

**Towards a platform organism for terpenoid production –  
*in silico* analysis of metabolic networks of potential hosts  
and *in vivo* validation**

Zur Erlangung des akademischen Grades eines

**Dr. rer. nat.**

von der Fakultät Bio- und Chemieingenieurwesen  
der Technischen Universität Dortmund

genehmigte Dissertation

vorgelegt von

**Dipl.-Biol. Evamaria Gruchattka**

aus

Homberg (Efze)

Tag der mündlichen Prüfung: 27.05.2016

1. Gutachter: Prof. Dr. Oliver Kayser

2. Gutachter: Prof. Dr. Jack T. Pronk

**Dortmund 2016**



## Acknowledgements

I would like to express my appreciation and gratitude towards the people who supported and encouraged me over the last years.

Firstly, I would like to express my gratitude to my supervisor **Prof. Dr. Oliver Kayser** (Chair of Technical Biochemistry, TU Dortmund) for his support and faith in me. He encouraged me to orientate my research to the field of my scientific preference and always supported me in my scientific endeavors.

I would like to thank **Dr. Verena Schütz** for guidance, support and friendship in the first two years. She introduced me to *in silico* design and aroused and shared my enthusiasm in the field.

Further, I thank **Dr. Steffen Klamt** and **Oliver Hädicke** (MPI Magdeburg) for support with the *in silico* computations, **Profs. Eve Wurtele and Ling Li** (Iowa State University) for supplying us with the cDNAs of ATP-citrate lyase, **Prof. Jens Nielsen** (Chalmers University of Technology) for supplying us with the plasmid pSP-GM1 and **Jonas Krause** (Laboratory of Plant and Process Design, TU Dortmund) for sharing the GC/MS.

I would like to thank the Ministry of Innovation, Science and Research of the German Federal State of North Rhine-Westphalia (NRW) and TU Dortmund for awarding me with a scholarship from the **CLIB-Graduate Cluster Industrial Biotechnology** (CLIB<sup>2021</sup>).

My special thanks goes to my colleagues from the Chair of Technical Biochemistry. I enjoyed very productive discussions, mental support and very good times, especially with **Kathleen Pamplaniyil, Torsten Tobias Arndt, Friederike Degenhardt, Bastian Zirpel, Friederike Ullrich** and **Laura Kohnen**.

Further I would like to express my gratitude to **my PhD colleagues** from the faculty (BT, BMP, BVT, ATP, DYN), from the CLIB-GC, especially Robin Wördenweber and Dr. Kyle J. Lauersen (CeBiTec, Bielefeld), as well as Sarah Zimmet (Goethe Universität Frankfurt) for fruitful discussions, feedback and assistance in various phases of my PhD.

Finally, I would like to particularly thank **my friends, my family** and **Marc** for the best support I could wish for.

## Table of Contents

Abstract	V
Zusammenfassung	VI
<b>CHAPTER 1 INTRODUCTION</b>	<b>1</b>
1.1 Preface	2
1.2 Terpenoids	2
1.3 Microbial strain improvement and design	13
1.4 Microbial terpenoid production	22
1.5 Scope of the thesis	25
1.6 References	26
<b>CHAPTER 2 <i>IN SILICO</i> PROFILING OF <i>ESCHERICHIA COLI</i> AND <i>SACCHAROMYCES CEREVISIAE</i> AS TERPENOID FACTORIES</b>	<b>37</b>
2.1 Abstract	38
2.2 Background	39
2.3 Methods	42
2.3.1 Elementary modes and minimal cut sets	42
2.3.2 Metabolic networks	43
2.4 Results and discussion	44
2.4.1 Terpenoid pathway and metabolic network analysis	46
2.4.2 Target identification for metabolic engineering	59
2.4.3 Conclusions	67
2.5 Supplementary information	68
2.5.1 Metabolic networks of <i>S. cerevisiae</i>	68
2.5.2 Metabolic networks of <i>E. coli</i>	78
2.6 References	84
<b>CHAPTER 3 <i>IN VIVO</i> VALIDATION OF <i>IN SILICO</i> PREDICTED METABOLIC ENGINEERING STRATEGIES IN YEAST: DISRUPTION OF A-KETOGLUTARATE DEHYDROGENASE AND EXPRESSION OF ATP-CITRATE LYASE FOR TERPENOID PRODUCTION</b>	<b>91</b>
3.1 Abstract	92
3.2 Background	93

<b>3.3 Methods</b>	<b>96</b>
3.3.1 Strains, media and growth conditions	96
3.3.2 Plasmid construction and transformation	96
3.3.3 Protein extraction, SDS-PAGE and Western blot analysis	98
3.3.4 Yeast strain construction	99
3.3.5 2-phase cultivations	99
3.3.6 OD and cell dry weight determination	100
3.3.7 Metabolite and sugar analysis	101
3.3.8 Analysis of sesquiterpenoids in the organic layer	101
3.3.9 Sterol extraction and analysis	102
3.3.10 Fatty acid extraction, derivatization and analytics	102
<b>3.4 Results</b>	<b>104</b>
3.4.1 Sesquiterpenoid production in <i>S. cerevisiae</i> : evaluation of two-phase cultivation conditions and sesquiterpenoid spectrum	104
3.4.2 Engineering of terpenoid pathway: <i>tHMG1</i> overexpression and fusion of FPP synthase with patchoulol synthase	107
3.4.3 Choice of carbon source: impact on terpenoid production	111
3.4.4 Metabolic engineering to redirect carbon flux towards terpenoids. Disruption of <i>KGD1</i> of citric acid cycle	113
3.4.5 Heterologous pathway to synthesize cytosolic acetyl-CoA: ATP-citrate lyase	116
<b>3.5 Discussion</b>	<b>121</b>
3.5.1 Sesquiterpenoid spectrum	121
3.5.2 <i>tHMG1</i> overexpression and fusion of FPP synthase with patchoulol synthase have a small effect on sesquiterpenoid production	122
3.5.3 Carbon source strongly influences terpenoid production	123
3.5.4 <i>In vivo</i> validation and considerations of <i>in silico</i> predicted strategies	124
3.5.5 Conclusions	127
<b>3.6 Supplementary information</b>	<b>129</b>
<b>3.7 References</b>	<b>136</b>
<b>CHAPTER 4 DISCUSSION AND OUTLOOK</b>	<b>141</b>
<b>4.1 <i>In silico</i> analysis</b>	<b>142</b>
<b>4.2 <i>In vivo</i> validation</b>	<b>144</b>
<b>4.3 Conclusion and outlook</b>	<b>146</b>
<b>4.4 References</b>	<b>149</b>
<b>APPENDIX</b>	<b>151</b>
<b>List of Abbreviations</b>	<b>152</b>
<b>List of Publications</b>	<b>156</b>
<b><i>Curriculum vitae</i></b>	<b>158</b>

## Abstract

Terpenoids are mostly plant derived compounds with industrial and medicinal applications. These compounds can be provided via biotechnological production as environmental friendly alternative to chemical synthesis or plant extraction. However, productivities from microbial hosts require improvement in order to be economically competitive. Therefore, microbial hosts are compared and new metabolic engineering targets aimed at increasing terpenoid production are investigated by stoichiometric metabolic network analysis and selected strategies are validated *in vivo*.

The *in silico* analysis of the metabolic networks of *Escherichia coli* and *Saccharomyces cerevisiae* (yeast) led to several promising metabolic engineering targets with potential to increase the theoretical maximum terpenoid yield. Production of the sesquiterpenoid patchoulol in yeast was chosen for the *in vivo* validation study. A two-phase cultivation method for terpenoid capture with dodecane overlay was established and the produced spectrum of sesquiterpenoids was determined. Precursor metabolite flux towards the desired product was increased via overexpression of a truncated HMG-CoA reductase and fusion of FPP synthase with patchoulol synthase. Two *in silico* identified strategies were implemented: (i) disruption of  $\alpha$ -ketoglutarate dehydrogenase gene redirected the metabolic flux as predicted, however, the intermediate acetate was produced in high amounts instead of the desired product; (ii) expression of ATP-citrate lyase from *Arabidopsis* did not increase terpenoid production due to insufficient *in vivo* activity.

The findings of this thesis contribute to an increased knowledge about enhancing terpenoid production in both *E. coli* and *S. cerevisiae*, as well as metabolic behavior of yeast. The *in silico* stoichiometric metabolic network analysis can be used successfully as a metabolic prediction tool. This study highlights that kinetics, regulation and cultivation conditions may interfere with predictions, resulting in poor *in vivo* performance. These findings promote developments of metabolic modelling to increase the predictive power and accelerate microbial strain improvement.

## Zusammenfassung

Terpenoide sind vor allem pflanzliche Verbindungen, welche Anwendungen in der Industrie und als Arzneistoffe haben und deren biotechnologische Produktion eine umweltfreundliche Alternative zur chemischen Synthese oder der Extraktion aus Pflanzenmaterial darstellt. Die Produktivität der Zellen muss jedoch gesteigert werden, um eine konkurrenzfähige Produktionsalternative realisieren zu können. Die vorliegende Arbeit umfasst die Analyse und den Vergleich von mikrobiellen Plattformorganismen und die Identifizierung innovativer Ansatzpunkte für das *Metabolic Engineering* durch die Nutzung einer stöchiometrischen Netzwerkanalyse, sowie die Validierung ausgewählter Strategien *in vivo*, um die Terpenproduktion zu verbessern.

Die metabolische Netzwerkanalyse von *Escherichia coli* und *Saccharomyces cerevisiae* (Hefe) liefert einige vielversprechende Ansatzpunkte für beide Organismen, welche die theoretisch maximale Terpenausbeute erhöhen. Die gewählten Strategien wurden zur Validierung in Hefe umgesetzt, um den metabolischen Fluss in Richtung des Sesquiterpenoids Patchoulol umzuleiten. Dazu wurde zunächst eine Zwei-Phasen-Kultivierung mit Dodecan etabliert, das gebildete Sesquiterpenoidspektrum wurde analysiert und der metabolische Fluss, durch Überexpression einer gekürzten HMG-CoA Reduktase und Fusion von FPP-Synthase mit Patchoulolsynthase, zu dem gewünschten Produkt erhöht. Zwei *in silico* identifizierte Strategien wurden umgesetzt: (i) die Ausschaltung des  $\alpha$ -Ketoglutarat-Dehydrogenase-Gens führt zu der gewünschten Flussverlagerung, allerdings wurde statt des gewünschten Produktes das Intermediat Acetat in hohen Mengen gebildet; (ii) die Überexpression der ATP-Citrat-Lyase aus *Arabidopsis* erhöhte die Terpenproduktion auf Grund der geringen *in vivo* Aktivität nicht.

Die Ergebnisse dieser Arbeit mehren die Kenntnisse über Ansatzpunkte zur Verbesserung der Terpenproduktion in *E. coli* und *S. cerevisiae*, sowie über das Stoffwechselverhalten von Hefe. Die stöchiometrische Netzwerkanalyse kann erfolgreich als metabolisches Vorhersagewerkzeug genutzt werden. Allerdings hebt diese Arbeit hervor, dass Kinetik, Regulation und Kultivierungsbedingungen die Vorhersagen beeinträchtigen können. Die Erkenntnisse dieser Arbeit fördern zudem die Verbesserung von metabolischen Modellierungen, mit dem Ziel die Genauigkeit getroffener Vorhersagen zu erhöhen und die mikrobielle Stammentwicklung zu beschleunigen.

# **CHAPTER 1**

## **Introduction**

**Evamaria Gruchattka, Oliver Kayser**

**EG designed and wrote the chapter receiving input and supervision from OK**



## 1.1 Preface

The demand for the development of more sustainable processes is growing due to an increase in environmental awareness and consumer demand for environmentally conscious products. The field of industrial biotechnology includes the microbial production of chemicals from renewable feedstocks and is a desired, reliable and more environmentally friendly alternative to chemical synthesis or plant extraction. Terpenoids are mostly plant derived compounds with diverse industrial and medicinal applications, which could be supplied by microbial fermentation. However, productivities of microbial processes are often low leading to the necessity of improving the performance of microorganisms. Metabolic engineering is an enabling technology and includes the *in silico* analysis and design as well as the genetic modification of production hosts with the aim to increase productivity of desired products. This thesis will cover the *in silico* analysis of metabolic networks of the potential hosts *Escherichia coli* and *Saccharomyces cerevisiae*, and the *in vivo* validation of selected predicted strategies to yield increased production of a target terpenoid.

This chapter will review the topic of terpenoids from classification and diversity to biosynthesis, different production approaches and the case study of patchoulol (Chapter 1.2). This section will also discuss microbial strain improvement and design including its development, strain design via *in silico* modeling, microbial strain selection (Chapter 1.3), and microbial terpenoid production in the last 20 years (Chapter 1.4). Finally, summarizing remarks with respect to the scope of this thesis will conclude this chapter (Chapter 1.5).

## 1.2 Terpenoids

### Definition and Diversity

Terpenes are natural hydrocarbon compounds composed of isoprene units. Terpenoids, sometimes also called isoprenoids, additionally contain functional groups (Withers and Keasling 2007). However, the terms are often interchangeably used, the term 'terpenoids' will be used hereafter. Terpenoids form a large and diverse class of natural products. Tens of thousands of compounds are known to date (Pateraki, Heskes, and Hamberger 2015; Vickers et al. 2014) and the Nobel prize in medicine was awarded

for the discovery of an anti-malarial terpenoid, artemisinin, in 2015 (Callaway and Cyranoski 2015).

Terpenic metabolites are produced by all living organisms, however, their diversity and abundance is highest in plants (Rodríguez-Concepción 2014). Several terpenoids have essential physiological functions, and are also referred to as primary metabolites. Some are important for respiration as part of the respiratory chain like ubiquinone (coenzyme Q<sub>10</sub>) and menaquinone (vitamin K<sub>2</sub>) (Meganathan and Kwon 2009), while sterols such as cholesterol, ergosterol and campesterol are important membrane stabilizers or precursors for steroid hormones (Bishop and Yokota 2001; Miller 1988; Mouritsen and Zuckermann 2004). Some compounds are relevant for plant photosynthesis including the carotenoids, the terpenic side chain of chlorophylls, and plastoquinone (Cazzonelli 2011; McGarvey and Croteau 1995). These compounds are also important for plant growth and development, for example gibberellins, brassinosteroids and abscisic acid (Bishop and Yokota 2001; Razem, Baron, and Hill 2006). Other metabolites are nonessential for the organism but confer some advantage, also referred to as secondary metabolites. Those terpenoids have ecological roles in antagonistic or mutualistic interactions among organisms. They can play a role in plant defense via protecting plants against herbivores or pathogens, or attraction of pollinators and seed-dispersing animals. They can act as intra- or interspecies signaling molecules (Gershenzon and Dudareva 2007). Nevertheless, the ecological role of many terpenoids is not yet known. Several secondary terpenoid metabolites have great economic value. As illustrated in Figure 1.1, applications range from pharmaceuticals such as paclitaxel (Taxol<sup>®</sup>, an anti-cancer drug) or artemisinin (an anti-malarial drug) (Howat et al. 2014; Paddon and Keasling 2014) to flavor and fragrance compounds commonly occurring in essential oils like patchoulol (oriental, masculine), nootkatone (grapefruit), citral (lemon), menthol (peppermint, cooling), limonene (lemon-like), linalool (flowery-fresh) and eucalyptol (also called 1,8-cineol; eucalyptus, camphor) (Berger 2007; Surburg and Panten 2006). Some terpenoids are potential biofuel precursors like bisabolene and farnesene (diesel and jet fuel replacement) (George et al. 2015) or food additives like astaxanthin and lycopene - both red pigments with antioxidant function (Naguib 2000).

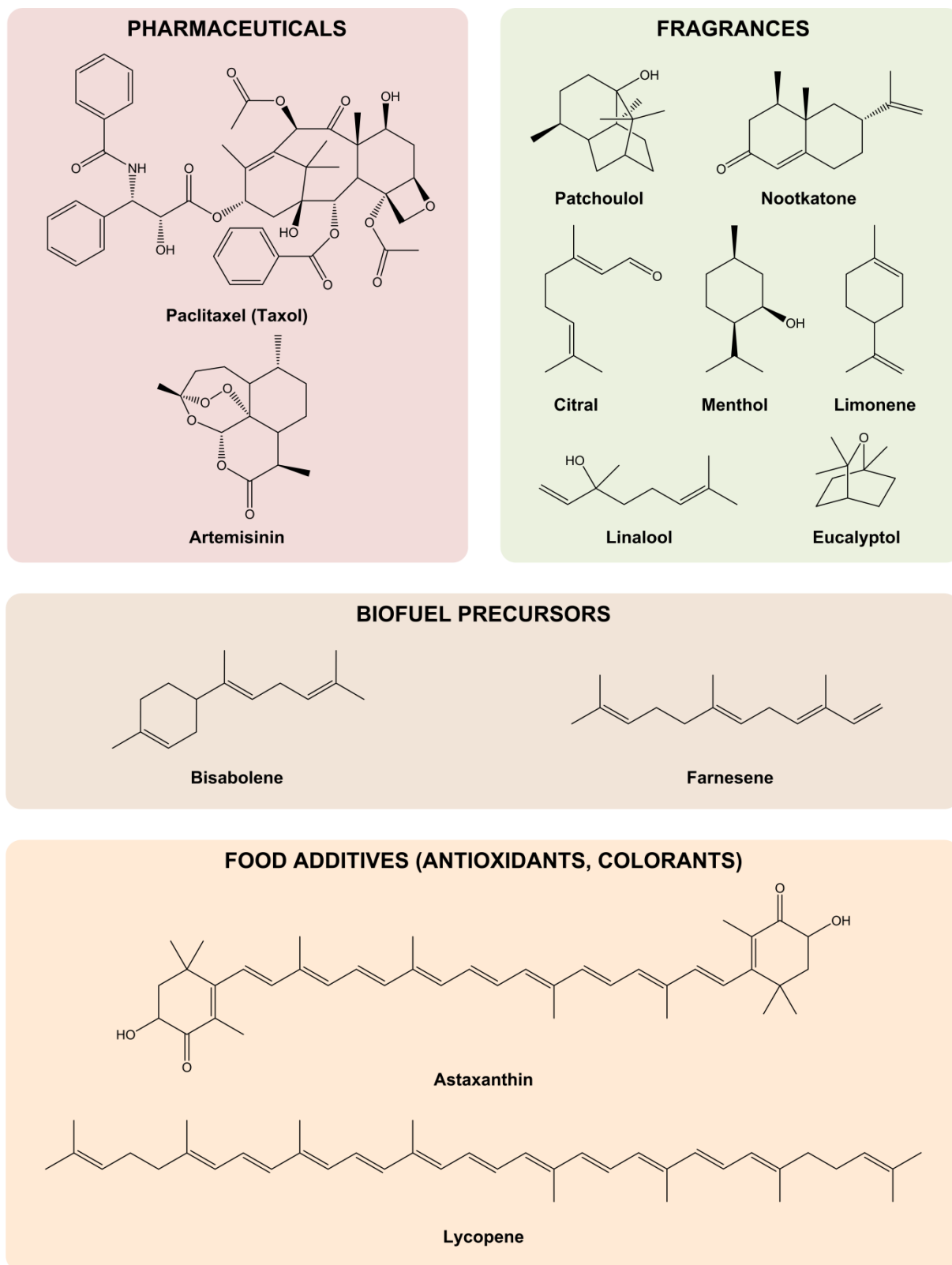


Figure 1.1 Overview of applications and structural diversity of terpenoids.

## Terpenoid Biosynthesis

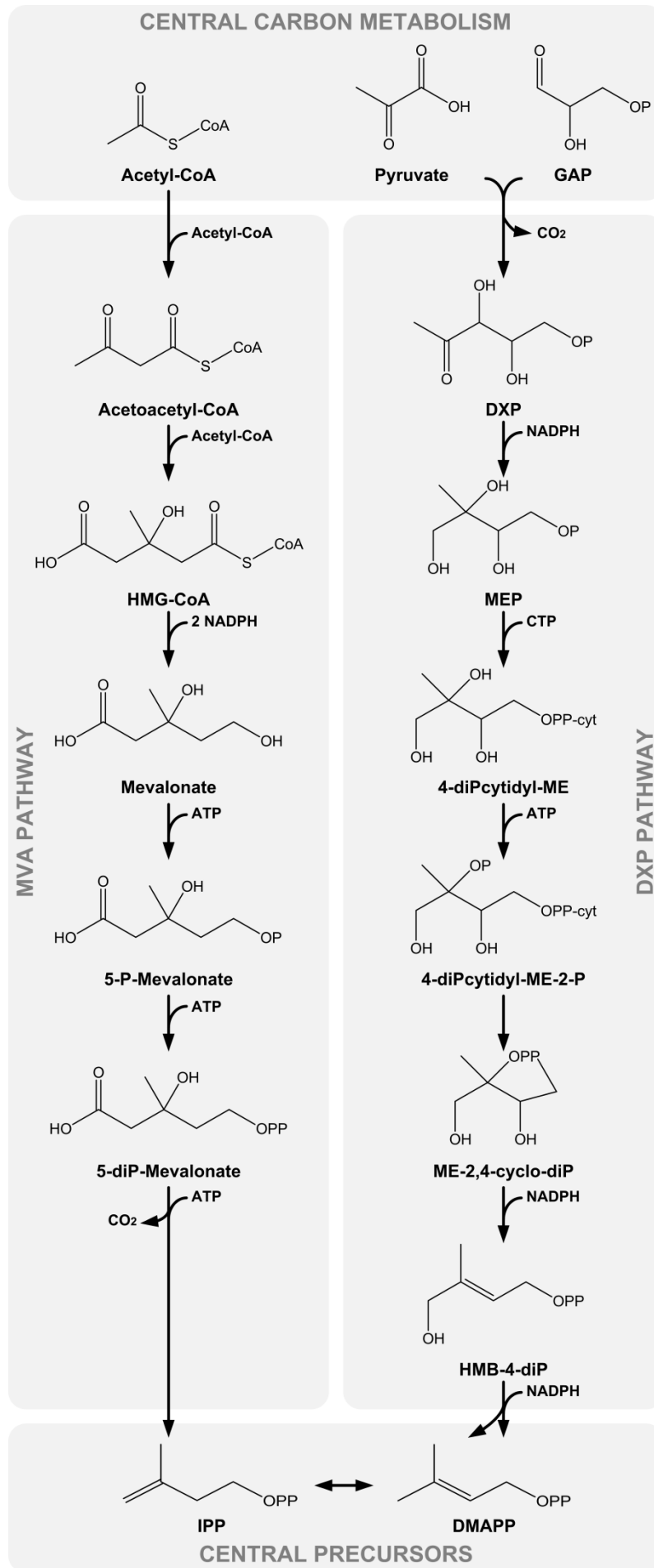
Even though terpenoids are structurally diverse, they are all derived from a common metabolic precursor, the 5-carbon (C<sub>5</sub>) isoprene unit, isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP). The central precursors IPP and DMAPP can be synthesized via two different pathways – the mevalonate (MVA) and the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway, see Figure 1.2.

Mevalonate and its role as a precursor of cholesterol was already discovered in the 1950s (Tavormina, Gibbs, and Huff 1956; Wolf et al. 1956), and the complete MVA pathway was elucidated shortly after (Bucher, Overath, and Lynen 1960; Chaykin et al. 1958; Henning, Möslein, and Lynen 1959; Rudney and Ferguson 1957) – named after the committed intermediate mevalonate. The pathway is initiated by the successive condensation of three molecules of acetyl-CoA to form acetoacetyl-CoA by the action of acetyl-CoA C-acetyltransferase and 3-hydroxy-3-methyl-glutaryl-coenzyme A (HMG-CoA) by the action of HMG-CoA synthase. HMG-CoA is then reduced by HMG-CoA reductase, which consumes two equivalents of NADPH to form mevalonate (Dickschat 2011). This reaction is the only step in the entire pathway that utilizes a redox cofactor and is described as the rate-limiting step (Brown and Goldstein 1980; Chappell et al. 1995). The enzymatic step is highly regulated at multiple levels of feedback inhibition on the level of transcription, translation, enzyme activity and protein stability, thus it represents the major site of regulation of the pathway (Maury et al. 2005). Subsequently, mevalonate is phosphorylated to 5-phosphomevalonate and 5-diphosphomevalonate by the actions of mevalonate and 5-phosphomevalonate kinase, which consumes one mol ATP per reaction, and is finally decarboxylated to IPP by the action of diphosphomevalonate decarboxylase, which also consumes one mol of ATP (Dickschat 2011). The terpenoid pathways are distributed between all living organisms. The MVA pathway is found in animals, plants, fungi, algae, archaea and some bacteria (Eisenreich, Rohdich, and Bacher 2001; Kuzuyama and Seto 2003). However, some differences occur, for example in cofactor requirements of HMG-CoA reductase (Kim, Stauffacher, and Rodwell 2000). The last two reactions of the pathway are different in archaea. 5-phosphomevalonate is converted to isopentenyl phosphate and then phosphorylated to IPP by the action of isopentenyl phosphate kinase (Grochowski, Xu, and White 2006).

The DXP pathway, often also called 2-C-methyl-D-erythritol-4-phosphate pathway (MEP), or non-mevalonate pathway, was only discovered in the 1990s via <sup>13</sup>C labelling experiments (Rodríguez-Concepción and Boronat 2002; Rohmer et al. 1996; Rohmer

1999). DXP formation is the initial step in the pathway and is generated by the condensation of pyruvate and glyceraldehyde-3-phosphate (GAP) by DXP synthase with concomitant loss of CO<sub>2</sub>, see Figure 1.2. DXP is then reduced to 2-C-methyl-D-erythritol-4-phosphate (MEP), the committed intermediate of the pathway. This reaction consumes NADPH and is mediated by DXP reductoisomerase (Takahashi et al. 1998). MEP is further converted to 4-diphosphocytidyl-2-C-methyl-D-erythritol, 4-diphosphocytidyl-2-C-methyl-D-erythritol-2-phosphate, 2-C-methyl-D-erythritol-2,4-cyclodiphosphate and 1-hydroxy-2-methyl-2-butenyl-4-diphosphate, which is then converted to IPP and DMAPP. The process consumes CTP, ATP, and two mol NADPH (Dickschat 2011; Rohdich et al. 2003; Seemann et al. 2002; Wolff et al. 2003). The DXP pathway is regulated strongly (Wolfertz et al. 2004), for example feedback inhibition of DXP synthase occurs from IPP and DMAPP accumulation (Banerjee et al. 2013). However, further details regarding the regulation are not yet well understood. The DXP pathway is found in bacteria and also in plants and algae (Eisenreich, Rohdich, and Bacher 2001; Kuzuyama and Seto 2003). Bacteria and algae use one and/or the other pathway while plants possess both pathways - the MVA is located in the cytosol, endoplasmic reticulum and peroxisomes, while the DXP pathway is located in chloroplasts, owing to the plastids origin as a prokaryotic endosymbiont. The MVA and DXP pathway are responsible for delivering precursors for different kinds of terpenoids. Though, immediate precursors like IPP can be exchanged between compartments resulting in mixed, MVA and DXP derived isoprene units in terpenoids (Eisenreich, Rohdich, and Bacher 2001; Rohmer 1999; Vickers et al. 2014).

Organisms that possess the MVA pathway require the conversion of IPP to DMAPP by an IPP isomerase (Mayer et al. 1992). However, organisms that possess the DXP pathway do not necessarily require the isomerase (Hahn, Hurlburt, and Poulter 1999). IPP and DMAPP are generated in the last step of the DXP pathway in a ratio of about 5:1 (Adam et al. 2002). Nevertheless, the isomerase performs a fine-tuning of the IPP:DMAPP ratio to the specific requirements of the downstream terpenoid metabolism (Dickschat 2011).

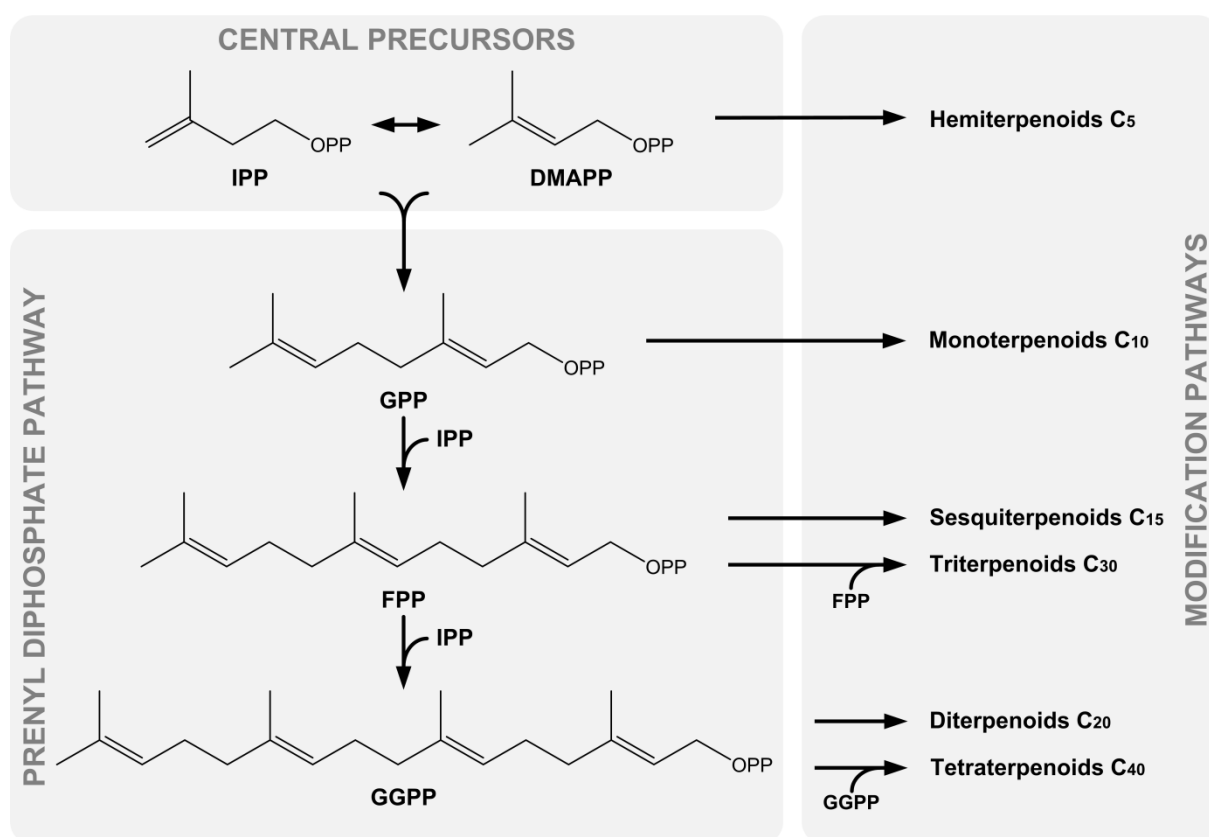


◀ **Figure 1.2 Biosynthesis of terpenoid precursors.** The mevalonate (MVA) pathway and the 1-deoxyxylulose-5-phosphate (DXP) pathway are fed by central carbon metabolites acetyl-CoA or glyceraldehyde 3-phosphate (GAP) and pyruvate, respectively. They form the central precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). HMG-CoA: 3-hydroxy-3-methyl-glutaryl-coenzyme A, MEP: 2-C-methyl-D-erythritol-4-phosphate, 4-diP-cytidyl-ME: 4-diphosphocytidyl-2-C-methyl-D-erythritol, 4-diP-cytidyl-ME-2-P: 4-diphosphocytidyl-2-C-methyl-D-erythritol-2-phosphate, ME-2,4-cyclo-diP: 2-C-methyl-D-erythritol-2,4-cyclodiphosphate, HMB-4-diP: 1-hydroxy-2-methyl-2-butenyl-4-diphosphate.

The isoprene rule states that the carbon backbone of terpenoids is composed of isoprene units (Ruzicka 1953) and terpenoids are classified according to their number of isoprene units and thus carbon atoms. Hemiterpenoids like isoprene are directly derived from DMAPP ( $C_5$ ), while the central precursors IPP and DMAPP are condensed in head-to-tail prenyltransferase reactions in the prenyl diphosphate pathway to generate the immediate prenyl diphosphate precursors of terpenoids, see Figure 1.3. IPP and DMAPP generate geranyl diphosphate (GPP), the immediate  $C_{10}$  precursor of monoterpenoids like limonene or of the terpenic moiety of cannabinoids (Fellermeier et al. 2001). GPP is extended (head-to-tail) by another molecule IPP to generate farnesyl diphosphate (FPP), the immediate  $C_{15}$  precursor of sesquiterpenoids like patchoulol. Two molecules of FPP can be condensed in a head-to-head manner to generate triterpenoids ( $C_{30}$ ) like sterols. FPP can again be extended by another molecule IPP (head-to-tail) to generate geranylgeranyl diphosphate (GGPP), the immediate  $C_{20}$  precursor of diterpenoids like taxadiene, the precursor of paclitaxel. GGPP can be extended by another molecule of GGPP in a head-to-head manner to generate tetraterpenoids ( $C_{40}$ ) like lycopene, a carotenoid. Terpenoids with more than eight isoprene units such as natural rubber are called polyterpenoids (Dickschat 2011).

The enormous diversity of chemical structures is generated by terpenoid synthases, which convert the linear prenyl diphosphates into complex, acyclic to multicyclic terpenoid scaffolds. The large number of terpenoid synthases, many of which produce multiple products, contribute to terpenoid diversity (Degenhardt, Kollner, and Gershenzon 2009; Gao, Honzatko, and Peters 2012).  $\gamma$ -humulene synthase of *Abies grandies*, for example, is able to produce 52 different sesquiterpenoids (Steele et al. 1998). Terpenoid synthases share a common carbocationic driven reaction mechanism and the ability to form multiple products could result from the aromatic residues in the active site of the enzyme, which can stabilize the highly reactive carbocationic intermediates in more than one way (Christianson 2006; Degenhardt, Kollner, and

Gershenzon 2009; Gao, Honzatko, and Peters 2012). The core terpenoids produced by terpenoid synthases can be further functionalized. A common reaction is hydroxylation by cytochrome P450 enzymes, however, cyclization reactions, rearrangements and glycosylations also occur – to produce the striking array of terpenoid compounds (Pateraki, Heskes, and Hamberger 2015; Weitzel and Simonsen 2013).



**Figure 1.3 Generation of terpenoid diversity.** The central precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) are successively condensed to generate phosphorylated terpenoid backbones with different numbers of carbon atoms. The backbones are then modified to generate the diversity of terpenoids, which are classified according to the number of carbon atoms. GPP: geranyl diphosphate, FPP: farnesyl diphosphate, GGPP: geranylgeranyl diphosphate.

### Terpenoid Production

As mentioned above, plants produce most terpenoids of human interest. Essential oils or single compounds of commercial interest are traditionally produced by steam distillation or solvent extraction of their native plant. This approach is possible if the plant is easy to grow and accumulates sufficient amounts, for example as in the case of menthol (Dai et al. 2010; Surburg and Panten 2006). However, several compounds of economic value are rare and produced in low amounts. Paclitaxel (Taxol®) for example,

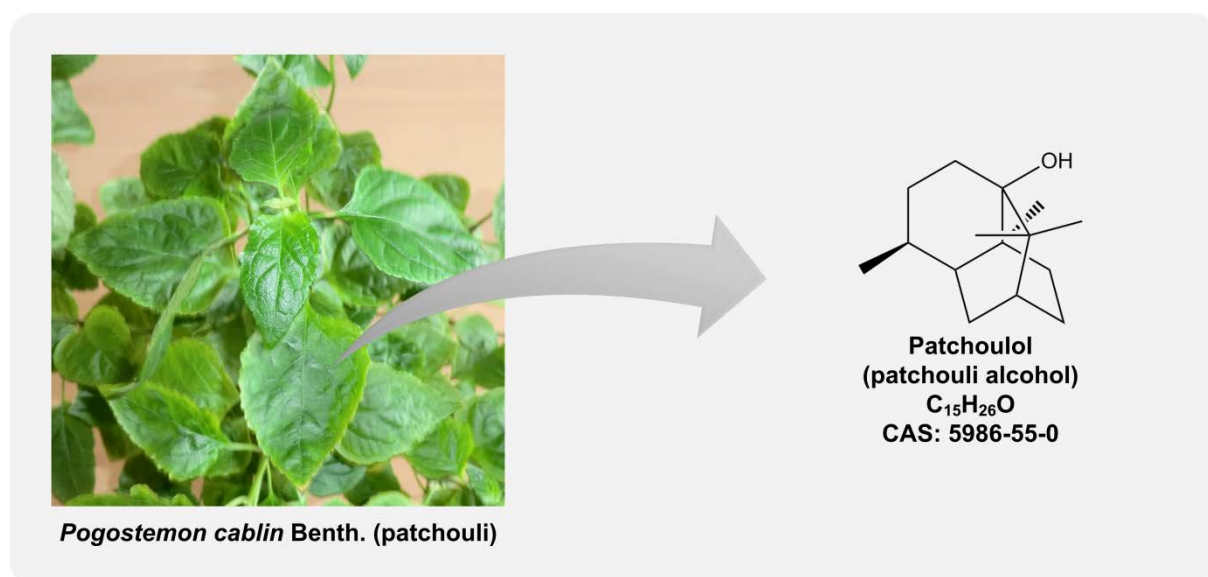


an FDA approved anticancer drug, is produced in the bark of the very slowly growing and rare Pacific yew tree (*Taxus brevifolia*, on the IUCN Red List of Threatened Species) with product yields of about 0.015 % of dry weight, ecological constraints limits its production (Cragg et al. 1993; Thomas 2013). Sandalwood oil is a valuable ingredient used in perfumery and is produced by steam distillation of dried wood of *Santalum album* L. Though, sandalwood is considered as one of the most valuable trees in the world as trees at the age of 30-50 years produce the best oil leading to a limited supply (Burdock and Carabin 2008). Moreover, geographical location and weather influence the terpenoid composition in plants (Gil et al. 2002) leading to a dependency on specific climate zones for cultivation and eventually fluctuating market prices. The price of artemisinin, a potent antimalarial drug, has varied drastically due to the effect of poor weather on crop yields (Wang 2010). Chemical synthesis might be an alternative to ensure supply. However, several terpenoids are structurally too complex so that synthetic routes suffering from low yields are too expensive. Chemical synthesis of paclitaxel, containing eleven chiral centers, has been described, however, would be too expensive with an overall yield of 0.4 % (Walji and MacMillan 2007). Additionally, another drawback is poor sustainability due to toxic byproducts or waste and dependency on nonrenewable resources (Gavrilescu and Chisti 2005). An alternative production strategy is to use living organisms. Plants can evolve to efficient producers via selective breeding and back-crossing (Ohlendorf 1996). Though, this is time-consuming and only worthwhile if the producing plant is suitable in terms of growth behavior. If this is not the case, as for example in the case of the Pacific yew tree, genetic modification of a suitable plant can be used to access genes from those species of interest. However, this approach is hampered by complex socio-political issues and regulatory requirements for genetically modified plants (Hails 2000; Masood 2003; Potrykus 2012). A solution is the use of a microbial platform. Advantages include the speed of growth and engineering as highly advanced tools are available. Microbial systems also present the possibility of large-scale and environmentally friendly industrial production in a controlled fermentation environment, which avoids the risk of pesticide residues from plant cultivation. Furthermore, such bioprocesses are independent from season, climate and cultivation risks, leading to more stable production, and thereby stable market prices. The terpenoid pathways, which generate the central precursors of all terpenoids, are also present in microorganisms. Thus, a microbial platform, which produces those precursors efficiently, could be used for the

production of a broad array of terpenoids of interest. Indeed, engineered microorganisms could also enable the production of novel or unusual terpenoids. (Maury et al. 2005; Vickers et al. 2014)

### Case Study: Patchoulol

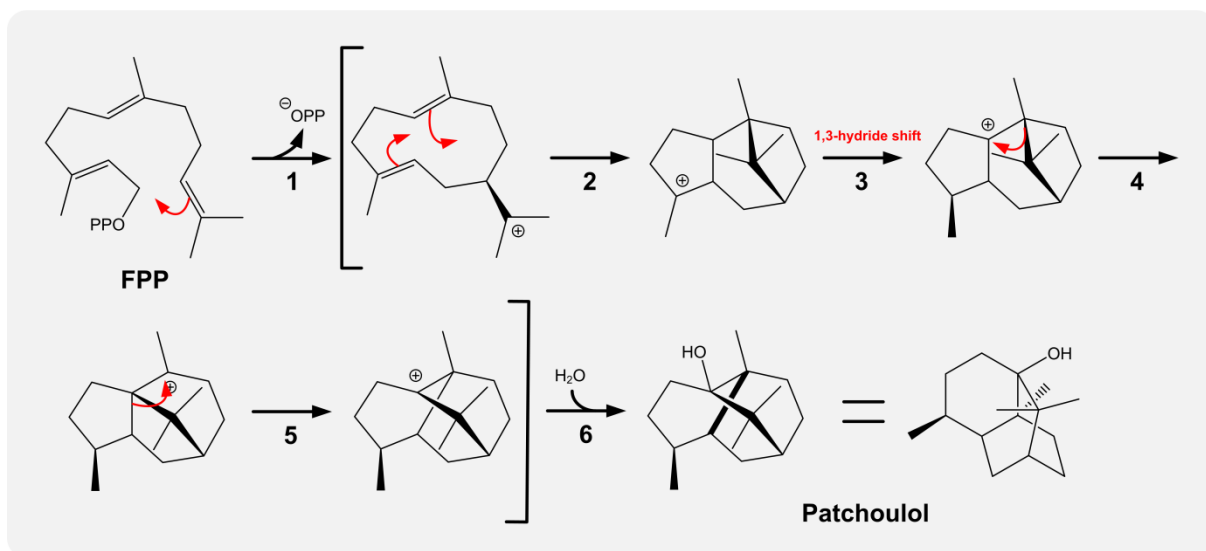
Patchoulol or patchouli alcohol (CAS: 5986-55-0) is a volatile tricyclic sesquiterpenoid naturally produced by *Pogostemon cablin* Benth. (patchouli), see Figure 1.4, a shrub belonging to the family Lamiaceae (Swamy and Sinniah 2015). As recently discovered it is also produced in *Valeriana jatamansi* Jones (Caprifoliaceae) from sub-temperature Himalaya and other *Pogostemon* species (Raina and Negi 2015; Sundaresan et al. 2009).



**Figure 1.4** *Pogostemon cablin* Benth. (patchouli) produces the volatile sesquiterpenoid patchoulol.

Patchouli is an aromatic shrub of commercial importance, native to the Philippines, however it grows wild in many South Asian countries. The plant has been used for its essential oil and in traditional medicine in India and China for a long time. Its oil accumulates in glandular trichomes of leaves, and has unique flavor and fragrance properties and a long lasting woody, earthy odor. The oil has several biological activities ranging from antioxidant, anti-inflammatory, antidepressant, antimutagenic and cytotoxic to antimicrobial and antiviral. Patchouli oil contains many ingredients, with patchoulol as the main component (~32-41 %). Several biological activities of the oil are associated with this compound. However, patchoulol and patchouli oil are mainly used

in perfumery for oriental, masculine notes in soaps, detergents, fabric conditioners, cosmetic products and perfume itself or in lower concentrations for the flavoring of beverages and food. The plant is commercially cultivated in Indonesia, India, Malaysia, China, Singapore, West Africa, and Vietnam and the oil is obtained by steam distillation of dried leaves with a yield of 2-3.5 %. About 800-1,000 t of oil are produced per year worldwide. Though, as already described as a drawback of plant production, the chemical composition of patchouli oil varies with geographic location and harvest time (Liu et al. 2015; Ramya, Palanimuthu, and Rachna 2013; Surburg and Panten 2006; Swamy and Sinniah 2015). The use of the plant as source material leads to fluctuating prices of the essential oil (McCarthy 2002). As alternative, chemical synthesis of patchoulol has been described to be possible but no synthetic route was commercially competitive (Kraft, Weymuth, and Nussbaumer 2006). The terpenoid cyclase patchoulol synthase, responsible for the conversion of FPP to patchoulol, was purified from plant material and characterized in 1990 (Munck and Croteau 1990). However, the cDNA (complementary DNA) coding the enzyme was only discovered 2006 (Deguerry et al. 2006) and a cDNA variant in 2015 (Frister et al. 2015). This discovery raised the possibility to heterologously express the patchoulol synthase and characterize the recombinant enzyme thoroughly. The proposed reaction mechanism involves a Mg-triggered departure of the diphosphate group of FPP, which generates a carbocation and a cyclization (electrophilic addition after Markovnikov), five tertiary cationic intermediates (electrophilic addition, 1,3-hydride shift, Wagner-Meerwein rearrangements) and a nucleophilic addition of water, which generates a hydroxy group. The mechanism has been validated (Faraldos et al. 2010; Munck and Croteau 1990), see Figure 1.5.



**Figure 1.5 Proposed reaction mechanism from FPP to patchoulol.** 1: Mg-triggered departure of diphosphate group, carbocation formation and cyclization (electrophilic addition after Markovnikov), 2: electrophilic addition, 3: 1,3-hydride shift, 4 and 5: Wagner-Meerwein rearrangement, 6: nucleophilic addition of water generating the hydroxy group. Modified after Faraldos and coworkers (Faraldos et al. 2010)

Furthermore, the production of patchoulol in heterologous hosts has been demonstrated to be feasible. Since then, patchoulol synthase was expressed in tobacco plants (Wu et al. 2006), *Escherichia coli* (Deguerry et al. 2006), *Saccharomyces cerevisiae* (Asadollahi et al. 2008), and the moss *Physcomitrella patens* (Zhan et al. 2014). A two-phase cultivation was established for the microbial production (Asadollahi 2008) in order to capture the volatile patchoulol and alleviate potential toxicity of the compound as lipophilic terpenoids (especially monoterpenoids) tend to integrate into membranes and influence fluidity, permeability and thus membrane function (Brennan et al. 2012). After all, the production yield of patchoulol remained comparably low.

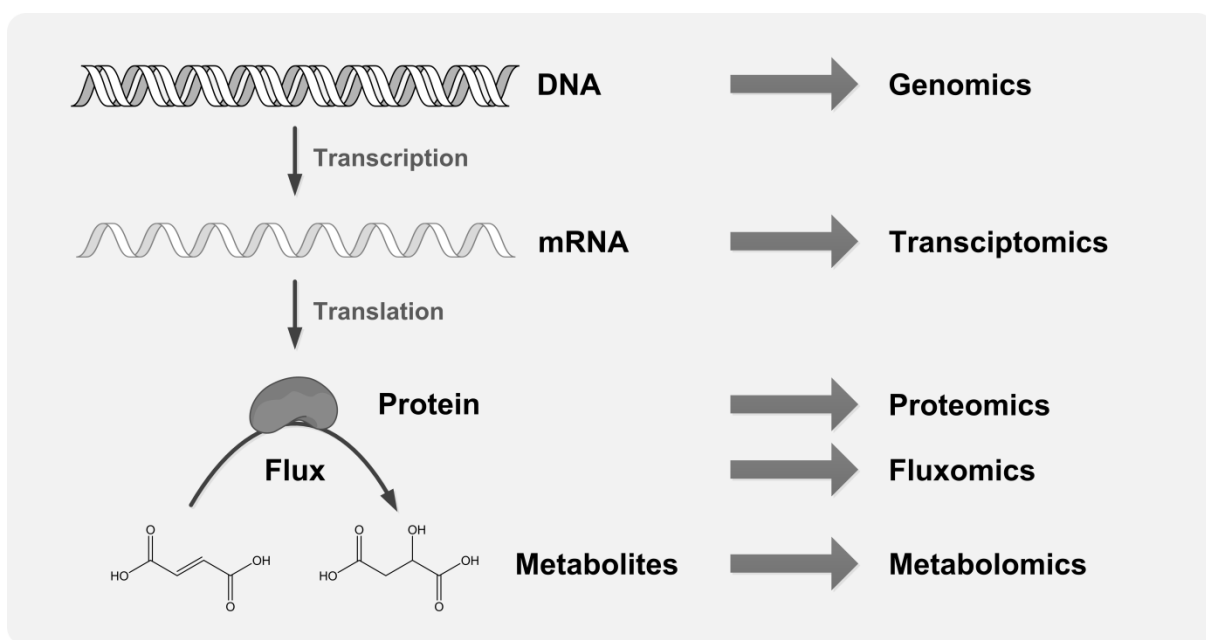
## 1.3 Microbial strain improvement and design

### Development of strain improvement

In addition to the optimization of fermentation and downstream processing, the improvement of microbial strains for the overproduction of desired products is an essential step towards commercial fermentation processes. Such optimization is important to boost inherent yields of the system, minimize production costs, and achieve competitiveness. Strain improvement includes the manipulation of microbial

strains with the aim to enhance their productivity (usually measured as titer, yield or rate). The traditional approach relies on random mutagenesis, screening and selection of superior production strains. Random mutagenesis includes the generation of mutations (DNA damage) using physical or chemical mutagens. Since the approach is non-targeted and non-specific, a very large number of strains needs to be screened and successive cycles of mutagenesis are usually required to obtain a strain with the desired phenotype. An additional approach to improve strains is genetic recombination of different strains with desired properties by protoplast fusion or sexual reproduction. This strategy leads to the creation of new gene combinations from those present in the fusion partners. These approaches have been shown to be effective and have resulted in low-cost production processes, for example, the production of penicillin (Thykaer and Nielsen 2003). An advantage of random mutational approaches is that they enable the improvement of strains without the necessity of knowing the biosynthetic pathways or genetics of the organism used. However, no or minimal mechanistic understanding is gained, which could be applied to future strain improvement, unless strains are sequenced and the information gained is used for reverse metabolic engineering. Screening is also labor-intensive unless improved via miniaturization and automation (Julleson et al. 2015; Parekh, Vinci, and Strobel 2000; Patnaik 2008; Rowlands 1984; Oud et al. 2012).

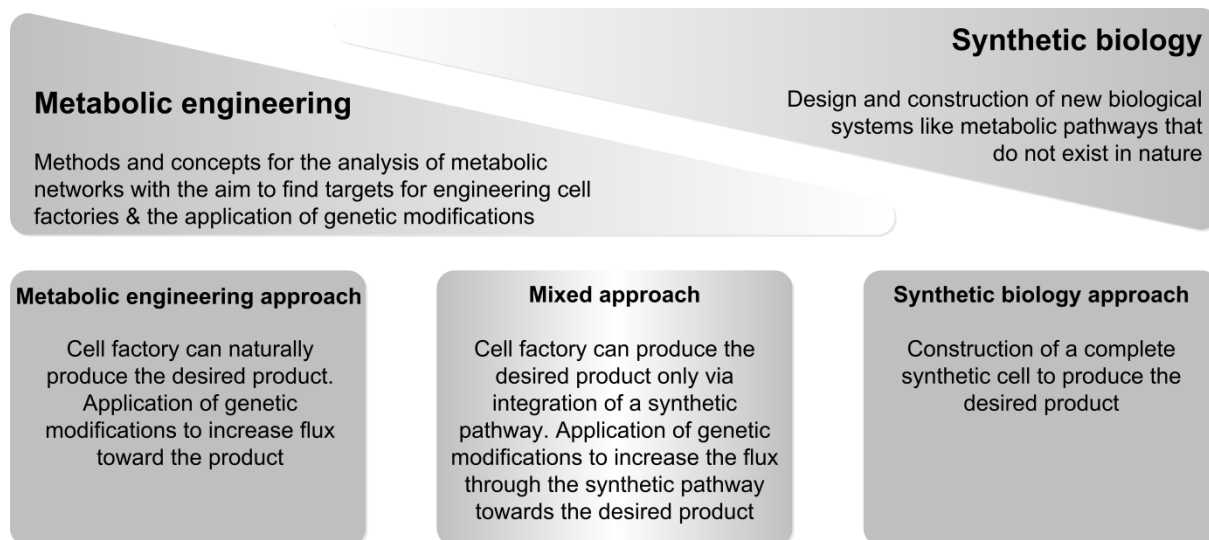
Due to advances in recombinant DNA technology in the last 20-30 years, not only can recombinant enzymes be expressed in non-native heterologous hosts, but a more targeted and rational approach towards developing microbial production strains is possible. Due to the complexity of biological systems, such approaches were initially limited to straightforward manipulations such as the elimination of side products or competing pathways. The development of high-throughput experimental techniques supported by bioinformatics has resulted in a wide range of 'omics data (see Figure 1.6) becoming available to promote rational engineering.



**Figure 1.6 Overview of different 'omics-technologies.** The central dogma of molecular biology is shown on the left. DNA is transcribed to mRNA, which in turn is translated into proteins. Enzymes (catalytic active proteins) catalyze the reactions of metabolites at a certain flux (rate or turnover of molecules). Genomics: complete genetic information of an organism; transcriptomics: complete set of mRNA transcribed from a genome at a given time and condition; proteomics: complete profile of proteins expressed in a biological system at a given time and condition; fluxomics: whole set of fluxes that are measured or calculated for a given metabolic reaction network; metabolomics: whole set of metabolites in a biological system at a given time and condition. Modified from (Lee, Lee, and Kim 2005)

The knowledge generated from large scale 'omics has provided a foundation for *in silico* modeling and simulation. These techniques can be aimed at deciphering the functional characteristics of biological systems in order to understand and predict cellular behavior. This leads also to the possibility to identify non-intuitive targets outside the biosynthetic pathway and generate new, potentially superior, production strains. Through these approaches, the field of metabolic engineering arose, in which engineering concepts are applied to the design and construction of cell factories. Metabolic engineering includes *in silico* modeling and simulation as well as the use of molecular biology tools for genetic modifications and is often employing the techniques of synthetic biology (see Figure 1.7). Strain improvement involves a cycle of design in which the successive steps seek to build on previous improvements in the host strain. The cycle begins with choice of the target for genetic engineering, followed by construction of the recombinant strain and its productivity analysis, elimination or reduction of by-product formation or expansion of the substrate range, followed by

implementation of further engineering targets until a desired characteristic is achieved (Horvat, Koller, and Braunegg 2015; Julleson et al. 2015; Lee, Lee, and Kim 2005; Nielsen 2001; Nielsen et al. 2014).



**Figure 1.7 Definitions and overlap of metabolic engineering and synthetic biology.** Modified from (Nielsen and Keasling 2011; Nielsen et al. 2014)

### Strain design via *in silico* modeling

The core concept of modeling cellular metabolism is the metabolic network, the quantitative phenotypic description of the biological system. The reconstruction of (genome-scale) metabolic networks was facilitated through genome sequencing and annotation of those sequences. Functional genomics linked gene products (enzymes) to gene functions (reaction stoichiometry) and thus enabled the development of metabolic network reconstructions. The first genome sequences were published 1995 and the number of sequenced genomes has since increased exponentially. Several databases have been developed from genomics, other 'omics and experimental data (NCBI<sup>1</sup>, KEGG<sup>2</sup>, BioCyc<sup>3</sup>, SGD<sup>4</sup>, BRENDA<sup>5</sup>, SWISS-PROT<sup>6</sup>, etc.) providing crucial information on several organisms. A metabolic network includes information on substrates and products of biochemical reactions catalyzed by enzymes. In addition, stoichiometric coefficients for each metabolite that participates in the reaction, reversibility of reactions, location of

<sup>1</sup> [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)

<sup>2</sup> [www.genome.jp/kegg](http://www.genome.jp/kegg)

<sup>3</sup> [www.biocyc.org](http://www.biocyc.org)

<sup>4</sup> [www.yeastgenome.org](http://www.yeastgenome.org)

<sup>5</sup> [www.brenda-enzymes.org](http://www.brenda-enzymes.org)

<sup>6</sup> [www.expasy.org](http://www.expasy.org)

the reaction in terms of compartmentation, and a pseudo reaction for biomass synthesis - representing the ability of the organism to produce biomass from central metabolites as well as redox- and energy-co-factors. (Feist et al. 2009; Lee, Lee, and Kim 2005; Otero and Nielsen 2010)

Stoichiometric models are based on the structure of metabolic networks, including stoichiometry of biochemical reactions, thermodynamic constraints (reaction reversibility), and the steady state assumption (internal metabolites cannot accumulate). Dynamic models are predicted to provide more accurate pictures of metabolic behavior, however, are still limited due to the lack of sufficient intracellular kinetic data. Stoichiometric modeling omitting kinetic data does, however, allow the study of important functional properties of metabolic networks and delivers verifiable predictions (Klamt, Hädicke, and von Kamp 2014; Llaneras and Picó 2008).

