BRNO UNIVERSITY OF TECHNOLOGY

Faculty of Electrical Engineering and Communication

MASTER'S THESIS

Brno, 2021
Bc. Monika Křápková

BRNO UNIVERSITY OF TECHNOLOGY

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

DEPARTMENT OF BIOMEDICAL ENGINEERING

DYNAMIC MODEL FOR PRODUCTION OF POLYHYDROXYALKANOATES BY THERMOPHILIC BACTERIUM S. THERMODEPOLYMERANS

MASTER'S THESIS
DIPLOMOVÁ PRÁCE

AUTHOR   Bc. Monika Křápková
AUTOR PRÁCE

SUPERVISOR   Mgr. Ing. Karel Sedlář, Ph.D.
VEDOUCÍ PRÁCE

BRNO 2021

BRNO FACULTY OF ELECTRICAL
UNIVERSITY ENGINEERING
OF TECHNOLOGY AND COMMUNICATION

Master's Thesis

Master's study program **Biomedical Engineering and Bioinformatics**

Department of Biomedical Engineering

Student: Bc. Monika Křápková   *ID:* 195191

Year of study: 2   *Academic year:* 2020/21

TITLE OF THESIS:

Dynamic Model for Production of Polyhydroxyalkanoates by Thermophilic Bacterium S. thermodepolymerans

INSTRUCTION:

1) Prepare a literature review of dynamic modeling in systems biology, including commonly used tools for a dynamic analysis of complex systems. 2) Study the possibility of using lab data for an inference of dynamic models. Aim on wet-lab techniques commonly used in metabolomics. 3) Using a suitable tool, e.g. Cell Colective, and the knowledge gathered from literature and public databases, design a basic dynamic model describing production of polyhydroxyalkanoates in thermophilic bacteria. 4) Refine the model using the results of a metabolomic analysis of the strain Schlegelella thermodepolymerans DSM 15344. 5) Perform a static and mainly dynamic analysis of the proposed model. 6) Discuss the results.

RECOMMENDED LITERATURE:

[1] KARASAVVAS, Evgenios a Christos CHATZIDOUKAS. Model-based dynamic optimization of the fermentative production of polyhydroxyalkanoates (PHAs) in fed-batch and sequence of continuously operating bioreactors. Biochemical Engineering Journal. 2020, 162.

[2] KOURILOVA, Xenie, Iva PERNICOVA, Karel SEDLAR, Jana MUSILOVA, Petr SEDLACEK, Michal KALINA, Martin KOLLER a Stanislav OBRUCA. Production of polyhydroxyalkanoates (PHA) by a thermophilic strain of Schlegelella thermodepolymerans from xylose rich substrates. Bioresource Technology. 2020, 315.

Date of project specification: 8.2.2021   *Deadline for submission:* 21.5.2021

Supervisor: Mgr. Ing. Karel Sedlář, Ph.D.

prof. Ing. Ivo Provazník, Ph.D.
Chair of study program board

ABSTRACT

This master's thesis deals with the reconstruction of a dynamic model for production of polyhydroxyalkanoates (PHA) by thermophilic bacterium *Schlegelella thermodepolymerans*. The first chapter provides readers with a brief introduction into the systems biology and mathematical graph theory. It is followed by Chapter Two dealing with different approaches in dynamic modelling, including the commonly used tools for dynamic analysis of complex systems. The third chapter then pursues further terms and possibilities regarding the model analysis. The following chapter focuses on metabolomics and the frequently used laboratory techniques and the fifth chapter is then occupied with polyhydroxyalkanoates, especially their chemical structure and properties. In Chapter Six, a general Boolean model for PHA production by thermophilic bacteria is proposed. Chapter Seven then aims at model refinement with focus on *S. thermodepolymerans*. The final dynamic model is analysed and the results are discussed.

KEYWORDS

Schlegelella, PHA, Dynamic modelling, Metabolomics, Flux balance analysis, Attractor analysis

ABSTRAKT

Tato diplomová práce se zabývá rekonstrukcí dynamického modelu produkce polyhydroxyalkanoátů (PHA) termofilní bakterií *Schlegelella thermodepolymerans*. První kapitola poskytuje čtenářům krátký úvod do systémové biologie a matematické teorie grafů. Na ni navazuje druhá kapitola zabývající se různými přístupy v dynamickém modelování, včetně běžně používaných nástrojů pro dynamickou analýzu komplexních systémů. Třetí kapitola pak sleduje další pojmy a možnosti týkající se analýzy modelu. Následující kapitola se zaměřuje na metabolomiku a často používané laboratorní techniky a pátá kapitola je pak věnována polyhydroxyalkanoátům, zejména jejich chemické struktuře a vlastnostem. V kapitole šesté je navržen obecný booleovský model pro produkci PHA termofilními bakteriemi. Kapitola sedmá se poté zaměřuje na zdokonalení modelu se zaměřením na *S. thermodepolymerans*. Výsledný dynamický model je podroben analýze a výsledky jsou diskutovány.

KLÍČOVÁ SLOVA

Schlegelella, PHA, Dynamické modelování, Metabolomika, Analýza rovnováhy toku, Analýza atraktorů

ROZŠÍŘENÝ ABSTRAKT

Tato diplomová práce se zabývá konstrukcí dynamického modelu pro produkci polyhydroxyalkanoátů (PHA) termofilní bakterií *Schlegelella thermodepolymerans*. PHA jsou polymerní látky vyznačující se podobnými vlastnostmi jako mají běžné plasty. Narozdíl od těchto umělých látek, které se vyrábějí především z ropy, jsou však produkovány bakteriemi, a navíc jsou biodegradabilní [1].

Navzdory těmto výhodám je jejich komercializace dlouhodobě problematická z důvodu vysoké ceny jejich produkce. Motivací za výzkumem termofilních bakterií je jejich základní potřeba vyšší teploty, v důsledku které by mohly být sníženy náklady spojené se sterilizací v průběhu kultivace. Bakterie *S. thermodepolymerans* je navíc unikátní svou vlastností efektivního zpracovávání xylózy, která je levným a dostupným zdrojem uhlíku [2]. Díky těmto skutečnostem představuje *S. thermodepolymerans* slibného PHA producenta mezi ostatními bakteriemi.

První část této práce se věnuje nezbytným teoretickým základům. V první kapitole je představena systémová biologie, jakožto holistická vědní disciplína, s přesahem do základních pojmů a vlastností matematické teorie grafů. Na ni navazuje kapitola druhá, která se zabývá různými přístupy v dynamickém modelování a přináší podrobnější popis třech vybraných metod, včetně běžně používaných nástrojů pro analýzu dynamických modelů. Třetí kapitola pak uvádí rozšiřující pojmy a možnosti týkající se analýzy komplexních systémů.

Kapitola čtvrtá přináší čtenáři úvod do metabolomiky s přehledem nejčastěji používaných laboratorních technik této vědní disciplíny. V rámci následující kapitoly jsou pak představeny polyhydroxyalkanoáty, především z hlediska jejich chemické struktury a vlastností, a v závěru jsou pak uvedeny tři základní metabolické dráhy pro syntézu PHA bakteriemi. Šestá kapitola se konečně věnuje termofilním bakteriím a rekonstrukci obecného dynamického modelu.

Doposud byla syntéza PHA popsána celkem u pěti termofilních bakterií, těmi jsou *Thermus thermophilus* [59, 60], *Caldimonas taiwanensis* [61], *Chelatococcus thermostellatus* [62], *Aneurinibacillus sp.* [51] a *Schlegelella thermodepolymerans* [2]. Ze srovnání v tabulce 6.1 vyplývá, že čtyři z pěti termofilů byli schopni produkce poly(3-hydroxybutyrátu), neboli P(3HB), z glycerolu a nějakého typu cukru. Na základě těchto znalostí bylo odvozeno základní schéma modelu.

Dynamický model byl vytvořen pomocí webového nástroje Cell Collective [21], vychází z Booleovské logiky a jeho chování odráží základní prvky reálných experimentů. Simulace uvedené na obrázku 6.2 demonstrují vliv teploty na celkový růst buněk a akumulaci P(3HB), a degradaci polymerů poté, co je většina výchozího substrátu zkonzumována. Nevhodná teplota při kultivaci obvykle vyústí v nižší růst i syntézu PHA. Degradace je pak dána tím, že buňky začnou nahromaděné zásoby PHA využívat pro vlastní potřebu, neboť jim slouží jako zásoba uhlíku a energie.

Poslední, sedmá kapitola, se věnuje úpravě a zdokonalení modelu se zaměřením na *S. thermodepolymerans*. První část se zabývá především zkoumáním jejích metabolických drah na základě anotovaného genomu. Získané informace jsou doplněny teoretickými znalostmi a případně srovnáním s modelovým organismem *E. coli*. Druhá část je pak věnována přímo úpravě dynamického modelu, jeho analýze a diskuzi výsledků.

Jak už bylo naznačeno, *S. thermodepolymerans* je unikátní tím, že velmi efektivně zpracovává xylózu [2]. Při konzumaci tohoto substrátu vykazuje větší růst a vyšší produkci P(3HB) než při spotřebě glukózy, a dokonce preferuje xylózu i v případě dostupnosti obou cukrů současně; teprve až je většina xylózy spotřebována, začne konzumovat glukózu. Toto sekvenční zpracovávání cukrů, známé jako diauxické chování, je u bakterií obvyklé, nicméně většina z nich dává přednost glukóze. V dynamickém modelu jsou zohledněny právě tyto dva výchozí monosacharidy.

Na základě analýzy genomu bylo zjištěno, že geny kódující proteiny pro transport i degradaci xylózy jsou součástí jednoho operonu. To znamená, že jsou regulovány společně, a tudíž ihned po transportu je xylóza degradována. Výsledná xylulóza 5-fosfát dále vstupuje do neoxidativní části pentózového cyklu, který se pak napojuje na glykolýzu. Transportní mechanismus glukózy zatím nebyl přesně stanoven, nicméně po vstupu do cytoplazmy je metabolizována glykolýzou.

Po fosforylaci může glukóza také vstoupit do oxidativní části pentózového cyklu. Zajímavostí je, že v genomu bakterie *S. thermodepolymerans* pravděpodobně chybí dva významné geny kódující část této metabolické dráhy. Její význam spočívá především v produkci redukovaného kofaktoru NADPH, který se uplatňuje při syntetických reakcích, a ribulózy 5-fosfát, která dále vstupuje do neoxidativní části. Jestliže oxidativní část v metabolismu *Schlegelelly* opravdu chybí, musí také existovat nějaké kompenzační mechanismy pro dodatečnou tvorbu kofaktoru, ribulóza 5-fosfát pak může být získána zpětně z neoxidativní části.

Výše uvedené poznatky byly zohledněny při následných úpravách modelu; oxidativní část pentózového cyklu byla vynechána. Dále byla v modelu zahrnuta produkce metabolických prekurzorů tak, aby byly kladeny stejné nároky na jejich produkci při konzumaci xylózy i glukózy. Vyšší efektivita xylózy byla nakonec zajištěna pomocí uměle vytvořené komponenty, která zpomaluje degradaci glukózy.

Ke stanovení směru toků reakcí v neoxidativní části pentózového cyklu byla použita analýza rovnováhy toků. Její výsledky také poukázaly na to, že glukóza má pravděpodobně vyšší metabolickou kapacitu ve srovnání s xylózou. Nicméně vytvořený model nebyl natolik komplexní, aby mohly být výsledky brány s jistotou.

V rámci dynamické analýzy byly simulovány podmínky reálných experimentů, které model dokáže napodobit. Opět byl demonstrován vliv teploty, dále pak rychlejší růst a produkce P(3HB) při konzumaci xylózy, a také specifické diauxické chování.

Výsledky simulace jsou uvedeny na obrázcích 7.3 a 7.4. Přestože model dobře napodobuje provedené experimenty, příčiny unikátního chování bakterie jsou stále předmětem diskuze. Hlavním faktorem přispívajícím k preferenci a vyšší efektivitě xylózy je patrně dříve zmíněná společná regulace genů pro její transport i degradaci.

Dalším zajímavým faktem je, že v genomu bakterie nebyl nalezen žádný specifický TPS transporter pro přenos glukózy. Ten by mohl rovněž přispět k absenci typické preference tohoto monosacharidu, neboť má také globální regulační funkci – v případě dostupnosti preferovaného cukru nejsou exprimovány geny pro zpracování jiných cukrů. Dále by mohla hrát roli i nepřítomnost oxidativní části pentózového cyklu, v případě dodatečné tvorby ribulózy 5-fosfát je totiž dráha k jejímu vytvoření kratší při zpracování xylózy.

Poslední část sedmé kapitoly se pak zabývá dlouhodobým chováním vytvořeného modelu. Ten musel být částečně zjednodušen, neboť diauxické chování, stejně jako degradace polymerů, byly zajištěny manuální úpravou externích komponent v průběhu simulace, a to zde nebylo možné. Výsledky analýzy demonstrovaly stabilní stav produkce P(3HB) při nepřetržitém dodávání živin a cukrů. Taková situace odpovídá kontinuální bakteriální kultivaci v chemostatu.

Závěrem lze shrnout, že *Schlegelella thermodepolymerans* je díky svým vlastnostem nadějným producentem polyhydroxyalkanoátů, a rozhodně by tak měla být předmětem dalšího výzkumu. Tato práce by tak mohla být podkladem pro její další studium.

DECLARATION

I declare that I have written the Master's Thesis titled "Dynamic Model for Production of Polyhydroxyalkanoates by Thermophilic Bacterium S. thermodepolymerans" independently, under the guidance of the advisor and using exclusively the technical references and other sources of information cited in the thesis and listed in the comprehensive bibliography at the end of the thesis.

As the author I furthermore declare that, with respect to the creation of this Master's Thesis, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll., Section 2, Head VI, Part 4.

Brno   . . . .

author's signature

ACKNOWLEDGEMENT

Contents

List of Figures

Introduction

This master's thesis focuses on production of polyhydroxyalkanoates (PHA) by thermophilic bacterium *Schlegelella thermodepolymerans*. PHA are a family of biodegradable polymers produced by many bacterial species. The history of their discovery dates back to mid-twenties and since then, they have become the subject of interest of many researchers [1]. The motivation behind the extensive studies concerning PHA is that they represent a "green" alternative to petrochemical plastics.

Despite the great advantage of biodegradability, the commercialisation of PHA is still out of sight since production costs are five to ten times higher in comparison with standard petrochemicals [1]. The initiative to study PHA synthesis in thermophilic bacteria is based on the assumption of reducing the costs associated with sterilisation, as the bacterial cultivations undergo higher temperatures. Furthermore, in case of thermophilic bacterium *S. thermodepolymerans*, the costs could be even more decreased since the species is able of PHA synthesis from xylose which is an inexpensive source of carbon [2]. The aim of this master's thesis, thus, is to explore the underlying biochemical mechanisms and reconstruct the dynamic model for PHA production by *S. thermodepolymerans*.

At the beginning, theoretical foundations need to be given. The first chapter of the thesis introduces systems biology as a holistic scientific discipline. Since all biological networks are based on mathematical graph theory, the basic graph terminology, representations and properties are briefly described. The subsequent chapter then focuses on different approaches in dynamic modelling and three regular modelling techniques are introduced; these are ordinary differential equations (ODEs), Boolean networks and Petri nets. In addition, three commonly used tools are mentioned at each of the methods.

The next chapter pursues further terms and possibilities related to model analysis. Basic concepts of abstract phase space, attractors or bifurcation are clarified, and the constraint-based modelling is introduced with the flux balance analysis (FBA) as the typical representative. The fourth chapter then focuses on regular approaches used in metabolomics and brings a brief comparison of the most frequently used methods; these are NMR spectroscopy and mass spectrometry coupled with different separation techniques.

The aim of the Chapter Five is to briefly wander into the history of the discovery of polyhydroxyalkanoates and introduce them in terms of chemical structure and properties. In addition, three natural pathways for PHA synthesis are described. The following chapter then deals with the actual production of PHA by thermophilic bacteria. To date, the PHA synthesis has been described just in five thermophiles. Based on the gathered knowledge, a general dynamic model is created.

The model follows the Boolean logic and reflects the common experiments concerning PHA production by thermophilic bacteria. The model's behaviour and the effect of temperature are demonstrated via simulation in time.

The Chapter Seven is finally devoted to PHA production by *S. thermodepolymerans*. In the first part, the underlying biochemical mechanisms are explored. The investigation is based on the *Schlegelella*'s annotated genome and literature and focuses mainly on transport mechanisms and sugar degradation pathways.

The second part is then concerned with the refinement of the general dynamic model. Xylose and glucose are determined as the input carbon sources and the flux balance analysis is used to help with the final arrangement of the model. The model's dynamics then follows the diauxic behaviour and the higher efficiency of xylose utilisation observed in *S. thermodepolymerans*. In addition, the long-term behaviour of the model is examined using an analysis of attractors. The results from individual sections are continuously discussed and finally summarised in conclusions.

1 Biological Networks

As reductionist researchers investigate constituent parts of an organism in detail, systems biology focuses on the individual as a complex system. This opposite approach is referred to as holistic [3]. In contrast to scientific reductionism, holism adheres to the statement that a system is not only a sum of its individual parts. In other words, the behaviour of biological systems cannot be explained only by properties of their individual parts.

The unpredictable behaviour arising from complex systems is called an emergent property. It emerges from the system as a whole and so it cannot be determined by reductionist methods. Systems biology in cooperation with many other scientific disciplines investigate biological processes by creating and analysing complex interaction networks. To build such a network, a given biological problem need to be expressed in terms of mathematical graph theory.

1.1 Biological Process as a Mathematical Graph

Before describing biological networks of different scales, one should be aware of basic graph theory concepts.

A graph $G = (V, E)$ is an ordered pair of the set of vertices $V = \{v_1, v_2, ...v_m\}$ and the set of edges $E = \{e_1, e_2, ...v_n\}$ [4]. A subgraph $G' = (V', E')$ of the graph G is then defined by a subset V' of the set V and a subset E' of the set E.

A graph can be directed or undirected. In a directed graph, the edges are oriented and direct from one vertex to another. Expressed mathematically, the graph would be defined by an ordered triple $G = (V, E, f)$, where f is a function mapping the edges from E to ordered pairs of vertices in V. An undirected graph does not provide any information about the edge direction, i.e. edges have no orientation. The connection between vertices i and j is defined as $E = \{\{i, j\}|i, j \in V\}$ and the vertices in question are called direct neighbours. An edge starting and ending in the same vertex is called a loop.

If there are more edges between any two vertices (meaning two unoriented edges or two edges of the same direction), the graph is called multi-edge or multigraph. In case the graph does not contain any multiple edges or loops, it is called simple. A complete graph or clique then contains one unique edge between any pair of vertices. If the number (density) of edges is low, the graph is called sparse. Another term is a weighted graph, where edges are associated with a weight function, mathematically expressed as $w : E \rightarrow R$, where R is the set of real numbers.
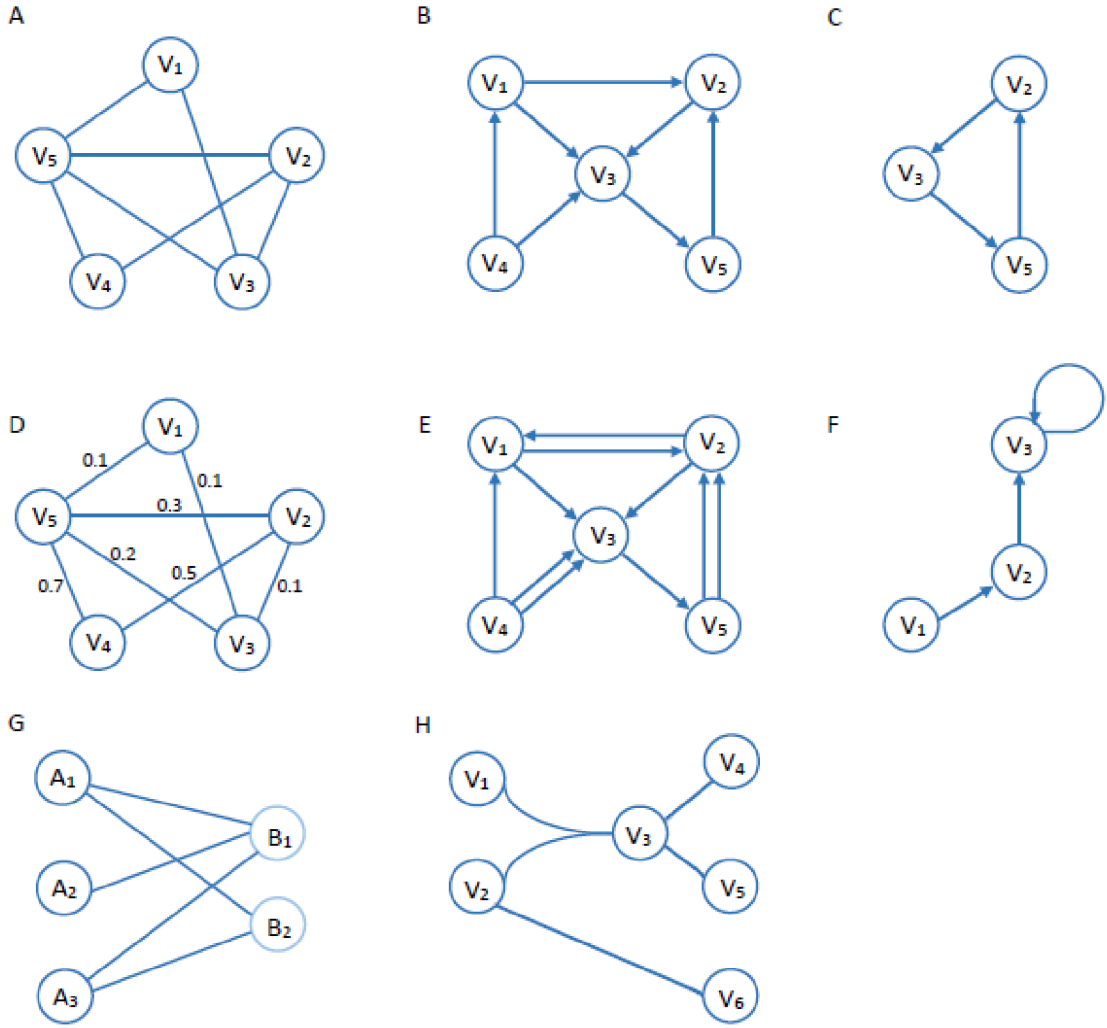
Fig. 1.1: Various types of graphs.

