Chemical Organizations in Natural Reaction Networks

Dissertation zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.) vorgelegt dem Rat der Fakultät für Mathematik und Informatik der Friedrich-Schiller-Universität Jena

von Dipl.-Systemwiss. Florian Centler,
geboren am 14.11.1975 in Dortmund

Gutachter:

1. PD Dr. Peter Dittrich
2. Prof. Dr. Stefan Schuster
3. Ao. Univ. Prof. Dr. Christoph Flamm

Tag der letzten Prüfung des Rigorosums: 11.09.2008

Tag der öffentlichen Verteidigung: 15.09.2008

To my loving parents.

Acknowledgements

This thesis owes its existence to the help and support of numerous people. Foremost, I am deeply indebted to my supervisor

Peter Dittrich.

I am grateful to:

Clemens Beckstein,
Christoph Flamm,
Gerd Grünert,
Christoph Kaleta,
Ralph Lehmann,
Naoki Matsumaru,
Hari Nair,
Stefan Schuster,
Pietro Speroni di Fenizio, and
Klaus-Peter Zauner.

Abstract

A living cell consists of a tremendous number of components that interact in complicated ways sustaining the processes of life. Knowledge of the inner workings of such systems is commonly portrayed as networks on different levels: gene regulatory networks, metabolic networks, and signal transduction networks. With the advent of the new field "systems biology", whole cell models come into reach that integrate these different networks and further cellular processes.

In this thesis, the theory of chemical organizations, which has recently been developed by Dittrich and Speroni di Fenizio (2007) extending ideas from Fontana and Buss (1994), is applied to biochemical reaction network models as a novel analysis technique that can deal with such integrated whole cell models. As kinetic data is not required for the analysis, the method is well suited for biological systems where such data is often scarce and hard to come by. The reaction network model is decomposed into subnetworks that are algebraically closed and self-maintaining. Being algebraically closed, such subnetworks cannot produce any novel species that are not yet part of the subnetwork. As they are self-maintaining, all species that are consumed are produced within the subnetwork at sufficient rates for their maintenance. These two properties make such subnetworks, termed *organizations*, likely to persist in time. They embody all potential steady state and growth state species combinations of the model. The dynamics of the system in state space can be mapped into the space of organizations, providing a new perspective on the system.

Applying the method to an atmospheric reaction network model of Mars, a model of bacteriophage lambda, and models of *Escherichia coli* of varying sizes shows that these natural reaction

networks contain non-trivial organization hierarchies. Organizations are often found to be related to biological functions and states. The method is proven to be a useful tool in the analysis and validation of biochemical reaction network models and the prediction of their potential dynamic behavior.

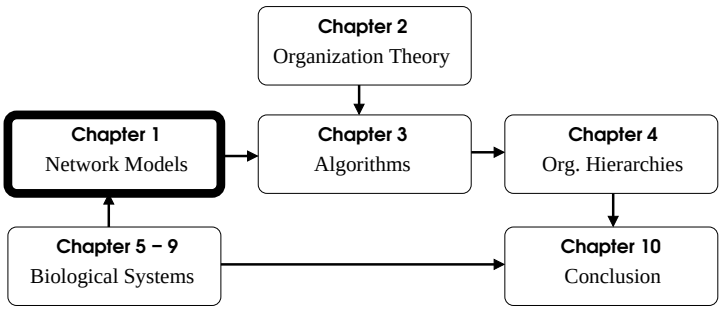
Deutsche Zusammenfassung

Eine lebende Zelle besteht aus einer sehr großen Anzahl von Komponenten, die in vielfältiger Art und Weise miteinander interagieren und damit die Prozesse des Lebens aufrecht erhalten. Erkenntnisse über die inneren Wirkungszusammenhänge solcher Systeme werden gewöhnlich in Netzwerkmodellen auf verschiedenen Ebenen dargestellt: Genregulationsnetzwerke, metabolische Netzwerke und Signaltransduktionsnetzwerke. Mit dem Entstehen der neuen Forschungsrichtung "Systembiologie" werden ganzheitliche Modelle denkbar, die diese verschiedenen Netzwerkmodelle und weitere zelluläre Prozesse integrieren.

In dieser Arbeit wird die kürzlich von Dittrich und Speroni di Fenizio (2007), aufbauend auf Ideen von Fontana und Buss (1994) entwickelte Theorie chemischer Organisationen als neue Analysemethode, die mit solchen ganzheitlichen Modellen zurecht kommt, auf verschiedene biochemische Reaktionsnetzwerkmodelle angewandt. Kinetische Informationen sind für die Analyse nicht notwendig. Dadurch ist die Methode gut für biologische Systeme geeignet, wo kinetische Details selten bekannt und schwer zu beschaffen sind. Das Reaktionsnetzwerkmodell wird in Teilnetze zerlegt, die algebraisch abgeschlossen und selbsterhaltend sind. Abgeschlossenheit bedeutet, dass das Teilnetz nicht in der Lage ist neue Spezies zu produzieren, die nicht bereits zum Teilnetz gehören. Aufgrund der Selbsterhaltung werden alle Spezies, die verbraucht werden, von dem Teilnetz in hinreichend hohen Raten für ihre Erhaltung produziert. Diese beiden Eigenschaften erlauben solchen Teilnetzen, genannt *Organisationen*, über die Zeit bestehen bleiben zu können. Sie stellen alle potentiellen Kombinationen von Spezies in Fließgleichgewichten und Wachstumszuständen dar. Die Dynamik im Zustandsraum kann in den Raum der Organisationen abgebildet werden, wodurch eine neue Perspektive auf das System ermöglicht wird.

Bei der Anwendung der Methode auf ein Reaktionsnetzwerkmodell der Marsatmosphäre, auf ein Modell der Bakteriophage Lambda und auf verschieden große Modelle von *Escherichia coli* zeigt sich, dass diese natürlichen Reaktionsnetzwerke nichttriviale Organisationshierarchien besitzen. In vielen Fällen entsprechen Organisationen biologischen Funktionen und Zuständen. Es zeigt sich, dass die Methode ein hilfreiches Werkzeug für die Analyse und Validierung von biochemischen Reaktionsnetzwerkmodellen und die Vorhersage ihrer potentiellen Dynamik ist.

Contents



Chapter 1

Introduction

Network models have become common in a wide range of disciplines ranging from microbiology to social sciences (Strogatz, 2001). The focus of research has moved from studying small parts of the system to a wider perspective, considering the whole system at once. Emergent phenomena arising from the local interplay of the system elements can only be understood by employing this perspective. Studying the local interactions separately is not sufficient (Kitano, 2002). In the biological sciences, this trend towards a more holistic approach has given rise to a new field termed "systems biology" (Ideker et al., 2001; Kitano, 2001; Klipp et al., 2005a; Palsson, 2006).

1.1 Network Models in Biology

In the biological sciences, network models are used over a wide range of scales, from the microbial level up to the ecosystem level. On the small end of the scale, a living cell is a complex system consisting of a tremendous number of components that interact in complicated ways sustaining the processes of life. Knowledge about these interactions is commonly portrayed in network models (Alm and Arkin, 2003; Alon, 2003; Barabási and Oltvai, 2004; Bower and Bolouri, 2000; Bray, 2003). Three types of intracellular networks are usually distinguished: *gene regulatory networks*, *signal transduction networks*, and *metabolic networks*.

Gene Regulatory Networks. Genes are stretches of DNA that are transcribed into mRNA. The mRNA fragments then are translated into proteins that carry out the cellular processes. The activation of genes depends on the binding of specific proteins called transcription factors to the promoter region of a gene. Within gene regulatory networks, the nodes represent genes, and the links specify how a gene activates or inactivates other genes. Gene regulatory networks are modeled using a variety of different techniques. See the review by de Jong (2002) for an overview.

Signal Transduction Networks. Cells detect external signals by receptor proteins permeating the cell membrane. The signal is then propagated inside the cell, for example by means of a phosphorylation cascade, and finally reaches the nucleus where a change in transcription constitutes the response to the extracellular signal. In such networks, the nodes represent proteins and compounds that are involved in signal detection, propagation, and processing. The links can detail biochemical reactions or on a more abstract level functional and causal relationships between the nodes. A review on the reconstruction and analysis of signal transduction networks is provided by Papin et al. (2005).

Metabolic Networks. The uptake and utilization of nutrients within the cell is detailed in metabolic networks. Nutrients enter the cell and are consecutively transformed into different metabolites by enzymes. These processes can be modeled in networks where the nodes represent the metabolites and the links the enzymatic reactions. An overview on modeling and analyzing metabolic networks is given in Heinrich and Schuster (1996).

Beside these intracellular networks, many more biological networks exist. Considering a slightly larger scale, cells communicate with each other in intercellular networks as the endocrine network (Potapov et al., 2006) and the immune network (Coutinho, 1995; Varela and Coutinho, 1991). On an even larger scale, the species within an ecosystem form foodwebs (Dunne, 2006), a special network describing which species depend on which others as their food source. And finally,

on a more abstract level, the evolutionary lineage of species can be depicted in phylogenetic trees and networks (Legendre and Makarenkov, 2002).

Although the theory of chemical organization can be applied to network models on all scales mentioned, this thesis focuses on the application to chemical and especially intracellular networks. The following section gives an overview of current methods to study such network models.

1.2 Current Methods to Analyze Biological Networks

Models of biological systems and their analysis can generally be separated into three domains: kinetic models, stoichiometric models, and network-based analysis methods (Deville et al., 2003; Stelling, 2004). The domains differ on two distinct scales: required knowledge and level of abstraction (see Figure 1.1). For kinetic models, a precise knowledge of the biological system and all its relevant components is required. Exact reaction mechanisms and kinetic rate constants must be known. As such a model tries to resemble the original model as close as possible, the level of abstraction is low. For stoichiometric models, only the reaction mechanisms including stoichiometric coefficients must be known. As kinetics are not required, less information is needed to create such models. They are more abstract than precise kinetic models. Even less knowledge is required for network models where only the relation between biochemical compounds is considered. Here, the abstraction level is the highest. The following sections summarizes the different types of models and current approaches to analyze them.

1.2.1 Kinetic Models

Using kinetic models, the biological system is considered as a dynamical system (Tyson et al., 2001). For each system component, an ordinary differential equation details its evolution over time in dependence on the other system components. All processes and interactions within the system are formulated within these differential equations. If spatial processes like diffusion and transport are additionally considered, partial differential equations are required to describe the

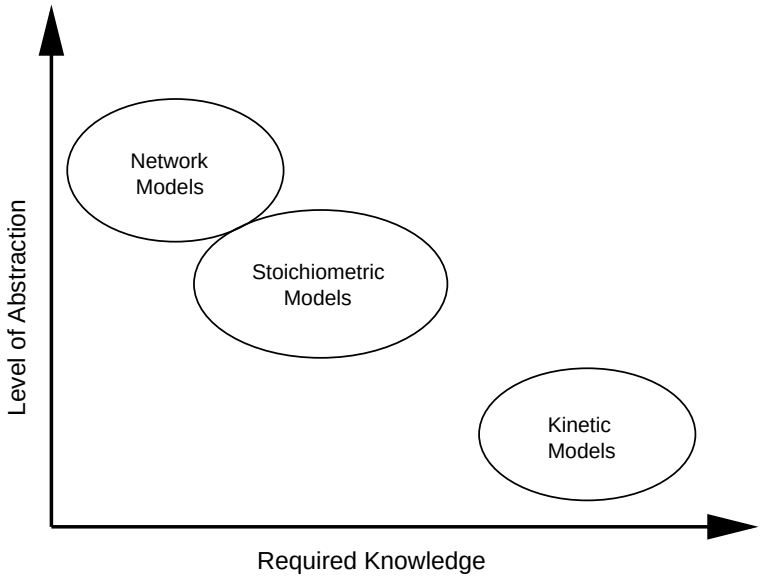


Figure 1.1: Model types differ in the knowledge about the real system that is required to build the model, and the level of abstraction the model represents.

system's dynamics. The well-established dynamical systems theory (Jetschke, 1989; Strogatz, 1994; von Bertalanffy, 1969) allows the analytical analysis of such systems. Steady states can be determined, and their stability analyzed using linear stability analysis. Bifurcation analysis helps to understand how the system's behavior changes when parameters are varied.

If the biological system contains elements having low numbers of copies, such that random fluctuations play a role, the deterministic approach is no longer valid. Stochastic models can be used in this case to account for random effects and noise (Arkin et al., 1998; McAdams and Arkin, 1999).

Kinetic models have been widely used to study all three types of biological networks: gene regulation (*e.g.*, Gardner et al., 2000; Keller, 1995), signal transduction networks (*e.g.*, Asthagiri and Lauffenburger, 2001; Ferrell and Xiong, 2001), and metabolic networks (*e.g.*, Teusink et al., 2000).

1.2.2 Stoichiometric Methods

Starting with Clarke (1980), several stoichiometric analysis methods have been developed that do not require detailed knowledge of the kinetic mechanisms and parameters that drive the network dynamics. Especially for large networks these information is often not available. But even without knowledge of the kinetic details, profound results concerning the potential dynamic behavior of the network are obtainable (Bailey, 2001).

The deficiency theory developed by Feinberg and Horn (1974) allows one to make further statements concerning the dynamical repertoire of a given network with respect to the number of steady states and potential oscillatory behavior. For certain types of networks (having deficiency zero and being weakly reversible) there exists a unique positive and stable steady state for all positive parameter values.

Methods belonging to the class of stoichiometric network analysis only consider the topology of the network, which is usually well-known. The network stoichiometry details for each reaction, which educts are transformed into which products, including how many copies of each educt are transformed into how many copies of each product. For example, the reaction $2\mathrm{A} + 3\mathrm{B} \rightarrow 4\mathrm{C}$ denotes that two copies of species A and three copies of B react to four copies of species C. The topology of a reaction network with m species and n reactions can be written as the stoichiometric $m \times n$ matrix $\mathbf{S}$. Each row of $\mathbf{S}$ corresponds to one species and each coloumn to one reaction. The entry $s_{i,j}$ details the net production of species i in reaction j. It is consumed if $s_{i,j} < 0$ and produced if $s_{i,j} > 0$. For example, consider the reaction network consisting of species A, B, and C, and the following two reactions:

$$2\mathrm{A} \rightarrow 3\mathrm{B} \tag{R1}$$

$$1\mathrm{A} + 1\mathrm{B} \rightarrow 1\mathrm{A} + 1\mathrm{C}. \tag{R2}$$

The corresponding stoichiometric matrix is:

$$\mathbf{S} = \begin{array}{c} \begin{array}{cc} \text{R1} & \text{R2} \end{array} \\ \begin{pmatrix} -2 & 0 \\ 3 & -1 \\ 0 & 1 \end{pmatrix} \end{array} \begin{array}{l} \text{A} \\ \text{B} \\ \text{C.} \end{array} \tag{1.1}$$

Note that species A acts as a catalyst in reaction R2. However, this information is lost, since only consumptions and productions are reflected in the stoichiometric matrix. The dynamics of such a system can be described by

$$d\mathbf{c}/dt = \mathbf{S} \cdot \mathbf{v}(t) \tag{1.2}$$

with concentration vector $\mathbf{c} \in \mathbb{R}^m_{\geq 0}$ and time dependent flux vector $\mathbf{v}(t) \in \mathbb{R}^n$. The flux vector assigns to each reaction a flux that describes the turnover of this reaction. If we model reversible reactions as a separate forward and backward reaction, all fluxes are nonnegative.

Stoichiometric methods usually assume that the network is in quasi steady state, demanding

$$\mathbf{S} \cdot \mathbf{v} = \mathbf{0}. \tag{1.3}$$

For example, if metabolic networks are considered, one can argue that metabolic turnover is fast in comparison to regulatory events. Hence, on longer times scales metabolite concentrations and reaction fluxes can be regarded as constant and the system as in steady state. All flux vectors $\mathbf{v}$ fulfilling the steady state condition 1.3 form the solution space containing all possible steady state flux distributions of the system. It has the shape of a convex cone originating in the point of origin.

Metabolic Flux Analysis tries to shrink the solution space defined by equation 1.3 by measuring certain fluxes (Stephanopoulos et al., 1998). However, it is mostly not possible to measure so many fluxes that the remaining fluxes can be computed and are unique (der Heijden et al., 1994; Klamt et al., 2002).

In *Flux Balance Analysis*, a linear optimization problem under constraints is solved to find a flux distribution representing an optimal function of the network (Edwards et al., 2001). The first constraint is the steady state assumption 1.3. Each flux can (but does not have to) be additionally restricted by an

upper and lower bound. The objective function is usually constructed to maximize growth or product yield. Solving this linear optimization problem helps to predict the capabilities of the network (Varma et al., 1993). And additionally, effects of gene deletions can be studied (Edwards and Palsson, 2000; Fong and Palsson, 2004).

Metabolic Pathway Analysis is concerned with the detection of pathways within reaction networks. Whereas in the previous methods a specific flux distribution was determined according to some optimality criterion, here, the whole space of admissible flux distributions is of interest. To characterize the solution space, two concepts have been used: *elementary modes* (Schuster et al., 1999) and *extreme pathways* (Schilling et al., 2000). Both concepts are linked closely to each other and share certain properties. They describe a chain of reactions that can operate at steady state. All metabolites that are included in the chain are neither produced nor consumed in the overall stoichiometry. Furthermore, all fluxes in the chain have to be thermodynamically feasible. In other words, irreversible reactions must proceed in the permitted direction. And finally, the chains have to be non-decomposable, meaning that no reaction(s) of an extreme pathway or elementary mode can be removed without violating the steady state condition.

For each network, its decomposition into extreme pathways and elementary modes is unique. The difference between both concepts was explored in detail by Klamt and Stelling (2003), and Papin et al. (2004). The set of extreme pathways is a subset of the set of elementary modes. The extreme pathways represent the edges of the convex solution space covering all admissible steady state flux distributions in the network. They are a convex basis for the network. For computing them, all internal reactions have first to be decoupled into two separate reactions for the forward and backward direction. While extreme pathways are minimal in the sense that no extreme pathway can be represented as a non-negative linear combination of other extreme pathways, elementary modes are minimal in the sense that the mode cannot operate as a functional unit anymore as soon as any of its reactions is removed. This "genetic independence" or "non-decomposability" property (Schuster et al., 2002a) is unique to elementary modes. However, combining extreme pathways can lead to elementary modes.

The concept of a minimal non-decomposable set of transitions appeared also in the domain of Petri nets (Petri, 1962). They have been used as another means to study and simulate biochemical networks (*e.g.*, Lautenbach, 1973; Voss et al., 2003). It has been shown by Zevedei-Oancea and Schuster (2003) how concepts of Petri net theory are closely related to concepts of standard stoichiometric modelling. For example, the incidence matrix of a Petri net is identical to the stoichiometric matrix of a metabolic network. Furthermore, minimal T-invariants correspond to elementary flux modes.

Although stoichiometric methods have mainly been used to study metabolic networks, they have recently also been applied to signal transduction networks (*e.g.*, Papin and Palsson, 2004a,b).

1.2.3 Network-Based Methods

Analysis methods inspired by graph theory have recently been applied to biological networks. Databases like BIND (Bader et al., 2001) store information on interactions between proteins. This information can be used to create protein-protein interaction networks, in which the nodes represent the proteins and a link connects two proteins if experimental evidence was found (*e.g.*, by yeast-2-hybrid assays) that the two proteins interact. Graph theoretical methods can be directly applied to such networks (*e.g.*, Bader and Hogue, 2003; Bu et al., 2003).

However, when networks detailing biochemical reactions are considered, the reaction network has first to be transformed into a regular graph. This can be done in two ways. Firstly, in a *substrate graph*, all molecular species become nodes. An edge between two species A and B indicates that they both participate in the same reaction (*e.g.*, used by Wagner and Fell, 2001). Alternatively, a directed link from A to B can mean that A is a substrate in a reaction in which B is a product (*e.g.*, used by Ma and Zeng, 2003). More information about the network structure is retained by using a *bipartite graph*. Here, species *and* reactions become nodes. Each reaction has incoming links from its substrates and outgoing links to its products. This modeling approach has for example been used by Jeong et al. (2000).

Analysis of the connectivity of metabolic networks has shown that they are scale free, following a power law (Jeong et al., 2000; Ma and Zeng, 2003; Wagner and Fell, 2001). In such a topology, only few species have a large number of links, serving as hubs for the network, while most species have low connectivities. The average path length between any two nodes is short in these networks. They have small world characteristics. These topological properties have been associated with robustness and error-tolerance (Albert et al., 2000). Csete and Doyle (2004) have found that metabolic networks feature a bow-tie structure in which a wide variety of nutrients are transformed into relative few core species from which a great variety of biomolecules is generated. Going beyond static network analysis, Luscombe et al. (2004) have shown how the topology of a genome scale network changes under different stimuli.

Applying graph theoretic approaches to biological networks has been very helpful in unveiling the intrinsic structure and organization of such systems. However, mostly only general properties like the overall network topology including clusters, robustness, or error-tolerance are considered. A link to the dynamic potential of the network, including for example the different functional states the system can adopt, is still missing.

1.3 Motivation and Aim of this Study

It has been recognized that when studying biological systems, the network characteristics of the system has to be taken into special account (Alm and Arkin, 2003). In many cases, studying biochemical pathways in isolation is not sufficient to elucidate the functionality of the whole system. In cell signalling, for example, distinct signal transduction pathways are coupled to each other by cross-talk (see *e.g.*, Genoud and Métraux, 1999; Houslay and Kolch, 2000).

With electronic databases like BIND (Bader et al., 2001), DIP (Xenarios et al., 2002), BioCyc (Karp et al., 2005), and KEGG (Kanehisa and Goto, 2000) collecting more and more knowledge about biochemical interactions, it has become feasible to construct genome-scale networks. Not only are the reconstructed networks becoming larger in size, but also first attempts are made to integrate different types of networks into one model (see *e.g.*, Ge et al., 2001; Grigoriev,

2001; Ideker et al., 2002; Klipp et al., 2005b). Even whole cell models have become conceivable that encompass all aspects of cellular processes (Kremling et al., 2003; Loew and Schaff, 2001; Puchalka and Kierzek, 2004; Tomita, 2001; Tomita et al., 1999). These two trends, network models becoming larger and larger, and the integration of different types of networks, make analysis using current approaches difficult. Usually not all kinetic details in large-scale models of cellular processes are known, making simulations using differential equations problematic. Stoichiometric methods like elementary flux modes face problems with large networks due to a combinatorial explosion of elementary modes (Klamt and Stelling, 2002). Network based methods can cope with large and heterogeneous networks, but the obtainable results remain somewhat vague. New approaches are required to analyze and understand the emerging large-scale network models of biological processes.

Biochemical networks can be decomposed into functional units or modules (Gagneur et al., 2003, 2004; Hartwell et al., 1999; Ravasz et al., 2002; Spirin and Mirny, 2003; von Mering et al., 2003). Such modules performing a certain cellular function are separated from each other spatially or by chemical specificity. A module within a network can be defined in many ways, for example, a subset of network species having more connections between themselves than with the remaining network species can define a module. Other definitions employ information theoretic measures (Ziv et al., 2005).

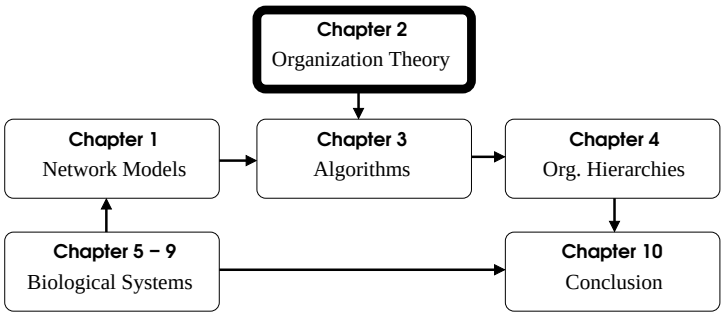
The inherent modular topology of biochemical networks can on the one hand be exploited in the modelling process when creating large-scale models (Kremling et al., 2000), and on the other hand it can be used to decompose large network models into more manageable parts (Schuster et al., 2002c).

In this thesis, the theory of chemical organizations, developed by Dittrich and Speroni di Fenizio (2007) extending ideas from Fontana and Buss (1994), will be used as another technique to decompose large reaction network models into subnetworks. Here, the modules are sets of network species that fulfill two properties: (algebraic) *closure* and *self-maintenance*. Such species sets are called *organizations*. They are defined in a strict mathematical way. A set of species is closed if the reaction network does not contain any reaction that would allow the creation of a species not yet present in the set from the set species. The set

cannot produce any novel species. The self-maintenance property ensures, that all species that are consumed within the species set are recreated from within the set at a sufficient rate for their maintenance. Hence, no species of the set vanishes. Self-maintenance can be seen as one reasonable criterion for living organisms as they have to constantly recreate their components as autopoietic systems (Varela et al., 1974). Formal definitions of the concepts closure and self-maintenance will be given in Chapter 2. All organizations of a network form an overlapping, hierarchical structure that can be visualized in a Hasse diagram. This not only gives an overview of the structure of the reaction network, but furthermore has implications for the potential dynamic behavior of the system. As an organization cannot create anything new and its components do not vanish, it can prevail over time, not changing the set of species present in the system. It can be shown that all steady states of the network coincide with organizations. While not all organizations represent steady states, each steady state can be mapped to a corresponding organization. The theory can be applied to any reaction network in which species react with each other to create new species. In particular, biochemical reaction networks like gene regulatory networks, signal transduction networks, metabolic networks, and combinations of these can be studied. Only relying on stoichiometric data and not requiring kinetic details of the reaction processes makes the method well-suited for the biological domain, where kinetic data is often hard to obtain. While other approaches assume that all network species are known in advance and present in the system, chemical organization theory explicitly allows one to study constructive dynamical systems. In such systems, species can continually vanish and novel species can appear.

Within this thesis, the theory of chemical organizations will be applied to study chemical and biochemical networks in order to assess the potential of this novel approach in the ultimate endeavor to understand the cellular processes of life within biological cells. First, the theory including several extensions (*e.g.*, connected organizations) that proved useful in the analysis of network models is introduced in Chapter 2. The link between chemical organizations and elementary modes is also elaborated in this chapter. Then, an algorithm that is able to compute all organizations for a given reaction network is described and its runtime complexity analyzed in Chapter 3. Once all organizations are determined, the

hierarchy of organizations can be studied. This is the focus of Chapter 4. Several concepts (*e.g.*, unit species sets as sets of species that always appear together in organizations) that help to "make sense" of organization hierarchies, which can become quite large, are introduced. With the concept of organization intensities, this chapter also presents a link between organizations as discrete sets of species and continuous concentration vectors as obtained by measurements or simulations. Finally, Chapters 5 to 9 apply the presented concepts and tools to several reaction network models of real systems. A photochemical reaction network model of the Martian atmosphere is analyzed in Chapter 5. A Petri net model of the genetic switch of bacteriophage lambda is studied in Chapter 6. Chapter 7 deals with a network model of the central sugar metabolism of *Escherichia coli* that contains not only metabolism, but also gene regulation and signal transduction. A more comprehensive network of *E. coli*, including the regulation of all involved genes, is studied in Chapter 8. Here, a concept is introduced that allows one to also consider inhibitory interactions within the framework of organization theory. As a last model, a genome-scale metabolic model of *E. coli* is considered in Chapter 9. Finally, Chapter 10 closes this thesis with the conclusion.



Chapter 2

Theory of Chemical Organizations

This chapter introduces the concepts and formal definitions of the theory of chemical organizations (Dittrich and Speroni di Fenizio, 2007). First, the basic concepts are presented in Section 2.1. Then, the notion of *connected organization* is introduced as an extension to the theory in Section 2.2, and its consequences are discussed. Finally, the concept of chemical organizations is related to elementary flux modes in Section 2.3. Both methods are used to study reaction networks and only require knowledge of the network structure and stoichiometry.

2.1 Basic Concepts

The state of a dynamical system at a certain time t is characterized by the values of its state variables at time t. The state space of all admissible system states is usually given by a subset of $\mathbb{R}^n$, where n is the number of state variables of the system. Under certain circumstances, this quantitative characterization of the system state might be inadequate (or even impossible due to missing kinetic information) and a qualitative description might be more appropiate. In a qualitative analysis, one could simply ask which system species are present at a certain time t. Then, the state of the system is characterized by the set of species present at that time. The state space becomes the power set $\mathcal{P}(\mathcal{M})$ over the set $\mathcal{M}$ of all system species. When considering the dynamics of the system, some species sets from $\mathcal{P}(\mathcal{M})$ will be more important than others. Extending ideas by Fontana and Buss (1994), the theory of chemical organizations identifies those sets that are most interesting with respect to their potential to persist in time. These sets are termed *organizations*. They have to fulfill two properties: algebraic closure and self-maintenance. The first property – closure – ensures that given the molecular species of an organization, there is no reaction within the reaction network that could create a species not yet present in the organization. No novel species can be generated. The second property – self-maintenance – guarantees that every molecular species that is used-up within the organization can be recreated from organization species at a sufficient rate at least for its maintenance and possibly for its accumulation.

In the remaining part of this section, formal definitions of the already mentioned main (and some further) concepts of chemical organization theory are presented. They will be illustrated on two example networks as depicted in Figure 2.1. First, the objects of study – reaction networks – are formalized as reaction networks.

2.1.1 Reaction Network (Dittrich and Speroni di Fenizio, 2007)

Let $\mathcal{M}$ be a set of elements (called species, molecular species, or just molecules). $\mathcal{P}_M(\mathcal{M})$ denotes the set of all multisets with elements from $\mathcal{M}$. A multiset differs from a set in the fact that it can contain the same element more than once. The

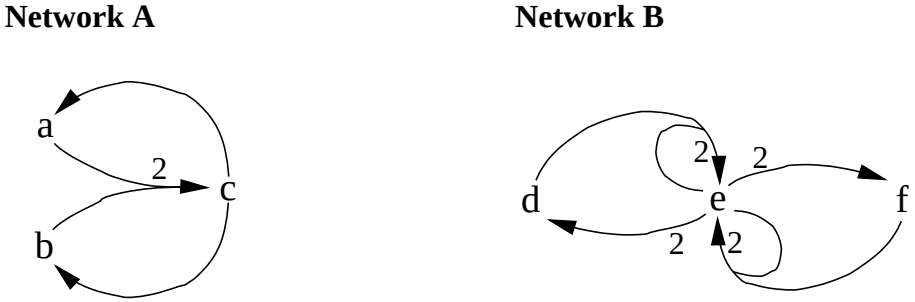


Figure 2.1: Two example networks are used to illustrate the concepts of chemical organization theory.

set of reactions $\mathcal{R}$ occuring among the species $\mathcal{M}$ can then be defined by the relation $\mathcal{R} \subseteq \mathcal{P}_M(\mathcal{M}) \times \mathcal{P}_M(\mathcal{M})$. We call the pair $\langle \mathcal{M}, \mathcal{R} \rangle$ a *reaction network*.

For simplicity, we adopt the notation from chemistry to write reactions. When $A \in \mathcal{P}_M(\mathcal{M})$ is the multiset of reactants and $B \in \mathcal{P}_M(\mathcal{M})$ is the multiset of products for a given reaction, we write $(A \to B) \in \mathcal{R}$ instead of $(A, B) \in \mathcal{R}$. When considering concrete reactions, we will use the common notation using stoichiometric coefficients instead of multisets. For example, instead of writing $(\{a, b, b\} \to \{c\}) \in \mathcal{R}$ with $a, b, c \in \mathcal{M}$, we write $a + 2b \to c$.

Note that all reactions are assumed to be irreversible. To model a reversible reaction, two separate reactions have to be explicitly defined, one for the forward direction and the other for the backward direction.

Input and output or decay reactions can be modeled by using the empty set $\emptyset$ as the multiset describing the reactants, respectively the products. For example, $\emptyset \to a$ defines a constant influx of species a into the system. Species a becomes an input species for the network. The reaction $a \to \emptyset$ indicates that species a is spontaneously removed from the system, for example by an explicit outflow, diffusion, or spontaneous decay.

The Networks A and B (*cf.* Figure 2.1) can be written as $\langle \{a, b, c\}, \{a + b \to 2c, c \to a, c \to b\} \rangle$, and $\langle \{d, e, f\}, \{2e \to d, 2e \to f, d + e \to 2e, e + f \to 2e\} \rangle$, respectively.

2.1.2 Closure (Dittrich and Speroni di Fenizio, 2007)

A set of species $S \subseteq \mathcal{M}$ is *closed*, if for all reactions $(A \to B) \in \mathcal{R}$ with $A \in \mathcal{P}_M(S)$, also $B \in \mathcal{P}_M(S)$. In other words: if the educts of a reaction are contained in S, then also its products must be in S. There is no reaction in $\mathcal{R}$ that could create a new species not yet in S from species contained in S.

Note that we define the closure in a pure algebraic way. The concept is not related to the notion of closure in the thermodynamical sense. Rather, it is closely related to the catalytic closure of autocatalytic sets (Kauffman, 1993), the closure as defined by Fontana and Buss (1994), and more generally to the concept of autopoiesis (Maturana and Varela, 1991).

Another concept related to the closure is the scope as defined by Handorf et al. (2005). Starting from a species set acting as a seed, the set is expanded until it becomes closed. The authors use this method to study the robustness and evolution of metabolic networks.

The closed sets of Network A are: $\{\}$, $\{a\}$, $\{b\}$, $\{a, b, c\}$. The closed sets of Network B are: $\{\}$, $\{d\}$, $\{f\}$, $\{d, f\}$, $\{d, e, f\}$.

2.1.3 Semi-self-maintenance (Dittrich and Speroni di Fenizio, 2007)

A set of species $S \subseteq \mathcal{M}$ is *semi-self-maintaining*, if all species that are consumed within S are also produced within S. A species s is consumed (produced) within S, if there exists a reaction $(A \to B) \in \mathcal{R}$ with both A and $B \in \mathcal{P}_M(S)$, such that s appears more often (less often) in A than in B.

The semi-self-maintaining sets of Network A are: $\{\}$, $\{a\}$, $\{b\}$, $\{a, b, c\}$. The semi-self-maintaining sets of Network B are: $\{\}$, $\{d\}$, $\{f\}$, $\{d, e\}$, $\{d, f\}$, $\{e, f\}$, $\{d, e, f\}$.

2.1.4 Semi-organization (Dittrich and Speroni di Fenizio, 2007)

A set of species $S \subseteq \mathcal{M}$ that is closed and semi-self-maintaining is called a *semi-organization*.

The semi-organizations of Network A are: $\{\}$, $\{a\}$, $\{b\}$, $\{a, b, c\}$. The semi-organizations of Network B are: $\{\}$, $\{d\}$, $\{f\}$, $\{d, f\}$, $\{d, e, f\}$.

2.1.5 Self-maintenance (Dittrich and Speroni di Fenizio, 2007)

Given a reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$ with $m = |\mathcal{M}|$ species and $n = |\mathcal{R}|$ reactions, its stoichiometric matrix $\mathbf{S}$, a set of species $S \subseteq \mathcal{M}$ is called *self-maintaining* if a flux vector $\mathbf{v} \in \mathbb{R}^n_{\geq 0}$ exists, such that the following three conditions are fulfilled:

(1) For every reaction $(A \to B) \in \mathcal{R}$ with $A \in \mathcal{P}_M(S)$, its corresponding flux is $v_{A \to B} > 0$.

(2) For every reaction $(A \to B) \in \mathcal{R}$ with $A \notin \mathcal{P}_M(S)$, its corresponding flux is $v_{A \to B} = 0$.

(3) For every species $i \in S$, its concentration change is nonnegative: $(\mathbf{Sv})_i \geq 0$.

In other words: if we consider only the subnetwork made up by the species of S and additionally the species that can be created from S (but are not in S) (conditions (1) and (2)), we can find a positive flux vector, such that no species of S decays (condition (3)).

Note that every set that is self-maintaining is also semi-self-maintaining. The property of self-maintenance is stronger than that of semi-self-maintenance.

The self-maintaining sets of Network A are: $\{\}$, $\{a\}$, $\{b\}$, $\{a, b, c\}$. The self-maintaining sets of Network B are: $\{\}$, $\{d\}$, $\{f\}$, $\{d, f\}$.

2.1.6 Organization (Dittrich and Speroni di Fenizio, 2007)

A set of species $S \subseteq \mathcal{M}$ that is closed and self-maintaining is called an *organization*.

Note that every organization is also a semi-organization, while the opposite is not true. It is important to note that it is only in principle that an organization is able to prevail in time. Albeit a flux vector exists that allows for the persistence of all organization species, it is not guaranteed that this flux vector can be realized in the real system or model. Further kinetic information including reaction rates are usually necessary to decide this question. However, organizations represent an exhaustive enumeration of all species combinations that might have the potential for persistence over time.

Note that if input species are defined for a network, all these input species will be part of all organizations of the network.

The organizations of Network A are: $\{\}$, $\{a\}$, $\{b\}$, $\{a, b, c\}$. The organizations of Network B are: $\{\}$, $\{d\}$, $\{f\}$, $\{d, f\}$.

2.1.7 Balanced Organization

An organization $O \subseteq \mathcal{M}$ is a *balanced organization*, if there exists a flux vector according to conditions (1) and (2) of the definition of self-maintenance (see Section 2.1.5), so that the concentration change of all organization species is zero. Hence, condition (3) becomes:

(3') For every species $i \in O$, its concentration change is zero: $(\mathbf{Sv})_i = 0$.

All organizations in the example Networks A and B are balanced. However, if we add the reaction $a \longrightarrow 2a$ in Network A, the organizations $\{a\}$ and $\{a, b, c\}$ are no longer balanced.

2.1.8 Generator

Given an organization $O \subseteq \mathcal{M}$, a species set $G \subseteq \mathcal{M}$ is called a *generator* of O, if the closure of G contains O. The closure of G is the smallest closed set containing G[1]. Note that a generator does not need to be a subset of the organization it generates. This can be illustrated with Network A: when inspecting species c, we find that it directly creates species a and b. Hence, the closure of $\{c\}$ contains the whole network, and with it all of its organizations. Species set $\{c\}$ is therefore a generator for all organizations of the network, including the organizations $\{a\}$ and $\{b\}$.

The organizations and their generators in Network A are: Organization $\{\}$: $\mathcal{P}\{a, b, c\}$; Organization $\{a\}$: $\{a\}$, $\{c\}$, $\{a, b\}$, $\{a, c\}$, $\{b, c\}$, $\{a, b, c\}$; Organization $\{b\}$: $\{b\}$, $\{c\}$, $\{a, b\}$, $\{a, c\}$, $\{b, c\}$, $\{a, b, c\}$; Organization $\{a, b, c\}$: $\{c\}$, $\{a, b\}$,

[1] Note that a stricter definition of the generator is possible: G is a generator of O, if O is the largest self-maintaining set (or one of the largest, if it is not unique) contained in the closure of G. However, we will use the concept of generator, respectively seed, in Chapter 4 to assess how difficult it is to remove an organization from a reaction vessel. For this purpose, the given more lax definition is more appropriate.

$\{a, c\}$, $\{b, c\}$, $\{a, b, c\}$. The organizations and their generators in Network B are: Organization $\{\}$: $\mathcal{P}\{d, e, f\}$; Organization $\{d\}$: $\{d\}$, $\{e\}$, $\{d, e\}$, $\{d, f\}$, $\{e, f\}$, $\{d, e, f\}$; Organization $\{f\}$: $\{e\}$, $\{f\}$, $\{d, e\}$, $\{d, f\}$, $\{e, f\}$, $\{d, e, f\}$; Organization $\{d, f\}$: $\{e\}$, $\{d, e\}$, $\{d, f\}$, $\{e, f\}$, $\{d, e, f\}$.

2.1.9 Seed

Given an organization $O \subseteq \mathcal{M}$, a set of species $S \subseteq \mathcal{M}$ is called a *seed* of O, if S is a generator of O and there is no other generator G with $G \subset S$.

As seeds are generators, they also do not need to be subsets of the organization they generate. Note that an organization can have more than one seed. For example in Network A, the organization $\{a\}$ has two generators that do not contain smaller generators as subsets. The organization has the two seeds $\{a\}$ and $\{c\}$.

The organizations and their seeds in Network A are: Organization $\{\}$: $\{\}$; Organization $\{a\}$: $\{a\}$, $\{c\}$; Organization $\{b\}$: $\{b\}$, $\{c\}$; Organization $\{a, b, c\}$: $\{c\}$, $\{a, b\}$. The organizations and their seeds in Network B are: Organization $\{\}$: $\{\}$; Organization $\{d\}$: $\{d\}$, $\{e\}$; Organization $\{f\}$: $\{e\}$, $\{f\}$; Organization $\{d, f\}$: $\{e\}$, $\{d, f\}$.

2.1.10 Hierarchy of Organizations

Since organizations may share the same species, the set of organizations together with the set inclusion $\subseteq$ form a partially ordered set that can be visualized in a Hasse diagram providing a hierarchical view on the network under consideration (see Figure 5.1 on Page 70 for an example). Organizations are vertically arranged according to their size, with small organizations at the bottom. Two organizations are connected by a line if the upper contains the lower organization and no other organization exists between them. The label of an organization in the Hasse diagram contains a list of species contained in that organization. To keep the labels short, only those species are usually listed that are not already contained in organizations to which a downlink exists. Hence, to get the complete list of molecular species of an organization, it is necessary to collect the molecular species contained in organizations to which a downlink exists plus the species

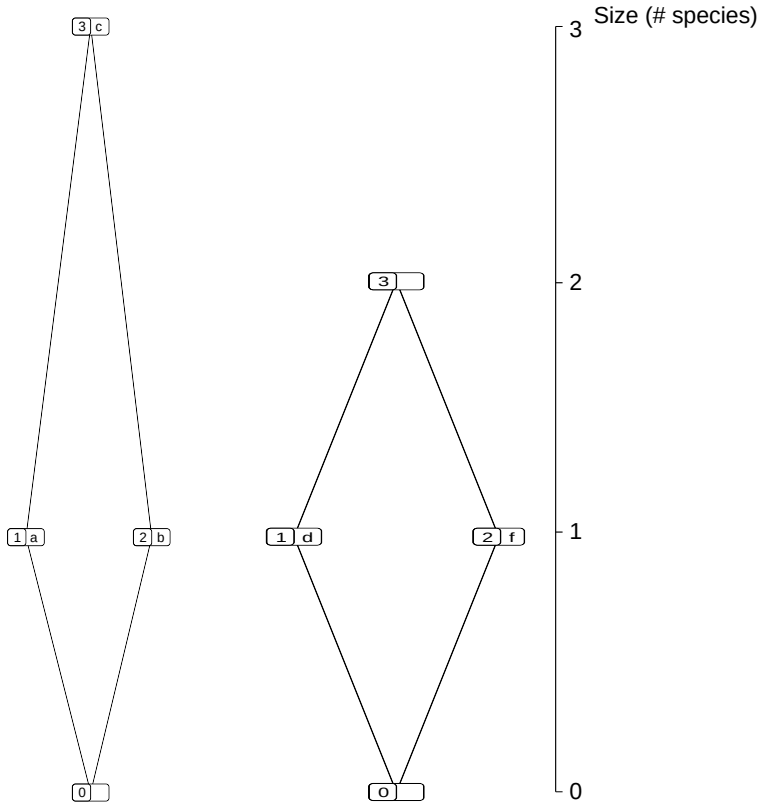


Figure 2.2: Hasse diagrams of the organizations in the two example Networks A (left) and B (right). Network A contains the organizations $\{\}$, $\{a\}$, $\{b\}$, and $\{a, b, c\}$. Network B contains the organizations $\{\}$, $\{d\}$, $\{f\}$, and $\{d, f\}$.

denoted in the organization label. The organizations of the example Networks A and B are depicted as Hasse diagrams in Figure 2.2. The Hasse diagram provides a hierarchical view on the network under consideration. If trajectories are available, the dynamic movement of the system in state space can be mapped to a movement in the space of organizations, as represented by the Hasse diagram (Dittrich and Speroni di Fenizio, 2007). For an important class of reaction networks, namely consistent reaction networks, including catalytic flow networks, and reactive flow networks with and without persistent molecules, the hierarchy of organizations forms a lattice (Dittrich and Speroni di Fenizio, 2007).

2.2 Connected Organizations

The computation of organizations for a given reaction network can yield a tremendous number of organizations. One mechanism leading to a huge number of organizations for certain reaction networks is a simple combinatorial explosion. Consider a reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$ with $m = |\mathcal{M}|$ chemical species but without any reactions: $\mathcal{R} = \emptyset$. In this case, any combination of species from $\mathcal{M}$ is an organization. The property of closure is satisfied as there is no reaction able to create novel species. Without any reactions, the self-maintenance property is also fulfilled as no species concentration decreases. As any species combination forms an organizations, there are $\sum_{i=0}^{m} \binom{m}{i} = 2^m$ organizations. However, none of these organizations contains chemically reacting species.

In order to avoid this combinatorial explosion of organizations with non-interacting species, it is appropriate to only consider those organizations, in which all species are connected to each other by reactions. These organizations form connected subnetworks without any isolated species. We term them *connected organizations*.

2.2.1 Definition of Connected Organizations

Given a reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$ and an organization O, the organization is a *connected organization*, if it is empty, or there exists a species $s \in O$ so that all species of O are connected to s. Two species s_i and $s_j \in O$ are connected to each other, if there exists a sequence of n species $s_1, \cdots, s_n$ with $s_k \in O$ for $k = 1, \cdots, n$, such that s_i and s_1, s_k and s_{k+1} for $k = 1, \cdots, n-1$, and s_n and s_j are directly connected. Two species s_o and $s_p \in O$ are directly connected, if there exists a reaction $(A \to B) \in \mathcal{R}$ with $A \in \mathcal{P}_M(O)$, $s_o \in A \cup B$, and $s_p \in A \cup B$.

All input species of the network (*i.e.*, all species that appear as products in reactions in which the reactant side is the empty set) are defined as being connected to each other.

In the example Network A (see Figure 2.1), all four organizations {}, {a}, {b}, and {a,b,c} are connected organizations. For Network B, only the organizations {}, {d}, and {f} are connected. Organization {d,f} is not a connected organization.

2.2.2 Connected Organizations as a Basis for all Organizations

The connected organizations can be used to construct all organizations of the network. They can be viewed as a basis for the complete hierarchy of organizations. If the network does not contain input species, every organization is a combination of connected or basis organizations. If n is the number of basis organizations, $\sum_{i=0}^{n} \binom{n}{i} = 2^n$ different set combinations exist. However, not every combination of basis organizations gives an organization. For example, consider the simple reaction network containing three species and one reaction $\langle\{a, b, c\}, \{a + b \rightarrow c\}\rangle$. Species $\{a\}$ and $\{b\}$ are two connected organizations. As such they are part of the basis, but their combination $\{a, b\}$ lacks the properties of closure and self-maintenance and hence is not an organization. Consequently, to obtain all organizations from the connected organizations, set unions of all combinations of basis organizations have to be considered and tested for the organization properties.

If no input species are defined for the reaction network, the basis organizations are exactly the connected organizations. In the presence of input species, the basis is larger. Firstly, again all connected organizations are basis organizations. Secondly, the inflow reactions of the input species must be removed from the network. The connected organizations of the resulting network are additionally basis organizations. This step is required to find connected subnetworks that are not connected to input species. In this case, not all set union combinations must be tested, since all organizations contain at least the input species. The whole procedure can be summarized in four steps:

1. For the given reaction network, compute the set of connected organizations O_{init}.

2. Remove all inflow reactions and compute the set of connected organizations for the modified network $O_{withoutinput}$.

3. The set of basis organizations is $O_{basis} = O_{init} \cup O_{withoutinput}$.

4. Make set unions of all possible combinations of organizations from O_{basis} such that exactly one organization from O_{init} is contained in every combination. (If a combination of organizations from O_{init} is already an organization,

it is already an element of O_{init}.) Test the species set of each combination for the closure and self-maintenance property. With $|O_{init}| = m$ and $|O_{withoutinput}| = n$, there are $m \cdot \sum_{i=0}^{n} \binom{n}{i} = m \cdot 2^n$ species sets to be tested.

To show that this procedure is sufficient to create all organizations, we need to proof that any organization is a combination of basis organizations. For this purpose, networks with and without input species will be discussed separately.

Networks without input. If the network has no input species, the basis organizations are exactly the connected organizations. Taking any organization O, we find that it is either connected or not. In the former case, it is a basis organization. In the later case, it consists of two or more parts that are not connected to each other. When inspecting each isolated part separately, we find that each part is closed and self-maintaining. In other words, each part is an organization. Even more, each part is a connected organization and hence a basis organization. Therefore, the unconnected organization O is equal the set union of these basis organizations.

Networks with input. Again, taking any organization O of the network, we find that it is either connected or not. If it is connected, it is already a basis organization. If not, we again inspect the isolated parts of the organization. Like in the case without input species, all parts are closed, self-maintaining, and connected. Some parts contain input species and others not. Recall that in the presence of input species, all input species are present in all organizations. Hence, the union of all isolated parts that contain at least one input species will be an organization (and contained in O_{init}). Parts without input species are only organizations in the absence of input species, and hence contained in $O_{withoutinput}$. We find that all isolated parts of organization O can be associated to basis organizations in O_{init} and $O_{withoutinput}$. Consequently, O is equal the set union of these basis organizations.

We conclude that all organizations are created using the described procedure. It must be noted that the basis organizations do not form a basis for all organizations that is minimal. Consider the reaction network consisting of two species

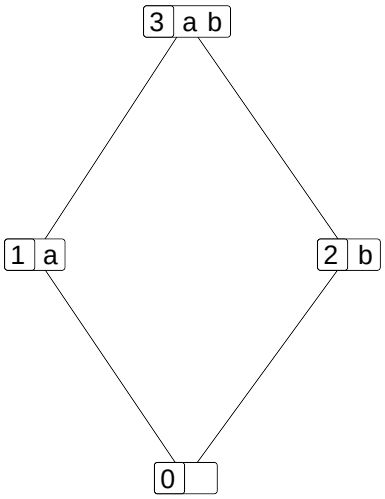


Figure 2.3: For the chemical reaction network $\langle\{a, b\}, \{a + b \rightarrow 2a + 2b\}\rangle$, all four organizations are connected and hence part of the basis. Since Organization 3 is the union of Organizations 1 and 2, the set of basis organizations is not minimal.

and one reaction $\langle\{a, b\}, \{a + b \rightarrow 2a + 2b\}\rangle$. This system contains four organizations as depicted in Figure 2.3. All four organizations are connected and therefore basis organizations. However, Organization 3 is the set union of Organizations 1 and 2, and hence would not be required in the basis. In this sense, the set of basis organizations does not form a basis for all organizations that is minimal.

2.3 Organizations and Elementary Modes

with Christoph Kaleta

Chemical organization theory and the concept of elementary flux modes both rely solely on network topology and neglect any kinetic data. In this section, we highlight the similiarities and differences of these two concepts (Kaleta et al., 2006).

Pathways are typically the central concept in the analysis of biochemical reaction networks. A pathway can be interpreted as a chain of enzymatical reactions performing a specific biological function. A common way to study metabolic networks is to identify minimal pathways that can operate at steady state called elementary modes (Schuster et al., 2000a). Steady states are broadly regarded as important system states of metabolic networks. Each steady state flux distribution of the system can be described as a combination of elementary modes. Furthermore, every steady state can be mapped to an organization (Dittrich and Speroni di Fenizio, 2007). This highlights the link between the two concepts: while elementary modes (or more precisely, the extreme pathways as a subset of all elementary modes, see Section 1.2.2) represent the boundaries of admissible steady state flux distributions of the network, organizations define metabolite compositions that are likely to be present at the same time in the network in biological feasible situations. On one hand, balanced organizations consist of combinations of elementary modes. On the other hand, it is possible to assign to each elementary mode a unique (and possibly empty) set of organizations, indicating the metabolites accompanying the active pathway in a feasible steady state and even in growth situations.

2.3.1 Elementary Modes

Elementary modes (Schuster et al., 2000a) have proven to be a powerful means in the analysis of metabolic networks and their underlying properties (*e.g.*, Poolman et al., 2003; Schwender et al., 2004). They have been used to assess network flexibility (Stelling et al., 2002), to find pathways with optimal yields for certain metabolites (Schuster et al., 2002b) and to study enzyme deficiencies (Schuster

and Kenanov, 2005). Since the number of elementary modes can grow exponentially with the size of the network, the study of elementary modes in larger systems is difficult (Gagneur and Klamt, 2004).

Elementary modes represent minimal sets of reactions that can operate at steady state with all reactions proceeding in their appropriate direction (Schuster et al., 2000a). The reaction set is minimal in the sense that there is no subset of reactions that could also operate at steady state. Metabolites are classified as either internal or external. While internal metabolites are required to be in steady state, external metabolites are considered to be buffered by reactions not contained in the model. They are the potential substrates and products of the pathway. Schuster et al. (2000b) have shown the close relation between elementary modes and minimal T-invariants in Petri nets (Lautenbach, 1973; Murata, 1989; Starke, 1990).

Generally, steady state solutions for a metabolic network containing n reactions can be determined in the n-dimensional flux space of the system. Each flux vector $\mathbf{v} \in \mathbb{R}^n$ in the flux space assigns to each reaction a value that represents the reaction's turnover rate. The steady state condition imposes constraints in the flux space so that the solution space containing all possible steady state flux distributions forms a convex polyhedral cone (Gagneur and Klamt, 2004). The edges of this cone are the extreme pathways. All extreme pathways are also elementary modes, and hence they contain a basis for the solution space. Accordingly, every feasible steady state flux distribution can be expressed as a linear combination of elementary modes. In this sense, elementary modes describe the boundaries of the network's potential steady state behavior.

2.3.2 Linking Elementary Modes and Organizations

In contrast to organizations, elementary modes are defined as (multi-) sets of reactions, not species. However, the concept of closure can be expanded to reactions easily. Since we intend to relate organizations to pathways made up by reaction chains, we do not consider organizations here that contain isolated species not participating in any reaction of the organization. The concept connecting elementary modes and chemical organizations is the self-maintenance property. To

elucidate this connection we have to inspect the definition of self-maintenance more closely. Self-maintenance is defined with respect to a set of species. In order to show the self-maintenance of such a set, a flux vector must exist fulfilling certain conditions. If the solution space of these conditions is empty, the set is not self-maintaining. The union of the solution spaces of all species sets lies within a convex polyhedral cone in flux space as will be shown in the following. Taking the set $\mathcal{M}$ of all m species of the network and its stoichiometric matrix $\mathbf{S}$ defining the n reactions among these species, the self-maintenance condition $\mathbf{Sv} \geq \mathbf{0}$ defines a set of m linear inequalities for the complete network. The restriction to nonnegative fluxes $\mathbf{v} \geq \mathbf{0}$ defines another set of n inequalities. The solution space of these $m + n$ linear inequalities is a convex polyhedral cone in the n-dimensional flux space. This cone, encompassing all flux distributions fulfilling the self-maintenance property, can serve as input to an algorithm that computes all organizations. As mentioned in Section 2.3.1, the elementary modes contain the edges of another convex polyhedral cone: the solution space of the equalities $\mathbf{Sv} = \mathbf{0}$ and $\mathbf{v} \geq \mathbf{0}$[1], representing all steady state flux distributions of the system. Obviously, the steady state cone lies within the self-maintenance cone.

With the steady state condition being the stricter constraint, flux vectors exist fulfilling the self-maintenance property but not the steady state condition. In such a case mass is produced and accumulates in the network. But if there exists a decay reaction of the form $a \rightarrow \emptyset$ for all metabolites $a \in \mathcal{M}$, the overproduction of species can be compensated for by the decay reactions. In such a setting we find that if a flux vector exists fulfilling the self-maintenance constraint, also a flux vector fulfilling the steady state condition exists as a linear combination of elementary modes.

This leads to the conclusion, that

1. If all metabolites decay spontaneously, we can find organizations by using the convex polyhedral cone that is spanned by the elementary modes.

[1]While elementary mode analysis allows reversible reactions, organization theory assumes that reactions can only proceed in one direction. Therefore, we assume that an explicit backward reaction is added for all reversible reaction. With this modelling, all reactions are irreversible and all reaction rates are hence nonnegative.

2. Elementary modes can be used in general reaction networks to search for balanced organizations fulfilling the steady state condition. Such balanced organizations are composed of a combination of elementary modes[2].

3. An elementary mode implies a unique set of organizations. The smallest organizations containing the mode constitute this set. If it is empty, the elementary mode cannot be present in any steady state of the system.

4. Organizations need not to contain elementary modes since they also account for positive productions of metabolites.

5. The set of metabolites taking part in an elementary mode is not necessarily self-maintaining.

The differences between elementary modes and organizations follow from the assumptions both approaches make. For example, while in elementary mode analysis it is possible to shut reactions off, in organization theory, a reaction is always assumed to be performed as long as its educts are present. Several examples of increasing complexity will be used in the following to illuminate and clarify these results.

2.3.3 Branching and Cycling Pathways

A simple linear metabolic network with one branching point is depicted in Figure 2.4 (A). An external substrate metabolite is transformed into internal metabolite A which is in turn transformed into B. From B, one path leads over C and another over D to external product metabolites. This network contains two elementary modes as depicted in Figure 2.4 (A). The first elementary mode uses the pathway including metabolite C while the second uses metabolite D. Up to the branching point B both modes are identical. Since the sets of metabolites making up the two pathways are not closed (it is possible to create C *and* D from B) they do not form organizations on their own. Indeed, the only organization of this

[2]Note that since elementary modes only consider reactions, this approach cannot find organizations that contain species that do not participate in any reaction within the organization. However, such organizations can be determined in a second computational step.

network contains the whole network as seen in Figure 2.4 (C). This is an example for a balanced organization composed of two elementary modes (Section 2.3.2, Statement 2). The smallest organization containing the first elementary mode is the same as for the second; it is the only organization containing the whole network.

The second simple metabolic network shown in Figure 2.4 (B) features a loop consisting of metabolites B, C, and D. Metabolite A is created from an external substrate and reacts with D to form B, which is transformed to C. Finally, C is transformed into D and E. Both metabolites D and E are transformed into external product metabolites. Only one elementary mode exists in this network. It contains all metabolites and all reactions except the transformation of D into an external metabolite. Although the set of all metabolites is closed, it is not an organization. The set is not self-maintaining since within the loop, D is transformed into an external metabolite, leaving the network (Section 2.3.2, Statement 5). Metabolite D is required to keep the loop running, but no reaction compensates for the outflow of D. We find that the set only containing A is the only organization in this network as seen in Figure 2.4 (D). Since A is accumulating in the organizational reaction network (just consisting of A and its creation reaction), it is not a balanced organization consisting of elementary modes (Section 2.3.2, Statement 4). Here, we find that there is no organization containing the elementary mode. Consequently, this mode cannot be present in a steady state of the network (Section 2.3.2, Statement 3), unless the decay of species D can be switched off.

2.3.4 Pathways with Catalysts

A more complex metabolic reaction network is shown in Figure 2.5 (A). An external substrate is transformed into metabolite A. With metabolite E acting as a catalyzing enzyme, A can react to form B. Then, B can be transformed into E via two reactions. One is catalyzed by metabolite C, while the other by metabolite D. The metabolites A, B, and E are transformed into external product metabolites. Note that in general, each reaction in a metabolic network is implicitly

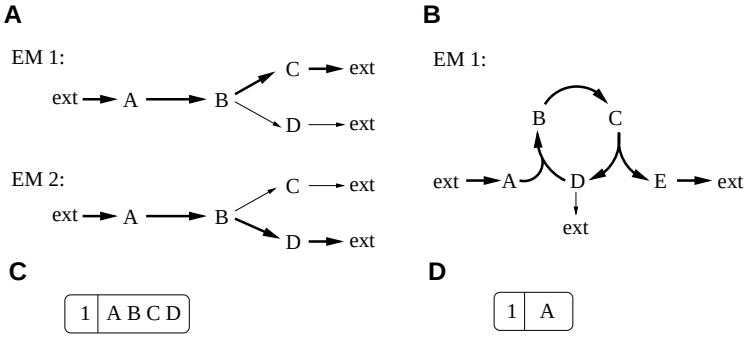


Figure 2.4: Elementary modes and organizations in simple linear branching and looping networks. External substrate and product metabolites are denominated with "ext". (**A**) The linear branching pathway contains two elementary modes. (**C**) The only organization consists of the whole pathway. (**B**) The network contains only one elementary mode consisting of all metabolites and all reactions except the outflow of D. (**D**) The only organization of this network solely contains A.

catalyzed by an enzyme. In this example, three metabolites are explicitly modeled as catalysts. The network contains four elementary modes as depicted in Figure 2.5 (A). The first mode just uses metabolite A to transform the external substrate metabolite into an external product metabolite. In the second mode, A is transformed into B with the help of enzyme metabolite E, and B is transformed into the external product. The third mode also transforms A to B using E as a catalyst. But additionally, C acts as a second catalyst to transform B into E. Finally, E is transformed into an external product. The fourth mode is similar to the third one with the exception that here, the reaction catalyzed by D is used to transform B into E.

The hierarchy of organizations is shown in Figure 2.5 (B). The network contains seven organizations. The smallest one just contains metabolite A. This organization coincides with the first elementary mode, it is a balanced organization. The three organizations above the first organization (2, 3, and 4) all have in common that they contain species that do not participate in any reaction of the organization. As mentioned in Section 2.3.2, we are not concerned with such organizations here. Organizations 5 and 6 contain exactly the metabolites

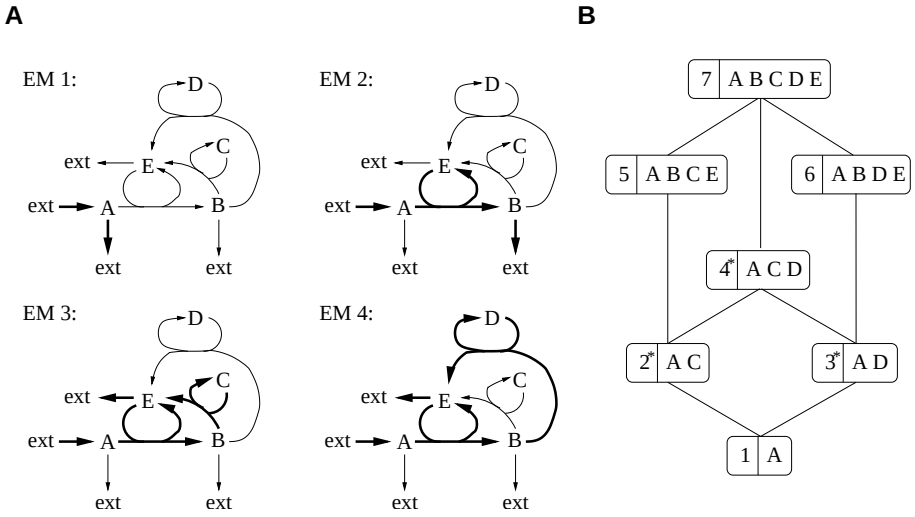


Figure 2.5: Comparing elementary modes with organizations in a more complex reaction network with five species. (**A**) The network contains four elementary modes. (**B**) The hierarchy of organizations consists of seven organizations. The starred organizations (Organization 2, 3, and 4) contain isolated species that do not react with each other.

making up elementary modes 3 and 4. Both are balanced organizations. Here, the sets of metabolites making up the elementary modes are already closed, and hence the smallest organizations containing the modes are already the very same sets. The smallest organization containing elementary mode 2 is not unique in this example: both Organizations 5 and 6 contain the mode and are of equal size (Section 2.3.2, Statement 3). Such an elementary mode can exist in different steady state network configurations and hence might be of particular importance. The largest Organization 7 comprises the whole network. It is a balanced organization combining elementary modes 3 and 4. Table 2.1 summarizes the relationship between elementary modes, the smallest organizations containing them, and all modes contained in organizations.

2.3.5 Central Sugar Metabolism of E. coli

As a more realistic example we finally analyze a reaction network introduced by Puchalka and Kierzek (2004) modeling the central sugar metabolism of *E. coli*

Table 2.1: Organizations and corresponding elementary modes for which the organization is the smallest enclosing one (Column 2), and all elementary modes contained in the organization (Column 3) for the example network with catalysts.

Organization	**EMs implying Org.**	**Contained EMs**
7	-	EM 1, EM 2, EM 3, EM 4
6	EM 2, EM 4	EM 1, EM 2, EM 4
5	EM 2, EM 3	EM 1, EM 2, EM 3
4	-	EM 1
3	-	EM 1
2	-	EM 1
1	EM 1	EM 1

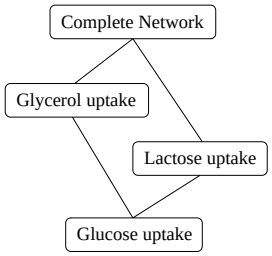


Figure 2.6: Hierarchy of organizations for the network modeling the central sugar metabolism of *E. coli*. The lower three organizations are associated with the uptake of glucose, lactose, and glycerol, respectively. The largest organization contains the whole network.

including gene expression, signal transduction, transport processes, and enzymatic activities. This model and its organizational structure in different growth media will be discussed in detail in Chapter 7. Here, we just consider the scenario in which all carbon sources (glucose, lactose, and glycerol) are present in the growth medium.

Modelling inducers and activators as required for gene transcription, the model network contains four organizations. Figure 2.6 depicts the hierarchy of organizations. The whole network constitutes an organization, and the three remaining organizations are associated with the uptake of glucose, lactose, and glycerol, respectively.

Table 2.2: Organizations and corresponding numbers of elementary modes for which the organization is the smallest enclosing one (Column 2), and number of total elementary modes contained in the organization (Column 3) for the network modelling the central sugar metabolism of *E. coli*.

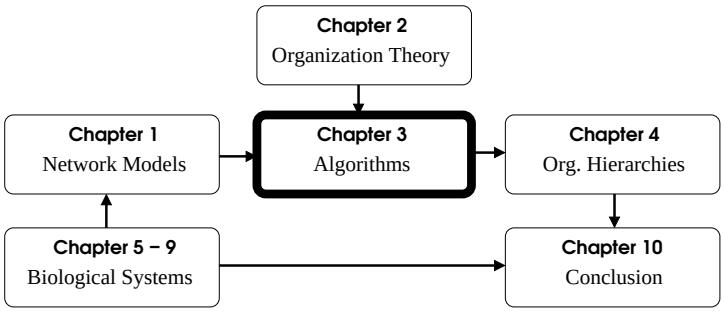
Organization	**EMs implying Org.**	**Contained EMs**
Complete Network	16	550
Glycerol uptake	37	209
Lactose uptake	325	497
Glucose uptake	172	172

Computing the elementary modes of the network revealed 550 modes. Determining the smallest organization containing each mode, they can be assigned to the organizations as shown in Table 2.2. Grouping elementary modes according to their enclosing organization helps to deal with the vast number of elementary modes usually found in large networks. In this example, the organizations can give a first hint on the physiological function an elementary mode plays a role in. While organizations specify species compositions required for physiological steady states (or states with increasing species concentrations), the elementary modes within organizations define the admissible flux distributions for the corresponding state.

2.3.6 Discussion

Combining elementary mode analysis with organization analysis gives a more complete picture of the potential dynamical behavior of metabolic networks. On the one hand, elementary modes represent pathways that can operate at steady state. Since the metabolite set associated with a mode needs not to be closed, single elementary modes are not expected to be observed in feasible system states. Organizations, on the other hand, specify metabolite combinations that are likely to be observed in feasible system states, taking a more global perspective on the system. Such a state can be a steady state or a state in which species have positive production rates. With elementary modes defining the boundaries of the potential steady state behavior of the metabolic network and balanced organizations rep-

resenting all metabolite compositions that allow for steady states, both concepts complement each other. Organizations help to identify all potential steady state metabolite combinations and then elementary modes help to define all admissible steady state flux distributions within the organizational network. And taking the opposite direction, classifying elementary modes according to their enclosing organization helps to deal with the typical vast number of elementary modes. With organizations also allowing for positive metabolite productions, bacterial growth can additionally be considered.



Chapter 3

Computing Organizations

Given a reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$, what are its organizations? This chapter introduces an algorithm to address this question. The problem is solved in two steps. First, all semi-organizations are computed for the reaction network. This can be done by just considering the network structure as defined by the reaction rules. As all organizations are also semi-organizations, the first step of the computation delivers all candidate species sets for organizations. In the second step, all these candidates are tested for the property of self-maintenance. It must be shown that a flux vector exists fulfilling the self-maintenance condition. This is equivalent to solving a linear programming problem. All semi-organizations fulfilling the self-maintenance condition, as discovered in the second step of the computation, form the set of organizations for the network.

To compute the organizations for a given reaction network, one could simply test all possible species combinations for the properties of closure and self-maintenance in a brute force fashion. However, such an approach is only feasible for networks with few species (*i.e.*, less than 30 species) as the number of sets to test equals 2^n, with n being the number of network species. Here, a more elaborate algorithm to compute organizations is proposed. First, the set of all semi-organizations is computed. This is done in a recursive manner: given an already determined semi-organization *so*, the semi-organizations above *so* are computed in the next step. To find a larger semi-organization that contains *so*, the network structure is taken into account to select species that, when added to *so*, are likely to give rise to a larger semi-organization. In this constructive

fashion, the hierarchy of semi-organizations is computed from bottom up. Then, in the second step, all semi-organizations have to be identified that are also organizations. The property of self-maintenance is the property distinguishing both organization types. All semi-organizations, for which it can be shown that a flux vector in accordance with the self-maintenance condition (see Section 2.1.5) exists, are organizations. For this, a linear programming problem is solved for each semi-organization.

3.1 Step 1: Computing Semi-Organizations

The algorithm starts with the smallest semi-organization and creates the whole hierarchy of semi-organizations from the bottom up. The function `closure(set)` computes the smallest closed set that contains the species set set. This is done by iteratively adding all species to set that, according to the reaction rules, can be produced from the species in set. Each addition might enable new reactions. The iteration stops as soon as no new novel species can be created. Taking the closure of the empty set delivers the smallest semi-organization of the network. If the network contains no input species, it is the empty set. If input species are present, the smallest semi-organization contains all input species and their closure.

The Main Loop. The main loop is shown as pseudo code in Figure 3.1. The central function is `SOsDirectlyAbove(so)`. It delivers the set of all semi-organizations that are directly above semi-organization so, that means all semi-organizations that contain so and that do not contain any other semi-organization that contains so[1]. In the Hasse diagram, these are exactly those semi-organizations to which so has uplinks.

[1] More precisely, *at least* those semi-organizations are delivered. Under certain circumstances, also semi-organizations are computed that are not directly above semi-organization so but contain another semi-organization above so. Consider for example the reaction network $\langle\{a, b\}, \{a \to a + b\}\rangle$. The system contains three organizations: $\{a, b\}$ above $\{b\}$ above $\{\}$. Applied to the empty set, the function `SOsDirectlyAbove(∅)` returns both $\{b\}$ and $\{a, b\}$ here, since function `SOsDirectlyAboveContaining()` first creates the closure of its argument (see below).

Function `computeSemiOrganizations`

Input: reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$
Output: set of all semi-organizations
result $\leftarrow \emptyset$; processedSets $\leftarrow \emptyset$;
SOsToCheck $\leftarrow$ { `closure( ∅ )` };
while *SOsToCheck* $\neq \emptyset$ **do**
  current $\leftarrow$ `getSmallestSO(`*SOsToCheck* `)`;
  processedSets $\leftarrow$ processedSets
  $\setminus$ { $p \in$ processedSets with $|p| \leq |$current$|$} ;
  SOsToCheck $\leftarrow$ SOsToCheck $\cup$ `SOsDirectlyAbove(`*current* `)`;
  SOsToCheck $\leftarrow$ SOsToCheck $\setminus$ { current };
  result $\leftarrow$ result $\cup$ { current };
end
`return(`*result* `)`;

Figure 3.1: Computing all semi-organizations for a given reaction network; the main loop.

