Novel strategies for engineering redox metabolism in *Saccharomyces cerevisiae*

Víctor Gabriel Guadalupe Medina

2013

Novel strategies for engineering redox metabolism in *Saccharomyces cerevisiae*

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.Ch.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
maandag 14 oktober 2013 om 15:00 uur

door

Víctor Gabriel GUADALUPE MEDINA

Magíster en Ciencias de la Ingeniería,
Pontificia Universidad Católica de Chile,
geboren te Rancagua, Chili.

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. J.T. Pronk

Copromotor: Dr. ir. A.J.A. van Maris

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. J.T. Pronk,	Technische Universiteit Delft, promotor
Dr. ir. A.J.A. van Maris,	Technische Universiteit Delft, copromotor
Prof. dr. J.G. Kuenen,	Technische Universiteit Delft
Prof. dr. R.A.L. Bovenberg,	Rijksuniversiteit Groningen / DSM
Prof. dr. B.M. Bakker,	Rijksuniversiteit Groningen
Prof. dr. J. Förster,	Technical University of Denmark / Novo Nordisk Foundation Center for Biosustainability
Dr. R.A. Weusthuis,	Wageningen University
Prof. dr. M.C.M. van Loosdrecht,	Technische Universiteit Delft, reservelid

The studies presented in this thesis were performed at the Industrial Microbiology section, Department of Biotechnology, Delft University of Technology, The Netherlands and part of Program 1 'Yeast for chemicals, fuels and chemicals' of the Kluyver Centre for Genomics of Industrial Fermentation, which is supported by the Netherlands Genomics Initiative.

The cover of this thesis was designed by Manuel Toledo Otaegui (www.toledotaegui.com). The photograph was kindly provided by Michael Grab, rock balancing artist (www.gravityglue.com).

Printed by Ipskamp Drukkers BV, Enschede, The Netherlands.

ISBN 978-94-6186-218-1

Table of Contents

Chapter 1

General Introduction

Industrial biotechnology and the production of fuels and chemicals

Over the past three centuries, the Earth's human population has increased nearly exponentially, from 350 million at the start of the 18th century to the current number of 7 billion. This population growth trend is predicted to continue for at least the next 50 years[1] and requires, amongst other adaptations, an increased capacity to produce food and materials. Using science as a catalyst, mankind has previously experienced several "revolutions" in the production of food, material and energy carriers[2]. The most recent of these revolutions started with the use of fossil feed-stocks and chemical engineering, which enabled automation, mass production of goods, and a large increase in agriculture production by the use of fertilizers and improved high-yielding crops (also called the "Green Revolution"). These changes, helped to fulfill the growing needs of society in the last century although, unfortunately, not to the same extent in all countries. As a consequence, 20% of the current world population, mostly inhabitants of developed countries, use *ca.* 80% of the natural resources[3]. As the other 80% of the world population develop, such as the populations of China, Brazil, India and Chile, their consumption of fossil feedstocks increases rapidly[4]. Simultaneously, due to concerns related to energy security and the effects of increased atmospheric carbon dioxide concentrations, developed countries (United States of America and Western Europe) are trying to improve the efficacy of the use of energy and materials[4]. Nowadays, more than 90% of the global transportation fuels are derived from fossil oil and this level of use is not expected to decrease below 75% of the current level within the next 20 years[5]. As a consequence of continuously increasing demand, fewer discoveries of new oil wells, and instability of geo-political regions where oil is extracted, the crude oil price has risen over the past years[6]. As the oil price becomes more volatile and tends to increase, both developed and developing countries are looking for alternative resources. In summary, for reasons of sustainability and economics, a future with high living standard for the entire world population requires alternative sources of energy and chemicals[7].

On the road to a paradigm change in the way society obtains energy and materials, several alternatives are being proposed to provide energy for transport or production in a sustainable manner[8]. For instance, the sunlight that reaches our planet in one hour represents enough energy to cover one year worth of world-wide current energy use, if it would be technologically feasible to catch, transform, store, and distribute all that energy[9]. However, for production of chemicals and materials, the main alternative is to return to the way materials where made before the

beginning of the 20th century: using biomass as raw material[10]. Biomass offers a sustainable alternative to limited fossil feedstocks[11]. However, an increased use of biomass as raw material for the production of chemicals and materials in a so-called bio-based economy involves a whole array of technical and socioeconomic challenges[12,13]. Fortunately, over the last 50 years, the capacity to understand and modify life has increased rapidly, driven to a significant extent by the emergence of molecular genetics during the 1970s and the enormous advances in genomics in the past two decades[14]. The field of industrial biotechnology seeks to address the societal need for sustainable alternatives to petrochemistry by harnessing these recent developments in life sciences. In particular, industrial biotechnology aims to use microorganisms and/or enzymes to produce chemicals, food and energy from biomass[15]. The archaeological record has provided evidence of the use of microorganisms in food products, especially in fermented beverages as far back as 7000 years BC[16]. Several microorganisms have been "domesticated" during the course of human history, and nowadays, microorganisms are widely used to extend food shelf life[17] and increase their nutritional content and flavor[17], and for the production of medicines (*e.g.* antibiotics and vaccines[18]). Moreover, microbial biodiversity is a generous fountain of interesting genes and enzymes for new industrial applications.

In many cases, industrial biotechnology offers several advantages over chemical synthesis, such as milder conditions (lower temperature and neutral pH), less energy intensive, specific molecular chirality and higher efficiency[19]. Secondly, the time constant of carbon recycling in industrial biotechnology is several orders of magnitude smaller than that typically found in the petrochemical industry[20], which can make production of materials and chemicals from biomass more sustainable than production from oil (Fig. 1.1). Since the introduction of, for example, the fermentation process for citric acid production in the 1920s, and industrial penicillin production in the 1950s[21], the number of processes involving microbial cultures has increased drastically, resulting in an impressive list of bio-based chemicals[22]. Of the many microorganisms used in industrial biotechnology, the yeast *Saccharomyces cerevisiae* stands out by combining a long history of safe use, simple nutritional requirements for growth, an ability to grow fast under anaerobic conditions and robustness under industrial process conditions. In addition, this microorganism has become one of the most studied eukaryotic cell systems[23].

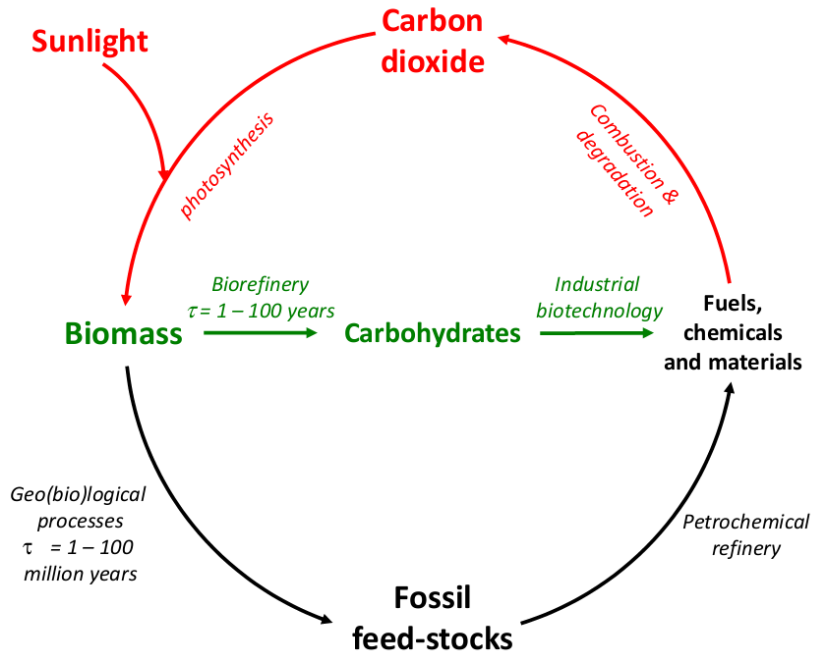

Figure 1.1: Carbon cycles for the production of fuels, chemicals and materials. If industrial biotechnology is used, a possible reduction in 6 orders of magnitude in turnover rate can be achieved, making the process more sustainable than petrochemical manufacture. Figure adapted from van Maris *et al.*[20]

Physiology of *Saccharomyces cerevisiae*

Saccharomyces cerevisiae, latin for "Sugar fungus (*Saccharomyces*) from beer (*cerevisiae*)", was one of the first microorganisms observed under the microscope by Antonie van Leeuwenhoek in Delft, The Netherlands[24]. *S. cerevisiae* can be found in nature on damaged grapes[25] and in the gastrointestinal tract of insects[26]. For a long time, *S. cerevisiae* has been widely used in the production of fermented beverages for the conversion of sugar to ethanol, and in bread baking for carbon dioxide production during dough leavening. Due to this important role in food industry, its physiology has been studied intensively since 1856, the year that Louis Pasteur showed that yeast was responsible for the conversion of grape juice into wine[27]. In the past 50 years, *S. cerevisiae* has become one of the main cell factories for the production of chemicals from biomass in industrial biotechnology[28-30]. Moreover, it has also become an important eukaryotic cell model, since its biochemical and signal transduction pathways are similar to those of higher eukaryotic organisms, while remaining much simpler to culture[23]. The study of yeast has contributed to the

development of an impressive variety of recombinant DNA technologies. Moreover, functional analysis of the genes present in its genome, the first complete eukaryotic genome published[31], has helped to better understand a variety of cellular processes. The accumulated knowledge on yeast physiology and functional genetics has a direct impact on the improvement of the bioprocesses in which *S. cerevisiae* cultures are used, thereby contributing to the development of industrial biotechnology as an alternative to petrochemical synthesis[29].

In view of the many industrial applications of *S. cerevisiae*, this introduction and this thesis will focus mainly on anaerobic glucose metabolism. Anaerobic conditions provide economic benefits compared to aerobic processes by avoiding costly air compression, simplifying process control and allowing cheaper reactor design[32]. Additionally glucose is the prevailing carbon source in industrial biotechnology. In alcoholic fermentation carried out by yeasts, the available 6-carbon sugars (glucose, fructose and galactose) and some disaccharides (maltose and sucrose) are catabolized to ethanol and carbon dioxide by fermentation (*i.e.* a redox-neutral catabolic process where transfer of electrons occurs from substrates to the products in the absence of an external electron acceptor[32]). Alcoholic fermentation starts by cytoplasmic conversion of one molecule of glucose to two molecules of pyruvate through the Embden-Meyerhof pathway (further referred as glycolysis). This pathway consists of 10 enzymatic reactions and its description was completed in 1940[33]. Glycolysis makes free energy available to the cell in the form of 2 moles of adenosine triphosphate (ATP) for each mole of glucose that is converted to 2 moles of pyruvate. This ATP, the free-energy currency of the cell, is formed by substrate-level phosphorylation of adenosine diphosphate (ADP). The capacity to 'move around' free energy within the cell and to couple ATP de-phosphorylation to thermodynamically unfavorable reactions is crucial for growth and cellular maintenance. Nonetheless, glycolysis does not only yield ATP. The conversion of glucose to pyruvate is an oxidative process and the electrons released in this process enter the metabolism in the form of two 2 moles of reduced pyridine-nucleotide (NADH), formed by reduction (accepting electrons) of 2 moles of NAD^+.

The transfer of electrons between reduction and oxidation reactions by pyridine-nucleotide (NAD^+) and pyridine-nucleotide phosphate ($NADP^+$) plays an important role in cellular metabolism[34]. In terms of biological function, catabolic reactions often use NAD^+, whereas anabolic reactions often use $NADP^+$. Cellular metabolism requires the regeneration of conserved moieties, such as AMP/ADP/ATP and $NAD(P)^+$/NAD(P)H. During alcoholic fermentation, the pyruvate formed by glycolysis is subsequently converted to ethanol and carbon

dioxide in two cytosolic reactions. First pyruvate decarboxylase converts pyruvate to CO_2 and acetaldehyde, which is afterwards reduced to ethanol by alcohol dehydrogenase. This reduction of acetaldehyde re-oxidizes both molecules of NADH generated by glycolysis, making alcoholic fermentation of glucose a redox neutral process (Fig. 1.2).

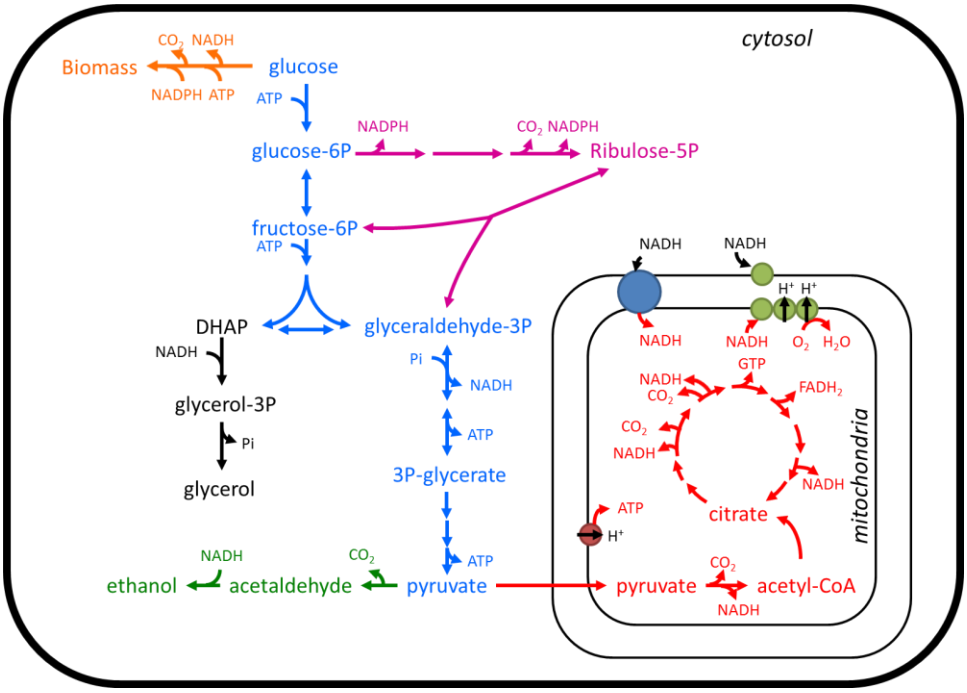

Figure 1.2: Schematic representation of glucose metabolism of *Saccharomyces cerevisiae* with emphasis on redox co-factor balance and ATP metabolism. In orange, biomass generation is summarized in one reaction consuming ATP and NADPH, and producing NADH[35]. NADPH is provided by the oxidative branch of pentose phosphate pathway (in purple), while glycolysis (in blue) provides ATP and NADH, the latter oxidized in pyruvate reduction to ethanol in green (redox neutral) alcoholic fermentation. To restore redox balance, (in black) glycerol is formed by reduction of dihydroxyacetone phosphate to oxidize biomass-derived NADH. In the presence of oxygen (red), pyruvate can be oxidized to carbon dioxide and water by the tricarboxylic acid cycle, and the formed NADH/$FADH_2$ are reoxidized by the respiratory chain (green dot in mitochondria inner membrane) and proton-gradient-driven ATP-synthase (red circle in mitochondria inner membrane) subsequently provides the cells with additional ATP. In the presence of oxygen, cytosolic NADH can also be reoxidized by action of the two external NADH dehydrogenases in *S. cerevisiae* (green circles in mitochondria outer membrane) or transferred via different shuttle mechanisms[36] into the mitochondria (blue circle).

Besides the reactions in alcoholic fermentation, there are *ca.* 200 other redox reactions that occur within metabolism that depend on the ratio between the redox cofactors NAD^+/NADH and $NADP^+$/NADPH[37]. These ratios together determine the redox state of the cell and influence the concentrations of other cellular metabolites and/or the directionality of other reactions. Since these redox cofactors are only present at low concentration in living cells[34], maintaining an exact redox-cofactor balance between oxidizing and reducing reactions is an absolute requirement of cellular metabolism. In eukaryotic organisms such as *S. cerevisiae*, cellular compartmentalization plays an important role in the maintenance of different redox states within the cell[36]. In the case of NAD^+/NADH ratio, several shuttle mechanisms exist to transfer electrons between intracellular compartments, *e.g.* between the cytosol and the mitochondria[36].

Pyruvate is at the branch point between fermentative and oxidative metabolism[38]. During respiratory dissimilation pyruvate is transported from the cytoplasm into the mitochondria, where it is converted to acetyl coenzyme A (Acetyl-CoA) by the pyruvate dehydrogenase complex. Acetyl-CoA enters the TriCarboxylic Acid cycle (TCA cycle), also known as the citric acid or Szent-Györgyi-Krebbs cycle, and is further dissimilated to carbon dioxide and water. The TCA cycle provides redox equivalents NADH and $FADH_2$ within the mitochondria. These reduced redox equivalents, together with the 2 NADH molecules generated during glycolysis, are oxidized to NAD^+ and FAD by the electron transport chain, a series of protein complexes present in the mitochondrial inner membrane that transfer the electrons to oxygen[39]. These protein complexes couple the transfer of electrons with the export of H^+ across the inner mitochondrial membrane, creating an electrochemical potential. Another protein complex located in the mitochondrial membrane, ATP synthase[40], then uses this potential to generate ATP from ADP and P_i.

Whereas anaerobically alcoholic fermentation is the only mode of sugar dissimilation in *S. cerevisiae*, in the presence of oxygen the type of metabolism depends on both the prevailing glucose concentration and the growth rate of the culture. At high glucose concentrations and/or high growth rates, such as observed during the exponential phase in aerobic batch fermentation, but also during aerobic chemostat cultivation at high dilution rates, a mixture between respiration and fermentation is observed. This type of metabolism is known as respiro-fermentative metabolism, and it is characterized by ethanol production even when oxygen is present[41]. Only at low glucose concentrations and at low specific growth rates, glucose dissimilation in aerobic cultures of *S. cerevisiae* is fully respiratory. These

conditions occur for instance in aerobic chemostat cultures at dilution rates below the so-called critical growth rate[42], and during fed-batch cultivation at low specific growth rates[43].

Formation of biomass from sugars generates a net amount of NADH as a consequence of, amongst others, oxidative decarboxylation reactions that occur during amino-acid and lipid biosynthesis[44,45]. Oxidative reactions in biosynthesis also lead to a net generation of NADH. In the presence of oxygen, this "excess" of NADH is oxidized together with the other redox equivalents produced in glycolysis and the TCA cycle to conserve free energy. However, in the absence of oxygen the cells need a different final electron acceptor to maintain its redox balance. In baker's yeast, this role is taken by the glycolytic metabolite dihydroxyacetone phosphate (DHAP), which is reduced to glycerol-3-phosphate (G3P) by glycerol-3-phosphate dehydrogenase. In *S. cerevisiae* two genes, *GPD1*[46,47] and *GPD2*[48], encode for isoenzymes of glycerol-3-phosphate dehydrogenase. These isoenzymes have a different role and regulation in yeast physiology[49]. Subsequently, G3P is de-phosphorylated by glycerol-3-phosphate phosphatase to produce glycerol[50], the main by-product in alcoholic fermentation.

Bioethanol production as an industrial biotechnological example

Ethanol production by alcoholic fermentation of sugars under anaerobic conditions and the use of ethanol as an alternative fuel is probably one of the best known examples of the potential role of industrial biotechnology in society[15]. The concept of using ethanol as a car fuel is, however, anything but novel. Already at the beginning of the automobile industry, ethanol was considered as a possible fuel in internal combustion engines (Fig. 1.3). Henry Ford originally designed his model T car to use ethanol, but switched to gasoline in 1908 when this became cheaper than ethanol[51].

Figure 1.3: Advertisement from the early 1900's for transportation vehicles using ethanol as fuel. The vehicle was constructed by Motorfahrzeug-und Motorenfabrik Berlin AG, a German company dedicated to the manufacture of motors, automobiles, and trucks. In 1902, the company merged with Daimler Motoren Gesellschaft, nowadays Daimler AG, the manufacturer of amongst others Mercedes-Benz cars[52].

However, since 1973, when the 'Oil Crisis' took place as a consequence of the embargo by the Organization of Arab Petroleum Exporting Countries (OAPEC), ethanol has regained interest as alternative fuel[51]. Currently, Brazil and the USA are the two main ethanol producers in the world. In 1975, Brazil established its National Alcohol Program (PROÁLCOOL) to promote the use of ethanol as gasoline additive or alternative fuel[51]. Also the USA started programs to promote the use of mixtures of alcohol and gasoline during the 70's, known as "Gasohol", in an attempt to become less dependent on OAPEC's oil[51]. Today, ethanol production is the largest process by volume in industrial biotechnology, reaching a global production of 110 billion litters[53]. Most, if not all, of this large volume of production is achieved by using *S. cerevisiae* as a microbial work horse to ferment sugar-rich crops, such as sugar cane, corn, wheat and sugar beet, to ethanol (1st generation ethanol)[51]. The use of these crops, which also play a role in the human food chain, in the ethanol industry has been blamed by some to have partly caused the rise in food prices in the last 6 years, leading to the so-called "Food *vs.* Fuel" discussion in 2007/2008[54]. For this reason, because of the relatively contribution of the use of first-generation biofuels on carbon dioxide emissions and because of land availability issues, the use of alternatives biomass sources such as

non-food crops or agricultural and forestry residues (2nd generation ethanol[55]), and algae (3rd generation biofuels[56]) have gained much research interest in industrial biotechnology. These biomass sources possess several advantages over food-crops, such as more efficient use of land and fresh water resources and higher fermentable biomass content. However, to make the production of ethanol from these alternative substrates feasible, multiple technological challenges have to be overcome (extensively reviewed in [20,55-58]).

Reduction of by-product formation and maximizing the ethanol yield under relevant industrial conditions, *i.e.* anaerobic and using glucose as main carbon source, is relevant for both 1st and 2nd generation ethanol production. Even relatively small increases in ethanol productivity and yield can have a significant impact given the high global production levels.

Process optimization related to redox metabolism

S. cerevisiae strains used for the production of ethanol have long been selected for better productivity, higher ethanol yields on sugar, and tolerance to the demanding conditions occurring during the production process[59,60]. These conditions include high initial glucose concentration[61], presence of organic acids from bacterial contamination[62], high dissolved carbon dioxide and high final ethanol concentrations[62]. Besides glycerol formation, biomass and carbon dioxide are the other (main) by-products in alcoholic fermentation. Several strategies related to process conditions have been studied to reduce by-product(s) formation and optimize the ethanol yield and productivity.

Decreasing the biomass yield on sugar not only directly increases the fraction of sugar that goes to ethanol formation, but also decreases the formation of glycerol, since this is stoichiometrically coupled to the reoxidation of NADH formed in biosynthesis of cell material. The biomass yield can be decreased by increasing the fraction of sugar that is used for cellular maintenance, as exemplified by the effect of organic acids addition to growth media[63,64]. Organic acids exist in a balance between their dissociated and un-dissociated forms, determined by the ambient pH and the acid dissociation constant pK_a. The un-dissociated form of organic acids can diffuse through the cell membrane and subsequently quickly dissociate into the anion and a proton (H^+) at the near neutral intracellular pH[65]. Both charged particles (H^+ and organic anion) need to be pumped out of the cell to maintain cellular homeostasis, at the expense of ATP. This free-energy expense increases the ATP requirement for maintenance of cellular functions. The direct consequence is that more carbon is directed towards ethanol formation to obtain the required extra ATP, which increases ethanol yield and reduces biomass and

glycerol yields[64]. However, high concentrations of organic acids can inhibit cellular metabolism or can even lead to cell death, as the ATP burden for pumping out the protons from within the cell becomes too high, leading to intracellular pH acidification[65]. Also recycling of yeast reduces growth and increases the fraction of sugar that is used for cellular maintenance. In laboratory studies, this can be mimicked in continuous fermentation set-ups that use a filter to retain biomass within the reactor or that recycle biomass back into the fermenter. Such a set-up, known as retentostat or recyclostat, can in theory lead to a situation where the growth rate becomes zero and substrate consumption matches maintenance energy requirements[66-68]. Using this fermentation set-up, it was observed that zero-growth yeast cells remained metabolically active and, importantly, increased their tolerance to difference environmental stress factors[66].

