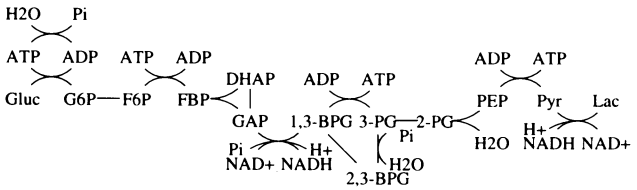

Figure 1: Base set of reactions comprising the glycolytic pathway. Abbreviations for the compounds are: Gluc - glucose, G6P - glucose-6-phosphate, F6P - fructose-6-phosphate, FBP - fructose-1,6-bisphosphate, DHAP - dihydroxyacetone phosphate, GAP - glyceraldehyde-3-phosphate, 1,3-BPG - 1,3-bisphosphoglycerate, 2,3-BPG - 2,3-bisphosphoglycerate, 3-PG - 3-phosphoglycerate, 2-PG - 2-phosphoglycerate, PEP - phosphoenolpyruvate, Pyr - pyruvate, Lac - lactate.

3.1 Selection Rule

We analyse network expansion on this set of reactions under the following conditions. First, we assume that the network in each stage depends on as few as possible external metabolites. Second, we assume that the network produces as many as possible internal metabolites since these compounds may serve as building blocks for other metabolic processes. Third, we consider it disadvantageous for a network to possess reactions that are not required for conversions under steady-state conditions. These reactions are characterised by the fact that the corresponding entries in all solution vectors $\mathbf{V}$ of Eq. (2) vanish.

The selection rule can be formalised as follows: From the candidate reactions, those reactions are chosen maximising a selection value for which we use the expression

$$f = \alpha x - \beta s - \gamma z, \tag{3}$$

where x and s are the numbers of internal and external metabolites, respectively, and z is the number of reactions that are not required in steady-state and α, β, and γ are weight factors. The presented simulations have been performed for $\alpha = \beta$ and $\gamma < \beta$. The first choice ensures that the value of f remains unchanged if the acquisition of a new reaction involves the incorporation of a new external compound into the network and simultaneously an external compound becomes internal. The second choice entails that an additional reaction which is not required for conversions is less disadvantageous in terms of the selection value f than an additional external compound.

Additional criteria to be fulfilled during the process are: a) an internal compound cannot become external in later generations; b) whenever a reaction is attached to one or more external compounds, at least one of these compounds must become internal; c) a maximal set of external compounds becomes internal under the condition that there exists a steady-state solution with a net conversion between the remaining external metabolites ($\mathbf{N_S} \cdot \mathbf{V} \neq 0$).

In the set of glycolytic reactions there are certain compounds whose production and consumption are strictly coupled. These are ATP and ADP as well as NAD^+, NADH and H^+. Therefore, the stoichiometric matrix is reduced by considering only one compound from each group. We choose to omit the rows corresponding to the compounds ADP, NADH and H^+.

3.2 Model Scenarios

Using the algorithm described in section 2, we analyse three scenarios. All of them are based on the weak expansion condition C1, using glucose as the only initial compound and considering water as highly abundant. The three scenarios differ in the choice of compounds which, in addition to water, are forced to be external and cannot become internal during the expansion. These compounds are selected from the metabolites ATP, NAD^+ and inorganic phosphate.

Case 1: ATP, NAD^+ and Inorganic Phosphate as Forced External Compounds

In the first step of the algorithm, there is only one candidate reaction (hexokinase – HK) which can be attached to glucose. In this stage the networks contains three new compounds, all of which are external. Two of them (ATP and ADP) are forced to be external whereas G6P can become internal later (see generation 1 in Figure 2). Like in the first step, for each of the following three steps there