Several computational tools to analyze cellular metabolism have been developed and one of the most distinguished metabolic pathway analyses is the elementary mode analysis (EMA). A simplified metabolic network and the concept of EMA are depicted in Figure 1.8. The example network consists of six reactions, one reversible, five irreversible reactions, and six metabolites, three of which are internal. The stoichiometric matrix ( $N$ ) illustrates the stoichiometric coefficients of each internal metabolite in each reaction. The reaction rate flux vector ( $r$ ) comprises the reaction rates and represents the metabolic flux distribution. The thermodynamic constraint implies that irreversible reactions ( $q_{irr}$ ) have non-negative fluxes, while reversible reactions can have positive or negative fluxes. Due to the mass conservation principle and the steady state assumption, which implies that internal metabolites cannot accumulate, the problem can be formulated as follows:

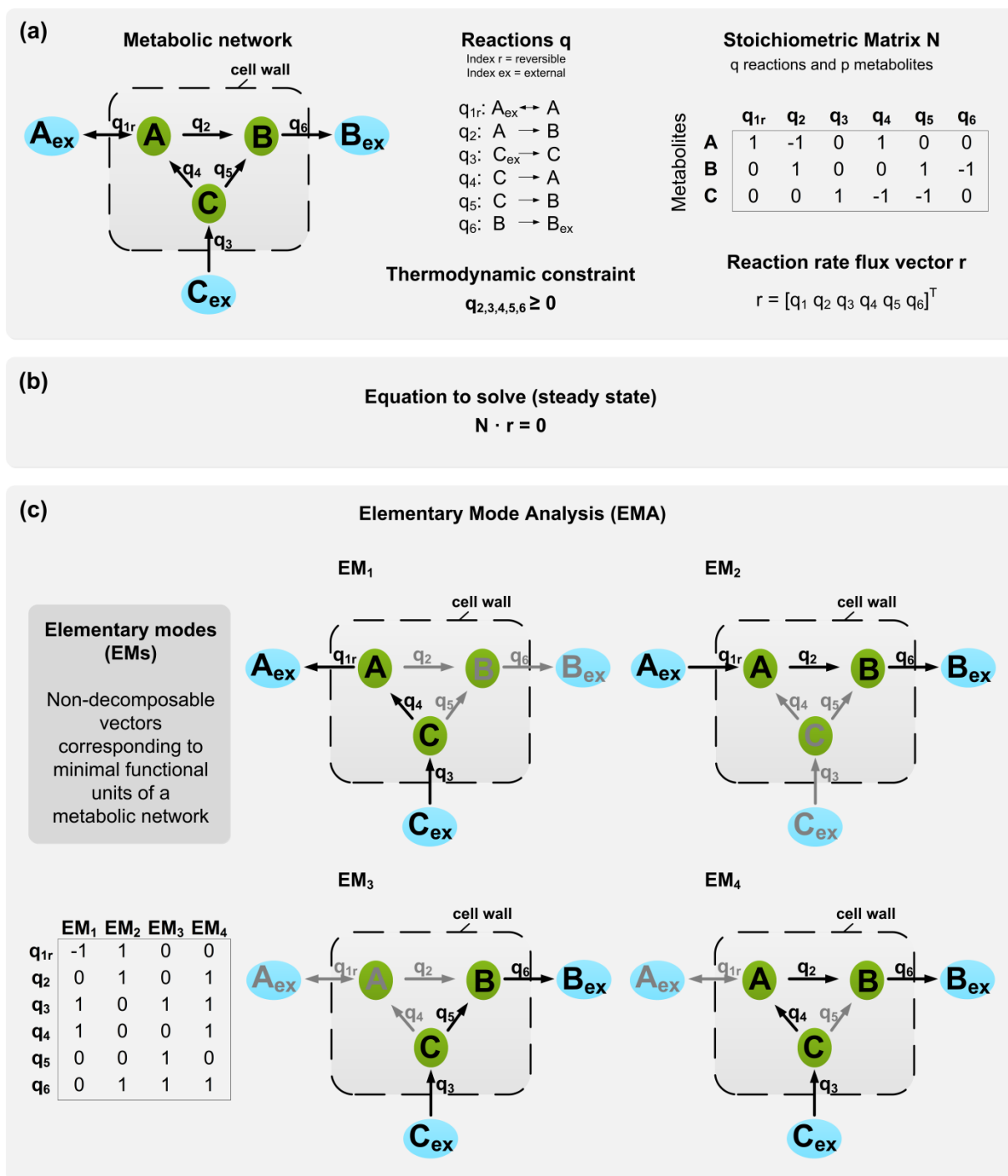
$$N \cdot r = 0$$

$$q_{irr} \geq 0$$

EMA identifies all metabolic flux vectors existing in a given metabolic network and applies an additional constraint, the non-decomposability of vectors. An elementary mode (EM) is thus a non-decomposable metabolic flux vector and corresponds to a minimal functional unit of the given metabolic network. Any real flux distribution of a cell can be represented as a positive superposition of EMs. Four EMs exist for the example metabolic network. EMs are unique, removal of any reaction will disrupt the functional unit and thus the EM. Finally, each EM and every positive superposition of EMs represents a solution to the formulated problem. The values of an elementary mode



flux vector are relative; normalized fluxes, for example, with respect to substrate flux allows the comparison of molar or C-molar yields of a metabolite in different EMs, thus the different pathways used. Therefore, theoretically efficient and inefficient pathways can be distinguished and theoretical optimal flux distributions, for example, with respect to product yield, can be determined. The effect of eliminating a reaction or of adding additional reactions (heterologous/synthetic pathways) can be analyzed. Consequently, promising metabolic engineering strategies can be identified (Horvat, Koller, and Braunegg 2015; Schuster, Dandekar, and Fell 1999; Trinh, Wlaschin, and Sreenc 2009).



**Figure 1.8 Elementary mode analysis of a simple metabolic network.** (a) A simple metabolic network is shown consisting of three internal metabolites (A, B, C) connected via internal reactions ( $q_2, q_4, q_5$ ). A steady state assumption implies that internal metabolites cannot accumulate in the system. The system is confined by dashed lines. External metabolites ( $A_{ex}, B_{ex}, C_{ex}$ ) can be produced or consumed by exchange reactions ( $q_{1r}, q_3, q_6$ ). Reaction  $q_{1r}$  is reversible and can thus have positive or negative fluxes while irreversible reactions ( $q_2, q_3, q_4, q_5, q_6$ ) have non-negative fluxes ( $q \geq 0$ ). The stoichiometric matrix  $N$  can be generated from the stoichiometric reactions  $q$ . Rows represent internal metabolites  $p$  and columns represent reactions  $q$ . The elements in the matrix represent the stoichiometric coefficients of a metabolite  $p$  in a reaction  $q$ . These coefficients are positive if a metabolite is produced and negative if a metabolite is consumed. The reaction rate flux vector  $r$  comprises the reaction rates / fluxes (defined as mol/L/h) of the reactions  $q$ . (b) Due to the principle

◀ of mass conservation of internal metabolites and the steady state assumption the equation  $N \cdot r = 0$  can be formulated and is to be solved given the thermodynamic constraint that irreversible reactions have non-negative fluxes. (c) Elementary mode analysis (EMA) can be used to solve the equation. EMA identifies all metabolic flux vectors existing in a given metabolic network. An elementary (flux) mode is a non-decomposable vector corresponding to a minimal functional unit of the network. Four elementary modes exist for the given example metabolic network. Elementary modes can be listed in a matrix; rows represent the reactions and columns represent elementary modes. The elements in the matrix represent the reaction rates / fluxes (defined as mol/L/h) of each reaction in each elementary mode (relative values). The four elementary modes are illustrated. Modified after Trinh and coworkers (Trinh, Wlaschin, and Srienc 2009).

The elimination of inefficient pathways forces the metabolism of an organism to operate within efficient pathways. Moreover efficient pathways can be coupled to cell growth/biomass formation via eliminating all pathways/EMs that enable growth without or with little product formation. Thus, the cell cannot grow without producing the desired product via efficient pathways and thus in a high yield (Trinh, Wlaschin, and Srienc 2009). A convenient approach to compute such intervention strategies was introduced by Hädicke and Klamt (Hädicke and Klamt 2011): constrained minimal cut sets (cMCS), implemented in the software package *CellNetAnalyzer* (Klamt, Saez-Rodriguez, and Gilles 2007). The cMCS approach allows computing such knockout strategies, which lead to coupling of product and biomass formation while simultaneously specifying the functionalities to be disabled or preserved. The set of target modes  $\tilde{\mathbf{T}}$  comprises those EMs to be disabled (low product yield) and the set of desired modes  $\tilde{\mathbf{D}}$  comprises those EMs to be preserved (high product yield, including modes with coupled product to biomass production). Interventions are computed with a specified minimal number of desired modes that will be preserved. Equivalent strategies to reach the same engineering goal are computed. Thus, the best combination of gene deletions in terms of practical realization can be chosen. The cMCS approach comprises and extends e.g. the approach of minimal metabolic functionality (Trinh et al. 2006), which has been applied successfully for the identification of metabolic engineering strategies (Trinh, Unrean, and Srienc 2008; Trinh and Srienc 2009; Unrean, Trinh, and Srienc 2010).

EMA was developed in the 1990s (Schuster and Hilgetag 1994; Schuster, Dandekar, and Fell 1999) and algorithms to compute EMs have been since advanced to decrease computational time and memory demand (Horvat, Koller, and Braunegg 2015; Trinh, Wlaschin, and Srienc 2009). Though, algorithms still cannot cope with genome-scale

metabolic networks as EMA computes all metabolic flux vectors, indeed, the number of EMs explodes when using more complex networks (Klamt and Stelling 2002; Klamt, Hädicke, and von Kamp 2014) - with the exception that shortest EMs or derivatives of EMs might be computed (de Figueiredo et al. 2009; Kaleta, de Figueiredo, and Schuster 2009). The analysis of medium-scale metabolic networks comprised of the central carbon metabolism (omitting less important biosynthetic pathways characterized by a naturally low metabolic flux, or incomplete pathways – with consideration that genome-scale metabolic networks are often not complete) may be sufficient, depending on the question to answer (Klamt, Hädicke, and von Kamp 2014). Indeed, EMA has been applied several times to calculate the capacity of a given metabolic network or to compute metabolic engineering strategies for the production of amino acids, ethanol, isobutanol, 1,3-propanediol, higher alcohols, biomass, lipids, muconic acid, poly- $\beta$ -hydroxybutyrate (PHB), diapolycopendioic acid, and lycopene for different microorganisms including: *Corynebacterium glutamicum*, *Escherichia coli*, *Bacillus subtilis*, *Shewanella oneidensis*, *Klebsiella pneumoniae*, *Saccharomyces cerevisiae*, *Scheffersomyces (Pichia) stipitis* and *Aspergillus niger* (Averesch and Krömer 2014; Boghigian et al. 2010; Carlson, Fell, and Srienc 2002; Carlson and Srienc 2004; Chen, Liu, and Liu 2011; Flynn et al. 2012; Hädicke and Klamt 2011; Kenanov et al. 2010; Krömer et al. 2006; Lai et al. 2012; Li et al. 2012; Matsuda et al. 2011; Melzer et al. 2009; Trinh 2012; Unrean, Trinh, and Srienc 2010; Unrean and Nguyen 2012).

### **Microbial strain selection**

The choice of a microbial host strain for the production of valuable compounds is critical. Desirable traits include robust properties such as stress tolerance, fermentation performance and substrate utilization. Moreover, accessibility of molecular techniques and a well-characterized genetic background are preferable for metabolic engineering approaches. An additional advantage is the use of organisms categorized as GRAS (generally recognized as safe, U.S. Food and Drug Administration, FDA) and accordingly QPS (qualified presumption of safety, European Food Safety Authority, EFSA) due to regulatory and environmental issues. The concept of a platform strain is based upon the use of a strain with basal desirable traits for the insertion of different product formation pathways deriving from specific common building blocks in order to reduce development time (Chen, Zhou, et al. 2015; Julleson et al. 2015). Terpenoids are derived

from IPP and DMAPP that both originate from acetyl-CoA or pyruvate and glyceraldehyde-3-phosphate. Therefore, a platform strain supplying those precursors efficiently could be used for the production of a wide array of products.

Different organisms have been analyzed for heterologous terpenoid production including but not limited to: *Corynebacterium glutamicum*, *Bacillus subtilis*, *Pseudomonas putida*, *Xanthophyllomyces dendrorhous*, *Schizophyllum commune*, *Candida utilis* and *Aspergillus nidulans* (Frohwitter et al. 2014; Guan et al. 2015; Lubertozzi and Keasling 2008; Melillo et al. 2013; Mi et al. 2014; Miura et al. 1998; Scholtmeijer et al. 2014). However, the most prominent hosts are the model organisms *Escherichia coli* and *Saccharomyces cerevisiae*. Both organisms have the advantage of inherently containing desirable traits and are well-established in commercial fermentation processes. *E. coli* is a very fast growing bacterium, still, the functional expression of membrane-bound plant cytochrome P450 enzymes along with their respective reductases, required for the functionalization of several terpenoids, is often challenging (Chang et al. 2007; Kirby and Keasling 2009). The yeast *S. cerevisiae* is not susceptible to phage infections and can withstand reduced pH and high osmotic pressure. The homologous recombination machinery of yeast can be exploited for the stable incorporation of genetic fragments into the genome (Chen, Zhou, et al. 2015; Julleson et al. 2015). These traits have led to the development of *E. coli* and *S. cerevisiae* as platform organisms for terpenoid production.

## 1.4 Microbial terpenoid production

Early studies of heterologous terpenoid production in microorganisms mostly focused on the expression of terpenoid synthases in *E. coli* for the characterization of the enzyme itself rather than terpenoid production (Cane et al. 1994; Colby et al. 1993; Mercke et al. 2000). The *in vivo* production of terpenoids was initially dominated by carotenoids in the heterologous hosts *E. coli* and *S. cerevisiae* (Chamovitz et al. 1992; Martínez-Férez et al. 1994; Yamano et al. 1994) and expanded to lower molecular weight terpenoids like sesquiterpenoids (Martin, Yoshikuni, and Keasling 2001). However, these strategies resulted in relatively low production levels and continued research focus was set on engineering the supply of precursors that are native to the terpenoid pathway in order to increase the production of desired compounds.

In *E. coli*, production of terpenoids was increased via the overexpression of prenyl diphosphate synthases and the combinatorial overexpression of probable rate limiting enzymes in order to deregulate the native DXP pathway. DXP synthase, DXP reductoisomerase, IPP isomerase and further enzymes that convert MEP to IPP and DMAPP were overexpressed (Ajikumar et al. 2010; Huang et al. 2001; Kim and Keasling 2001; Wang, Oh, and Liao 1999; Yuan et al. 2006). Another approach to circumvent the unidentified regulation of the native DXP pathway included the heterologous expression of the MVA pathway from yeast, at the same time decoupling the MVA pathway from its native control mechanisms. Heterologous expression of a MVA pathway and balancing of the pathway to circumvent HMG-CoA accumulation led to an increase in terpenoid formation (Kizer et al. 2008; Martin et al. 2003; Morrone et al. 2010; Pitera et al. 2007; Rodriguez-Villalon, Perez-Gil, and Rodriguez-Concepcion 2008; Tsuruta et al. 2009; Wang et al. 2010; Willrodt et al. 2014).

In *S. cerevisiae*, several strategies to deregulate the native MVA pathway and increase the flux towards desired terpenoids were pursued. In order to target the major site of regulation, a truncated, and thus soluble, version of HMG-CoA reductase (*tHMG1*) that is devoid of feedback inhibition by FPP was overexpressed (Jackson, Hart-Wells, and Matsuda 2003; Polakowski, Stahl, and Lang 1998; Ro et al. 2006; Verwaal et al. 2007). Moreover, a mutant transcription factor *upc2-1*, which upregulates several MVA pathway genes, has been overexpressed successfully (Dai et al. 2012; Jackson, Hart-Wells, and Matsuda 2003; Ro et al. 2006). Squalene synthase (*ERG9*) forms squalene from FPP, the general sesquiterpenoid precursor. In order to increase the flux towards sesquiterpenoids, squalene synthase was downregulated via different approaches: replacing the native promoter with a methionine repressible *MET3* promoter (Asadollahi et al. 2008; Paradise et al. 2008; Ro et al. 2006) or with the promoter of a hexose transporter that has low level expression at low glucose concentrations (*HXT1*) and thus leads to a downregulation of *ERG9* gene expression in fed-batch fermentations (Scalcinati, Knuf, et al. 2012). In order to increase the flux towards monoterpenoids, heterologous GPP synthases or a FPP synthase mutant (*erg20-2*) generating GPP were expressed because yeast does not possess a native GPP synthase (Oswald et al. 2007). In order to increase terpenoid formation and generate stable strains, all genes of the MVA pathway were overexpressed via stable integration into the genome (Westfall et al. 2012) and a focus was set on pathway balancing (Rodriguez et al. 2014).

More recent studies have focused on modifications within the central carbon metabolism of *S. cerevisiae* to redirect the metabolic flux towards an increased cytosolic acetyl-CoA production, the precursor of the MVA pathway for terpenoid production or other acetyl-CoA derived compounds, including lipids, polyketides, polyhydroxyalkanoates or *n*-butanol (Krivoruchko et al. 2014). The pyruvate dehydrogenase bypass was engineered (*ALD6* and a mutant *acs* from *Salmonella*) (Shiba et al. 2007), ammonium assimilation was modified (*gdh1Δ* and *GDH2*) (Asadollahi et al. 2009; Scalcinati, Partow, et al. 2012), and engineering of the pyruvate dehydrogenase bypass (*ALD6*, a mutant *acs* from *Salmonella* and *ADH2*) was combined with deletions in the glyoxylate cycle (*cit2Δ* and *mIs1Δ*) for terpenoid, biobutanol, polyhydroxybutyrate, 3-hydroxypropionic acid and fatty acid production (Chen et al. 2013; Chen et al. 2014; de Jong et al. 2014; Kocharin et al. 2012; Krivoruchko et al. 2013). Furthermore, the phosphoketolase pathway was expressed for fatty acid and polyhydroxybutyrate production (de Jong et al. 2014; Kocharin, Siewers, and Nielsen 2013; Papini et al. 2012), a bacterial pyruvate dehydrogenase complex was expressed functionally in the cytosol (Kozak et al. 2014), and the acetylating acetaldehyde dehydrogenase and pyruvate-formate lyase were expressed functionally (Kozak et al. 2013). Additionally, a model for an indirect acetyl-CoA transport from mitochondria to the cytosol, including acetate:succinyl-CoA transferase (*Ach1*) was proposed (Chen, Zhang, et al. 2015).

In *E. coli*, little effort has been made towards the engineering of the central carbon metabolism to redirect the metabolic flux towards the precursors of the terpenoid pathways - possibly due to the strong regulations of the native terpenoid pathway. Only few studies have been performed all focusing on carotenoid production (Alper, Miyaoku, and Stephanopoulos 2005; Alper et al. 2005; Choi et al. 2010; Farmer and Liao 2001; Jin and Stephanopoulos 2007; Vadali et al. 2005; Zhao et al. 2013).

## 1.5 Scope of the thesis

The biotechnological production of terpenoids of medicinal or industrial interest could be an environmentally friendly and reliable alternative to chemical synthesis or plant extraction. However, productivity needs to be increased to be economically competitive. Only few studies have focused on the improvement of strains based on modifications in the central carbon metabolism. Rational strain design based on elementary mode analysis has not been performed for the two model organisms *Escherichia coli* and *Saccharomyces cerevisiae*. Hence, the aim of this thesis is:

- I. To compare the two most prominently used heterologous hosts regarding their metabolic network capability to support efficient terpenoid production. In addition, identify new promising metabolic engineering targets in the central carbon metabolism for increased terpenoid production using elementary mode analysis and constrained minimal cut sets (Chapter 2)
- II. To establish terpenoid production in a promising host and deregulate the terpenoid pathway. The production strain will be used for *in vivo* validation of promising metabolic engineering targets in the central carbon metabolism in order to redirect the metabolic flux towards terpenoids (Chapter 3)
- III. *In vivo* validation of promising metabolic engineering strategies in order to evaluate the predictive power of the *in silico* analysis with all possibilities and limitations. This is used to identify possible considerations in order to improve *in silico* stoichiometric metabolic network predictions and reliably increase productivity *in vivo* (Chapter 3 and 4)



## 1.6 References

- Adam, P., S. Hecht, W. Eisenreich, J. Kaiser, T. Grawert, D. Arigoni, A. Bacher, and F. Rohdich. 2002. Biosynthesis of terpenes: studies on 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase. *Proc Natl Acad Sci U S A* 99 (19):12108-13.
- Ajikumar, P. K., W. H. Xiao, K. E. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer, and G. Stephanopoulos. 2010. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* 330 (6000):70-4.
- Alper, H., Y. S. Jin, J. F. Moxley, and G. Stephanopoulos. 2005. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metabolic engineering* 7 (3):155-64.
- Alper, H., K. Miyaoku, and G. Stephanopoulos. 2005. Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nat Biotechnol* 23 (5):612-6.
- Asadollahi, M. A. 2008. Establishment of yeast platform for isoprenoid production. Ph.D. Thesis, Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark.
- Asadollahi, M. A., J. Maury, K. Moller, K. F. Nielsen, M. Schalk, A. Clark, and J. Nielsen. 2008. Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: effect of *ERG9* repression on sesquiterpene biosynthesis. *Biotechnol Bioeng* 99 (3):666-77.
- Asadollahi, M. A., J. Maury, K. R. Patil, M. Schalk, A. Clark, and J. Nielsen. 2009. Enhancing sesquiterpene production in *Saccharomyces cerevisiae* through *in silico* driven metabolic engineering. *Metab Eng* 11 (6):328-34.
- Averesch, Nils J. H., and Jens O. Krömer. 2014. Tailoring strain construction strategies for muonic acid production in *S. cerevisiae* and *E. coli*. *Metabolic Engineering Communications* 1 (0):19-28.
- Banerjee, A., Y. Wu, R. Banerjee, Y. Li, H. Yan, and T. D. Sharkey. 2013. Feedback inhibition of deoxy-D-xylulose-5-phosphate synthase regulates the methylerythritol 4-phosphate pathway. *J Biol Chem* 288 (23):16926-36.
- Berger, Ralf Günter. 2007. *Flavours and fragrances: chemistry, bioprocessing and sustainability*: Springer Science & Business Media.
- Bishop, G. J., and T. Yokota. 2001. Plants steroid hormones, brassinosteroids: current highlights of molecular aspects on their synthesis/metabolism, transport, perception and response. *Plant Cell Physiol* 42 (2):114-20.
- Boghigian, B. A., H. Shi, K. Lee, and B. A. Pfeifer. 2010. Utilizing elementary mode analysis, pathway thermodynamics, and a genetic algorithm for metabolic flux determination and optimal metabolic network design. *BMC Syst Biol* 4:49.
- Brennan, T. C., C. D. Turner, J. O. Krömer, and L. K. Nielsen. 2012. Alleviating monoterpene toxicity using a two-phase extractive fermentation for the bioproduction of jet fuel mixtures in *Saccharomyces cerevisiae*. *Biotechnol Bioeng*.
- Brown, M. S., and J. L. Goldstein. 1980. Multivalent feedback regulation of HMG CoA reductase, a control mechanism coordinating isoprenoid synthesis and cell growth. *J Lipid Res* 21 (5):505-17.
- Bucher, Nancy L. R., Peter Overath, and Feodor Lynen. 1960.  $\beta$ -hydroxy- $\beta$ -methylglutaryl coenzyme a reductase, cleavage and condensing enzymes in relation to cholesterol formation in rat liver. *Biochimica et Biophysica Acta* 40:491-501.
- Burdock, G. A., and I. G. Carabin. 2008. Safety assessment of sandalwood oil (*Santalum album* L.). *Food Chem Toxicol* 46 (2):421-32.

- Callaway, E., and D. Cyranoski. 2015. Anti-parasite drugs sweep Nobel prize in medicine 2015. *Nature* 526 (7572):174-5.
- Cane, D. E., J. K. Sohng, C. R. Lamberson, S. M. Rudnicki, Z. Wu, M. D. Lloyd, J. S. Oliver, and B. R. Hubbard. 1994. Pentalenene synthase. Purification, molecular cloning, sequencing, and high-level expression in *Escherichia coli* of a terpenoid cyclase from *Streptomyces* UC5319. *Biochemistry* 33 (19):5846-57.
- Carlson, R., D. Fell, and F. Sreenc. 2002. Metabolic pathway analysis of a recombinant yeast for rational strain development. *Biotechnol Bioeng* 79 (2):121-34.
- Carlson, R., and F. Sreenc. 2004. Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: identification of reactions. *Biotechnol Bioeng* 85 (1):1-19.
- Cazzonelli, Christopher I. 2011. Goldacre review: carotenoids in nature: insights from plants and beyond. *Functional Plant Biology* 38 (11):833-847.
- Chamovitz, Daniel, Norihiko Misawa, Gerhard Sandmann, and Joseph Hirschberg. 1992. Molecular cloning and expression in *Escherichia coli* of a cyanobacterial gene coding for phytoene synthase, a carotenoid biosynthesis enzyme. *FEBS Letters* 296 (3):305-310.
- Chang, M. C., R. A. Eachus, W. Trieu, D. K. Ro, and J. D. Keasling. 2007. Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat Chem Biol* 3 (5):274-7.
- Chappell, J., F. Wolf, J. Proulx, R. Cuellar, and C. Saunders. 1995. Is the Reaction Catalyzed by 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase a Rate-Limiting Step for Isoprenoid Biosynthesis in Plants? *Plant Physiol* 109 (4):1337-1343.
- Chaykin, S., J. Law, A. H. Phillips, T. T. Tchen, and K. Bloch. 1958. Phosphorylated intermediates in the synthesis of squalene. *Proc Natl Acad Sci U S A* 44 (10):998-1004.
- Chen, Y., J. Bao, I. K. Kim, V. Siewers, and J. Nielsen. 2014. Coupled incremental precursor and co-factor supply improves 3-hydroxypropionic acid production in *Saccharomyces cerevisiae*. *Metabolic engineering* 22:104-109.
- Chen, Y., L. Daviet, M. Schalk, V. Siewers, and J. Nielsen. 2013. Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. *Metabolic engineering* 15:48-54.
- Chen, Y., Y. Zhang, V. Siewers, and J. Nielsen. 2015. *Ach1* is involved in shuttling mitochondrial acetyl units for cytosolic C2 provision in *Saccharomyces cerevisiae* lacking pyruvate decarboxylase. *FEMS Yeast Res* 15 (3).
- Chen, Y., Y. J. Zhou, V. Siewers, and J. Nielsen. 2015. Enabling technologies to advance microbial isoprenoid production. *Adv Biochem Eng Biotechnol* 148:143-60.
- Chen, Zhen, Hongjuan Liu, and Dehua Liu. 2011. Metabolic pathway analysis of 1, 3-propanediol production with a genetically modified *Klebsiella pneumoniae* by overexpressing an endogenous NADPH-dependent alcohol dehydrogenase. *Biochemical Engineering Journal* 54 (3):151-157.
- Choi, H. S., S. Y. Lee, T. Y. Kim, and H. M. Woo. 2010. *In silico* identification of gene amplification targets for improvement of lycopene production. *Appl Environ Microbiol* 76 (10):3097-105.
- Christianson, D. W. 2006. Structural biology and chemistry of the terpenoid cyclases. *Chem Rev* 106 (8):3412-42.
- Colby, S. M., W. R. Alonso, E. J. Katahira, D. J. McGarvey, and R. Croteau. 1993. 4S-limonene synthase from the oil glands of spearmint (*Mentha spicata*). cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpene cyclase. *J Biol Chem* 268 (31):23016-24.

- Cragg, Gordon M., Saul A. Schepartz, Matthew Suffness, and Michael R. Grever. 1993. The Taxol Supply Crisis. New NCI Policies for Handling the Large-Scale Production of Novel Natural Product Anticancer and Anti-HIV Agents. *Journal of Natural Products* 56 (10):1657-1668.
- Dai, Jianming, Valerie Orsat, G. S. Vijaya Raghavan, and Varoujan Yaylayan. 2010. Investigation of various factors for the extraction of peppermint (*Mentha piperita* L.) leaves. *Journal of Food Engineering* 96 (4):540-543.
- Dai, Z., Y. Liu, L. Huang, and X. Zhang. 2012. Production of miltiradiene by metabolically engineered *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 109 (11):2845-53.
- de Figueiredo, L. F., A. Podhorski, A. Rubio, C. Kaleta, J. E. Beasley, S. Schuster, and F. J. Planes. 2009. Computing the shortest elementary flux modes in genome-scale metabolic networks. *Bioinformatics* 25 (23):3158-65.
- de Jong, B. W., S. Shi, V. Siewers, and J. Nielsen. 2014. Improved production of fatty acid ethyl esters in *Saccharomyces cerevisiae* through up-regulation of the ethanol degradation pathway and expression of the heterologous phosphoketolase pathway. *Microb Cell Fact* 13 (1):39.
- Degenhardt, J., T. G. Kollner, and J. Gershenzon. 2009. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* 70 (15-16):1621-37.
- Deguerry, F., L. Pastore, S. Wu, A. Clark, J. Chappell, and M. Schalk. 2006. The diverse sesquiterpene profile of patchouli, *Pogostemon cablin*, is correlated with a limited number of sesquiterpene synthases. *Arch Biochem Biophys* 454 (2):123-36.
- Dickschat, J. S. 2011. Isoprenoids in three-dimensional space: the stereochemistry of terpene biosynthesis. *Nat Prod Rep* 28 (12):1917-36.
- Eisenreich, Wolfgang, Felix Rohdich, and Adelbert Bacher. 2001. Deoxyxylulose phosphate pathway to terpenoids. *Trends in plant science* 6 (2):78-84.
- Faraldos, J. A., S. Wu, J. Chappell, and R. M. Coates. 2010. Doubly deuterium-labeled patchouli alcohol from cyclization of singly labeled [2-(2)H(1)]farnesyl diphosphate catalyzed by recombinant patchoulol synthase. *J Am Chem Soc* 132 (9):2998-3008.
- Farmer, W. R., and J. C. Liao. 2001. Precursor balancing for metabolic engineering of lycopene production in *Escherichia coli*. *Biotechnol Prog* 17 (1):57-61.
- Feist, Adam M., Markus J. Herrgard, Ines Thiele, Jennie L. Reed, and Bernhard O. Palsson. 2009. Reconstruction of biochemical networks in microorganisms. *Nat Rev Micro* 7 (2):129-143.
- Fellermeier, M., W. Eisenreich, A. Bacher, and M. H. Zenk. 2001. Biosynthesis of cannabinoids. Incorporation experiments with (13)C-labeled glucoses. *Eur J Biochem* 268 (6):1596-604.
- Flynn, C. M., K. A. Hunt, J. A. Gralnick, and F. Sreenc. 2012. Construction and elementary mode analysis of a metabolic model for *Shewanella oneidensis* MR-1. *Biosystems* 107 (2):120-8.
- Frister, T., S. Hartwig, S. Alemdar, K. Schnatz, L. Thons, T. Scheper, and S. Beutel. 2015. Characterisation of a Recombinant Patchoulol Synthase Variant for Biocatalytic Production of Terpenes. *Appl Biochem Biotechnol* 176 (8):2185-201.
- Frohwitter, J., S. A. Heider, P. Peters-Wendisch, J. Beekwilder, and V. F. Wendisch. 2014. Production of the sesquiterpene (+)-valencene by metabolically engineered *Corynebacterium glutamicum*. *J Biotechnol* 191:205-213.
- Gao, Y., R. B. Honzatko, and R. J. Peters. 2012. Terpene synthase structures: a so far incomplete view of complex catalysis. *Nat Prod Rep* 29 (10):1153-75.

- Gavrilescu, Maria, and Yusuf Chisti. 2005. Biotechnology - a sustainable alternative for chemical industry. *Biotechnology Advances* 23 (7–8):471-499.
- George, K. W., J. Alonso-Gutierrez, J. D. Keasling, and T. S. Lee. 2015. Isoprenoid Drugs, Biofuels, and Chemicals-Artemisinin, Farnesene, and Beyond. *Adv Biochem Eng Biotechnol* 148:355-389.
- Gershenson, J., and N. Dudareva. 2007. The function of terpene natural products in the natural world. *Nat Chem Biol* 3 (7):408-14.
- Gil, Alejandra, Elba B. de la Fuente, Adriana E. Lenardis, Mónica López Pereira, Susana A. Suárez, Arnaldo Bandoni, Catalina van Baren, Paola Di Leo Lira, and Claudio M. Ghersa. 2002. Coriander Essential Oil Composition from Two Genotypes Grown in Different Environmental Conditions. *Journal of Agricultural and Food Chemistry* 50 (10):2870-2877.
- Grochowski, L. L., H. Xu, and R. H. White. 2006. *Methanocaldococcus jannaschii* uses a modified mevalonate pathway for biosynthesis of isopentenyl diphosphate. *J Bacteriol* 188 (9):3192-8.
- Guan, Z., D. Xue, Abdallah, II, L. Dijkshoorn, R. Setroikromo, G. Lv, and W. J. Quax. 2015. Metabolic engineering of *Bacillus subtilis* for terpenoid production. *Appl Microbiol Biotechnol*.
- Hädicke, O., and S. Klamt. 2011. Computing complex metabolic intervention strategies using constrained minimal cut sets. *Metab Eng* 13 (2):204-13.
- Hahn, F. M., A. P. Hurlburt, and C. D. Poulter. 1999. *Escherichia coli* open reading frame 696 is *idi*, a nonessential gene encoding isopentenyl diphosphate isomerase. *J Bacteriol* 181 (15):4499-504.
- Hails, R. S. 2000. Genetically modified plants - the debate continues. *Trends Ecol Evol* 15 (1):14-18.
- Henning, U., E. M. Möslein, and F. Lynen. 1959. Biosynthesis of terpenes. V. Formation of 5-pyrophosphomevalonic acid by phosphomevalonic kinase. *Archives of Biochemistry and Biophysics* 83 (1):259-267.
- Horvat, P., M. Koller, and G. Brauneegg. 2015. Recent advances in elementary flux modes and yield space analysis as useful tools in metabolic network studies. *World J Microbiol Biotechnol*.
- Howat, S., B. Park, I. S. Oh, Y. W. Jin, E. K. Lee, and G. J. Loake. 2014. Paclitaxel: biosynthesis, production and future prospects. *N Biotechnol* 31 (3):242-5.
- Huang, Q., C. A. Roessner, R. Croteau, and A. I. Scott. 2001. Engineering *Escherichia coli* for the synthesis of taxadiene, a key intermediate in the biosynthesis of taxol. *Bioorg Med Chem* 9 (9):2237-42.
- Jackson, B. E., E. A. Hart-Wells, and S. P. Matsuda. 2003. Metabolic engineering to produce sesquiterpenes in yeast. *Org Lett* 5 (10):1629-32.
- Jin, Y. S., and G. Stephanopoulos. 2007. Multi-dimensional gene target search for improving lycopene biosynthesis in *Escherichia coli*. *Metabolic engineering* 9 (4):337-47.
- Julleson, D., F. David, B. Pflieger, and J. Nielsen. 2015. Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. *Biotechnol Adv.*
- Kaleta, C., L. F. de Figueiredo, and S. Schuster. 2009. Can the whole be less than the sum of its parts? Pathway analysis in genome-scale metabolic networks using elementary flux patterns. *Genome Res* 19 (10):1872-83.
- Kenanov, D., C. Kaleta, A. Petzold, C. Hoischen, S. Diekmann, R. A. Siddiqui, and S. Schuster. 2010. Theoretical study of lipid biosynthesis in wild-type *Escherichia coli* and in a protoplast-type L-form using elementary flux mode analysis. *FEBS J* 277 (4):1023-34.

- Kim, D. Y., C. V. Stauffacher, and V. W. Rodwell. 2000. Dual coenzyme specificity of *Archaeoglobus fulgidus* HMG-CoA reductase. *Protein Sci* 9 (6):1226-34.
- Kim, S. W., and J. D. Keasling. 2001. Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production. *Biotechnol Bioeng* 72 (4):408-15.
- Kirby, J., and J. D. Keasling. 2009. Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu Rev Plant Biol* 60:335-55.
- Kizer, L., D. J. Pitera, B. F. Pfleger, and J. D. Keasling. 2008. Application of functional genomics to pathway optimization for increased isoprenoid production. *Appl Environ Microbiol* 74 (10):3229-41.
- Klamt, S., J. Saez-Rodriguez, and E. D. Gilles. 2007. Structural and functional analysis of cellular networks with *CellNetAnalyzer*. *BMC Syst Biol* 1:2.
- Klamt, S., and J. Stelling. 2002. Combinatorial complexity of pathway analysis in metabolic networks. *Mol Biol Rep* 29 (1-2):233-6.
- Klamt, Steffen, Oliver Hädicke, and Axel von Kamp. 2014. Stoichiometric and constraint-based analysis of biochemical reaction networks. In *Large-Scale Networks in Engineering and Life Sciences*, edited by P. Benner, R. Findeisen, D. Flockerzi, U. Reichl and K. Sundmacher: Springer.
- Kocharin, K., Y. Chen, V. Siewers, and J. Nielsen. 2012. Engineering of acetyl-CoA metabolism for the improved production of polyhydroxybutyrate in *Saccharomyces cerevisiae*. *AMB Express* 2 (1):52.
- Kocharin, K., V. Siewers, and J. Nielsen. 2013. Improved polyhydroxybutyrate production by *Saccharomyces cerevisiae* through the use of the phosphoketolase pathway. *Biotechnol Bioeng*.
- Kozak, B. U., H. M. van Rossum, K. R. Benjamin, L. Wu, J. M. Daran, J. T. Pronk, and A. J. van Maris. 2013. Replacement of the *Saccharomyces cerevisiae* acetyl-CoA synthetases by alternative pathways for cytosolic acetyl-CoA synthesis. *Metabolic engineering*.
- Kozak, B. U., H. M. van Rossum, M. A. Luttkik, M. Akeroyd, K. R. Benjamin, L. Wu, S. de Vries, J. M. Daran, J. T. Pronk, and A. J. van Maris. 2014. Engineering Acetyl Coenzyme A Supply: Functional Expression of a Bacterial Pyruvate Dehydrogenase Complex in the Cytosol of *Saccharomyces cerevisiae*. *MBio* 5 (5).
- Kraft, Philip, Christophe Weymuth, and Cornelius Nussbaumer. 2006. Total Synthesis and Olfactory Evaluation of (1R\*,3S\*,6S\*,7S\*,8S\*)- 3-Hydroxy-6,8-dimethyltricyclo[5.3.1.0<sup>3,8</sup>]undecan-2-one: A New Synthetic Route to the Patchoulol Skeleton. *European Journal of Organic Chemistry* 2006 (6):1403-1412.
- Krivoruchko, A., C. Serrano-Amatriain, Y. Chen, V. Siewers, and J. Nielsen. 2013. Improving biobutanol production in engineered *Saccharomyces cerevisiae* by manipulation of acetyl-CoA metabolism. *J Ind Microbiol Biotechnol* 40 (9):1051-6.
- Krivoruchko, A., Y. Zhang, V. Siewers, Y. Chen, and J. Nielsen. 2014. Microbial acetyl-CoA metabolism and metabolic engineering. *Metabolic engineering* 28C:28-42.
- Krömer, J. O., C. Wittmann, H. Schröder, and E. Heinzle. 2006. Metabolic pathway analysis for rational design of L-methionine production by *Escherichia coli* and *Corynebacterium glutamicum*. *Metab Eng* 8 (4):353-69.
- Kuzuyama, T., and H. Seto. 2003. Diversity of the biosynthesis of the isoprene units. *Nat Prod Rep* 20 (2):171-83.
- Lai, ShuJuan, Yun Zhang, ShuWen Liu, Yong Liang, XiuLing Shang, Xin Chai, and TingYi Wen. 2012. Metabolic engineering and flux analysis of *Corynebacterium glutamicum* for L-serine production. *Science China Life Sciences* 55 (4):283-290.
- Lee, Sang Yup, Dong-Yup Lee, and Tae Yong Kim. 2005. Systems biotechnology for strain improvement. *Trends in Biotechnology* 23 (7):349-358.

- Li, S., D. Huang, Y. Li, J. Wen, and X. Jia. 2012. Rational improvement of the engineered isobutanol-producing *Bacillus subtilis* by elementary mode analysis. *Microb Cell Fact* 11 (1):101.
- Liu, X. C., Q. Liu, H. Chen, Q. Z. Liu, S. Y. Jiang, and Z. L. Liu. 2015. Evaluation of Contact Toxicity and Repellency of the Essential Oil of *Pogostemon cablin* Leaves and Its Constituents Against *Blattella germanica* (Blattodea: Blattellidae). *J Med Entomol* 52 (1):86-92.
- Llaneras, Francisco, and Jesús Picó. 2008. Stoichiometric modelling of cell metabolism. *Journal of Bioscience and Bioengineering* 105 (1):1-11.
- Lubertozzi, D., and J. D. Keasling. 2008. Expression of a synthetic Artemesia annua amorphadiene synthase in *Aspergillus nidulans* yields altered product distribution. *J Ind Microbiol Biotechnol* 35 (10):1191-8.
- Martin, V. J., D. J. Pitera, S. T. Withers, J. D. Newman, and J. D. Keasling. 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol* 21 (7):796-802.
- Martin, V. J., Y. Yoshikuni, and J. D. Keasling. 2001. The *in vivo* synthesis of plant sesquiterpenes by *Escherichia coli*. *Biotechnol Bioeng* 75 (5):497-503.
- Martínez-Férez, Isabel, Blanca Fernández-González, Gerhard Sandmann, and Agustín Vioque. 1994. Cloning and expression in *Escherichia coli* of the gene coding for phytoene synthase from the cyanobacterium *Synechocystis sp.* PCC6803. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1218 (2):145-152.
- Masood, Ehsan. 2003. GM crops: A continent divided. *Nature* 426 (6964):224-226.
- Matsuda, F., C. Furusawa, T. Kondo, J. Ishii, H. Shimizu, and A. Kondo. 2011. Engineering strategy of yeast metabolism for higher alcohol production. *Microb Cell Fact* 10:70.
- Maury, J., M. A. Asadollahi, K. Moller, A. Clark, and J. Nielsen. 2005. Microbial isoprenoid production: an example of green chemistry through metabolic engineering. *Adv Biochem Eng Biotechnol* 100:19-51.
- Mayer, M. P., F. M. Hahn, D. J. Stillman, and C. D. Poulter. 1992. Disruption and mapping of *IDI1*, the gene for isopentenyl diphosphate isomerase in *Saccharomyces cerevisiae*. *Yeast* 8 (9):743-8.
- McCarthy, John F. 2002. Turning in Circles: District Governance, Illegal Logging, and Environmental Decline in Sumatra, Indonesia. *Society and Natural Resources* 15:867-886.
- McGarvey, D. J., and R. Croteau. 1995. Terpenoid metabolism. *Plant Cell* 7 (7):1015-26.
- Meganathan, R, and O Kwon. 2009. Biosynthesis of Menaquinone (Vitamin K2) and Ubiquinone (Coenzyme Q). *EcoSal Plus* 3.
- Melillo, E., R. Setroikromo, W. J. Quax, and O. Kayser. 2013. Production of alpha-cuprenene in *Xanthophyllomyces dendrorhous*: a step closer to a potent terpene biofactory. *Microb Cell Fact* 12 (1):13.
- Melzer, G., M. E. Esfandabadi, E. Franco-Lara, and C. Wittmann. 2009. Flux Design: *In silico* design of cell factories based on correlation of pathway fluxes to desired properties. *BMC Syst Biol* 3:120.
- Mercke, P., M. Bengtsson, H. J. Bouwmeester, M. A. Posthumus, and P. E. Brodelius. 2000. Molecular cloning, expression, and characterization of amorpho-4,11-diene synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L. *Arch Biochem Biophys* 381 (2):173-80.

- Mi, J., D. Becher, P. Lubuta, S. Dany, K. Tusch, H. Schewe, M. Buchhaupt, and J. Schrader. 2014. *De novo* production of the monoterpenoid geranic acid by metabolically engineered *Pseudomonas putida*. *Microb Cell Fact* 13 (1):170.
- Miller, W.L. 1988. Molecular biology of steroid hormone synthesis. *Endocrine reviews* 9 (3):295.
- Miura, Y., K. Kondo, T. Saito, H. Shimada, P. D. Fraser, and N. Misawa. 1998. Production of the carotenoids lycopene, beta-carotene, and astaxanthin in the food yeast *Candida utilis*. *Appl Environ Microbiol* 64 (4):1226-9.
- Morrone, D., L. Lowry, M. K. Determan, D. M. Hershey, M. Xu, and R. J. Peters. 2010. Increasing diterpene yield with a modular metabolic engineering system in *E. coli*: comparison of MEV and MEP isoprenoid precursor pathway engineering. *Appl Microbiol Biotechnol* 85 (6):1893-906.
- Mouritsen, O. G., and M. J. Zuckermann. 2004. What's so special about cholesterol? *Lipids* 39 (11):1101-13.
- Munck, S. L., and R. Croteau. 1990. Purification and characterization of the sesquiterpene cyclase patchoulol synthase from *Pogostemon cablin*. *Arch Biochem Biophys* 282 (1):58-64.
- Naguib, Y. M. 2000. Antioxidant activities of astaxanthin and related carotenoids. *J Agric Food Chem* 48 (4):1150-4.
- Nielsen, J. 2001. Metabolic engineering. *Appl Microbiol Biotechnol* 55 (3):263-83.
- Nielsen, J., M. Fussenegger, J. Keasling, S. Y. Lee, J. C. Liao, K. Prather, and B. Palsson. 2014. Engineering synergy in biotechnology. *Nat Chem Biol* 10 (5):319-22.
- Nielsen, J., and J. D. Keasling. 2011. Synergies between synthetic biology and metabolic engineering. *Nat Biotechnol* 29 (8):693-5.
- Ohlendorf, Wulf. 1996. Domestication and crop development of *Duboisia* spp.(Solanaceae). *Domestication and commercialization of non-timber forest products* 176:11-103.
- Oswald, M., M. Fischer, N. Dirninger, and F. Karst. 2007. Monoterpenoid biosynthesis in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 7 (3):413-21.
- Otero, J. M., and J. Nielsen. 2010. Industrial systems biology. *Biotechnol Bioeng* 105 (3):439-60.
- Oud, B., A. J. van Maris, J. M. Daran, and J. T. Pronk. 2012. Genome-wide analytical approaches for reverse metabolic engineering of industrially relevant phenotypes in yeast. *FEMS Yeast Res* 12 (2):183-96.
- Paddon, C. J., and J. D. Keasling. 2014. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat Rev Microbiol*.
- Papini, M., I. Nookaew, V. Siewers, and J. Nielsen. 2012. Physiological characterization of recombinant *Saccharomyces cerevisiae* expressing the *Aspergillus nidulans* phosphoketolase pathway: validation of activity through (13)C-based metabolic flux analysis. *Appl Microbiol Biotechnol*.
- Paradise, E. M., J. Kirby, R. Chan, and J. D. Keasling. 2008. Redirection of flux through the FPP branch-point in *Saccharomyces cerevisiae* by down-regulating squalene synthase. *Biotechnol Bioeng* 100 (2):371-8.
- Parekh, S., V. A. Vinci, and R. J. Strobel. 2000. Improvement of microbial strains and fermentation processes. *Applied Microbiology and Biotechnology* 54 (3):287-301.
- Pateraki, I., A. M. Heskes, and B. Hamberger. 2015. Cytochromes P450 for terpene functionalisation and metabolic engineering. *Adv Biochem Eng Biotechnol* 148:107-39.
- Patnaik, R. 2008. Engineering complex phenotypes in industrial strains. *Biotechnol Prog* 24 (1):38-47.

- Pitera, D. J., C. J. Paddon, J. D. Newman, and J. D. Keasling. 2007. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metab Eng* 9 (2):193-207.
- Polakowski, T., U. Stahl, and C. Lang. 1998. Overexpression of a cytosolic hydroxymethylglutaryl-CoA reductase leads to squalene accumulation in yeast. *Appl Microbiol Biotechnol* 49 (1):66-71.
- Potrykus, Ingo. 2012. "Golden Rice", a GMO-product for public good, and the consequences of GE-regulation. *Journal of Plant Biochemistry and Biotechnology* 21 (1):68-75.
- Raina, A. P., and K. S. Negi. 2015. Essential oil composition of *Valeriana jatamansi* Jones from himalayan regions of India. *Indian J Pharm Sci* 77 (2):218-22.
- Ramya, HG, V Palanimuthu, and Singla Rachna. 2013. An Introduction to Patchouli (*Pogostemon cablin* Benth.) - A Medicinal and Aromatic Plant: It's Importance to Mankind. *Agricultural Engineering International: CIGR Journal* 15 (2):243-250.
- Razem, F. A., K. Baron, and R. D. Hill. 2006. Turning on gibberellin and abscisic acid signaling. *Curr Opin Plant Biol* 9 (5):454-9.
- Ro, D. K., E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Chang, S. T. Withers, Y. Shiba, R. Sarpong, and J. D. Keasling. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440 (7086):940-3.
- Rodríguez-Concepción, M, and A Boronat. 2002. Elucidation of the Methylerythritol Phosphate Pathway for Isoprenoid Biosynthesis in Bacteria and Plastids. A Metabolic Milestone Achieved through Genomics. *Plant Physiology* 130 (3):1079-1089.
- Rodríguez-Concepción, M. 2014. Plant isoprenoids: a general overview. *Methods Mol Biol* 1153:1-5.
- Rodriguez-Villalon, A., J. Perez-Gil, and M. Rodriguez-Concepcion. 2008. Carotenoid accumulation in bacteria with enhanced supply of isoprenoid precursors by upregulation of exogenous or endogenous pathways. *J Biotechnol* 135 (1):78-84.
- Rodriguez, S., J. Kirby, C. M. Denby, and J. D. Keasling. 2014. Production and quantification of sesquiterpenes in *Saccharomyces cerevisiae*, including extraction, detection and quantification of terpene products and key related metabolites. *Nat Protoc* 9 (8):1980-96.
- Rohdich, F., F. Zepeck, P. Adam, S. Hecht, J. Kaiser, R. Laupitz, T. Grawert, S. Amslinger, W. Eisenreich, A. Bacher, and D. Arigoni. 2003. The deoxyxylulose phosphate pathway of isoprenoid biosynthesis: studies on the mechanisms of the reactions catalyzed by IspG and IspH protein. *Proc Natl Acad Sci U S A* 100 (4):1586-91.
- Rohmer, M. 1999. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Nat Prod Rep* 16 (5):565-74.
- Rohmer, Michel, Myriam Seemann, Silke Horbach, Stephanie Bringer-Meyer, and Hermann Sahn. 1996. Glyceraldehyde 3-phosphate and pyruvate as precursors of isoprenic units in an alternative non-mevalonate pathway for terpenoid biosynthesis. *Journal of the American Chemical Society* 118 (11):2564-2566.
- Rowlands, R. T. 1984. Industrial strain improvement: Mutagenesis and random screening procedures. *Enzyme and Microbial Technology* 6 (1):3-10.
- Rudney, Harry, and James J. Ferguson. 1957. The biosynthesis of  $\beta$ -hydroxy- $\beta$ -methylglutaryl coenzyme A. *Journal of the American Chemical Society* 79 (20):5580-5581.
- Ruzicka, L. 1953. The isoprene rule and the biogenesis of terpenic compounds. *Experientia* 9 (10):357.



- Scalcinati, G., C. Knuf, S. Partow, Y. Chen, J. Maury, M. Schalk, L. Daviet, J. Nielsen, and V. Siewers. 2012. Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquiterpene alpha-santalene in a fed-batch mode. *Metabolic engineering* 14 (2):91-103.
- Scalcinati, G., S. Partow, V. Siewers, M. Schalk, L. Daviet, and J. Nielsen. 2012. Combined metabolic engineering of precursor and co-factor supply to increase alpha-santalene production by *Saccharomyces cerevisiae*. *Microb Cell Fact* 11 (1):117.
- Scholtmeijer, K., K. Cankar, J. Beekwilder, H. A. Wosten, L. G. Lugones, and D. Bosch. 2014. Production of (+)-valencene in the mushroom-forming fungus *S. commune*. *Appl Microbiol Biotechnol* 98 (11):5059-68.
- Schuster, S., T. Dandekar, and D. A. Fell. 1999. Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol* 17 (2):53-60.
- Schuster, S., and C. Hilgetag. 1994. On elementary flux modes in biochemical reaction systems at steady state. *Journal of Biological Systems* 02 (02):165-182.
- Seemann, M., B. T. Bui, M. Wolff, D. Tritsch, N. Campos, A. Boronat, A. Marquet, and M. Rohmer. 2002. Isoprenoid biosynthesis through the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GcpE) is a [4Fe-4S] protein. *Angew Chem Int Ed Engl* 41 (22):4337-9.
- Shiba, Y., E. M. Paradise, J. Kirby, D. K. Ro, and J. D. Keasling. 2007. Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* 9 (2):160-8.
- Steele, C. L., J. Crock, J. Bohlmann, and R. Croteau. 1998. Sesquiterpene synthases from grand fir (*Abies grandis*). Comparison of constitutive and wound-induced activities, and cDNA isolation, characterization, and bacterial expression of delta-selinene synthase and gamma-humulene synthase. *J Biol Chem* 273 (4):2078-89.
- Sundaresan, V., S. P. Singh, A. N. Mishra, Ajit K. Shasany, Mahendra P. Darokar, Alok Kalra, and A. A. Naqvi. 2009. Composition and Comparison of Essential Oils of *Pogostemon cablin* (Blanco) Benth. (Patchouli) and *Pogostemon travancoricus* Bedd. var. *travancoricus*. *Journal of Essential Oil Research* 21 (3):220-222.
- Surburg, H., and J. Panten. 2006. *Common Fragrance and Flavour Materials. Preparation, Properties and Uses*. 5 ed. Weinheim: WILEY-VCH.
- Swamy, M. K., and U. R. Sinniah. 2015. A Comprehensive Review on the Phytochemical Constituents and Pharmacological Activities of *Pogostemon cablin* Benth.: An Aromatic Medicinal Plant of Industrial Importance. *Molecules* 20 (5):8521-8547.
- Takahashi, S., T. Kuzuyama, H. Watanabe, and H. Seto. 1998. A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc Natl Acad Sci U S A* 95 (17):9879-84.
- Tavormina, Peter A., Margaret H. Gibbs, and Jesse W. Huff. 1956. The utilization of  $\beta$ -hydroxy- $\beta$ -methyl- $\delta$ -valerolactone in cholesterol biosynthesis. *Journal of the American Chemical Society* 78 (17):4498-4499.
- Thomas, P. *Taxus brevifolia*. *The IUCN Red List of Threatened Species. Version 2015.2* 2013 [cited 25.08.2015. Available from [www.iucnredlist.org](http://www.iucnredlist.org).
- Thykaer, J., and J. Nielsen. 2003. Metabolic engineering of beta-lactam production. *Metabolic engineering* 5 (1):56-69.
- Trinh, C. T. 2012. Elucidating and reprogramming *Escherichia coli* metabolisms for obligate anaerobic *n*-butanol and isobutanol production. *Appl Microbiol Biotechnol* 95 (4):1083-1094.