A) A simple undirected graph $G = (V, E)$ with the set of vertices $V = \{V_1, V_2, V_3, V_4, V_5\}$ and the set of edges $E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7\} = \{\{V_1, V_3\}, \{V_1, V_5\}, \{V_2, V_3\}, \{V_2, V_4\}, \{V_2, V_5\}, \{V_3, V_5\}, \{V_4, V_5\}\}$.

B) A simple directed graph $G = (V, E)$ with the vertices $V = \{V_1, V_2, V_3, V_4, V_5\}$ and the edges $E = \{(V_1, V_2), (V_1, V_3), (V_2, V_3), (V_3, V_5), (V_4, V_1), (V_4, V_3), (V_5, V_2)\}$.

C) A subgraph $G' = (V', E')$ of the graph G from **B)**.

D) A simple weighted graph $G = (V, E)$, where $E = \{\{V_1, V_3, 0.1\}, \{V_1, V_5, 0.1\}, \{V_2, V_3, 0.1\}, \{V_2, V_4, 0.5\}, \{V_2, V_5, 0.3\}, \{V_3, V_5, 0.2\}, \{V_4, V_5, 0.7\}\}$.

E) A multigraph $G = (V, E)$ with two double edges (V_4, V_3) and (V_5, V_2). Note that the links between vertices V_1 and V_2 are not equally oriented, i.e. these are two simple edges (V_1, V_2), (V_2, V_1) and do not make the graph multi-edge.

F) A simple oriented graph $G = (V, E)$ with loop (V_3, V_3).

G) A bipartite graph $G = (V, E)$, where V is composed of the sets A and B. Note that there are no links between the two vertices from the same set.

H) A hypergraph $G = (V, E)$, where $E = \{\{V_1, V_2, V_3\}, \{V_2, V_6\}, \{V_3, V_4\}, \{V_3, V_5\}\}$.

A bipartite graph $G = (V, E)$ is an undirected graph, where the set of vertices V is divided into two separate sets, A and B, and only the edges connecting these two sets are allowed. The last type of graph to mention is a hypergraph. The typical property of hypergraph is that the edges can connect any number of vertices. Certainly, there are other various types of graphs, but it is not the aim of this chapter introducing them all. Some of the mentioned above are depicted in Fig. 1.1.

A biological network, arising from the graph theory as a mathematical abstraction, describes the relations between biological entities interacting at different structural levels. The vertices, often called nodes, usually represent genes, proteins, enzymes or metabolites, and the edges then gene regulations, protein interactions, enzyme reactions etc [4]. Generally, biological networks are simple oriented or unoriented sparse graphs, their edges are often weighted and regularly create loops. Commonly, they also contain statistically significant subgraphs called motifs.

1.2   Basic Graph Representation

Probably the most intuitive way of representing graphs is a drawing. There is a variety of layouts to be used and the final form always depends on the specific application [4]. The point is that different layouts can highlight different network properties, so it is always important to choose an appropriate style. Vertices of graph are usually depicted as circles or points and the edges as their links. When modelling biological processes, oriented edges can be drawn as positive $\rightarrow$ (performing activation) or negative $\dashv$ (performing inhibition) [5].

The more practical way of representing a graph (especially for further data analysis) is data structures. The common are an adjacency matrix, an adjacency list, or a sparse matrix [4].

An adjacency matrix A of a graph $G = (V, E)$ is a square matrix of dimension $N \times N$, where $N = |V|$ [6]. The matrix is defined as

$$A_{ij} = \begin{cases} 1 & (i, j) \in E \\ 0 & (i, j) \notin E. \end{cases} \tag{1.1}$$

For undirected graphs, the matrix is symmetric – otherwise it is not. In case of simple graph, the diagonal contains only zeros. The problem of adjacency matrices is that they require $O(|V|^2)$ memory according to the O notation [1] [4]. Thus, this data structure brings difficulties into storing larger networks and becomes useless especially for storing sparse networks.

---

[1]The big O notation is a theoretical calculation that refers to the running time or memory needed for a given algorithm in connection to the growing size of input data. [4]

An alternative form to the matrix can be the adjacency list [4]. The adjacency list A is an array of vertices, where each vertex is accompanied by a list of all adjacent vertices. This data structure requires $O(V+E)$ memory, which is noticeably less memory-consuming compared to the adjacency matrix. Similarly to the lists, an efficient solution can be a sparse matrix storing only the non-zero values and their coordinates. A drawing, the adjacency matrix and list are shown in the Fig. 1.2.
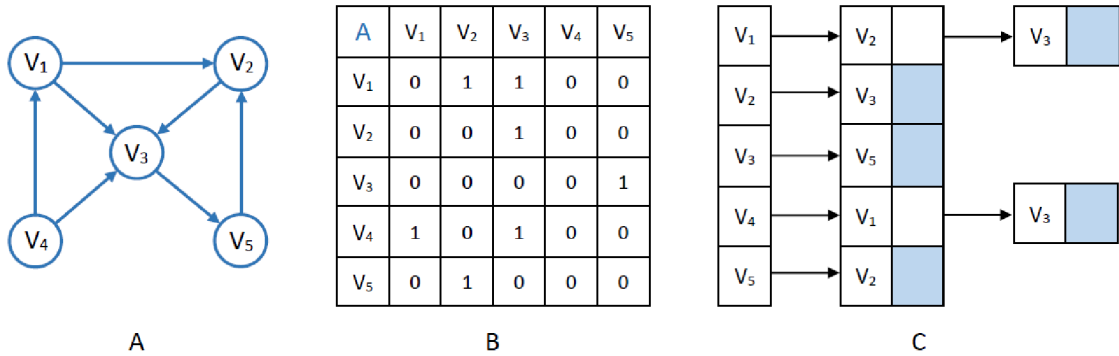
Fig. 1.2: Various representations of a simple directed graph G. **A)** A drawing. **B)** The adjacency matrix. **C)** The adjacency list.

1.3   Network Properties

Equally important part to introduce is formed by the network properties. The various types of graphs and their representations were already mentioned, now some of the basic network characteristics will be described.

A few definitions from the start. The size of the graph G is the number of all its edges [7]. The vertices linked to the vertex i, i.e. the neighbours or adjacent vertices, form a subgraph called the neighbourhood N_i of the vertex i. A degree deg_i of the vertex i is the size of its neighbourhood N_i (with the exception that loops containing i are counted twice) [4]. In case of directed graphs, incoming and out-coming edges can be distinguished as in-degree deg_i^{in} and out-degree deg_i^{out} (giving the deg_i in the sum).

A probability, that a randomly selected vertex from the graph G has a degree k, is described by the degree distribution $P(k)$ [4]. Another significant characteristic of a graph is the clustering coefficient. It can be determined for a single node or for the whole graph and it provides the information about the tendency of the node or the graph to form cluster(s).

A walk in a graph is a sequence of vertices, where there is an edge between any of two consecutive vertices [7]. In the walk, both edges and nodes can be freely repeated. The length of a walk is then the number of edges it uses. More specific terms are a trail or a path.

The trail is a walk, where no edge can be used twice. Conversely, the path is a walk, where no vertex can be used twice. The closed trail including all the edges of a graph is called Eulerian and the path going through all the vertices in a graph is called Hamiltonian.

A cycle (circuit) is a closed trail, it starts and ends in the same vertex (representing the only vertex repetition) [7]. Based on this statement, the definition of loop would be a cycle of length one. A graph containing cycles is called cyclic, otherwise it is acyclic. The graphs depicted in Fig. 1.1 are all cyclic except for the hypergraph in **H)**.

2 From Structure to Dynamics

The issue discussed in previous chapter lays the foundations for static analysis of the model. However, to discover the dynamics of the system, the relations between interacting components need to be further investigated and defined precisely.

Concerning the methods used in dynamic modelling, there are two different approaches to consider [8]. The first one, deterministic, works with the substance concentrations and disregards any molecule fluctuations. The second one, stochastic, conversely weighs every single molecule and works with the small particle numbers. The apparent difference between these two approaches is, thus, the level of approximation and model complexity.

The matter of deterministic methods is that they disregard the heterogeneity as one of the main properties of all biological systems [9]. The slightly different development of two separate individuals is caused by genetic, extrinsic (environmental) and intrinsic (stochastic) heterogeneities. The third one, intrinsic, arises from random thermal fluctuations at molecular level and thus, manifests itself especially when considering lower molecule numbers. In such a case, deterministic methods fail to cope with the intrinsic noise and hence, the stochastic are required.

A great field for comparison of these two approaches is performing a simulation [9]. Considering deterministic algorithm and unchanging parameter settings, the result after several simulation runs will always be the same. In case of stochastic simulation, the result after each run is slightly different. This fact enables stochastic algorithms to reveal the full probability distribution, whereas the deterministic are able to define only one point of it. On the other hand, stochastic algorithms are much slower to solve. Thus, the advantage of deterministic methods is their speed.

Despite some imperfections of deterministic methods, they have a broad range of use and generally provide adequate results. The most common are ordinary differential equations (ODEs), Boolean networks or Petri nets [8]. The most favourite stochastic methods are then chemical master equation or rule-based formalisms [8]. Interestingly, all the given deterministic methods can be modified for stochastic modelling as well, only the involvement of a stochastic component is required. A specific area is then rule-based models which can be simulated either stochastically or deterministically using an appropriate algorithm.

This thesis focuses on metabolic networks and works with experimental data of higher volumes. Hence, the mentioned deterministic methods will be next introduced. In addition, three relevant tools will be proposed for each of the methods since there are many.

2.1 Ordinary Differential Equations

The most common deterministic method are ordinary differential equations (ODEs). It enables continuous modelling in time by solving the equations either analytically, or numerically. When applying ODEs to biological systems, the two assumptions must be considered: firstly, the contents of biological compartment is ideally mixed and the concentrations are high (so that there is no time delay caused by the internal transport of species); and secondly, the transport between two different compartments is slow and observable [10]. Otherwise, the spatial information must be included by using the partial differential equations (PDEs). In such a case, solving equations becomes much harder since the functions are multivariable.

Generally, a dynamic system can be described as a system of state variables and development equations. The variables are represented by a state vector $\boldsymbol{x_t}$, where each of the elements is determined in time t by the function $F(\boldsymbol{x_t})$ [11]. Concerning biological models using ODEs, the state vector is comprised of concentrations of particular species, usually given in units of mols per volume. The change of concentration in time is then modelled as a function of the reaction fluxes and thus, it depends on chemical kinetics of the reactions and concentrations of the other species. Hence, forming the equations requires precise knowledge of the modelled reactions and the underlying mechanisms.

In the real world, any reaction comprising more than two reactants occurs as a sequence of simple interactions involving only the two of the reacting entities at the time. These single-step reactions between two chemical species are called elementary reactions and form the basis of every ODE model [12]. The following forms can be considered:

$$\emptyset \xrightarrow{k} products, \tag{2.1}$$

$$reactants \xrightarrow{k} products, \tag{2.2}$$

$$reactants \xrightarrow{k} \emptyset, \tag{2.3}$$

where $\emptyset$ represents all entities out of the model's ambit and k represents the rate constant as it describes the speed of the reaction.

The Eq. (2.1) denotes the inflow of entities into the system. This reaction does not depend on concentrations of other species and only the rate constant k is engaged. Hence, it is called zero-order reaction.

The reaction described by Eq. (2.2) represents the transformation of two reactants at most into any number of products. In case of one unique reactant, it is first-order reaction and describes the change of state of an entity.

An example can be a movement of species from one compartment to another, e.g. from extracellular to intracellular space, or a simple conversion, e.g. the protein autophosphorylation. Staying with the second example, the transformation can be depicted as $A \xrightarrow{k} A^p$. Such a reaction must be described by two equations as there are two different species participated (protein A and phosphorylated protein A^p):

$$\frac{d[A]}{dt} = -k \cdot [A], \tag{2.4}$$

$$\frac{d[A^p]}{dt} = k \cdot [A], \tag{2.5}$$

where $[A]$, $[A^p]$ are the concentrations of proteins A and A^p [12]. The Eq. (2.4), denoting the change of concentration of protein A in time, is tied with the minus sign as the concentration of protein A is decreasing.

The reaction with one unique reactant is called unimolecular [13]. In case of two reactants, it is called bimolecular and it is the reaction of second order since the reaction order is associated with molecularity. An elementary second-order reaction can be depicted as $A + B \xrightarrow{k} AB$ and is described by following equations:

$$\frac{d[A]}{dt} = -k \cdot [A] \cdot [B], \tag{2.6}$$

$$\frac{d[B]}{dt} = -k \cdot [A] \cdot [B], \tag{2.7}$$

$$\frac{d[AB]}{dt} = k \cdot [A] \cdot [B]. \tag{2.8}$$

The reaction flux here is controlled by the law of mass action and the reaction is often depicted in a reversible form as $A + B \xleftrightarrow{k_1, k_2} AB$. The mass action law determines that the overall reaction rate is proportional to the product of concentrations of the reacting species [12].

The final form in Eq. (2.3) denotes the degradation reaction or the outflow of an entity from the modelled system [12]. In case of one reactant involved, it is the first-order reaction and the rate constant k is tied with the negative sign.

Aiming at metabolic networks, enzyme kinetics and Michaelis-Menten equation need to be mentioned. Consider enzymatic reaction:

$$E + S \xleftrightarrow{k_1, k_2} ES \xrightarrow{k_3} E + P, \tag{2.9}$$

where E is an enzyme, S is a substrate, SE is the substrate-enzyme complex and P is the product [13]. This reaction can be modelled according to the mass action law by the set of four equations as there are four species involved.

However, Michaelis and Menten came with the simplified form based on the mass action kinetics, but only with substrate and product included. The simplified form of enzymatic reaction is depicted as

$$S \xrightarrow{v_{max}, K_m} P, \qquad (2.10)$$

where P, S represent product and substrate, v_{max} is the maximum reaction rate and K_m is the Michaelis constant [14]. For this simplified form, the following equation was derived

$$v = \frac{d[P]}{dt} = \frac{v_{max} \cdot [S]}{[S] + K_m}, \qquad (2.11)$$

where v is the reaction rate and $[P]$, $[S]$ are the concentrations of product and substrate [14]. This equation is called Michaelis-Menten equation and the two constants are defined as

$$v_{max} = k_3 \cdot E_{tot}, \;\; K_m = \frac{k_2 + k_3}{k_1}, \qquad (2.12)$$

where E_{tot} represents the total amount of enzyme in the reaction and k_1, k_2, k_3 are the rate constants depicted in (2.9) [14]. The Michaelis constant represents the concentration of substrate in case of half-maximum reaction rate.

Thus, the enzymatic reaction according to the Michaelis and Menten is described by the following two equations

$$\frac{d[S]}{dt} = -\frac{v_{max} \cdot [S]}{[S] + K_m}, \qquad (2.13)$$

$$\frac{d[P]}{dt} = \frac{v_{max} \cdot [S]}{[S] + K_m}. \qquad (2.14)$$

Besides the mass action and Michaelis-Menten kinetics, there exist other kinetic laws for enzymatic reactions, e.g. Hill or Goldbeter-Koshland kinetics [14].

Commonly Used Tools

CellDesigner [15]. A pathway editor for construction and simulation of gene-regulatory and biochemical networks. Besides the creation, it enables user to import models in SBML format or from online databases such as BioModels [16]. The simulation is based on SBML ODE Solver and is performed through a graphical user interface.

COPASI (COmplex PAthway SImulator) [17]. A tool supporting deterministic, stochastic, or even hybrid simulation of ODE-based models. The user is allowed to construct the model via specifying the reactions and relevant mathematical equations directly, or via selecting the pre-defined rate laws which results in automatically defined mathematical model. Importing a model in SBML format is supported as well.

DMPy [18]. A Python package proposing an automated pipeline for translation of static networks into the dynamic mathematical models. The input static network is parsed and the kinetic rates of all reactions involved are being searched for in online databases. The tool focuses especially on large-scale metabolic networks.

2.2   Boolean Networks

Another method for dynamic modelling is a Boolean network [19]. A great advantage here is that there is no need of defining any rate constants or forming the set of differential equations. Hence, in comparison with quantitative ODE models, establishing the Boolean model does not require such a detailed knowledge of the given problem. The character of Boolean models is rather qualitative.

This simplification makes Boolean networks relatively easy to use. However, one needs to learn how to express the real situation using only ones and zeroes, since the state variables here can only reach the two states – true or false. In case the state variable is ON (true), it is considered to be activated or expressed, or it signifies that the concentration of the relevant species is above the given threshold. In case the state is OFF (false), the species is inhibited or not expressed, or the concentration is under the given threshold. The Boolean model with the state vector $\mathbf{x_t}$ of the length n can reach 2^n possible states in total [19].

The individual entities of Boolean model are represented by the nodes and the interactions between them by the links of a directed graph [19]. The future state x_i^* of the node i is given by the Boolean transition function f_i which involves the current states of the regulating species and logic operators AND, OR and NOT [20]. An example can be the function $f_5 = (x_2 \, AND \, x_3) \, OR \, (NOT \, x_4)$ denoting the future state of the variable x_5. The verbal interpretation would be that entity represented by variable x_5 (node 5) is ON only if x_2 and x_3 are both ON or if x_4 is OFF. In this example, the variables x_2, x_3, x_4 are the regulators of x_5.

All the possible future states of relevant variables can be expressed by the truth table [20]. Considering the function f_i with k variables, the truth table would have 2^k rows (all the variations of the regulators' states) and $k + 1$ columns. The Tab.2.1 shows the truth table of three functions, f_C, f_D and f_E, denoting the possible future states of variables x_C, x_D and x_E. Hence, the table has $k + 3$ columns.

In contrast to the continuous ODE modelling, simulation of Boolean networks flows in discrete time, i.e. the network is updated at each discrete time t by applying the Boolean functions to the state variables [20]. Considering deterministic models, the network update is synchronous. It means that at each time step, all transition functions are applied. This paradigm should demonstrate that all components require the same time for transition to the future state.

However, there exist also asynchronous or probabilistic updating, both of stochastic character. In asynchronous, only one randomly selected variable is updated per step. In probabilistic, state variables can be associated with more functions of different probability. At each step, only one of the functions for each variable is applied.

Tab. 2.1: The truth table of three functions, f_C, f_D and f_E, denoting all the possible future states of variables x_C, x_D and x_E. The regulators here are x_A and x_B [20].

x_A	x_B	$f_C = NOT\ x_A$	$f_D = x_A\ OR\ x_B$	$f_E = x_A\ AND\ x_B$
0	0	1	0	0
0	1	1	1	0
1	0	0	1	0
1	1	0	1	1

Commonly Used Tools

Cell Collective [21]. A web-based platform focused on construction and analysis of Boolean networks. In contrast to other common tools, it enables to share the created models with other users and so gather the acquired complex knowledge at one place. Besides the research, Cell Collective is supposed to assist in teaching and learning of fundamental biological processes.

BoolNet [22]. An R package supporting construction and analysis of synchronous, asynchronous, and probabilistic Boolean networks. It further enables analysis and visualisation of attractors, including visualisation of state transitions and basins of attraction.

BooleanNet [23]. A Python source code for modelling rather smaller biological sub-systems based on Boolean logic. The user input consists in simple text-based model description. Besides the synchronous and asynchronous simulation, it enables hybrid modelling via piecewise linear formalisms and ODEs.

2.3 Petri Nets

The two significant methods were introduced, now the third one is going to be presented. In the previous subchapter, the difference between the quantitative and qualitative model was briefly mentioned, considering ODE models and Boolean networks as typical quantitative and qualitative representatives. As for the Petri nets, both types of models can be defined [24].

The Petri net model is based on bipartite directed and weighted graph [24]. The two sets of nodes are composed of places and transitions and the arrows, connecting these sets, denote the arcs of the net. Every place is allowed to contain a number of marks called tokens. The weights of edges then represent the number of tokens required for enabling the transitions. The actual distribution of tokens in the net is called marking and defines the actual state of the system. The change of state is related to the flow of tokens across the network which is called firing.

A simple Petri net model describing water formation is depicted in Fig. 2.1 [25]. The reaction follows the equation $2\,H_2 + O_2 \rightarrow 2\,H_2O$. The marking on the left denotes the situation before firing the reaction, the transition t is enabled since the places H_2 and O_2 contain adequate number of tokens. On the right, there is a situation after firing the reaction. The number of tokens differs here as they were consumed by the firing.
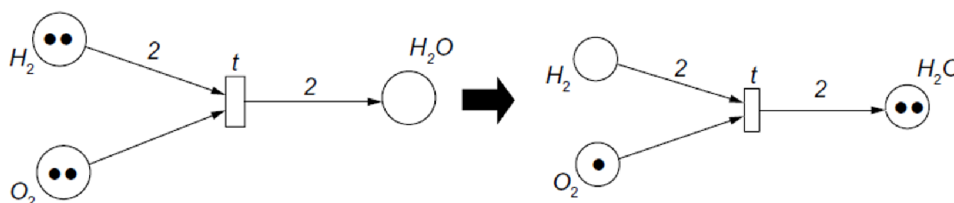
Fig. 2.1: A simple Petri net based model describing water formation [25]. The marking on the left denotes the situation before firing the reaction, the marking on the right refers to the situation after firing the reaction.