Differences in biomass yield on sugar between laboratory and industrial conditions already decrease the amount of sugar used for glycerol formation from 10% to 4%[69]. To further reduce glycerol formation in anaerobic ethanol production, alternative electron acceptors are required. One option is the use of oxygen as electron acceptor, known as micro-oxygenation or micro-aeration, which would require limited air compression and mixing. Previous studies on limiting oxygen availability in glucose-excess continuous cultures showed increased ethanol and biomass productivity coupled to decrease in glycerol productivity, though a decrease in specific ethanol production rate was also observed[70-73]. Addition of oxygen also increased cell viability and specific growth rate in fed-batch cultures at very high ethanol concentrations[74]. The addition of an alternative electron acceptor to the media, *i.e.* any other compound that can be taken up easily by yeast and that can be reduced by NADH, also leads to a net glycerol reduction. Examples of this are the addition of acetoin, pyruvate or acetaldehyde resulting in the formation of 2,3 butanediol, ethanol/CO_2 and ethanol respectively. However, the addition of these compounds only replaces glycerol formation by another by-product and/or increases the operational cost. A second strategy to reduce glycerol is the use of (costly) rich media containing mixtures of amino acids. Several studies have shown that not only glycerol production is reduced, but also higher maximum specific growth rates are (μ_{max}) are obtained, when a mixture of amino-acids or just one (glutamic acid) are used as nitrogen source[44,75].

Another strategy used to increase ethanol productivity and reduce operation costs is the use of very high gravity fermentations (VHG), where the concentration of sugars at the start of the batch process is very high (> 300 g l^{-1})[61]. Besides its important role in redox-cofactor balancing, under high gravity conditions glycerol is the main compatible solute in *S. cerevisiae*. A compatible solute is a compound that it

is accumulated within the cell to compensate for extracellular osmotic pressure[76]. The reaction of baker's yeast to osmotic stress is a coordinated response involving membrane modification, protein degradation, and intracellular accumulation of glycerol, a process which is regulated by the High Osmotic Glycerol response (HOG) pathway[77,78]. This sensory and induction pathway is the most studied mitogen activated protein (MAP) kinase signal cascade in eukaryotes, and it is used as a model of how cells sense and react to their environments. Modifications of this signaling pathway or of the capacity of the cell to produce glycerol in order to reduce the formation of this by-product in ethanol fermentation, *e.g.* the deletion of *GPD1* and *GPD2* in *S. cerevisiae* genome, severely impacts the tolerance of yeast towards osmotic stress[78,79].

As indicated above, low concentrations of organic acids can be beneficial for ethanol production by increasing the ethanol yield on sugar. However, lignocellulosic hydrolysates, the desired feedstocks for 2nd generation ethanol production, often contain much higher levels of organic acids[80]. One of these, acetic acid, is a structural component of hemicellulose and reaches a final concentration up to 15 g l^{-1} [81,82], depending of the biomass source[83]. The use of fed-batch fermentation of hydrolysates has proven effective to keep inhibitors concentration low by controlling the feeding rate and allowing *in vivo* detoxification[84,85]. Nevertheless, detoxifying these hydrolysates or operation at neutral pH adds operational cost and increases the risk for contamination. Therefore, the use of robust and tolerant *S. cerevisiae* strains is more desirable.

Yeast metabolic engineering as a way to improve product formation

Although process optimization and classical strain improvement have resulted in higher ethanol yields and lower by-product formation, the advances in molecular genetics allow the use of metabolic engineering for further strain improvement. Metabolic engineering was defined in 1991 by Professor James E. Bailey as "the improvement of cellular activities by manipulation of enzymatic, transport and regulatory functions of the cell with the use of recombinant DNA technology"[86]. This is illustrated by the metabolic engineering cycle, which provides industrial biotechnology with microorganisms having new desired characteristics for the production of fuels and chemicals (Fig. 1.4).

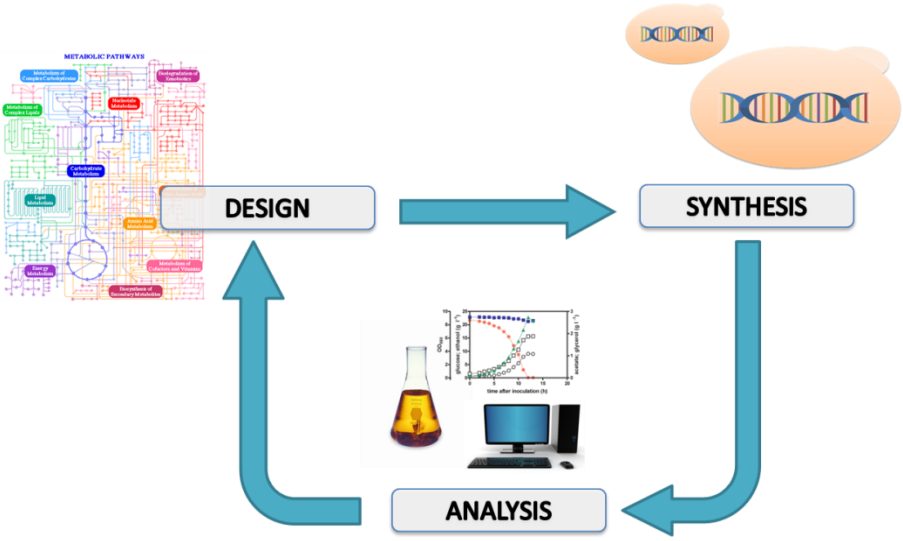

Figure 1.4: The metabolic engineering cycle as adapted from Nielsen (2001)[87], a recursive process used in the development of microbial cell factories. After the definition of a particular aim for improvement, the design of the metabolic engineering strategy is the first phase, using current metabolic pathways (pathway picture obtained from Kyoto Encyclopedia of Genes and Genomes[88]), regulatory, and genetic knowledge of the chosen microbial workhorse. In the synthesis phase, genetic modification and different selection strategies are used to bring into reality the designed metabolic engineering strategy. After the synthesis step, physiological characterization under laboratory controlled conditions and different omics-technologies can provide information to determine if the initial objective was achieved, and if not, how the design can be improved.

The main aims of metabolic engineering can be classified as follows:

- Expanding substrate range: microorganisms isolated from nature are capable of using a certain number of sources for carbon, nitrogen, or any other type of required substrates, and this capacity is different from species to species. The introduction of heterologous genes, which encode enzymes that catalyze reactions linking a substrate that a microorganism cannot normally use to its main catabolic route, may enable the consumption of this new substrate by the genetically modified microorganism and its use for the production of biomass or any other product of interest. The first step in such a novel pathway is often the uptake of the substrate from the extracellular environment. A good example of metabolic engineering for the

expansion of the substrate range is the use of pentose sugars xylose and arabinose by *S. cerevisiae* in alcoholic fermentation[20,89,90].

- Expanding product range: By the introduction of novel, heterologous genes that encode enzymes involved in a product pathway from a native metabolite to a product of interest, strains capable to produce a new compound can be constructed. Usually additional research efforts are required to re-distribute metabolic fluxes to the product of interest from other (catabolic) products, and in this process the use of driving forces within the metabolism, such as ATP generation or the redox cofactor balance, are of importance. If necessary, a final step in the production of new compounds is its export from the cytosol to the extracellular environment. Lactic acid production in aerobic cultures of *S. cerevisiae* is a good example for the production of new catabolic products[91].
- Robustness under process conditions: In order to avoid loss of productivity it is essential to maintain cellular viability, *i.e.* to keep cells metabolically active and capable of growth, under demanding industrial process conditions. For this, the transfer of mechanisms to degrade toxic compounds, tolerate a broad range of temperature or pH, and/or high concentration of substrates from more tolerant microorganisms to cell factories is of industrial interest. An example of this objective is the tolerance to furfural, a compound present in lignocellulosic hydrolysates, by the overexpression of glucose-6-phosphate dehydrogenase gene *ZWF1* in *S. cerevisiae*[92].
- By-product elimination and increased product yield: By-product(s) formation in industrial biotechnology results in product yields below the theoretical maximum. This is especially important when the cost of raw materials plays a major role in the overall process economy, as is the case in ethanol production. The redirection of metabolic fluxes from by-product formation towards product of interest can be achieved by genetic modifications that reduce or eliminate the formation of the by-product. However, it is of great importance to understand the mechanisms behind by-product formation. A fine example of this aim is the elimination of oxalate formation in penicillin production using the fungi *Penicillium chrysogenum*. Oxalate formation is associated with lower product yields and harder product recovery. The deletion of the *PcoahA* gene encoding for oxaloacetate hydrolase completely eliminated oxalate formation, leading to

higher production yields of the penicillin precursor metabolite adipoyl-6-aminopenicillinic acid[93].

- Higher titers: For an industrial process to be feasible, not only the capacity to produce a (new) chemical compound is required, but also the final product concentration is important for product purification and thereby for process economics. Because of this, genetic modifications leading to higher titers, and to tolerance to higher concentration of product, are also necessary. For instance, increased ethanol tolerance is required to reach higher final ethanol concentration in VHG fermentations, which reduces operation costs for ethanol distillation. Several strategies have been used for genetically engineering ethanol tolerance in yeast, as reviewed by Zhao & Bai[62].
- Increased productivity: Increasing product yield on substrate is important for increasing productivity but other parameters must also be considered, such as biomass specific production rates, volumetric production rate and the growth kinetics of the microorganisms. The development of metabolic engineering strategies to achieve higher production rates and growth kinetics can directly affect the process economy. The selection of a *S. cerevisiae* strain for faster consumption of different sugars is a good example for increasing ethanol volumetric productivity[94].

After the selection of a goal for improvement, a metabolic engineering strategy is designed relying on the existing knowledge of the industrial host cell and of any heterologous proteins and/or pathways that are to be expressed. Considering the thermodynamics of relevant reactions under biologically relevant conditions is of paramount importance to assess whether the envisaged strategies are realistic. The use of mathematical models representing the entire metabolism of a microbe and linked to its genome, *i.e.* genomic scale models, allows to predict the stoichiometric impact of genetic modifications on cell factories[37,95,96] and even explore possible novel metabolic pathways that do not exist within nature[97,98]. After the design phase, the main tool to execute these metabolic strategies is the ability to genetically modify a cell factory. Genetic modification is based on the fact that under specific conditions, cells readily take up deoxyribonucleic acid (DNA) sequences, and incorporate it into their own genetic repertoire[99]. This process was described even before the discovery of the double-helix DNA structure[100] and the role of DNA as hereditary material passed from progenitors to offspring[101]. Since these discoveries, the ability to construct specific DNA sequences and to incorporate them in a rational manner into microorganisms has continuously

increased. Two techniques have been crucial in the development of molecular biology: the specific amplification of DNA sequences by polymerase chain reaction (PCR[102]), and the use of restriction and ligation enzymes to "cut" and "paste" DNA sequences[103,104]. With these techniques, scientists were able to generate significant amounts of specific DNA fragments and insert them in foreign DNA sequences (*i.e.* vectors), which later are incorporated into the cellular machinery through transformation and selection for a particular selection pressure (*i.e.* by using specific marker genes[104-106]). One type of vectors consists of self-replicating circular DNA sequences that exist outside the genomic DNA (*i.e.* centromeric or multicopy plasmids). This approach can be used to increase the gene copy number (overexpression), but has as main disadvantage that once the selection pressure disappears, the plasmids tend to be quickly lost[104,107]. A second type of vectors widely used are linear or circular DNA sequences that lack an origin of replication but that can be directly inserted into genomic DNA and selected using a marker. Different mechanisms are responsible for the direct insertion of these linear vectors into the genome. One such mechanism is homologous recombination, a type of genetic exchange that occurs between DNA that contains similar or identical regions, and that it is used by the cell to repair double strand breaks in the genome, during meiosis or horizontal transfer of genes[108]. Using this technique, stable genetic modifications (gene insertion, gene deletion, or promoter exchange) can be achieved, which make the use of this strategy more appealing for industrial purposes. All together, these techniques make a powerful toolbox for modification and construction of new microorganisms. The most impressive example of the power of molecular biology is the recent construction of the first synthetic genome of *Mycoplasma genitalum*[109]. These developments are part of a fast-moving, largely technology-driven field, the field of Synthetic Biology, which greatly speeds up design and engineering of microorganisms for industrial purposes[110].

However, targeted genetic modification of microorganisms by recombinant-DNA technology is not the only tool that metabolic engineers have to obtain microorganisms with desired properties. The use of evolution to select for a desired phenotype, by imposing a selection pressure when culturing microorganisms, is a principle used since the beginning of animal domestication and their breeding. When applied to microorganisms, this approach is known as evolutionary engineering and is a powerful tool for the selection of improved phenotypes[111]. Nevertheless, evolutionary engineering does not, by itself, allow the rapid transfer of the improved phenotype to other strains and, without detailed analysis of the underlying mutations, is essentially a black-box approach. Recent advances in DNA sequencing technology and rapidly decreasing costs for

sequencing whole genomes have increased the capacity to unravel the genetic differences responsible for a desired phenotype. Once these differences have been identified and confirmed, they can be transferred to new strains. This process is known as reverse metabolic engineering[112].

Metabolic engineering in redox metabolism in *S. cerevisiae*

For many years, reduction or elimination of glycerol as an inevitable by-product has been a priority in yeast metabolic engineering[69,96,113-116]. Inactivation of glycerol formation, such as by the deletion of both genes encoding glycerol-3-phosphate dehydrogenase (*GPD1* and *GPD2*) from the *S. cerevisiae* genome, forces anaerobically grown yeast cells to use alternative routes for reoxidizing the excess biosynthetic NADH[49]. In aerobic batch experiments, a *gpd1*Δ *gpd2*Δ strain can use oxygen as electron acceptor, but grows at half the maximum specific growth rate of a *GPD1 GPD2* strain and shows a changed ratio between respiratory and fermentative metabolism[117]. Anaerobically, *gpd1*Δ *gpd2*Δ strains cannot grow in the absence of externally added alternative redox sinks, such as acetoin. However, genetic modifications that introduce heterologous reaction(s) that enable (net) NADH oxidation in the metabolism of the engineered yeast can also restore growth. Examples of this are the production of sorbitol[118], mannitol[118], or 1,2 propanediol[119] as reduced products through heterologous expression of different dehydrogenase enzymes that can reoxidize NADH. However, these strategies lead to the formation of by-products other than glycerol and do not always increase ethanol yield. Additionally, some of these reduced products can also act as compatible solutes against osmotic stress, though they do not provide the same protective effect as glycerol accumulation[118].

An alternative strategy to decrease the impact of glycerol formation on the ethanol yield is the fine-tuning of its formation. This idea has been studied using the replacement of *GPD1* or *GPD2* promoter region with constitutive, weak versions of the promoter *TEF1* in *S. cerevisiae* strains[120], while its paralog was deleted[121]. These studies resulted in a significant decrease of glycerol formation, without a strong effect on the growth rate of the constructed strains, but also did not demonstrate a significant increase in the ethanol yield. Nevertheless, the most important advantage of this intermediate approach is that osmotolerance of *S. cerevisiae* is not affected as the cells are still able to form and use glycerol as compatible solute[122].

A third strategy to decrease glycerol formation relies on changing the co-factor specificity of any reaction(s) of yeast metabolism that results in the replacement of an NADPH-oxidizing reaction by an NADH-oxidizing reaction or by replacing an NADH-forming reaction by an NADPH-forming reaction. This is

also known as redox cofactor engineering[123-125]. A good example of this strategy is the work done by Nissen and collaborators (2000) on redox metabolism of ammonia assimilation in *S. cerevisiae*[69]. In their work, the predominantly NADPH-dependent reductive amination of 2-oxoglutarate to glutamate, catalyzed by $NADP^+$-dependent glutamate dehydrogenase, was replaced by an alternative pathway that oxidized NADH and consumed ATP. Their final result was an increase by 10% of the ethanol yield and a reduction by 38% of the glycerol yield on glucose. However, use of this elegant strategy is limited to media containing ammonia as nitrogen source[69]. A second example of redox cofactor engineering is the engineering of the co-factor specificity (NADPH instead of NADH) of the oxidation of glyceraldehyde 3-phosphate (GAP) in glycolysis. Depending on the enzyme used, the non-phosphorylating $NADP^+$-dependent glyceraldehyde-3-phosphate dehydrogenase (GAPN) or the $NADP^+$-dependent glyceraldehyde-3-phosphate dehydrogenase (GAPDH), GAP is reduced towards 3-phosphoglycerate or 1,3-bisphosphoglycerate, which results in a glycolytic net-ATP formation of zero or two moles respectively. Heterologous expression of GAPN gene from *Bacillus cereus* in a *gpd1*⊿ background together with the overexpression of the genes that encode the first step in trehalose formation led to an increase of the ethanol yield (by *ca.* 8%), and a significant decrease of the glycerol yield (49%)[126].

Unsuccessful attempts to reduce glycerol formation have also helped to understand redox metabolism. One example is the expression of a functional transhydrogenase in *S. cerevisiae*, an enzyme that catalyzes the interconversion of NADH and $NADP^+$ to NADPH and NAD^+. In nature, several microorganisms use this enzyme to interconnect the levels of their redox equivalents, but *S. cerevisiae* does not harbor such an enzyme. When a heterologous transhydrogenase was expressed in *S. cerevisiae*, under the prevailing intracellular concentrations thermodynamically favourable, the reaction worked in the opposite direction ($NADPH + NAD^+ \rightarrow NADH + NADP^+$), leading to an increased 'excess' of NADH in anaerobic conditions and consequently to increased formation of glycerol[127-129]. Even if a free-energy coupled transhydrogenase reaction could drive the desired reaction, this would only provide a partial solution, since the amount of NADH generated is much larger than the NADPH that is required in biosynthesis[35]. An additional study also made a connection between NADH and NADPH[130], by overexpressing the *S. cerevisiae POS5* gene, encoding a mitochondrial NADH kinase ($NADH + ATP \rightarrow NADPH$), and studying product formation of the resulting strain in batch cultures. Even though these authors claim an increased final ethanol concentration in aerobic batches, physically converting the conserved moiety NADH into NADPH, which fundamentally differs from transferring the electrons of NADH to $NADP^+$, can by

itself never solve the stoichiometric problem of excess NADH from biosynthesis. The observation of increased ethanol concentrations therefore seem an artifact of their experimental set-up or side-effect of their genetic intervention. Another study attempted to decrease intracellular NADH availability by heterologous expression of different types of NADH-oxidase ($NADH + ½\ O_2 + H^+ \rightarrow NAD^+ + H_2O$)[131]. Obviously, such oxygen-dependent systems have little relevance for large-scale anaerobic bioethanol production. Finally, also related to redox metabolism engineering although not desirable in alcoholic fermentation, increasing the production of glycerol has been studied by providing the cells with additional 'excess' NADH, for instance through the oxidation of formate by formate dehydrogenase ($formate + NAD^+ \rightarrow CO_2 + NADH + H^+$)[132,133].

The examples discussed above illustrate that engineering of redox metabolism is a powerful tool to modify the direction of metabolic fluxes towards a desired product, and that a consideration of the effects of redox cofactor balance must be taken into account in the design of metabolic engineering strategies. As a research objective in industrial biotechnology, the exploration of new metabolic engineering strategies that attempt complete elimination of glycerol formation and maximize metabolic fluxes towards ethanol, or any other product of interest, remains important.

Scope of this thesis

Redox metabolism plays a crucial role in (by)-product formation in *Saccharomyces cerevisiae*. This is exemplified by the formation of glycerol as the main by-product in alcoholic fermentation. The requirement for redox cofactor balancing in anaerobic fermentation can also be used for redox-driven formation of products such as lactic acid and/or succinic acid. Increased flexibility in the redox metabolism of *S. cerevisiae* would even further enhance the potential of this already versatile microbial cell factory.

After a general introduction to industrial biotechnology and *S. cerevisiae*, **Chapter 1** describes the concept of metabolic engineering and how this has been applied to eliminate glycerol formation in alcoholic fermentation. Although deletion of *GPD1* and *GPD2* results in *S. cerevisiae* strains that no longer produce glycerol, these strains can only grow anaerobically in the presence of an alternative electron acceptor, such as acetoin. However, addition of acetoin is not a viable option for economically competitive industrial alcoholic fermentation. Interestingly, acetate, which is more oxidized than ethanol, is abundantly available in lignocellulosic hydrolysates, which are the desired feedstocks for 2nd generation alcoholic fermentation. Although the reduction of acetic acid to ethanol is theoretically

possible, the reaction is thermodynamically unfavorable due to the low acetic acid concentration resulting from the pK_a of acetate and the near neutral intracellular pH. Some microorganisms other than *S. cerevisiae* are able to use acetate as electron acceptor, by first converting it to acetyl-CoA and investing 2 moles of ATP per mol of acetyl-CoA in this process, and by subsequently reducing acetyl-CoA towards ethanol in 2 NADH-oxidizing enzyme-catalyzed reactions. Only one of these reactions, catalyzed by an (acetylating) acetaldehyde dehydrogenase, is not present in *S. cerevisiae*. In **Chapter 2**, the use of acetic acid as external electron acceptor by a *gpd1*Δ *gpd2*Δ *S. cerevisiae* strain that expresses an (acetylating) acetaldehyde dehydrogenase (encoded by the *Escherichia coli mhpF* gene) is investigated. After a theoretical analysis of the impact of this strategy on the ethanol yield on glucose, this strategy was tested *in vivo*.

Glycerol production by *S. cerevisiae* is not only important for the maintenance of the anaerobic redox-cofactor balance, but also protects yeasts against high osmotic pressure. Complete elimination of glycerol formation might therefore negatively affect the fermentation characteristics under process conditions, where high initial sugar concentrations create an environment with high osmotic pressure. In **Chapter 3**, evolutionary engineering of a *gpd1*Δ *gpd2*Δ *S. cerevisiae* strain expressing an (acetylating) acetaldehyde dehydrogenase for anaerobic growth at high sugar concentrations (1 molar [M] glucose) using sequential batch reactors (SBRs) is described. In the selection for relevant phenotypes in industrial biotechnology, the use of reverse metabolic engineering is of great importance not only for the transfer of desired phenotypes, but also for the elimination of un-wanted phenotypes. During the selection for an osmotolerant *gpd1*Δ *gpd2*Δ *S. cerevisiae* strain, a strong flocculation phenotype was observed. Flocculation and/or cell aggregation regularly occurs during evolutionary engineering experiments using SBRs. Although flocculation can be a desired phenotype in beer brewing, where it facilitates separation of the yeast, it is not desirable during evolutionary engineering, where it diverts the selective pressure away from the target of interest and furthermore complicates the selection of single colony isolate cell lines. In **Chapter 4**, the reverse engineering of a cell aggregation phenotype observed in a *S. cerevisiae* strain evolved for faster galactose consumption is described. The knowledge obtained from this work can be used to optimize future evolutionary engineering experiments or to repair strains that display an aggregating phenotype.

The redirection of carbon and a more efficient use of substrate is a permanent metabolic engineering objective. Carbon dioxide is an abundant product of anaerobic and aerobic sugar metabolism of yeasts and other microorganisms. An hitherto unexplored metabolic engineering strategy is the use of carbon dioxide as

electron acceptor in yeast by the heterologous expression of Calvin Cycle enzymes, phosphoribulokinase (PRK) and ribulose-1,5-bis-phosphate carboxylase/oxygenase (Rubisco). **Chapter 5** describes the use of state-of-the-art synthetic biology methods to investigate the physiological impact of this strategy and the requirement for co-expression of bacterial chaperones for functional expression of Rubisco in yeast. The impact of the use of CO_2 as electron acceptor for the reoxidation of excess NADH from biosynthetic reactions was subsequently studied during alcoholic fermentation of *S. cerevisiae*. This new strategy paves the way for more efficient use of carbon in yeast metabolism, creates possibilities for the use of CO_2 fixation in other metabolic engineering strategies and contributes to the metabolic engineering toolbox with the use of chaperone proteins for functional expression of bacterial enzymes in the cytosol of eukaryotic microbial cell factory hosts.