The set SOsToCheck contains all semi-organizations that have already been found and that still have to be processed. It is initialized with the smallest semi-organization computed by `closure(∅)`. In each iteration step, the smallest semi-organization[1] current is taken from SOsToCheck. All semi-organizations returned by `SOsDirectlyAbove(`*current*`)` are then added to SOsToCheck, and current is removed from the set. The iteration stops when no semi-organization is left in SOsToCheck. In order to avoid processing the same set of species twice, the global variable processedSets keeps track of the processed sets (*e.g.,* in a hash structure). In order to save memory, entries pointing to sets of sizes up to the size of the just processed semi-organization can be removed.

Functions for Computing Semi-Organizations. The function `SOsDirectlyAbove(`*so*`)` computes all semi-organizations that are directly above semi-organization so. All

[1] If the smallest semi-organization is not unique, the choice is random.

Function `SOsDirectlyAbove`

Input: semi-organization so, reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$
Output: set of all semi-organizations directly above so
result $\leftarrow \emptyset$;
usableSpecies $\leftarrow \mathcal{M} \setminus$ so ;
foreach *s* $\in$ *usableSpecies* **do**
    result $\leftarrow$ result $\cup$ `SOsDirectlyAboveContaining(`*so*, { *s* } `)`;
end
`return(`*result* `)`;

Figure 3.2: Computing all semi-organizations for a given reaction network. Finding all semi-organizations that are directly above the semi-organization *so*.

such semi-organizations contain so and additional species. For all species not in so it is tested, whether a semi-organization above so containing that specific species exists. The pseudo code of this function is detailed in Figure 3.2.

The main work is done in the function `SOsDirectlyAboveContaining(`*so*`,` *speciesSet*`)`. All semi-organizations that are directly above so and contain the species in speciesSet are returned by this function. Figure 3.3 contains the pseudo code. First, the closure of the union of so and speciesSet is computed. If it is identical to a previously computed closure, the function simply returns in order to avoid duplicated computations. If the computed closure is semi-self-maintaining, a semi-organization with the desired properties is found and the function returns. If not, those species in the closure are identified that are consumed but not produced. In order to become a semi-organization, these species must be produced somehow. The function `producerSets(`*speciesSet*`)` returns all possible species combinations that can produce all species in speciesSet. For each such combination, `SOsDirectlyAboveContaining()` is recursively called again. This time, the producer combination is additionally required to be present in the new semi-organizations.

Function `SOsDirectlyAboveContaining`

Input: semi-organization so, species set species to be contained in new semi-organizations, reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$, set of already considered sets processedSets

Output: set of all semi-organizations directly above so that contain species

```
result ← ∅ ;
closure ← closure(so ∪ species );
if closure ∈ processedSets then
    return(∅);
else
    processedSets ← processedSets ∪ { closure };
end
if closure is semi-self-maintaining then
    result ← { closure } ;
else
    speciesToProduce ← {s ∈ closure | s is consumed but not produced in
    closure } ;
    producingSets ← setsOfProducers(speciesToProduce );
    foreach set ∈ producingSets do
        result ← result ∪ SOsDirectlyAboveContaining(so, species ∪ set
        );
    end
end
return(result );
```

Figure 3.3: Computing all semi-organizations for a given reaction network. Finding all semi-organizations that are directly above the semi-organization *so* and contain species set species.

Function to Compute Producer Species. The function `setsOfProducers(speciesSet)` computes all species combinations that produce the species contained in speciesSet. Its pseudo code is detailed in Figure 3.4. In the first step, for each species s

Function `setsOfProducers`

Input: set of species to produce speciesSet, reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$
Output: set of all species sets that can produce all species in speciesSet
result $\leftarrow \emptyset$;
foreach *s* $\in$ *speciesSet* **do**
  productionSets$_s$ $\leftarrow \emptyset$;
  foreach *reaction* $\in \mathcal{R}$ **do**
    if *s has positive stoichiometric coefficient in reaction* **then**
      productionSets$_s$ $\leftarrow$ productionSets$_s$ $\cup$ { `educts(`*reaction* `)` } ;
    end
  end
end
repeat
  current $\leftarrow \emptyset$;
  foreach *s* $\in$ *speciesSet* **do**
    select a set setProducingS from productionSets$_s$;
    current $\leftarrow$ current $\cup$ setProducingS ;
  end
  result $\leftarrow$ result $\cup$ { current } ;
until *all possible set combinations have been considered*;
`return(`*result* `)`;

Figure 3.4: Computing all semi-organizations for a given reaction network. Finding all species combinations that produce all species in speciesSet.

in speciesSet, a set of species sets productionSets$_s$ is generated. This set contains all species combinations that can produce s. The sets are computed by inspecting all reactions. For each reaction in which s is produced (having a positive stoichiometric coefficient), the set of the reaction educts forms a producing set. In the second step, all possible combinations of species sets from the sets productionSets are generated. Each combination contains exactly one set of productionSets$_s$ for each species s in speciesSet.

3.2 Step 2: Test for Self-Maintenance

Every semi-organization determined in the first step of the algorithm is a candidate for an organization. For a semi-organization to also be an organization, it must be shown that a flux vector exists in accordance with the self-maintenance condition. Let O be a semi-organization with $n = |O|$ species implying m reactions. With $\mathbf{S}_O$ being the stoichiometric matrix for this subnetwork, we must show that a flux vector $\mathbf{v} \in \mathbb{R}^m$ exists with $\mathbf{v} > \mathbf{0}$ and $\mathbf{S}_O\mathbf{v} \geq \mathbf{0}$ (*cf.* Section 2.1.5). With the solution space of these inequalities forming a convex polyhedral cone in the positive orthant, originating in the point of origin, the problem is equivalent to finding a flux vector $\mathbf{v}$ with $\mathbf{v} > \mathbf{1}$ and $\mathbf{S}_O\mathbf{v} \geq \mathbf{0}$. This is a linear programming problem that can be solved using the simplex method (Dantzig, 1963). Since only the existence of such a flux vector $\mathbf{v}$ is of concern, only the first phase of the simplex method needs to be performed. In this phase, an initial feasible solution is determined. The original problem is transformed into restricted normal form by introducing slack variables and artificial variables. Then, a new linear programming problem is formulated. For this problem, an initial feasible solution can be directly determined. The second phase of the simplex method then optimizes the solution by searching along the edges of the solution space. If the optimal solution of the newly formulated linear programming problem fulfills certain criteria, an initial feasible solution is found for the original problem. In this case, the semi-organization O is also an organization.

3.3 Connected Organizations

In order to compute connected semi-organizations, only the function `SOsDirectlyAbove()` needs to be modified. Now, only those species are added to an already discovered semi-organization that are directly connected to a member species of that semi-organization. Figure 3.5 contains the pseudo code.

Function `ConnectedSOsDirectlyAbove`

Input: semi-organization so, reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$
Output: set of all connected semi-organizations directly above so
result $\leftarrow \emptyset$; usableSpeciesSets $\leftarrow \emptyset$;
if *so* $= \emptyset$ **then**
    usableSpeciesSets $\leftarrow \cup_{s \in \mathcal{M}} \{\{s\}\}$;
else
    foreach *reaction* $\in \mathcal{R}$ *with* ***educts(*** reaction ***)*** $\not\subseteq$ *so* **do**
        if $\exists$ *s* $\in$ *so with s* $\in$ ***educts(*** reaction ***)*** $\cup$ ***products(*** reaction ***)***
        then
            usableSpeciesSets $\leftarrow$ usableSpeciesSets $\cup$ { `educts(`*reaction* `)` $\setminus$
            so } ;
        end
    end
end
foreach *set* $\in$ *usableSpeciesSets* **do**
    result $\leftarrow$ result $\cup$ `SOsDirectlyAboveContaining(`*so*, *set* `)`;
end
`return(`*result* `)`;

Figure 3.5: Computing all connected semi-organizations for a given reaction network. Finding all connected semi-organizations that are directly above the connected semi-organization *so*.

3.4 Runtime Complexity

To study the runtime complexity of the presented algorithm we first consider the computation of semi-organizations. The function `SOsDirectlyAboveContaining()` is the innermost function which is called recursively. Hence, the number of its invocations serves as a measure for the runtime. The crucial loop in `SOsDirectlyAbove()` runs over the network species. The number of species n in the network is hence used to characterize the size of the input. It is difficult to define how a "typical" reaction network looks like. However, given the number of species $n = |\mathcal{M}|$ of a network we can construct a worst case

scenario reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$ to compute the worst case runtime. In order to maximize the calls to `SOsDirectlyAboveContaining()`, all species sets should be closed. A closure of a given set being much larger than the set itself would speed up the computation. Furthermore, to maximize the recursive calls to `SOsDirectlyAboveContaining()`, as many species sets as possible should be producer sets for a given species. All species decay spontaneously to prevent one species semi-organizations. The desired network can be described as follows:

1. $\forall s \in \mathcal{M} : (\{s\} \rightarrow \emptyset) \in \mathcal{R}$
2. $\forall s \in \mathcal{M}$ and $\forall S \in \mathcal{P}(\mathcal{M})$ with $s \in S$ and $|S| > 1 : (S \rightarrow \{s, s\}) \in \mathcal{R}$.

The empty set and all species sets containing at least two species are semi-organizations in this network. The only organization is the empty set.

To evaluate the number of invocations of `SOsDirectlyAboveContaining()`, we start in the main loop. The empty set is processed first. In `SOsDirectlyAbove(∅)`, `SOsDirectlyAboveContaining()` is called for each species of the network, leading to n invocations. Once in `SOsDirectlyAboveContaining()`, it is found that a species set containing a single species cannot be a semi-organization due to the decay reactions. Species are added that are able to produce the species in question. The network has been designed such that any species set with two or more species containing species s is a producer for s. There are 2^{n-1} sets that contain s. Since the set just containing s is not a producing set, we arrive at $(2^{n-1}) - 1$ producing sets. Hence, the function is called $n \cdot (2^{n-1} - 1)$ times recursively. In each recursive call, it will directly return the passed species sets as a semi-organization.

After processing the empty set in the main loop, all possible species sets with at least two species are stored in SOsToCheck. Each invocation of `SOsDirectlyAbove()` will lead to further invocations of `SOsDirectlyAboveContaining()`. If the processed semi-organization has i species, the function is called $n - i$ times. This is the number of species that can be added to the processed semi-organization. However, at this stage no further recursive invocations occur, as all passed sets are already semi-organizations. We arrive at $\sum_{i=2}^{n} \binom{n}{i} \cdot (n - i)$

further function invocations. Summing up, the total number of invocations of `SOsDirectlyAboveContaining()` computes to

$$
\begin{aligned}
f(n) = & \quad n + n(2^{n-1} - 1) + \sum_{i=2}^{n} \binom{n}{i} \cdot (n - i) && (3.1)\\
= & \quad n2^{n-1} + \sum_{i=2}^{n} \binom{n}{i} \cdot (n - i) && (3.2)\\
\leq & \quad n2^{n-1} + n \cdot \sum_{i=2}^{n} \binom{n}{i} && (3.3)\\
= & \quad n2^{n-1} + n \cdot \left(2^n - \binom{n}{0} - \binom{n}{1}\right) && (3.4)\\
= & \quad n2^{n-1} + n \cdot (2^n - 1 - n) && (3.5)\\
\leq & \quad n2^{n-1} + n2^n && (3.6)\\
\leq & \quad 2n2^n && (3.7)\\
\leq & \quad n2^n. && (3.8)
\end{aligned}
$$

This gives an exponential runtime with $O(n2^n)$. This is not surprising as the number of semi-organizations also increases exponentially with network size. The empty set and all species sets containing at least two species are semi-organizations.

However if the result is constant, the worst case runtime is still exponential as will be shown now. We construct a network that contains the empty set as the only semi-organization, independent of the network size. The network contains n numbered species $s_1, \cdots, s_n$. All species decay spontaneously. For a given species s_j, all species combinations from the set $\{s_{j+1}, \cdots, s_n\}$ can produce s_j. Species s_n is the only species for which no production reaction exists in the network. It is added as a reactant to every reaction in which it not yet appears as a reactant. This limits the recursion in the algorithm and makes the computation of function invocations easier. The network can be formalized as:

1. $\mathcal{M} = \{s_1, \cdots, s_n\}$
2. $\forall j = 1, \cdots, n : (\{s_j\} \rightarrow \emptyset) \in \mathcal{R}$

3. $\forall j = 1, \cdots, n-1 : \forall P \in \mathcal{P}(\{s_{j+1}, \cdots, s_n\})$ with $P \neq \emptyset : (P \cup \{s_n\} \rightarrow s_j) \in \mathcal{R}$

Starting in the main loop, the empty set is processed first. In `SOsDirectlyAbove(∅)`, `SOsDirectlyAboveContaining()` is called for each species s_j of the network, leading to n invocations. In function `SOsDirectlyAboveContaining(`s_j`)`, it is discovered that single species are not semi-organizations due to the spontaneous decay. Species sets must be added that produce s_j. The network was constructed such that all subsets of species set $\{s_{j+1}, \cdots, s_n\}$ except the empty set produce s_j. The set contains $n - j$ species. However, as s_n is a reactant in every production reaction and therewith contained in all producing sets, only 2^{n-j-1} different producing sets exist for species s_j. Accordingly, `SOsDirectlyAboveContaining()` is recursively called 2^{n-j-1} times for all species s_j except for s_n, for which no further calls occur as no production reaction for s_n exists in the network. In the second invocation of the function, s_n is part of the considered species set, which is again found not to be a semi-organization. As s_n cannot be produced in the network, no viable producing sets are found and no further recursive function calls are made. Summing up, the total number of invocations of `SOsDirectlyAboveContaining()` computes to

$$
\begin{aligned}
f(n) &= n + \sum_{j=1}^{n-1} 2^{n-j-1} && (3.9)\\
&= n + \sum_{j=2}^{n} 2^{n-j} && (3.10)\\
&= n + \sum_{j=0}^{n-2} 2^{j} && (3.11)\\
&= n + \frac{2^{n-1} - 1}{2 - 1} && (3.12)\\
&= n + 2^{n-1} - 1 && (3.13)\\
&\leq n + 2^{n-1} && (3.14)\\
&\leq n + 2^{n}. && (3.15)
\end{aligned}
$$

Although slightly better than in the former case, this still gives an exponential runtime with $O(n + 2^n)$.

The best case network contains input reactions for all n species. In this case the smallest semi-organization is already the whole network. Function `SOsDirectlyAbove()` returns directly and `SOsDirectlyAboveContaining()` is never called, leading to constant runtime $O(1)$.

To compute organizations, a linear programming problem must be solved for each semi-organization using the simplex algorithm. Although quite fast for typical problems, the worst case runtime is exponential for this algorithm (Klee and Minty, 1972). However, we only use the first phase of the algorithm. If we take the first worst case network and remove the decay reactions, all 2^n sets of species are semi-organizations. Hence, the overall worst case complexity to compute the organizations of a reaction network using the presented algorithm is $O(n2^n + 2^n \cdot 2^n) = O(n2^n + 2^{2n})$. However, for realistic networks the runtime of the constructive algorithm is typically much shorter. It not only depends on the number of species of the network, but also on its specific structure. Usually, the algorithm is much faster than a brute force approach with a guaranteed runtime of $O(2^n \cdot 2^n) = O(2^{2n})$.

3.5 Implementation

The presented algorithm was implemented in C++ and Java. For solving the linear programming problem, the `lp_solve` package (Berkelaar et al.) is used. Reaction networks specified in the Systems Biology Markup Language (Finney and Hucka, 2003) can be processed by the tool. The Java version plugs into the Systems Biology Workbench (Sauro et al., 2003), an open source framework connecting different software tools and application for creating, simulating, and analyzing biological models.

3.6 Runtime

We compare the runtime, measured in the number of invocations of function `SOsDirectlyAboveContaining()` and in elapsed real time, for several network models that are analyzed in this thesis. Table 3.1 relates the size of the networks to the required runtime to compute the organizations and the connected

organizations. It must be noted that the real time runtime also includes the time for computing the hierarchy relations between organizations, required for the visualization in the Hasse diagram.

Computations were performed on an Intel Pentium 4 processor at 1.80 GHz with 1 GB RAM, running Linux.

Contrary to the assumption that larger networks with more species and reactions lead to longer runtimes, we find that the most critical property determining the runtime is the network structure. Taking the network of Mars at night with 31 species and 103 reactions, we notice that adding one single reaction brings down the runtime from more than 9 days to less than a minute. The reaction that is added to the night network is an input reaction that creates a species from nothing. Species that are supplied as input are always part of the smallest organization, and therewith of all organizations. If the smallest organization contains already many species, there are fewer possibilities to expand this organization using the remaining species. Hence, a high number of input reactions results in faster computation.

Computing the hierarchy of connected organizations is usually much faster than computing the hierarchy of organizations. For connected organizations, only connected species are considered as candidate species for organization expansions, resulting in fewer expansion possibilities.

That network size alone is not necessarily determining the runtime can be illuminated by the fact that for any network size, a best case and a worst case network can be constructed as detailed in Section 3.4. It is an open problem how to parameterize the network structure in order to estimate the runtime for a given network. For certain networks, the runtime to compute organizations using the presented algorithm exceeds practical limits, even for connected organizations (*cf.* Chapters 8 and 9). Kaleta (2005) has developed an algorithm to compute organizations that is based on extreme pathways. For some networks, this approach is faster than the constructive approach, while it is slower for others (unpublished data). For networks, for which both algorithms fail, a random based heuristic approach can be used to compute at least a subset of the organizations of the network. A species set is picked at random, and the closure of this set is checked for the property of self-maintenance.

Table 3.1: Comparing network size, number of organizations, and the runtime to compute organizations (top) and connected organizations (bottom) using the constructive algorithm. "Invocations" refers to calls to function `SOsDirectlyAboveContaining()`. Runtime is the sum of user and system time as reported by `memtime`, rounded up to seconds. Runtime also includes the computation of hiearchy relations between organizations required for representing the organization hierarchy in a Hasse diagram.

Organizations

Network	**Spec./React.**	**Semi-org./ Org.**	**Invocations**	**Runtime**
Dry Mars, day	7/ 16	6/ 6	24	$1s$
Dry Mars, night	7/ 15	22/ 22	144	$1s$
Mars, day	31/104	1.496/ 1.484	37.918	$50s$
Mars, night	31/103	1.089.330/1.088.640	27.381.810	$9,2d$
Lambda	55/ 81	7/ 7	501	$1s$
Central *E. coli*	92/168	30/ 4	839	$2s$

Connected Organizations

Network	**Spec./React.**	**Semi-org./ Org.**	**Invocations**	**Runtime**
Dry Mars, day	7/ 16	3/ 3	19	$1s$
Dry Mars, night	7/ 15	7/ 7	76	$1s$
Mars, day	31/104	14/ 8	549	$1s$
Mars, night	31/103	41/ 27	1.985	$2s$
Lambda	55/ 81	2/ 2	12	$1s$
Central *E. coli*	92/168	4/ 1	76	$1s$

Runtime on Networks of Increasing Size

Network size alone is not sufficient to estimate the required runtime to compute organizations. To eliminate the effect of network structure, we consider here a sequence of growing reaction networks. In each enlargement step, species and reactions are added to the network. By this procedure, the structure of the network is preserved and the influence of network size on computation time can be studied in isolation.

To generate biological feasible networks of increasing size, we take a reaction network model of signaling events by Blinov et al. (2006), containing 356 species and 3749 reactions. To create the model, the authors used BioNetGen (Blinov et al., 2004), a tool that implements a domain- and rule-based approach for modeling signal transduction. In order to tackle the combinatorial complexity of signaling molecules, BioNetGen allows one to define molecules, propertiers, and interaction rules. These definitions are then transformed into reaction network models in an iterative fashion. In each iteration step, species and reactions are added to the previous model. We take the reaction network models of each iteration step and compute their organizations. The model by Blinov et al. (2006) is generated in 10 iteration steps. Figure 3.6 shows the sizes of the resulting 10 networks that will be analyzed.

Organizations could only be computed in feasible time for the first four networks and the last network. Figure 3.7 allows one to compare the number of (semi-) organizations and the runtime for the five networks. The number of (semi-) organizations and the runtime in invocations of `SOsDirectlyAboveContaining()` and in real time increases for the first four networks. The largest network contains fewer (semi-) organizations. Although the runtime measured in function invocation also decreases, the real time runtime increases slightly. The algorithm spends more time here for the tasks maintenance (*e.g.*, keeping track of processed sets), function `setsOfProducers()`, solving the linear programming problems, and computing the hierarchy relations between organizations.

The connected organizations could be computed for all 10 reaction networks. Figure 3.8 shows the results. The number of (semi-) organizations peaks at the eighth network. Smaller and larger networks have fewer organizations. The

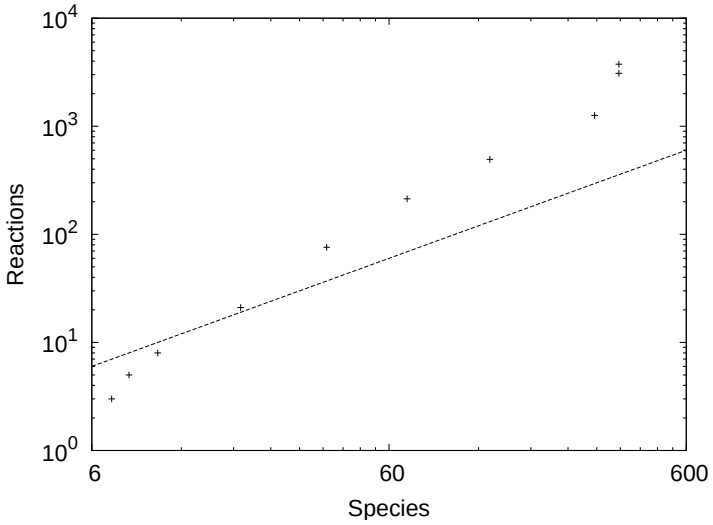


Figure 3.6: Sizes of the 10 networks iteratively generated by BioNetGen. In each iteration step, the number of species and reactions increases, except for the last iteration. Only the number of reactions increases in the last iteration. The final network contains 356 species and 3749 reactions and models signaling events (Blinov et al., 2006). Identity as dashed line for reference. Both axis in logarithmic scaling.

runtime increases steadily with network size. While the number of invocations increase almost linear, the runtime increases almost exponential. Again, time spent in maintenance and other tasks is responsible for this discrepancy.

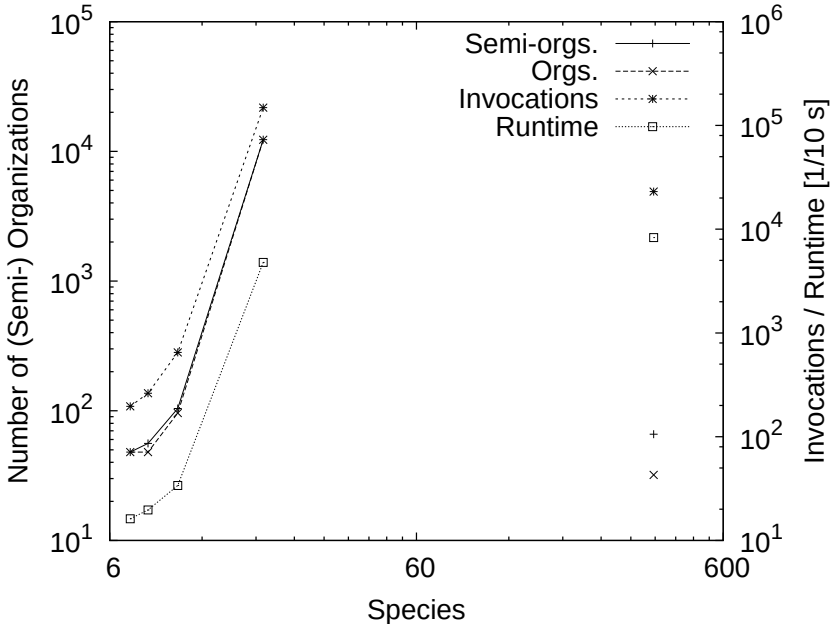


Figure 3.7: Number of semi-organizations and organizations (left ordinate), and required runtime in invocations of `SOsDirectlyAboveContaining()` and real time including computation of the Hasse diagram (right ordinate) for the five networks, for which the computation was feasible. Lines connecting data points do not imply a linear relationship, they appear only for better readability. All axis in logarithmic scaling.

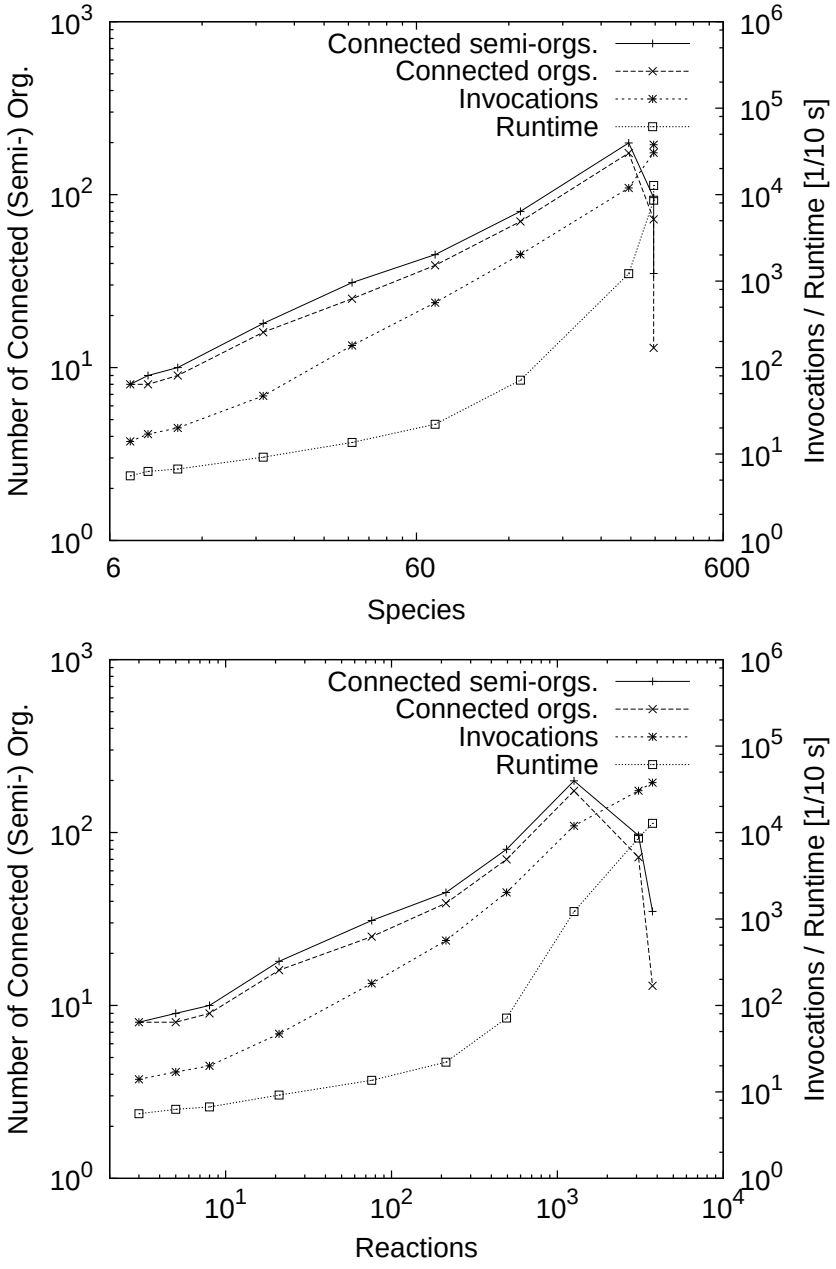
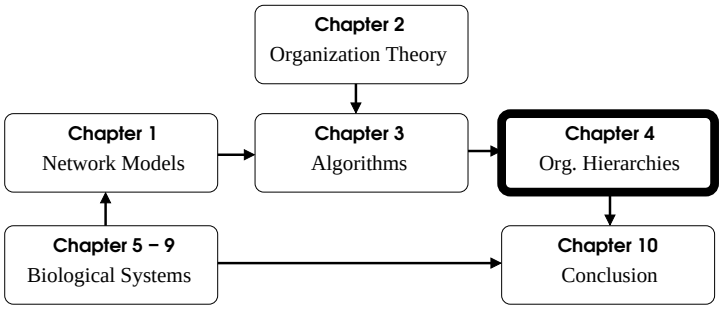


Figure 3.8: Number of connected semi-organizations and connected organizations (left ordinate), and required runtime in invocations of `SOsDirectlyAboveContaining()` and real time, including computation of the Hasse diagram, (right ordinate) for the ten networks generated by BioNetGen. Network size in species (top) and reactions (bottom). Lines connecting data points do not imply a linear relationship, they appear only for better readability. All axis in logarithmic scaling.



Chapter 4

Analyzing Hierarchies of Organizations

The theory of chemical organizations delivers the set of organizations for a given reaction network. As organizations can overlap, the whole set of organizations can be visualized as a hierarchical structure in a Hasse diagram. In this chapter, this interleaved structure will be used to define several species sets that have certain properties with respect to the organizational hierarchy. These sets lead to further insights regarding the modular structure of the network.

To bridge the gap between qualitative systems analysis using chemical organizations, and quantitative approaches, intensity values based on concentration vectors will be assigned to organizations.

But first, we start with the building blocks of organizational hierarchies: the chemical organizations themselves. A simple example demonstrates what we mean when stating that it is only in principle that organizations have the ability to prevail in time (see Section 2.1.6).

4.1 Organizations Prevailing in Time

Organizations represent all potential species combinations that allow for steady states and states featuring growth. They are potential in the sense that a steady state (or a growth state) can be (at least) realized if concentrations and reaction rate laws can be arbitrarily chosen. Whether such a state is feasible in the

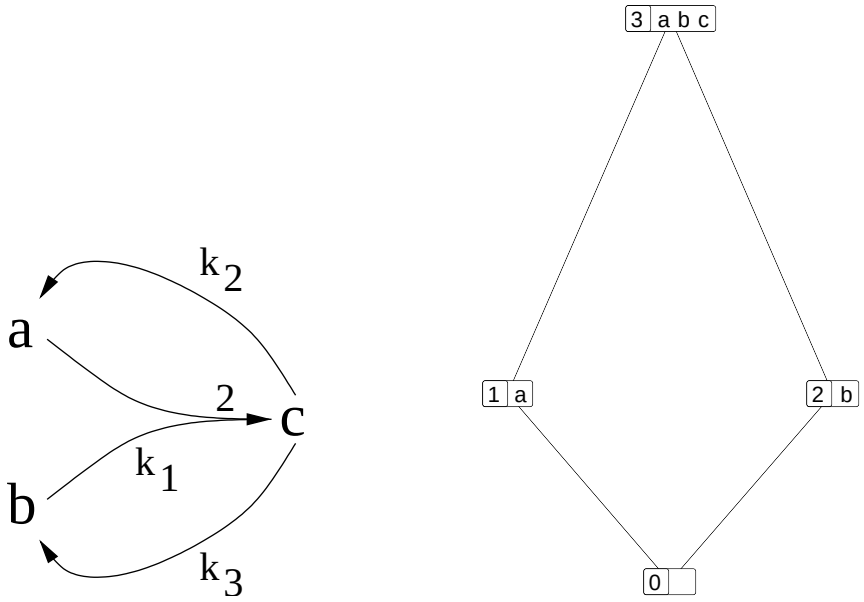


Figure 4.1: A simple reaction network (left) containing four organizations (right).

dynamic model and in the real system, or not depends on the kinetics of the network. Even for mass-conserving networks (see Chapter 5), a steady state might not be feasible. For example, consider the network depicted in Figure 4.1, left.

The chemical reactions are:

$$a + b \xrightarrow{k_1} 2\ c \tag{4.1}$$

$$c \xrightarrow{k_2} a \tag{4.2}$$

$$c \xrightarrow{k_3} b. \tag{4.3}$$

Since the reactions are irreversible, we demand the reaction rate constants k_1, k_2, and k_3 to be nonnegative.

Using mass action kinetics, the differential equations describing the dynamics of the system can be written as the product of its stoichiometric matrix and its flux vector:

$$\begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 2 & -1 & -1 \end{pmatrix} \begin{pmatrix} [a][b]k_1 \\ [c]k_2 \\ [c]k_3 \end{pmatrix} = \begin{pmatrix} d[a]/dt \\ d[b]/dt \\ d[c]/dt \end{pmatrix}. \tag{4.4}$$

The network contains four organizations as depicted in Figure 4.1, right. The empty Organization 0 and one species Organizations 1 and 2 contain no reactions. Hence, they can prevail in time independent of any kinetic data. Organization 3 contains the whole network with all three reactions. To find a flux vector in accordance with the self-maintenance property, we demand nonnegative production rates for all three species:

$$d[a]/dt = -[a][b]k_1 + [c]k_2 \qquad \geq 0 \tag{4.5}$$
$$d[b]/dt = -[a][b]k_1 + [c]k_3 \qquad \geq 0 \tag{4.6}$$
$$d[c]/dt = 2[a][b]k_1 - [c]k_2 - [c]k_3 \geq 0. \tag{4.7}$$

It follows:

$$[c]k_2 \geq [a][b]k_1 \tag{4.8}$$
$$[c]k_3 \geq [a][b]k_1 \tag{4.9}$$
$$2[a][b]k_1 \geq [c]k_2 + [c]k_3 \tag{4.10}$$

$$\Rightarrow 2[c]k_2 \geq 2[a][b]k_1 \geq [c]k_2 + [c]k_3 \tag{4.11}$$
$$2[c]k_3 \geq 2[a][b]k_1 \geq [c]k_2 + [c]k_3 \tag{4.12}$$

$$\Rightarrow 2[c]k_2 \geq [c]k_2 + [c]k_3 \tag{4.13}$$
$$2[c]k_3 \geq [c]k_2 + [c]k_3 \tag{4.14}$$

$$\Rightarrow [c]k_2 \geq [c]k_3 \tag{4.15}$$
$$[c]k_3 \geq [c]k_2 \tag{4.16}$$

$$\Rightarrow k_2 = k_3. \tag{4.17}$$

It follows for k_1:

$$d[a]/dt = d[b]/dt = -[a][b]k_1 + [c]k_2 \geq 0 \tag{4.18}$$
$$d[c]/dt = 2[a][b]k_1 - 2[c]k_2 \geq 0 \tag{4.19}$$

$$\Rightarrow [c]k_2 \geq [a][b]k_1 \tag{4.20}$$

$$[a][b]k_1 \geq [c]k_2 \tag{4.21}$$

$$\Rightarrow [a][b]k_1 = [c]k_2 \tag{4.22}$$

$$\Rightarrow k_1 = k_2 \frac{[c]}{[a][b]}. \tag{4.23}$$

We find that the condition of self-maintenance can only be fulfilled for $k_1 = k_2 \frac{[c]}{[a][b]}$ and $k_2 = k_3$. The production rates of all three species become zero for this choice. Hence, this organization can only feature a steady state but not a state related to growth. Whether this steady state is feasible in the simulation model, respectively the real system, or not depends on the values of the reaction rate constants (in general: on the kinetic laws). Only if they fulfill the stated conditions, a steady state is possible and the organization can prevail in time.

4.2 Species Groups

The set $\mathcal{O}$ of all organizations for a given reaction network $\langle \mathcal{M}, \mathcal{R} \rangle$ will be used to define several species groups with special properties as follows.

Reachable Species Set. The set of *reachable species* $\mathcal{M}_r \subseteq \mathcal{M}$ contains all species that appear at least in one organization:

$$\mathcal{M}_r := \bigcup_{O \in \mathcal{O}} O. \tag{4.24}$$

Unreachable Species Set. The set of *unreachable species* $\mathcal{M}_u \subseteq \mathcal{M}$ contains all species that are not part of any organization. For example, species that are reactants in unimolecular reactions and that do not have a production reaction belong to this set. They cannot be part of any organization. The set $\mathcal{M}_u$ contains all species that are never part of a closed and self-maintaining subnetwork. In this sense, they are unreachable. More formal:

$$\mathcal{M}_u := \mathcal{M} \setminus \bigcup_{O \in \mathcal{O}} O. \tag{4.25}$$

Accompanying Species Set. Given a species set $S \subseteq \mathcal{M}$, the set of *accompanying species* $\mathcal{A}(S)$ contains all species $a \in \mathcal{M}$ for which is true: if $S \subseteq O$ for any organization $O \in \mathcal{O}$, then also $a \in O$. In other words: if species set S is contained in an organization, it follows that also its accompanying species $\mathcal{A}(S)$ are members of this organization. More formal:

$$\mathcal{A}(S) := \bigcap_{O \in \mathcal{O} \text{ with } S \subseteq O} O. \tag{4.26}$$

The concept of accompanying species is equivalent to the closure of item sets in data-mining (*e.g.,* Han and Kamber, 2006). Accompanying species sets are closed item sets.

Unit Species Set. Given a species $s \in \mathcal{M}$, the set of *unit species* $\mathcal{U}(s)$ is the largest species set containing s for which is true: if and only if any species $u \in \mathcal{U}(s)$ is contained in an organization $O \in \mathcal{O}$, then the whole set of unit species is in the organization as well: $\mathcal{U}(s) \subseteq O$. If one species of $\mathcal{U}(s)$ is not contained in an organization, all other species of the set will not be part of the organization neither. How can $\mathcal{U}(s)$ be constructed? First, the presence of s must imply the presence of all species in $\mathcal{U}(s)$. Hence, the candidate species for the unit species set are $\mathcal{A}(\{s\})$. Second, a successful candidate species $a \in \mathcal{A}(\{s\})$ must imply the presence of all other members of $\mathcal{U}(s)$, in particular also the presence of s. Hence, it must hold: $s \in \mathcal{A}(\{a\})$. This leads to:

$$\mathcal{U}(s) := \bigcup_{a \in \mathcal{A}(\{s\}) \text{ with } s \in \mathcal{A}(\{a\})} \{a\}. \tag{4.27}$$

All unit species sets form a cover for the set of reachable species:

$$\mathcal{M}_r = \bigcup_{s \in \mathcal{M}} \mathcal{U}(s). \tag{4.28}$$

Unit Add-on Species Set. Given a unit species set $\mathcal{U}$, the set of *unit add-on species* $\mathcal{U}_a(\mathcal{U})$ contains all species that are additionally present in any organization in which $\mathcal{U}$ is contained. Whenever $\mathcal{U}$ is contained in an organization, $\mathcal{U}_a$ will also be part of the organization. More formal:

$$\mathcal{U}_a(\mathcal{U}) := \mathcal{A}(\mathcal{U}) \setminus \mathcal{U}. \tag{4.29}$$

Versatile Species Sets. A set of species $S \in \mathcal{M}$ is called *versatile*, if it holds true: $\mathcal{A}(S) = S$. In other words, the presence of S does not imply the presence of any other species. Hence, S can appear together with disjunct sets of species in organizations. In this sense, it is not bound to other species, but it is versatile.

A tool to compute the defined species sets has been implemented in C++. In the analysis of a photochemical reaction network model of the Martian atmosphere in Chapter 5, we will show that chemical species of similar molecular structure make up the unit species sets of that network.

4.3 Intensities of Organizations

Organizations represent qualitative states of the reaction system under study. If quantitative data is available, for example a measured or a simulated trajectory, the dynamic evolution of the system can be mapped to the space of organizations. As a first step towards such a mapping, Dittrich and Speroni di Fenizio (2007) introduced the concepts *abstraction* and *instance* to map the system state described by a quantitative concentration vector to a set of species, and back. Here, we focus on the question: given the quantitative system state, what are the most important organizations that dominate the system and determine the overall system behavior? Given the quantitative state $\mathbf{c}(t)$ of the system at time t, *intensity values* will be assigned to each organization of the system, indicating its importance at time t. First, a rather simple approach to assign intensity values to organizations will be introduced. Then, a more elaborate seed based method will be described that takes the ability of organizations into account to regenerate themselves.

4.3.1 Assigning Intensity Values to Organizations

By simulating a network model over time, a concentration trajectory for each chemical species of the network can be obtained. The vector holding the relative concentrations of all species $\mathbf{c}(t)$ then details the species composition within the reaction vessel at time t. It is now interesting to analyze which organizational subnetworks are present in the vessel at time t. The notion of "being present" can

be interpreted as a gradual quality. For example, one organization can be "more present" than another, because its species occur at higher concentrations than the species of the other organization. To measure the degree of "being present" in the reaction system, we assign intensity values to organizations. Generally, a function is needed for each organization O, that given a relative concentration vector $\mathbf{c} \in C$ computes the intensity value: $int_O : C \to [0, 1]$.

4.3.2 Concentration Based Approach

In a first step, we demand $int_O(\mathbf{c}) = 0$ in case there is a species $s \in O$ with zero concentration $c_s = 0$. In other words, for an organization to have an intensity value greater zero, all species of this organization must be present in the system. We use a rather simple function fulfilling this requirement by just summing up over the relative concentrations of the organization species for nonempty organizations:

$$int_O(\mathbf{c}(t)) := \begin{cases} 0 & \text{if } \exists\, i \in O \text{ with } c_i = 0, \\ \sum_{j \in O} c_j(t) & \text{else.} \end{cases} \tag{4.30}$$

Just taking into account the concentrations of organization species to measure the presence or intenseness of an organization is of course a crude approach. Only the total concentration of all organization species is considered, neglecting the internal distribution within the organization. For example, an organization will get the same intensity value no matter if the total concentration is equally distributed over all organization species or if it is concentrated on one species only, while the other species have low concentrations.

If the largest organization O contains all species of the network, int_O will obviously always equal one as long as all species are present within the reaction vessel (even if only at very low concentrations). If organization O_{above} contains organization O_{below}, and all species of O_{above} are present, it is always true: $int_{O_{above}} > int_{O_{below}}$, since O_{above} contains all species of O_{below} and at least one more species that is also present in the vessel. Due to this dependency on organization size, this method cannot be used to directly compare the intensity values of different organizations. Instead, the intensity time-course of an organization

can be analyzed and the dynamics of different organizations can be compared to each other. In order to make an easy comparison possible, we scale the intensity trajectories so that their maximum equals one:

$$int_O^S(\mathbf{c}(t)) := int_O(\mathbf{c}(t)) \,/\, \max_t(int_O(\mathbf{c}(t)). \tag{4.31}$$

Besides summation, other functions based on species concentrations were tested to compute intensity values for organizations. However, when applied to simple test cases (as will be used in Section 4.3.4), they led to inconsistent results (data not shown). When the multiplication of normalized concentration values is considered, large organizations with many species tend to get smaller intensity values. Taking the shannon entropy led to futile results when the reactor is populated by one species only, as the logarithm of 1 is zero. A function taking the distribution of species concentrations within the organization into account, akin to the variance formula, produced promising results on simple test cases but failed on slightly more complex cases.

4.3.3 Seed Based Approach

In certain reaction systems, species with small concentrations can play a crucial role for the whole system. For example, small amounts of HO_x radicals are responsible for the recycling of large amounts of CO_2 in the Martian atmosphere (see Chapter 5). Having small concentrations, these species will not be considered by the simple intensity approach in an appropriate way.

In order to remedy this problem, we introduce a second method to assign intensity values to organizations. Given the current state of a reaction reactor, we relate the intensity value to the effort that is necessary to remove an organization completely from the reactor. This gives a better indicator for the "presence" of an organization and relates to its stability or robustness. Removing one (or more) arbitrarily chosen organization species completely might not be enough to lead to the collapse of an organization, as other species might recreate the removed species. The smallest species sets that can (re-)create an organization are its

seeds[1]. Hence, we must remove all seeds of an organization to make it collapse. We have to consider all species sets $D \subseteq \mathcal{M}$ whose removal leads to the removal of all seeds. To remove a seed, it is sufficient to eliminate one of its species. For each such set D, the relative concentrations of its member species are summed up. This value indicates, how many molecules must be removed from the reactor (or how difficult it is) to remove the organization. The smallest such value will be used as the intensity value for the organization. Formally for organizations with nonempty seeds:

$$int'_O(\mathbf{c}(t)) := \min_D \left(\sum_{j \in D} c_j(t) \right), D \subseteq \mathcal{M} \text{ with } D \cap S \neq \emptyset \; \forall \text{ seeds } S \text{ of } O. \quad (4.32)$$

A final scaling can be applied as in the simple approach (Equation 4.31). Unfortunately, an efficient algorithm to compute the seeds of an organization is still lacking.

4.3.4 Comparing Intensity Concepts

To compare the two intensity approaches, we consider two small networks A and B. Both networks consist of two species a and b, and two reactions. They are depicted together with their hierarchies of organizations and the seeds for all organizations in Figure 4.2. For the comparison, we examine all relative species compositions possible within a reactor and two species, excluding an empty reactor. Starting with a reactor only populated by molecules of a (reactor composition 0), more and more molecules of b are introduced until b is the only remaining species within the reactor (reactor composition 1). The relative concentrations for both species in all possible reactor compositions are depicted in Figure 4.3, top. Note that this is not a time dependent trajectory of the system, but simply all possible relative species compositions within the reactor.

For the concentration based intensity approach, the intensities of organizations containg one species only (Org. 1 in network A, and Org. 1 and Org. 2 in

[1]Note that we defined the seed, respectively the generator of an organization, with this application in mind (see Section 2.1.8). The definition of generator differs thererfore from a more common definition (*cf.* footnote on Page 18).

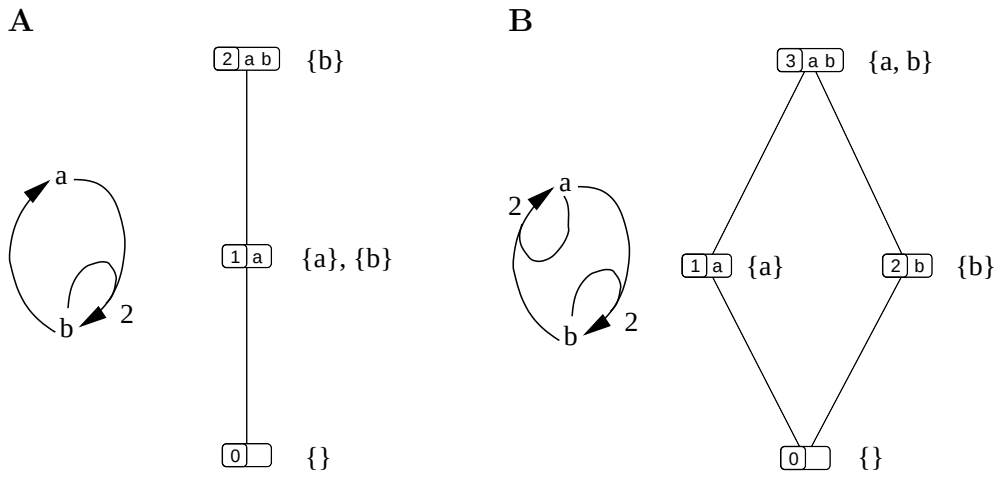


Figure 4.2: Two networks A (left) and B (right) for comparing the concentration based and seed based intensity concepts. The networks, the corresponding hierarchies of organizations, and the seeds for all organizations are shown.

network B) follow exactly the relative concentration of their species (Figure 4.3, top and middle). For the organizations containing both species a and b (Org. 2 in network A, and Org. 3 in network B), the concentration based intensity is constant equal 1, as the sum over both species is taken. Note however, that the intensity drops to zero at the boundaries (reactor composition 0 and 1), as in these two settings, one of the species vanishes.

Using the seed based approach for network A, we find that the intensity of Organization 1 equals the concentration based intensity of Organization 2, except for the boundaries (Figure 4.3, left middle and bottom). As Organization 1 has two seeds, $\{a\}$ and $\{b\}$, both species have to be removed from the reactor to remove the organization. Hence, the sum of both species is used for the intensity. For Organization 2, there exists only one seed: $\{b\}$. Hence, the seed based intensity for this organization follows exactly the relative concentration of b. For network B, the seeds are identical to the organizations. Hence, the seed based intensities are identical to the concentration based intensities, except for the largest Organization 3 at the boundary (Figure 4.3, right middle and bottom).

Both approaches to assign intensity values to organizations differ for network A. While the concentration based approach considers the largest Organi-

zation 2 to be of constant high importance throughout all compositions (except for the boundaries), the seed based approach deems this organization as of rising importance as b becomes more and more dominant in the reactor. This might be more appropriate as b is the more important species for this organization. Being a seed, it can (re-)create the whole organization, independent of species a. Organization 1 only containing a, on the other hand is considered to be of declining importance with declining relative concentration of a by the concentration based approach. The seed based approach assigns a constant high importance to this organization, as both a and b are seeds for this organization. Both have to be removed from the reactor to eliminate this organization.

While the concentration based approach does not assign a value to the empty organization, the seed based approach does not assign a value to the smallest organization, which is empty if the network does not contain any input species. In the presence of input species, the smallest organization is the closure of the input species. Since the empty set is a seed for this organization, the seed based approach cannot assign an intensity value to this organization.

Concluding this section, the two proposed approaches to assign intensity values to organizations have been applied to two simple test networks in order to compare them. The concentration based approach is easy to use, it only considers the concentrations of all organization species. It focuses on the current reactor composition only. The seed based approach additionally considers the regenerative capabilities of organizations. It assigns high values to organizations that are difficult to remove from the reactor, indicating robust and, with respect to the random removal of molecules, stable organizations. This approach might give a truer picture. However, an efficient algorithm to compute organization seeds is still missing. Hence, the concentration based approach will later be used in Chapter 5 to compare and group the organizations of a reaction network modeling the Martian atmosphere.

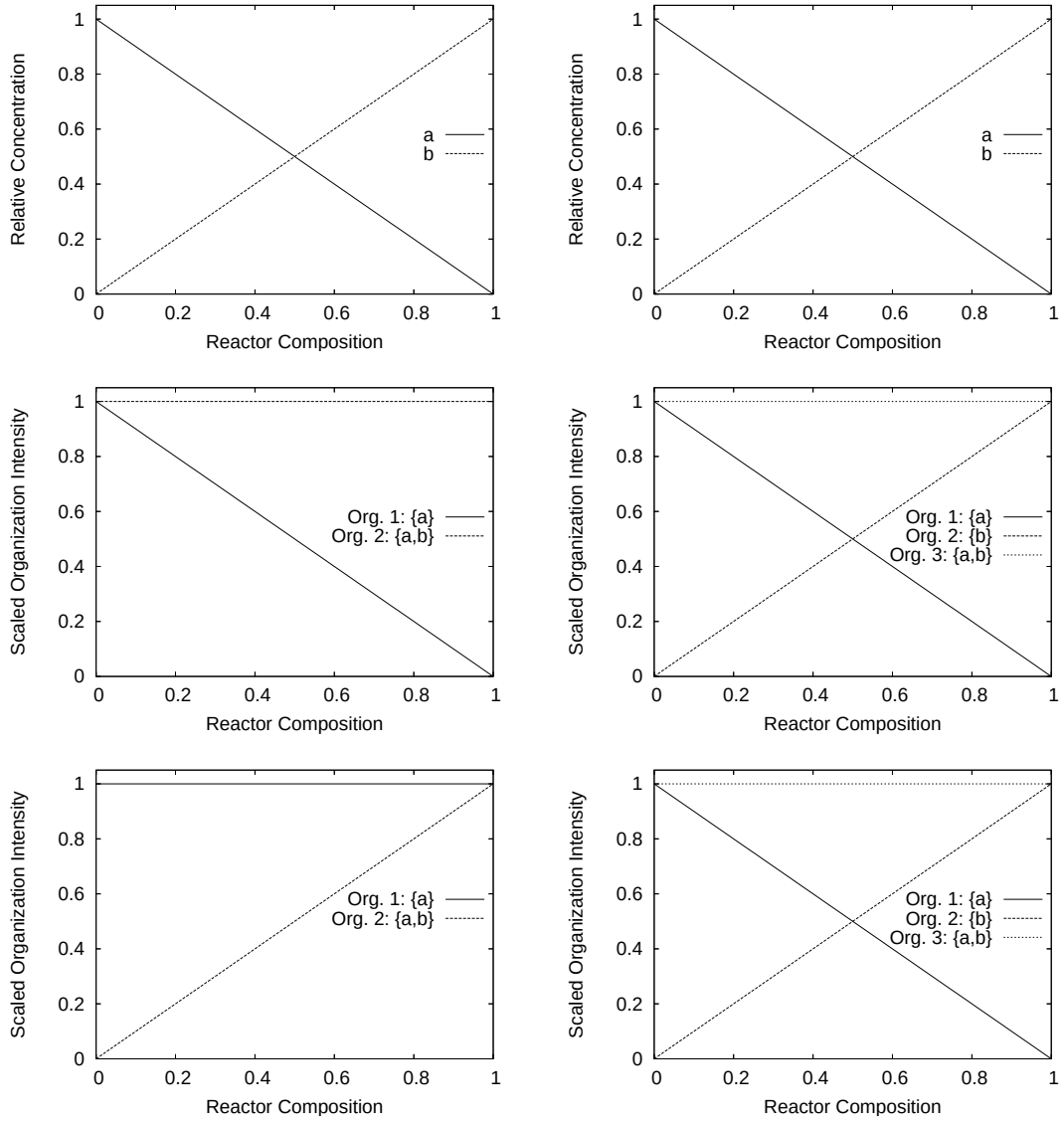
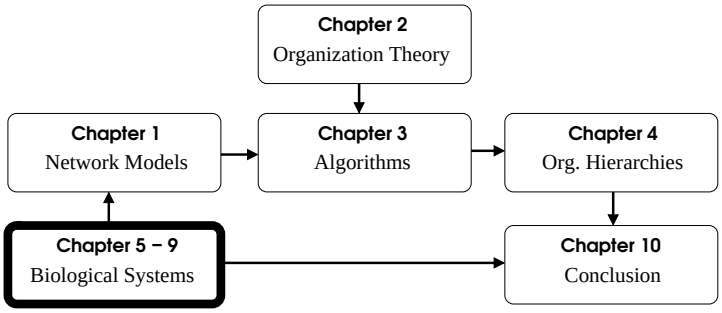


Figure 4.3: Comparing the concentration based and seed based intensity approaches on two networks A (left) and B (right). Top: relative species concentrations within the reactor. Middle: intensity values according to the concentration based approach. Bottom: intensity values according to the seed based approach. The concentration based intensity for organizations containing a and b (Org. 2 in network A and Org. 3 in network B) is zero at the boundaries (reactor composition 0 and 1).



Chapter 5

Photochemical Model of the Mars Atmosphere

As a first application of the theory of chemical organizations for network analysis, a chemical reaction network is investigated that models the atmospheric chemistry of Mars. Chemical networks have the advantage over more general network models that mass is explicitly conserved in all reactions. In such networks, organizations are even more meaningful: it will be shown that for each organization not only a flux vector exists that ensures nonnegative production of all organization species, but that a flux vector exists that leads to zero production of all organization species. Hence, every organization is a balanced organization and describes a potential steady state.

5.1 Models of Planetary Atmospheres

With space probes collecting more and more data on composition and dynamics of planetary atmospheres, it has become feasible to build models of atmospheric chemistries (Yung and DeMore, 1999). The structure of planetary reaction networks has recently been analyzed by methods taken from graph theory (Gleiss et al., 2001; Solé and Munteanu, 2004). Random graph models have further been employed to study the evolution and complexity of planetary atmospheres (Dobrijevic and Dutour, 2006)

Such novel methods are required besides the classical dynamic systems approach as atmospheric models are increasing in size and complexity. This makes them also a good candidate for analysis using the theory of chemical organizations.

In this chapter it will be shown that models of atmospheric chemistries have an inherent structure and how these structures can be uncovered by chemical organization theory. As an example, a photochemical reaction network of the Martian atmosphere by Nair et al. (1994) is analyzed. First, a small subnetwork of the whole system is considered that models a dry atmosphere. Then, the complete network is studied.

5.2 Atmospheric Reaction Networks

Chemical species reacting with each other within an atmosphere form a network. The reaction network can be represented by a bipartite graph, in which one node type represents the molecular species populating the atmosphere and the other the chemical reactions occuring among these species (Zeigarnik and Temkin, 1994). Each reaction node is connected to its educts by incoming links and to its products by outgoing links (*cf.* Sections 1.1 and 2.1.1).

In the model considered, every reaction preserves the number of atoms, that means the mass of the educts is equal to the mass of the products. Accordingly, the whole atmospheric reaction network is mass-conserving.

Typically, the reactions are of second order. For example, hydrogen and oxygen react to form hydroperoxyl: $H + O_2 \rightarrow HO_2$. Some reactions need an additional interaction partner that does not directly take part in the reaction but delivers energy needed for a reaction to occur or absorbs surplus energy released by a reaction. This additional interacting species is labeled M. It acts as a catalyst, appearing on both educt and product side of the reaction.

Apart from direct molecular interactions, solar radiation can lead to the dissociation and ionization of molecules. To describe these processes as ordinary reactions, a pseudo molecule $h\nu$ is introduced that represents energy supplied by solar radiation. As an example, the photodissociation of CO_2 can be written as $CO_2 + h\nu \rightarrow CO + O$. Since pseudo molecule $h\nu$ is not created by any reaction in

the network it is easy to model the day- and nightside of an atmosphere: defining $h\nu$ as an external input corresponds to the dayside. If $h\nu$ is not an input, the nightside situation is modeled.

Planetary atmospheres are spatial inhomogeneous systems. Diffusion, convection and advection play an important role in the dynamics of atmospheres. However, by just considering the topology of the reaction network, we treat atmospheric chemistries as if their reactions would take place in a well-stirred reaction vessel, neglecting any spatial structure present in the real system.

5.3 Chemical Organizations in an Atmospheric Model of Mars

As an example for an atmospheric reaction model, a photochemical model of the Martian atmosphere by Nair et al. (1994) is considered. Since pseudo molecule M, representing an additional, unspecified reaction partner, always acts as a catalyst, it can be removed from the network without loosing information about its structure. Dust particles blown from the suface by wind lead to aerosols forming dust storms. These aerosol particles can adsorb HO_2 molecules (reaction R103, see Appendix A.2). In the adsorbed state, HO_2 reacts with OH to form H_2O and O_2 (reaction R104). For consistency, grain is added as a product in this reaction, which is missing in the original model by Nair et al. (1994).

After these modifications, the network contains 31 molecular species and 103 reactions, excluding the input reaction providing $h\nu$ to simulate the dayside. The most connected species is O with 42 appearances as educt and product. Next are O_2 and OH with 35 appearances each. Among the less connected species are N_2O_5, CO_2H^+, $(HO_2)_{\text{grain}}$, and grain with two appearances each. The majority of all species has a low connectivity with less than 10 appearances in reactions. See Appendix A.2 for a complete list of all reactions and A.3 for species connectivities. An exponential degree distribution for the network is reported by Solé and Munteanu (2004).

We assume that the reactions contained in the model are exactly the reactions that can occur within the atmosphere under the respective physical conditions. If

these conditions (*e.g.*, temperature and pressure) change during the atmospheric evolution and new chemical reactions become possible or reactions are no longer feasible, the analysis of the modified reaction network model will most likely lead to a different organizational structure.

As detailed above, all reactions are mass-conserving, preserving the number of atoms for all atomic species in the system. Reactions including pseudo molecule $h\nu$ seem to be an exception to this rule. This species is used up or created from "nothing" in the respective reactions. But with this species only indicating additional photonic energy required for specific reactions to occur, the overall mass-conservation of atomic species is still fulfilled. Under such mass-conserving circumstances without external influx, even more can be said for organizations. Not only does a positive flux vector exist, so that the concentration change for all species within the organization is greater or equal to zero ($dc_i/dt \equiv \dot{c}_i \geq 0$ for all organization species i), but for all such flux vectors the concentration change is exactly equal to zero ($\dot{c}_i = 0$ for all organization species i), representing a steady state. Proof: having an organization O, we know that just considering the reaction network made up by the organization species, at least one flux vector exists with $\dot{c}_i \geq 0$ for all organization species i. Let v be such a flux vector and assume that there is one organization species j with $\dot{c}_j > 0$. Accordingly, the concentration of species j increases in time. More and more atoms will be allocated to this molecular species. But since the number of atoms is constant for all atomic species over time (all reactions are mass-conserving), there must be another molecular species k in the organization whose concentration is declining with $\dot{c}_k < 0$. This violates our initial assumption that v fulfills $\dot{c}_i \geq 0$ for all organization species i. Hence, all flux vectors fulfilling $\dot{c}_i \geq 0$ for all organization species i must also fulfill $\dot{c}_i = 0$. All organizations are balanced organizations.

Consequently, organizations represent all potential steady state species compositions for mass-conserving networks without external influx. The steady states denoted by an organization are characterized by having a concentration greater zero for all species contained in the organization and zero for all other species. There are no other steady states possible, in which the set of species with concentrations greater zero does not exactly coincide with a corresponding organization. The steady states defined by organizations are only potential, since it is assumed

that the reaction fluxes can be arbitrarily chosen. Whether the required fluxes can be realized or not depends on further kinetic data which is not considered here (*e.g.*, reaction rates).

5.3.1 Dry Atmosphere

In order to demonstrate how organization theory can be applied to the Martian atmospheric model, we first consider a small part of the complete network model. Nair et al. (1994) used a dry atmosphere without H_2O molecules as a test case for their model. We use this scenario for a first analysis. Without water, only six chemical species (CO_2, CO, O, O_2, $O(^1D)$, and O_3) and pseudo species $h\nu$ are part of the reaction network. The species react in 16 reactions: R1-6, R14-16, R18-23, and R25 (see Appendix A.2 for a complete reaction list).

5.3.1.1 Dayside

To model the dayside atmosphere, reaction R1 supplies pseudo species $h\nu$ to the system, representing incident sunlight. The system contains six organizations as depicted in Figure 5.1 (A). The smallest Organization 0 just comprises pseudo species $h\nu$. It is closed as $h\nu$ alone cannot produce any other species, and self-maintaining as it is defined as an input. Adding CO or $O(^1D)$ gives rise to two organizations of size two: Organizations 1 and 2. As there is no reaction between the members of these organizations, the species sets are closed, and additionally self-maintaining as no species decays. The same is true for Organization 3 which is the combination of Organizations 1 and 2. Adding O and O_2 to Organization 2 gives Organization 4. This is the first organization in which reactions occur. The middle circle in Figure 5.2 encloses the organizational network of this organization. Finally, the whole network is contained in the largest Organization 5. The corresponding network (*i.e.*, the complete network) is also portrayed in Figure 5.2.

The six organizations of this network represent all steady state species compositions of the system. From this analysis we can already conclude that CO and $O(^1D)$ are the only candidates for stable atmospheres just containing a single species, as they are the only single species organizations. Any other species

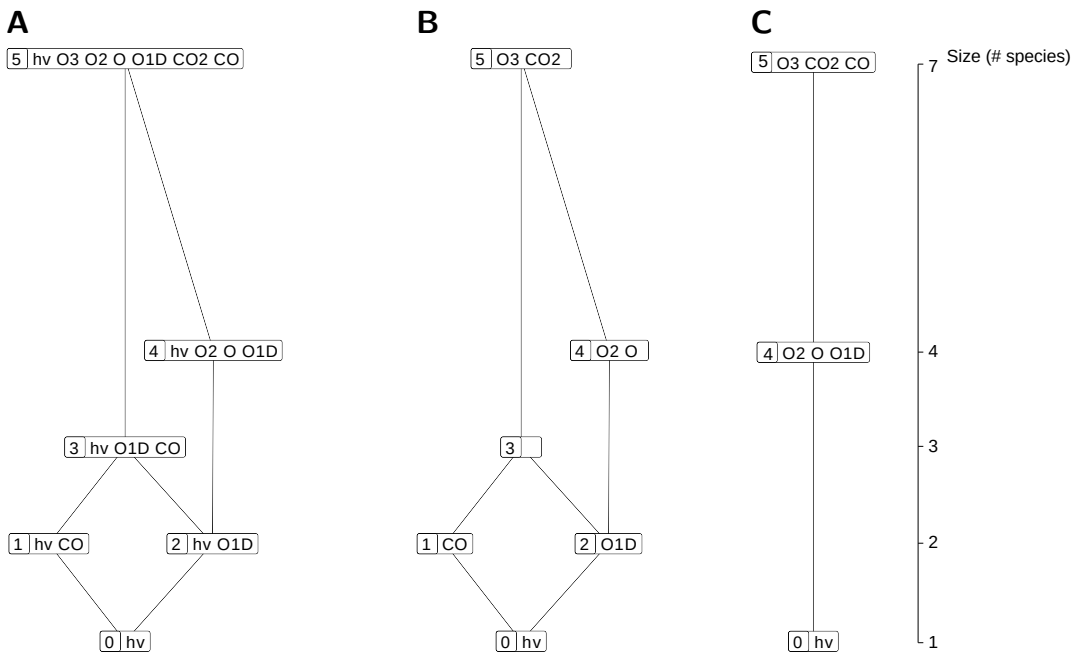


Figure 5.1: Hierarchy of organizations for the dry atmosphere model, dayside. (A) Hasse diagram with all species listed in organization labels. The network contains six organizations. Organization 5 contains the whole network. (B) Compact form of the Hasse diagram. Only those species are listed in organization labels, that are not already contained in organizations to which a downlink exists. All following Hasse diagrams will use this compact form. (C) Hierarchy of connected organizations. Only two (non-trivial) organizations are connected. See Appendix A.1 for a list of species symbols.

would dissociate under solar radiation (O_2, O_3, CO_2) or react spontaniously (O), giving rise to new product species. Hence, these single species sets are not closed.

The hierarchy of organizations can be presented in a more compact form as shown in Figure 5.1 (B). Here, the organization labels contain only those species, which are additionally present in the organization apart from the species in organizations to which a downlink exists. This compact representation of the Hasse diagram of organizations will be used throughout the remaining part of the thesis. Empty labels indicate that the organization is a simple union of smaller organizations and no new species are generated (*e.g.*, Organization 3).

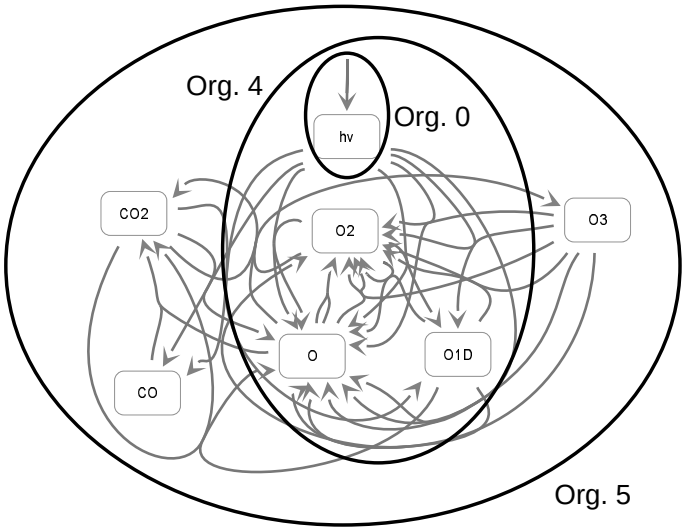


Figure 5.2: Connected organizations in the reaction network of the dry atmosphere model, dayside. The network contains seven species and 16 reactions. Species $h\nu$ is supplied as an input, representing incident sunlight. The smallest connected organization only contains $h\nu$. Connected Organization 4 contains $h\nu$, O, O(^{1}D), and O_2. The largest connected Organization 5 additionally contains O_3, CO, and CO_2. This is the complete network. See Appendix A.1 for a list of species symbols.

In Organization 5, CO from Organization 3 and O from Organization 4 come together making the synthesis of CO_2 possible. Hence this species appears in the label. An organization with non-empty label that has just one downlink indicates, that the lower organization can be expanded. Adding species specified in the label allows the creation of species beyond the closure of the smaller organization (*e.g.*, Figure 5.1 (B), Organization 4).

The hierarchy of connected organizations for the dry atmosphere model is shown in Figure 5.1 (C). Only two of the original six organizations are connected: the complete network and one smaller subnetwork. Both corresponding reaction networks are shown in Figure 5.2. Note, how Organization 4 is expanded to the enclosing Organization 5. Only one reaction can create O_3. For this, reaction R18 requires CO_2 as a catalyst. If CO_2 is present, also CO can be generated (R14), and vice versa (R20).

5.3.1.2 Nightside

To model the nightside, the $h\nu$ supplying reaction R1 is removed from the network model. The resulting nightside hierarchy of organizations is shown in Figure 5.3 (A). The network contains 22 organizations. The smallest organization is empty. If no input species are present, the empty set is always an (trivial) organization. All species except species O form single species organizations. Specis O would react spontaneously to from O_2 (R16), hence the species set just containing O is not closed. Moreover, species O does not appear in any organization. Without incident sunlight, $O(^1D)$ is required to form O (R21, 23, and 25). But to create $O(^1D)$, sunlight is required (R3, 5, and 15). Hence, there is no production pathway for species O and with spontaneous O_2 formation (R16), the concentration of O will decay. Hence, species O cannot be part of any organization. The hierarchy of connected organizations is shown in Figure 5.3 (B). It is identical to the lower part of the complete organizational hierarchy. Only the empty organization and the one species organizations are (trivially) connected. It follows that all other organizations are simply combinations of species that do not react with each other. No organization contains any reactions. With organizations representing all species compositions which allow for steady states, we conclude that for the nightside, all reactions have to come to a halt in any steady state. Steady state atmospheres on the nightside contain only one single chemical species: $O(^1D)$, CO, CO_2, O_2, or O_3.

5.3.2 Complete Atmospheric Model

Finally, the whole network model of the Martian atmosphere is considered. Some results obtained from the reduced dry model coincide with results for the complete model, especially for the nightside.

5.3.2.1 Dayside

To analyze the dayside of the full Martian atmosphere model, reaction R1 defines pseudo species $h\nu$, representing sunlight, as an input to the system. We find 1484 organizations (see Appendix A.4 for a complete list). The Hasse diagram is

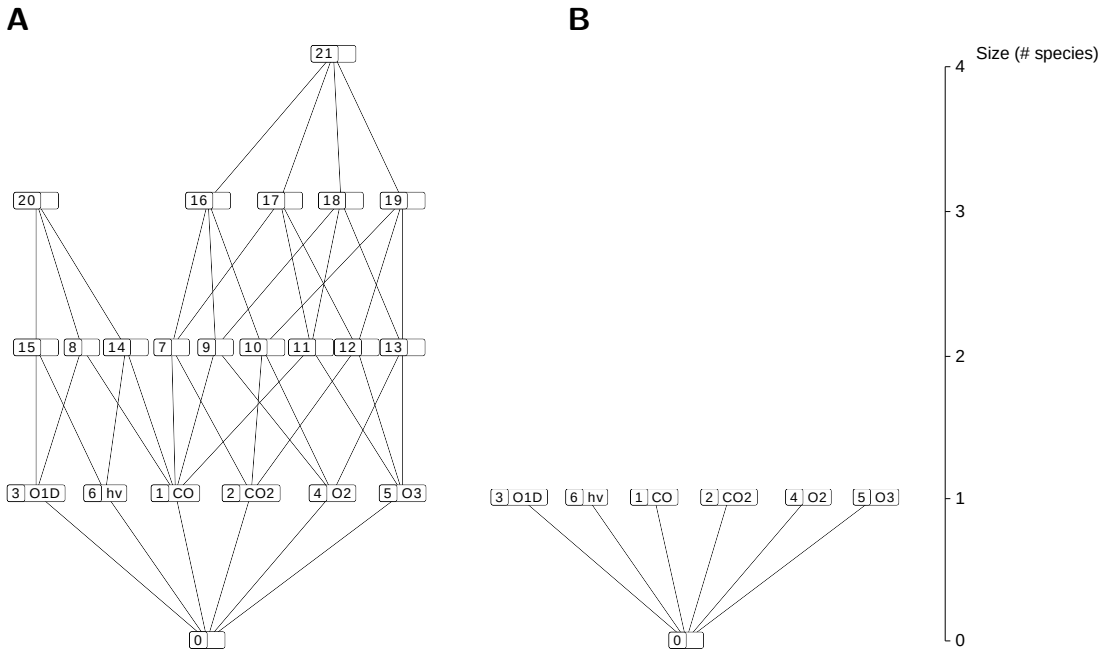


Figure 5.3: Hierarchy of organizations for the dry atmosphere model, nightside. (A) The network contains 22 organizations. No organization contains any reactions, hence all organizations are simply containing non-interacting species. Species O is not contained in any organization. (B) Hierarchy of connected organizations. Only the empty organization and all single species organizations are (trivially) connected. See Appendix A.1 for a list of species symbols.

depicted in Figure 5.4 (A). The smallest organization just contains one species: $h\nu$, the input species. There are 11 organizations directly above this organization containing two species each: $h\nu$ and one additional species. A combinatorial explosion leads to a multitude of organizations of sizes 3 up to 10 species. The mechanism behind this combinatorial complexity is simple: if the species of one organization have no interactions with the species of another organization, these two organizations can be merged to form a new organization. However, there are only 14 organizations with more than 10 species, indicating that only few species combinations containing many species are feasible for steady states.