In relation with metabolic networks, the places of Petri net represent metabolites and enzymes and the transitions represent metabolic reactions. Arc weights correspond to the stoichiometric coefficients and the occurrence of tokens in places to the occurrence of relevant chemical species.

Based on the initial mathematical concept, various extended versions of Petri nets were developed. In a very simple way, Petri nets can be classified as low-level or high-level [26]. In low-level systems, places are represented either by Boolean or integer tokens which means they can be marked by one at most or by a number of tokens. In high-level systems, tokens are carrying an additional information.

Commonly Used Tools

Snoopy [27]. A software tool which focuses on designing, animation and simulation of Petri nets. It enables the time-free or time-dependent execution dealing with qualitative, stochastic, continuous or hybrid Petri nets.

Cell Illustrator [28]. A software tool supporting the construction and simulation of biological systems based on the Hybrid Functional Petri nets with Extensions. With an intuitive user-friendly interface and no need of higher mathematical skills, Cell Illustrator is a suitable tool for a wide scientific community.

GreatSPN [29]. A well-established tool supporting coloured Generalised Stochastic Petri Nets (GSPN), GSPN with deterministic and/or general transitions and Markov Decision Petri nets (MDWNs). It went through many changes from its beginning and currently proposes many tools for qualitative and quantitative model analysis.

3   Besides the Simulation

In previous chapter, different modelling techniques were introduced in order to explore the dynamics of the system. The two approaches, deterministic and stochastic, were clarified as well as the qualitative and quantitative essence of the model. Using any of these strategies enables biologists to analyse the model's behaviour via the simulation in time.

In this chapter, further terms attending the dynamic analysis, such as attractors or bifurcation, will be introduced. Moreover, a particular approach laying between the purely static and dynamic modelling will be presented – a brief introduction into constraint-based modelling.

3.1   Phase Space, Attractors and Bifurcation

Before introducing attractors or bifurcation, the term phase space needs to be explained. The phase space is an abstract space whose dimensions are comprised of state variables [30]. The system development in time is then represented as a sequence of points (in case of discrete time) or a continuous curve (in case of continuous time) through this space. When considering a deterministic system and given initial conditions, the system will always perform the same, i.e. it will always copy the same trajectory in the phase space.

An important property of complex systems is that they are able to resist adequate parameter perturbations – they are said to be robust [6]. It means the dynamic development of the system from two different initial points can still end up in the same state. The final behaviour into which these points converge is referred to as attractor and the initial distinct trajectories as transient [31]. All the initial states converging to the same attractor form the basin of attraction. A dynamic system can have several attractors whose basins of attraction do not overlap.

There are three types of attractors to be distinguished – a fixed point, a limit cycle, and strange attractors [31]. A fixed point is an attractor represented by a steady point in a phase space. Since its location (i.e. parameters) remains stable, it denotes the steady state of the system. Limit cycles represent attractors defined by a periodical repetition of the set of states. The number of recurrent states then determine the period or length of the attractor [32]. Finally, strange attractors are those that evince a chaotic behaviour. As the trajectories of two similar initial points are close during the whole development at fixed points and limit cycles, at strange attractors, two similar initial points have diverse trajectories [31].

Staying with the parameter perturbations, the qualitative changes in the structure of solutions due to a tiny parameter shift are called a bifurcation [33]. In practice it means that exceeding the critical value of a control parameter results in significantly distinct development of the system. Some fixed points may arise and some may vanish or become unstable, to illustrate. The point in the phase space giving rise such a change in the system's behaviour is called a bifurcation point.

3.2   Constraint-based Modelling

According to Raman and Chandra [34], constraint-based modelling fits somewhere between the interaction-based static models, represented by mathematical graphs, and mechanistic dynamic models, defined by stoichiometric equations and kinetic parameters. Thus, it is based on the stoichiometry, but no kinetic parameters are involved.

Once the metabolic network is reconstructed, the information about stoichiometry can be stored in the form of stoichiometric matrix [35]. The rows of the matrix correspond to the metabolites and the columns to the reactions involved. The matrix elements then agree with stoichiometric coefficients of relevant species and reactions. In addition, all reactants in the matrix are tied with the negative sign and thus, all products are represented by positive values. In case the species is not present in the relevant reaction, its value is zero.

S	v_1	v_2	v_3	v_4	b_1	b_2
A	-1	-1	1	0	1	0
B	1	0	0	1	0	-1
C	0	1	-1	-1	0	0
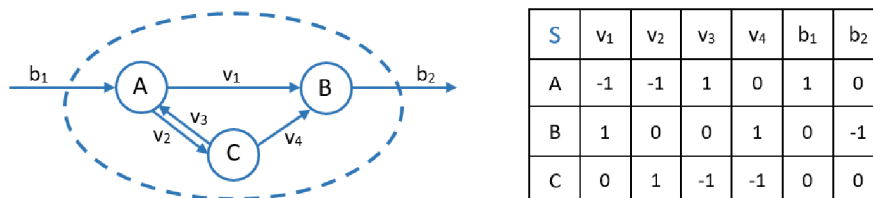
Fig. 3.1: A drawing of a very simple metabolic network and its representation by stoichiometric matrix S. The network consists of three metabolites A, B and C, internal reactions v_1-v_4 and exchange fluxes b_1 and b_2 (the ellipse denotes the boundary). In the matrix, reactants are represented by negative, products by positive values.

Further constraints, as they are needed to form the solution space, derive from thermodynamics and enzyme capacity. The typical representatives of constraint-based modelling are metabolic flux analysis or flux balance analysis [36].

Flux Balance Analysis

Flux balance analysis (FBA) is a predictive constraint-based method used for determining the steady-state flux distribution in metabolic networks [34]. From mathematical point of view, it is a linear optimisation of an objective function subject to given constraints. The idea of biologists is to discover the metabolic capabilities of the system with respect to the given biochemical goal, e.g. the maximisation of growth or biomass production [37].

Consider the example of a simple metabolic network given in Fig. 3.1. The steady state of this network can be described by the following set of dynamic mass balance equations:

$$\frac{d[A]}{dt} = -v_1 - v_2 + v_3 + b_1 = 0, \tag{3.1}$$

$$\frac{d[B]}{dt} = v_1 + v_4 - b_2 \qquad\quad = 0, \tag{3.2}$$

$$\frac{d[C]}{dt} = v_2 - v_3 - v_4 \qquad\quad = 0. \tag{3.3}$$

The same can be expressed using the stoichiometric matrix S and the vector of fluxes $\boldsymbol{v}$:

$$\begin{bmatrix} \frac{d[A]}{dt} \\ \frac{d[B]}{dt} \\ \frac{d[C]}{dt} \end{bmatrix} = S \cdot \boldsymbol{v} = \begin{bmatrix} -1 & -1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & -1 & -1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ b_1 \\ b_2 \end{bmatrix} = 0. \tag{3.4}$$

It follows that the steady-state mass balance of any metabolic network can be defined as

$$\frac{d[\boldsymbol{x}]}{dt} = S \cdot \boldsymbol{v} = 0, \tag{3.5}$$

where $\frac{d[\boldsymbol{x}]}{dt}$ represents the time derivatives of metabolite concentrations, S is the stoichiometric matrix and $\boldsymbol{v}$ is the desired vector of fluxes in which v_1-v_m are the internal and b_1-b_n the exchange fluxes [34].

The expression in Eq. (3.5) represents the equality constraints referred to as balances [38]. In practice, further constraints must be given in order to restrict the solution space. In the real-world metabolic networks, there are commonly many more reactions than species which makes the system under-determined. The additional constraints are referred to as bounds [38] and usually are determined as the lower and upper boundaries of the reaction fluxes.

Once the constraints are established, the crucial step is to form an objective function Z reflecting the requirements for the system. Mathematically, the objective function is defined as

$$Z = \boldsymbol{c}^T \cdot \boldsymbol{v}, \tag{3.6}$$

where $\boldsymbol{c}$ is the vector of weights defining how much the reactions contribute to the objective function [37]. In fact, it is nothing but a linear combination of the vector of fluxes $\boldsymbol{v}$. In case the goal would be to maximise only one reaction, $\boldsymbol{c}$ would be a vector of zeroes with one at the position of the desired reaction flux.

Finally, the whole analysis becomes a linear optimisation problem defined as

$$\max_{v} \ \boldsymbol{c}^T \cdot \boldsymbol{v} \qquad subject \ to \ S \cdot \boldsymbol{v} = 0. \tag{3.7}$$

FBA is a widespread method with many uses [34]. Since it does not require the exact values of kinetic parameters and its principal is basically simple, the computation is feasible even for large-scale metabolic networks. Typically, it can be used for discovering the metabolic capabilities when using different substrates or tightening the constraints in order to simulate diverse preconditions. Furthermore, FBA has many extensions such as regulatory FBA (rFBA), integrated or integrated dynamic FBA (iFBA or idFBA), to name a few [34].

4 Experimental Methods

Systems biology pursues the interactions of chemical species at different functional levels of the organism. At the beginning of each biochemical model are theoretical knowledge and experimental results gained by the means of different "omics" sciences. The four major are genomics, transcriptomics, proteomics and metabolomics focusing on genome, transcriptome, proteome or metabolome determination [39]. Since this thesis aims at metabolic networks, metabolomics will be further discussed.

Metabolites are low molecular weight species produced by metabolic reactions [39]. The term metabolome then refers to a complete set of metabolites contained in a cell, tissue, organ, or the whole organism. Metabolome is closely related to the phenotype as it reflects the perturbations at lower functional levels. There are two main approaches when concerning the metabolomic analysis – targeted and untargeted [40]. The aim of the untargeted analysis is to discover the various metabolites present in the metabolome. Targeted analysis then focuses on quantitative analysis of the specific metabolites. It is based on the prior knowledge and has rather hypothesis-testing character based on the previous untargeted survey.

Furthermore, three regular terms are being used – metabolic fingerprinting, footprinting and metabolic profiling [39]. Metabolic fingerprinting aims at comparing the patterns ("fingerprints") reflecting the changes within intracellular metabolites. This approach searches for the changes in metabolome caused by the exposure of the system to different conditions. Metabolic footprinting then focuses on monitoring the extracellular environment – the metabolites produced or (not) accepted by the cells. Thus, it brings the information about the exo-metabolome. Finally, metabolic profiling aims at predefined group of metabolites in order to identify and quantify the metabolites related to the specific metabolic pathway.

There are two most commonly used techniques in metabolomics – nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) [41]. The latter is usually coupled with gas or liquid chromatography (GC-MS and LC-MS), choosing between the GC and LC is then conditioned by the sample to be analysed. GC is a suitable tool for small and volatile molecules, whereas LC can be used for wider range of metabolites.

Alternatively, capillary electrophoresis (CE) can represent the separation technique for MS. However, it struggles with a low repeatability and with balancing the temperature changes [41]. An alternative spectroscopic technique is then Fourier-transform infrared spectroscopy (FTIR). In the following sections, the most frequently used metabolomic techniques will be introduced.

4.1   NMR Spectroscopy

NMR spectroscopy represents a powerful and commonly used tool for the determination of the chemical composition of the given sample [42]. Generally, NMR aims at certain type of nuclei, typically ^{1}H, ^{13}C or ^{31}P, having the optimal magnetic properties for NMR experiments. The sample is placed into the strong magnetic field and exposed to electromagnetic pulses. Subsequently, the energy absorbed by the nuclei is emitted back and detected. The basic principle, which enables the method to distinguish between nuclei adjacent to different atoms, is that the resonance frequency of nuclei differs according to the atom surroundings.

Various NMR spectroscopic techniques were developed including one, two or three-dimensional or solid-state NMR [42]. The advantage of NMR spectroscopy is that it requires only minimal sample preparation and enables to identify various compounds simultaneously. Nevertheless, the number of detected species is limited by the spectral resolution [43]. The technique is non-invasive and non-destructive and enables quantitative analysis as well. In addition, the accurate chemical structure and spatial decomposition of individual molecules can be revealed [44]. The whole analysis is fast and well reproducible.

4.2   Mass Spectrometry

Mass spectrometry is another important technique dealing with the identification and quantification of molecules [45]. The whole analysis has three consecutive parts. At first, molecules are vaporised and ionised so that they form gas-phase ions. Next, ions are separated according to the mass to charge ratio (m/z) in the mass analyser and finally, different m/z ratios are detected. Data from the analyser and detector are combined to form the resulting mass spectrum.

The system development covers the application of various ionisation techniques and different types of mass analysers [43]. In comparison with NMR, the MS device is less expensive and the requirements for the installation are not so demanding. However, the technique is surely less reproducible. Unlike NMR, a great advantage of MS is the high sensitivity – even nanomolar concentrations of species can be detected. In general, MS efficiency depends on the proper ionisation which is affected by the sample composition [44]. Hence, MS is usually coupled with different separation techniques. Once used, the sample cannot be recovered as MS is a destructive tool.

4.3 Chromatography-Mass Spectrometry

Chromatography is a well-established technique of analytical chemistry used to separate, identify and purify the individual components of the mixture [46]. The method utilises two phases for the separation – the stationary and the mobile one. The constituents of the mixture are reacting with the stationary phase while being carried through the system by the mobile phase. Based on the different molecule properties, such as their size and shape or total charge, the individual components pass through the system with different velocity as they are reacting with the stationary phase differently.

The mobile phase can be either liquid or gas which results in the basic division into liquid chromatography (LC) and gas chromatography (GC). In CG, the mobile phase is comprised of an inert gas such as He or N_2. Hence, the sample needs to be vaporised which entails the requirement for the volatility of the compounds. In metabolomics, GC is used when coupled with MS. At first, the individual constituents of the given sample are separated using GC and subsequently detected by the highly sensitive MS. Combining these two techniques results in the accurate determination of the metabolites present in the sample. The problem of using GC alone is that some species can have very similar properties so that they pass through the system simultaneously – that makes the species indistinguishable [47].

Another commonly used technique in metabolomics is LC-MS combining high-pressure liquid chromatography (HPLC) and MS. In a standard column chromatography, passing the mobile phase through the system is ensured by gravity. However, in HPLC, the mobile phase is pushed through the column under 10–400 atmospheric pressure and thus, smaller particles can be used to form the stationary phase [46]. In result, high separation efficiency can be obtained in a short time.

In comparison with GC-MS systems, LC-MS does not require any demanding sample preparation and it can be used for a wide range of metabolites [45]. The point of GC-MS is that it focuses on small volatile molecules and thus, compounds with no such properties need to be derivatised (if possible).

5 Biosynthesis of Polyhydroxyalkanoates

The discovery of polyhydroxyalkanoates (PHA) dates back to 1920s [48]. In 1926, bacteriologist Maurice Lemoigne first isolated and characterised poly(3-hydroxybuty-rate) [P(3HB)] from *Bacillus megaterium* [49]. In the following years, the granules of P(3HB) were identified in variety of bacteria as a reserve material. However, no one was further engaged in more detailed description of their function until the late 1950s and P(3HB) rediscovery [50].

Since then, many further studies were concerned with the PHA biosynthesis and hand in hand with the general scientific progress they formed the great knowledge we have now. Besides better comprehension of P(3HB) function, further PHA were discovered in various bacteria (even in plants), the synthesis was investigated under different cultivation conditions and its basic metabolic pathways were revealed. The PHA production was analysed in genetically modified bacterial strains and found in recombinant strain of the model *E. coli*, some PHA started being produced industrially and many other goals were achieved [1].

Nowadays, it is known that bacteria synthetise PHA due to a kind of metabolic pressure or stress, such as an excess of carbon sources or a lack of vital nutrients (e.g. oxygen, phosphorus or nitrogen) [1]. PHA stored in the form of granules (see Fig. 5.1) are then supposed to provide cells with carbon and energy reserves. Anyway, the issue which makes the topic of polyhydroxyalkanoates so popular is that, in terms of chemical structure and properties, they represent a great biodegradable substitute to synthetic plastics. The chemical essence of PHA is the point of the following subchapter.
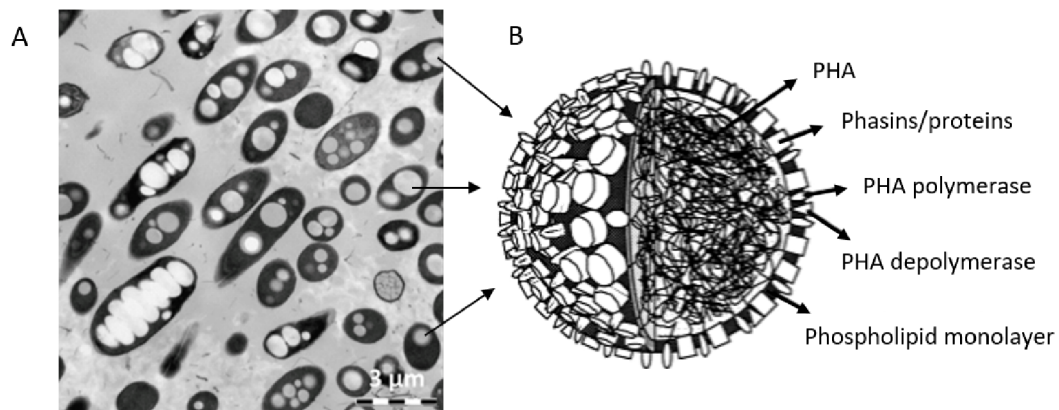

Fig. 5.1: The accumulation of PHA granules in bacteria and their basic structure. **A)** TEM image of PHA granule formation in *Aneurinibacillus sp.* H1 [51]. **B)** A general scheme of PHA granule [1]. PHA are encased in phospholipid monolayer containing various enzymes and structural proteins (phasins).

5.1 PHA in Terms of Chemistry

Polyhydroxyalkanoates are linear polyesters composed of hydroxyalkanoic acid monomers [52]. The polyester synthesis is based on the polycondensation reaction which generally includes the species with two functional groups at least and small molecules as side products. In case of polyesters, the ester bond arises via reaction between carboxyl and hydroxyl groups of reacting monomers accompanied by water as the side product [53]. The general structure of PHA is depicted in Fig. 5.2.
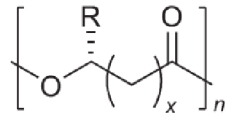

Fig. 5.2: The general structure of PHA, n determines the degree of polymerisation (usually 100-1000) and x the hydroxyl group position [54]. Most commonly, x is equal to one which means the hydroxyl group is localised at the third carbon atom.

To be exact, the Fig. 5.2 shows the general structure of PHA homopolymer. The typical example of such a polymer is P(3HB) which is composed only from 3HB monomers. The polymers containing different types of monomers are called copolymers (or heteropolymers) and the typical example here is poly(3-hydroxy-butyrate-co-3-hydroxyvalerate) [P(3HB-co-3HV)] depicted in Fig. 5.3.
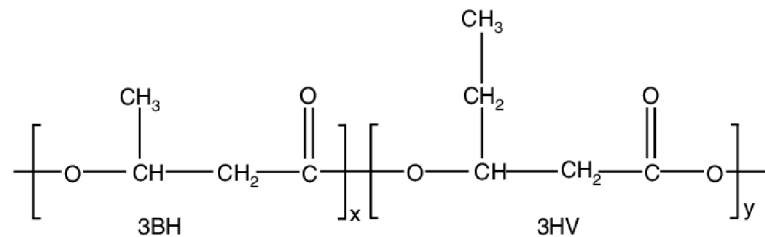
Fig. 5.3: The chemical structure of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) [P(3HB-co-3HV)], the typical PHA copolymer composed of 3HB (3-hydroxybutyrate) and 3HV (3-hydroxyvalerate) monomers [55].

According to the carbon number of monomers present in biopolymer, PHA can be divided into two major groups – short chain length (scl-PHA) and medium chain length (mcl-PHA) [56]. Monomeric units of scl-PHA contain from three to five carbon atoms and the ones of mcl-PHA from six to fourteen carbon atoms. The various carbon number provides the biopolymers with different properties. Scl-PHA are usually crystalline, brittle and have high melting points, while mcl-PHA are less crystalline, elastomeric, and their melting point is lower in comparison with scl-PHA. Typical PHA monomers of different chain lengths are depicted in Fig. 5.4.

The polymerisation reaction in bacteria is catalysed by enzymes called PHA synthases [56]. Based on the bacterium species, enzyme specificity and carbon sources, various types of biopolymers can be produced. As a result, PHA with diverse properties are obtained since PHA structure determines the substance properties. To date, over 150 different monomers were recognised [54].
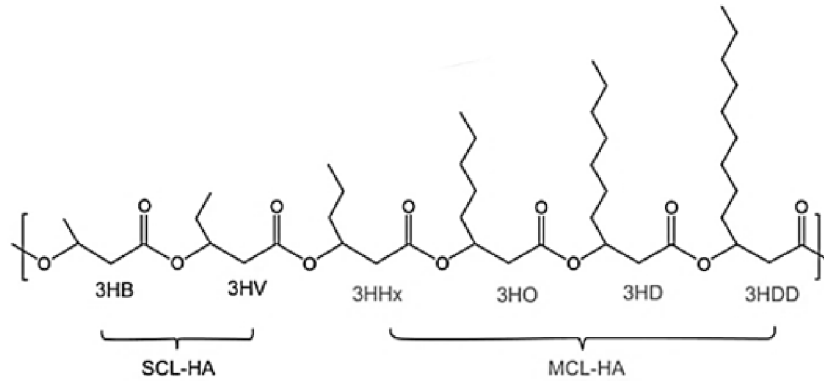

Fig. 5.4: An example of common scl-PHA and mcl-PHA monomers [54]. 3HB: 3-hydroxybutyrate, 3HV: 3-hydroxyvalerate, 3HHx: 3-hydroxyhexanoate, 3HO: 3-hydroxyoctanoate, 3HD: 3-hydroxydecanoate, 3HDD: 3-hydroxydodecanoate.