- Trinh, C. T., R. Carlson, A. Wlaschin, and F. S. S. 2006. Design, construction and performance of the most efficient biomass producing *E. coli* bacterium. *Metab Eng* 8 (6):628-38.
- Trinh, C. T., and F. S. S. 2009. Metabolic engineering of *Escherichia coli* for efficient conversion of glycerol to ethanol. *Appl Environ Microbiol* 75 (21):6696-705.
- Trinh, C. T., P. Unrean, and F. S. S. 2008. Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. *Appl Environ Microbiol* 74 (12):3634-43.
- Trinh, C. T., A. Wlaschin, and F. S. S. 2009. Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl Microbiol Biotechnol* 81 (5):813-26.
- Tsuruta, H., C. J. Paddon, D. Eng, J. R. Lenihan, T. Horning, L. C. Anthony, R. Regentin, J. D. Keasling, N. S. Renninger, and J. D. Newman. 2009. High-level production of amorphadiene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS One* 4 (2):e4489.
- Unrean, P., and N. H. Nguyen. 2012. Metabolic pathway analysis of *Scheffersomyces (Pichia) stipitis*: effect of oxygen availability on ethanol synthesis and flux distributions. *Appl Microbiol Biotechnol* 94 (5):1387-98.
- Unrean, P., C. T. Trinh, and F. S. S. 2010. Rational design and construction of an efficient *E. coli* for production of diapolycopendioic acid. *Metab Eng* 12 (2):112-22.
- Vadali, R. V., Y. Fu, G. N. Bennett, and K. Y. San. 2005. Enhanced lycopene productivity by manipulation of carbon flow to isopentenyl diphosphate in *Escherichia coli*. *Biotechnol Prog* 21 (5):1558-61.
- Verwaal, R., J. Wang, J. P. Meijnen, H. Visser, G. Sandmann, J. A. van den Berg, and A. J. van Ooyen. 2007. High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous*. *Appl Environ Microbiol* 73 (13):4342-50.
- Vickers, C. E., M. Bongers, Q. Liu, T. Delatte, and H. Bouwmeester. 2014. Metabolic engineering of volatile isoprenoids in plants and microbes. *Plant Cell Environ* 37 (8):1753-75.
- Walji, Abbas M, and David WC MacMillan. 2007. Strategies to bypass the taxol problem. Enantioselective cascade catalysis, a new approach for the efficient construction of molecular complexity. *Synlett* 18 (10):1477-1489.
- Wang, A. *Artemisinin: Vagaries of weather and the market hamper deliveries* 2010 [cited 27.08.2015. Available from [www.ft.com](http://www.ft.com).
- Wang, C. W., M. K. Oh, and J. C. Liao. 1999. Engineered isoprenoid pathway enhances astaxanthin production in *Escherichia coli*. *Biotechnol Bioeng* 62 (2):235-41.
- Wang, C., S. H. Yoon, A. A. Shah, Y. R. Chung, J. Y. Kim, E. S. Choi, J. D. Keasling, and S. W. Kim. 2010. Farnesol production from *Escherichia coli* by harnessing the exogenous mevalonate pathway. *Biotechnol Bioeng* 107 (3):421-9.
- Weitzel, Corinna, and Henrik Toft Simonsen. 2013. Cytochrome P450-enzymes involved in the biosynthesis of mono- and sesquiterpenes. *Phytochemistry Reviews*:1-18.
- Westfall, P. J., D. J. Pitera, J. R. Lenihan, D. Eng, F. X. Woolard, R. Regentin, T. Horning, H. Tsuruta, D. J. Melis, A. Owens, S. Fickes, D. Diola, K. R. Benjamin, J. D. Keasling, M. D. Leavell, D. J. McPhee, N. S. Renninger, J. D. Newman, and C. J. Paddon. 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci U S A* 109 (3):E111-8.

- Willrodt, C., C. David, S. Cornelissen, B. Bühler, M. K. Julsing, and A. Schmid. 2014. Engineering the productivity of recombinant *Escherichia coli* for limonene formation from glycerol in minimal media. *Biotechnol J* 9 (8):1000-12.
- Withers, S. T., and J. D. Keasling. 2007. Biosynthesis and engineering of isoprenoid small molecules. *Appl Microbiol Biotechnol* 73 (5):980-90.
- Wolf, Donald E., Carl H. Hoffman, Paul E. Aldrich, Helen R. Skeggs, Lemuel D. Wright, and Karl Folkers. 1956.  $\beta$ -Hydroxy- $\beta$ -methyl- $\delta$ -valerolactone (divalonic acid), a new biological factor. *Journal of the American Chemical Society* 78 (17):4499-4499.
- Wolfertz, M., T. D. Sharkey, W. Boland, and F. Kuhnemann. 2004. Rapid regulation of the methylerythritol 4-phosphate pathway during isoprene synthesis. *Plant Physiol* 135 (4):1939-45.
- Wolff, M., M. Seemann, B. Tse Sum Bui, Y. Frapart, D. Tritsch, A. G. Estrabot, M. Rodríguez-Concepción, A. Boronat, A. Marquet, and M. Rohmer. 2003. Isoprenoid biosynthesis via the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase (LytB/IspH) from *Escherichia coli* is a [4Fe-4S] protein. *FEBS Letters* 541 (1-3):115-120.
- Wu, S., M. Schalk, A. Clark, R. B. Miles, R. Coates, and J. Chappell. 2006. Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nat Biotechnol* 24 (11):1441-7.
- Yamano, S., T. Ishii, M. Nakagawa, H. Ikenaga, and N. Misawa. 1994. Metabolic engineering for production of beta-carotene and lycopene in *Saccharomyces cerevisiae*. *Biosci Biotechnol Biochem* 58 (6):1112-4.
- Yuan, L. Z., P. E. Rouviere, R. A. Larossa, and W. Suh. 2006. Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. *Metab Eng* 8 (1):79-90.
- Zhan, X., Y. H. Zhang, D. F. Chen, and H. T. Simonsen. 2014. Metabolic engineering of the moss *Physcomitrella patens* to produce the sesquiterpenoids patchoulol and alpha/beta-santalene. *Front Plant Sci* 5:636.
- Zhao, Jing, Qingyan Li, Tao Sun, Xinna Zhu, Hongtao Xu, Jinlei Tang, Xueli Zhang, and Yanhe Ma. 2013. Engineering central metabolic modules of *Escherichia coli* for improving  $\beta$ -carotene production. *Metab Eng* 17:42-50.

## CHAPTER 2

### ***In silico* profiling of *Escherichia coli* and *Saccharomyces cerevisiae* as terpenoid factories**

Evamaria Gruchattka, Oliver Hädicke, Steffen Klamt, Verena Schütz, Oliver Kayser

EG did all the computations and interpretations. OH and SK supported the constrained minimal cut sets computations. EG and VS designed the study. EG, VS, OH and SK wrote and drafted the manuscript. OK supervised the research. VS conceived of the study and supervised the research.

Published in: *Microbial Cell Factories*. 2013. 12(1):84

## 2.1 Abstract

Heterologous microbial production of rare plant terpenoids of medicinal or industrial interest is attracting more and more attention but terpenoid yields are still low. *Escherichia coli* and *Saccharomyces cerevisiae* are the most widely used heterologous hosts; a direct comparison of both hosts based on experimental data is difficult though. Hence, the terpenoid pathways of *E. coli* (via 1-deoxy-D-xylulose 5-phosphate, DXP) and *S. cerevisiae* (via mevalonate, MVA), the impact of the respective hosts metabolism as well as the impact of different carbon sources were compared *in silico* by means of elementary mode analysis. The focus was set on the yield of isopentenyl diphosphate (IPP), the general terpenoid precursor, to identify new metabolic engineering strategies for an enhanced terpenoid yield.

Starting from the respective precursor metabolites of the terpenoid pathways (pyruvate and glyceraldehyde-3-phosphate for the DXP pathway and acetyl-CoA for the MVA pathway) and considering only carbon stoichiometry, the two terpenoid pathways are identical with respect to carbon yield. However, with glucose as substrate, the MVA pathway has a lower potential to supply terpenoids in high yields than the DXP pathway if the formation of the required precursors is taken into account, due to the carbon loss in the formation of acetyl-CoA. This maximum yield is further reduced in both hosts when the required energy and reduction equivalents are considered. Moreover, the choice of carbon source (glucose, xylose, ethanol or glycerol) has an effect on terpenoid yield with non-fermentable carbon sources being more promising. Both hosts have deficiencies in energy and redox equivalents for high yield terpenoid production leading to new overexpression strategies (heterologous enzymes/pathways) for an enhanced terpenoid yield. Finally, several knockout strategies are identified using constrained minimal cut sets enforcing a coupling of growth to a terpenoid yield which is higher than any yield published in scientific literature so far.

This study provides for the first time a comprehensive and detailed *in silico* comparison of the most prominent heterologous hosts *E. coli* and *S. cerevisiae* as terpenoid factories giving an overview on several promising metabolic engineering strategies paving the way for an enhanced terpenoid yield.

## 2.2 Background

Terpenoids are a class of natural products with important medicinal and industrial applications such as artemisinin (antimalarial) or paclitaxel (anticancer) (Hong and Nielsen 2012; Kirby and Keasling 2009). However, several of these bioactive compounds are scarce and produced only in low amounts in plants, which makes the production in their native hosts uneconomical and environmentally destructive (Kirby and Keasling 2009). The total chemical synthesis of many terpenoids is challenging due to their complex structure and neither ecologically nor economically efficient (Davies 2009; Jiang, Stephanopoulos, and Pfeifer 2012; Woodrow, Haynes, and Krishna 2005). Alternatively, the use of a microbial platform organism for the production of terpenoids may offer the possibility of large-scale, cost-effective and environmentally friendly industrial production independent from climate or cultivation risks. Today, *Escherichia coli* and *Saccharomyces cerevisiae* are the most widely used microorganisms for heterologous terpenoid production. They are preferably used due to advanced molecular biology tools, growth speed and their well-established use in industrial biotechnology (Jiang, Stephanopoulos, and Pfeifer 2012; Kirby and Keasling 2008). While most studies for the production of carotenoid compounds of lower complexity such as lycopene (an antioxidative carotenoid) have been carried out in *E. coli*, the trends concerning the production of artemisinin or paclitaxel show a more even distribution between these two hosts with an increasingly more prevalent use of *S. cerevisiae* (Jiang, Stephanopoulos, and Pfeifer 2012). This trend might partly be attributed to the fact that *S. cerevisiae* as an eukaryotic organism has the advantage of being more suitable than *E. coli* for the expression of membrane-bound plant cytochrome P450 enzymes along with their respective reductase (Kirby and Keasling 2009) required for the functionalization of a series of terpenoids (Chang et al. 2007; Ro et al. 2006; Westfall et al. 2012). A further advantage is the possibility to harness different subcellular compartments (Farhi et al. 2011). Moreover, yeast can withstand reduced pH and high osmotic pressure (Kampranis and Makris 2012) and is not susceptible to phage infections.

The production of terpenoids in both organisms requires a sufficient supply of their common precursor isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP), which are then condensed, cyclized etc. leading to the structural diversity of terpenoids. *E. coli* and *S. cerevisiae* differ in their metabolic routes to supply these precursors. The 1-deoxy-D-xylulose 5-phosphate (DXP) pathway of *E. coli* (often

also called 2C-methyl-D-erythritol 4-phosphate pathway, MEP) is fed by pyruvate and glyceraldehyde-3-phosphate whereas in the mevalonate (MVA) pathway of *S. cerevisiae* IPP derives from acetyl-CoA (Maury et al. 2005; Partow et al. 2012).

Attempts to engineer the DXP pathway in *E. coli* mainly focused on the combinatorial overexpression of the rate limiting enzymes encoded by *dxs*, *idi*, *dxr*, *ispD* and *ispF* with various ranges of expression levels. Those studies stressed the need to carefully balance the expression of the DXP pathway genes with the heterologous terpenoid forming pathway as well as with the host's overall metabolism (Ajikumar et al. 2010; Huang et al. 2001; Kim and Keasling 2001; Martin et al. 2003; Morrone et al. 2010; Rodriguez-Villalon, Perez-Gil, and Rodriguez-Concepcion 2008; Yuan et al. 2006; Zhou et al. 2011). Alternatively, various examples show that the introduction of an optimized MVA pathway efficiently increases terpenoid production in *E. coli*, circumventing the largely unidentified regulations of its native DXP pathway and also decoupling the MVA pathway from its native control in yeast (Kizer et al. 2008; Martin et al. 2003; Pitera et al. 2007; Tsuruta et al. 2009; Yoon et al. 2009).

Numerous engineering strategies have been applied especially to the MVA pathway in *S. cerevisiae* (Kampranis and Makris 2012). Three particularly successful interventions have been described in literature: i) overexpression of a truncated version of 3-hydroxy-3-methylglutaryl-coenzyme A reductase (*HMG1*), the key regulatory enzyme of the pathway, devoid of feedback inhibition by farnesyl diphosphate (FPP) (Donald, Hampton, and Fritz 1997; Gardner and Hampton 1999; Polakowski, Stahl, and Lang 1998); ii) downregulation of squalene synthase (*ERG9*) to reduce the flux draining of FPP to biosynthetically related sterols (Asadollahi et al. 2008; Paradise et al. 2008); and iii) expression of a mutant transcription factor (*upc2-1*) to transcriptionally up-regulate several MVA pathway genes (Ro et al. 2006; Scalcinati et al. 2012; Westfall et al. 2012).

A direct *in vivo* head-to-head comparison of both organisms and terpenoid pathways is difficult to interpret, especially if it is considered that the expression of enzymes of both IPP forming pathways has to be balanced carefully for optimal precursor supply and that *E. coli* and *S. cerevisiae* differ in their central metabolic networks. However, *in silico* methods can reveal their theoretical potential and give further insight into the properties of the underlying metabolic networks, thus providing a rational basis for metabolic engineering.

Elementary mode analysis (EMA) is a key methodology to study essential properties of a metabolic network. Elementary modes (EMs) characterize the space of feasible

steady-state flux distributions; any such flux distribution can be represented as a non-negative (conic) linear combination of elementary modes. Based on the stoichiometric matrix, the steady-state assumption, thermodynamic constraints (reaction reversibilities) and a non-decomposability constraint, elementary modes can be calculated without the need for kinetic data, *a priori* measurements, or an objective function (Melzer et al. 2009; Papin et al. 2004; Schuster, Fell, and Dandekar 2000; Schuster et al. 2002; Trinh, Wlaschin, and Sreenc 2009). EMA has been applied to calculate the overall capacity of a given metabolic network, e.g. the theoretical maximum yield under a given genetic constitution or on different substrates (Krömer et al. 2006; Melzer et al. 2009) in order to estimate the potential efficiency of a biotechnological process. Moreover, it has been successfully used as a basis for the computation of intervention strategies to obtain superior production strains (Carlson, Fell, and Sreenc 2002; Carlson and Sreenc 2004; Hädicke and Klamt 2011; Melzer et al. 2009; Unrean, Trinh, and Sreenc 2010). Recently, Hädicke *et al.* (Hädicke and Klamt 2011) introduced the concept of constrained minimal cut sets (cMCSs) to compute knockout strategies for coupled product and biomass synthesis allowing not only the specification of functionalities that are to be disabled in the metabolic network but also of those that have to be preserved. Furthermore, all equivalent knockout strategies to accomplish the same engineering goal are computed, giving the researcher the flexibility to choose the best combination of gene deletions in terms of practical realization. cMCSs comprise and extend e.g. the approach of minimal metabolic functionality (Trinh et al. 2006) that has been successfully applied multiple times in order to identify metabolic engineering strategies (Trinh et al. 2006; Trinh, Unrean, and Sreenc 2008; Trinh and Sreenc 2009; Unrean, Trinh, and Sreenc 2010).

In this study, EMA is applied to compare the metabolic networks of *E. coli* and *S. cerevisiae* for the production of the common terpenoid precursor IPP by means of carbon stoichiometry, energy and redox equivalent requirements. Further, the potential of the two hosts' central metabolic networks to meet those requirements if the MVA, the DXP or both pathways are operating is analyzed. Moreover, the impact of different industrially relevant carbon sources on IPP production is investigated *in silico*.

While most studies on terpenoid production mainly focused on engineering the DXP and MVA pathway, only a few studies deal with the identification of engineering targets in the two host's central metabolism for an improved precursor supply (Alper, Miyaoku, and Stephanopoulos 2005; Asadollahi et al. 2009; Choi et al. 2010; Farmer and Liao



2001; Scalcinati et al. 2012; Shiba et al. 2007; Unrean, Trinh, and Srienc 2010; Zhao et al. 2013) leaving tremendous room for further improvements. Therefore, the impact of heterologously introduced enzymes is analyzed *in silico* and knockout targets are computed based on constrained minimal cut sets to give guidelines for metabolic engineering in order to obtain a highly efficient IPP platform organism.

## 2.3 Methods

### 2.3.1 Elementary modes and minimal cut sets

Consider a given metabolic reaction network with  $q \times p$  stoichiometric matrix  $\mathbf{N}$  (with  $q$  reactions and  $p$  metabolites) and a set  $Irrev$  of irreversible reactions. The set of steady-state flux vectors  $\mathbf{r}$  form the convex (flux) cone reads:

$$\mathbf{F} = \{ \mathbf{r} \in \mathbb{R}^q \mid \mathbf{N}\mathbf{r} = \mathbf{0}, r_i \geq 0 \ \forall i \in Irrev \}$$

An Elementary Mode is a non-decomposable vector  $\mathbf{e} \in \mathbf{F}$ , i.e. for every vector  $\mathbf{r} \in \mathbf{F}$ ,  $\mathbf{r} \neq \mathbf{0}$ , one has  $\text{supp}(\mathbf{e}) \subseteq \text{supp}(\mathbf{r})$ , where the support of a vector is defined as the index set of all reactions carrying a non-zero flux ( $\text{supp}(\mathbf{r}) = \{i: r_i \neq 0\}$ ). EMs correspond to minimal functional units (pathways or cycles) of a metabolic network and are useful to study various functional network properties (Schuster, Fell, and Dandekar 2000; Trinh, Wlaschin, and Srienc 2009). EMs were computed using *efmtool* and the software package *CellNetAnalyzer* (Klamt, Saez-Rodriguez, and Gilles 2007; Terzer and Stelling 2008)

In contrast, a minimal cut set (MCS) is a (support-) minimal set  $\mathbf{C}$  of reactions, the removal of which will block (disrupt) a given set  $\mathbf{M}$  of target flux vectors,  $\mathbf{M} \subseteq \mathbf{F}$  (Klamt and Gilles 2004; Klamt 2006):

$$\{ \mathbf{r} \in \mathbf{F} \mid r_i = 0, \forall i \in \mathbf{C} \} \cap \mathbf{M} = \emptyset$$

One drawback of the original definition of MCSs is that besides the targeted functionalities some desired functionalities can be lost as well. Therefore, a generalization to constrained MCSs was introduced (Hädicke and Klamt 2011) which can be described as follows (in the following, EMs are represented by their support, i.e. by the set of reactions carrying a non-zero flux in the EM):

1.) Define a set of target modes  $\tilde{\mathbf{T}}$  (where each target mode  $\mathbf{t}$  is represented by its support  $T = \text{supp}(\mathbf{t})$ )

2.) Define a set of desired modes  $\tilde{\mathbf{D}}$  (where each desired mode  $\mathbf{d}$  is represented by its support  $\mathbf{D} = \text{supp}(\mathbf{d})$ )

3.) Specify a number  $n$  quantifying the minimal number of EMs from  $\tilde{\mathbf{D}}$  that must be kept in the network, i.e. a cMCS  $\mathbf{C}$  fulfills:  $\forall \mathbf{T} \in \tilde{\mathbf{T}} : \mathbf{C} \cap \mathbf{T} \neq \emptyset \wedge |\{\mathbf{D} \in \tilde{\mathbf{D}} \mid \mathbf{D} \cap \mathbf{C} = \emptyset\}| \geq n$ .

cMCSs provide a flexible and convenient approach for formulating and solving intervention problems. The approach also allows one to compute interventions that lead to robust coupling of product and biomass formation (coupling even for suboptimal growth). To compute the cMCSs, we applied an adapted Berge algorithm (allowing the computation of cMCSs as minimal hitting sets of  $\tilde{\mathbf{T}}$  subject to  $\tilde{\mathbf{D}}$  and  $n$ ), implemented in the software package *CellNetAnalyzer* (Klamt, Saez-Rodriguez, and Gilles 2007).

### 2.3.2 Metabolic networks

A comprehensive analysis of the complete set of elementary modes is still infeasible with genome-scale metabolic models. Therefore, stoichiometric models of the central metabolism of *E. coli* and *S. cerevisiae* were built up manually. The basic glucose network for *E. coli* consists of 65 reactions, of which 21 are reversible, and 50 metabolites (12 external). For *S. cerevisiae* the basic glucose network comprises 69 reactions, thereof 30 reversible, and 60 metabolites (8 external). The P/O-ratios (mol ATP produced per 0.5 mol O<sub>2</sub> reduced) representing the efficiency of ATP production are assumed to be 1.2 for *S. cerevisiae* (Förster, Gombert, and Nielsen 2002; Nookaew et al. 2007) and for *E. coli* 2 from NADH and 1 from quinol (Carlson and Srienc 2004; Puustinen et al. 1991; Trinh, Unrean, and Srienc 2008). The metabolic networks of *S. cerevisiae* and *E. coli* have been constructed considering the current knowledge from genome-scale models and literature data (Cherry et al. 2012; Herrgard et al. 2008; Kanehisa et al. 2012; Keseler et al. 2011; Nookaew et al. 2008; Schomburg et al. 2013). The most important reactions are depicted in Figures 2.3 and 2.4, respectively. Details of the models (including the complete list of enzymes, genes and considered literature) are given in supplementary information (Tables 2.5 – 2.12).

## 2.4 Results and discussion

This section is conceptually divided into two parts. The first subsection focuses on the analysis of the two terpenoid pathways starting with a comparison of the DXP and MVA pathway independent from their host organism. Thereafter, the stoichiometric potential of *E. coli* and *S. cerevisiae* to produce the precursor IPP is analyzed. For this purpose, models of the central carbon metabolism were built up and analyzed (65 reactions and 50 metabolites for *E. coli* and 69 reactions and 60 metabolites for *S. cerevisiae*, both for growth on glucose, see methods section and supplementary information (Tables 2.5 – 2.12) for details). Subsequently, the optimal flux distributions are characterized and the impact of an interchange of the pathways of *E. coli* and *S. cerevisiae* is investigated *in silico*. The pathway analysis part is completed by analyzing alternative substrates and comparing the identified theoretically optimal yields with available literature data. The results of all these calculations are summarized in Table 2.1.

The second part of the results section is related to target identification for metabolic engineering. Suitable heterologous overexpression candidates are identified and promising knockout strategies are computed by means of cMCSs.

**Table 2.1 Overview of *in silico* computations.** Number of elementary modes (EMs) and maximal IPP and biomass carbon yields on different carbon sources [C-mol/C-mol] for computations for *E. coli* and *S. cerevisiae*. GAP-DH (NADP<sup>+</sup>) = NADP<sup>+</sup>-dependent glyceraldehyde-3-phosphate dehydrogenase; PDH = cytosolic pyruvate dehydrogenase complex; ACL = ATP-citrate-lyase; TH = soluble and energy-independent transhydrogenase; XI = xylose isomerase pathway; XR-XDH = xylose reductase and xylitol dehydrogenase pathway.

Organism	Carbon source	Characteristics	EMs	Max IPP yield at zero growth	Max IPP yield with growth	Max biomass carbon yield
<i>E. coli</i>	Glucose	Wild type	36,590	0.67	0.57	0.85
<i>E. coli</i>	Glucose	Wild type, anaerobic	8,343	0.48	0.41	0.35
<i>E. coli</i>	Glucose	ATP source	48,603	0.71	0.57	0.99
<i>E. coli</i>	Glucose	NADH source	45,570	0.83	0.57	0.99
<i>E. coli</i>	Glucose	NADPH source	44,842	0.83	0.57	0.99
<i>E. coli</i>	Glucose	GAP-DH (NADP <sup>+</sup> )	74,651	0.71	0.62	0.99
<i>E. coli</i>	Glucose	MVA only	35,968	0.56	0.54	0.85
<i>E. coli</i>	Glucose	MVA + DXP	44,850	0.67	0.59	0.85
<i>E. coli</i>	Glycerol	Wild type	26,160	0.78	0.77	0.99
<i>E. coli</i>	Glycerol	ATP source	29,088	0.83	0.77	0.99
<i>E. coli</i>	Glycerol	NADH source	29,126	0.83	0.77	0.99
<i>E. coli</i>	Glycerol	NADPH source	28,252	0.83	0.77	0.99
<i>E. coli</i>	Glycerol	GAP-DH (NADP <sup>+</sup> )	62,665	0.83	0.77	0.99
<i>E. coli</i>	Glycerol	MVA only	22,225	0.56	0.45	0.99

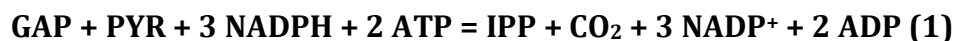
**Table 2.1** continued

Organism	Carbon source	Characteristics	EMs	Max IPP yield at zero growth	Max IPP yield with growth	Max biomass carbon yield
<i>E. coli</i>	Glycerol	MVA + DXP	28,326	0.78	0.77	0.99
<i>S. cerevisiae</i>	Glucose	Wild type	9,844	0.53	0.39	0.80
<i>S. cerevisiae</i>	Glucose	Wild type, anaerobic	495	0.10	0.09	0.21
<i>S. cerevisiae</i>	Glucose	ATP source	15,326	0.56	0.39	0.89
<i>S. cerevisiae</i>	Glucose	NADH source	12,080	0.56	0.39	0.89
<i>S. cerevisiae</i>	Glucose	PDH	17,943	0.56	0.51	0.81
<i>S. cerevisiae</i>	Glucose	Mitochondrial MVA only	10,232	0.56	0.51	0.80
<i>S. cerevisiae</i>	Glucose	Mitochondrial + cytosolic MVA	128,427	0.56	0.53	0.82
<i>S. cerevisiae</i>	Glucose	ACL	17,155	0.55	0.39	0.80
<i>S. cerevisiae</i>	Glucose	DXP only	8,865	0.64	0.57	0.80
<i>S. cerevisiae</i>	Glucose	DXP + MVA	11,738	0.64	0.57	0.80
<i>S. cerevisiae</i>	Glucose	DXP + ATP source	13,731	0.66	0.57	0.89
<i>S. cerevisiae</i>	Glucose	DXP + NADH source	10,714	0.66	0.57	0.89
<i>S. cerevisiae</i>	Glucose	DXP + NADPH source	16,754	0.80	0.57	0.89
<i>S. cerevisiae</i>	Glucose	DXP + TH	25,254	0.67	0.57	0.81
<i>S. cerevisiae</i>	Glucose	DXP + GAP-DH (NADP <sup>+</sup> )	28,383	0.67	0.57	0.81
<i>S. cerevisiae</i>	Galactose	Wild type	9,844	0.53	0.39	0.80
<i>S. cerevisiae</i>	Fructose	Wild type	9,844	0.53	0.39	0.80
<i>S. cerevisiae</i>	Xylose	XI	6,330	0.53	0.39	0.80
<i>S. cerevisiae</i>	Xylose	XR-XDH	16,911	0.53	0.51	0.80
<i>S. cerevisiae</i>	Xylose	XI + DXP only	7,364	0.64	0.42	0.80
<i>S. cerevisiae</i>	Xylose	XI + DXP + MVA	9,088	0.64	0.44	0.80
<i>S. cerevisiae</i>	Xylose	XR-XDH + DXP only	16,513	0.64	0.42	0.80
<i>S. cerevisiae</i>	Xylose	XR-XDH + DXP + MVA	20,337	0.64	0.51	0.80
<i>S. cerevisiae</i>	Glycerol	Wild type	2,648,133	0.56	0.47	0.87
<i>S. cerevisiae</i>	Glycerol	Wild type w/o glyoxylate cycle	5,377	0.56	0.35	0.87
<i>S. cerevisiae</i>	Glycerol	DXP only	2,401,411	0.67	0.60	0.87
<i>S. cerevisiae</i>	Glycerol	DXP + MVA	4,038,007	0.67	0.60	0.87
<i>S. cerevisiae</i>	Glycerol	DXP + NADPH source	3,735,984	0.83	0.80	0.94
<i>S. cerevisiae</i>	Glycerol	DXP + GAP-DH (NADP <sup>+</sup> )	6,153,971	0.76	0.63	0.90
<i>S. cerevisiae</i>	Ethanol	Wild type	2,874,235	0.68	0.59	0.78
<i>S. cerevisiae</i>	Ethanol	ATP source	3,559,725	0.83	0.59	0.81
<i>S. cerevisiae</i>	Ethanol	NADH source	3,635,994	0.83	0.62	0.81
<i>S. cerevisiae</i>	Ethanol	DXP only	2,846,217	0.63	0.57	0.78
<i>S. cerevisiae</i>	Ethanol	DXP + MVA	3,133,493	0.68	0.59	0.78

## 2.4.1 Terpenoid pathway and metabolic network analysis

### Stoichiometry of terpenoid pathways

The stoichiometry of both pathways is initially analyzed independently from the respective host organism. The precursors of the DXP pathway are glyceraldehyde-3-phosphate (GAP) and pyruvate (PYR) and the overall stoichiometry of this pathway is given by equation 1.



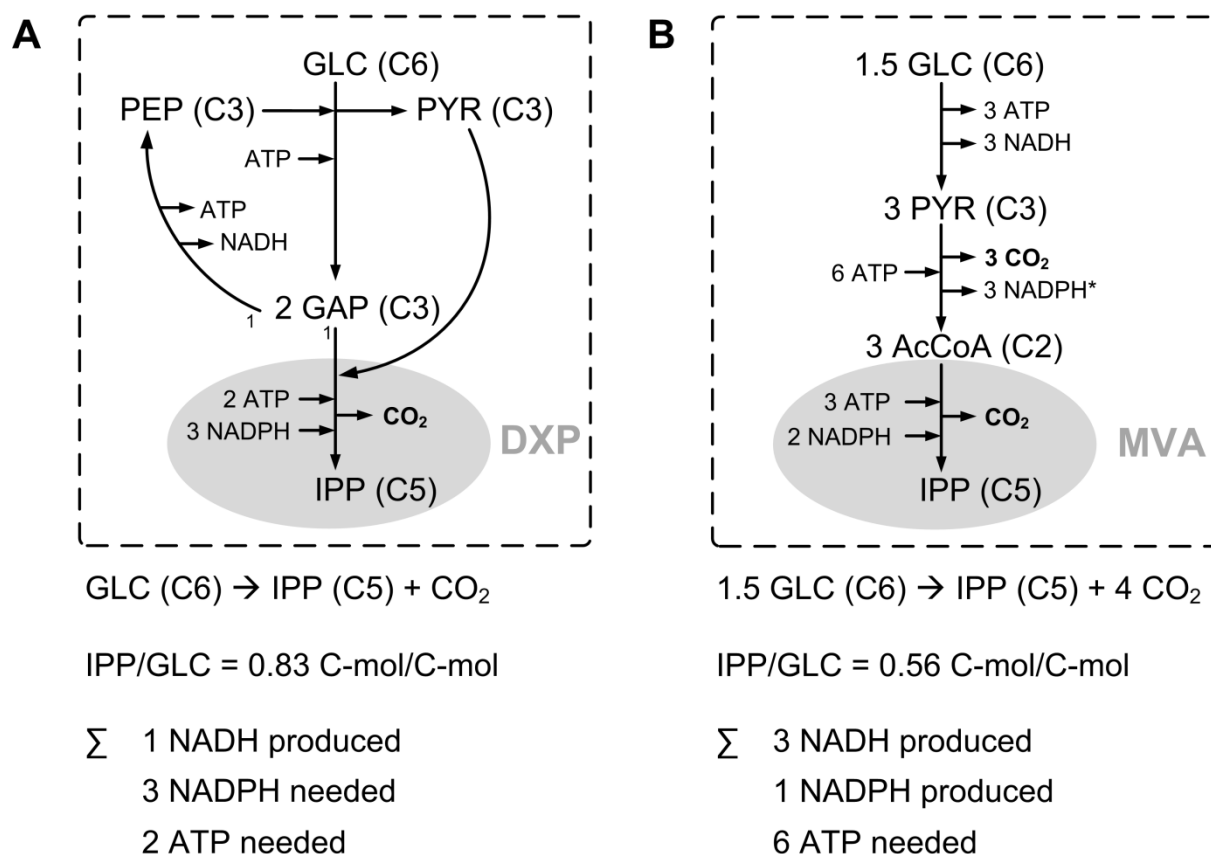
The MVA pathway is fed by the carbon precursor acetyl-CoA (AcCoA) and its stoichiometry is given by equation 2.



Both terpenoid pathways lose one mol CO<sub>2</sub> per mol IPP (C<sub>5</sub>); therefore they are identical with respect to carbon yield ( $5/6 = 0.83$ ) when starting from their precursors GAP (C<sub>3</sub>) and PYR (C<sub>3</sub>) or AcCoA (C<sub>2</sub>). Regarding the demand of ATP and NADPH, they show only minor differences as the DXP pathway needs one mol more NADPH but one mol less ATP than the MVA pathway. Hence, *a priori*, one could assume that no significant differences in achievable IPP yields occur between *E. coli* and *S. cerevisiae*.

### Metabolic potential of wild type *E. coli* and *S. cerevisiae* to supply IPP

Both terpenoid pathways are now analyzed with respect to the metabolic background of their respective host. A first comparison of the maximal IPP yields purely based on carbon stoichiometry, ignoring energy and redox equivalent requirements (i.e. when ATP and NAD(P)H are initially considered to be external metabolites in the model), shows a significant difference. With energy and redox equivalents in excess, the maximum IPP yield on glucose via the DXP pathway in *E. coli* is  $5/6 = 0.83$  C-mol/C-mol compared to  $5/9 = 0.56$  C-mol/C-mol in *S. cerevisiae* if the MVA pathway is used (Figure 2.1). This is due to different amounts of carbon loss via CO<sub>2</sub> in the formation of the terpenoid pathway precursors. The formation of one mol acetyl-CoA, the precursor of the MVA pathway, includes the generation (and thus carbon loss) of one mol CO<sub>2</sub>. In contrast, the formation of one mol pyruvate and one mol glyceraldehyde-3-phosphate, the precursors of the DXP pathway, does not involve a carbon loss via CO<sub>2</sub>.



**Figure 2.1 Comparison of the DXP and MVA pathway.** A section of the central carbon metabolism from glucose (GLC) to IPP: **A:** of *E. coli* including glycolysis and DXP pathway, and **B:** of *S. cerevisiae* including glycolysis and the MVA pathway. The maximal IPP yields based on carbon stoichiometry and ignoring energy and redox equivalent requirements ( $5/6 = 0.83$  and  $5/9 = 0.56$  C-mol/C-mol) differ due to different carbon loss via CO<sub>2</sub> in the formation of the precursors pyruvate (PYR) and glyceraldehyde-3-phosphate (GAP) or acetyl-CoA (AcCoA). The sum shows energy requirements and redox equivalents that have to be balanced by the organism's metabolism. \*: Note that cytosolic NADP<sup>+</sup>-dependent aldehyde dehydrogenase (*ALD6*) is constitutive while cytosolic NAD<sup>+</sup>-dependent aldehyde dehydrogenases (*ALD2*, *ALD3*) are stress-induced, glucose-repressed [82] and were not considered.

Besides the different carbon use efficiencies, the energy and redox equivalent requirements for IPP synthesis differ in *E. coli* and *S. cerevisiae* as well (Figure 2.1). For the formation of one mol IPP from glucose (GLC) in *E. coli*, the overall stoichiometry is given by equation 3.



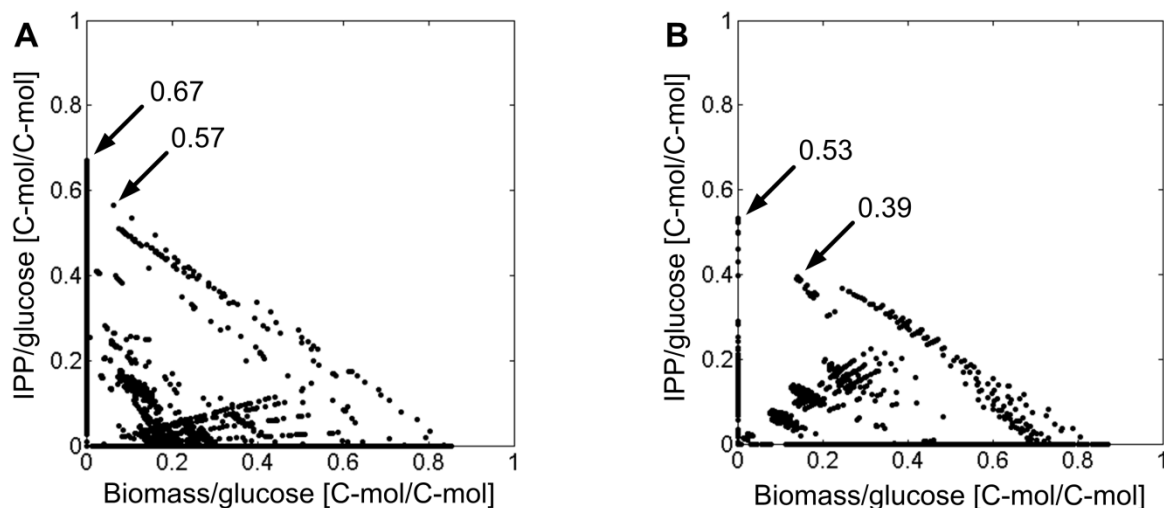
In *S. cerevisiae*, the overall stoichiometry for synthesizing one mol IPP with glucose as substrate via the MVA pathway (considering only the constitutive NADP<sup>+</sup>-dependent aldehyde dehydrogenase, *ALD6*) is given by equation 4.



Now, in the context of the metabolic background, significant differences are observed. When the pathways are considered isolated from the network, they can both serve as redox sinks as they require reduced redox equivalents (see equations 1 and 2). However, due to the synthesis of the specific pathway precursors, different energy and redox equivalent requirements have to be balanced by the organism's metabolism. In *E. coli*, the synthesis of IPP is still accompanied by oxidation of NADPH (equation 3) whereby in *S. cerevisiae* reduced redox equivalents are generated (equation 4). In both organisms, additional ATP has to be generated from the substrate for IPP synthesis, which leads to a further reduction of the maximum IPP yield.

Taking the cofactors explicitly into account requires a network-wide view of all pathways. Therefore, elementary mode analysis was performed on the metabolic networks of *E. coli* (DXP pathway) and *S. cerevisiae* (MVA pathway) wild types with glucose as the single carbon source. In total, 36,590 elementary modes were obtained for the *E. coli* network and 9,844 for the *S. cerevisiae* network (Figure 2.2). Of these, 2,597 modes (26.4 %) of *S. cerevisiae* and even 7,675 modes (21.0 %) of *E. coli* include IPP formation. Computed elementary modes exhibit different yields for IPP and biomass in both organisms (Figure 2.2). The maximum IPP yield for *E. coli* is 0.67 C-mol/C-mol and 0.53 C-mol/C-mol for *S. cerevisiae*. If only elementary modes that couple IPP synthesis with biomass formation are considered, the maximum IPP yield drops to 0.57 C-mol/C-mol in *E. coli* and 0.39 C-mol/C-mol in *S. cerevisiae*, respectively.

Hence, *E. coli* has a higher theoretical potential to provide IPP, the common terpenoid precursor.

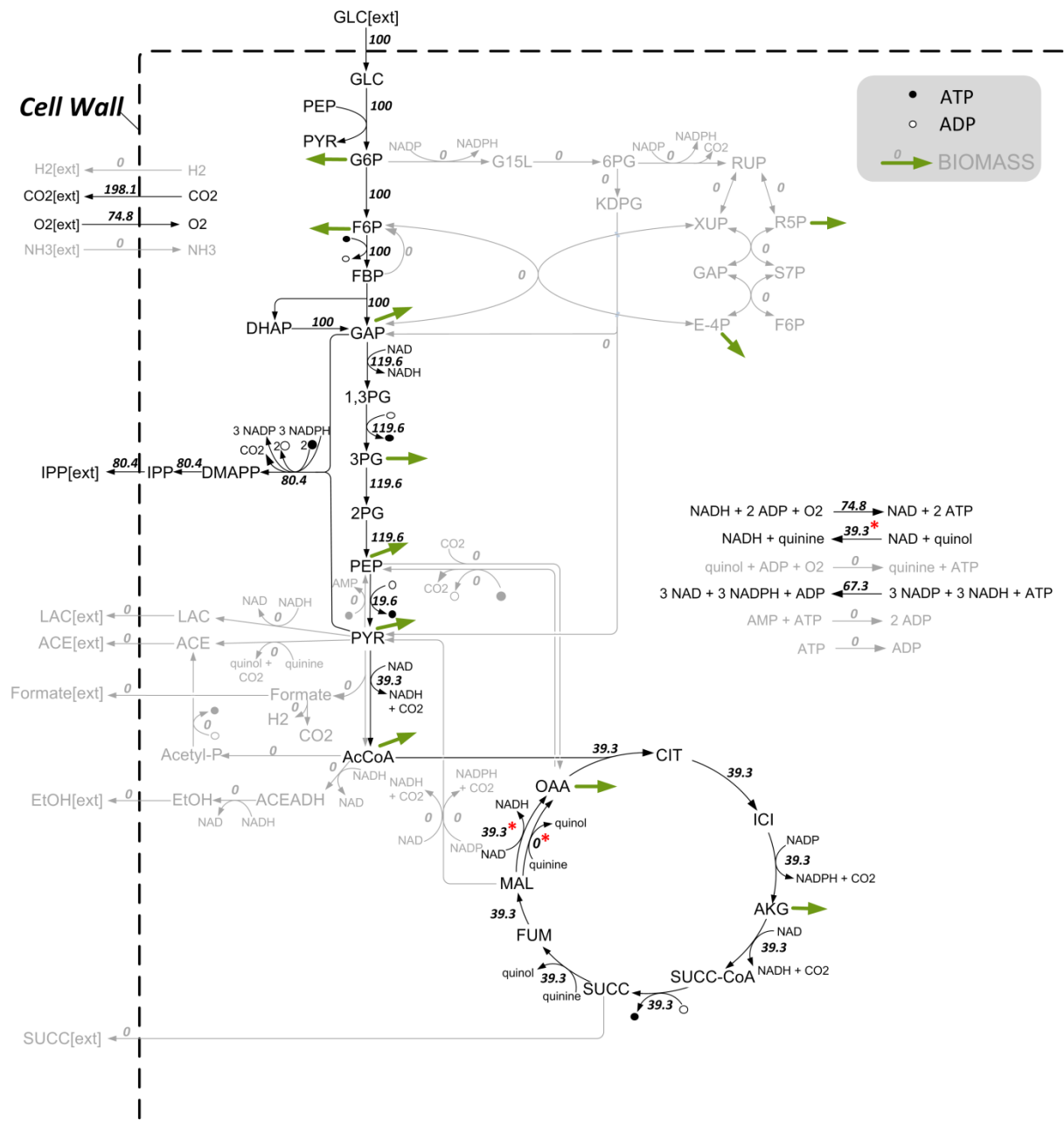


**Figure 2.2 Comparison of the solution space spanned by elementary modes.** **A:** Wild type network of *E. coli* (36,590 EMs). **B:** Wild type network of *S. cerevisiae* (9,844 EMs). The biomass yield on glucose is plotted against the IPP yield on glucose [C-mol/C-mol] for all computed elementary modes. Elementary modes on the axes represent modes producing either only biomass (x-axis) or only IPP (y-axis). Theoretical maximal IPP yields at zero growth as well as with biomass formation are highlighted.

### Comparison of flux distributions with maximum IPP production

The differences in carbon stoichiometry of both IPP forming pathways, in energy and redox equivalent requirements as well as in the two hosts' central carbon metabolisms lead to different theoretical maximal IPP yields. A comparison of the theoretical flux distributions in both organisms provides a detailed picture of the reactions involved and reveals key pathways contributing to maximum IPP production as well as pathways that could be dispensable. Furthermore, they show limitations and provide valuable starting points for metabolic engineering. For *E. coli* and *S. cerevisiae*, two and five elementary modes enable the optimum yield, respectively. In this subsection, main characteristics of these optimal flux distributions are discussed. The flux maps of one particular optimal mode for each case are shown in Figures 2.3 and 2.4, respectively.



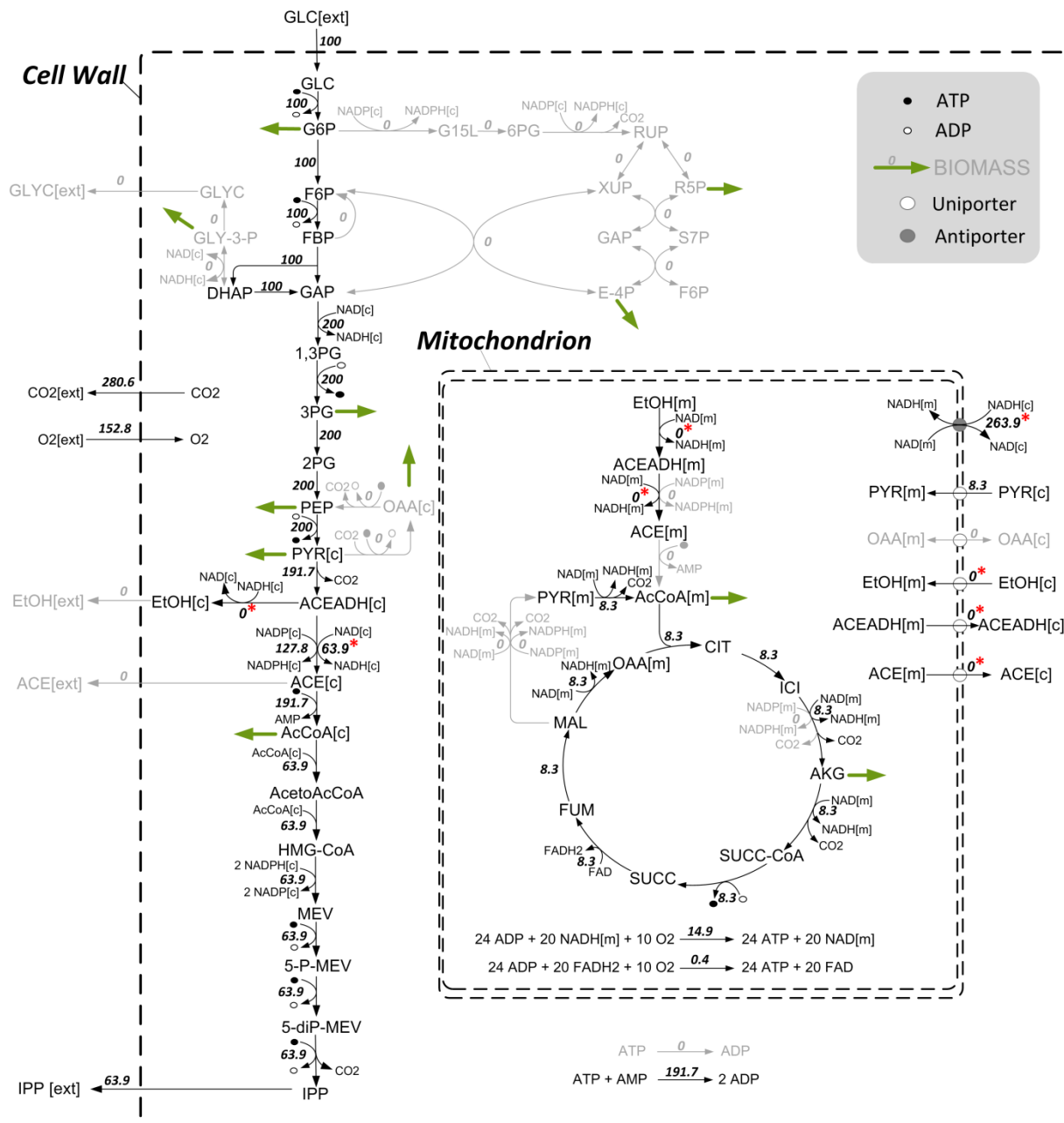


**Figure 2.3 Optimal flux distribution of *E. coli*.** Wild type (DXP pathway) on glucose as single carbon source without biomass formation (one of two modes with maximum theoretical IPP carbon yield). Numbers indicate relative molar fluxes (mmol/(gCDW h), CDW = cell dry weight) normalized to glucose uptake. Grey reactions are not active (flux = 0). Variable fluxes within two optimal modes are marked with an asterisk (\*). Bold green arrows indicate biomass precursors.

**Yield optimal flux distributions in *E. coli*** In *E. coli*, 3 NADPH as well as 2 ATP are needed for the production of one mol IPP from one mol glucose (see above). In the optimal case, NADPH is provided by a transhydrogenase, an enzyme which transfers reducing equivalents between two nucleotide acceptors to balance the intracellular redox potential ( $6 \text{ NADH} + 6 \text{ NADP}^+ + 2 \text{ ATP} = 6 \text{ NADPH} + 6 \text{ NAD}^+ + 2 \text{ ADP}$ ). However, this leads to an increased requirement for ATP and NADH. Thus, the cell is forced to

branch off some carbon flux to the citric acid cycle to generate NADH and ATP as shown in the optimal flux distributions (Figure 2.3). NADPH is not provided by the oxidative part of the pentose phosphate pathway in this case due to increased carbon loss via CO<sub>2</sub> which leads to decreased IPP yields. The main flux passes through glycolysis and DXP pathway, while a portion passes pyruvate dehydrogenase complex and citric acid cycle to generate ATP, NADH and NADPH along with respiratory chain and transhydrogenase. The high branch-off leads to a carbon loss and thus to a reduction in the maximum IPP yield from 0.83 (if carbon stoichiometry is considered only) to about 0.67 C-mol/C-mol in the optimal flux distributions (Figure 2.3, see also Table 2.1). Some fluxes are flexible within optimal elementary modes namely those of malate dehydrogenase, malate quinone oxidoreductase as well as NADH dehydrogenase II indicating a certain flexibility in redox balancing which does not influence the IPP yield. Optimal IPP production does not involve any by-product formation except CO<sub>2</sub> and is purely respirative. However, fermentative modes show high product yields as well, up to 0.48 C-mol/C-mol.

The limitation in redox equivalents and energy for high yield IPP production can be validated *in silico* via the introduction of additional artificial ATP, NADH or NADPH sources in the metabolic network model. The introduction of an artificial ATP source increases the maximum yield slightly to 0.71, while artificial NADH or NADPH sources lead to the theoretical maximum value of about 0.83 C-mol/C-mol (as discussed in the previous section).



**Figure 2.4 Optimal flux distribution of *S. cerevisiae*.** Wild type (MVA pathway) on glucose as single carbon source without biomass formation (one out of five modes with maximum theoretical IPP carbon yield). Numbers indicate relative molar fluxes (mmol/(gCDW h), CDW = cell dry weight) normalized to glucose uptake. Grey reactions are not active (flux = 0). Variable fluxes within the 5 optimal modes are marked with an asterisk (\*). Bold green arrows indicate biomass precursors.

**Yield optimal flux distributions in *S. cerevisiae*** In *S. cerevisiae*, 3 NADH and one NADPH are produced via glycolysis while 6 ATP are needed for the production of one mol IPP from 1.5 moles glucose as described above. Generated NADH can be reoxidized via the respiratory chain to produce ATP. However, not enough ATP can be produced due to the stoichiometry of the respiratory chain. Thus, the cell would have a deficiency in energy for high yield IPP production and would be forced to branch some carbon

flux to the citric acid cycle to generate ATP via NADH. This branch-off is shown in the computed optimal flux distributions (Figure 2.4). The main flux passes glycolysis, pyruvate dehydrogenase bypass and MVA pathway, while a small part is branched off to pyruvate dehydrogenase complex and citric acid cycle in the mitochondria to generate NADH which is then reoxidized via the respiratory chain in order to produce ATP. This branch-off leads to a further carbon loss and thus to a reduction of the IPP yield to about 0.53 C-mol/C-mol in the optimal flux distributions. As shown above, the energy deficiency preventing high yield IPP production of the MVA pathway can be validated via the introduction of an additional artificial ATP or NADH source *in silico* which leads to the theoretical maximum value of about 0.56 C-mol/C-mol. Some fluxes are flexible within optimal elementary modes namely those of mitochondrial and cytosolic alcohol dehydrogenases and NAD<sup>+</sup>-dependent aldehyde dehydrogenases as well as mitochondrial ethanol, acetaldehyde and acetate transport and a redox shuttle indicating a certain flexibility in NADH redox balancing which does not influence the IPP yield. The optimal flux distributions are purely respirative modes. Respiro-fermentative modes (modes that include respiration as well as the formation of fermentation products) show high product yields as well. However, pure fermentative modes are characterized by a very low theoretical maximum product yield of about 0.1 C-mol/C-mol. High ATP requirements of the MVA pathway cannot be delivered by fermentative metabolism. Additionally, the formation of CO<sub>2</sub>, the only side-product, is much higher than in optimal flux distributions in *E. coli* reflecting the high carbon loss in the formation of acetyl-CoA.

Even though theoretical maximal yields differ in *E. coli* and *S. cerevisiae*, the main flux in computed optimal flux distributions at maximum IPP yield is similar. Key pathways are glycolysis and terpenoid pathway and a portion of the flux is directed to citric acid cycle and respiratory chain. Transhydrogenase is active in *E. coli* while pyruvate dehydrogenase bypass is active in yeast to supply NADPH. Dispensable reactions include by-product formation except CO<sub>2</sub>, gluconeogenic and anaplerotic reactions as well as pentose phosphate pathway if no biomass is produced. Limitations in redox equivalents and energy for high yield IPP production could be validated *in silico*.

### **Exchange of terpenoid pathways**

The MVA pathway from *S. cerevisiae* has been functionally introduced in *E. coli* and this strain produced substantially more terpenoid than strains with the native or even

engineered DXP pathway (Martin et al. 2003). Due to this observation the MVA pathway was introduced *in silico* into *E. coli* to analyze its effects on IPP production. The introduction of the MVA and the removal of the DXP pathway (35,968 EMs) lead to a theoretical maximum value of about 0.56 C-mol/C-mol (which is slightly higher than the value of 0.53 obtained for *S. cerevisiae*). No energy or redox equivalent limitations exist for high yield IPP production. *E. coli* possesses a cytosolic pyruvate dehydrogenase complex instead of a pyruvate dehydrogenase bypass (NADPH source). However, transhydrogenase, Entner-Doudoroff pathway as well as malic enzyme can act as NADPH source. Nevertheless, the sole use of the MVA pathway in *E. coli* is stoichiometrically not efficient as the maximum yield is lower than with the native DXP pathway. The coexistence of both pathways in *E. coli* (44,850 EMs, 1,207 of which that use both pathways simultaneously) leads to a theoretical maximum IPP yield of 0.67 C-mol/C-mol, which is identical to the value if only the DXP pathway is active. However, the theoretical maximum value in an elementary mode, which includes biomass formation, is slightly enhanced to 0.59 C-mol/C-mol. Thus, the coexistence of both pathways is not only beneficial due to the decoupling of the MVA pathway from its native control in yeast but also in terms of stoichiometry.

*In silico*, the introduction of the DXP and the removal of the MVA pathway in *S. cerevisiae* (8,865 EMs) lead to an increase in the theoretical maximum IPP yield to about 0.64 C-mol/C-mol, and 0.57 C-mol/C-mol including biomass formation. The coexistence of both pathways in yeast (11,738 EMs, 276 of which that use both pathways simultaneously) does not lead to a further enhanced IPP yield. Solely the number of elementary modes and thus the flexibility is enhanced. The IPP carbon yield is slightly lower than that one of *E. coli* as yeast does not possess a transhydrogenase to supply NADPH. Hence, pentose phosphate pathway has to deliver NADPH which involves a carbon loss. Pyruvate dehydrogenase bypass can deliver NADPH in the formation of acetyl-CoA; however, acetyl-CoA is not needed for the DXP pathway. The limitation in energy and redox equivalents can be validated *in silico* by the introduction of artificial ATP, NADH and especially NADPH sources in the model, which lead to an enhanced theoretical maximum IPP yield of 0.66-0.8 C-mol/C-mol (see Table 2.1).

In conclusion, the heterologous introduction of the MVA pathway, in addition to the DXP pathway, into *E. coli* as well as the introduction of the DXP pathway into *S. cerevisiae* are theoretically beneficial for high yield IPP production. Although, the DXP pathway has

not been functionally expressed in *S. cerevisiae* yet (Carlsen et al. 2013; Partow et al. 2012), it is a promising target for further research.