Chapter 2

Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor

Víctor Guadalupe Medina , Marinka J. H. Almering, Antonius J. A. van Maris and Jack T. Pronk

Published in Applied and Environmental Microbiology, January 2010, Vol. 76, No. 1, p. 190-195 (doi:10.1128/AEM.01772-09)

Abstract

In anaerobic cultures of wild-type *Saccharomyces cerevisiae*, glycerol production is essential to reoxidize NADH produced in biosynthetic processes. Consequently, glycerol is a major by-product during anaerobic production of ethanol by *S. cerevisiae*, the single largest fermentation process in industrial biotechnology. The present study investigates the possibility of completely eliminating glycerol production by engineering *S. cerevisiae* such that it can reoxidize NADH by the reduction of acetic acid to ethanol via NADH-dependent reactions. Acetic acid is available at significant amounts in lignocellulosic hydrolysates of agricultural residues. Consistent with earlier studies, deletion of the two genes encoding NAD^+-dependent glycerol-3-phosphate dehydrogenase (*GPD1* and *GPD2*) led to elimination of glycerol production and an inability to grow anaerobically. However, when the *E. coli mhpF* gene, encoding the acetylating NAD^+-dependent acetaldehyde dehydrogenase (EC 1.2.1.10; acetaldehyde + NAD^+ + coenzyme A ↔ acetyl coenzyme A + NADH + H^+), was expressed in the *gpd1*Δ *gpd2*Δ strain, anaerobic growth was restored by supplementation with 2.0 g l^{-1} acetic acid. The stoichiometry of acetate consumption and growth was consistent with the complete replacement of glycerol formation by acetate reduction to ethanol as the mechanism for NADH reoxidation. This study provides a proof of principle for the potential of this metabolic engineering strategy to improve ethanol yields, eliminate glycerol production, and partially convert acetate, which is a well-known inhibitor of yeast performance in lignocellulosic hydrolysates, to ethanol. Further research should address the kinetic aspects of acetate reduction and the effect of the elimination of glycerol production on cellular robustness (*e.g.*, osmotolerance).

Introduction

Bioethanol production by *Saccharomyces cerevisiae* is currently, by volume, the single largest fermentation process in industrial biotechnology. A global research effort is under way to expand the substrate range of *S. cerevisiae* to include lignocellulosic hydrolysates of non-food feedstocks (*e.g.* energy crops and agricultural residues) and to increase productivity, robustness and product yield (for reviews see [20,115]). A major challenge relating to the stoichiometry of yeast-based ethanol production is that substantial amounts of glycerol are invariably formed as a by-product[134]. It has been estimated that, in typical industrial ethanol processes, up to 4% of the sugar feedstock is converted into glycerol[134]. Although glycerol also serves as a compatible solute at high extracellular osmolarity[135], glycerol production under anaerobic conditions is primarily linked to redox metabolism[34].

During anaerobic growth of *S. cerevisiae*, sugar dissimilation occurs via alcoholic fermentation. In this process, the NADH formed in the glycolytic glyceraldehyde-3-phosphate dehydrogenase reaction is reoxidized by converting acetaldehyde, formed by decarboxylation of pyruvate to ethanol via NAD^+-dependent alcohol dehydrogenase. The fixed stoichiometry of this redox-neutral dissimilatory pathway causes problems when a net reduction of NAD^+ to NADH occurs elsewhere in metabolism. Such a net production of NADH occurs in assimilation when yeast biomass is synthesized from glucose and ammonia[34]. Under anaerobic conditions, NADH reoxidation in *S. cerevisiae* is strictly dependent on reduction of sugar to glycerol[34]. Glycerol formation is initiated by reduction of the glycolytic intermediate dihydroxyacetone phosphate to glycerol-3-phosphate, a reaction catalyzed by NAD^+-dependent glycerol-3-phosphate dehydrogenase. Subsequently, the glycerol-3-phosphate formed in this reaction is hydrolyzed by glycerol-3-phosphatase to yield glycerol and inorganic phosphate.

The importance of glycerol production for fermentative growth of yeasts was already observed in the 1960's during studies on non-*Saccharomyces* yeasts that exhibit a so-called "Custers effect". In such yeast species, which are naturally unable to produce glycerol, fermentative growth on glucose is only possible in the presence of an external electron acceptor that can be reduced via an NADH-dependent reaction (*e.g.* the reduction of acetoin to butanediol via NAD^+-dependent butanediol dehydrogenase[136]). It was later shown that *gpd1*Δ *gpd2*Δ strains of *S. cerevisiae*, which are also unable to produce glycerol, are similarly unable to grow under anaerobic conditions unless provided with acetoin as an external electron acceptor[117].

In view of its large economic significance, several metabolic engineering strategies have been explored to reduce or eliminate glycerol production in

anaerobic cultures of *S. cerevisiae*. Nissen *et al.*[69] changed the cofactor specificity of glutamate dehydrogenase, the major ammonia fixing enzyme of *S. cerevisiae*, thereby increasing NADH consumption in biosynthesis. This approach significantly reduced glycerol production in anaerobic cultures grown with ammonia as the nitrogen source. Attempts to further reduce glycerol production by expression of a heterologous transhydrogenase, with the aim to convert NADH and $NADP^+$ into NAD^+ and NADPH, were unsuccessful[134] because intracellular concentrations of these pyridine nucleotide cofactor couples favor the reverse reaction[129].

The goal of the present study is to investigate whether the engineering of a linear pathway for the NADH dependent reduction of acetic acid to ethanol can replace glycerol formation as a redox sink in anaerobic, glucose-grown cultures of *S. cerevisiae* and thus provides a stoichiometric basis for elimination of glycerol production during industrial ethanol production. Significant amounts of acetic acid are released upon hydrolysis of lignocellulosic biomass and, in fact, acetic acid is studied as an inhibitor of yeast metabolism in lignocellulosic hydrolysates[63,81,83]. The *S. cerevisiae* genome already contains genes encoding acetyl-coenzyme A synthetase[137] and NAD^+-dependent alcohol dehydrogenases (ADH1-5[138]). To complete the linear pathway for acetic acid reduction, we expressed an NAD^+-dependent (acetylating) acetaldehyde dehydrogenase (EC 1.2.1.10) from *Escherichia coli* into a *gpd1*Δ *gpd2*Δ strain of *S. cerevisiae*. This enzyme, encoded by the *E. coli mhpF* gene[139], catalyzes the reaction acetaldehyde + NAD^+ + coenzyme A ↔ acetyl coenzyme A + NADH + H^+. Growth and product formation of the engineered strain were then compared in the presence and absence of acetic acid and compared to that of a congenic reference strain.

Materials and Methods

Strains construction and maintenance. The *Saccharomyces cerevisiae* strains used in this study (Table 2.1) originate from the CEN.PK family, which was previously identified as a suitable background for combined genetic and physiological studies[140]. Strain RWB0094, carrying deletions in the open reading frames of the *GPD1* and *GPD2* genes of strain CEN.PK102-3A (*MAT***a** *ura3 leu2*) were replaced by the loxP-*KanMX*-loxP cassette from pUG6[105] and the *hphMX4* cassette from pAG32[141], respectively, was acquired from BIRD Engineering, Rotterdam, The Netherlands. The KanMX marker of strain RWB0094 was removed by expression of the Cre recombinase[105] and its leucine auxotrophy was complemented by transformation with the *LEU2*-bearing plasmid YEPlac181[142], yielding strain IMZ008. Transformation of strain IMZ008 with the *URA3*-bearing *mhpF* expression plasmid pUDE43 (see below) yielded the prototrophic, *mhpF*-expressing

strain IMZ132, transformation with the *URA3*-bearing 'empty' vector p426_GPD yielded strain IMZ127. Finally, transformation of strain CEN.PK113-5D (*ura3-53*) with p426_GPD yielded the prototrophic *GPD1 GPD2* reference strain IME076. Cultures transformed with deletion cassettes were plated on YPD complex medium[143] containing G418 (200 mg l^{-1}) or hygromycin (200 mg l^{-1}). Successful integration of the deletion cassettes was confirmed by diagnostic PCR.

Stock cultures of all strains were grown in shake flasks containing 100 ml of synthetic medium (see below) with 20 g l^{-1} glucose as the carbon source. After adding 30% (v/v) glycerol, 1-ml aliquots of stationary phase cultures were stored at -80 °C.

Table 2.1 *Saccharomyces cerevisiae* strains used in this study.

Strain	Relevant genotype	Source/Reference
CEN.PK113-5D	*MAT***a** *ura3 GPD1 GPD2*	P. Kötter, Frankfurt
IME076	*MAT***a** *ura3 GPD1 GPD2* p426_GPD(*URA3*)	This study
CEN.PK102-3A	*MAT***a** *ura3 leu2 GPD1 GPD2*	P. Kötter, Frankfurt
RWB0094	*MAT***a** *ura3 leu2 gpd1*(-1,1133)::loxP-*KanMX*-loxP *gpd2*(-2,1281)::*hphMX4*	Bird Engineering, Rotterdam
IMZ008	*MAT***a** *ura3 leu2 gpd1*(-1,1133)::loxP-*KanMX*-loxP *gpd2*(-2,1281)::*hphMX4* YEplac181(*LEU2*)	This study
IMZ132	*MAT***a** *ura3 leu2 gpd1*(-1,1133)::loxP-*KanMX*-loxP *gpd2*(-2,1281)::*hphMX4* YEplac181(*LEU2*) pUDE43(*URA3* pTDH3::*mhpF* (*E. coli*)::CYC1t)	This study
IMZ127	*MAT***a** *ura3 leu2 gpd1*(-1,1133)::loxP-*KanMX*-loxP *gpd2*(-2,1281)::*hphMX4* YEplac181(*LEU2*) p426_GPD(*URA3*)	This study

Plasmid construction. The *E. coli mhpF* gene (EMBL accession number Y09555.7[139]) was PCR amplified from *E. coli* K12 strain JM109 genomic DNA using primer pairs *mhpF*-FW (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAGTAAGCGTAAAGTC GCCATTATCGG -3') and *mhpF*-RV (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTGTTCATGCCGCTTCTCCTG CCTTGC-3'), which contained attB1 and attB2 sequences, respectively. The polymerase chain reaction[141] was performed using Phusion Hot Start high-fidelity DNA polymerase (Finnzymes Oy, Espoo, Finland) according to manufacturer specifications and in a Biometra TGradient thermocycler (Biometra, Göttingen, Germany) with the following settings: 25 cycles of 10 s denaturation at 98 °C and

30 s annealing and extension at 72 °C. The 1011-bp PCR product was cloned using Gateway cloning technology (Invitrogen, Carlsbad, CA, USA). Plasmid pDONR221, using the BP reaction, was used to create the entry clone, designated as plasmid pUD64. From this entry clone and the multicopy plasmid pAG426GPD-ccdB (Addgene, Cambridge, MA, USA) the yeast expression plasmid pUDE43, was constructed employing the LR reaction. Transformations of recombination reaction products into competent *E. coli* K12 strain JM109 were performed according to the Z-Competent *E. coli* Transformation Kit (Zymoresearch Corporation, Orange, USA) and plated on LB media containing either ampicillin (100 mg l^{-1}) or kanamycin (50 mg l^{-1}). Yeast transformations were performed according to Burke *et al.*[143]. After transformations with the yeast expression plasmid, cells were plated on synthetic media. Successful insertion of multicopy plasmid pUDE43 was confirmed by diagnostic PCR using the primer pairs for cloning.

Cultivation and media. Shake-flask cultivation was performed at 30 °C in a synthetic medium[35]. The pH of the medium was adjusted to 6.0 with 2 M KOH prior to sterilization. Precultures were prepared by inoculating 100 ml medium containing 20 g l^{-1} glucose in a 500-ml shake-flask with 1 ml frozen stock culture. After 24 h incubation at 30 ºC in an Innova® incubator shaker (200 rpm; New Brunswick Scientific, NJ), cultures were transferred to bioreactors.

Anaerobic batch fermentations were carried out at 30 °C in 2-liter laboratory fermenters (Applikon, Schiedam, The Netherlands) with a working volume of 1 liter. Synthetic medium with 20 g l^{-1} glucose[35] was used for all fermentations and supplemented with 100 µl l^{-1} of silicone antifoam (Silcolapse 5020; Caldic Belgium, Bluestar Silicones) as well as with the anaerobic growth factors, ergosterol (0.01 g l^{-1}) and Tween 80 (0.42 g l^{-1}) dissolved in ethanol. This resulted in 0.5 to 0.6 g l^{-1} ethanol in the medium. Where indicated, acetic acid was added at a concentration of 2 g l^{-1} and the pH was readjusted to 5.0 prior to inoculation. Culture pH was maintained at 5.0 by the automatic addition of 2 M KOH. Cultures were stirred at 800 rpm and sparged with 0.5 liter min^{-1} nitrogen (<10 ppm oxygen). Dissolved oxygen was monitored with an autoclavable oxygen electrode (Applisens, Schiedam, The Netherlands). To minimize diffusion of oxygen, fermenters were equipped with Norprene tubing (Cole Palmer Instrument Company, Vernon Hills). All fermentations were carried out at least in duplicate.

Determination of culture dry weight and optical density. Culture samples (10 ml) at selected time intervals were filtered over pre-weighed nitrocellulose filters (pore size 0.45 µm; Gelman Laboratory, Ann Arbor, USA). After removal of medium the filters were washed with demineralized water and dried in a microwave oven (Bosch, Stuttgart, Germany) for 20 min at 350 W and

weighed. Duplicate determinations varied by less than 1%. Culture growth was also monitored via optical density readings at a wavelength of 660 nm on a Novaspec II spectrophotometer.

Gas analysis. Exhaust gas was cooled in a condenser (2 ºC) and dried with a Permapure dryer type MD-110-48P-4 (Permapure, Toms River, NJ). Oxygen and carbon dioxide concentrations were determined with a NGA 2000 analyzer (Rosemount Analytical, Orrville, OH). Exhaust gas-flow rate and carbon dioxide production rates were determined as described previously[144]. In calculating these biomass-specific rates, a correction was made for volume changes caused by withdrawing culture samples.

Metabolite analysis. Supernatant obtained by centrifugation of culture samples was analyzed for glucose, acetic acid, succinic acid, lactic acid, glycerol and ethanol via HPLC analysis on a Waters Alliance 2690 HPLC (Waters, Milford, MA) containing a Biorad HPX 87H column (Biorad, Hercules, USA). The column was eluted at 60 ºC with 0.5 g l^{-1} H_2SO_4 at a flow rate of 0.6 ml min^{-1}. Detection was by means of a Waters 2410 refractive-index detector and a Waters 2487 UV detector. Initial and final glycerol concentrations were further determined using an enzymatic determination kit (R-Biopharm AG, Darmstadt, Germany). During cultivation in bioreactors that are sparged with nitrogen gas, a significant fraction of the ethanol is lost through the off gas[89]. To correct for this, ethanol evaporation kinetics were analyzed in bioreactors operated under identical conditions at different working volumes with sterile synthetic medium. The resulting volume-dependent ethanol evaporation constants (for this set-up equal to 0.008 divided by the volume in liters, expressed in h^{-1}) were used to correct HPLC measurements of ethanol concentrations in culture supernatants, taking into account changes in volume that were caused by sampling.

Enzyme activity assays. Cell extracts for activity assays of NAD^+-dependent (acetylating) acetaldehyde dehydrogenase were prepared from exponentially growing anaerobic batch cultures and analyzed for protein content as described previously[145]. NAD^+-dependent (acetylating) acetaldehyde dehydrogenase activity was measured at 30 °C by monitoring the oxidation of NADH at 340 nm. The reaction mixture (total volume 1 ml) contained 50 mM potassium phosphate buffer (pH 7.5), 0.15 mM NADH and cell extract. The reaction was started by addition of 0.5 mM acetyl coenzyme A. For glycerol-3-phosphate dehydrogenase[146] activity determination, cell extracts were prepared as described above except that the phosphate buffer was replaced by triethanolamine buffer (10 mM, pH 5)[132,146]. Glycerol-3-phosphate dehydrogenase activities were assayed in cell extracts at 30 ºC

as described previously[76]. Reaction rates were proportional to the amounts of cell extract added.

Results

Theoretical analysis of the introduction of a linear acetate reduction pathway. For a theoretical prediction of the impact of the introduction of a linear acetate reduction pathway into a *gpd1*⊿ *gpd2*⊿ strain of *S. cerevisiae*, we used published data on the yields of biomass, ethanol, carbon dioxide and acetate in glucose-grown anaerobic batch cultures of the reference strain *S. cerevisiae* CEN.PK113-7D[89]. According to these data, the anaerobic growth stoichiometry of this strain is described by equation (1):

56 mmol glucose → 1 g biomass + 88 mmol ethanol + 95 mmol CO_2 + 11 mmol glycerol + 1.7 mmol acetate (1)

In this scenario, 10 C-mol% of the glucose is converted to glycerol and 52 C-mol% to ethanol. These yields are higher and lower, respectively, than predicted values for large scale anaerobic ethanol production[69]. A number of factors may contribute to this difference, including the often lower biomass yield under industrial conditions, the identity of the nitrogen source[44], a redox-sparing effect of acetic acid in hydrolysates[147] and the presence of alternative electron acceptors such as furfural in hydrolysates[148].

Production of glycerol for redox balancing occurs according to equation (2):

0.5 glucose + NADH + H^+ + ATP → glycerol + NAD^+ + ADP + P_i (2)

The metabolic engineering strategy in the present study is based on the replacement of this reaction by the reaction:

Acetic acid + 2 NADH + 2 H^+ + ATP → ethanol + 2 NAD^+ + AMP + PP_i (3)

If it is assumed that the formation of AMP and pyrophosphate in the acetyl-CoA synthetase reaction is equivalent to the hydrolysis of 2 ATP to ADP and inorganic phosphate due to hydrolysis of pyrophosphate to orthophosphate[149], the formation of one mole of glycerol from glucose is equivalent to the reduction of 0.5 mol of acetate to ethanol, in terms of NADH oxidation as well as in terms of ATP hydrolysis. If it is furthermore assumed that the acetate produced by the reference

strain is reconsumed (equivalent to not formed), the growth stoichiometry can be rewritten as:

50 mmol glucose + 3.9 mmol acetate → 1 g biomass + 93 mmol ethanol + 95 mmol CO_2 (4)

In this new situation, the glycerol yield has decreased to zero, while the apparent ethanol yield on glucose has increased to 62 C-mol%, which represents a theoretical 18% increase relative to the ethanol yield of the reference strain grown on glucose as the sole carbon source.

Growth and product formation in anaerobic batch cultures. When cultures of the prototrophic reference strain *S. cerevisiae* IME076 (*GPD1 GPD2*) were supplemented with 2.0 g l^{-1} acetic acid, the specific growth rate (0.32 h^{-1}) was identical to that reported for cultures grown in the absence of acetic acid (0.34 h^{-1})[89]. Consistent with results from a recent study, the addition of acetic acid led to a slight decrease of the biomass yield and, consequently a decrease of the glycerol yield on glucose relative to cultures grown in the absence of acetic acid (Figure 2.1, Table 2.2). This effect has been attributed to the higher rate of glucose dissimilation for intracellular pH homeostasis due to diffusion of acetic acid into the cell, which in turn results in a lower biomass yield on glucose[81]. Under the same conditions, an isogenic *gpd1*Δ *gpd2*Δ strain, in which absence of NAD^+-dependent glycerol-3-phosphate dehydrogenase activity was confirmed in cell extracts (Table 2.2), was completely unable to grow anaerobically (data not shown), consistent with the notion that glycerol production via Gpd1 and Gpd2 is essential for NADH reoxidation in anaerobic cultures of *S. cerevisiae*[117].

Expression of the *E. coli mhpF* gene in a *gpd1*Δ *gpd2*Δ strain, resulting in acetyl-CoA dependent rates of NADH reduction in cell extracts of 0.020 μmol min^{-1} (mg protein)$^{-1}$ (Table 2.2), did not enable anaerobic growth when glucose was the sole carbon source. However, when the medium was supplemented with 2.0 g l^{-1} acetic acid, either before or after inoculation, exponential growth was observed at a specific growth rate of 0.14 h^{-1}. No formation of glycerol occurred during cultivation. (Figure 2.1, Table 2.2). The trace amounts (< 0.1 g l^{-1}) of glycerol present in cultures of *gpd1*Δ *gpd2*Δ strains originate from the inoculum cultures which were started from frozen glycerol stocks. Ethanol was the major organic product and the small amounts of succinate and lactate produced were similar to those observed in cultures of the reference strain grown under the same conditions (data not shown). Acetate addition did not rescue the anaerobic growth defect of a

congenic *gpd1*⊿ *gpd2*⊿ *S. cerevisiae* reference strain that did not express the *E. coli mhpF* gene (data not shown).

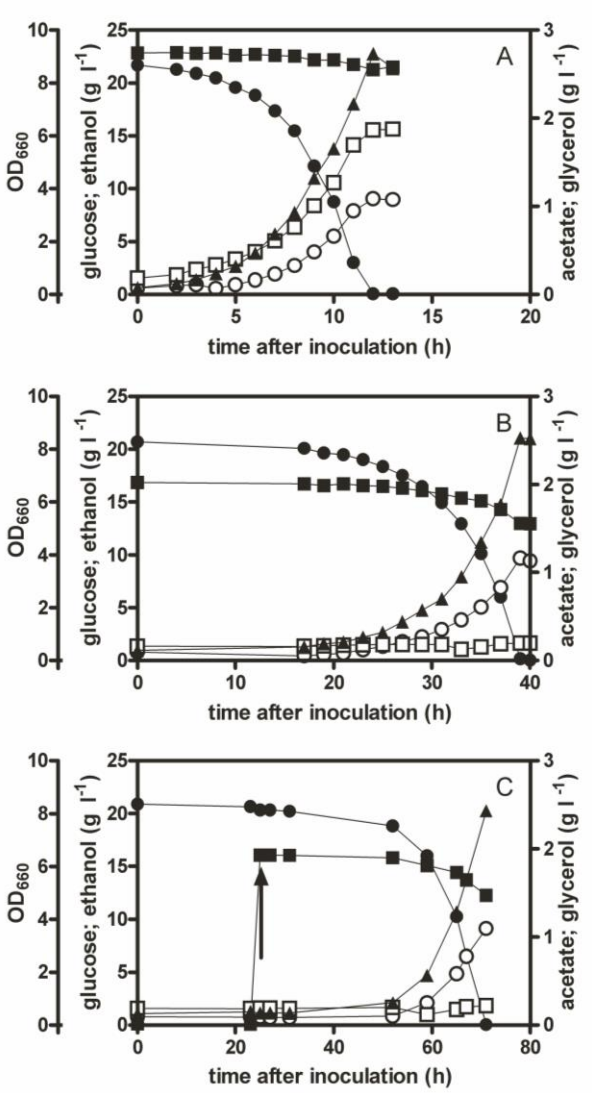

Figure 2.1 Concentrations of biomass and products in anaerobic batch cultures of different *S. cerevisiae* strains on glucose (20 g l^{-1}). Acetic acid (2.0 g l^{-1}) was present from the start of the fermentation (panel A, B) or added at the time point indicated by the arrow (panel C). Growth conditions: T = 30 °C, pH 5.0. Symbols: ▲, optical density at 660 nm; ●, glucose; ○, ethanol; ■, acetate; □, glycerol. Each graph represents values from one of two independent replicates, yielding data that differed by less than 5%. Panel A: *S. cerevisiae* IME076 (*GPD1 GPD2*). Panel B: *S. cerevisiae* IMZ132 (*gpd1*⊿ *gpd2*⊿ overexpressing the *E. coli mhpF* gene). Panel C: *S. cerevisiae* IMZ132 (*gpd1*⊿ *gpd2*⊿ overexpressing the *E. coli mhpF* gene).

In contrast to the theoretical predictions described above, only a slight difference in ethanol yield on glucose was observed based on HPLC data (Table 2.2). However, since the IMZ132 fermentations (40 h) lasted longer than the wild type strain (15 h) and the anaerobic batch cultures were sparged with nitrogen gas, the fraction of ethanol lost through evaporation was higher for strain IMZ132. After determination of the kinetics of ethanol evaporation in sterile control experiments

and correction of the ethanol yields, a 13% higher apparent ethanol yield on glucose was shown for the engineered strain using the linear pathway for NADH dependent reduction of acetic acid to ethanol (Table 2.2).

Table 2.2 Physiology of the engineered *S. cerevisiae* strain IMZ132 and the empty-vector reference strain IME076 during anaerobic batch cultivation on synthetic medium (pH 5) with glucose-acetate mixtures[a].

Yeast strain			**IME076**	**IMZ132**
Relevant genotype			*GPD1 GPD2*	*gpd1⊿ gpd2⊿ + mhpF*
Glycerol-3-phosphate dehydrogenase (μmol mg protein^{-1} min^{-1})			0.034 ± 0.003	< 0.002
Acetaldehyde dehydrogenase (acetylating) (μmol mg protein^{-1} min^{-1})			< 0.002	0.020 ± 0.004
Specific growth rate (h^{-1})			0.32 ± 0.01	0.14 ± 0.01
	Biomass (g mol^{-1})		15.0 ± 0.1	14.8 ± 1.6
	Glycerol (mol mol^{-1})		0.14 ± 0.01	< 0.002
Yield on glucose	Ethanol (mol mol^{-1})	Not corrected for evaporation	1.54 ± 0.02	1.67 ± 0.04
		Corrected for evaporation	1.61 ± 0.03	1.82 ± 0.05

[a] Results are represented as average and mean deviations of data from two independent batch cultures. Glycerol yields on glucose were calculated using glycerol concentrations obtained by enzymatic determination.