5.2   Three Fundamental Pathways

In general, studies of PHA synthesis mechanism in bacteria revealed three major natural pathways [57] (see Fig. 5.5). The pathway I represents the basic scl-PHA synthesis. At first, two molecules of acetyl-CoA are produced from sugar, amino acids, or fatty acids as initial carbon sources. Next, with the help of enzyme β-ketothiolase, they are condensed into acetoacetyl-CoA which is further reduced to 3-hydroxybutyryl-CoA by enzyme acetoacetyl-CoA reductase. The final polymerisation is then catalysed by PHA synthase.

The pathway II is related to fatty acids utilisation and leads to mcl-PHA synthesis. Fatty acids are catabolised via β-oxidation cycle. One of the cycle's products is acetyl-CoA which can be further used in pathway I or III. However, pathway II utilises the intermediate cycle product, R-3-hydroxyacyl-CoA, which is the substrate for PHA synthase catalysing mcl-PHA synthesis.

The pathway III uses one of the intermediate products of *in situ* fatty acid synthesis from acetyl-CoA. As in the previous case, this pathway leads to mcl-PHA synthesis. It is conditioned by the presence of enzyme 3-hydroxyacyl-acyl carrier protein-CoA transferase which catalyses the transformation of R-3-hydroxyacyl-ACP, obtained from fatty acid synthesis cycle, to R-3-hydroxy-acyl-CoA.
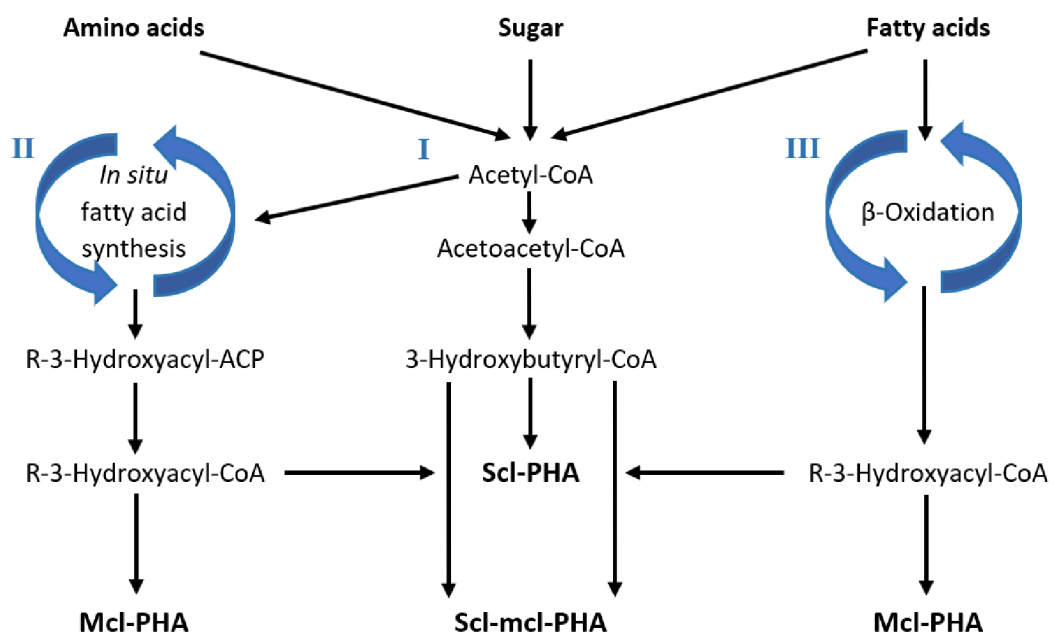
Fig. 5.5: Three natural metabolic pathways for PHA synthesis in bacteria. The pathway I is the most common, it leads to the short chain length PHA (scl-PHA) synthesis from sugars, amino acids, or fatty acids. The other two pathways result in medium chain length PHA (mcl-PHA) formation. When combined with pathway I, scl-mcl-PHA can be produced.

6 Production by Thermophilic Bacteria

The major factor impeding the commercial production of PHA are the five to ten times higher costs compared to petrochemical plastics [1]. This difficulty accompanies the topic of PHA from the very beginning and many studies took various steps in order to deal with this problem. Nevertheless, none of them revealed to be efficient enough. The issue of thermophilic bacteria is that their cultivation undergoes higher temperatures which eliminate the risk of contamination by ubiquitous microorganisms [2]. Hence, the costs related to sterilisation are reduced.

Thermophilic bacteria are a particular type of bacteria whose optimal growth temperature is 45 °C and more [58]. To date, PHA production was described at five thermophilic bacteria, namely *Thermus thermophilus* [59, 60], *Caldimonas taiwanensis* [61], *Chelatococcus thermostellatus* [62], *Aneurinibacillus sp.* [51] and *Schlegelella thermodepolymerans* [2]. The Tab. 6.1 summarises the basic results from studies describing PHA synthesis from various carbon sources. In the table, only significant carbon sources were mentioned. Further, cell dry mass (CDM) as the biomass content, percentage of PHA per CDM and total amount of PHA were given, if available.

Besides the PHA production from various substrates, the effect of temperature was commonly investigated in attendant experiments. Hence, the optimal temperatures were also given for comparison. Note that the given optimal temperatures do not have to correspond to the temperatures used for the cultivations on various substrates.

All bacteria, except for *T. thermophilus*, accumulated PHA in the form of homopolymer P(3HB). PHA produced by *T. thermophilus* were all copolymers composed of diverse monomer units. This species also differ in the optimal growth temperature which is about 70 °C. Such a temperature is typical for extreme thermophiles [58].

According to the Tab. 6.1 summarising the PHA production by thermophilic bacteria, four out of five thermophiles were able to use some type of sugar and glycerol to produce scl-PHA, namely P(3HB). This knowledge leads to the assumption that the mechanism of PHA synthesis in thermophilic bacteria mostly agrees with pathway I. Hence, the basic dynamic model describing PHA production by thermophilic bacteria will be based on this pathway with a sugar at the input.

The sugar catabolism in bacteria has already been well described [63]. In fact, different bacteria are able to utilise different types of monosaccharides, but the basic principle usually remains the same. The part, where the individual cases differ, is the sugar degradation into glyceraldehyde 3-phosphate (GAP). From this point, GAP is further transformed via the trunk pathway into pyruvate which is then converted into acetyl-CoA.

Tab. 6.1: The PHA production by thermophilic bacteria from various carbon sources and the optimal growth temperatures.

Thermus thermophilus

Substrate	CDM [g/l]	PHA [% p CDM]	PHA [g/l]	T_{opt} [°C]	Ref.
Sodium gluconate	-	35.00	-		[59]
Sodium octanoate	-	40.00	-	70	[59]
Whey supernatant*	1.60 ± 0.05	35.60	0.57 ± 0.07		[60]

Caldimonas taiwanensis

Substrate	CDM [g/l]	PHA [% p CDM]	PHA [g/l]	T_{opt} [°C]	Ref.
Sodium gluconate	-	70.00	-		
Fructose	-	62.00	-	55	[61]
Maltose	-	60.00	-		
Glycerol	-	52.00	-		

Chelatococcus thermostellatus

Substrate	Strain	CDM [g/l]	PHA [% p CDM]	PHA [g/l]	T_{opt} [°C]	Ref.
Glucose	MW9	3.04 ± 0.22	64.47	1.96 ± 0.20		
	MW10	3.34 ± 0.14	65.57	2.19 ± 0.05		
	MW13	2.52 ± 0.16	66.27	1.67 ± 0.20	50	[62]
	MW14	3.14 ± 0.26	64.01	2.01 ± 0.10		
Glycerol	MW11	2.04 ± 0.09	42.16	0.86 ± 0.20		
	MW12	3.18 ± 0.31	61.64	1.96 ± 0.40		

Aneurinibacillus sp. H1

Substrate	CDM [g/l]	PHA [% p CDM]	PHA [g/l]	T_{opt} [°C]	Ref.
Glucose	2.00 ± 0.36	27.74 ± 0.18	0.55 ± 0.10	45	[51]
Glycerol	2.19 ± 0.07	45.95 ± 0.10	1.00 ± 0.04		

Schlegelella thermodepolymerans

Substrate	CDM [g/l]	PHA [% p CDM]	PHA [g/l]	T_{opt} [°C]	Ref.
Glucose	3.34 ± 0.01	37.15 ± 0.96	1.24 ± 0.03		
Fructose	3.40 ± 0.02	29.36 ± 0.55	1.00 ± 0.02		
Lactose	3.27 ± 0.06	40.41 ± 5.75	1.32 ± 0.19	55	[2]
Xylose	5.35 ± 0.04	53.20 ± 0.11	2.85 ± 0.01		
Glycerol	3.08 ± 0.03	54.17 ± 0.70	1.67 ± 0.02		

*A more complex substrate containing lactose.

As for the glycerol uptake, the situation is similar as in the case of different monosaccharides. Glycerol degradation leads to dihydroxyacetone phosphate which is next converted into GAP [64]. Then, it is metabolised via the trunk pathway as in the case of sugars.

6.1 Dynamic Model Reconstruction

The general model was reconstructed by using the web-based modelling platform Cell Collective [21]. Using this tool enables to create and analyse qualitative dynamic Boolean models. The model reconstruction consists of defining the qualitative information by the user which is next automatically transformed into the Boolean expressions. Formed expressions are encoded to C++ files and these are together compiled into the dynamic library which is finally executable by the simulation engine, ChemChains.

The development of individual components in time is expressed by the activity level measured as % ON. In principle, the user is allowed to choose the length of a sliding window and the level activity of a component then corresponds to the fraction of active/inactive states over the chosen window. All simulations given in this thesis will be synchronous and performed with a sliding window of the length 1000.

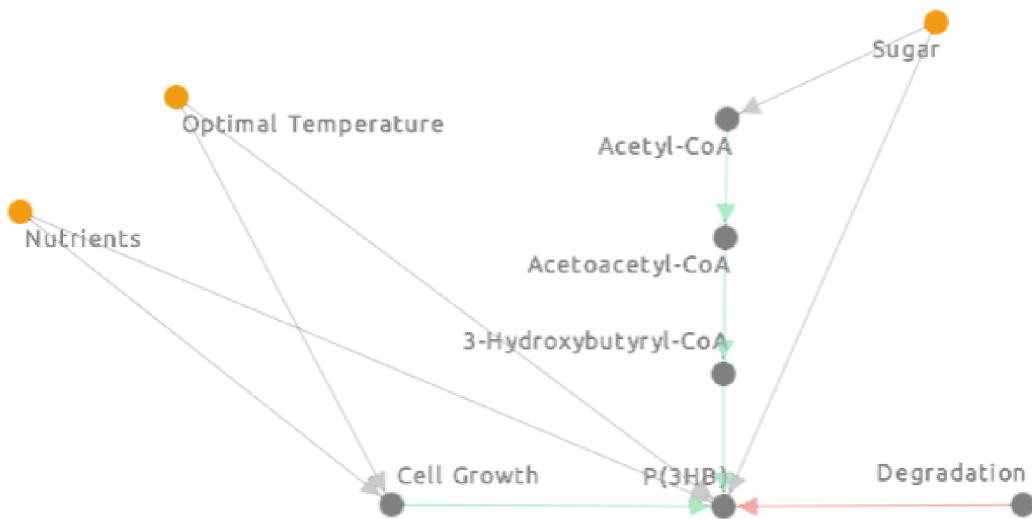The final scheme of the created general model is depicted in Fig. 6.1.

Fig. 6.1: The general model of PHA production by thermophilic bacteria. The backbone of the model is formed by the acetyl-CoA to 3-hydroxybutyryl-CoA pathway (pathway I) resulting in P(3HB) production.

There are three external components (highlighted in orange) – Nutrients, Optimal Temperature and Sugar. Cell Collective allows users to set the activity of external components in the range from zero to 100 %. This extention enables to apply finer changes in the initial conditions and thus, simulate more diverse situations. The green links then denote positive and the red links negative regulations. The grey ones represent conditioned relations.

The Nutrients component in the model represents standard cultivation medium needed for the cell growth – its lower activity results in lower Cell Growth activity. Optimal Temperature specifies the optimal temperature during the cultivation – reduced component activity denotes that the temperature is too low/high which again results in lower Cell Growth activity. The third component, Sugar, represents the initial sugar concentration in the cultivation medium.

The metabolic pathway describing the sugar degradation into acetyl-CoA was omitted since it does not play any key role here. The sugar could also be replaced with glycerol, however, sugar is the conventional input component when considering the pathway I. Another name for pathway I is acetyl-CoA to 3-hydroxybutyryl-CoA pathway and here, it creates the backbone of the model.

The experimental results from studies of PHA production by thermophilic bacteria refer to one important factor – the time of cultivation. Usually, at the beginning of cultivation, a dose of carbon source is applied into the cultivation medium. In response to such an excess of carbon source, bacteria start to consume it in order to create reserves in the form of PHA granules. Thus, the CDM and so the PHA content increase.

However, due to such abundant consumption, the concentration of carbon source in medium gradually decreases until it achieves a sort of limit value. At this point, bacteria stop consuming the given carbon source and start to degrade their own reserves. In the model, this behaviour is assured by the Degradation component which is dependent on the Sugar activity. The Degradation is 100 % active from the very beginning, however, it becomes a negative regulator of PHA activity only when the Sugar is inactive. Such a situation can be simulated by manual Sugar activity decrease after a while simulation.

At the top of the Fig. 6.2, there is a simulation plot denoting the situation with ideal growth conditions (Nutrients, Sugar and Optimal Temperature at 100 %). In such a case, cells are growing and accumulating PHA granules at maximum. In performed experiments, after some time, the concentration of sugar greatly decreased, and bacteria started to degrade the accumulated reserves. This is simulated by decreasing the Sugar activity to zero which enables the Degradation component to negatively regulate P(3HB) activity.

In contrast, the simulation plot at the bottom denotes the situation with unsuitable growth temperature. The Optimal Temperature activity here is at 60 %, other conditions are the same as in the first case. Due to the inconvenient temperature, the accumulation of P(3HB) is not so efficient, and bacteria do not grow so fast. Temperature plays a crucial role when estimating the metabolic capabilities of thermophilic bacteria. Usually, the effect of temperature is investigated only after the initial experiment of PHA production from different carbon sources.
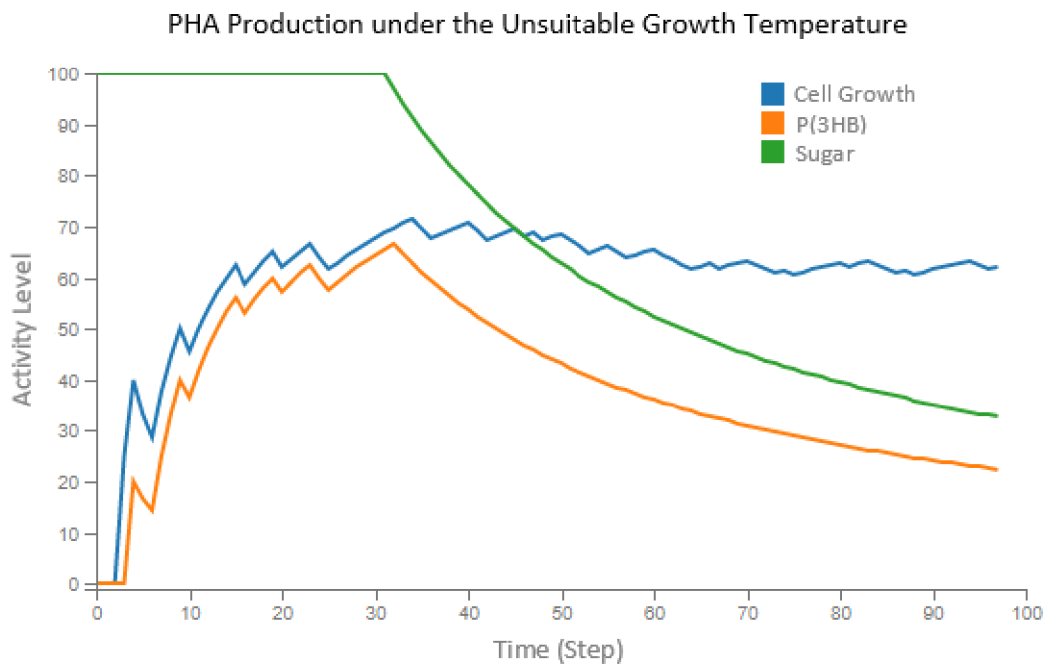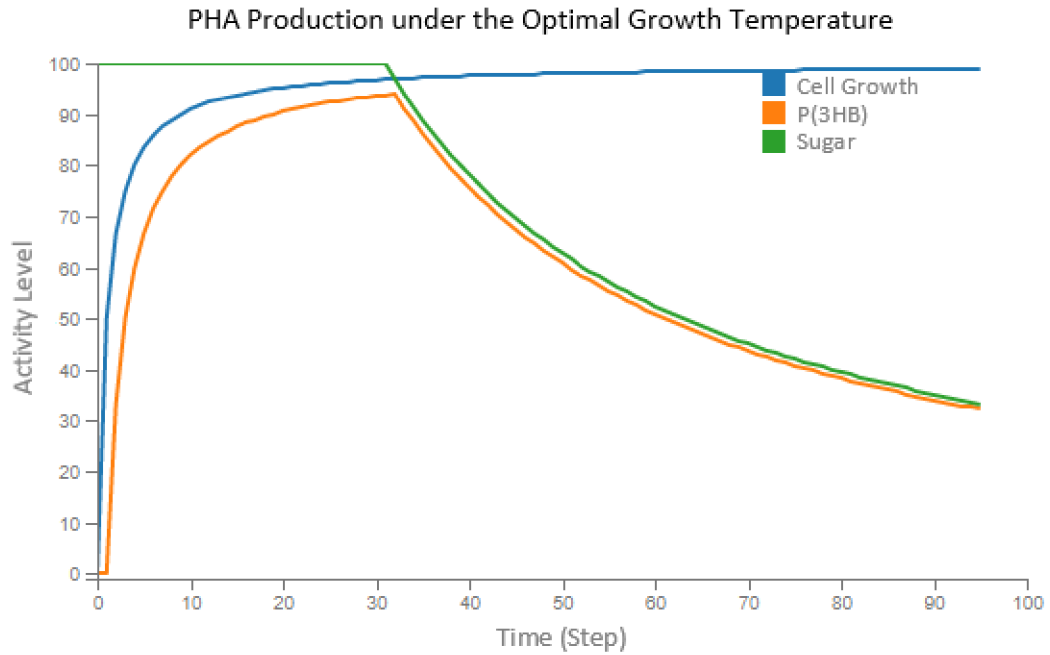
Fig. 6.2: Simulation of PHA production by thermophilic bacteria under various temperatures. In the first case (top), the cultivation is performed under the optimal growth temperature (Optimal Temperature at 100 %). In the second case (bottom), the temperature is unsuitable (Optimal Temperature at 60 %) and hence, the resulting cell growth and P(3HB) accumulation are lower. Nutrients and Sugar are adjusted to 100 % in both cases. After a while simulation, the Sugar activity is reduced to zero and thus, the Degradation component starts to negatively regulate P(3HB) activity. At this point, bacteria start to degrade their reserves.

7 Production by *S. thermodepolymerans*

The *Schlegelella* genus was first named in 2003 in honour of H. G. Schlegel who made a great contribution to the primal PHA research [65]. It was described as Gram-negative non-sporulating aerobic rod-shaped bacterium genus with the growth temperatures between 37 °C and 60 °C (optimum about 50 °C). In terms of phylogenetics, it is a member of the β-subclass of the *Proteobacteria*.

The species designation, *Schlegelella thermodepolymerans*, originated from its ability to depolymerise P(3HB) at high temperatures [65]. Several substances, such as gluconate, lactate, valerate or 3HB, were reported to be utilised by the isolated strains. Nevertheless, it was also stated that xylose or mannose are not employed which today proves to be wrong.

According to Kourilova et al. [2], *S. thermodepolymerans* is not only able to degrade, but also synthesise P(3HB). Furthermore, it transpired that the most efficient carbon source is xylose. The significant results from cultivation at 50 °C and PHA production on various carbon sources are given in the Tab. 6.1. The optimal cultivation temperature for PHA production was later experimentally determined as 55 °C. At this temperature, the cultivation on xylose resulted in CDM 6.27 g/l with the total amount of P(3HB) about 5.47 g/l.

Following xylose as the most efficient carbon source, the utilisation of xylose over glucose was further investigated. It was discovered that *S. thermodepolymerans* shows so called diauxic behaviour which means it utilises individual substrates sequentially. At first, bacterium employed xylose at the consumption rate about 0.19 l/h (glucose consumption rate was about 0.03 l/h). However, as the concentration of xylose decreased after 48 hours of cultivation, the substrate preference switched over to glucose at the consumption rate about 0.17 l/h (xylose consumption rate decreased to 0.06 l/h).