### Comparison of carbon sources

Studies showed that different carbon sources might have a different potential to supply high product yields (Martin et al. 2003; Westfall et al. 2012; Yoon et al. 2009). At present, glucose is the most widely used feedstock in bioprocesses (Dobson, Gray, and Rumbold 2012). However, other carbon sources are of industrial interest and could serve as substrate for terpenoid production. Molasses consists mainly of sucrose, a disaccharide composed of glucose and fructose, and is a commonly used carbon source (Randez-Gil, Sanz, and Prieto 1999). Hemicellulosic hydrolysates of agricultural by-products like sugarcane bagasse or wood consist mainly of xylose and glucose as well as other sugars like galactose (Carvalho, Duarte, and Gírio 2008) and could be a cheap alternative carbon source for bioprocesses (Pandey et al. 2000). Pure glycerol is used in different industries too and crude glycerol could be a cheap alternative as it is a by-product from biodiesel production with decreasing price (da Silva, Mack, and Contiero 2009; Dobson, Gray, and Rumbold 2012). Thus, EMA was further performed for glycerol for *E. coli*. Moreover, the carbon sources galactose, fructose and xylose were analyzed for *S. cerevisiae* as well as glycerol and ethanol because high terpenoid yields have been described with this substrate (Westfall et al. 2012). The metabolic networks of the non-fermentable carbon sources ethanol and glycerol involve the glyoxylate cycle and an increased number of mitochondrial shuttles in yeast.

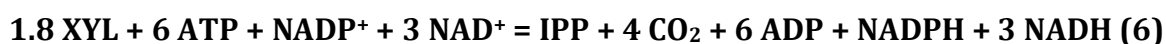
**Alternative substrates for *E. coli*** The maximum IPP yield in *E. coli* on the non-fermentable carbon source glycerol purely based on carbon stoichiometry (ATP and NAD(P)H considered as external metabolites) is identical to that on glucose (0.83 C-mol/C-mol). The overall stoichiometry for the formation of one mol IPP from glycerol (GLYC) is given by equation 5.



The carbon loss for the formation of IPP from glycerol is identical to glucose. However, less NADPH is required as it is supplied by glycerol-3-phosphate dehydrogenase. If the balances of ATP, NADH and NADPH are taken into account, the yields for *E. coli* and the native DXP pathway on glycerol (26,160 EMs) were enhanced in comparison to glucose *in silico* (0.78 C-mol/C-mol at zero growth and 0.77 C-mol/C-mol in an elementary mode that includes biomass formation). Yields were slightly lower for

the heterologous MVA pathway (0.56 C-mol/C-mol at zero growth and 0.45 C-mol/C-mol including biomass formation). Thus, glycerol and the native DXP pathway have the highest potential for *E. coli* to support high IPP production. The main flux in optimal flux distributions passes lower part of glycolysis and DXP pathway, while only a small portion is diverted to citric acid cycle, respiratory chain and transhydrogenase to generate NADPH, NADH and ATP. The limitation in energy and redox equivalents for high yield IPP production can be validated *in silico* by the introduction of artificial ATP, NADH or NADPH sources in the model, which lead to the maximum IPP yield of 0.83 C-mol/C-mol (see Table 2.1).

**Alternative substrates for *S. cerevisiae*** The carbon as well as the overall stoichiometry for the formation of one mol IPP from galactose or fructose is identical to glucose (see equation 4). Accordingly, the numbers of elementary modes as well as the IPP and biomass yields with galactose and fructose are identical to those with glucose. A heterologous pathway has to be introduced into *S. cerevisiae* for xylose as wild type strains are not able to grow on the pentose xylose. Two different pathways have been used for ethanol production based on xylose with yeast: the xylose isomerase (XI) pathway as well as the xylose reductase and xylitol dehydrogenase (XR-XDH) pathway (Karhumaa et al. 2007). The carbon stoichiometry for the formation of one mol IPP from xylose is identical to glucose (maximum 0.56 C-mol/C-mol). Likewise, the overall stoichiometry is identical to the one on glucose for the XI pathway except that 1.8 mol xylose (XYL) is needed per mol IPP (equation 6).



The overall stoichiometry is identical for the XR-XDH pathway if xylose reductase uses NADH (as xylitol dehydrogenase delivers NADH) and slightly changes if the enzyme uses NADPH. The theoretical maximal IPP and biomass yields are identical to the values on glucose for both the XI (6,330 EMs) and the XR-XDH pathway (16,911 EMs). However, the IPP yield including biomass formation is enhanced to 0.51 C-mol/C-mol with the XR-XDH pathway. The number of elementary modes and thus the flexibility of the metabolic network is greatly enhanced if the XR-XDH pathway is chosen. The optimal flux distributions were basically identical to those on glucose except that xylose enters glycolysis via the non-oxidative part of pentose phosphate pathway.

The maximum IPP yield on the non-fermentable carbon source glycerol (GLYC) purely based on carbon stoichiometry (ignoring energy and redox equivalents) is with

0.56 C-mol/C-mol identical to glucose. However, the overall stoichiometry is slightly modified (equation 7).



Less energy is required and more NADH is delivered for the production of one mol IPP from glycerol. Glycerol is channeled into glycolysis via glyceraldehyde-3-phosphate and the main flux passes lower part of glycolysis, pyruvate dehydrogenase bypass and the MVA pathway. The carbon flux is not diverted to citric acid cycle as enough energy is produced via NADH and respiratory chain. Thus, the maximum IPP yield of 0.56 C-mol/C-mol is achieved, which is identical to the maximum IPP yield based on carbon stoichiometry only. No limitations in energy or redox equivalent requirements thus exist for high yield IPP production. The optimal flux distributions further reveal that pentose phosphate pathway, side product formation, the first part of citric acid cycle as well as glyoxylate cycle are not active. Some fluxes are flexible namely those of alcohol dehydrogenases, NAD<sup>+</sup>-dependent aldehyde dehydrogenases, second part of citric acid cycle as well as mitochondrial transport systems indicating a certain flexibility in NADH redox balancing which does not influence the IPP yield. The very high number of elementary modes (2,648,133 EMs) implies an increased redundancy and thus network flexibility and is due to the activity of the glyoxylate cycle. The glyoxylate cycle showed only a weak activity in yeast cells grown on glycerol in our experiments (data not shown). If the cycle is excluded from the computations, the number of elementary modes decreases (5,377 EMs) but the theoretical maximum IPP yield remains 0.56 C-mol/C-mol. Nevertheless, the IPP yield including biomass formation is much higher if the cycle is included (0.47 C-mol/C-mol versus 0.35 C-mol/C-mol) indicating that an active glyoxylate cycle could be of benefit.

The maximum IPP yield on the non-fermentable carbon source ethanol purely based on carbon stoichiometry is with 0.83 C-mol/C-mol as high as the IPP yield in *E. coli*. The overall stoichiometry for the formation of one mol IPP from ethanol (EtOH) is given by equation 8.



In contrast to glucose or glycerol as the carbon source, the formation of IPP from ethanol requires more energy but does not involve an additional carbon loss via CO<sub>2</sub> (except the carbon loss in the terpenoid pathway itself). In fact, similarly to glycerol and *E. coli*, ethanol (2,874,235 EMs) reveals an increased theoretical maximum product yield of 0.68 C-mol/C-mol at zero growth as well as 0.59 C-mol/C-mol including biomass



formation for the MVA pathway. The high number of elementary modes is again due to the activity of the glyoxylate cycle. The optimal flux distributions of yeast (without biomass formation) with only the MVA pathway reveal that the glycolysis, gluconeogenesis, pentose phosphate pathway and side product formation are not active. The flux passes pyruvate dehydrogenase bypass, glyoxylate cycle as well as citric acid cycle and respiratory chain to generate ATP. The energy deficiency preventing high IPP yields can be validated *in silico* by introducing artificial ATP or NADH sources which lead to the maximum yield of 0.83 C-mol/C-mol. As expected from these computations, the sesquiterpenoid yield of the same yeast strain was increased significantly *in vivo* by changing the carbon source from glucose (0.0111 C-mol/C-mol) to ethanol (0.1867 C-mol/C-mol) (Westfall et al. 2012).

The introduction of the DXP pathway *in silico* and the growth on galactose, fructose or xylose lead to similar yields to those on glucose, except that the theoretical maximal IPP yields including biomass formation are slightly increased on xylose. With ethanol as substrate, the heterologous DXP pathway shows a lower potential compared to the native MVA pathway, with yields of 0.63 C-mol/C-mol at zero growth and 0.57 C-mol/C-mol including biomass formation. Lower yields are due to energy-consuming gluconeogenic reactions which are essential to provide the precursors of the DXP pathway from ethanol. Nevertheless, the introduction of the DXP pathway could be beneficial for growth on glycerol. The theoretical maximum IPP yield is enhanced to 0.67 C-mol/C-mol at zero growth and 0.6 C-mol/C-mol including biomass formation. The cells would have a limitation in NADPH for high yield IPP production, which can be validated via the introduction of an artificial NADPH source leading to the maximum value of 0.83 C-mol/C-mol.

To summarize this subsection, sugars like glucose, galactose, fructose and even xylose lead to identical or similar theoretical maximal IPP yields in *S. cerevisiae* while the non-fermentable carbon sources ethanol and glycerol show the highest potential with respect to yield and network flexibility. Similarly, glycerol shows the highest potential in *E. coli* to supply IPP in high yields.

### **Comparison of theoretical yields with experimental yields reported in literature**

A comparison of computed theoretical maximal IPP yields with terpenoid yields published in literature is shown in Table 2.2. The experimental carbon yields are still very low, even though strains used in these studies had been genetically modified within

the respective terpenoid pathway. This highlights the tremendous potential of yield improvement. The terpenoid yield for yeast on ethanol was quite high in a study (Westfall et al. 2012) indicating that the capacity of the terpene synthase or the terpenoid pathway (MVA) are not the only limiting factors but that the flux into these pathways is a limiting factor on other carbon sources like glucose. This shows that, besides the already very well-studied modifications within the terpenoid pathway, genetic modifications within the central carbon metabolism are indispensable to efficiently redirect the carbon flux to the terpenoid pathway ('push and pull' strategy).

**Table 2.2 Comparison of theoretical maximal IPP yields with published terpenoid yields.** IPP yields are given as C-mol/C-mol. Experimental yields of sesquiterpenoids are partly estimated and converted into IPP yields (only to that part of IPP that has been converted to the respective terpenoid); + indicates complex media.

Organism	Pathway	Carbon source	Maximum IPP yield (this study)	Experimental IPP yield estimated from published yield	
<i>E. coli</i>	DXP	glycerol+	0.79	0.04	(Meng et al. 2011)
<i>E. coli</i>	DXP	glycerol+	0.79	0.06	(Ajikumar et al. 2010)
<i>E. coli</i>	DXP+MVA	glycerol+	0.79	0.04	(Wang et al. 2010)
<i>S. cerevisiae</i>	MVA	glucose+	0.53	0.003	(Muramatsu et al. 2009)
<i>S. cerevisiae</i>	MVA	glucose+	0.53	0.00006	(Lindahl et al. 2006)
<i>S. cerevisiae</i>	MVA	galactose	0.53	0.004	(Asadollahi et al. 2008)
<i>S. cerevisiae</i>	MVA	glucose	0.53	0.0052	(Scalcinati et al. 2012)
<i>S. cerevisiae</i>	MVA	glucose	0.53	0.01	(Westfall et al. 2012)
<i>S. cerevisiae</i>	MVA	glucose & ethanol	0.53/0.68	0.06	(Westfall et al. 2012)
<i>S. cerevisiae</i>	MVA	ethanol	0.68	0.19	(Westfall et al. 2012)

## 2.4.2 Target identification for metabolic engineering

### Identification of overexpression targets

The identified limitations in the generation of sufficient energy and redox equivalents for high yield IPP production with both organisms can be used as starting point for metabolic engineering. Evidently, a reasonable approach would be to implement heterologous enzymes that are more efficient with respect to NAD(P)H or ATP consumption.

For *E. coli*, an increase in the maximum yield could be accomplished by the introduction of alternative pathways with less energy requirements or additional NAD(P)H sources. For instance, an additional heterologous NADP<sup>+</sup>-dependent glyceraldehyde-3-phosphate dehydrogenase enhances the theoretical maximum IPP yield to 0.71 C-mol/C-mol on glucose and to 0.83 C-mol/C-mol on glycerol.

For *S. cerevisiae*, an increase in the maximum IPP yield could be accomplished by the introduction of alternative pathways with less energy requirements. *S. cerevisiae* possesses a mitochondrial pyruvate dehydrogenase complex and a cytosolic pyruvate dehydrogenase bypass to produce acetyl-CoA. Solely acetyl-CoA produced via the bypass, which requires ATP, can be used for terpenoid production since acetyl-CoA cannot be exported from the mitochondria. In contrast, *E. coli* possesses a cytosolic pyruvate dehydrogenase complex, which does not require energy in the form of ATP. *In silico*, a heterologous introduction of cytosolic pyruvate dehydrogenase complex from *E. coli* into *S. cerevisiae* leads to the theoretical maximum IPP yield of 0.56 C-mol/C-mol and to a mode enabling biomass formation coupled to an increased IPP yield of 0.51 C-mol/C-mol on glucose. A cytosolic pyruvate dehydrogenase complex has to our knowledge not been introduced into yeast with the aim of enhancing terpenoid production. Though, a patent application indicates that the complex of *E. coli* (*lpdA*, *aceA*, *aceF*) might increase the cytosolic acetyl-CoA pool for butanol production (Gunawardena et al. 2010).

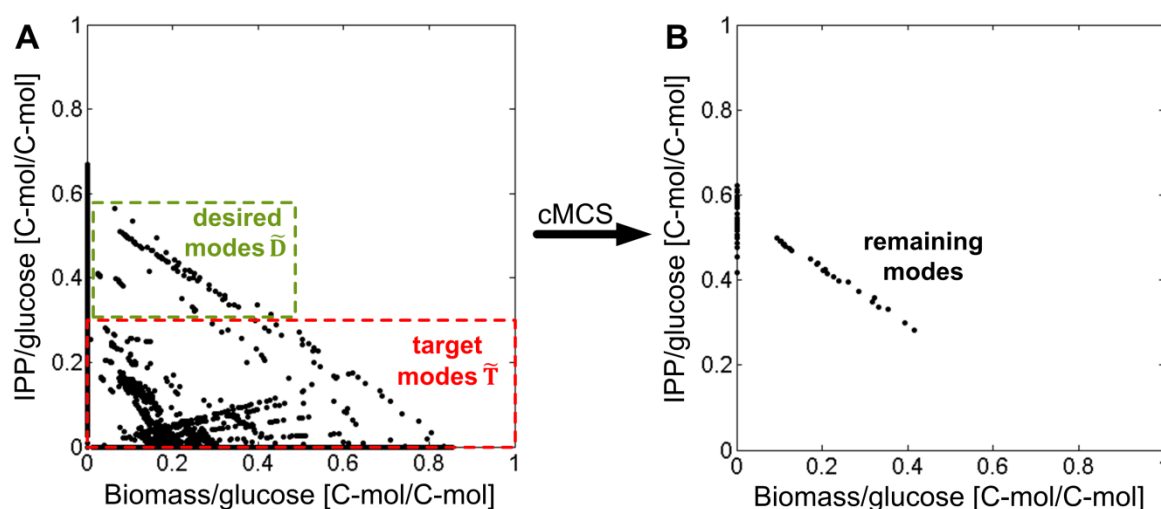
A different approach would be to transfer the native MVA pathway from *S. cerevisiae* into the mitochondria. Likewise, on glucose, the theoretical maximum IPP yield at zero growth is enhanced to 0.56 C-mol/C-mol and the IPP yield including biomass formation is enhanced to 0.51 C-mol/C-mol, if only the MVA in the mitochondria is active, or to 0.56 and 0.53 C-mol/C-mol if the MVA is active in the cytosol and the mitochondria. The mitochondrial IPP and DMAPP as well as the FPP pool have been harnessed successfully for terpenoid production (Farhi et al. 2011). However, the whole MVA pathway has not been transferred to the mitochondria. The advantage of this approach would be that the mitochondrial acetyl-CoA pool, which is formed via the pyruvate dehydrogenase complex that does not require ATP, would be used.

Alternatively, an ATP-citrate-lyase can be expressed heterologously in the cytosol. The enzyme forms cytosolic acetyl-CoA from citrate that has been exported from the mitochondria and thus circumvents pyruvate dehydrogenase bypass. The reaction is accompanied by the conversion of one mol ATP to ADP per acetyl-CoA, while the cytosolic pyruvate dehydrogenase bypass converts in the net one mol ATP to AMP. Thus, ATP-citrate-lyase is more energy-efficient and thus promising. The theoretical maximum IPP yield at zero growth is indeed enhanced to 0.55 C-mol/C-mol. Moreover, a patent application indicates that this approach might be promising (Hansen 2013).

If the DXP pathway is introduced into yeast, a heterologous soluble transhydrogenase leads to a higher theoretical maximum IPP yield of 0.67 C-mol/C-mol on glucose. However, the introduction of a heterologous soluble transhydrogenase from *Azotobacter vinelandii* as well as the introduction of a membrane-bound enzyme from *E. coli* into *S. cerevisiae* led to a reduction of the NADPH pool including an increase in glycerol formation (Anderlund et al. 1999; Jeun et al. 2003; Nissen et al. 2001), which is the opposite of the desired effect. Though, the introduction of a heterologous NADP<sup>+</sup>-dependent glyceraldehyde-3-phosphate dehydrogenase leads *in silico* to the same improvement in IPP yield on glucose and to a yield of 0.76 C-mol/C-mol on glycerol.

### Identification of knockout strategies

The constrained minimal cut sets (cMCSs) approach allows the identification of all possible knockout combinations that eliminate undesired (low-yield) elementary modes and keep desired modes, which (i) have a specified minimum product yield and (ii) allow at least some biomass formation (Hädicke and Klamt 2011). Hence, a selection pressure imposed by gene deletions forces the cell to produce a predefined minimum IPP yield to be able to grow. However, optimal biological functions are acquired through evolutionary processes. For *in vivo* validation, it has been shown that adaptive evolution is essential to improve production capabilities from sub-optimal metabolic states to optimal ones as predicted from *in silico* analysis (Fong et al. 2005; Ibarra, Edwards, and Palsson 2002).



**Figure 2.5** Concept of cMCSs illustrated with an example of *E. coli*. **A**: Solution space formed by the elementary modes of the wild type network of *E. coli* (36,590 EMs; compare figure 2.2). The biomass

yield on glucose is plotted against the IPP yield on glucose [C-mol/C-mol] of all computed elementary modes. The deletion task is defined by collecting all modes exhibiting an IPP yield lower than 0.25 C-mol/C-mol in the set of target modes  $\tilde{\mathbf{T}}$  (33,087 EMs). The set of desired modes  $\tilde{\mathbf{D}}$  comprises all EMs having (i) a minimum IPP yield of 0.25 C-mol/C-mol; and (ii) a simultaneous biomass yield strictly higher than zero (268 EMs). **B:** Example of one specific cMCS. The *in silico* knockout of transhydrogenase and phosphoglycerate mutase (strategy 2 in table 3) leads to a reduction of elementary modes. Remaining elementary modes (191) display a minimum IPP yield of 0.28 and a maximum of 0.62 C-mol/C-mol as well as a maximum biomass yield of 0.41 C-mol/C-mol on glucose. In those modes, biomass formation is coupled to a minimum yield of IPP production.

In this study, cMCSs are used to identify combinations of gene deletions for both *E. coli* and *S. cerevisiae* to present new intervention targets for metabolic engineering.

**cMCSs for *E. coli*** Constrained minimal cut sets were computed for the *E. coli* wild type metabolic network on glucose (36,590 EMs) as well as glycerol as carbon source (26,160 elementary modes). For glucose, the deletion task was defined by collecting all elementary modes exhibiting an IPP yield lower than 0.25 C-mol/C-mol in the set of target modes  $\tilde{\mathbf{T}}$  (33,087 EMs). The set of desired modes  $\tilde{\mathbf{D}}$  comprises all EMs having (i) a minimum IPP yield of 0.25 C-mol/C-mol; and (ii) a simultaneous biomass yield necessarily higher than zero (268 modes) (see Figure 2.5). When at least one EM of the set of  $\tilde{\mathbf{D}}$  should be preserved ( $n=1$ ; see methods), a total of 488 cMCSs accomplishing the specified engineering goal were computed. Cut sets including diffusion processes or parts of respiratory chain were excluded from further considerations due to poor feasibility leaving 28 cMCSs with two or four interventions (see Table 2.3). One group of cut sets forces a coupling of IPP production with biomass formation via redox equivalents. This is achieved by preventing oxygen uptake and thus respiration as well as side product formation and reactions within citric acid cycle that can reoxidize NADH (see Table 2.3). Hence, NADH has to be reoxidized via transhydrogenase which leads to excess NADPH production. Excess NADPH is then reoxidized via the DXP pathway. Applying these cut sets, the remaining elementary modes are anaerobic, exhibit a minimum IPP carbon yield of 0.38 and the maximum anaerobic IPP yield of 0.48. However, the maximum possible biomass carbon yield is very low. A second group of cut sets includes the knockout of the lower part of glycolysis to force the cell to use the Entner-Doudoroff pathway instead, which results in less ATP but more NADPH. The cell cannot reoxidize excess NADPH by the additional knockout of transhydrogenase. Excess NADPH is then reoxidized via the DXP pathway. Remaining elementary modes are

aerobic, exhibit a slightly lower minimum IPP carbon yield but have a much higher maximum biomass carbon yield than the anaerobic modes.

**Table 2.3 cMCSs for *E. coli*.** Number of remaining elementary modes (EMs), minimum and maximum IPP yield as well as maximum biomass yield on glucose or glycerol for different knockout strategies. All yields given as C-mol/C-mol. DH = dehydrogenase; ex = excretion.

cMCS knockout strategy	EMs	Min. IPP yield	Max. IPP yield	Max. biomass yield
<b>Wild type on glucose</b>	36,590	-	0.67	0.85
<b>Strategy 1:</b> O <sub>2</sub> uptake + {lactate ex OR lactate formation} + {ethanol ex OR ethanol formation OR acetaldehyde formation} + {succinate ex OR succinate OR malate DH OR fumarate hydratase}	60-86	0.38	0.48	0.08
<b>Strategy 2:</b> transhydrogenase + {phosphoglycerate mutase OR enolase}	191	0.28	0.62	0.41
<b>Strategy 3:</b> transhydrogenase + {GAP DH OR phosphoglycerate kinase}	191	0.35	0.62	0.32
<b>Wild type on glycerol</b>	26,160	-	0.78	0.99
<b>Strategy 1:</b> transhydrogenase + pyruvate kinase + pyruvate DH complex + ribulose-phosphate 3-epimerase + acetate formation + glyoxylate shunt + {succinate DH OR fumarate hydratase}	22	0.38	0.68	0.40

Another cut set included glucose-6-phosphate-isomerase, transhydrogenase, Entner-Doudoroff pathway as well as one specific part of respiratory chain. The cMCS is not shown in Table 2.3 due to poor feasibility of the last target. However, it serves as an example because it has been implemented. The coupling is again due to excess NADPH production. The knockout of glucose-6-phosphate-isomerase ( $\Delta pgi$ ) has been done in *E. coli in vivo* and led to an increased flux through pentose phosphate pathway and thus to excess NADPH formation. However, an additional knockout of transhydrogenase ( $\Delta pgi \Delta udhA$ ) was lethal due to NADPH imbalance in this wild type strain (Sauer et al. 2004). This example illustrates that the principle of the method is effective but regulatory constraints may interfere with proposed knockout strategies. One can assume that the overexpression of a terpenoid synthase together with the DXP pathway as a sink and eventually adaptive evolution could balance the excess NADPH production via a ‘push and pull’ strategy and lead to the predicted coupling.

For glycerol, the deletion task was set up by (i) defining  $\tilde{\mathbf{T}}$  as the set of elementary modes exhibiting an IPP yield of not more than 0.33 C-mol/C-mol (22,801 EMs); (ii) defining  $\tilde{\mathbf{D}}$  as all EMs with a minimum IPP yield of 0.33 C-mol/C-mol and a simultaneous

biomass formation strictly higher than zero (570 EMs); and (iii) demanding that at least one EM of  $\tilde{\mathbf{D}}$  is preserved ( $n=1$ ). A total of 769 cMCSs were computed, while cut sets that cannot be realized in practice (including diffusion processes, parts of respiratory chain, ATP maintenance etc.) were excluded from further considerations leaving only 8 cMCSs with eight interventions each (see Table 2.3). Feasible minimal cut sets include the knockout of transhydrogenase, pyruvate kinase, pyruvate dehydrogenase complex, ribulose-phosphate 3-epimerase, acetate formation, glyoxylate shunt and citric acid cycle (succinate dehydrogenase or fumarate hydratase). A minimum IPP carbon yield of 0.38 C-mol/C-mol can be guaranteed but only two of the remaining elementary modes enable biomass formation.

The presented minimal cut set strategies can also be applied to *E. coli* strains carrying the additional MVA pathway. The IPP and biomass yields of remaining elementary modes are only slightly modified (results not shown).

The cMCSs approach points to new and promising combinations of metabolic engineering targets. The predicted minimal IPP yields are much higher than the yields that have been published in literature so far. Nevertheless, kinetics and regulation, especially of the DXP pathway, may impose additional challenges to be solved in order to achieve high terpenoid yields with *E. coli*.

**cMCSs for *S. cerevisiae*** Constrained minimal cut sets were computed based on the wild type metabolic network with glucose as carbon source (9,844 EMs). Modes exhibiting an IPP yield of not more than 0.25 C-mol/C-mol were collected in the set of target modes  $\tilde{\mathbf{T}}$  (9,453 EMs). All EMs exhibiting (i) an IPP yield of not less than 0.25 C-mol/C-mol; and (ii) a concurrent biomass yield strictly higher than zero were defined as desired modes  $\tilde{\mathbf{D}}$  (322 EMs). While preserving at least one of the desired modes ( $n=1$ ), a total of 407 cMCSs were computed. Cut sets including mitochondrial transport systems, diffusion, ATP maintenance etc. were excluded from further considerations leaving 8 feasible cMCSs with three to six interventions (Table 2.4). All remaining cMCS require blocking of acetate secretion and either of ethanol secretion or of all alcohol dehydrogenases (cytosolic and mitochondrial). The prevention of acetate and ethanol production is crucial as the engineering goal cannot be accomplished without those two targets. An additional target is the partial knockout of citric acid cycle after  $\alpha$ -ketoglutarate, namely  $\alpha$ -ketoglutarate dehydrogenase or succinyl-CoA ligase (succinate dehydrogenase or fumarate hydratase would be an alternative as well but were not considered due to possible succinate secretion). Alternatively malate dehydrogenase

and malic enzyme can be deleted for a partial knockout of citric acid cycle. A second alternative is malate dehydrogenase, NAD<sup>+</sup>-dependent isocitrate dehydrogenase as well as pyruvate dehydrogenase complex.

**Table 2.4 cMCSs for *S.cerevisiae*.** Number of remaining elementary modes (EMs), minimum and maximum IPP carbon yield as well as maximum biomass carbon yield on glucose [C-mol/C-mol] for different knockout strategies. DH = dehydrogenase; ex = excretion.

cMCS knockout strategy	EMs	Min. IPP yield	Max. IPP yield	Max. biomass yield
<b>Wild type</b>	9,844	-	0.53	0.80
<b>Strategy 1:</b> acetate ex + {ethanol ex OR alcohol DH} + { $\alpha$ -ketoglutarate DH OR succinyl-CoA ligase}	142/48	0.33	0.53	0.29
<b>Strategy 2:</b> acetate ex + {ethanol ex OR alcohol DH} + malate DH + malic enzyme	60/21	0.33	0.53	0.29
<b>Strategy 3:</b> acetate ex + {ethanol ex OR alcohol DH} + malate DH + NAD <sup>+</sup> -dependent isocitrate DH + pyruvate DH complex	51/19	0.27	0.53	0.37
<b>Wild type plus DXP pathway</b>	11,738	-	0.64	0.80
<b>Strategy 1:</b> glucose-6-phosphate-isomerase	631	0.37	0.63	0.35

Possible knockout combinations were simulated and Table 2.4 shows the number of remaining elementary modes and the minimal and maximal IPP and biomass carbon yields when applying the respective cMCS. All minimal cut sets lead to a reduced maximum biomass yield and a reduced number of elementary modes. However, a certain flexibility is preserved. The IPP carbon yield can reach the maximum value of the wild type, and a minimum value of 0.27 or 0.33 is computed. Due to the partial disruption of citric acid cycle the cell is not able to convert glucose completely to CO<sub>2</sub>. Hence, the cell is forced to secrete products like ethanol or acetate. The repression of acetate and ethanol excretion forces the cell to find another sink like IPP. Targets identified for glucose can also be applicable to other sugars like galactose or fructose.

The knockout of succinyl-CoA ligase ( $\Delta LSC1$ ) was implemented in *S. cerevisiae* and an enhanced acetate secretion was observed (Blank, Kuepfer, and Sauer 2005). This indicates that the principle of the method is effective even though acetate is produced instead of channeling the flux to acetyl-CoA, the precursor of the MVA pathway. Nevertheless, acetate secretion can be lowered by the overexpression of acetyl-CoA synthetase (Shiba et al. 2007). Moreover, an overexpression of the MVA pathway together with a terpenoid synthase might channel the flux to IPP via the ‘push and pull’ strategy.



cMCSs including a heterologous DXP pathway are basically identical to those of the wild type. Thus, presented strategies can also be applied to a strain carrying an additional DXP pathway if the pathway was active *in vivo*. However, the minimal IPP yields are unchanged and thus the heterologous DXP pathway would be of no benefit. Nevertheless, one additional interesting intervention strategy was computed. The single knockout of glucose-6-phosphate-isomerase leads to a coupling of biomass formation to a minimal IPP yield of 0.37 C-mol/C-mol. The knockout of glucose-6-phosphate-isomerase forces the cell to direct the flux completely to pentose phosphate pathway, which leads to excess NADPH production. Wild type yeast cells are not able to compensate excess NADPH as they do not possess a transhydrogenase. Further, NADH dehydrogenases (Ndi1p, Nde1p, Nde2p) do not oxidize NADPH and malic enzyme reaction seems to be irreversible *in vivo* (Boles, de Jong-Gubbels, and Pronk 1998; Luttik et al. 1998). Yeast cells are not able to grow on glucose after the knockout of *PGI1* gene encoding glucose-6-phosphate-isomerase (Aguilera 1986) and no elementary modes could be computed for the wild type network including only the MVA pathway with the single knockout of glucose-6-phosphate-isomerase. Nevertheless, an active DXP pathway could serve as a NADPH sink as it requires more NADPH and less ATP per mol IPP than the MVA pathway.

Minimal cut sets were also computed for the metabolic network with xylose as carbon source including either the XI or the XR-XDH pathway together with the MVA pathway. Identified knockout targets as well as yields after the implementation of a minimal cut set for the XI pathway are identical to those of glucose. Knockout targets for the XR-XDH pathway are different. Minimal cut sets only exist with an additional target, ribulose-5-phosphate-3-epimerase in pentose phosphate pathway. Cut sets including pyruvate dehydrogenase complex do not exist. Moreover, possible IPP yields after the implementation of a minimal cut set are slightly different and can vary between 0.30 and 0.50 C-mol/C-mol.

To our knowledge, a disruption of citric acid cycle has not been done with the aim to enhance terpenoid production. Predicted minimal IPP yields are much higher than yields that have been achieved in literature so far. Thus, targets identified by the minimal cut sets approach are completely new and promising for an enhanced terpenoid yield on sugars with *S. cerevisiae*. Moreover, the knockout of glucose-6-phosphate-isomerase appears as a very promising strategy to enhance terpenoid production if the DXP pathway can be functionally expressed in yeast.

### 2.4.3 Conclusions

The use and optimization of microorganisms for an ecologically and economically efficient production of plant terpenoids is a contemporary issue in academic research as well as industry. The most widely used model organisms *E. coli* and *S. cerevisiae* as well as their respective terpenoid pathway, heterologous enzymes and different carbon sources have been analyzed *in silico* in this study with respect to their potential as terpenoid factories. The DXP pathway of *E. coli* has a tremendously higher potential for high IPP yields (nearly 50 % higher) than the MVA pathway of yeast with respect to carbon stoichiometry due to their different precursors. However, kinetics, regulation and a balanced overexpression have to be considered for *in vivo* production. Further research should focus on the optimization of the DXP pathway to tap its full potential. *E. coli* wild type shows a higher theoretical potential to supply IPP in high yields than *S. cerevisiae* wild type. However, the theoretical maximal yields can be enhanced in both organisms by introducing heterologous enzymes or whole pathways or by simply changing the carbon source. Moreover, new constrained minimal cut set strategies can be applied to both hosts to enforce a coupling of growth to a minimum IPP yield which is higher than any yield published in scientific literature so far.

## 2.5 Supplementary information

The metabolic networks of *S. cerevisiae* and *E. coli* have been constructed considering the current knowledge from genome-scale models and literature (Cherry et al. 2012; Herrgard et al. 2008; Kanehisa et al. 2012; Keseler et al. 2011; Nookaew et al. 2008; Schomburg et al. 2013) as well as the literature cited in the following tables. Since H<sub>2</sub>O and P<sub>i</sub> are assumed to be ubiquitous and CoA-SH is part of a conservation relation, these species are not considered explicitly. For *S. cerevisiae*, reactions and metabolites are compartmentalized between cytosolic [c] and mitochondrial [m] compartments.

### 2.5.1 Metabolic networks of *S. cerevisiae*

**Table 2.5** Wild type metabolic network of *S. cerevisiae* on glucose as single carbon source.

Enzyme/Description	Gene (systematic name)	Gene (standard name)	Reaction
<b>Transport</b>			
	YFL011W	HXT10	
	YOL156W	HXT11	
	YEL069C	HXT13	
	YNL318C	HXT14	
	YDL245C	HXT15	
	YJR158W	HXT16	
	YNR072W	HXT17	
Transport: hexose transporter (HXT 1-11, 13-17) and further genes (GAL2, MAL11, MPH2,3) (Wieczorke et al. 1999)	YDR345C	HXT3	==> glucose
	YHR092C	HXT4	
	YHR096C	HXT5	
	YDR343C	HXT6	
	YDR342C	HXT7	
	YJL214W	HXT8	
	YJL219W	HXT9	
	YHR094C	HXT1	
	YMR011W	HXT2	
	YLR081W	GAL2	
	YGR289C	MAL11	
	YDL247W	MPH2	
	YJR160C	MPH3	
	Transport	-	
Transport	-	-	glycerol ==>
Transport	-	-	acetate[c] ==>
Transport	-	-	CO <sub>2</sub> <==>
Transport	-	-	==> O <sub>2</sub>
Artificial IPP production	-	-	isopentenyl-diphosphate ==>
Growth in mmol*gCDW <sup>-1</sup> *h <sup>-1</sup>	-	-	0.1 BIOMASS ==>
<b>Glycolysis</b>			

**Table 2.5** continued

Enzyme/Description	Gene (systematic name)	Gene (standard name)	Reaction
Glucose kinase [EC 2.7.1.2], hexokinase [EC:2.7.1.1]	<i>YCL040W</i> <i>YFR053C</i> <i>YGL253W</i>	<i>GLK1</i> <i>HXK1</i> <i>HXK2</i>	ATP + glucose ==> ADP + glucose-6-P
Glucose-6-phosphate isomerase [EC:5.3.1.9]	<i>YBR196C</i>	<i>PGI1</i>	glucose-6-P <==> fructose-6-P
6- phosphofructokinase [EC:2.7.1.11]	<i>YGR240C</i> <i>YMR205C</i>	<i>PFK1</i> <i>PFK2</i>	ATP + fructose-6-P ==> ADP + fructose-1,6-bis-P
Fructose- bisphosphate aldolase, class II [EC:4.1.2.13]	<i>YKL060C</i>	<i>FBA1</i>	fructose-1,6-bis-P <==> DHAP + GAP
Triosephosphate isomerase (TIM) [EC:5.3.1.1]	<i>YDR050C</i>	<i>TPI1</i>	DHAP <==> GAP
Glyceraldehyde 3- phosphate dehydrogenase [EC:1.2.1.12]	<i>YJL052W</i> <i>YJR009C</i> <i>YGR192C</i>	<i>TDH1</i> <i>TDH2</i> <i>TDH3</i>	GAP + NAD[c] <==> 1,3-P-glycerate + NADH[c]
Phosphoglycerate kinase [EC:2.7.2.3]	<i>YCR012W</i>	<i>PGK1</i>	1,3-P-glycerate + ADP <==> 3-P-glycerate + ATP
Phosphoglycerate mutase [EC:5.4.2.1]	<i>YKL152C</i> <i>YDL021W</i> <i>YOL056W</i>	<i>GPM1</i> <i>GPM2</i> <i>GPM3</i>	3-P-glycerate <==> 2-P-glycerate
Enolase [EC:4.2.1.11]	<i>YGR254W</i> <i>YHR174W</i>	<i>ENO1</i> <i>ENO2</i>	2-P-glycerate <==> PEP
Pyruvate kinase [EC:2.7.1.40]	<i>YAL038W</i> <i>YOR347C</i>	<i>PYK1</i> <i>PYK2</i>	ADP + PEP ==> ATP + pyruvate[c]
<b>Pentose phosphate pathway</b>			
Glucose-6-phosphate 1-dehydrogenase [EC:1.1.1.49]	<i>YNL241C</i>	<i>ZWF1</i>	NADP[c] + glucose-6-P ==> G15L + NADPH[c]
6- phosphogluconolacto nase [EC:3.1.1.31]	<i>YNR034W</i> <i>YCR073W-A</i> <i>YHR163W</i> <i>YGR248W</i>	<i>SOL1</i> - <i>SOL3</i> <i>SOL4</i>	G15L ==> gluconate-6-P
6-phosphogluconate dehydrogenase [EC:1.1.1.44]	<i>YGR256W</i> <i>YHR183W</i>	<i>GND2</i> <i>GND1</i>	NADP[c] + gluconate-6-P ==> CO <sub>2</sub> + NADPH[c] + ribulose-5-P
Ribose 5-phosphate isomerase A [EC:5.3.1.6]	<i>YOR095C</i>	<i>RKI1</i>	ribulose-5-P <==> ribose-5-P
Ribulose-phosphate 3-epimerase [EC:5.1.3.1]	<i>YJL121C</i>	<i>RPE1</i>	ribulose-5-P <==> xylulose-5-P
Transketolase [EC:2.2.1.1]	<i>YBR117C</i> <i>YPR074C</i>	<i>TKL2</i> <i>TKL1</i>	ribose-5-P + xylulose-5-P <==> GAP + sedoheptulose-7-P

Table 2.5 continued

Enzyme/Description	Gene (systematic name)	Gene (standard name)	Reaction
Transaldolase [EC:2.2.1.2]	<i>YGR043C</i>	<i>NOM1</i>	GAP + sedoheptulose-7-P $\rightleftharpoons$ erythrose-4-P + fructose-6-P
Transketolase [EC:2.2.1.1]	<i>YBR117C</i> <i>YPR074C</i>	<i>TKL2</i> <i>TKL1</i>	erythrose-4-P + xylulose-5-P $\rightleftharpoons$ GAP + fructose-6-P
<b>Gluconeogenesis, anaplerotic reactions, pyruvate dehydrogenase bypass and side product formation</b>			
Fructose-1,6-bisphosphatase [EC:3.1.3.11]	<i>YLR377C</i>	<i>FBP1</i>	fructose-1,6-bis-P $\Rightarrow$ fructose-6-P
Pyruvate carboxylase [EC:6.4.1.1]	<i>YGL062W</i> <i>YBR218C</i>	<i>PYC1</i> <i>PYC2</i>	ATP + CO <sub>2</sub> + pyruvate[c] $\Rightarrow$ ADP + oxaloacetate[c]
Phosphoenolpyruvate carboxykinase [EC:4.1.1.49]	<i>YKR097W</i>	<i>PCK1</i>	ATP + oxaloacetate[c] $\Rightarrow$ ADP + CO <sub>2</sub> + PEP
Pyruvate decarboxylase [EC:4.1.1.1]	<i>YLR044C</i> <i>YLR134W</i> <i>YGR087C</i>	<i>PDC1</i> <i>PDC5</i> <i>PDC6</i>	pyruvate[c] $\Rightarrow$ CO <sub>2</sub> + acetaldehyde[c]
Alcohol dehydrogenase [EC:1.1.1.1]	<i>YOL086C</i> <i>YMR303C</i> <i>YGL256W</i> <i>YBR145W</i>	<i>ADH1</i> <i>ADH2</i> <i>ADH4</i> <i>ADH5</i>	NADH[c] + acetaldehyde[c] $\rightleftharpoons$ NAD[c] + ethanol[c]
Mg(2+)-ACDH; Mg(2+)-activated acetaldehyde dehydrogenase [EC:1.2.1.3]	<i>YPL061W</i>	<i>ALD6</i>	NADP[c] + acetaldehyde[c] $\Rightarrow$ NADPH[c] + acetate[c]
Aldehyde dehydrogenase 1 and 2 [EC:1.2.1.5]	<i>YMR170C</i> <i>YMR169C</i>	<i>ALD2</i> <i>ALD3</i>	NAD[c] + acetaldehyde[c] $\Rightarrow$ NADH[c] + acetate[c]
Acetyl-CoA synthetase [EC:6.2.1.1]	<i>YLR153C</i>	<i>ACS2</i>	ATP + acetate[c] $\Rightarrow$ AMP + AcCoA[c]
Glycerol-3-phosphate dehydrogenase [EC 1.1.1.8]	<i>YDL022W</i> <i>YOL059W</i>	<i>GPD1</i> <i>GPD2</i>	DHAP + NADH[c] $\rightleftharpoons$ NAD[c] + glycerol-P
Glycerol-3- phosphatase 2 [EC:3.1.3.21]	<i>YER062C</i> <i>YIL053W</i>	<i>HOR2</i> <i>RHR2</i>	glycerol-P $\Rightarrow$ glycerol
<b>Pyruvate dehydrogenase complex, citric acid cycle and mitochondrial reactions</b>			
Pyruvate dehydrogenase complex:E2 (Lat1p) core [EC:2.3.1.12], E1 (Pda1p and Pdb1p) [EC:1.2.4.1], E3 (Lpd1p) and Protein X (Pdx1p)	<i>YBR221C</i> <i>YNL071W</i> <i>YGR193C</i> <i>YER178W</i> <i>YFL018C</i>	<i>PDB1</i> <i>LAT1</i> <i>PDX1</i> <i>PDA1</i> <i>LPD1</i>	NAD[m] + pyruvate[m] $\Rightarrow$ AcCoA[m] + CO <sub>2</sub> + NADH[m]
Citrate synthase [EC:2.3.3.1]	<i>YNR001C</i> <i>YPR001W</i>	<i>CIT1</i> <i>CIT3</i>	AcCoA[m] + oxaloacetate[m] $\Rightarrow$ citrate

**Table 2.5** continued

<b>Enzyme/Description</b>	<b>Gene (systematic name)</b>	<b>Gene (standard name)</b>	<b>Reaction</b>
Aconitate hydratase 1 [EC:4.2.1.3]	<i>YLR304C</i>	<i>ACO1</i>	citrate <=> isocitrate
Isocitrate dehydrogenase [NAD] [EC:1.1.1.41]	<i>YNL037C</i> <i>YOR136W</i>	<i>IDH1</i> <i>IDH2</i>	NAD[m] + isocitrate ==> CO2 + NADH[m] + a- ketoglutarate
Isocitrate dehydrogenase [NADP] [EC:1.1.1.42]	<i>YDL066W</i>	<i>IDP1</i>	NADP[m] + isocitrate ==> CO2 + NADPH[m] + a- ketoglutarate
Alpha-ketoglutarate dehydrogenase, Dihydrolipoyl dehydrogenase [EC:1.2.4.2], [EC:1.8.1.4]	<i>YIL125W</i> <i>YDR148C</i> <i>YFL018C</i>	<i>KGD1</i> <i>KGD2</i> <i>LPD1</i>	NAD[m] + a-ketoglutarate ==> CO2 + NADH[m] + succinyl-CoA
Succinyl-CoA ligase [EC:6.2.1.5]	<i>YOR142W</i> <i>YGR244C</i>	<i>LSC1</i> <i>LSC2</i>	ADP + succinyl-CoA <=> ATP + succinate
Succinate dehydrogenase complex [EC:1.3.5.1]	<i>YDR178W</i> <i>YKL141W</i> <i>YKL148C</i> <i>YLL041C</i> <i>YJL045W</i>	<i>SDH4</i> <i>SDH3</i> <i>SDH1</i> <i>SDH2</i> -	FAD + succinate <=> FADH2 + fumarate
Fumarate hydratase [EC:4.2.1.2]	<i>YPL262W</i>	<i>FUM1</i>	fumarate <=> malate
Malate dehydrogenase [EC:1.1.1.37]	<i>YKL085W</i>	<i>MDH1</i>	NAD[m] + malate <=> NADH[m] + oxaloacetate[m]
Malic enzyme [EC:1.1.1.38]	<i>YKL029C</i>	<i>MAE1</i>	NADP[m] + malate ==> CO2 + NADPH[m] + pyruvate[m]
Malic enzyme [EC:1.1.1.38]	<i>YKL029C</i>	<i>MAE1</i>	NAD[m] + malate ==> CO2 + NADH[m] + pyruvate[m]
Alcohol dehydrogenases <i>ADH1-5</i> [EC:1.1.1.1]	<i>YOL086C</i> <i>YMR303C</i> <i>YMR083W</i> <i>YGL256W</i> <i>YBR145W</i>	<i>ADH1</i> <i>ADH2</i> <i>ADH3</i> <i>ADH4</i> <i>ADH5</i>	NAD[m] + ethanol[m] <=> NADH[m] + acetaldehyde[m]
Potassium-activated aldehyde dehydrogenase [EC:1.2.1.3]	<i>YOR374W</i>	<i>ALD4</i>	NAD[m] + acetaldehyde[m] ==> NADH[m] + acetate[m]
Potassium-activated aldehyde dehydrogenase, Aldehyde dehydrogenase 5 [EC:1.2.1.3]	<i>YOR374W</i> <i>YER073W</i>	<i>ALD4</i> <i>ALD5</i>	NADP[m] + acetaldehyde[m] ==> NADPH[m] + acetate[m]
Acetyl-CoA synthetase 1-2 [EC:6.2.1.1]	<i>YAL054C</i>	<i>ACS1</i>	ATP + acetate[m] ==> AMP + AcCoA[m]

Table 2.5 continued

Enzyme/Description	Gene (systematic name)	Gene (standard name)	Reaction
<b>Mitochondrial shuttles</b>			
NAD/NADH shuttle (Förster, Gombert, and Nielsen 2002)	-		$\text{NADH}[c] + \text{NAD}[m] \Rightarrow \text{NADH}[m] + \text{NAD}[c]$
Transport: mitochondrial pyruvate carrier	<i>YGL080W</i>	<i>FMP37</i>	$\text{pyruvate}[c] \Leftrightarrow \text{pyruvate}[m]$
Transport: mitochondrial inner membrane transporter	<i>YKL120W</i>	<i>OAC1</i>	$\text{oxaloacetate}[c] \Leftrightarrow \text{oxaloacetate}[m]$
Transport	-		$\text{ethanol}[c] \Leftrightarrow \text{ethanol}[m]$
Transport	-		$\text{acetaldehyde}[c] \Leftrightarrow \text{acetaldehyde}[m]$
Transport	-		$\text{acetate}[c] \Leftrightarrow \text{acetate}[m]$
<b>MVA pathway</b>			
Acetyl-CoA C- acetyltransferase [EC:2.3.1.9]	<i>YPL028W</i>	<i>ERG10</i>	$2 \text{ AcCoA}[c] \Leftrightarrow \text{acetoacetyl-CoA}$
Hydroxymethylglutar yl-CoA synthase [EC:2.3.3.10]	<i>YML126C</i>	<i>ERG13</i>	$\text{AcCoA}[c] + \text{acetoacetyl-CoA} \Leftrightarrow \text{3-hydroxy-3-methylglutaryl-CoA}$
Hydroxymethylglutar yl-CoA reductase (NADPH) [EC:1.1.1.34]	<i>YML075C</i> <i>YLR450W</i>	<i>HMG1</i> <i>HMG2</i>	$\text{3-hydroxy-3-methylglutaryl-CoA} + 2 \text{ NADPH}[c] \Leftrightarrow 2 \text{ NADP}[c] + \text{mevalonate}$
Mevalonate kinase [EC:2.7.1.36]	<i>YMR208W</i>	<i>ERG12</i>	$\text{ATP} + \text{mevalonate} \Rightarrow \text{5-phosphomevalonate} + \text{ADP}$
Phosphomevalonate kinase [EC:2.7.4.2]	<i>YMR220W</i>	<i>ERG8</i>	$\text{5-phosphomevalonate} + \text{ATP} \Rightarrow \text{5-diphosphomevalonate} + \text{ADP}$
Diphosphomevalonat e decarboxylase [EC:4.1.1.33]	<i>YNR043W</i>	<i>ERG19</i>	$\text{5-diphosphomevalonate} + \text{ATP} \Leftrightarrow \text{ADP} + \text{CO}_2 + \text{isopentenyl-diphosphate}$
<b>Oxidative phosphorylation and ATP maintenance</b>			
	<i>YMR145C</i>	<i>NDE1</i>	
	<i>YDL085W</i>	<i>NDE2</i>	
	<i>YBL099W</i>	<i>ATP1</i>	
Respiratory chain: NADH	<i>YDL004W</i>	<i>ATP16</i>	
dehydrogenase	<i>YDR377W</i>	<i>ATP17</i>	
[EC:1.6.5.3] and F1F0	<i>YJR121W</i>	<i>ATP2</i>	
ATP synthase	<i>YKL016C</i>	<i>ATP7</i>	
[EC:3.6.3.14]	<i>YLR295C</i>	<i>ATP14</i>	
	<i>Q0085</i>	<i>ATP6</i>	
	<i>Q0130</i>	<i>ATP9</i>	
	<i>YPL271W</i>	<i>ATP15</i>	
			$24 \text{ ADP} + 20 \text{ NADH}[m] + 10 \text{ O}_2 \Rightarrow 24 \text{ ATP} + 20 \text{ NAD}[m]$

**Table 2.5** continued

Enzyme/Description	Gene (systematic name)	Gene (standard name)	Reaction
Respiratory chain: Succinate dehydrogenase complex [EC:1.3.5.1] and F1F0 ATP synthase [EC:3.6.3.14]	<i>YDR178W</i>	<i>SDH4</i>	
	<i>YKL141W</i>	<i>SDH3</i>	
	<i>YKL148C</i>	<i>SDH1</i>	
	<i>YLL041C</i>	<i>SDH2</i>	
	<i>YJL045W</i>	-	
	<i>YBL099W</i>	<i>ATP1</i>	
	<i>YDL004W</i>	<i>ATP16</i>	24 ADP + 20 FADH2 + 10 O2 ==> 24 ATP + 20
	<i>YDR377W</i>	<i>ATP17</i>	FAD
	<i>YJR121W</i>	<i>ATP2</i>	
	<i>YKL016C</i>	<i>ATP7</i>	
	<i>YLR295C</i>	<i>ATP14</i>	
	<i>Q0085</i>	<i>ATP6</i>	
<i>Q0130</i>	<i>ATP9</i>		
<i>YPL271W</i>	<i>ATP15</i>		
ATP maintenance	-	-	ATP ==> ADP
Adenylate kinase [EC:2.7.4.3]	<i>YER170W</i> <i>YDR226W</i>	<i>ADK2</i> <i>ADK1</i>	AMP + ATP ==> 2 ADP
<b>Biomass formation</b>			
Biomass formation (Förster, Gombert, and Nielsen 2002; Gombert et al. 2001)	-	-	6 3-P-glycerate + 254 ATP + 24 AcCoA[c] + 3 AcCoA[m] + 90 NADPH[c] + 22 NADPH[m] + 16 NAD[c] + 6 NAD[m] + 6 PEP + 11 a- ketoglutarate + 3 erythrose-4-P + 25 glucose-6- P + glycerol-P + 10 oxaloacetate[c] + 18 pyruvate[c] + 3 ribose-5-P ==> 254 ADP + BIOMASS + 16 NADH[c] + 6 NADH[m] + 90 NADP[c] + 22 NADP[m]

**Table 2.6** Artificial sources for the metabolic network of *S. cerevisiae*.

Enzyme/Description	Gene	Reaction
Artificial ATP source	-	ADP ==> ATP
Artificial NADPH source	-	NADP[c] ==> NADPH[c]
Artificial NADH source	-	NAD[c] ==> NADH[c]



**Table 2.7** Heterologous enzymes/pathways for the metabolic network of *S. cerevisiae*.

Enzyme/Description	Gene	Reaction
Heterologous soluble transhydrogenase from <i>E. coli</i>	<i>sthA/udhA</i>	$\text{NADH[c]} + \text{NADP[c]} \rightleftharpoons \text{NADPH[c]} + \text{NAD[c]}$
Cytosolic pyruvate dehydrogenase complex from <i>E. coli</i>	<i>lpdA</i> <i>aceEF</i>	$\text{NAD[c]} + \text{pyruvate[c]} \Rightarrow \text{AcCoA[c]} + \text{CO}_2 + \text{NADH[c]}$
NADP+-dependent glyceraldehyde-3-phosphate dehydrogenase [EC:1.2.1.13]; e.g. from <i>K. lactis</i> (Verho et al. 2003)	<i>GDP1</i>	$\text{GAP} + \text{NADP[c]} \rightleftharpoons \text{1,3-P-glycerate} + \text{NADPH[c]}$
Transfer of the native MVA pathway into mitochondria	<i>YPL028W / ERG10</i>	$2 \text{ AcCoA[m]} \rightleftharpoons \text{acetoacetyl-CoA}$
	<i>YML126C / ERG13</i>	$\text{AcCoA[m]} + \text{acetoacetyl-CoA} \rightleftharpoons \text{3-hydroxy-3-methylglutaryl-CoA}$
	<i>YML075C / HMG1</i>	$\text{3-hydroxy-3-methylglutaryl-CoA} + 2 \text{ NADPH[m]} \rightleftharpoons 2 \text{ NADP[m]} + \text{mevalonate}$
	<i>YLR450W / HMG2</i>	$\text{ATP} + \text{mevalonate} \Rightarrow \text{5-phosphomevalonate} + \text{ADP}$
	<i>YMR208W / ERG12</i>	$\text{5-phosphomevalonate} + \text{ATP} \Rightarrow \text{5-diphosphomevalonate} + \text{ADP}$
	<i>YMR220W / ERG8</i>	$\text{5-diphosphomevalonate} + \text{ATP} \rightleftharpoons \text{ADP} + \text{CO}_2 + \text{isopentenyl-diphosphate}$
Heterologous ATP-citrate-lyase [EC:2.3.3.8], e.g. from <i>Yarrowia lipolytica</i>	<i>ACL1</i>	$\text{Cit} + \text{ATP} \Rightarrow \text{oxaloacetate[c]} + \text{AcCoA[c]} + \text{ADP}$
Introduction of the DXP pathway from <i>E. coli</i>	<i>dxs, dxr/ispC, ispDEFGH</i>	$2 \text{ ATP} + \text{GAP} + 3 \text{ NADPH[c]} + \text{pyruvate[c]} \Rightarrow 2 \text{ ADP} + \text{CO}_2 + 3 \text{ NADP[c]} + \text{DMAPP}$
	<i>idi</i>	$\text{DMAPP} \rightleftharpoons \text{isopentenyl-diphosphate}$