All organizations with more than 10 species are listed in Table 5.1. The largest organization contains all species of the network, representing the whole reaction

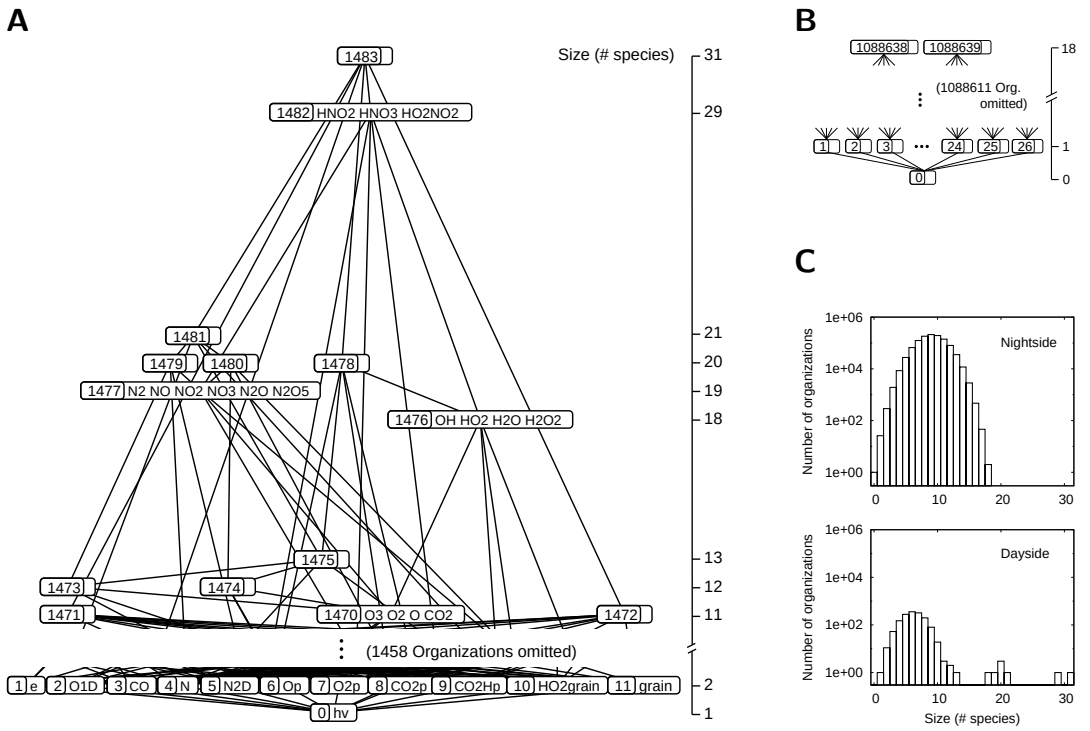


Figure 5.4: Hierarchy of organizations for the Martian photochemistry model. (A) Dayside. 1484 organizations total, see Table 5.1 for a full list of organizations with more than 10 species. See Appendix A.1 for a list of species symbols. (B) Nightside. 1088640 organizations total. Organizations 1 to 26 contain one species each; every species except O, H, OH, HO_2, and N_2 appear in these organizations. Organization 1088638 contains 18 species: O_3, O_2, H_2, H_2O, H_2O_2, CO, $N(^2D)$, NO_3, N_2O, N_2O_5, HNO_2, HNO_3, HO_2NO_2, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, and grain. Organization 1088639 contains the same 18 species, with H_2 replaced by CO_2^+. Smaller terminal organizations are omitted. Species information in labels is omitted for clarity. (C) Distribution of organization sizes for nightside (top) and dayside (bottom).

system. This is in accordance with the assumption that the atmosphere is in a steady state on the dayside with all species present.

When comparing the hierarchy of organizations of the complete model with the introductory dry atmosphere example, it becomes apparent that the four smallest Organizations 0-3 of the smaller model are also organizations in the

Table 5.1: Organizations with more than 10 species for the dayside. Connected organizations are marked with *; all their species are connected to each other by reactions forming a single subnetwork.

ID	# Species/ # Reactions	Species
1483*	31 / 104	$h\nu$, e, O_3, O_2, O, $O(^1D)$, H_2, H, OH, HO_2, H_2O, H_2O_2, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, HNO_2, HNO_3, HO_2NO_2, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1482*	29 / 102	$h\nu$, e, O_3, O_2, O, $O(^1D)$, H_2, H, OH, HO_2, H_2O, H_2O_2, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, HNO_2, HNO_3, HO_2NO_2, O^+, O_2^+, CO_2^+, CO_2H^+
1481	21 / 55	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1480	20 / 55	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, O^+, O_2^+, CO_2^+, grain
1479	20 / 55	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1478*	20 / 57	$h\nu$, e, O_3, O_2, O, $O(^1D)$, H_2, H, OH, HO_2, H_2O, H_2O_2, CO_2, CO, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1477*	19 / 55	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, O^+, O_2^+, CO_2^+
1476*	18 / 55	$h\nu$, e, O_3, O_2, O, $O(^1D)$, H_2, H, OH, HO_2, H_2O, H_2O_2, CO_2, CO, O^+, O_2^+, CO_2^+, CO_2H^+
1475	13 / 25	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1474	12 / 25	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, O^+, O_2^+, CO_2^+, grain
1473	12 / 25	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1472	11 / 3	$h\nu$, H_2, H, CO, N, $N(^2D)$, O^+, O_2^+, CO_2H^+,

ID	# Species/ # Reactions	Species
		$(HO_2)_{grain}$, grain
1471	11 / 1	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1470*	11 / 25	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, O^+, O_2^+, CO_2^+

complete model. However, Organizations 4 and 5 in the small model do not appear in the full model. The corresponding species sets are not closed anymore if the reactions of the complete model are considered. Starting from Organization 4 with $h\nu$, $O(^1D)$, O, and O_2, the ionization reactions R92 and R93 lead to the formation of O^+, O_2^+, and e. In order to consume O^+, species CO_2 is required to facilitate the only O^+ consuming reaction R98 (species sinks cannot exist in organizations of mass-conserving networks). Then, also CO is present in the species set. And furthermore, CO_2^+ can be formed (R94) and O_3 (R18). The resulting species set is Organization 1470 in the large model. This is the smallest organization containing Organization 4 and 5 of the smaller model.

One interesting (and anticipated) observation is that all species involved in the main CO_2 recycling pathway appear together in an organization. The smallest organization encompassing the whole CO_2 recycling is Organization 1476 (see Table 5.1). The atmosphere of Mars consists mainly of CO_2 (95%). In a pure CO_2 atmosphere under solar radiation, CO_2 would quickly dissociate to CO and O, and two O molecules would easily react to form O_2 according to the overall reaction:

$$2\ CO_2 + h\nu \rightarrow 2\ CO + O_2. \qquad \text{(R14, 16)}$$

Under these circumstances, high amounts of CO and O_2 would be expected. The lack of such high amounts in the Martian atmosphere and the question how CO_2 can be constantly maintained at such high levels is known as the CO_2 stability problem. It was solved independently by McElroy and Donahue (1972) and Parkinson and Hunten (1972). The main idea is simple: the Martian atmosphere contains small amounts of water vapor, most likely originating from the northern polar cap during spring time. Its concentration fluctuates strongly with

season, altitude, and latitude. Water can easily be photolyzed by solar radiation:

$$H_2O + h\nu \rightarrow OH + H. \tag{R10}$$

The OH group formed in reaction R10 can be used to form CO_2 from CO and O in the following reaction scheme (R43, 28, 35):

$$CO + OH \rightarrow CO_2 + H \tag{R43}$$
$$H + O_2 + M \rightarrow HO_2 + M \tag{R28'}$$
$$O + HO_2 \rightarrow O_2 + OH. \tag{R35}$$

Being used as catalysts for the recycling of CO_2, the HO_x radicals (H, OH, HO_2) are not used up in this reaction sequence. This is the most important recycling pathway for CO_2 in the Martian atmosphere. Variants of this pathway synthesize two CO_2 molecules from two CO and one O_2 molecule using additionally NO_x molecules as catalysts. As the catalysts are not consumed in the recycling pathway, small amounts of these molecules are sufficient to recycle large amounts of CO_2 leading to the observed high and stable concentration of CO_2. The smallest organization containing CO_2 is Organization 1470. Adding species H_2, H, OH, HO_2, H_2O, H_2O_2, and CO_2H^+ to this organization gives Organization 1476. This is the smallest organization containing CO_2 and additionally all HO_x species. Here, all species involved in the major CO_2 recycling pathway are present. In Organization 1482, the NO_x species are additionally present. Hence, this organization contains this alternative CO_2 recycling pathway. It is interesting to note that Organization 1470 — containing CO_2 but not the HO_x species necessary for CO_2 recyling — is still a self-maintaining species set. That means that CO_2 can in principle be maintained even without the radical catalysts. This has already been shown in the introductory dry atmosphere model where CO_2 was contained in the largest organization. Formation of CO_2 was facilitated by reacting O and CO in reaction R20. Besides this reaction, another CO_2 formation reaction is possible in Organization 1470: electrons can be transfered from O to CO_2^+ in reaction R100. Since CO_2 maintenance in the absence of HO_x radicals is not observed in the real atmosphere, the fact is highlighted that it is only in

principle that organizations are capable of self-maintenance. Further kinetic information not considered in the analysis is necessary to determine whether an organization is able to prevail over time or not.

Another interesting observation should be noted: computing the nontrivial unit species sets (see Section 4.2) reveals that several species always appear together in organizations of the dayside. There are five such species groups as listed in Table 5.2. If one species of such a group is present in an organization, then all other species of that group will be present, too. If an organization does not contain one species of a group, the other species of that group will be absent as well. Table 5.2 reveals that species of similar molecular structure form the groups. The first group contains H and H_2. In presence of incident sunlight, these two species can easily be converted into each other (R7, 27). The second group consists of O_x species and CO_2. If O is present, O_2 can be formed and vice versa (R2, 16). From these two species, CO_2 catalyzes the creation of O_3 (R18). And conversely, O_3 can be decomposed into O and O_2 (R4). The NO_x species are lumped together in Group 3. The 4th group contains the H_xO_x species and finally, Group 5 collects the HNO_x species. This grouping effect is caused by the fact that it is more likely to have reactions transforming species into species of similar molecular structure. The species within one group can mostly be transformed from one into the other easily. It should be noted that the groups are not exhaustive. For example, $O(^1D)$, O^+, O_2^+, CO, and CO_2^+ clearly would fit into Group 2, and indeed if an organization contains the species of Group 2, also these species will be present in the organization (they are in the unit add-on species set of Group 2). But since there are other organizations in which Group 2 is not present but one or more of the listed species, they do not belong to the group. With species of similar molecular structure grouped together, it is tempting to suspect that these groups also form organizations. But this is only true for the first group: H, H_2, and $h\nu$ together form organization 28. The other groups are not closed and hence cannot be organizations. Yet, computing the closure for the remaining groups shows a close resemblance to certain organizations. Computing the closure of a species set is done by consecutively adding all species to the set that can be produced from the set species. The closure of species Group 2 is identical to the species of Orga-

Table 5.2: Species always appearing together in organizations, dayside. Species of similar molecular structure are grouped together in the unit species sets.

Group	Species	Occurences in Organizations
1	H, H_2	324
2	O_3, O_2, O, CO_2	12
3	N_2, NO, NO_2, NO_3, N_2O, N_2O_5	6
4	OH, HO_2, H_2O, H_2O_2	4
5	HNO_2, HNO_3, HO_2NO_2	2

nization 1470 (see Table 5.1). The closure of Group 3 does not contain CO, CO_2, and CO_2^+ but is otherwise exactly Organization 1477. Generating the closure of Group 4 leads to Organization 1476 but excluding CO, CO_2, CO_2^+, CO_2H^+, and O_3. And finally, Group 5 gives Organization 1482 but without CO, CO_2, CO_2^+, and CO_2H^+. The mentioned organizations are for all cases the smallest organizations containing the group species. All organizations above these organizations will naturally contain the group species, too. This grouping of species according to similarity in molecular structure is not based on knowledge about the real structure of the species. Rather, the reaction network structure is exploited to achieve this grouping. If the system can be assumed to be in a steady state, the species groups show which species are always present simultaneously. If one species of a species group is detected, all other species of the group are expected to be present as well. Such predictions can be trivial as in Group 1. A not so trivial prediction can be found in Group 2. Here, the presence of O_x species implies the presence of CO_2. Interestingly, N_2 can also be used as a catalyst for O_3 formation like CO_2, yet is not part of this species group.

5.3.2.2 Nightside

For simulating the nightside, $h\nu$ is not supplied as an input species. In this case, many more organzations exist: 1088640. Figure 5.4 (B) gives a schematic overview of the organizational hierarchy. The results are similar to the results of the smaller dry atmosphere model. Since no input molecules are present, the smallest organization is the empty organization. Above this organization, 26

organizations exist that contain only a single species. All species except O, H, OH, HO_2, and N_2 appear as single species organizations. For the mentioned species, reactions exist that transform these molecules without the help of any other species (*e.g.*, 2 O $\rightarrow$ O_2, R16). Hence, they cannot be organizations just by themselves since they are not closed. They also do not appear in any other organization of the nightside.

The two largest organizations contain 18 species each (see caption of Figure 5.4 for species lists). Other terminal organizations, that means organizations that have no organizations above them, are not shown in the diagram. When analyzing the organizations more closely, it turns out that all organizations consist of species that are not interacting[1]. Every organization is an organization simply because all species contained in it do not interact in any way. Hence, no reactions occur within the nightside organizations, which is in accordance with the model being a photochemical model of atmospheric reaction processes, relying on sunlight as an external energy source.

5.3.3 Connected Organizations

Although having more organizations compared to the dayside, the hierarchy of organizations is much simpler at night: the combination of non-interacting species gives rise to a combinatorial explosion leading to a large number of organizations. The higher complexity of the dayside hierarchy becomes apparent when comparing the size distributions of organizations for the day- and nightside. The size histograms in Figure 5.4 (C) show, how many organizations of a particular size exist in the organizational hierarchy of the day- and nightside. The combinatorial explosion leads to a bell-shaped distribution of organization sizes for the

[1]The two largest organizations have no reactions. Consequently, all organizations below them cannot contain any reactions. But there might be organizations not below the two largest organizations that have reactions. However, when inspecting the connected organizations (see Section 5.3.3), there are only one species organizations. If there were a partially connected organization not below the two largest organization, we could remove the not connected species and end up with a connected organization. But this organization would then show up in the connected hierarchy. Consequently, no organization of the nightside contains any reactions.

nightside[1]. The dayside also contains a bell-shaped distributed component for organization sizes between approximately 0 and 10 species. But additionally, we find 14 organizations with more than 10 species.

To eliminate the observed combinatorial effects, it is appropriate to consider only the connected organizations, in which all molecular species are connected to each other by reactions. These organizations form coherent networks without isolated species or subnetworks. Figure 5.5, panels (A) and (C) depict the hierarchies of connected organizations for the day- and nightside. On the dayside, eight organizations are connected. Their sizes range from one (just the input species $h\nu$) to 31 species (the complete reaction network). On the nightside, 27 connected organizations exist: the empty organization and 26 organizations containing just a single species. This hierarchy is identical with the lower part of the hierarchy of all nightside organizations as depicted in Figure 5.4 (B).

For the dayside, the hierarchy of connected organizations reveals how species of similar chemical composition form organizational subnetworks. Starting from the smallest organization, two different principles lead to larger organizations. First, a set of species can be added to an organization to give a larger one. And second, the union of two organizations can lead to a new organization. Here, such an organization always contains more species than the union of the species contained in the two constituting organizations (with one exception: the merger of Organzations 1482 and 1478 to form the largest Organization 1483 encompassing the whole network does lead to more species than the set union). Interactions between the merged species sets lead to the creation of novel species. The smallest organization can be expanded in two ways: adding H_x species gives Organization 28, while adding the O_x and CO_x species, and e leads to Organization 1470. Combining these two organizations brings together hydrogen and oxygen atoms, and hence we additionally find the H_xO_x species and CO_2H^+ in Organization 1476. Organization 1470 can be expanded by adding the N_x and N_xO_x species to form Organization 1477. When this organization is merged with Organization 1476, hydrogen, nitrogen, and oxygen atoms are available. Hence, the resulting Organization 1482 contains species made up by these atoms, namely the HNO_x

[1]The Kolmogorov-Smirnov test (with Lillefors modification) reveals that it is not a normal distribution.

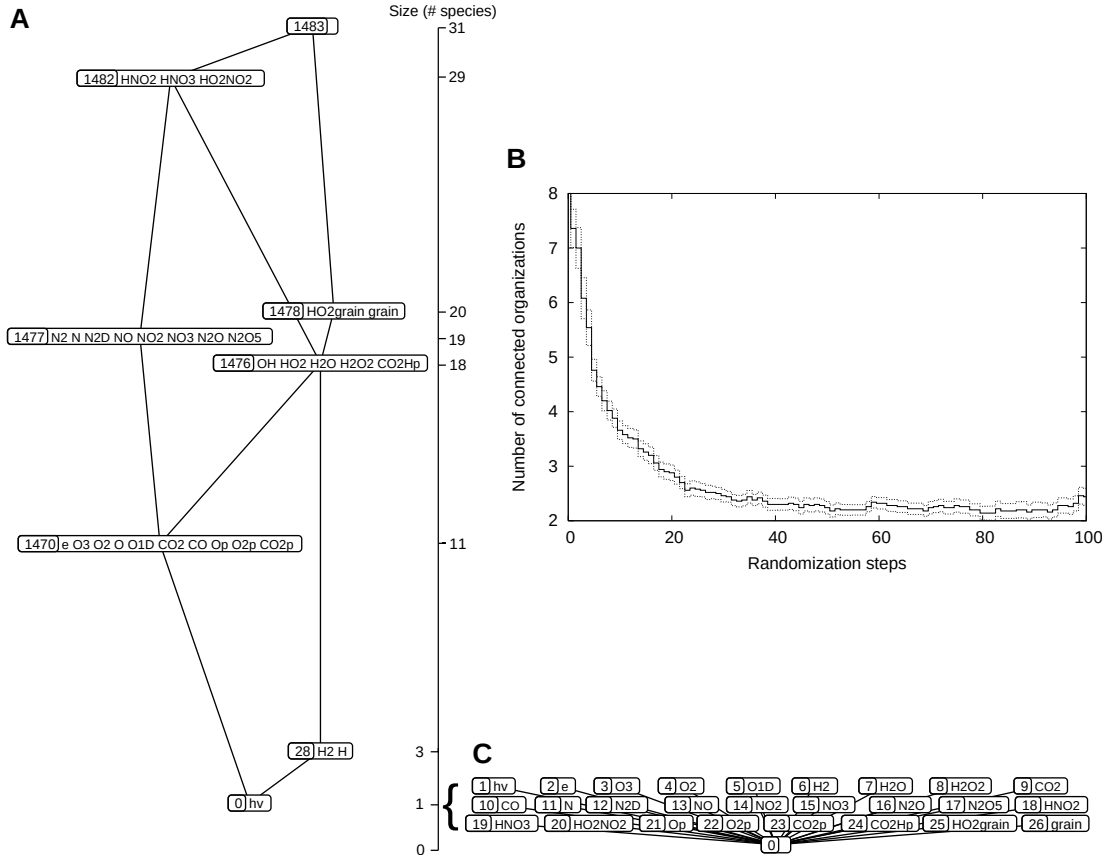


Figure 5.5: Hierarchy of connected organizations for the Martian photochemistry model. (A) Dayside with eight connected organizations. (C) Nightside with 27 connected organizations. Organizations 1 to 26 all contain one single molecular species. See Appendix A.1 for a list of species symbols. (B) Randomizing the reaction network, dayside. Number of connected organizations, averaged over 50 runs and standard deviation of the mean.

species and HO_2NO_2. Adding the two species grain and $(HO_2)_{grain}$ to Organization 1476 gives Organization 1478. This organization can finally be merged with Organization 1482, resulting in the complete reaction network as the largest Organization 1483. Note that in this case the merger does not give rise to novel species. There are reactions between the two organizations but they do not create species not yet present in the two organizations. The connected organizations group the molecular species according to their atomic constituents. For hydrogen,

all species made up solely by hydrogen even form an organization (Organization 28). In the presence of incident sunlight, H_2 can be split into two H atoms and two H atoms can react back to form H_2. Since these species cannot produce any other species just by themselves, they form not only a self-maintaining but also a closed set, resulting in an organization. This is not the case for N. Here, N_2 can react to N (R46), but to create N_2 from N, additional NO is required (R61). Hence, N and N_2 do not form an organization. Also there is no organization containing solely the O_x species. Organization 1470 not only contains O_x species but also the CO_x species. When inspecting the O_x species, we find that there is only one reaction that consumes O^+. Namely, O^+ together with CO_2 react to CO and O_2^+ (R98). If we only consider the network made up by the O_x species, O^+ would therefore represent a sink. Since organizations in mass-conserving networks cannot contain sinks, an O_x species organization is not feasible. If we add the reaction $O^+ + e \rightarrow O$, so that O^+ is no longer a sink, we find that then $h\nu$ and all O_x species except O_3 indeed form an organization. Species O_3 is not included since N_2 or CO_2 is required as a catalyst for its creation (R17, 18).

This exemplifies how organizations can be used to validate a reaction network model. If an expected organization is missing, or the other way round, an organization seems implausible, closely inspecting the reaction network will uncover the cause. This in turn can give hints on missing or incorrect reactions in the model. Whether the suggested reaction in this case is chemical feasible or should be added to the model or not is not at the focus of this study.

5.3.4 Randomizing the Reaction Network

For the dayside, the reaction network of the Martian atmosphere exhibits a nontrivial hierarchy of 1484 organizations. It is reasonable to assume that such nontrivial hierarchies are typical for natural reaction systems. To test this hypothesis, we analyze more and more randomized versions of the Martian network. The original reaction network is randomized step by step. In each step, two molecules are taken from the reaction list at random and swapped, preserving the network structure including arity of educts and products for all reactions. As an example we consider following two reactions:

$$\mathrm{O} + \mathrm{O_3} \rightarrow 2\ \mathrm{O_2} \tag{R19}$$

$$\mathrm{O} + \mathrm{CO} \rightarrow \mathrm{CO_2}. \tag{R20}$$

If O in reaction R19 and CO_2 in reaction R20 are selected, the two reactions will be changed to:

$$\mathrm{CO_2} + \mathrm{O_3} \rightarrow 2\ \mathrm{O_2} \tag{R19'}$$

$$\mathrm{O} + \mathrm{CO} \rightarrow \mathrm{O}. \tag{R20'}$$

Note that the network's property of mass-conservation is not preserved by this procedure. Starting from the original network, we successively apply 100 such randomization steps and count the organizations found in the resulting networks at each step. This procedure is independently repeated 50 times. Since a large number of organizations does not necessarily indicate a complicated organization hierarchy due to combinatorial effects, we do not count the number of all organizations for this analysis. Instead, we count the number of connected organizations. Figure 5.5 (B) depicts the result of the network randomization, averaged over the 50 independent runs. In the original network, eight connected organizations exist. With increasing randomization, the hierarchy of connected organizations quickly breaks down. After 50 randomization steps, an asymptotic value of approximately 2.3 connected organizations is reached and maintained. Since the network contains one input species (pseudo species $h\nu$), there is always at least one connected organization, namely the organization containing the input species (plus possibly further follow up species). As expected, the randomized reaction networks feature a significantly lower number of connected organizations compared to the original reaction network of Mars.

5.3.5 Organization Intensities in a Simulated Mars Profile

Using a simulated concentration profile of the Martian dayside atmosphere based on an augmented mars model that also includes the methane chemistry (Nair

et al., 2005) obtained from the authors, we calculate the height dependent distribution of organizations. For the analysis, the data on molecular species not included in our reaction network is ignored (notably the methane species). Additionally, the simulated model does not contain information on the species $(HO_2)_{grain}$, grain, and pseudo species $h\nu$. Hence, these species will be omitted by treating them as if they were not part of the organization when calculating intensity values. To calculate the organization intensities, the concentration based approach is used (see Section 4.3.2, equations 4.30 and 4.31). For height dependent concentration profiles, the altitude takes the role of the time in concentration time courses.

The altitude profile of the relative species concentrations for the six most abundant species is shown in Figure 5.6, left. For altitudes up to roughly 140 km, CO_2 is the most abundant component making up more than 95% of the atmosphere. Above 140 km, the CO_2 fraction is steadily declining. The loss of CO_2 is compensated for by increasing relative concentrations of O, N_2, CO, H_2, and N. Relative concentrations of N_2 and CO peak at approximately 230 km and slightly decline afterwards. The absolute concentrations of nearly all species decline with higher altitudes as the atmosphere becomes thinner and thinner.

The scaled intensity trajectories for all 1484 dayside organizations are shown in Figure 5.6, right. The trajectories can be grouped into four distinct groups: (i) jumping from zero to one at approximately 60 km and slightly declining later on; (ii) increasing and reaching a maximum at approximately 230 km and slightly declining afterwards; (iii) increasing sigmoidal and reaching a maximum at the upper boundary of the model at 240 km; and (iv) exponential-like increase beginning at approximately 140 km, peaking at the upper model boundary.

A common feature of almost all trajectories is the increasing behavior. This can be explained by the fact that CO_2 is the only considerably decreasing molecular species. Since CO_2 is contained in only 12 organizations, these are the only organizations that can feature a decreasing intensity value. Coincidently, these are exactly the organizations making up Group (i). All organizations listed in Table 5.1 except organization 1471 belong to this group. All connected organizations except organizations 0 and 28 (see Figure 5.5 (A)) are included. The jump at 60 km is caused by the presence of CO_2^+ in these organizations. This species

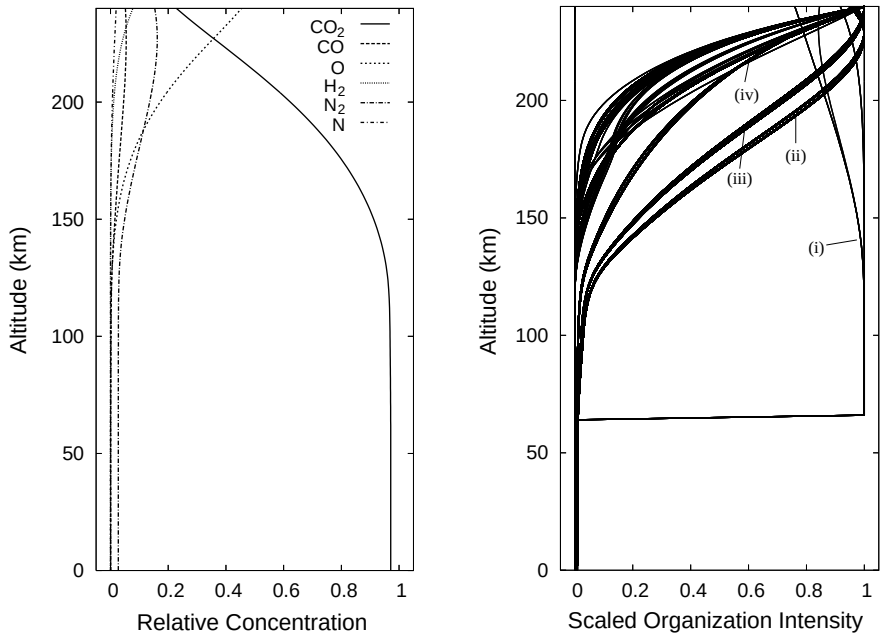


Figure 5.6: Relative concentration profile of the Martian atmosphere according to a simulation run showing the six most abundant species (left), and profile of scaled intensity values for all 1484 dayside organizations. The intensity profiles are divided into four groups: (i) organizations containing CO_2, (ii) organizations containing CO but not N, H_2, and O, (iii) organizations containing CO and N, but not H_2 and O, and (iv) all remaining organizations with positive concentrations. See text for details.

only occurs at altitudes above 60 km. The trajectory of the largest organization containing all species is also in this group. With all species being present above 60 km, it jumps from zero to one at this altitude and remains there up to the upper model boundary.

The 288 organizations belonging to Group (ii) all contain CO but not N, H_2, and O. The increase and slight decrease is caused by the relative concentration profile of CO having the same characteristics. Group (iii) contains 288 organizations as well. They all include CO and N, but not H_2 and O. These are exactly the same organizations as in Group (ii) with species N added to each organization. Here, the decrease of CO is compensated for by further increases in species N.

Organization 1471 (see Table 5.1) is a member of this group.

Group (iv) is constituted by the remaining 893 organizations. Species that have increasing relative concentrations are part of many organizations. For example, H_2 is contained in 324 organizations and N in 742 organizations. Ions are another class of species with increasing relative concentrations at high altitudes. They are also present in many organizations, for example O^+ in 748. Species that have declining relative concentrations in high altitudes are found in much fewer organizations, for example CO_2 in 12 organizations as mentioned, or N_2O in six organizations. Hence, the intensity values of the majority of all organizations increases at high altitudes. The connected Organization 28 (see Figure 5.5 (A)) belongs to this group.

The intensity value for organizations containing $h\nu$ only, $h\nu$ and grain, and $h\nu$ and $(HO_2)_{\text{grain}}$ is constantly zero, since no data for these species is contained in the concentration profile. An exhaustive list with all organizations sorted according to their intensity profile characteristics can be found in Appendix A.4.

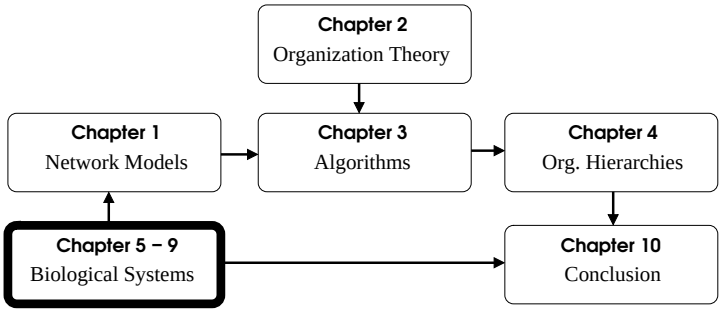
Summing up, we find that the CO_2 containing organizations of Group (i), which roughly coincide with the twelve largest organizations, are very dominant with high intensity values above 60 km (where all species are present), and decline slightly at higher altitudes. Most of the smaller organizations have very low intensity values at low altitudes and become more and more important at higher altitudes, beginning at approximately 140 km. These organizations reach their maximum intensity value close to, or at the upper boundary of the model. The acquired intensity profiles help to identify those organizational subnetworks that prevail at certain altitudes. When studying atmospheric events taking place at specific altitudes, those organizations with high intensity values at this altitude should be in the focus of the investigation.

5.4 Discussion

Applying the theory of chemical organizations to photochemical reaction networks reveals that such networks have an intriguing internal structure. The method identifies this structure which can then be visualized in a Hasse diagram. Organizations, the structural key elements, represent closed and self-maintaining

subnetworks whose instances coincide with potential steady states of the system. While generally, not every organization harbors a potential steady state, for mass-conserving networks such as atmospheric reaction networks, this is true as every organization is balanced. The observed organizational structure breaks down when the network is randomized. Hence, properties of the natural, real system (or more precise: properties of its network model) must give rise to the observed structure. One such property is that all species consist of atoms. Although knowledge about the specific molecular makeup of species is not used, network structure alone is enough to group species of similar molecular structure together in organizations. The connected organizations of the dayside show clearly, how such species groups create the organization hierarchy. The comparison between night- and dayside confirms that incident sunlight is required as a driving force to keep the atmospheric chemistry running. Without light, all organizations consist of unreactive species only. Only few organizations with many species exist. Apparently, only very specific species combinations allow for steady states having a high species diversity. The organization intensity profiles give first hints at which subnetworks play a dominant role at certain heights.

When the evolution of atmospheric atmospheres is considered, the movement from one steady state species composition to another can be interpreted as a movement between organizations. While the movement to a smaller organization can happen spontaneously, an up movement into a larger organization always requires a constructive perturbation. That means, novel species must be injected into the system. Catastrophic events like volcano eruptions releasing massive amounts of SO_2 or comet impacts might cause novel species to appear in the Martian atmosphere, potentially leading to a movement into a larger organization. Also, if a desired atmopsheric composition is to be achieved, the theory of organization helps to determine which species need to be added or removed from the atmosphere to cause the system to move into the desired organization.



Chapter 6

Lambda-Phage Model

In this chapter, a model of the genetic switch of bacteriophage lambda is analyzed and its organizational structure studied. First, a Hybrid Functional Petri Net (HFPN) model is transformed into a reaction network. Then, the organizations are determined. The original model contains inhibitory interactions that are difficult to transform into chemical reactions. They are omitted and their effects are afterwards discussed separately for each organization. A general procedure to deal with inhibitions will later be introduced in Chapter 8. We find two organizations representing the two possible states of the switch: a lysogenic and a lytic organization. Using a HFPN simulator, the persistence of these organizations in time is verified in dynamic simulations.

6.1 The Genetic Switch of Bacteriophage Lambda

Bacteriophage lambda is a well studied virus that infects *E. coli* cells. After injection into the host cell, the phage chooses between two courses of infection: lysogeny or lysis. In the lysogenic phase, the viral DNA is integrated into the host DNA. Thus, the viral DNA will also be replicated in any subsequent cell division. In the lytic phase, the genes of the phage are expressed and its structural proteins are synthesized. New phages assemble in the host cell and spread after the host cell membrane is finally disintegrated. The decision between lysogeny and lysis is based on the state of the host cell. In ailing hosts (indicated by high concentration levels of protein CII), the lysogenic phase is initiated. In a healthy

cell (low levels of CII), the lytic phase is prefered. Characteristic for lysogeny are high levels of protein CI and low levels of protein Cro. For lysis, CI is low and Cro is high. Once the decision is made, feedback loops ensure that the chosen path is followed. Figure 6.1 depicts the simplified mechanism of the genetic switch that is responsible for deciding between the lysogenic and lytic phase. The promoter of genes *cI* (P_{RM}) and *cro* (P_R) share three operator sites: O_{R1}, O_{R2}, and O_{R3}. CI and Cro bind competitively to these sites in their dimerized forms. Due to differing affinities, CI first binds to O_{R1} and O_{R2}, leading to inhibition of *cro* and activation of its own synthesis. This positive feedback loop is complemented by a negative feedback: with high CI concentrations, CI will also bind to O_{R3}, leading to inhibition of *cI*. Gene *cro* is active when nothing is bound at O_{R1}. Cro has highest affinity to operator O_{R3} where it inhibits *cI* on binding. With rising levels of Cro it will also bind to O_{R2} and O_{R1}, leading to inhibition of its own synthesis. Concluding, the positive and negative feedback loops described here ensure that either CI or Cro is exclusively synthesized, corresponding to locking the system into the lysogenic or lytic phase, respectively. These feedback loops and their importance for phage development were investigated in more detail by Thieffry and Thomas (1995).

A more detailed description of the genetic switch can be found in Ptashne (1986). The system has been intensively studied using a variety of different modelling approaches. Continuous models (Shea and Ackers, 1985), models inspired by electric engineering (McAdams and Shapiro, 1995), stochastic models (Arkin et al., 1998), models combining continuous and discrete modelling (Kiehl et al., 2004), qualitative models (Heidtke and Schulze-Kremer, 1998), and Petri net models (Matsuno et al., 2000) have been devised to understand the dynamics of this genetic switch.

6.2 The Reaction Model

Hybrid Functional Petri Nets (HFPNs) have been used to represent and simulate biopathways (Matsuno et al. (2003)). A HFPN consists of three different types of components: places, transitions, and arcs connecting these elements. There are three different types of arcs: normal, inhibitory, and test arcs. Places represent

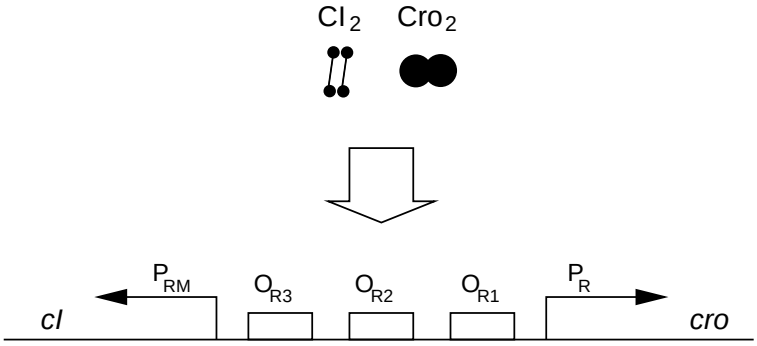


Figure 6.1: The genetic switch of bacteriophage lambda. Promoters P_{RM} and P_R share three operator sites: O_{R1}, O_{R2}, and O_{R3}. The dimerized form of CI and Cro bind competitively to each of the three sites. Positive and negative feedback loops (not shown here) ensure that either CI or Cro is exclusively synthesized, corresponding to the lysogenic and lytic phase.

biological entities (*e.g.,* mRNA and proteins) and can also be used on a more abstract level to model binding states of DNA binding sites or the binding of RNA polymerase at a specific DNA position. Biological reactions and activities are modeled by transitions (*e.g.,* translation, transcription, binding, and decay of proteins). Transitions have input and output connections to places. A transition defines a firing condition that is fulfilled when the concentrations at the input places reach a certain threshold. When firing, the transition transfers tokens from its input places to its output places. Places and transitions can be discrete or continuous. During a simulation, discrete places hold a discrete numbers of tokens, whereas continuous places contain concentration levels. Continuous transitions change the concentrations at their input and output places accordingly. Test and inhibitory arcs do not transfer tokens or concentrations, but are used to change the firing behavior of transitions. Test arcs make transitions fire only if the concentration at the corresponding place is above a threshold, without actually changing the concentration at this place. The inverse is true for inhibitory arcs: the transition can only fire if the concentration at the corresponding place is below a threshold.

We take a HFPN model of bacteriophage lambda created by Doi (2005) and convert it to a chemical reaction network. The HFPN model consists of 56 places

Table 6.1: Structure of the HFPN model of bacteriophage lambda.

Places representing ...	**#**	**Transitions modeling ...**	**#**
mRNA	13	mRNA and protein degradation	26
protein	11	RNA translation	10
position of RNA polymerase	25	unbinding of RNA polymerase	4
binding state of O_{R1}, O_{R2}, O_{R3}	6	(un-) binding of O_{R1}, O_{R2}, O_{R3}	10
UV	1	binding and movement of	
		RNA polymerase and transcription	31
total	**56**	**total**	**81**

and 81 transitions, covering gene regulation including the genetic switch as described in Section 6.1, transcription, and translation. The model was created by Doi (2005) in "Cell Illustrator" (Nagasaki et al., 2004). Its structure is summed up in Table 6.1.

The binding of CI and Cro at the operator sites O_{R1}, O_{R2}, and O_{R3} is modeled by six places: OR1_CI refers to CI bound at O_{R1} and OR1_Cro to Cro at that operator, and so forth. If no tokens are present in all six places, neither CI nor Cro is bound at any operator site. Following three places act as input to the system: promoter $P_{R'}$ is activated unconditionally, whereas P_R is only activated in the absence of OR1_CI and OR1_Cro, and P_L is only activated in the absence of CI and Cro due to inhibitory arcs.

For converting the HFPN model to a chemical reaction network, one reaction is created for each transition. Input places are used as educts and output places as products. Test links are modeled as catalysts: the corresponding place appears as both educt and product. Inhibitory links cannot be converted to chemical reaction equations in a straightforward way (see Chapter 8). They are therefore omitted here and their influence will later be discussed for each organization separately. The place representing exposure to UV light is removed from the model, and a decay reaction for head and tail proteins is added. Promoter $P_{R'}$ is treated as input: $\emptyset \rightarrow P_{R'}$. The other two input species can be inhibited: P_R by OR1_CI and OR1_Cro, and P_L by CI and Cro. Hence, they cannot be

modeled as regular input species. Instead, they are modeled as self-replicators: $P_L \rightarrow 2\ P_L, P_R \rightarrow 2\ P_R$. Whether they should be present or not in a specific situation depends on the presence of the species that inhibit them. This will be discussed in detail in Section 6.4.

The resulting chemical reaction network consists of 55 molecular species and 81 reactions (see Appendix B for a list of all reactions).

6.3 Hierarchy of Organizations

The reaction network modeling the bacteriophage lambda contains seven organizations as depicted in Figure 6.2. The used species set labels are detailed in Table 6.2. The smallest organization just contains the input species PR', representing the unconditionally activated promoter $P_{R'}$ (Org. 0). The two organizations next in size contain the feedback mechanisms of CI (Org. 1) and Cro (Org. 2). These organizations correspond to the lysogenic and lytic phase of the phage. Their reaction networks are shown in Figure 6.3. Organization 3 contains all species of Orgs. 1 and 2. There are no interactions between these two organizations that create novel species when merged. Organization 4 is next in size and contains the expression of all genes controlled by promoter P_L: N, CIII, Xis, and Int. The union of Org. 4 with Org. 1 gives Org. 5. Again, the merger does not lead to the creation of new species. Finally, Org. 6 is the largest organization and contains all molecular species of the network. This organization comprises the whole reaction network.

6.4 Inhibitory Interactions

Neglecting inhibitory interactions, the analysis has revealed seven sets of molecular species that are closed and self-maintaining within the reaction network. How can this result change if we wish to include the inhibitory interactions in our considerations? First, the existence of inhibitions within an organizational network can lead to the violation of the self-maintenance property of the corresponding set of molecular species. An inhibition could limit the fluxes of reactions that create an organizational species, such that its maintenance is no longer feasible.

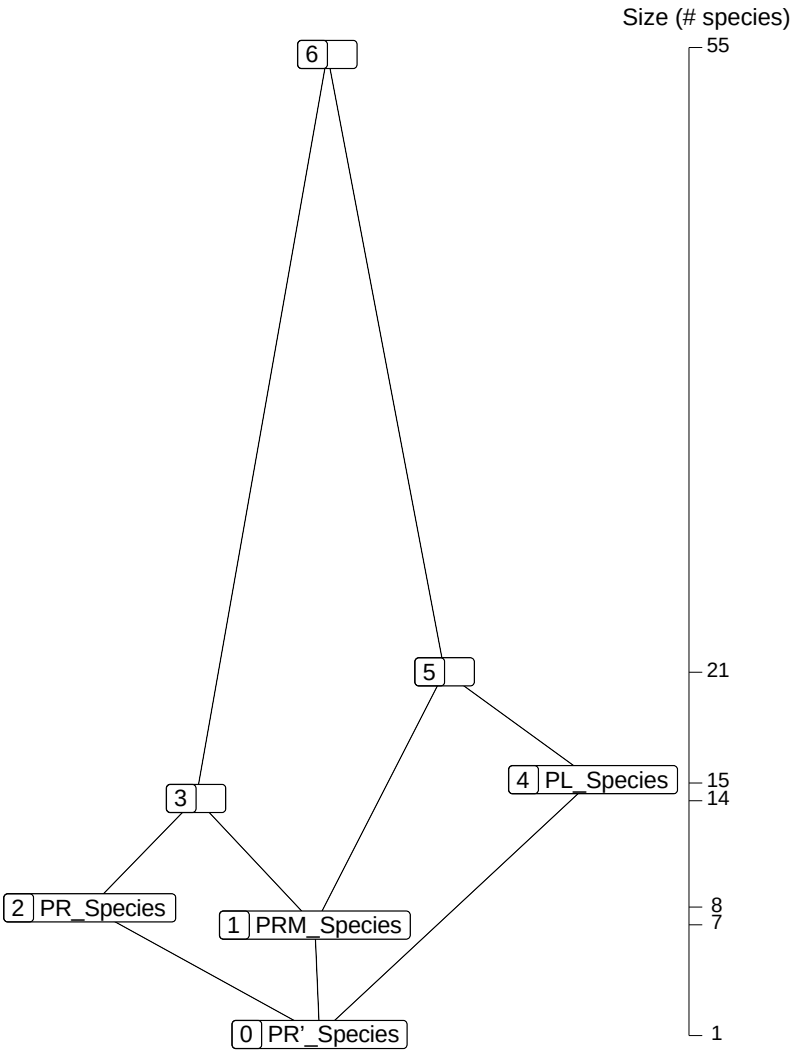


Figure 6.2: Hasse diagram of seven organizations of the lambda model. Organization 1 contains the CI feedback loop and corresponds to the lysogenic phase. Organization 2 contains the Cro feedback loop and corresponds to the lytic phase. See Figure 6.3 for the reaction networks of these organizations. The largest Org. 6 contains the complete network. Labels only detail additional species not yet contained in organizations to which a downlink exists. See Table 6.2 for species set labels.

Second, we modeled two input species as self-replicators in our model as they can be inhibited by other species. In case the input species are considered as absent, the disappearance of their inhibiting species leads to the creation of the

Table 6.2: For a compact representation, the Hasse diagram in Figure 6.2 does not list the species in organizations individually. Instead, four sets of species as defined here are used.

PL_Species	:=	{PL, N_mRNA, N, P_m16, P_m7, cIII_mRNA, CIII, P_m17, P_m12, xis_mRNA, Xis, P_m18, P_m13, int_mRNA_m14}
PR_Species	:=	{PR, cro_mRNA, Cro, P_m28, OR1_Cro, OR2_Cro, OR3_Cro}
PRM_Species	:=	{PRM, cI_mRNA, CI, OR1_CI, OR2_CI, OR3_CI}
PR'_Species	:=	{PR'}

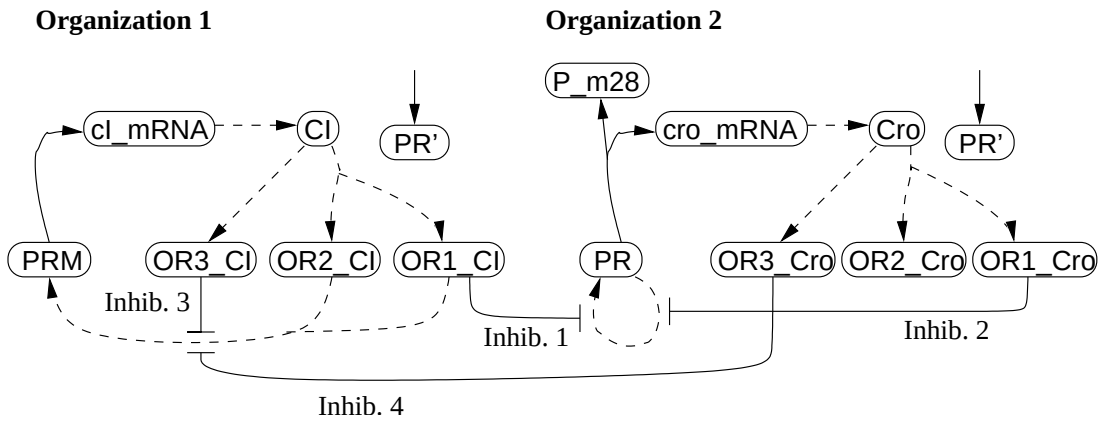


Figure 6.3: Organizational reaction networks belonging to Orgs. 1 and 2. Dotted arrows mark catalytic reactions. For clarity, decay reactions are omitted. All species except P_{RM} and P_R decay. Organization 1 corresponds to the lysogenic phase, Org. 2 to the lytic phase. Shown inhibitions are not part of the analyzed model, but part of the original HFPN model. Both feedback loops inhibit each other by Inhibitions 1 and 4.

input species. Hence, novel species are created, violating the closure property. Consequently, it is necessary to examine each organization separately with respect to these possible effects of inhibitory interactions. For the lambda model, we will find that only two of the seven organizations remain when inhibition is

considered[1].

Following six inhibitory interactions must be considered: 1. OR1_CI inhibits input P_R, 2. OR1_Cro inhibits input P_R, 3. OR3_CI inhibits P_{RM}, 4. OR3_Cro inhibits P_{RM}, 5. CI inhibits input P_L, and 6. Cro inhibits input P_L. See Table 6.3 for a list of inhibitions and organizations in which they are active. The smallest Org. 0 just contains the unconditional input species $P_{R'}$. The species that inhibit the other two input species P_L and P_R are not present in the organization. As these other two inputs depend on the absence of inhibitor species, they were not modeled as unconditional input but as self-replicators. However, if only the species of Org. 0 is present, the absence of the inhibitors leads to the production of the input species P_L and P_R. Hence, if inhibition (or rather the lack of inhibition) is considered, the system does not stay in Org. 0. New molecular species appear, corresponding to an up-movement in the hierarchy of organizations. Organizations 1 and 2 each contain one inhibitory interaction (Inhibition 3, resp. 2). The corresponding reaction networks are shown in Figure 6.3. In both cases positive feedback loops are coupled with additional negative feedback. The inhibitions do not lead to the disappearance of molecular species, rather they stabilize the concentrations of CI and Cro, respectively. The property of self-maintenance still holds true for both organizations. Organization 3 combines these organizations. Besides the two inhibitions mentioned above, Inhibitions 1 and 4 are present in this organization leading to a mutual inhibition of the CI and Cro feedback loops (*cf.* Figure 6.3). Due to this mutual inhibition, it is likely that one loop suppresses the other sooner or later, leading to the extinction of the molecular species of the suppressed loop. This corresponds to a down-movement in the hierarchy of organizations. No inhibitions are present in Org. 4. However, the species inhibiting input P_R, which was modeled as a self-replicator, are also not present. Consequently, if this inhibitory interaction (resp. its absence) is considered, we find that P_R will be produced. The system cannot stay in Org. 4 as P_R is created as a novel species, resulting in an up-movement. Organization 5 contains two inhibitions: Inhibitions 3 and 5. Inhibition of P_L by CI is likely to

[1]Note that the inclusion of inhibitory interactions can generally also lead to the creation of new organizations: for example, the decay of a species might be inhibited, making its maintenance feasible. However, this does not play a role in this lambda model.

Table 6.3: Inhibitory interactions and their presence in organizations.

Id	Inhibition
1	OR1_CI inhibits P_R
2	OR1_Cro inhibits P_R
3	OR3_CI inhibits P_{RM}
4	OR3_Cro inhibits P_{RM}
5	CI inhibits P_L
6	Cro inhibits P_L

	Inhibition Id					
Org.	**1**	**2**	**3**	**4**	**5**	**6**
0						
1			X			
2		X				
3	X	X	X	X		
4						
5			X		X	
6	X	X	X	X	X	X

Table 6.4: Movement in the hierarchy of organizations if inhibitory interactions are considered.

Initial State	Movement
Org. 0	up-movement to Org. 6
Org. 1	-
Org. 2	-
Org. 3	down-movement to Org. 1 or 2
Org. 4	up-movement to Org. 6
Org. 5	down-movement to Org. 1
Org. 6	down-movement to Org. 1, 2, or 3

remove molecular species P_L, resulting in a down-movement. The largest Org. 6 containing the complete network contains all six inhibitions. A down-movement is likely to occur with the same argument as for Org. 3.

Table 9 summarizes our findings. When taking inhibitory interactions into account, only two organizations remain: Organization 1 and 2. Interestingly, exactly these correspond to the lysogenic and lytic phase, the two potential behavior patterns of bacteriophage lambda.

6.5 Dynamic Simulation

Considering network structure and stoichiometric information only, the analysis of the lambda reaction network has revealed two organizations that are possibly capable of dynamic permanence. To test whether they are indeed able to prevail in time or not, dynamic simulations of the original HFPN model using the provided kinetics were performed using the simulation software "Cell Illustrator" (Nagasaki et al., 2004). To simulate the lytic phase, all places were initialized with zero. For the lysogenic phase, all places were set to zero except for CII, which was set to 250. Tokens are generated in the network by input transitions that create tokens in the places of the three promoters $P_{R'}$, P_R, and P_L. The trajectories for both scenarios are depicted in Figure 6.4. After a short transient phase, the system settles in a periodic pattern in both cases. The periodic behavior is caused by the competing activation and inhibition of the promoter P_{RM} in Org. 1 and P_R in Org. 2 by the feedback mechanisms as discussed in Section 6.1. In the periodic phase, only a very small part of the network is active. The active places correspond exactly to the species of Org. 1 for the lysogenic phase and to the species of Org. 2 for the lytic phase. The concentrations at all other places are permanently less equal 1 token for discrete places and less equal 1 arbitrary concentration unit for continuous places. Since transitions only fire if the concentrations at their input places are greater than 1, the corresponding part of the network is inactive. This shows that both organizations indeed are able to persist over time in the model. Surprisingly, the head_tail proteins required for phage assembly are never synthesized in the lytic phase in which the virus proliferates. The concentration of protein Q does not reach sufficient levels to enable synthesis of this protein from promoter $P_{R'}$. Hence, the model fails to simulate the last stage of phage proliferation.

6.6 Discussion

Analyzing a reaction network model of bacteriophage lambda derived from a HFPN model revealed a relative simple hiearchy of seven organizations. When carefully considering the effects of inhibitory interactions which had to be omitted

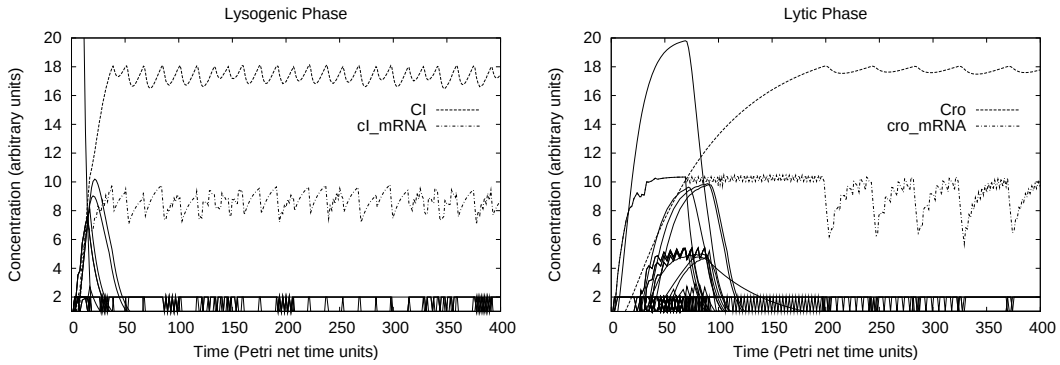
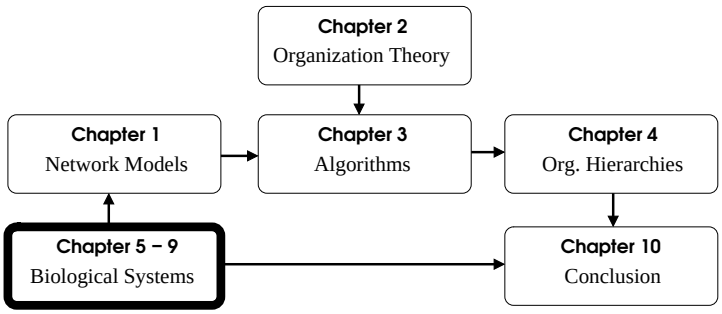


Figure 6.4: Typical simulation runs of the original HFPN model for the lysogenic phase (left) and lytic phase (right) of bacteriophage lambda using the "Cell Illustrator" software package. After a transient phase, a periodic pattern emerges. The active species in the lysogenic simulation run are identical to the species of Org. 1, those active in the lytic run are identical to Org. 2.

during the conversion to a reaction network, only two organizations remain that have the potential to persist in time. That these organization indeed can prevail in time could be verified in simulations. The two organizations coincide with the two potential behavior patterns of bacteriophage lambda: lysogeny and lysis. However, both organizations only contain the feedback loops that lock the system into the respective phases. Other species that are expected to be present in the lytic phase, for example protein Q and the head_tail proteins necessary for phage assembly, are not part of the lytic organization. The antiterminator protein N is required for the synthesis of Q and further proteins related to lysogeny and lysis. However, N is not part of both organizations as both CI in lysogeny and Cro in lysis inhibit its synthesis, and inhibition is considered as completely shutting the synthesis down, independent of the inhibitor concentration. In the HFPN model, inhibition is depending on the inhibitor concentration, and indeed, Q is temporarily synthesized in the transient phase of the model simulation, but not in the final periodic phase.

Although quite encouraging, the detection of organizations corresponding to the lysogenic and lytic phase of bacteriophage lambda might be trivial considering that the studied model was intended to exhibit these two behavior patterns. The

question arises, what can be said about the real system? Models covering the complete life-cycle, including host replication and further interactions between phage and host are highly desirable to address this question. Inhibitory interactions are common in biology, but difficult to be realized in reaction networks. However, in this relatively small network their effects could be assessed "by hand". A general framework to deal with inhibitions will be presented in Chapter 8.



Chapter 7

Central Sugar Metabolism of Escherichia coli

Puchalka and Kierzek (2004) constructed a model of the central sugar metabolism of *E. coli* including gene expression, signal transduction, and enzymatic activities. The model is able to reproduce the preferential and exclusive uptake of glucose by the organism in the presence of other sugars, known as diauxic shift. Puchalka and Kierzek (2004) used this model to test a new stochastic simulation method called the "maximal timestep method", which can deal with processes depending on very large and very small numbers of molecules at the same time. As this model integrates different molecular networks that are at play in a cell (*cf.* Section 1.1), it is an interesting candidate for analysis using the theory of chemical organizations.

In the resulting hierarchy of organizations, some organizations represent states corresponding to growth and some correspond to starvation states. The inducible pathways responsible for lactose and glycerol uptake can clearly be identified in the organizational hierarchy. They appear as the difference between organizations representing starvation and growth on lactose and glycerol, respectively.

7.1 The Sugar Metabolism of E. coli

Within the central sugar metabolism of *E. coli*, external sugars are taken up by the cell and transformed into pyruvate which is fed into further metabolic processes downstream. If several sugars are available in the growth medium, *E. coli* first

exclusively metabolizes its preferred carbon source glucose. Only after depletion of glucose, the bacterium will begin to utilize other available sugars. This diauxic growth phenomenon has been extensively studied in experiments and by mathematical modeling (Kremling et al., 2001; Monod, 1942; Thattai and Shraiman, 2003; Wang et al., 2001), leading to a good understanding of the molecular mechanisms at work. The two main mechanisms facilitating the switch-like behavior are inducer exclusion and catabolite repression. Briefly, if external glucose is available, the glucose uptake by the phosphotransferase system (PTS) leads to the dephosphorylation of enzyme EIIA. In the unphosphorylated form, EIIA binds to the lactose permease lacY, inhibiting it. Hence, lactose cannot enter the cell. Intracellular lactose induces, after transformation to allolactose, the genes necessary for metabolizing lactose. Consequently, while glucose is metabolized the lactose system is not induced, since the inducer cannot enter the cell. Catabolite repression refers to the fact that cyclic adenosine monophosphate (cAMP) only accumulates in the cell in the absence of external glucose. Then, cAMP forms a complex with the catabolite repressor protein (Crp) and Crp-cAMP substantially increases the transcription rate of genes encoding alternative sugar uptake systems. In the presence of external glucose, cAMP levels are low and the alternative sugar genes are not expressed. See the referenced literature for more details on these mechanisms.

Extending models by Kremling et al. (2001) and Wang et al. (2001), Puchalka and Kierzek (2004) constructed a reaction network modeling the sugar metabolism of *E. coli* including gene expression, signal transduction, and transport and enzymatic activities. We take this network as an example to demonstrate how the theory of organizations can be applied to intracellular networks of non-trivial size. First, the network is adapted as described in the next section. Then, organizations are analyzed for several scenarios representing bacterial growth on different sugar sources.

7.2 Reaction Network Model

The original network by Puchalka and Kierzek (2004) consists of 92 substances reacting with each other in 120 reactions. The model contains reactions modeling

transcription and translation of 21 genes. The uptake and utilization of external glucose, lactose, and glycerol is included in the model as well as catabolic repression and inducer exclusion, allowing the model to exhibit diauxic growth. Each reaction of the network consists of up to three different types of species: educts, products, and modifiers. If a reaction occurs, the educt species are transformed into the product species while the modifiers are not affected:

$$educts \overset{modifiers}{\longrightarrow} products \tag{7.1}$$

Modifier species only change the reaction rate. Two types of modifiers are used in the model: *enzymes*, that are required for a reaction to take place (the rate equation is a product with the enzyme concentration as one factor), and *effectors*, which increase the reaction rate acting as an activator, or decrease the reaction rate acting as an inhibitor or repressor. As modifiers cannot be directly specified in reaction networks, they have to be handled separately as follows. If a reaction does not have modifiers, we take the reaction exactly as it is. The original model contains six reactions that are reversible. An explicit back reaction for each of them is added in our model. In the presence of modifiers, we inspect the reaction rate equation. In case the modifier species concentration has to be greater zero for the reaction rate to become greater zero, we add the modifier species on both educt and product side of the reaction. This is the typical case for enzymes. Only in their presence, the reaction in question can be performed. If the reaction rate can be greater zero even in the absence of the modifier species, we simply ignore them, as they are not necessary for the reaction to take place. They merely increase (or decrease) the reaction rate, acting as nonessential activators (resp. repressors or inhibitors). It is important to note that all inhibitory or negative interactions are ignored by this procedure.

7.2.1 Modeling Gene Expression

The handling of modifiers as described above cannot be applied to reactions modeling gene expression. Negative interactions can be ignored as before, but activators need special treatment. The model contains five transcription reactions that have activating and/or repressing effectors, as detailed in Table 7.1.

Table 7.1: Transcription reactions with activating (+) and/or repressing (-) effectors. For species abbreviations, see Appendix C.

Reaction	**Effectors**
1. RNAP + PromCrp →Tscription + PromCrp + CrpmRNA	Crp(∓), cAMP(∓)
2. RNAP + PromCya →Tscription + PromCya + CyamRNA	Crp(-), cAMP(-)
3. RNAP + PromLacZY →Tscription + PromLacZY + LacZYmRNA	Crp(+), cAMP(+), LacI(-), Allo(+)
4. RNAP + PromGlpFK →Tscription + PromGlpFK + GlpFKmRNA	Crp(+), cAMP(+), GlpR(-), G3P(+)
5. RNAP + PromGlpD →Tscription + PromGlpD + GlpDmRNA	Crp(+), cAMP(+), GlpR(-), G3P(+)

With activator concentrations being zero, the transcription reaction rates in the original model are still positive but small. This corresponds to the basal transcription rate of a gene: even if activators are not present RNA polymerase occasionally binds to the promoter and transcription is initiated, leading to a basal concentration of the respective protein. Applying the procedure as described above to these reactions (*i.e.*, ignoring all activators) leads to an unconditional transcription of all genes, representing the basal activity. But as shown below for the transcription of the *lac* genes, basal concentration of proteins is not sufficient to perform certain metabolic tasks. Consequently, a protein having only basal concentration should be regarded as not being present in our analysis. Only if activators are present increasing the transcription rate so that protein concentrations reach levels significantly above basal level – effectively switching the gene on – the corresponding protein should be regarded as being present. Activators and inducers for gene transcription should therefore be modeled as necessary catalysts in gene transcription reactions. The five transcription reactions having effectors are discussed separately.

1. Transcription of crp: *effectors Crp, cAMP.* Crp is activated by the binding of cAMP. The activated Crp–cAMP complex negatively regulates the transcription of *crp*. It was shown by Hanamura and Aiba (1992) that with further increasing concentration of Crp–cAMP this inhibition is overcome and an upregulation occurs. The inhibition is ignored and since the activation only occurs at high concentrations, it is ignored as well (since the reaction can take place in the

absence of the effector species). Hence the effectors Crp and cAMP are ignored for this transcription reaction and we model the reaction as:

$$RNAP + PromCrp \rightarrow Tscription + PromCrp + CrpmRNA. \tag{7.2}$$

2. Transcription of cya: *effectors Crp, cAMP.* Crp–cAMP downregulates transcription of *cya*. Being an inhibition, the effector species Crp and cAMP are ignored for this reaction. The reaction reads:

$$RNAP + PromCya \rightarrow Tscription + PromCya + CyamRNA. \tag{7.3}$$

3-5. Transcription of lacZY, glpFK, and glpD: *effectors Crp, cAMP, LacI/GlpR, and allolactose (Allo)/glycerol-3-phosphate (G3P).* These genes code for enzymes necessary for lactose and glycerol uptake and utilization. The transcription regulation is similar for both sugars. Two mechanisms are at work for transcription regulation of *lacZY* (*glpFK*, *glpD*). Firstly, repressor LacI (GlpR) represses transcription. If inducer Allo (G3P) is present, it binds to LacI (GlpR) and by this inactivates the repressor, enabling transcription. Secondly, Crp–cAMP complex acts as an activator. Both mechanisms are modeled in one reaction equation in the model (see above, reactions 3-5). We ignore the inhibiting effect of effector species LacI (GlpR). Instead, by adding the inducer Allo (G3P) on both educt and product side of the reaction, we require the inducer to be present for transcription. This is in accordance with biological knowledge: only in the presence of the inducer, the corresponding gene products are synthesized at above basal concentration levels. Mutants not able to synthesize Crp or cAMP were found unable to grow on several carbon sources (Botsford and Harman, 1992; Postma et al., 1993). Hence we require the presence of Crp and cAMP for the synthesis of enzymes necessary for carbon uptake and utilization. Accordingly, effectors Crp and cAMP are also added on both educt and product side of the reactions. The reactions in our model are:

$$\begin{aligned} &RNAP + PromLacZY + Allo + Crp + cAMP \rightarrow \\ &\quad Tscription + PromLacZY + LacZYmRNA + Allo + Crp + cAMP \end{aligned} \tag{7.4}$$

$$RNAP + PromGlpFK + G3P + Crp + cAMP \rightarrow$$
$$Tscription + PromGlpFK + GlpFKmRNA + G3P + Crp + cAMP \quad (7.5)$$

$$RNAP + PromGlpD + G3P + Crp + cAMP \rightarrow$$
$$Tscription + PromGlpD + GlpDmRNA + G3P + Crp + cAMP \quad (7.6)$$

7.2.2 Transcription Rates for lac and glp in the Model

In order to (computer-) experimentally test whether the mentioned effectors are necessary for expression of the *lac* and *glp* genes or not in the model, we compute the reaction rates for gene expression under different effector concentrations. Using the standard simulation settings from Puchalka and Kierzek (2004) and soley supplying external glucose (2×10^{12} molecules) yields typical concentrations for all involved species in the uninduced condition. To derive typical concentrations for the induced states, solely external lactose (5×10^{11} molecules) or external glycerol (2×10^{12} molecules) is supplied. By computing the gene transcription reaction rate using a mixture of uninduced and induced concentrations for the different effectors, the influence of each single effector on the overall gene transcription can be estimated. Table 7.2 summarizes the results for the induction of the *lac* genes and Table 7.3 sums up the results for induction of the *glp* genes. The induction of the *lac* genes corresponds to a thousandfold increase of the reaction rate. The concentration level of Crp and LacI is the same for the uninduced and the induced state. If, starting from the uninduced state, only the concentration of cAMP (Allo) is raised to the induced concentration level, only an approximately 36-fold (34-fold) increase in the reaction rate can be achieved. Hence, it is reasonable to require in our model cAMP and Allo both to be present for the transcription reaction to be performed. If taking the induced state and setting the concentration of Crp to zero, the reaction rate becomes smaller even than in the uninduced state. Hence, also Crp is required for the *lac* genes to be induced.

The results for the glycerol uptake system is not as clear as for the lactose system (Table 7.3). The reaction rate in the induced state is only approximately 80-fold higher than in the uninduced state. Levels of Crp and GlpR are the same for the uninduced and induced state. The induction solely depends on cAMP.

Table 7.2: Induction of the *lac* genes and dependency of the transcription rate on the effectors Crp, cAMP, Allo, and LacI. For five different concentration combinations of the effector species, the rate of the transcription reaction and the relative rate compared to the uninduced state is shown. Uninduced and induced states are shown in (1) and (2). In (3) and (4), only the concentration of cAMP, respectively Allo, is set to the induced level, while keeping the other concentrations at the uninduced level. In (5), the induced state is modified by setting the Crp concentration to zero. The concentration of the promoter is set to one copy and that of RNAP to 60 copies.

State	**Crp**	**cAMP**	**Allo**	**LacI**	**trans. rate**	**rel. to (1)**
(1) uninduced	1500	400	0	15	2.962×10^{-5}	1.0
(2) induced	1500	2500	1400	15	3.412×10^{-2}	1152
(3) cAMP ind.	1500	2500	0	15	1.086×10^{-3}	36.66
(4) Allo ind.	1500	400	1400	15	1.015×10^{-3}	34.27
(5) ind. w/o Crp	0	2500	1400	15	6.757×10^{-6}	0.2281

Only setting the concentration of cAMP to the induced level is sufficient for obtaining the transcription rate of the induced state. The concentration of G3P rises from zero in the uninduced state to three molecules in the induced state. This increase does not have a measurable effect on the transcription rate in the model. If starting from the induced state the concentration of Crp is set zero, the transcritpion rate again becomes smaller even than for the uninduced state. We see that in the model, Crp and cAMP is required for induction of the *glp* genes while the inducer G3P does not play a significant role. The difference in reaction rate speed for the uninduced and induced state is not as pronounced as for the *lac* system. However, we still model Crp, cAMP, and G3P as required for transcription, as G3P removes DNA-bound repressor GlpR (Larson et al., 1987) and is the exclusive inducer for *glp* (Lin, 1976).

7.2.3 Modeling Growth and Defining Input

Cell growth and cell division is accounted for in the original model by dividing all species concentrations by two on cell division, except for the DNA species.

Table 7.3: Induction of the *glp* genes and dependency of the transcription rate on the effectors Crp, cAMP, G3P, and GlpR. For five different concentration combinations of the effector species, the rate of the *glpFK* transcription reaction and the relative rate compared to the uninduced state is shown. The reaction rate for *glpD* transcription is four times smaller than for *glpFK* transcription. Uninduced and induced states are shown in (1) and (2). In (3) and (4), only the concentration of cAMP, respectively G3P, is set to the induced level, while keeping the other concentrations at the uninduced level. In (5), the induced state is modified by setting the Crp concentration to zero. The concentration of the promoter is set to one copy and that of RNAP to 60 copies.