The last experiment performed was concerned with the bacterium ability to incorporate 3-hydroxyvalerate (3HV) and 4-hydroxybutyrate (4HB) into the polymer chain. Various 3HV precursors were used and the most efficient proved to be valeric acid. When using this precursor, the crystallinity of resultant P(3HB-co-3HV) titres was almost zero. However, the presence of valeric acid had an inhibitory effect on bacterial cultivation and hence, the titers were very low. This problem can be dealt with by managing the time of application and valeric acid concentration at the expense of lower 3HV yield. As for the other copolymer synthesis, it transpired that the capability of *S. thermodepolymerans* to produce poly(3-hydroxybutyrate-co-4-hydroxybutyrate) [P(3HB-co-4HB)] is too low to be further considered.

Based on these recent findings, it is obvious that *S. thermodepolymerans* represents a promising PHA producent among other bacteria. The preference of xylose together with the higher cultivation temperature are two key properties which could help to reduce the costs related to PHA production by bacteria. The following subchapter of this thesis will be concerned with deeper understanding of transport mechanisms and metabolic pathways for sugar degradation in *S. thermodepolymerans*.

7.1   Extending the Qualitative Base

In order to better comprehend the system's behaviour, it is favourable to start with a proper qualitative base. The metabolic pathways and reactions involved in the bacterium were derived from the annotated genome of *S. thermodepolymerans* DSM 15344. The complete genome sequence was assembled by Musilova et al. and it is available under the accession number CP064338. To explore *Schlegelella*'s metabolic network, the Pathway Tools software was used.

The Pathway Tools [66] allows to create, query and analyse the model-organism databases called Pathway/Genome Databases (PGDB). The software consists of a graphical environment for visualisation and some interactive tasks and an ontology and database API for the complex queries and computations. An important part of the tool is PathoLogic [67] component which enables to automatically create a new PGDB based on the annotated genome and MetaCyc [68] reference database. The PGDB for *S. thermodepolymerans* DSM 15344 was created by Dalimil Bujdoš from Microbial Bioengineering Laboratory, Department of Experimental Biology, Masaryk University lead by Dr. Pavel Dvořák.

Sugar Transport Mechanisms

When trying to understand the preferences in sugar uptake, one needs to deal with transport first. *S. thermodepolymerans* is a Gram-negative bacterium and thus, it has two membranes. Transport through the outer membrane is ensured by porins [69], channel-forming proteins which allow sugars to pass into the space between the two membranes, periplasm. Transport through the inner membrane into cytoplasm is then provided by more complex means.

Concerning xylose, the transport mechanism and metabolic degradation pathway have already been described [2]. Xylose import is ensured by ATP-binding cassette (ABC) transporter XylFGH, where XylF represents the high-affinity periplasmic binding protein, XylG is ATP-binding subunit and XylH is the integral membrane component of the uptake system [70]. The import via the ABC transporter requires one molecule of ATP as it belongs to active transports.

Usually, genes encoding XylFGH transporter form a single operon. In case of *S. thermodepolymerans*, this operon consists also of genes encoding the xylose degradation pathway, *xylA* and *xylB*. It means, by the time xylose enters cytosol via the ABC transporter, it is immediately metabolised via the xylose isomerase pathway. The most similar sequences to these genes were found in *Phaeobacter inhibiens* DSM 17395 [71]. However, the genes *xylF*, *xylG*, *xylH* and *xylA*, *xylB* formed two separate operons and also, the XylFGH transporter probably enabled to import glucose and sucrose as well.

In *Escherichia coli*, the model Gram-negative bacterium, xylose is usually imported via the xylose/proton symporter XylE or ABC transporter XylFGH [70]. The XylE symporter is a representative of major facilitator superfamily (MFS) and in *E. coli*, it is considered to be less efficient and lower-affinity compared to XylFGH. No similar genes to *xylE* were found in *Schlegelella*'s genome.

Besides the XylFGH, a non-specific ABC sugar transporter was predicted in *Schlegelella*'s transport system. It is nothing surprising since the specificity of substrate-binding protein is often unclear and varies in different species. Anyway, there is one more type of transporter in the game.

The third important group of transport proteins is PEP-dependent phosphotransferase system (PTS) [69]. It is characterised by three unique features compared to ABC transporters or MFS: first, it is found only in bacteria; second, the sugars are phosphorylated during the transport and third, it affects carbon and nitrogen metabolism in relation to available sugars.

A PTS transporter consists of three components from which two are general and one is substrate-specific. The two common phosphotransferase proteins are enzyme I (EI) and histidine protein (HPr), the third one is sugar-specific complex called enzyme II (EII). The sugar transport is ensured by sequential transfer of phosphoryl group from phosphoenolpyruvate (PEP) to sugar. In *S. thermodepolymerans*, a kind of TPS transporter was predicted with either non-specific, or fructose transporter subunit IIA. The basic scheme of PTS is depicted in Fig. 7.1.

Carbon Catabolite Repression Phenomenon

It is nothing new that bacteria can utilise multiple carbon sources. After all, the PHA production by thermophilic bacteria from various substrates has already been discussed in this thesis (see Tab. 6.1). In most cases, however, only one substrate was examined at the time. Interestingly, what happens when multiple substrates are available at the same time?

In such a case, most bacteria usually start to use the individual substrates sequentially according to their preferences. This phenomenon is called carbon catabolite repression (CCR) and it is typically glucose which is preferred over the others [72]. Nevertheless, *S. thermodepolymerans* shows different behaviour: it prefers xylose over glucose. This occurrence can be termed as reverse CCR.

The reason why only one preferred substrate is consumed at the time is that the genes for the use of secondary carbon sources are not currently expressed. The transcription is usually controlled by global and operon-specific regulations. Concerning xylose utilisation in *E. coli*, the transcription of genes encoding both, the transport via XylFGH and the degradation pathway, are regulated via the global cyclic AMP (cAMP) receptor protein (CRP) and a type of specific regulator XylR [70]. In *S. thermodepolymerans*, no similar genes to *xylR* were found.

As regards the global control of sugar uptake in *E. coli*, the key role plays the PEP-dependent phosphotransferase system [72]. The regulation mainly depends on phosphorylation of EIIA domain of the EII complex which is influenced by the activity of PTS transporter and the phosphoenolpyruvate (PEP) to pyruvate ratio.
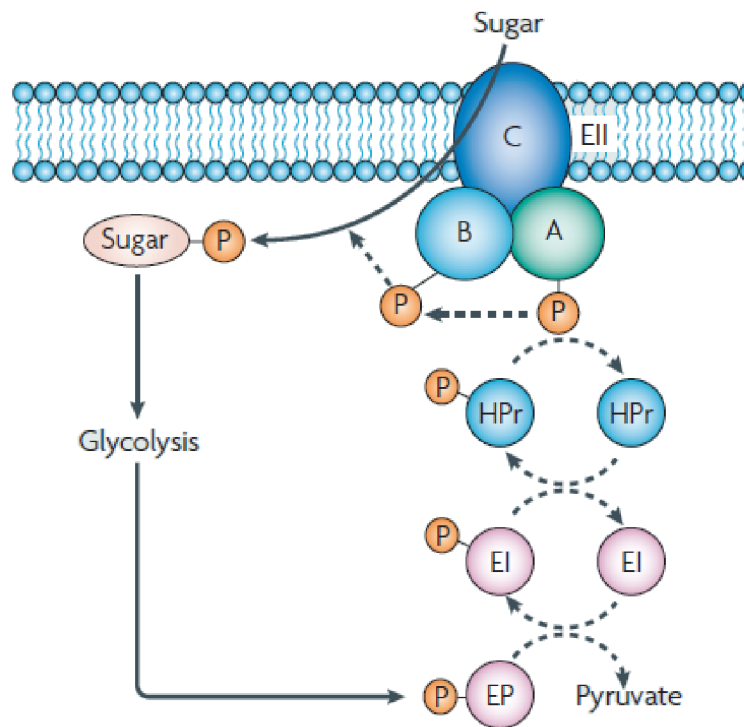

Fig. 7.1: The basic scheme of PEP-dependent phosphotransferase system (PTS) [72]. The phosphoryl group is sequentially transferred from phosphoenolpyruvate (PEP) to sugar. P: phosphate, EP: enolpyruvate, EI: enzyme I, HPr: histidine protein, EII: enzyme II and its domains A, B, C.

If a PTS sugar (typically glucose) is available, EIIA subunit remains in dephosphorylated state as it transfers the phosphoryl group to sugar during the transport. In dephosphorylated form, EIIA cannot activate adenylate cyclase (AC). After majority of glucose is consumed, EIIA remains mostly phosphorylated and hence, it can activate AC which further catalyses the conversion of ATP into cAMP. Cyclic AMP then binds CRP and together they activate the expression of catabolic genes for other non-PTS sugars.

Sugar Degradation Pathways

As already mentioned, genes *xylA*, *xylB* encoding xylose degradation pathway are regulated together with the genes for xylose transport. In other words, once it is transported into cytoplasm, it is immediately degraded via xylose isomerase pathway. At first, xylose is transformed into xylulose by xylose isomerase (*xylA*) and xylulose is next converted into xylulose 5-phosphate by xylulokinase (*xylB*).

Xylulose 5-phosphate can be further metabolised via the non-oxidative branch of pentose phosphate pathway (PPP). The purpose of this fully reversible part is simply to rearrange carbon atoms in molecules in order to produce C4, C5 and C7 sugars [73]. These can be further used for biosynthesis. Regarding production of polyhydroxyalkanoates, xylulose 5-phosphate together with erythrose 4-phosphate can be converted into fructose 6-phosphate and glyceraldehyde 3-phosphate. GAP can further engage in glycolysis which leads to pyruvate formation and after pyruvate decarboxylation to acetyl-CoA, P(3HB) can be formed via acetyl-CoA to 3-hydroxybutyryl-CoA pathway (pathway I in Fig. 5.5).

Concerning glucose, it is probably transported via a non-PTS transporter and thus, the degradation starts with phosphorylation by glucokinase in cytoplasm. Glucose 6-phosphate is converted into fructose 6-phosphate by glucose-6-phosphate isomerase, another phosphoryl group is joined by 6-phosphofructokinase and the resulting fructose 1,6-bisphosphate can be then transformed into GAP and dihydroxyacetone phosphate by fructose-bisphosphate aldolase. At this point, different sugar degradation pathways usually converge. The basic scheme denoting xylose and glucose degradation pathways in *S. thermodepolymerans* is depicted in Fig. 7.2.

Glucose 6-phosphate can also enter the oxidative branch of PPP. Unlike the non-oxidative part, it is unidirectional and it aims at production of NADPH and ribulose 5-phosphate [73]. The latter further engages in the non-oxidative branch. Reduced nicotinamide adenine dinucleotide phosphate, shortly NADPH, is produced by reduction of $NADP^+$ and mainly acts as an electron donor in anabolic reactions. In *S. thermodepolymerans*, the presence of the oxidative part of PPP is questionable as orthologues of two important genes cannot be identified in the genome.
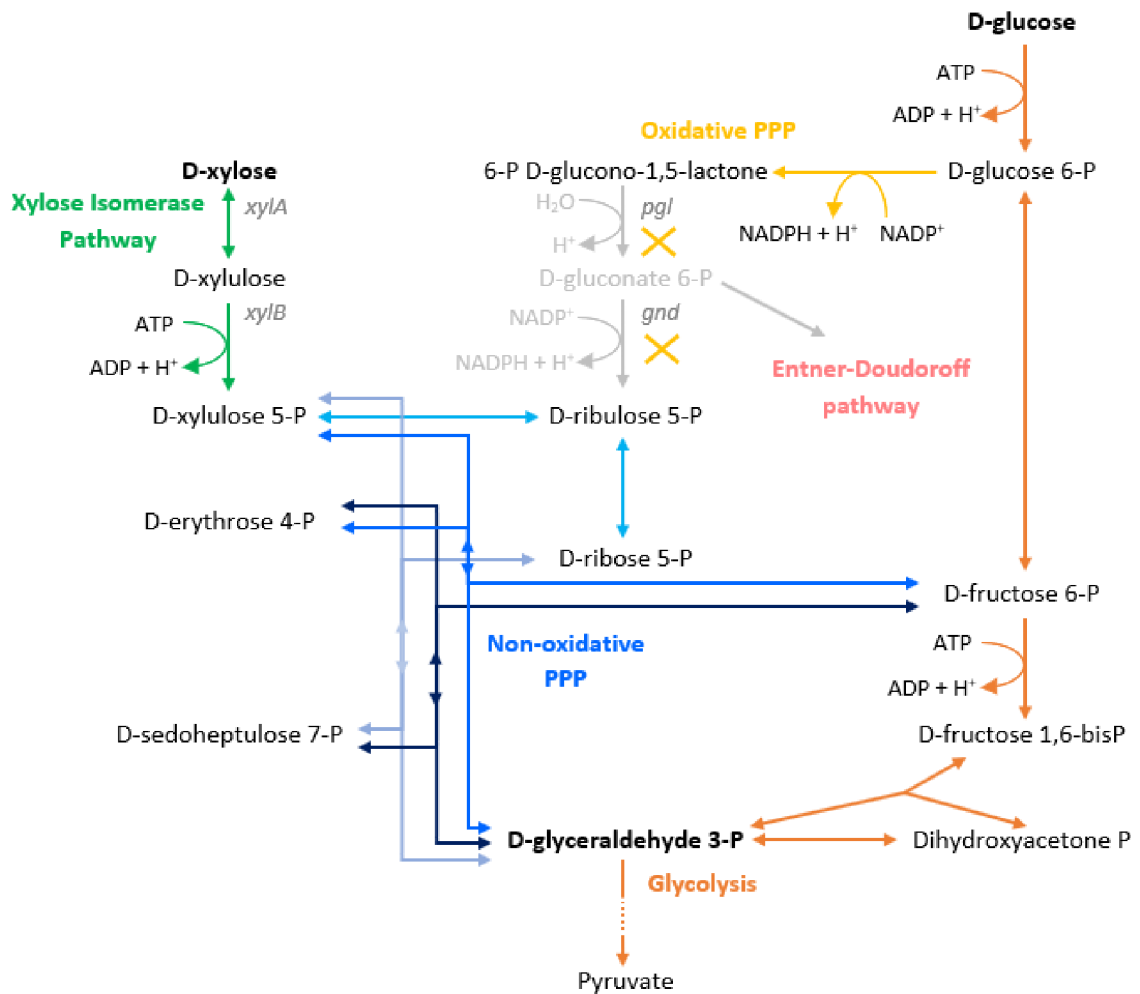
Fig. 7.2: The basic scheme of xylose and glucose degradation pathways in *S. thermodepolymerans*. Xylose is transformed via xylose isomerase pathway (green) into xylulose 5-P which further enters the non-oxidative branch (blue shades) of pentose phosphate pathway (PPP). Glucose is metabolised via glycolysis, aka the Embden-Meyerhof-Parnas pathway (orange). The presence of the oxidative branch (yellow) of PPP is uncertain as orthologues of *pgl* and *gnd* genes cannot be identified in the *Schlegelella*'s genome.

The first one is *pgl*, the gene encoding 6-phosphogluconolactonase which cataly-ses hydrolysis of 6-phosphoglucono-1,5-lactone to 6-phosphogluconate. In the past, the reaction was thought to occur spontaneously, but later study [74] provided new insights. It was revealed that the enzyme not only accelerates the reaction, but it also prevents lactones from conversion into the gamma form. This intramolecular rearrangement would lead to a biochemical dead end.

The second gene probably missing is *gnd*. It encodes 6-phosphogluconate de-hydrogenase which catalyses the conversion of gluconate 6-phosphate into ribu-lose 5-phosphate with NADPH as a side product. A reference protein sequence from *E. coli* was aligned to *Schlegelella*'s proteome using BLASTP [75] and a kind of NAD(P)-dependent oxidoreductase was found. However, only about 26 % was identical with the 57 % of the query sequence and hence, the presence of this enzyme in *S. thermodepolymerans* cannot be clearly determined without further exploration.

The deletion of *gnd* has already been investigated in *E. coli* and as a result, some compensatory mechanisms have been described [76]. One of them is that gluconate 6-phosphate, which cannot be further converted into ribulose 5-phosphate, takes the Entner-Doudoroff (ED) shunt instead. The ED pathway is an alternative to glycolysis and leads to GAP and pyruvate formation. Thanks to the gluconate 6-phosphate consumption by ED, half of the NADPH yield could be preserved. In *S. thermodepolymerans*, *pgl* gene is absent as well and thus, the use of ED shunt is uncertain.

Except for the alternative ways for NADPH production, changes in direction of fluxes were observed in the non-oxidative part of PPP. Since ribulose 5-phosphate cannot be produced from the oxidative branch, it must be obtained additionally via the carbon rearrangement of intermediates from the latter part of PPP.

7.2 Dynamic Model Refinement

A dynamic model for PHA production by thermophilic bacteria has already been established (see Chapter 6). Now, it will be refined with a focus on *S. thermode-polymerans* and the gathered knowledge discussed in the previous subchapter.

Starting with the carbon sources, the general term "sugar" at the input of the ini-tial model will be replaced by two specific substrates, xylose and glucose. Xylose is degraded by the xylose isomerase pathway into xylulose 5-phosphate which next enters the non-oxidative part of PPP. Glucose is metabolised only via glycolysis assuming the oxidative branch of PPP is absent. The sequence of metabolic reac-tions describing GAP transformation into pyruvate is not included in order to keep the model's arrangement clear. Pyruvate is then converted into acetyl-CoA which enters the pathway I resulting in P(3HB) synthesis.

The intracellular and extracellular sugar are distinguished by the addition of "int" or "ext" to the name of the sugar component. The diauxic behaviour of the system is then ensured by negative regulation of glucose import (i.e. Glucose int) which is conditioned by the presence of xylose in the cultivation medium (Xylose ext). Only after most of the xylose has already been consumed (the activity of Xylose ext decreases), glucose starts to be imported (the activity of Glucose int increases).

Other components from the initial model remain, those are Nutrients, Optimal Temperature, Cell Growth and Degradation. Thus, P(3HB) synthesis can now be simulated with two different carbon sources and various temperature and nutrient settings. Nevertheless, the higher efficiency of xylose utilisation as a sole carbon source is still not included in the model. Actually, P(3HB) production from xylose is even slower than from glucose due to the higher number of steps involved in its metabolic pathway. To correct this behaviour, the production of four metabolic precursors is taken into account in the model.

The central carbon metabolism (i.e. glycolysis, pentose phosphate pathway and tricarboxylic acid cycle) generates 12 metabolic precursors which ensure the entire biomass production [77]. Four of them are relevant to involve into the model, these are: ribose 5-phosphate, necessary for biosynthesis of nucleotides; erythrose 4-phosphate, used for the formation of aromatic amino acids; glucose 6-phosphate, precursor of glycogen and lipopolysaccharides; and fructose 6-phosphate, needed for the cell wall formation.

The production of the four precursors was uneven when using the two carbon sources separately, which favoured the use of glucose. Since these metabolites are essential for the proper functioning of cells, the activity of Cell Growth component was conditioned by their activity. It means that cells cannot grow and thus, P(3HB) cannot be synthetised until all four metabolic precursors are present.

To be able to use the proposed compensatory mechanism, one remaining problem needs to be solved – the direction of reaction fluxes in the non-oxidative PPP. A flux balance analysis will be used to deal with this issue.

Performing Flux Balance Analysis

To perform flux balance analysis, MetaFlux component [78] of Pathway Tools can be used. The input model is specified by a simple .fba text file which enables users to determine various parameters. Based on these instructions, an FBA model is generated from the relevant Pathway/Genome Database. After running the input model, MetaFlux provides the user with a .log file reporting whether the specified reactions were employed or not. This is helpful especially during the initial model formation when there might be no fluxes found for any reaction.

To create a new FBA model, a template FBA file and several examples are available. The first parameter to be specified is the relevant PGDB. Next, there are four basic sections to be completed: reactions, biomass, nutrients and secretions. To build a feasible model, it is important to consider all substances consumed or produced by the specified reactions. Once there is one of them missing, the MetaFlux solving mode will probably generate zero fluxes and the user is dependent on the .log file. To facilitate the creation of more complex models, MetaFlux provides users with a development mode.

The FBA model for *S. thermodepolymerans* consists of four metabolic pathways: xylose degradation pathway, non-oxidative part of PPP, glycolysis and pyruvate decarboxylation to acetyl-CoA. Involvement of pathway I for P(3HB) synthesis is problematic since the polymer formation is defined by a single step polymeric reaction. Hence, the product to be maximised (placed in the biomass compartment) is acetyl-CoA which can further enter the acetyl-CoA to 3-hydroxybutyryl-CoA pathway.

Nutrients and secretions were defined according to the requirements of the reactions. The important nutrients are xylose and glucose whose fluxes were limited by the same upper bound. This should help to discover and compare the metabolic capabilities of both sugars. The fluxes of other nutrients, namely NAD, phosphate and coenzymeA, were not limited[1] (in practice, these would be obtained as products of other reactions which are not part of the model).

The crucial secreted metabolites were the four metabolic precursors discussed earlier: ribose 5-phosphate, erythrose 4-phosphate, glucose 6-phosphate and fructose 6-phosphate. The boundaries for their secretion were adjusted according to the previously mentioned study [76] concerning *gnd* deletion in *E. coli*. A part of the study was to examine the chemostat culture in a steady state in order to determine the distribution of reaction fluxes in the central carbon metabolism. Thus, the secretion boundaries for the *Schlegelella*'s FBA model were set with respect to the ratio among the relative fluxes of the four metabolites in *E. coli*. The fluxes of other secretions, such as NADH, protons, water and carbon dioxide, were not additionally limited.