**Table 2.8** Alternative carbon sources for the metabolic network of *S. cerevisiae*

Enzyme/Description	Gene (systematic name)	Gene (standard name)	Reaction
<b>Xylose</b>			
Transport by <i>HXT</i> genes or heterologous transporters (Saloheimo et al. 2007)	-	-	==> Xylose
xylulose kinase [EC:2.7.1.17]	<i>YGR194C</i>	<i>XKS1</i>	ATP + Xylulose ==> ADP + xylulose-5-P
XI pathway: xylose isomerase [EC:5.3.1.5], e.g. from <i>Piromyces sp.</i> (Karhumaa et al. 2007)	-	<i>xylA</i>	Xylose ==> Xylulose
XR-XDH pathway: xylose reductase (XR) [EC:1.1.1.307] and xylitol dehydrogenase (XDH) [EC:1.1.1.9], e.g. from <i>Pichia stipitis</i> (Karhumaa et al. 2007)	-	<i>XYL1</i>	NADH[c] + Xylose ==> NAD[c] + xylitol
	-	<i>XYL1</i>	NADPH[c] + Xylose ==> NADP[c] + xylitol
	-	<i>XYL2</i>	NAD[c] + xylitol ==> NADH[c] + Xylulose
<b>Galactose</b>			
Transport: see glucose	-	-	==> galactose
Galactokinase [EC:2.7.1.6], galactose-1-P-uridylyltransferase [EC:2.7.7.12], UDP-glucose-4-epimerase [EC:5.1.3.2], phosphoglucomutase-1 & 2 [EC:5.4.2.2]	<i>YBR020W</i> <i>YDR009W</i> <i>YBR018C</i> <i>YBR019C</i> <i>YKL127W</i> <i>YMR105C</i> <i>YMR278W</i>	<i>GAL1</i> <i>GAL3</i> <i>GAL7</i> <i>GAL10</i> <i>PGM1</i> <i>PGM2</i> <i>PRM15</i>	galactose + ATP ==> glucose-6-P + ADP
<b>Fructose</b>			
Transport: see glucose	-	-	==> fructose
Hexokinase [EC:2.7.1.1]	<i>YGL253W</i>	<i>HXK2</i>	fructose + ATP ==> ADP + fructose-6-P
<b>Ethanol</b>			

Table 2.8 continued

Enzyme/Description	Gene (systematic name)	Gene (standard name)	Reaction
Transport by diffusion (Guijarro and Lagunas 1984)	-	-	==> ethanol[c]
<b>Glycerol</b>			
Transport	-	-	==> glycerol
glycerol kinase [EC:2.7.1.30]	<i>YHL032C</i>	<i>GUT1</i>	glycerol + ATP ==> glycerol-P + ADP
<b>glyoxylate cycle plus mitochondrial shuttles plus compartmentalization of intermediates of citric acid cycle for growth on ethanol and glycerol</b>			
Citrate synthase [EC:2.3.3.1]	<i>YCR005C</i>	<i>CIT2</i>	AcCoA[c] + oxaloacetate[c] ==> citrate[c]
Aconitate hydratase 1 [EC:4.2.1.3]	<i>YLR304C</i>	<i>ACO1</i>	citrate[c] <==> isocitrate[c]
Isocitrate lyase [EC:4.1.3.1]	<i>YER065C</i> <i>YPR006C</i>	<i>ICL1</i> <i>ICL2</i>	isocitrate[c] ==> succinate[c] + glyoxylate
Malate synthase 1 [EC:2.3.3.9]	<i>YNL117W</i> <i>YIRO31C</i>	<i>MLS1</i> <i>MLS2</i>	AcCoA[c] + glyoxylate ==> malate[c]
Malate dehydrogenase [EC:1.1.1.37]	<i>YOL126C</i> <i>YDL078C</i>	<i>MDH2</i> <i>MDH3</i>	NAD[c] + malate[c] <==> NADH[c] + oxaloacetate[c]
Fumarate hydratase [EC:4.2.1.2]	<i>YPL262W</i>	<i>FUM1</i>	malate[c] <==> fumarate[c]
Fumarate reductase [EC:1.3.1.6]	<i>YJR051W</i> <i>YEL047C</i>	<i>OSM1</i> <i>FRD1</i>	NADH[c] + fumarate[c] ==> NAD[c] + succinate[c]
Isocitrate dehydrogenase [EC:1.1.1.42]	<i>YNL009W</i> <i>YLR174W</i>	<i>IDP3</i> <i>IDP2</i>	NADP[c] + isocitrate[c] ==> CO <sub>2</sub> + NADPH[c] + a-ketoglutarate[c]
Transport: mitochondrial dicarboxylate transporter	<i>YLR348C</i>	<i>DIC1</i>	succinate[c] ==> succinate[m]
Transport: mitochondrial citrate- oxoglutarate carrier	<i>YMR241W</i>	<i>YHM2</i>	a-ketoglutarate[c] + citrate[m] ==> citrate[c] + a-ketoglutarate[m]
Transport: mitochondrial succinate-fumarate transporter	<i>YJR095W</i>	<i>SFC1</i>	succinate[c] + fumarate[m] ==> fumarate[c] + succinate[m]
Transport: mitochondrial inner membrane citrate transporter	<i>YBR291C</i>	<i>CTP1</i>	citrate[c] + malate[m] <==> malate[c] + citrate[m]

**Table 2.8** continued

Enzyme/Description	Gene (systematic name)	Gene (standard name)	Reaction
Transport: mitochondrial inner membrane citrate transporter	<i>YBR291C</i>	<i>CTP1</i>	$\text{citrate}[\text{c}] + \text{isocitrate}[\text{m}] \rightleftharpoons \text{isocitrate}[\text{c}] + \text{citrate}[\text{m}]$
Transport: mitochondrial dicarboxylate transporter	<i>YLR348C</i>	<i>DIC1</i>	$\text{malate}[\text{c}] \rightleftharpoons \text{malate}[\text{m}]$
Transport: carnitine O-acetyltransferase [EC:2.3.1.7]	<i>YOR100C</i> <i>YAR035W</i> <i>YER024W</i> <i>YML042W</i>	<i>CRC1</i> <i>YAT1</i> <i>YAT2</i> <i>CAT2</i>	$\text{AcCoA}[\text{c}] \rightleftharpoons \text{AcCoA}[\text{m}]$
Citrate synthase [EC:2.3.3.1]	<i>YNR001C</i> <i>YPR001W</i>	<i>CIT1</i> <i>CIT3</i>	$\text{AcCoA}[\text{m}] + \text{oxaloacetate}[\text{m}] \rightleftharpoons \text{citrate}[\text{m}]$
Aconitate hydratase 1 [EC:4.2.1.3]	<i>YLR304C</i>	<i>ACO1</i>	$\text{citrate}[\text{m}] \rightleftharpoons \text{isocitrate}[\text{m}]$
Isocitrate dehydrogenase [NAD] [EC:1.1.1.41]	<i>YNL037C</i> <i>YOR136W</i>	<i>IDH1</i> <i>IDH2</i>	$\text{NAD}[\text{m}] + \text{isocitrate}[\text{m}] \rightleftharpoons \text{CO}_2 + \text{NADH}[\text{m}] + \text{a-ketoglutarate}[\text{m}]$
Isocitrate dehydrogenase [NADP] [EC:1.1.1.42]	<i>YDL066W</i>	<i>IDP1</i>	$\text{NADP}[\text{m}] + \text{isocitrate}[\text{m}] \rightleftharpoons \text{CO}_2 + \text{NADPH}[\text{m}] + \text{a-ketoglutarate}[\text{m}]$
Alpha-ketoglutarate dehydrogenase, Dihydrolipoyl dehydrogenase [EC:1.2.4.2], [EC:1.8.1.4]	<i>YIL125W</i> <i>YDR148C</i> <i>YFL018C</i>	<i>KGD1</i> <i>KGD2</i> <i>LPD1</i>	$\text{NAD}[\text{m}] + \text{a-ketoglutarate}[\text{m}] \rightleftharpoons \text{CO}_2 + \text{NADH}[\text{m}] + \text{succinyl-CoA}$
Succinyl-CoA ligase [EC:6.2.1.5]	<i>YOR142W</i> <i>YGR244C</i>	<i>LSC1</i> <i>LSC2</i>	$\text{ADP} + \text{succinyl-CoA} \rightleftharpoons \text{ATP} + \text{succinate}[\text{m}]$
Succinate dehydrogenase [EC:1.2.4.2], [EC:1.8.1.4], [EC:2.3.1.61]	<i>YDR178W</i> <i>YKL141W</i> <i>YJL045W</i> <i>YKL148C</i> <i>YLL041C</i>	<i>SDH4</i> <i>SDH3</i> <i>SDH1</i> <i>SDH2</i> -	$\text{FAD} + \text{succinate}[\text{m}] \rightleftharpoons \text{FADH}_2 + \text{fumarate}[\text{m}]$
Fumarate hydratase [EC:4.2.1.2]	<i>YPL262W</i>	<i>FUM1</i>	$\text{fumarate}[\text{m}] \rightleftharpoons \text{malate}[\text{m}]$
Malate dehydrogenase [EC:1.1.1.37]	<i>YKL085W</i>	<i>MDH1</i>	$\text{NAD}[\text{m}] + \text{malate}[\text{m}] \rightleftharpoons \text{NADH}[\text{m}] + \text{oxaloacetate}[\text{m}]$
Malic enzyme [EC:1.1.1.38]	<i>YKL029C</i>	<i>MAE1</i>	$\text{NADP}[\text{m}] + \text{malate}[\text{m}] \rightleftharpoons \text{CO}_2 + \text{NADPH}[\text{m}] + \text{pyruvate}[\text{m}]$

**Table 2.8** continued

Enzyme/Description	Gene (systematic name)	Gene (standard name)	Reaction
Malic enzyme [EC:1.1.1.38]	<i>YKL029C</i>	<i>MAE1</i>	NAD[m] + malate[m] ==> CO <sub>2</sub> + NADH[m] + pyruvate[m]
Biomass formation (Förster, Gombert, and Nielsen 2002; Gombert et al. 2001)	-	-	6 3-P-glycerate + 254 ATP + 24 AcCoA[c] + 3 AcCoA[m] + 90 NADPH[c] + 22 NADPH[m] + 16 NAD[c] + 6 NAD[m] + 6 PEP + 11 a-ketoglutarate [m] + 3 erythrose-4-P + 25 glucose-6-P + glycerol-P + 10 oxaloacetate[c] + 18 pyruvate[c] + 3 ribose-5-P ==> 254 ADP + BIOMASS + 16 NADH[c] + 6 NADH[m] + 90 NADP[c] + 22 NADP[m]

## 2.5.2 Metabolic networks of *E. coli*

**Table 2.9** Wild type metabolic network of *E. coli* on glucose as single carbon source.

Enzyme/Description	Gene	Reaction
<b>Transport</b>		
Transport	-	==> glucose
Glucose uptake by phosphoenolpyruvate:glucose transferase system	<i>ptsGHI, crr</i>	PEP + glucose ==> glucose-6-P + pyruvate
Transport	-	==> O <sub>2</sub>
Transport	-	==> NH <sub>3</sub>
Transport	-	ethanol ==>
Transport	-	lactate ==>
Transport	-	succinate ==>
Transport	-	acetate ==>
Transport	-	formate ==>
Transport	-	CO <sub>2</sub> <==>
Transport	-	H <sub>2</sub> ==>
Growth in mmol*gCDW <sup>-1</sup> *h <sup>-1</sup>	-	0.001 BIOMASS ==>
Artificial IPP production	-	isopentenyl-diphosphate ==>
<b>Glycolysis</b>		
Glucose-6-phosphate isomerase [EC:5.3.1.9]	<i>pgi</i>	glucose-6-P <==> fructose-6-P
6-phosphofruktokinase [EC:2.7.1.11]	<i>pfkAB</i>	ATP + fructose-6-P ==> ADP + fructose-1,6-bis-P
Fructose-bisphosphate aldolase, class I [EC:4.1.2.13]	<i>fbaAB</i>	fructose-1,6-bis-P <==> DHAP + GAP

**Table 2.9** continued

Enzyme/Description	Gene	Reaction
Triosephosphate isomerase [EC:5.3.1.1]	<i>tpiA</i>	DHAP <=> GAP
Glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12]	<i>gapA</i>	GAP + NAD <=> 1,3-P-glycerate + NADH
Phosphoglycerate kinase [EC:2.7.2.3]	<i>pgk</i>	1,3-P-glycerate + ADP <=> 3-P-glycerate + ATP
Phosphoglycerate mutase [EC:5.4.2.1]	<i>gpmM, ytcJ, gpmA</i>	3-P-glycerate <=> 2-P-glycerate
Enolase [EC:4.2.1.11]	<i>eno</i>	2-P-glycerate <=> PEP
Pyruvate kinase [EC:2.7.1.40]	<i>pykAF</i>	ADP + PEP ==> ATP + pyruvate
<b>Pentose phosphate pathway</b>		
Glucose-6-phosphate 1- dehydrogenase [EC:1.1.1.49]	<i>zwf</i>	NADP + glucose-6-P ==> G15L + NADPH
6-phosphogluconolactonase [EC:3.1.1.31]	<i>pgl</i>	G15L ==> gluconate-6-P
6-phosphogluconate dehydrogenase [EC:1.1.1.44]	<i>gnd</i>	NADP + gluconate-6-P ==> CO <sub>2</sub> + NADPH + ribulose-5-P
Ribose 5-phosphate isomerase A [EC:5.3.1.6]	<i>rpiAB</i>	ribulose-5-P <=> ribose-5-P
Ribulose-phosphate 3- epimerase [EC:5.1.3.1]	<i>rpe</i>	ribulose-5-P <=> xylulose-5-P
Transketolase [EC:2.2.1.1]	<i>tktAB</i>	ribose-5-P + xylulose-5-P <=> GAP + sedoheptulose-7-P
Transaldolase [EC:2.2.1.2]	<i>talAB</i>	GAP + sedoheptulose-7-P <=> erythrose-4-P + fructose-6-P
Transketolase [EC:2.2.1.1]	<i>tktAB</i>	erythrose-4-P + xylulose-5-P <=> GAP + fructose-6-P
<b>Entner-Doudoroff/KDPG pathway</b>		
Phosphogluconate dehydratase [EC:4.2.1.12]	<i>edd</i>	gluconate-6-P ==> 2-keto-3-deoxy-6-phospho- gluconate
2-keto-3-deoxy-6- phosphogluconate aldolase [EC:4.1.3.16]	<i>eda</i>	2-keto-3-deoxy-6-phospho-gluconate ==> GAP + pyruvate
<b>Gluconeogenesis</b>		
Fructose-1,6-biphosphatase [EC:3.1.3.11]	<i>glpX, fbp</i>	fructose-1,6-bis-P ==> fructose-6-P
Phosphoenolpyruvate synthase [EC:2.7.9.2]	<i>pps</i>	ATP + pyruvate ==> AMP + PEP
<b>Pyruvate dehydrogenase complex and citric acid cycle</b>		
Pyruvate dehydrogenase complex [EC:1.2.4.1] [EC:2.3.1.12] [EC:1.8.1.4]	<i>lpdA, aceEF</i>	NAD + pyruvate ==> AcCoA + CO <sub>2</sub> + NADH
Citrate synthase [EC:2.3.3.1]	<i>prpC, gltA</i>	AcCoA + oxaloacetate ==> citrate

**Table 2.9** continued

Enzyme/Description	Gene	Reaction
Aconitate hydratase 1 [EC:4.2.1.3]	<i>acnAB</i>	citrate $\rightleftharpoons$ isocitrate
Isocitrate dehydrogenase [NADP] [EC:1.1.1.42]	<i>icd</i>	NADP + isocitrate $\Rightarrow$ CO <sub>2</sub> + NADPH + a-ketoglutarate
Alpha-ketoglutarate dehydrogenase complex [EC:1.2.4.2] [EC:1.8.1.4] [EC:2.3.1.61]	<i>lpdA, sucAB</i>	NAD + a-ketoglutarate $\Rightarrow$ CO <sub>2</sub> + NADH + succinyl-CoA
Succinyl-CoA synthetase [EC:6.2.1.5]	<i>sucCD</i>	ADP + succinyl-CoA $\rightleftharpoons$ ATP + succinate
Succinate dehydrogenase [EC:1.3.5.1] and fumarate reductase [EC:1.3.5.4]	<i>sdhABCD</i> <i>frdABCD</i>	succinate + quinone $\rightleftharpoons$ fumarate + quinol
Fumarate hydratase [EC:4.2.1.2]	<i>fumABC</i>	fumarate $\rightleftharpoons$ malate
Malate dehydrogenase [EC:1.1.1.37]	<i>mdh</i>	NAD + malate $\rightleftharpoons$ NADH + oxaloacetate
Malate:quinone oxidoreductase [EC:1.1.99.16]	<i>mqr</i>	malate + quinone $\Rightarrow$ oxaloacetate + quinol
<b>Anaplerotic reactions</b>		
Malic enzyme [EC 1.1.1.38]	<i>maeA</i>	NAD + malate $\Rightarrow$ CO <sub>2</sub> + NADH + pyruvate
Malic enzyme [EC 1.1.1.40]	<i>maeB</i>	NADP + malate $\Rightarrow$ CO <sub>2</sub> + NADPH + pyruvate
Phosphoenolpyruvate carboxylase [EC:4.1.1.31]	<i>ppc</i>	CO <sub>2</sub> + PEP $\Rightarrow$ oxaloacetate
Phosphoenolpyruvate carboxykinase [EC:4.1.1.49]	<i>pck</i>	ATP + oxaloacetate $\Rightarrow$ ADP + CO <sub>2</sub> + PEP
<b>Mixed acid fermentation</b>		
D-lactate dehydrogenase [EC:1.1.1.28]	<i>ldhA</i>	NADH + pyruvate $\Rightarrow$ NAD + lactate
Acetaldehyde dehydrogenase/alcohol dehydrogenase [EC:1.2.1.10] [EC:1.1.1.1]	<i>adhE</i>	AcCoA + NADH $\Rightarrow$ NAD + acetaldehyde
Alcohol dehydrogenase [EC:1.1.1.1]	<i>adhE</i> <i>adhP</i>	NADH + acetaldehyde $\Rightarrow$ NAD + ethanol
Phosphate acetyltransferase [EC:2.3.1.8]	<i>pta</i>	AcCoA $\Rightarrow$ acetylphosphate
Acetate kinase [EC:2.7.2.1]	<i>ackAB</i>	ADP + acetylphosphate $\Rightarrow$ ATP + acetate
pyruvate dehydrogenase (pyruvate oxidase) [EC:1.2.2.2]	<i>poxB</i>	pyruvate + quinone $\Rightarrow$ CO <sub>2</sub> + acetate + quinol
pyruvate formate-lyase [EC:2.3.1.54]	<i>pflB, tdcE</i>	pyruvate $\Rightarrow$ AcCoA + formate

Table 2.9 continued

Enzyme/Description	Gene	Reaction
Formate hydrogenylase complex [EC:1.12.99.6] [EC:1.1.99.33]	<i>hycBCDEFG, fdhF</i>	formate ==> CO2 + H2
<b>DXP pathway</b>		
1-deoxyxylulose-5-phosphate synthase [EC:2.2.1.7]		
1-deoxy-D-xylulose 5-phosphate reductoisomerase (uses NADPH (Takahashi et al. 1998)) [EC:1.1.1.267]		
4-diphosphocytidyl-2C-methyl-D-erythritol synthase (uses CTP) [EC:2.7.7.60]		
4-diphosphocytidyl-2-C-methylerythritol kinase (uses ATP) [EC:2.7.1.148]		
2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase [EC:4.6.1.12]	<i>dxs, dxr/ispC, ispDEFGH</i>	2 ATP + GAP + 3 NADPH + pyruvate ==> 2 ADP + CO2 + 3 NADP + dimethylallyl-diphosphate
1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (uses NADPH (Seemann et al. 2002)) [EC:1.17.7.1]		
1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase (uses NADPH (Rohdich et al. 2003; Wolff et al. 2003)) [EC:1.17.1.2]		
isopentenyl-diphosphate delta-isomerase [EC:5.3.3.2]	<i>idi</i>	dimethylallyl-diphosphate <==> isopentenyl-diphosphate
<b>Oxidative phosphorylation and ATP maintenance</b>		
Respiratory chain: NADH dehydrogenase I [EC:1.6.5.3] and ATP synthase [EC:3.6.3.14]	<i>nuoAHJKLMNEFG BCI, atpABCDEFGHI</i>	2 ADP + NADH + O2 ==> 2 ATP + NAD
Respiratory chain: NADH dehydrogenase II [EC:1.6.5.8]	<i>ndh</i>	NADH + quinone <==> NAD + quinol
Respiratory chain: Cytochrome <i>bo</i> terminal oxidase [EC:1.10.3.10] and ATP synthase [EC:3.6.3.14]	<i>cyoABCD, atpABCDEFGHI</i>	ADP + O2 + quinol ==> ATP + quinone



**Table 2.9** continued

Enzyme/Description	Gene	Reaction
Pyridine nucleotide transhydrogenase, membrane-bound [EC:1.6.1.2] ( <i>sthA/udhA</i> [EC:1.6.1.1] is ignored)	<i>pntAB</i>	$ADP + 3 NADPH + 3 NAD \rightleftharpoons ATP + 3 NADH + 3 NADP$
Adenylate kinase [EC:2.7.4.3]	<i>adk</i>	$AMP + ATP \rightleftharpoons 2 ADP$
ATP maintenance	-	$ATP \rightleftharpoons ADP$
<b>Biomass formation</b>		
Biomass formation (Carlson and Srienc 2004; Trinh, Unrean, and Srienc 2008)	-	$1642 \text{ 3-P-glycerate} + 40680 \text{ ATP} + 1207 \text{ AcCoA} + 31 \text{ GAP} + 18320 \text{ NADPH} + 4079 \text{ NAD} + 960 \text{ PEP} + 1427 \text{ a-ketoglutarate} + 512 \text{ erythrose-4-P} + 17 \text{ fructose-6-P} + 49 \text{ glucose-6-P} + 2355 \text{ oxaloacetate} + 3920 \text{ pyruvate} + 860 \text{ ribose-5-P} + 12502 \text{ NH}_3 \rightleftharpoons 40680 \text{ ADP} + \text{BIOMASS} + 4079 \text{ NADH} + 18320 \text{ NADP}$

**Table 2.10** Artificial sources for the metabolic network of *E. coli*.

Enzyme/Description	Gene	Reaction
Artificial ATP source	-	$ADP \rightleftharpoons ATP$
Artificial NADPH source	-	$NADP[c] \rightleftharpoons NADPH[c]$
Artificial NADH source	-	$NAD[c] \rightleftharpoons NADH[c]$

**Table 2.11** Heterologous enzymes/pathways for the metabolic network of *E. coli*.

Enzyme/Description	Gene	Reaction
NADP <sup>+</sup> -dependent glyceraldehyde-3-phosphate dehydrogenase [EC:1.2.1.13]; e.g. from <i>K. lactis</i> (Verho <i>et al.</i> 2003)	<i>GDP1</i>	GAP + NADP $\rightleftharpoons$ 1,3-P-glycerate + NADPH
	<i>YPL028W / ERG10</i>	2 AcCoA $\rightleftharpoons$ acetoacetyl-CoA
	<i>YML126C / ERG13</i>	AcCoA + acetoacetyl-CoA $\rightleftharpoons$ 3-hydroxy-3-methylglutaryl-CoA
	<i>YML075C / HMG1</i>	3-hydroxy-3-methylglutaryl-CoA + 2 NADPH
	<i>YLR450W / HMG2</i>	$\rightleftharpoons$ 2 NADP + mevalonate
MVA pathway from yeast	<i>YMR208W / ERG12</i>	ATP + mevalonate $\Rightarrow$ 5-phosphomevalonate + ADP
	<i>YMR220W / ERG8</i>	5-phosphomevalonate + ATP $\Rightarrow$ 5-diphosphomevalonate + ADP
	<i>YNR043W / ERG19</i>	5-diphosphomevalonate + ATP $\rightleftharpoons$ ADP + CO <sub>2</sub> + isopentenyl-diphosphate

**Table 2.12** Alternative carbon sources for the metabolic network of *E. coli*.

Enzyme/Description	Gene	Reaction
Glycerol plus glyoxylate shunt		
Transport	-	$\Rightarrow$ glycerol
Glycerol uptake by glycerol kinase [EC:2.7.1.30]	<i>glpK</i>	ATP + glycerol $\Rightarrow$ ADP + glycerol-P
Glycerol-3-phosphate dehydrogenase [EC:1.1.1.94]	<i>gpsA</i>	NADP + glycerol-P $\rightleftharpoons$ DHAP + NADPH
Isocitrate lyase [EC:4.1.3.1]	<i>aceA</i>	isocitrate $\Rightarrow$ succinate + glyoxylate
Malate synthase [EC:2.3.3.9]	<i>aceB, glcB</i>	AcCoA + glyoxylate $\Rightarrow$ malate

## 2.6 References

- Aguilera, A. 1986. Deletion of the phosphoglucose isomerase structural gene makes growth and sporulation glucose dependent in *Saccharomyces cerevisiae*. *Mol Gen Genet* 204 (2):310-6.
- Ajikumar, P. K., W. H. Xiao, K. E. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer, and G. Stephanopoulos. 2010. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* 330 (6000):70-4.
- Alper, H., K. Miyaoku, and G. Stephanopoulos. 2005. Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nat Biotechnol* 23 (5):612-6.
- Anderlund, M., T. L. Nissen, J. Nielsen, J. Villadsen, J. Rydstrom, B. Hahn-Hagerdal, and M. C. Kielland-Brandt. 1999. Expression of the *Escherichia coli pntA* and *pntB* genes, encoding nicotinamide nucleotide transhydrogenase, in *Saccharomyces cerevisiae* and its effect on product formation during anaerobic glucose fermentation. *Appl Environ Microbiol* 65 (6):2333-40.
- Asadollahi, M. A., J. Maury, K. Moller, K. F. Nielsen, M. Schalk, A. Clark, and J. Nielsen. 2008. Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: effect of *ERG9* repression on sesquiterpene biosynthesis. *Biotechnol Bioeng* 99 (3):666-77.
- Asadollahi, M. A., J. Maury, K. R. Patil, M. Schalk, A. Clark, and J. Nielsen. 2009. Enhancing sesquiterpene production in *Saccharomyces cerevisiae* through *in silico* driven metabolic engineering. *Metab Eng* 11 (6):328-34.
- Blank, L. M., L. Kuepfer, and U. Sauer. 2005. Large-scale <sup>13</sup>C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol* 6 (6):R49.
- Boles, E., P. de Jong-Gubbels, and J. T. Pronk. 1998. Identification and characterization of *MAE1*, the *Saccharomyces cerevisiae* structural gene encoding mitochondrial malic enzyme. *J Bacteriol* 180 (11):2875-82.
- Carlsen, S., P. K. Ajikumar, L. R. Formenti, K. Zhou, T. H. Phon, M. L. Nielsen, A. E. Lantz, M. C. Kielland-Brandt, and G. Stephanopoulos. 2013. Heterologous expression and characterization of bacterial 2-C-methyl-D-erythritol-4-phosphate pathway in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 97:5753-5769.
- Carlson, R., D. Fell, and F. Srienc. 2002. Metabolic pathway analysis of a recombinant yeast for rational strain development. *Biotechnol Bioeng* 79 (2):121-34.
- Carlson, R., and F. Srienc. 2004. Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: identification of reactions. *Biotechnol Bioeng* 85 (1):1-19.
- Carvalho, Florbela, Luís C. Duarte, and Francisco M Gírio. 2008. Hemicellulose biorefineries: a review on biomass pretreatments. *J Sci Ind Res* 67:849-864.
- Chang, M. C., R. A. Eachus, W. Trieu, D. K. Ro, and J. D. Keasling. 2007. Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat Chem Biol* 3 (5):274-7.
- Cherry, J. M., E. L. Hong, C. Amundsen, R. Balakrishnan, G. Binkley, E. T. Chan, K. R. Christie, M. C. Costanzo, S. S. Dwight, S. R. Engel, D. G. Fisk, J. E. Hirschman, B. C. Hitz, K. Karra, C. J. Krieger, S. R. Miyasato, R. S. Nash, J. Park, M. S. Skrzypek, M. Simison, S. Weng, and E. D. Wong. 2012. *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res* 40 (Database issue):D700-5.
- Choi, H. S., S. Y. Lee, T. Y. Kim, and H. M. Woo. 2010. *In silico* identification of gene amplification targets for improvement of lycopene production. *Appl Environ Microbiol* 76 (10):3097-105.

- da Silva, G. P., M. Mack, and J. Contiero. 2009. Glycerol: a promising and abundant carbon source for industrial microbiology. *Biotechnol Adv* 27 (1):30-9.
- Davies, H. M. 2009. Organic chemistry: Synthetic lessons from nature. *Nature* 459 (7248):786-7.
- Dobson, R., V. Gray, and K. Rumbold. 2012. Microbial utilization of crude glycerol for the production of value-added products. *J Ind Microbiol Biotechnol* 39 (2):217-26.
- Donald, K. A., R. Y. Hampton, and I. B. Fritz. 1997. Effects of overproduction of the catalytic domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase on squalene synthesis in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 63 (9):3341-4.
- Farhi, M., E. Marhevka, T. Masci, E. Marcos, Y. Eyal, M. Ovadis, H. Abeliovich, and A. Vainstein. 2011. Harnessing yeast subcellular compartments for the production of plant terpenoids. *Metab Eng* 13 (5):474-81.
- Farmer, W. R., and J. C. Liao. 2001. Precursor balancing for metabolic engineering of lycopene production in *Escherichia coli*. *Biotechnol Prog* 17 (1):57-61.
- Fong, S. S., A. P. Burgard, C. D. Herring, E. M. Knight, F. R. Blattner, C. D. Maranas, and B. O. Palsson. 2005. *In silico* design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol Bioeng* 91 (5):643-8.
- Förster, J., A. K. Gombert, and J. Nielsen. 2002. A functional genomics approach using metabolomics and in silico pathway analysis. *Biotechnol Bioeng* 79 (7):703-12.
- Gardner, R. G., and R. Y. Hampton. 1999. A highly conserved signal controls degradation of 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase in eukaryotes. *J Biol Chem* 274 (44):31671-8.
- Gombert, A. K., M. Moreira dos Santos, B. Christensen, and J. Nielsen. 2001. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J Bacteriol* 183 (4):1441-51.
- Guijarro, J. M., and R. Lagunas. 1984. *Saccharomyces cerevisiae* does not accumulate ethanol against a concentration gradient. *J Bacteriol* 160 (3):874-8.
- Gunawardena, U., P. Meinhold, M. W. Peters, R. M. Urano, and R. M. R. Feldmann. 2010. Butanol production by metabolically engineered yeast. US 2010/0062505 A1.
- Hädicke, O., and S. Klamt. 2011. Computing complex metabolic intervention strategies using constrained minimal cut sets. *Metab Eng* 13 (2):204-13.
- Hansen, J. 2013. Method of producing isoprenoid compounds in yeast. US 2013/0137138 A1.
- Herrgard, M. J., N. Swainston, P. Dobson, W. B. Dunn, K. Y. Arga, M. Arvas, N. Bluthgen, S. Borger, R. Costenoble, M. Heinemann, M. Hucka, N. Le Novere, P. Li, W. Liebermeister, M. L. Mo, A. P. Oliveira, D. Petranovic, S. Pettifer, E. Simeonidis, K. Smallbone, I. Spasic, D. Weichart, R. Brent, D. S. Broomhead, H. V. Westerhoff, B. Kirdar, M. Penttila, E. Klipp, B. O. Palsson, U. Sauer, S. G. Oliver, P. Mendes, J. Nielsen, and D. B. Kell. 2008. A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat Biotechnol* 26 (10):1155-60.
- Hong, K. K., and J. Nielsen. 2012. Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cell Mol Life Sci* 69 (16):2671-90.
- Huang, Q., C. A. Roessner, R. Croteau, and A. I. Scott. 2001. Engineering *Escherichia coli* for the synthesis of taxadiene, a key intermediate in the biosynthesis of taxol. *Bioorg Med Chem* 9 (9):2237-42.

- Ibarra, R. U., J. S. Edwards, and B. O. Palsson. 2002. *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth. *Nature* 420 (6912):186-9.
- Jeun, Y.-S., M.-D. Kim, Y.-C. Park, T.-H. Lee, M.-S. Yoo, Y.-W. Ryu, and J.-H. Seo. 2003. Expression of *Azotobacter vinelandii* soluble transhydrogenase perturbs xylose reductase-mediated conversion of xylose to xylitol by recombinant *Saccharomyces cerevisiae*. *J Mol Catal. B: Enzym* 26 (3-6):251-256.
- Jiang, M., G. Stephanopoulos, and B. A. Pfeifer. 2012. Toward biosynthetic design and implementation of *Escherichia coli*-derived paclitaxel and other heterologous polyisoprene compounds. *Appl Environ Microbiol* 78 (8):2497-504.
- Kampranis, S. C., and A. M. Makris. 2012. Developing a yeast cell factory for the production of terpenoids. *Computational and Structural Biotechnology Journal* 3 (4).
- Kanehisa, M., S. Goto, Y. Sato, M. Furumichi, and M. Tanabe. 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40 (Database issue):D109-14.
- Karhumaa, K., R. Garcia Sanchez, B. Hahn-Hagerdal, and M. F. Gorwa-Grauslund. 2007. Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant *Saccharomyces cerevisiae*. *Microb Cell Fact* 6:5.
- Keseler, I. M., J. Collado-Vides, A. Santos-Zavaleta, M. Peralta-Gil, S. Gama-Castro, L. Muniz-Rascado, C. Bonavides-Martinez, S. Paley, M. Krummenacker, T. Altman, P. Kaipa, A. Spaulding, J. Pacheco, M. Latendresse, C. Fulcher, M. Sarker, A. G. Shearer, A. Mackie, I. Paulsen, R. P. Gunsalus, and P. D. Karp. 2011. EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Res* 39 (Database issue):D583-90.
- Kim, S. W., and J. D. Keasling. 2001. Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production. *Biotechnol Bioeng* 72 (4):408-15.
- Kirby, J., and J. D. Keasling. 2008. Metabolic engineering of microorganisms for isoprenoid production. *Nat Prod Rep* 25 (4):656-61.
- . 2009. Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu Rev Plant Biol* 60:335-55.
- Kizer, L., D. J. Pitera, B. F. Pfleger, and J. D. Keasling. 2008. Application of functional genomics to pathway optimization for increased isoprenoid production. *Appl Environ Microbiol* 74 (10):3229-41.
- Klamt, S. 2006. Generalized concept of minimal cut sets in biochemical networks. *Biosystems* 83 (2-3):233-47.
- Klamt, S., and E. D. Gilles. 2004. Minimal cut sets in biochemical reaction networks. *Bioinformatics* 20 (2):226-34.
- Klamt, S., J. Saez-Rodriguez, and E. D. Gilles. 2007. Structural and functional analysis of cellular networks with *CellNetAnalyzer*. *BMC Syst Biol* 1:2.
- Krömer, J. O., C. Wittmann, H. Schröder, and E. Heinzle. 2006. Metabolic pathway analysis for rational design of L-methionine production by *Escherichia coli* and *Corynebacterium glutamicum*. *Metab Eng* 8 (4):353-69.
- Lindahl, A. L., M. E. Olsson, P. Mercke, O. Tollbom, J. Schelin, M. Brodelius, and P. E. Brodelius. 2006. Production of the artemisinin precursor amorpha-4,11-diene by engineered *Saccharomyces cerevisiae*. *Biotechnol Lett* 28 (8):571-80.
- Luttik, M. A., K. M. Overkamp, P. Kotter, S. de Vries, J. P. van Dijken, and J. T. Pronk. 1998. The *Saccharomyces cerevisiae* *NDE1* and *NDE2* genes encode separate

- mitochondrial NADH dehydrogenases catalyzing the oxidation of cytosolic NADH. *J Biol Chem* 273 (38):24529-34.
- Martin, V. J., D. J. Pitera, S. T. Withers, J. D. Newman, and J. D. Keasling. 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol* 21 (7):796-802.
- Maury, J., M. A. Asadollahi, K. Moller, A. Clark, and J. Nielsen. 2005. Microbial isoprenoid production: an example of green chemistry through metabolic engineering. *Adv Biochem Eng Biotechnol* 100:19-51.
- Melzer, G., M. E. Esfandabadi, E. Franco-Lara, and C. Wittmann. 2009. Flux Design: *In silico* design of cell factories based on correlation of pathway fluxes to desired properties. *BMC Syst Biol* 3:120.
- Meng, Hailin, Yong Wang, Qiang Hua, Siliang Zhang, and Xiaoning Wang. 2011. *In silico* analysis and experimental improvement of taxadiene heterologous biosynthesis in *Escherichia coli*. *Biotechnol Bioprocess Eng* 16 (2):205-215.
- Morrone, D., L. Lowry, M. K. Determan, D. M. Hershey, M. Xu, and R. J. Peters. 2010. Increasing diterpene yield with a modular metabolic engineering system in *E. coli*: comparison of MEV and MEP isoprenoid precursor pathway engineering. *Appl Microbiol Biotechnol* 85 (6):1893-906.
- Muramatsu, M., C. Ohto, S. Obata, E. Sakuradani, and S. Shimizu. 2009. Alkaline pH enhances farnesol production by *Saccharomyces cerevisiae*. *J Biosci Bioeng* 108 (1):52-5.
- Nissen, T. L., M. Anderlund, J. Nielsen, J. Villadsen, and M. C. Kielland-Brandt. 2001. Expression of a cytoplasmic transhydrogenase in *Saccharomyces cerevisiae* results in formation of 2-oxoglutarate due to depletion of the NADPH pool. *Yeast* 18 (1):19-32.
- Nookaew, I., M. C. Jewett, A. Meechai, C. Thammarongtham, K. Laoteng, S. Cheevadhanarak, J. Nielsen, and S. Bhumiratana. 2008. The genome-scale metabolic model iIN800 of *Saccharomyces cerevisiae* and its validation: a scaffold to query lipid metabolism. *BMC Syst Biol* 2:71.
- Nookaew, I., A. Meechai, C. Thammarongtham, K. Laoteng, V. Ruanglek, S. Cheevadhanarak, J. Nielsen, and S. Bhumiratana. 2007. Identification of flux regulation coefficients from elementary flux modes: A systems biology tool for analysis of metabolic networks. *Biotechnol Bioeng* 97 (6):1535-49.
- Pandey, Ashok, Carlos R. Soccol, Poonam Nigam, and Vanete T. Soccol. 2000. Biotechnological potential of agro-industrial residues. I: sugarcane bagasse. *Bioresour Technol* 74 (1):69-80.
- Papin, J. A., J. Stelling, N. D. Price, S. Klamt, S. Schuster, and B. O. Palsson. 2004. Comparison of network-based pathway analysis methods. *Trends Biotechnol* 22 (8):400-5.
- Paradise, E. M., J. Kirby, R. Chan, and J. D. Keasling. 2008. Redirection of flux through the FPP branch-point in *Saccharomyces cerevisiae* by down-regulating squalene synthase. *Biotechnol Bioeng* 100 (2):371-8.
- Partow, S., V. Siewers, L. Daviet, M. Schalk, and J. Nielsen. 2012. Reconstruction and Evaluation of the Synthetic Bacterial MEP Pathway in *Saccharomyces cerevisiae*. *PLoS One* 7 (12):e52498.
- Pitera, D. J., C. J. Paddon, J. D. Newman, and J. D. Keasling. 2007. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metab Eng* 9 (2):193-207.

- Polakowski, T., U. Stahl, and C. Lang. 1998. Overexpression of a cytosolic hydroxymethylglutaryl-CoA reductase leads to squalene accumulation in yeast. *Appl Microbiol Biotechnol* 49 (1):66-71.
- Puustinen, A., M. Finel, T. Haltia, R. B. Gennis, and M. Wikstrom. 1991. Properties of the two terminal oxidases of *Escherichia coli*. *Biochemistry* 30 (16):3936-42.
- Randez-Gil, F., P. Sanz, and J. A. Prieto. 1999. Engineering baker's yeast: room for improvement. *Trends Biotechnol* 17 (6):237-44.
- Ro, D. K., E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Chang, S. T. Withers, Y. Shiba, R. Sarpong, and J. D. Keasling. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440 (7086):940-3.
- Rodriguez-Villalon, A., J. Perez-Gil, and M. Rodriguez-Concepcion. 2008. Carotenoid accumulation in bacteria with enhanced supply of isoprenoid precursors by upregulation of exogenous or endogenous pathways. *J Biotechnol* 135 (1):78-84.
- Rohdich, F., F. Zepeck, P. Adam, S. Hecht, J. Kaiser, R. Laupitz, T. Grawert, S. Amslinger, W. Eisenreich, A. Bacher, and D. Arigoni. 2003. The deoxyxylulose phosphate pathway of isoprenoid biosynthesis: studies on the mechanisms of the reactions catalyzed by IspG and IspH protein. *Proc Natl Acad Sci U S A* 100 (4):1586-91.
- Saloheimo, A., J. Rauta, O. V. Stasyk, A. A. Sibirny, M. Penttila, and L. Ruohonen. 2007. Xylose transport studies with xylose-utilizing *Saccharomyces cerevisiae* strains expressing heterologous and homologous permeases. *Appl Microbiol Biotechnol* 74 (5):1041-52.
- Sauer, U., F. Canonaco, S. Heri, A. Perrenoud, and E. Fischer. 2004. The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*. *J Biol Chem* 279 (8):6613-9.
- Scalcinati, G., S. Partow, V. Siewers, M. Schalk, L. Daviet, and J. Nielsen. 2012. Combined metabolic engineering of precursor and co-factor supply to increase alpha-santalene production by *Saccharomyces cerevisiae*. *Microb Cell Fact* 11 (1):117.
- Schomburg, I., A. Chang, S. Placzek, C. Sohngen, M. Rother, M. Lang, C. Munaretto, S. Ulas, M. Stelzer, A. Grote, M. Scheer, and D. Schomburg. 2013. BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. *Nucleic Acids Res* 41 (Database issue):D764-72.
- Schuster, S., D. A. Fell, and T. Dandekar. 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat Biotechnol* 18 (3):326-32.
- Schuster, S., C. Hilgetag, J. H. Woods, and D. A. Fell. 2002. Reaction routes in biochemical reaction systems: algebraic properties, validated calculation procedure and example from nucleotide metabolism. *J Math Biol* 45 (2):153-81.
- Seemann, M., B. T. Bui, M. Wolff, D. Tritsch, N. Campos, A. Boronat, A. Marquet, and M. Rohmer. 2002. Isoprenoid biosynthesis through the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GcpE) is a [4Fe-4S] protein. *Angew Chem Int Ed Engl* 41 (22):4337-9.
- Shiba, Y., E. M. Paradise, J. Kirby, D. K. Ro, and J. D. Keasling. 2007. Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* 9 (2):160-8.
- Takahashi, S., T. Kuzuyama, H. Watanabe, and H. Seto. 1998. A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc Natl Acad Sci U S A* 95 (17):9879-84.

- Terzer, M., and J. Stelling. 2008. Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics* 24 (19):2229-35.
- Trinh, C. T., R. Carlson, A. Wlaschin, and F. Sreenc. 2006. Design, construction and performance of the most efficient biomass producing *E. coli* bacterium. *Metab Eng* 8 (6):628-38.
- Trinh, C. T., and F. Sreenc. 2009. Metabolic engineering of *Escherichia coli* for efficient conversion of glycerol to ethanol. *Appl Environ Microbiol* 75 (21):6696-705.
- Trinh, C. T., P. Unrean, and F. Sreenc. 2008. Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. *Appl Environ Microbiol* 74 (12):3634-43.
- Trinh, C. T., A. Wlaschin, and F. Sreenc. 2009. Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl Microbiol Biotechnol* 81 (5):813-26.
- Tsuruta, H., C. J. Paddon, D. Eng, J. R. Lenihan, T. Horning, L. C. Anthony, R. Regentin, J. D. Keasling, N. S. Renninger, and J. D. Newman. 2009. High-level production of amorpho-4,11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS One* 4 (2):e4489.
- Unrean, P., C. T. Trinh, and F. Sreenc. 2010. Rational design and construction of an efficient *E. coli* for production of diapolycopendioic acid. *Metab Eng* 12 (2):112-22.
- Verho, R., J. Londesborough, M. Penttila, and P. Richard. 2003. Engineering redox cofactor regeneration for improved pentose fermentation in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 69 (10):5892-7.
- Wang, C., S. H. Yoon, A. A. Shah, Y. R. Chung, J. Y. Kim, E. S. Choi, J. D. Keasling, and S. W. Kim. 2010. Farnesol production from *Escherichia coli* by harnessing the exogenous mevalonate pathway. *Biotechnol Bioeng* 107 (3):421-9.
- Westfall, P. J., D. J. Pitera, J. R. Lenihan, D. Eng, F. X. Woolard, R. Regentin, T. Horning, H. Tsuruta, D. J. Melis, A. Owens, S. Fickes, D. Diola, K. R. Benjamin, J. D. Keasling, M. D. Leavell, D. J. McPhee, N. S. Renninger, J. D. Newman, and C. J. Paddon. 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci U S A* 109 (3):E111-8.
- Wieczorke, R., S. Krampe, T. Weierstall, K. Freidel, C. P. Hollenberg, and E. Boles. 1999. Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. *FEBS Lett* 464 (3):123-8.
- Wolff, M., M. Seemann, B. Tse Sum Bui, Y. Frapart, D. Tritsch, A. G. Estrabot, M. Rodríguez-Concepción, A. Boronat, A. Marquet, and M. Rohmer. 2003. Isoprenoid biosynthesis via the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase (LytB/IspH) from *Escherichia coli* is a [4Fe-4S] protein. *FEBS Letters* 541 (1-3):115-120.
- Woodrow, C. J., R. K. Haynes, and S. Krishna. 2005. Artemisinins. *Postgrad Med J* 81 (952):71-8.
- Yoon, S. H., S. H. Lee, A. Das, H. K. Ryu, H. J. Jang, J. Y. Kim, D. K. Oh, J. D. Keasling, and S. W. Kim. 2009. Combinatorial expression of bacterial whole mevalonate pathway for the production of beta-carotene in *E. coli*. *J Biotechnol* 140 (3-4):218-26.
- Yuan, L. Z., P. E. Rouviere, R. A. Larossa, and W. Suh. 2006. Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. *Metab Eng* 8 (1):79-90.



- Zhao, Jing, Qingyan Li, Tao Sun, Xinna Zhu, Hongtao Xu, Jinlei Tang, Xueli Zhang, and Yanhe Ma. 2013. Engineering central metabolic modules of *Escherichia coli* for improving  $\beta$ -carotene production. *Metab Eng* 17:42-50.
- Zhou, K., L. Zhou, Q. Lim, R. Zou, G. Stephanopoulos, and H. P. Too. 2011. Novel reference genes for quantifying transcriptional responses of *Escherichia coli* to protein overexpression by quantitative PCR. *BMC Mol Biol* 12:18.

## CHAPTER 3

### ***In vivo* validation of *in silico* predicted metabolic engineering strategies in yeast: disruption of $\alpha$ -ketoglutarate dehydrogenase and expression of ATP-citrate lyase for terpenoid production**

Evamaria Gruchattka, Michael Felten, Nicole Jurisch, Verena Schütz, Oliver Kayser

EG was responsible for all experiments and interpretations, drafted and wrote the manuscript. Master student MF was involved in preliminary studies and the construction of pSP-P. Master student NJ was involved in the construction of the *KGD1* knockout strain. VS was involved in the design of the preliminary studies. OK supervised the research.

Published in: POLS ONE. 2015. 10(12):e0144981

### 3.1 Abstract

Engineering of the central carbon metabolism of *Saccharomyces cerevisiae* to redirect metabolic flux towards cytosolic acetyl-CoA has become a central topic in yeast biotechnology. A cell factory with increased flux into acetyl-CoA can be used for heterologous production of terpenoids for pharmaceuticals, biofuels, fragrances, or other acetyl-CoA derived compounds. In a previous study, we identified promising metabolic engineering targets in *S. cerevisiae* using an *in silico* stoichiometric metabolic network analysis. Here, we validate selected *in silico* strategies *in vivo*.

Patchoulol was produced by yeast via a heterologous patchoulol synthase of *Pogostemon cablin*. To increase the metabolic flux from acetyl-CoA towards patchoulol, a truncated HMG-CoA reductase was overexpressed and farnesyl diphosphate synthase was fused with patchoulol synthase. The highest increase in production could be achieved by modifying the carbon source; sesquiterpenoid titer increased from glucose to ethanol by a factor of 8.4. Two strategies predicted *in silico* were chosen for validation in this work. Disruption of  $\alpha$ -ketoglutarate dehydrogenase gene (*KGD1*) was predicted to redirect the metabolic flux via the pyruvate dehydrogenase bypass towards acetyl-CoA. The metabolic flux was redirected as predicted, however, the effect was dependent on cultivation conditions and the flux was interrupted at the level of acetate. High amounts of acetate were produced. As an alternative pathway to synthesize cytosolic acetyl-CoA, ATP-citrate lyase was expressed as a polycistronic construct, however, *in vivo* performance of the enzyme needs to be optimized to increase terpenoid production.

Stoichiometric metabolic network analysis can be used successfully as a metabolic prediction tool. However, this study highlights that kinetics, regulation and cultivation conditions may interfere, resulting in poor *in vivo* performance. Main sites of regulation need to be released and improved enzymes are essential to meet the required activities for an increased product formation *in vivo*.

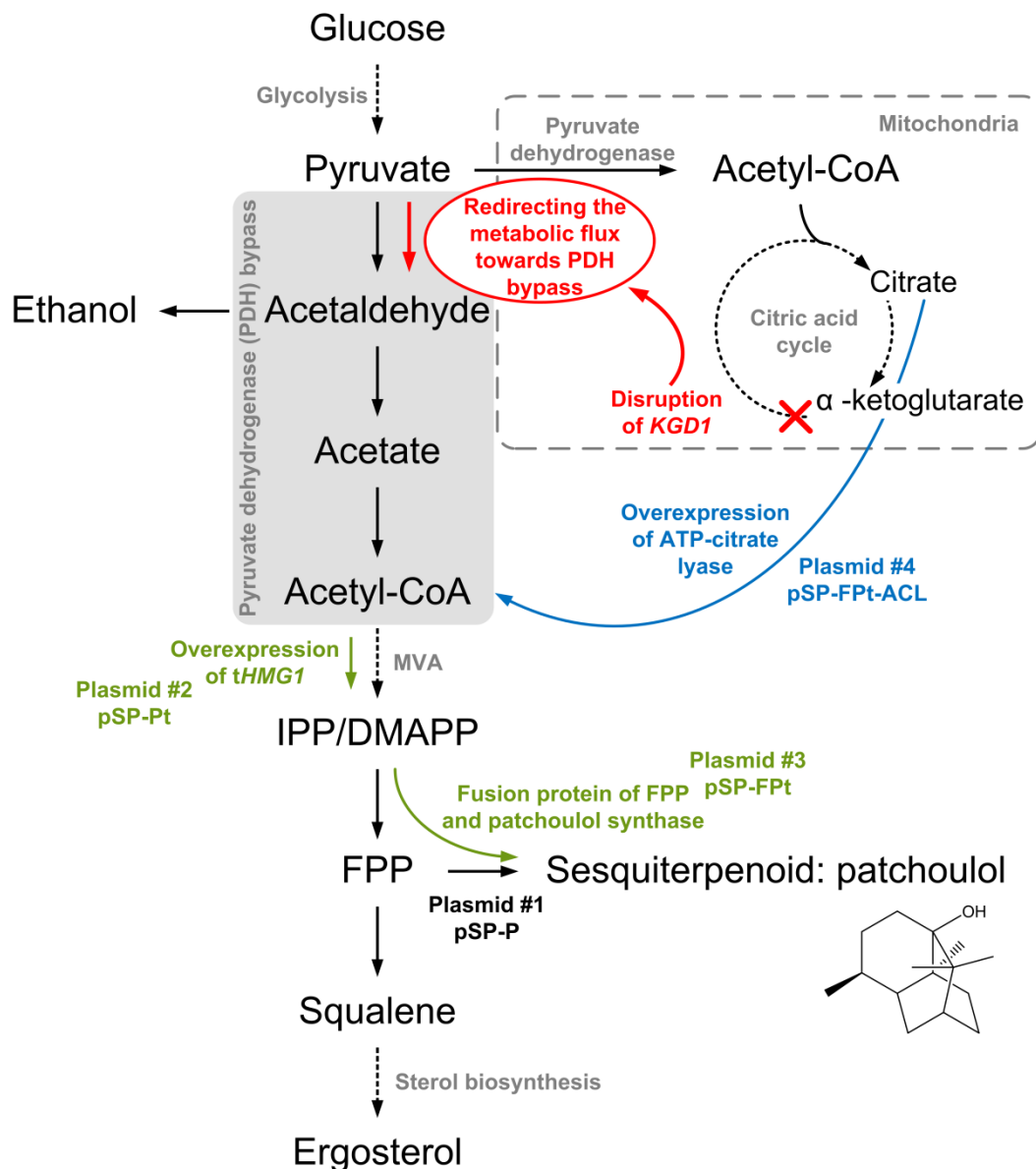
## 3.2 Background

Terpenoids are one of the largest classes of natural products comprising of tens of thousands of compounds. These hydrocarbon compounds have multiple applications from pharmaceuticals to biofuels and fragrances (George et al. 2015). Heterologous production of terpenoids with the yeast *Saccharomyces cerevisiae* has become popular due to advances in molecular biology tools and the general robustness of this organism in fermentation processes (Tippmann et al. 2013). Metabolic engineering, which involves detailed metabolic analysis to identify targets, followed by directed genetic modifications for improvement of cells via recombinant DNA technology (Nielsen and Keasling 2011), represents a strategy to increase heterologous terpenoid formation from the yeast host. Early heterologous terpenoid production studies in yeast focused on engineering the native terpenoid biosynthetic pathway, the mevalonate pathway (MVA) by several strategies including: overexpressing a truncated and thus soluble as well as non-regulated version of HMG-CoA reductase (*tHMG1*) (Jackson, Hart-Wells, and Matsuda 2003; Ro et al. 2006), down-regulating squalene synthase (*ERG9*), which leads via several steps to the major end product ergosterol (Asadollahi et al. 2008; Paradise et al. 2008) as well as overexpression of a mutant transcription factor *upc2-1* upregulating several MVA pathway genes (Ro et al. 2006). However, little research had been done concerning modifications within the central carbon metabolism to redirect the metabolic flux towards terpenoids. Thus, our group has previously analyzed the metabolic networks of *S. cerevisiae* and also *Escherichia coli*, the most prominently used hosts for heterologous terpenoid production, and identified promising metabolic engineering strategies *in silico* using a stoichiometric metabolic network analysis (Gruchattka et al. 2013). Based on our analysis, both hosts have the potential to produce terpenoids to higher yields than previously achieved with these newly identified targets for metabolic engineering. *S. cerevisiae* was chosen here for this validation study due to its robustness and the higher probability to functionally express plant P450 enzymes, which are necessary for the biosynthesis of several terpenoids. Different carbon sources were analysed for their potential to supply terpenoids in high yields and ethanol was identified as the most promising carbon source for yeast (Gruchattka et al. 2013). However, ethanol is an expensive carbon substrate, and less expensive alternatives in the form of sugars are preferable. Molasses and hemicellulosic hydrolysates of agricultural by-products represent inexpensive carbon substrates, which contain high amounts of simple sugars (Carvalho, Duarte, and Gírio 2008). Metabolic engineering

strategies identified *in silico*, which employ glucose as substrate, can as well be applied to other sugars like fructose, galactose or even xylose (Gruchattka et al. 2013). Thus, the focus here was set on the *in vivo* validation of metabolic engineering strategies based on glucose as carbon source in the yeast *S. cerevisiae*. Patchoulol synthase (Pts), which produces the sesquiterpenoid patchoulol as its main product, was chosen as reporter in this study.

In this report, we analyze one metabolic engineering strategy *in vivo*, which was based on the constrained minimal cut sets approach (Hädicke and Klamt 2011). In this approach a gene knockout within the citric acid cycle ( $\alpha$ -ketoglutarate dehydrogenase) in addition to inhibiting ethanol and acetate production by the cell was predicted to yield high productivities of terpenoids. The rationale behind this strategy is that the metabolic flux is redirected from citric acid cycle via the pyruvate dehydrogenase bypass towards acetyl-CoA for increased terpenoid production (see Figure 3.1).