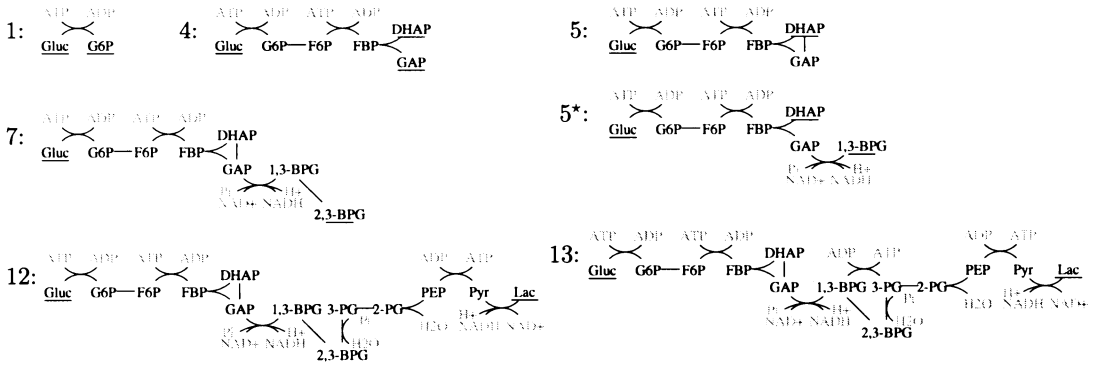

Figure 2: Expansion process on the limited set of the glycolytic reactions (case 1). It is assumed that water, ATP, NAD^+, and inorganic phosphate are highly abundant in the environment and forced to be external. These compounds are marked in grey. ADP, NADH, and H^+ are strictly coupled to other compounds and also do not contribute to the selection value f and are therefore marked in grey as well. Compounds which are external in one generation but may become internal in a later generation are underlined. The calculations were performed with the parameters $\alpha = \beta = 1$ and $\gamma = \beta/2$.

exists only one candidate reaction. This leads to the generation 4 which includes the reactions HK, glucosephosphate isomerase (PGI), phosphofructokinase (PFK) and aldolase (Ald). In the fifth step there are two candidate reactions, triose phosphate isomerase (TIM) and glyceraldehyde phosphate dehydrogenase (GAPDH). Incorporation of TIM leads to generation 5 whereas incorporation of GAPDH leads to a network depicted by 5* in Figure 2. These two possibilities differ in their relations between

the numbers of external and internal compounds. Whereas both networks contain four internal compounds, networks 5 and 5* include two and three external compounds, respectively. Therefore, the expansion proceeds with network 5.

In step 6 there is only one candidate reaction, leading to the attachment of GAPDH. In step 7 there are two candidate reactions, namely bisphosphoglycerate mutase (BPGM) and phosphoglycerate kinase (PGK). Both resulting networks have the properties $s = 2$, $x = 6$, and $z = 0$ resulting in the same selection value $f = 4$. The algorithm randomly chooses the first possibility.

In step 8, from the two candidate reactions PGK and BPGase, the latter is selected since in this case the network expands increasing the number of internal as well as external compounds by one, whereas in the other case only one external compound is added.

In the subsequent steps 9–12, the metabolic chain is expanded until the incorporation of lactate dehydrogenase (LDH).

In this stage only two reactions are not yet incorporated (PGK and ATPase). Since in this scenario ATPase has only forced external compounds, its attachment cannot result in compounds becoming internal and therefore this reaction is not one of the candidate reactions. Consequently, the last reaction added is PGK leading to generation 13 representing the full glycolytic pathway.

Case 2: Inorganic Phosphate as Forced External Compound

As in case 1, the first step represents the incorporation of HK. Since in this scenario ATP can become an internal compound, the expansion continues in a different fashion compared to case 1. In fact, there are five candidate reactions in step 2, four of them proceed with ATP turnover and are therefore suited to make ATP an internal substrate when combined HK. The fifth candidate reaction is PGI. The algorithm selects ATPase since no new external compounds are incorporated (note that in this scenario inorganic phosphate does not contribute to the selection value f). From this step the expansion proceeds in the same way as in case 1.

Case 3: No Additional Forced External Compound

In this scenario in principle all compounds except water can become internal during the expansion. The first six steps proceed in the same way as in case 2 resulting in the network 6 in Figure 3. This networks converts glucose into two molecules DHAP and consumes two molecules of ATP which are produced by ATPase working in the opposite direction by the expense of inorganic phosphate which in this stage is external.