State	**Crp**	**cAMP**	**G3P**	**GlpR**	**trans. rate**	**rel. to (1)**
(1) uninduced	1500	400	0	70	3.938×10^{-4}	1.0
(2) induced	1500	4000	3	70	3.117×10^{-2}	79.15
(3) cAMP ind.	1500	4000	0	70	3.117×10^{-2}	79.15
(4) G3P ind.	1500	400	3	70	3.938×10^{-4}	1.0
(5) ind. w/o Crp	0	4000	3	70	1.028×10^{-5}	0.02610

Hence, decay reactions for all non-DNA species that do not already decay in the original model are added. The remaining species that do not decay are: all 21 promoter species, RNAP, Tscription, Glcex, Lacex, and Glyex. Several species are not produced from within the original network model. Among them are all 21 promoter species, ATP, ADP, and AMP. Assuming that they are nevertheless present in the cell at all times, they are provided as external input. We add a reaction of the form

$$\emptyset \rightarrow \text{Species} \tag{7.7}$$

for each of them. Additionally, RNAP is provided as input. Finally, our network model consists of 92 species and 168 reactions. See Appendix C for a complete list of species and reactions. Glucose, lactose, and glycerol in the growth medium are represented by the species Glcex, Lacex, and Glyex. By adding additional input reactions for these species, growth on different sugar sources can be modeled.

7.3 Hierarchies of Organizations

We compute the hierarchy of organizations of the network for five different scenarios. The scenarios only differ in which external sugars are supplied as input, resembling bacterial growth on different sugar sources. First, no external sugars are supplied at all. Then, one of the three sugars glucose, lactose, and glycerol is consecutively supplied as the exclusive carbon source. And finally, all three sugars are provided simultaneously. Supplying a sugar source is accomplished simply by adding an input reaction of the form $\emptyset \to$ Sugarex to the reaction network. Changing the reaction network also changes the hierarchy of organizations. The resulting hierachies are depicted in Figure 7.1. They all consist of four organizations. The labels within organizations refer to sets of species as detailed in Table 7.4.

The network model covers the transformation of external sugar into pyruvate, which is then fed into further metabolic processes not considered by the model. These follow-up processes enabling cellular survival are represented by pseudo species Metabolism. Species set Metabolites contains all relevant species of this pathway and its presence in an organization hence represents a cell being able to maintain its metabolism and grow.

Starvation. No external sugars are supplied as input. The resulting hierarchy of organizations is depicted in Figure 7.1(a). The smallest Org. 1 contains all input species (21 promoter species, ATP, ADP, AMP, and RNAP). In the presence of the promoters and RNA polymerase, all unregulated genes are transcribed and translated, so that all mRNA and protein species of all 18 unregulated genes are additionally contained in the smallest organization (*cf.* Genes+Enzymes, Table 7.4). Organizations 2 and 3 contain all species of Org. 1 and additionally Glyex and Lacex, respectively. This seems surprising since these species are not supplied as input in this scenario. But recall that an organization is a set of species that is algebraically closed and self-maintaining. Although the species Glyex and Lacex are not supplied as input, they are still a regular part of the reaction network. Inspecting the networks making up Orgs. 2 and 3 reveals that Glyex and Lacex do not participate in any reaction. They are isolated nodes

Table 7.4: Sets of species as used in Figure 7.1.

Genes+Enzymes	:=	{PromCrp, PromCya, PromEIIA, PromEIIBC, PromEI, PromFbp, PromFda, PromGap, Prom GlcT, PromGlk, PromGlpD, PromGlpFK, PromGlpR, PromGpm, PromHPr, PromLacI, PromLacZY, PromPfk, PromPgi, PromPyk, PromTpi, RNAP, Tscription, CrpmRNA, CyamRNA, EIIAmRNA, EIIBCmRNA, EImRNA, FbpmRNA, FdamRNA, GapmRNA, GlcTmRNA, GlkmRNA, GlpRmRNA, GpmmRNA, HPrmRNA, LacImRNA, PfkmRNA, PgimRNA, PykmRNA, TpimRNA, Crp, Cya, EIIA, EIIBC, EI, Fbp, Fda, Gap, GlcT, Glk, GlpR, Gpm, HPr, LacI, Pfk, Pgi, Pyk, Tpi, AMP, ATP, ADP, cAMP}
Metabolites	:=	{Glc, Glc6P, Fru6P, FBP, DHAP, T3P, 3PG, PEP, Pyr, Metabolism, EIIAP, HPrP}
Metabolites*	:=	Metabolites\{Glc}
Glcex	:=	{Glcex}
Lacex	:=	{Lacex}
Glyex	:=	{Glyex}
LacSpecies	:=	{Lac, Allo, LacZYmRNA, LacZYmRNA1, LacZ, LacY}
GlySpecies	:=	{Gly, G3P, GlpDmRNA, GlpFKmRNA, GlpFKmRNA1, GlpD, GlpF, GlpK}

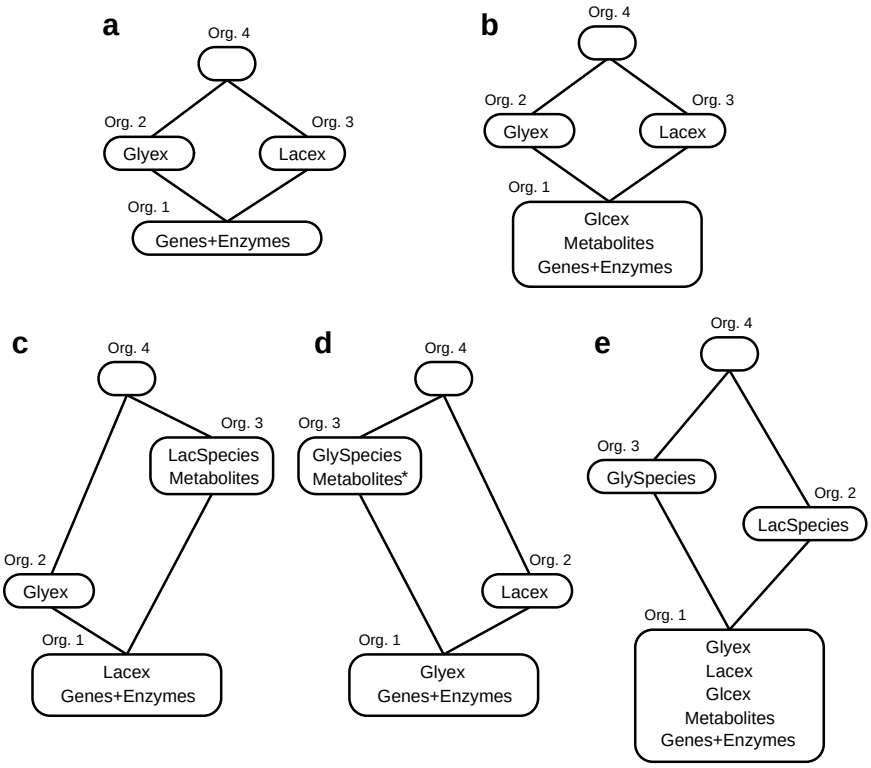


Figure 7.1: Hierarchies of organizations of the *E. coli* network for five scenarios differing in supplied external sugars, resembling growth on different carbon sources. Organizations consist of the species sets contained in their lower organization(s) plus the species set(s) denoted in their label. Species set labels are detailed in Table 7.4. (**a**) starvation; (**b**) growth on glucose only; (**c**) growth on lactose only; (**d**) growth on glycerol only; (**e**) growth on glucose, lactose, and glycerol. See text for details.

in the subnetworks of Orgs. 2 and 3. As such, they do not decay, neither are produced, fulfilling the requirements of closure and self-maintenance. The two organizations represent a state in which a fixed amount of Glyex, respectively Lacex entered the system "by accident" and the uptake systems are not induced. In this case, the concentration of the external sugars will not change. For the real system, this state is not a steady state. In a transient phase, the uptake systems would be induced and the external sugars would be used up completely. Finally, the system would reach the steady state corresponding to Org. 1. The

largest Org. 4 combines Orgs. 2 and 3. All species of the smallest organization, and Glyex and Lacex are contained. In this scenario, we find no organization containing the metabolites of the network. This indicates that with no external sugar source, the network cannot sustain its metabolism, which means that the cell is starving.

Growth on glucose. After adding the reaction $\emptyset \rightarrow$ Glcex, the hierarchy of organizations contains again four organizations as shown in Figure 7.1(b). The smallest Org. 1 contains all unregulated genes and enzymes as in the first scenario and additionally Glcex. With Glcex present, all metabolites can be created and maintained. Consequently, all these species are part of Org. 1, too. With species set Metabolism present in the smallest organization, the cell can maintain its metabolism when external glucose is supplied. The remaining part of the organization hierarchy is equivalent to the first scenario without any sugar input.

Growth on lactose. When lactose is supplied as the exclusive external sugar source, the resulting hierarchy of organizations again contains four organizations as depicted in Figure 7.1(c). The smallest organization contains all unregulated genes and enzymes and additionally Lacex. In Org. 2, only Glyex is added as in the previous cases. Organization 3 contains the species of the smallest organization, all species necessary for taking up and metabolizing external lactose, and the species belonging to the metabolism. Being an organization, the network made up by all these species is algebraically closed and self-maintaining, representing a cell that has switched its *lac* genes on and utilizes external lactose. Figure 7.2(a) details schematically, how Org. 1 is expanded to form Org. 3. Once inducer allolactose is present, the *lac* genes are switched on and LacY and LacZ are synthesized. LacY facilitates the uptake of external lactose while LacZ transforms intracellular lactose and allolactose to glucose and glucose–6–phosphate. Additionally, LacZ transforms lactose to allolactose, closing a positive feedback loop. Glucose then enters the metabolic pathway leading to pyruvate and further metabolic processes. Adding Glyex to Org. 3 results in the largest organization Org. 4. This scenario shows that bacterial growth is possible on lactose as the

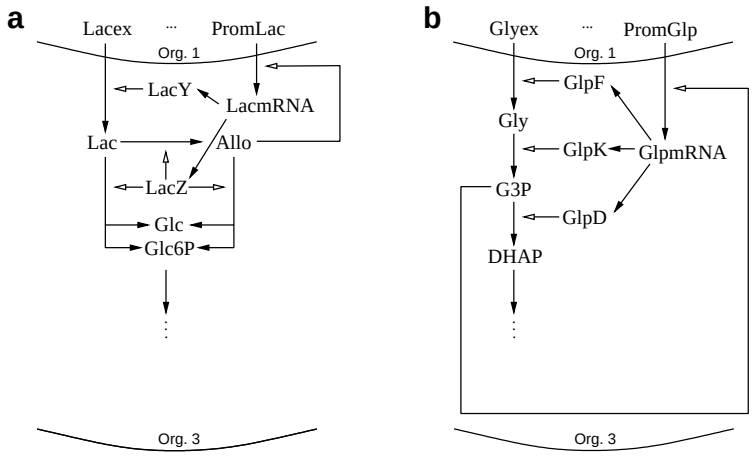


Figure 7.2: Induction of sugar uptake systems. When lactose or glycerol is the exclusive carbon source, Org. 1 corresponds to the state in which the respective uptake systems are not activated and the bacterium is starving (upper part). In Org. 3, the systems are induced and the external sugar is utilized. A schematic sketch of the reaction network of Org. 3 responsible for utilization of (**a**) external lactose and (**b**) external glycerol is shown. Open arrows point from species acting as catalysts to the reactions that are catalyzed. See text for details.

only carbon source after induction of the lactose uptake system (in Orgs. 3 and 4).

Growth on glycerol. Now glycerol is provided as the exclusive carbon source. The resulting hierarchy of organizations is visualized in Figure 7.1(d). The result is equivalent to the lactose scenario. The smallest Org. 1 contains the unconditionally transcribed genes and resulting enzymes, and external glycerol. Organization 3 additionally contains the molecular species necessary for utilizing external glycerol and the metabolism species. Figure 7.2(b) shows, how this organization is formed by expanding Org. 1. Once inducer G3P is present, the genes corresponding to glycerol utilization are switched on and GlpF, GlpK, and GlpD are synthesized. GlpF then enables uptake of external glycerol, GlpK transforms internal glycerol to G3P closing a positive feedback loop, and GlpD transforms G3P to DHAP which in turn fuels the pathway ending in pyruvate and further metabolic processes. Adding Lacex to this organization leads to the largest Org. 4.

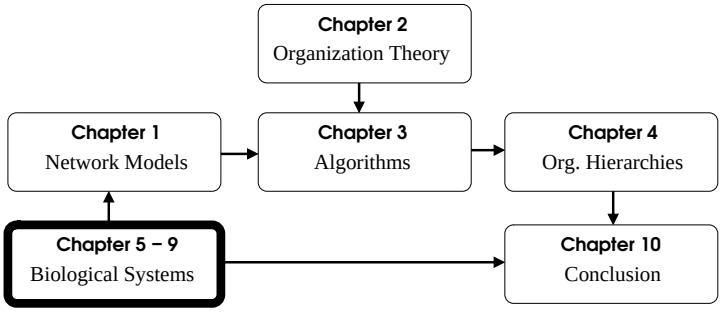
Again we find that once the uptake system for the external sugar is induced, the cell can maintain its metabolism in Orgs. 3 and 4.

Growth on all sugars. In the last scenario, all three external sugars are supplied as input simultaneously. Figure 7.1(e) depicts the resulting hierarchy of organizations. With external glucose being input, the smallest organization resembles the smallest organization of the glucose scenario, with external lactose und glycerol added. Glucose alone is sufficient for growth, hence the smallest organization already represents a state in which the cell grows (on glucose). The two organizations above the smallest one contain the species necessary for utilizing lactose (Org. 2) and glycerol (Org. 3). They represent states in which the cell metabolizes lactose, respectively glycerol, in addition to glucose. The largest Org. 4 finally merges Orgs. 2 and 3, containing all species of the model. Here, all three sugars are metabolized simultaneously. From a biological point of view, only Org. 1 is meaningful since the uptake of lactose and glycerol is repressed in the presence of glucose. The existence of the remaining organizations will be discussed in the next section.

7.4 Discussion

In all five analyzed scenarios the hierarchy of organizations consists of four organizations, representing four feasible states of the system. Some organizations just contain a lower organization and a new species that does not interact with the species of the lower organization (*e.g.*, Orgs. 2 and 3 in the starvation scenario and in the glucose scenario). In other cases, exactly those species performing a specific cellular function make up the difference between an organization and its lower neighbor (*e.g.*, Orgs. 2 and 3 in the scenario with all sugars supplied). In these cases a modularity of the analyzed network model is uncovered by organization theory. In this example, the uncovered modules correspond to the inducible uptake systems for lactose and glycerol. Only those organizations that contain the metabolic species correspond to system states facilitating bacterial growth. As expected, such an organization is not found in the scenario without any supplied sugar. For glucose as the exclusive carbon source, all organizations contain

the metabolites. For lactose and glycerol, only those organizations contain the metabolites that also contain the species of the respective uptake systems. This result confirms that glucose can be unconditionally utilized, while lactose and glycerol can only be utilized after their respective uptake systems have been induced. The diauxic growth behavior of *E. coli* is not revealed by the hierarchy of organizations. In the scenario with three sugars supplied as input, organizations are found that correspond to states where glucose and other sugars are utilized simultaneously. Firstly, this highlights the fact that organizations only represent potential steady states of the system. Further kinetic information is required to determine whether an organization indeed contains steady states or not. And secondly, inhibitory interactions play a crucial role in diauxic growth, but had to be ignored in the conversion of the original network model. Since inhibitory interactions in the original network only decrease reaction rates, they in principle cannot be captured by the theory of organizations in which only the presence or absence of molecular species is considered. The inhibitory interactions not considered in our model (inducer exclusion) ensure that the system moves down to Org. 1 in the scenario with all sugars present (see Figure 7.1(e)). This organization represents growth on glucose only. If then for example glucose and glycerol were removed from the medium (switching to Org. 1 in the lactose scenario, Figure 7.1(c)), the inhibition would be removed and the lactose uptake system could be induced. Note that a basal concentration of LacY and LacZ is required so that external lactose can be taken up by the cell and transformed into allolactose which induces the uptake system (see Figure 7.2(a)). Since we model species at basal levels as not being present, we have to add the respective mRNA species in form of a constructive perturbation to the system in order to move it from Org. 1 up to Org. 3 and thus inducing the lactose uptake system.



Chapter 8

Regulated Metabolic Network of E. coli

with Christoph Kaleta

In the previous chapter, a network model of the central sugar metabolism of *E. coli* was analyzed. Organizations were found to coincide with growth on different carbon sources. However, some organizations were related to the simultaneous uptake of all carbon sources. Such biological infeasible states appeared as organizations, because inhibitions had to be neglected. Negative interactions were expressed within the kinetics in the original network, making their representation in a reaction network infeasible. Inhibitory interactions, and regulation in general, however, can also be described by other means, for example by boolean statements. In this chapter, a model of the regulated central metabolic network of *E. coli* by Covert and Palsson (2002) will be analyzed. In this model, the regulatory interaction are expressed by boolean expressions. First, a procedure will be introduced to translate these logic statements into the language of chemical reactions. A small network example will be used to illustrate the procedure. Then, after the metabolic network of *E. coli* and its regulatory boolean expressions are merged into one reaction network, its organizational structure is studied. We find that here at last, only biological feasible organizations exist. They represent all known growth states on (combinations of) the different carbon sources. The network is finally used to predict the lethality of gene knockouts. The method is able to predict 100 out of 116 cases correctly. While regulatory flux balance anal-

ysis (Covert et al., 2001), another method to study regulated metabolic networks, predicts six more cases correctly, these cases highlight the differing assumptions both approaches make.

8.1 Regulatory Interactions

Regulation has not yet been considered in the analysis using the theory of chemical organizations. The aim of this section is to elaborate a concept that allows the consideration of regulation within this framework. As a result, regulated (metabolic) networks are made accessible to organization analysis.

8.1.1 Types of Regulation

To examine the effects of regulation on chemical organizations we first need to discuss the general types of regulatory interactions that occur in biological systems in more detail.

Regulation appears on different levels in the cell, being carried out by a variety of biological entities (*e.g.*, small molecules, proteins, RNA) acting on other biological target entities. As we are considered with metabolic networks, we focus here on the regulation of reactions. Two different types of regulation have to be considered. The first type of regulation only changes the flux of the regulated reaction slightly. This leaves the species composition of the system unchanged. Certain types of autoregulation fall into this category. This kind of regulation does not change the reaction network and hence does not affect the organizational hierarchy. The second type of regulation is more drastic: it turns a reaction completely off or enables a formerly unavailable reaction. This is the case, for example, when the expression of a protein that catalyzes a reaction is suddenly repressed. As a consequence, the catalyzed reaction is not available to the network anymore. The induction of uptake pathways (*e.g.*, the lac uptake system, see Section 7.1) is an example for the enabling of novel reactions. Such changes in network structure can possibly lead to a change in the hierarchy of organizations.

Note that this kind of categorization of regulation leads to meaningful models (Covert et al., 2001) and also generalizes to discretizations using more than two levels as used for example by Espinosa-Soto et al. (2004).

Regulatory interactions do not happen instantly. The time delay between the onset of a regulatory event and its measurable effect in the system can vary between milliseconds (*e.g.*, phosphorylation of proteins in signal cascades (Segall et al., 1982)) and minutes (*e.g.*, changes in gene expression (Hargrove et al., 1991)). However, as we are here interested in the longterm behavior of the system, we do not consider different time scales of regulation.

8.1.2 Modeling Regulatoy Interactions

Several approaches exist to represent regulatory interactions (de Jong, 2002), for example, boolean logic (Covert et al., 2001; Kauffman, 1993; Thomas, 1973), stochastic modeling (McAdams and Arkin, 1997), and differential equations (Smolen et al., 2000). Whereas some approaches require very detailed knowledge about the mechanism and the kinetics behind the regulation, the representation of regulatory interactions by boolean logic can be useful if such information is not completely available (Espinosa-Soto et al., 2004). In this approach, the states *on* and *off* are assigned to regulated reactions (Thomas, 1973). We adopt this notion to model regulatory interactions. Two types of regulatory events have to be considered: *activation*, in which a species is required in order to perform a certain reaction, and *inhibition*, in which a species inhibits a certain reaction and makes it unavailable to the system.

Activation. The activation or turning on of a reaction by a specific species can be simply modeled by considering this species as a catalyst. By this, the reaction can only take place as long as the activating species is present. Being a catalyst, the activating species is not used up by the reaction. Let us consider the general case in which species E activates a reaction that transforms an educt A into product B. In the absence of E, the reaction shall have a zero flux, while the flux shall become positive in the presence of E and A. A reaction $A \rightarrow B$ activated

by E then becomes:

$$E + A \rightarrow E + B. \tag{8.1}$$

Note that adding E as a catalyst on both the educt and product side of the reaction equation does not change the stoichiometric matrix $\mathbf{S}$. Therefore, any flux vector that guarantees self-maintenance for a set of species including E but without considering E as an activator, will also guarantee self-maintenance when E is added as a catalyst to the reaction to model activation.

Inhibition. Handling inhibition is a little bit more difficult. If inhibitor I inhibits a reaction, we could add an if-statement to each reaction that guarantees that the reaction is only available in the absence of I. However, as we intend to model regulation within the language of reactions, this approach is not feasible. Instead, we consider inhibition as another type of activation: the absence of the inhibitor activates the reaction. For achieving this, we have to introduce a pseudo species $\overline{I}$ that represents the absence of inhibitor I. That means that for each inhibitor I, $\mathcal{M}$ contains two species: I and $\overline{I}$. A reaction $A \rightarrow B$ inhibited by I becomes:

$$\overline{I} + A \rightarrow \overline{I} + B. \tag{8.2}$$

Only in the absence of I, represented by pseudo species $\overline{I}$, educt A can react to form product B.

Consistent Organizations. Because now, two species refer to the same molecular entity, – one indicating its presence, and the other its absence – we have to ensure that both are not present at the same time in an organization, and both are not absent at the same time. In both situations, the presence of the respective molecular species would not be clearly defined. Hence, we restrict our analysis to those organizations, in which either I or pseudo species $\overline{I}$ is contained. We call such organizations *consistent*.

Definition of Consistent Organization: An organization $O \subseteq \mathcal{M}$ is called a *consistent organization*, if for all species $S \in \mathcal{M}$ for which a pseudo species $\overline{S} \in \mathcal{M}$ exists that indicates its absence, either S or $\overline{S}$ is part of the organization.

In passing we note that this approach allows one to model even more than two states of a molecule, for example different phosphorylation states.

8.1.3 Modeling Boolean Logic

There are few cases where a reaction is regulated by a single molecular species alone. In most cases regulation is more complex, with many different proteins playing a role in the activation of a reaction. In such cases we need to model the regulation by a set of boolean functions. This section presents an approach to account for such functions on the level of regulation (see also Matsumaru et al., 2007).

All binary boolean functions can be reduced to either *AND* or *OR*, and the negation *NOT*. How a negation can be realized in a reaction network has been outlined above. In principle, it would be sufficient to present a method to implement *AND* or *OR*. However, we present methods for both to ease the process of converting logic statements describing regulation to chemical reactions.

First, we consider the *AND* function. A typical regulatory example is the required presence of two activators to perform a reaction. If we consider activator $E1$ and activator $E2$ to be necessary for a reaction converting educt A into product B, we write:

$$E1 + E2 + A \rightarrow E1 + E2 + B. \tag{8.3}$$

Next, the *OR* function is considered. A biochemical example is a reaction transforming educt A into product B that can alternatively be activated by two activators $E1$ and $E2$. The presence of one of the activators is sufficient to perform the reaction. In this case, the reaction is split into two parts: one that accounts for the presence of activator $E1$, and one the other accounts for the presence of activator $E2$:

$$E1 + A \rightarrow E1 + B \tag{8.4}$$

$$E2 + A \rightarrow E2 + B. \tag{8.5}$$

Taking the two presented basic functions, it is possible to model all regulatory interactions in metabolic networks that are represented by boolean rules (Matsumaru et al., 2007).

8.1.4 Example: A Regulatory Switch

As an example for the presented procedure, we analyze a simple reaction network comprising – apart from inflow and outflow – two reactions forming a switch as depicted in Figure 8.1 (A). The product of one reaction inhibits the other reaction and vice versa. Additionally, inhibitor I shuts down both reactions. Thus, we have an *AND* function that requires for both reactions that both I *and* $P2$, respectively $P1$, are absent. A model without regulation would contain only reactions transforming A to $P1$ and $P2$, the influx to A and the outflux from the products: $\mathcal{R}' = \{\ \emptyset \to A,\ A \to P1,\ A \to P2,\ P1 \to \emptyset,\ P2 \to \emptyset\}$.

The boolean expression for the regulated reactions are:

$$A \to P1 \quad \text{if} \quad \neg I \wedge \neg P2 \tag{8.6}$$

$$A \to P2 \quad \text{if} \quad \neg I \wedge \neg P1. \tag{8.7}$$

Applying the presented procedure, these expressions are transformed into chemical reactions. The resulting reaction network contains the following reactions: $\mathcal{R} = \{$

$$\emptyset \to A, \tag{8.8}$$

$$\overline{P2} + \overline{I} + A \to \overline{P2} + \overline{I} + P1, \tag{8.9}$$

$$\overline{P1} + \overline{I} + A \to \overline{P1} + \overline{I} + P2, \tag{8.10}$$

$$P1 \to \emptyset, \tag{8.11}$$

$$P2 \to \emptyset \quad \}. \tag{8.12}$$

The network contains 16 organizations as listed in Table 8.1. Three organizations are consistent organizations: $O_{10} = \{A,\ I,\ \overline{P1},\ \overline{P2}\}$, $O_{11} = \{A,\ \overline{I},\ P1,\ \overline{P2}\}$, and $O_{12} = \{A,\ \overline{I},\ \overline{P1},\ P2\}$. In the remaining organizations it is at least for one species not clearly defined whether it is present or not. In Organization 2 for example, the presence of A and I is determined with A present

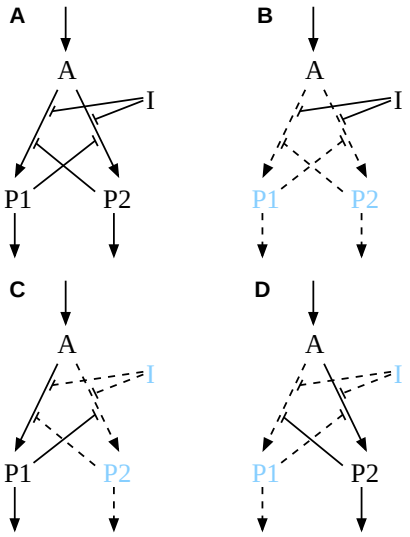


Figure 8.1: Regulatory switch network (**A**) and the reaction networks belonging to its three consistent organizations (**B**, **C**, and **D**). Absent species appear in gray. Inactive reactions and interactions are dashed. Panel **B** represents Organization 10 $= \{A, I\}$, where inhibitor I represses both reactions from A to $P1$ and $P2$. Panels **C** and **D** represent Organizations 11 $= \{A, P1\}$ and Organization 12 $= \{A, P2\}$, where one pathway is active, either over $P1$ or $P2$.

and I absent, but there is no information concerning species $P1$ and $P2$. In Organization 6, inhibitor I is present and absent at the same time. Figure 8.1 (B, C, D) depicts the reaction networks belonging to the three consistent organizations. The consistent organizations represent the three states of the switch. In Organization 10, inhibitor I is present and shuts down reactions 8.9 and 8.10. In the other two consistent organizations I is not present and there is either a flux through reaction 8.9 (Organization 11) or through reaction 8.10 (Organization 12). They represent the two other states of the switch.

8.2 Application to a Metabolic Model of E. coli

The introduced method will be applied to a model of the regulated central metabolism of *E. Coli* by Covert and Palsson (2002). The regulatory interactions are described by a set of boolean expressions in this model. A total of 73

Table 8.1: All organizations of the regulatory switch network (Figure 8.1 (A)). Three organizations are consistent: 10, 11, and 12 (marked bold).

Org.	Species	Real Species
1	A	-
2	A, $\overline{I}$	-
3	A, I	-
4	A, $\overline{P2}$	-
5	A, $\overline{P1}$	-
6	A, I, $\overline{I}$	-
7	A, I, $\overline{P1}$	-
8	A, I, $\overline{P2}$	-
9	A, $\overline{P1}$, $\overline{P2}$	-
10	A, I, $\overline{P1}$, $\overline{P2}$	A, I
11	A, $\overline{I}$, $P1$, $\overline{P2}$	A, $P1$
12	A, $\overline{I}$, $\overline{P1}$, $P2$	A, $P2$
13	A, $\overline{I}$, I, $\overline{P1}$, $\overline{P2}$	-
14	A, $\overline{I}$, I, $P1$, $\overline{P2}$	-
15	A, $\overline{I}$, $\overline{P1}$, $P1$, $\overline{P2}$, $P2$	-
16	A, $\overline{I}$, I, $\overline{P1}$, $P1$, $\overline{P2}$, $P2$	-

enzymes catalyze 113 reactions. Of these reactions, 43 are regulated by 16 regulatory proteins. The unregulated proteins are assumed to be present in the cell at all times, and hence we add an inflow for all of them in our analysis. To incorporate the regulation into the reaction network, we add the proteins that catalyze reactions explicitly as catalysts in the reactions as described in Section 8.1.2. The regulatory logic is incorporated by introducing pseudo species and adapting the reactions accordingly, as described in Sections 8.1.2 and 8.1.3. The activity of several genes is described by boolean statements. Appropriate chemical reactions are added to model this gene regulation. We analyze two variants of the network model by Covert and Palsson (2002): a simplified *core network* to study growth on different carbon sources, and the *complete network* for predicting knockout

experiments. Both reaction networks including a list with species abbreviations can be found in Appendix D.

Core Network Model. For studying growth on different carbon sources including diauxic shift, the network is reduced to the set of reactions that lead from external glucose, lactose, and glycerol to pyruvate via glycolysis. Additionally, the pentose-phosphate pathway reactions and the reactions leading from glucose-6-phosphate to this pathway are removed. The resulting network comprises 49 reactions of the original network. The considered part of the network does not contain any ATP production. However, ATP is used up by some reactions, for example, in glucose uptake. Therefore, ATP is provided as input. Furthermore, UTP, NAD, NADP, Ubiquinone, and external hydrogen ions are necessary for other uptake and transformation reactions and cannot be provided by this part of the network. These species are added as input as well. To model growth, an outflow is added for every biomass precursor, as in the original network. Since we consider proteins as being active only when they are produced, an outflow for every protein is added as well, modeling degradation. In order to model different growth media and conditions, self-replicator reactions for external glucose, lactose, glycerol, and oxygen are added of the form $M \to 2\ M$. These reactions ensure that a constant supply of the respective species is available, whenever it is considered as being present. Using self-replicator reactions, all $2^4 = 16$ different growth conditions can be modeled in a single network and can be simultaneously considered in one analyis.

The final model comprises 95 species (including 15 pseudo species representing the absence of a species) and 168 reactions.

Complete Network Model. For predicting the lethality of gene knockouts, we use the complete network model of the regulated central metabolism of *E. coli* by Covert and Palsson (2002). Depending on the availability of oxygen and the different carbon sources in the growth medium, influxes are added for the respective external species. The currency metabolites HEXT, PI, ADP, ATP, NAD, NADH, Q, QH2, NADP, NADPH, FAD, FADH, UTP, and COA are considered

to be unconditionally available in the cell. Input reactions are added for all these species.

Without the influxes for external carbon sources and oxygen, the network contains 206 species and 463 reactions.

8.2.1 Growth on Carbon Sources

The core network model contains 16 consistent organizations. They are listed in Table 8.2. A graphical represention is provided in Table 8.3. The consistent organizations coincide with the 16 possible growth conditions. The smallest Organization 1 just contains the input species plus the products of the hydrolyzation of ATP, ADP, and phosphate. When analyzing the genes that are active in this organization, we find that the response regulators for glucose, lactose, and glycerol are active, indicating that the respective carbon sources are not present. Due to the absence of oxygen, the aerobic response regulators ArcA and Fnr are also active.

In Organization 2, external oxygen is available. Consequently, the aerobic response regulators ArcA and Fnr are absent here. This is the only difference to Organization 1.

Glucose uptake. The first organization that utilizes an external carbon source is Organization 3 which contains the reactions for glucose uptake. Consequently, the metabolites of the central metabolism are present in this organization. When examining the proteins of the organization, we find that the glucose response regulator Mlc is absent. The organization next in size is Organization 4. Here, lactose is additionally available in the medium. Although the repressor of the *lac* genes, lacI, is absent in the organization, no uptake reactions for external lactose are contained in the organization. The lactose permease LacY, a product of the *lac* genes, is missing. As glucose is available in the medium, the lactose uptake system is not induced by the presence of external lactose. This effect, known as inducer exclusion, leading to the diauxic shift behavior of *E. coli* was already discussed in detail in Section 7.1. Organization 5 represents a similar case in which glycerol is available in the growth medium but not taken up. All

external carbon sources and oxygen are available in Organization 10, but the cell is still exclusively utilizing glucose. Organizations 6 to 9 represent further input combinations defining growth conditions with external glucose available. The availability of oxygen does not change the reactions in the part of the central metabolism that is considered in the core network model.

Lactose uptake. In Organization 13, lactose is the exclusive external carbon source. Consequently, LacI is absent as it is bound by allolactose, a derivative of lactose. Hence, it cannot repress the genes necessary for lactose uptake and utilization. We find the corresponding gene products present in this organization, namely LacZ and LacY. Additionally, derivatives of lactose like galactose are contained in the organization. These metabolites are created in the pathway leading from lactose to the central metabolism. Another diauxic shift effect can be observed in Organization 14. Here, external lactose and glycerol are present as carbon sources, but as in the case with glucose and lactose, only lactose is taken up. Organizations 15 and 16 represent further growth conditions in which lactose is taken up. Once again, the availability of oxygen does not change the reactions in the modeled part of the central metabolism.

Glycerol uptake. Glycerol is the exclusive external carbon source in Organization 11. As all proteins necessary for glycerol uptake are present, glycerol is taken up. For glycerol uptake, three different enzymes catalyze the reaction from glycerol-3-phosphate to dihydroxyacetone-phospate, a metabolite of glycolysis. One of these enzymes, glycerol-3-phosphate-dehydrogenase, is constitutively expressed in the model. The other two proteins, glycerol-3-phosphate kinases GlpABC and GlpD are specific for anaerobic, respectively aerobic growth conditions. Therefore, GlpABC is present and GlpD absent in Organization 11, where no oxygen is available. When oxygen is available as in Organization 12, GlpD is present and GlpABC absent.

Table 8.2: Consistent organizations in the core network model of the regulated central metabolism of *E. Coli*, ordered by size.

Cons. Org.	Species	Growth medium	Uptake
1	Input metabolites, ADP, PI, ArcA, Fnr, GalR, GalS, GlpR, LacI, Mlc, PykF, Ubiquitous proteins	-	-
2	Input metabolites, ADP, O2, O2xt, PI, GalR, GalS, GlpR, LacI, Mlc, PykF, Ubiquitous proteins	O2	-
3	Input metabolites, Glycolysis metabolites, ADP, G1P, GLC, GLCxt, LCTSxt, NADH, PI, PPI, UDPG, ArcA, Crr, FadR, Fnr, Food, GalP, GalR, GalS, GlpR, LacI, Pgk, PtsGHI, PykF, Ubiquitous proteins	GLC	GLC
4	Input metabolites, Glycolysis metabolites, ADP, G1P, GLC, GLCxt, LCTSxt, NADH, O2, O2xt, PI, PPI, UDPG, ArcA, Crr, FadR, Fnr, Food, GalP, GlpR, Pgk, PtsGHI, PykF, Ubiquitous proteins	GLC, LCTS	GLC
5	Input metabolites, Glycolysis metabolites, ADP, G1P, GLC, GLCxt, GLxt, LCTSxt, NADH, PI, PPI, UDPG, ArcA, Crr, FadR, Fnr, Food, GalP, GalR, GalS, LacI, Pgk, PtsGHI, PykF, Ubiquitous proteins	GLC, GL	GLC
6	Input metabolites, Glycolysis metabolites, ADP, G1P, GLC, GLCxt, GLxt, LCTSxt, NADH, O2, O2xt, PI, PPI, UDPG, ArcA, Crr, FadR, Fnr, Food, GalP, Pgk, PtsGHI, PykF, Ubiquitous proteins	GLC, GL, LCTS	GLC
7	Input metabolites, Glycolysis metabolites, ADP, G1P, GLC, GLCxt, NADH, PI, PPI, UDPG, Crr, FadR, Food, GalP, GalR, GalS, GlpR, LacI, Pgk, PtsGHI, PykF, Ubiquitous proteins	GLC, O2	GLC
8	Input metabolites, Glycolysis metabolites, ADP, G1P, GLC, GLCxt, NADH, O2, O2xt, PI, PPI, UDPG, Crr, FadR, Food, GalP, GlpR, Pgk, PtsGHI, PykF, Ubiquitous proteins	GLC, LCTS, O2	GLC
9	Input metabolites, Glycolysis metabolites, ADP, G1P, GLC, GLCxt, GLxt, NADH, PI, PPI, UDPG, Crr, FadR, Food, GalP, GalR, GalS, LacI, Pgk, PtsGHI, PykF, Ubiquitous proteins	GLC, GL, O2	GLC
10	Input metabolites, Glycolysis metabolites, ADP, G1P, GLC, GLCxt, GLxt, NADH, O2, O2xt, PI, PPI, UDPG, Crr, FadR, Food, GalP, Pgk, PtsGHI, PykF, Ubiquitous proteins	GLC, GL, LCTS, O2	GLC
11	Input metabolites, Glycolysis metabolites, ADP, G1P, GL, GL3P, GLxt, NADH, NADPH, O2, O2xt, PI, PPI, QH2, UDPG, ArcA, Crr, Fnr, Food, GalP, GalR, GalS, GlpABC, GlpF, GlpK, LacI, Mlc, Pgk, PtsGHI, PykF, Ubiquitous proteins	GL	GL
12	Input metabolites, Glycolysis metabolites, ADP, G1P, GL, GL3P, GLxt, NADH, NADPH, PI, PPI, QH2, UDPG, Crr, Food, GalP, GalR, GalS, GlpD, GlpF, GlpK, LacI, Mlc, Pgk, PtsGHI, PykF, Ubiquitous proteins	GL, O2	GL
13	Input metabolites, Glycolysis metabolites, Lactose derivatives, ADP, G1P, GLC, LCTS, LCTSxt, NADH, PI, PPI, UDPG, ArcA, Crr, Fnr, Food, GalE, GalK, GalM, GalP, GalT, GlpR, LacY, LacZ, Mlc, Pgk, PtsGHI, PykF, Ubiquitous proteins	LCTS	LCTS

Cons. Org.	Species	Growth medium	Uptake
14	Input metabolites, Glycolysis metabolites, Lactose derivatives, ADP, G1P, GLC, LCTS, LCTSxt, NADH, O2, O2xt, PI, PPI, UDPG, ArcA, Crr, Fnr, Food, GalE, GalK, GalM, GalP, GalT, LacY, LacZ, Mlc, Pgk, PtsGHI, PykF, Ubiquitous proteins	GL, LCTS	LCTS
15	Input metabolites, Glycolysis metabolites, Lactose derivatives, ADP, G1P, GLC, GLxt, LCTS, LCTSxt, NADH, PI, PPI, UDPG, Crr, Food, GalE, GalK, GalM, GalP, GalT, GlpR, LacY, LacZ, Mlc, Pgk, PtsGHI, PykF, Ubiquitous proteins	LCTS, O2	LCTS
16	Input metabolites, Glycolysis metabolites, Lactose derivatives, ADP, G1P, GLC, GLxt, LCTS, LCTSxt, NADH, O2, O2xt, PI, PPI, UDPG, Crr, Food, GalE, GalK, GalM, GalP, GalT, LacY, LacZ, Mlc, Pgk, PtsGHI, PykF, Ubiquitous proteins	GL, LCTS, O2	LCTS

For brevity, pseudo species indicating the absence of a species are not listed. A list of abbrevations can be found in Appendix D. A species followed by 'xt' denotes its extra-cellular form. "Ubiquitous proteins" include the proteins that are considered ubiquitously present in the cell and therefore are not listed separately. They are: Eno, Fba, Fbp, GalU, GapA, Glk, GpmA, GpmB, GpsA, PfkA, PfkB, Pgi, Pgm, PykA, and TpiA. "Input metabolites" denotes the metabolites provided as input to the system: HEXT (external hydrogen), Q (Ubiquinone), ATP, UTP, NAD, and NADP. "Glycolysis metabolites" denotes the metabolites of the Glycolysis: G6P, F6P, FDP, T3P2, T3P1, 13PDG, 3PG, 2PG, PEP, and PYR. "Lactose derivatives" denotes the derivatives of lactose in the central metabolism: GAL1P, GLAC, UDPGAL, bDGLAC, bDGLC.

Table 8.3: Consistent organizations in the core network model of the regulated central metabolism of *E. coli*, ordered by size. The presence of a metabolite in an organization is indicated by dark gray, its absence by light gray. A list of abbrevations can be found in the Appendix D. A species followed by 'xt' denotes its extra-cellular form. "Ubiquitous proteins" include the proteins that are considered ubiquitously present in the cell and therefore are not listed separately. They are: Eno, Fba, Fbp, GalU, GapA, Glk, GpmA, GpmB, GpsA, PfkA, PfkB, Pgi, Pgm, PykA, and TpiA. "Input metabolites" denotes the metabolites provided as input to the system: HEXT (external hydrogen), Q (Ubiquinone), ATP, UTP, NAD, and NADP. "Glycolysis metabolites" denotes the metabolites of the Glycolysis: G6P, F6P, FDP, T3P2, T3P1, 13PDG, 3PG, 2PG, PEP, and PYR. "Lactose derivatives" denotes the derivatives of lactose in the central metabolism: GAL1P, GLAC, UDPGAL, bDGLAC, bDGLC.

8.2.2 Predicting Gene Knockout Experiments

Knockout experiments are performed using the complete network model. Gene knockouts are modeled by deleting all reactions in which the corresponding protein takes part as educt or product. The set of consistent organizations is determined for each knockout experiment. In several cases, the deterministic algorithm to compute organizations (see Chapter 3) did not finish in reasonable time, and a heuristic approach had to be used instead (cases marked 'h' in Table 8.4). The lethality of a knockout can be predicted by the existence of organizations containing all biomass precursor metabolites. If such an organizations is not found, the knockout is predicted to be lethal. We use organization theory (OT) to predict the same gene knockouts as Covert and Palsson (2002). The authors used regulated flux balance analysis (rFBA) for gene knockout predictions. This publication is also our source for *in vivo* data and predictions by flux balance analysis (FBA) and regulatory flux balance analysis (rFBA). The results are presented in Table 8.4. Out of 116 experiments, the predictions by FBA are correct in 97 cases (83,6%). The predictions by rFBA are correct in 106 cases (91,4%) and improve the results of FBA in nine cases. Organization theory predicts the lethality of knockouts correctly in 100 cases (86,2%). The predictions are identical to rFBA predictions except for six cases in which rFBA predictions match the *in vivo* data but OT predictions do not. The reason for these discrepancies will be discussed in detail.

Assumption that accumulation of mass is lethal. In two cases, OT predicts a lethal knockout to be nonlethal (*rpiA*, and *rpiA* + *rpiB* on glucose). The self-maintenance property allows for the accumulation of internal metabolites, while in rFBA, only steady states are considered, and any accumulation of metabolites is regarded as lethal. In these two cases, the organizations containing all biomass precursors contain metabolites with positive productions[1]. Hence, OT predicts the knockout to be nonlethal while rFBA predicts it to be lethal as no steady state

[1] Note that all species except the pseudo species indicating the absence of species decay in the network model. Hence, all organizations are balanced organizations. However, accumulation of metabolites occur, if the decay reactions (which are not present in the original network) are removed.

exists. Restricting our analysis to balanced organizations only, the predictions for these two knockout experiments are identical to rFBA.

Assumption that secreted molecules have no effect. Further three incorrect predictions by OT (*ackA* and *pta* on acetate, and *ppc* on glycerol) yield deeper insights into the differences between chemical organization theory and regulatory flux balance analysis. In the case of acetate uptake, there are two pathways that enable the utilization of this carbon source. One pathway leads directly from acetate to acetyl-CoA, and the other takes the route via acetyl phosphate. The first pathway is catalyzed by the Acetyl-CoA synthethase (gene *acs*). According to the model, *acs* is only transcribed if no carbon source is available or at most acetate or formate, or both. The second pathway is catalyzed by acetate kinase A (gene *ackA*) and phosphotransacetylase (gene *pta*). If one of these genes is knocked out, the first pathway can still support the central metabolism, given that acetate is the exclusive external carbon source. In this case, chemical organization theory predicts both knockouts as lethal, which is not the case *in vivo* and correctly predicted by rFBA. The reason for this discrepancy is that in any network containing the biomass precursor metabolite pyruvate, this metabolite will be secreted. Therefore, such a network also comprises the external form of pyruvate which is an inhibitor for the only remaining uptake reaction for acetate. Consequently, there exists no organization containing all biomass precursor metabolites when acetate is the exclusive carbon source in the growth medium and the second pathway is knocked out. Since the presence of metabolites is not explicitly considered in rFBA, this inhibition is not detected by rFBA. However, since the knockout is nonlethal in *in vivo* experiments, the levels of secreted pyruvate might not be sufficient to have an effect on the expression of *acs*. Or, the cell switches its uptake from acetate to pyruvate until it is depleted and then switches back to acetate again. Since organization theory does not consider different concentration levels of metabolites (only the presence or absence is considered), concentration dependent inhibition cannot be taken into account. The wrong prediction of the knockout of *ppc* on glycerol as nonlethal can be explained by the same argument. Gene *ppc* codes for the phosphoenolpyruvate carboxylase which supplies the citric acid cycle with oxaloacetate (OA). When *ppc* is knocked

out, the only alternative for OA production is the glyoxylate shunt, consisting of the isocitrate lyase (gene *aceA*) and the malate synthase A (gene *aceB*). However, the glyoxylate shunt is only active if *E. coli* grows on acetate or fatty acids as the sole carbon source (Maloy and Nunn, 1982). Hence, the knockout of *ppc* on glycerol is lethal *in vivo*, as the glyoxylate shunt is not actived. In the model, the regulation of the glyoxylate shunt is implemented as follows. If no glucose but acetate is present in the growth medium, the fatty acid and acetate response regulator genes *fadR* and *iclR* are not active. IclR leads to the downregulation of the expression of *aceA* and *aceB*. But as *iclR* is not expressed, the glyoxylate shunt is actived. Any organization containing the biomass precursor metabolite acetyl coenzyme A also contains acetate which is secreted. Hence, any organization containing the biomass precursors also contains the external form of acetate. Consequently, when glycerol is the only supplied carbon source, the condition for the glyoxylate shunt activation is fulfilled for any organization containing the biomass precursors, and the glyoxylate shunt is activated. The secreted form of acetate activates the glyoxylate shunt, enabling the survival of the cell even if *ppc* is knocked out. Again, secreted material from the cell is not considered by rFBA, as concentrations are assumed to be too low for having further effects.

However, the discussed wrong predictions can be easily corrected within the organization theory framework. If external carbon species are modeled as two separate species, with one for the supply from the growth medium and one for the secreted material from the cell (having too low concentrations to trigger further cellular responses, as assumed in rFBA), the predictions of OT match those of rFBA.

Regulatory rules based on concentration difference. One case (*mdh* on succinate) is wrongly predicted as lethal due to the lack of an approach to handle regulation on the flux level. Activation of catabolite activator protein Cra depends on whether FDP or F6P is available in excess or not. In rFBA, this is determined by requiring production and consumption reactions to proceed in appropriate directions. Hence, concentration levels are infered from flux conditions. Cra activates the expression of *ppsA*. Since the production of Cra could not be modelled, *ppsA* is never active, so that the knockout of *mdh* in the case of

Table 8.4: Comparing *in vivo* knockout experiment results with predictions made by FBA, rFBA, and by OT. A '+' indicates growth, a '-' no growth of the mutants on the indicated substrate(s). For cases denoted as 'N', data was not available. Results and predictions are derived from *in vivo*/FBA/rFBA/OT. *In vivo* data and references, FBA and rFBA predictions are taken from Covert and Palsson (2002). A heuristic approach to determine organizations had to be used in cases marked with 'h'. In six instances, predictions made by OT deviate from rFBA predictions (shaded boxes). See text for discussion. The growth medium contained glucose (glc), glycerol (gl), succinate (suc), acetate (ac), or ribose (rib). Anaerobic condition is denoted by '$-O_2$'.

	glc	**gl**	**suc**	**ac**	**rib**	**glc ($-O_2$)**	**Dual Substr.**	**Reference**
aceA	+/+/+/+		+/+/+/+^{h}	-/-/-/-^{h}		+/+/+/+		Creaghan and Guest (1978)
aceB				-/-/-/-^{h}				Cronan and LaPorte (1996)
aceEF	-/+/-/-		-/+/-/-^{h}	+/+/+/+^{h}		+/+/+/+	+/+/+/+ (glc-ac)	Langley and Guest (1977)
ackA				+/+/+/-^{h}				Kumari et al. (1995)
ackA + *pta* + *acs*				-/-/-/-^{h}				Kumari et al. (1995)
acnA	+/+/+/+	+/+/+/+^{h}	+/+/+/+^{h}	+/+/+/+^{h}		+/+/+/+		Cronan and LaPorte (1996); Gruer et al. (1997)
acnB	+/+/+/+	+/+/+/+^{h}	+/+/+/+^{h}	-/+/+/+^{h}		+/+/+/+		Gruer et al. (1997)
acnA + *acnB*	-/-/-/-	-/-/-/-^{h}	-/-/-/-^{h}	-/-/-/-^{h}		-/-/-/-		Gruer et al. (1997)
acs				+/+/+/+^{h}				Kumari et al. (1995)
adh	+/+/+/+					-/+/+/+		Cunningham and Clark (1986)
cyd	+/+/+/+							Calhoun et al. (1993)
cyo	+/+/+/+							Calhoun et al. (1993)
eno	-/-/-/-	-/-/-/-^{h}	-/-/-/-^{h}				+/+/+/+ (gl-suc)	Irani and Maitra (1977)
fbaA	-/+/+/+							Fraenkel (1996)
fbp	+/+/+/+	-/-/-/-^{h}	-/-/-/-^{h}	-/-/-/-^{h}				Fraenkel and Horecker (1965)
frdA	+/+/+/+		+/+/+/+^{h}	+/+/+/+^{h}		+/+/+/+		Creaghan and Guest (1978)
fumA				-/+/-/-^{h}		+/+/+/+		Cronan and LaPorte (1996)
gap	-/-/-/-	-/-/-/-^{h}	-/-/-/-^{h}				+/+/+/+^{h} (gl-suc)	Irani and Maitra (1977)
glk	+/+/+/+							Fraenkel (1996)
glk + *pfkA*	+/+/+/+							Fraenkel (1996)
glk + *pts*	-/-/-/-							Fraenkel (1996)

	glc	gl	suc	ac	rib	glc ($-O_2$)	Dual Substr.	Reference
gltA	-/-/-/-			-/-/-/-^{h}				Gruer et al. (1997)
gnd	+/+/+/+							Fraenkel (1996)
icd	-/-/-/-			-/-/-/-^{h}				Gruer et al. (1997)
mdh	+/+/+/+	+/+/+/+^{h}	+/+/+/-^{h}			+/+/+/+		Courtright and Henning (1970)
ndh	+/+/+/+	+/+/+/+^{h}						Tran et al. (1997)
nuo	+/+/+/+	+/+/+/+^{h}						Tran et al. (1997)
pfl						+/+/+/+		Mat-Jan et al. (1989)
pgi	+/+/+/+	+/-/-/-^{h}	+/-/-/-^{h}					Fraenkel (1996)
pgi + *gnd*	-/-/-/-							Fraenkel (1996)
pgi + *zwf*	-/-/-/-							Fraenkel (1996)
pgk	-/-/-/-	-/-/-/-^{h}	-/-/-/-^{h}				+/+/+/+^{h} (gl-suc)	Irani and Maitra (1977)
pgl	+/+/+/+							Fraenkel (1996)
ppc	-/+/-/-	-/+/-/+^{h}	+/+/+/+^{h}				+/+/+/+^{h} (gl-suc) +/+/+/+ (glc-suc)	Courtright and Henning (1970); Fraenkel (1996)
pta				+/+/+/-^{h}				Kumari et al. (1995)
pts	+/+/+/+							Fraenkel (1996)
pykA	+/+/+/+							Fraenkel (1996)
pykA + *pykF*	+/+/+/+							Fraenkel (1996)
pykF	+/+/+/+							Fraenkel (1996)
rpiA	-/+/-/+				+/+/+/+^{h}		+/+/+/+ (glc-rib)	Sørensen and Hove-Jensen (1996)
rpiA + *rpiB*	-/-/-/+				-/+/+/+^{h}		+/+/+/+ (glc-rib)	Sørensen and Hove-Jensen (1996)
rpiB	+/+/+/+				+/+/+/+^{h}		+/+/+/+ (glc-rib)	Sørensen and Hove-Jensen (1996)
rpiR + *rpiA*	+/N/+/+				+/N/+/+^{h}		+/N/+/+ (glc-rib)	Sørensen and Hove-Jensen (1996)
sdhABCD	+/+/+/+		-/-/-/-^{h}	-/-/-/-^{h}		+/+/+/+		Creaghan and Guest (1978)
sucAB-lpd	-/+/+/+		-/+/+/+^{h}	-/+/+/+^{h}		+/+/+/+	+/+/+/+ (glc-suc)	Creaghan and Guest (1978); Langley and Guest (1977)
tpi	-/+/+/+	-/-/-/-^{h}	-/-/-/-^{h}	-/-/-/-^{h}			+/+/+/+ (glc-suc) +/+/+/+ (glc-gl)	Anderson and Cooper (1969); Irani and Maitra (1977)
zwf	+/+/+/+							Fraenkel (1996)

succinate uptake is predicted to be lethal. Not all biomass precursor metabolites

of the glycolysis can be produced. An approach to deal with this kind of regulation on the flux level might be found by constraining the fluxes of certain reactions. Currently, this cannot be done easily within the theory of chemical organizations.

8.2.3 Regulatory Effects of Gene Knockouts

When a gene is knocked out, this perturbation of the system can lead to further effects in the system, resulting in additional changes in gene expression. For the knockout experiments, the network was changed by removing appropriate reactions. If the knockout is nonlethal, the network still contains an organization containing all biomass precursors. And even for some lethal knockouts, the network still contains an organization. This organization then lacks one or more essential biomass precursors. We compare these knockout organizations with the wildtype organization for all knockout experiments, for which an organization was still found. We test if other species representing genes or gene products appear or are missing compared to the wildtype organization. This differences represent further changes in gene expression that are induced by the gene knockout. Table 8.5 lists all knockout experiments, in which further genes (besides the knocked out gene) disappear, respectively appear, in the knockout organization.

The observed changes in gene expression are caused by several different causes. A typical example is a metabolite that was present in the wildtype organization, but is missing in the knockout organization. By inspecting the reation network, the causes for all observed regulatory effects can be identified. Table 8.6 lists them together with their regulatory effects.

It must be noted that *dcuR* and *dcuS* are only active in *E. coli*, if succinate is present in the growth medium. Succinate can also be secreted by the cell (at low rates). Since the single species $SUCCxt$ represents both secreted and externally added succinate, the secreted low concentration succinate is able to turn *dcuR* and *dcuS* on in the organization theory framework. The problem of modeling the externally supplied and the secreted version of a species by a single network species was already discussed for pyruvate in Section 8.2.2.

We finally summarize the differences between wildtype and knockout organizations that lead to the observed change in gene expression in Table 8.7. Out of

Table 8.5: Knockout experiments in which further genetic species are missing or additionally present in the knockout organization compared to the wildtype organization. Pseudo species indicating the absence of species are omitted, as well as the gene that was knocked out. The lethality as predicted by organization theory is additionally indicated in the column "Knockout".

Substr.	**Knockout**	**Missing Species**	**Additional S.**
glc	*rpiR + rpiA* (+)	-	*rpiB*
glc	*eno* (-)	*aceEF*, *dcuR*, *dcuS*, *frdABCD*, *lpdA*, *sucAB*	*pdhR*
glc	*ppc* (-)	*dcuR*, *dcuS*, *frdABCD*	-
gl	*eno* (-)	*aceA*, *aceB*, *aceEF*, *dcuR*, *dcuS*, *frdABCD*, *lpdA*, *sucAB*	*fadR*, *iclR*, *pdhR*
gl	*gap* (-)	*aceA*, *aceB*, *aceEF*, *dcuR*, *dcuS*, *frdABCD*, *lpdA*, *sucAB*	*fadR*, *iclR*, *pdhR*
gl	*pgk* (-)	*aceA*, *aceB*, *aceEF*, *dcuR*, *dcuS*, *frdABCD*, *lpdA*, *sucAB*	*fadR*, *iclR*, *pdhR*
suc	*aceEF* (-)	*aceA*, *aceB*	*fadR*, *iclR*
suc	*pgi* (-)	*crr*, *ptsGHI*, *pykF*	*cra*, *ppsA*
ac	*fumA* (-)	*aceEF*, *dcuR*, *dcuS*, *frdABCD*, *lpdA*, *sucAB*	*acs*, *pdhR*
ac	*gltA* (-)	*aceEF*, *dcuR*, *dcuS*, *frdABCD*, *lpdA*, *sucAB*	*acs*, *pdhR*
glc (-O_2)	*acnA + acnB* (-)	*dcuR*, *dcuS*	-
glc (-O_2)	*sucAB-lpd* (+)	*dcuR*, *dcuS*	-

Table 8.6: The observed regulatory effects and their cause as derived from the reaction network.

Missing Species	**Additional Species**	**Caused by**
aceEF, *lpdA*, *sucAB*	*pdhR*	$\overline{PYR}$
aceA, *aceB*	*fadR*, *iclR*	$GLCxt$ or $\overline{ACxt}$
dcuR, *dcuS*	-	$\overline{SUCCxt}$ $(-O_2)$
dcuR, *dcuS*, *frdABCD*	-	$\overline{SUCCxt}$
-	*rpiB*	$\overline{rpiR}$
crr, *ptsGHI*, *pykF*	*cra*, *ppsA*	$\overline{G6P}$ and FDP and fbp
-	*acs*	$\overline{Food}$ and $ACxt$

the 12 considered knockout experiments, 10 are lethal. The absence of central metabolites like pyruvate and glucose 6-phosphate are an important cause for the change in gene expression in these cases.

8.3 Discussion

By transforming the boolean formalism that represents the regulation of a metabolic network into reaction rules, we were able to demonstrate how chemical organization theory can be applied to regulated metabolic networks. Using a model of the central metabolism of *E. coli*, each of the 16 wildtype growth scenarios were correctly predicted down to the expression of each protein. Each external condition could be directly mapped to an organization implying a distinct state of the network (*i.e.*, a set of molecular species present). Without specific assumptions, organization theory was able to predict the lethality of knockout experiments correctly in 100 out of 116 cases (86.2%).

In comparison to (r)FBA (Covert and Palsson, 2002), the predictions by organization theory differ from rFBA predictions in six cases. Five of these cases can be resolved in a straightforward way by taking assumptions into account also made by Covert and Palsson, leading to 105 (out of 116) correctly classified cases (90.5%). In the remaining case, a specific constraint, based on a flux condition, was used that cannot be easily considered within organization theory. In

Table 8.7: Differences between wildtype and knockout organizations that induce the observed change in gene expression.

Substrate	Knockout	Difference to Wildtype Organization
glc	*rpiR + rpiA* (+)	$\overline{rpiR}$
glc	*eno* (-)	$\overline{PYR}$, $\overline{SUCCxt}$
glc	*ppc* (-)	$\overline{SUCCxt}$
gl	*eno* (-)	$\overline{PYR}$, $\overline{SUCCxt}$, $\overline{ACxt}$
gl	*gap* (-)	$\overline{PYR}$, $\overline{SUCCxt}$, $\overline{ACxt}$
gl	*pgk* (-)	$\overline{PYR}$, $\overline{SUCCxt}$, $\overline{ACxt}$
suc	*aceEF* (-)	$\overline{ACxt}$
suc	*pgi* (-)	$\overline{G6P}$
ac	*fumA* (-)	$\overline{PYR}$, $\overline{SUCCxt}$, $\overline{Food}$
ac	*gltA* (-)	$\overline{PYR}$, $\overline{SUCCxt}$, $\overline{Food}$
glc (-O_2)	*acnA + acnB* (-)	$\overline{SUCCxt}$
glc (-O_2)	*sucAB-lpd* (+)	$\overline{SUCCxt}$

particular, the deviation between rFBA and organization theory has uncovered three critical aspects:

First, (r)FBA only considers steady states. Any system state with accumulating metabolites is regarded as lethal. In organization theory, accumulating metabolites are explicitly allowed to also cover system states related to growth. To adopt the steady state assumption in organization theory however, one simply can restrict the analysis to balanced organizations.

Second, material secreted by the organism was assumed to be too low in concentration to trigger further cellular responses. In organization theory, only the presence or absence of metabolites is considered. Hence, even smallest concentrations of species will potentially trigger further responses. This problem can easily be resolved by introducing a new molecular species representing the secreted version of a species.

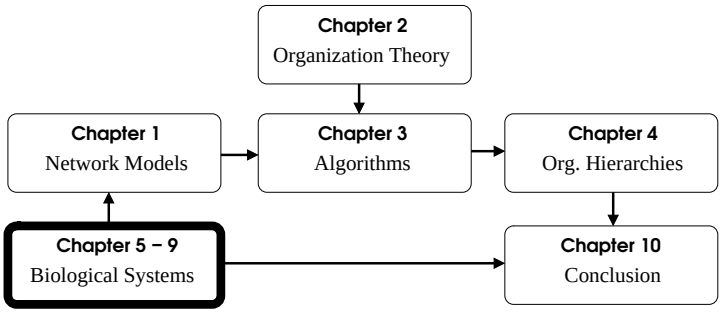
Third, the regulated metabolic model contains regulatory rules that do not refer to concentrations but to concentration differences, or rather, reaction fluxes. Such regulatory mechanisms can easily be incorporated in rFBA, but are difficult to handle within the organization theory framework. However, Kaleta (2007) has

recently proposed a method to also consider this kind of regulation by augmenting the reaction network.

Another reason why organization theory was unable to predict all knockout experiments correctly is that it requires the network model to contain all reactions that can be carried out in the real system. When this complete knowledge is not available, the results are limited. This is both a strength and a weakness. As a strength on one hand, organization theory helps to test if a model that claims to be complete is actually so. On the other hand, the necessity to have a complete model covering all possible reactions to gain fully reliable results is an obvious weakness. Typically, a model is very precise with respect to certain aspects of the real system while other aspects are simplified, in accordance with the intended purpose of the model.

Because chemical organization theory does not rely on the sometimes extensive kinetics, it can serve as a first step to analyze the potential behavior of a regulated system. The analysis delivers all potential network states, described by the sets of molecular species that can coexist over a long time. The further analysis of the network can then focus on interesting states. Taking the other direction, it is possible to validate *in silico* network models. All network states of interest observed *in vivo* should have corresponding organizations in the network model.

When regulation is considered in metabolic networks, the presented approach offers the advantage that both the metabolism and its regulation are modeled within one single framework: chemical reaction rules forming a network. The unification comes at the expense of introducing a set of pseudo species to represent the absence of species. However, this approach allows one to model and consider inhibitory interactions within the framework of organization theory.



Chapter 9

Genome-Scale Metabolic Model of E. coli

The previous chapter pointed out that somehow complete models of biological systems are required for the analysis using the theory of chemical organizations. Are such models of the required scope and quality available? To address this question, a genome-scale metabolic network of *E. coli* by Reed et al. (2003) will be analyzed in this chapter. Although the model is not complete due to 67 dead-end metabolites[1], we discover a rich hierarchy of organizations. However, when this hierarchy is studied in detail we find only few biological meaningful organizations. The species that give rise to these organizations are species, for which the biosynthesis is not contained in the model. Hence, these organizations exist merely due to the incompleteness of the network model which does not account for the synthesis of certain model species. More complete network models of *E. coli*, not only encompassing metabolism, but also gene regulation, signal transduction, and further cellular processes are required to tackle the ultimate question: do living organisms contain a hierarchy of organizations?

9.1 Reaction Network Model

According to the GenProtEC database (Serres et al., 2004), in version from February 1, 2007, the genome of *E. coli* consists of 4485 genes. Of these, 2557 could be

[1] 70, if the biomass production reaction is not considered.

functionally characterized by experiments. Using bioinformatics methods, functions could be assigned to a further 1115 genes. Several years before this comprehensive knowledge was available, Reed et al. (2003) created a metabolic network that accounts for 904 genes. We will analyze this network as a genome-scale metabolic network and study its organizational hierarchy. The reaction network contains 761 species and 931 reactions. All reactions are elementally balanced. 143 species of the network represent molecular species existing outside the cell. For our analysis, we add explicit backward reactions for all 245 reversible reactions. Additionally, we add pseudo species Biomass representing the production of biomass and its production reaction as proposed by Reed et al. (2003). The reaction transforms the 49 biomass precursor species 5mthf, accoa, ala-L, amp, arg-L, asn-L, asp-L, atp, clpn_EC, coa, ctp, cys-L, datp, dctp, dgtp, dttp, fad, gln-L, glu-L, gly, glycogen, gtp, h2o, his-L, ile-L, leu-L, lps_EC, lys-L, met-L, nad, nadh, nadp, nadph, pe_EC, peptido_EC, pg_EC, phe-L, pro-L, ps_EC, ptrc, ser-L, spmd, succoa, thr-L, trp-L, tyr-L, udpg, utp, and val-L into adp, h, pi, ppi, and Biomass. For a complete list of species abbreviations, see Appendix E. With these modifications, the network contains 762 species and 1177 irreversible reactions. None of the species decays spontaneously. Some reactions have non-integer stoichiometric coefficients in the original network. We scale them up so that all coefficients become integer values. To model growth on a rich medium, we add input reactions for 16 external species: ac_ex, co2_ex, fe2_ex, glc-D_ex, glyc_ex, h2o_ex, h_ex, k_ex, lac-D_ex, lac-L_ex, na1_ex, nh4_ex, o2_ex, pi_ex, so4_ex, and succ_ex. The suffix '_ex' denotes the extracellular version of the species. All reactions of the model are listed in the additional data files in Reed et al. (2003).

9.2 Hierarchy of Organizations

The reaction network model contains 67 internal deadend metabolites. While 34 metabolites only appear as products in reactions, 33 metabolites are only used up in reactions as substrates. For organization analysis, product deadends pose no problem as organizations allow for species to accumulate. Substrate deadends however will never show up in any organization as they cannot be produced, prohibiting their (self)-maintenance.

The network is too large to be processed by the constructive approach to compute organizations (see Chapter 3) in reasonable time. Instead, we use the heuristic approach. This method only delivers reactive organizations, that are organizations that only contain species that participate in reactions. Organizations containing isolated species are not considered.

After a runtime of 16 hours, 249 organizations are discovered. The smallest organization contains 31 species and the largest 559 species out of the total 762 species. Figure 9.1 gives an overview of the hierarchy of organizations. Two jumps are noticeable. The first jump occurs between relative small organizations containing 31 to 44 species and Org. C containing 345 species. This is the smallest organization containing the central parts of the metabolism of *E. coli*. All larger organizations contain all species of this organization. The second jump occurs at the top of the organization hierarchy ending in Org. B, containing 532 species. This is the smallest organization to contain pseudo species Biomass, representing the creation of biomass. Again, all larger organizations contain all species of this organization.

9.3 Organizations Containing Novel Species

In order to cope with the large number of organizations in this network, we focus on a specific subset of organizations. In the following, we only consider organizations that contain more species than the mere union of organizations to which downlinks exist. By this, we neglect organizations that are simply the combination of other organizations. Only when a merger leads to novel species, the organization is considered. These are exactly those organizations that have a non-empty label in the simplified notation (see Section 2.1.10). The *E. coli* network contains 65 such organizations as depicted in Figure 9.2. Org. C (Org. 22) and B (Org. 62) again spawn all organizations above them.

9.3.1 Organizations and Biological Functions

In order to study what functions the organizations encapsulate, we sort the network species according to the biological functions and pathways they are associ-

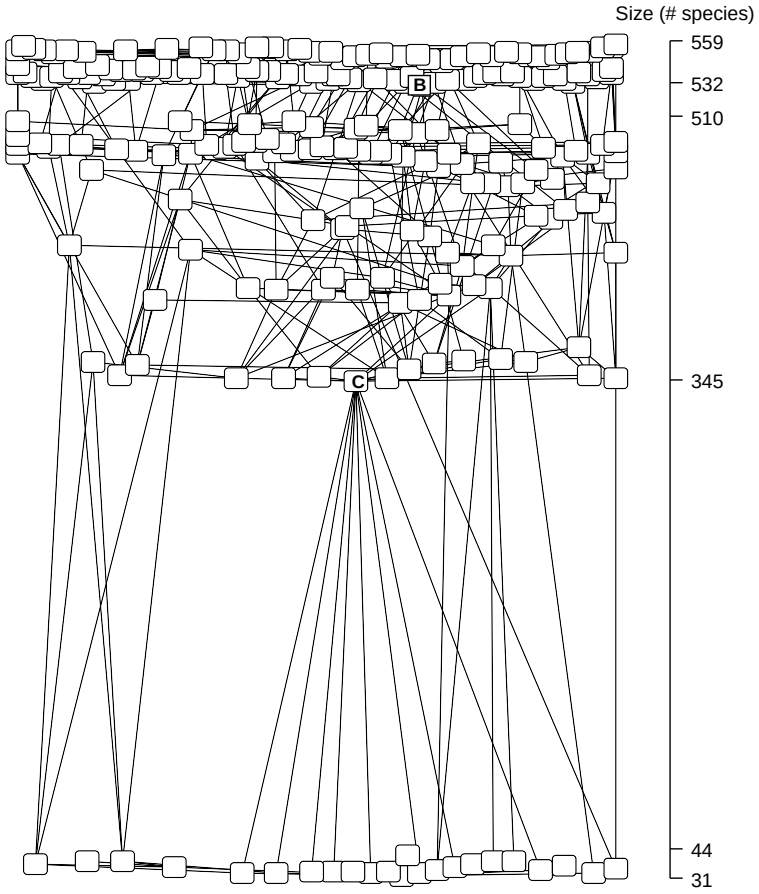


Figure 9.1: Hierarchy of 249 reactive organizations of the genome-scale metabolic network of *E. coli* found by the heuristic approach. Org. C contains the central part of the metabolism; all larger organizations contain this organization. Org. B is the smallest organization containing pseudo species Biomass, indicating the production of biomass. Again, all organizations above this contain all species of this organization. Some lines appear to be horizontal due to the limited vertical resolution.

ated with. We use the network maps provided by Reed et al. (2003) to assign species to eight categories: alternate carbon sources, amino acid metabolism, cell membrane constituents, central metabolism, cofactor biosynthesis, miscellaneous, nucleotide metabolism, and pyruvate metabolism. A species can appear in more than one category. Next, we count for every organization how many species of

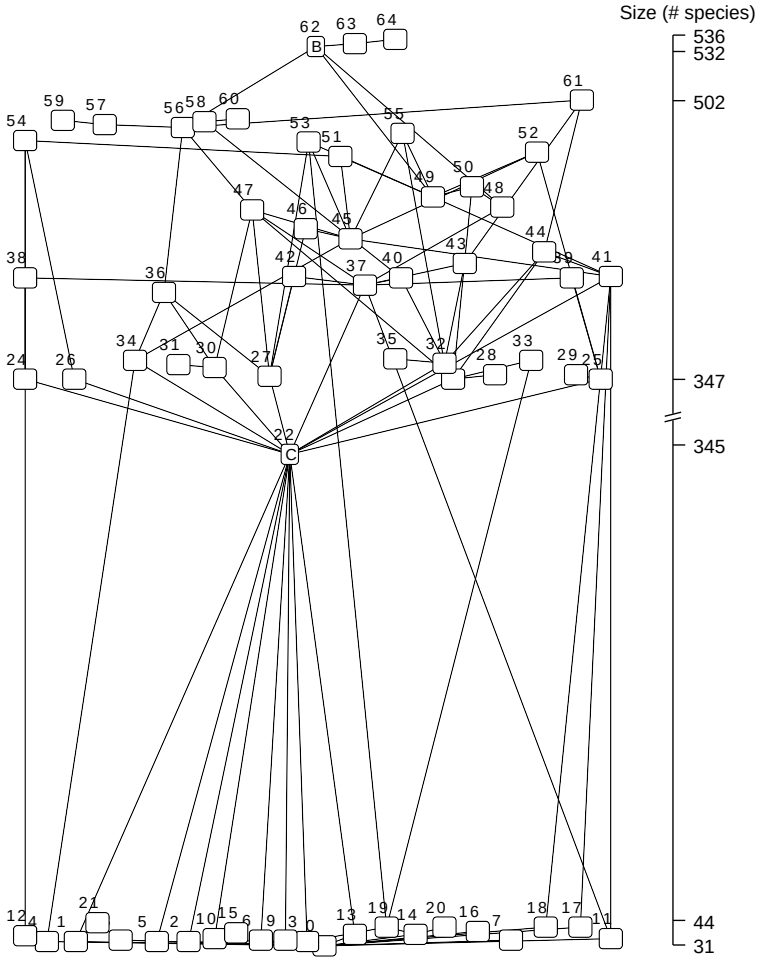


Figure 9.2: Hierarchy of 65 organizations that contain novel species compared to downlink organizations. The smallest organization Org. 0 contains 31 species, the largest Org. 64 contains 536 species. Org. C (Org. 22) containing the central part of the metabolism, and Org. B (Org. 62) containing pseudo species Biomass spawn all larger organizations. The central Org. 22 is placed below its actual position to better reveal the fan-like structure. Some lines appear to be horizontal due to the limited vertical resolution.

each of the eight categories are present. The results are summed up in Figure 9.3. The discovered jumps can be clearly seen here. While Orgs. 0 to 21 have very few species in all categories, Org. 22 contains almost the whole central metabolism, the whole pyruvate metabolism, and almost half of all species in the remaining

six categories. The second jump is not so evident. Org. 62 is the first organization that has almost all species in all categories with the exception of alternate carbon sources and miscellaneous (roughly 50% in both cases). Smaller organizations have similar contents but only feature 50% of all species in the category cell membrane constituents (Org. 56–61), or nucleotide metabolism (Org. 48–55), or both (Org. 37-47).