Discussion

The present study provides a proof of principle that, stoichiometrically, the role of glycerol as a redox sink for anaerobic growth of *S. cerevisiae* can be fully replaced by a linear pathway for NADH-dependent reduction of acetate to ethanol. This offers interesting perspectives for large-scale ethanol production from feedstocks that contain acetic acid, such as lignocellulosic hydrolysates. The 'bacterial' pathway from acetyl-CoA to ethanol has previously been integrated into *S. cerevisiae* as part of a strategy to ferment xylose via a phosphoketolase pathway[150]. While this approach indeed resulted in the conversion of xylose, the impact on glycerol production was not evaluated and the pathway was not tested in a *gpd1⊿ gpd2⊿* strain.

In addition to reducing the organic carbon content of spent media and increasing the ethanol yield, the reduction of acetic acid to ethanol may at least partially alleviate acetate inhibition of yeast growth and metabolism, which is especially problematic at low pH and during the consumption of pentose sugars by engineered yeast strains[81]. However, before industrial implementation can be

contemplated, several issues remain to be addressed. First, growth and product formation in the engineered strain were significantly slower than in the reference strain. Several factors may contribute to this. First, the *in vivo* kinetics of acetyl coenzyme A synthetase may limit the rate of acetate reduction. In anaerobic, glucose-grown batch cultures, the high-affinity Acs1 isoenzyme is not expressed due to a combination of transcriptional repression and glucose catabolite inactivation[137,151]. The constitutively expressed Acs2 enzyme has a low affinity for acetate (K_m = *ca.* 20 mM[137]), which may limit *in vivo* rates of acetate activation. Despite its uncoupling effect, the intracellular acetic acid concentration can be low due to active export of acetic acid from yeast cells[152]. Obviously, the *in vivo* activity of the heterologous acetaldehyde dehydrogenase may also control the rate of acetate reduction.

A further factor that may affect growth kinetics is that glycerol-3-phosphate is a key precursor of glycerolipids. Although it has been demonstrated that, in the absence of Gpd1 and Gpd2, glycerolipids can be produced by reduction of the corresponding dihydroxyacetone esters[153], the kinetics of this process have not been studied in anaerobic cultures. Analysis and optimization of the kinetics of acetate reduction and growth by further metabolic engineering and/or evolutionary approaches is essential to make the rates of ethanol production of *gpd1*⊿ *gpd2*⊿ strains expressing the *E. coli mhpF* gene compatible with industrial production.

In addition to the kinetics of growth and ethanol formation, another important aspect of the physiology of the engineered strains requires further investigation. The role of glycerol as a compatible solute, which protects yeast cells at high extracellular osmolarity, is likely to be relevant in industrial fermentations with high initial sugar concentrations. The metabolic engineering strategy described in this paper provides an excellent experimental platform to analyze osmotic stress in anaerobic cultures unable to produce glycerol. Such research should, ultimately, address the question whether robust industrial yeast strains can be constructed that do not produce glycerol.

Acknowledgements

We are grateful to our colleagues Eline Huisjes and Bianca van Leeuwen for the determination of the ethanol evaporation kinetics. In addition, we thank Marijke Luttik, Erik de Hulster, Giang Huong Duong and Jeremiah Wright for experimental assistance. The PhD project of VGM is sponsored by The Netherlands Genomics Initiative through its support for the Kluyver Centre for Genomics of Industrial Fermentation.

Chapter 3

Evolutionary engineering of a glycerol-3-phosphate dehydrogenase-negative, acetate-reducing *Saccharomyces cerevisiae* strain enables anaerobic growth at high glucose concentrations

Víctor Guadalupe-Medina, Benjamin Metz, Bart Oud, Charlotte M. van der Graaf, Robert Mans, Jack T. Pronk and Antonius J.A. van Maris

Essentially as published in Microbial Biotechnology, September 2013 (doi:10.1111/1751-7915.12080)

Abstract

Glycerol production by *Saccharomyces cerevisiae*, which is required for redox-cofactor balancing in anaerobic cultures, causes a loss of sugar in industrial bioethanol production. Recently, glycerol formation in anaerobic *S. cerevisiae* cultures was eliminated by expressing *Escherichia coli* (acetylating) acetaldehyde dehydrogenase (encoded by *mhpF*) and simultaneously deleting the *GPD1* and *GPD2* genes encoding glycerol-3-phosphate dehydrogenase, thus coupling NADH reoxidation to reduction of acetate to ethanol. Gpd$^-$ strains are, however, sensitive to high sugar concentrations, which complicates industrial implementation of this metabolic engineering concept. In this study, laboratory evolution was used to improve osmotolerance of a Gpd$^-$ *mhpF*-expressing *S. cerevisiae* strain. Serial batch cultivation at increasing osmotic pressure enabled isolation of an evolved strain that grew anaerobically at 1 M glucose, at a specific growth rate of 0.12 h^{-1}. The evolved strain produced glycerol at low concentrations (0.64 $\pm$ 0.33 g l^{-1}). However, these glycerol concentrations were below 10% of those observed with a Gpd$^+$ reference strain and the ethanol yield on sugar of the evolved strains reached 92% of the theoretical maximum. Genetic analysis indicated that osmotolerance under aerobic conditions required a single dominant chromosomal mutation, and one further mutation in the plasmid-borne *mhpF* gene for anaerobic growth.

Introduction

Bioethanol production with *Saccharomyces cerevisiae* is the single largest fermentation process in industrial biotechnology with an annual global product volume of *ca.* 8.6×10^{10} litres[53]. This puts *S. cerevisiae* at the centre of a global research effort to improve its productivity, robustness under process conditions, substrate range and product yield[20]. Anaerobic fermentation of sugars to ethanol and CO_2 is a redox-neutral process. However, in anaerobic cultures of *S. cerevisiae*, an 'excess' of NADH is generated from biosynthetic reactions such as oxidative decarboxylations in amino-acid and lipid synthesis[34,36]. In anaerobic yeast cultures, this 'excess' NADH is reoxidized through glycerol formation via NADH-dependent reduction of dihydroxyacetone phosphate to glycerol-3-phosphate, which is subsequently dephosphorylated to glycerol. Glycerol production has been estimated to account for a loss of 4% of the consumed sugar in industrial ethanol production[69]. Under tightly controlled laboratory growth conditions, where biomass yields are typically higher than in industrial yeast fermentation processes, this percentage can be as high as 10%[35,69,154]. Elimination of glycerol formation via metabolic engineering strategies has therefore attracted significant interest[69,78,119,154].

In *S. cerevisiae*, deletion of the *GPD1* and *GPD2* genes encoding NAD^+-dependent glycerol-3-phosphate dehydrogenase (EC 1.1.1.8) eliminates glycerol formation[117]. However, such a double deletion also completely blocks growth under anaerobic conditions unless an external electron acceptor for NADH reoxidation, such as acetoin or acetaldehyde, is provided[49,117,136]. We recently proposed a metabolic engineering strategy for eliminating glycerol production in anaerobic *S. cerevisiae* cultures that is based on the use of acetic acid, a common inhibitor present in plant biomass hydrolysates, as electron acceptor[154]. This strategy encompasses expression of a NAD^+-dependent (acetylating) acetaldehyde dehydrogenase (EC 1.2.1.10) *mhpF* gene (EMBL: CAA70751) from *Escherichia coli* in a *gpd1*Δ *gpd2*Δ (Gpd^-) *S. cerevisiae* strain. After activation of acetate by *S. cerevisiae* acetyl coenzyme A synthetase[137], the resulting acetyl coenzyme A can be reduced to ethanol by the combined activity of the NAD^+-dependent (acetylating) acetaldehyde dehydrogenase and yeast alcohol dehydrogenases. Anaerobic growth of the resulting engineered yeast strain on glucose was coupled to acetate reduction, glycerol production was eliminated and the ethanol yield increased by 13% relative to that of a *GPD1 GPD2* (Gpd^+) reference strain[154].

Glycerol formation is not only crucial for redox balancing in anaerobic cultures of wild-type *S. cerevisiae* but, as its main compatible solute, is also required for osmotolerance. Osmotolerance is essential in industrial ethanol production due

to the high sugar concentrations present at the start of fermentation processes[46,76,78]. The response of *S. cerevisiae* to high osmolarity is regulated by the High-Osmolarity Glycerol (HOG) pathway and involves not only intracellular glycerol accumulation but also regulation of other stress-related genes[79]. The osmosensitivity of Gpd$^-$ strains of *S. cerevisiae*[49] can be partly alleviated by introduction of sorbitol-6-P-dehydrogenase and mannitol-1-P-dehydrogenase encoding genes. In such engineered strains, mannitol or sorbitol act as alternative compatible solutes, although growth rates are lower than in wild-type strains[118].

Evolutionary engineering is a powerful approach to select for microbial strains with industrially relevant traits. In evolutionary engineering, regimes for prolonged cultivation are designed such that a selective advantage is conferred to spontaneous mutants that express the trait of interest[111,112]. Evolutionary engineering not only generates strains with industrially relevant phenotypes, but subsequent analysis of molecular mechanisms responsible for their improved performance also enables their reverse engineering into non-evolved strains[111,112].

The goal of the present study was to investigate whether evolutionary engineering enables the isolation of osmotolerant mutants of Gpd$^-$ *S. cerevisiae* expressing an *E. coli* (acetylating) acetaldehyde dehydrogenase. The ability of these strains to grow anaerobically with acetic acid as electron acceptor makes it possible to specifically focus anaerobic evolutionary engineering experiments on improvement of osmotolerance. To this end, sequential batch cultivation of the engineered strains was performed under anaerobic conditions and osmotic pressures deemed relevant for industrial cultivation[155-157]. After prolonged cultivation under selective conditions, which involved glucose concentrations of up to 1 M, single-cell lines were isolated and characterized in anaerobic bioreactors.

Materials and Methods

Strain construction and maintenance. All *Saccharomyces cerevisiae* strains in this study (Table 3.1) originate from the CEN.PK family[140,158,159]. Stock cultures and precultures were grown as described previously[154]. *S. cerevisiae* IMK006, obtained by removing the *KanMX* marker from the *gpd1*Δ *gpd2*Δ strain RWB0094[154] by expression of Cre recombinase[105], was transformed with the *LEU2*-bearing plasmid pRS405, which was linearized with BstEII (NEB, Massachusetts, USA), yielding strain IMX031. Transformation of strain IMX031 with the *URA3*-bearing *mhpF*-expression plasmid pUDE43[154] yielded the prototrophic, Gpd$^-$ *mhpF*-expressing strain IMZ160. Plasmid(s) were isolated from *S. cerevisiae* IMZ333 with the Sigma GenElute plasmid miniprep kit (Sigma-Aldrich Chemie Gmbh, Munich, Germany) according to manufacturer's instructions. Plasmids were transformed into *E. coli*

One Shot TOP10 Z-competent cells (Invitrogen, Paisley, United Kingdom) and transformants were selected on LB medium plates containing ampicillin (100 mg l^{-1}). Restriction analysis of isolated plasmids was done with XmnI (Fermentas Gmbh, Germany). Plasmid sequencing was performed by BaseClear (Leiden, The Netherlands). *URA3*-bearing plasmids were cured from strain IMZ333 by growth in complex medium with 20 g l^{-1} glucose and subsequent selection on complex medium with 5-fluoro-orotic acid (5-FOA). A single colony was isolated on synthetic medium containing 20 g l^{-1} glucose, 5-FOA (1 g l^{-1}) and uracil and named IMS343. Strain IMJ004 was constructed by transforming strain IMS343 with the original pUDE43 plasmid. Strains IMZ380 and IMZ381, and IMJ005 and IMJ006 were constructed by transforming pUDE043ev1 and pUDE043ev2 into the unevolved parent strain IMX031 and the evolved plasmid-cured strain IMS343 respectively. Strain IMJ009 was constructed by transforming plasmid p426_GPD (*URA3*) into IMS343. Strains bearing plasmids with auxotrophic markers were plated on synthetic media[35,106] agar plates (1% w/v) using 20 g l^{-1} glucose as carbon source. Confirmation of correct genetic modification and transformations were performed as described earlier[154].

Table 3.1 *Saccharomyces cerevisiae* strains used in this study.

Strain	Relevant genotype/description	Source/reference
IME076	*MAT***a** *ura3 LEU2 GPD1 GPD2* p426_GPD(*TDH3*$_p$::*CYC1*$_t$ *URA3* 2μ)	Guadalupe-Medina *et al.* (2010)[154].
RWB0094	*MAT***a** *ura3 leu2 gpd1::loxP-KanMX-loxP gpd2::hphMX4*	BIRD Engineering, Rotterdam, Guadalupe-Medina *et al.* (2010)[154].
IMK006	*MAT***a** *ura3 leu2 gpd1::loxP gpd2⊥::hphMX4*	This study.
IMX031	*MAT***a** *ura3 leu2::LEU2*[pRS405] *gpd1::loxP gpd2::hphMX4*	This study.
IMZ132	*MAT***a** *ura3 leu2 gpd1::loxP gpd2::hphMX4* pUDE43(*TDH3*$_p$::*mhpF(E. coli)*::*CYC1*$_t$ *URA3* 2μ) YEplac181(*LEU2*)	Guadalupe-Medina *et al.* (2010)[154]
IMZ160	*MAT***a** *ura3 leu2::LEU2*[pRS405] *gpd1::loxP gpd2::hphMX4* pUDE43(*TDH3*$_p$::*mhpF(E. coli)*::*CYC1*$_t$ *URA3* 2μ)	This study.
IMZ333	IMZ160 evolved for anaerobic growth at 1 M glucose	This study.
IMS343	IMZ333 cured of plasmid	This study.
IMZ380	IMX031 with pUDE43ev1(*TDH3*$_p$:: *mhpF*(*E. coli*)::*CYC1*$_t$ *URA3* 2μ evolved)	This study.
IMZ381	IMX031 with pUDE43ev2(*TDH3*$_p$:: *mhpF*(*E. coli*)::*CYC1*$_t$ *URA3* 2μ evolved)	This study.
IMJ004	IMS343 with pUDE43(*TDH3*$_p$:: *mhpF*(*E. coli*)::*CYC1*$_t$ *URA3* 2μ)	This study.
IMJ005	IMS343 with pUDE43ev1(*TDH3*$_p$:: *mhpF*(*E. coli*)::*CYC1*$_t$ *URA3* 2μ evolved)	This study.
IMJ006	IMS343 with pUDE43ev2 (*TDH3*$_p$::*mhpF*(*E. coli*)::*CYC1*$_t$ *URA3* 2μ evolved)	This study.
IMJ009	IMS343 with p426_GPD(*TDH3*$_p$::*CYC1*$_t$ *URA3* 2μ)	This study.
IMK527	*MAT*α *ura3::loxP-kanMX-loxP leu2 gpd1::loxP gpd2::hphMX4*	This study.
IMD011	Diploid strain resulting from IMS343×IMK527	This study.
IMD012	Diploid strain resulting from IMZ333×IMK527	This study.

Shake flask cultivation. Shake flask cultivations were performed as described previously[154] using synthetic media[35]. For serial shake-flask cultivations, synthetic media with urea as the nitrogen source were used[35]. Glucose and sorbitol were autoclaved separately at 110 °C and sterile, 10-fold concentrated synthetic medium was added afterwards. Three parallel evolution experiments were performed by serial transfer in aerobic shake flasks. 1 ml from a shake flask pre-culture of IMZ160 was used to inoculate a first shake flask containing 1 M sorbitol. Serial transfer was done with 1 ml inocula from shake flasks that had reached stationary phase. During the evolution experiment, increasing concentrations of

sorbitol were used in 28 serial shake flasks: 4 at 1 M sorbitol, 8 at 1.5 M sorbitol, and 16 at 2 M sorbitol. At the end of the evolution experiment, a sample of the evolving population was stored at -80 °C.

Anaerobic shake-flask cultures for inoculum or characterization were incubated in a BactronX anaerobic chamber (Shell Lab, Oregon, USA) at 30 °C and 200 rpm (Heidolph Unimax 2010 shaker).

Sequential batch reactors (SBR) and batch characterizations. Anaerobic bioreactor batch cultures, off-gas and metabolite analysis, enzymatic glycerol determination, optical density readings, determination of dry weight and enzymatic activity measurements for NAD^+-dependent (acetylating) acetaldehyde dehydrogenase and glycerol-3-phosphate dehydrogenase were performed as described previously[154]. All fermentations were carried out at least in duplicate. To correct for ethanol evaporation during cultivation in nitrogen sparged bioreactors, evaporation kinetics were analyzed as described previously[154].

Sequencing batch reactors were operated as described previously[94]. The medium vessels were prepared by autoclaving 18 l of demineralized water containing sugar (1.11 M), and subsequently adding 2 l of 10-fold concentrated synthetic media containing acetic acid (20 g l^{-1}), antifoam (0.2 g l^{-1}, Emulsion C [Sigma-Aldrich, Zwijndrecht, The Netherlands]), ergosterol (0.1 g l^{-1}) and Tween 80 (4.2 g l^{-1}). The pH of the 10-fold concentrated synthetic medium containing acetic acid was adjusted to pH 4.8 with KOH before autoclaving. During sequential batch cultivation in bioreactors, carbon dioxide concentrations in the exhaust gas were used to determine the moment to start a new batch. A control routine was programmed in MFCS/win 3.0 (Sartorius AG, Göttingen, Germany) to initiate the switch to a new batch cycle. Fermenters were automatically emptied, leaving *ca.* 1.5 ml of remaining culture volume, and refilled when the CO_2 % in the off gas reached 1.2%. When growth accelerated after the first three cycles, this threshold was gradually increased to 3.2% CO_2.

Single colony isolates were obtained by streaking a sample taken from the sequential batch reactors on synthetic media agar plates (1% w/v) containing 1 M glucose as carbon source, 2 g l^{-1} acetic acid and anaerobic growth factors. The plates were placed under anaerobic environment in a BactronX anaerobic chamber (Shell Lab, Oregon, USA) and kept at 30 °C. After two transfers of single isolates to fresh agar plates, one colony was inoculated in 1 M glucose synthetic media for stock and named IMZ333. Before characterization in bioreactors, the evolved strains IMZ333 and IMJ006 were precultured anaerobically in synthetic media shake flasks with 1 M glucose as carbon source.

Spot assays experiments. Growth under high osmotic stress was assessed by spotting 5 µl of serial dilution of 10^6, 10^5, 10^4 cells ml^{-1} of exponentially growing cultures onto 0.1, 0.5, and/or 1 M glucose synthetic media agar plates (1% w/v). The plates were incubated at 30 °C under anaerobic and aerobic conditions for 7 days and pictures were taken at 3 and 7 days.

Backcrossing and sporulation. To enable crossing, *MAT***a** strain IMK006 was transformed with the marker gene KanMX using primers for the *ura3* locus (URA3-KanMXF TTCTTAACCCAACTGCACAGAACAAAAACCTGCAGGAAACGAAGATAA ATCCAGCTGAAGCTTCGTACGC and URA3-KanMXR AGCTCTAATTTGTGAGTTTAGTATACATGCATTTACTTATAATACAGTT TTCTTTAAACACGGCCGCATAG) before the mating was switched by transforming plasmid pHO[160,161] into this strain into this strain. The resulting diploid strain was sporulated yielding a *MAT*α strain IMK527. Sporulation was performed as described by Bahalul *et al.* (2010)[162]. Strains were inoculated in complex medium with 10 g l^{-1} acetate as carbon source. After incubation at 30 °C for 24 h, cultures were washed and resuspended in sporulation medium (20 g l^{-1} potassium acetate). After 48 h at 30 °C, spore formation was checked microscopically. Prior to dissection, a culture sample (1-ml) was incubated with 2 µl zymolyase (1000 U ml^{-1}) in a 200 µl 0.5M sorbitol solution at 37 °C for 10 min. Tetrad dissection on complex media plates with 20 g l^{-1} glucose was performed with a dissection microscope (Singer MSM System 300, Singer Instruments, Somerset, United Kingdom). Plates were incubated at 30 °C. IMK527 (*MAT*α) was crossed with haploid strains IMS343 (*MAT***a**) and IMZ333 (*MAT***a**) by streaking cultures over each other on selective synthetic medium agar plates containing G418 (100 mg l^{-1}) on which only diploids could grow. For auxotrophic diploids uracil (20 g l^{-1}) was added to the medium. The resulting diploids were re-streaked and single colonies were isolated, yielding strain IMD011 (IMS343×IMK527) and IMD012 (IMZ333×IMK527). Sporulation and tetrad dissection of IMD011 and IMD012 were performed as described above. Dissected spores were replica plated on 1 M glucose synthetic medium agar plates to score for osmotolerant segregants.

Results

Evolutionary engineering for improved osmotolerance. The ability of *S. cerevisiae* IMZ160 (*gpd1*⊿ *gpd2*⊿ *mhpF*) to grow at industrially relevant osmotic pressures was assessed with spot assays on synthetic medium plates containing 0.1, 0.5 and 1.0 M glucose. In line with previous research on Gpd$^-$ strains[49], growth of strain IMZ160 was severely inhibited at 0.5 M glucose, both under aerobic and

anaerobic conditions, and completely abolished at 1.0 M glucose. Growth of the Gpd⁺ reference strain *S. cerevisiae* IME076 was not inhibited at these glucose concentrations (Fig. 3.1).

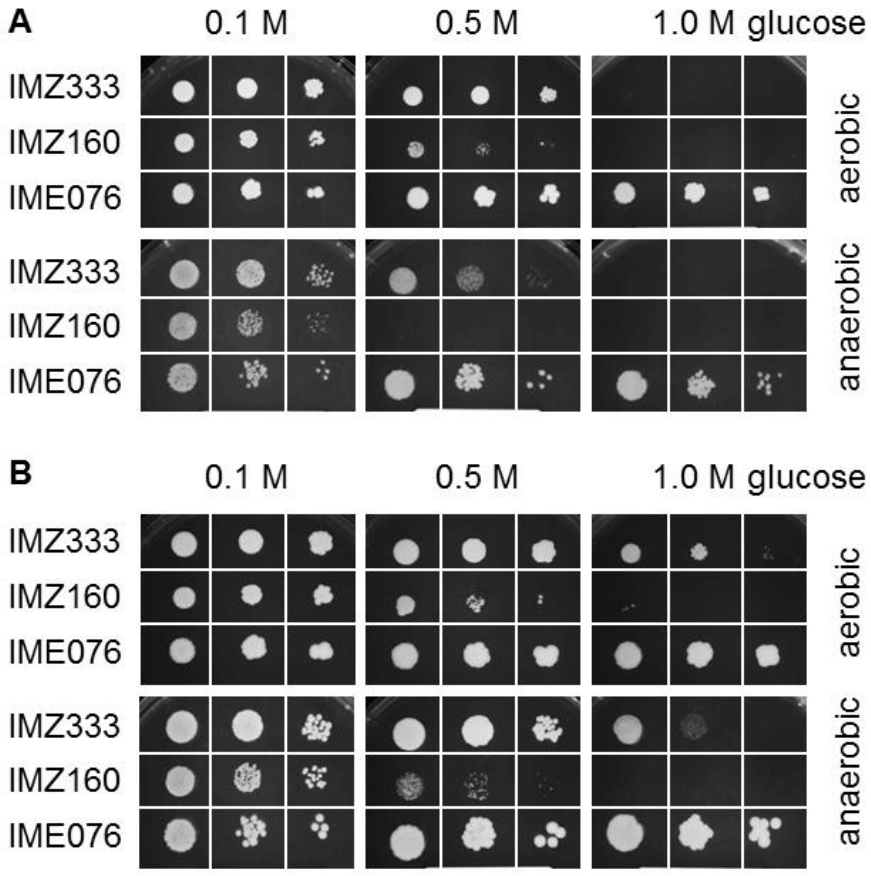

Figure 3.1: Osmotolerance of evolved strain IMZ333 (evolved Gpd⁻), ancestral strain IMZ160 (unevolved Gpd⁻) and the reference strain IME076 (Gpd⁺). Spot assay experiments were performed on synthetic medium agar plates with 0.1-1.0 M glucose under aerobic and anaerobic conditions. Pictures were taken after 3 days (panel A) and 7 days (panel B) of incubation at 30 °C.