A second engineering target was a heterologous pathway to synthesize cytosolic acetyl-CoA. Pyruvate dehydrogenase bypass is the major source of cytosolic acetyl-CoA in *S. cerevisiae*. However, most other organisms possess ATP-citrate lyase as alternative pathway to synthesize cytosolic acetyl-CoA (see Figure 3.1). The heterologous enzyme was identified as promising metabolic engineering target for increased terpenoid production in our *in silico* computation (Gruchattka et al. 2013). The rationale behind this is that citrate is the substrate for acetyl-CoA synthesis and is produced via the pyruvate dehydrogenase complex and citric acid cycle, circumventing pyruvate dehydrogenase bypass and especially acetyl-CoA synthetase, which requires one additional mol ATP per acetyl-CoA.



**Figure 3.1 - Metabolic engineering approaches in yeast for increased terpenoid production.**

A simplified metabolic network of yeast depicting most important reactions is shown. Production of the sesquiterpenoid patchoulol is enabled via expression of patchoulol synthase (via plasmid #1 pSP-P). Overexpression of truncated HMG-CoA reductase (*tHMG1*, via plasmid #2 pSP-Pt) and fusion of FPP synthase with patchoulol synthase (via plasmid #3 pSP-FPt) to increase the flux from the central carbon precursor acetyl-CoA to the terpenoid product are highlighted in green. The disruption of  $\alpha$ -ketoglutarate dehydrogenase (*KGD1*) to redirect the metabolic flux from citric acid cycle via pyruvate dehydrogenase bypass towards terpenoids is shown in red. Overexpression of ATP-citrate lyase (via plasmid #4 pSP-FPt-ACL) to produce cytosolic acetyl-CoA via pyruvate dehydrogenase and citric acid cycle instead of pyruvate dehydrogenase bypass is shown in blue.

This work was conducted in order to confirm these predictions with all discussed possibilities and limitations and identify possible considerations in order to improve stoichiometric metabolic network predictions *in silico* and reliably increase productivity *in vivo*. In addition, this study includes prerequisites important for validating metabolic

engineering strategies within the central carbon metabolism for enhanced terpenoid production, *i.e.* a two-phase cultivation and identification of the product spectrum of the terpenoid synthase. Furthermore, the flux in the mevalonate pathway was increased towards the desired product via overexpression of a truncated HMG-CoA reductase (*tHMG1*) and fusion of FPP synthase with patchoulol synthase to ensure that a modified flux distribution in the central carbon metabolism is resulting in increased terpenoid formation.

### 3.3 Methods

#### 3.3.1 Strains, media and growth conditions

Plasmids were amplified in *E. coli* DH5 $\alpha$  (LifeTechnologies). Cells were cultured at 37 °C and 200 rpm in LB medium (5 g/L yeast extract, 10 g/L tryptone, 10 g/L NaCl, pH 7, for solid medium 20 g/L agar) containing 100 mg/L ampicillin when necessary. Yeast strain *S. cerevisiae* CEN.PK2-1C (*MATa*; *ura3-52 MAL2-8C SUC2*; Euroscarf; the auxotrophic markers *HIS3*, *TRP1* and *LEU2* had been reintroduced) was used as background in all experiments. Yeast strains were cultured at 30 °C and 200 rpm in unbaffled shake flasks. Yeast strains not carrying a plasmid were grown in YM (11 g/L glucose x H<sub>2</sub>O, 3 g/L yeast extract, 5 g/L peptone and 3 g/L malt extract, for solid medium 20 g/L agar). Yeast strains carrying a plasmid were grown in synthetic dextrose (SD) medium containing 20 g/L glucose x H<sub>2</sub>O, 6.7 g/L yeast nitrogen base without amino acids, 10.2 g/L KH-phthalate, adjusted to pH 5.5 with KOH and for solid medium 20 g/L agar. Strains were stored in 30 % glycerol at -80 °C in cryo-vials.

#### 3.3.2 Plasmid construction and transformation

Plasmids used and constructed in this study are depicted in Table 3.1, maps of expression plasmids, gene sequences and primers used in this study are shown in supplementary information (Figure 3.16 and Tables 3.2-3.3). Plasmids were constructed exploiting yeasts homologous recombination machinery. Therefore fragments containing approx. 30-60 bp homologous overlap regions were amplified via PCR and transformed into yeast (250 ng per insert) together with the linearized target plasmid (500 ng). Yeast transformation was performed using a standard lithium acetate procedure (Gietz and Schiestl 2007). Plasmids were isolated from yeast, transformed into *E. coli* for amplification using a standard procedure for chemically competent cells

and heat shock transformation (Inoue, Nojima, and Okayama 1990) and isolated for verification via PCR and sequencing. Phusion® High-Fidelity DNA Polymerase (Finnzymes/Thermo Scientific) was used for amplifications, *Taq* DNA Polymerase 1.1x Master Mix Red (Ampliqon, biomol) was used to verify correct integrations.

**Table 3.1 - Plasmids used in this study.**

Plasmid	Characteristics	Source
pUG6	Plasmid carrying loxP-flanked marker gene deletion cassette: loxP- $P_{TEF1}$ -kanMX-T $_{TEF1}$ -loxP	Euroscarf
pSH47	Plasmid carrying Cre recombinase for loxP mediated marker gene removal	Euroscarf
P423H7	Plasmid carrying truncated HXT7 promoter (-392 bp to -1 bp)	(Becker and Boles 2003)
pYX042- <i>ACLA-1</i>	Plasmid carrying cDNA <i>ACLA-1</i>	(Fatland et al. 2002)
pYX012- <i>ACLB-2</i>	Plasmid carrying cDNA <i>ACLB-2</i>	(Fatland et al. 2002)
pSP-GM1	Yeast expression plasmid containing bidirectional promoters $P_{TEF1}$ and $P_{PGK1}$	(Chen et al. 2012)
pSP-P	pSP-GM1 carrying the sequence optimized patchoulol synthase cDNA ( <i>PTS</i> ) from <i>Pogostemon cablin</i> Benth. ( $P_{TEF1}$ - <i>PTS</i> )	this study
pSP-Pt	pSP-P carrying the truncated (first 1590 bp are cut, start codon was reintroduced) HMG-Co reductase gene ( <i>tHMG1</i> ) from <i>S. cerevisiae</i> CEN.PK2-1C ( $P_{TEF1}$ - <i>PTS</i> , $P_{PGK1}$ - <i>tHMG</i> )	this study
pSP-FPt	pSP-Pt carrying the farnesol diphosphate synthase gene ( <i>ERG20</i> ) from <i>S. cerevisiae</i> CEN.PK2-1C fused to <i>PTS</i> via a short flexible linker ( <i>GSG</i> ) ( $P_{TEF1}$ - <i>FPPS-PTS</i> , $P_{PGK1}$ - <i>tHMG</i> )	this study
pSP-FPt-ACL	pSP-FPt carrying the ATP-citrate lyase cDNAs ( <i>ACLA-1</i> and <i>ACLB-2</i> ) from <i>Arabidopsis</i> (Fatland et al. 2002) fused via T2A sequences from <i>Thosea asigna</i> virus ( $P_{TEF1}$ - <i>FPPS-PTS</i> , $P_{PGK1}$ - <i>tHMG</i> , $P_{HXT7^*}$ - <i>ACLA-1-T2A-ACLB-2</i> )	this study

The cDNA encoding patchoulol synthase (*PTS*) was sequence optimized for expression in *S. cerevisiae* (synthesized by Geneart) and is based on the cDNA of *Pogostemon cablin* Benth. [GenBank: AY508730.1]. *PTS* was amplified using primers MF0018 and MF0019 adding overlaps homologous to plasmid pSP-GM1, while pSP-GM1 was linearized using restriction enzymes *PacI* and *NotI*. Both were transformed into yeast, the plasmid pSP-P was isolated and analyzed for correct PCR product size and sequenced using primers MF0016\_S and MF0017\_S. The truncated version (first 1,590 bp were removed, start codon was reintroduced) of the endogenous yeast gene of HMG-CoA reductase (*tHMG1*) was PCR amplified using primers EG\_tHMG\_Fw and EG\_tHMG\_Rv



and genomic DNA of yeast as template. The fragment was transformed into yeast together with pSP-P (linearized with *Xho*I) and pSP-Pt was isolated and verified using primers MF0014\_S and MF\_0015\_S. The endogenous yeast gene of farnesyl diphosphate synthase (*ERG20*) was amplified from genomic yeast DNA using primers EMG\_FPPS\_Fw and EMG\_FPPS\_Rv. The fragment was transformed into yeast together with pSP-Pt (linearized with *Not*I) generating a fusion of farnesyl diphosphate synthase and patchoulol synthase via a short flexible linker (GSG) (Albertsen et al. 2011). pSP-FPt was isolated and verified using primers EMG\_FPPS\_S and MF0017\_S. The yeast strain *S. cerevisiae* CEN.PK2-1C was transformed with the constructed plasmids as described above.

ATP-citrate lyase of *Arabidopsis* sp. consists of two subunits. The cDNAs of both subunits, *ACLA-1* and *ACLB-2*, were amplified from pYX042-*ACLA-1* and pYX012-*ACLB-2* (Fatland et al. 2002) using primers EMG\_ACL\_3\_PHXT7+*ACLA-1* and EMG\_ACL\_4\_ACLA-1+T2A for *ACLA-1* and primers EMG\_ACL\_5\_T2A+*ACLB-2* and EMG\_ACL\_6\_ACLB-2+TTEF for *ACLB-2* which generates a separation of both cDNAs by a T2A sequence from *Thosea asigna* virus (Beekwilder et al. 2014). The truncated *HXT7* promoter (-392 bp to -1 bp) was amplified from p423H7 using primers EMG\_ACL\_1\_pSP+PHXT7 and EMG\_ACL\_2\_PHXT7+*ACLA-1*. Moreover, the *TEF* terminator was amplified from pUG6 using primers EMG\_ACL\_7\_ACLB-2+TTEF and EMG\_ACL\_8\_TTEF+*Asc*I+pSP. The four fragments were transformed into yeast together with pSP-FPt (linearized with *Asc*I) for homologous recombination. The new plasmid pSP-FPt-ACL was isolated from yeast and verified using primers EMG\_ACL\_9\_S and EMG\_ACL\_10\_S, EMG\_ACL\_11\_S and EMG\_ACL\_12\_S, EMG\_ACL\_13\_S and EMG\_ACL\_14\_S.

### 3.3.3 Protein extraction, SDS-PAGE and Western blot analysis

Cells corresponding to 100 mg CDW were resuspended in 1 mL 50 mM Tris, 1 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 1 mM DTT, 1 mM PMSF, 1 µg/mL aprotinin, 1 µg/mL leupeptin, pH 7.2 and approx. 500 µL glass beads (0.75-1 mm diameter) were added before vortexing twice for 10 min. Samples were centrifuged (10 min, 16,000 *g*, 4 °C) and the protein concentration of the supernatant was determined using a colorimetric assay after Bradford (Bradford 1976) using bovine serum albumin as reference. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to Laemmli (Laemmli 1970) using a 4 % stacking gel and 12.5 % separating gel. Samples according to 20 µg protein were loaded per slot on a gel together with a protein marker

(Page Ruler™ Plus Prestained Protein Ladder, Thermo Scientific). Proteins were transferred to a nitrocellulose membrane at 0.8 mA/cm<sup>2</sup> for 1 h. After blocking for 1 h with PBS buffer containing Tween 20 and milk powder, the membranes were incubated with the primary antibody (1:1000 dilution of Anti-2A Peptide Antibody, polyclonal from rabbit, Cat. #ABS31, Lot # 2446870, Merck Millipore; validated previously (Beekwilder et al. 2014)) at 4 °C overnight, washed and incubated with the secondary antibody (Anti-Rabbit IgG, Sigma Aldrich) for 1 h. After washing, proteins were detected using a colorimetric assay with 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium (BCIP/NBT) in 50 mM sodium carbonate buffer.

### 3.3.4 Yeast strain construction

The disruption of  $\alpha$ -ketoglutarate dehydrogenase gene (*KGD1*) was performed using recyclable marker gene disruption cassettes as described by Güldener *et al.* (Güldener et al. 1996). The loxP flanked resistance cassette (loxP-P<sub>TEF1</sub>-kanMX-T<sub>TEF1</sub>-loxP) was amplified via PCR (Phusion® High-Fidelity DNA Polymerase, Finnzymes/Thermo Scientific) from pUG6 using primers EMG\_KGD\_Fw and EMG\_KGD\_Rv (see Additional file 2 for primer sequences) generating approx. 50 bp overlap regions homologous to the genomic locus to be disrupted. Yeast was transformed with 2  $\mu$ g PCR product and strains carrying the *kgd1* disruption were selected based on their resistance to G418. Moreover, genomic DNA was isolated and the correct integration was validated via PCR (*Taq* DNA Polymerase 1.1x Master Mix Red, Ampliqon, biomol) and sequencing using primers EMG\_KGD\_S1 and EMG\_KGD\_S2. The strain was transformed with pSH47 and Cre recombinase expression was induced in an overnight culture with 2 % galactose for 30 min. Subsequently, cells were selected on agar plates with and without G418 as efficient recombination leads again to sensitivity to G418. The loss of the kanMX cassette was verified as well via PCR using primers EMG\_KGD\_S1 and EMG\_KGD\_S2. Finally, cells were streaked on SD agar plates containing 2 g/L 5-fluoroorotic acid (5-FOA) and 80 mg/L uracil to select for those strains that have lost the plasmid pSH47. The newly generated yeast strain carrying the  $\alpha$ -ketoglutarate dehydrogenase gene disruption (*S. cerevisiae* CEN.PK2-1C: *MATa*; *ura3-52 MAL2-8<sup>c</sup> SUC2 kgd1 $\Delta$ ::loxP*) was transformed with plasmids as described above.

### 3.3.5 2-phase cultivations

For characterizations of terpenoid production, yeast strains carrying plasmids were cultivated in SD medium. Cells were streaked from glycerol stocks onto agar plates,

incubated for 3-4 days and used as inoculum for the first preculture (1/10 to 1/5 medium filling). The cells were harvested via centrifugation (5 min, 2,000 *g*, 4 °C) after approx. 24 h, resuspended in 0.9 % KCl and used to inoculate the second preculture to OD<sub>660</sub> of 0.2 – 0.4 (1/5 medium filling). After incubation overnight, the second preculture was harvested and cells were resuspended in 0.9 % KCl and used to inoculate the main cultures. For strain characterization during growth on glucose in batch mode as well as during growth on the produced side products ethanol, glycerol and acetate, main cultures (250 mL flasks with 1/5 medium filling) were inoculated to OD<sub>660</sub> of 0.4 and 7.5 % dodecane was added immediately to the cultures. As the fragrance compound patchoulol, the chosen sesquiterpenoid reporter, has volatile and potentially toxic properties, a two-phase cultivation was established as *in situ* product removal tool. Dodecane was chosen as it is widely used as second organic phase (Beekwilder et al. 2014; Peralta-Yahya et al. 2011; Scalcinati, Knuf, et al. 2012) due to its hydrophobicity (log P<sub>0/w</sub> of 6.6 is regarded as biocompatible (Heipieper et al. 1994)), low volatility and good phase separation. Carbon source consumption was assessed to harvest the cultures when glucose was abolished (14-17 h) or ethanol, glycerol and acetate (4 d and 6 d). For glucose-limited experiments, a third preculture (250 mL flasks with 1/5 medium filling) was inoculated to OD<sub>660</sub> of 0.4 and cultured for 15 h. Cells were harvested, washed in 0.9 % KCl, resuspended in SD medium without glucose and used for the main culture (OD<sub>660</sub> of 4.5, 500 mL flasks with 1/10 initial filling, 1/5 final filling). 7.5 % dodecane was added immediately and carbon limiting conditions were realized by feeding glucose (SD medium containing 0.2 % glucose x H<sub>2</sub>O) to the cultures at a rate of 51.5 µL/min for approx. 16 h using a multichannel peristaltic pump (IPC, ISM937, Ismatec). Cultivations were performed at least in triplicate.

### 3.3.6 OD and cell dry weight determination

The optical density (OD) of cultivation samples was determined at 660 nm in duplicate by using a UV-1800 Shimadzu UV spectrometer. Cell dry weight (CDW) was determined by using an OD-CDW correlation:  $0.4939 \times OD_{660} = \text{g/L CDW}$ . For this purpose, the OD<sub>660</sub> was determined in duplicate of 5 different samples of a cultivation and pellets of defined aliquots of each sample (triplicates) were dried at 75 °C for 6 days and weighed.

### 3.3.7 Metabolite and sugar analysis

Samples from the aqueous phase from cultivations were centrifuged (10 min, 16,000 *g*, 4 °C) and the supernatant was frozen at -20 °C if not measured directly. Glucose concentrations were estimated using Combur3Test® (Roche) test stripes. Glucose and extracellular metabolites were quantified using HPLC/RID. An Agilent Technologies HPLC system, equipped with an RID detector at 35 °C, was used. Compounds were separated on a Hi-Plex H column (300 x 7.7 mm, ID 6.5 mm, particle size 8 µm) with 5 mM sulfuric acid as isocratic mobile phase with a flow of 0.5 mL/min at 65 °C for 38 minutes. The injection volume was 20 µL. Glucose, ethanol, glycerol and acetate were quantified using standard curves of the reference compounds.

### 3.3.8 Analysis of sesquiterpenoids in the organic layer

Cultures were harvested by centrifugation (10 min at 5,000 *g*, 4 °C); a sample of the dodecane layer was taken and dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, centrifuged for a second time and stored at -20 °C if not measured directly. 100 µL of sample were diluted in 400 µL of n-hexane including 250 µM α-humulene (Sigma Aldrich) as internal standard and measured using GC/FID. An Agilent Technologies 7890A GC system equipped with a flame ionization detector (FID; 300 °C, H<sub>2</sub> flow set to 30 mL/min, air flow set to 400 mL/min, make-up gas (N<sub>2</sub>) flow set to 30 mL/min) was used for terpenoid quantification. A FactorFour VF-5ms column (CP8944; 30 m x 0.25 mm, ID 0.25 µm) was used and 1 µL sample was injected in splitless mode (inlet at 250 °C). The carrier gas (H<sub>2</sub>) flow was set to 1.4 mL/min. The initial oven temperature was set to 80 °C; after one min the oven temperature was increased to 120 °C at a rate of 10 °C/min. Thereafter, the temperature was increased to 160 °C at a rate of 3 °C/min and then to 180 °C at a rate of 10 °C/min. The oven temperature was finally increased to 270 °C at a rate of 30 °C/min and held for 2.5 min. Patchoulol was quantified using standard curves of a reference compound (patchouli alcohol, PhytoLab). Measurements were performed in duplicate. Additional sesquiterpenoids formed by the patchoulol synthase with nominal masses of 222 and 204 were identified by GC/MS and the comparative chemical composition of sesquiterpenoids was performed using FID using a 7890A GC system (Agilent Technologies). A HP-5MS column (30 m x 0.25 mm, ID 0.25 µm) was used and 1 µL sample was injected in splitless mode (inlet at 250 °C). The carrier gas (He) flow was set to 1.1 mL/min. The oven temperature program described above was used. The ion

source was kept at 230 °C, the detection was performed in electron impact ionization mode at 70 eV and the mass spectrum was recorded from 50 to 550 m/z.

### 3.3.9 Sterol extraction and analysis

A modified protocol after Asadollahi *et al.* (Asadollahi *et al.* 2008) was used. Cells were harvested by centrifugation (4,000 *g*, 10 min) and the cell pellet was washed and then resuspended in dH<sub>2</sub>O. Cells corresponding to 30 mg CDW were pelleted again and resuspended in 4 mL 0.2 M HCl and incubated for 1 hour at 85°C in a water bath. Samples were cooled to room temperature and centrifuged (10 min, 4,000 *g*). The cell pellet was resuspended in 2 mL methanol and 1 mL 4 M KOH, transferred to a dark glass vials and incubated at 85 °C for 2 hours in a water bath. Samples were cooled to room temperature, 5 mL heptane including 250 µM α-tocopherol (Sigma Aldrich) as internal standard was added and vortexed for 2 min. The heptane phase was taken for analytics. Ergosterol was quantified using HPLC/DAD and squalene was quantified using GC/FID, both via standard curves of the reference compounds (Sigma Aldrich). An Agilent Technologies HPLC system equipped with a 1260 Infinity DAD was used. Compounds were separated on a Poroshell 120 EC-C18 column (2.1 x 100 mm, 2.7 µM) with acetonitrile as isocratic mobile phase with a flow of 7 mL/min at 35 °C for 6 minutes. The injection volume was 10 µL. Ergosterol and α-tocopherol were detected at 280 nm. The GC/FID system described in 'Analysis of sesquiterpenoids in the organic layer' was used with a modified oven temperature program. The initial oven temperature was set to 80 °C; after one min the oven temperature was increased to 270 °C at a rate of 10 °C/min and held for 20 min. GC measurements were performed in duplicate.

### 3.3.10 Fatty acid extraction, derivatization and analytics

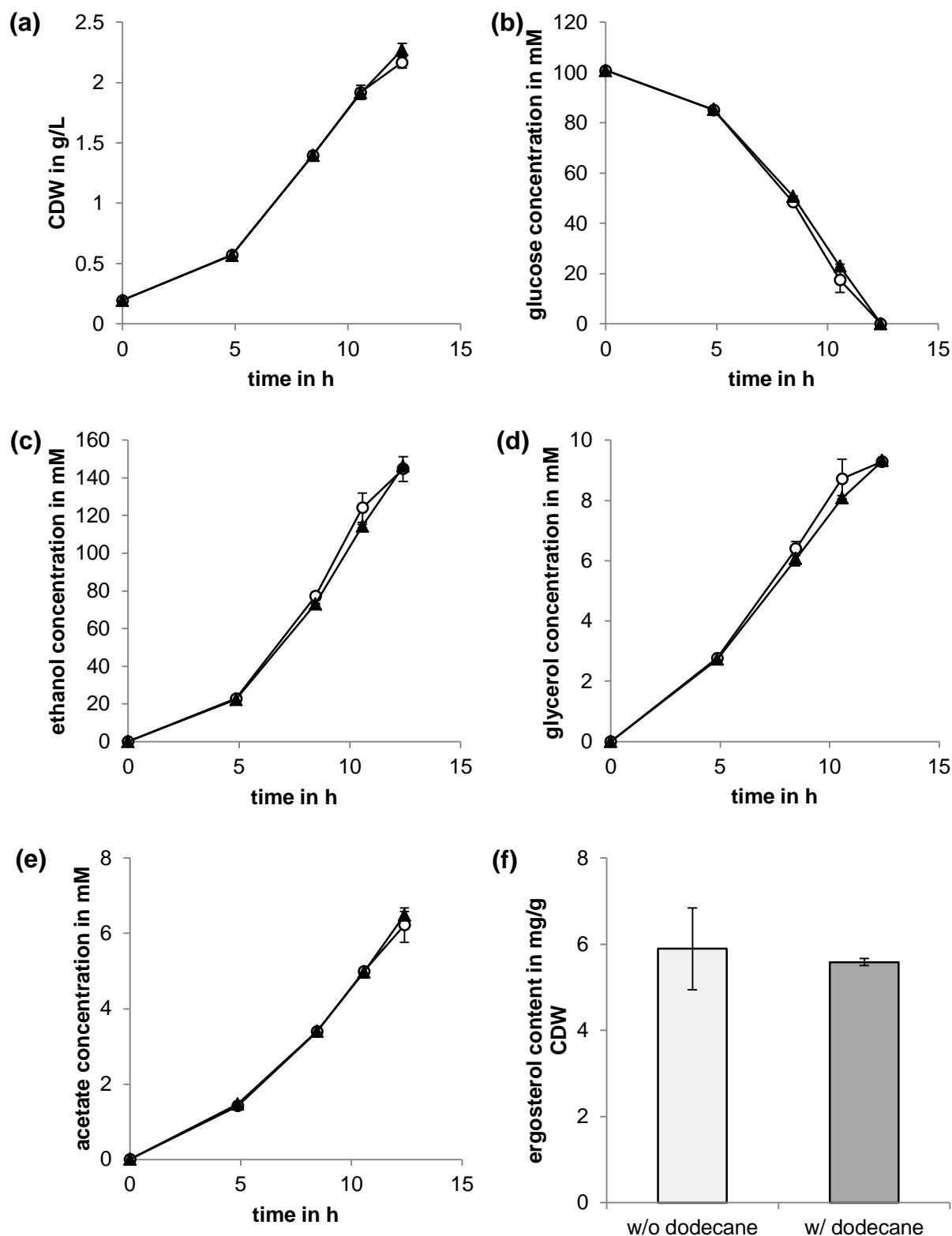
A modified protocol after Chen *et al.* (Chen, Zhang, and Chen 2014) was used. Cells corresponding to 5 mg CDW were resuspended in 500 µL 10 mM Tris pH 7.5, 10 µL 2-mercaptoethanol and 200 U lyticase (from *Arthrobacter luteus*, Sigma Aldrich) and incubated at 30 °C for 30 min. Approx. 500 µL of glass beads (0.75-1 mm diameter), another 500 µL buffer and 200 µL acetic acid were added and vortexed two times for 5 min. 10 µL of 10 mg/mL heptadecanoic acid (Sigma Aldrich) dissolved in chloroform were added as internal standard. 3 mL of a chloroform-methanol (2:1) mixture was added, vortexed vigorously for 30 sec, centrifuged (10 min, 5,000 *g*, 4 °C) and the generated chloroform layer was collected. Additional 3 mL of the chloroform-methanol mixture were added to the aqueous and cell debris layer, vortexed and centrifuged

again. The chloroform layer were combined and evaporated to dryness using a rotary evaporator (Rotavapor® R-210, Büchi) and afterwards a desiccator with silica gel orange overnight. The dried lipid residue was dissolved in 500  $\mu$ L 10 %  $\text{BF}_3$ -methanol (Sigma Aldrich) and incubated in sealed screw cap tube at 90 °C for 20 min. After cooling down to room temperature, 300  $\mu$ L of a saturated NaCl solution was added and vortexed. Fatty acid methyl esters were extracted with 600  $\mu$ L n-hexane via harsh vortexing and analyzed using GC/FID. The GC/FID system described in 'Analysis of sesquiterpenoids in the organic layer' was used with a modified oven temperature program. The initial oven temperature was set to 50 °C; after one min the oven temperature was increased to 270 °C at a rate of 7 °C/min and held for 10 min. Fatty acid methyl esters (F.A.M.E.) mix  $\text{C}_8 - \text{C}_{24}$  (CRM18918 Supelco) as well as GC-MS (system described in 'Analysis of sesquiterpenoids in the organic layer', oven temperature program identical to the one from GC/FID) were used for identification of fatty acid methyl esters.

## 3.4 Results

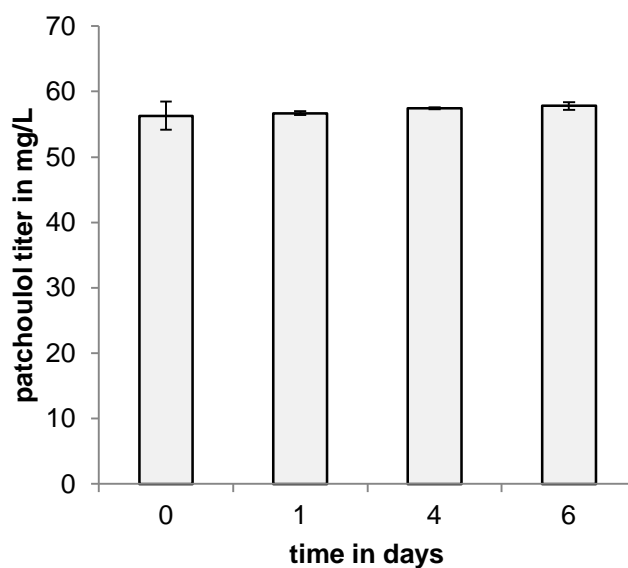
### 3.4.1 Sesquiterpenoid production in *S. cerevisiae*: evaluation of two-phase cultivation conditions and sesquiterpenoid spectrum

Since the fragrance compound patchoulol, the chosen reporter sesquiterpenoid, has volatile and potentially toxic properties, a two-phase cultivation using dodecane was established. Dodecane showed no detrimental effects on physiological parameters (cell growth, glucose consumption, ethanol, glycerol and acetate formation) or ergosterol production when 7.5 % were added to yeast cultures carrying pSP-GM1 as empty vector (see Figure 3.2). Moreover, dodecane captures patchoulol efficiently over time and prevents loss in the air (Figure 3.3). Thus, a two-phase cultivation using dodecane is well suited. *S. cerevisiae* was transformed with plasmid pSP-P carrying a synthetic gene coding for patchoulol synthase under control of *TEF1* promoter and was analyzed for patchoulol production using the two-phase cultivation procedure with dodecane. Patchoulol was produced and its accumulation in the organic phase was monitored over time. Patchoulol formation curve and biomass formation curve run in parallel and both stagnate as soon as glucose is consumed (Figure 3.4). Moreover, additional sesquiterpenoids were detected. Using a long-time experiment, high patchoulol concentrations were achieved and within these samples 22 peaks with nominal masses of 222 (sesquiterpenoids with hydroxy-group) and 204 (sesquiterpenoids without hydroxy-group) could be detected (Figure 3.5). Based on a comparative chemical composition of sesquiterpenoids produced, patchoulol represents 34.3 % +/- 1.6 of total sesquiterpenoids. The patchoulol content determined here is used hereafter to estimate the production of total sesquiterpenoids in this work.

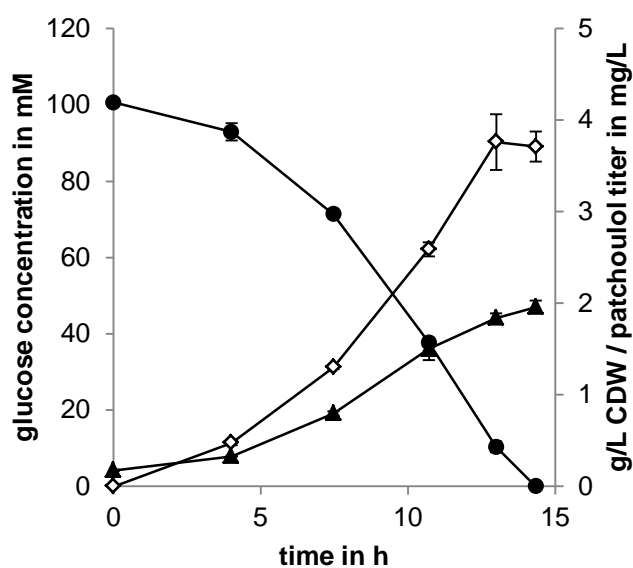


**Figure 3.2 - Effect of dodecane on yeast cultures.** Profiles of (a) biomass formation (cell dry weight, CDW), (b) glucose consumption, (c) ethanol formation, (d) glycerol formation and (e) acetate formation as a function of time as well as (f) ergosterol content of cells after glucose was exhausted. No second phase (○) or 7.5 % dodecane (▲) were added to cultures of the yeast strain carrying pSP-GM1 grown in batch mode in shake flasks. Mean values and standard deviations of 3 experiments are shown.

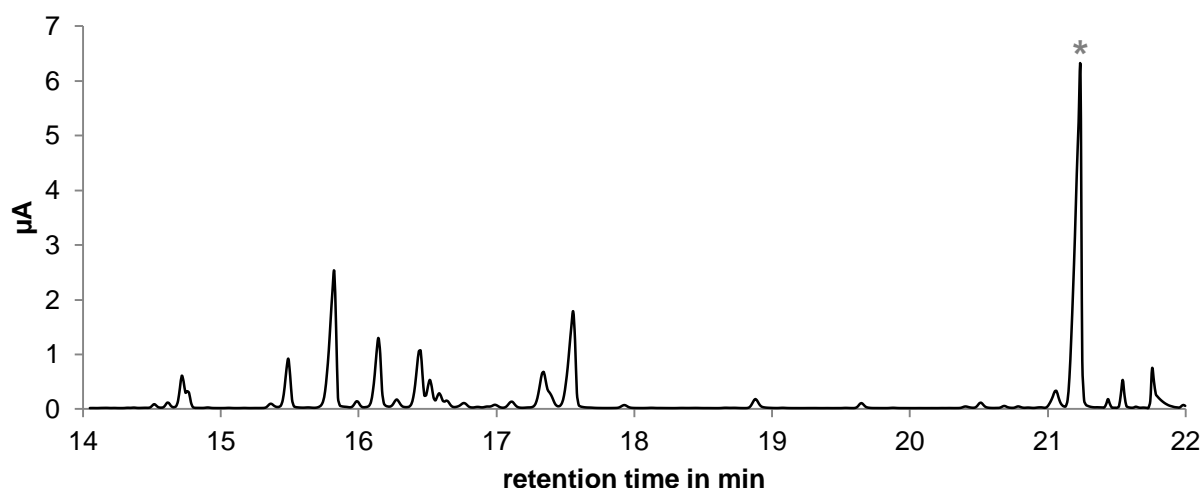




**Figure 3.3 - Effectiveness of dodecane to capture patchoulol over time.** 7.5 % dodecane including patchoulol was added to sterile culture medium, flasks were incubated analogous to yeast strains and harvested. Patchoulol titre was determined at different time points. Mean values and standard deviations of 3 experiments are shown.



**Figure 3.4 - Profiles of patchoulol, glucose and biomass.** Patchoulol formation ( $\diamond$ ) was determined as a function of time together with biomass formation (cell dry weight, CDW) ( $\blacktriangle$ ) and glucose consumption ( $\bullet$ ) for the patchoulol producing yeast strain carrying pSP-P. Cells were grown in batch mode in shake flasks. Mean values and standard deviations of 3 experiments are shown.



Retention time in min	14.72	14.76	15.36	15.49	15.82	15.99	16.14	16.28	16.44	16.52	16.59	16.64	16.76	17.11	17.34	17.56	17.93	18.88	20.4	20.51*	21.06	21.23*
Nominal mass	204																			222		

**Figure 3.5 - Sesquiterpenoid spectrum of patchouli synthase produced by yeast.** A representative GC/FID analysis of the patchouli synthase products in the organic phase of a yeast two-phase cultivation heterologously expressing patchouli synthase is shown. 22 Peaks with nominal masses of 222 (sesquiterpenoids with hydroxy-group) and 204 (sesquiterpenoids without hydroxy-group) were detected. The peak marked with an asterisk (\*) was identified as patchouli.

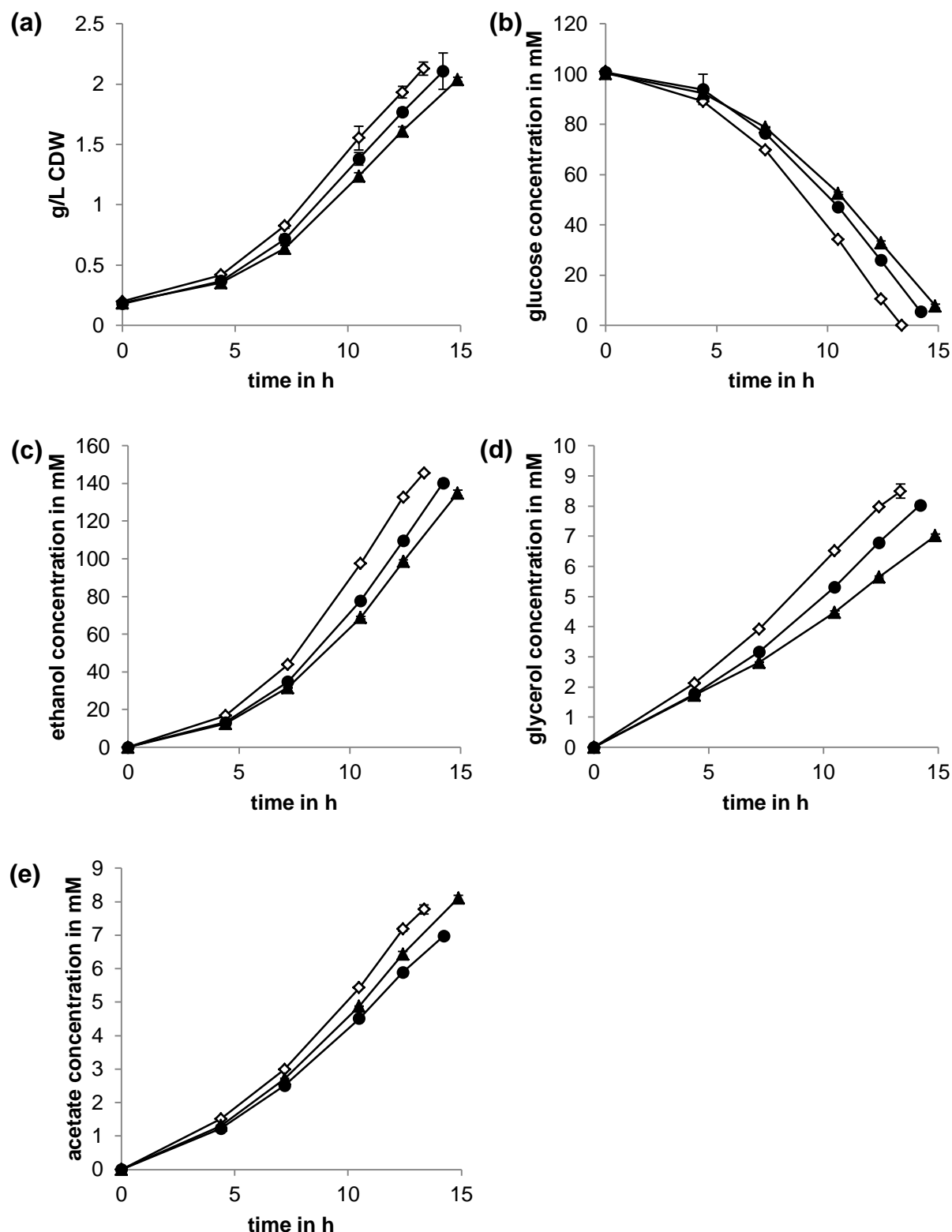
### 3.4.2 Engineering of terpenoid pathway: *tHMG1* overexpression and fusion of FPP synthase with patchouli synthase

Deregulating the MVA pathway is a prerequisite for metabolic engineering approaches to redirect the carbon flux within the central carbon metabolism towards terpenoids. The HMG-CoA reductase (*Hmg1/2p*) is described as the most important site of regulation in the MVA pathway. The overexpression of a truncated, and thus soluble, version (*tHMG1*) that is devoid of feedback inhibition by FPP (farnesyl diphosphate) has been shown to enhance terpenoid production several times (Donald, Hampton, and Fritz 1997; Paradise et al. 2008; Polakowski, Stahl, and Lang 1998; Ro et al. 2006; Verwaal et al. 2007). Thus, *tHMG1* was overexpressed using pSP-Pt carrying *tHMG1* under control of *PGK1* promoter next to *PTS* under control of *TEF1* promoter.

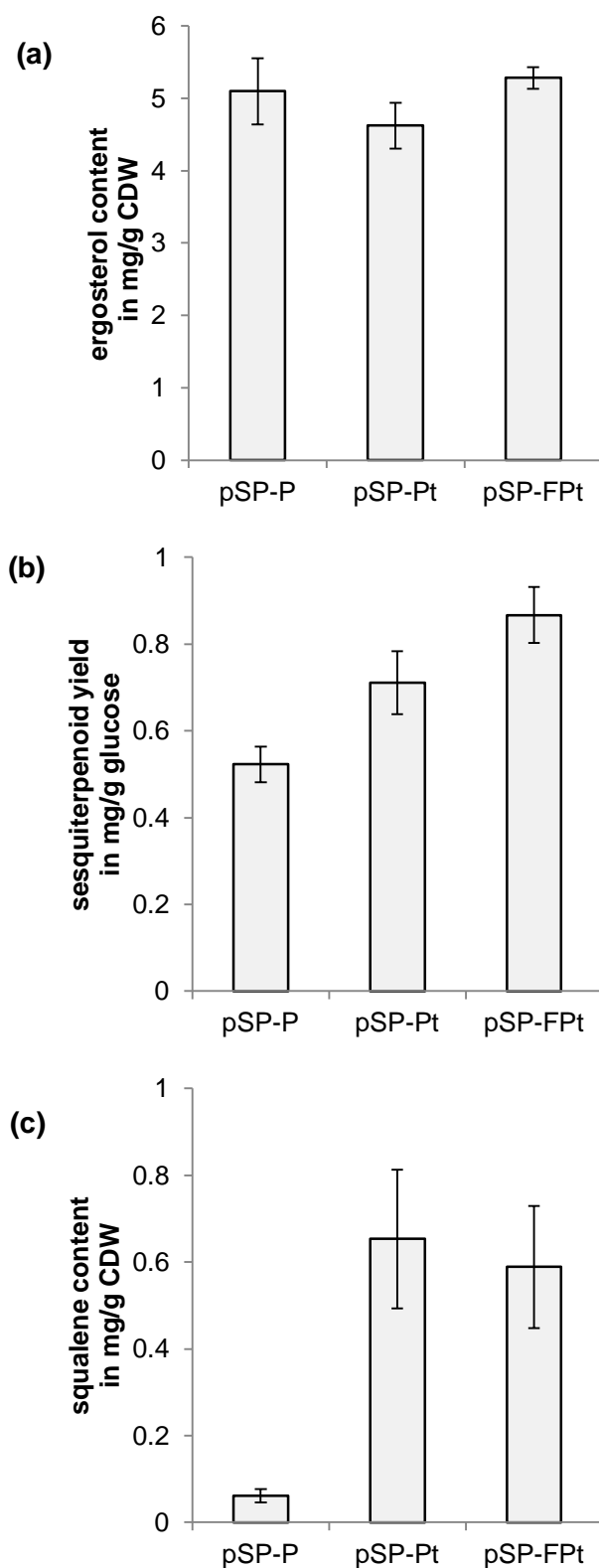
The generated yeast strain showed a slightly reduced growth rate accompanied by a slightly reduced glucose consumption rate as well as ethanol, glycerol and acetate formation rate (Figure 3.6). The ergosterol content was basically unchanged but the

sesquiterpenoid yield on glucose was indeed increased by 36 % (see Figure 3.7 a-b). Nevertheless, the strongest effect was visible for squalene. The squalene content was increased about 10-fold (see Figure 3.7 c).

Patchoulol synthase competes with squalene synthase (Erg9p) for their common substrate FPP. However, Erg9p has a lower  $K_m$  and a significantly higher  $k_{cat}$  than Pts ( $K_m$  of Erg9p for FPP is 2.5  $\mu\text{M}$  while  $K_m$  of Pts for FPP is 4.5 - 6.8  $\mu\text{M}$ ;  $k_{cat}$  of Erg9p is 0.53  $\text{s}^{-1}$  while  $k_{cat}$  of Pts is 0.00043 - 0.026  $\text{s}^{-1}$  (Deguerry et al. 2006; LoGrasso, Soltis, and Boettcher 1993; Munck and Croteau 1990)). The physical fusion of FPP synthase and Pts is thought to increase the flux towards sesquiterpenoids due to close proximity of the active sites of the enzymes which could bypass loss of the intermediate FPP by diffusion, degradation or conversion through competitive pathways via substrate channeling (Brodelius et al. 2002). Thus, a fusion construct was generated according to Albertsen et al. (Albertsen et al. 2011) and overexpressed under the control of *TEF1* promoter using pSP-FPt which carries additionally *tHMG1* under control of *PGK1* promoter. The generated strain exhibits a further decreased growth and glucose consumption rate as well as metabolite formation rate (Figure 3.6). The use of the fusion protein had basically no effect on squalene and ergosterol levels but it led to an increase in sesquiterpenoid yield on glucose by 22 % (see Figure 3.7).



**Figure 3.6 - Effect of *tHMG1* overexpression and fusion of FPP synthase and patchouliol synthase on physiological parameters of yeast.** Profiles of physiological parameters of yeast strains carrying pSP-P (◇), pSP-Pt (●) and pSP-FPt (▲) grown in batch mode in shake flasks: (a) biomass formation (cell dry weight, CDW), (b) glucose consumption, (c) ethanol formation, (d) glycerol formation and (e) acetate formation as a function of time. Mean values and standard deviations of 3 experiments are shown.

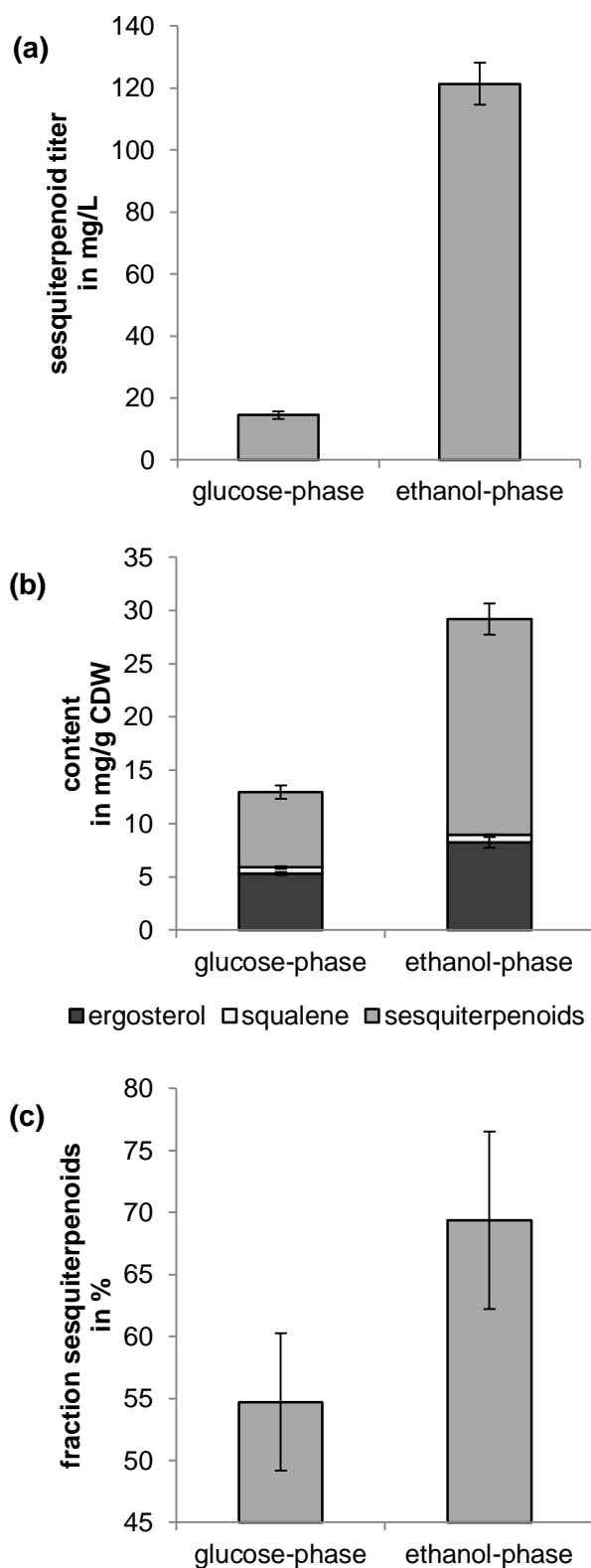


**Figure 3.7 - Effect of *tHMG1* overexpression and fusion of FPP synthase and patchouliol synthase on terpenoid production.** Terpenoid formation of yeast strains carrying pSP-P, pSP-Pt and pSP-FPt after growth on glucose in batch conditions: (a) ergosterol content, (b) sesquiterpenoid yield on glucose and (c) squalene content. Shown are mean values and standard deviations of six experiments.

### 3.4.3 Choice of carbon source: impact on terpenoid production

The focus of this study was on terpenoid production based on glucose. However, yeast is producing ethanol, glycerol and acetate while consuming and growing on glucose in batch mode (hereafter: glucose-phase). These non-fermentable carbon sources can then be consumed and used as carbon source by yeast after the diauxic shift in a second growth phase (hereafter: ethanol-phase). Thus, terpenoid production was not only analyzed during growth on glucose but as well during growth on those non-fermentable carbon sources. The yeast strain expressing the fusion protein of FPP synthase and patchoulol synthase was cultured on glucose in batch mode. As expected, cells grew first on glucose while producing ethanol, glycerol and acetate and then cells shifted to growth on the fermentation metabolites and the carbon sources were consumed completely.

Interestingly, sesquiterpenoid titer was increased tremendously by a factor of 8.4 from the glucose- to the ethanol-phase (Figure 3.8 a). To account for different biomass formation, the sesquiterpenoid yield was calculated per biomass (mg sesquiterpenoid/g CDW, cell dry weight). The yield was increased by a factor 2.9 in the ethanol-phase (Figure 3.8 b). The ergosterol and squalene contents were also increased during ethanol-phase by a factor 1.6 and 1.2 respectively. Furthermore, the fraction of sesquiterpenoids in total terpenoids (sesquiterpenoids plus squalene plus ergosterol) was increased from 55 % in glucose-phase to even 69 % in ethanol-phase (Figure 3.8 c).



**Figure 3.8 - Effect of carbon source on terpenoid production.** Terpenoid production of yeast strain carrying pSP-FPt during glucose-phase (cells were harvested as soon as glucose was exhausted) and during ethanol-phase (cells were harvested after four days when ethanol, glycerol and acetate were exhausted). (a) sesquiterpenoid titer; (b) ergosterol, squalene and sesquiterpenoid content per cell dry weight, CDW; (c) fraction of sesquiterpenoids in total terpenoids (ergosterol, squalene and sesquiterpenoids). Shown are mean values and standard deviations of three experiments.

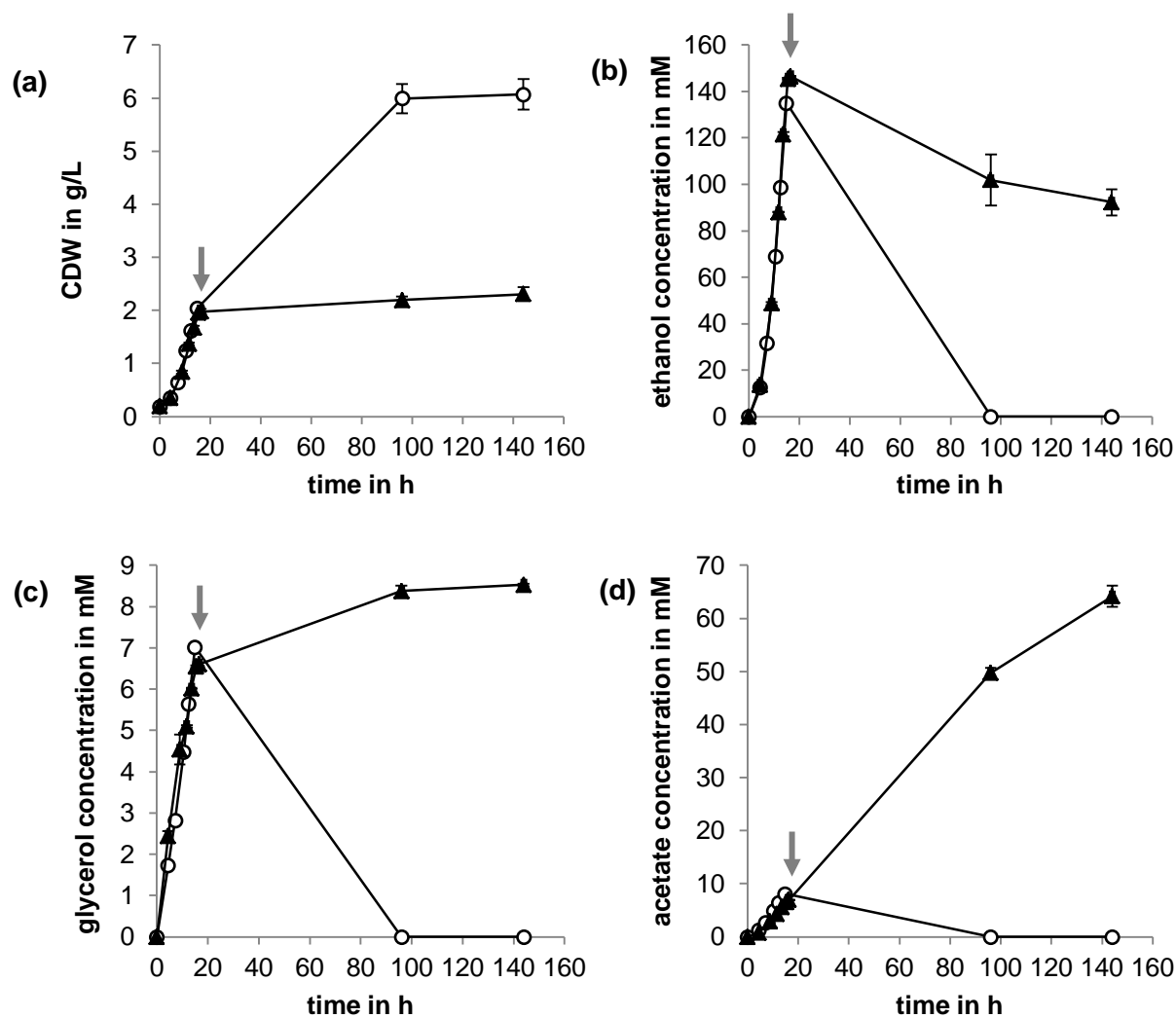
### 3.4.4 Metabolic engineering to redirect carbon flux towards terpenoids.

#### Disruption of *KGD1* of citric acid cycle

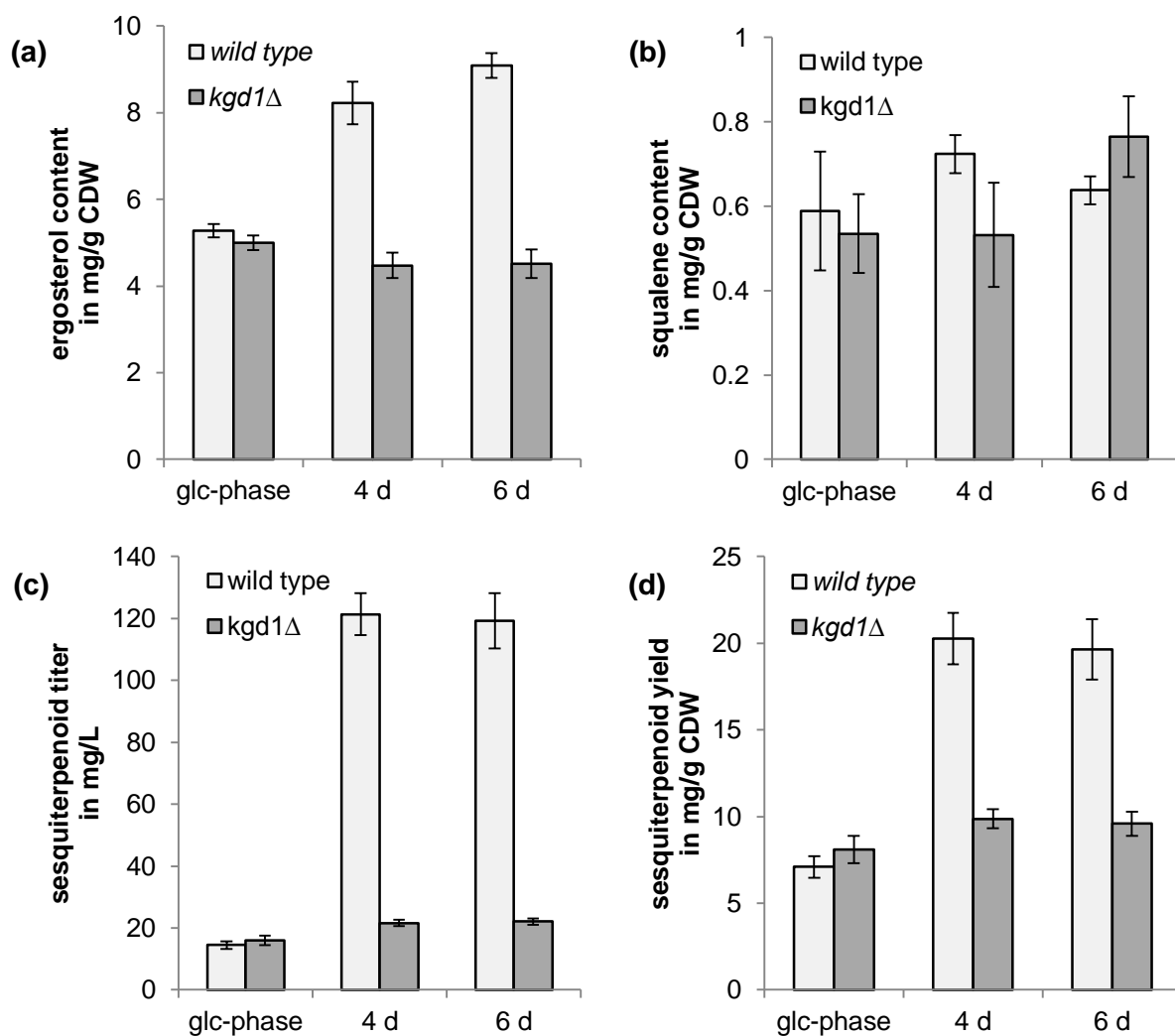
Based on our *in silico* computation (Gruchattka et al. 2013), a gene knockout within the citric acid cycle in combination with avoidance of ethanol and acetate production is a promising metabolic engineering target. The  $\alpha$ -ketoglutarate dehydrogenase gene (*KGD1*) of  $\alpha$ -ketoglutarate dehydrogenase complex was chosen as knockout target. The gene was disrupted in *S. cerevisiae* using the Cre-LoxP system via integration of a resistance cassette at the genomic locus and thus replacement of the first 1,702 bp of the gene. The resistance cassette was cut out using Cre recombinase leaving a single LoxP site at the genomic locus. The strain carrying the *KGD1* disruption was transformed with pSP-FPt for characterization in the two-phase cultivation.