The specific numbers of all boundaries were adjusted partly by trial and error in order to obtain fluxes about between zero and 100. The input constraints together with the results from the analysis using the two different carbon sources are depicted in Fig. 7.3 and Fig. 7.4. In both cases, the given constraints led to the same production of acetyl-CoA and the same secretion fluxes. Furthermore, the consumption of nutrients, except for sugars, was equal as well.

---

[1]In fact, all fluxes are limited by default from zero to 3000. Nevertheless, these values are not even achievable with respect to the other chosen constraints.
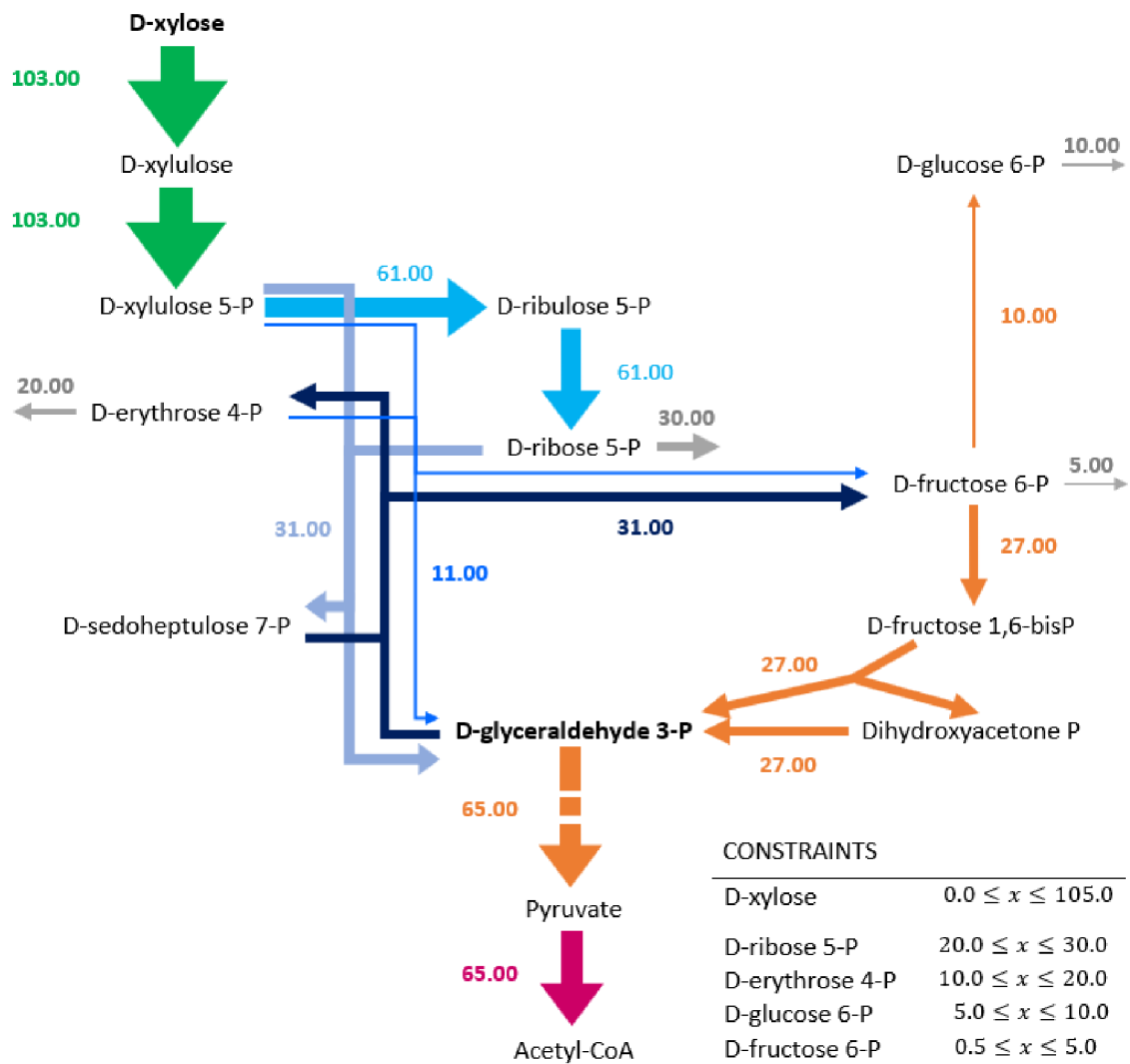
Fig. 7.3: The distribution of reaction fluxes in xylose degradation pathway based on the flux balance analysis. The product to maximise is acetyl-CoA which further enters the pathway I for P(3HB) synthesis. The constraints were defined so that the chosen metabolic precursors were produced in the specific ratio. Note that the non-oxidative PPP carries a considerable amount of flux.
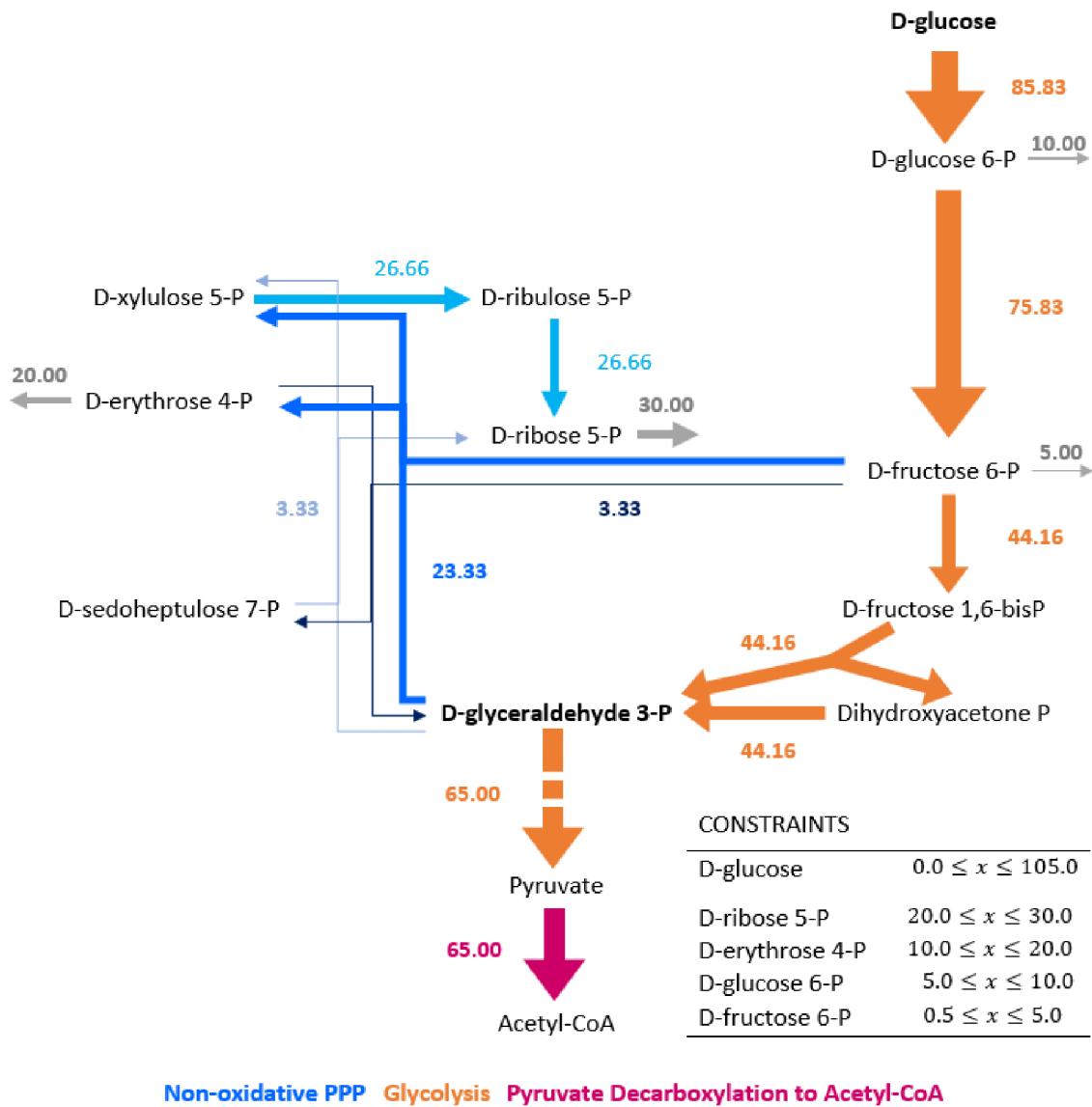
Fig. 7.4: The distribution of reaction fluxes in glucose degradation pathway based on the flux balance analysis. The parameters of the model are the same as in the previous case. In comparison to xylose, several reactions flow in the opposite direction and PPP carries lower fluxes, whereas glycolysis is greatly employed. Importantly, the overall yield of acetyl-CoA is the same in both cases. However, the lower amount of glucose is required compared to xylose.

As shown in Fig. 7.3 and Fig. 7.4, the consumption of xylose was about 103 whereas the consumption of glucose about 86. Since both sugars allowed to produce the same amount of acetyl-CoA with the same amount of nutrients consumed and secretions produced, it can be deduced that glucose has probably higher metabolic potential than xylose. In other words, a smaller amount of glucose leads to the same production in comparison to xylose.

As for the other reaction fluxes, different values and directions can be observed in the two cases. With xylose consumption, significant fluxes are detected in PPP, whereas glycolysis is less employed. One of the reactions of glycolysis is reversed in order to additionally produce the remaining precursor, glucose 6-phosphate. With glucose consumption, the fluxes in PPP are lower and three out of five are heading the opposite direction. The fluxes in glycolysis are noticeably higher.

When the directions of reaction fluxes are determined, the dynamic model can be finally completed. The resulting scheme is depicted in Fig. 7.5.
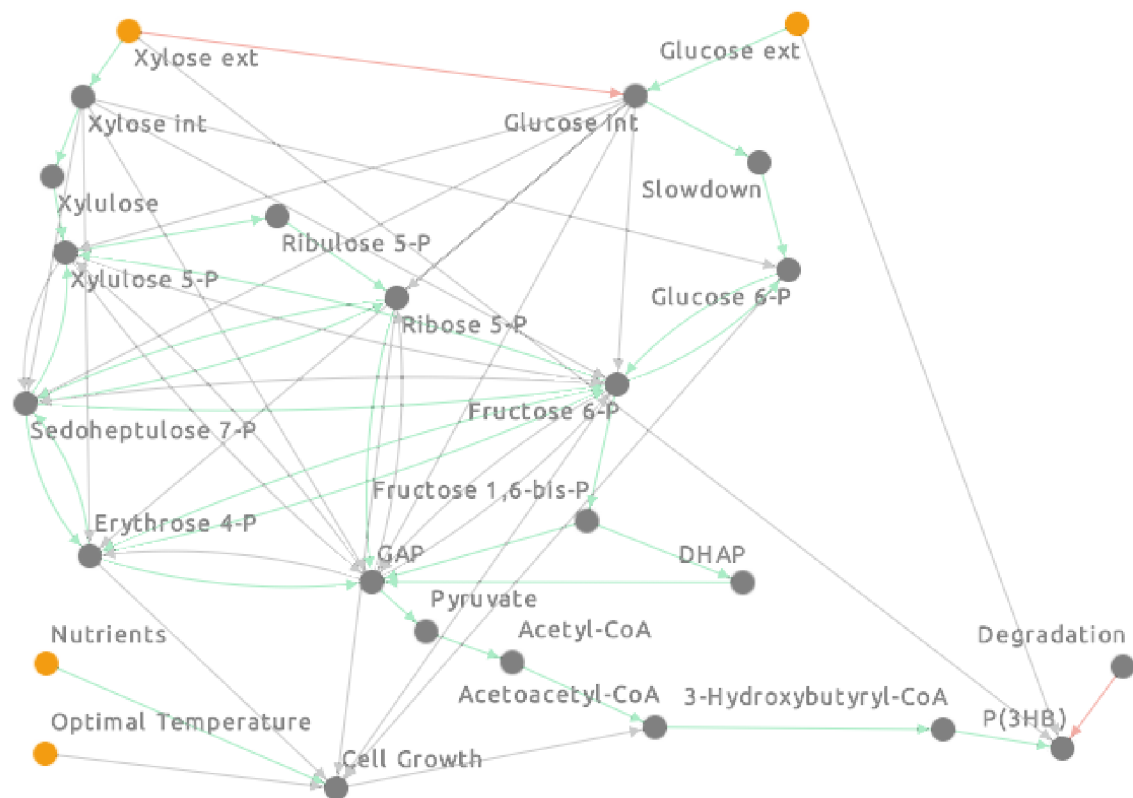

Fig. 7.5: The final scheme of the dynamic model for P(3HB) synthesis in *S. thermod-epolymerans*. The model consists mainly of xylose isomerase pathway, non-oxidative pentose phosphate pathway, glycolysis and acetyl-CoA to 3-hydroxybutyryl-CoA pathway. There are four adjustable external components (highlighted in orange): Xylose ext, Glucose ext, Nutrients and Optimal Temperature.

Exploring the Dynamics

To recapitulate, the general model for PHA production by thermophilic bacteria was refined with respect to *S. thermodepolymerans*. The input carbon sources are xylose and glucose. The model demonstrates the sugar degradation pathways, reflects the diauxic behaviour and follows the assumption of the necessary production of metabolic precursors. Once the last condition is met, cells are growing and P(3HB) is synthetised. Both can be further influenced by the activity of external components such as Optimal Temperature or Nutrients (besides the input sugar). After majority of sugar is consumed, cells start to degrade the accumulated reserves.

The employment of the production of metabolic precursors resulted in equal P(3HB) synthesis from both input sugars. The situations are balanced now, but the higher efficiency of xylose is still not reproduced. Hence, an artificial component "Slowdown" was added to the model in order to demonstrate the lower efficiency of P(3HB) synthesis from glucose. This one more step slows down the glucose degradation which results in delayed P(3HB) formation compared to xylose.

To demonstrate the model's behaviour, three model situations were examined. The two of them are shown in Fig. 7.6 and are supposed to depict the difference between the xylose and glucose utilisation. Thus, the preconditions were equal in both cases. Nutrients and the relevant sugar were at 100 %, the Optimal Temperature was adjusted to 90 % in order to simulate the initial experiment with different carbon sources. Usually, the effect of temperature is studied only after this primary investigation. In *S. thermodepolymerans*, the various carbon sources were examined under the temperature 50 °C, while the optimal temperature was later experimentally determined as 55 °C [2].

As can be seen in the simulation plots, the P(3HB) production from glucose is one step slower. This occurrence is ensured by the artificial component Slowdown which delays the glucose degradation. The real reason why the use of xylose is more effective is still the subject of discussion. A significant role plays probably the joint regulation of genes encoding xylose transport and degradation pathway. As previously mentioned, xylose is imported by the XylFGH transporter and as soon as it enters cytosol, it is immediately degraded. Concerning glucose there is not much information about its transport. It can only be said that there was no glucose-specific TPS transporter found in the *Schlegelella*'s genome.

The absence of glucose-specific TPS could support the occurrence of specific diauxic behaviour. As described in section 7.1, TPS contributes to CCR phenomenon as a global regulator. When TPS sugar is available, the genes for transport of other sugars are not expressed. In many bacteria, glucose is preferred over the others, but *Schlegelella* rather uses xylose. Such behaviour is depicted in Fig. 7.7.
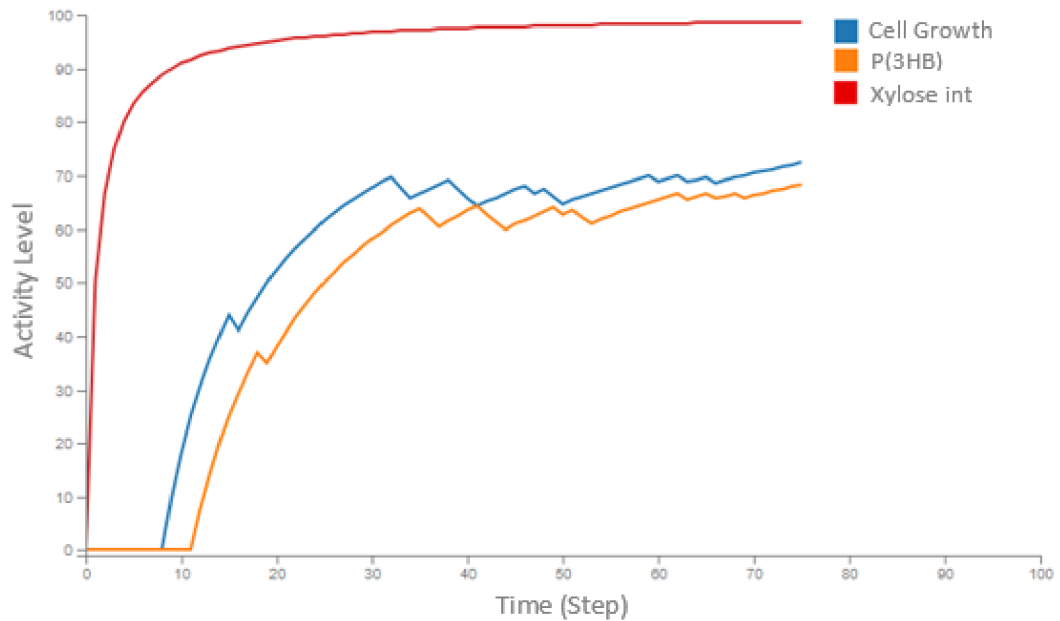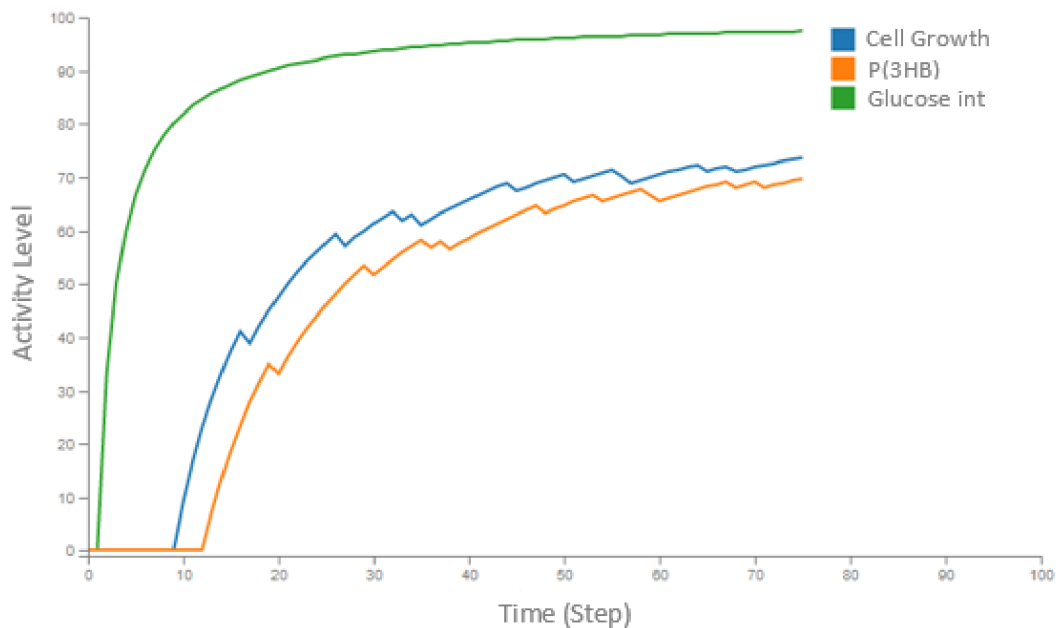
Fig. 7.6: The simulation of P(3HB) production by *S. thermodepolymerans*. The upper plot denotes the production from xylose, the lower from glucose. As can be seen, P(3HB) formation from glucose is one step behind compared to xylose. This occurrence should demonstrate the higher efficiency of xylose utilisation. As the cultivation temperature is not ideal (Optimal Temperature at 90 %) during both experiments, cell growth and P(3HB) yield cannot reach the maximal values.

Fig. 7.7: The simulation of the specific diauxic behaviour of *S. thermodepolymerans*. When xylose and glucose are both available, bacteria first use xylose and only after most of the xylose is consumed, they start to consume glucose. Note that glucose is present in the cultivation medium from the very beginning (Glucose ext at 100 %), but it is imported into cells (Glucose int) later due to the xylose decrease. The initial setting of all external components was 100 %.
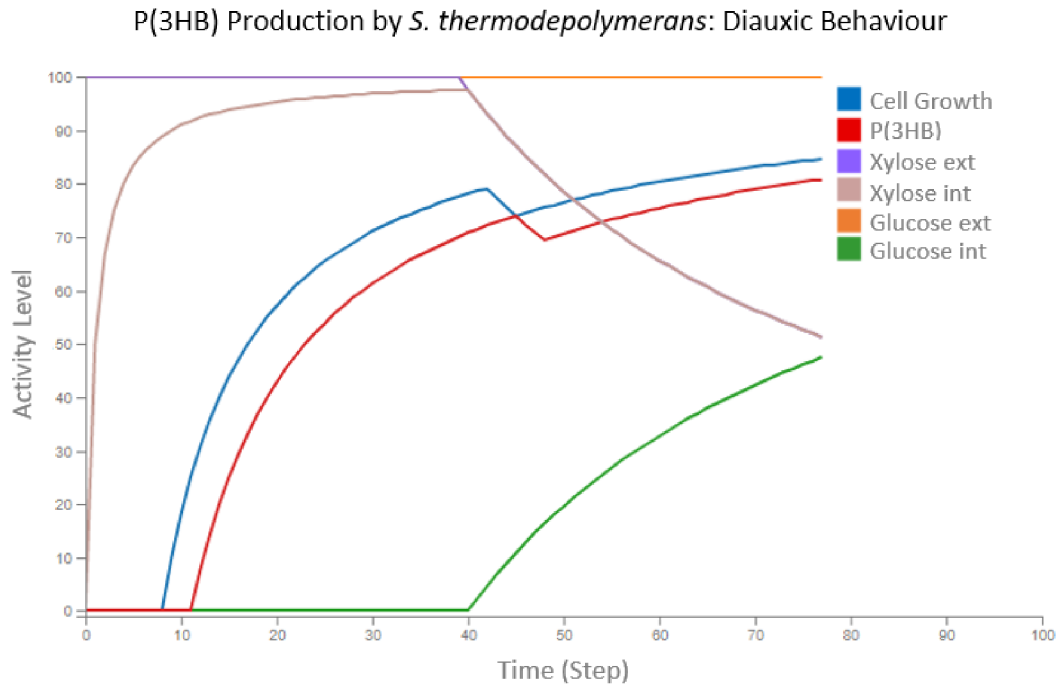
The simulation plot above denotes the sequential use of available sugars. Since *Schlegelella* prefers xylose, it is immediately imported and catabolised as previously demonstrated at the top of Fig. 7.6. Glucose is present in the cultivation medium, but it has not been used yet. After about 40 steps, Xylose ext activity is decreased from 100 % to zero to simulate the significant decline in the external xylose concentration. Such a fall finally triggers the transition to glucose consumption. Regarding the cultivation conditions, the simulation was performed with ideal temperature and enough nutrients (both external components at 100 %).