Evolutionary engineering for improved osmotolerance was initiated in shake-flask cultures with 20 g l^{-1} glucose as the carbon source supplemented with sorbitol to increase osmolarity. At an initial concentration of 1.0 M sorbitol, the Gpd⁻ strain IMZ160 showed a specific growth rate of 0.06 ± 0.00 h^{-1}, as compared

to 0.37 ± 0.00 h^{-1} for the Gpd^+ reference strain IME076. After 4 serial transfers at 1.0 M sorbitol, 8 transfers at 1.5 M sorbitol and 16 transfers at 2.0 M sorbitol, the cultures showed a maximum specific growth rate of 0.18 ± 0.01 h^{-1} at 1.0 M sorbitol and 0.15 ± 0.01 h-1 at 2.0 M sorbitol. To achieve anaerobic growth at 1.0 M glucose, a shake-flask culture adapted for growth in 2.0 M sorbitol was used as inoculum for an anaerobic bioreactor batch culture at low osmotic pressure in synthetic media with 20 g l^{-1} glucose and 2 g l^{-1} acetate. After 10 days, an anaerobic specific growth rate of 0.12 h^{-1} ± 0.00 was observed. Subsequently, anaerobic sequential batch cultivation was performed on synthetic media supplemented with 2 g l^{-1} acetic acid and with 1.0 M glucose as source of carbon and to increase osmolarity. Starting with an initial anaerobic specific growth rate of 0.05 h^{-1}, the sequential batch culture showed a continuously increasing specific growth rate until, after 187 sequential batch cultures, a specific growth rate of 0.13 ± 0.01 h^{-1} was measured as the average of the last 10 sequential batches. Glycerol, which was not detected during the initial cycles, was detected in culture supernatants later in the evolution experiments, albeit at much lower levels than in cultures of the Gpd^+ reference strain grown under identical conditions (data not shown).

After 187 sequential batch cultures, individual single colony isolates were obtained, whose growth rates were analyzed in anaerobic batch cultures on synthetic glucose supplemented with 1.0 M glucose and 2 g l^{-1} acetate. A single-colony isolate that exhibited the highest maximum specific growth rate of 0.12 ± 0.00 h^{-1} and the lowest final extracellular glycerol concentration of 0.64 ± 0.33 g l^{-1} was named IMZ333 (evolved *gpd1*Δ *gpd2*Δ *mhpF*). Spot assay experiments under aerobic and anaerobic conditions at 0.1, 0.5 and 1.0 M glucose confirmed that, in contrast to the ancestral Gpd^- strain *S. cerevisiae* IMZ160, the evolved strain IMZ333 was able to grow at a concentration of 1 M glucose, albeit slower than the Gpd^+ reference strain IME076 (Fig. 3.1).

Growth and product formation in anaerobic batch cultures at high glucose concentrations. To quantitatively characterize *S. cerevisiae* IMZ333 (evolved *gpd1*Δ *gpd2*Δ *mhpF*), this strain was grown in anaerobic bioreactors on synthetic medium supplemented with 2 g l^{-1} acetic acid. At a glucose concentration of 20 g l^{-1}, the specific growth rate of strain IMZ333 was 0.21 ± 0.01 h^{-1}, which is significantly higher than that of the ancestral strain IMZ160 (0.13 ± 0.01 h^{-1}), but still lower than the Gpd^+ reference strain IME076 (0.32 ± 0.01 h^{-1}). Under these conditions, no glycerol formation was observed for either IMZ333 or IMZ160, whereas the Gpd^+ reference strain produced up to 1.75 ± 0.20 g l^{-1} glycerol.

At an initial glucose concentration of 1.0 M glucose, the evolved strain IMZ333 grew with a specific growth rate of 0.12 ± 0.01 h^{-1} (Fig. 3.2A). Under

identical conditions, its ancestral strain IMZ160 did not grow during a 10 day incubation, while the Gpd⁺ reference strain grew at 0.24 ± 0.01 h^{-1} (Fig. 3.2B), which was in accordance with the spot plate experiments (Fig. 3.1).

Growth of the evolved Gpd⁻ strain IMZ333 was clearly coupled to the use of acetic acid as electron acceptor to reoxidize the excess NADH generated during growth (Fig. 3.2A). During the growth phase, no glycerol was formed by the Gpd⁻ strain. Upon depletion of acetic acid, growth stopped and glucose consumption slowed down. Increasing the acetic acid concentration resulted in a continuation of glucose consumption and in drastically shortened fermentation times (Fig. 3.2C). When all glucose was consumed, low amounts of glycerol, up to 0.64 ± 0.33 g l^{-1} at 182 h, appeared in the supernatant of cultures of the evolved Gpd⁻ strain IMZ333 (Fig. 3.2A). However, the glycerol concentration in these cultures remained at least 10-fold lower than those observed in cultures of the Gpd⁺ reference strain (7.4 ± 0.37 g l^{-1} at 19.5 h) (Fig. 3.2B). The concentration of glycerol measured at the end of the batch cultures supplemented with 3 g l^{-1} acetate (0.53 ± 0.02 g l^{-1} at 93.6 h) was lower than the glycerol concentration measured when 2 g l^{-1} acetate was used (Fig. 3.2). Enzyme activity assays in cell extracts of strain IMZ333 confirmed that glycerol-3-phosphate dehydrogenase activity remained below the detection level of 0.002 μmol min^{-1} mg $protein^{-1}$. The activity of (acetylating) acetaldehyde dehydrogenase in IMZ333 was 0.011 ± 0.005 μmol $(mg\ protein)^{-1}$ min^{-1}, which is not significantly different from the value previously observed for the non-evolved strain IMZ132 (0.020 ± 0.004 μmol min^{-1} mg $protein^{-1}$)[154].

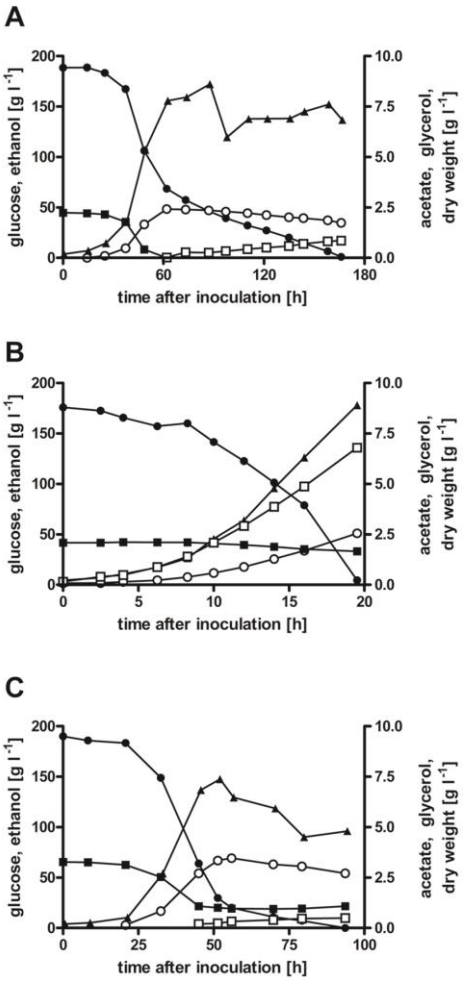

Figure 3.2: Anaerobic batch cultivation of the evolved osmotolerant strain *S. cerevisiae* IMZ333 (evolved Gpd$^-$) and the reference strain IME076 (Gpd$^+$) on synthetic medium with 1 M glucose. Both strains were grown at pH 5.0 and at 30 °C. Panel A: IMZ333, 2 g l^{-1} acetic acid. Panel B: IME076, 2 g l^{-1} acetic acid. Panel C: IMZ333, 3 g l^{-1} acetic acid. Symbols: ▲, Dry weight; ●, glucose; ○, ethanol (not corrected for evaporation); ■, acetate; □, glycerol. Each graph represents values for one of two independent replicates, which differ less than 5% in growth kinetics.

The ultimate goal of eliminating glycerol formation in anaerobic yeast cultures is to increase the ethanol yield on sugar. In the nitrogen-sparged anaerobic bioreactors, a significant amount of ethanol is lost through evaporation. Since ethanol loss via evaporation is time dependent, it will be higher for cultures with a lower specific growth rate[154]. After correction for ethanol evaporation, the apparent ethanol yield on glucose of strain IMZ333 (1.77 ± 0.09 mol mol^{-1}) was 11% higher than that of the Gpd$^+$ reference strain IME076 (1.59 ± 0.02 mol mol^{-1}) in cultures grown on 1 M glucose and 2 g l^{-1} acetic acid. At 3 g l^{-1} acetic acid, the apparent ethanol yield of IMZ333 was 1.84 ± 0.01 mol mol^{-1}, which represents 92% of the theoretical ethanol yield on glucose.

Genetic analysis of the mutations in IMZ333 through mating and sporulation. Mating and sporulation, followed by analysis of segregants, is a powerful approach to investigate the number and nature of mutation(s) in evolved haploid yeast strains[163]. To further investigate the evolved osmotolerant genotype, the evolved strain IMZ333 was mated with its osmosensitive ancestral strain IMK006, after transformation with a selection marker and mating-type switching of the latter. Interpretation of results from crossing and segregation requires that causal mutation(s) reside on the chromosomes rather than on plasmids and, secondly, that the evolved strain should contain the same plasmids as the ancestral strain to avoid random segregation of different plasmids in the spores. A subset of ten plasmids isolated from IMZ333 was characterized by restriction analysis with XmnI. This indicated that there were at least two types of plasmids in this strain: one that had lost the XmnI restriction site in the *TDH3* promoter upstream of the *mhpF* gene and a second that resembles the restriction pattern of the original pUDE043 plasmid. These plasmids were named pUDE043ev1 and pUDE043ev2 respectively. Reinserting these two plasmids and the original pUDE043 plasmid into a plasmid-free ancestral strain IMX031 and in the plasmid-cured evolved strain IMS343 indicated that causal mutations for aerobic osmotolerance were chromosomal, since only the evolved strain, transformed with either of the three plasmids, was able to grow aerobically on 1 M glucose plates. Further analysis showed that only the evolved strain with the reintroduced pUDE043ev2 was able to grow anaerobically, albeit at a lower specific growth rate of (0.07 ± 0.01 h^{-1}), than the original evolved IMZ333 strain. This observation indicated that anaerobic growth of the evolved strain on 1 M glucose required chromosomal as well as (a) plasmid-borne mutation(s) (Fig. 3.3). Sequencing of the *mhpF* gene on this plasmid revealed a point mutation at base pair position 111 of the open reading frame.

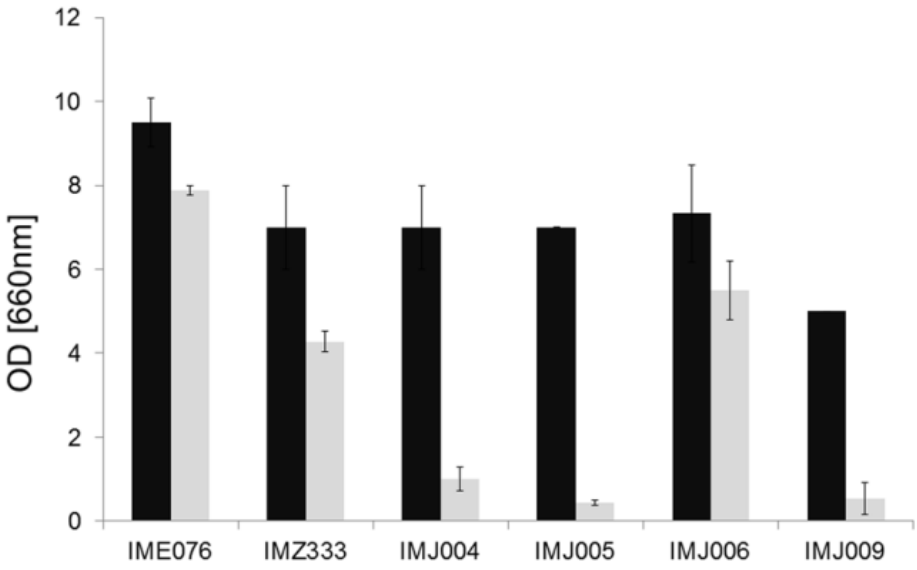

Figure 3.3: Analysis of the contributions of genomic and/or plasmid based mutations to the evolved osmotolerant phenotype of Gpd$^-$ *S. cerevisiae*. Aerobic (black bars) and anaerobic (grey-bars) shake-flask cultures were both incubated at 30 °C and at 200 rpm with an initial glucose concentration of 1 M. The optical density (OD 660 nm) was measured after 48 h for strains IME076 (Gpd$^+$ with empty-vector p426_GPD) and IMZ333 (evolved Gpd$^-$ with evolved pUDE043 population) or after 72 h for strains IMJ004 (evolved Gpd$^-$ pUDE043), IMJ005 (evolved Gpd$^-$ and pUDE043ev1), IMJ006 (evolved Gpd$^-$ and pUDE043ev2) and IMJ009 (evolved Gpd$^-$ with empty-vector p426_GPD).

To prevent interference of plasmids in the backcross analysis, the backcross was performed with a plasmid-free ancestral (IMK527) and a plasmid-cured evolved strain (IMS343) and the osmotolerance was tested under aerobic conditions only. The resulting diploid IMD011 was able to grow on a 1 M glucose plate, indicating that the causal mutation(s) conferring aerobic osmotolerance was (were) dominant. Sporulation of this diploid strain revealed a 2:2 segregation of growth on 1 M glucose plates in 19 out of 19 tetrads, indicating that aerobic osmotolerance in the evolved strain is caused by a single mutation (assuming that there are no linked mutations). Also the mutation(s) enabling anaerobic osmotolerance was/were dominant, since the diploid IMD012, resulting from a cross between the ancestral strain IMK527 with the plasmid containing evolved strain IMZ333, was able to

grow anaerobically on 1 M glucose. Sporulation of this diploid strain yielded very few viable spores and no complete tetrads on non-selective medium, which precluded an accurate analysis of the number of causal mutations underlying anaerobic osmotolerance.

Discussion

We recently proposed a metabolic engineering strategy that enables the use of acetic acid as an electron acceptor for reoxidation of the excess NADH generated in biosynthetic reactions by *S. cerevisiae* and thereby obviates the need for glycerol production for redox balancing[154]. This strategy, which enables increased ethanol yields on sugar, is especially attractive for conversion of lignocellulosic feedstocks, in which acetic acid is invariably present and inhibits yeast fermentation performance. However, the deletion of the *GPD1* and *GPD2* genes encoding glycerol-3-phosphate dehydrogenase proposed by Guadalupe Medina *et al.* (2010) renders *S. cerevisiae* osmosensitive[49,154] (Fig. 3.1). The present study provides a proof of principle that evolutionary engineering can be successfully applied to enable growth of acetate-reducing, *gpd1Δ gpd2Δ* strains to levels that are compatible with industrial bioethanol production. Although further increases in rate are definitely required, this represents an important step towards industrial implementation of an acetate-reducing *gpd1Δ gpd2Δ S. cerevisiae* strain with (acetylating) acetaldehyde dehydrogenase.

The evolved osmotolerant strain IMZ333 was not only able to grow at 1 M glucose, but converted this sugar to ethanol at increased yields relative to a *GPD1 GPD2* reference strain (11 and 15% increases in cultures grown at 2 and 3 g l^{-1} acetic acid, respectively). Part of the increased ethanol yield arises from the elimination of glycerol formation (80 mM in the reference strain), which frees up additional glucose that can be converted to ethanol. A further increase of the ethanol yield can be attributed to the slower growth and longer duration of the fermentation, which increases the fraction of sugar that is converted to ethanol by increasing the cellular maintenance energy requirement[66,164]. Finally, additional ethanol is produced during the reduction of acetic acid, leading to maximal increases of 33 and 55 mM for the fermentations containing 2 and 3 g l^{-1} respectively. The complete consumption of acetate in the cultures grown at 2 g l^{-1} acetate (Fig. 3.2) illustrates that expression the (acetylating) acetaldehyde dehydrogenase strategy not only increases ethanol yields, but also enables the detoxification of significant amounts of inhibiting acetic acid. This detoxifying effect may be particularly relevant when high initial acetate concentrations are prevented by gradual feeding, for example by simultaneous

saccharification and fermentation or by fed-batch feeding of hydrolysates to yeast fermentation processes[165,166].

Even though both genes encoding glycerol-3-phosphate dehydrogenase were deleted and the enzymatic activity of glycerol-3-phosphate dehydrogenase was confirmed to be below the detection limit, low concentrations of glycerol were observed upon cessation of growth of the evolved strain IMZ333. Glycerolipids are essential for the growth of *S. cerevisiae* and are formed by acylation of glycerol-3-phosphate[167] (G3P). However, glycerolipids can also be obtained by acylation of dihydroxyacetone phosphate (DHAP) by the same G3P/DHAP acyltransferase, producing acyl-DHAP, which is later reduced to acyl-G3P by 1-acyldihydroxyacetone-phosphate reductase (EC 1.1.1.101), encoded by *AYR1*[168]. Through this route Gpd$^-$ *S. cerevisiae* strains are able to form glycerolipids and grow. The low concentration of glycerol that was observed in the evolved Gpd$^-$ strain at the end of the fermentation, might be formed by deacylation of the glycerolipids and subsequently released when growth stopped and/or cells lysed.

Analysis of the evolved strain IMZ333 indicated that few mutations were required to increase osmotolerance in *gpd1*Δ *gpd2*Δ acetate-reducing *S. cerevisiae* strain. Detailed analysis of the molecular basis of improved osmotolerance in evolved acetate-reducing strains by whole genome resequencing, thereby enabling its reverse engineering into industrial strains, will require additional, independent evolution experiments[112]. Such experiments should also reveal whether evolutionary engineering always leads to the low-level *GPD1/GPD2*-independent glycerol formation found in strain IMZ333, or that alternative pathways, for example involving trehalose or proline as alternative compatible solutes[169-171] can also contribute to evolution of increased osmotolerance in these strain backgrounds.

Although analysis of the molecular basis of improved tolerance was outside its scope, the present study provides valuable information for experimental design towards this goal. Firstly, our results indicate that mutations which confer osmotolerance under aerobic conditions are not necessarily sufficient to enable growth at high glucose concentrations in anaerobic cultures. In view of the envisaged application of strains in anaerobic bioethanol processes, future evolution experiments should therefore preferably be performed under anaerobic conditions. Secondly, our results indicate that mutations on the *mhpF* expression plasmid contributed to anaerobic osmotolerance. To facilitate the application of classical genetics and whole genome resequencing, it is therefore preferable to perform future evolution experiments with engineered strains in which the *mhpF* expression cassette has been integrated into the *S. cerevisiae* genome. Moreover, this result indicates that mutagenesis of *mhpF* and/or expression of other (acetylating)

acetaldehyde dehydrogenase genes may contribute to osmotolerance in anaerobic cultures.

Acknowledgements

We are grateful to Erik de Hulster, Susan Weening and Frank Fluitman for experimental assistance. The PhD projects of VGM and BO are sponsored by The Netherlands Genomics Initiative through its support for the Kluyver Centre for Genomics of Industrial Fermentation. The work of VGM, BM and BO presented in this manuscript is financially supported by DSM.

Chapter 4

Genome duplication and mutations in *ACE2* cause multicellular, fast-sedimenting phenotypes in evolved *Saccharomyces cerevisiae*

Bart Oud, Víctor Guadalupe-Medina, Jurgen F. Nijkamp, Dick de Ridder, Jack T. Pronk, Antonius J. A. van Maris and Jean-Marc Daran.

Submitted for publication.

Abstract

Laboratory evolution of the yeast *Saccharomyces cerevisiae* in bioreactor batch cultures has, under different selection pressures, yielded variants that grow as multicellular, fast-sedimenting clusters. Knowledge on the molecular basis of this phenomenon may contribute to the understanding of the natural evolution of multicellularity and to manipulating cell sedimentation in laboratory and industrial applications of *S. cerevisiae*. Multicellular, fast-sedimenting lineages obtained from two independent evolution experiments with a haploid *S. cerevisiae* strain were analyzed by whole-genome resequencing. The two independent cell lines showed different frameshift mutations in a stretch of eight adenosines in *ACE2*, which encodes a transcriptional regulator implicated in cell cycle control. Introduction of the two *ace2* mutant alleles into the haploid parental strain led to slow-sedimenting cell clusters that consisted of just a few cells, thus representing only a partial reconstruction of the evolved phenotype. In addition to single-nucleotide mutations, a whole-genome duplication event had occurred in both evolved multicellular strains. Construction of a diploid reference strain with two mutant *ace2* alleles led to complete reconstruction of the multicellular-fast sedimenting phenotype. This study shows that whole-genome duplication and a frameshift mutation in *ACE2* are sufficient to generate a fast-sedimenting, multicellular phenotype in *S. cerevisiae*. The nature of the *ace2* mutations and their occurrence in two independent evolution experiments encompassing fewer than 500 generations of selective growth suggest that switching between unicellular and multicellular phenotypes may be relevant for competitiveness of *S. cerevisiae* in natural environments.

Introduction

Ease of cultivation and genome analysis, short generation times and large population sizes have contributed to the popularity of micro-organisms as model systems in experimental evolution. In addition to providing insights into evolutionary adaptation mechanisms and strategies, laboratory evolution of microorganisms provides a powerful tool to improve characteristics that are relevant to microbial biotechnology. This application of laboratory evolution, known as evolutionary engineering[111] has, for example, contributed to expanding substrate range[94,154,172,173], functional implementation of alternative product pathways[174,175] and increased tolerance to inhibitors[154,176,177] in various production organisms (reviewed in[178]). Recent advances in DNA sequencing and genetic modification facilitate characterization and reconstruction of the molecular basis of evolved phenotypes obtained in laboratory evolution, thus enabling experimental testing of hypotheses on evolutionary strategies and underlying molecular mechanisms[179]. This approach has generated new insights into mutation rates[180,181], genetic drift[180,182], epistasis[183], clonal interference[184] and other important aspects of evolution by natural selection (reviewed in[185]). In microbial biotechnology, reverse engineering of evolved phenotypes, known as inverse metabolic engineering[186], has similarly benefited from the availability of these genomic methodologies[112]. In this applied research context, knowledge on the genetic basis of an industrially relevant phenotype not only increases understanding, but also enables its reconstruction and improvement in other microbial strains and species[112,187,188].

In unicellular organisms such as the yeast *Saccharomyces cerevisiae*, laboratory evolution is facilitated by the ease with which single-cell lines can be isolated from evolving cultures. Recently, however, Ratcliff *et al.* (2012) described evolution of multicellularity in *S. cerevisiae* cells within a single long-term cultivation experiment[189]. The multicellular variant, in which daughter cells did not separate from the mother cell upon cell division, dominated the population within a few generations when fast sedimentation was selected for in test tubes. Evolution of these multicellular clusters of *S. cerevisiae*, which even showed signs of cellular differentiation, was proposed to be a laboratory model for the origin of multicellularity in eukaryotes[189].

At least 25 occurrences of the shift from unicellular to multicellular life forms have been recognized in the evolution of life on Earth[190-192]. However, knowledge on the evolutionary pressures resulting in the selection of multicellular life forms and the underlying molecular mechanisms is far from complete. It has been proposed that multicellularity can contribute to phenotypes as diverse as stress tolerance[193,194], affinity for substrates[195] and relief of predatory pressure[196].

Knowledge on the mutations that cause the switch from unicellular to multicellular growth in yeast may contribute to understanding of the events leading to the transition to multicellular life forms. Moreover, such knowledge can contribute to a better modulation of biomass sedimentation in laboratory research and industrial application of *S. cerevisiae*. In our research on evolutionary engineering of *S. cerevisiae*, we frequently observed multicellular, fast-sedimenting clusters that, upon microscopic examination, resemble the phenotype described by Ratcliff *et al.* [189]. The goal of the present study was to elucidate mutations that are responsible for the generation of multicellular variants. To this end, we monitored the formation of multicellular variants in two independent laboratory evolution experiments with a haploid laboratory strain of *S. cerevisiae*. Subsequently, representative mutants from the two experiments evolutions were characterized. Genetic changes identified by whole-genome resequencing were reverse engineered in the unicellular parental strain, enabling the identification of two changes that, together, were sufficient to reproduce the multicellular, fast-sedimenting phenotype.