We find that for most categories, the number of related species contained in an organization has generally either two or three levels. Either almost none of the species are present, or roughly half of them or almost all are present in an organization. With organizations getting larger and larger, there is no smooth increase of organization species related to certain biological functions. The categories cofactor biosynthesis and miscellaneous are the two exceptions. For them, the number of contained species increases in smaller amounts starting from Org. 22. The observed stepwise increase of related species hints to a modular structure of the network. In a feasible state of the system, either the whole set of species belonging to a specific function is present, or none of the species are present (or a certain subset). The species associated with a specific function act as an unit that cannot be divided arbitrarily.

9.3.2 Organization of the Organization Hierarchy

What gives rise to the observed organization hierarchy? To answer this question, we will consider the 65 organizations containing novel species. Three mechanisms can lead to larger organizations containing novel species compared to its downlink organizations. First, if more than one downlink organization exists, the combination of these organizations can facilitate the creation of novel species. For example, if species A and B create species A, B, and C, and both educts are in different organizations, the union of both will enable the production of C. In the second mechanism, the novel species must be added to an organization (respectively to the union of the downlink organizations) in order to get a larger organization containing novel species. This novel species can allow for the production of further species. These further species must take care of the production of the novel species to fulfill the self-maintenance condition. And third, both

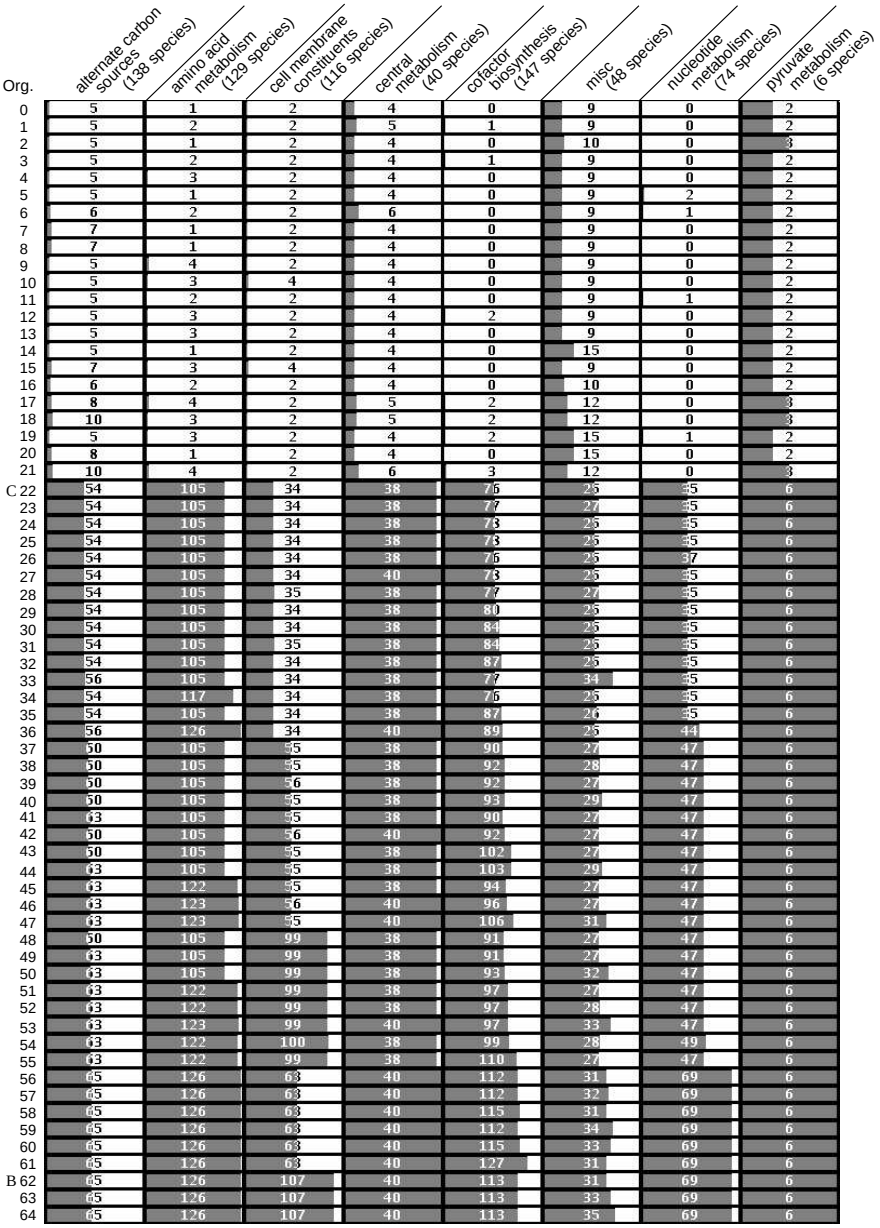


Figure 9.3: Relating organizations to biological function. Numbers refer to how many species of the category are contained in the organization. An empty bar indicates that no species are contained in the organization, a full bar indicates the presence of all category species in the organization. The organization size increases from top to bottom. Biomass is only created in the three largest Orgs. 62–64.

mechanisms can be combined. Interactions between downlink organizations can lead to novel species at the expense of the self-maintenance property. Taking the example from above, this might be the case if species A and B only produce C. As soon as this reaction is possible, species A and B might not be maintainable anymore. Hence, further novel species have to be added to create an organization, for example species that allow the recreation of A and B from species C. Examples for all three mechanisms will be given in the following.

Note that to create a larger organization, it is under certain circumstances not necessary to unite *all* downlink organizations. Interactions between two downlink organizations might already lead to the creation of all species of a third downlink organization. In such cases, the merger of the first two organizations is sufficient to create the larger organization. However, in this analysis we will not determine the minimal sets of downlink organizations capable of creating the larger organization.

1. Interactions between downlink organizations lead to novel species.

Organizations that contain novel species due to interactions between downlink organizations are listed in Table 9.1. For these 12 organizations, the closure of the union of their downlink organizations is already self-maintaining.

Organization 62 ist the smallest to contain pseudo species Biomass, indicating biomass production. It has three downlink organizations. While Orgs. 48 and 49 contain the same 42 required precursors for biomass out of 49 in total, Org. 56 contains 44. To create biomass, Orgs. 48 and 49 need seven more precursor species: cys-L, datp, dctp, dgtp, dttp, met-L, and spmd. These are contained in Org. 56. Conversely, Org. 56 needs five more precursor species: clpn_EC, lps_EC, pre_EC, pg_EC, and ps_EC. These are part of Orgs. 48 and 49. Hence, when the organizations are united all precursor species come together and biomass can be created.

Table 9.1: Organizations with novel species due to interactions between downlink organizations.

Org.	Novel Species	Downlink Org.
34	2kmb, 5mdr1p, 5mdru1p, 5mta, 5mtr, N1aspmd, dkmpp, met-L, n8aspmd, spmd	4, 22
41	2mecdp, 2ohph, 2oph, 2p4c2me, 3ophb, 4c2me, 5dglcn, cdp, ckdo, cmp, csn, ctp, cytd, cytd_ex, dmpp, frdp, gal, gal1p, grdp, h2mb4p, idon-L, idon-L_ex, ipdp, octdp, orot, orot5p, peptido_EC, tre, tre6p, uaagmda, uaccg, uacgam, uacmam, uacmamu, uagmda, uama, uamag, uamr, udcpdp, udcpp, udp, udpg, udpgal, udpgalfur, udpglcur, ugmd, ugmda, ump, unaga, unagamu, uri, uri_ex, utp	11, 17, 18, 22
43	hemeO	32, 37
44	hemeO	23, 32, 41
45	2ombzl, 2omhmbl, 2ommbl, 2omph, ahcys, dhptd, hcys-L, hmfurn, rhcys	34, 41
47	gtspmd	23, 27, 30, 37, 45
53	gtspmd	19, 27, 45, 49
55	hemeO, shcl, sheme, srch	32, 45, 49
56	4ppcys, dcdp, dcmp, dctp, dcyt, dpcoa, dtdp, dtdp4aaddg, dtdp4addg, dtdp4d6dg, dtdp4d6dm, dtdpglu, dtdprmn, dtmp, dttp, dudp, dump, duri, dutp, eca_EC, pan4p, thym, thymd, thymd_ex, unagamuf	36, 47
58	btn	40, 56
61	shcl, sheme, srch	43, 44, 56
B 62	Biomass	48, 49, 56

2. Organizations created by the addition of novel species.

For some organizations, the union of its downlink organizations does not give rise to novel species. In particular, all organizations with only one downlink organization belong into this group. For these organizations, species need to be added to the organization species to give rise to a larger organization. Typically, not all novel species need to be added. Rather, the addition of some species is sufficient to create the larger organization. Often, it is possible to create the same larger organization by adding different species sets to the organization. In Table 9.2, the smallest such sets are listed together with the 48 organizations, for which the mere union of downlink organizations does not give rise to novel species.

As an example for this group of organizations, we inspect Org. 59. The novel species are crncoa and ctbtcoa. Adding any of them to the only downlink Org. 57 is sufficient to create the organization. In Org. 59, the two species only take part in one reversible reaction:

$$\text{crncoa} \rightleftharpoons \text{h2o} + \text{ctbtcoa}. \tag{9.1}$$

Table 9.2: Organizations with no interactions between downlink organizations leading to novel species. Species have to be added to the union of downlink organizations to create a larger organization. For starred organizations, any of the novel species is enough to be added to create the organization.

Org.	Novel Species	Creator Sets	Down. Org.
0	ac, ac_ex, co2, co2_ex, fe2_ex, glc-D, glc-D_ex, glyc, glyc_ex, h, h2o, h2o_ex, h_ex, hco3, k, k_ex, lac-D, lac-D_ex, lac-L, lac-L_ex, na1, na1_ex, nh4, nh4_ex, o2, o2_ex, pi, pi_ex, so4_ex, succ, succ_ex	input species: {ac_ex, co2_ex, fe2_ex, glc-D_ex, glyc_ex, h2o_ex, h_ex, k_ex, lac-D_ex, lac-L_ex, na1_ex, nh4_ex, o2_ex, pi_ex, so4_ex, succ_ex}	-
1*	akg, akg_ex	{akg}, {akg_ex}	0
2*	etoh, etoh_ex	{etoh}, {etoh_ex}	0
3*	tyr-L, tyr-L_ex	{tyr-L}, {tyr-L_ex}	0
4*	amet, ametam	{amet}, {ametam}	0
5*	cbasp, dhor-S	{cbasp}, {dhor-S}	0
6*	fum, fum_ex, mal-L	{fum}, {fum_ex}, {mal-L}	0
7*	galur, galur_ex, tagur	{galur}, {galur_ex}, {tagur}	0
8*	fcl-L, fuc-L, fuc-L_ex,	{fcl-L}, {fuc-L}, {fuc-L_ex},	0
9*	2ippm, 3c2hmp, 3c3hmp	{2ippm}, {3c2hmp}, {3c3hmp}	0
10*	ala-D, ala-D_ex, ala-L, ala-L_ex	{ala-D}, {ala-D_ex}, {ala-L}, {ala-L_ex}	0
11	pro-L, pro-L_ex, ura, ura_ex	{pro-L, ura}, {pro-L, ura_ex}, {pro-L_ex, ura}, {pro-L_ex, ura_ex}	0
12	8aonn, amob, dann	{8aonn}	4
13	leu-L, leu-L_ex, ptrc, ptrc_ex, urea, urea_ex	{leu-L, ptrc, urea}, {leu-L, ptrc, urea_ex}, {leu-L, ptrc_ex, urea}, {leu-L, ptrc_ex, urea_ex}, {leu-L_ex, ptrc, urea_ex}, {leu-L_ex, ptrc_ex, urea_ex}, {leu-L_ex, urea, ptrc}, {leu-L_ex, urea, ptrc_ex}	0
14	bbtcoa, crn, crncoa, ctbt, ctbtcoa, gbbtn	{crn, bbtcoa}, {crncoa, gbbtn}, {ctbtcoa, gbbtn}	0
15*	fruur, glcur, glcur_ex	{fruur}, {glcur}, {glcur_ex}	0
16	idon-L, idon-L_ex, no2, no2_ex, ptrc, ptrc_ex, urea, urea_ex	{idon-L, no2, ptrc, urea}, {idon-L, no2, ptrc, urea_ex}, {idon-L, no2, ptrc_ex, urea}, {idon-L, no2, ptrc_ex, urea_ex}, {idon-L, no2_ex, ptrc, urea}, {idon-L, no2_ex, ptrc, urea_ex}, {idon-L, no2_ex, ptrc_ex, urea}, {idon-L, no2_ex, ptrc_ex, urea_ex}, {idon-L_ex, no2, ptrc, urea}, {idon-L_ex, no2, ptrc, urea_ex}, {idon-L_ex, no2, ptrc_ex, urea}, {idon-L_ex, no2, ptrc_ex, urea_ex}, {idon-L_ex, no2_ex, ptrc, urea}, {idon-L_ex, no2_ex, ptrc, urea_ex}, {idon-L_ex, no2_ex, ptrc_ex, urea}, {idon-L_ex, no2_ex, ptrc_ex, urea_ex}	0
17	alac-S, glcn, glcn_ex, micit, pro-L, pro-L_ex, pyr, pyr_ex, q8, q8h2	{pro-L, q8}, {pro-L, q8h2}, {pro-L_ex, q8}, {pro-L_ex, q8h2}	0
18	2h3oppan, alac-S, glcn, glcn_ex, hpyr, micit, pyr, pyr_ex, q8, q8h2	{2h3oppan, q8}, {2h3oppan, q8h2}, {hpyr, q8}, {hpyr, q8h2}	0
19*	gly, gly_ex	{gly}, {gly_ex}	3, 14
20	2ddglcn, 2ddglcn_ex, rml, rmn	{2ddglcn, rml}, {2ddglcn, rmn}, {2ddglcn_ex, rml}, {2ddglcn_ex, rmn}	14

Org.	Novel Species	Creator Sets	Down. Org.
21	alac-S, glcn, glcn_ex, micit, pyr, pyr_ex, q8, q8h2	{q8}, {q8h2}	1, 8
23*	2dmmq8, 2dmmql8	{2dmmq8}, {2dmmql8}	22
24*	8aonn, pmcoa	{8aonn}, {pmcoa}	22
25*	adocbl, cbl1	{adocbl}, {cbl1}	22
26*	idp, itp	{idp}, {itp}	22
27*	lgt-S, gthox, gthrd	{lgt-S}, {gthox}, {gthrd}	22
28*	chol, chol_ex	{chol}, {chol_ex}	23
29*	dann, dtbt	{dann}, {dtbt}	25
30	2shchc, dhna, sbzcoa, ssaltpp, sucbz, thm, thmmp, thmpp	{ssaltpp}, {thm}, {thmmp}, {thmpp}	22
31*	chol, chol_ex	{chol}, {chol_ex}	30
32	5aop, cpppg3, glu1sa, glutrna, hmbil, pheme, ppbng, ppp9, pppg9, trnaglu, uppg3	{glutrna}, {trnaglu}	22
33	gal, gal1p, glyb, glyb_ex	{gal, glyb}, {gal, glyb_ex}, {gal1p, glyb}, {gal1p, glyb_ex}	19, 23
35*	glyb, glyb_ex	{glyb}, {glyb_ex}	32
38*	glyb, glyb_ex	{glyb}, {glyb_ex}	24, 37
39*	chol, chol_ex	{chol}, {chol_ex}	25, 37
40*	dann, dtbt	{dann}, {dtbt}	23, 37
42*	chol, chol_ex	{chol}, {chol_ex}	27, 37
48	12dgr_EC, 3hmrsACP, ACP, acACP, actACP, agpe_EC, agpg_EC, apg_EC, cdpdag1, cdpea, clpn_EC, ddcaACP, etha, g3pe, g3pg, hdca, hdcea, hdeACP, kdo2lipid4, kdo2lipid4L, kdo2lipid4p, kdolipid4, lipa, lipa_cold, lipidA, lipidAds, lipidX, lps_EC, malACP, myrsACP, ocdcea, octeACP, pa_EC, palmACP, pe_EC, pg_EC, pgp_EC, ps_EC, tdeACP, ttdca, ttdcea, u23ga, u3aga, u3hga	{3hmrsACP}, {ACP}, {acACP}, {malACP}	37
49	12dgr_EC, 3hmrsACP, ACP, acACP, actACP, agpe_EC, agpg_EC, apg_EC, cdpdag1, cdpea, clpn_EC, ddcaACP, etha, g3pe, g3pg, hdca, hdcea, hdeACP, kdo2lipid4, kdo2lipid4L, kdo2lipid4p, kdolipid4, lipa, lipa_cold, lipidA, lipidAds, lipidX, lps_EC, malACP, myrsACP, ocdcea, octeACP, pa_EC, palmACP, pe_EC, pg_EC, pgp_EC, ps_EC tdeACP, ttdca, ttdcea, u23ga, u3aga, u3hga,	{3hmrsACP}, {ACP}, {acACP}, {malACP}	41
50*	glyb, glyb_ex	{glyb}, {glyb_ex}	23, 48, 49
51*	dann, dtbt	{dann}, {dtbt}	45, 49
52*	glyb, glyb_ex	{glyb}, {glyb_ex}	25, 45, 49

Org.	Novel Species	Creator Sets	Down. Org.
54	chol, chol_ex, glyb, glyb_ex	{chol, glyb}, {chol, glyb_ex}, {chol_ex, glyb}, {chol_ex, glyb_ex}	12, 24, 26, 51
57*	glyb, glyb_ex	{glyb}, {glyb_ex}	56
59*	crncoa, ctbtcoa	{crncoa}, {ctbtcoa}	57
60*	crncoa, ctbtcoa	{crncoa}, {ctbtcoa}	58
63*	crncoa, ctbtcoa	{crncoa}, {ctbtcoa}	62
64*	crn, ctbt	{crn}, {ctbt}	63

As h2o is present in the downlink organization, adding any of both species will also create the other. All organizations, for which any of the novel species is enough to create the larger organization are starred in Table 9.2.

3. Organizations that need the addition of novel species for self-maintenance.

The last group of organizations combines features of the two former groups. Interactions between their downlink organizations create novel species. However, the created species set is not self-maintaining. Hence, further species need to be added to create a closed self-maintaining set. The four organizations belonging to this group are listed in Table 9.3.

As an example for these organizations, we inspect Org. 36. The novel species gtspmd is created by interactions between the downlink organizations. For gtspmd production, the species atp, gthrd, and spmd are required. While all three downlink Orgs. 27, 30, and 34 provide atp, only Org. 27 provides gthrd, and only Org. 34 provides spmd. Hence, the merger of these organizations leads to the production of gtspmd. However, gthrd is then consumed and no longer maintainable. The species trdrd or trdox need to be added to facilitate the formation of gthrd via cys-L and glucys. Both gthrd and gthox refer to the same polypeptide thioredoxin, in reduced and oxidized form. The reaction network model contains 10 reactions transforming one form into the other and back using varying electron acceptors and donors. These reactions enable the creation of the remaining novel species of Org. 36.

Table 9.3: Organizations, for which interactions between downlink organizations lead to novel species, but not to a self-maintaining species set. Additional species have to be added to create the organization. Starred species are created by interactions between downlink organizations.

Org.	Novel Species	Creator Sets	Down. Org.
C 22	10fthf, 12ppd-S, 12ppd-S_ex, 13dpg, 15dap, 15dap_ex, 1pyr5c, 23ddhb, 23dhb, 23dhba, 23dhdp, 23dhmb, 23dhmp, 25aics, 25drapp, 26dap-LL, 26dap-M, 2ahbut, 2aobut, 2cpr5p, 2dda7p, 2ddg6p, 2dh3dgal6p, 2dhp, 2h3oppan, 2mahmp, 2mcacn, 2mcit, 2me4p, 2obut, 2pg, 34hpp*, 3c4mop, 3dhq, 3dhsk, 3ig3p, 3mob, 3mop, 3pg, 3php, 3psme, 4abut*, 4abut_ex*, 4abutn*, 4abz, 4adcho, 4ampm, 4hbz, 4hthr, 4mop, 4pasp, 4per, 4ppan, 4r5au, 5aizc, 5aprbu, 5apru, 5caiz, 5mthf, 6hmhpt, 6hmhptpp, 6pgc, 6pgl, aacoa, acald, acald_ex, accoa, acg5p, acg5sa, acgam1p, acglu, acorn, acser, actp, ade, ade_ex, adn, adn_ex, adp, adpglc, adphep-D,D, adphep-L,D, agm, ahdt, aicar, air, ala-B, alaala, alac-S*, amp, anth, aps, ara5p, arg-L, arg-L_ex, argsuc, asn-L, asn-L_ex, asp-L, aspsa, atp, camp, cbp, chor, cit, citr-L, coa, db4p, dcamp, dha, dha_ex, dhap, dhf, dhnpt, dhpmp, dhpt, dmlz, dnad, dxyl5p, e4p, eig3p, enter, f6p, fad, fadh2, fc1p, fdp, fe2, fgam, fmn, for, for_ex, fpram, fprica, g1p, g3p, g6p, gam1p, gam6p, gar, gcald, gdp, gdpddman, gdpfuc, gdpmann, gdpofuc, gln-L, glu-D*, glu-L*, glu-L_ex*, glu5p, glu5sa, glx, gly, gly_ex, glyald, glyald_ex, glyc-R, glyc3p, glyclt, glyclt_ex, glycogen, gmhep17bp, gmhep1p, gmhep7p, gmp, gsn, gtp, gua, gua_ex, h2, h2o2, his-L, his-L_ex, hisp, histd, hom-L, hpyr, hxan, hxan_ex, iasp, ichor, icit, ile-L, ile-L_ex, imacp, imp, indole, indole_ex, ins, ins_ex, kdo, kdo8p, lald-L, lys-L, lys-L_ex, malcoa, man1p, man6p, methf, micit*, mlthf, mmcoa-R, mmcoa-S, mnl1p, mthgxl, nac, nad, nadh, nadp, nadph, ncam, nicrnt, nmn, oaa, ohpb, orn, orn_ex, pant-R, paps, pdx5p, pep, phe-L, phe-L_ex, phom, phpyr, phthr, pnto-R, ppa, ppap, ppcoa, pphn, ppi, pppi, pram, pran, prbamp, prbatp, prfp, prlp, pro-L, pro-L_ex, prpp, pser-L, pyam5p, pydam, pydx, pydx5p, pydxn, pyr*, pyr_ex*, quln, r1p, r5p, ribflv, rml1p, ru5p-D, ru5p-L, s7p, sbt6p, ser-L, ser-L_ex, seramp, skm, skm5p, sl26da, sl2a6o, so4, sucarg, succoa, sucglu, sucgsa, suchms, sucorn, sucsal*, tagdp-D, thdp, thf, thr-L, thr-L_ex, trp-L, trp-L_ex, val-L, val-L_ex, xan, xan_ex, xmp, xtsn, xtsn_ex, xu5p-D	{accoa, adp}, {accoa, adpglc}, {accoa, atp}, {adp, mmcoa-R}, {adp, mmcoa-S}, {adp, coa}, {adp, malcoa}, {adp, ppcoa}, {adp, succoa}, {adpglc, coa}, {adpglc, malcoa}, {adpglc, mmcoa-R}, {adpglc, mmcoa-S}, {adpglc, ppcoa}, {adpglc, succoa}, {atp, coa}, {atp, malcoa}, {atp, mmcoa-R}, {atp, mmcoa-S}, {atp, ppcoa}, {atp, succoa}	1, 2, 3, 5, 6, 9, 10, 13
36	2dr1p, 2dr5p, 4hba, 4mpetz, ahcys, cys-L, cyst-L, dad-2, dadp, damp, datp, dgdp, dgmp, dgsn, dgtp, dhptd, din, glucys, gtspmd*, h2s, hcys-L, hmfurn, pap, rhcys, so3, trdox, trdrd	{trdox}, {trdrd}	27, 30, 34
37	2mecdp*, 2ohph*, 2oph*, 2p4c2me*, 3ophb*, 4c2me*, cdp*, ckdo*, cmp*, csn*, ctp*, cytd*, cytd_ex*, dmpp*, frdp*, gal*, gal1p*, grdp*, h2mb4p*, ipdp*, mql8, mqn8, octdp*, orot, orot5p, peptido_EC*, tre,* tre6p*, uaagmda*, uaccg*, uacgam*, uacmam*, uacmamu*, uagmda*, uama*, uamag*, uamr*, udcpdp*, udcpp*, udp*, udpg*, udpgal*, udpgalfur*, udpglcur*, ugmd*, ugmda*, ump*, unaga*, unagamu*, uri*, uri_ex*, utp*	{mql8}, {mqn8}	11, 22
46	chol, chol_ex, gtspmd*	{chol}, {chol_ex}	27, 45

9.3.3 Creator Species Leading to Novel Species

It has been shown how the hierarchy of organizations is constructed from bottom up. Reactions that become available when downlink organizations are merged, or the addition of species can give rise to larger organizations. However, in many cases the expansion of an organization is trivial. Often, the list of novel species contains species that can be converted into each other by reversible reaction. For example, this is true for the transport of metabolites across the cell membrane. The external form of a metabolite is transformed into its internal form and back. In this case, any of the novel species is a creator for the organization (*cf.* starred organizations in Table 9.2). If the set of novel species contains several groups of species that have reversible reactions between them, all species combinations containg one species of each group are creator sets (*e.g.*, Orgs. 13 and 16, Table 9.2). In these cases, all novel species appear at least in one creator set.

In 12 organizations, the list of novel species contains more species than appear as creators: Orgs. 12, 14, 17, 18, 21, 22, 30, 32, 36, 37, 48, and 49. In the simplest case, the remaining novel species are created directly from the creator species. This is the case for Orgs. 12, 14, 30, and 37. Another mechanism is at work for Orgs. 32, 36, 48, and 49. The creator sets consist of single species that refer to the same biological entity in different configurations or states. It is tRNA (trnaglu, glutrna) for Org. 32, thioredoxin in reduced and oxidized form (trdrd, trdox) for Org. 36, and the acyl carrier protein (ACP, 3hmrsACP, malACP, acACP) for Org. 48 and 49. The common feature for all these species is that their biosynthesis is not contained in the reaction network model. They can be converted into their different states, but never created or destroyed in the model. For Org. 17, 18, and 21, ubiquinone8 and ubiquinol8 (q8 and q8h2) play an important role. Although the biosynthesis pathway for these species is contained in the model, it is not present in these organizations. They represent a state, in which the synthesis pathway was shut down after a certain amount of the species was created. The smallest organization containing the biosynthesis pathway is Org. 45. The organization representing the central metabolism Org. 22 has only creator sets containing two species. One of them is atp, adp, or adpglc, and the other aacoa, accoa, coa, malcoa, mmcoa-R, mmcoa-S, ppcoa, or succoa. It is obvious that both

groups again refer to related biological entities. When inspecting the reaction network of Org. 22, we find that the species of the second group can be converted into each other. However, the species are never produced *de novo*, or consumed. The biosynthesis pathway for Coenzyme A (coa) is contained in the reaction network model but not present in this organization. The smallest organization containing coa and its synthesis is Org. 56. The species atp, adp, and adpglc of the first group can also be converted into each other. Another species that is closely associated with this group is amp, which is also easily converted into these species. For the species of the group however, reactions exist that transform the species into other, unrelated species. This consumption is compensated for by the *de novo* synthesis of amp from the central metabolism.

9.4 Flow Conditions

Is the model suited to model cellular growth? In order to investigate this question, we require all species of the network to be produced at positive rates. We implement this by applying flow conditions. Decay reactions are added for all model species, except for those species, for which the network does not contain the *de novo* biosynthesis. These are: 3hmrsACP, ACP, acACP, actACP, apoACP, ddcaACP, hdeACP, malACP, myrsACP, octeACP, palmACP, tdeACP, trdox, trdrd, glutrna, and trnaglu (16 species). The resulting hierarchy of organizations as computed by the heuristic method is depicted in Figure 9.4. Table 9.4 sums up the species of all organizations and their respective creator sets.

Compared to non-flow conditions, the hierarchy collapses to four organizations that are arranged on top of each other. The smallest Org. 0 is unchanged. Organization next in size Org. 1 is identical to Org. 56 under non-flow conditions. This organization contains all species of Org. 22, including the species of the central metabolism. Org. 2 equals Org. 62. Here, pseudo species Biomass is contained for the first time. The acyl carrier protein, which is involved in lipid synthesis, is in several variants a creator for this organization. The novel species of this organization consequently include further species related to lipid synthesis. The largest Org. 3 has no counterpart in the organization hierarchy under non-flow conditions. However, it contains all species of the largest Org. 64

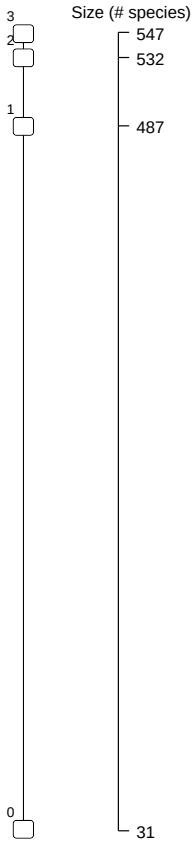


Figure 9.4: Hierarchy of four organizations of the genome-scale metabolic network of *E. coli* under flow conditions, computed using heuristics. Organization 0 is identical to Org. 0 under non-flow conditions, Org. 1 is identical to Org. 56, and Org. 2 to Org. 62. Organization 3 contains Org. 64 except for species crn, crncoa, ctbtcoa, and ctbt, and additionally 15 more species. Pseudo species Biomass representing the production of biomass is contained in Orgs. 2 and 3.

except the species crn, crncoa, ctbtcoa, and ctbt. Several species related to heme biosynthesis appear as novel species in this organization. The synthesis is enabled by the creator species trnaglu. Figure 9.5 sums up the findings.

When flow conditions are applied, all species contained in organizations can be produced at positive rates solely from input species. Hence, with two organizations containing pseudo species Biomass, the analysis shows that the model is capable of converting the input species into all metabolites required for biomass

Table 9.4: Novel species compared to the downlink organization in the four organizations under flow conditions. Species have to be added to create larger organizations.

Org.	Novel Species	Creator Sets	Down. Org.
0	ac, ac_ex, co2, co2_ex, fe2_ex, glc-D, glc-D_ex, glyc, glyc_ex, h, h2o, h2o_ex, h_ex, hco3, k, k_ex, lac-D, lac-D_ex, lac-L, lac-L_ex, na1, na1_ex, nh4, nh4_ex, o2, o2_ex, pi, pi_ex, so4_ex, succ, succ_ex	input species: {ac_ex, co2_ex, fe2_ex, glc-D_ex, glyc_ex, h2o_ex, h_ex, k_ex, lac-D_ex, lac-L_ex, na1_ex, nh4_ex, o2_ex, pi_ex, so4_ex, succ_ex}	-
1	10fthf, 12ppd-S, 12ppd-S_ex, 13dpg, 15dap, 15dap_ex, 1pyr5c, 23ddhb, 23dhb, 23dhba, 23dhdp, 23dhmb, 23dhmp, 25aics, 25drapp, 26dap-LL, 26dap-M, 2ahbut, 2aobut, 2cpr5p, 2dda7p, 2ddg6p, 2dh3dgal6p, 2dhp, 2dmmq8, 2dmmql8, 2dr1p, 2dr5p, 2h3oppan, 2ippm, 2kmb, 2mahmp, 2mcacn, 2mcit, 2me4p, 2mecdp, 2obut, 2ohph, 2ombzl, 2omhmbl, 2ommbl, 2omph, 2oph, 2p4c2me, 2pg, 2shchc, 34hpp, 3c2hmp, 3c3hmp, 3c4mop, 3dhq, 3dhsk, 3ig3p, 3mob, 3mop, 3ophb, 3pg, 3php, 3psme, 4abut, 4abut_ex, 4abutn, 4abz, 4adcho, 4ampm, 4c2me, 4hba, 4hbz, 4hthr, 4mop, 4mpetz, 4pasp, 4per, 4ppan, 4ppcys, 4r5au, 5aizc, 5aprbu, 5apru, 5caiz, 5dglcn, 5mdr1p, 5mdru1p, 5mta, 5mthf, 5mtr, 6hmhpt, 6hmhptpp, 6pgc, 6pgl, N1aspmd, aacoa, acald, acald_ex, accoa, acg5p, acg5sa, acgam1p, acglu, acorn, acser, actp, ade, ade_ex, adn, adn_ex, adp, adpglc, adphep-D,D, adphep-L,D, agm, ahcys, ahdt, aicar, air, akg, akg_ex, ala-B, ala-D, ala-D_ex, ala-L, ala-L_ex, alaala, alac-S, amet, ametam, amp, anth, aps, ara5p, arg-L, arg-L_ex, argsuc, asn-L, asn-L_ex, asp-L, aspsa, atp, camp, cbasp, cbp, cdp, chor, cit, citr-L, ckdo, cmp, coa, csn, ctp, cys-L, cyst-L, cytd, cytd_ex, dad-2, dadp, damp, datp, db4p, dcamp, dcdp, dcmp, dctp, dcyt, dgdp, dgmp, dgsn, dgtp, dha, dha_ex, dhap, dhf, dhna, dhnpt, dhor-S, dhpmp, dhpt, dhptd, din, dkmpp, dmlz, dmpp, dnad, dpcoa, dtdp, dtdp4aaddg, dtdp4addg, dtdp4d6dg, dtdp4d6dm, dtdpglu, dtdprmn, dtmp, dttp, dudp, dump, duri, dutp, dxyl5p, e4p, eca_EC, eig3p, enter, etoh, etoh_ex, f6p, fad, fadh2, fc1p, fdp, fe2, fgam, fmn, for, for_ex, fpram, fprica, frdp, fum, fum_ex, g1p, g3p, g6p, gal, gal1p, gam1p, gam6p, gar, gcald, gdp, gdpddman, gdpfuc, gdpmann, gdpofuc, glcn, glcn_ex, gln-L, glu-D, glu-L, glu-L_ex, glu5p, glu5sa, glucys, glx, gly, gly_ex, glyald, glyald_ex, glyc-R, glyc3p, glyclt, glyclt_ex, glycogen, gmhep17bp, gmhep1p, gmhep7p, gmp, grdp, gsn, gthox, gthrd, gtp, gtspmd, gua, gua_ex, h2, h2mb4p, h2o2, h2s, hcys-L, his-L, his-L_ex, hisp, histd, hmfurn, hom-L, hpyr, hxan, hxan_ex,	{atp, mmcoa-R, nad, q8, trdrd}[1]	0

[1]Computed using heuristics. Due to combinatorial complexity, not all creator sets could be determined. No subset of the stated set is a creator set. However, other creator sets with less than five species might exist.

Org.	Novel Species	Creator Sets	Down. Org.
	iasp, ichor, icit, idon-L, idon-L_ex, ile-L, ile-L_ex, imacp, imp, indole, indole_ex, ins, ins_ex, ipdp, kdo, kdo8p, lald-L, leu-L, leu-L_ex, lgt-S, lys-L, lys-L_ex, mal-L, malcoa, man1p, man6p, met-L, methf, micit, mlthf, mmcoa-R, mmcoa-S, mnl1p, mql8, mqn8, mthgxl, n8aspmd, nac, nad, nadh, nadp, nadph, ncam, nicrnt, nmn, oaa, octdp, ohpb, orn, orn_ex, orot, orot5p, pan4p, pant-R, pap, paps, pdx5p, pep, peptido_EC, phe-L, phe-L_ex, phom, phpyr, phthr, pnto-R, ppa, ppap, ppcoa, pphn, ppi, pppi, pram, pran, prbamp, prbatp, prfp, prlp, pro-L, pro-L_ex, prpp, pser-L, ptrc, ptrc_ex, pyam5p, pydam, pydx, pydx5p, pydxn, pyr, pyr_ex, q8, q8h2, quln, r1p, r5p, rhcys, ribflv, rml1p, ru5p-D, ru5p-L, s7p, sbt6p, sbzcoa, ser-L, ser-L_ex, seramp, skm, skm5p, sl26da, sl2a6o, so3, so4, spmd, ssaltpp, sucarg, sucbz, succoa, sucglu, sucgsa, suchms, sucorn, sucsal, tagdp-D, thdp, thf, thm, thmmp, thmpp, thr-L, thr-L_ex, thym, thymd, thymd_ex, trdox, trdrd, tre, tre6p, trp-L, trp-L_ex, tyr-L, tyr-L_ex, uaagmda, uaccg, uacgam, uacmam, uacmamu, uagmda, uama, uamag, uamr, udcpdp, udcpp, udp, udpg, udpgal, udpgalfur, udpglcur, ugmd, ugmda, ump, unaga, unagamu, unagamuf, ura, ura_ex, urea, urea_ex, uri, uri_ex, utp, val-L, val-L_ex, xan, xan_ex, xmp, xtsn, xtsn_ex, xu5p-D		
2	12dgr_EC, 3hmrsACP, ACP, Biomass, acACP, actACP, agpe_EC, agpg_EC, apg_EC, cdpdag1, cdpea, clpn_EC, ddcaACP, etha, g3pe, g3pg, hdca, hdcea, hdeACP, kdo2lipid4, kdo2lipid4L, kdo2lipid4p, kdolipid4, lipa, lipa_cold, lipidA, lipidAds, lipidX, lps_EC, malACP, myrsACP, ocdcea, octeACP, pa_EC, palmACP, pe_EC, pg_EC, pgp_EC, ps_EC, tdeACP, ttdca, ttdcea, u23ga, u3aga, u3hga	{3hmrsACP}, {ACP}, {acACP}, {malACP}	1
3	5aop, cpppg3, glu1sa, glutrna, hemeO, hmbil, pheme, ppbng, ppp9, pppg9, shcl, sheme, srch, trnaglu, uppg3	{glutrna}, {trnaglu}	2

production, except for the 16 species for which the model does not contain the *de novo* biosynthesis and for which no decay was added. The hierarchy of organizations collapses under flow conditions, featuring much fewer organizations. With all species required to be produced at positive rates under flow conditions, organizations containing certain cycles are no longer organizations. Such cycles can for example be formed by species that are simply transformed into different states. In the simplest case, it can be a reversible reactions, transforming for example a species in its external form into its internal form and back. Applying flow conditions helps to filter out organizations built on such cycles.

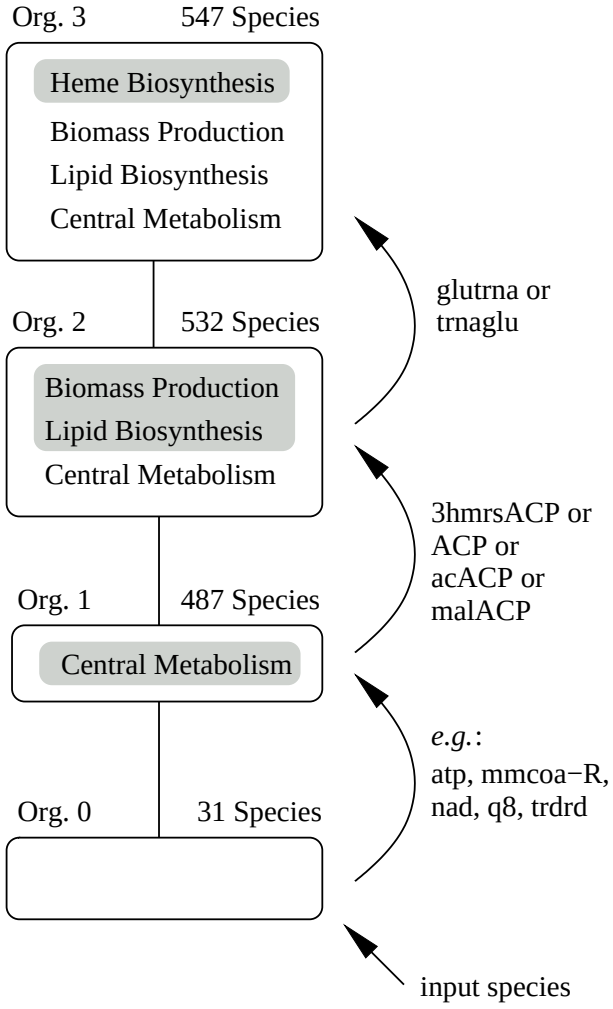


Figure 9.5: Hierarchy of four organizations of the genome-scale metabolic network of *E. coli* under flow conditions, computed using heuristics. The pathways contained in organizations are shown. Gray boxes indicate novel pathways not present in the downlink organization. Creator species are listed on the right hand side. They must be added to the species of the downlink organization to create the larger organization.

9.5 Discussion

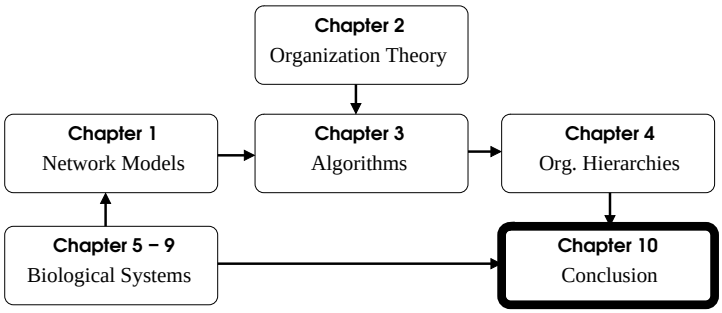
The analysis of a genome-scale metabolic network of *E. coli* has revealed a complicated hierarchy of organizations. Focusing on organizations containing novel

species compared to downlink organizations, it was shown how metabolites related to certain biological functions are distributed in the organizations. For most biological functions, an organization either contains almost none of the relevant species, or a large amount of these species (*cf.* Figure 9.3). A modular organization of the reaction network model is revealed. In any feasible state of the system either the whole module, for example the whole central metabolism, or the whole amino acid metabolism, is contained or not. Modules are not divisible. This modularity however could also be caused by a bias during the creation of the network model. Typically, efforts to elucidate mechanisms of biological systems are focused on specific parts of the whole system. By this approach, reaction network models of high accuracy are attainable for the distinct parts. However, links connecting the different subsystem might not be discovered by this approach. Consequently, links connecting different subsystems might be absent in large scale models spanning many subsystems, suggesting a modular structure of the system.

The organization of the organization hierarchy was analyzed and the mechanisms revealed, by which organizations can give rise to larger organizations. However, many organizations turned out to be trivial due to reversible reactions or species representing the same biological entity in different states forming specific cycles. Applying flow conditions removed these trivial organizations. Only four organizations on top of each other remained under flow conditions: a minimal one, one containing the central metabolism (and several more species), one related to lipid biosynthesis, and one related to heme biosynthesis. However, the creator species for these biological meaningful organizations coincide exactly with those species, for which the reaction network model does not contain the *de novo* biosynthesis. Hence, these organizations could be interpreted as artefacts stemming from the incompleteness of the model. If the species without synthesis were considered to be present at all times (by modeling them as input species), the only remaining organization of the system would be the largest Org. 3.

The largest organization does not contain the whole reaction network as expected, due to 67 deadend metabolites in the model. Species only taking part as educts in reactions cannot be members of any organization. Again, the incompleteness of the model makes the analysis difficult.

The results from analysing a genome-scale metabolic network of *E. coli* stress the need for complete models for applying the theory of chemical organizations. The findings so far suggest however, that even for complete metabolic models containing the synthesis pathways for all network species, the hierarchy of organizations might turn out to be trivial, for example containing only one organization encompassing the whole network. Although such a hierarchy would be trivial, its implications are not. It would indicate that the whole metabolism acts as a unit that cannot be devided. In order to get a more faithful picture of the whole biological system, the integration of metabolism, gene regulation, cell signaling, and further cellular processes into one model is highly desirable.



Chapter 10

Conclusion

Within this thesis, the theory of chemical organizations has been applied as a novel analysis technique to study biochemical reaction network models. As kinetic information is not required for the analysis, the method is well suited for microbiological systems, where such data is often scarce and difficult to come by. The method takes more information about the network structure into account than comparable approaches like elementary flux modes which are solely based on the stoichiometric matrix of the system. The stoichiometric matrix does not contain information about whether a catalyst is required for a reaction to be performed or not. Furthermore, the reactions $1\ A \rightarrow \emptyset$ and $2\ A \rightarrow 1\ A$, two fundamentally different processes, lead to the same entries in the stoichiometric matrix.

The application to models of natural reaction systems has revealed that such models have non-trivial hierarchies of organizations. In several cases, organizations were found to correspond to biological functions or states (*e.g.*, inducible uptake pathways in *E. coli*, Chapter 7). However, for a genome-scale metabolic network, only one biological meaningful organization encompassing almost the whole network was found (Chapter 9). This might seem trivial, but not finding a rich organizational structure also tells something interesting about the system. It indicates that the system is indivisible and only acts as a unit. Inhibitory interactions play an important role in biological systems, but are difficult to implement in a reaction network lacking kinetic information. This problem was first solved for the model of lambda by assessing the consequences of inhibitory interactions

"by hand" after the organizational analysis (Chapter 6). A more elaborate procedure to handle inhibitions in a straightforward manner was then introduced for studying regulated metabolic networks (Chapter 8).

The analysis of a diverse set of reaction network models has opened up several interesting applications for organization theory. With organizations representing all potential steady states and growth conditions of the network, a first assessment of the potential dynamics of the system is possible. The organizational subnetworks then can be separately analyzed by available standard methods to determine steady states. Obviously, analyzing small subnetworks is more feasible than studying the complete reaction network as a whole. To get a trajectory for the whole system, connected organizations that do not overlap can be simulated separately. The trajectory for the complete system can then be assembled from the simulations.

Organization theory can be used to predict the lethality of gene knockouts and to determine metabolites that can be produced from a given substrate species set (Chapter 8). For all genes considered in the network model, organization theory clearly indicates which genes are activated and which are not in different system states.

Another important application of organization theory is the validation of reaction network models. Implausible organizations or the absence of expected organizations can help to identify inconsistencies in the reaction model (Chapter 5). Furthermore, during the modeling process of large networks, the hierarchy of organizations can give hints on inconsistencies in the network, for example on unintended species sinks or sources.

Since species contained in an organization tend to have more reactions among each other than with other species, organizations can help in the visualization of reaction networks. Species of organizations should be grouped closely together in order to obtain a clearer graphical representation of the network (*cf.* Figure 5.2 on Page 71). From a didactical point of view, when large reaction network models are to be explained, it might prove beneficial to first discuss the smaller connected organizations before presenting the complete network as a whole.

Determining the unit species sets containing species that always appear together in organizations (Section 4.2) leads to groups of species that likely share

certain properties, for example molecular structure as in the example of the Martian network (Chapter 5). If one species of a unit species group is found to be present, all other species of this group are present as well. This can be helpful if the presence of species in the real system is to be detected. Only one species of a unit species group needs to be measured in order to assess the presence of all other species of that group.

Reaction networks are dynamical systems. The movement of the system through state space can be mapped to a movement in the space of its organizations (Dittrich and Speroni di Fenizio, 2007), reducing the dimensionality for systems having fewer organizations than species. This mapping from state space to the space of organizations provides a new perspective on the model under consideration as demonstrated for photochemical models (*cf.* Figure 5.6 on Page 86). The introduced intensity functions are a first step in this direction (Section 4.3). The movement between organizations can also be seen as a movement from one system state to another. While the movement to a smaller organization can happen spontaneously, an up movement into a larger organization always requires a constructive perturbation, that means novel species must be injected into the system. If a desired system state is to be achieved, the theory of organizations helps to determine which species need to be added or removed from the system in order to move it into the desired organization.

The results presented in this thesis suggest that the theory of chemical organizations is a valuable tool in the analysis of biological systems. However, many open questions remain. As biological data is often uncertain, it is important to investigate the stability of organization hierarchies against such uncertainties. How does the hierarchy change when the educts or products of a reaction change? What happens when the reversibility is changed? To tackle larger networks as they become available, faster algorithms to compute organizations are desirable. While in this thesis, only single networks of single biological species were analyzed in isolation, it is interesting to compare the organization hierarchies of network models of different biological species. Phylogenetic trees based on the organizational structure are conceivable, exploring the evolution of organizations. To fully exploit the potential of organization theory, integrated models including metabolism, gene regulation, signal transduction, and further cellular processes

are required. Such whole cell models are not yet available. However, the rapid developments in the -omics fields, along with several current research initiatives in systems biology will help to create models of ever increasing scope and accuracy, suited for organizational analysis.

Objects of the analysis using the theory of chemical organizations within this thesis have always been models of biochemical systems, and one photochemical model. Hence, all findings first apply to these models. Whether they also apply to the real system depends on how faithful the models represent the real systems. It remains an open question, whether and how the organization hierarchy of a model is mirrored in the real biological systems.

Appendix A

Photochemical Model of the Mars Atmosphere

A.1 List of Species

	Symbol	Molecular Species
1	hv	incident sunlight
	O	O
	O2	O_2
	O3	O_3
5	O1D	$O(^1D)$, O in excited state
	H	H
	H2	H_2
	OH	OH
	HO2	HO_2
10	H2O	H_2O
	H2O2	H_2O_2
	CO	CO
	CO2	CO_2
	N	N
15	N2	N_2
	N2D	$N(^2D)$, N in excited state
	NO	NO
	NO2	NO_2
	NO3	NO_3
20	N2O	N_2O

	Symbol	Molecular Species
	N2O5	N_2O_5
	HNO2	HNO_2
	HNO3	HNO_3
	HO2NO2	HO_2NO_2
25	e	free electron
	Op	O^+
	O2p	O_2^+
	CO2p	CO_2^+
	CO2Hp	CO_2H^+
30	HO2grain	$(HO_2)_{\text{grain}}$, adsorbed HO_2
	grain	aerosol particle

A.2 List of Reactions (Dayside)

R1	$\emptyset \rightarrow h\nu$
R2	$O_2 + h\nu \rightarrow 2\ O$
R3	$O_2 + h\nu \rightarrow O + O(^1D)$
R4	$O_3 + h\nu \rightarrow O_2 + O$
R5	$O_3 + h\nu \rightarrow O_2 + O(^1D)$
R6	$O_3 + h\nu \rightarrow 3\ O$
R7	$H_2 + h\nu \rightarrow 2\ H$
R8	$OH + h\nu \rightarrow O + H$
R9	$HO_2 + h\nu \rightarrow OH + O$
R10	$H_2O + h\nu \rightarrow H + OH$
R11	$H_2O + h\nu \rightarrow H_2 + O(^1D)$
R12	$H_2O + h\nu \rightarrow 2\ H + O$
R13	$H_2O_2 + h\nu \rightarrow 2\ OH$
R14	$CO_2 + h\nu \rightarrow CO + O$
R15	$CO_2 + h\nu \rightarrow CO + O(^1D)$
R16	$2\ O \rightarrow O_2$
R17	$O + O_2 + N_2 \rightarrow O_3 + N_2$
R18	$O + O_2 + CO_2 \rightarrow O_3 + CO_2$
R19	$O + O_3 \rightarrow 2\ O_2$
R20	$O + CO \rightarrow CO_2$
R21	$O(^1D) + O_2 \rightarrow O + O_2$
R22	$O(^1D) + O_3 \rightarrow 2\ O_2$

R23	$O(^1D) + O_3 \rightarrow O_2 + 2\ O$
R24	$O(^1D) + H_2 \rightarrow H + OH$
R25	$O(^1D) + CO_2 \rightarrow O + CO_2$
R26	$O(^1D) + H_2O \rightarrow 2\ OH$
R27	$2\ H \rightarrow H_2$
R28	$H + O_2 \rightarrow HO_2$
R29	$H + O_3 \rightarrow OH + O_2$
R30	$H + HO_2 \rightarrow 2\ OH$
R31	$H + HO_2 \rightarrow H_2 + O_2$
R32	$H + HO_2 \rightarrow H_2O + O$
R33	$O + H_2 \rightarrow OH + H$
R34	$O + OH \rightarrow O_2 + H$
R35	$O + HO_2 \rightarrow OH + O_2$
R36	$O + H_2O_2 \rightarrow OH + HO_2$
R37	$2\ OH \rightarrow H_2O + O$
R38	$2\ OH \rightarrow H_2O_2$
R39	$OH + O_3 \rightarrow HO_2 + O_2$
R40	$OH + H_2 \rightarrow H_2O + H$
R41	$OH + HO_2 \rightarrow H_2O + O_2$
R42	$OH + H_2O_2 \rightarrow H_2O + HO_2$
R43	$OH + CO \rightarrow CO_2 + H$
R44	$HO_2 + O_3 \rightarrow OH + 2\ O_2$
R45	$2\ HO_2 \rightarrow H_2O_2 + O_2$
R46	$N_2 \rightarrow 2\ N$
R47	$N_2 \rightarrow 2\ N(^2D)$
R48	$NO + h\nu \rightarrow N + O$
R49	$NO_2 + h\nu \rightarrow NO + O$
R50	$NO_3 + h\nu \rightarrow NO_2 + O$
R51	$NO_3 + h\nu \rightarrow NO + O_2$
R52	$N_2O + h\nu \rightarrow N_2 + O(^1D)$
R53	$N_2O_5 + h\nu \rightarrow NO_2 + NO_3$
R54	$HNO_2 + h\nu \rightarrow OH + NO$
R55	$HNO_3 + h\nu \rightarrow NO_2 + OH$
R56	$HO_2NO_2 + h\nu \rightarrow HO_2 + NO_2$
R57	$N + O_2 \rightarrow NO + O$
R58	$N + O_3 \rightarrow NO + O_2$
R59	$N + OH \rightarrow NO + H$

R60	$N + HO_2 \rightarrow NO + OH$
R61	$N + NO \rightarrow N_2 + O$
R62	$N + NO_2 \rightarrow N_2O + O$
R63	$N(^2D) + O \rightarrow N + O$
R64	$N(^2D) + CO_2 \rightarrow NO + CO$
R65	$N(^2D) + N_2 \rightarrow N + N_2$
R66	$N(^2D) + NO \rightarrow N_2 + O$
R67	$O + NO \rightarrow NO_2$
R68	$O + NO_2 \rightarrow NO + O_2$
R69	$O + NO_2 \rightarrow NO_3$
R70	$O + NO_3 \rightarrow O_2 + NO_2$
R71	$O + HO_2NO_2 \rightarrow OH + NO_2 + O_2$
R72	$O(^1D) + N_2 \rightarrow O + N_2$
R73	$O(^1D) + N_2 \rightarrow N_2O$
R74	$O(^1D) + N_2O \rightarrow 2\ NO$
R75	$O(^1D) + N_2O \rightarrow N_2 + O_2$
R76	$NO + O_3 \rightarrow NO_2 + O_2$
R77	$NO + HO_2 \rightarrow NO_2 + OH$
R78	$NO + NO_3 \rightarrow 2\ NO_2$
R79	$H + NO_2 \rightarrow OH + NO$
R80	$H + NO_3 \rightarrow OH + NO_2$
R81	$OH + NO \rightarrow HNO_2$
R82	$OH + NO_2 \rightarrow HNO_3$
R83	$OH + NO_3 \rightarrow HO_2 + NO_2$
R84	$OH + HNO_2 \rightarrow H_2O + NO_2$
R85	$OH + HNO_3 \rightarrow H_2O + NO_3$
R86	$OH + HO_2NO_2 \rightarrow H_2O + NO_2 + O_2$
R87	$HO_2 + NO_2 \rightarrow HO_2NO_2$
R88	$HO_2 + NO_3 \rightarrow O_2 + HNO_3$
R89	$NO_2 + O_3 \rightarrow NO_3 + O_2$
R90	$NO_2 + NO_3 \rightarrow N_2O_5$
R91	$NO_2 + NO_3 \rightarrow NO + NO_2 + O_2$
R92	$O + h\nu \rightarrow O^+ + e$
R93	$O_2 + h\nu \rightarrow O_2^+ + e$
R94	$CO_2 + h\nu \rightarrow CO_2^+ + e$
R95	$CO_2 + h\nu \rightarrow CO + O^+ + e$
R96	$O_2^+ + e \rightarrow 2\ O$

R97	$CO_2^+ + e \rightarrow CO + O$
R98	$O^+ + CO_2 \rightarrow O_2^+ + CO$
R99	$O + CO_2^+ \rightarrow O_2^+ + CO$
R100	$O + CO_2^+ \rightarrow O^+ + CO_2$
R101	$CO_2^+ + H_2 \rightarrow CO_2H^+ + H$
R102	$CO_2H^+ + e \rightarrow CO_2 + H$
R103	$HO_2 + \text{grain} \rightarrow (HO_2)_{\text{grain}}$
R104	$(HO_2)_{\text{grain}} + OH \rightarrow H_2O + O_2 + \text{grain}$

A.3 Species Connectivity

The connectivity of a molecular species refers to how often it occurs as an educt and product.

Species	Connectivity
O	42
O_2	35
OH	35
$h\nu$	28
NO_2	25
NO	20
H	20
HO_2	19
$O(^1D)$	15
O_3	14
CO_2	14
NO_3	13
N_2	13
H_2O	13
N	10
CO	9
H_2	8
e	7
N_2O	5
$N(^2D)$	5
H_2O_2	5
CO_2^+	5
O^+	4
O_2^+	4

Species	**Connectivity**
HO_2NO_2	4
HNO_3	4
HNO_2	3
N_2O_5	2
CO_2H^+	2
$(HO_2)_{grain}$	2
grain	2

A.4 List of Dayside Organizations

The list of all organizations is divided into five groups according to the characteristics of the scaled intensity profiles (see Section 5.3.5).

Group (i)

ID	# Species	Species
1470	11	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, O^+, O_2^+, CO_2^+
1473	12	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1474	12	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, O^+, O_2^+, CO_2^+, grain
1475	13	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1476	18	$h\nu$, e, O_3, O_2, O, $O(^1D)$, H_2, H, OH, HO_2, H_2O, H_2O_2, CO_2, CO, O^+, O_2^+, CO_2^+, CO_2H^+
1477	19	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, O^+, O_2^+, CO_2^+
1478	20	$h\nu$, e, O_3, O_2, O, $O(^1D)$, H_2, H, OH, HO_2, H_2O, H_2O_2, CO_2, CO, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1479	20	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1480	20	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, O^+, O_2^+, CO_2^+, grain
1481	21	$h\nu$, e, O_3, O_2, O, $O(^1D)$, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1482	29	$h\nu$, e, O_3, O_2, O, $O(^1D)$, H_2, H, OH, HO_2, H_2O, H_2O_2, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, HNO_2, HNO_3, HO_2NO_2, O^+, O_2^+, CO_2^+, CO_2H^+
1483	31	$h\nu$, e, O_3, O_2, O, $O(^1D)$, H_2, H, OH, HO_2, H_2O, H_2O_2, CO_2, CO, N_2, N, $N(^2D)$, NO, NO_2, NO_3, N_2O, N_2O_5, HNO_2, HNO_3, HO_2NO_2, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain

Group (ii)

ID	# Species	Species
3	2	$h\nu$, CO
13	3	$h\nu$, e, CO
19	3	$h\nu$, O(^{1}D), CO
30	3	$h\nu$, CO, N(^{2}D)
31	3	$h\nu$, CO, O^+
32	3	$h\nu$, CO, O_2^+
33	3	$h\nu$, CO, CO_2^+
34	3	$h\nu$, CO, CO_2H^+
35	3	$h\nu$, CO, $(HO_2)_{grain}$
36	3	$h\nu$, CO, grain
71	4	$h\nu$, e, O(^{1}D), CO
73	4	$h\nu$, e, CO, N(^{2}D)
74	4	$h\nu$, e, CO, O^+
75	4	$h\nu$, e, CO, $(HO_2)_{grain}$
76	4	$h\nu$, e, CO, grain
88	4	$h\nu$, O(^{1}D), CO, N(^{2}D)
89	4	$h\nu$, O(^{1}D), CO, O^+
90	4	$h\nu$, O(^{1}D), CO, O_2^+
91	4	$h\nu$, O(^{1}D), CO, CO_2^+
92	4	$h\nu$, O(^{1}D), CO, CO_2H^+
93	4	$h\nu$, O(^{1}D), CO, $(HO_2)_{grain}$
94	4	$h\nu$, O(^{1}D), CO, grain
138	4	$h\nu$, CO, N(^{2}D), O^+
139	4	$h\nu$, CO, N(^{2}D), O_2^+
140	4	$h\nu$, CO, N(^{2}D), CO_2^+
141	4	$h\nu$, CO, N(^{2}D), CO_2H^+
142	4	$h\nu$, CO, N(^{2}D), $(HO_2)_{grain}$
143	4	$h\nu$, CO, N(^{2}D), grain
144	4	$h\nu$, CO, O^+, O_2^+
145	4	$h\nu$, CO, O^+, CO_2^+
146	4	$h\nu$, CO, O^+, CO_2H^+
147	4	$h\nu$, CO, O^+, $(HO_2)_{grain}$
148	4	$h\nu$, CO, O^+, grain
149	4	$h\nu$, CO, O_2^+, CO_2^+
150	4	$h\nu$, CO, O_2^+, CO_2H^+
151	4	$h\nu$, CO, O_2^+, $(HO_2)_{grain}$
152	4	$h\nu$, CO, O_2^+, grain
153	4	$h\nu$, CO, CO_2^+, CO_2H^+
154	4	$h\nu$, CO, CO_2^+, $(HO_2)_{grain}$
155	4	$h\nu$, CO, CO_2^+, grain
156	4	$h\nu$, CO, CO_2H^+, $(HO_2)_{grain}$
157	4	$h\nu$, CO, CO_2H^+, grain

ID	# Species	Species
158	4	$h\nu$, CO, $(HO_2)_{grain}$, grain
215	5	$h\nu$, e, $O(^1D)$, CO, $N(^2D)$
216	5	$h\nu$, e, $O(^1D)$, CO, O^+
217	5	$h\nu$, e, $O(^1D)$, CO, $(HO_2)_{grain}$
218	5	$h\nu$, e, $O(^1D)$, CO, grain
240	5	$h\nu$, e, CO, $N(^2D)$, O^+
241	5	$h\nu$, e, CO, $N(^2D)$, $(HO_2)_{grain}$
242	5	$h\nu$, e, CO, $N(^2D)$, grain
243	5	$h\nu$, e, CO, O^+, $(HO_2)_{grain}$
244	5	$h\nu$, e, CO, O^+, grain
245	5	$h\nu$, e, CO, $(HO_2)_{grain}$, grain
263	5	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+
264	5	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O_2^+
265	5	$h\nu$, $O(^1D)$, CO, $N(^2D)$, CO_2^+
266	5	$h\nu$, $O(^1D)$, CO, $N(^2D)$, CO_2H^+
267	5	$h\nu$, $O(^1D)$, CO, $N(^2D)$, $(HO_2)_{grain}$
268	5	$h\nu$, $O(^1D)$, CO, $N(^2D)$, grain
269	5	$h\nu$, $O(^1D)$, CO, O^+, O_2^+
270	5	$h\nu$, $O(^1D)$, CO, O^+, CO_2^+
271	5	$h\nu$, $O(^1D)$, CO, O^+, CO_2H^+
272	5	$h\nu$, $O(^1D)$, CO, O^+, $(HO_2)_{grain}$
273	5	$h\nu$, $O(^1D)$, CO, O^+, grain
274	5	$h\nu$, $O(^1D)$, CO, O_2^+, CO_2^+
275	5	$h\nu$, $O(^1D)$, CO, O_2^+, CO_2H^+
276	5	$h\nu$, $O(^1D)$, CO, O_2^+, $(HO_2)_{grain}$
277	5	$h\nu$, $O(^1D)$, CO, O_2^+, grain
278	5	$h\nu$, $O(^1D)$, CO, CO_2^+, CO_2H^+
279	5	$h\nu$, $O(^1D)$, CO, CO_2^+, $(HO_2)_{grain}$
280	5	$h\nu$, $O(^1D)$, CO, CO_2^+, grain
281	5	$h\nu$, $O(^1D)$, CO, CO_2H^+, $(HO_2)_{grain}$
282	5	$h\nu$, $O(^1D)$, CO, CO_2H^+, grain
283	5	$h\nu$, $O(^1D)$, CO, $(HO_2)_{grain}$, grain
389	5	$h\nu$, CO, $N(^2D)$, O^+, O_2^+
390	5	$h\nu$, CO, $N(^2D)$, O^+, CO_2^+
391	5	$h\nu$, CO, $N(^2D)$, O^+, CO_2H^+
392	5	$h\nu$, CO, $N(^2D)$, O^+, $(HO_2)_{grain}$
393	5	$h\nu$, CO, $N(^2D)$, O^+, grain
394	5	$h\nu$, CO, $N(^2D)$, O_2^+, CO_2^+
395	5	$h\nu$, CO, $N(^2D)$, O_2^+, CO_2H^+
396	5	$h\nu$, CO, $N(^2D)$, O_2^+, $(HO_2)_{grain}$
397	5	$h\nu$, CO, $N(^2D)$, O_2^+, grain
398	5	$h\nu$, CO, $N(^2D)$, CO_2^+, CO_2H^+
399	5	$h\nu$, CO, $N(^2D)$, CO_2^+, $(HO_2)_{grain}$
400	5	$h\nu$, CO, $N(^2D)$, CO_2^+, grain

ID	# Species	Species
401	5	$h\nu$, CO, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$
402	5	$h\nu$, CO, N(^{2}D), CO_2H^+, grain
403	5	$h\nu$, CO, N(^{2}D), $(HO_2)_{grain}$, grain
404	5	$h\nu$, CO, O^+, O_2^+, CO_2^+
405	5	$h\nu$, CO, O^+, O_2^+, CO_2H^+
406	5	$h\nu$, CO, O^+, O_2^+, $(HO_2)_{grain}$
407	5	$h\nu$, CO, O^+, O_2^+, grain
408	5	$h\nu$, CO, O^+, CO_2^+, CO_2H^+
409	5	$h\nu$, CO, O^+, CO_2^+, $(HO_2)_{grain}$
410	5	$h\nu$, CO, O^+, CO_2^+, grain
411	5	$h\nu$, CO, O^+, CO_2H^+, $(HO_2)_{grain}$
412	5	$h\nu$, CO, O^+, CO_2H^+, grain
413	5	$h\nu$, CO, O^+, $(HO_2)_{grain}$, grain
414	5	$h\nu$, CO, O_2^+, CO_2^+, CO_2H^+
415	5	$h\nu$, CO, O_2^+, CO_2^+, $(HO_2)_{grain}$
416	5	$h\nu$, CO, O_2^+, CO_2^+, grain
417	5	$h\nu$, CO, O_2^+, CO_2H^+, $(HO_2)_{grain}$
418	5	$h\nu$, CO, O_2^+, CO_2H^+, grain
419	5	$h\nu$, CO, O_2^+, $(HO_2)_{grain}$, grain
420	5	$h\nu$, CO, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
421	5	$h\nu$, CO, CO_2^+, CO_2H^+, grain
422	5	$h\nu$, CO, CO_2^+, $(HO_2)_{grain}$, grain
423	5	$h\nu$, CO, CO_2H^+, $(HO_2)_{grain}$, grain
495	6	$h\nu$, e, O(^{1}D), CO, N(^{2}D), O^+
496	6	$h\nu$, e, O(^{1}D), CO, N(^{2}D), $(HO_2)_{grain}$
497	6	$h\nu$, e, O(^{1}D), CO, N(^{2}D), grain
498	6	$h\nu$, e, O(^{1}D), CO, O^+, $(HO_2)_{grain}$
499	6	$h\nu$, e, O(^{1}D), CO, O^+, grain
500	6	$h\nu$, e, O(^{1}D), CO, $(HO_2)_{grain}$, grain
535	6	$h\nu$, e, CO, N(^{2}D), O^+, $(HO_2)_{grain}$
536	6	$h\nu$, e, CO, N(^{2}D), O^+, grain
537	6	$h\nu$, e, CO, N(^{2}D), $(HO_2)_{grain}$, grain
538	6	$h\nu$, e, CO, O^+, $(HO_2)_{grain}$, grain
565	6	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, O_2^+
566	6	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, CO_2^+
567	6	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, CO_2H^+
568	6	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, $(HO_2)_{grain}$
569	6	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, grain
570	6	$h\nu$, O(^{1}D), CO, N(^{2}D), O_2^+, CO_2^+
571	6	$h\nu$, O(^{1}D), CO, N(^{2}D), O_2^+, CO_2H^+
572	6	$h\nu$, O(^{1}D), CO, N(^{2}D), O_2^+, $(HO_2)_{grain}$
573	6	$h\nu$, O(^{1}D), CO, N(^{2}D), O_2^+, grain
574	6	$h\nu$, O(^{1}D), CO, N(^{2}D), CO_2^+, CO_2H^+
575	6	$h\nu$, O(^{1}D), CO, N(^{2}D), CO_2^+, $(HO_2)_{grain}$