As denotes the Fig. 7.7, there is a little decline in growth and P(3HB) synthesis due to the switching to glucose consumption. In practice, the amount of P(3HB) was constantly increasing during the whole cultivation [2]. However, the cell dry mass measured at the time of switching was slightly lower than before, suggesting the growth had probably stopped for a while due to the decreasing amount of xylose. After switching to glucose, an increase in CDM was observed again. Additionally, there is only one sugar used at the time in the model. In actual fact, the other sugar is consumed as well, but at a very low consumption rate.

Even though the model's behaviour in the previous three situations reflects the results from real experiments, the knowledge behind the scenes is still not sufficient. The joint regulation of xylose transport and utilisation is suggested to be the key factor contributing to both, xylose preference over glucose and its higher efficiency as a sole carbon source. These two curiosities observed in *Schlegelella*'s behaviour can be considered to go hand in hand. Further, it was stated that there is no glucose-specific TPS found in the bacterium's genome which also supports the absence of typical glucose preference.

Another fact to consider is that the oxidative branch of PPP is probably missing in *Schlegelella*'s metabolism. In case the pathway is really absent, there have to be some compensatory mechanisms for production of NADPH (which is usually formed mostly via this pathway) and thus, consumption of glucose does not have to be such an advantage. Moreover, ribulose 5-phosphate must be produced alternatively by the means of the non-oxidative part.

Concerning the latter consequence, the benefit of xylose is that it is degraded in two steps into xylulose 5-phosphate which can be directly converted into ribulose 5-phosphate. Glucose, on the contrary, has to be catabolised into fructose 6-phosphate and then GAP, which can together form xylulose 5-phosphate and erythrose 4-phosphate. Xylulose 5-phosphate is then transformed into ribulose 5-phosphate. Thus, the pathway to compensate the ribulose 5-phosphate formation is shorter when using xylose.

On the other hand, glucose 6-phosphate, as one of the metabolic precursors, must be obtained additionally when using xylose. However, there might not be such requirements for its production compared to ribulose 5-phosphate (or ribose 5-phosphate, more precisely) as follows from the *gnd* deletion study in *E. coli* [76].

The fact which plays in favour of glucose is that it has probably higher metabolic potential in comparison with xylose. According to the flux balance analysis, which was previously performed, to produce the same amount of acetyl-CoA and other metabolites, a smaller amount of glucose is required. Nevertheless, the model was not sufficiently comprehensive and thus, the results cannot be entirely convincing.

There is another effective carbon source, which has not been mentioned yet, and that is cellobiose; two molecules of glucose joined by the $\beta(1 \rightarrow 4)$ glycosidic bond. The P(3HB) production by *S. thermodepolymerans* from this substrate was even higher than from xylose (data not published). This is not very surprising since cellobiose provides cells with double amount of energy compared to glucose. Thus, it is likely that it can overcome the certain benefits of xylose.

Other substrates, such as glycerol, lactose or fructose, provided reasonable results. However, these are not so attractive in terms of reducing the production costs and hence, this master's thesis focused mainly on xylose utilisation.
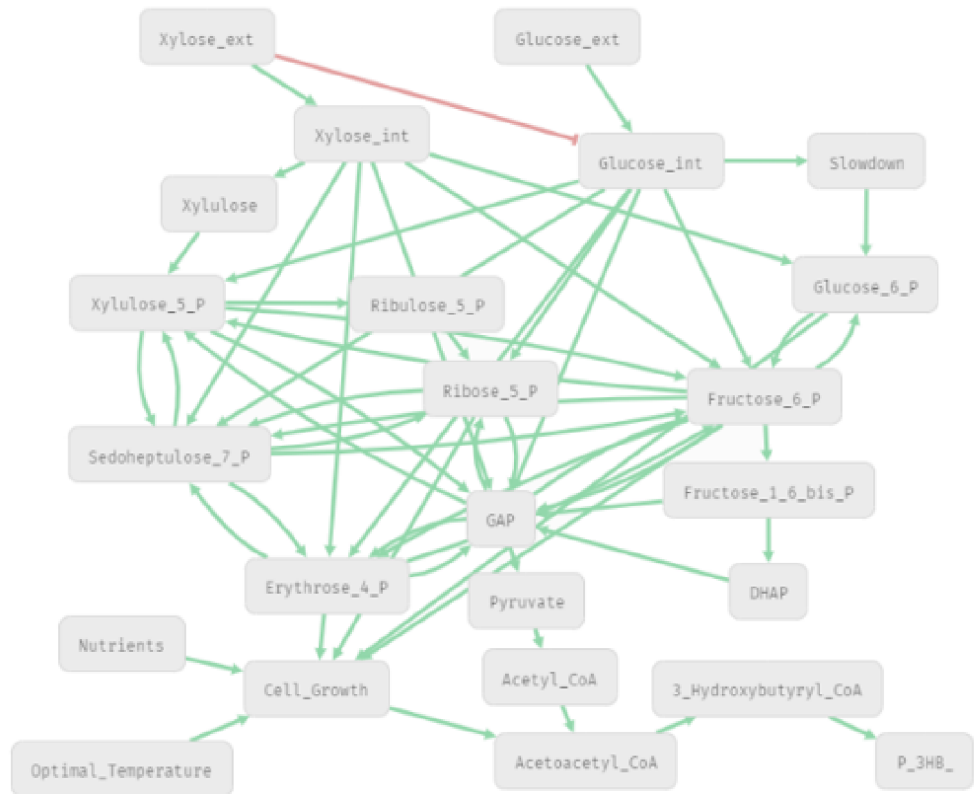
Attractor Analysis

The last analysis was concerned with the long-term behaviour of the model. It was performed using the AEON [79] software which is designed for the attractor bifurcation analysis of asynchronous parametrised Boolean models. The parametrisation of the model allows researchers to define update functions as logical parameters which comes in useful when some of the functions are not completely known. The created model for *S. thermodepolymerans* is non-parametrised as all parameters are fixed. Thus, it can be seen as one instantiation of a parametrised Boolean model.

The software includes two basic components, the client and the compute engine. The client provides users with a web-based graphical interface. Besides creating a new model, it enables to import and edit an existing model in .sbml format. The compute engine is then used to perform the analysis.

Several problem were coupled with the analysis of the *Schlegelella*'s model. In Cell Collective, the application and subsequent vanishing of carbon source from the cultivation medium was demonstrated by manual adjustment of the relevant external sugar component. In other words, the diauxic behaviour or the degradation of accumulated granules could be simulated only by the manual intervention into the course of simulation. However, such actions cannot be made in attractor analysis and thus, the key elements of *Schlegelella*'s behaviour cannot be included in the AEON model.

This fact leads to the considerable simplification. The Degradation component and relevant links were deleted as they were useless here. The Optimal Temperature and Nutrients were set as true (since nothing between zero and one can be set in AEON) and the sugar input components were not specified. Such model has four initial states – both sugars are present, only one, or any of them. Three of these situations lead to the cell growth and P(3HB) formation, the fourth is the case when no sugar is present. When both sugars were available, the bacterium used only xylose, which is in agreement with the model's disposition. Such situation is denoted in Fig. 7.8.

The three applicable situations denote the cultivation with ideal temperature and constant supply of nutrients and carbon source. Resulting attractors for such preconditions show stability. In practice, such conditions are typical for chemostat cultivations, where nutrients are continuously supplied, and the bacterial culture is regularly harvested. This method is considered to be the most reproducible technique for PHA production by bacteria [80].

Fig. 7.8: Screenshots from the analysis of attractors in AEON. Four initial states were considered (both sugars are present, only one, or any of them) and all resulted in a stable state. If at least one of the sugars is available, cells grow and P(3HB) is synthetised. The resulting stable state is described via binary notation or colour-coded names of components. The one in the picture corresponds to the situation when both sugars are present, but only xylose is consumed as it is preferred.

Conclusions

The aim of this master's thesis was to reconstruct a dynamic model for PHA production by thermophilic bacterium *Schlegelella thermodepolymerans*. At the beginning, the theoretical foundations were given. The Chapter One focuses on mathematical graph theory as the basis of biological networks. It is followed by Chapter Two which deals with different approaches in dynamic modelling and commonly used tools for dynamic analysis of complex systems. In the third chapter, further terms and possibilities related to model analysis are introduced and the fourth chapter then aims at metabolomics and the frequently used laboratory techniques.

The fifth chapter is finally dedicated to polyhydroxyalkanoates and in the end, the three natural metabolic pathways for PHA synthesis in bacteria are introduced. The subsequent chapter then deals with PHA production by thermophilic bacteria. Before focusing on *S. thermodepolymerans*, a general dynamic model is proposed. The Boolean model was created in Cell Collective [21] and its behaviour reflects the usual course of basic experiments. The simulation plots depict the PHA degradation due to the large decrease of carbon source in cultivation medium and the effect of temperature on cell growth and polymer yields.

Once the general model was established, it was further refined with respect to *S. thermodepolymerans*. Since the bacterium is specific by the effective use of xylose and its preference over glucose, the first part of Chapter Seven is focused on transport mechanisms and degradation pathways of the two sugars. The actual refinement of the general dynamic model is then discussed in the second part.

The final model includes mainly xylose isomerase pathway, non-oxidative pentose phosphate pathway, glycolysis and pathway I for P(3HB) production. It is proposed that the oxidative branch of PPP is probably missing in the *Schlegelella*'s metabolic network. To determine the direction of fluxes in the non-oxidative part, flux balance analysis was used. Besides assisting in the model's arrangement, results from FBA indicate to the likely higher metabolic capacity of glucose. The attractor analysis of simplified model resulted in the stable state of P(3HB) production with continuous supply of nutrients.

In conclusion, the final model for PHA production by *S. thermodepolymerans* demonstrates the higher efficiency of xylose utilisation and the specific diauxic behaviour. Nevertheless, the reason behind such occurrences is still the subject of discussion. The joint regulation of genes encoding xylose transport and degradation is proposed to be the key factor. Since xylose is an inexpensive substrate and higher cultivation temperature reduces the costs related to sterilisation, *S. thermodepolymerans* represents a promising PHA producent among other bacteria. The knowledge gathered in this thesis should lay the foundations for further research.

Bibliography

[1] SHAHID, Salma, Sadia RAZZAQ, Robina FAROOQ, Zill-i-Huma NAZLI. Polyhydroxyalkanoates: Next generation natural biomolecules and a solution for the world's future economy. *International Journal of Biological Macromolecules* [online]. 2020 [cit. 2020-11-22]. ISSN 01418130. DOI: 10.1016/j.ijbiomac.2020.10.187.

[2] KOURILOVA, Xenie, Iva PERNICOVA, Karel SEDLAR, Jana MUSILOVA, Petr SEDLACEK, Michal KALINA, Martin KOLLER, Stanislav OBRUCA. Production of polyhydroxyalkanoates (PHA) by a thermophilic strain of *Schlegelella thermodepolymerans* from xylose rich substrates. *Bioresource Technology* [online]. 2020, 315 [cit. 2020-11-25]. ISSN 09608524. DOI: 10.1016/j.biortech.2020.123885.

[3] KESIĆ, Srdjan. Systems biology, emergence and antireductionism. *Saudi Journal of Biological Sciences* [online]. 2016, 23(5), 584-591 [cit. 2020-09-03]. ISSN 1319562X. DOI: 10.1016/j.sjbs.2015.06.015.

[4] KOUTROULI, Mikaela, Evangelos KARATZAS, David PAEZ-ESPINO, Georgios A. PAVLOPOULOS. A Guide to Conquer the Biological Network Era Using Graph Theory. *Frontiers in Bioengineering and Biotechnology* [online]. 2020, 8 [cit. 2020-09-11]. ISSN 2296-4185. DOI: 10.3389/fbioe.2020.00034.

[5] BARZEL, Baruch, Amitabh SHARMA, Albert-László BARABÁSI. Graph Theory Properties of Cellular Networks. *Handbook of Systems Biology* [online]. Elsevier, 2013, 2013, s. 177-193 [cit. 2020-10-11]. ISBN 9780123859440. DOI: 10.1016/B978-0-12-385944-0.00009-5.

[6] CHAO, Shih-Yi. Graph Theory and Analysis of Biological Data in Computational Biology. JAYANTHAKUMARAN, Kankesu, ed. *Advanced Technologies* [online]. InTech, 2009, 2009-10-01 [cit. 2020-10-11]. ISBN 978-953-307-009-4. DOI: 10.5772/8205.

[7] HSU, Lih-Hsing, Cheng-Kuan LIN. *Graph theory and interconnection networks*. Boca Raton: CRC Press, c2009. ISBN 978-1-4200-4481-2.

[8] PALANIAPPAN, Sucheendra K., Ayako YACHIE-KINOSHITA, Samik GHOSH. Computational Systems Biology. *Encyclopedia of Bioinformatics and Computational Biology* [online]. Elsevier, 2019, 2019, s. 789-795 [cit. 2020-09-17]. ISBN 9780128114322. DOI: 10.1016/B978-0-12-809633-8.20287-2.

[9] SZÉKELY, Tamás, Kevin BURRAGE. Stochastic simulation in systems biology. *Computational and Structural Biotechnology Journal* [online]. 2014, 12(20-21), 14-25 [cit. 2020-09-17]. ISSN 20010370. DOI: 10.1016/j.csbj.2014.10.003.

[10] ALDRIDGE, Bree B., John M. BURKE, Douglas A. LAUFFENBURGER, Peter K. SORGER. Physicochemical modelling of cell signalling pathways. *Nature Cell Biology* [online]. 2006, 8(11), 1195-1203 [cit. 2020-09-20]. ISSN 1465-7392. DOI: 10.1038/ncb1497.

[11] ELSNER, Wolfram, Torsten HEINRICH, Henning SCHWARDT. Dynamics, Complexity, Evolution, and Emergence—The Roles of Game Theory and Simulation Methods. *The Microeconomics of Complex Economies* [online]. Elsevier, 2015, 2015, s. 277-304 [cit. 2020-09-17]. ISBN 9780124115859. DOI: 10.1016/B978-0-12-411585-9.00011-7.

[12] LECCA, Paola, Angela RE, Adaoha IHEKWABA, Ivan MURA, Thanh-Phuong NGUYEN. Deterministic Differential Equations. *Computational Systems Biology* [online]. Elsevier, 2016, 2016, s. 67-98 [cit. 2020-09-27]. ISBN 9780081000953. DOI: 10.1016/B978-0-08-100095-3.00005-6.

[13] LECCA, Paola, Ian LAURENZI, Ferenc JORDAN. Deterministic chemical kinetics. *Deterministic Versus Stochastic Modelling in Biochemistry and Systems Biology* [online]. Elsevier, 2013, 2013, s. 1-34 [cit. 2020-09-27]. ISBN 9781907568626. DOI: 10.1533/9781908818218.1.

[14] GRATIE, Diana-Elena, Bogdan IANCU, Ion PETRE. ODE Analysis of Biological Systems. BERNARDO, Marco, Erik DE VINK, Alessandra DI PIERRO a Herbert WIKLICKY, ed. *Formal Methods for Dynamical Systems* [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, 2013, s. 29-62 [cit. 2020-09-21]. Lecture Notes in Computer Science. ISBN 978-3-642-38873-6. DOI: 10.1007/978-3-642-38874-3_2.

[15] FUNAHASHI, Akira, Mineo MOROHASHI, Yukiko MATSUOKA, Akiya JOURAKU, Hiroaki KITANO. CellDesigner: A Graphical Biological Network Editor and Workbench Interfacing Simulator. CHOI, Sangdun, ed. *Introduction to Systems Biology* [online]. Totowa, NJ: Humana Press, 2007, s. 422-434 [cit. 2020-12-30]. ISBN 978-1-58829-706-8. DOI: 10.1007/978-1-59745-531-2_21.

[16] LE NOVERE, Nicolas. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Research* [online]. 2006, 34(90001), D689-D691 [cit. 2020-12-30]. ISSN 0305-1048. DOI: 10.1093/nar/gkj092.

[17] HOOPS, Stefan, Sven SAHLE, Ralph GAUGES, et al. COPASI – a COmplex PAthway SImulator. *Bioinformatics* [online]. 2006, 22(24), 3067-3074 [cit. 2020-12-30]. ISSN 1367-4803. DOI: 10.1093/bioinformatics/btl485.

[18] SMITH, Robert W., Rik P. VAN ROSMALEN, Vitor A. P. MARTINS DOS SANTOS, Christian FLECK. DMPy: a Python package for automated mathematical model construction of large-scale metabolic systems. *BMC Systems Biology* [online]. 2018, 12(1) [cit. 2020-12-30]. ISSN 1752-0509. DOI: 10.1186/s12918-018-0584-8.

[19] SCHWAB, Julian D., Silke D. KÜHLWEIN, Nensi IKONOMI, Michael KÜHL, Hans A. KESTLER. Concepts in Boolean network modeling: What do they all mean? *Computational and Structural Biotechnology Journal* [online]. 2020, 18, 571-582 [cit. 2020-10-17]. ISSN 20010370. DOI: 10.1016/j.csbj.2020.03.001.

[20] ALBERT, Réka, Raina ROBEVA. Signaling Networks. *Algebraic and Discrete Mathematical Methods for Modern Biology* [online]. Elsevier, 2015, 2015, s. 65-91 [cit. 2020-10-17]. ISBN 9780128012130. DOI: 10.1016/B978-0-12-801213-0.00004-6.

[21] HELIKAR, Tomáš, Bryan KOWAL, Sean MCCLENATHAN, et al. The Cell Collective: Toward an open and collaborative approach to systems biology. *BMC Systems Biology* [online]. 2012, 6(1) [cit. 2020-12-28]. ISSN 1752-0509. ISSN 1751-7915. DOI: 10.1186/1752-0509-6-96.

[22] MÜSSEL, Christoph, Martin HOPFENSITZ, Hans A. KESTLER. BoolNet—an R package for generation, reconstruction and analysis of Boolean networks. *Bioinformatics* [online]. 2010, 26(10), 1378-1380 [cit. 2020-12-30]. ISSN 1460-2059. DOI: 10.1093/bioinformatics/btq124.

[23] ALBERT, István, Juilee THAKAR, Song LI, Ranran ZHANG, Réka ALBERT. Boolean network simulations for life scientists. *Source Code for Biology and Medicine* [online]. 2008, 3(1) [cit. 2020-12-30]. ISSN 1751-0473. DOI: 10.1186/1751-0473-3-16.

[24] CHAOUIYA, Claudine. Petri net modelling of biological networks. *Briefings in Bioinformatics* [online]. 2007, 8(4), 210-219 [cit. 2020-10-20]. ISSN 1467-5463. DOI: 10.1093/bib/bbm029.

[25] OYELADE, Jelili, Itunuoluwa ISEWON, Solomon ROTIMI, Ifeoluwa OKUNOREN. Modeling of the Glycolysis Pathway in Plasmodium falciparum using Petri Nets. *Bioinformatics and Biology Insights* [online]. 2016, 10 [cit. 2020-12-29]. ISSN 1177-9322. DOI: 10.4137/BBI.S37296.

[26] JENSEN, Kurt, Grzegorz ROZENBERG. *High-level Petri Nets: Theory and Application* [online]. 1. Springer-Verlag Berlin Heidelberg, 1991 [cit. 2020-12-30]. ISBN 978-3-642-84524-6. DOI: 10.1007/978-3-642-84524-6.

[27] MARWAN, Wolfgang, Christian ROHR, Monika HEINER. Petri Nets in Snoopy: A Unifying Framework for the Graphical Display, Computational Modelling, and Simulation of Bacterial Regulatory Networks. VAN HELDEN, Jacques, Ariane TOUSSAINT a Denis THIEFFRY, ed. *Bacterial Molecular Networks* [online]. New York, NY: Springer New York, 2012, 2012-10-28, s. 409-437 [cit. 2020-12-29]. Methods in Molecular Biology. ISBN 978-1-61779-360-8. DOI: 10.1007/978-1-61779-361-5_21.

[28] NAGASAKI, Masao, Ayumu SAITO, Euna JEONG, Chen LI, Kaname KOJIMA, Emi IKEDA, Satoru MIYANO. Cell Illustrator 4.0: A Computational Platform for Systems Biology. *In Silico Biology* [online]. 2010, 10(1,2), 5-26 [cit. 2020-12-30]. ISSN 13866338. DOI: 10.3233/ISB-2010-0415.