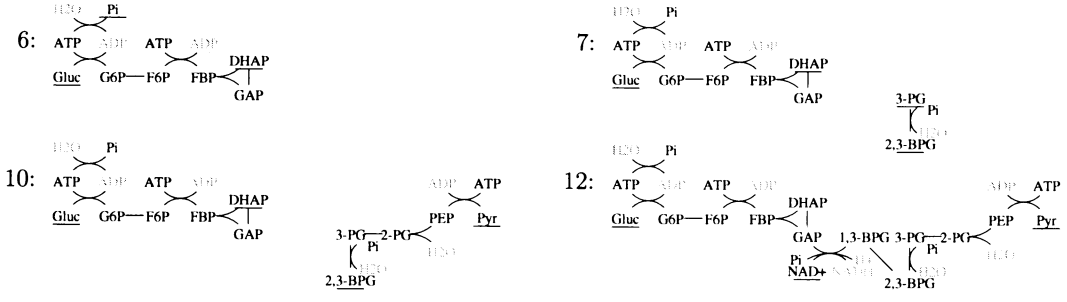

Figure 3: Expansion on the glycolytic pathway (case 3). Conditions, labels, and parameter values are the same as for Figure 2 with the exception that only water is considered to be highly abundant.

From this step on, the expansion differs dramatically from both previously considered cases. In step 7 the sugar converting chain is not extended by further metabolising GAP but rather the reaction BPGase is attached allowing inorganic phosphate to become an internal compound.

The expansion proceeds by prolongation of the chain from 3PG to pyruvate. This leads in step 10 to a pathway consuming besides glucose also 2,3-BPG with a net consumption of one ATP. ATPase still works in the opposite direction where the rate is halved compared to network 6.

In steps 11 and 12 the two sugar converting chains are connected by incorporation of BPGM and GAPDH. This network transforms glucose into pyruvate with no net consumption of ATP. Therefore, ATPase occurs with a zero rate resulting in a value $z = 1$.

In step 13 LDH is attached allowing for the internalisation of NAD^+ and pyruvate which were external in step 12.

Only with the incorporation of PGK in the last step, the glycolytic pathway achieves the ability to produce ATP. Consequently, ATPase now operates in the forward direction.

We have chosen the well known glycolytic pathway to demonstrate in detail the procedure of network expansion. Despite its simplicity the expansion may proceed in many modes differing in the order of attachment of the reactions. Expansion conditions and selection rules restrict the number of possibilities significantly. We could demonstrate that applying a small number of such rules is sufficient to ensure that intermediary generations do not represent arbitrary collections of reactions but rather networks with preliminary capabilities. Interestingly, the same criteria did not lead to a premature end but the expansion evolved until all reactions of the glycolytic pathway were incorporated.

4 Network Expansion on a Large Scale

We will now apply the expansion methods on a set of 5161 reactions and 4450 compounds as provided by the KEGG database. Whereas in section 3 we examined specific expansion modes, we now characterise the processes by general properties.

4.1 Strong Expansion Condition without Selection Rules

First, we investigate the expansion on this global set of reactions according to the following rules: a) Several randomly chosen metabolites serve as initial compounds; b) the expansion condition C2 is used which is more restricted than condition C1; c) all candidate reactions are attached when proceeding from one generation to the next.

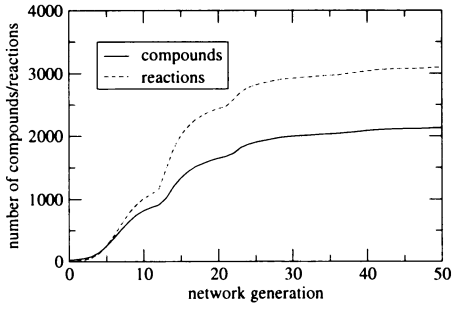

Figure 4: Evolution of the number of compounds and reactions during network expansion. The process has been started with 20 randomly chosen compounds.

Figure 4 shows the development of the total numbers of compounds and reactions during one particular expansion. The curve is characterised by a strong initial increase and saturation in later generations. Interestingly, the algorithm stops after 50 generations resulting in a network of about 2100 compounds which is considerably smaller than the global network. This behaviour is due to the fact that at this stage no candidate reaction can be found fulfilling the strong expansion condition C2.

Similar results were obtained for other sets of initial compounds leading to final networks of different sizes. The strong initial increase of the network size results from the fact that in this phase with each additional reaction a large number of new compounds are recruited which in turn increase the number of candidate reactions for the next generation. Another interesting feature is the tendency of the curve to level off in intermediate stages followed by another strong increase. We explain this by the existence of "bottlenecks" in the global network where only few reactions link highly connected domains.