ID	# Species	Species
576	6	$h\nu$, O(^{1}D), CO, N(^{2}D), CO_2^+, grain
577	6	$h\nu$, O(^{1}D), CO, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$
578	6	$h\nu$, O(^{1}D), CO, N(^{2}D), CO_2H^+, grain
579	6	$h\nu$, O(^{1}D), CO, N(^{2}D), $(HO_2)_{grain}$, grain
580	6	$h\nu$, O(^{1}D), CO, O^+, O_2^+, CO_2^+
581	6	$h\nu$, O(^{1}D), CO, O^+, O_2^+, CO_2H^+
582	6	$h\nu$, O(^{1}D), CO, O^+, O_2^+, $(HO_2)_{grain}$
583	6	$h\nu$, O(^{1}D), CO, O^+, O_2^+, grain
584	6	$h\nu$, O(^{1}D), CO, O^+, CO_2^+, CO_2H^+
585	6	$h\nu$, O(^{1}D), CO, O^+, CO_2^+, $(HO_2)_{grain}$
586	6	$h\nu$, O(^{1}D), CO, O^+, CO_2^+, grain
587	6	$h\nu$, O(^{1}D), CO, O^+, CO_2H^+, $(HO_2)_{grain}$
588	6	$h\nu$, O(^{1}D), CO, O^+, CO_2H^+, grain
589	6	$h\nu$, O(^{1}D), CO, O^+, $(HO_2)_{grain}$, grain
590	6	$h\nu$, O(^{1}D), CO, O_2^+, CO_2^+, CO_2H^+
591	6	$h\nu$, O(^{1}D), CO, O_2^+, CO_2^+, $(HO_2)_{grain}$
592	6	$h\nu$, O(^{1}D), CO, O_2^+, CO_2^+, grain
593	6	$h\nu$, O(^{1}D), CO, O_2^+, CO_2H^+, $(HO_2)_{grain}$
594	6	$h\nu$, O(^{1}D), CO, O_2^+, CO_2H^+, grain
595	6	$h\nu$, O(^{1}D), CO, O_2^+, $(HO_2)_{grain}$, grain
596	6	$h\nu$, O(^{1}D), CO, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
597	6	$h\nu$, O(^{1}D), CO, CO_2^+, CO_2H^+, grain
598	6	$h\nu$, O(^{1}D), CO, CO_2^+, $(HO_2)_{grain}$, grain
599	6	$h\nu$, O(^{1}D), CO, CO_2H^+, $(HO_2)_{grain}$, grain
761	6	$h\nu$, CO, N(^{2}D), O^+, O_2^+, CO_2^+
762	6	$h\nu$, CO, N(^{2}D), O^+, O_2^+, CO_2H^+
763	6	$h\nu$, CO, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$
764	6	$h\nu$, CO, N(^{2}D), O^+, O_2^+, grain
765	6	$h\nu$, CO, N(^{2}D), O^+, CO_2^+, CO_2H^+
766	6	$h\nu$, CO, N(^{2}D), O^+, CO_2^+, $(HO_2)_{grain}$
767	6	$h\nu$, CO, N(^{2}D), O^+, CO_2^+, grain
768	6	$h\nu$, CO, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$
769	6	$h\nu$, CO, N(^{2}D), O^+, CO_2H^+, grain
770	6	$h\nu$, CO, N(^{2}D), O^+, $(HO_2)_{grain}$, grain
771	6	$h\nu$, CO, N(^{2}D), O_2^+, CO_2^+, CO_2H^+
772	6	$h\nu$, CO, N(^{2}D), O_2^+, CO_2^+, $(HO_2)_{grain}$
773	6	$h\nu$, CO, N(^{2}D), O_2^+, CO_2^+, grain
774	6	$h\nu$, CO, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$
775	6	$h\nu$, CO, N(^{2}D), O_2^+, CO_2H^+, grain
776	6	$h\nu$, CO, N(^{2}D), O_2^+, $(HO_2)_{grain}$, grain
777	6	$h\nu$, CO, N(^{2}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$
778	6	$h\nu$, CO, N(^{2}D), CO_2^+, CO_2H^+, grain
779	6	$h\nu$, CO, N(^{2}D), CO_2^+, $(HO_2)_{grain}$, grain
780	6	$h\nu$, CO, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$, grain

ID	# Species	Species
781	6	$h\nu$, CO, O^+, O_2^+, CO_2^+, CO_2H^+
782	6	$h\nu$, CO, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
783	6	$h\nu$, CO, O^+, O_2^+, CO_2^+, grain
784	6	$h\nu$, CO, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
785	6	$h\nu$, CO, O^+, O_2^+, CO_2H^+, grain
786	6	$h\nu$, CO, O^+, O_2^+, $(HO_2)_{grain}$, grain
787	6	$h\nu$, CO, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
788	6	$h\nu$, CO, O^+, CO_2^+, CO_2H^+, grain
789	6	$h\nu$, CO, O^+, CO_2^+, $(HO_2)_{grain}$, grain
790	6	$h\nu$, CO, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
791	6	$h\nu$, CO, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
792	6	$h\nu$, CO, O_2^+, CO_2^+, CO_2H^+, grain
793	6	$h\nu$, CO, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
794	6	$h\nu$, CO, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
795	6	$h\nu$, CO, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
852	7	$h\nu$, e, $O(^1D)$, CO, $N(^2D)$, O^+, $(HO_2)_{grain}$
853	7	$h\nu$, e, $O(^1D)$, CO, $N(^2D)$, O^+, grain
854	7	$h\nu$, e, $O(^1D)$, CO, $N(^2D)$, $(HO_2)_{grain}$, grain
855	7	$h\nu$, e, $O(^1D)$, CO, O^+, $(HO_2)_{grain}$, grain
891	7	$h\nu$, e, CO, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
928	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, CO_2^+
929	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, CO_2H^+
930	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$
931	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, grain
932	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, CO_2^+, CO_2H^+
933	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$
934	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, CO_2^+, grain
935	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$
936	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, CO_2H^+, grain
937	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
938	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+
939	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O_2^+, CO_2^+, $(HO_2)_{grain}$
940	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O_2^+, CO_2^+, grain
941	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O_2^+, CO_2H^+, $(HO_2)_{grain}$
942	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O_2^+, CO_2H^+, grain
943	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O_2^+, $(HO_2)_{grain}$, grain
944	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
945	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, CO_2^+, CO_2H^+, grain
946	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, CO_2^+, $(HO_2)_{grain}$, grain
947	7	$h\nu$, $O(^1D)$, CO, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$, grain
948	7	$h\nu$, $O(^1D)$, CO, O^+, O_2^+, CO_2^+, CO_2H^+
949	7	$h\nu$, $O(^1D)$, CO, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
950	7	$h\nu$, $O(^1D)$, CO, O^+, O_2^+, CO_2^+, grain
951	7	$h\nu$, $O(^1D)$, CO, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$

ID	# Species	Species
952	7	$h\nu$, $O(^1D)$, CO, O^+, O_2^+, CO_2H^+, grain
953	7	$h\nu$, $O(^1D)$, CO, O^+, O_2^+, $(HO_2)_{grain}$, grain
954	7	$h\nu$, $O(^1D)$, CO, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
955	7	$h\nu$, $O(^1D)$, CO, O^+, CO_2^+, CO_2H^+, grain
956	7	$h\nu$, $O(^1D)$, CO, O^+, CO_2^+, $(HO_2)_{grain}$, grain
957	7	$h\nu$, $O(^1D)$, CO, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
958	7	$h\nu$, $O(^1D)$, CO, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
959	7	$h\nu$, $O(^1D)$, CO, O_2^+, CO_2^+, CO_2H^+, grain
960	7	$h\nu$, $O(^1D)$, CO, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
961	7	$h\nu$, $O(^1D)$, CO, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
962	7	$h\nu$, $O(^1D)$, CO, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1124	7	$h\nu$, CO, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+
1125	7	$h\nu$, CO, $N(^2D)$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1126	7	$h\nu$, CO, $N(^2D)$, O^+, O_2^+, CO_2^+, grain
1127	7	$h\nu$, CO, $N(^2D)$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1128	7	$h\nu$, CO, $N(^2D)$, O^+, O_2^+, CO_2H^+, grain
1129	7	$h\nu$, CO, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$, grain
1130	7	$h\nu$, CO, $N(^2D)$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1131	7	$h\nu$, CO, $N(^2D)$, O^+, CO_2^+, CO_2H^+, grain
1132	7	$h\nu$, CO, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$, grain
1133	7	$h\nu$, CO, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1134	7	$h\nu$, CO, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1135	7	$h\nu$, CO, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, grain
1136	7	$h\nu$, CO, $N(^2D)$, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1137	7	$h\nu$, CO, $N(^2D)$, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1138	7	$h\nu$, CO, $N(^2D)$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1139	7	$h\nu$, CO, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1140	7	$h\nu$, CO, O^+, O_2^+, CO_2^+, CO_2H^+, grain
1141	7	$h\nu$, CO, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1142	7	$h\nu$, CO, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1143	7	$h\nu$, CO, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1144	7	$h\nu$, CO, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1173	8	$h\nu$, e, $O(^1D)$, CO, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
1230	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+
1231	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1232	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, CO_2^+, grain
1233	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1234	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, CO_2H^+, grain
1235	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$, grain
1236	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1237	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, CO_2^+, CO_2H^+, grain
1238	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$, grain
1239	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1240	8	$h\nu$, $O(^1D)$, CO, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$

ID	# Species	Species
1241	8	$h\nu$, O(^{1}D), CO, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, grain
1242	8	$h\nu$, O(^{1}D), CO, N(^{2}D), O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1243	8	$h\nu$, O(^{1}D), CO, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1244	8	$h\nu$, O(^{1}D), CO, N(^{2}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1245	8	$h\nu$, O(^{1}D), CO, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1246	8	$h\nu$, O(^{1}D), CO, O^+, O_2^+, CO_2^+, CO_2H^+, grain
1247	8	$h\nu$, O(^{1}D), CO, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1248	8	$h\nu$, O(^{1}D), CO, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1249	8	$h\nu$, O(^{1}D), CO, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1250	8	$h\nu$, O(^{1}D), CO, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1356	8	$h\nu$, CO, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1357	8	$h\nu$, CO, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, grain
1358	8	$h\nu$, CO, N(^{2}D), O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1359	8	$h\nu$, CO, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1360	8	$h\nu$, CO, N(^{2}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1361	8	$h\nu$, CO, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1362	8	$h\nu$, CO, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1399	9	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1400	9	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, grain
1401	9	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1402	9	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1403	9	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1404	9	$h\nu$, O(^{1}D), CO, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1405	9	$h\nu$, O(^{1}D), CO, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1449	9	$h\nu$, CO, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1459	10	$h\nu$, O(^{1}D), CO, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain

Group (iii)

ID	# Species	Species
29	3	$h\nu$, CO, N
72	4	$h\nu$, e, CO, N
87	4	$h\nu$, O(^{1}D), CO, N
131	4	$h\nu$, CO, N, N(^{2}D)
132	4	$h\nu$, CO, N, O^+
133	4	$h\nu$, CO, N, O_2^+
134	4	$h\nu$, CO, N, CO_2^+
135	4	$h\nu$, CO, N, CO_2H^+
136	4	$h\nu$, CO, N, $(HO_2)_{grain}$
137	4	$h\nu$, CO, N, grain
230	5	$h\nu$, e, O(^{1}D), CO, N
236	5	$h\nu$, e, CO, N, N(^{2}D)
237	5	$h\nu$, e, CO, N, O^+
238	5	$h\nu$, e, CO, N, $(HO_2)_{grain}$
239	5	$h\nu$, e, CO, N, grain
256	5	$h\nu$, O(^{1}D), CO, N, N(^{2}D)
257	5	$h\nu$, O(^{1}D), CO, N, O^+
258	5	$h\nu$, O(^{1}D), CO, N, O_2^+
259	5	$h\nu$, O(^{1}D), CO, N, CO_2^+
260	5	$h\nu$, O(^{1}D), CO, N, CO_2H^+
261	5	$h\nu$, O(^{1}D), CO, N, $(HO_2)_{grain}$
262	5	$h\nu$, O(^{1}D), CO, N, grain
368	5	$h\nu$, CO, N, N(^{2}D), O^+
369	5	$h\nu$, CO, N, N(^{2}D), O_2^+
370	5	$h\nu$, CO, N, N(^{2}D), CO_2^+
371	5	$h\nu$, CO, N, N(^{2}D), CO_2H^+
372	5	$h\nu$, CO, N, N(^{2}D), $(HO_2)_{grain}$
373	5	$h\nu$, CO, N, N(^{2}D), grain
374	5	$h\nu$, CO, N, O^+, O_2^+
375	5	$h\nu$, CO, N, O^+, CO_2^+
376	5	$h\nu$, CO, N, O^+, CO_2H^+
377	5	$h\nu$, CO, N, O^+, $(HO_2)_{grain}$
378	5	$h\nu$, CO, N, O^+, grain
379	5	$h\nu$, CO, N, O_2^+, CO_2^+
380	5	$h\nu$, CO, N, O_2^+, CO_2H^+
381	5	$h\nu$, CO, N, O_2^+, $(HO_2)_{grain}$
382	5	$h\nu$, CO, N, O_2^+, grain
383	5	$h\nu$, CO, N, CO_2^+, CO_2H^+
384	5	$h\nu$, CO, N, CO_2^+, $(HO_2)_{grain}$
385	5	$h\nu$, CO, N, CO_2^+, grain
386	5	$h\nu$, CO, N, CO_2H^+, $(HO_2)_{grain}$
387	5	$h\nu$, CO, N, CO_2H^+, grain

ID	# Species	Species
388	5	$h\nu$, CO, N, $(HO_2)_{grain}$, grain
494	6	$h\nu$, e, $O(^1D)$, CO, N, grain
513	6	$h\nu$, e, $O(^1D)$, CO, N, $N(^2D)$
514	6	$h\nu$, e, $O(^1D)$, CO, N, O^+
515	6	$h\nu$, e, $O(^1D)$, CO, N, $(HO_2)_{grain}$
529	6	$h\nu$, e, CO, N, $N(^2D)$, O^+
530	6	$h\nu$, e, CO, N, $N(^2D)$, $(HO_2)_{grain}$
531	6	$h\nu$, e, CO, N, $N(^2D)$, grain
532	6	$h\nu$, e, CO, N, O^+, $(HO_2)_{grain}$
533	6	$h\nu$, e, CO, N, O^+, grain
534	6	$h\nu$, e, CO, N, $(HO_2)_{grain}$, grain
544	6	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+
545	6	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O_2^+
546	6	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, CO_2^+
547	6	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, CO_2H^+
548	6	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, $(HO_2)_{grain}$
549	6	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, grain
550	6	$h\nu$, $O(^1D)$, CO, N, O^+, O_2^+
551	6	$h\nu$, $O(^1D)$, CO, N, O^+, CO_2^+
552	6	$h\nu$, $O(^1D)$, CO, N, O^+, CO_2H^+
553	6	$h\nu$, $O(^1D)$, CO, N, O^+, $(HO_2)_{grain}$
554	6	$h\nu$, $O(^1D)$, CO, N, O^+, grain
555	6	$h\nu$, $O(^1D)$, CO, N, O_2^+, CO_2^+
556	6	$h\nu$, $O(^1D)$, CO, N, O_2^+, CO_2H^+
557	6	$h\nu$, $O(^1D)$, CO, N, O_2^+, $(HO_2)_{grain}$
558	6	$h\nu$, $O(^1D)$, CO, N, O_2^+, grain
559	6	$h\nu$, $O(^1D)$, CO, N, CO_2^+, CO_2H^+
560	6	$h\nu$, $O(^1D)$, CO, N, CO_2^+, $(HO_2)_{grain}$
561	6	$h\nu$, $O(^1D)$, CO, N, CO_2^+, grain
562	6	$h\nu$, $O(^1D)$, CO, N, CO_2H^+, $(HO_2)_{grain}$
563	6	$h\nu$, $O(^1D)$, CO, N, CO_2H^+, grain
564	6	$h\nu$, $O(^1D)$, CO, N, $(HO_2)_{grain}$, grain
726	6	$h\nu$, CO, N, $N(^2D)$, O^+, O_2^+
727	6	$h\nu$, CO, N, $N(^2D)$, O^+, CO_2^+
728	6	$h\nu$, CO, N, $N(^2D)$, O^+, CO_2H^+
729	6	$h\nu$, CO, N, $N(^2D)$, O^+, $(HO_2)_{grain}$
730	6	$h\nu$, CO, N, $N(^2D)$, O^+, grain
731	6	$h\nu$, CO, N, $N(^2D)$, O_2^+, CO_2^+
732	6	$h\nu$, CO, N, $N(^2D)$, O_2^+, CO_2H^+
733	6	$h\nu$, CO, N, $N(^2D)$, O_2^+, $(HO_2)_{grain}$
734	6	$h\nu$, CO, N, $N(^2D)$, O_2^+, grain
735	6	$h\nu$, CO, N, $N(^2D)$, CO_2^+, CO_2H^+
736	6	$h\nu$, CO, N, $N(^2D)$, CO_2^+, $(HO_2)_{grain}$
737	6	$h\nu$, CO, N, $N(^2D)$, CO_2^+, grain

ID	# Species	Species
738	6	$h\nu$, CO, N, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$
739	6	$h\nu$, CO, N, N(^{2}D), CO_2H^+, grain
740	6	$h\nu$, CO, N, N(^{2}D), $(HO_2)_{grain}$, grain
741	6	$h\nu$, CO, N, O^+, O_2^+, CO_2^+
742	6	$h\nu$, CO, N, O^+, O_2^+, CO_2H^+
743	6	$h\nu$, CO, N, O^+, O_2^+, $(HO_2)_{grain}$
744	6	$h\nu$, CO, N, O^+, O_2^+, grain
745	6	$h\nu$, CO, N, O^+, CO_2^+, CO_2H^+
746	6	$h\nu$, CO, N, O^+, CO_2^+, $(HO_2)_{grain}$
747	6	$h\nu$, CO, N, O^+, CO_2^+, grain
748	6	$h\nu$, CO, N, O^+, CO_2H^+, $(HO_2)_{grain}$
749	6	$h\nu$, CO, N, O^+, CO_2H^+, grain
750	6	$h\nu$, CO, N, O^+, $(HO_2)_{grain}$, grain
751	6	$h\nu$, CO, N, O_2^+, CO_2^+, CO_2H^+
752	6	$h\nu$, CO, N, O_2^+, CO_2^+, $(HO_2)_{grain}$
753	6	$h\nu$, CO, N, O_2^+, CO_2^+, grain
754	6	$h\nu$, CO, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$
755	6	$h\nu$, CO, N, O_2^+, CO_2H^+, grain
756	6	$h\nu$, CO, N, O_2^+, $(HO_2)_{grain}$, grain
757	6	$h\nu$, CO, N, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
758	6	$h\nu$, CO, N, CO_2^+, CO_2H^+, grain
759	6	$h\nu$, CO, N, CO_2^+, $(HO_2)_{grain}$, grain
760	6	$h\nu$, CO, N, CO_2H^+, $(HO_2)_{grain}$, grain
867	7	$h\nu$, e, O(^{1}D), CO, N, N(^{2}D), O^+
868	7	$h\nu$, e, O(^{1}D), CO, N, N(^{2}D), $(HO_2)_{grain}$
869	7	$h\nu$, e, O(^{1}D), CO, N, N(^{2}D), grain
870	7	$h\nu$, e, O(^{1}D), CO, N, O^+, $(HO_2)_{grain}$
871	7	$h\nu$, e, O(^{1}D), CO, N, O^+, grain
872	7	$h\nu$, e, O(^{1}D), CO, N, $(HO_2)_{grain}$, grain
887	7	$h\nu$, e, CO, N, N(^{2}D), O^+, $(HO_2)_{grain}$
888	7	$h\nu$, e, CO, N, N(^{2}D), O^+, grain
889	7	$h\nu$, e, CO, N, N(^{2}D), $(HO_2)_{grain}$, grain
890	7	$h\nu$, e, CO, N, O^+, $(HO_2)_{grain}$, grain
893	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, O_2^+
894	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, CO_2^+
895	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, CO_2H^+
896	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, $(HO_2)_{grain}$
897	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, grain
898	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O_2^+, CO_2^+
899	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O_2^+, CO_2H^+
900	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O_2^+, $(HO_2)_{grain}$
901	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O_2^+, grain
902	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), CO_2^+, CO_2H^+
903	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), CO_2^+, $(HO_2)_{grain}$

ID	# Species	Species
904	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), CO_2^+, grain
905	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$
906	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), CO_2H^+, grain
907	7	$h\nu$, O(^{1}D), CO, N, N(^{2}D), $(HO_2)_{grain}$, grain
908	7	$h\nu$, O(^{1}D), CO, N, O^+, O_2^+, CO_2^+
909	7	$h\nu$, O(^{1}D), CO, N, O^+, O_2^+, CO_2H^+
910	7	$h\nu$, O(^{1}D), CO, N, O^+, O_2^+, $(HO_2)_{grain}$
911	7	$h\nu$, O(^{1}D), CO, N, O^+, O_2^+, grain
912	7	$h\nu$, O(^{1}D), CO, N, O^+, CO_2^+, CO_2H^+
913	7	$h\nu$, O(^{1}D), CO, N, O^+, CO_2^+, $(HO_2)_{grain}$
914	7	$h\nu$, O(^{1}D), CO, N, O^+, CO_2^+, grain
915	7	$h\nu$, O(^{1}D), CO, N, O^+, CO_2H^+, $(HO_2)_{grain}$
916	7	$h\nu$, O(^{1}D), CO, N, O^+, CO_2H^+, grain
917	7	$h\nu$, O(^{1}D), CO, N, O^+, $(HO_2)_{grain}$, grain
918	7	$h\nu$, O(^{1}D), CO, N, O_2^+, CO_2^+, CO_2H^+
919	7	$h\nu$, O(^{1}D), CO, N, O_2^+, CO_2^+, $(HO_2)_{grain}$
920	7	$h\nu$, O(^{1}D), CO, N, O_2^+, CO_2^+, grain
921	7	$h\nu$, O(^{1}D), CO, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$
922	7	$h\nu$, O(^{1}D), CO, N, O_2^+, CO_2H^+, grain
923	7	$h\nu$, O(^{1}D), CO, N, O_2^+, $(HO_2)_{grain}$, grain
924	7	$h\nu$, O(^{1}D), CO, N, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
925	7	$h\nu$, O(^{1}D), CO, N, CO_2^+, CO_2H^+, grain
926	7	$h\nu$, O(^{1}D), CO, N, CO_2^+, $(HO_2)_{grain}$, grain
927	7	$h\nu$, O(^{1}D), CO, N, CO_2H^+, $(HO_2)_{grain}$, grain
1089	7	$h\nu$, CO, N, N(^{2}D), O^+, O_2^+, CO_2^+
1090	7	$h\nu$, CO, N, N(^{2}D), O^+, O_2^+, CO_2H^+
1091	7	$h\nu$, CO, N, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$
1092	7	$h\nu$, CO, N, N(^{2}D), O^+, O_2^+, grain
1093	7	$h\nu$, CO, N, N(^{2}D), O^+, CO_2^+, CO_2H^+
1094	7	$h\nu$, CO, N, N(^{2}D), O^+, CO_2^+, $(HO_2)_{grain}$
1095	7	$h\nu$, CO, N, N(^{2}D), O^+, CO_2^+, grain
1096	7	$h\nu$, CO, N, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$
1097	7	$h\nu$, CO, N, N(^{2}D), O^+, CO_2H^+, grain
1098	7	$h\nu$, CO, N, N(^{2}D), O^+, $(HO_2)_{grain}$, grain
1099	7	$h\nu$, CO, N, N(^{2}D), O_2^+, CO_2^+, CO_2H^+
1100	7	$h\nu$, CO, N, N(^{2}D), O_2^+, CO_2^+, $(HO_2)_{grain}$
1101	7	$h\nu$, CO, N, N(^{2}D), O_2^+, CO_2^+, grain
1102	7	$h\nu$, CO, N, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$
1103	7	$h\nu$, CO, N, N(^{2}D), O_2^+, CO_2H^+, grain
1104	7	$h\nu$, CO, N, N(^{2}D), O_2^+, $(HO_2)_{grain}$, grain
1105	7	$h\nu$, CO, N, N(^{2}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1106	7	$h\nu$, CO, N, N(^{2}D), CO_2^+, CO_2H^+, grain
1107	7	$h\nu$, CO, N, N(^{2}D), CO_2^+, $(HO_2)_{grain}$, grain
1108	7	$h\nu$, CO, N, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$, grain

ID	# Species	Species
1109	7	$h\nu$, CO, N, O^+, O_2^+, CO_2^+, CO_2H^+
1110	7	$h\nu$, CO, N, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1111	7	$h\nu$, CO, N, O^+, O_2^+, CO_2^+, grain
1112	7	$h\nu$, CO, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1113	7	$h\nu$, CO, N, O^+, O_2^+, CO_2H^+, grain
1114	7	$h\nu$, CO, N, O^+, O_2^+, $(HO_2)_{grain}$, grain
1115	7	$h\nu$, CO, N, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1116	7	$h\nu$, CO, N, O^+, CO_2^+, CO_2H^+, grain
1117	7	$h\nu$, CO, N, O^+, CO_2^+, $(HO_2)_{grain}$, grain
1118	7	$h\nu$, CO, N, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1119	7	$h\nu$, CO, N, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1120	7	$h\nu$, CO, N, O_2^+, CO_2^+, CO_2H^+, grain
1121	7	$h\nu$, CO, N, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1122	7	$h\nu$, CO, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1123	7	$h\nu$, CO, N, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1183	8	$h\nu$, e, $O(^1D)$, CO, N, $N(^2D)$, O^+, $(HO_2)_{grain}$
1184	8	$h\nu$, e, $O(^1D)$, CO, N, $N(^2D)$, O^+, grain
1185	8	$h\nu$, e, $O(^1D)$, CO, N, $N(^2D)$, $(HO_2)_{grain}$, grain
1186	8	$h\nu$, e, $O(^1D)$, CO, N, O^+, $(HO_2)_{grain}$, grain
1194	8	$h\nu$, e, CO, N, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
1195	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2^+
1196	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2H^+
1197	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$
1198	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, grain
1199	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+
1200	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$
1201	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, CO_2^+, grain
1202	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$
1203	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, CO_2H^+, grain
1204	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
1205	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+
1206	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O_2^+, CO_2^+, $(HO_2)_{grain}$
1207	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O_2^+, CO_2^+, grain
1208	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1209	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O_2^+, CO_2H^+, grain
1210	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O_2^+, $(HO_2)_{grain}$, grain
1211	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1212	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, CO_2^+, CO_2H^+, grain
1213	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, CO_2^+, $(HO_2)_{grain}$, grain
1214	8	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$, grain
1215	8	$h\nu$, $O(^1D)$, CO, N, O^+, O_2^+, CO_2^+, CO_2H^+
1216	8	$h\nu$, $O(^1D)$, CO, N, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1217	8	$h\nu$, $O(^1D)$, CO, N, O^+, O_2^+, CO_2^+, grain
1218	8	$h\nu$, $O(^1D)$, CO, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$

ID	# Species	Species
1219	8	$h\nu$, $O(^1D)$, CO, N, O^+, O_2^+, CO_2H^+, grain
1220	8	$h\nu$, $O(^1D)$, CO, N, O^+, O_2^+, $(HO_2)_{grain}$, grain
1221	8	$h\nu$, $O(^1D)$, CO, N, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1222	8	$h\nu$, $O(^1D)$, CO, N, O^+, CO_2^+, CO_2H^+, grain
1223	8	$h\nu$, $O(^1D)$, CO, N, O^+, CO_2^+, $(HO_2)_{grain}$, grain
1224	8	$h\nu$, $O(^1D)$, CO, N, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1225	8	$h\nu$, $O(^1D)$, CO, N, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1226	8	$h\nu$, $O(^1D)$, CO, N, O_2^+, CO_2^+, CO_2H^+, grain
1227	8	$h\nu$, $O(^1D)$, CO, N, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1228	8	$h\nu$, $O(^1D)$, CO, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1229	8	$h\nu$, $O(^1D)$, CO, N, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1335	8	$h\nu$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+
1336	8	$h\nu$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1337	8	$h\nu$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2^+, grain
1338	8	$h\nu$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1339	8	$h\nu$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2H^+, grain
1340	8	$h\nu$, CO, N, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$, grain
1341	8	$h\nu$, CO, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1342	8	$h\nu$, CO, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+, grain
1343	8	$h\nu$, CO, N, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$, grain
1344	8	$h\nu$, CO, N, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1345	8	$h\nu$, CO, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1346	8	$h\nu$, CO, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, grain
1347	8	$h\nu$, CO, N, $N(^2D)$, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1348	8	$h\nu$, CO, N, $N(^2D)$, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1349	8	$h\nu$, CO, N, $N(^2D)$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1350	8	$h\nu$, CO, N, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1351	8	$h\nu$, CO, N, O^+, O_2^+, CO_2^+, CO_2H^+, grain
1352	8	$h\nu$, CO, N, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1353	8	$h\nu$, CO, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1354	8	$h\nu$, CO, N, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1355	8	$h\nu$, CO, N, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1376	9	$h\nu$, e, $O(^1D)$, CO, N, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
1378	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+
1379	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1380	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2^+, grain
1381	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1382	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, CO_2H^+, grain
1383	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$, grain
1384	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1385	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+, grain
1386	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$, grain
1387	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1388	9	$h\nu$, $O(^1D)$, CO, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$

ID	# Species	Species
1389	9	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, grain
1390	9	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1391	9	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1392	9	$h\nu$, O(^{1}D), CO, N, N(^{2}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1393	9	$h\nu$, O(^{1}D), CO, N, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1394	9	$h\nu$, O(^{1}D), CO, N, O^+, O_2^+, CO_2^+, CO_2H^+, grain
1395	9	$h\nu$, O(^{1}D), CO, N, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1396	9	$h\nu$, O(^{1}D), CO, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1397	9	$h\nu$, O(^{1}D), CO, N, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1398	9	$h\nu$, O(^{1}D), CO, N, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1442	9	$h\nu$, CO, N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1443	9	$h\nu$, CO, N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, grain
1444	9	$h\nu$, CO, N, N(^{2}D), O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1445	9	$h\nu$, CO, N, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1446	9	$h\nu$, CO, N, N(^{2}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1447	9	$h\nu$, CO, N, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1448	9	$h\nu$, CO, N, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1452	10	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1453	10	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, grain
1454	10	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1455	10	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1456	10	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1457	10	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1458	10	$h\nu$, O(^{1}D), CO, N, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1469	10	$h\nu$, CO, N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1471	11	$h\nu$, O(^{1}D), CO, N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain

Group (iv)

ID	# Species	Species
1	2	$h\nu$, e
2	2	$h\nu$, $O(^1D)$
4	2	$h\nu$, N
5	2	$h\nu$, $N(^2D)$
6	2	$h\nu$, O^+
7	2	$h\nu$, O_2^+
8	2	$h\nu$, CO_2^+
9	2	$h\nu$, CO_2H^+
12	3	$h\nu$, e, $O(^1D)$
14	3	$h\nu$, e, N
15	3	$h\nu$, e, $N(^2D)$
16	3	$h\nu$, e, O^+
17	3	$h\nu$, e, $(HO_2)_{grain}$
18	3	$h\nu$, e, grain
20	3	$h\nu$, $O(^1D)$, N
21	3	$h\nu$, $O(^1D)$, $N(^2D)$
22	3	$h\nu$, $O(^1D)$, O^+
23	3	$h\nu$, $O(^1D)$, O_2^+
24	3	$h\nu$, $O(^1D)$, CO_2^+
25	3	$h\nu$, $O(^1D)$, CO_2H^+
26	3	$h\nu$, $O(^1D)$, $(HO_2)_{grain}$
27	3	$h\nu$, $O(^1D)$, grain
28	3	$h\nu$, H_2, H
37	3	$h\nu$, N, $N(^2D)$
38	3	$h\nu$, N, O^+
39	3	$h\nu$, N, O_2^+
40	3	$h\nu$, N, CO_2^+
41	3	$h\nu$, N, CO_2H^+
42	3	$h\nu$, N, $(HO_2)_{grain}$
43	3	$h\nu$, N, grain
44	3	$h\nu$, $N(^2D)$, O^+
45	3	$h\nu$, $N(^2D)$, O_2^+
46	3	$h\nu$, $N(^2D)$, CO_2^+
47	3	$h\nu$, $N(^2D)$, CO_2H^+
48	3	$h\nu$, $N(^2D)$, $(HO_2)_{grain}$
49	3	$h\nu$, $N(^2D)$, grain
50	3	$h\nu$, O^+, O_2^+
51	3	$h\nu$, O^+, CO_2^+
52	3	$h\nu$, O^+, CO_2H^+
53	3	$h\nu$, O^+, $(HO_2)_{grain}$
54	3	$h\nu$, O^+, grain
55	3	$h\nu$, O_2^+, CO_2^+

ID	# Species	Species
56	3	$h\nu$, O_2^+, CO_2H^+
57	3	$h\nu$, O_2^+, $(HO_2)_{grain}$
58	3	$h\nu$, O_2^+, grain
59	3	$h\nu$, CO_2^+, CO_2H^+
60	3	$h\nu$, CO_2^+, $(HO_2)_{grain}$
61	3	$h\nu$, CO_2^+, grain
62	3	$h\nu$, CO_2H^+, $(HO_2)_{grain}$
63	3	$h\nu$, CO_2H^+, grain
64	3	$h\nu$, $(HO_2)_{grain}$, grain
65	4	$h\nu$, e, $O(^1D)$, N
66	4	$h\nu$, e, $O(^1D)$, $N(^2D)$
67	4	$h\nu$, e, $O(^1D)$, O^+
68	4	$h\nu$, e, $O(^1D)$, $(HO_2)_{grain}$
69	4	$h\nu$, e, $O(^1D)$, grain
70	4	$h\nu$, e, H_2, H
77	4	$h\nu$, e, N, $N(^2D)$
78	4	$h\nu$, e, N, O^+
79	4	$h\nu$, e, N, $(HO_2)_{grain}$
80	4	$h\nu$, e, N, grain
81	4	$h\nu$, e, $N(^2D)$, O^+
82	4	$h\nu$, e, $N(^2D)$, $(HO_2)_{grain}$
83	4	$h\nu$, e, $N(^2D)$, grain
84	4	$h\nu$, e, O^+, $(HO_2)_{grain}$
85	4	$h\nu$, e, O^+, grain
86	4	$h\nu$, e, $(HO_2)_{grain}$, grain
95	4	$h\nu$, $O(^1D)$, N, $N(^2D)$
96	4	$h\nu$, $O(^1D)$, N, O^+
97	4	$h\nu$, $O(^1D)$, N, O_2^+
98	4	$h\nu$, $O(^1D)$, N, CO_2^+
99	4	$h\nu$, $O(^1D)$, N, CO_2H^+
100	4	$h\nu$, $O(^1D)$, N, $(HO_2)_{grain}$
101	4	$h\nu$, $O(^1D)$, N, grain
102	4	$h\nu$, $O(^1D)$, $N(^2D)$, O^+
103	4	$h\nu$, $O(^1D)$, $N(^2D)$, O_2^+
104	4	$h\nu$, $O(^1D)$, $N(^2D)$, CO_2^+
105	4	$h\nu$, $O(^1D)$, $N(^2D)$, CO_2H^+
106	4	$h\nu$, $O(^1D)$, $N(^2D)$, $(HO_2)_{grain}$
107	4	$h\nu$, $O(^1D)$, $N(^2D)$, grain
108	4	$h\nu$, $O(^1D)$, O^+, O_2^+
109	4	$h\nu$, $O(^1D)$, O^+, CO_2^+
110	4	$h\nu$, $O(^1D)$, O^+, CO_2H^+
111	4	$h\nu$, $O(^1D)$, O^+, $(HO_2)_{grain}$
112	4	$h\nu$, $O(^1D)$, O^+, grain
113	4	$h\nu$, $O(^1D)$, O_2^+, CO_2^+

ID	# Species	Species
114	4	$h\nu$, $O(^1D)$, O_2^+, CO_2H^+
115	4	$h\nu$, $O(^1D)$, O_2^+, $(HO_2)_{grain}$
116	4	$h\nu$, $O(^1D)$, O_2^+, grain
117	4	$h\nu$, $O(^1D)$, CO_2^+, CO_2H^+
118	4	$h\nu$, $O(^1D)$, CO_2^+, $(HO_2)_{grain}$
119	4	$h\nu$, $O(^1D)$, CO_2^+, grain
120	4	$h\nu$, $O(^1D)$, CO_2H^+, $(HO_2)_{grain}$
121	4	$h\nu$, $O(^1D)$, CO_2H^+, grain
122	4	$h\nu$, $O(^1D)$, $(HO_2)_{grain}$, grain
123	4	$h\nu$, H_2, H, CO
124	4	$h\nu$, H_2, H, N
125	4	$h\nu$, H_2, H, $N(^2D)$
126	4	$h\nu$, H_2, H, O^+
127	4	$h\nu$, H_2, H, O_2^+
128	4	$h\nu$, H_2, H, CO_2H^+
129	4	$h\nu$, H_2, H, $(HO_2)_{grain}$
130	4	$h\nu$, H_2, H, grain
159	4	$h\nu$, N, $N(^2D)$, O^+
160	4	$h\nu$, N, $N(^2D)$, O_2^+
161	4	$h\nu$, N, $N(^2D)$, CO_2^+
162	4	$h\nu$, N, $N(^2D)$, CO_2H^+
163	4	$h\nu$, N, $N(^2D)$, $(HO_2)_{grain}$
164	4	$h\nu$, N, $N(^2D)$, grain
165	4	$h\nu$, N, O^+, O_2^+
166	4	$h\nu$, N, O^+, CO_2^+
167	4	$h\nu$, N, O^+, CO_2H^+
168	4	$h\nu$, N, O^+, $(HO_2)_{grain}$
169	4	$h\nu$, N, O^+, grain
170	4	$h\nu$, N, O_2^+, CO_2^+
171	4	$h\nu$, N, O_2^+, CO_2H^+
172	4	$h\nu$, N, O_2^+, $(HO_2)_{grain}$
173	4	$h\nu$, N, O_2^+, grain
174	4	$h\nu$, N, CO_2^+, CO_2H^+
175	4	$h\nu$, N, CO_2^+, $(HO_2)_{grain}$
176	4	$h\nu$, N, CO_2^+, grain
177	4	$h\nu$, N, CO_2H^+, $(HO_2)_{grain}$
178	4	$h\nu$, N, CO_2H^+, grain
179	4	$h\nu$, N, $(HO_2)_{grain}$, grain
180	4	$h\nu$, $N(^2D)$, O^+, O_2^+
181	4	$h\nu$, $N(^2D)$, O^+, CO_2^+
182	4	$h\nu$, $N(^2D)$, O^+, CO_2H^+
183	4	$h\nu$, $N(^2D)$, O^+, $(HO_2)_{grain}$
184	4	$h\nu$, $N(^2D)$, O^+, grain
185	4	$h\nu$, $N(^2D)$, O_2^+, CO_2^+

ID	# Species	Species
186	4	$h\nu$, N(^{2}D), O_2^+, CO_2H^+
187	4	$h\nu$, N(^{2}D), O_2^+, $(HO_2)_{grain}$
188	4	$h\nu$, N(^{2}D), O_2^+, grain
189	4	$h\nu$, N(^{2}D), CO_2^+, CO_2H^+
190	4	$h\nu$, N(^{2}D), CO_2^+, $(HO_2)_{grain}$
191	4	$h\nu$, N(^{2}D), CO_2^+, grain
192	4	$h\nu$, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$
193	4	$h\nu$, N(^{2}D), CO_2H^+, grain
194	4	$h\nu$, N(^{2}D), $(HO_2)_{grain}$, grain
195	4	$h\nu$, O^+, O_2^+, CO_2^+
196	4	$h\nu$, O^+, O_2^+, CO_2H^+
197	4	$h\nu$, O^+, O_2^+, $(HO_2)_{grain}$
198	4	$h\nu$, O^+, O_2^+, grain
199	4	$h\nu$, O^+, CO_2^+, CO_2H^+
200	4	$h\nu$, O^+, CO_2^+, $(HO_2)_{grain}$
201	4	$h\nu$, O^+, CO_2^+, grain
202	4	$h\nu$, O^+, CO_2H^+, $(HO_2)_{grain}$
203	4	$h\nu$, O^+, CO_2H^+, grain
204	4	$h\nu$, O^+, $(HO_2)_{grain}$, grain
205	4	$h\nu$, O_2^+, CO_2^+, CO_2H^+
206	4	$h\nu$, O_2^+, CO_2^+, $(HO_2)_{grain}$
207	4	$h\nu$, O_2^+, CO_2^+, grain
208	4	$h\nu$, O_2^+, CO_2H^+, $(HO_2)_{grain}$
209	4	$h\nu$, O_2^+, CO_2H^+, grain
210	4	$h\nu$, O_2^+, $(HO_2)_{grain}$, grain
211	4	$h\nu$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
212	4	$h\nu$, CO_2^+, CO_2H^+, grain
213	4	$h\nu$, CO_2^+, $(HO_2)_{grain}$, grain
214	4	$h\nu$, CO_2H^+, $(HO_2)_{grain}$, grain
219	5	$h\nu$, e, O(^{1}D), N, N(^{2}D)
220	5	$h\nu$, e, O(^{1}D), N, O^+
221	5	$h\nu$, e, O(^{1}D), N, $(HO_2)_{grain}$
222	5	$h\nu$, e, O(^{1}D), N, grain
223	5	$h\nu$, e, O(^{1}D), N(^{2}D), O^+
224	5	$h\nu$, e, O(^{1}D), N(^{2}D), $(HO_2)_{grain}$
225	5	$h\nu$, e, O(^{1}D), N(^{2}D), grain
226	5	$h\nu$, e, O(^{1}D), O^+, $(HO_2)_{grain}$
227	5	$h\nu$, e, O(^{1}D), O^+, grain
228	5	$h\nu$, e, O(^{1}D), $(HO_2)_{grain}$, grain
229	5	$h\nu$, e, H_2, H, CO
231	5	$h\nu$, e, H_2, H, N
232	5	$h\nu$, e, H_2, H, N(^{2}D)
233	5	$h\nu$, e, H_2, H, O^+
234	5	$h\nu$, e, H_2, H, $(HO_2)_{grain}$

ID	# Species	Species
235	5	$h\nu$, e, H_2, H, grain
246	5	$h\nu$, e, N, N(^{2}D), O^+
247	5	$h\nu$, e, N, N(^{2}D), $(HO_2)_{grain}$
248	5	$h\nu$, e, N, N(^{2}D), grain
249	5	$h\nu$, e, N, O^+, $(HO_2)_{grain}$
250	5	$h\nu$, e, N, O^+, grain
251	5	$h\nu$, e, N, $(HO_2)_{grain}$, grain
252	5	$h\nu$, e, N(^{2}D), O^+, $(HO_2)_{grain}$
253	5	$h\nu$, e, N(^{2}D), O^+, grain
254	5	$h\nu$, e, N(^{2}D), $(HO_2)_{grain}$, grain
255	5	$h\nu$, e, O^+, $(HO_2)_{grain}$, grain
284	5	$h\nu$, O(^{1}D), N, N(^{2}D), O^+
285	5	$h\nu$, O(^{1}D), N, N(^{2}D), O_2^+
286	5	$h\nu$, O(^{1}D), N, N(^{2}D), CO_2^+
287	5	$h\nu$, O(^{1}D), N, N(^{2}D), CO_2H^+
288	5	$h\nu$, O(^{1}D), N, N(^{2}D), $(HO_2)_{grain}$
289	5	$h\nu$, O(^{1}D), N, N(^{2}D), grain
290	5	$h\nu$, O(^{1}D), N, O^+, O_2^+
291	5	$h\nu$, O(^{1}D), N, O^+, CO_2^+
292	5	$h\nu$, O(^{1}D), N, O^+, CO_2H^+
293	5	$h\nu$, O(^{1}D), N, O^+, $(HO_2)_{grain}$
294	5	$h\nu$, O(^{1}D), N, O^+, grain
295	5	$h\nu$, O(^{1}D), N, O_2^+, CO_2^+
296	5	$h\nu$, O(^{1}D), N, O_2^+, CO_2H^+
297	5	$h\nu$, O(^{1}D), N, O_2^+, $(HO_2)_{grain}$
298	5	$h\nu$, O(^{1}D), N, O_2^+, grain
299	5	$h\nu$, O(^{1}D), N, CO_2^+, CO_2H^+
300	5	$h\nu$, O(^{1}D), N, CO_2^+, $(HO_2)_{grain}$
301	5	$h\nu$, O(^{1}D), N, CO_2^+, grain
302	5	$h\nu$, O(^{1}D), N, CO_2H^+, $(HO_2)_{grain}$
303	5	$h\nu$, O(^{1}D), N, CO_2H^+, grain
304	5	$h\nu$, O(^{1}D), N, $(HO_2)_{grain}$, grain
305	5	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+
306	5	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2^+
307	5	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2H^+
308	5	$h\nu$, O(^{1}D), N(^{2}D), O^+, $(HO_2)_{grain}$
309	5	$h\nu$, O(^{1}D), N(^{2}D), O^+, grain
310	5	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2^+
311	5	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2H^+
312	5	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, $(HO_2)_{grain}$
313	5	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, grain
314	5	$h\nu$, O(^{1}D), N(^{2}D), CO_2^+, CO_2H^+
315	5	$h\nu$, O(^{1}D), N(^{2}D), CO_2^+, $(HO_2)_{grain}$
316	5	$h\nu$, O(^{1}D), N(^{2}D), CO_2^+, grain

ID	# Species	Species
317	5	$h\nu$, O(^{1}D), N(^{2}D), CO_2H^+, $(HO_2)_{grain}$
318	5	$h\nu$, O(^{1}D), N(^{2}D), CO_2H^+, grain
319	5	$h\nu$, O(^{1}D), N(^{2}D), $(HO_2)_{grain}$, grain
320	5	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2^+
321	5	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2H^+
322	5	$h\nu$, O(^{1}D), O^+, O_2^+, $(HO_2)_{grain}$
323	5	$h\nu$, O(^{1}D), O^+, O_2^+, grain
324	5	$h\nu$, O(^{1}D), O^+, CO_2^+, CO_2H^+
325	5	$h\nu$, O(^{1}D), O^+, CO_2^+, $(HO_2)_{grain}$
326	5	$h\nu$, O(^{1}D), O^+, CO_2^+, grain
327	5	$h\nu$, O(^{1}D), O^+, CO_2H^+, $(HO_2)_{grain}$
328	5	$h\nu$, O(^{1}D), O^+, CO_2H^+, grain
329	5	$h\nu$, O(^{1}D), O^+, $(HO_2)_{grain}$, grain
330	5	$h\nu$, O(^{1}D), O_2^+, CO_2^+, CO_2H^+
331	5	$h\nu$, O(^{1}D), O_2^+, CO_2^+, $(HO_2)_{grain}$
332	5	$h\nu$, O(^{1}D), O_2^+, CO_2^+, grain
333	5	$h\nu$, O(^{1}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$
334	5	$h\nu$, O(^{1}D), O_2^+, CO_2H^+, grain
335	5	$h\nu$, O(^{1}D), O_2^+, $(HO_2)_{grain}$, grain
336	5	$h\nu$, O(^{1}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$
337	5	$h\nu$, O(^{1}D), CO_2^+, CO_2H^+, grain
338	5	$h\nu$, O(^{1}D), CO_2^+, $(HO_2)_{grain}$, grain
339	5	$h\nu$, O(^{1}D), CO_2H^+, $(HO_2)_{grain}$, grain
340	5	$h\nu$, H_2, H, CO, N
341	5	$h\nu$, H_2, H, CO, N(^{2}D)
342	5	$h\nu$, H_2, H, CO, O^+
343	5	$h\nu$, H_2, H, CO, O_2^+
344	5	$h\nu$, H_2, H, CO, CO_2H^+
345	5	$h\nu$, H_2, H, CO, $(HO_2)_{grain}$
346	5	$h\nu$, H_2, H, CO, grain
347	5	$h\nu$, H_2, H, N, N(^{2}D)
348	5	$h\nu$, H_2, H, N, O^+
349	5	$h\nu$, H_2, H, N, O_2^+
350	5	$h\nu$, H_2, H, N, CO_2H^+
351	5	$h\nu$, H_2, H, N, $(HO_2)_{grain}$
352	5	$h\nu$, H_2, H, N, grain
353	5	$h\nu$, H_2, H, N(^{2}D), O^+
354	5	$h\nu$, H_2, H, N(^{2}D), O_2^+
355	5	$h\nu$, H_2, H, N(^{2}D), CO_2H^+
356	5	$h\nu$, H_2, H, N(^{2}D), $(HO_2)_{grain}$
357	5	$h\nu$, H_2, H, N(^{2}D), grain
358	5	$h\nu$, H_2, H, O^+, O_2^+
359	5	$h\nu$, H_2, H, O^+, CO_2H^+
360	5	$h\nu$, H_2, H, O^+, $(HO_2)_{grain}$

ID	# Species	Species
361	5	$h\nu$, H_2, H, O^+, grain
362	5	$h\nu$, H_2, H, O_2^+, CO_2H^+
363	5	$h\nu$, H_2, H, O_2^+, $(HO_2)_{grain}$
364	5	$h\nu$, H_2, H, O_2^+, grain
365	5	$h\nu$, H_2, H, CO_2H^+, $(HO_2)_{grain}$
366	5	$h\nu$, H_2, H, CO_2H^+, grain
367	5	$h\nu$, H_2, H, $(HO_2)_{grain}$, grain
424	5	$h\nu$, N, $N(^2D)$, O^+, O_2^+
425	5	$h\nu$, N, $N(^2D)$, O^+, CO_2^+
426	5	$h\nu$, N, $N(^2D)$, O^+, CO_2H^+
427	5	$h\nu$, N, $N(^2D)$, O^+, $(HO_2)_{grain}$
428	5	$h\nu$, N, $N(^2D)$, O^+, grain
429	5	$h\nu$, N, $N(^2D)$, O_2^+, CO_2^+
430	5	$h\nu$, N, $N(^2D)$, O_2^+, CO_2H^+
431	5	$h\nu$, N, $N(^2D)$, O_2^+, $(HO_2)_{grain}$
432	5	$h\nu$, N, $N(^2D)$, O_2^+, grain
433	5	$h\nu$, N, $N(^2D)$, CO_2^+, CO_2H^+
434	5	$h\nu$, N, $N(^2D)$, CO_2^+, $(HO_2)_{grain}$
435	5	$h\nu$, N, $N(^2D)$, CO_2^+, grain
436	5	$h\nu$, N, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$
437	5	$h\nu$, N, $N(^2D)$, CO_2H^+, grain
438	5	$h\nu$, N, $N(^2D)$, $(HO_2)_{grain}$, grain
439	5	$h\nu$, N, O^+, O_2^+, CO_2^+
440	5	$h\nu$, N, O^+, O_2^+, CO_2H^+
441	5	$h\nu$, N, O^+, O_2^+, $(HO_2)_{grain}$
442	5	$h\nu$, N, O^+, O_2^+, grain
443	5	$h\nu$, N, O^+, CO_2^+, CO_2H^+
444	5	$h\nu$, N, O^+, CO_2^+, $(HO_2)_{grain}$
445	5	$h\nu$, N, O^+, CO_2^+, grain
446	5	$h\nu$, N, O^+, CO_2H^+, $(HO_2)_{grain}$
447	5	$h\nu$, N, O^+, CO_2H^+, grain
448	5	$h\nu$, N, O^+, $(HO_2)_{grain}$, grain
449	5	$h\nu$, N, O_2^+, CO_2^+, CO_2H^+
450	5	$h\nu$, N, O_2^+, CO_2^+, $(HO_2)_{grain}$
451	5	$h\nu$, N, O_2^+, CO_2^+, grain
452	5	$h\nu$, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$
453	5	$h\nu$, N, O_2^+, CO_2H^+, grain
454	5	$h\nu$, N, O_2^+, $(HO_2)_{grain}$, grain
455	5	$h\nu$, N, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
456	5	$h\nu$, N, CO_2^+, CO_2H^+, grain
457	5	$h\nu$, N, CO_2^+, $(HO_2)_{grain}$, grain
458	5	$h\nu$, N, CO_2H^+, $(HO_2)_{grain}$, grain
459	5	$h\nu$, $N(^2D)$, O^+, O_2^+, CO_2^+
460	5	$h\nu$, $N(^2D)$, O^+, O_2^+, CO_2H^+

ID	# Species	Species
461	5	$h\nu$, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$
462	5	$h\nu$, N(^{2}D), O^+, O_2^+, grain
463	5	$h\nu$, N(^{2}D), O^+, CO_2^+, CO_2H^+
464	5	$h\nu$, N(^{2}D), O^+, CO_2^+, $(HO_2)_{grain}$
465	5	$h\nu$, N(^{2}D), O^+, CO_2^+, grain
466	5	$h\nu$, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$
467	5	$h\nu$, N(^{2}D), O^+, CO_2H^+, grain
468	5	$h\nu$, N(^{2}D), O^+, $(HO_2)_{grain}$, grain
469	5	$h\nu$, N(^{2}D), O_2^+, CO_2^+, CO_2H^+
470	5	$h\nu$, N(^{2}D), O_2^+, CO_2^+, $(HO_2)_{grain}$
471	5	$h\nu$, N(^{2}D), O_2^+, CO_2^+, grain
472	5	$h\nu$, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$
473	5	$h\nu$, N(^{2}D), O_2^+, CO_2H^+, grain
474	5	$h\nu$, N(^{2}D), O_2^+, $(HO_2)_{grain}$, grain
475	5	$h\nu$, N(^{2}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$
476	5	$h\nu$, N(^{2}D), CO_2^+, CO_2H^+, grain
477	5	$h\nu$, N(^{2}D), CO_2^+, $(HO_2)_{grain}$, grain
478	5	$h\nu$, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$, grain
479	5	$h\nu$, O^+, O_2^+, CO_2^+, CO_2H^+
480	5	$h\nu$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
481	5	$h\nu$, O^+, O_2^+, CO_2^+, grain
482	5	$h\nu$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
483	5	$h\nu$, O^+, O_2^+, CO_2H^+, grain
484	5	$h\nu$, O^+, O_2^+, $(HO_2)_{grain}$, grain
485	5	$h\nu$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
486	5	$h\nu$, O^+, CO_2^+, CO_2H^+, grain
487	5	$h\nu$, O^+, CO_2^+, $(HO_2)_{grain}$, grain
488	5	$h\nu$, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
489	5	$h\nu$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
490	5	$h\nu$, O_2^+, CO_2^+, CO_2H^+, grain
491	5	$h\nu$, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
492	5	$h\nu$, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
493	5	$h\nu$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
501	6	$h\nu$, e, O(^{1}D), N, N(^{2}D), O^+
502	6	$h\nu$, e, O(^{1}D), N, N(^{2}D), $(HO_2)_{grain}$
503	6	$h\nu$, e, O(^{1}D), N, N(^{2}D), grain
504	6	$h\nu$, e, O(^{1}D), N, O^+, $(HO_2)_{grain}$
505	6	$h\nu$, e, O(^{1}D), N, O^+, grain
506	6	$h\nu$, e, O(^{1}D), N, $(HO_2)_{grain}$, grain
507	6	$h\nu$, e, O(^{1}D), N(^{2}D), O^+, $(HO_2)_{grain}$
508	6	$h\nu$, e, O(^{1}D), N(^{2}D), O^+, grain
509	6	$h\nu$, e, O(^{1}D), N(^{2}D), $(HO_2)_{grain}$, grain
510	6	$h\nu$, e, O(^{1}D), O^+, $(HO_2)_{grain}$, grain
511	6	$h\nu$, e, H_2, H, CO, N

ID	# Species	Species
512	6	$h\nu$, e, H_2, H, CO, $N(^2D)$
516	6	$h\nu$, e, H_2, H, CO, O^+
517	6	$h\nu$, e, H_2, H, CO, $(HO_2)_{grain}$
518	6	$h\nu$, e, H_2, H, CO, grain
519	6	$h\nu$, e, H_2, H, N, $N(^2D)$
520	6	$h\nu$, e, H_2, H, N, O^+
521	6	$h\nu$, e, H_2, H, N, $(HO_2)_{grain}$
522	6	$h\nu$, e, H_2, H, N, grain
523	6	$h\nu$, e, H_2, H, $N(^2D)$, O^+
524	6	$h\nu$, e, H_2, H, $N(^2D)$, $(HO_2)_{grain}$
525	6	$h\nu$, e, H_2, H, $N(^2D)$, grain
526	6	$h\nu$, e, H_2, H, O^+, $(HO_2)_{grain}$
527	6	$h\nu$, e, H_2, H, O^+, grain
528	6	$h\nu$, e, H_2, H, $(HO_2)_{grain}$, grain
539	6	$h\nu$, e, N, $N(^2D)$, O^+, $(HO_2)_{grain}$
540	6	$h\nu$, e, N, $N(^2D)$, O^+, grain
541	6	$h\nu$, e, N, $N(^2D)$, $(HO_2)_{grain}$, grain
542	6	$h\nu$, e, N, O^+, $(HO_2)_{grain}$, grain
543	6	$h\nu$, e, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
600	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+
601	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2^+
602	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2H^+
603	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, $(HO_2)_{grain}$
604	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, grain
605	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2^+
606	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2H^+
607	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, $(HO_2)_{grain}$
608	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, grain
609	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2^+, CO_2H^+
610	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2^+, $(HO_2)_{grain}$
611	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2^+, grain
612	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$
613	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2H^+, grain
614	6	$h\nu$, $O(^1D)$, N, $N(^2D)$, $(HO_2)_{grain}$, grain
615	6	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2^+
616	6	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2H^+
617	6	$h\nu$, $O(^1D)$, N, O^+, O_2^+, $(HO_2)_{grain}$
618	6	$h\nu$, $O(^1D)$, N, O^+, O_2^+, grain
619	6	$h\nu$, $O(^1D)$, N, O^+, CO_2^+, CO_2H^+
620	6	$h\nu$, $O(^1D)$, N, O^+, CO_2^+, $(HO_2)_{grain}$
621	6	$h\nu$, $O(^1D)$, N, O^+, CO_2^+, grain
622	6	$h\nu$, $O(^1D)$, N, O^+, CO_2H^+, $(HO_2)_{grain}$
623	6	$h\nu$, $O(^1D)$, N, O^+, CO_2H^+, grain
624	6	$h\nu$, $O(^1D)$, N, O^+, $(HO_2)_{grain}$, grain

ID	# Species	Species
625	6	$h\nu$, O(^{1}D), N, O_2^+, CO_2^+, CO_2H^+
626	6	$h\nu$, O(^{1}D), N, O_2^+, CO_2^+, $(HO_2)_{grain}$
627	6	$h\nu$, O(^{1}D), N, O_2^+, CO_2^+, grain
628	6	$h\nu$, O(^{1}D), N, O_2^+, CO_2H^+, $(HO_2)_{grain}$
629	6	$h\nu$, O(^{1}D), N, O_2^+, CO_2H^+, grain
630	6	$h\nu$, O(^{1}D), N, O_2^+, $(HO_2)_{grain}$, grain
631	6	$h\nu$, O(^{1}D), N, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
632	6	$h\nu$, O(^{1}D), N, CO_2^+, CO_2H^+, grain
633	6	$h\nu$, O(^{1}D), N, CO_2^+, $(HO_2)_{grain}$, grain
634	6	$h\nu$, O(^{1}D), N, CO_2H^+, $(HO_2)_{grain}$, grain
635	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+, CO_2^+
636	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+, CO_2H^+
637	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$
638	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+, grain
639	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2^+, CO_2H^+
640	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2^+, $(HO_2)_{grain}$
641	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2^+, grain
642	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$
643	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2H^+, grain
644	6	$h\nu$, O(^{1}D), N(^{2}D), O^+, $(HO_2)_{grain}$, grain
645	6	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2^+, CO_2H^+
646	6	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2^+, $(HO_2)_{grain}$
647	6	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2^+, grain
648	6	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$
649	6	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2H^+, grain
650	6	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, $(HO_2)_{grain}$, grain
651	6	$h\nu$, O(^{1}D), N(^{2}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$
652	6	$h\nu$, O(^{1}D), N(^{2}D), CO_2^+, CO_2H^+, grain
653	6	$h\nu$, O(^{1}D), N(^{2}D), CO_2^+, $(HO_2)_{grain}$, grain
654	6	$h\nu$, O(^{1}D), N(^{2}D), CO_2H^+, $(HO_2)_{grain}$, grain
655	6	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2^+, CO_2H^+
656	6	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
657	6	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2^+, grain
658	6	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
659	6	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2H^+, grain
660	6	$h\nu$, O(^{1}D), O^+, O_2^+, $(HO_2)_{grain}$, grain
661	6	$h\nu$, O(^{1}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
662	6	$h\nu$, O(^{1}D), O^+, CO_2^+, CO_2H^+, grain
663	6	$h\nu$, O(^{1}D), O^+, CO_2^+, $(HO_2)_{grain}$, grain
664	6	$h\nu$, O(^{1}D), O^+, CO_2H^+, $(HO_2)_{grain}$, grain
665	6	$h\nu$, O(^{1}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
666	6	$h\nu$, O(^{1}D), O_2^+, CO_2^+, CO_2H^+, grain
667	6	$h\nu$, O(^{1}D), O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
668	6	$h\nu$, O(^{1}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain

ID	# Species	Species
669	6	$h\nu$, O(^{1}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
670	6	$h\nu$, H_2, H, CO, N, N(^{2}D)
671	6	$h\nu$, H_2, H, CO, N, O^+
672	6	$h\nu$, H_2, H, CO, N, O_2^+
673	6	$h\nu$, H_2, H, CO, N, CO_2H^+
674	6	$h\nu$, H_2, H, CO, N, $(HO_2)_{grain}$
675	6	$h\nu$, H_2, H, CO, N, grain
676	6	$h\nu$, H_2, H, CO, N(^{2}D), O^+
677	6	$h\nu$, H_2, H, CO, N(^{2}D), O_2^+
678	6	$h\nu$, H_2, H, CO, N(^{2}D), CO_2H^+
679	6	$h\nu$, H_2, H, CO, N(^{2}D), $(HO_2)_{grain}$
680	6	$h\nu$, H_2, H, CO, N(^{2}D), grain
681	6	$h\nu$, H_2, H, CO, O^+, O_2^+
682	6	$h\nu$, H_2, H, CO, O^+, CO_2H^+
683	6	$h\nu$, H_2, H, CO, O^+, $(HO_2)_{grain}$
684	6	$h\nu$, H_2, H, CO, O^+, grain
685	6	$h\nu$, H_2, H, CO, O_2^+, CO_2H^+
686	6	$h\nu$, H_2, H, CO, O_2^+, $(HO_2)_{grain}$
687	6	$h\nu$, H_2, H, CO, O_2^+, grain
688	6	$h\nu$, H_2, H, CO, CO_2H^+, $(HO_2)_{grain}$
689	6	$h\nu$, H_2, H, CO, CO_2H^+, grain
690	6	$h\nu$, H_2, H, CO, $(HO_2)_{grain}$, grain
691	6	$h\nu$, H_2, H, N, N(^{2}D), O^+
692	6	$h\nu$, H_2, H, N, N(^{2}D), O_2^+
693	6	$h\nu$, H_2, H, N, N(^{2}D), CO_2H^+
694	6	$h\nu$, H_2, H, N, N(^{2}D), $(HO_2)_{grain}$
695	6	$h\nu$, H_2, H, N, N(^{2}D), grain
696	6	$h\nu$, H_2, H, N, O^+, O_2^+
697	6	$h\nu$, H_2, H, N, O^+, CO_2H^+
698	6	$h\nu$, H_2, H, N, O^+, $(HO_2)_{grain}$
699	6	$h\nu$, H_2, H, N, O^+, grain
700	6	$h\nu$, H_2, H, N, O_2^+, CO_2H^+
701	6	$h\nu$, H_2, H, N, O_2^+, $(HO_2)_{grain}$
702	6	$h\nu$, H_2, H, N, O_2^+, grain
703	6	$h\nu$, H_2, H, N, CO_2H^+, $(HO_2)_{grain}$
704	6	$h\nu$, H_2, H, N, CO_2H^+, grain
705	6	$h\nu$, H_2, H, N, $(HO_2)_{grain}$, grain
706	6	$h\nu$, H_2, H, N(^{2}D), O^+, O_2^+
707	6	$h\nu$, H_2, H, N(^{2}D), O^+, CO_2H^+
708	6	$h\nu$, H_2, H, N(^{2}D), O^+, $(HO_2)_{grain}$
709	6	$h\nu$, H_2, H, N(^{2}D), O^+, grain
710	6	$h\nu$, H_2, H, N(^{2}D), O_2^+, CO_2H^+
711	6	$h\nu$, H_2, H, N(^{2}D), O_2^+, $(HO_2)_{grain}$
712	6	$h\nu$, H_2, H, N(^{2}D), O_2^+, grain

ID	# Species	Species
713	6	$h\nu$, H_2, H, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$
714	6	$h\nu$, H_2, H, $N(^2D)$, CO_2H^+, grain
715	6	$h\nu$, H_2, H, $N(^2D)$, $(HO_2)_{grain}$, grain
716	6	$h\nu$, H_2, H, O^+, O_2^+, CO_2H^+
717	6	$h\nu$, H_2, H, O^+, O_2^+, $(HO_2)_{grain}$
718	6	$h\nu$, H_2, H, O^+, O_2^+, grain
719	6	$h\nu$, H_2, H, O^+, CO_2H^+, $(HO_2)_{grain}$
720	6	$h\nu$, H_2, H, O^+, CO_2H^+, grain
721	6	$h\nu$, H_2, H, O^+, $(HO_2)_{grain}$, grain
722	6	$h\nu$, H_2, H, O_2^+, CO_2H^+, $(HO_2)_{grain}$
723	6	$h\nu$, H_2, H, O_2^+, CO_2H^+, grain
724	6	$h\nu$, H_2, H, O_2^+, $(HO_2)_{grain}$, grain
725	6	$h\nu$, H_2, H, CO_2H^+, $(HO_2)_{grain}$, grain
796	6	$h\nu$, N, $N(^2D)$, O^+, O_2^+, CO_2^+
797	6	$h\nu$, N, $N(^2D)$, O^+, O_2^+, CO_2H^+
798	6	$h\nu$, N, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$
799	6	$h\nu$, N, $N(^2D)$, O^+, O_2^+, grain
800	6	$h\nu$, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+
801	6	$h\nu$, N, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$
802	6	$h\nu$, N, $N(^2D)$, O^+, CO_2^+, grain
803	6	$h\nu$, N, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$
804	6	$h\nu$, N, $N(^2D)$, O^+, CO_2H^+, grain
805	6	$h\nu$, N, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
806	6	$h\nu$, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+
807	6	$h\nu$, N, $N(^2D)$, O_2^+, CO_2^+, $(HO_2)_{grain}$
808	6	$h\nu$, N, $N(^2D)$, O_2^+, CO_2^+, grain
809	6	$h\nu$, N, $N(^2D)$, O_2^+, CO_2H^+, $(HO_2)_{grain}$
810	6	$h\nu$, N, $N(^2D)$, O_2^+, CO_2H^+, grain
811	6	$h\nu$, N, $N(^2D)$, O_2^+, $(HO_2)_{grain}$, grain
812	6	$h\nu$, N, $N(^2D)$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
813	6	$h\nu$, N, $N(^2D)$, CO_2^+, CO_2H^+, grain
814	6	$h\nu$, N, $N(^2D)$, CO_2^+, $(HO_2)_{grain}$, grain
815	6	$h\nu$, N, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$, grain
816	6	$h\nu$, N, O^+, O_2^+, CO_2^+, CO_2H^+
817	6	$h\nu$, N, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
818	6	$h\nu$, N, O^+, O_2^+, CO_2^+, grain
819	6	$h\nu$, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
820	6	$h\nu$, N, O^+, O_2^+, CO_2H^+, grain
821	6	$h\nu$, N, O^+, O_2^+, $(HO_2)_{grain}$, grain
822	6	$h\nu$, N, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
823	6	$h\nu$, N, O^+, CO_2^+, CO_2H^+, grain
824	6	$h\nu$, N, O^+, CO_2^+, $(HO_2)_{grain}$, grain
825	6	$h\nu$, N, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
826	6	$h\nu$, N, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$

ID	# Species	Species
827	6	$h\nu$, N, O_2^+, CO_2^+, CO_2H^+, grain
828	6	$h\nu$, N, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
829	6	$h\nu$, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
830	6	$h\nu$, N, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
831	6	$h\nu$, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+
832	6	$h\nu$, N(^{2}D), O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
833	6	$h\nu$, N(^{2}D), O^+, O_2^+, CO_2^+, grain
834	6	$h\nu$, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
835	6	$h\nu$, N(^{2}D), O^+, O_2^+, CO_2H^+, grain
836	6	$h\nu$, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$, grain
837	6	$h\nu$, N(^{2}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
838	6	$h\nu$, N(^{2}D), O^+, CO_2^+, CO_2H^+, grain
839	6	$h\nu$, N(^{2}D), O^+, CO_2^+, $(HO_2)_{grain}$, grain
840	6	$h\nu$, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$, grain
841	6	$h\nu$, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
842	6	$h\nu$, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, grain
843	6	$h\nu$, N(^{2}D), O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
844	6	$h\nu$, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
845	6	$h\nu$, N(^{2}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
846	6	$h\nu$, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
847	6	$h\nu$, O^+, O_2^+, CO_2^+, CO_2H^+, grain
848	6	$h\nu$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
849	6	$h\nu$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
850	6	$h\nu$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
851	6	$h\nu$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
856	7	$h\nu$, e, O(^{1}D), N, N(^{2}D), O^+, $(HO_2)_{grain}$
857	7	$h\nu$, e, O(^{1}D), N, N(^{2}D), O^+, grain
858	7	$h\nu$, e, O(^{1}D), N, N(^{2}D), $(HO_2)_{grain}$, grain
859	7	$h\nu$, e, O(^{1}D), N, O^+, $(HO_2)_{grain}$, grain
860	7	$h\nu$, e, O(^{1}D), N(^{2}D), O^+, $(HO_2)_{grain}$, grain
861	7	$h\nu$, e, H_2, H, CO, N, N(^{2}D)
862	7	$h\nu$, e, H_2, H, CO, N, O^+
863	7	$h\nu$, e, H_2, H, CO, N, $(HO_2)_{grain}$
864	7	$h\nu$, e, H_2, H, CO, N, grain
865	7	$h\nu$, e, H_2, H, CO, N(^{2}D), O^+
866	7	$h\nu$, e, H_2, H, CO, N(^{2}D), $(HO_2)_{grain}$
873	7	$h\nu$, e, H_2, H, CO, N(^{2}D), grain
874	7	$h\nu$, e, H_2, H, CO, O^+, $(HO_2)_{grain}$
875	7	$h\nu$, e, H_2, H, CO, O^+, grain
876	7	$h\nu$, e, H_2, H, CO, $(HO_2)_{grain}$, grain
877	7	$h\nu$, e, H_2, H, N, N(^{2}D), O^+
878	7	$h\nu$, e, H_2, H, N, N(^{2}D), $(HO_2)_{grain}$
879	7	$h\nu$, e, H_2, H, N, N(^{2}D), grain
880	7	$h\nu$, e, H_2, H, N, O^+, $(HO_2)_{grain}$