[29] BAARIR, Soheib, Marco BECCUTI, Davide CEROTTI, Massimiliano DE PIERRO, Susanna DONATELLI, Giuliana FRANCESCHINIS. The GreatSPN tool. *ACM SIGMETRICS Performance Evaluation Review* [online]. 2009, 36(4), 4-9 [cit. 2020-12-29]. ISSN 0163-5999. DOI: 10.1145/1530873.1530876.

[30] ELSNER, Wolfram, Torsten HEINRICH, Henning SCHWARDT. Dynamics, Complexity, Evolution, and Emergence—The Roles of Game Theory and Simulation Methods. *The Microeconomics of Complex Economies* [online]. Elsevier, 2015, 2015, s. 277-304 [cit. 2020-11-08]. ISBN 9780124115859. DOI: 10.1016/B978-0-12-411585-9.00011-7.

[31] RICKLES, Dean, Penelope HAWE, Alan SHIELL. A simple guide to chaos and complexity. *Journal of Epidemiology & Community Health* [online]. 2007, 61(11), 933-937 [cit. 2020-11-03]. ISSN 0143-005X. DOI: 10.1136/jech.2006.054254.

[32] WANG, Rui-Sheng, Assieh SAADATPOUR, Réka ALBERT. Boolean modeling in systems biology: an overview of methodology and applications. *Physical Biology* [online]. 2012, 9(5) [cit. 2020-11-05]. ISSN 1478-3967. DOI: 10.1088/1478-3975/9/5/055001.

[33] AIHARA, Kazuyuki, Jun-ichi IMURA, Tetsushi UETA, ed. *Analysis and Control of Complex Dynamical Systems* [online]. Tokyo: Springer Japan, 2015 [cit. 2020-11-08]. Mathematics for Industry. ISBN 978-4-431-55012-9. DOI: 10.1007/978-4-431-55013-6.

[34] RAMAN, Karthik, Nagasuma CHANDRA. Flux balance analysis of biological systems: applications and challenges. *Briefings in Bioinformatics* [online]. 2009, 10(4), 435-449 [cit. 2020-10-25]. ISSN 1467-5463. DOI: 10.1093/bib/bbp011.

[35] SINGH, Shalini, Areejit SAMAL, Varun GIRI, Sandeep KRISHNA, Nandula RAGHURAM, Sanjay JAIN. Flux-based classification of reactions reveals a functional bow-tie organization of complex metabolic networks. *Physical Review E* [online]. 2013, 87(5) [cit. 2020-12-29]. ISSN 1539-3755. DOI: 10.1103/PhysRevE.87.052708.

[36] MACHADO, Daniel, Rafael S. COSTA, Miguel ROCHA, Eugénio C. FERREIRA, Bruce TIDOR, Isabel ROCHA. Modeling formalisms in Systems Biology. *AMB Express* [online]. 2011, 1(1) [cit. 2020-11-21]. ISSN 2191-0855. DOI: 10.1186/2191-0855-1-45.

[37] ORTH, Jeffrey D., Ines THIELE, Bernhard Ø. PALSSON. What is flux balance analysis? *Nature Biotechnology* [online]. 2010, 28(3), 245-248 [cit. 2020-10-31]. ISSN 1087-0156. DOI: 10.1038/nbt.1614.

[38] PRICE, Nathan D., Jennifer L. REED, Bernhard Ø. PALSSON. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature Reviews Microbiology* [online]. 2004, 2(11), 886-897 [cit. 2020-11-01]. ISSN 1740-1526. DOI: 10.1038/nrmicro1023.

[39] BUJAK, Renata, Wiktoria STRUCK-LEWICKA, Michał J. MARKUSZEWSKI, Roman KALISZAN. Metabolomics for laboratory diagnostics. *Journal of Pharmaceutical and Biomedical Analysis* [online]. 2015, 113, 108-120 [cit. 2021-01-02]. ISSN 07317085. DOI: 10.1016/j.jpba.2014.12.017.

[40] NALBANTOGLU, Sinem. Metabolomics: Basic Principles and Strategies. NALBANTOGLU, Sinem, Hakima AMRI, ed. *Molecular Medicine* [online]. IntechOpen, 2019, 2019-11-6 [cit. 2021-01-03]. ISBN 978-1-83962-759-0. DOI: 10.5772/intechopen.88563.

[41] BAHARUM, Syarul N., Kamalrul A. AZIZAN. Metabolomics in Systems Biology. AIZAT, Wan Mohd, Hoe-Han GOH a Syarul Nataqain BAHARUM, ed. *Omics Applications for Systems Biology* [online]. Cham: Springer International Publishing, 2018, 2018-11-01, s. 51-68 [cit. 2021-01-03]. Advances in Experimental Medicine and Biology. ISBN 978-3-319-98757-6. DOI: 10.1007/978-3-319-98758-3_4.

[42] KRUK, Joanna, Marek DOSKOCZ, Elżbieta JODŁOWSKA, Anna ZACHARZEWSKA, Joanna ŁAKOMIEC, Kornelia CZAJA, Jacek KU-JAWSKI. NMR Techniques in Metabolomic Studies: A Quick Overview on Examples of Utilization. *Applied Magnetic Resonance* [online]. 2017, 48(1), 1-21 [cit. 2020-12-31]. ISSN 0937-9347. DOI: 10.1007/s00723-016-0846-9.

[43] EMWAS, Abdul-Hamid, Raja ROY, Ryan T. MCKAY, et al. NMR Spectroscopy for Metabolomics Research. *Metabolites* [online]. 2019, 9(7) [cit. 2020-12-31]. ISSN 2218-1989. DOI: 10.3390/metabo9070123.

[44] SEGERS, Karen, Sven DECLERCK, Debby MANGELINGS, Yvan Vander HEYDEN, Ann Van EECKHAUT. Analytical techniques for metabolomic studies: a review. *Bioanalysis* [online]. 2019, 11(24), 2297-2318 [cit. 2021-01-02]. ISSN 1757-6180. DOI: 10.4155/bio-2019-0014.

[45] EMWAS, Abdul-Hamid M. The Strengths and Weaknesses of NMR Spectroscopy and Mass Spectrometry with Particular Focus on Metabolomics Research. BJERRUM, Jacob T., ed. *Metabonomics* [online]. New York, NY: Springer New York, 2015, 2015-1-27, s. 161-193 [cit. 2020-12-31]. Methods in Molecular Biology. ISBN 978-1-4939-2376-2. DOI: 10.1007/978-1-4939-2377-9_13.

[46] COSKUN, Ozlem. Separation Techniques: CHROMATOGRAPHY. *Northern Clinics of Istanbul* [online]. 2016 [cit. 2021-01-01]. ISSN 21484902. DOI: 10.14744/nci.2016.32757.

[47] MCMASTER, Marvin C. *GC/MS* [online]. Hoboken, NJ, USA: John Wiley, 2008 [cit. 2021-01-01]. ISBN 9780470228357. DOI: 10.1002/9780470228357.

[48] LENZ, Robert W., Robert H. MARCHESSAULT. Bacterial Polyesters: Biosynthesis, Biodegradable Plastics and Biotechnology. *Biomacromolecules* [online]. 2005, 6(1), 1-8 [cit. 2020-11-22]. ISSN 1525-7797. DOI: 10.1021/bm049700c.

[49] LEMOIGNE, Maurice. *Products of dehydration and of polymerization of Polyhydroxybutyric acid.* Bull. Soc. Chem. Biol., 8 (1926), pp. 770-782

[50] WILLIAMSON, D. H., J. F. WILKINSON. The Isolation and Estimation of the Poly-hydroxy-butyrate Inclusions of Bacillus Species. *Journal of General Microbiology* [online]. 1958, 19(1), 198-209 [cit. 2020-11-22]. ISSN 0022-1287. DOI: 10.1099/00221287-19-1-198.

[51] PERNICOVA, Iva, Ivana NOVACKOVA, Petr SEDLACEK, et al. Introducing the Newly Isolated Bacterium *Aneurinibacillus* sp. H1 as an Auspicious Thermophilic Producer of Various Polyhydroxyalkanoates (PHA) Copolymers–1. Isolation and Characterization of the Bacterium. *Polymers* [online]. 2020, 12(6) [cit. 2020-11-25]. ISSN 2073-4360. DOI: 10.3390/polym12061235.

[52] FUESSL, Andreas, Motonori YAMAMOTO, Arnold SCHNELLER. Opportunities in Bio-Based Building Blocks for Polycondensates and Vinyl Polymers.*Polymer Science: A Comprehensive Reference* [online]. Elsevier, 2012, 2012, s. 49-70 [cit. 2020-12-04]. ISBN 9780080878621. DOI: 10.1016/B978-0-444-53349-4.00132-1.

[53] RAZA, Zulfiqar Ali, Shahina RIAZ, Ibrahim M. BANAT. Polyhydroxyalkanoates: Properties and chemical modification approaches for their functionalization. *Biotechnology Progress* [online]. 2018, 34(1), 29-41 [cit. 2020-12-04]. ISSN 87567938. DOI: 10.1002/btpr.2565.

[54] LI, Zibiao, Jing YANG, Xian Jun LOH. Polyhydroxyalkanoates: opening doors for a sustainable future. *NPG Asia Materials* [online]. 2016, 8(4), e265-e265 [cit. 2020-12-05]. ISSN 1884-4049. DOI: 10.1038/am.2016.48.

[55] KESKIN, Gülsah, Gülnur KIZIL, Mikhael BECHELANY, Céline POCHAT-BOHATIER, Mualla ÖNER. Potential of polyhydroxyalkanoate (PHA) polymers family as substitutes of petroleum based polymers for packaging applications and solutions brought by their composites to form barrier materials. *Pure and Applied Chemistry* [online]. 2017, 89(12), 1841-1848 [cit. 2020-12-05]. ISSN 1365-3075. DOI: 10.1515/pac-2017-0401.

[56] BASNETT, Pooja, Sujata RAVI, Ipsita ROY. Natural bacterial biodegradable medical polymers. *Science and Principles of Biodegradable and Bioresorbable Medical Polymers* [online]. Elsevier, 2017, 2017, s. 257-277 [cit. 2020-12-04]. ISBN 9780081003725. DOI: 10.1016/B978-0-08-100372-5.00008-8.

[57] CHEN, Guo-Qiang, Ivan HAJNAL, Hong WU, Li LV, Jianwen YE. Engineering Biosynthesis Mechanisms for Diversifying Polyhydroxyalkanoates. *Trends in Biotechnology* [online]. 2015, 33(10), 565-574 [cit. 2020-12-18]. ISSN 01677799. DOI: 10.1016/j.tibtech.2015.07.007.

[58] ZELDES, Benjamin M., Matthew W. KELLER, Andrew J. LODER, Christopher T. STRAUB, Michael W. W. ADAMS, Robert M. KELLY. Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals. *Frontiers in Microbiology* [online]. 2015, 6 [cit. 2020-12-05]. ISSN 1664-302X. DOI: 10.3389/fmicb.2015.01209.

[59] PANTAZAKI, Anastasia A., Maria G. TAMBAKA, Valerie LANGLOIS, Philippe GUERIN, Dimitrios A. KYRIAKIDIS. Polyhydroxyalkanoate (PHA) biosynthesis in *Thermus thermophilus*: Purification and biochemical properties of PHA synthase. *Molecular and Cellular Biochemistry* [online]. 2003, 254(1/2), 173-183 [cit. 2020-11-25]. ISSN 03008177. DOI: 10.1023/A:1027373100955.

[60] PANTAZAKI, Anastasia A., Christos P. PAPANEOPHYTOU, Agathi G. PRITSA, Maria LIAKOPOULOU-KYRIAKIDES, Dimitrios A. KYRIAKIDIS. Production of polyhydroxyalkanoates from whey by *Thermus thermophilus* HB8. *Process Biochemistry* [online]. 2009, 44(8), 847-853 [cit. 2020-11-25]. ISSN 13595113. DOI: 10.1016/j.procbio.2009.04.002.

[61] SHEU, Der-Shyan, Wen-Ming CHEN, Jr-Yung YANG, Rey-Chang CHANG. Thermophilic bacterium *Caldimonas taiwanensis* produces poly(3-hydroxybutyrate-co-3-hydroxyvalerate) from starch and valerate as carbon sources. *Enzyme and Microbial Technology* [online]. 2009, 44(5), 289-294 [cit. 2020-11-25]. ISSN 01410229. DOI: 10.1016/j.enzmictec.2009.01.004.

[62] IBRAHIM, Mohammad H. A., Anne WILLEMS, Alexander STEINBÜCHEL. Isolation and characterization of new poly(3HB)-accumulating star-shaped cell-aggregates-forming thermophilic bacteria. *Journal of Applied Microbiology* [online]. 2010, [cit. 2020-11-25]. ISSN 13645072. DOI: 10.1111/j.1365-2672.2010.04786.x.

[63] COOPER, Ronald A. Convergent Pathways of Sugar Catabolism in Bacteria. MORGAN, Michael J., ed. *Carbohydrate Metabolism in Cultured Cells* [online]. Boston, MA: Springer US, 1986, 1986, s. 461-491 [cit. 2020-12-27]. ISBN 978-1-4684-7681-1. DOI: 10.1007/978-1-4684-7679-8_12.

[64] POBLETE-CASTRO, Ignacio, Christoph WITTMANN, Pablo I. NIKEL. Biochemistry, genetics and biotechnology of glycerol utilization in *Pseudomonas* species. *Microbial Biotechnology* [online]. 2020, 13(1), 32-53 [cit. 2020-12-27]. ISSN 1751-7915. DOI: 10.1111/1751-7915.13400.

[65] ELBANNA, Khaled, Tina LÜTKE-EVERSLOH, Stefanie VAN TRAPPEN, Joris MERGAERT, Jean SWINGS, Alexander STEINBÜCHEL. *Schlegelella thermodepolymerans* gen. nov., sp. nov., a novel thermophilic bacterium that degrades poly(3-hydroxybutyrate-co-3-mercaptopropionate). *International Journal of Systematic and Evolutionary Microbiology* [online]. 2003, 53(4), 1165-1168 [cit. 2020-12-05]. ISSN 1466-5026. DOI: 10.1099/ijs.0.02562-0.

[66] KARP, Peter D., Peter E. MIDFORD, Richard BILLINGTON, et al. Pathway Tools version 23.0 update: software for pathway/genome informatics and systems biology. *Briefings in Bioinformatics* [online]. 2021, 22(1), 109-126 [cit. 2021-5-9]. ISSN 1467-5463. DOI: 10.1093/bib/bbz104.

[67] KARP, Peter D., Mario LATENDRESSE, Ron CASPI. The Pathway Tools Pathway Prediction Algorithm. *Standards in Genomic Sciences* [online]. 2011, 5(3), 424-429 [cit. 2021-5-9]. ISSN 1944-3277. DOI: 10.4056/sigs.1794338.

[68] CASPI, Ron, Richard BILLINGTON, Ingrid M KESELER, et al. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Research* [online]. 2020, 48(D1), D445-D453 [cit. 2021-5-9]. ISSN 0305-1048. DOI: 10.1093/nar/gkz862.

[69] JECKELMANN, Jean-Marc, Bernhard ERNI. Transporters of glucose and other carbohydrates in bacteria. *Pflügers Archiv - European Journal of Physiology* [online]. 2020, 472(9), 1129-1153 [cit. 2021-04-13]. ISSN 0031-6768. DOI: 10.1007/s00424-020-02379-0.

[70] ZHAO, Zhe, Mo XIAN, Min LIU, Guang ZHAO. Biochemical routes for uptake and conversion of xylose by microorganisms. *Biotechnology for Biofuels* [online]. 2020, 13(1) [cit. 2021-04-14]. ISSN 1754-6834. DOI: 10.1186/s13068-020-1662-x.

[71] WIEGMANN, Katharina, Michael HENSLER, Lars WÖHLBRAND, Marcus ULBRICH, Dietmar SCHOMBURG, Ralf RABUS, F. E. LÖFFLER. Carbohydrate Catabolism in Phaeobacter inhibens DSM 17395, a Member of the Marine Roseobacter Clade. *Applied and Environmental Microbiology* [online]. 2014, 80(15), 4725-4737 [cit. 2021-04-13]. ISSN 0099-2240. DOI: 10.1128/AEM.00719-14.

[72] GÖRKE, Boris, Jörg STÜLKE. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nature Reviews Microbiology* [online]. 2008, 6(8), 613-624 [cit. 2021-04-16]. ISSN 1740-1526. DOI: 10.1038/nrmicro1932.

[73] WAMELINK, Mirjam M. C., Eduard A. STRUYS, Cornelis JAKOBS. The biochemistry, metabolism and inherited defects of the pentose phosphate pathway: A review. *Journal of Inherited Metabolic Disease* [online]. 2008, 31(6), 703-717 [cit. 2021-5-16]. ISSN 0141-8955. DOI: 10.1007/s10545-008-1015-6.

[74] MICLET, Emeric, Véronique STOVEN, Paul A.M. MICHELS, Fred R. OP-
PERDOES, Jean-Yves LALLEMAND, Francis DUFFIEUX. NMR Spectro-
scopic Analysis of the First Two Steps of the Pentose-Phosphate Pathway Elu-
cidates the Role of 6-Phosphogluconolactonase. *Journal of Biological Chem-
istry* [online]. 2001, 276(37), 34840-34846 [cit. 2021-5-1]. ISSN 00219258. DOI:
10.1074/jbc.M105174200.

[75] ALTSCHUL, Stephen. Gapped BLAST and PSI-BLAST: a new generation
of protein database search programs. *Nucleic Acids Research* [online]. 25(17),
3389-3402 [cit. 2021-5-2]. ISSN 13624962. DOI: 10.1093/nar/25.17.3389.

[76] JIAO, Zhao, Tomoya BABA, Hirotada MORI, Kazuyuki SHIMIZU. Analysis
of metabolic and physiological responses to gnd knockout in Escherichia coli by
using C-13 tracer experiment and enzyme activity measurement. *FEMS Micro-
biology Letters* [online]. 2003, 220(2), 295-301 [cit. 2021-04-18]. ISSN 03781097.
DOI: 10.1016/S0378-1097(03)00133-2.

[77] NOOR, Elad, Eran EDEN, Ron MILO, Uri ALON. Central Carbon Metabolism
as a Minimal Biochemical Walk between Precursors for Biomass and Energy.
Molecular Cell [online]. 2010, 39(5), 809-820 [cit. 2021-5-9]. ISSN 10972765.
DOI: 10.1016/j.molcel.2010.08.031.

[78] LATENDRESSE, Mario, Markus KRUMMENACKER, Miles TRUPP, Peter
D. KARP. Construction and completion of flux balance models from pathway
databases. *Bioinformatics* [online]. 2012, 28(3), 388-396 [cit. 2021-5-7]. ISSN
1460-2059. DOI: 10.1093/bioinformatics/btr681.

[79] BENEŠ, Nikola, Luboš BRIM, Jakub KADLECAJ, Samuel PASTVA, David
ŠAFRÁNEK. AEON: Attractor Bifurcation Analysis of Parametrised Boolean
Networks. LAHIRI, Shuvendu K. and Chao WANG, ed. *Computer Aided Veri-
fication* [online]. Cham: Springer International Publishing, 2020, 2020-07-14, s.
569-581 [cit. 2021-5-14]. Lecture Notes in Computer Science. ISBN 978-3-030-
53287-1. DOI: 10.1007/978-3-030-53288-8_28.

[80] AMSTUTZ, Véronique, Nils HANIK, Julien POTT, Camila UTSUNOMIA,
Manfred ZINN. Tailored biosynthesis of polyhydroxyalkanoates in chemo-
stat cultures. *Enzymatic Polymerizations* [online]. Elsevier, 2019, 2019, s.
99-123 [cit. 2021-5-15]. Methods in Enzymology. ISBN 9780128170953. DOI:
10.1016/bs.mie.2019.08.018.

List of Abbreviations

ABC	ATP-binding cassette (transporter)
AC	adenylate cyclase
ATP	adenosine triphosphate
cAMP	cyclic adenosine monophosphate
CCR	carbon catabolite repression
CDM	cell dry mass
CE	capilar electrophoresis
CoA	coenzyme A
CRP	cAMP receptor protein
ED	Entner-Doudoroff (pathway)
EI	enzyme I
EII	enzyme II
FBA	flux balance analysis
FTIR	Fourier-transform infrared spectroscopy
GAP	glyceraldehyde 3-phosphate
GC	gas chromatography
GSPN	generalised stochastic Petri nets
HPLC	high-pressure liquid chromatography
HPr	histidine protein
idFBA	integrated dynamic flux balance analysis
iFBA	integrated flux balance analysis
LC	liquid chromatography
mcl-PHA	medium chain length polyhydroxyalkanoates
MDWNs	Markov decision Petri nets

MFS	major facilitator superfamily
MS	mass spectrometry
NAD	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
NADPH	reduced nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
ODE(s)	ordinary differential equation(s)
P(3HB)	poly(3-hydroxybutyrate)
P(3HB-co-3HV)	poly(3-hydroxybutyrate-co-3-hydroxyvalerate)
P(3HB-co-4HB)	poly(3-hydroxybutyrate-co-4-hydroxybutyrate)
PHA	polyhydroxyalkanoates
PDE(s)	partial differential equation(s)
PEP	phosphoenolpyruvate
PGDB	pathway/genome database
PPP	pentose phosphate pathway
PTS	PEP-dependent phosphotransferase system
rFBA	regulatory flux balance analysis
SBML	systems biology markup language
scl-PHA	short chain length polyhydroxyalkanoates