Results

Selection of multicellular clusters in sequential bioreactor batch cultures. Formation of large multicellular clusters of *S. cerevisiae* in sequential batch cultures has been reported previously[189]. We reproducibly observed a similar phenotype during prolonged anaerobic cultivation of the haploid *S. cerevisiae* strain CEN.PK113-7D[159] in sequential bioreactor batch cultures. To facilitate identification of mutations contributing to the multicellular phenotype[112,197], two identical independent anaerobic evolution experiments were started on a mixture of 20 g l^{-1} glucose and 20 g l^{-1} galactose. . While the specific growth rate on galactose doubled in both evolution experiments (from 0.11 to 0.22 h^{-1} and 0.20 h^{-1}; Fig. 4.1A and 4.S1A) and the length of the batch cultivation cycles decreased by at least 35% (Fig. 4.S1H and Fig. 4.S1I), the morphology of *S. cerevisiae* changed dramatically as large, multicellular clusters became dominant in both evolution experiments (Fig. 4.1B-F and Fig. 4.S1B-G). The sedimentation index, calculated from the time-dependent decrease of the optical density of statically incubated cell suspensions, strongly increased, in parallel with the increasing abundance of multicellular clusters (Fig. 4.1B-F and 4.S1B-G). Culture samples taken at the end of the two evolution runs (after 4200 h (*ca.* 900 generations) or 2880 h (*ca.* 500 generations)) sedimented almost completely within 5 min of static incubation (Fig. 4.1C).

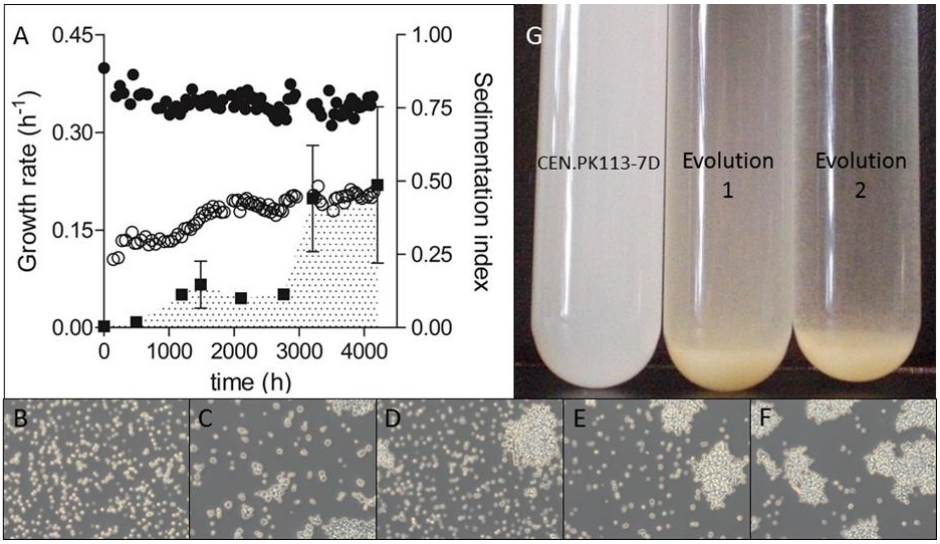

Figure 4.1: Sequential batch cultivation in bioreactors on glucose-galactose mixtures resulted in evolution of multicellular *S. cerevisiae*. **A**- Maximum specific growth rate (μ_{max}) estimated from CO_2 production during glucose consumption in the glucose-galactose batch cultures (●) and the μ_{max} on galactose estimated from galactose batch cultures (○) in evolution experiment 1. Culture samples were taken at different stages of the evolution experiment, grown to stationary phase in shake flasks containing YP medium with 20 g l^{-1} glucose and were left to settle for 30 min in a 1 ml cuvette. Sedimentation indices (■) represent the difference in OD_{660} within 30 min. The data represent the average and the mean deviation of duplicate experiments. **B**- Microscopic pictures of evolution line 1 after 0, **C**- 1196 h, **D**- 2105 h, **E**- 3209 h and **F**- 4200 h of evolution. **G**- Sedimentation of the reference strain CEN.PK113-7D and a culture sample of evolution line 1 and 2 at 4200 and 2877 hours respectively, photographed after 5 min of static incubation.

In *S. cerevisiae*, reversible aggregation of individual cells into fast-sedimenting clusters can also occur via flocculation, which involves a Ca^{2+}-dependent interaction of yeast cell wall proteins and carbohydrates[198]. However, the multicellular clusters observed in the evolved cultures could not be reverted to a single cell morphology by incubation with anti-flocculant agents such as EDTA (0.5 M)[199] or protease (trypsin 1500 units ml^{-1}). This indicated that the phenotype did not result from interaction of unicellular yeasts, but rather from an incomplete cell division.

Whole-genome sequence analysis of two evolved multicellular isolates. To investigate the molecular basis of the evolved multicellular phenotype, a fast-sedimenting mutant was isolated from each of the two evolution experiments. Strains IMS0267 and IMS0386 originated from evolution 1 and from evolution 2,

respectively. To verify the genetic stability of the mutations responsible for the multicellular phenotype, strains IMS0267 and IMS0386 were grown for at least 10 generations on glucose in shake flask cultures. This did not result in observable changes in multicellularity or sedimentation behavior. Genomic DNA of strains IMS0267 and IMS0386 was sequenced at high genome coverage (81.6-fold and 38.5-fold coverage for IMS0267 and IMS0386, respectively) and compared to the reference genome of the parental strain CEN.PK113-7D[159]. The high coverage enabled accurate analysis of genome-wide copy number variation (CNV) by co-assembly of the evolved and the reference strains[200] as well as identification of single nucleotide variations (SNV) and indels.

To estimate the ploidy of the evolved strains we *de novo* co-assembled sequence reads of the evolved and the CEN.PK113-7D reference strains. Copy numbers of the assembled contigs were estimated using the Poisson mixture model-based algorithm Magnolya[200]. Surprisingly, this analysis revealed that both evolved mutants had undergone a whole-genome duplication event relative to the haploid *MAT***a** ancestor CEN.PK113-7D (Fig. 4.2A and 4.2B). Both IMS0267 and IMS0386 were for the most part diploid with triplicated genome islands. IMS0267 exhibited triplication of parts of CHRII, XIII and XVI while IMS0386, besides triplication of parts of CHRIII, VIII and quadruplication of XIII, had a complete trisomy of CHRII and XI (Fig. 4.2A). However, IMS0267 and IMS0386 kept haploid characteristics as the strains were not able to form tetrads, but were able to mate with a *MAT*α strain (IMI081) and sporulate at a low rate.

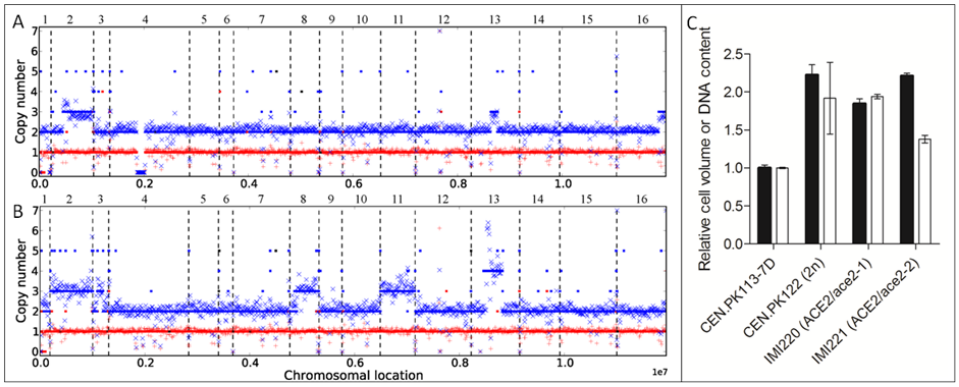

Figure 4.2: Ploidy of the evolved mutants IMS0267 and IMS0386. A- Prediction of DNA content in the evolved strains IMS0267 and **B**- IMS0386, using the Magnolya algorithm [200]. The numbers indicate chromosome position. + (red) indicates the ploidy of the ancestral genome and x (blue) indicates the ploidy of the evolved genome **C**- Determination of cell size (white bar) and DNA content measurements (black bar) of strains CEN.PK113-7D (*MAT***a**), CEN.PK122 (MAT**a**/*MAT*α), IMI220 (*ACE2*/*ace2-1-HphNT1*) and IMI221 (*ACE2*/*ace2-2-HphNT1*). The strains IMI220 and IMI221 are unicellular strains derived from IMS0267 and IMI0386 by reintroduction of a wild-type *ACE2* allele. The data represented are presented as average ± mean deviation of duplicate biological replicates.

Mapping of sequence reads of the evolved strains onto the genome sequence of CEN.PK113-7D revealed no single nucleotide variations (SNVs) and only two high-probability indels (Table 4.1). Only a single gene, *ACE2*, was affected by mutation in both evolved strains (Table 4.1). *ACE2* encodes a transcriptional regulator of, amongst others, *CTS1*, a gene involved in the latest phase of the cell cycle and more specifically required for septum destruction after cytokinesis[201-203]. Interestingly, two differently mutated *ACE2* alleles were identified in the evolved isolates. These mutations were found in the same region of *ACE2*: in IMS0267 an adenosine was introduced at position 1112 while in IMS0386 an adenosine was deleted at the same position. The resulting alleles were named *ace2-1* and *ace2-2*. Both mutations caused the introduction of a premature stop codon, at position 1165 or position 1114 in IMS0267 and IMS0386 respectively (Fig. 4.S2). Based on its occurrence in both evolved strains and its role in the yeast cell cycle, we hypothesized that the mutations in *ACE2* contributed to the evolved multicellular phenotype.

Table 4.1: Insertions (INS) and deletions (DEL) detected in the genomes of two evolved multicellular mutant isolates (IMS0267 and IMS0386) relative to the sequence of the parental haploid reference strain CEN.PK113-7D.

Gene	Description	Nucleotide change	Type mutation	Amino acid change
	Genes mutated in IMS0267			
ACE2	Transcription factor that activates expression of early G1-specific genes.	*1112A	INS	frame-shift
	Genes mutated in IMS0386			
ACE2	Transcription factor that activates expression of early G1-specific genes.	A1112*	DEL	frame-shift

Ace2-1 and Ace2-2 exhibit reduced transcriptional activation of Ace2 targets. To investigate the role of the mutations found in *ACE2* in the evolved multicellular strains, the transcriptional activity of *ace2-1* and *ace2-2* alleles was evaluated. The predicted proteins encoded by *ace2-1* and *ace2-2* alleles were 388 and 371 amino acids long instead of 770 amino acids for original protein (Fig. 4.S2). As a result, the truncated proteins have lost the three C2H2-type zinc finger domains and the Nuclear Localization Signal sequence (NLS) located at the C-terminus of the proteins. Conversely, the truncated Ace2 versions retained the nuclear export signal sequence and the interaction domain with Cbk1, a protein kinase involved in the regulation and localization of Ace2. Quantification of the transcripts of the previously characterized *ACE2* targets, *EGD1*, *EGD2*, *CTS1* and *SCW11*[203,204], by real-time RT PCR revealed that the expression of the four marker genes was reduced by at least 90% in the evolved strains compared to the CEN.PK113-7D reference strain (Fig. 4.3A). Among these Ace2 target genes, *CTS1* is of special interest, since it encodes an endo-chitinase required for the degradation of the mother-daughter septum[205]. Cell wall staining with Calcofluor White, which specifically stains chitin[206], confirmed that within the multicellular clusters, the cells remained attached at the chitin bud neck site (Fig. 4.3B). Moreover, treatment with chitinase led to dispersal of the clusters into single cells (Fig. 4.3C and Fig. 4.3D).

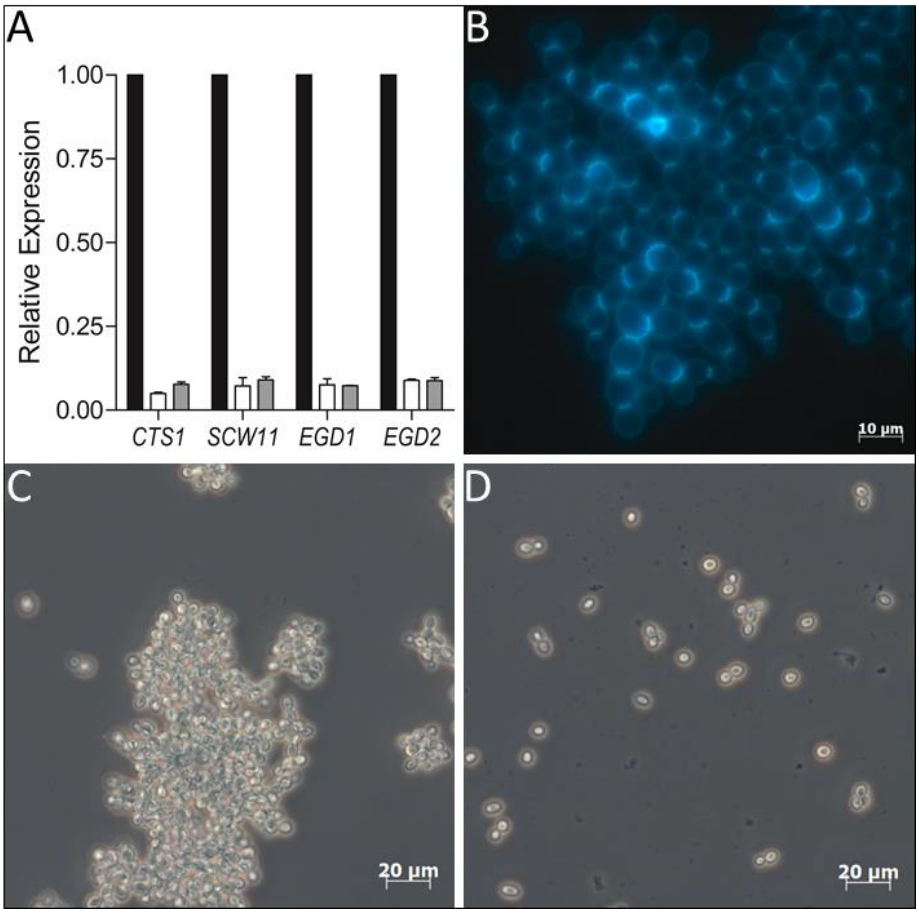

Figure 4.3: Effect of mutations in *ACE2* on gene expression and multicellularity. **A**- Quantification of the expression of characterized Ace2 regulated genes (*CTS1*, *SCW11*, *EGD1* and *EGD2*) in the strains CEN.PK113-7D (black bar; *ACE2*) IMS0267 (white bar; *ace2-1/ace2-1*) and IMS0386 (grey bar; *ace2-2/ace2-2*). The samples were sampled in mid exponential phase from a shake flask culture grown on YP medium with 20 g l^{-1} glucose. Relative gene expression data represent the expression of *CTS1*, *SCW11*, *EGD1* and *EGD2* normalized to *ACT1*. The expression ratios were further normalized relative to CEN.PK113-7D. The data represented are average ± mean deviation of duplicate biological replicates. **B**- Calcofluor White staining of an IMS0267 multicellular cluster. This picture is representative for the entire culture as well as for the two other single-colony isolates obtained from evolved hypersedimenting cultures. **C**- Microscopic observations of a multicellular cluster of IMS0386 resuspended in 100 KPB (potassium phosphate buffer) prior to and **D**- after 7 h incubation with 60 units of chitinase at 25 °C, respectively.

Reverse engineering of different *ace2* alleles in unicellular strains. To further investigate the importance of the *ace2-1* and *ace2-2* mutations in evolution of multicellular, fast-sedimenting *S. cerevisiae* strains, the wild type *ACE2* allele in the haploid ancestor strain CEN.PK113-7D was replaced by either of the two mutant versions. Neither reverse engineering of these mutant *ace2* alleles nor complete deletion of *ACE2* in CEN.PK113-7D (strain IMK395) resulted in complete reconstruction of the multicellular phenotype of the evolved strains (Fig. 4.4). The clusters formed by strains IMK395, IMI197 (*ace2-1-HphNT1*) and IMK245 (*ace2-2-KanMX*) were much smaller and their sedimentation indices, although significantly higher than that of CEN.PK113-7D, were 10-fold lower than those of the evolved isolates IMS0267 and IMS0386. Conversely, replacement of one of the *ace2-1* or *ace2-2* copies in IMS0267 and IMS0386, respectively, by the wild type *ACE2* allele led to a complete reversion of the phenotype to single cells (Fig. 4.4). This observation confirmed the recessive character of the *ace2* mutations (IMI220 and IMI221) that was expected based on the loss of transcriptional activation activity (Fig. 4.3A).

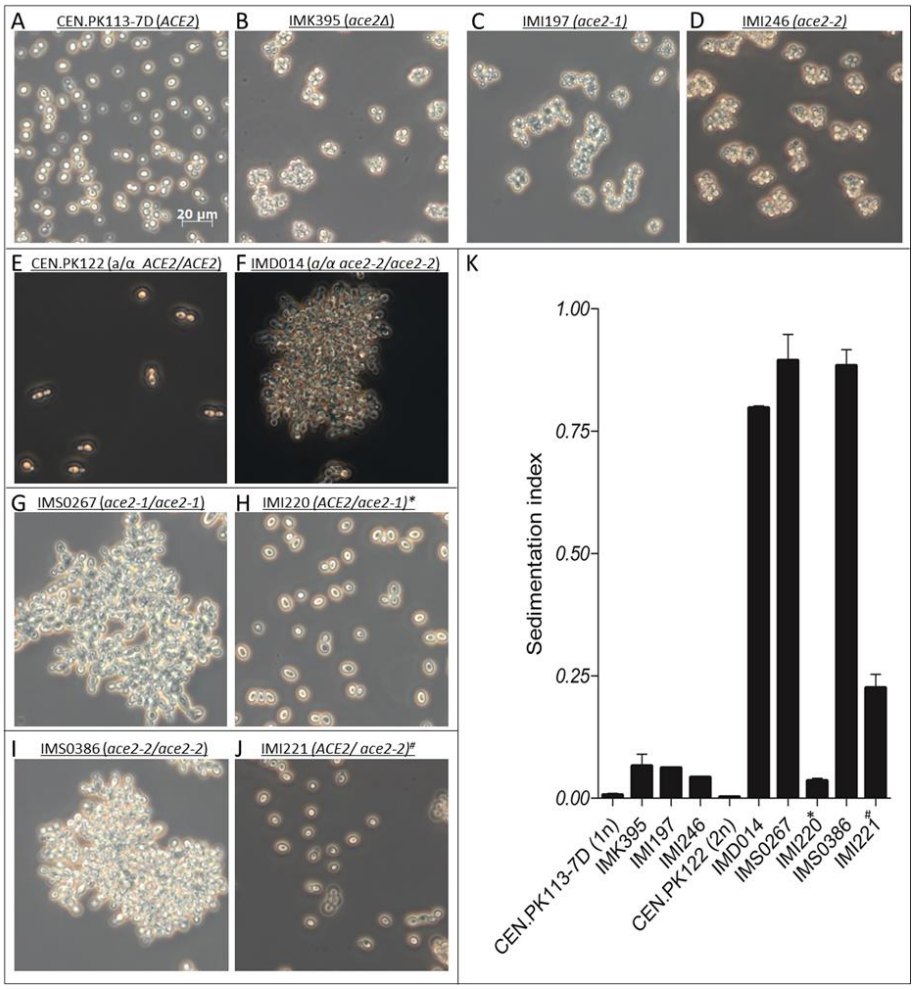

Figure 4.4: Reverse engineering of the multicellular phenotype. Cellular morphology of strains **A**- CEN.PK113-7D (*MAT***a** *ACE2*), **B**- IMK395 (*MAT***a** *ace2*Δ::*loxP-HphNT1-loxP*), **C**- IMI197 (*MAT***a** *ace2-1-loxP-HphNT1-loxP*), **D**- IMI246 (*ace2-2*), **E**- CEN.PK122 (*MAT***a**/*MAT*α *ACE2*/*ACE2*), **F**- IMD014 (*MAT***a**/*MAT*α *ace2-2-loxP-HphNT1-loxP*/*ace2-2-loxP-KanMX-loxP*), **G**- IMS0267 (*ace2-1*/*ace2-1*), **H**- IMI220* (*ACE2*/*ace2-1-loxP-HphNT1-loxP*), **I**- IMS0386 (*ace2-2*/*ace2-2*), **J**- IMI221# (*ACE2*/*ace2-2-loxP-HphNT1-loxP*). **K**- Sedimentation indices of the reference haploid strain CEN.PK113-7D, of the diploid reference CEN.PK122 (*MAT***a**/*MAT*α), the evolved multicellular fast-sedimenting strains IMS0267 and IMS0386 and the reverse engineered mutants IMK395, IMI197, IMI246, IMD014, IMI220* and IMI221#. The sedimentation index indicates the fraction of cells that sediment in a cuvette after 30 min. The data represented are average ± mean deviation of duplicate biological replicates. * denotes strains constructed in the IMS0267 background, # denotes strains constructed in the IMS0386 strain background.

While DNA-content analysis by flow cytometry was not possible with the multicellular evolved strains IMS0267 and IMS0386, similar analysis with strains IMI220 (*ace2-1/ACE2*) and IMI221 (*ace2-2/ACE2*) confirmed the Magnolya prediction (Fig 4.2A and Fig 4.2B). IMI220 and IMI221 exhibited a 2-fold and a 1.4 fold increase in DNA content relative to the haploid reference CEN.PK113-7D (Fig. 4.2C). To exclude the possibility of transformation-associated selection of unicellular mutants, we confirmed that re-exchanging the *ACE2* wild-type allele introduced in IMI220 and IMI221 by *ace2-1* (IMW064 (*ace2-1/ace2-1*) and IMW066 (*ace2-1/ace2-2*)) restored formation of large clusters (Fig. 4.S3).

Since introduction of the *ace2-1* or *ace2-2* alleles in a haploid strain was not sufficient to reconstruct the multicellular phenotype observed in the evolved strains, we investigated the impact of the change in ploidy of the evolved strains on the multicellular phenotype. To this end, the *MAT*α strain IMI246 (*ace2-2-KanMX*) was constructed by replacing *ACE2* in CEN.PK113-13D and crossed with the *MAT***a** strain IMI197 (*ace2-2-HphNT1*). Strikingly, the resulting diploid strain IMD014 (*ace2-2-KanMX/ace2-2-HphNT1*) formed large multicellular clusters (Fig. 4.4) and exhibited a sedimentation index similar to that of the evolved strains IMS0267 and IMS0386 (Fig. 4.4).

These results demonstrate complete reverse engineering of an evolved multicellular, fast sedimenting phenotype by introduction of a specific recessive mutation in *ACE2* that drastically reduces the transcriptional activity of Ace2p in diploid *S. cerevisiae*.

Discussion

This study provides the first complete identification of a molecular mechanism by which the unicellular eukaryote *S. cerevisiae* can evolve into a multicellular, fast-sedimenting phenotype. Although our laboratory evolution experimental set up differed from that used in another recent study evolution of multicellularity in *S. cerevisiae*, a multicellular phenotype appeared within the first 20 days of selective growth in both studies[189]. In contrast to the extraordinary phenotypical impact of this morphological transition in the history of life, the molecular events underlying the shift from a unicellular yeast morphology to multicellular aggregates were simple, involving a mutation in only a single gene and a genome duplication. The recessive characteristic of the *ace2-1* and *ace2-1* mutations strongly suggests that the *ace2* mutations preceded the genome duplication event that occurred during laboratory evolution of strains IMS0267 and IMS0386.

Alongside multicellularity, the evolved cultures showed accelerated diauxic consumption of glucose-galactose mixtures (Fig. 4.1A and Fig. 4.S1). This faster diauxic galactose consumption cannot be completely attributed to the mutations that caused multicellularity (Fig. 4.S4), suggesting that additional mutations contribute to this characteristic. Analysis of these mutations, which is outside the scope of this study, is complicated by the possible occurrence of allelic variations after genome duplication. Our study underlines the importance of analyzing whole or partial genome duplication in the analysis of evolved strains[207-209]. In addition to facilitating the identification of key mutations, research on genome duplication and subsequent further evolution in laboratory experiments may lead to further insight in the evolutionary past of *S. cerevisiae*, in which a whole-genome duplication has played an important role[210].