ID	# Species	Species
881	7	$h\nu$, e, H_2, H, N, O^+, grain
882	7	$h\nu$, e, H_2, H, N, $(HO_2)_{grain}$, grain
883	7	$h\nu$, e, H_2, H, $N(^2D)$, O^+, $(HO_2)_{grain}$
884	7	$h\nu$, e, H_2, H, $N(^2D)$, O^+, grain
885	7	$h\nu$, e, H_2, H, $N(^2D)$, $(HO_2)_{grain}$, grain
886	7	$h\nu$, e, H_2, H, O^+, $(HO_2)_{grain}$, grain
892	7	$h\nu$, e, N, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
963	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, CO_2^+
964	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, CO_2H^+
965	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$
966	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, grain
967	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+
968	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$
969	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2^+, grain
970	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$
971	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2H^+, grain
972	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
973	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+
974	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2^+, $(HO_2)_{grain}$
975	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2^+, grain
976	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2H^+, $(HO_2)_{grain}$
977	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2H^+, grain
978	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, $(HO_2)_{grain}$, grain
979	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
980	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2^+, CO_2H^+, grain
981	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2^+, $(HO_2)_{grain}$, grain
982	7	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$, grain
983	7	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2^+, CO_2H^+
984	7	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
985	7	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2^+, grain
986	7	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
987	7	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2H^+, grain
988	7	$h\nu$, $O(^1D)$, N, O^+, O_2^+, $(HO_2)_{grain}$, grain
989	7	$h\nu$, $O(^1D)$, N, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
990	7	$h\nu$, $O(^1D)$, N, O^+, CO_2^+, CO_2H^+, grain
991	7	$h\nu$, $O(^1D)$, N, O^+, CO_2^+, $(HO_2)_{grain}$, grain
992	7	$h\nu$, $O(^1D)$, N, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
993	7	$h\nu$, $O(^1D)$, N, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
994	7	$h\nu$, $O(^1D)$, N, O_2^+, CO_2^+, CO_2H^+, grain
995	7	$h\nu$, $O(^1D)$, N, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
996	7	$h\nu$, $O(^1D)$, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
997	7	$h\nu$, $O(^1D)$, N, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
998	7	$h\nu$, $O(^1D)$, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+
999	7	$h\nu$, $O(^1D)$, $N(^2D)$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$

ID	# Species	Species
1000	7	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+, CO_2^+, grain
1001	7	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1002	7	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+, CO_2H^+, grain
1003	7	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$, grain
1004	7	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1005	7	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2^+, CO_2H^+, grain
1006	7	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2^+, $(HO_2)_{grain}$, grain
1007	7	$h\nu$, O(^{1}D), N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1008	7	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1009	7	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2^+, CO_2H^+, grain
1010	7	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1011	7	$h\nu$, O(^{1}D), N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1012	7	$h\nu$, O(^{1}D), N(^{2}D), CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1013	7	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1014	7	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2^+, CO_2H^+, grain
1015	7	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1016	7	$h\nu$, O(^{1}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1017	7	$h\nu$, O(^{1}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1018	7	$h\nu$, O(^{1}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1019	7	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+
1020	7	$h\nu$, H_2, H, CO, N, N(^{2}D), O_2^+
1021	7	$h\nu$, H_2, H, CO, N, N(^{2}D), CO_2H^+
1022	7	$h\nu$, H_2, H, CO, N, N(^{2}D), $(HO_2)_{grain}$
1023	7	$h\nu$, H_2, H, CO, N, N(^{2}D), grain
1024	7	$h\nu$, H_2, H, CO, N, O^+, O_2^+
1025	7	$h\nu$, H_2, H, CO, N, O^+, CO_2H^+
1026	7	$h\nu$, H_2, H, CO, N, O^+, $(HO_2)_{grain}$
1027	7	$h\nu$, H_2, H, CO, N, O^+, grain
1028	7	$h\nu$, H_2, H, CO, N, O_2^+, CO_2H^+
1029	7	$h\nu$, H_2, H, CO, N, O_2^+, $(HO_2)_{grain}$
1030	7	$h\nu$, H_2, H, CO, N, O_2^+, grain
1031	7	$h\nu$, H_2, H, CO, N, CO_2H^+, $(HO_2)_{grain}$
1032	7	$h\nu$, H_2, H, CO, N, CO_2H^+, grain
1033	7	$h\nu$, H_2, H, CO, N, $(HO_2)_{grain}$, grain
1034	7	$h\nu$, H_2, H, CO, N(^{2}D), O^+, O_2^+
1035	7	$h\nu$, H_2, H, CO, N(^{2}D), O^+, CO_2H^+
1036	7	$h\nu$, H_2, H, CO, N(^{2}D), O^+, $(HO_2)_{grain}$
1037	7	$h\nu$, H_2, H, CO, N(^{2}D), O^+, grain
1038	7	$h\nu$, H_2, H, CO, N(^{2}D), O_2^+, CO_2H^+
1039	7	$h\nu$, H_2, H, CO, N(^{2}D), O_2^+, $(HO_2)_{grain}$
1040	7	$h\nu$, H_2, H, CO, N(^{2}D), O_2^+, grain
1041	7	$h\nu$, H_2, H, CO, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$
1042	7	$h\nu$, H_2, H, CO, N(^{2}D), CO_2H^+, grain
1043	7	$h\nu$, H_2, H, CO, N(^{2}D), $(HO_2)_{grain}$, grain

ID	# Species	Species
1044	7	$h\nu$, H_2, H, CO, O^+, O_2^+, CO_2H^+
1045	7	$h\nu$, H_2, H, CO, O^+, O_2^+, $(HO_2)_{grain}$
1046	7	$h\nu$, H_2, H, CO, O^+, O_2^+, grain
1047	7	$h\nu$, H_2, H, CO, O^+, CO_2H^+, $(HO_2)_{grain}$
1048	7	$h\nu$, H_2, H, CO, O^+, CO_2H^+, grain
1049	7	$h\nu$, H_2, H, CO, O^+, $(HO_2)_{grain}$, grain
1050	7	$h\nu$, H_2, H, CO, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1051	7	$h\nu$, H_2, H, CO, O_2^+, CO_2H^+, grain
1052	7	$h\nu$, H_2, H, CO, O_2^+, $(HO_2)_{grain}$, grain
1053	7	$h\nu$, H_2, H, CO, CO_2H^+, $(HO_2)_{grain}$, grain
1054	7	$h\nu$, H_2, H, N, $N(^2D)$, O^+, O_2^+
1055	7	$h\nu$, H_2, H, N, $N(^2D)$, O^+, CO_2H^+
1056	7	$h\nu$, H_2, H, N, $N(^2D)$, O^+, $(HO_2)_{grain}$
1057	7	$h\nu$, H_2, H, N, $N(^2D)$, O^+, grain
1058	7	$h\nu$, H_2, H, N, $N(^2D)$, O_2^+, CO_2H^+
1059	7	$h\nu$, H_2, H, N, $N(^2D)$, O_2^+, $(HO_2)_{grain}$
1060	7	$h\nu$, H_2, H, N, $N(^2D)$, O_2^+, grain
1061	7	$h\nu$, H_2, H, N, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$
1062	7	$h\nu$, H_2, H, N, $N(^2D)$, CO_2H^+, grain
1063	7	$h\nu$, H_2, H, N, $N(^2D)$, $(HO_2)_{grain}$, grain
1064	7	$h\nu$, H_2, H, N, O^+, O_2^+, CO_2H^+
1065	7	$h\nu$, H_2, H, N, O^+, O_2^+, $(HO_2)_{grain}$
1066	7	$h\nu$, H_2, H, N, O^+, O_2^+, grain
1067	7	$h\nu$, H_2, H, N, O^+, CO_2H^+, $(HO_2)_{grain}$
1068	7	$h\nu$, H_2, H, N, O^+, CO_2H^+, grain
1069	7	$h\nu$, H_2, H, N, O^+, $(HO_2)_{grain}$, grain
1070	7	$h\nu$, H_2, H, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1071	7	$h\nu$, H_2, H, N, O_2^+, CO_2H^+, grain
1072	7	$h\nu$, H_2, H, N, O_2^+, $(HO_2)_{grain}$, grain
1073	7	$h\nu$, H_2, H, N, CO_2H^+, $(HO_2)_{grain}$, grain
1074	7	$h\nu$, H_2, H, $N(^2D)$, O^+, O_2^+, CO_2H^+
1075	7	$h\nu$, H_2, H, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$
1076	7	$h\nu$, H_2, H, $N(^2D)$, O^+, O_2^+, grain
1077	7	$h\nu$, H_2, H, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$
1078	7	$h\nu$, H_2, H, $N(^2D)$, O^+, CO_2H^+, grain
1079	7	$h\nu$, H_2, H, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
1080	7	$h\nu$, H_2, H, $N(^2D)$, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1081	7	$h\nu$, H_2, H, $N(^2D)$, O_2^+, CO_2H^+, grain
1082	7	$h\nu$, H_2, H, $N(^2D)$, O_2^+, $(HO_2)_{grain}$, grain
1083	7	$h\nu$, H_2, H, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$, grain
1084	7	$h\nu$, H_2, H, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1085	7	$h\nu$, H_2, H, O^+, O_2^+, CO_2H^+, grain
1086	7	$h\nu$, H_2, H, O^+, O_2^+, $(HO_2)_{grain}$, grain
1087	7	$h\nu$, H_2, H, O^+, CO_2H^+, $(HO_2)_{grain}$, grain

ID	# Species	Species
1088	7	$h\nu$, H_2, H, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1145	7	$h\nu$, N, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+
1146	7	$h\nu$, N, $N(^2D)$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1147	7	$h\nu$, N, $N(^2D)$, O^+, O_2^+, CO_2^+, grain
1148	7	$h\nu$, N, $N(^2D)$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1149	7	$h\nu$, N, $N(^2D)$, O^+, O_2^+, CO_2H^+, grain
1150	7	$h\nu$, N, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$, grain
1151	7	$h\nu$, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1152	7	$h\nu$, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+, grain
1153	7	$h\nu$, N, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$, grain
1154	7	$h\nu$, N, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1155	7	$h\nu$, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1156	7	$h\nu$, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, grain
1157	7	$h\nu$, N, $N(^2D)$, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1158	7	$h\nu$, N, $N(^2D)$, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1159	7	$h\nu$, N, $N(^2D)$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1160	7	$h\nu$, N, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1161	7	$h\nu$, N, O^+, O_2^+, CO_2^+, CO_2H^+, grain
1162	7	$h\nu$, N, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1163	7	$h\nu$, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1164	7	$h\nu$, N, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1165	7	$h\nu$, N, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1166	7	$h\nu$, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1167	7	$h\nu$, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+, grain
1168	7	$h\nu$, $N(^2D)$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1169	7	$h\nu$, $N(^2D)$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1170	7	$h\nu$, $N(^2D)$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1171	7	$h\nu$, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1172	7	$h\nu$, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1174	8	$h\nu$, e, $O(^1D)$, N, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
1175	8	$h\nu$, e, H_2, H, CO, N, $N(^2D)$, O^+
1176	8	$h\nu$, e, H_2, H, CO, N, $N(^2D)$, $(HO_2)_{grain}$
1177	8	$h\nu$, e, H_2, H, CO, N, $N(^2D)$, grain
1178	8	$h\nu$, e, H_2, H, CO, N, O^+, $(HO_2)_{grain}$
1179	8	$h\nu$, e, H_2, H, CO, N, O^+, grain
1180	8	$h\nu$, e, H_2, H, CO, N, $(HO_2)_{grain}$, grain
1181	8	$h\nu$, e, H_2, H, CO, $N(^2D)$, O^+, $(HO_2)_{grain}$
1182	8	$h\nu$, e, H_2, H, CO, $N(^2D)$, O^+, grain
1187	8	$h\nu$, e, H_2, H, CO, $N(^2D)$, $(HO_2)_{grain}$, grain
1188	8	$h\nu$, e, H_2, H, CO, O^+, $(HO_2)_{grain}$, grain
1189	8	$h\nu$, e, H_2, H, N, $N(^2D)$, O^+, $(HO_2)_{grain}$
1190	8	$h\nu$, e, H_2, H, N, $N(^2D)$, O^+, grain
1191	8	$h\nu$, e, H_2, H, N, $N(^2D)$, $(HO_2)_{grain}$, grain
1192	8	$h\nu$, e, H_2, H, N, O^+, $(HO_2)_{grain}$, grain

ID	# Species	Species
1193	8	$h\nu$, e, H_2, H, $N(^2D)$, O^+, $(HO_2)_{grain}$, grain
1251	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+
1252	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$
1253	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, CO_2^+, grain
1254	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1255	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, CO_2H^+, grain
1256	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, O_2^+, $(HO_2)_{grain}$, grain
1257	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1258	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2^+, CO_2H^+, grain
1259	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2^+, $(HO_2)_{grain}$, grain
1260	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1261	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1262	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, grain
1263	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1264	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1265	8	$h\nu$, $O(^1D)$, N, $N(^2D)$, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1266	8	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1267	8	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2^+, CO_2H^+, grain
1268	8	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1269	8	$h\nu$, $O(^1D)$, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1270	8	$h\nu$, $O(^1D)$, N, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1271	8	$h\nu$, $O(^1D)$, N, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1272	8	$h\nu$, $O(^1D)$, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1273	8	$h\nu$, $O(^1D)$, $N(^2D)$, O^+, O_2^+, CO_2^+, CO_2H^+, grain
1274	8	$h\nu$, $O(^1D)$, $N(^2D)$, O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1275	8	$h\nu$, $O(^1D)$, $N(^2D)$, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1276	8	$h\nu$, $O(^1D)$, $N(^2D)$, O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1277	8	$h\nu$, $O(^1D)$, $N(^2D)$, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1278	8	$h\nu$, $O(^1D)$, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1279	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, O^+, O_2^+
1280	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, O^+, CO_2H^+
1281	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, O^+, $(HO_2)_{grain}$
1282	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, O^+, grain
1283	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, O_2^+, CO_2H^+
1284	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, O_2^+, $(HO_2)_{grain}$
1285	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, O_2^+, grain
1286	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, CO_2H^+, $(HO_2)_{grain}$
1287	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, CO_2H^+, grain
1288	8	$h\nu$, H_2, H, CO, N, $N(^2D)$, $(HO_2)_{grain}$, grain
1289	8	$h\nu$, H_2, H, CO, N, O^+, O_2^+, CO_2H^+
1290	8	$h\nu$, H_2, H, CO, N, O^+, O_2^+, $(HO_2)_{grain}$
1291	8	$h\nu$, H_2, H, CO, N, O^+, O_2^+, grain
1292	8	$h\nu$, H_2, H, CO, N, O^+, CO_2H^+, $(HO_2)_{grain}$
1293	8	$h\nu$, H_2, H, CO, N, O^+, CO_2H^+, grain

ID	# Species	Species
1294	8	$h\nu$, H_2, H, CO, N, O^+, $(HO_2)_{grain}$, grain
1295	8	$h\nu$, H_2, H, CO, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1296	8	$h\nu$, H_2, H, CO, N, O_2^+, CO_2H^+, grain
1297	8	$h\nu$, H_2, H, CO, N, O_2^+, $(HO_2)_{grain}$, grain
1298	8	$h\nu$, H_2, H, CO, N, CO_2H^+, $(HO_2)_{grain}$, grain
1299	8	$h\nu$, H_2, H, CO, N(^{2}D), O^+, O_2^+, CO_2H^+
1300	8	$h\nu$, H_2, H, CO, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$
1301	8	$h\nu$, H_2, H, CO, N(^{2}D), O^+, O_2^+, grain
1302	8	$h\nu$, H_2, H, CO, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$
1303	8	$h\nu$, H_2, H, CO, N(^{2}D), O^+, CO_2H^+, grain
1304	8	$h\nu$, H_2, H, CO, N(^{2}D), O^+, $(HO_2)_{grain}$, grain
1305	8	$h\nu$, H_2, H, CO, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$
1306	8	$h\nu$, H_2, H, CO, N(^{2}D), O_2^+, CO_2H^+, grain
1307	8	$h\nu$, H_2, H, CO, N(^{2}D), O_2^+, $(HO_2)_{grain}$, grain
1308	8	$h\nu$, H_2, H, CO, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$, grain
1309	8	$h\nu$, H_2, H, CO, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1310	8	$h\nu$, H_2, H, CO, O^+, O_2^+, CO_2H^+, grain
1311	8	$h\nu$, H_2, H, CO, O^+, O_2^+, $(HO_2)_{grain}$, grain
1312	8	$h\nu$, H_2, H, CO, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1313	8	$h\nu$, H_2, H, CO, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1314	8	$h\nu$, H_2, H, N, N(^{2}D), O^+, O_2^+, CO_2H^+
1315	8	$h\nu$, H_2, H, N, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$
1316	8	$h\nu$, H_2, H, N, N(^{2}D), O^+, O_2^+, grain
1317	8	$h\nu$, H_2, H, N, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$
1318	8	$h\nu$, H_2, H, N, N(^{2}D), O^+, CO_2H^+, grain
1319	8	$h\nu$, H_2, H, N, N(^{2}D), O^+, $(HO_2)_{grain}$, grain
1320	8	$h\nu$, H_2, H, N, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$
1321	8	$h\nu$, H_2, H, N, N(^{2}D), O_2^+, CO_2H^+, grain
1322	8	$h\nu$, H_2, H, N, N(^{2}D), O_2^+, $(HO_2)_{grain}$, grain
1323	8	$h\nu$, H_2, H, N, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$, grain
1324	8	$h\nu$, H_2, H, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1325	8	$h\nu$, H_2, H, N, O^+, O_2^+, CO_2H^+, grain
1326	8	$h\nu$, H_2, H, N, O^+, O_2^+, $(HO_2)_{grain}$, grain
1327	8	$h\nu$, H_2, H, N, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1328	8	$h\nu$, H_2, H, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1329	8	$h\nu$, H_2, H, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1330	8	$h\nu$, H_2, H, N(^{2}D), O^+, O_2^+, CO_2H^+, grain
1331	8	$h\nu$, H_2, H, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$, grain
1332	8	$h\nu$, H_2, H, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1333	8	$h\nu$, H_2, H, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1334	8	$h\nu$, H_2, H, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1363	8	$h\nu$, N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1364	8	$h\nu$, N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, grain
1365	8	$h\nu$, N, N(^{2}D), O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain

ID	# Species	Species
1366	8	$h\nu$, N, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1367	8	$h\nu$, N, N(^{2}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1368	8	$h\nu$, N, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1369	8	$h\nu$, N, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1370	8	$h\nu$, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1371	9	$h\nu$, e, H_2, H, CO, N, N(^{2}D), O^+, $(HO_2)_{grain}$
1372	9	$h\nu$, e, H_2, H, CO, N, N(^{2}D), O^+, grain
1373	9	$h\nu$, e, H_2, H, CO, N, N(^{2}D), $(HO_2)_{grain}$, grain
1374	9	$h\nu$, e, H_2, H, CO, N, O^+, $(HO_2)_{grain}$, grain
1375	9	$h\nu$, e, H_2, H, CO, N(^{2}D), O^+, $(HO_2)_{grain}$, grain
1377	9	$h\nu$, e, H_2, H, N, N(^{2}D), O^+, $(HO_2)_{grain}$, grain
1406	9	$h\nu$, O(^{1}D), N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$
1407	9	$h\nu$, O(^{1}D), N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, grain
1408	9	$h\nu$, O(^{1}D), N, N(^{2}D), O^+, O_2^+, CO_2^+, $(HO_2)_{grain}$, grain
1409	9	$h\nu$, O(^{1}D), N, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1410	9	$h\nu$, O(^{1}D), N, N(^{2}D), O^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1411	9	$h\nu$, O(^{1}D), N, N(^{2}D), O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1412	9	$h\nu$, O(^{1}D), N, O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1413	9	$h\nu$, O(^{1}D), N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1414	9	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, O_2^+, CO_2H^+
1415	9	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$
1416	9	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, O_2^+, grain
1417	9	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$
1418	9	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, CO_2H^+, grain
1419	9	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, $(HO_2)_{grain}$, grain
1420	9	$h\nu$, H_2, H, CO, N, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$
1421	9	$h\nu$, H_2, H, CO, N, N(^{2}D), O_2^+, CO_2H^+, grain
1422	9	$h\nu$, H_2, H, CO, N, N(^{2}D), O_2^+, $(HO_2)_{grain}$, grain
1423	9	$h\nu$, H_2, H, CO, N, N(^{2}D), CO_2H^+, $(HO_2)_{grain}$, grain
1424	9	$h\nu$, H_2, H, CO, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1425	9	$h\nu$, H_2, H, CO, N, O^+, O_2^+, CO_2H^+, grain
1426	9	$h\nu$, H_2, H, CO, N, O^+, O_2^+, $(HO_2)_{grain}$, grain
1427	9	$h\nu$, H_2, H, CO, N, O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1428	9	$h\nu$, H_2, H, CO, N, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1429	9	$h\nu$, H_2, H, CO, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1430	9	$h\nu$, H_2, H, CO, N(^{2}D), O^+, O_2^+, CO_2H^+, grain
1431	9	$h\nu$, H_2, H, CO, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$, grain
1432	9	$h\nu$, H_2, H, CO, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1433	9	$h\nu$, H_2, H, CO, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1434	9	$h\nu$, H_2, H, CO, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1435	9	$h\nu$, H_2, H, N, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1436	9	$h\nu$, H_2, H, N, N(^{2}D), O^+, O_2^+, CO_2H^+, grain
1437	9	$h\nu$, H_2, H, N, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$, grain
1438	9	$h\nu$, H_2, H, N, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$, grain

ID	# Species	Species
1439	9	$h\nu$, H_2, H, N, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1440	9	$h\nu$, H_2, H, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1441	9	$h\nu$, H_2, H, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1450	9	$h\nu$, N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1451	10	$h\nu$, e, H_2, H, CO, N, N(^{2}D), O^+, $(HO_2)_{grain}$, grain
1460	10	$h\nu$, O(^{1}D), N, N(^{2}D), O^+, O_2^+, CO_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1461	10	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$
1462	10	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, O_2^+, CO_2H^+, grain
1463	10	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, O_2^+, $(HO_2)_{grain}$, grain
1464	10	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, CO_2H^+, $(HO_2)_{grain}$, grain
1465	10	$h\nu$, H_2, H, CO, N, N(^{2}D), O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1466	10	$h\nu$, H_2, H, CO, N, O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1467	10	$h\nu$, H_2, H, CO, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1468	10	$h\nu$, H_2, H, N, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain
1472	11	$h\nu$, H_2, H, CO, N, N(^{2}D), O^+, O_2^+, CO_2H^+, $(HO_2)_{grain}$, grain

Group of organizations having constant intensity value zero

ID	# Species	Species
0	1	$h\nu$
10	2	$h\nu$, $(HO_2)_{grain}$
11	2	$h\nu$, grain

A. PHOTOCHEMICAL MODEL OF THE MARS ATMOSPHERE

Appendix B

Reaction Network Model of Bacteriophage Lambda

1. Input reactions:

$$\begin{array}{rcl} \emptyset & \rightarrow & \text{PR'} \\ \text{PR} & \rightarrow & \text{2 PR} \\ \text{PL} & \rightarrow & \text{2 PL} \end{array}$$

2. Model reactions:

$$\begin{array}{rcl} \text{PRM} & \rightarrow & \text{cI\_mRNA} \\ \text{P\_m17} & \rightarrow & \text{P\_m12} \\ \text{P\_m21} & \rightarrow & \text{P\_m31} \\ \text{P\_m39} & \rightarrow & \text{cI\_mRNA} \\ \text{P\_m45} & \rightarrow & \text{P\_m26} \\ \text{P\_m26} & \rightarrow & \text{head\_tail} \\ \text{P\_m53} & \rightarrow & \text{P\_m19} \\ \text{P\_m13} & \rightarrow & \text{int\_mRNA} \\ \text{P\_m18} & \rightarrow & \text{P\_m13} \\ \text{P\_m24} & \rightarrow & \text{Q\_mRNA} \\ \text{PInt} & \rightarrow & \text{int\_mRNA\_m41} \\ \text{Panti-Q} & \rightarrow & \text{anti-Q} \\ \text{P\_m20} & \rightarrow & \text{P\_m46} \\ \text{cI\_mRNA} & \rightarrow & \text{cI\_mRNA + CI} \\ \text{PL} & \rightarrow & \text{N\_mRNA + P\_m16} \\ \text{cIII\_mRNA} & \rightarrow & \text{cIII\_mRNA + CIII} \\ \text{P\_m12} & \rightarrow & \text{xis\_mRNA + P\_m18} \\ \text{N + P\_m19} & \rightarrow & \text{N + P\_m24} \end{array}$$

$$
\begin{aligned}
\text{O\_mRNA} &\rightarrow \text{O\_mRNA} + \text{O} \\
\text{P\_m37} &\rightarrow \text{P\_m21} + \text{cII\_mRNA} \\
\text{PR} &\rightarrow \text{cro\_mRNA} + \text{P\_m28} \\
\text{N} + \text{P\_m28} &\rightarrow \text{N} + \text{P\_m37} \\
\text{CII} &\rightarrow \text{CII} + \text{PRE} \\
\text{Cro} &\rightarrow \text{Cro} + \text{OR2\_Cro} + \text{OR1\_Cro} \\
\text{Cro} &\rightarrow \text{Cro} + \text{OR3\_Cro} \\
\text{PR'} + \text{Q} &\rightarrow \text{P\_m45} + \text{Q} \\
\text{P\_mRNA} &\rightarrow \text{P\_mRNA} + \text{Prot\_P} \\
\text{OR1\_CI} + \text{OR2\_CI} &\rightarrow \text{OR1\_CI} + \text{OR2\_CI} + \text{PRM} \\
\text{int\_mRNA\_m41} &\rightarrow \text{int\_mRNA\_m41} + \text{Int} \\
\text{N\_mRNA} &\rightarrow \text{N\_mRNA} + \text{N} \\
\text{P\_m7} &\rightarrow \text{cIII\_mRNA} + \text{P\_m17} \\
\text{xis\_mRNA} &\rightarrow \text{xis\_mRNA} + \text{Xis} \\
\text{N} + \text{P\_m16} &\rightarrow \text{N} + \text{P\_m7} \\
\text{Q\_mRNA} &\rightarrow \text{Q\_mRNA} + \text{Q} \\
\text{P\_m31} &\rightarrow \text{O\_mRNA} + \text{P\_m20} \\
\text{cII\_mRNA} &\rightarrow \text{cII\_mRNA} + \text{CII} \\
\text{cro\_mRNA} &\rightarrow \text{cro\_mRNA} + \text{Cro} \\
\text{PRE} &\rightarrow \text{P\_m39} + \text{anti-Cro} \\
\text{CII} &\rightarrow \text{CII} + \text{PInt} + \text{Panti-Q} \\
\text{CI} &\rightarrow \text{CI} + \text{OR1\_CI} + \text{OR2\_CI} \\
\text{CI} &\rightarrow \text{CI} + \text{OR3\_CI} \\
\text{P\_m46} &\rightarrow \text{P\_mRNA} + \text{P\_m53}
\end{aligned}
$$

3. Antisense reactions:

$$
\begin{aligned}
\text{cro\_mRNA} + \text{anti-Cro} &\rightarrow \emptyset \\
\text{Q\_mRNA} + \text{anti-Q} &\rightarrow \emptyset
\end{aligned}
$$

4. Decay reactions for species cI_mRNA, N, CIII, int_mRNA, Q, cII_mRNA, P_m28, int_mRNA_m41, OR2_Cro, OR3_Cro, anti-Q, CI, P_m16, xis_mRNA, P_m19, O_mRNA, CII, Cro, Int, OR2_CI, OR3_CI, P_mRNA, N_mRNA, cIII_mRNA, Xis, Q_mRNA, O, cro_mRNA, anti-Cro, OR1_Cro, OR1_CI, PR', Prot_P, and head_tail have the form:

$$\text{Species} \rightarrow \emptyset$$

Appendix C

Model of the Central Sugar Metabolism of E. coli

C.1 List of Species

Species Names	Substances
ATP, ADP, AMP, cAMP	ATP, ADP, AMP, and cyclic AMP
RNAP, Tscription	RNA polymerase and RNAP bound to DNA
Crp, PromCrp, CrpmRNA	catabolite repressor protein, gene, and mRNA
Cya, PromCya, CyamRNA	adenylate cyclase, gene, and mRNA
EIIA, PromEIIA, EIIAmRNA	PTS system enzyme IIA^{Glc}, gene, and mRNA
EIIAP	phosphorylated PTS system enzyme IIA^{Glc}
EIIBC, PromEIIBC, EIIBCmRNA	PTS system enzyme $IIBC^{Glc}$, gene, and mRNA
EI, PromEI, EImRNA	PTS system enzyme I, gene, and mRNA
Fbp, PromFbp, FbpmRNA	fructose bis–phosphatase, gene, and mRNA
Fda, PromFda, FdamRNA	fructose bisphosphate aldolase, gene, and mRNA
Gap, PromGap, GapmRNA	glyceraldehyde–3–phosphate dehydrogenase, gene, and mRNA
GlcT, PromGlcT, GlcTmRNA	glucose transporter, gene, and mRNA
Glk, PromGlk, GlkmRNA	glucokinase, gene, and mRNA
GlpD, PromGlpD, GlpDmRNA	glycerol–3–phosphate dehydrogenase, gene, and

Species Names	Substances
	mRNA
GlpFKmRNA, GlpFKmRNA1	glpFK operon mRNA
GlpR, PromGlpR, GlpRmRNA	glp regulon repressor, gene, and mRNA
Gpm, PromGpm, GpmmRNA	phosphoglycerate mutase, gene, and mRNA
HPr, PromHPr, HPrmRNA	PTS system HPr protein, gene, and mRNA
HPrP	phosphorylated PTS system HPr protein
LacI, PromLacI, LacImRNA	lac operon repressor, gene, and mRNA
LacZYmRNA, LacZYmRNA1	lac operon mRNA
Pfk, PromPfk, PfkmRNA	phosphofructokinase, gene, and mRNA
Pgi, PromPgi, PgimRNA	phosphoglucose isomerase, gene, and mRNA
Pyk, PromPyk, PykmRNA	pyruvate kinase, gene, and mRNA
Tpi, PromTpi, TpimRNA	triose phosphate isomerase, gene, and mRNA
PromGlpFK, GlpF, GlpK	glpFK operon, glycerol faciliator and kinase
PromLacZY, LacZ, LacY	lac operon, β–galactosidase, and lactose permease
Glcex, Glyex, Lacex	extracellular glucose, glycerol and lactose
Glc, Gly, Lac	intracellular glucose, glycerol and lactose
Allo	Allolactose
Glc6P	glucose–6–phoshpate
G3P	glycerol–3–phosphate
Fru6P	fructose–6–phosphate
FBP	fructose–1,6–bisphosphate
DHAP	dihydroxy–acetone–phosphate
T3P	glyceraldehyde–3–phosphate
3PG	3–phospho–glycerate
PEP	phosphoenolopyruvate
Pyr	pyruvate
Metabolism	further metabolic processes

C.2 Reaction Network

1. Synthesis and decay is identical for species Crp, Cya, EIIA, EIIBC, EI, Fbp, Fda, Gap, GlcT, Glk, GlpR, Gpm, HPr, LacI, Pfk, Pgi, Pyk, and Tpi:

$$\begin{aligned} \text{RNAP} + \text{PromSpecies} &\rightarrow \text{Tscription} + \text{PromSpecies} + \text{SpeciesmRNA} \\ \text{SpeciesmRNA} &\rightarrow \text{SpeciesmRNA} + \text{Species} \\ \text{SpeciesmRNA} &\rightarrow \emptyset \\ \text{Species} &\rightarrow \emptyset \end{aligned}$$

2. Synthesis and decay of inducible species LacZY, GlpFK, and GlpD:

$$
\begin{aligned}
\text{RNAP} + \text{PromLacZY} + \\
\text{Allo} + \text{Crp} + \text{cAMP} &\rightarrow \text{Tscription} + \text{PromLacZY} + \\
&\quad \text{LacZYmRNA} + \text{Allo} + \text{Crp} + \text{cAMP} \\
\text{LacZYmRNA} &\rightarrow \text{LacZYmRNA1} + \text{LacZ} \\
\text{LacZYmRNA1} &\rightarrow \text{LacZYmRNA} + \text{LacY} \\
\text{LacZYmRNA} &\rightarrow \emptyset \\
\text{LacZYmRNA1} &\rightarrow \emptyset \\
\text{LacZ} &\rightarrow \emptyset \\
\text{LacY} &\rightarrow \emptyset \\
\text{RNAP} + \text{PromGlpFK} + \\
\text{G3P} + \text{Crp} + \text{cAMP} &\rightarrow \text{Tscription} + \text{PromGlpFK} + \\
&\quad \text{GlpFKmRNA} + \text{G3P} + \text{Crp} + \text{cAMP} \\
\text{GlpFKmRNA} &\rightarrow \text{GlpFKmRNA1} + \text{GlpF} \\
\text{GlpFKmRNA1} &\rightarrow \text{GlpFKmRNA} + \text{GlpK} \\
\text{GlpFKmRNA} &\rightarrow \emptyset \\
\text{GlpFKmRNA1} &\rightarrow \emptyset \\
\text{GlpF} &\rightarrow \emptyset \\
\text{GlpK} &\rightarrow \emptyset \\
\text{RNAP} + \text{PromGlpD} + \\
\text{G3P} + \text{Crp} + \text{cAMP} &\rightarrow \text{Tscription} + \text{PromGlpD} + \\
&\quad \text{GlpDmRNA} + \text{G3P} + \text{Crp} + \text{cAMP} \\
\text{GlpDmRNA} &\rightarrow \text{GlpDmRNA} + \text{GlpD} \\
\text{GlpDmRNA} &\rightarrow \emptyset \\
\text{GlpD} &\rightarrow \emptyset
\end{aligned}
$$

3. Unbinding of RNAP:

$$\text{Tscription} \rightarrow \text{RNAP}$$

4. Signal transduction, transport and metabolic reactions:

$$
\begin{aligned}
\text{ATP} + \text{Cya} &\rightarrow \text{cAMP} + \text{Cya} \\
\text{PEP} + \text{EI} + \text{HPr} &\rightarrow \text{Pyr} + \text{EI} + \text{HPrP} \\
\text{Pyr} + \text{EI} + \text{HPrP} &\rightarrow \text{PEP} + \text{EI} + \text{HPr} \\
\text{EIIA} + \text{HPrP} &\rightarrow \text{EIIAP} + \text{HPr} \\
\text{EIIAP} + \text{HPr} &\rightarrow \text{EIIA} + \text{HPrP} \\
\text{Glcex} + \text{EIIAP} + \text{EIIBC} &\rightarrow \text{Glc6P} + \text{EIIA} + \text{EIIBC} \\
\text{Glc} + \text{EIIAP} + \text{EIIBC} &\rightarrow \text{Glc6P} + \text{EIIA} + \text{EIIBC} \\
\text{Glcex} + \text{GlcT} &\rightarrow \text{Glc} + \text{GlcT} \\
\text{Lacex} + \text{LacY} &\rightarrow \text{Lac} + \text{LacY} \\
\text{Lac} + \text{LacZ} &\rightarrow \text{Allo} + \text{LacZ} \\
\text{Lac} + \text{LacZ} &\rightarrow \text{Glc} + \text{Glc6P} + \text{LacZ} \\
\text{Allo} + \text{LacZ} &\rightarrow \text{Glc} + \text{Glc6P} + \text{LacZ} \\
\text{Glc} + \text{Glk} &\rightarrow \text{Glc6P} + \text{Glk} \\
\text{Glc6P} + \text{Pgi} &\rightarrow \text{Fru6P} + \text{Pgi} \\
\text{Fru6P} + \text{Pgi} &\rightarrow \text{Glc6P} + \text{Pgi} \\
\text{Fru6P} + \text{Fbp} &\rightarrow \text{FBP} + \text{Fbp} \\
\text{FBP} + \text{Fbp} &\rightarrow \text{Fru6P} + \text{Fbp} \\
\text{Fru6P} + \text{Pfk} &\rightarrow \text{FBP} + \text{Pfk} \\
\text{FBP} + \text{Fda} &\rightarrow \text{T3P} + \text{DHAP} + \text{Fda} \\
\text{T3P} + \text{DHAP} + \text{Fda} &\rightarrow \text{FBP} + \text{Fda} \\
\text{Glyex} + \text{GlpF} &\rightarrow \text{Gly} + \text{GlpF} \\
\text{Gly} + \text{GlpF} &\rightarrow \text{Glyex} + \text{GlpF} \\
\text{Gly} + \text{GlpK} &\rightarrow \text{G3P} + \text{GlpK} \\
\text{G3P} + \text{GlpD} &\rightarrow \text{DHAP} + \text{GlpD} \\
\text{DHAP} + \text{Tpi} &\rightarrow \text{T3P} + \text{Tpi} \\
\text{T3P} + \text{Tpi} &\rightarrow \text{DHAP} + \text{Tpi} \\
\text{T3P} + \text{Gap} &\rightarrow \text{3PG} + \text{Gap} \\
\text{3PG} + \text{Gap} &\rightarrow \text{T3P} + \text{Gap} \\
\text{3PG} + \text{Gpm} &\rightarrow \text{PEP} + \text{Gpm} \\
\text{PEP} + \text{Gpm} &\rightarrow \text{3PG} + \text{Gpm} \\
\text{PEP} + \text{FBP} + \text{Pyk} &\rightarrow \text{Pyr} + \text{FBP} + \text{Pyk} \\
\text{Pyr} &\rightarrow \text{Metabolism}
\end{aligned}
$$

5. Decay reactions for species ATP, ADP, AMP, cAMP, EIIAP, HPrP, Glc, Gly, Lac, Allo, Glc6P, G3P, Fru6P, FBP, DHAP, T3P, 3PG, PEP, Pyr, and Metabolism have the form:

$$\text{Species} \rightarrow \emptyset$$

6. Input reactions for ATP, ADP, AMP, RNAP, PromCrp, PromCya, PromEIIA, PromEIIBC, PromEI, PromFbp, PromFda, PromGap, PromGlcT, PromGlk,

PromGlpD, PromGlpR, PromGpm, PromHPr, PromLacI, PromPfk, PromPgi, PromPyk, PromTpi, PromGlpFK, and PromLacZY have the form:

$$\emptyset \quad \rightarrow \quad \text{Species}$$

C. MODEL OF THE CENTRAL SUGAR METABOLISM OF *E. COLI*

Appendix D

Reaction Network Model of the Regulated Central Metabolism of E. coli

Species and reactions marked with '*' make up the core network model.

D.1 Metabolites

Abbr.	Metabolite	Abbr.	Metabolite
$13PDG^*$	1,3-bis-Phosphoglycerate	$GLxt^*$	External glycerol
$2PG^*$	2-Phosphoglycerate	$HEXT^*$	External H+
$3PG^*$	3-Phosphoglycerate	$ICIT$	Isocitrate
AC	Acetate	LAC	D-Lactate
$ACCOA$	Acetyl-CoA	$LACxt$	External lactate
$ACTP$	Acetyl-phosphate	$LCTS^*$	Lactose
$ACxt$	External acetate	$LCTSxt^*$	External Lactose
ADP^*	Adenosine diphosphate	MAL	Malate
AKG	a-Ketoglutarate	NAD^*	Nicotinamide adenine dinucleotide
AMP^*	Adenosine monophosphate	$NADH^*$	Nicotinamide adenine dinucleotide red.
ATP^*	Adenosine triphosphate	$NADP^*$	Nicotinamide adenine dinucleotide phosphate
$bDGLAC^*$	b-D-Galactose	$NADPH^*$	Dihydronicotinamide adenine dinucleotide phosphate reduced
$bDGLC^*$	b-D-Glucose	$O2^*$	Oxygen
$Biomass$	Cell biomass	$O2xt^*$	External Oxygen
CIT	Citrate	OA	Oxaloacetate
$CO2$	Carbon dioxide	PEP^*	Phosphoenolpyruvate
$CO2xt$	External carbon dioxide		

Abbr.	Metabolite	Abbr.	Metabolite
COA	Coenzyme A	PI^*	Phosphate (inorganic)
$D6PGC$	D-6-Phosphate-gluconate	$PIxt$	External phosphate
$D6PGL$	D-6-Phosphate-glucono-delta-lactone	PPI^*	Pyrophosphate
$E4P$	Erythrose 4-phosphate	PYR^*	Pyruvate
ETH	Ethanol	$PYRxt$	External pyruvate
$ETHxt$	External ethanol	Q^*	Ubiquinone
$F6P^*$	Fructose 6-phosphate	$QH2^*$	Ubiquinol
FAD	Flavin adenine dinucleotide	$R5P$	Ribose 5-phosphate
$FADH$	Flavin adenine dinucleotide reduced	RIB	Ribose
FDP^*	Fructose 1,6-diphosphate	$RIBxt$	External ribose
FOR	Formate	$RL5P$	Ribulose 5-phosphate
$FORxt$	External Formate	$S7P$	sedo-Heptulose
FUM	Fumarate	$SUCC$	Succinate
$G1P^*$	Glucose 1-phosphate	$SUCCOA$	Succinate CoA
$G6P^*$	Glucose 6-phosphate	$SUCCxt$	External succinate
$GAL1P^*$	Galactose 1-Phosphate	$T3P1^*$	Glyceraldehyde 3-phosphate
GL^*	Glycerol	$T3P2^*$	Dihydroxyacetone phosphate
$GL3P^*$	Glycerol 3-phosphate	$UDPG^*$	UDP Glucose
$GLAC^*$	Galactose	$UDPGAL^*$	UDP Galactose
GLC^*	a-D-Glucose	UTP^*	Uridine triphosphate
$GLCxt^*$	External glucose	$X5P$	Xylulose-5-phosphate
GLX	Glyoxylate	$food^*$	carbon source present in medium
$Oxid-Radicals$	superoxid radicals	$LactateUP$	lactate uptake activated

D.2 Genes and Proteins

Gene	Protein	Gene	Protein
aceA	Isocitrate lyase	pgm^*	Phosphoglucomutase
aceB	Malate synthase A	*pntAB*	Pyridine nucleotide transhydrogenase
aceEF, *lpdA*	Pyruvate dehydrogenase	*ppa*	Inorganic pyrophosphatase
ackA	Acetate kinase A	*ppc*	Phosphoenolpyruvate carboxylase
acnA	Aconitase A	$ppsA^*$	Phosphoenolpyruvate synthase
acnB	Aconitase B	*pta*	Phosphotransacetylase
acs	Acetyl-CoA synthetase	$pykA^*$	Pyruvate Kinase II
adhE	Acetaldehyde dehydrogenase	$pykF^*$	Pyruvate Kinase I
adk	Adenylate kinase	*rbsK*	Ribokinase
atpABC-DEFGHI	F0F1-ATPase	*rpe*	Ribulose phosphate 3-epimerase
cydAB	Cytochrome oxidase bd	*rpiA*	Ribose-5-phosphate isomerase A

Gene	Protein	Gene	Protein
cyoABCD	Cytochrome oxidase bo3	*rpiB*	Ribose-5-phosphate isomerase B
dld	D-Lactate dehydrogenase 1	*sdhABCD*	Succinate dehydrogenase complex
*eno**	Enolase	*sfcA*	Malic enzyme (NAD)
*fba**	Fructose-1,6-bisphosphatate aldolase	*sucAB*, *lpdA*	2-Ketoglutarate dehyrogenase
*fbp**	Fructose-1,6-bisphosphatase	*sucCD*	Succinyl-CoA synthetase
fdnGHI	Formate dehydrogenase-N	*talA*	Transaldolase A
fdoIHG	Formate dehydrogenase-O	*talB*	Transaldolase B
frdABCD	Fumarate reductase	*tktA*	Transketolase I
fumA	Fumarase A		
fumB	Fumarase B	*tktB*	Transketolase II
fumC	Fumarase C		
*galE**	UDP-glucose 4-epimerase	*tpiA**	Triosphosphate Isomerase
*galK**	Galactokinase	*zwf*	Glucose 6-phosphate-1-dehydrogenase
*galM**	Aldose 1-epimerase (mutorotase)	*focA*	Formate transport
		*ptsGHI**, *crr**	Glucose transport
*galT**	Galactose-1-phosphate uridylyltransferase	*galP**	Glucose transport (low affinity)
*galU**	UDP-glucose-1-phosphate uridylyltransferase	*glpF**	Glycerol transporter
*gapA**	Glyceraldehyde-3-phosphate dehydrogenase-A complex	*lacY**	Lactose permease
*glk**	Glucokinase	*pitAB*	Phosphate transport
*glpABC**	Glycerol-3-phosphate dehydrogenase (anaerobic)	*rbsABCD*	Ribose transport
*glpD**	Glycerol-3-phosphate dehydrogenase (aerobic)	*dctA*	Succinate transport
*glpK**	Glycerol kinase	*dcuA*	Succinate transport
gltA	Citrate synthase	*dcuB*	Succinate transport
gnd	6-Phosphogluconate dehydrogenase (decarboxylating)	*dcuC*	Succinate efflux
*gpmA**	Phosphoglycerate mutase 1	*arcA**	Aerobic/Anaerobic response regulator
*gpmB**	Phosphoglycerate mutase 2	*cra**(fruR)	Catabolite activator protein
*gpsA**	Glycerol-3-phosphate-dehydrogenase-[NAD(P)+]		
icdA	Isocitrate dehydrogenase	*dcuR*	Dicarboxylate response regulator
*lacZ**	Beta-galactosidase (LACTase)	*dcuS*	Dicarboxylate response sensor
maeB	Malic enzyme (NADP)	*fadR**	Fatty acid/Acetate response regulator
mdh	Malate dehydrogenase	*fnr**	Aerobic/Anaerobic response regulator

Gene	Protein	Gene	Protein
ndh	NADH dehydrogenase II	$galR^*$	Galactose operon repressor
nuoABEF– GHI-JKLMN	NADH dehydrogenase I	$galS^*$	Galactose operon repressor
pckA	Phosphoenolpyruvate carboxykinase	$glpR^*$	Glycerol response regulator
$pfkA^*$	Phosphofructokinase	*iclR*	Fatty acid/Acetate response regulator
$pfkB^*$	Phosphofructokinase B	$lacI^*$	Lactose operon repressor
pflAB	Pyruvate formate lyase 1	mlc^*	Glucose response regulator
pflCD	Pyruvate formate lyase 2	*pdhR*	Pyruvate response regulator
pgi^*	Phosphoglucose isomerase	*rbsR*	Ribose response regulator
pgk^*	Phosphoglycerate kinase	*rpiR*	Ribose response regulator
pgl	6-Phosphogluconolactonase		

D.3 Pseudo Species

$\overline{G6P}^*$, $\overline{ACxt}$, $\overline{GLAC}^*$, $\overline{GLCxt}^*$, $\overline{GLxt}^*$, $\overline{GalR}^*$, $\overline{GalS}^*$, $\overline{GlpR}^*$, $\overline{IclR}$, $\overline{LACxt}$, $\overline{LCTSxt}^*$, $\overline{O2xt}^*$, $\overline{PYR}$, $\overline{PYRxt}$, $\overline{RIBxt}$, $\overline{SUCCxt}$, $\overline{ETHxt}$, $\overline{ArcA}^*$, $\overline{Cra}^*$, $\overline{Fnr}^*$, $\overline{Food}^*$, $\overline{LacI}^*$, $\overline{Mlc}^*$, $\overline{PdhR}$, $\overline{RbsR}$, $\overline{RpiR}$

D.4 Spontaneously Created Species

ADP, ATP, AckA, Adk, AtpABCDEFGHI, COA, DcuA, Dld, Eno*, FAD, FADH, Fba*, Fbp*, FdoIHG, GalU*, GapA*, Glk*, GltA, Gnd, GpmA*, GpmB*, GpsA*, HEXT*, IcdA, MaeB, NAD*, NADH, NADP*, NADPH, NuoABEFGHIJKLMN, PI, PckA, PfkA*, PfkB*, Pgi*, Pgl, Pgm*, PitAB, PntAB, Ppa, Ppc, Pta, PykA*, Q*, QH2, Rpe, RpiA, SfcA, SucCD, TalA, TalB, TktA, TktB, TpiA*, UTP*, Zwf*

Species used as input species (in the complete network), respectively self-replicators (in the core network): *GLCxt*, *LCTSxt*, *GLxt*, *O2xt*

D.5 Spontaneously Decaying Species

13PDG, 2PG, 3PG, AC, ACCOA, ACTP, ACxt, ADP*, AKG, AMP*, ATP*, AceA, AceB, AceEF, AckA, AcnA, AcnB, Acs, AdhE, Adk, ArcA*, AtpABCDEFGHI, BDGLAC, BDGLC, Biomass, CIT, CO2, CO2xt, COA, Cra*, Crr*, CydAB, CyoABCD, D6PGC, D6PGL, DctA, DcuA, DcuB, DcuC, DcuR, DcuS, Dld, E4P, ETH, ETHxt, Eno*, F6P, FAD, FADH, FDP, FOR, FORxt, FUM, FadR*, Fba*, Fbp*, FdnGHI, FdoIHG, Fnr*, FocA, Food*, FrdABCD, FumA, FumB, FumC, G1P, G6P*, GAL1P*, GL*, GL3P, GLAC, GLC, GLCxt*, GLX, GLxt, GalE*, GalK*, GalM*, GalP*, GalR*, GalS*, GalT*, GalU*, GapA*, Glk*, GlpABC*, GlpD*, GlpF*, GlpK*, GlpR*, GltA, Gnd, GpmA*, GpmB*, GpsA*, HEXT*, ICIT, IcdA, IclR, LAC, LACxt, LCTS, LCTSxt*, LacI*, LacY*, LacZ*, LactateUP, LpdA, MAL, MaeB, Mdh, Mlc*, NAD*, NADH*, NADP*, NADPH*, Ndh, NuoABEFGHIJKLMN, O2*, O2xt*, OA, PEP*,*

PI, PIxt, PPI*, PYR*, PYRxt, PckA, PdhR, PfkA*, PfkB*, PflAB, PflCD, Pgi, Pgk*, Pgl, Pgm*, PitAB, PntAB, Ppa, Ppc, PpsA*, Pta, PtsGHI*, PykA*, PykF*, Q*, QH2*, R5P, RIB, RIBxt, RL5P, RbsABCD, RbsK, RbsR, Rpe, RpiA, RpiB, RpiR, S7P, SOxidRadicals, SUCC, SUCCOA, SUCCxt, SdhABCD, SfcA, SucAB, SucCD, T3P1*, T3P2, TalA, TalB, TktA, TktB, TpiA*, UDPG, UDPGAL, UTP*, X5P, Zwf*

D.6 Transport Reactions

$$
\begin{aligned}
HEXT{+}ACxt &\rightarrow AC \\
AC &\rightarrow HEXT{+}ACxt \\
CO2xt &\rightarrow CO2 \\
CO2 &\rightarrow CO2xt \\
HEXT{+}ETHxt &\rightarrow ETH \\
ETH &\rightarrow HEXT{+}ETHxt \\
FocA{+}FORxt &\rightarrow FOR{+}FocA \\
FOR{+}FocA &\rightarrow FocA{+}FORxt \\
GLCxt{+}PEP{+}Crr{+}PtsGHI &\rightarrow PYR{+}G6P{+}Crr{+}PtsGHI^{*} \\
GLCxt{+}HEXT{+}GalP &\rightarrow GLC{+}GalP^{*} \\
GLxt{+}GlpF &\rightarrow GL{+}GlpF^{*} \\
GL{+}GlpF &\rightarrow GLxt{+}GlpF^{*} \\
HEXT{+}LACxt{+}LactateUP &\rightarrow LAC{+}LactateUP \\
LAC &\rightarrow HEXT{+}LACxt \\
HEXT{+}LCTSxt{+}LacY &\rightarrow LCTS{+}LacY^{*} \\
LCTS{+}LacY &\rightarrow HEXT{+}LCTSxt{+}LacY^{*} \\
O2xt &\rightarrow O2^{*} \\
O2 &\rightarrow O2xt^{*} \\
HEXT{+}PIxt{+}PitAB &\rightarrow PI{+}PitAB \\
PI{+}PitAB &\rightarrow HEXT{+}PIxt{+}PitAB \\
HEXT{+}PYRxt &\rightarrow PYR \\
PYR &\rightarrow HEXT{+}PYRxt \\
ATP{+}RbsABCD{+}RIBxt &\rightarrow ADP{+}PI{+}RIB{+}RbsABCD \\
SUCCxt{+}HEXT{+}DctA &\rightarrow SUCC{+}DctA \\
SUCC{+}DctA &\rightarrow SUCCxt{+}HEXT{+}DctA \\
SUCCxt{+}HEXT{+}DcuA &\rightarrow SUCC{+}DcuA \\
SUCC{+}DcuA &\rightarrow SUCCxt{+}HEXT{+}DcuA \\
SUCCxt{+}HEXT{+}DcuB &\rightarrow SUCC{+}DcuB \\
SUCC{+}DcuB &\rightarrow SUCCxt{+}HEXT{+}DcuB \\
SUCC{+}DcuC &\rightarrow SUCCxt{+}HEXT{+}DcuC
\end{aligned}
$$

D.7 Metabolic reactions

$$
\begin{aligned}
&ACCOA{+}PYR{+}PEP{+}T3P1{+}F6P{+}G6P \\
&{+}OA{+}3PG{+}AKG{+}R5P{+}SUCCOA{+}E4P \rightarrow Biomass
\end{aligned}
$$

D. REACTION NETWORK MODEL OF THE REGULATED CENTRAL METABOLISM OF *E. COLI*

$$
\begin{aligned}
AceA+ICIT &\rightarrow AceA+SUCC+GLX\\
GLX+ACCOA+AceB &\rightarrow COA+AceB+MAL\\
NAD+COA+AceEF+PYR+LpdA &\rightarrow NADH+ACCOA+AceEF+LpdA+CO2\\
ADP+AckA+ACTP &\rightarrow ATP+AckA+AC\\
ATP+AckA+AC &\rightarrow ADP+AckA+ACTP\\
CIT+AcnA &\rightarrow ICIT+AcnA\\
ICIT+AcnA &\rightarrow CIT+AcnA\\
CIT+AcnB &\rightarrow ICIT+AcnB\\
ICIT+AcnB &\rightarrow CIT+AcnB\\
ATP+COA+AC+Acs &\rightarrow ACCOA+Acs+PPI+AMP\\
ACCOA+AdhE+2.0\ NADH &\rightarrow COA+AdhE+ETH+2.0\ NAD\\
COA+AdhE+ETH+2.0\ NAD &\rightarrow ACCOA+AdhE+2.0\ NADH\\
ATP+AMP+Adk &\rightarrow Adk+2.0\ ADP\\
Adk+2.0\ ADP &\rightarrow ATP+AMP+Adk\\
ATP+AtpABCDEFGHI &\rightarrow ADP+AtpABCDEFGHI+PI+4.0\ HEXT\\
ADP+AtpABCDEFGHI+PI+4.0\ HEXT &\rightarrow ATP+AtpABCDEFGHI\\
CydAB+O2+2.0\ QH2 &\rightarrow CydAB+2.0\ Q+4.0\ HEXT\\
O2+CyoABCD+2.0\ QH2 &\rightarrow CyoABCD+2.0\ Q+5.0\ HEXT\\
NADH+PYR+Dld &\rightarrow NAD+Dld+LAC\\
NAD+Dld+LAC &\rightarrow NADH+PYR+Dld\\
Q+Dld+LAC &\rightarrow QH2+PYR+Dld\\
2PG+Eno &\rightarrow Eno+PEP^{*}\\
Eno+PEP &\rightarrow 2PG+Eno^{*}\\
Fba+FDP &\rightarrow Fba+T3P2+T3P1^{*}\\
Fba+T3P2+T3P1 &\rightarrow Fba+FDP^{*}\\
FDP+Fbp &\rightarrow PI+Fbp+F6P^{*}\\
Q+FdnGHI+FOR &\rightarrow QH2+CO2+FdnGHI+2.0\ HEXT\\
Q+FOR+FdoIHG &\rightarrow QH2+CO2+FdoIHG+2.0\ HEXT\\
FADH+FrdABCD+FUM &\rightarrow FAD+SUCC+FrdABCD\\
FUM+FumA &\rightarrow MAL+FumA\\
MAL+FumA &\rightarrow FUM+FumA\\
FUM+FumB &\rightarrow MAL+FumB\\
MAL+FumB &\rightarrow FUM+FumB\\
FUM+FumC &\rightarrow MAL+FumC\\
MAL+FumC &\rightarrow FUM+FumC\\
GalE+UDPGAL &\rightarrow GalE+UDPG^{*}\\
GalE+UDPG &\rightarrow GalE+UDPGAL^{*}\\
ATP+GLAC+GalK &\rightarrow ADP+GalK+GAL1P^{*}\\
ADP+GalK+GAL1P &\rightarrow ATP+GLAC+GalK^{*}\\
BDGLAC+GalM &\rightarrow GLAC+GalM^{*}\\
GLAC+GalM &\rightarrow BDGLAC+GalM^{*}\\
GalM+BDGLC &\rightarrow GalM+GLC^{*}\\
GalM+GLC &\rightarrow GalM+BDGLC^{*}\\
UTP+GAL1P+GalT &\rightarrow PPI+UDPGAL+GalT^{*}\\
PPI+UDPGAL+GalT &\rightarrow UTP+GAL1P+GalT^{*}\\
UTP+GalU+G1P &\rightarrow PPI+UDPG+GalU^{*}\\
PPI+UDPG+GalU &\rightarrow UTP+GalU+G1P^{*}\\
NAD+PI+T3P1+GapA &\rightarrow NADH+GapA+13PDG^{*}\\
NADH+GapA+13PDG &\rightarrow NAD+PI+T3P1+GapA^{*}
\end{aligned}
$$

$$
\begin{aligned}
ATP+GLC+Glk &\rightarrow ADP+Glk+G6P^{*}\\
Q+GlpABC+GL3P &\rightarrow QH2+T3P2+GlpABC^{*}\\
Q+GL3P+GlpD &\rightarrow QH2+T3P2+GlpD^{*}\\
ATP+GL+GlpK &\rightarrow ADP+GL3P+GlpK^{*}\\
ACCOA+OA+GltA &\rightarrow COA+CIT+GltA\\
NADP+Gnd+D6PGC &\rightarrow NADPH+CO2+Gnd+RL5P\\
3PG+GpmA &\rightarrow 2PG+GpmA^{*}\\
2PG+GpmA &\rightarrow 3PG+GpmA^{*}\\
3PG+GpmB &\rightarrow 2PG+GpmB^{*}\\
2PG+GpmB &\rightarrow 3PG+GpmB^{*}\\
NADP+GL3P+GpsA &\rightarrow NADPH+T3P2+GpsA^{*}\\
NADPH+T3P2+GpsA &\rightarrow NADP+GL3P+GpsA^{*}\\
NADP+ICIT+IcdA &\rightarrow NADPH+CO2+IcdA+AKG\\
NADPH+CO2+IcdA+AKG &\rightarrow NADP+ICIT+IcdA\\
LCTS+LacZ &\rightarrow bDGLAC+GLC+LacZ^{*}\\
NADP+MAL+MaeB &\rightarrow NADPH+PYR+CO2+MaeB\\
NAD+MAL+Mdh &\rightarrow NADH+OA+Mdh\\
NADH+OA+Mdh &\rightarrow NAD+MAL+Mdh\\
NADH+Q+Ndh &\rightarrow NAD+QH2+Ndh\\
NuoABEFGHIJKLMN+2.0\ NADH+2.0\ Q &\rightarrow NuoABEFGHIJKLMN+2.0\ NAD+2.0\ QH2+\\
&\qquad 7.0\ HEXT\\
ATP+OA+PckA &\rightarrow ADP+CO2+PEP+PckA\\
ATP+F6P+PfkA &\rightarrow ADP+FDP+PfkA^{*}\\
ATP+F6P+PfkB &\rightarrow ADP+FDP+PfkB^{*}\\
COA+PYR+PflAB &\rightarrow ACCOA+FOR+PflAB\\
COA+PYR+PflCD &\rightarrow ACCOA+FOR+PflCD\\
G6P+Pgi &\rightarrow F6P+Pgi^{*}\\
F6P+Pgi &\rightarrow G6P+Pgi^{*}\\
ADP+13PDG+Pgk &\rightarrow ATP+3PG+Pgk^{*}\\
ATP+3PG+Pgk &\rightarrow ADP+13PDG+Pgk^{*}\\
Pgl+D6PGL &\rightarrow D6PGC+Pgl\\
G1P+Pgm &\rightarrow G6P+Pgm^{*}\\
G6P+Pgm &\rightarrow G1P+Pgm^{*}\\
NAD+NADPH+PntAB &\rightarrow NADH+NADP+PntAB\\
NADH+NADP+PntAB+2.0\ HEXT &\rightarrow NAD+NADPH+PntAB\\
PPI+Ppa &\rightarrow Ppa+2.0\ PI\\
CO2+PEP+Ppc &\rightarrow PI+OA+Ppc\\
ATP+PYR+PpsA &\rightarrow AMP+PI+PEP+PpsA^{*}\\
ACCOA+PI+Pta &\rightarrow COA+ACTP+Pta\\
COA+ACTP+Pta &\rightarrow ACCOA+PI+Pta\\
ADP+PEP+PykA &\rightarrow ATP+PYR+PykA^{*}\\
ADP+PEP+PykF &\rightarrow ATP+PYR+PykF^{*}\\
ATP+RIB+RbsK &\rightarrow ADP+RbsK+R5P\\
RL5P+Rpe &\rightarrow Rpe+X5P\\
Rpe+X5P &\rightarrow RL5P+Rpe\\
RL5P+RpiA &\rightarrow R5P+RpiA\\
R5P+RpiA &\rightarrow RL5P+RpiA\\
RL5P+RpiB &\rightarrow R5P+RpiB\\
R5P+RpiB &\rightarrow RL5P+RpiB
\end{aligned}
$$

$$
\begin{aligned}
FAD+SUCC+SdhABCD &\rightarrow FADH+FUM+SdhABCD\\
Q+FADH+SdhABCD &\rightarrow QH2+FAD+SdhABCD\\
QH2+FAD+SdhABCD &\rightarrow Q+FADH+SdhABCD\\
NAD+MAL+SfcA &\rightarrow NADH+PYR+CO2+SfcA\\
NAD+COA+LpdA+AKG+SucAB &\rightarrow NADH+CO2+SUCCOA+LpdA+SucAB\\
ADP+PI+SUCCOA+SucCD &\rightarrow ATP+COA+SUCC+SucCD\\
ATP+COA+SUCC+SucCD &\rightarrow ADP+PI+SUCCOA+SucCD\\
T3P1+S7P+TalA &\rightarrow F6P+TalA+E4P\\
F6P+TalA+E4P &\rightarrow T3P1+S7P+TalA\\
T3P1+S7P+TalB &\rightarrow F6P+E4P+TalB\\
F6P+E4P+TalB &\rightarrow T3P1+S7P+TalB\\
R5P+X5P+TktA &\rightarrow T3P1+S7P+TktA\\
T3P1+S7P+TktA &\rightarrow R5P+X5P+TktA\\
X5P+E4P+TktA &\rightarrow T3P1+F6P+TktA\\
T3P1+F6P+TktA &\rightarrow X5P+E4P+TktA\\
R5P+X5P+TktB &\rightarrow T3P1+S7P+TktB\\
T3P1+S7P+TktB &\rightarrow R5P+X5P+TktB\\
X5P+E4P+TktB &\rightarrow T3P1+F6P+TktB\\
T3P1+F6P+TktB &\rightarrow X5P+E4P+TktB\\
T3P1+TpiA &\rightarrow T3P2+TpiA^{*}\\
T3P2+TpiA &\rightarrow T3P1+TpiA^{*}\\
NADP+G6P+Zwf &\rightarrow NADPH+D6PGL+Zwf\\
NADPH+D6PGL+Zwf &\rightarrow NADP+G6P+Zwf\\
ATP &\rightarrow ADP+PI^{*}
\end{aligned}
$$

D.8 Regulatory Reactions

$$
\begin{aligned}
\overline{IclR} &\rightarrow \overline{IclR}+AceA\\
\overline{IclR}+\overline{ArcA} &\rightarrow \overline{IclR}+\overline{ArcA}+AceB\\
\overline{PdhR} &\rightarrow \overline{PdhR}+AceEF+LpdA\\
Food &\rightarrow AcnA+Food\\
Food &\rightarrow AcnB+Food\\
\overline{IclR}+\overline{GLCxt}+\overline{LCTSxt}+\overline{RIBxt}&\\
+\overline{GLxt}+\overline{LACxt}+\overline{PYRxt}+\overline{SUCCxt}+\overline{ETHxt} &\rightarrow \overline{IclR}+\overline{GLCxt}+\overline{LCTSxt}+\overline{RIBxt}\\
&\quad +\overline{GLxt}+\overline{LACxt}+\overline{PYRxt}+\overline{SUCCxt}+\\
&\quad \overline{ETHxt}+Acs\\
\overline{O2xt} &\rightarrow \overline{O2xt}+AdhE\\
\overline{Fnr} &\rightarrow \overline{Fnr}+CydAB\\
ArcA &\rightarrow CydAB+ArcA\\
\overline{ArcA}+\overline{Fnr} &\rightarrow \overline{ArcA}+\overline{Fnr}+CyoABCD\\
Fnr &\rightarrow FdnGHI+Fnr\\
Fnr &\rightarrow FrdABCD+Fnr\\
DcuR &\rightarrow FrdABCD+DcuR\\
\overline{ArcA}+\overline{Fnr} &\rightarrow \overline{ArcA}+\overline{Fnr}+FumA\\
Fnr &\rightarrow FumB+Fnr\\
SOxidRadicals &\rightarrow FumC+SOxidRadicals
\end{aligned}
$$

$$\overline{GLCxt}+\overline{GalR}+\overline{GalS} \rightarrow \overline{GLCxt}+\overline{GalR}+\overline{GalS}+GalE^*$$
$$\overline{GLCxt}+\overline{GalR}+\overline{GalS} \rightarrow \overline{GLCxt}+\overline{GalR}+\overline{GalS}+GalK^*$$
$$\overline{GLCxt}+\overline{GalR}+\overline{GalS} \rightarrow \overline{GLCxt}+\overline{GalR}+\overline{GalS}+GalM^*$$
$$\overline{GLCxt}+\overline{GalR}+\overline{GalS} \rightarrow \overline{GLCxt}+\overline{GalR}+\overline{GalS}+GalT^*$$
$$\overline{GLCxt}+\overline{GlpR}+\overline{LCTSxt}+\overline{RIBxt}+Fnr \rightarrow \overline{GLCxt}+\overline{GlpR}+\overline{LCTSxt}+\overline{RIBxt}+GlpABC+Fnr^{*\,1}$$
$$\overline{GLCxt}+\overline{GlpR}+\overline{LCTSxt}+\overline{RIBxt}+\overline{ArcA} \rightarrow \overline{GLCxt}+\overline{GlpR}+\overline{LCTSxt}+\overline{RIBxt}+\overline{ArcA}+GlpD^{*\,1}$$
$$\overline{GLCxt}+\overline{GlpR}+\overline{LCTSxt}+\overline{RIBxt} \rightarrow \overline{GLCxt}+\overline{GlpR}+\overline{LCTSxt}+\overline{RIBxt}+GlpK^{*\,1}$$
$$\overline{GLCxt}+\overline{LacI} \rightarrow \overline{GLCxt}+\overline{LacI}+LacZ^*$$
$$\overline{ArcA} \rightarrow \overline{ArcA}+Mdh$$
$$\overline{Fnr} \rightarrow \overline{Fnr}+Ndh$$
$$ArcA \rightarrow PflAB+ArcA$$
$$Fnr \rightarrow PflAB+Fnr$$
$$ArcA \rightarrow PflCD+ArcA$$
$$Fnr \rightarrow PflCD+Fnr$$
$$Food \rightarrow Pgk+Food^*$$
$$Cra \rightarrow PpsA+Cra^*$$
$$\overline{Cra} \rightarrow \overline{Cra}+PykF^*$$
$$\overline{GLCxt}+\overline{LCTSxt}+\overline{RbsR} \rightarrow \overline{GLCxt}+\overline{LCTSxt}+\overline{RbsR}+RbsK$$
$$\overline{RpiR} \rightarrow \overline{RpiR}+RpiB$$
$$\overline{ArcA}+\overline{Fnr} \rightarrow \overline{ArcA}+\overline{Fnr}+SdhABCD$$
$$\overline{PdhR} \rightarrow \overline{PdhR}+LpdA+SucAB$$
$$ArcA \rightarrow FocA+ArcA$$
$$Fnr \rightarrow FocA+Fnr$$
$$\overline{Mlc}+Food \rightarrow \overline{Mlc}+Crr+PtsGHI+Food^*$$
$$\overline{Cra}+Food \rightarrow \overline{Cra}+Crr+PtsGHI+Food^*$$
$$Food \rightarrow GalP+Food^*$$
$$\overline{GLCxt}+\overline{GlpR}+\overline{LCTSxt}+\overline{RIBxt} \rightarrow \overline{GLCxt}+\overline{GlpR}+\overline{LCTSxt}+\overline{RIBxt}+GlpF^{*\,1}$$
$$\overline{GLCxt}+\overline{GLxt}+\overline{LCTSxt}+\overline{RIBxt} \rightarrow \overline{GLCxt}+\overline{GLxt}+\overline{LCTSxt}+\overline{RIBxt}+LactateUP$$
$$\overline{GLCxt}+\overline{LacI} \rightarrow \overline{GLCxt}+\overline{LacI}+LacY^*$$
$$\overline{GLCxt}+\overline{LCTSxt}+\overline{RbsR} \rightarrow \overline{GLCxt}+\overline{LCTSxt}+\overline{RbsR}+RbsABCD$$
$$\overline{GLCxt}+\overline{GLxt}+\overline{LACxt}+\overline{LCTSxt}+\overline{PYRxt}+\overline{RIBxt}+\overline{ArcA}+DcuR \rightarrow \overline{GLCxt}+\overline{GLxt}+\overline{LACxt}+\overline{LCTSxt}+\overline{PYRxt}+\overline{RIBxt}+\overline{ArcA}+DctA+DcuR$$
$$\overline{GLCxt}+\overline{GLxt}+\overline{LACxt}+\overline{LCTSxt}+\overline{PYRxt}+\overline{RIBxt}+Fnr+DcuR \rightarrow \overline{GLCxt}+\overline{GLxt}+\overline{LACxt}+\overline{LCTSxt}+\overline{PYRxt}+\overline{RIBxt}+DcuB+Fnr+DcuR$$
$$\overline{ArcA}+\overline{Fnr} \rightarrow \overline{ArcA}+\overline{Fnr}+DcuC$$
$$\overline{O2xt} \rightarrow \overline{O2xt}+ArcA^*$$
$$\overline{G6P}+FDP+fbp \rightarrow \overline{G6P}+FDP+fbp+cra$$
$$DcuS \rightarrow DcuR+DcuS$$
$$SUCCxt \rightarrow SUCCxt+DcuS$$
$$GLCxt \rightarrow GLCxt+FadR^*$$
$$\overline{ACxt} \rightarrow \overline{ACxt}+FadR$$

[1] For the core network model, $\overline{RIBxt}$ was removed from the reaction.

$$
\begin{aligned}
\overline{O2xt} &\rightarrow \overline{O2xt}+Fnr^{*}\\
\overline{LCTSxt} &\rightarrow \overline{LCTSxt}+GalR^{*}\\
\overline{LCTSxt} &\rightarrow \overline{LCTSxt}+GalS^{*}\\
\overline{GLxt} &\rightarrow \overline{GLxt}+GlpR^{*}\\
FadR &\rightarrow FadR+IclR\\
\overline{LCTSxt} &\rightarrow \overline{LCTSxt}+LacI^{*}\\
\overline{GLCxt} &\rightarrow \overline{GLCxt}+Mlc^{*}\\
\overline{PYR} &\rightarrow \overline{PYR}+pdhR\\
\overline{RIBxt} &\rightarrow \overline{RIBxt}+RbsR\\
\overline{RIBxt} &\rightarrow \overline{RIBxt}+RpiR\\
GLCxt &\rightarrow GLCxt+Food^{*}\\
LCTSxt &\rightarrow LCTSxt+Food^{*}\\
RIBxt &\rightarrow RIBxt+Food\\
GLxt &\rightarrow GLxt+Food^{*}\\
LACxt &\rightarrow LACxt+Food\\
PYRxt &\rightarrow PYRxt+Food\\
SUCCxt &\rightarrow SUCCxt+Food\\
ETHxt &\rightarrow ETHxt+Food\\
ACxt &\rightarrow ACxt+Food\\
FORxt &\rightarrow FORxt+Food
\end{aligned}
$$