An intriguing question is in which generation a certain compound is acquired. Naturally, the particular value is highly dependent on the choice of initial compounds. We have therefore run a large number of simulations, each starting from a random initial set of 20 compounds. Figure 5 represents the results for the two compounds NAD^+ and α-D-glucose. The diagram shows the frequency of acquisition in a certain generation. It is observed that glucose is generally acquired earlier during the expansion than NAD^+. The mean generations of acquisition for NAD^+ and glucose are 12.4 and 6.8, the standard deviations are 3.3 and 3.8, respectively.

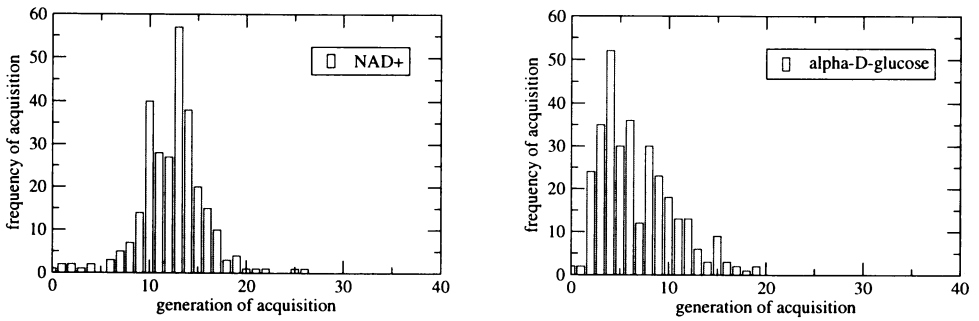

Figure 5: Histograms of the network generations at which the compounds glucose and NAD^+ are acquired during the expansion.

We have calculated corresponding histograms and their mean values for all compounds. Figure 6 shows the number of compounds having a certain mean generation of acquisition.

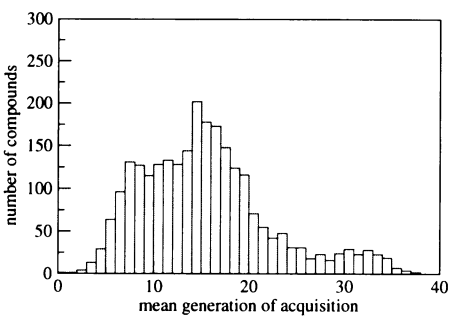

Figure 6: Histogram of mean generations of acquisition. Simulations have been performed for all compounds as explained for NAD^+ and glucose in the legend of Figure 5. The diagram includes the values of those compounds which were acquired in at least 5% of the simulations.

In Table 1, left part, are listed those ten compounds having the lowest mean generation of acquisition. For each compound the number of reactions is listed in which it participates as substrate or product (connectivity). In the right part we list those ten compounds having the highest connectivity.

Table 1: The ten compounds with the lowest mean generation of acquisition (left side) and the ten compounds with the highest connectivity (right side).

compound	generation	conn.
Water	0	1804
Oxygen	1,83	648
H^+	2,11	1400
CO_2	2,2	333
H_2O_2	2,33	147
NH_3	2,75	274
Orthophosphate	3,03	329
O_2^-	3,25	9
Carbonic acid	3,42	3
Pyruvate	3,56	146

compound	conn.	generation
Water	1804	0
H^+	1400	2,11
Oxygen	648	1,83
NAD^+	582	12,41
NADH	574	13,44
$NADP^+$	556	13,46
NADPH	554	14,39
ATP	421	10,4
CO_2	333	2,2
Orthophosphate	329	3,03

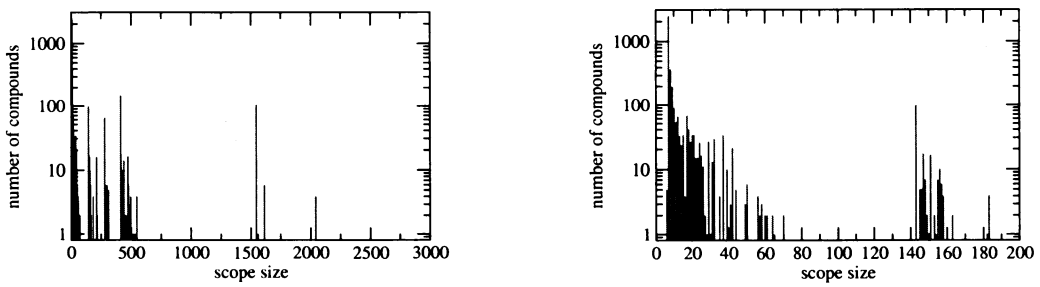

Figure 7: Numbers of compounds leading to a scope of a certain size. A section of the histogram for low sizes is shown on the right. The calculations were performed under the assumption that water is highly abundant.