Disruption of *KGD1* did not have any significant effect on physiological parameters (growth rates, or ethanol, glycerol and acetate formation) (Figure 3.9), ergosterol and squalene content or sesquiterpenoid production during growth on glucose in batch mode until glucose was consumed (see Figure 3.10). However, cell growth was interrupted after glucose was consumed while the wild type strain grew further on ethanol, glycerol and acetate as carbon sources. After 4 days, the wild type strain had consumed ethanol, glycerol and acetate completely. The strain carrying the *KGD1* disruption, however, consumed ethanol only partially while producing glycerol and especially high amounts of acetate (Figure 3.9 d). This trend continued after 6 days, where some ethanol was consumed, while acetate was produced in high amounts. Acetate concentrations were increased nearly 10-fold from glucose-phase to ethanol-phase after 6 days with this mutant. Squalene content was not significantly affected while ergosterol content was decreased in the strain carrying the *KGD1* disruption in the ethanol-phase (Figure 3.10 a-b). Surprisingly, sesquiterpenoid titer or yield per CDW was reduced in this strain (Figure 3.10 d-e).





**Figure 3.9 - Effect of disruption of *KGD1* on physiological parameters of yeast.** Influence of *KGD1* disruption (▲, *kgd1Δ*) in comparison to the wild type (O) on: (a) biomass formation (cell dry weight, CDW), (b) ethanol formation, (c) glycerol formation and (d) acetate formation as a function of time during glucose-phase as well as ethanol-phase, arrows indicate the point in time when glucose was exhausted. Shown are mean values and standard deviations of three experiments.

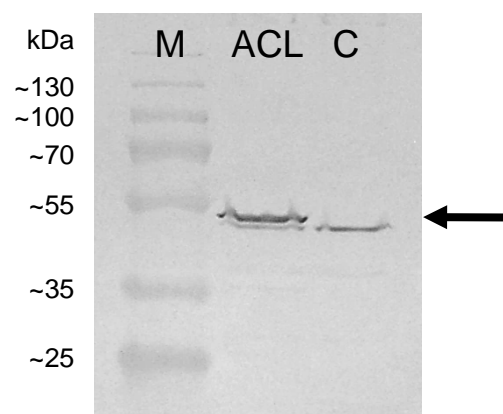


**Figure 3.10 - Effect of disruption of *KGD1* on terpenoid production.** Influence of *KGD1* disruption (*kgd1Δ*) in comparison to the wild type on: (a) ergosterol content, (b) squalene content, (c) sesquiterpenoid titer and (d) sesquiterpenoid yield on biomass during glucose-phase (glc-phase) and ethanol-phase at four and six days (4 d, 6 d). Shown are mean values and standard deviations of three experiments.

### 3.4.5 Heterologous pathway to synthesize cytosolic acetyl-CoA: ATP-citrate lyase

The pyruvate dehydrogenase bypass, especially acetate activation, appears to be a central issue *in vivo* in supplying significant acetyl-CoA for terpenoid production. The enzyme ATP-citrate lyase was identified as promising metabolic engineering target for increased terpenoid production due to increased energy-efficiency in comparison to the pyruvate dehydrogenase bypass (Gruchattka et al. 2013). Moreover, the heterologous expression of this enzyme could circumvent acetate accumulation.

Only few data on kinetic properties of ATP-citrate lyase from different organisms are available. Even though no data are available for ATP-citrate lyase from *Arabidopsis*, this enzyme has been shown to be active in *S. cerevisiae* (Fatland et al. 2002), and was therefore chosen for expression in this study. The enzyme is made up of two different polypeptide chains and consequently two genes; the cDNAs of *ACLA-1* and *ACLB-2* were used within this study. To reduce cloning time and effort and avoid excessive promoter and terminator use, the cDNAs were separated by a T2A sequence from *Thosea asigna* virus for polycistronic expression and introduced to pSP-FPt generating pSP-FPt-ACL. The *in vivo* functionality of the T2A sequence to generate two separate proteins was assessed via western blotting using 2A-peptide specific antibodies. Figure 3.11 shows a western blot of cell lysates of *S. cerevisiae* carrying pSP-FPt-ACL and pSP-FPt as a control.

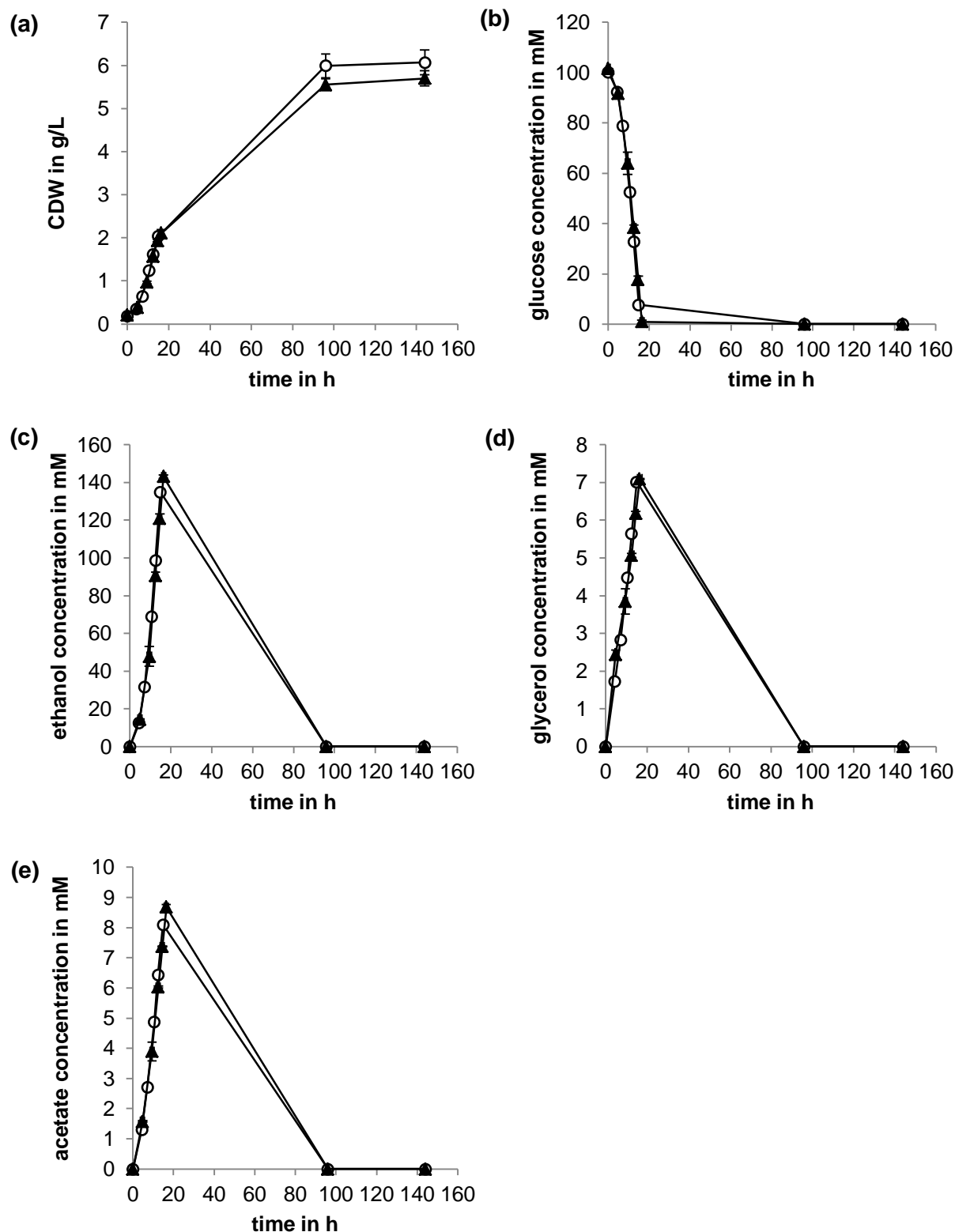


**Figure 3.11 - Western blot analysis of yeast proteins.** A representative western blot of protein from yeast strain carrying pSP-FPt-ACL (ACL) and pSP-FPt as control (C) is shown. The protein ladder is shown on the left (M). The predicted position of AclA-1-T2A is indicated with an arrow.

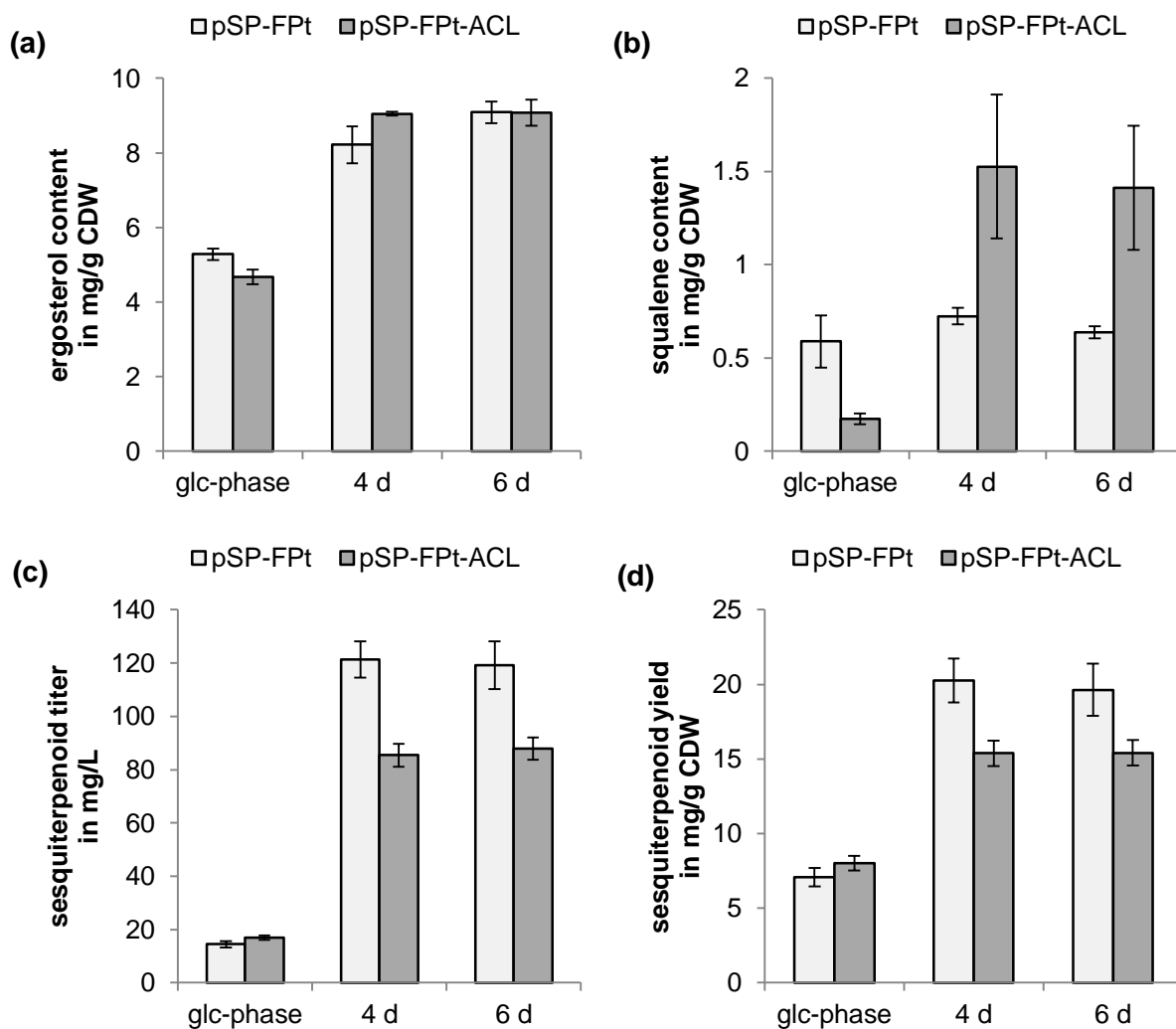
The blot should reveal a band corresponding to a molecular weight of 51 kDa if the polypeptide was cleaved correctly to AclA-1-T2A and AclB-2. A band corresponding to a

slightly lower molecular weight was visible in protein from both the strain expressing the polycistronic ATP-citrate lyase construct and the control strain not expressing the enzyme. Yet western blotting indeed revealed an additional band at the correct height only for the strain expressing the polycistronic ATP-citrate lyase construct.

The *S. cerevisiae* strain expressing the polycistronic ATP-citrate lyase construct which generates these two separate polypeptides was characterized in the two-phase cultivation conditions. ATP-citrate lyase overexpression did not have any effect on physiological parameters (growth, glucose consumption, ethanol, glycerol, and acetate formation) at the cultivation conditions investigated: batch conditions in glucose- and ethanol-phase (Figure 3.12). The effect on terpenoid production was ambivalent. Only a minor effect could be detected on ergosterol content (Figure 3.13 a), while squalene content was decreased in glucose-phase but increased in ethanol-phase (Figure 3.13 b). Sesquiterpenoid production was basically not affected in glucose-phase, however, was even reduced during ethanol-phase (Figure 3.13 c-d).



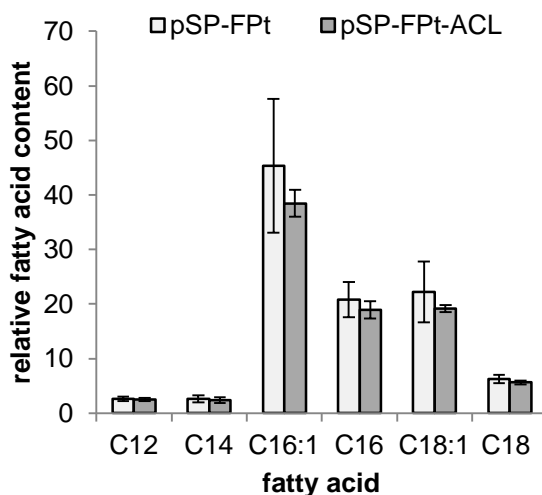
**Figure 3.12** - Effect of ATP-citrate lyase overexpression on physiological parameters of yeast. Profiles of physiological parameters of yeast strains carrying pSP-FPt (O) and pSP-FPt-ACL (▲) in shake flask experiments: (a) biomass formation (cell dry weight, CDW) as a function of time, (b) glucose consumption, (c) ethanol formation, (d) glycerol formation and (e) acetate formation as a function of time. Mean values and standard deviations of 3 experiments are shown.



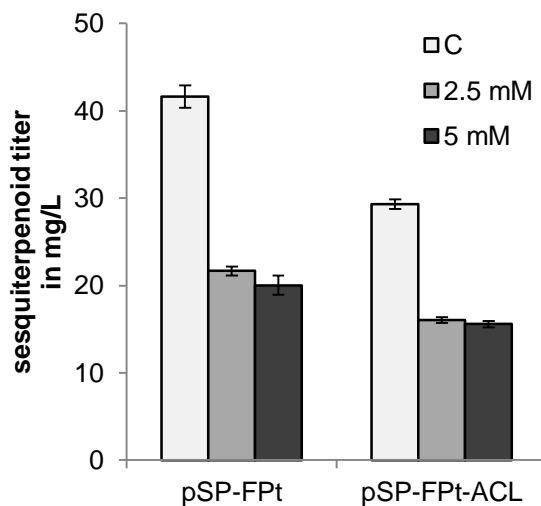
**Figure 3.13 - Effect of ATP-citrate lyase (ACL) overexpression on terpenoid production.** Terpenoid formation, *i.e.* (a) ergosterol content, (b) squalene content, (c) sesquiterpenoid titer and (d) sesquiterpenoid yield on biomass of yeast strains carrying pSP-FPt and pSP-FPt-ACL during glucose-phase (glc-phase) and ethanol-phase at four and six days (4 d, 6 d). Shown are mean values and standard deviations of three experiments.

To evaluate whether the reduced sesquiterpenoid production during ethanol-phase could be explained by a redirection of the carbon flux from acetyl-CoA towards fatty acids instead of terpenoids, fatty acid content of cells was determined. However, no increase in fatty acid content of the yeast strain expressing the polycistronic ATP-citrate lyase construct could be detected (Figure 3.14). To analyze whether activity and affinity of ATP-citrate lyase to its substrate are the limiting factors, citrate supplementation was tested. Supplementation of citrate to the medium (2.5 and 5 mM) to cells carrying pSP-FPt and pSP-FPt-ACL, caused a strong decrease in sesquiterpenoid production in

glucose-phase (hardly detectable amounts) as well as in ethanol-phase (reduction to about half) while physiological parameters were not affected (Figure 3.15).



**Figure 3.14 - Effect of ATP-citrate lyase overexpression on relative fatty acid content.** Shown is the relative fatty acid content of C12 to C18 of cells after 4 days of cultivation of yeast strains carrying pSP-FPt-ACL and pSP-FPt as control. Mean values and standard deviations of 3 experiments are shown.



**Figure 3.15 - Effect of citrate supplementation on sesquiterpenoid production.** Sesquiterpenoid titer (in mg/L) after 4 days is shown for yeast strains carrying pSP-FPt and pSP-FPt-ACL in shake flask experiments without (C) and with 2.5 and 5 mM citrate added to the culture medium. Mean values and standard deviations of 3 experiments are shown.

### 3.5 Discussion

Enhancing the intracellular capacity for heterologous production of acetyl-CoA derived products is not limited to terpenoids; metabolic engineering strategies could as well be applied to other acetyl-CoA derived products such as lipids, polyketides, polyhydroxyalkanoates or *n*-butanol. Currently, there is still room for improvement via metabolic engineering to boost product yield for all classes of these products. Recent studies have focused on modifications within the central carbon metabolism of *S. cerevisiae* to redirect the metabolic flux towards an increased cytosolic acetyl-CoA production, the precursor of the MVA pathway. Strategies employed in these studies include: engineering of the pyruvate dehydrogenase bypass (*ALD6* and a mutant *acs* from *Salmonella*) (Shiba et al. 2007); modification of ammonium assimilation (*gdh1Δ* and *GDH2*) (Asadollahi et al. 2009; Scalcinati, Partow, et al. 2012); combining engineering of the pyruvate dehydrogenase bypass (*ALD6*, a mutant *acs* from *Salmonella* and *ADH2*) with deletions in the glyoxylate cycle (*cit2Δ* and *mls1Δ*) for terpenoid, biobutanol, polyhydroxybutyrate, 3-hydroxypropionic acid and fatty acid production (Chen et al. 2013; Chen et al. 2014; de Jong et al. 2014; Kocharin et al. 2012; Krivoruchko et al. 2013); expression of the phosphoketolase pathway for fatty acid and polyhydroxybutyrate production (de Jong et al. 2014; Kocharin, Siewers, and Nielsen 2013; Papini, Nookaew, Siewers, et al. 2012); functional expression of a bacterial pyruvate dehydrogenase complex in the cytosol (Kozak et al. 2014); functional expression of the acetylating acetaldehyde dehydrogenase and pyruvate-formate lyase (Kozak et al. 2013); and a model for an indirect acetyl-CoA transport from mitochondria to the cytosol, including acetate:succinyl-CoA transferase (*Ach1*) was proposed (Chen et al. 2015). Several strategies to improve terpenoid production were also predicted *in silico* in our previous study (Gruchattka et al. 2013). Here, we report the *in vivo* validation of selected strategies and identify general considerations important for the *in vivo* validation of such stoichiometric metabolic network predictions.

#### 3.5.1 Sesquiterpenoid spectrum

A two-phase cultivation procedure was established for the volatile sesquiterpenoid patchoulol. 21 additional sesquiterpenoids in addition to patchoulol could be detected as product of the patchoulol synthase. Here, patchoulol represents about 34.3 % of total sesquiterpenoids produced by this enzyme, slightly lower than the previously reported value of 36.9 % (Deguerry et al. 2006). This difference is likely because we could detect



more sesquiterpenoids due to a higher production titer. The value was used to estimate the production of total sesquiterpenoids using patchoulol as a marker. As prerequisites for validating metabolic engineering strategies within the central carbon metabolism, the flux from acetyl-CoA towards the desired product was increased and carbon sources were analyzed for their capacity to support sesquiterpenoid production.

### **3.5.2 *tHMG1* overexpression and fusion of FPP synthase with patchoulol synthase have a small effect on sesquiterpenoid production**

Two well described engineering strategies were employed to increase the flux from the central carbon precursor acetyl-CoA to the terpenoid product. A truncated HMG-CoA reductase (*tHMG1*) that is devoid of feedback inhibition by FPP was overexpressed to deregulate the MVA pathway. In addition, endogenous FPP synthase and patchoulol synthase were fused to increase the flux towards the desired product and to bypass loss of the intermediate FPP via substrate channeling. Surprisingly, overexpression of *tHMG1* led only to an increase in sesquiterpenoid yield by 36 % and protein fusion led only to a further increase by 22 %. However, squalene levels were increased approximately 10-fold when *tHMG1* was overexpressed. HMG-CoA reductase was identified as the rate-limiting enzyme in early sterol biosynthesis in yeast in the late 1990s when its overexpression was shown to increase squalene levels by 10-fold and 40-fold (Donald, Hampton, and Fritz 1997; Polakowski, Stahl, and Lang 1998). However, these apparently strong improvements in squalene correspond to very low absolute amounts when compared to ergosterol and sesquiterpenoid amounts (Figure 3.8 b). For most mono-, di-, and sesquiterpenoids the effect was comparably low with an increase of 1.4 to 5-fold (Asadollahi et al. 2009; Asadollahi et al. 2010; Engels, Dahm, and Jennewein 2008; Farhi et al. 2011; Paradise et al. 2008; Jackson, Hart-Wells, and Matsuda 2003; Rico, Pardo, and Orejas 2010; Ro et al. 2006) and in some cases, no effect or even an adverse effect was detected (Asadollahi et al. 2010; Albertsen et al. 2011; Jackson, Hart-Wells, and Matsuda 2003) depending on terpenoid synthase and previous genetic modifications. Several fusion proteins consisting of terpenoid synthase and FPP or GPP (geranyl diphosphate) synthase were shown to increase terpenoid production (Brodelius et al. 2002; Ignea et al. 2014; Sarria et al. 2014). Albertsen and co-workers (Albertsen et al. 2011) showed that a fusion protein consisting of patchoulol synthase and endogenous FPP synthase can increase patchoulol production up to 2-fold, again depending on

additional genetic modifications. Thus, our findings that *tHMG1* overexpression and fusion of FPP synthase with patchoulol synthase have small positive effects on sesquiterpenoid production are in good agreement with literature. Therefore, further production improvements observed from these modifications likely depend strongly on additional genetic modifications of the yeast strain used or even on cultivation conditions.

### 3.5.3 Carbon source strongly influences terpenoid production

Yeast shows two separate growth phases in batch mode: a phase of growth based on glucose as carbon source accompanied by the production of ethanol, glycerol and acetate (glucose-phase), and a second growth phase based on ethanol, glycerol and acetate as carbon source (ethanol-phase). Sesquiterpenoid as well as total terpenoid (sesquiterpenoids plus squalene plus ergosterol) titer and yield on biomass were increased tremendously during growth ethanol-phase in comparison to glucose-phase (Figure 3.8 a-b). While Chen and co-workers (Chen et al. 2013) showed that at least 50 % of sesquiterpenoids were produced during the glucose-phase in comparison to the following ethanol-phase, 88 % of sesquiterpenoids formed after four days were produced during the ethanol-phase in our experiments (see Figure 3.8 a). A patchoulol titer of 41.6 +/- 1.3 mg/L was achieved, which is to our knowledge the highest final titer of patchoulol in shake flasks and which is as high as the highest reported patchoulol titer in a bioreactor (40.9 +/- 2.7 mg/L) (Albertsen et al. 2011). The higher potential of ethanol as a carbon source to increase terpenoid yield was indeed predicted in our *in silico* analysis (Gruchattka et al. 2013) and demonstrated previously with ethanol feeding (Westfall et al. 2012). The results presented here indicate that the use of glucose as a cheap carbon source together with ethanol, which was produced by the yeast itself, is a promising approach to produce terpenoids in high yields.

Moreover, we could show that the fraction of sesquiterpenoids in total terpenoids was increased from 55 % in glucose-phase to 69 % in ethanol-phase after four days (Figure 3.8 c). This increase indicates that not only the flux into the mevalonate pathway was increased during growth on ethanol, glycerol, and acetate but the partition of the flux at the intersection of FPP was affected positively, which has to our knowledge not been shown before.

### 3.5.4 *In vivo* validation and considerations of *in silico* predicted strategies

The disruption of  $\alpha$ -ketoglutarate dehydrogenase gene (*KGD1*) of  $\alpha$ -ketoglutarate dehydrogenase complex did not lead to the desired increase in terpenoid production. The citric acid cycle exhibits only a weak activity during growth on high glucose concentrations (Papini, Nookaew, Uhlen, et al. 2012) and the flux is redirected mainly towards ethanol (Crabtree effect) due to a complex regulatory machinery involving repression of *LPD1* of pyruvate dehydrogenase complex (Bowman et al. 1992), aldehyde dehydrogenases *ALD2-4* (Navarro-Avino et al. 1999), acetyl-CoA synthetase *ACS1* (van den Berg et al. 1996) and alcohol dehydrogenase *ADH2* which converts ethanol to acetaldehyde (Ciriacy 1975) at high glucose concentrations while pyruvate decarboxylase *PDC1* is presumed to be activated by glucose (or repressed by ethanol) (Liesen, Hollenberg, and Heinisch 1996; Pronk, Yde Steensma, and Van Dijken 1996). Hence the disruption of *KGD1* resulted in no effect. Our previous *in silico* prediction did not take this regulation into account (Gruchattka et al. 2013). However, during ethanol-phase, the use of a different carbon source circumvents this regulation and leads to a higher flux into the citric acid cycle. Here, disruption of  $\alpha$ -ketoglutarate dehydrogenase gene under these conditions did lead to an effect (see Figure 3.9). Cell growth was inhibited, high amounts of acetate were secreted, and terpenoid formation was decreased. In addition, glucose-limited conditions were analyzed to determine the effects of this modification during conditions, which normally increase the flux into in the citric acid cycle (Papini, Nookaew, Uhlen, et al. 2012). Similarly to ethanol-phase cultures, the strain carrying the *KGD1* disruption exhibited a reduced cell growth and production of acetate during glucose limitation, while the wild type strain did not produce any fermentation products. However, terpenoid formation was not affected under these conditions (Figure 3.17, supplementary information).

Disruption of *KGD1* was predicted to redirect the metabolic flux from citric acid cycle via pyruvate dehydrogenase bypass towards terpenoids. Although the metabolic flux was indeed redirected to the bypass, the redirection was interrupted at the level of acetate (see Figure 3.1). Activation of acetate and acetyl-CoA formation seems to be the rate-limiting step in this process due to regulation and activity of acetyl-CoA synthetase (*ACS1*, *ACS2*). Shiba *et al.* (Shiba et al. 2007) reported the same issue, and succeeded in reducing acetate formation while at the same time increasing terpenoid formation by overexpressing a non-post-translationally regulated acetyl-CoA synthetase from *Salmonella* together with a native aldehyde dehydrogenase. However, acetate formation

could not be completely abolished with this strategy and acetate levels were more than three times higher in our experiments. Lian *et al.* (Lian *et al.* 2014) also reported similar issues of increased acetate formation to cytotoxic levels, which was accompanied by decreased product formation (*n*-butanol).

The disruption of  $\alpha$ -ketoglutarate dehydrogenase led to no effect during batch conditions and to three consistent effects during glucose-limited conditions or ethanol-phase: (i) reduced growth during glucose-limited conditions or abolished growth during ethanol-phase, (ii) acetate formation, which was especially high in the ethanol-phase, and (iii) reduced terpenoid formation in ethanol-phase only. Strains carrying the gene disruption should theoretically be able to grow on glucose as well as on ethanol according to our *in silico* computations. However, a reduced growth of citric acid cycle knockouts on non-fermentable carbon sources has been described in literature (Przybyla-Zawislak *et al.* 1999). An explanation for the effect has not been given in previous literature. Acetate formation could be the reason for the growth inhibition as acetate acts as a weak acid and dissipates the pH gradient over the plasma membrane, which in turn has to be balanced by ATP-consuming proton pumps (Verduyn 1991). This additional ATP requirement could as well be a reason for the reduced terpenoid production during the ethanol-phase, as terpenoid synthesis requires high amounts of ATP. Moreover, acetic acid has been described to trigger apoptosis (programmed cell death) in yeast even at pH levels above neutral. Apoptosis is likely triggered via reactive oxygen species, especially increased  $\text{H}_2\text{O}_2$  levels, modulated by active superoxide ( $\text{O}_2^-$ ) dismutase (which converts  $2 \text{O}_2^- + 2 \text{H}^+$  into  $\text{H}_2\text{O}_2$  and  $\text{O}_2$ ) and via acetic acid inactivated catalase (which normally converts  $2 \text{H}_2\text{O}_2$  into  $\text{H}_2\text{O}$  and  $\text{O}_2$ ). Additionally, increased proteolytic activity has been shown upon acetic acid stress (Guaragnella *et al.* 2011; Knorre, Smirnova, and Severin 2005). Acetate formation could also be the reason for the negative effect observed on terpenoid formation due to the role of acetate as a signal, which triggers multiple effects. Another weak organic acid is citrate. Citrate supplementation in our experiments led also to a strong reduction in terpenoid formation. Thus, there could be a common mechanism. Nevertheless, a yet unknown mechanism could add to the effects seen. Acetate was produced during glucose-limited conditions, however, concentrations were comparable to those in batch conditions. Thus, one can assume that an additional mechanism is responsible for the inhibition of growth. Further studies need to address the issue of reduced growth, reduced terpenoid formation and the effects of weak organic acids such as acetate to elucidate this issue.

It can be proposed from these findings that acetate formation should be avoided in order to increase terpenoid production. An alternative solution could be the use of heterologous pathways, which circumvent the pyruvate dehydrogenase bypass. ATP-citrate lyase converts citrate that has been produced via pyruvate dehydrogenase complex and the citric acid cycle in the mitochondria to acetyl-CoA. The reaction is more energy-efficient and avoids the pyruvate dehydrogenase bypass. ATP-citrate lyase from *Arabidopsis* was overexpressed as a polycistronic construct in this study. A self-cleaving T2A sequence from *Thosea asigna* virus was used for the generation of two separate polypeptides from a polycistronic mRNA. This represents only the second time such a strategy has been used in yeast biotechnology (Beekwilder et al. 2014), and validates the efficiency of self-cleaving peptides to reduce cloning time and effort for expression of multiple transgenes from a single promoter.

Overexpression of ATP-citrate lyase from *Arabidopsis* did not lead to the desired circumvention of acetate formation or an increase in terpenoid production during batch cultivation conditions in either glucose- or ethanol-phase or even glucose-limited conditions (Figure 3.18, supplementary information). In the ethanol-phase, squalene levels were indeed increased, however, sesquiterpenoid formation was decreased. Unfortunately, those decreased levels could not be explained by modified fatty acid formation. A negative influence of the enzyme tag from the viral sequence on activity seems improbable as ATP-citrate lyase genes from animals arose from gene fusion of the two genes (as found in *Arabidopsis sp.*) and thus possess a natural linker of the size of the tag used between the two polypeptide chains homologous to the plant ones (Fatland et al. 2002).

ATP-citrate lyase has been overexpressed in yeast before to increase *n*-butanol or fatty acid production. The enzyme from mouse led to an increase in fatty acid formation of 15-20 % (Tang, Feng, and Chen 2013). The enzyme from *Arabidopsis* did not have an effect while the enzyme from *Yarrowia lipolytica*, an oleaginous yeast, led to a 2-fold increase in *n*-butanol production (Lian et al. 2014). While no kinetic properties of ATP-citrate lyase from those organisms are available, we assume that the kinetic properties are unfavorable and *in vivo* activities are not sufficient of the enzyme from *Arabidopsis* or mouse to influence formation of acetyl-CoA derived products tremendously. Previously, Lian and coworkers (Lian et al. 2014) assumed that the *in vivo* activity and affinity of ATP-citrate lyase of *Arabidopsis* is not sufficient. Citrate supplementation in their experiments resulted in an increase in *n*-butanol production, an acetyl-CoA derived

product, in the case of ATP-citrate lyase of *Arabidopsis*, which indicates that activity and affinity are issues for this enzyme in the yeast host. ATP-citrate lyase from an oleaginous yeast or carotenoid producing yeast like *Yarrowia lipolytica* or *Xanthophyllomyces dendrorhous* may be more promising as those organisms use ATP-citrate lyase as main acetyl-CoA source for the production of acetyl-CoA derived carotenoids and lipids (Chavez-Cabrera et al. 2010; Koch, Schmidt, and Daum 2014). Indeed, other studies confirm that *in vivo* performance of heterologous enzymes needs to be optimized. Kozak et al. (Kozak et al. 2014) functionally expressed a bacterial pyruvate dehydrogenase complex in *S. cerevisiae*. However, the addition of lipoate to growth media was required for activity and an increase in product formation of acetyl-CoA derived substances due to the bacterial complex has not yet been shown. In a further study, Kozak et al. (Kozak et al. 2013) demonstrated that acetylating acetaldehyde dehydrogenase and pyruvate-formate lyase can replace acetyl-CoA synthetase of yeast but indicated that further research is essential for significant increases in the cytosolic acetyl-CoA supply. These studies confirm the necessity of further optimization of alternative reactions or pathways for a sufficient *in vivo* performance leading eventually to an increased product formation.

### 3.5.5 Conclusions

Our previous *in silico* analysis (Gruchattka et al. 2013) identified promising metabolic engineering targets, however, regulation and kinetics were not taken into account. This study demonstrates that regulation can interfere, such as observed with *KGD1* disruption, which had no effect in glucose-phase but in ethanol-phase as well as in glucose-limited conditions. Furthermore, *in vivo* activity of central carbon metabolism enzymes can interfere, such as demonstrated with ATP-citrate lyase or acetyl-CoA synthetase, resulting in poor *in vivo* performance. This necessity of further optimization was confirmed by other studies (Kozak et al. 2013; Kozak et al. 2014). Main sites of regulation need to be released and different sources for enzymes, rational protein design, or directed evolution might be essential to thoroughly fulfil desired activities of central carbon metabolism enzymes *in vivo* for an increased product formation from a heterologous host. These are issues to be addressed in the future.

Although the stoichiometric model was limited due to regulation and unpredicted enzyme activity, stoichiometric metabolic network analyses such as the one applied

previously (Gruchattka et al. 2013) can be used successfully as a metabolic flux prediction tool. Here, predictions were confirmed that ethanol is the more promising carbon source for sesquiterpenoid formation, and the metabolic flux in the strain carrying the  $\alpha$ -ketoglutarate dehydrogenase disruption was redirected as predicted. Unfortunately, the intermediate acetate was produced instead of terpenoids. However, if the produced acetate could be efficiently converted to sesquiterpenoids, a titer in the g/L range could be achieved. This highlights that the *in silico* predicted strategy is indeed valuable and should be pursued in the future. Nevertheless, the inclusion of regulatory circuits and *in vivo* kinetics of enzymatic reactions into *in silico* metabolic network analyses are required to increase the predictive power of these models and thus reduce the time of strain improvement in the future.

### 3.6 Supplementary information

Plasmid maps, gene sequences and primers used in this study.

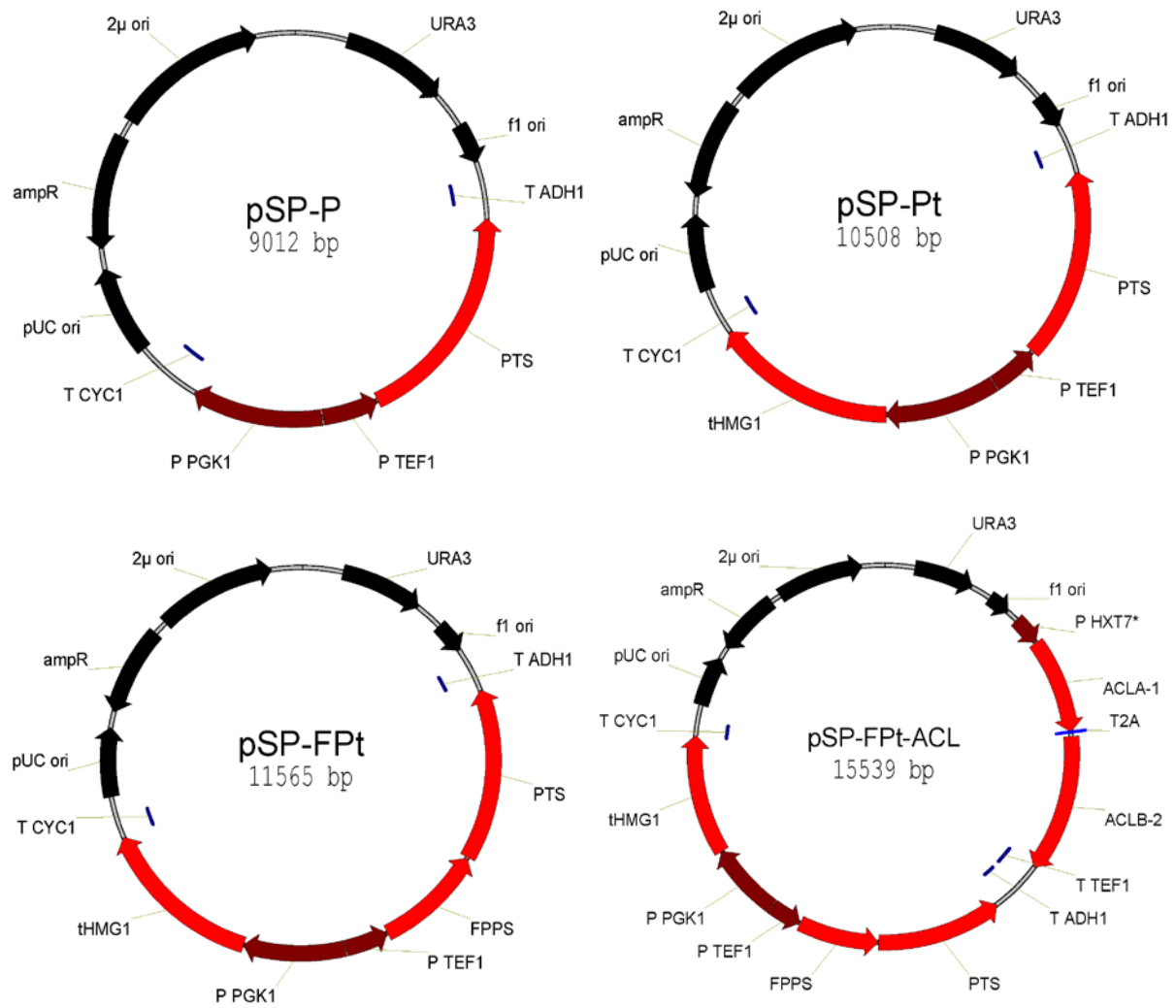


Figure 3.16 - Maps of expression plasmids used in this study.



**Table 3.2 - Gene sequences used in this study.** T2A sequence from *Thosea asigna* virus is highlighted in bold.

Name	Sequence
<i>PTS</i>	ATGGAATTATACGCCCAATCCGTTGGTGTGGTGCTGCTTCTAGACCATTGGCTAATTTT CATCCATGTGTTGGGGTGATAAGTTCATCGTTTACAACCCACAATCTTGTCAAGCTGGT GAAAGAGAAGAAGCTGAAGAATTGAAGGTCAATTGAAGAGAGAATTGAAAGAAGCCTCC GACAACTACATGAGACAATTGAAAATGGTTGACGCCATCCAAAGATTGGGTATCGATTAC TTGTTTCGTTGAAGATGTTGATGAAGCCTTGAAGAACTTGTTCGAAATGTTTCGATGCTTTC TGCAAGAACCAACCATGATATGCATGCTACTGCTTTGTCCTTCAGATTATTGAGACAACAC GGTTACAGAGTTTCTTTCGGAAGTCTTTGAAAAGTTCAAGGATGGTAAGGACGGTTTCAAG GTTCCAAATGAAGATGGTGTGCTGTTGCTGTTTGGAAATCTTTGAAGCTACCCATTTGAGA GTTCATGGTGAAGATGTATTGGATAACGCTTTCGATTTCCAGAAACTACTTGGAAATC GTTTACGCTACTTTGAACGATCCAACCTGTAAGCAAGTTCATAAGTTCGTTGAACGAAATC TCATTCAGAAGAGGTTTGGCCAAAGAGTTGAAGCCAGAAAGTACATTTCCATCTACGAACAA TATGCCCTCCATCATAAGGGTTTGTGAAAATTGGCTAAGTTGGACTTCAATTTGGTTCAA GCCTTGACAGAAGAGAATTGTCTGAAGATTCTAGATGGTGGAAAACCTTGCAAGTTCCA ACTAAGTTGTCTTTCGTTAGAGATAGATTGGTCAAGTCTTACTTTTGGGCTTCTGGTTCT TACTTCGAACCTAATTTATTCGTTGGCAGAATGATTTGGCTAAAGGTTTGGCTGTTTTG TCCTTGATGGATGATGTTTATGATGCCCTACGGTACTTTCGAAGAATTGCAAAATGTTCCAC GATGCCATTGAAAGATGGGATGCTTCTTGTGTTGGATAAGTTGCCAGATTACATGAAGATT GTCTACAAGGCTTGTGTTGGACGTATTGGAAGAAGTTGACGAAGAATTGATTAAGTTGGGT GCTCCATATAGAGCCTACTATGGTAAAGAAGCTATGAAGTACGCTGCTAGAGCTTATATG GAAGAAGCTCAATGGAGAGAACAAGCAAGCCAACCTACCAAGAATATATGAAGTTG GCTACTAAGACCTGCGGTTACATTACCTTGATTATCTTGTTCATGCTTGGGTGTCGAAGAA GGTATCGTTACAAAAGAAGCTTTTACTGGGTTTTCTCTAGACCACCTTTTATTGAAGCC ACCTTGATCATTGCTAGATTGGTTAACGATATCACCGGTACGAATTGAAAAAAGAGA GAACACGTTAGAACCCTGTTGAATGTTACATGGAAGAACATAAGGTCGTTAAGCAAGAA GTTGTCTCCGAATCTACAATCAAATGGAATCTGCTTGGAAAGGACATCAACGAAGGTTTT TTAAGACCAGTCGAATCCCAATCCCTTGTGTTGACTTGATCTTGAACCTGTGAGAAC TTGGAAGTCATCTACAAGAAGGTGATTCTTACACTCATGTTGGTCCAGCTATGCAAAAC ATTATCAAGCAATTATACTTGCACCCAGTTCATATTA
<i>tHMG</i>	ATGGACCAATTGGTGAAAACCTGAAGTCACCAAGAAGTCTTTTACTGCTCCTGTACAAAAG GCTTCTACACCAGTTTTAACCAATAAAACAGTCATTTCTGGATCGAAAGTCAAAAAGTTTA TCATCTGCGCAATCGAGCTCATCAGGACCTTCATCATCTAGTGAGGAAGATGATTTCCGCG GATATTGAAAGCTTGGATAAGAAAATACGTCTTTTGAAGAATTAGAAGCATTATTAAGT AGTGAAAATACAAAACAATTGAAGAACAAGAGGTGCTGCTTGGTTATTACCGGTAAG TTACCTTTGTACGCTTTGGAGAAAAAATTAGGTGATACTACGAGAGCGGTTGCGGTACGT AGGAAGGCTCTTTCAATTTTGGCAGAAGCTCCTGTATTAGCATCTGATCGTTTACCATAT AAAAAATTATGACTACGACCGGATTTTGGCGCTTGTGTTGAAAATGTTATAGGTTACATG CCTTTGCCGTTGGTGTATAGGCCCTTGGTTATCGATGGTACATCTTATCATATACCA ATGGCAACTACAGAGGGTGTGTTGGTAGCTTCTGCCATGCGTGGCTGTAAGGCAATCAAT GCTGGCGGTGGTGCAACAACCTGTTTTAACTAAGGATGGTATGACAAGAGGCCAGTAGTC CGTTTCCCAACTTTGAAAAGATCTGGTGCTGTAAGATATGGTTAGACTCAGAAGAGGGA CAAAACGAATTAATAAAGCTTTTAACTCTACATCAAGATTTGCACGCTGCAACATATT CAACTTGTCTAGCAGGAGATTTACTCTTCATGAGATTTAGAACAACCTACTGGTGACGCA ATGGGTATGAATATGATTTCTAAGGGTGTGCAATACTCATTAAGCAAAATGGTAGAAGAG TATGGCTGGGAAGATATGGAGGTTGTCTCCGTTTCTGGTAACTACTGTACCAGCAAAAAA CCAGCTGCCATCAACTGGATCGAAGGTCGTGGTAAGAGTGTGCTGCGAGAAGCTACTATT CCTGGTGATGTTGTGAGAAAAGTGTAAAAAGTGATGTTTCCGCATTGGTTGAGTTGAAC ATTGCTAAGAATTTGGTTGGATCTGCAATGGCTGGGTCTGTTGGTGGATTTAACGCACAT GCAGCTAATTTAGTGACAGCTGTTTTCTTGGCATTAGGACAAGATCCTGCACAAAATGTC GAAAGTTCCAACCTGTATAACATTGATGAAAAGAAGTGGACGGTGATTTGAGAAATTTCCGTA TCCATGCCATCCATCGAAGTAGTACCATCGGTGGTGGTACTGTTCTAGAACCACCAAGGT GCCATGTTGGACTTATTAGGTGTAAGAGGCCACATGCTACCGCTCCTGGTACCAACGCA CGTCAATTAGCAAGAATAGTTGCCTGTGCCGCTTGGCAGGTGAATTTATCTTATGTGCT GCCCTAGCAGCCGGCATTTGGTTCAAAGTCATATGACCACAACAGGAAACCTGCTGAA CCAACAAAACCTAACAATTTGGACGCCACTGATATAAAATCGTTTGAAGATGGGTCCGTC ACCTGCATTAATCCTAA

**Table 3.2** continued

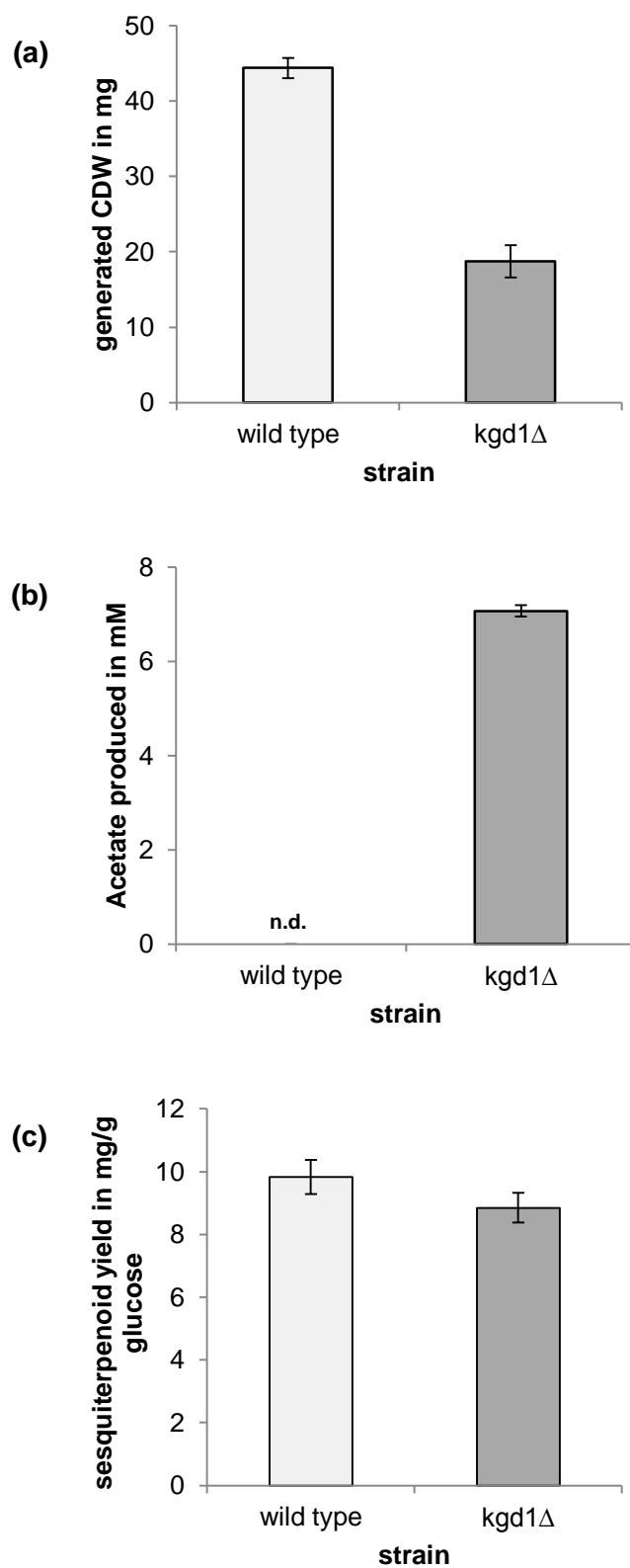
Name	Sequence
FPP synthase: <i>FPPS</i>	ATGGCTTCAGAAAAAGAAATTAGGAGAGAGAGATTCTTGAACGTTTTCCCTAAATTAGTA GAGGAATTGAACGCATCGCTTTTGGCTTACGGTATGCCCTAAGGAAGCATGTGACTGGTAT GCCCACTCATTGAACTACAACACTCCAGGCGGTAAGTTAAATAGAGGTTTGTCCGTTGTG GACACGTATGCTATTCTCTCCAACAAGACCGTTGAACAATTGGGGCAAGAAGAATACGAA AAGTTGCTATTCTAGGTTGGTGCATTGAGTTGTTGCAGGCTTACTTCTTGGTCGCCGAT GATATGATGGACAAGTCCATTACCAGAAGAGGCCAACCATGTTGGTACAAGGTTCTTGAA GTTGGGGAAATTGCCATCAATGACGCATTCATGTTAGAGGCTGCTATCTACAAGCTTTTG AAATCTCACTTCAGAAACGAAAAATACTACATAGATATCACCGAATTGTTCCATGAAGTC ACCTTCAAACC GAATTGGGCCAATTGATGGACTTAATCACTGCACCTGAAGACAAAGTC GACTTGAGTAAGTTCTCCCTAAAGAAGCACTCCTTCATAGTTACTTTCAAGACTGCTTAC TATTCTTTCTACTTGCCTGTGCGATTGGCTATGTACGTTGCCGGTATCACAGATGAAAAG GATTTGAAACAAGCCAGAGATGTCTTGATTCCATTGGGTGAATATTTCCAAATTC AAGAT GACTACTTAGACTGCTTCGGTACCCAGAACAGATCGGTAAGATCGGTACAGATATCCAA GATAACAAATGTTCTTGGGTAATCAACAAGGCATTAGA ACTTGCTTCCGCAGAACAAAGA AAGACTTAGACGAAAATTACGGTAAGAAGGACTCAGTCGCAGAAGCCAAATGCAAAAAG ATTTTCAATGACTTGAAAATCGACCAGTTATACCACGAATATGAAGAGTCTGTTGCCAAG GATTTGAAGGCCAAGATCTCCAAGTCGACGAGTCTCGTGGCTTCAAAGCCGACGCTTTA ACTGCGTTTTTGAACAAAGTTTACAAGAGAAGCAA

**Table 3.2** continued

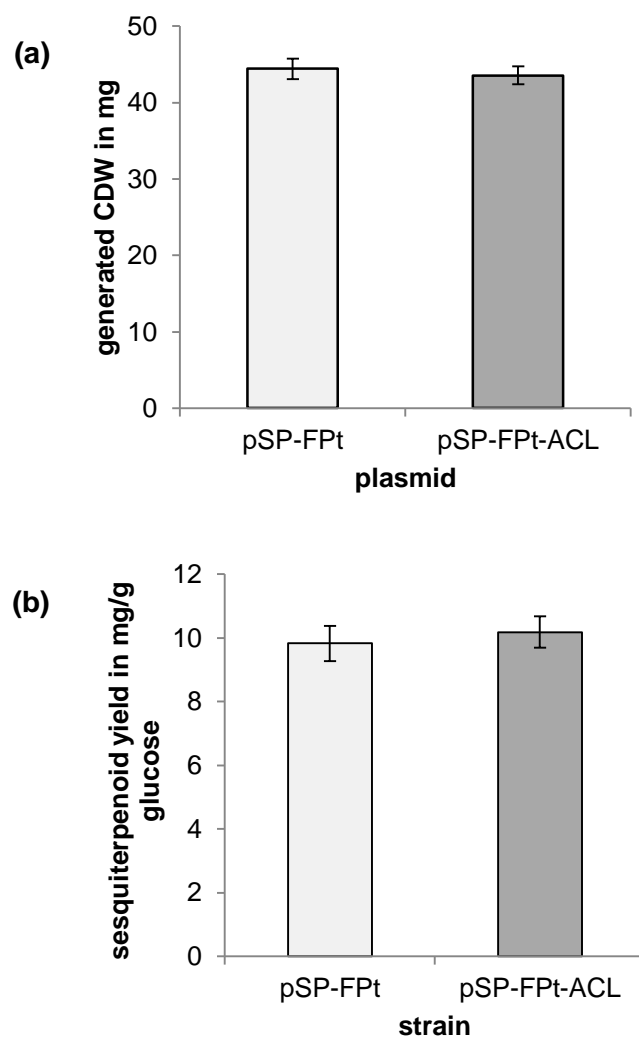
Name	Sequence
<i>ACLA-1_T2A_ACLB-2</i>	<p>ATGGCGAGGAAGAAGATCAGAGAGTATGACTCAAAGAGGTTGGTGAAGGAACATTTCAA  AGGCTTTCTGGCAAAGAGCTTCTATCAGATCCGTTCCAGGCAGATGATGTTTTCTGGTTA  CATTCTTGCTTTGGTTTCTCCGACATTAGGCACCATGTGATTAATGAAACAACCTGATCTA  AATGAGCTAGTTGAAAAGGAACCTTGGCTCTCGTCTGAGAAGCTGGTGGTGAACCTGAC  ATGTTGTTTGGAAAGCGTGGCAAGAGTGGTTTGGTTGCCCTTGAATTAGATTTTGCTGAT  GTTGCCACTTTTGTAAAGAACGTTTGGGAAAAGAGGTAGAGATGAGTGGATGCAAAGGA  CCCATAACAACATTCATAGTTGAACCATTTGTTCCACACAATGAGGAGTATTTATCTCAAT  GTTGTCTCGGATCGGCTTGGTTGCAGCATAAGCTTTTCTGAGTGTGGAGGAATTGAGATC  GAGGAGAACTGGGACAAGGTCAAGACAAATTTTTTACCAACAGGTGCTTCCCTGACACCT  GAAATATGTGCACCTCTTGTGCGCAACTTCCCTTAGAGATCAAAGCTGAAATTGAAGAA  TTTATCAAAGTCATTTTCAACCTATTTCCAAGATCTTGATTTCACTTTCTTGAGATGAAT  CCTTTCACCTAGTTGATGGAAGTCTTATCCTCTGGATATGAGGGGTGAGCTTGATGAT  ACTGCTGCCTTCAAACCTTTAAAAAATGGGGCGACATTGAATTTCCCTGCCCATTGGGA  AGAGTAATGAGTCTTACAGAAAGCTTTATCCACGGACTGGATGAGAAGACAAGTGCCTCT  TTGAAGTTTACCGTTCTGAACCCCAAGGGACGGATTGGACAATGGTAGCTGGTGGAGGA  GCAAGTGTCTATGCGGATACGGTTGGAGATCTCGGGTATGCATCTGAACTTGGCAAC  TATGCTGAATACAGTGGAGCACCCAAAGAAGATGAGGTTTTGACAGTACGCCAGAGTCCGTT  ATTGATTTGTGCTACAGCAAACCCGGATGAAAAAGCAGAGCCCTTGTCTCGGAGGCGGA  ATTGCCAACTTCACTGACGTTGCTGCTACTTTCAATGGCATAATCCGCGCTCTTAAAGAA  AAGGAAGCAAAGCTGAAAGCAGCAAGGATGCATATATTTGTGAGGAGAGGAGGACCAAAC  TACCAAAGGGACTTGCTAAAATGCGAGCCCTTGGAGATGATATCGGTGTCCCATCGAG  GTCTATGGCCAGAAGCAACCATGACAGGTATCTGCAAGGAGGCAATCCAGTACATCACA  GCAGCAGCA<b>AGAGCAGAAGGAAGGGGTTCTTTGTTGACTTGTGGAGATGTT</b>  <b>GAGGAGAATCCAGGACCA</b>GCAACGGGACAGCTTTTTTCTCGTACCACACAAGCTTT  GTTCTACAACATAAGCAGCTTCTGTTCAACGAATGCTCGATTTGACTTCTCTGTGG  ACGTGAAACGCCTTCTGTTGCTGGAATCATAAATCCTGGTTCTGAAGTTTTCAAAGCT  CTTTTTCGGGCAGGAGGAAATCGCTATCCCTGTTTATGCGCCATTGAGGCAGCTTGTGC  TGCGCATCCAACAGCGGATGTATTCACTCAACTTTGCATCTTTTAGGAGTGTCTGCTGCTT  ATCCATGGCTGCTTTGAAGCAGCCGACTATTAAGTTTGGCAATTATAGCTGAAGGTG  TCCAGAATCAGACACTAAGCAGCTGATTGCGTATGCTCGTGCAAACAATAAGGTTGTTAT  TGGACCGGCTACTGTTGGAGGTATTCAAGCTGGAGCCTTAAGATTGGTGATACTGCAGG  AACAAATGATAACATTTACAGTGAAGCTATACAGACCTGGATCTGTTGGTTTTGTCTC  CAAATCTGGTGGAAATGTCTAATGAAATGTACAATACTGTTGCCCGTGTGACTGATGGGAT  CTACGAAGGCATTGCTATTTGGTGGAGACGTGTTCCAGGATCGACTTTATCTGACCACAT  CCTTCGGTTTAAACAACATCCACAGATAAAAATGATGGTTGTACTGGAGACTTGGAGG  AAGAGATGAATACTCTCTTGTGGAAGCTTTGAAAGAGGGAAAAGTCAATAAACCTGTGGT  TGCTTGGGTCAGTGGAACTTGTGCACGACTCTTCAAGTCTGAAGTACAGTTTGGTCTATGC  AGGTGCCAAAAGTGGCGCGAGATGGAGTCTGCACAAGCCAAGAATCAAGCTCTCATAGA  TGCTGGAGCTATTGTTCCCACTTCATTTGAAGCTCTAGAATCTGCAATCAAAGAGACTTT  TGAGAACTGGTTGAAGAAGGAAAGGTCCTCTATCAAGGAAGTCATTCCTCCACAAAT  CCCTGAGGATCTCAATCTGCAATTAAGAGTGGGAAAGTCCGGGCTCTACTCACATCAT  CTCCACCATATCTGATGACAGAGGGGAGGAACCATGCTATGCTGGTGTTCATGCTCTC  CATCATCGAACAAAGGCTATGGAGTGGGTGATGTCTTTCCCTTCTATGGTTCAAACGTAG  TCTACCTCGTTACTGTACAAAATTCATTGAGATATGCATAATGCTGTGTGCTGATCACGG  TCCATGCGTCTCCGGCGCTCACAACACCATTTGTAACAGCAAGAGCAGGCAAGACCTCGT  CTCAAGTCTTGTCTCAGGTTTATTGACCATTTGGTCCCGGATTTGGTGGTGCCATTGATGA  CGCTGCTCGATACTTCAAAGACGCGTGTGACAGGAATCTCACACCTTATGAATTTGTTGA  GGAAATGAAGAAAAGGGAATCCGAGTCCCGGGATTGGACACAGGATCAAGAGCAGAGA  CAACAGAGACAAAAGAGTGGAGCTTCTCAGAAATTTGCTCGGTCCAACCTCCCATCAGT  GAAGTACATGGAGTACGCACTGACAGTGGAGACATACACGCTCTCAAAGGCAACAACCT  CGTACTCAACGTTGATGGAGCCATTGGAATCTCTCTTCTTGGACCTTCTAGCTGGAAGTGG  GATGTTCACTAAACAAGAGATTGACGAGATTGTTGAGATCGGTTATCTCAACGGTCTGTT  TGTTCTTGCTCGCTCCATCGGTTTGAATCGGGCACACGTTTATCAGAAGAGATTGAAGCA  GCCACTGTATCGTACCATGGGAAGATGTGTTGTACACCAAGTAA</p>

**Table 3.3 - Primers used in this study.** Flanking homologous regions are marked in bold, asterisk (\*) indicates PAGE purified primers.