Appendix E

Genome-Scale Metabolic Network of E. coli

Abbrevation	Metabolite Name
10fthf	10-Formyltetrahydrofolate
12dgr_EC	1,2-Diacylglycerol (E.coli)
12ppd-S	(S)-Propane-1,2-diol
12ppd-S_ex	(S)-Propane-1,2-diol (Extracellular)
13dpg	3-Phospho-D-glyceroyl phosphate
15dap	1,5-Diaminopentane
15dap_ex	1,5-Diaminopentane (Extracellular)
1pyr5c	1-Pyrroline-5-carboxylate
23ddhb	2,3-Dihydro-2,3-dihydroxybenzoate
23dhb	2,3-Dihydroxybenzoate
23dhba	(2,3-Dihydroxybenzoyl)adenylate
23dhdp	2,3-Dihydrodipicolinate
23dhmb	(R)-2,3-Dihydroxy-3-methylbutanoate
23dhmp	(R)-2,3-Dihydroxy-3-methylpentanoate
23doguln	2,3-Dioxo-L-gulonate
25aics	(S)-2-[5-Amino-1-(5-phospho-D-ribosyl)imidazole-4-carboxamido]succinate
25dkglcn	2,5-diketo-D-gluconate
25drapp	2,5-Diamino-6-(ribosylamino)-4-(3H)-pyrimidinone 5'-phosphate
26dap-LL	LL-2,6-Diaminoheptanedioate
26dap-M	meso-2,6-Diaminoheptanedioate
26dap-M_ex	meso-2,6-Diaminoheptanedioate (Extracellular)
2ahbut	(S)-2-Aceto-2-hydroxybutanoate
2aobut	L-2-Amino-3-oxobutanoate
2cpr5p	1-(2-Carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate
2dda7p	2-Dehydro-3-deoxy-D-arabino-heptonate 7-phosphate
2ddg6p	2-Dehydro-3-deoxy-D-gluconate 6-phosphate
2ddglcn	2-Dehydro-3-deoxy-D-gluconate
2ddglcn_ex	2-Dehydro-3-deoxy-D-gluconate (Extracellular)
2dh3dgal	2-Dehydro-3-deoxy-D-galactonate
2dh3dgal6p	2-Dehydro-3-deoxy-D-galactonate 6-phosphate

Abbrevation	Metabolite Name
2dhglcn	2-Dehydro-D-gluconate
2dhguln	2-Dehydro-L-gulonate
2dhp	2-Dehydropantoate
2dmmq8	2-Demethylmenaquinone 8
2dmmql8	2-Demethylmenaquinol 8
2dr1p	2-Deoxy-D-ribose 1-phosphate
2dr5p	2-Deoxy-D-ribose 5-phosphate
2h3oppan	2-Hydroxy-3-oxopropanoate
2ippm	2-Isopropylmaleate
2kmb	2-keto-4-methylthiobutyrate
2mahmp	2-Methyl-4-amino-5-hydroxymethylpyrimidine diphosphate
2mcacn	cis-2-Methylaconitate
2mcit	2-Methylcitrate
2me4p	2-C-methyl-D-erythritol 4-phosphate
2mecdp	2-C-methyl-D-erythritol 2,4-cyclodiphosphate
2obut	2-Oxobutanoate
2ohph	2-Octaprenyl-6-hydroxyphenol
2ombzl	2-Octaprenyl-6-methoxy-1,4-benzoquinol
2omhmbl	2-Octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinol
2ommbl	2-Octaprenyl-3-methyl-6-methoxy- 1,4-benzoquinol
2omph	2-Octaprenyl-6-methoxyphenol
2oph	2-Octaprenylphenol
2p4c2me	2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol
2pg	D-Glycerate 2-phosphate
2pglyc	2-Phosphoglycolate
2shchc	2-Succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate
34hpp	3-(4-Hydroxyphenyl)pyruvate
3c2hmp	3-Carboxy-2-hydroxy-4-methylpentanoate
3c3hmp	3-Carboxy-3-hydroxy-4-methylpentanoate
3c4mop	3-Carboxy-4-methyl-2-oxopentanoate
3dgulnp	3-keto-L-gulonate-6-phosphate
3dhguln	3-Dehydro-L-gulonate
3dhq	3-Dehydroquinate
3dhsk	3-Dehydroshikimate
3hcinnm	3-hydroxycinnamic acid
3hcinnm_ex	3-hydroxycinnamic acid (Extracellular)
3hmrsACP	R-3-hydroxy-myristoyl-ACP
3hpppn	3-(3-hydroxy-phenyl)propionate
3hpppn_ex	3-(3-hydroxy-phenyl)propionate (Extracellular)
3ig3p	C'-(3-Indolyl)-glycerol 3-phosphate
3mob	3-Methyl-2-oxobutanoate
3mop	(S)-3-Methyl-2-oxopentanoate
3ophb	3-Octaprenyl-4-hydroxybenzoate
3pg	3-Phospho-D-glycerate
3php	3-Phosphohydroxypyruvate
3psme	5-O-(1-Carboxyvinyl)-3-phosphoshikimate
4abut	4-Aminobutanoate
4abut_ex	4-Aminobutanoate (Extracellular)

Abbrevation	Metabolite Name
4abutn	4-Aminobutanal
4abz	4-Aminobenzoate
4adcho	4-amino-4-deoxychorismate
4ahmmp	4-Amino-5-hydroxymethyl-2-methylpyrimidine
4ampm	4-Amino-2-methyl-5-phosphomethylpyrimidine
4c2me	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol
4h2opntn	4-Hydroxy-2-oxopentanoate
4hba	4-Hydroxy-benzyl alcohol
4hbz	4-Hydroxybenzoate
4hthr	4-Hydroxy-L-threonine
4mhetz	4-Methyl-5-(2-hydroxyethyl)-thiazole
4mop	4-Methyl-2-oxopentanoate
4mpetz	4-Methyl-5-(2-phosphoethyl)-thiazole
4pasp	4-Phospho-L-aspartate
4per	4-Phospho-D-erythronate
4ppan	D-4'-Phosphopantothenate
4ppcys	N-((R)-4-Phosphopantothenoyl)-L-cysteine
4r5au	4-(1-D-Ribitylamino)-5-aminouracil
5aizc	5-amino-1-(5-phospho-D-ribosyl)imidazole-4-carboxylate
5aop	5-Amino-4-oxopentanoate
5aprbu	5-Amino-6-(5'-phosphoribitylamino)uracil
5apru	5-Amino-6-(5'-phosphoribosylamino)uracil
5caiz	5-phosphoribosyl-5-carboxyaminoimidazole
5dglcn	5-Dehydro-D-gluconate
5dh4dglc	5-Dehydro-4-deoxy-D-glucarate
5mdr1p	5-Methylthio-5-deoxy-D-ribose 1-phosphate
5mdru1p	5-Methylthio-5-deoxy-D-ribulose 1-phosphate
5mta	5-Methylthioadenosine
5mthf	5-Methyltetrahydrofolate
5mtr	5-Methylthio-D-ribose
5prdmbz	N1-(5-Phospho-alpha-D-ribosyl)-5,6-dimethylbenzimidazole
6hmhpt	6-hydroxymethyl dihydropterin
6hmhptpp	6-hydroxymethyl-dihydropterin pyrophosphate
6pgc	6-Phospho-D-gluconate
6pgl	6-phospho-D-glucono-1,5-lactone
8aonn	8-Amino-7-oxononanoate
aacald	Aminoacetaldehyde
aacoa	Acetoacetyl-CoA
ac	Acetate
ac_ex	Acetate (Extracellular)
acac	Acetoacetate
acac_ex	Acetoacetate (Extracellular)
acACP	Acetyl-ACP
acald	Acetaldehyde
acald_ex	Acetaldehyde (Extracellular)
accoa	Acetyl-CoA
acg5p	N-Acetyl-L-glutamyl 5-phosphate
acg5sa	N-Acetyl-L-glutamate 5-semialdehyde

Abbrevation	Metabolite Name
acgam_ex	N-Acetyl-D-glucosamine (Extracellular)
acgam1p	N-Acetyl-D-glucosamine 1-phosphate
acgam6p	N-Acetyl-D-glucosamine 6-phosphate
acglu	N-Acetyl-L-glutamate
acmana	N-Acetyl-D-mannosamine
acmana_ex	N-Acetyl-D-mannosamine (Extracellular)
acmanap	N-Acetyl-D-mannosamine 6-phosphate
acnam	N-Acetylneuraminate
acnam_ex	N-Acetylneuraminate (Extracellular)
aconm	E-3-carboxy-2-pentenedioate 6-methyl ester
acon-T	trans-Aconitate
acorn	N2-Acetyl-L-ornithine
ACP	acyl carrier protein
acser	O-Acetyl-L-serine
actACP	Acetoacetyl-ACP
actp	Acetyl phosphate
ade	Adenine
ade_ex	Adenine (Extracellular)
adn	Adenosine
adn_ex	Adenosine (Extracellular)
adocbi	Adenosyl cobinamide
adocbip	Adenosyl cobinamide phosphate
adocbl	Adenosylcobalamin
adp	ADP
adpglc	ADPglucose
adphep-D,D	ADP-D-glycero-D-manno-heptose
adphep-L,D	ADP-L-glycero-D-manno-heptose
agdpcbi	Adenosine-GDP-cobinamide
agm	Agmatine
agpc_EC	acyl-glycerophosphocholine (E.coli)
agpe_EC	acyl-glycerophosphoethanolamine (E.coli)
agpg_EC	acyl-glycerophosphoglycerol (E.coli)
ahcys	S-Adenosyl-L-homocysteine
ahdt	2-Amino-4-hydroxy-6-(erythro-1,2,3-trihydroxypropyl)dihydropteridine triphosphate
aicar	5-Amino-1-(5-Phospho-D-ribosyl)imidazole-4-carboxamide
air	5-amino-1-(5-phospho-D-ribosyl)imidazole
akg	2-Oxoglutarate
akg_ex	2-Oxoglutarate (Extracellular)
alaala	D-Alanyl-D-alanine
ala-B	beta-Alanine
alac-S	(S)-2-Acetolactate
ala-D	D-Alanine
ala-D_ex	D-Alanine (Extracellular)
ala-L	L-Alanine
ala-L_ex	L-Alanine (Extracellular)
alltn	Allantoin
alltn_ex	Allantoin (Extracellular)

Abbrevation	Metabolite Name
alltt	Allantoate
altrn	D-Altronate
amet	S-Adenosyl-L-methionine
ametam	S-Adenosylmethioninamine
amob	S-Adenosyl-4-methylthio-2-oxobutanoate
amp	AMP
amp_ex	AMP (Extracellular)
anth	Anthranilate
ap4a	P1,P4-Bis(5'-adenosyl) tetraphosphate
ap5a	P1,P5-Bis(5'-adenosyl) pentaphosphate
apg_EC	acyl phosphatidylglycerol (E.coli)
apoACP	apoprotein [acyl carrier protein]
aps	Adenosine 5'-phosphosulfate
ara5p	D-Arabinose 5-phosphate
arab-L	L-Arabinose
arab-L_ex	L-Arabinose (Extracellular)
arbt6p	Arbutin 6-phosphate
arg-L	L-Arginine
arg-L_ex	L-Arginine (Extracellular)
argsuc	N(omega)-(L-Arginino)succinate
asn-L	L-Asparagine
asn-L_ex	L-Asparagine (Extracellular)
asp-L	L-Aspartate
asp-L_ex	L-Aspartate (Extracellular)
aspsa	L-Aspartate 4-semialdehyde
atp	ATP
bbtcoa	gamma-butyrobetainyl-CoA
betald	Betaine aldehyde
btcoa	Butanoyl-CoA
btn	Biotin
btnso	d-biotin d-sulfoxide
but	Butyrate (n-C4:0)
but_ex	Butyrate (n-C4:0) (Extracellular)
camp	cAMP
cbasp	N-Carbamoyl-L-aspartate
cbi	Cobinamide
cbl1	Cob(I)alamin
cbl1_ex	Cob(I)alamin (Extracellular)
cbp	Carbamoyl phosphate
cdp	CDP
cdpdag1	CDPdiacylglycerol (E coli)
cdpea	CDPethanolamine
cechddd	cis-3-(3-carboxyethyl)-3,5-cyclohexadiene-1,2-diol
cenchddd	cis-3-(3-carboxyethenyl)-3,5-cyclohexadiene-1,2-diol
chol	Choline
chol_ex	Choline (Extracellular)
chor	Chorismate
cinnm	trans-Cinnamate

Abbrevation	Metabolite Name
cit	Citrate
cit_ex	Citrate (Extracellular)
citr-L	L-Citrulline
ckdo	CMP-3-deoxy-D-manno-octulosonate
clpn_EC	Cardiolipin (Ecoli)
cmp	CMP
co2	CO2
co2_ex	CO2 (Extracellular)
coa	Coenzyme A
cpppg3	Coproporphyrinogen III
crn	L-Carnitine
crn_ex	L-Carnitine (Extracellular)
crncoa	Carnitinyl-CoA
csn	Cytosine
csn_ex	Cytosine (Extracellular)
ctbt	crotonobetaine
ctbtcoa	crotonobetainyl-CoA
ctp	CTP
cyan	Cyanide
cynt	Cyanate
cynt_ex	Cyanate (Extracellular)
cys-L	L-Cysteine
cys-L_ex	L-Cysteine (Extracellular)
cyst-L	L-Cystathionine
cytd	Cytidine
cytd_ex	Cytidine (Extracellular)
dad-2	Deoxyadenosine
dad-2_ex	Deoxyadenosine (Extracellular)
dadp	dADP
damp	dAMP
dann	7,8-Diaminononanoate
datp	dATP
db4p	3,4-dihydroxy-2-butanone 4-phosphate
dcamp	N6-(1,2-Dicarboxyethyl)-AMP
dcdp	dCDP
dcmp	dCMP
dctp	dCTP
dcyt	Deoxycytidine
dcyt_ex	Deoxycytidine (Extracellular)
ddcaACP	Dodecanoyl-ACP (n-C12:0ACP)
dgdp	dGDP
dgmp	dGMP
dgsn	Deoxyguanosine
dgsn_ex	Deoxyguanosine (Extracellular)
dgtp	dGTP
dha	Dihydroxyacetone
dha_ex	Dihydroxyacetone (Extracellular)
dhap	Dihydroxyacetone phosphate

Abbrevation	Metabolite Name
dhcinnm	2,3-dihydroxicinnamic acid
dhf	7,8-Dihydrofolate
dhna	1,4-Dihydroxy-2-naphthoate
dhnpt	2-Amino-4-hydroxy-6-(D-erythro-1,2,3-trihydroxypropyl)-7,8-dihydropteridine
dhor-S	(S)-Dihydroorotate
dhpmp	Dihydroneopterin monophosphate
dhpppn	3-(2,3-Dihydroxyphenyl)propanoate
dhpt	Dihydropteroate
dhptd	4,5-dihydroxy-2,3-pentanedione
din	Deoxyinosine
din_ex	Deoxyinosine (Extracellular)
dkmpp	2,3-diketo-5-methylthio-1-phosphopentane
dmbzid	5,6-Dimethylbenzimidazole
dmlz	6,7-Dimethyl-8-(1-D-ribityl)lumazine
dmpp	Dimethylallyl diphosphate
dms	Dimethyl sulfide
dms_ex	Dimethyl sulfide (Extracellular)
dmso	Dimethyl sulfoxide
dmso_ex	Dimethyl sulfoxide (Extracellular)
dnad	Deamino-NAD+
dpcoa	Dephospho-CoA
dtbt	Dethiobiotin
dtdp	dTDP
dtdp4aaddg	dTDP-4-acetamido-4,6-dideoxy-D-galactose
dtdp4addg	dTDP-4-amino-4,6-dideoxy-D-glucose
dtdp4d6dg	dTDP-4-dehydro-6-deoxy-D-glucose
dtdp4d6dm	dTDP-4-dehydro-6-deoxy-L-mannose
dtdpglu	dTDPglucose
dtdprmn	dTDP-L-rhamnose
dtmp	dTMP
dttp	dTTP
dudp	dUDP
dump	dUMP
duri	Deoxyuridine
duri_ex	Deoxyuridine (Extracellular)
dutp	dUTP
dxyl	1-deoxy-D-xylulose
dxyl5p	1-deoxy-D-xylulose 5-phosphate
e4p	D-Erythrose 4-phosphate
eca_EC	Enterobacterial common antigen polysaccharide (Ecoli)
eig3p	D-erythro-1-(Imidazol-4-yl)glycerol 3-phosphate
enter	Enterochelin
etha	Ethanolamine
etoh	Ethanol
etoh_ex	Ethanol (Extracellular)
f1p	D-Fructose 1-phosphate
f6p	D-Fructose 6-phosphate

Abbrevation	Metabolite Name
fad	FAD
fadh2	FADH2
fc1p	L-Fuculose 1-phosphate
fcl-L	L-fuculose
fdp	D-Fructose 1,6-bisphosphate
fe2	Fe2+
fe2_ex	Fe2+ (Extracellular)
fgam	N2-Formyl-N1-(5-phospho-D-ribosyl)glycinamide
fmn	FMN
for	Formate
for_ex	Formate (Extracellular)
fpram	2-(Formamido)-N1-(5-phospho-D-ribosyl)acetamidine
fprica	5-Formamido-1-(5-phospho-D-ribosyl)imidazole-4-carboxamide
frdp	Farnesyl diphosphate
fru	D-Fructose
fru_ex	D-Fructose (Extracellular)
fruur	D-Fructuronate
fuc1p-L	L-Fucose 1-phosphate
fuc1p-L_ex	L-Fucose 1-phosphate (Extracellular)
fuc-L	L-Fucose
fuc-L_ex	L-Fucose (Extracellular)
fum	Fumarate
fum_ex	Fumarate (Extracellular)
g1p	D-Glucose 1-phosphate
g3p	Glyceraldehyde 3-phosphate
g3pc	sn-Glycero-3-phosphocholine
g3pe	sn-Glycero-3-phosphoethanolamine
g3pg	Glycerophosphoglycerol
g3pi	sn-Glycero-3-phospho-1-inositol
g3ps	Glycerophosphoserine
g6p	D-Glucose 6-phosphate
g6p_ex	D-Glucose 6-phosphate (Extracellular)
gal	D-Galactose
gal_ex	D-Galactose (Extracellular)
gal1p	alpha-D-Galactose 1-phosphate
galct-D	D-Galactarate
galct-D_ex	D-Galactarate (Extracellular)
galctn-D	D-Galactonate
galctn-D_ex	D-Galactonate (Extracellular)
galt_ex	Galactitol (Extracellular)
galt1p	Galactitol 1-phosphate
galur	D-Galacturonate
galur_ex	D-Galacturonate (Extracellular)
gam_ex	D-Glucosamine (Extracellular)
gam1p	D-Glucosamine 1-phosphate
gam6p	D-Glucosamine 6-phosphate
gar	N1-(5-Phospho-D-ribosyl)glycinamide
gbbtn	gamma-butyrobetaine

Abbrevation	Metabolite Name
gbbtn_ex	gamma-butyrobetaine (Extracellular)
gcald	Glycolaldehyde
gdp	GDP
gdpddman	GDP-4-dehydro-6-deoxy-D-mannose
gdpfuc	GDP-L-fucose
gdpmann	GDP-D-mannose
gdpofuc	GDP-4-oxo-L-fucose
glc-D	D-Glucose
glc-D_ex	D-Glucose (Extracellular)
glcn	D-Gluconate
glcn_ex	D-Gluconate (Extracellular)
glcr	D-Glucarate
glcr_ex	D-Glucarate (Extracellular)
glcur	D-Glucuronate
glcur_ex	D-Glucuronate (Extracellular)
gln-L	L-Glutamine
gln-L_ex	L-Glutamine (Extracellular)
glu1sa	L-Glutamate 1-semialdehyde
glu5p	L-Glutamate 5-phosphate
glu5sa	L-Glutamate 5-semialdehyde
glucys	gamma-L-Glutamyl-L-cysteine
glu-D	D-Glutamate
glu-L	L-Glutamate
glu-L_ex	L-Glutamate (Extracellular)
glutrna	L-Glutamyl-tRNA(Glu)
glx	Glyoxylate
gly	Glycine
gly_ex	Glycine (Extracellular)
glyald	D-Glyceraldehyde
glyald_ex	D-Glyceraldehyde (Extracellular)
glyb	Glycine betaine
glyb_ex	Glycine betaine (Extracellular)
glyc	Glycerol
glyc_ex	Glycerol (Extracellular)
glyc3p	Glycerol 3-phosphate
glyc3p_ex	Glycerol 3-phosphate (Extracellular)
glyclt	Glycolate
glyclt_ex	Glycolate (Extracellular)
glycogen	glycogen
glyc-R	(R)-Glycerate
gmhep17bp	D-Glycero-D-manno-heptose 1,7-bisphosphate
gmhep1p	D-Glycero-D-manno-heptose 1-phosphate
gmhep7p	D-Glycero-D-manno-heptose 7-phosphate
gmp	GMP
gp4g	P1,P4-Bis(5'-guanosyl) tetraphosphate
grdp	Geranyl diphosphate
gsn	Guanosine
gsn_ex	Guanosine (Extracellular)

Abbrevation	Metabolite Name
gthox	Oxidized glutathione
gthrd	Reduced glutathione
gtp	GTP
gtspmd	Glutathionylspermidine
gua	Guanine
gua_ex	Guanine (Extracellular)
h	H+
h_ex	H+ (Extracellular)
h2	H2
h2mb4p	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
h2o	H2O
h2o_ex	H2O (Extracellular)
h2o2	Hydrogen peroxide
h2s	Hydrogen sulfide
hco3	Bicarbonate
hcys-L	L-Homocysteine
hdca	Hexadecanoate (n-C16:0)
hdca_ex	Hexadecanoate (n-C16:0) (Extracellular)
hdcea	hexadecenoate (n-C16:1)
hdeACP	Hexadecenoyl-ACP (n-C16:1ACP)
hemeO	Heme O
his-L	L-Histidine
his-L_ex	L-Histidine (Extracellular)
hisp	L-Histidinol phosphate
histd	L-Histidinol
hkndd	2-Hydroxy-6-oxonona-2,4-diene-1,9-dioate
hkntd	2-hydroxy-6-ketononatrienedioate
hmbil	Hydroxymethylbilane
hmfurn	4-hydroxy-5-methyl-3(2H)-furanone
hom-L	L-Homoserine
hpyr	Hydroxypyruvate
hqn	Hydroquinone
hxan	Hypoxanthine
hxan_ex	Hypoxanthine (Extracellular)
iasp	Iminoaspartate
ichor	Isochorismate
icit	Isocitrate
idon-L	L-Idonate
idon-L_ex	L-Idonate (Extracellular)
idp	IDP
ile-L	L-Isoleucine
ile-L_ex	L-Isoleucine (Extracellular)
imacp	3-(Imidazol-4-yl)-2-oxopropyl phosphate
imp	IMP
indole	Indole
indole_ex	Indole (Extracellular)
inost	myo-Inositol
ins	Inosine

Abbrevation	Metabolite Name
ins_ex	Inosine (Extracellular)
ipdp	Isopentenyl diphosphate
itp	ITP
k	K+
k_ex	K+ (Extracellular)
kdo	3-Deoxy-D-manno-2-octulosonate
kdo2lipid4	KDO(2)-lipid IV(A)
kdo2lipid4L	KDO(2)-lipid IV(A) with laurate
kdo2lipid4p	KDO(2)-lipid IV(A) with palmitoleoyl
kdo8p	3-Deoxy-D-manno-octulosonate 8-phosphate
kdolipid4	KDO-lipid IV(A)
lac-D	D-Lactate
lac-D_ex	D-Lactate (Extracellular)
lac-L	L-Lactate
lac-L_ex	L-Lactate (Extracellular)
lald-L	L-Lactaldehyde
lcts	Lactose
lcts_ex	Lactose (Extracellular)
leu-L	L-Leucine
leu-L_ex	L-Leucine (Extracellular)
lgt-S	(R)-S-Lactoylglutathione
lipa	KDO(2)-lipid (A)
lipa_cold	cold adapted KDO(2)-lipid (A)
lipidA	2,3,2'3'-Tetrakis(beta-hydroxymyristoyl)-D-glucosaminyl-1,6-beta-D-glucosamine 1,4'-bisphosphate
lipidAds	Lipid A Disaccharide
lipidX	2,3-Bis(3-hydroxytetradecanoyl)-beta-D-glucosaminyl 1-phosphate
lps_EC	lipopolysaccharide (Ecoli)
lys-L	L-Lysine
lys-L_ex	L-Lysine (Extracellular)
malACP	Malonyl-[acyl-carrier protein]
malcoa	Malonyl-CoA
mal-L	L-Malate
mal-L_ex	L-Malate (Extracellular)
malt	Maltose
malt_ex	Maltose (Extracellular)
malt6p	Maltose 6'-phosphate
malthp	Maltoheptaose
malthx	Maltohexaose
malthx_ex	Maltohexaose (Extracellular)
maltpt	Maltopentaose
maltpt_ex	Maltopentaose (Extracellular)
malttr	Maltotriose
malttr_ex	Maltotriose (Extracellular)
maltttr	Maltotetraose
maltttr_ex	Maltotetraose (Extracellular)
man_ex	D-Mannose (Extracellular)
man1p	D-Mannose 1-phosphate

Abbrevation	Metabolite Name
man6p	D-Mannose 6-phosphate
man6p_ex	D-Mannose 6-phosphate (Extracellular)
mana	D-Mannonate
melib	Melibiose
melib_ex	Melibiose (Extracellular)
met-D	D-Methionine
met-D_ex	D-Methionine (Extracellular)
methf	5,10-Methenyltetrahydrofolate
met-L	L-Methionine
met-L_ex	L-Methionine (Extracellular)
mi1p-D	1D-myo-Inositol 1-phosphate
micit	methylisocitrate
mlthf	5,10-Methylenetetrahydrofolate
mmcoa-R	(R)-Methylmalonyl-CoA
mmcoa-S	(S)-Methylmalonyl-CoA
mnl_ex	D-Mannitol (Extracellular)
mnl1p	D-Mannitol 1-phosphate
mql8	Menaquinol 8
mqn8	Menaquinone 8
mthgxl	Methylglyoxal
myrsACP	Myristoyl-ACP (n-C14:0ACP)
N1aspmd	N1-Acetylspermidine
n8aspmd	N8-Acetylspermidine
na1	Sodium
na1_ex	Sodium (Extracellular)
nac	Nicotinate
nac_ex	Nicotinate (Extracellular)
nad	Nicotinamide adenine dinucleotide
nad_ex	Nicotinamide adenine dinucleotide (Extracellular)
nadh	Nicotinamide adenine dinucleotide - reduced
nadp	Nicotinamide adenine dinucleotide phosphate
nadph	Nicotinamide adenine dinucleotide phosphate - reduced
ncam	Nicotinamide
nh4	ammonium
nh4_ex	ammonium (Extracellular)
nicrnt	Nicotinate D-ribonucleotide
nmn	NMN
nmn_ex	NMN (Extracellular)
no2	Nitrite
no2_ex	Nitrite (Extracellular)
no3	Nitrate
no3_ex	Nitrate (Extracellular)
o2	O2
o2_ex	O2 (Extracellular)
o2-	Superoxide anion
oaa	Oxaloacetate
ocdca	octadecanoate (n-C18:0)
ocdca_ex	octadecanoate (n-C18:0) (Extracellular)

Abbrevation	Metabolite Name
ocdcea	octadecenoate (n-C18:1)
octdp	all-trans-Octaprenyl diphosphate
octeACP	Octadecenoyl-ACP (n-C18:1ACP)
ohpb	2-Oxo-3-hydroxy-4-phosphobutanoate
op4en	2-Oxopent-4-enoate
orn	Ornithine
orn_ex	Ornithine (Extracellular)
orot	Orotate
orot5p	Orotidine 5'-phosphate
pa_EC	phosphatidate (E.coli)
pac	Phenylacetic acid
pacald	Phenylacetaldehyde
palmACP	Palmitoyl-ACP (n-C16:0ACP)
pan4p	Pantetheine 4'-phosphate
pant-R	(R)-Pantoate
pap	Adenosine 3',5'-bisphosphate
paps	3'-Phosphoadenylyl sulfate
pc_EC	Phosphatidylcholine (E.coli)
pdx5p	Pyridoxine 5'-phosphate
pe_EC	Phosphatidylethanolamine (Ecoli)
peamn	Phenethylamine
pep	Phosphoenolpyruvate
peptido_EC	Peptidoglycan subunit of Escherichia coli
pg_EC	Phosphatidylglycerol (Ecoli)
pgp_EC	Phosphatidylglycerophosphate (Ecoli)
phaccoa	Phenylacetyl-CoA
phe-L	L-Phenylalanine
phe-L_ex	L-Phenylalanine (Extracellular)
pheme	Protoheme
phom	O-Phospho-L-homoserine
phpyr	Phenylpyruvate
phthr	O-Phospho-4-hydroxy-L-threonine
pi	Phosphate
pi_ex	Phosphate (Extracellular)
pmcoa	Pimeloyl-CoA
pnto-R	(R)-Pantothenate
pnto-R_ex	(R)-Pantothenate (Extracellular)
ppa	Propionate
ppap	Propanoyl phosphate
ppbng	Porphobilinogen
ppcoa	Propanoyl-CoA
pphn	Prephenate
ppi	Diphosphate
ppp9	Protoporphyrin
pppg9	Protoporphyrinogen IX
pppi	Inorganic triphosphate
pppn	Phenylpropanoate
pppn_ex	Phenylpropanoate (Extracellular)

Abbrevation	Metabolite Name
pram	5-Phospho-beta-D-ribosylamine
pran	N-(5-Phospho-D-ribosyl)anthranilate
prbamp	1-(5-Phosphoribosyl)-AMP
prbatp	1-(5-Phosphoribosyl)-ATP
prfp	1-(5-Phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide
prlp	5-[(5-phospho-1-deoxyribulos-1-ylamino)methylideneamino]-1-(5-phosphoribosyl)imidazole-4-carboxamide
pro-L	L-Proline
pro-L_ex	L-Proline (Extracellular)
prpp	5-Phospho-alpha-D-ribose 1-diphosphate
ps_EC	phosphatidylserine (Ecoli)
pser-L	O-Phospho-L-serine
ptrc	Putrescine
ptrc_ex	Putrescine (Extracellular)
pyam5p	Pyridoxamine 5'-phosphate
pydam	Pyridoxamine
pydx	Pyridoxal
pydx5p	Pyridoxal 5'-phosphate
pydxn	Pyridoxine
pyr	Pyruvate
pyr_ex	Pyruvate (Extracellular)
q8	Ubiquinone-8
q8h2	Ubiquinol-8
quln	Quinolinate
r1p	alpha-D-Ribose 1-phosphate
r5p	alpha-D-Ribose 5-phosphate
rbl-L	L-Ribulose
rdmbzi	N1-(alpha-D-ribosyl)-5,6-dimethylbenzimidazole
rhcys	S-Ribosyl-L-homocysteine
rib-D	D-Ribose
rib-D_ex	D-Ribose (Extracellular)
ribflv	Riboflavin
rml	L-Rhamnulose
rml1p	L-Rhamnulose 1-phosphate
rmn	L-Rhamnose
rmn_ex	L-Rhamnose (Extracellular)
ru5p-D	D-Ribulose 5-phosphate
ru5p-L	L-Ribulose 5-phosphate
s7p	Sedoheptulose 7-phosphate
sbt6p	D-Sorbitol 6-phosphate
sbt-D_ex	D-Sorbitol (Extracellular)
sbzcoa	O-Succinylbenzoyl-CoA
seln	Selenide
selnp	Selenophosphate
seramp	L-seryl-AMP
ser-D	D-Serine
ser-D_ex	D-Serine (Extracellular)

Abbrevation	Metabolite Name
ser-L	L-Serine
ser-L_ex	L-Serine (Extracellular)
shcl	Sirohydrochlorin
sheme	Siroheme
skm	Shikimate
skm5p	Shikimate 5-phosphate
sl26da	N-Succinyl-LL-2,6-diaminoheptanedioate
sl2a6o	N-Succinyl-2-L-amino-6-oxoheptanedioate
so3	Sulfite
so4	Sulfate
so4_ex	Sulfate (Extracellular)
spmd	Spermidine
spmd_ex	Spermidine (Extracellular)
srch	Sirochlorin
ssaltpp	Succinate semialdehyde-thiamin diphosphate anion
suc6p	Sucrose 6-phosphate
sucarg	N2-Succinyl-L-arginine
sucbz	o-Succinylbenzoate
succ	Succinate
succ_ex	Succinate (Extracellular)
succoa	Succinyl-CoA
sucglu	N2-Succinyl-L-glutamate
sucgsa	N2-Succinyl-L-glutamate 5-semialdehyde
suchms	O-Succinyl-L-homoserine
sucorn	N2-Succinyl-L-ornithine
sucr_ex	Sucrose (Extracellular)
sucsal	Succinic semialdehyde
tag6p-D	D-Tagatose 6-phosphate
tagdp-D	D-Tagatose 1,6-biphosphate
tagur	D-Tagaturonate
tartr-L	L-tartrate
tartr-L_ex	L-tartrate (Extracellular)
taur	Taurine
taur_ex	Taurine (Extracellular)
tcynt	Thiocyanate
tdeACP	Tetradecenoyl-ACP (n-C14:1ACP)
thdp	2,3,4,5-Tetrahydrodipicolinate
thf	5,6,7,8-Tetrahydrofolate
thm	Thiamin
thm_ex	Thiamin (Extracellular)
thmmp	Thiamin monophosphate
thmpp	Thiamine diphosphate
thr-L	L-Threonine
thr-L_ex	L-Threonine (Extracellular)
thym	Thymine
thymd	Thymidine
thymd_ex	Thymidine (Extracellular)
tma	Trimethylamine

Abbrevation	Metabolite Name
tma_ex	Trimethylamine (Extracellular)
tmao	Trimethylamine N-oxide
tmao_ex	Trimethylamine N-oxide (Extracellular)
trdox	Oxidized thioredoxin
trdrd	Reduced thioredoxin
tre	Trehalose
tre_ex	Trehalose (Extracellular)
tre6p	alpha,alpha'-Trehalose 6-phosphate
trnaglu	tRNA (Glu)
trp-L	L-Tryptophan
trp-L_ex	L-Tryptophan (Extracellular)
tsul	Thiosulfate
tsul_ex	Thiosulfate (Extracellular)
ttdca	tetradecanoate (n-C14:0)
ttdca_ex	tetradecanoate (n-C14:0) (Extracellular)
ttdcea	tetradecenoate (n-C14:1)
tyr-L	L-Tyrosine
tyr-L_ex	L-Tyrosine (Extracellular)
u23ga	UDP-2,3-bis(3-hydroxytetradecanoyl)glucosamine
u3aga	UDP-3-O-(3-hydroxytetradecanoyl)-N-acetylglucosamine
u3hga	UDP-3-O-(3-hydroxytetradecanoyl)-D-glucosamine
uaagmda	Undecaprenyl-diphospho-N-acetylmuramoyl-(N-acetylglucosamine)-L-ala-D-glu-meso-2,6-diaminopimeloyl-D-ala-D-ala
uaccg	UDP-N-acetyl-3-O-(1-carboxyvinyl)-D-glucosamine
uacgam	UDP-N-acetyl-D-glucosamine
uacmam	UDP-N-acetyl-D-mannosamine
uacmamu	UDP-N-acetyl-D-mannosaminouronate
uagmda	Undecaprenyl-diphospho-N-acetylmuramoyl-L-alanyl-D-glutamyl-meso-2,6-diaminopimeloyl-D-alanyl-D-alanine
uama	UDP-N-acetylmuramoyl-L-alanine
uamag	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate
uamr	UDP-N-acetylmuramate
udcpdp	Undecaprenyl diphosphate
udcpp	Undecaprenyl phosphate
udp	UDP
udpg	UDPglucose
udpgal	UDPgalactose
udpgalfur	UDP-D-galacto-1,4-furanose
udpglcur	UDP-D-glucuronate
ugmd	UDP-N-acetylmuramoyl-L-alanyl-D-gamma-glutamyl-meso-2,6-diaminopimelate
ugmda	UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-meso-2,6-diaminopimeloyl-D-alanyl-D-alanine
ump	UMP
unaga	Undecaprenyl diphospho N-acetyl-glucosamine
unagamu	Undecaprenyl-diphospho-N-acetylglucosamine-N-acetylmannosaminuronate

Abbrevation	Metabolite Name
unagamuf	Undecaprenyl-diphospho N-acetylglucosamine-N-acetylmannosaminuronate-N-acetamido-4,6-dideoxy-D-galactose
uppg3	Uroporphyrinogen III
ura	Uracil
ura_ex	Uracil (Extracellular)
urdglyc	(-)-Ureidoglycolate
urea	Urea
urea_ex	Urea (Extracellular)
uri	Uridine
uri_ex	Uridine (Extracellular)
utp	UTP
val-L	L-Valine
val-L_ex	L-Valine (Extracellular)
xan	Xanthine
xan_ex	Xanthine (Extracellular)
xmp	Xanthosine 5'-phosphate
xtsn	Xanthosine
xtsn_ex	Xanthosine (Extracellular)
xu5p-D	D-Xylulose 5-phosphate
xu5p-L	L-Xylulose 5-phosphate
xyl-D	D-Xylose
xyl-D_ex	D-Xylose (Extracellular)
xylu-D	D-Xylulose

References

Albert, R., Jeong, H., Barabási, A.-L., 2000. Error and attack tolerance of complex networks. Nature 406 (6794), 378–382.
URL http://dx.doi.org/10.1038/35019019

Alm, E., Arkin, A. P., 2003. Biological networks. Curr Opin Struct Biol 13 (2), 193–202.

Alon, U., 2003. Biological networks: the tinkerer as an engineer. Science 301 (5641), 1866–1867.
URL http://dx.doi.org/10.1126/science.1089072

Anderson, A., Cooper, R., 1969. Gluconeogenesis in Escherichia coli The role of triose phosphate isomerase. FEBS Lett 4 (1), 19–20.

Arkin, A., Ross, J., McAdams, H. H., 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. Genetics 149 (4), 1633–1648.

Asthagiri, A. R., Lauffenburger, D. A., 2001. A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. Biotechnol Prog 17 (2), 227–239.

Bader, G. D., Donaldson, I., Wolting, C., Ouellette, B. F., Pawson, T., Hogue, C. W., 2001. BIND–The Biomolecular Interaction Network Database. Nucleic Acids Res 29 (1), 242–245.

Bader, G. D., Hogue, C. W. V., 2003. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 4, 2.

Bailey, J. E., 2001. Complex biology with no parameters. Nat Biotechnol 19 (6), 503–504.
URL http://dx.doi.org/10.1038/89204

Barabási, A.-L., Oltvai, Z. N., 2004. Network biology: understanding the cell's functional organization. Nat Rev Genet 5 (2), 101–113.
URL http://dx.doi.org/10.1038/nrg1272

Berkelaar, M., Eikland, K., Notebaert, P., 2005. lp solve: Open source (mixed-integer) linear programming system, version 5.5.
URL http://tech.groups.yahoo.com/group/lp\protect\unhbox\voidb@x\kern.06em\vbox{\hrulewidth.3em}solve/

Blinov, M. L., Faeder, J. R., Goldstein, B., Hlavacek, W. S., 2004. BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. Bioinformatics 20 (17), 3289–3291.
URL http://dx.doi.org/10.1093/bioinformatics/bth378

REFERENCES

Blinov, M. L., Faeder, J. R., Goldstein, B., Hlavacek, W. S., 2006. A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity. Biosystems 83 (2-3), 136–151.
URL http://dx.doi.org/10.1016/j.biosystems.2005.06.014

Botsford, J. L., Harman, J. G., 1992. Cyclic amp in prokaryotes. Microbiol Rev 56 (1), 100–122.

Bower, J. M., Bolouri, H. (Eds.), 2000. Computational Modeling of Genetic and Biochemical Networks. The MIT Press, London.

Bray, D., 2003. Molecular networks: the top-down view. Science 301 (5641), 1864–1865.
URL http://dx.doi.org/10.1126/science.1089118

Bu, D., Zhao, Y., Cai, L., Xue, H., Zhu, X., Lu, H., Zhang, J., Sun, S., Ling, L., Zhang, N., Li, G., Chen, R., 2003. Topological structure analysis of the protein-protein interaction network in budding yeast. Nucleic Acids Res 31 (9), 2443–2450.

Calhoun, M. W., Oden, K. L., Gennis, R. B., de Mattos, M. J., Neijssel, O. M., 1993. Energetic efficiency of Escherichia coli: effects of mutations in components of the aerobic respiratory chain. J Bacteriol 175 (10), 3020–3025.

Clarke, B. L., 1980. Stability of complex reaction networks. Adv Chem Phys 42, 1–213.

Courtright, J. B., Henning, U., 1970. Malate dehydrogenase mutants in Escherichia coli K-12. J Bacteriol 102 (3), 722–728.

Coutinho, A., 1995. The network theory: 21 years later. Scand J Immunol 42 (1), 3–8.

Covert, M. W., Palsson, B., 2002. Transcriptional regulation in constraints-based metabolic models of escherichia coli. J Biol Chem 277 (31), 28058–28064.

Covert, M. W., Schilling, C. H., Palsson, B., 2001. Regulation of gene expression in flux balance models of metabolism. J Theor Biol 213 (1), 73–88.

Creaghan, I. T., Guest, J. R., 1978. Succinate dehydrogenase-dependent nutritional requirement for succinate in mutants of Escherichia coli K12. J Gen Microbiol 107 (1), 1–13.

Cronan, J. E., LaPorte, D., 1996. Tricarboxylic acid cycle and glyoxylate bypass. In: Neidhardt, F. C., III, R. C., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., Umbarger, H. E. (Eds.), Escherichia coli and Salmonella: cellular and molecular biology. Vol. 1. ASM Press, Washington, D. C., pp. 206–216.

Csete, M., Doyle, J., 2004. Bow ties, metabolism and disease. Trends Biotechnol 22 (9), 446–450.
URL http://dx.doi.org/10.1016/j.tibtech.2004.07.007

Cunningham, P. R., Clark, D. P., 1986. The use of suicide substrates to select mutants of Escherichia coli lacking enzymes of alcohol fermentation. Mol Gen Genet 205 (3), 487–493.

Dantzig, G. B., 1963. Linear Programming and Extensions. Princeton University Press, Princeton.

de Jong, H., 2002. Modeling and simulation of genetic regulatory systems: a literature review. J Comput Biol 9 (1), 67–103.
URL http://dx.doi.org/10.1089/10665270252833208

der Heijden, R. T. J. M. V., Heijnen, J. J., Hellinga, C., Romein, B., Luyben, K. C. A. M., 1994. Linear constraint relations in biochemical reaction systems: I. classification of the calculability and the balancebility of conversion rates. Biotechnol Bioeng 43, 3–10.

Deville, Y., Gilbert, D., van Helden, J., Wodak, S. J., 2003. An overview of data models for the analysis of biochemical pathways. Brief Bioinform 4 (3), 246–259.

Dittrich, P., Speroni di Fenizio, P., 2007. Chemical organization theory. B Math Biol 69 (4), 1199–1231.

Dobrijevic, M., Dutour, I., 2006. A random graphs model for the study of chemical complexity in planetary atmospheres. Planet Space Sci 54 (3), 287–295.

Doi, A., 2005. A hfpn model of bacteriophage lambda, `http://www.genomicobject.net/member3/GONET/lambda.html`, accessed in Feb. 2005.

Dunne, J., 2006. The network structure of food webs. In: Pascual, M., Dunne, J. (Eds.), Ecological Networks: Linking Structure to Dynamics in Food Webs. Oxford University Press, pp. 27–86.

Edwards, J. S., Ibarra, R. U., Palsson, B. O., 2001. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. Nat Biotechnol 19 (2), 125–130.
URL http://dx.doi.org/10.1038/84379

Edwards, J. S., Palsson, B. O., 2000. The Escherichia coli MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. Proc Natl Acad Sci U S A 97 (10), 5528–5533.

Espinosa-Soto, C., Padilla-Longoria, P., Alvarez-Buylla, E. R., 2004. A gene regulatory network model for cell-fate determination during arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. Plant Cell 16 (11), 2923–2939.

Feinberg, M., Horn, F. J. M., 1974. Dynamics of open chemical systems and the algebraic structure of the underlying reaction network. Chem Eng Sci 29 (3), 775–787.

Ferrell, J. E., Xiong, W., 2001. Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. Chaos 11 (1), 227–236.
URL http://dx.doi.org/10.1063/1.1349894

Finney, A., Hucka, M., 2003. Systems biology markup language: Level 2 and beyond. Biochem Soc Trans 31 (Pt 6), 1472–1473.
URL http://dx.doi.org/10.1042/

Fong, S. S., Palsson, B. O., 2004. Metabolic gene-deletion strains of escherichia coli evolve to computationally predicted growth phenotypes. Nat Genet 36 (10), 1056–1058.
URL http://dx.doi.org/10.1038/ng1432

REFERENCES

Fontana, W., Buss, L. W., 1994. 'The arrival of the fittest': Towards a theory of biological organization. B Math Biol 56, 1–64.

Fraenkel, D. G., 1996. Glycolysis. In: Neidhardt, F. C., III, R. C., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., Umbarger, H. E. (Eds.), Escherichia coli and Salmonella: cellular and molecular biology. Vol. 1. ASM Press, Washington, D. C., pp. 189–198.

Fraenkel, D. G., Horecker, B. L., 1965. Fructose-1, 6-diphosphatase and acid hexose phosphatase of Escherichia coli. J Bacteriol 90 (4), 837–842.

Gagneur, J., Jackson, D. B., Casari, G., 2003. Hierarchical analysis of dependency in metabolic networks. Bioinformatics 19 (8), 1027–1034.

Gagneur, J., Klamt, S., 2004. Computation of elementary modes: a unifying framework an the new binary approach. BMC Bioinformatics 5, 175.

Gagneur, J., Krause, R., Bouwmeester, T., Casari, G., 2004. Modular decomposition of protein-protein interaction networks. Genome Biol 5 (8), R57.
URL http://dx.doi.org/10.1186/gb-2004-5-8-r57

Gardner, T. S., Cantor, C. R., Collins, J. J., 2000. Construction of a genetic toggle switch in Escherichia coli. Nature 403 (6767), 339–342.
URL http://dx.doi.org/10.1038/35002131

Ge, H., Liu, Z., Church, G. M., Vidal, M., 2001. Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. Nat Genet 29 (4), 482–486.
URL http://dx.doi.org/10.1038/ng776

Genoud, Métraux, 1999. Crosstalk in plant cell signaling: structure and function of the genetic network. Trends Plant Sci 4 (12), 503–507.

Gleiss, P. M., Stadler, P. F., Wagner, A., Fell, D. A., 2001. Relevant cycles in chemical reaction networks. Adv Complex Syst 4, 207–226.

Grigoriev, A., 2001. A relationship between gene expression and protein interactions on the proteome scale: analysis of the bacteriophage T7 and the yeast Saccharomyces cerevisiae. Nucleic Acids Res 29 (17), 3513–3519.

Gruer, M. J., Bradbury, A. J., Guest, J. R., 1997. Construction and properties of aconitase mutants of Escherichia coli. Microbiology 143 (Pt 6), 1837–1846.

Han, J., Kamber, M., 2006. Data mining: concepts and techniques. Elsevier/Morgan Kaufmann.

Hanamura, A., Aiba, H., 1992. A new aspect of transcriptional control of the *Escherichia coli* *crp* gene: positive autoregulation. Mol Microbiol 17 (6), 2489–2497.

Handorf, T., Ebenhöh, O., Heinrich, R., 2005. Expanding metabolic networks: scopes of compounds, robustness, and evolution. J Mol Evol 61 (4), 498–512.
URL http://dx.doi.org/10.1007/s00239-005-0027-1

Hargrove, J. L., Hulsey, M. G., Beale, E. G., 1991. The kinetics of mammalian gene expression. Bioessays 13 (12), 667–674.

Hartwell, L. H., Hopfield, J. J., Leibler, S., Murray, A. W., 1999. From molecular to modular cell biology. Nature 402, C47–C52.

Heidtke, K. R., Schulze-Kremer, S., 1998. Design and implementation of a qualitative simulation model of lambda phage infection. Bioinformatics 14 (1), 81–91.

Heinrich, R., Schuster, S., 1996. The Regulation of Cellular Systems. Chapman & Hall, New York.

Houslay, M. D., Kolch, W., 2000. Cell-type specific integration of cross-talk between extracellular signal-regulated kinase and cAMP signaling. Mol Pharmacol 58 (4), 659–668.

Ideker, T., Galitski, T., Hood, L., 2001. A new approach to decoding life: systems biology. Annu Rev Genomics Hum Genet 2, 343–372.
URL http://dx.doi.org/2.1.343

Ideker, T., Ozier, O., Schwikowski, B., Siegel, A. F., 2002. Discovering regulatory and signalling circuits in molecular interaction networks. Bioinformatics 18 Suppl 1, S233–S240.

Irani, M. H., Maitra, P. K., 1977. Properties of Escherichia coli mutants deficient in enzymes of glycolysis. J Bacteriol 132 (2), 398–410.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., Barabási, A. L., 2000. The large-scale organization of metabolic networks. Nature 407 (6804), 651–654.
URL http://dx.doi.org/10.1038/35036627

Jetschke, G., 1989. Mathematik der Selbstorganisation. Harri Deutsch.

Kaleta, C., 2005. personal communication.

Kaleta, C., 2007. personal communication.

Kaleta, C., Centler, F., Dittrich, P., 2006. Analyzing molecular reaction networks: from pathways to chemical organizations. Mol Biotechnol 34 (2), 117–123.

Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28 (1), 27–30.

Karp, P. D., Ouzounis, C. A., Moore-Kochlacs, C., Goldovsky, L., Kaipa, P., Ahrn, D., Tsoka, S., Darzentas, N., Kunin, V., López-Bigas, N., 2005. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. Nucleic Acids Res 33 (19), 6083–6089.
URL http://dx.doi.org/10.1093/nar/gki892

Kauffman, S. A., 1993. The Origins of Order: Self-Organization and Selection in Evolution. Oxford University Press.

Keller, A. D., 1995. Model genetic circuits encoding autoregulatory transcription factors. J Theor Biol 172 (2), 169–185.
URL http://dx.doi.org/10.1006/jtbi.1995.0014

Kiehl, T. R., Mattheyses, R. M., Simmons, M. K., 2004. Hybrid simulation of cellular behavior. Bioinformatics 20 (3), 316–322.

Kitano, H. (Ed.), 2001. Foundations of Systems Biology. The MIT Press.

Kitano, H., 2002. Systems biology: a brief overview. Science 295 (5560), 1662–1664. URL http://dx.doi.org/10.1126/science.1069492

Klamt, S., Schuster, S., Gilles, E. D., 2002. Calculability analysis in underdetermined metabolic networks illustrated by a model of the central metabolism in purple nonsulfur bacteria. Biotechnol Bioeng 77 (7), 734–751.

Klamt, S., Stelling, J., 2002. Combinatorial complexity of pathway analysis in metabolic networks. Mol Biol Rep 29 (1-2), 233–236.

Klamt, S., Stelling, J., 2003. Two approaches for metabolic pathway analysis? Trends Biotechnol 21 (2), 64–69.

Klee, V., Minty, G. J., 1972. How good is the simplex algorithm? In: Shisha, O. (Ed.), Inequalities. Vol. III. Academic Press, New York, pp. 159–175.

Klipp, E., Herwig, R., Kowald, A., Wierling, C., Lehrach, H., 2005a. Systems Biology in Practice. WILEY-VCH.

Klipp, E., Nordlander, B., Krüger, R., Gennemark, P., Hohmann, S., 2005b. Integrative model of the response of yeast to osmotic shock. Nat Biotechnol 23 (8), 975–982. URL http://dx.doi.org/10.1038/nbt1114

Kremling, A., Bettenbrock, K., Fischer, S., Ginkel, M., Sauter, T., Gilles, E. D., 2003. Towards whole cell "in silico" models for cellular systems: model set-up and model validation. Proceedings of the first multidisciplinary international symposium on Positive Systems: theory and applications, Rome, 95–102.

Kremling, A., Bettenbrock, K., Laube, B., Jahreis, K., Lengeler, J. W., Gilles, E. D., 2001. The organization of metabolic reaction networks, iii. application for diauxic growth on glucose and lactose. Metab Eng 3, 362–379.

Kremling, A., Jahreis, K., Lengeler, J. W., Gilles, E. D., 2000. The organization of metabolic reaction networks: a signal-oriented approach to cellular models. Metab Eng 2 (3), 190–200. URL http://dx.doi.org/10.1006/mben.2000.0159

Kumari, S., Tishel, R., Eisenbach, M., Wolfe, A. J., 1995. Cloning, characterization, and functional expression of acs, the gene which encodes acetyl coenzyme A synthetase in Escherichia coli. J Bacteriol 177 (10), 2878–2886.

Langley, D., Guest, J. R., 1977. Biochemical genetics of the alpha-keto acid dehydrogenase complexes of Escherichia coli K12: isolation and biochemical properties of deletion mutants. J Gen Microbiol 99 (2), 263–276.

Larson, T. J., Ye, S. Z., Weissenborn, D. L., Hoffmann, H. J., Schweizer, H., 1987. Purification and characterization of the repressor for the sn-glycerol 3-phosphate regulon of escherichia coli k12. J Biol Chem 262 (33), 15869–15874.

Lautenbach, K., 1973. Exact liveness conditions of a petri net class (in german). Berichte der GMD 82.

Legendre, P., Makarenkov, V., 2002. Reconstruction of biogeographic and evolutionary networks using reticulograms. Syst Biol 51 (2), 199–216.
URL http://dx.doi.org/10.1080/10635150252899725

Lin, E. C. C., 1976. Glycerol dissimilation and its regulation in bacteria. Ann Rev Microbiol 30, 535–578.

Loew, L. M., Schaff, J. C., 2001. The virtual cell: a software environment for computational cell biology. Trends Biotechnol 19 (10), 401–406.

Luscombe, N. M., Babu, M. M., Yu, H., Snyder, M., Teichmann, S. A., Gerstein, M., 2004. Genomic analysis of regulatory network dynamics reveals large topological changes. Nature 431 (7006), 308–312.
URL http://dx.doi.org/10.1038/nature02782

Ma, H., Zeng, A.-P., 2003. Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. Bioinformatics 19 (2), 270–277.

Maloy, S. R., Nunn, W. D., 1982. Genetic regulation of the glyoxylate shunt in escherichia coli k-12. J Bacteriol 149 (1), 173–180.

Mat-Jan, F., Alam, K. Y., Clark, D. P., 1989. Mutants of Escherichia coli deficient in the fermentative lactate dehydrogenase. J Bacteriol 171 (1), 342–348.

Matsumaru, N., Centler, F., di Fenizio, P. S., Dittrich, P., 2007. Chemical organization theory as a theoretical base for chemical computing. Int J Unconv Comp.

Matsuno, H., Doi, A., Nagasaki, M., Miyano, S., 2000. Hybrid petri net representation of gene regulatory network. In: Altman, R. B., Dunker, A. K., Hunter, L., Klein, T. E. (Eds.), Proceedings of the Pacific Symposium on Biocomputing (PSB). pp. 341–352.

Matsuno, H., Tanaka, Y., Aoshima, H., Doi, A., Matsui, M., Miyano, S., 2003. Biopathways representation and simulation on hybrid functional petri net. In Silico Biol 3, 32.

Maturana, H. R., Varela, F. J., 1991. Autopoiesis and Cognition : The Realization of the Living (Boston Studies in the Philosophy of Science). Springer.
URL http://www.amazon.ca/exec/obidos/redirect?tag=citeulike04-20{\&}path=ASIN/9027710163

McAdams, H. H., Arkin, A., 1997. Stochastic mechanisms in gene expression. Proc Natl Acad Sci U S A 94 (3), 814–819.

McAdams, H. H., Arkin, A., 1999. It's a noisy business! Genetic regulation at the nanomolar scale. Trends Genet 15 (2), 65–69.

McAdams, H. H., Shapiro, L., 1995. Circuit simulation of genetic networks. Science 269, 650–656.

REFERENCES

McElroy, M. B., Donahue, T. M., 1972. Stability of the martian atmosphere. Science 177, 986–988.

Monod, J., 1942. Recherches sur la croissance des cultures bactériennes. Ph.D. thesis, Université de Paris.

Murata, T., 1989. Petri nets: Properties, analysis and applications. P IEEE 77 (4), 541–580.

Nagasaki, M., Doi, A., Matsuno, H., Miyano, S., 2004. Genomic object net: I. a platform for modeling and simulating biopathways. Applied Bioinformatics 2, 181–184.

Nair, H., Allen, M., Anbar, A. D., Yung, Y. L., 1994. A photochemical model of the martian atmosphere. Icarus 111, 124–150.

Nair, H., Summers, M. E., Miller, C. E., Yung, Y. L., 2005. Isotopic fractionation of methane in the martian atmosphere. Icarus 175, 32–35.

Palsson, B. O., 2006. Systems Biology. Cambridge University Press.

Papin, J. A., Hunter, T., Palsson, B. O., Subramaniam, S., 2005. Reconstruction of cellular signalling networks and analysis of their properties. Nat Rev Mol Cell Bio 6 (2), 99–111.
URL http://dx.doi.org/10.1038/nrm1570

Papin, J. A., Palsson, B. O., 2004a. The JAK-STAT signaling network in the human B-cell: an extreme signaling pathway analysis. Biophys J 87 (1), 37–46.
URL http://dx.doi.org/10.1529/biophysj.103.029884

Papin, J. A., Palsson, B. O., 2004b. Topological analysis of mass-balanced signaling networks: a framework to obtain network properties including crosstalk. J Theor Biol 227 (2), 283–297.
URL http://dx.doi.org/10.1016/j.jtbi.2003.11.016

Papin, J. A., Stelling, J., Price, N. D., Klamt, S., Schuster, S., Palsson, B. O., 2004. Comparison of network-based pathway analysis methods. Trends Biotechnol 22 (8), 400–405.
URL http://dx.doi.org/10.1016/j.tibtech.2004.06.010

Parkinson, T. M., Hunten, D. M., 1972. Spectroscopy and aeronomy of O_2 on mars. J Atmos Sci 29, 1380–1390.

Petri, C. A., 1962. Kommunikation mit automaten. Ph.D. thesis, TU Darmstadt.

Poolman, M. G., Fell, D. A., Raines, C. A., 2003. Elementary modes analysis of photosynthate metabolism in the chloroplast stroma. Eur J Biochem 270 (3), 430–439.

Postma, P. W., Lengeler, J. W., Jacobson, G. R., 1993. Phosphoenolpyruvate: Carbohydrate phosphotransferase systems of bacteria. Microbiol Rev 57 (3), 543–594.

Potapov, A., Liebich, I., Dönitz, J., Schwarzer, K., Sasse, N., Schoeps, T., Crass, T., Wingender, E., 2006. EndoNet: an information resource about endocrine networks. Nucleic Acids Res 34 (Database issue), D540–D545.
URL http://dx.doi.org/10.1093/nar/gkj121

Ptashne, M., 1986. A Genetic Switch: Gene Control and Phage λ. Blackwell Scientific Publications, Oxford.

Puchalka, J., Kierzek, A. M., 2004. Bridging the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks. Biophys J 86 (3), 1357–1372.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., Barabási, A. L., 2002. Hierarchical organization of modularity in metabolic networks. Science 297 (5586), 1551–1555.
URL http://dx.doi.org/10.1126/science.1073374

Reed, J. L., Vo, T. D., Schilling, C. H., Palsson, B. O., 2003. An expanded genome-scale model of escherichia coli k-12 (ijr904 gsm/gpr). Genome Biol 4 (9), R54.
URL http://dx.doi.org/10.1186/gb-2003-4-9-r54

Sauro, H. M., Hucka, M., Finney, A., Wellock, C., Bolouri, H., Doyle, J., Kitano, H., 2003. Next generation simulation tools: the systems biology workbench and biospice integration. OMICS 7 (4), 355–372.
URL http://dx.doi.org/10.1089/153623103322637670

Schilling, C. H., Letscher, D., Palsson, B. O., 2000. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. J Theor Biol 203 (3), 229–248.
URL http://dx.doi.org/10.1006/jtbi.2000.1073

Schuster, S., Dandekar, T., Fell, D. A., 1999. Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. Trends Biotechnol 17 (2), 53–60.

Schuster, S., Fell, D., Dandekar, T., 2000a. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nat Biotechnol 18, 326–332.

Schuster, S., Hilgetag, C., Woods, J. H., Fell, D. A., 2002a. Reaction routes in biochemical reaction systems: algebraic properties, validated calculation procedure and example from nucleotide metabolism. J Math Biol 45 (2), 153–181.
URL http://dx.doi.org/10.1007/s002850200143

Schuster, S., Kenanov, D., 2005. Adenine and adenosine salvage pathways in erythrocytes and the role of s-adenosylhomocysteine hydrolase. a theoretical study using elementary flux modes. FEBS J 272 (20), 5278–5290.
URL http://dx.doi.org/10.1111/j.1742-4658.2005.04924.x

Schuster, S., Klamt, S., Weckwerth, W., Pfeiffer, T., 2002b. Use of network analysis of metabolic systems in bioengineering. Bioproc Biosyst Eng 24, 363–372.

Schuster, S., Pfeiffer, T., Moldenhauer, F., Koch, I., Dandekar, T., 2000b. Structural analysis of metabolic networks: Elementary flux modes, analogy to petri nets, and application to mycoplasma pneumoniae. In: Bauer, E.-B., Rost, U., Stoye, J., Vingron, M. (Eds.), Proc. Germ. Conf. Bioinf. Heidelberg, pp. 115–120.

Schuster, S., Pfeiffer, T., Moldenhauer, F., Koch, I., Dandekar, T., 2002c. Exploring the pathway structure of metabolism: decomposition into subnetworks and application to Mycoplasma pneumoniae. Bioinformatics 18 (2), 351–361.

Schwender, J., Goffman, F., Ohlrogge, J. B., Shachar-Hill, Y., 2004. Rubisco without the calvin cycle improves the carbon efficiency of developing green seeds. Nature 432 (7018), 779–782.
URL http://dx.doi.org/10.1038/nature03145

Segall, J. E., Manson, M. D., Berg, H. C., 1982. Signal processing times in bacterial chemotaxis. Nature 296 (5860), 855–857.

Serres, M. H., Goswami, S., Riley, M., 2004. Genprotec: an updated and improved analysis of functions of escherichia coli k-12 proteins. Nucleic Acids Res 32 (Database issue), D300–D302.
URL http://dx.doi.org/10.1093/nar/gkh087

Shea, M. A., Ackers, G. K., 1985. The O_R control system of bacteriophage lambda - a physical-chemical model for gene regulation. J Mol Biol 181, 211–230.

Smolen, P., Baxter, D. A., Byrne, J. H., 2000. Modeling transcriptional control in gene networks–methods, recent results, and future directions. B Math Biol 62 (2), 247–292.

Solé, R. V., Munteanu, A., 2004. The large-scale organization of chemical reaction networks in astrophysics. Europhys Lett 68 (2), 170–176.

Sørensen, K. I., Hove-Jensen, B., 1996. Ribose catabolism of Escherichia coli: characterization of the rpiB gene encoding ribose phosphate isomerase B and of the rpiR gene, which is involved in regulation of rpiB expression. J Bacteriol 178 (4), 1003–1011.

Spirin, V., Mirny, L. A., 2003. Protein complexes and functional modules in molecular networks. Proc Natl Acad Sci U S A 100 (21), 12123–12128.
URL http://dx.doi.org/10.1073/pnas.2032324100

Starke, P. H., 1990. Analyse von Petri-Netz-Modellen. B. G. Teubner, Stuttgart.

Stelling, J., 2004. Mathematical models in microbial systems biology. Curr Opin Microbiol 7 (5), 513–518.
URL http://dx.doi.org/10.1016/j.mib.2004.08.004

Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., Gilles, E., 2002. Metabolic network structure determines key aspects of functionality and regulation. Nature 420, 190–193.

Stephanopoulos, G. N., Aristidou, A. A., Nielsen, J., 1998. Metabolic Engineering. Academic Press, San Diego.

Strogatz, S. H., 1994. Nonlinear Dynamics and Chaos. Addison-Wesley.

Strogatz, S. H., 2001. Exploring complex networks. Nature 410 (6825), 268–276.
URL http://dx.doi.org/10.1038/35065725

Teusink, B., Passarge, J., Reijenga, C. A., Esgalhado, E., van der Weijden, C. C., Schepper, M., Walsh, M. C., Bakker, B. M., van Dam, K., Westerhoff, H. V., Snoep, J. L., 2000. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. Eur J Biochem 267 (17), 5313–5329.

Thattai, M., Shraiman, B. I., 2003. Metabolic switching in the sugar phosphotransferase system of *Escherichia coli*. Biophys J 85, 744–754.

Thieffry, D., Thomas, R., 1995. Dynamical behaviour of biological regulatory networks – II. immunity control in bacteriophage lambda. B Math Biol 57 (2), 277–97.

Thomas, R., 1973. Boolean formalization of genetic control circuits. J Theor Biol 42 (3), 563–585.

Tomita, M., 2001. Whole-cell simulation: a grand challenge of the 21st century. Trends Biotechnol 19 (6), 205–210.

Tomita, M., Hashimoto, K., Takahashi, K., Shimizu, T. S., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J. C., Hutchison, C. A., 1999. E-cell: software environment for whole-cell simulation. Bioinformatics 15 (1), 72–84.

Tran, Q. H., Bongaerts, J., Vlad, D., Unden, G., 1997. Requirement for the proton-pumping NADH dehydrogenase I of Escherichia coli in respiration of NADH to fumarate and its bioenergetic implications. Eur J Biochem 244 (1), 155–160.

Tyson, J. J., Chen, K., Novak, B., 2001. Network dynamics and cell physiology. Nat Rev Mol Cell Biol 2 (12), 908–916.
URL http://dx.doi.org/10.1038/35103078

Varela, F. G., Maturana, H. R., Uribe, R., 1974. Autopoiesis: the organization of living systems, its characterization and a model. Curr Mod Biol 5 (4), 187–196.

Varela, F. J., Coutinho, A., 1991. Second generation immune networks. Immunol Today 12 (5), 159–166.

Varma, A., Boesch, B. W., Palsson, B. O., 1993. Biochemical production capabilities of escherichia coli. Biotechnol Bioeng 42, 59–73.

von Bertalanffy, L., 1969. General System Theory. George Braziller, New York.

von Mering, C., Zdobnov, E. M., Tsoka, S., Ciccarelli, F. D., Pereira-Leal, J. B., Ouzounis, C. A., Bork, P., 2003. Genome evolution reveals biochemical networks and functional modules. Proc Natl Acad Sci U S A 100 (26), 15428–15433.
URL http://dx.doi.org/10.1073/pnas.2136809100

Voss, K., Heiner, M., Koch, I., 2003. Steady state analysis of metabolic pathways using petri nets. In Silico Biol 3 (3), 367–387.

Wagner, A., Fell, D. A., 2001. The small world inside large metabolic networks. Proc Biol Sci 268 (1478), 1803–1810.
URL http://dx.doi.org/10.1098/rspb.2001.1711

REFERENCES

Wang, J., Gilles, E. D., Lengeler, J. W., Jahreis, K., 2001. Modeling of inducer exclusion and catabolite repression based on a pts-dependent sucrose and non-pts-dependent glycerol transport systems in *Escherichia coli* k-12 and its experimental verification. J Biotechnol 92, 133–158.

Xenarios, I., Salwnski, L., Duan, X. J., Higney, P., Kim, S.-M., Eisenberg, D., 2002. DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. Nucleic Acids Res 30 (1), 303–305.

Yung, Y. L., DeMore, W. B., 1999. Photochemistry of Planetary Atmospheres. Oxford University Press, Oxford.

Zeigarnik, A., Temkin, O., 1994. A graph-theoretic model of complex reaction mechanisms: bipartite graphs and the stoichiometry of complex reactions. Kinet Catal Engl. Transl. 35, 691–701.

Zevedei-Oancea, I., Schuster, S., 2003. Topological analysis of metabolic networks based on petri net theory. In Silico Biol 3 (3), 323–345.

Ziv, E., Middendorf, M., Wiggins, C. H., 2005. Information-theoretic approach to network modularity. Phys Rev E Stat Nonlin Soft Matter Phys 71 (4 Pt 2), 046117.

About the Author

Scientific Curriculum Vitae

Name: Florian Centler
Date of Birth: November 14, 1975
Place of Birth: Dortmund, Germany
Nationality: German
Email: fcentler@gmx.net

1995-2002	Studies in Applied Systems Science, University of Osnabrück, Germany
2003-2007	Research Associate in Junior Research Group "Bio Systems Analysis", Friedrich-Schiller-University of Jena, Germany
2003-2007	Part-time Studies in Economics, Distance University of Hagen, Germany

Publications

1. F. Centler, P. Dittrich, L. Ku, N. Matsumaru, J. Pfaffmann, and K.-P. Zauner (2003): Artificial Life as an Aid to Astrobiology: Testing Life Seeking Techniques, in W. Banzhaf et al. (Eds.): Advances in Artificial Life, ECAL 2003, LNAI 2801, pp. 31–40, Springer, Berlin
2. N. Matsumaru, F. Centler, K.-P. Zauner, and P. Dittrich (2004): Self-Adaptive Scouting - Autonomous Experimentation for Systems Biology, in G.R. Raidl et al. (Eds.): Applications of Evolutionary Computing, EvoWorkshops 2004, LNCS 3005, pp. 52–62, Springer, Berlin
3. A. Kremling, R. Heermann, F. Centler, K. Jung, and E.D. Gilles (2004): Analysis of two-component signal transduction by mathematical modeling using the KdpD/KdpE system of Escherichia coli, BioSystems 78 (1-3), pp. 23–37
4. N. Matsumaru, F. Centler, P. Speroni di Fenizio, and P. Dittrich (2006): Chemical Organization Theory Applied to Virus Dynamics, it - Information Technology 48 (3), pp. 154–160, Oldenourg Wissenschaftsverlag
5. N. Matsumaru, P. Speroni di Fenizio, F. Centler, and P. Dittrich (2006): On the Evolution of Chemical Organizations, in S. Artmann and P. Dittrich (Eds.): Proceedings of the 7th German Workshop on Artificial Life, pp. 135–146, Aka, Berlin

6. C. Kaleta, F. Centler, and P. Dittrich (2006): Analyzing Molecular Reaction Networks: From Pathways to Chemical Organizations, Mol Biotechnol 34(2), pp. 117–124

7. F. Centler, P. Speroni di Fenizio, N. Matsumaru, and P. Dittrich (2007): Chemical Organizations in the Central Sugar Metabolism of Escherichia coli, in A. Deutsch et al. (Eds.): Mathematical Modeling of Biological Systems, Volume I, pp. 105–119, Birkhäuser, Boston

8. F. Centler and P. Dittrich (2007): Chemical Organizations in Atmospheric Photochemistries - A New Method to Analyze Chemical Reaction Networks, Planet Space Sci 55, pp. 413–428

9. N. Matsumaru, F. Centler, P. Speroni di Fenizio, and P. Dittrich (2007): Chemical Organization Theory as a Theoretical Base for Chemical Computing, Int J Unconv Comp 3(4), pp. 285–309

10. C. Kaleta, F. Centler, P. Speroni di Fenizio, and P. Dittrich (2008): Phenotype prediction in regulated metabolic networks, BMC Systems Biol, 2:37

11. F. Centler, C. Kaleta, P. Speroni di Fenizio, and P. Dittrich (2008): Computing chemical organizations in biological networks, Bioinformatics 24(14), pp. 1611–1618

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Florian Centler
Jena, den 1. Oktober 2008