It can be seen that five compounds appear on both sides meaning that compounds with a high connectivity tend to be acquired early in the expansion. Interestingly, there is no one to one correspondence between these two tables, an observation which will be discussed below.

4.2 Scope of Compounds

Network expansion can also start with a single initial compound (A). Again we apply the strong expansion condition C2 resulting in a final network containing besides A all possible compounds which can be synthesised from A. This set of compounds we denote by $\Sigma(A)$ and call it the scope of A. If a compound B lies in the scope of A, then the scope of B is a subset of the scope of A:

$$B \in \Sigma(A) \Rightarrow \Sigma(B) \subseteq \Sigma(A). \tag{4}$$

There exist compounds which can be interconverted. For two such compounds A and B, the following implication holds true due to relation (4):

$$B \in \Sigma(A) \text{ and } A \in \Sigma(B) \quad \Leftrightarrow \quad \Sigma(A) = \Sigma(B). \tag{5}$$

Figure 7 shows a histogram of the scope sizes. It can be observed that most scopes are of a size smaller than 500. There are only a few scopes of very large size. Not all scope sizes occur leading to gaps in the histogram. Scopes exist with almost all sizes below 50.

Of special interest are the scopes of large sizes. Closer inspection shows that the largest set of 2042 compounds represents one scope that can be reached from four initial compounds which are

3'-phosphoadenosine 5'-phosphosulfate (PAPS), adenosine 5'-phosphosulfate (APS), dephospho-CoA, and UDP-6-sulfoquinovose. From the property of scopes expressed by Eq. (5) it follows that these four compounds are interconvertible. Similarly, the set of 1545 compounds represents a single scope which can be reached by 105 initial compounds. Among these are central cofactors such as ATP, GTP, NADH, NADPH as well as the compound UDPglucose. Further investigation reveals that the latter scope is a subset of the largest scope. This means, for example, that ATP can be synthesised from PAPS whereas the opposite process is not possible.

4.3 Expansion with Selection Rules

So far we expanded the global network by adding at the transition from one generation to the next all candidate reactions allowed by the expansion condition C2. Now we perform expansion on the global network in a very similar way as it was done in section 3 on the glycolytic pathway. Again, we start from glucose, apply the weak expansion condition C1, and attach in each generation the one reaction which maximises the selection value f given in Eq. (3). We assume that during expansion all compounds can in principle be internalised.

We compare this process to the expansion where in each generation one randomly chosen candidate reaction is attached. In each case we characterise the expansion by the development of the numbers of external and internal compounds as well as the numbers of independent steady state fluxes and internal cycles. For the first 30 generations, the results are shown in Figure 8.

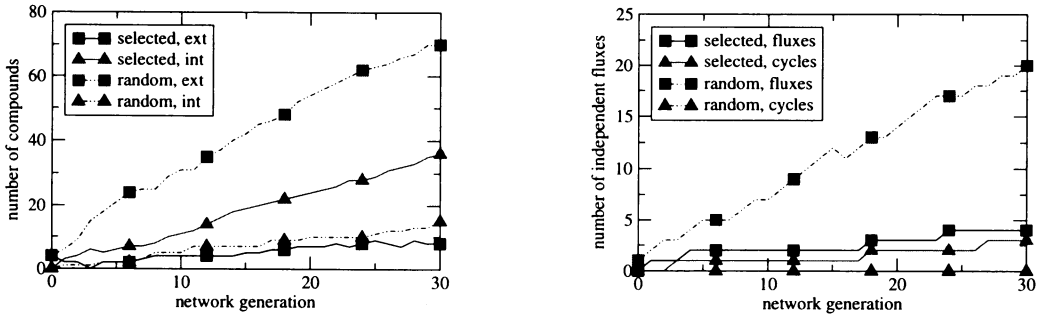

Figure 8: Initial phase of the expansion on the global network. Left: development of the numbers of external (squares) and internal (triangles) compounds for random (dashed lines) and selected (solid lines) expansion. Right: development of the numbers of independent steady state fluxes (squares) and internal cycles (triangles). Dashed lines are used for random and solid lines for selected expansion.