Post-division adhesion caused by imperfect degradation of the chitin septum between the mother and the daughter cells appears to be the main mechanism implicated in the formation of the multicellular clusters observed in our laboratory evolution study. This evolutionary mechanism may have played a role in the transition from unicellular yeast to dimorphic and further to filamentous organisms, since these organisms share a conserved role for chitin in cell wall architecture. Inactivation of the *ACE2* ortholog in the pathogenic yeast *Candida glabrata* led to cell clusters and hypervirulence in a murine model[211,212]. Similarly, *C. albicans* strains with an *ace2⊿/⊿* genotype showed altered separation and morphology and, moreover, resistance to azole anti-fungal drugs[212]. However, outbreaks of hypervirulent and/or antibiotic-resistant mutants of these pathogens have hitherto not been reported.

Intriguingly, although mutations in *CTS1* that encodes the endochitinase involved in cell separation or in genes composing the signalling pathway "regulation of Ace2 and morphogenesis (RAM)" might have yielded a similar phenotype, mutations in *ACE2* were found in two independent evolution experiments. The *ace2-1* and *ace2-2* mutations occurred in the same homopolymer of eight adenosine residues (Fig. 4.S4). Poly-(dA:dT) tracts occur frequently in *S. cerevisiae* genome[213,214], and these regions may participate in the yeast genome evolution by creating mutagenesis hot-spots[213], however it worth mentioning that poly-(dA:dT) tracts occurrence in coding regions remains significantly lower than in intergenic (Table 4.S4) maybe to prevent fast accumulation of deleterious mutations (*e.g. ace2-1* and *ace2-2*) in protein encoding DNA.

Evolvability of a fast-sedimentation phenotype may offer selective advantages in natural environments, for example in sugar-rich micro-environments such as flowers or fruits subjected to frequent intensive rainfall. Close inspection of

the nucleotide sequences of *Candida ACE2* orthologs, *S. cerevisiae CTS1* and the genes of the RAM pathway did not reveal homopolymers longer than five residues. In pathogenic *Candida* strains, this might limit the frequency with which hypervirulence occurs as a consequence of loss of function mutations in *ACE2*.

Knowledge on the mutations responsible for a multicellular, fast-sedimenting phenotype in *S. cerevisiae* allows modulation of this property by genetic engineering. The results presented in this study indicate that stable, fast-sedimenting yeast strains for use in cell retention systems can be constructed by inactivation of both copies of *ACE2* in diploid strains. The formation of multicellular clusters, as observed in the evolved strains investigated in this study, does not hinder cell growth. IMS0267, IMS0386 have higher growth rates than their ancestor CEN.PK113-7D in chemically defined medium with glucose and galactose (Fig. 4.1 and 4.S1). Additionally, it may be possible to prevent or delay occurrence of multicellular phenotypes in adaptive evolution experiments, where it is not always a desirable feature, by ectopic integration of multiple *ACE2* genes, since simultaneous mutations of all copies will be required to induce multicellularity.

Materials and Methods

Strain maintenance. *S. cerevisiae* strains used in this study (Table 4.2 and 4.S1) were all derived from CEN.PK113-7D[158]. Strains were maintained on YP medium (demineralized water; 10 g l^{-1} yeast extract [BD Difco, Franklin Lakes, NJ]; 20 g l^{-1} peptone [BD Difco],) with 20 g l^{-1} glucose. Culture stocks were prepared from shake flask cultures, which were incubated at 30 °C and stirred at 200 rpm, by the addition of 20% (v/v) glycerol and were stored at -80 °C.

Table 4.2: Strains used in this study. * denotes strains constructed in the IMS0267 and # strains constructed in the IMS0386 strain backgrounds.

Strain	Description and Genotype	Source
CEN.PK113-7D	*MAT***a** *ACE2*	Euroscarf
CEN.PK113-13D	*MAT*α *ura3-52*	Euroscarf
CEN.PK113-16B	*MAT*α *ACE2 leu2-3-112*	Euroscarf
CEN.PK122	*MAT***a**/*MAT*α *ACE2*/*ACE2*	Euroscarf
IMS0267	*ace2-1* /*ace2-1*	This study
IMS0386	*ace2-2* /*ace2-2*	This study
IMK395	*MAT***a** *ace2::loxP-HphNT1-loxP*	This study
IMI196	*MAT***a** *ACE2-loxP-HphNT1-loxP*	This study
IMK485	*MAT***a** *ACE2 loxP-KanMX-loxP*	This study
IMI197	*MAT***a** *ace2-2-loxP-HphNT1-loxP*	This study
IMK484	*MAT***a** *ace2-1-loxP-KanMX-loxP*	This study
IMI246	*MAT*α *ura3-52 ace2-2-loxP-KanMX-loxP*	This study
IMD014	*MAT***a**/*MAT*α *ura3-52*/*URA3 ace2-2-loxP-HphNT1-loxP*/*ace2-2-loxP-KanMX-loxP*	This study
IMI220	*ACE2*/*ace2-1-loxP-HphNT1-loxP**	This study
IMW064	*ace2-1*/*ace2-1-loxP-KanMX-loxP**	This study
IMI221	*ACE2*/*ace2-2-loxP-HphNT1-loxP*#	This study
IMW066	*ace2-1*/*ace2-2-loxP-KanMX-loxP*#	This study
IMI081	*MAT*α *ACE2 leu2-3-112 loxP-HphNT1-loxP*	This study

Laboratory evolution of CEN.PK113-7D and batch cultivations. Long-term cultivation in sequential batch reactors was the method used to improve the anaerobic growth characteristics of CEN.PK113-7D in a mixture of 20 g l^{-1} glucose

and 20 g l^{-1} galactose. Bioreactors were inoculated by adding a shake flask culture that has been incubated overnight in synthetic medium (SM) (5 g l^{-1} $(NH_4)_2SO_4$, 3 g l^{-1} KH_2PO_4, 0.5 g l^{-1} $MgSO_4 \cdot 7H_2O$, trace elements and vitamins as described in [215]), and 20 g l^{-1} glucose at 30 °C. An alternating batch regime was conducted with every first batch containing 20 g l^{-1} glucose and 20 g l^{-1} galactose medium and every second batch containing 20 g l^{-1} galactose as the sole carbon source in the medium. The batch on galactose only medium was performed to ensure having equal generations on galactose and on glucose[94].

The strains CEN.PK113-7D, CEN.PK122, IMS0267, IMS0386 and IMD014 were compared with respect to fermentation time by batch cultivation in a bioreactor. Bioreactors containing SM with 20 g l^{-1} glucose and 20 g l^{-1} galactose were inoculated by adding a shake flask culture that had been incubated overnight in synthetic medium and 20 g l^{-1} galactose at 30 °C.

Cultivation was carried out in 2 l laboratory bioreactors (Applikon, Schiedam, The Netherlands) with a working volume of 1 liter. SM supplemented with 0.01 g l^{-1} ergosterol, 0.42 g l^{-1} Tween 80 dissolved in ethanol and trace elements was used as the medium to which either 20 g l^{-1} glucose and 20 g l^{-1} galactose or only 20 g l^{-1} galactose was added. Antifoam Emulsion C (Sigma-Aldrich, Zwijndrecht, the Netherlands) was autoclaved separately (120 °C) as a 20% (w/ v) solution and added to a final concentration of 0.2 g l^{-1} to the bioreactor. Cultures were stirred at 800 rpm, cultures were kept anaerobic by sparging 0.5 l min^{-1} nitrogen gas (<10 ppm oxygen) and culture pH was kept at 5 by automatically adding 2M KOH. The bioreactor was equipped with Norprene tubing (Cole Palmer Instrument Company, Vernon Hills, Illinois, USA) to minimize oxygen diffusion. The bioreactor was automatically drained when off-gas CO_2 levels dropped below 0.05% after the CO_2 production peak, leaving 25 ml (evolution 1) or 5 ml (evolution 2) as inoculum for the next batch. The bioreactor was filled to 1 liter using a feed pump controlled by an electric level sensor. For each cycle the specific growth rate on either glucose or galactose was estimated from the off-gas CO_2 production in the exponential phase by fitting an exponential function through the data points. The amount of generations was estimated to be ranged from 7 to 10 per 2 batches based on dry-weight measurements. The culture was regularly checked for purity by plating on lithium-containing agar plates[216] and by microscopical analysis. Intermediate culture samples were stored by the addition of 20% (v/v) glycerol and kept at -80 °C.

Single colony isolation. Representative single colony isolates from both evolution lines were selected by streaking the final biomass sample of both evolutions on YP medium with 20 g l^{-1} galactose. Single colonies were restreaked twice before inoculating a 15 ml plastic tube containing 1 ml synthetic medium

supplemented with vitamins, trace elements and 20 g l^{-1} galactose, which was incubated at 30 °C. After 1 day, these cultures were used to inoculate shake flasks containing the same medium with a working volume of 100 ml. Fully grown cultures of these shake flasks were stocked. The mutant with the highest sedimentation index in evolution 1 was called IMS0267 and in evolution 2 was called IMS0386. IMS1384 was an isolate with the highest sedimentation index from evolution line 2. IMS1384 was isolated by streaking and culturing on YP medium supplemented with 20 g l^{-1} glucose similar to the isolation of IMS0267 and IMS0386 (Table 4.S2 and Fig. 4.S5).

Calcofluor white staining. 2 ml of a fully grown YPD culture was washed 3 times in PBS buffer (3.3 mM NaH_2PO_4, 6.7 mM Na_2HPO_4, 0.2 mM EDTA, 130 mM NaCl) and resuspended in 500 µl PBS buffer. 100 µl of that suspension was incubated for 15 minutes with 10 µl calcofluor white stain (Calcofluor White M2R 1 g l^{-1} and Evans blue 0.5 gl^{-1}; Fluka, Buchs, Switzerland). After incubation the cell suspension was washed once more. Directly thereafter, phase-contrast and fluorescence microscopy was performed with a Zeiss Imager.D1 microscope equipped with a 40× Plan Neofluor lens and Filter Set 01 (excitation bandpass filter width from 353 to 377 nm, emission longpass filter from 397 nm, 395 nm beam splitter filter) (Carl-Zeiss, Oberkochen, Germany). Images were taken with a Zeiss Axiocam MRc using the Axiovision 4.5 software.

Chitinase assay. 100 µl of an overnight YP medium shake flask culture supplemented with 20 g l^{-1} glucose was spun down and resuspended in either 100 µl KPB buffer (pH 6.0) or 100 µl KPB buffer (pH 6.0) with 1 mg ml^{-1} chitinase (chitinase from *Trichoderma viride*, >600 units mg^{-1} (Sigma-Aldrich)).

Sedimentation assay. To visualize the sedimentation in test tubes, yeast cells were harvested from fully grown shake flask cultures with YP medium with 20 g l^{-1} glucose before they were washed twice with synthetic medium. After washing, cells were resuspended in synthetic medium to a standardized concentration equivalent to a dry weight of 2 g l^{-1}. After vortexing thoroughly to ensure a homogeneous cell suspension, the samples were rapidly placed in test tubes, and the clock was immediately started. After 5 minutes, pictures were taken.

To quantitatively determine the speed of sedimentation, cell cultures were grown to stationary phase in YP medium containing 20 g l^{-1} glucose before they were washed twice with synthetic medium. The cell dry weight per volume was normalized to 0.42 g l^{-1}. The cell suspension was left to settle in a cuvette for 30 minutes and continuous recording of OD_{660} was performed using a Hitachi U-3010 spectrophotometer (Hitachi High-Technologies Europe GmbH, Mannheim, Germany). The sedimentation index was calculated by dividing the initial OD_{660} value by the difference of the final OD_{660} value and the initial OD_{660} value.

Whole genome sequencing. Genomic DNA from the two evolved isolates and CEN.PK113-7D was isolated using the Qiagen 100/G kit (Qiagen, Hilden, Germany). A library of 200-bp genomic fragments was created and sequenced paired-end (50-bp reads) using an Illumina HiSeq 2000 sequencer by Baseclear BV (Baseclear, Leiden, the Netherlands). The individual reads were mapped onto the reference genome of CEN.PK113-7D[159], using BWA[217] and further processed using SAMtools[218]. Single-nucleotide variations and small insertions and deletions were extracted from the mapping using SAMtools' varFilter. Default settings were used, except that the minimum and maximum read depth were set to 10× and 400× (-d10 -D400), respectively. To minimize false positive mutation calls, custom scripts and manual curation were used for further mutation filtering. First, mutation calls containing ambiguous bases in either reference or mapping consensus were filtered out. Second, only single-nucleotide variations with a quality of at least 20 and small insertions and deletions with a quality of at least 60 were kept. Variant quality is defined as the Phred-scaled probability that the mutation call is incorrect[219]. Third, mutations with a depth of coverage < 10x were discarded. Fourth, insertion and deletion mutation calls were only kept when at least 70% of the reads spanning the location confirmed the insertion or deletion. Fifth, small insertions and deletions that were close to an 'N' in the reference were removed, because this would complicate correct alignments and introduce false positive mutation calls. All variations were manually verified by comparing with raw sequencing data of CEN.PK113-7D.

Magnolya was used to analyze copy number variation, employing Newbler (454 Life Sciences, Branford, CT) for the co-assembly. Haploid settings were used for CEN.PK113-7D and diploid settings for either of the mutants to determine their ploidy levels[200]. The raw sequencing data were deposit as SRA under the BIOproject ID: PRJNA193417.

Flow cytometric analysis. The cell volume and the DNA content of the evolved isolates, a haploid and diploid reference strain (respectively, CEN.PK113-7D and CEN.PK122) were analyzed by flow cytometry. Therefore a culture volume corresponding to 1×10^7 cells ml^{-1}, determined with the aid of the Z2 Coulter® Particle Count & Size Analyzer (Beckman Coulter, Woerden, The Netherlands), was centrifuged (5 min, 7000 rpm). The pellet was washed once with phosphate buffer (NaH_2PO_4 3.3 mM, Na_2HPO_4 6.7 mM, NaCl 130 mM, EDTA 0.2 mM)[220], and resuspended in phosphate buffer again. Cells were briefly sonicated (~3 seconds) in an MSE Soniprep 150 sonicator (150 W output, 7 mm peak-to-peak amplitude) (MSE, London, UK) to disturb cell aggregation. The DNA content of the living cells were stained with the Vybrant DyeCycle Orange Stain (Invitrogen,

Grand Island, NY) and incubated in the dark for 30 minutes at 37 °C. Stained and unstained samples were analyzed on a Cell Lab Quanta™ SC MPL flow cytometer equipped with a 488 nm laser (Beckman Coulter). Quantification of the fluorescence intensity (DNA content) and electronic volume (EV, as a measure for cell volume) was performed by using the free CyFlogic software (version 1.2.1, CyFlo Ltd, Turku, Finland).

qPCR. To determine the RNA expression levels of Ace2 targets in CEN.PK113-7D, IMS0267 and IMS0386, shake flask cultures were grown in duplicate on YP medium with 20 g l^{-1} glucose. The cultures were incubated at 30 °C until they consumed between 5-15 grams of glucose. The culture was cooled and 20 ml of broth was harvested. Total RNA extraction was based on a method described previously[221]. Cells were centrifuged and resuspended in one pellet volume of TAE buffer, two pellet volumes acid chloroform (5:1, pH 4.5) and 1/10 pellet volume 10% (w/v) SDS. The tubes were placed in a water bath at 65 °C for 5 min before being aliquoted in three 1 ml tubes and stored at -80 °C. RNA extraction was performed by the method of Schmitt *et al.*[222]. cDNA was synthesized using the QuantiTect Reverse Transcription Kit (Qiagen, Dusseldorf, Germany). qPCRs were prepared using the QuantiTect SYBR Green PCR Kit (Qiagen, Dusseldorf, Germany). qPCRs were performed in triplicate on two dilutions in the Rotor-Gene Q (Qiagen, Dusseldorf, Germany). A primer concentration of 0.5 µM in a total volume of 20 µl was used. All qPCR primers are listed in Table 4.S3.

The expression of each transcript relative to the expression in CEN.PK13-7D and normalized to the transcript level of *ACT1* was calculated using the program REST (Qiagen, Dusseldorf, Germany) by inserting take-off values and amplification values. All transcript levels were significantly different to the values in CEN.PK113-7D. A 100% efficient reaction would give an amplification value of 2 for every sample, meaning that the amplicon doubled in every cycle. The actual amplification of the reactions was similar with that obtained using primers for actin *ACT1* (1.65 – 1.9). Outliers (<1.65) were manually removed. The take-off represents the cycle at which the second derivative is at 20% of the maximum level, indicating the end of the noise and the transition to the exponential phase. The take-off value was calculated for each gene of interest by the Rotor-Gene Q Series Software (Qiagen, Dusseldorf, Germany). The average relative transcript levels were determined from 2-4 technical replicates. All results presented are averages of at least two biological replicates, *i.e.* 2 shake flask cultivations

Strain construction. Transformation of linear DNA fragments into several *S. cerevisiae* mutants was done according to the lithium-based transformation protocol described by Gietz and Woods[223]. Transformants were selected on YP agar

medium containing 200 mg l^{-1} hygromycin B or 200 mg l^{-1} G418 and 20 g l^{-1} glucose. Tranformants were restreaked once before they were confirmed to have the correct integration by PCR on colony material suspended in 0.02 M NaOH and boiled for ten minutes. To confirm the presence of the correct allele(s) single read (Sanger) sequencing was performed on selected PCR products by Baseclear on the ABI3730XL sequencer (Life Technologies Ltd. Paisley, United Kingdom).

Disruption of *ACE2* in CEN.PK113-7D was done by integrating the ACE2KO construct, which was amplified by PCR from the plasmid pUG-*hphNT1*[175] with primers ACE2KOf and ACE2KOr. Correct replacement of the *ACE2* gene by the hygromycin B resistance gene was confirmed by PCR with primers sets ACE2fw - Hph NT1 fw, ACE2rv - Hph NT1 and ACE2fw - ACE2rv. The resulting strain was named IMK395 (*ace2Δ::loxP-HphNT1-loxP*).

Introduction of the ancestral *ACE2* allele (resulting in IMI196), the *ace2-1* allele (resulting in IMK245) and the *ace2-2* allele (resulting in IMI197) into CEN.PK113-7D or introduction of the *ace2-1* allele in CEN.PK113-13D (resulting in IMK484) was done by integrating two, partially overlapping, constructs into the appropriate genome (Fig. 4.S6A). The first construct contained either the ancestral *ACE2* allele or one of the *ace2* alleles from the isolates of the evolution flanked by a unique overlapping sequence to the second construct. This first construct was obtained by PCR on genomic DNA of CEN.PK113-7D or on genomic DNA of IMS0386 using primers ACE2idF and ACE2tagA. For IMS0267, the first construct was amplified from genomic DNA of IMS0267 using primers ACE2idf and ACE2tagB. The second construct also contained the unique sequence, together with the hygromycin B or kanamycin resistance gene and a sequence homologous to a sequence 204-bp upstream of *ACE2*. This second construct was obtained by PCR on the plasmid pUG-*hphNT1*[174] using primers tagApUG and pUGACE2r or by a PCR on pUG6[224] using primers tagBpUG and pUGACE2r. After integration of 2 constructs in the CEN.PK113-7D genome, correct insertion of the constructs was confirmed by PCR using primers pairs ACE2seqf - Hph NT1 rv or ACE2seqf-KanA, ACE2hygidrv - Hph NT1 fw or ACE2hygidrv - KanB and ACE2seqf - ACE2hygidrv. By sequencing the PCR product obtained from the primer pair ACE2seqf - Hph NT1 rv or ACE2seqf - KanA the insertion of the correct allele was confirmed using the primer ACE2seqf.

Since the introduction of two genetic elements into the multicellular mutants proved more difficult than in the unicellular ancestor, allele switching in these mutants was done by integrating one complete construct into the *ACE2* locus (Fig. 4.S6B). The construct was obtained by amplifying the complete ACE2-tagA-HphNT1-ACE2 construct from genomic DNA of the appropriate mutants

constructed in CEN.PK113-7D by PCR with primers ACE2seqf and ACE2hygidrv. After integration of this constructs in IMS0267 (resulting in IMI220) and IMS0386 (resulting in IMI221) the correct insertion of the construct was confirmed by PCR using primer pairs ACE2f - Hph NT1 rv, ACE2TARcheck - Hph NT1 fw and ACE2f - ACE2TARcheck. By sequencing the PCR product obtained from the primer pair ACE2f - Hph NT1 rv and by sequencing the smaller PCR product from the primer pair ACE2f - ACE2TARcheck using the primer ACE2seqf the presence of the expected alleles was confirmed.

Construction of a diploid *ace2-1/ace2-2* mutant was done by crossing strain IMI197 and strain IMI246 on YP agar medium supplemented with 20 g l^{-1} glucose. The resulting diploid strain was selected on synthetic agar medium with 200 mg l^{-1} G418 by restreaking twice on this medium. Correct insertion of the correct alleles was confirmed by sequencing the PCR product obtained from the primer pair ACE2f-Hph NT1 rv and by sequencing the PCR product obtained from the primer pair ACE2f-KanA.

Reintroduction of the appropriate mutated *ace2* allele back into IMI220 or IMI221 was done by integrating two overlapping constructs into the *ACE2* locus, thereby replacing the *ACE2*-tagA-*HphNTI-ACE2* construct (Fig. 4.S6C). The first construct contained one of the *ace2* alleles from IMS0267 or IMS1384 flanked by a unique overlapping sequence to the second construct. This first construct was obtained by PCR on genomic DNA of IMS0267 or IMS0386 using primers ACE2idf and ACE2tagB. The second construct also contained the unique sequence, together with the kanamycin resistance gene and a sequence homologous to a sequence 204 bp downstream of *ACE2*. This second construct was obtained by PCR on the plasmid pUG6[224] using primers tagBpUG and pUGACE2r. After integration of 2 constructs in the appropriate genome, correct insertion of the constructs was confirmed by PCR using primers pairs ACE2f-KanA, ACE2TARcheck-KanB and ACE2f-ACE2TARcheck as well as by observing growth on G418 plates and not on hygromycin containing plates. By sequencing the PCR product obtained from the primer pair ACE2f-Hph NT1 rv and by sequencing the smaller PCR product from the primer pair ACE2f-ACE2TARcheck using the primer ACE2seqf the presence of the expected alleles was confirmed.

Introduction of a hygromycin resistance gene into the *MAT*α CEN.PK113-16B strain was done by transforming in a genetic construct, which was obtained by PCR from the plasmid pUG-*hphNT1*[174] with primers MTH1markfw and MTH1markrv. The resulting strain was named IMI082 (*ACE2 loxP-HphNT1-loxP*)

Constructs were made by PCR amplification on genomic DNA by using Expand high fidelity Polymerase (Roche, Basel, Switzerland) according to

manufacturer's instructions in a Biometra TGradient Thermocycler (Biometra, Gottingen, Germany). Isolation of fragments from gel was done with the Zymoclean Gel DNA Recovery kit (Zymo Research, Orange, USA). PCR amplification on colony material was done using FastStart Taq DNA Polymerase (Roche) according to manufaturer's instructions on colony material suspended in 0.02 M NaOH, which was boiled for ten minutes at 100 °C.

ACE2 was replaced in a similar procedure in IMS1384 by *ace2-1*, resulting in IMI219. IMI219 was subsequently converted to a multicellular phenotype again by replacing the *ACE2 HphNT1* construct by *ace2-1 kanMX* contruct resulting in IMW062. This procedure was similar to the creation of IMW064 and IMW066.

Mating and Sporulation. Strains IMS0267 and IMS0386 were mated with IMI081 by streaking both strains on plates containing YP medium supplemented with 20 g l^{-1} glucose. After overnight incubation at 30 °C the strains were streaked over each other. After another 4 h of incubation at 30 °C diploids were selected by streaking the mixture of two strains on selective medium (SM medium with 20 g l^{-1} glucose and 200 mg l^{-1} hygromycin). Resulting single colonies were restreaked twice on the same medium.

Sporulation was performed by incubating a culture in YP medium supplemented with 10 g l^{-1} potassium acetate for 1 day at 30 °C. Subsequently the entire culture was washed twice, resuspended 20 g l^{-1} potassium acetate and incubated for 3-4 days at 30 °C. Spores were segregated on YP agar plates supplemented with 20 g l^{-1} glucose using a standard micromanipulator (Singer Instruments, Somerset, UK) and were incubated at 30 °C.