Name	Sequence (5'-3')
MF0018	<b>GCAATCTAATCTAAGTTTTAATTACAAG</b> CGGCCGCAACAAAATGGAATTAT ACG
MF0019	<b>CTTGACCAAACCTCTGGCGAAGAATTG</b> TTAATTAATTAATATGGAAGTGGG TGC
MF0016_S	CCATCTTTTCGTAAATTTTC
MF0017_S	CATTTTTCTTGTCTATTACAAC
EG_tHMG_Fw	<b>GTAATTATCTACTTTTTACAACA</b> AAATATAAAACAAAAACAATGGACCAAT TGGTGAAAACCTG
EG_tHMG_Rv	<b>CTAACTCCTTCCTTTTCGGTTAGAGCGGATCTT</b> ATTAGGATTTAATGCAGG TG
MF0014_S	ACAGATCATCAAGGAAG
MF0015_S	GTTACATGCGTACACGC
EMG_FPPS_Fw	<b>CATAGCAATCTAATCTAAGTTTTAATTACA</b> AAACAAAATGGCTTCAGAAAA AGAAATTAG
EMG_FPPS_Rv	<b>CAACACCAACGGATTGGGCGTATAATTC</b> CATACCAGAACCTTTGCTTCTCT TGTAACCTTTG
EMG_FPPS_S	CATGGATGAAAATTAGCC
EMG_ACL_1_pSP+ PHXT7	<b>CGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGATAAAGGCGCGCCT</b> ACGCCAAGCGCGCAATTAAC
EMG_ACL_2_PHXT7+ACLA-1	<b>CTTCCTCGCCATTGTTTTTTTTGATTAAAATTA</b> AAAAAACTTTTTGTTTTTTG
EMG_ACL_3_PHXT7+ACLA-1	<b>CAAAGAATAAACACAAAAACAAAAGTTTTTTT</b> AATTTTAATCAAAAAAAC AATGGCGAGGAAGAAGATCAG
EMG_ACL_4_ACLA-1+T2A*	<b>TGGTCC TGGATTCTCC TCAACATCTCCACAAGTCAACA</b> AGAACCCTTCC <b>TTCTGCTCTTGCTGCTGCTGTGATG</b> ACTG
EMG_ACL_5_T2A+ACLB-2*	<b>AGAGCAGAAGGAAGGGTTCTTTGTTGACTTGTGGAGATGTTGAGGAGAAT</b> <b>CCAGGACCAGCAACGGGACAGCTTTTTTTC</b>
EMG_ACL_6_ACLB-2+TTEF	<b>CAAATGACAAGTTCTTGAAAACAAGAATCTTTTT</b> TATTGTCAGTACTGATTA CTTGGTGTACAACACATCTTCCC
EMG_ACL_7_ACLB-2+TTEF	<b>CACCAAGTAATCAGTACTGACAATAAAAAGATTC</b>
EMG_ACL_8_TTEF+Ascl+pSP	<b>CTCAGGTATAGCATGAGGTCGCTCCAATTCCTAGGT</b> CGTTTGGCGGCCAA TACGACTCACTATAGGGAG
EMG_ACL_9_S	GAACGTGGCGAGAAAGGAA
EMG_ACL_10_S	TCTTTGGGTGCTCCACTGT
EMG_ACL_11_S	GGCGACATTGAATTTCTT
EMG_ACL_12_S	GGGATTTGTGGAGGAATGA
EMG_ACL_13_S	AGTTTGGTCATGCAGGTGCC
EMG_ACL_14_S	TGTCGGCTTGTCTACCTTGC
EMG_KGD_Fw	<b>CCCCAAATTAAGTTTTCGTTTTGAAAGAAACAACA</b> AAAGAGAAAGAAAGCTT CGTACGCTGCAGGTGCGAC
EMG_KGD_Rv	<b>GATGGGACGTAATCCTTTGCCTTTTCGAAAGCATCTT</b> CAAATAAGTTCCGC ATAGGCCACTAGTGGATCTG
EMG_KGD_S1	TCTCGTTCAGCATCATACT
EMG_KGD_S2	GCAGCAGTTAACCATTCT



**Figure 3.17 - Effect of disruption of *KGD1* on several parameters during glucose-limited conditions.** (a) Cell growth (in generated CDW per flask), (b) acetate production, (c) sesquiterpenoids yield on glucose. Shown are mean values and standard deviations of three experiments. n.d., not detectable.



**Figure 3.18** - Effect of APT-citrate lyase expression on growth and sesquiterpenoid production during glucose-limited conditions. (a) Cell growth (in generated CDW per flask), (b) sesquiterpenoids yield on glucose. Shown are mean values and standard deviations of three experiments.

### 3.7 References

- Albertsen, L., Y. Chen, L. S. Bach, S. Rattleff, J. Maury, S. Brix, J. Nielsen, and U. H. Mortensen. 2011. Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. *Appl Environ Microbiol* 77 (3):1033-40.
- Asadollahi, M. A., J. Maury, K. Moller, K. F. Nielsen, M. Schalk, A. Clark, and J. Nielsen. 2008. Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: effect of *ERG9* repression on sesquiterpene biosynthesis. *Biotechnol Bioeng* 99 (3):666-77.
- Asadollahi, M. A., J. Maury, K. R. Patil, M. Schalk, A. Clark, and J. Nielsen. 2009. Enhancing sesquiterpene production in *Saccharomyces cerevisiae* through *in silico* driven metabolic engineering. *Metab Eng* 11 (6):328-34.
- Asadollahi, M. A., J. Maury, M. Schalk, A. Clark, and J. Nielsen. 2010. Enhancement of farnesyl diphosphate pool as direct precursor of sesquiterpenes through metabolic engineering of the mevalonate pathway in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 106 (1):86-96.
- Becker, J., and E. Boles. 2003. A modified *Saccharomyces cerevisiae* strain that consumes L-Arabinose and produces ethanol. *Appl Environ Microbiol* 69 (7):4144-50.
- Beekwilder, J., H. M. van Rossum, F. Koopman, F. Sonntag, M. Buchhaupt, J. Schrader, R. D. Hall, D. Bosch, J. T. Pronk, A. J. van Maris, and J. M. Daran. 2014. Polycistronic expression of a beta-carotene biosynthetic pathway in *Saccharomyces cerevisiae* coupled to beta-ionone production. *J Biotechnol*.
- Bowman, S. B., Z. Zaman, L. P. Collinson, A. J. Brown, and I. W. Dawes. 1992. Positive regulation of the *LPD1* gene of *Saccharomyces cerevisiae* by the *HAP2/HAP3/HAP4* activation system. *Mol Gen Genet* 231 (2):296-303.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248-54.
- Brodelius, M., A. Lundgren, P. Mercke, and P. E. Brodelius. 2002. Fusion of farnesyl diphosphate synthase and epi-aristolochene synthase, a sesquiterpene cyclase involved in capsidiol biosynthesis in *Nicotiana tabacum*. *Eur J Biochem* 269 (14):3570-7.
- Carvalho, Florbela, Luís C. Duarte, and Francisco M. Gírio. 2008. Hemicellulose biorefineries: a review on biomass pretreatments. *J Sci Ind Res* 67:849-864.
- Chavez-Cabrera, C., Z. R. Flores-Bustamante, R. Marsch, C. Montes Mdel, S. Sanchez, J. C. Cancino-Diaz, and L. B. Flores-Cotera. 2010. ATP-citrate lyase activity and carotenoid production in batch cultures of *Phaffia rhodozyma* under nitrogen-limited and nonlimited conditions. *Appl Microbiol Biotechnol* 85 (6):1953-60.
- Chen, L., J. Zhang, and W. N. Chen. 2014. Engineering the *Saccharomyces cerevisiae* beta-oxidation pathway to increase medium chain fatty acid production as potential biofuel. *PLoS One* 9 (1):e84853.
- Chen, Y., J. Bao, I. K. Kim, V. Siewers, and J. Nielsen. 2014. Coupled incremental precursor and co-factor supply improves 3-hydroxypropionic acid production in *Saccharomyces cerevisiae*. *Metabolic engineering* 22:104-109.
- Chen, Y., L. Daviet, M. Schalk, V. Siewers, and J. Nielsen. 2013. Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. *Metabolic engineering* 15:48-54.
- Chen, Y., S. Partow, G. Scalcinati, V. Siewers, and J. Nielsen. 2012. Enhancing the copy number of episomal plasmids in *Saccharomyces cerevisiae* for improved protein production. *FEMS Yeast Res* 12 (5):598-607.

- Chen, Y., Y. Zhang, V. Siewers, and J. Nielsen. 2015. *Ach1* is involved in shuttling mitochondrial acetyl units for cytosolic C2 provision in *Saccharomyces cerevisiae* lacking pyruvate decarboxylase. *FEMS Yeast Res* 15 (3).
- Ciriacy, Michael. 1975. Genetics of alcohol dehydrogenase in *Saccharomyces cerevisiae*: II. Two loci controlling synthesis of the glucose-repressible ADHII. *Molecular and General Genetics MGG* 138 (2):157-164.
- de Jong, B. W., S. Shi, V. Siewers, and J. Nielsen. 2014. Improved production of fatty acid ethyl esters in *Saccharomyces cerevisiae* through up-regulation of the ethanol degradation pathway and expression of the heterologous phosphoketolase pathway. *Microb Cell Fact* 13 (1):39.
- Deguerry, F., L. Pastore, S. Wu, A. Clark, J. Chappell, and M. Schalk. 2006. The diverse sesquiterpene profile of patchouli, *Pogostemon cablin*, is correlated with a limited number of sesquiterpene synthases. *Arch Biochem Biophys* 454 (2):123-36.
- Donald, K. A., R. Y. Hampton, and I. B. Fritz. 1997. Effects of overproduction of the catalytic domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase on squalene synthesis in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 63 (9):3341-4.
- Engels, B., P. Dahm, and S. Jennewein. 2008. Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metabolic engineering* 10 (3-4):201-6.
- Farhi, M., E. Marhevka, T. Masci, E. Marcos, Y. Eyal, M. Ovadis, H. Abeliovich, and A. Vainstein. 2011. Harnessing yeast subcellular compartments for the production of plant terpenoids. *Metab Eng* 13 (5):474-81.
- Fatland, B. L., J. Ke, M. D. Anderson, W. I. Mentzen, L. W. Cui, C. C. Allred, J. L. Johnston, B. J. Nikolau, and E. S. Wurtele. 2002. Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in *Arabidopsis*. *Plant Physiol* 130 (2):740-56.
- George, K. W., J. Alonso-Gutierrez, J. D. Keasling, and T. S. Lee. 2015. Isoprenoid Drugs, Biofuels, and Chemicals-Artemisinin, Farnesene, and Beyond. *Adv Biochem Eng Biotechnol* 148:355-389.
- Gietz, R. D., and R. H. Schiestl. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2 (1):31-4.
- Gruchattka, E., O. Hädicke, S. Klamt, V. Schütz, and O. Kayser. 2013. *In silico* profiling of *Escherichia coli* and *Saccharomyces cerevisiae* as terpenoid factories. *Microb Cell Fact* 12 (1):84.
- Guaragnella, N., L. Antonacci, S. Passarella, E. Marra, and S. Giannattasio. 2011. Achievements and perspectives in yeast acetic acid-induced programmed cell death pathways. *Biochem Soc Trans* 39 (5):1538-43.
- Gueldener, U., S. Heck, T. Fielder, J. Beinhauer, and J. H. Hegemann. 1996. A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res* 24 (13):2519-24.
- Hädicke, O., and S. Klamt. 2011. Computing complex metabolic intervention strategies using constrained minimal cut sets. *Metab Eng* 13 (2):204-13.
- Heipieper, Hermann J., Frans J. Weber, Jan Sikkema, Heribert Keweloh, and Jan A. M. de Bont. 1994. Mechanisms of resistance of whole cells to toxic organic solvents. *Trends in Biotechnology* 12 (10):409-415.
- Ignea, C., M. Pontini, M. E. Maffei, A. M. Makris, and S. C. Kampranis. 2014. Engineering Monoterpene Production in Yeast Using a Synthetic Dominant Negative Geranyl Diphosphate Synthase. *ACS Synth Biol*.



- Inoue, H., H. Nojima, and H. Okayama. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96 (1):23-8.
- Jackson, B. E., E. A. Hart-Wells, and S. P. Matsuda. 2003. Metabolic engineering to produce sesquiterpenes in yeast. *Org Lett* 5 (10):1629-32.
- Knorre, D. A., E. A. Smirnova, and F. F. Severin. 2005. Natural conditions inducing programmed cell death in the yeast *Saccharomyces cerevisiae*. *Biochemistry (Mosc)* 70 (2):264-6.
- Koch, B., C. Schmidt, and G. Daum. 2014. Storage lipids of yeasts: a survey of nonpolar lipid metabolism in *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Yarrowia lipolytica*. *FEMS Microbiol Rev* 38 (5):892-915.
- Kocharin, K., Y. Chen, V. Siewers, and J. Nielsen. 2012. Engineering of acetyl-CoA metabolism for the improved production of polyhydroxybutyrate in *Saccharomyces cerevisiae*. *AMB Express* 2 (1):52.
- Kocharin, K., V. Siewers, and J. Nielsen. 2013. Improved polyhydroxybutyrate production by *Saccharomyces cerevisiae* through the use of the phosphoketolase pathway. *Biotechnol Bioeng*.
- Kozak, B. U., H. M. van Rossum, K. R. Benjamin, L. Wu, J. M. Daran, J. T. Pronk, and A. J. van Maris. 2013. Replacement of the *Saccharomyces cerevisiae* acetyl-CoA synthetases by alternative pathways for cytosolic acetyl-CoA synthesis. *Metabolic engineering*.
- Kozak, B. U., H. M. van Rossum, M. A. Luttk, M. Akeroyd, K. R. Benjamin, L. Wu, S. de Vries, J. M. Daran, J. T. Pronk, and A. J. van Maris. 2014. Engineering Acetyl Coenzyme A Supply: Functional Expression of a Bacterial Pyruvate Dehydrogenase Complex in the Cytosol of *Saccharomyces cerevisiae*. *MBio* 5 (5).
- Krivoruchko, A., C. Serrano-Amatriain, Y. Chen, V. Siewers, and J. Nielsen. 2013. Improving biobutanol production in engineered *Saccharomyces cerevisiae* by manipulation of acetyl-CoA metabolism. *J Ind Microbiol Biotechnol* 40 (9):1051-6.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227 (5259):680-5.
- Lian, J., T. Si, N. U. Nair, and H. Zhao. 2014. Design and construction of acetyl-CoA overproducing *Saccharomyces cerevisiae* strains. *Metabolic engineering* 24:139-49.
- Liesen, T., C. P. Hollenberg, and J. J. Heinisch. 1996. ERA, a novel cis-acting element required for autoregulation and ethanol repression of *PDC1* transcription in *Saccharomyces cerevisiae*. *Mol Microbiol* 21 (3):621-32.
- LoGrasso, P. V., D. A. Soltis, and B. R. Boettcher. 1993. Overexpression, purification, and kinetic characterization of a carboxyl-terminal-truncated yeast squalene synthetase. *Arch Biochem Biophys* 307 (1):193-9.
- Munck, S. L., and R. Croteau. 1990. Purification and characterization of the sesquiterpene cyclase patchoulol synthase from *Pogostemon cablin*. *Arch Biochem Biophys* 282 (1):58-64.
- Navarro-Avino, J. P., R. Prasad, V. J. Miralles, R. M. Benito, and R. Serrano. 1999. A proposal for nomenclature of aldehyde dehydrogenases in *Saccharomyces cerevisiae* and characterization of the stress-inducible *ALD2* and *ALD3* genes. *Yeast* 15 (10A):829-42.
- Nielsen, J., and J. D. Keasling. 2011. Synergies between synthetic biology and metabolic engineering. *Nat Biotechnol* 29 (8):693-5.
- Papini, M., I. Nookaew, V. Siewers, and J. Nielsen. 2012. Physiological characterization of recombinant *Saccharomyces cerevisiae* expressing the *Aspergillus nidulans* phosphoketolase pathway: validation of activity through (13)C-based metabolic flux analysis. *Appl Microbiol Biotechnol*.

- Papini, M., I. Nookaew, M. Uhlen, and J. Nielsen. 2012. *Scheffersomyces stipitis*: a comparative systems biology study with the Crabtree positive yeast *Saccharomyces cerevisiae*. *Microb Cell Fact* 11 (1):136.
- Paradise, E. M., J. Kirby, R. Chan, and J. D. Keasling. 2008. Redirection of flux through the FPP branch-point in *Saccharomyces cerevisiae* by down-regulating squalene synthase. *Biotechnol Bioeng* 100 (2):371-8.
- Peralta-Yahya, P. P., M. Ouellet, R. Chan, A. Mukhopadhyay, J. D. Keasling, and T. S. Lee. 2011. Identification and microbial production of a terpene-based advanced biofuel. *Nat Commun* 2:483.
- Polakowski, T., U. Stahl, and C. Lang. 1998. Overexpression of a cytosolic hydroxymethylglutaryl-CoA reductase leads to squalene accumulation in yeast. *Appl Microbiol Biotechnol* 49 (1):66-71.
- Pronk, Jack T., H. Yde Steensma, and Johannes P. Van Dijken. 1996. Pyruvate Metabolism in *Saccharomyces cerevisiae*. *Yeast* 12 (16):1607-1633.
- Przybyla-Zawislak, B., D. M. Gadde, K. Ducharme, and M. T. McCammon. 1999. Genetic and biochemical interactions involving tricarboxylic acid cycle (TCA) function using a collection of mutants defective in all TCA cycle genes. *Genetics* 152 (1):153-66.
- Rico, J., E. Pardo, and M. Orejas. 2010. Enhanced production of a plant monoterpene by overexpression of the 3-hydroxy-3-methylglutaryl coenzyme A reductase catalytic domain in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 76 (19):6449-54.
- Ro, D. K., E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Chang, S. T. Withers, Y. Shiba, R. Sarpong, and J. D. Keasling. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440 (7086):940-3.
- Sarria, S., B. Wong, H. G. Martin, J. D. Keasling, and P. Peralta-Yahya. 2014. Microbial Synthesis of Pinene. *ACS Synth Biol*.
- Scalcinati, G., C. Knuf, S. Partow, Y. Chen, J. Maury, M. Schalk, L. Daviet, J. Nielsen, and V. Siewers. 2012. Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquiterpene alpha-santalene in a fed-batch mode. *Metabolic engineering* 14 (2):91-103.
- Scalcinati, G., S. Partow, V. Siewers, M. Schalk, L. Daviet, and J. Nielsen. 2012. Combined metabolic engineering of precursor and co-factor supply to increase alpha-santalene production by *Saccharomyces cerevisiae*. *Microb Cell Fact* 11 (1):117.
- Shiba, Y., E. M. Paradise, J. Kirby, D. K. Ro, and J. D. Keasling. 2007. Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* 9 (2):160-8.
- Tang, X., H. Feng, and W. N. Chen. 2013. Metabolic engineering for enhanced fatty acids synthesis in *Saccharomyces cerevisiae*. *Metabolic engineering* 16:95-102.
- Tippmann, S., Y. Chen, V. Siewers, and J. Nielsen. 2013. From flavors and pharmaceuticals to advanced biofuels: production of isoprenoids in *Saccharomyces cerevisiae*. *Biotechnol J* 8 (12):1435-44.
- van den Berg, M. A., P. de Jong-Gubbels, C. J. Kortland, J. P. van Dijken, J. T. Pronk, and H. Y. Steensma. 1996. The two acetyl-coenzyme A synthetases of *Saccharomyces cerevisiae* differ with respect to kinetic properties and transcriptional regulation. *J Biol Chem* 271 (46):28953-9.
- Verduyn, C. 1991. Physiology of yeasts in relation to biomass yields. *Antonie Van Leeuwenhoek* 60 (3-4):325-53.

- Verwaal, R., J. Wang, J. P. Meijnen, H. Visser, G. Sandmann, J. A. van den Berg, and A. J. van Ooyen. 2007. High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous*. *Appl Environ Microbiol* 73 (13):4342-50.
- Westfall, P. J., D. J. Pitera, J. R. Lenihan, D. Eng, F. X. Woolard, R. Regentin, T. Horning, H. Tsuruta, D. J. Melis, A. Owens, S. Fickes, D. Diola, K. R. Benjamin, J. D. Keasling, M. D. Leavell, D. J. McPhee, N. S. Renninger, J. D. Newman, and C. J. Paddon. 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci U S A* 109 (3):E1111-8.

# **CHAPTER 4**

## **Discussion and Outlook**

**Evamaria Gruchattka, Oliver Kayser**

**EG designed and wrote the chapter receiving input and supervision from OK**

## 4.1 *In silico* analysis

The *in silico* analysis of metabolic networks for production organisms has become a popular tool for two reasons: (i) the inherent potential of a microbial host can be deciphered, and (ii) metabolic engineering targets can be identified in order to increase the production of a compound of interest. The two promising and most prominently used heterologous hosts *Escherichia coli* and *Saccharomyces cerevisiae* were analyzed and compared *in silico* using elementary mode analysis (Chapter 2) in order to identify new metabolic engineering strategies for strain improvement to reliably increase microbial terpenoid production. The terpenoid pathway of *E. coli* (DXP pathway) has a tremendously higher theoretical potential to supply terpenoids in high yields than the one of *S. cerevisiae* (MVA pathway). However, the use of the DXP pathway in *E. coli* is inherently limited by as of yet unidentified regulation, in addition, the full pathway has not been functionally expressed in yeast (Carlsen et al. 2013; Partow et al. 2012). The last two steps of the DXP pathway are catalyzed by iron-sulfur [4Fe-4S] cluster proteins, whose functional expression in the cytosol of yeast is challenging. Other heterologous iron-sulfur cluster proteins have not yet been functionally expressed to a level that reaches a sufficient enzyme activity. For example, 6-phosphogluconate dehydratase of the bacterial Entner-Doudoroff pathway could not be over-expressed in yeast (Benisch and Boles 2014). Functional expression still poses a challenge that will need to be addressed in the future in order to tap the full potential of heterologous enzymes for yeast metabolic engineering.

In both *Escherichia coli* and *Saccharomyces cerevisiae*, deficiencies in energy and redox equivalents for high yield terpenoid production have been identified *in silico* and are the source of new overexpression strategies. Promising heterologous enzymes for an increased terpenoid yield have been identified. NADP<sup>+</sup>-dependent glyceraldehyde-3-phosphate dehydrogenase enhances theoretical terpenoid yield in *E. coli* when the DXP pathway is operating. Indeed, this target has been validated by another research group and the NADP<sup>+</sup>-dependent enzyme from *Bacillus subtilis* increased lycopene production in *E. coli* (Wang, San, and Bennett 2013). In yeast, a cytosolic pyruvate dehydrogenase complex is a promising target as it increases the theoretical maximum terpenoid yield. Indeed, a bacterial pyruvate dehydrogenase complex has been functionally expressed in the cytosol of yeast (Kozak et al. 2014). However, media supplementation with lipoate (a cofactor produced in mitochondria that cannot readily leave the compartment) was required for activity and it has not yet been shown whether the enzyme activity is

sufficient to increase terpenoid production. Another promising approach included the transfer of yeasts native MVA pathway from cytosol into mitochondria to use the more energy-efficient mitochondrial pyruvate dehydrogenase complex for acetyl-CoA formation. However, this strategy was not validated yet likely due to the technical challenges of targeting six enzymes to mitochondria. Heterologous expression of ATP-citrate lyase is another promising approach as it increases the theoretical maximum terpenoid yield. The gene had not been previously expressed with the aim to increase terpenoid production in yeast and was thus chosen as a strategy to be validated *in vivo* as part of this thesis. Finally, several knockout strategies were identified using constrained minimal cut sets that enforce coupling of growth to a predefined terpenoid yield. None of those strategies have yet been validated, therefore, one of these strategies involving a gene disruption in the citric acid cycle, was chosen for investigation in this work.

Different carbon sources were also analyzed for their potential to supply terpenoids in high yields. Indeed, yeast and *E. coli* can reach higher theoretical maximum terpenoid yields using non-fermentable carbon sources, such as ethanol and glycerol, due to less carbon loss or favorable redox equivalent usage, respectively. The robust potential of ethanol as carbon source has been validated for yeast *in vivo*, and terpenoid yield was increased tremendously when comparing glucose to ethanol as carbon source (Westfall et al. 2012). Sugars are desired carbon sources, molasses and hemicellulosic hydrolysates of agricultural byproducts consist mainly of sugars, and are low-cost substrates for industrial fermentation (Carvalho, Duarte, and Gírio 2008). Glucose, fructose, galactose, and xylose have the same or a similar theoretical potential to supply terpenoids in high yields. Moreover, several metabolic engineering strategies could be applied for growth on these different sugars and experimental terpenoid yields achieved so far on sugars are very low (Table 2.2). Thus, the validation of metabolic engineering strategies to improve terpenoid yield from glucose as carbon source was chosen for analysis in this project.

Even though *E. coli* wild type has a theoretical higher potential than yeast wild type to supply terpenoids in high yields, metabolic engineering can be applied to both organisms to increase theoretical maximum terpenoid yield. Both organisms are suitable promising heterologous hosts. An advantage of yeast is its robustness and the higher probability to functionally express plant P450 enzymes (Julleson et al. 2015; Kirby and Keasling 2009), therefore, yeast was chosen for *in vivo* validation in this work.

In summary, the metabolic network capabilities of the two hosts *Escherichia coli* and *Saccharomyces cerevisiae* were analyzed and, indeed, new promising metabolic engineering targets in the central carbon metabolism were identified as envisaged (aim I., Chapter 1.5). Several strategies were validated by other research groups since then – either successfully or reaching some limitations concerning *in vivo* activity of certain enzymes. Two strategies were chosen for *S. cerevisiae* and glucose as a carbon source in the follow-up study for hands-on validation: disruption of  $\alpha$ -ketoglutarate dehydrogenase gene (*KGD1*) of  $\alpha$ -ketoglutarate dehydrogenase complex (citric acid cycle) and overexpression of a heterologous ATP-citrate lyase.

## 4.2 *In vivo* validation

Certain requirements had to be complied with in order to validate selected strategies to redirect the metabolic flux in the central carbon metabolism towards terpenoids. As the chosen reporter, the sesquiterpenoid patchoulol, has volatile properties, a two-phase cultivation was employed as an *in situ* product removal tool. Dodecane was chosen and shown to have no detrimental effects on yeast growth and captures patchoulol efficiently over time as previously described (Asadollahi et al. 2008). As many terpenoid synthases produce several terpenic products, the sesquiterpenoid spectrum of patchoulol synthase was analyzed to estimate the production of total sesquiterpenoids by the yeast strain. The product spectrum and the patchoulol content were determined to consist of ~ 34 %, well in accordance with previous studies (Deguerry et al. 2006).

A truncated HMG-CoA reductase was overexpressed to eliminate the main site of regulation in the MVA pathway and patchoulol synthase was fused to FPP synthase in order to encourage that the redirected metabolic flux was directly processed by the terpenoid synthase rather than spread through native FPP metabolizing pathways. Both strategies led to an increase in sesquiterpenoid formation, however, only 36 % and 22 % increases respectively. Nevertheless, the squalene content was increased 10-fold when *tHMG1* was overexpressed, which is well in accordance with literature (Donald, Hampton, and Fritz 1997) although the absolute increase was small. Even though HMG-CoA reductase is well known as the rate-limiting enzyme in the MVA pathway, the increase in sesquiterpenoids after *tHMG1* overexpression was comparably low in many studies including ours (Albertsen et al. 2011; Farhi et al. 2011; Jackson, Hart-Wells, and Matsuda 2003). Many inherent issues of comparison of these other studies may be

attributed to: different cultivation conditions, different previous genetic modifications of the strains, different expression levels due to different molecular strategies, or even the use of different terpenoid synthases.

After establishment of the terpenoid capture and analysis strategy for *S. cerevisiae*, as well as deregulation of the MVA pathway (aim II., Chapter 1.5), selected metabolic engineering strategies were implemented (aim III., Chapter 1.5). The disruption of  $\alpha$ -ketoglutarate dehydrogenase gene (*KGD1*) of  $\alpha$ -ketoglutarate dehydrogenase complex belongs to a constrained minimal cut sets approach and was thought to redirect the metabolic flux from citric acid cycle via pyruvate dehydrogenase bypass towards terpenoid precursors. Indeed, the flux was redirected, but an intermediate in the pyruvate dehydrogenase bypass, acetate, was secreted and accumulated in the cultivation broth due to insufficient activity of acetyl-CoA synthetase (Chapter 3.4.4). Moreover, the flux redirection was dependent on cultivation conditions and no effect was achieved in batch conditions. However, acetate accumulated during glucose-limited conditions and during growth on ethanol, which was produced during the glucose-phase. Further effects were detected: reduced terpenoid formation (ethanol-phase only) and inhibition of cell growth. Acetate itself might be a cause for the observed effects due to its role as a signal (Guaragnella et al. 2011). Additional research is needed to resolve these issues. This initial analysis could indicate the effect of different acetate concentrations on biomass and terpenoid formation. Certainly, acetate formation should be avoided to increase terpenoid formation.

The second strategy to increase the metabolic flux towards terpenoid production was the overexpression of a heterologous ATP-citrate lyase, which forms acetyl-CoA from citrate that has been produced via the pyruvate dehydrogenase complex and citric acid cycle in the mitochondria. The reaction is more energy-efficient and avoids the pyruvate dehydrogenase bypass. A self-cleaving T2A sequence from *Thosea asigna* virus was successfully used to generate two separate polypeptides (*ACLA-1* and *ACLB-2*) from a polycistronic mRNA. However, overexpressing ATP-citrate lyase did not lead to an increase in terpenoid formation. As ATP-citrate lyase from *Arabidopsis* was previously overexpressed with insufficient *in vivo* activity in an attempt to increase *n*-butanol production (Lian et al. 2014), we assume that the kinetic properties are unfavorable and *in vivo* activities of the enzyme from *Arabidopsis* are not sufficient to influence terpenoid production in *S. cerevisiae*. Indeed, other studies confirm that *in vivo* performance of heterologous enzymes needs to be optimized to increase the desired product, for



example, as in the case of a the bacterial pyruvate dehydrogenase complex as well as the acetylating acetaldehyde dehydrogenase and pyruvate-formate lyase (Kozak et al. 2013; Kozak et al. 2014).

Different carbon sources were also evaluated for their role as fuel for cellular metabolism and terpenoid production. Ethanol was identified as a very promising carbon source in the *in silico* analysis and indeed, terpenoid production was increased tremendously from glucose-phase to ethanol-phase by a factor of 8.4 (Chapter 3.4.3). The results presented here indicate that the use of glucose as an inexpensive carbon source combined with ethanol, which was produced by the yeast itself, is a promising approach to produce terpenoids to greater yields. The effect of the disruption of  $\alpha$ -ketoglutarate dehydrogenase was strongly dependent on cultivation conditions due to a complex regulation within the yeast. Therefore, this study highlights the importance of the use of different carbon sources or cultivation conditions for the characterization of genetic modifications aimed at increasing terpenoid production capacity.

In summary, patchoulol production was successfully established in *S. cerevisiae*, and the main sites of regulation in the MVA pathway were eliminated, which resulted in a subtle increase in target terpenoid production (aim II, Chapter 1.5). Cultivation conditions were analyzed and two selected metabolic engineering strategies based on *in silico* predictions were implemented: disruption of  $\alpha$ -ketoglutarate dehydrogenase and overexpression of ATP-citrate lyase. Some predictions could be positively validated (aim III, Chapter 1.5), including (i) the potential of ethanol as carbon source, and (ii) redirection of the metabolic flux after a knockout in citric acid cycle to the pyruvate dehydrogenase bypass. However, some limitations related to *in vivo* activity of certain enzymes were reached: (i) ATP-citrate lyase (insufficient activity *in vivo* to influence terpenoid production), and (ii) acetyl-CoA synthetase (Insufficient activity to activate the quantity of acetate and prevent acetate secretion).

### 4.3 Conclusion and outlook

In order to develop robust metabolic engineering strategies to boost terpenoid production from microbial hosts, several limitations will need to be overcome. The main sites of regulation must be elucidated and circumvented in order to maximize the pathways metabolite flow. In addition, enzyme variants, either from other source organisms, through semi-rational protein design, or directed evolution are needed to

ensure activity in the production host, which is true as well for native host enzymes. It may be possible to generate variants which allow un-regulated or increased flux towards terpenoid building blocks for increased target product generation.

A reduction in the acetate formation would be an important first step. Acetate can passively diffuse across the plasma membrane in its undissociated form - one approach could be to increase the conversion of acetate to acetyl-CoA via replacing (Shiba et al. 2007) or deregulating and increasing the activity of acetyl-CoA synthetase. This may be accomplished via rational protein design or directed evolution. In addition, an increase in acetate uptake via overexpression of transporters for acetate uptake, for example *ADY2* (Paiva et al. 2004), may be advantageous. ATP-citrate lyase from different organisms could be screened to identify a more suitable enzyme in terms of *in vivo* activity, like enzyme variants from the oleaginous yeast *Yarrowia lipolytica* or the carotenoid producing yeast *Xanthophyllomyces dendrorhous*. Additionally, expression levels could be optimized and enzyme engineering may increase relative activity.

Although the applied stoichiometric metabolic network analysis was limited due to regulation and unpredicted enzyme activity, such methods can be used successfully as a metabolic flux prediction tool. In this study, predictions regarding ethanol as carbon source were confirmed and the metabolic flux was indeed redirected as predicted according to the constrained minimal cut sets approach. Unfortunately, the intermediate acetate was produced instead of terpenoids. If the produced acetate could be efficiently converted to sesquiterpenoids, a titer in the g/L range could be achieved. Thus, the *in silico* predicted strategy is indeed valuable and should be pursued in the future.

The investigations conducted in this work highlight that the inclusion of regulatory circuits and *in vivo* kinetics of enzymatic reactions into the *in silico* metabolic network analysis could be beneficial to improve prediction quality. The applied constrained minimal cut sets approach has been developed further since our application to reduce the computation time, to enable the analysis of genome-scale metabolic networks, and to include up/downregulation of reaction rates together with reaction deletions (Mahadevan, von Kamp, and Klamt 2015; von Kamp and Klamt 2014). A combination with kinetic modelling has been suggested to increase the robustness of the prediction and the incorporation of kinetics has been discussed (Mahadevan, von Kamp, and Klamt 2015). Still, the implementation is pending.

In conclusion, the results of this thesis contribute to increased knowledge of the inherent metabolic behavior of yeast, and towards enhancing heterologous terpenoid

production in both *E. coli* and *S. cerevisiae*, using various carbon sources, terpenoid pathways, and metabolic engineering strategies. The applied elementary mode analysis and the constrained minimal cut sets approach have proven to yield accurate structural predictions. However, kinetics and regulation were not included in the model and interfered, thus resulting in poor *in vivo* performance. These findings promote the incorporation of regulation and kinetic parameters into metabolic modelling, in order to increase the predictive power and accelerate microbial strain improvement.

## 4.4 References

- Albertsen, L., Y. Chen, L. S. Bach, S. Rattleff, J. Maury, S. Brix, J. Nielsen, and U. H. Mortensen. 2011. Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. *Appl Environ Microbiol* 77 (3):1033-40.
- Asadollahi, M. A., J. Maury, K. Moller, K. F. Nielsen, M. Schalk, A. Clark, and J. Nielsen. 2008. Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: effect of *ERG9* repression on sesquiterpene biosynthesis. *Biotechnol Bioeng* 99 (3):666-77.
- Benisch, F., and E. Boles. 2014. The bacterial Entner-Doudoroff pathway does not replace glycolysis in *Saccharomyces cerevisiae* due to the lack of activity of iron-sulfur cluster enzyme 6-phosphogluconate dehydratase. *J Biotechnol* 171:45-55.
- Carlsen, S., P. K. Ajikumar, L. R. Formenti, K. Zhou, T. H. Phon, M. L. Nielsen, A. E. Lantz, M. C. Kielland-Brandt, and G. Stephanopoulos. 2013. Heterologous expression and characterization of bacterial 2-C-methyl-D-erythritol-4-phosphate pathway in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 97:5753-5769.
- Carvalho, Florbela, Luís C. Duarte, and Francisco M Gírio. 2008. Hemicellulose biorefineries: a review on biomass pretreatments. *J Sci Ind Res* 67:849-864.
- Deguerry, F., L. Pastore, S. Wu, A. Clark, J. Chappell, and M. Schalk. 2006. The diverse sesquiterpene profile of patchouli, *Pogostemon cablin*, is correlated with a limited number of sesquiterpene synthases. *Arch Biochem Biophys* 454 (2):123-36.
- Donald, K. A., R. Y. Hampton, and I. B. Fritz. 1997. Effects of overproduction of the catalytic domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase on squalene synthesis in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 63 (9):3341-4.
- Farhi, M., E. Marhevka, T. Masci, E. Marcos, Y. Eyal, M. Ovadis, H. Abeliovich, and A. Vainstein. 2011. Harnessing yeast subcellular compartments for the production of plant terpenoids. *Metab Eng* 13 (5):474-81.
- Guaragnella, N., L. Antonacci, S. Passarella, E. Marra, and S. Giannattasio. 2011. Achievements and perspectives in yeast acetic acid-induced programmed cell death pathways. *Biochem Soc Trans* 39 (5):1538-43.
- Jackson, B. E., E. A. Hart-Wells, and S. P. Matsuda. 2003. Metabolic engineering to produce sesquiterpenes in yeast. *Org Lett* 5 (10):1629-32.
- Jullesson, D., F. David, B. Pflieger, and J. Nielsen. 2015. Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. *Biotechnol Adv.*
- Kirby, J., and J. D. Keasling. 2009. Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu Rev Plant Biol* 60:335-55.
- Kozak, B. U., H. M. van Rossum, K. R. Benjamin, L. Wu, J. M. Daran, J. T. Pronk, and A. J. van Maris. 2013. Replacement of the *Saccharomyces cerevisiae* acetyl-CoA synthetases by alternative pathways for cytosolic acetyl-CoA synthesis. *Metabolic engineering.*
- Kozak, B. U., H. M. van Rossum, M. A. Luttik, M. Akeroyd, K. R. Benjamin, L. Wu, S. de Vries, J. M. Daran, J. T. Pronk, and A. J. van Maris. 2014. Engineering Acetyl Coenzyme A Supply: Functional Expression of a Bacterial Pyruvate Dehydrogenase Complex in the Cytosol of *Saccharomyces cerevisiae*. *MBio* 5 (5).
- Lian, J., T. Si, N. U. Nair, and H. Zhao. 2014. Design and construction of acetyl-CoA overproducing *Saccharomyces cerevisiae* strains. *Metabolic engineering* 24:139-49.
- Mahadevan, R., A. von Kamp, and S. Klamt. 2015. Genome-scale strain designs based on regulatory minimal cut sets. *Bioinformatics.*

- Paiva, S., F. Devaux, S. Barbosa, C. Jacq, and M. Casal. 2004. Ady2p is essential for the acetate permease activity in the yeast *Saccharomyces cerevisiae*. *Yeast* 21 (3):201-10.
- Partow, S., V. Siewers, L. Daviet, M. Schalk, and J. Nielsen. 2012. Reconstruction and Evaluation of the Synthetic Bacterial MEP Pathway in *Saccharomyces cerevisiae*. *PLoS One* 7 (12):e52498.
- Shiba, Y., E. M. Paradise, J. Kirby, D. K. Ro, and J. D. Keasling. 2007. Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* 9 (2):160-8.
- von Kamp, A., and S. Klamt. 2014. Enumeration of smallest intervention strategies in genome-scale metabolic networks. *PLoS Comput Biol* 10 (1):e1003378.
- Wang, Y., K. Y. San, and G. N. Bennett. 2013. Improvement of NADPH bioavailability in *Escherichia coli* by replacing NAD(+)-dependent glyceraldehyde-3-phosphate dehydrogenase GapA with NADP (+)-dependent GapB from *Bacillus subtilis* and addition of NAD kinase. *J Ind Microbiol Biotechnol* 40 (12):1449-60.
- Westfall, P. J., D. J. Pitera, J. R. Lenihan, D. Eng, F. X. Woolard, R. Regentin, T. Horning, H. Tsuruta, D. J. Melis, A. Owens, S. Fickes, D. Diola, K. R. Benjamin, J. D. Keasling, M. D. Leavell, D. J. McPhee, N. S. Renninger, J. D. Newman, and C. J. Paddon. 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinin acid, precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci U S A* 109 (3):E111-8.

# Appendix

## List of Abbreviations

1,3PG	1,3-bisphosphoglycerate
2PG	2-phosphoglycerate
3PG	3-phosphoglycerate
6PG	6-phosphogluconate
5-FOA	5-fluoroorotic acid
ACE	acetate
AcCoA	acetyl-CoA
ACEADH	acetaldehyde
AcetoAcCoA	acetoacetyl-CoA
Acetyl-P	acetyl phosphate
ACL	ATP-citrate-lyase
ADP	adenosine diphosphate
AKG	$\alpha$ -ketoglutarate
AMP	adenosine monophosphate
ATP	adenosine triphosphate
BCIP	5-bromo-4-chloro-3-indolyl phosphate
bp	base pairs
[c]	cytosolic
CIT	citrate
cDNA	complementary DNA
CDW	cell dry weight
cMCS	constrained minimal cut set
CoA	coenzyme A
Cre	cyclization recombinase
CTP	cytidine triphosphate
$\tilde{\mathbf{D}}$	desired modes
DAD	diode array detector
DH	dehydrogenase
DHAP	dihydroxyacetone phosphate
DMAPP	dimethylallyl diphosphate
DNA	deoxyribonucleic acid
DTT	dithiothreitol

---

DXP	1-deoxy-D-xylulose 5-phosphate
E4P	erythrose-4-phosphate
EDTA	ethylenediaminetetraacetic acid
EM	elementary mode
EMA	elementary mode analysis
EtOH	ethanol
ex	excretion
ext	external
F6P	fructose-6-phosphate
FAD/H <sub>2</sub>	oxidized/reduced flavin adenine dinucleotide
FBP	fructose-1,6-bisphosphate
F.A.M.E.	fatty acid methyl esters
FID	flame ionization detector
FPP	farnesyl diphosphate
FUM	fumarate
G	glycine, Gly
G15L	6-phosphogluconolacton
G418	geneticin
G6P	glucose-6-phosphate
GAP	glyceraldehyde-3-phosphate
GC	gas chromatography
GLC	glucose
GLYC	glycerol
GPP	geranyl diphosphate
GGPP	geranylgeranyl diphosphate
HPLC	high performance liquid chromatography
ICI	isocitrate
IgG	immunoglobulin G
IPP	isopentenyl diphosphate
IUCN	International Union for the Conservation of Nature
KDPG	2-keto-3-deoxy-6-phospho-gluconate
K <sub>m</sub>	Michaelis–Menten constant
k <sub>cat</sub>	constant for conversion to product
LAC	lactate



---

[m]	mitochondrial
M	marker or protein ladder
MAL	malate
MCS	minimal cut set
MEP	2C-methyl-D-erythritol 4-phosphate
MS	mass spectrometer/-metry
MVA	mevalonate
N	stoichiometric matrix
NAD(P) <sup>+</sup> /H	oxidized/reduced nicotinamide adenine dinucleotide (phosphate)
NBT	nitroblue tetrazolium
OAA	oxalacetate
OD	optical density
p	metabolites
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDH	cytosolic pyruvate dehydrogenase complex
PEP	phosphoenolpyruvate
PHB	poly- $\beta$ -hydroxybutyrate
PMSF	phenylmethylsulfonyl fluoride
PTS	patchoulol synthase
PYR	pyruvate
q	reactions
q <sub>r</sub>	reversible reaction
q <sub>irr</sub>	irreversible reaction
r	reaction rate flux vector
R5P	ribose-5-phosphate
RID	refractive index detector
rpm	rounds per minute
RUP	ribulose-5-phosphate
S	serine, Ser
S7P	sedoheptulose-7-phosphate
SD	synthetic dextrose medium
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SUCC	succinate

---

<i>Taq</i>	<i>Thermus aquaticus</i>
$\tilde{\mathbf{T}}$	target modes
TH	soluble and energy-independent transhydrogenase
tHMG	truncated HMG-CoA (3-hydroxy-3-methyl-glutaryl-coenzyme A) reductase
VWD	variable wave length detector
XI	xylose isomerase
XR-XDH	xylose reductase and xylitol dehydrogenase
XUP	xylose-5-phosphate
XYL	xylose

## List of Publications

### Peer Reviewed Articles

Gruchattka, E., Kayser, O., 2015, *In vivo* validation of *in silico* predicted metabolic engineering strategies in yeast: disruption of  $\alpha$ -ketoglutarate dehydrogenase and expression of ATP-citrate lyase for terpenoid production, *POLS ONE*, 10(12):e0144981

Gruchattka, E., Hädicke, O., Klamt, S., Schütz, V., Kayser, O., 2013, *In silico* profiling of *Escherichia coli* and *Saccharomyces cerevisiae* as terpenoid factories, *Microbial Cell Factories* 12(1):84

### Oral Presentations

Gruchattka, E., 2014, Towards a platform organism for terpenoid production – *in silico* analysis of potential hosts and *in vivo* validation, 5<sup>th</sup> Annual CLIB-Graduate Cluster Retreat, 12.02. – 14.02.2014, Lünen, Germany

Gruchattka, E., 2013, Towards a platform organism for terpenoid production – *in silico* analysis of central pathways of *Saccharomyces cerevisiae* for pathway optimization, 4<sup>th</sup> European Yeast Flavour Workshop Conference: Biotechnology for Natural Flavour Production, 22.07. – 23.07.2013, Freising, Germany

### Poster Presentations

Gruchattka, E., Kayser, O., 2015, *In vivo* validation of *in silico* predicted metabolic engineering strategies for terpenoid production in yeast, Bioflavour 2015, 09.09. – 11.09.2015, Frankfurt/Main, Germany

Gruchattka, E., Schütz, V., Kayser, O., 2014, Towards a platform organism for terpenoid production – *in silico* analysis of potential hosts and *in vivo* validation, 5<sup>th</sup> Annual CLIB-Graduate Cluster Retreat, 12.02. – 14.02.2014, Lünen, Germany

Gruchattka, E., Schütz, V., Kayser, O., 2013, Towards a platform organism for terpenoid production – *in silico* analysis of *Saccharomyces cerevisiae* as potential host, 26<sup>th</sup> International Conference on Yeast Genetics and Molecular Biology (Yeast), 29.08. – 03.09.2013, Frankfurt/Main, Germany

Gruchattka, E., Schütz, V., Kayser, O., 2013, Towards a platform organism for terpenoid production – *in silico* analysis of metabolic networks of potential hosts and *in vivo* validation, 4<sup>th</sup> Annual CLIB-Graduate Cluster Retreat, 13.02. – 15.02.2013, Lünen, Germany

Felten, M., Gruchattka, E., Kayser, O., Schütz, V., 2012, Patchoulol production in *S. cerevisiae* – Metabolic Engineering and process optimization, Biotrends: Sustainable industrial biocatalysis, 29.11. – 30.11.2012, Dortmund, Germany

Gruchattka, E., Kayser, O., Schütz, V., 2012, Towards a platform organism for terpenoid production – *in silico* comparison of *E. coli* and *S. cerevisiae* as potential hosts, Metabolic Engineering IX: Metabolic Engineering and Synthetic Biology, 03.06. – 07.06.2012, Biarritz, France

Gruchattka, E., Kayser, O., Schütz, V., 2012, Towards a platform organism for terpenoid production – *in silico* comparison of metabolic networks of *E. coli* and *S. cerevisiae* as potential hosts, Applied Synthetic Biology in Europe, 06.02. – 08.02.2012, Barcelona, Spain

Gruchattka, E., Schütz, V., Kayser, O., 2012, Towards a platform organism for terpenoid production – *in silico* analysis of metabolic networks of potential hosts and *in vivo* validation, 3<sup>rd</sup> Annual CLIB-Graduate Cluster Retreat, 22.02. – 24.02.2012, Bergisch Gladbach, Germany

Gruchattka, E., Schütz, V., Kayser, O., 2011, Metabolic and metabolomic profiling of *Xanthophyllomyces dendrorhous* as biosynthetic platform organism for isoprenoid drugs, 2<sup>nd</sup> Annual CLIB-Graduate Cluster Retreat, 07.04. – 08.04.2011, Bonn, Germany

## ***Curriculum vitae***

### **PERSONAL DETAILS**

Name	Evamaria Gruchattka
Date of Birth	18.07.1983
Place of Birth	Homburg (Efze)
Nationality	German

### **EDUCATION, QUALIFICATIONS AND PROFESSIONAL EXPERIENCE**

Since 04/2015	Scientific research assistant at the EU Project Office, Research Funding and Knowledge Transfer, TU Dortmund University
Since 04/2011	Ph.D. student at the Department of Biochemical and Chemical Engineering (BCI), TU Dortmund University Scholarship holder of the CLIB-Graduate Cluster "Industrial Biotechnology" and scientific research assistant at the Chair of Technical Biochemistry (TB)
11/2010 – 02/2011	Scientific research assistant at MicrobiMaris Biotec GmbH, Preetz
07/2007 – 06/2008	Study Abroad, University of Tasmania, Australia
03/2007 – 02/2009	Student research assistant, intermittent, at Kieler Wirkstoff-Zentrum at Leibniz-Institut für Meereswissenschaften (IFM-GEOMAR), Kiel and MicrobiMaris Biotec GmbH, Preetz
09/2003 – 9/2010	Christian-Albrechts-Universität zu Kiel, Leibniz-Institut für Meereswissenschaften (IFM-GEOMAR) and Kieler Wirkstoff-Zentrum (KiWiZ), Kiel Studies of Biology Degree: Diplom (very good) Major subject: Biological Oceanography Minor subjects: Microbiology, Geology
1994 – 2003	Jacob Grimm Schule, Rotenburg an der Fulda Degree: Abitur (very good, 1.5)
1990 – 1994	Lindenschule, Alheim Heinebach