Figure 8 demonstrates that the two expansions proceed in significantly different ways. With the selection rule, the number of external compounds always remains below the number of internal compounds. For the random expansion the opposite is true. It can be seen that for the random expansion there exist a higher number of fluxes which can be explained by the greater amount of external compounds. The number of internal compounds remains so low that in the observed phase of the random expansion no cycles appear. Interestingly, the networks appearing during the expansion with selection rule possess a small number of independent fluxes but a comparably high number of internal cycles. This means that these networks are rather independent from external compounds and additionally can interconvert them in a very flexible manner.

5 Discussion

The main results of our analysis can be summarised as follows. 1) Single metabolic pathways as well as large parts of the global network of cellular metabolism can be reconstructed by simulating evolu-

tionary processes resulting in series of networks of increasing size. 2) Depending on the availability of resources, different modes of expansion lead to the glycolytic pathway. 3) Expansion on the global network yields a ranking of the metabolites according to the mean generation in which they were acquired when starting from randomly selected initial metabolites. All compounds which are on average acquired early in the expansion are well known core components of the metabolism and moreover are rather simple in their chemical structure. 4) Compounds of a network may be characterised by their scope defining the set of substances that can be synthesised through the available reactions without further use of other resources. Compounds with a large scope are found to be composed of a rather large number of chemical groups. 5) Expansion processes obeying selection rules proceed significantly different to random expansion. Selected expansion results in networks being able to perform relatively few interconversions but flexibly via various different internal routes whereas random expansion results in networks which can perform many interconversions but are highly dependent on external compounds.

We observed that the compounds with a low mean generation of acquisition are not necessarily those participating in many reactions. Carbonic acid, for example, is acquired very early in the expansion but has an extremely low connectivity (see Table 1). This can be understood by the fact that it can be produced from water and carbon dioxide both having a very high connectivity and are acquired very early. Moreover, metabolites with high connectivities are not necessarily acquired early. Examples are NAD^+ and other cofactors. This can be explained by the fact that the synthesis of these compounds require a large number of steps.

The present investigation was performed by considering all reactions as reversible. For some reactions it is well known that they proceed only in one direction under physiological conditions. This information could easily be taken into account by applying a sign restrictions for the corresponding elements of the feasible solution vectors of Eq. (2). We refrained from including this information since the results would be dependent on the concentrations of external metabolites which do not represent a structural property.

We do not claim that the calculated expansion modes reflect the details of the evolutionary development during natural selection of metabolic pathways. This particularly concerns the results obtained using expansion condition C1 in which many intermediate generations depend on the abundance of external metabolites. This problem is less severe when using expansion condition C2 since networks get expanded only by such reactions which further convert metabolites made available by already existing reactions of the network.

Our studies use only structural information about the biochemical reactions. However, it is conceivable to include further biological information. For example, using the results from comparative genomics, one can identify enzymes which are closely related to enzymes catalysing a biochemical reaction which is already present in the network. Such additional information could be incorporated into the specific selection rules.

Further refinement of the methods presented here may contribute to reveal whether the history of network evolution is encoded in their present design.

References

[1] Ebenhöh, O. and Heinrich, R., Evolutionary optimization of metabolic pathways. Theoretical reconstruction of the stoichiometry of ATP and NADH producing systems, *Bull. Math. Biol.*, 63:21–55, 2001.

[2] Ebenhöh, O. and Heinrich, R., Stoichiometric design of metabolic networks: Multifunctionality, clusters, optimization, weak and strong robustness, *Bull. Math. Biol.*, 65:323–357, 2001.

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

[4] Kanehisa, M. and Goto, S., KEGG: Kyoto Encyclopedia of Genes and Genomes, *Nucleic Acids Res.*, 28:27–30, 2000.

[5] Schilling, C.H., Schuster, S., Palsson, B.O., and Heinrich, R., Metabolic pathway analysis: Basic concepts and scientific applications in the post-genomic era, *Biotechnol. Prog.*, 15:296–303, 1999.

[6] Stephani, A., Nuño, J.C., and Heinrich, R., Optimal stoichiometric design of ATP-producing systems as determined by an evolutionary algorithm, *J. Theor. Biol.*, 199:45–61, 1999.