Homopolymer stretch distribution. The *S. cerevisiae* reference genome and its annotation (release 64-1-1, February 3 2011) were downloaded from the *Saccharomyces* Genome Database (http://www.yeastgenome.org/)[225]. A file "domains.tab", containing domains predicted using InterProScan[226], was downloaded from the same site (March 10 2013). The number of occurrences of homopolymeric (dA:dT) stretches of length 8 or longer was counted in the overall genome, in genes (*i.e.* sequences annotated as "gene" in the reference genome), in coding sequences within genes, in introns and in domains. Stretches were considered present when all bases fell inside the genomic feature. For each of these features, a Fisher exact test (two-tailed) was then performed under the null hypothesis that the occurrence of homopolymeric stretches is independent of the underlaying genomic feature (genes, coding sequences, introns, domains).

Acknowledgements

We thank Mark Bisschops and Marijke Luttik for help with confocal microscopy and Marit Hebly for help with flow cytometry analysis. Erik de Hulster is acknowledged for expert advice and support on bioreactor operation. Carsten Blom, Edwin van der Pol and Vito Meulenberg are acknowledged for their contributions via student projects. We thank Tim Vos for critical reading of the manuscript. The PhD project of BO was carried out within the research program of the Kluyver Centre for Genomics of Industrial Fermentation, which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research.

Supplementary Material

Supplementary Tables

Table 4.S1: Additional strains used in this study (complementary to table 4.2). IMS1384 is an isolate from the evolution line number 2 as IMS0386. & denotes strains constructed in the IMS1384 background.

Strain	Description and Genotype	Source
IMS1384	*ace2-1 /ace2-1*	This study
IMI219	*ACE2/ace2-1 HphNT1*&	This study
IMW062	*ace2-1/ace2-1 KanMX*&	This study

Table 4.S2: Single Nucleotide Variations (SNVs), insertions (INS) and deletions (DEL) detected in the genomes of an additional evolved multicellular mutant isolate IMS1384 relative to the sequence of the ancestor haploid reference strain CEN.PK113-7D.

Gene	Description	Nucleotide change	Type mutation	Amino acid change
	Genes mutated in isolate 2 of evolution 2			
ACE2	Transcription factor that activates expression of early G1-specific genes.	*1112A	INS	frame-shift
MIH1	Protein tyrosine phosphatase involved in cell cycle control.	T356C	SNV	T->T
KEL3	Cytoplasmic protein of unknown function.	G980A	SNV	N->N
NPP1	Nucleotide pyrophosphatase/phosphodiesterase family member.	G446T	SNV	G->G
YHR033W	Putative protein of unknown function.	A+136T	SNV	Upstream

Table 4.S3: Oligonucleotides used in this study.

Name	Sequence (5'-> 3')
qPCR	
CTS1qf	CATAGCGCTTGGGTTTACATGG
CTS1qr	CTGAGTCGCCGGAGTCATAAG
DSE1qf	TCCCTTCTTACTGGCCTTAGTTGG
DSE1qr	ACTGGTTTACCGTCGCAGGATTG
DSE2qf	GCTCGGATGGCACTTGTTACG
DSE2qr	AGTTATAGTGGTGGCGGCATC
SCW11qf	GGTTACCCATCGTCTGGTATTC
SCW11qr	TAATCACCGGGAGCCTTCCAG
3' FW ACT1	GGCTTCTTTGACTACCTTCCA
3' RV ACT1	AGAAACACTTGTGGTGAACGA
Strain construction	
ACE2KOf	AAGAAATAACTAAAGAAATCTATAGGACCAAAAACGGTGT TAATACAATCCGTACGCTGCAGGTCGAC
ACE2KOr	TATTGTTACTATTATTTATTATGTTAATATCATGCATAGATA AATGTTCGGCATAGGCCACTAGTGGATCTG
ACE2fw	CAGGGAGACTCAAGCAACAG
ACE2rv	TGGCCCTTAAGACTACAGTG
Hph NT1 fw	AGACGTCGCGGTGAGTTCAG
Hph NT1 rv	CTCGCCGATAGTGGAAACCG
ACE2idf	CTTGGACGGCTTGACTTA
ACE2tagA	GTGCCTATTGATGATCTGGCGGAATGTCTGCCGTGCCATA GCCATGCCTTCACATATAGTGGCAATGTACCCTAAAGGTTG TG
ACE2tagB	GTTGAACATTCTTAGGCTGGTCGAATCATTTAGACACGGG CATCGTCCTCTCGAAAGGTGGGCAATGTACCCTAAAGGTT GTG
tagApUG	ACTATATGTGAAGGCATGGCTATGGCACGGCAGACATTCC GCCAGATCATCAATAGGCACCTTCGTACGCTGCAGGTCGAC
tagBpUG	CACCTTTCGAGAGGACGATGCCCGTGTCTAAATGATTCGAC CAGCCTAAGAATGTTCAACCTTCGTACGCTGCAGGTCGAC
pUGACE2r	TAGTGGGTTTTAGATCGTGTCTCCCAAAAGGTTGCGTCATT TAAAATCGTCACACTGTGTTTTACCAAAGGATGTGTGAAGC TGGTTTGTAGTAGTTAGCATAGGCCACTAGTGGATCTG
ACE2seqf	TCGCCTCGGATGTCAAATAC
ACE2hygidrv	TCTTCCGGTCTAACCAACAG
ACE2f	GAGAGCGGCTCGTCAGATAG
ACE2TARcheck	TCACTTGGAAACGCCTGCAATG
KanA	CGCACGTCAAGACTGTCAAG
KanB	TCGTATGTGAATGCTGGTCG

Table 4.S4: Relation between the number of homopolymeric (dA:dT)-stretches of length 8 or more and genomic features, compared to the background distribution. p-values result from applying Fisher's exact test under the null hypothesis that the occurrence of homopolymer stretches is independent of that of the genomic features.

In	Bases	%	Homopolymer (dA:dT) stretches ≥ 8 bp	%	p-value
Genome	12,071,326 bp	71.3%	6,254×	25.7%	$< 10^{-15}$
Genes	8,603,670 bp		1,607×		
Genome	12,071,326 bp	70.7%	6,254×	24.4%	$< 10^{-15}$
Coding sequence	8,528,636 bp		1,524×		
Genes	8,603,670 bp	99.2%	1,607×	95.4%	$< 10^{-15}$
Exons	8,537,723 bp		1,533×		
Genes	8,603,670 bp	23.4%	1,607×	18.8%	4.5×10^{-4}
Domains	2,009,787 bp		302×		

Supplementary Figures

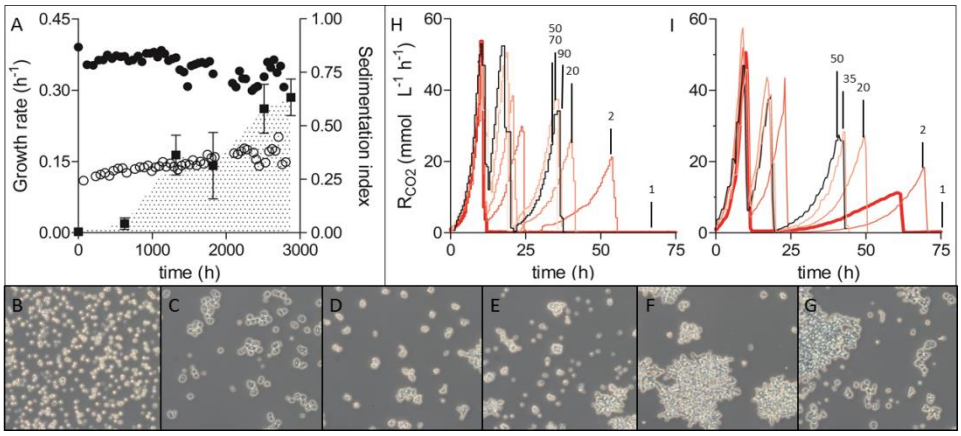

Figure 4.S1: Sequential batch cultivations in bioreactors on a mixture of glucose and galactose resulted in a culture containing multicellular *S. cerevisiae*. **A**- μ_{max} derived from the CO2 production of glucose consumption in the MY-glc/gal batches (●) and the μ_{max} of the MY-Gal batch (○) in evolution line 2. Intermediate biomass samples of the evolution were cultivated to stationary phase in shake flasks containing YP medium with 20 g l^{-1} glucose and were left to settle for 30 minutes in a 1 ml cuvette. Sedimentation indexes (■) represent the difference in OD_{660} within 30 min. The data represent the average and the mean deviation of duplicated experiments. **B**- Microscopic pictures of the evolution line 1 after 0, **C**- 626, **D**- 1321, **E**- 1824, **F**- 2513 and **G**- 2877 hours of evolution. **H**- Volumetric CO_2 production rate (mmol l^{-1} h^{-1}) of a cycle of two batches during evolution 1 and **I**- evolution 2. The numbers indicate the cycle in which the evolution is. Cycle 1 is after 0 hours, cycle 2 after 188 or 119 hours, cycle 20 after 1215 or 1168 hours, cycle 35 after 2133 hours, cycle 50 after 2408 or 2782 hours, cycle 70 after 3338 hours and cycle 90 after 4061 hours.

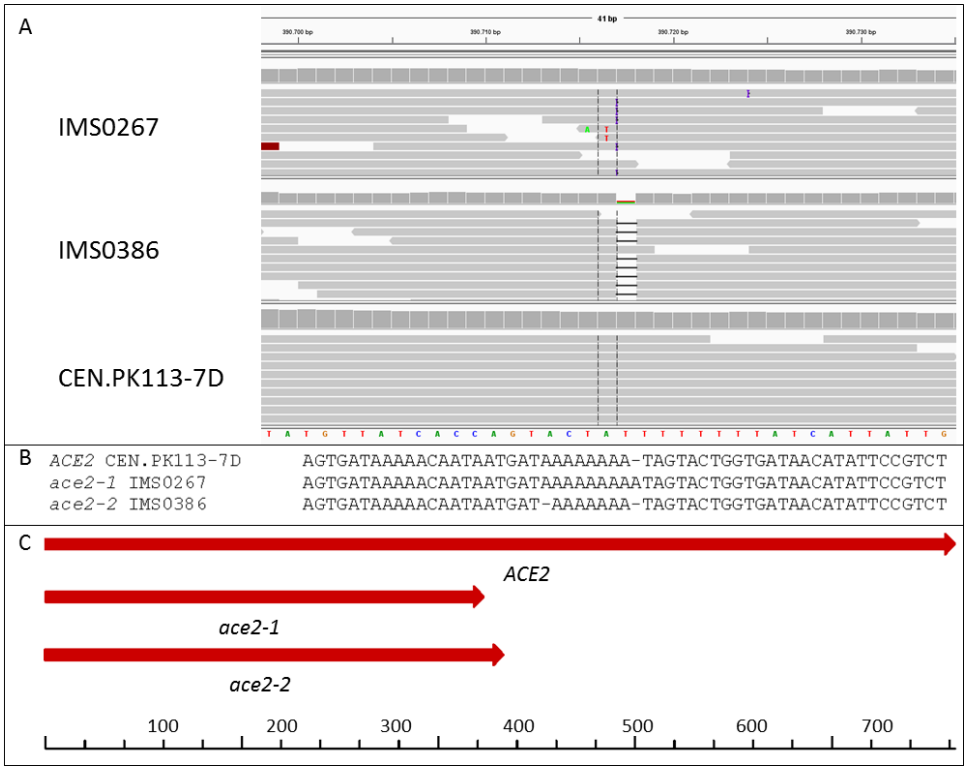

Figure 4.S2: Identification and characterization of *ACE2* mutations in IMS0267 and IMS0386 genome sequences. A- Visualization using IGV (http://www.broadinstitute.org/software/igv/home) of alignments of sequencing reads of CEN.PK113-7D, IMS0267 and IMS0386 mapped to the reference CEN.PK113-7D strains. **B**- Alignment of resequencing data (Sanger method) of the *ace2-1* and *ace2-2* alleles from IMS0267 and IMS0386 respectively. **C**- Schematic representation of the translated products of *ACE2*, *ace2-1* and *ace2-2*.

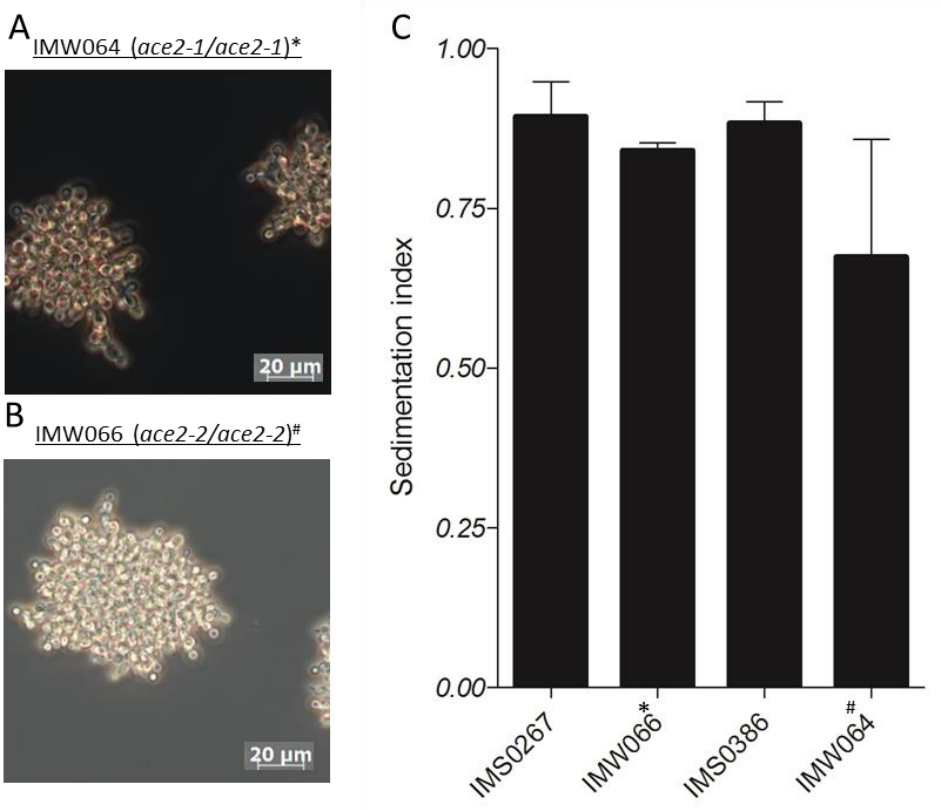

Figure 4.S3 Re-reverse engineering of the multicellular phenotype. Cellular morphology after replacing the *ACE2* allele with a mutated *ace2-2* allele in strains IMI220 and IMI221 resulted in **A**- IMW064 (*ace2-1/ace2-1*) and **B**- IMW066 (*ace2-1/ace2-2*) respectively. **C**- Sedimentation index of the evolved multicellular strains IMS0267 and IMS0386 and the re-reverse engineered mutants IMW064* and IMW066#. The sedimentation index indicates the fraction of cells that sediment in a cuvette after 30 min. The data represented are average ± mean deviation of duplicate biological replicates. * denotes strains constructed in the IMS0267 and # denotes strains constructed in the IMS0386 strain backgrounds.

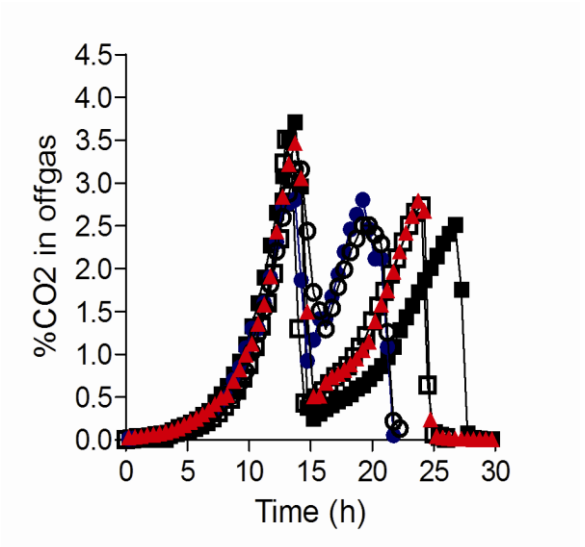

Figure 4.S4: Batch cultivations in bioreactors on a mixture of 20 g l^{-1} glucose and 20 g l^{-1} galactose revealed that additional mutations are responsible for the observed improved galactose consumption phenotype. Volumetric CO_2 production profiles of ■ CEN.PK113-7D (*MAT***a** *ACE2*), □ CEN.PK122 (*MAT***a**/*MAT*α *ACE2*/*ACE2*), ● IMS0267 (*ace2-1*/*ace2-1*), ○ IMS0386 (*ace2-2*/*ace2-2*) and ▲ IMD014 (*MAT***a**/*MAT*α *ace2-2*/*ace2-2*) cultivated in an anaerobic bioreactor containing synthetic medium supplemented with 20 g l^{-1} glucose and 20 g l^{-1} galactose. One representative experiment is shown of at least a duplicate experiment.

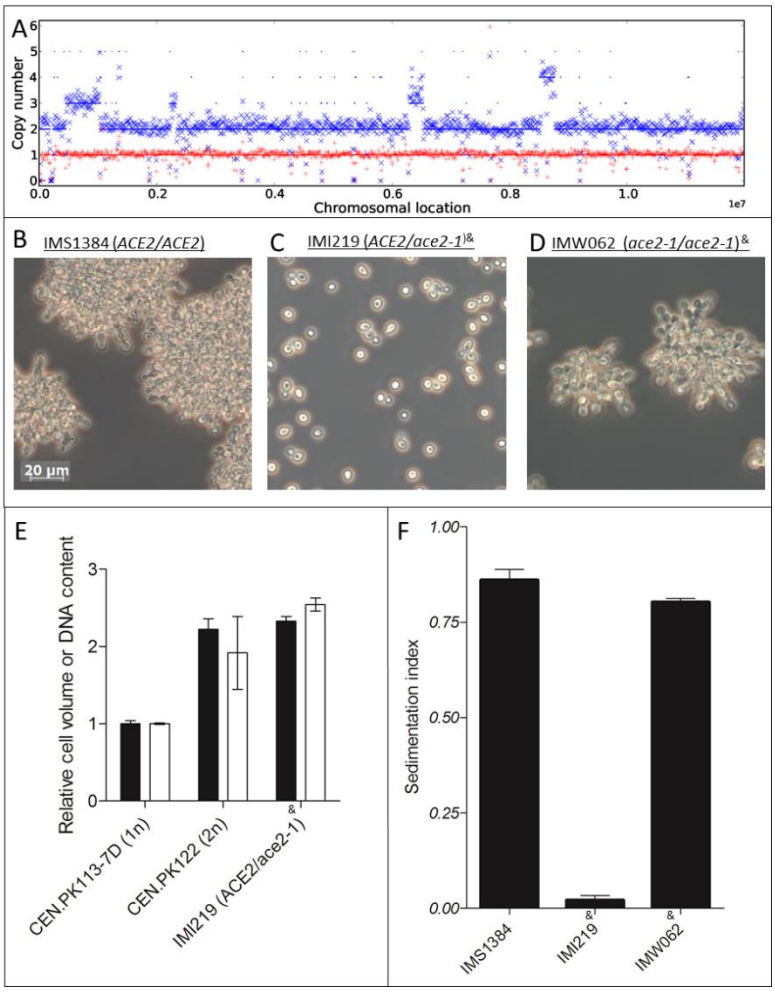

Figure 4.S5 Analysis of an additional isolate IMS1384 from evolution 2. A- Prediction of DNA content in the evolved strain IMS1384 using the Magnolya algorithm[200]. + (red) indicates the ploidy of the ancestral genome and x (blue) indicates the ploidy of the evolved genome **B**- Cellular morphology of strains IMS1384 (*MAT***a**/*MAT***a** *ace2-1*/*ace2-1*), **C**- IMI219 (*MAT***a**/*MAT***a** *ACE2*/*ace2-1*) and **D**- IMW062 (*MAT***a**/*MAT***a** *ACE2*/*ace2-1*). **E**- Determination of the cell size (white bar) and DNA content measurements (black bar) of strains CEN.PK113-7D (*MAT***a**), CEN.PK122 (*MAT***a**/*MAT*α) and IMI219 (*MAT***a**/*MAT***a** *ACE2*/*ace2-1-HphNT1*). **F**- Sedimentation index of the evolved multicellular strain IMS1384, of the reverse engineered strain IMI219 and of the re-reverse engineered mutant IMW062. The sedimentation index indicates the fraction of cells that sediment in a cuvette after 30 min. Strain IMI219 is a non sedimenting strain derived from IMS1384, strain IMW062 is again a multicellular strain derived from IMI219. The data represented are mean and mean deviation of duplicate biological replicates.

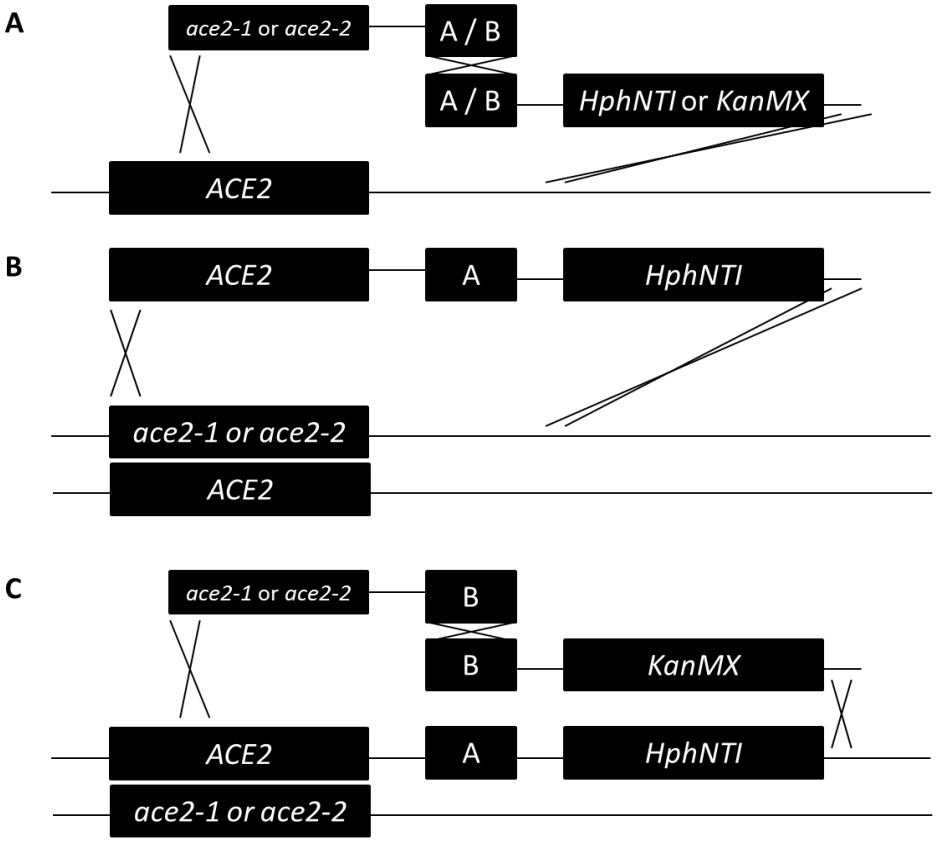

Figure 4.S6 Overview of mode of construction of different mutants. A- Mutants created in the CEN.PK113-7D background were created by transforming two constructs into yeast. The first construct can contain any *ACE2* allele with a linker sequence A or B and the second construct contains the same linker sequence together with a dominant marker and a sequence homologues to a region 204 bp upstream of *ACE2*. **B**- Mutants created in the evolved backgrounds of IMS0267, IMS0386 and also IMS1384 were created by transforming one construct into yeast containing the non-mutated *ACE2* allele with a linker sequence A, the hygromycin resistance gene and a sequence homologues to a region 204 bp upstream of *ACE2*. **C**- To exclude the possibility of transformation-associated selection of unicellular mutants, the construct inserted in A was re-exchanging by two constructs. The first construct can contain any of the mutated *ace2* alleles with a linker sequence B and the second construct contains the same linker sequence together with a G418 resistance gene and a sequence homologous to a region 204 bp upstream of *ACE2*.

Chapter 5

Carbon dioxide fixation by Calvin-cycle enzymes improves ethanol yield in yeast

Víctor Guadalupe-Medina, H. Wouter Wisselink, Marijke A.H. Luttik, Erik de Hulster, Jean-Marc Daran, Jack T. Pronk and Antonius J.A. van Maris

Essentially as published in Biotechnology for Biofuels, August 2013, Vol. 6, p. 125-195 (doi:10.1186/1754-6834-6-125).

Abstract

Redox-cofactor balancing constrains product yields in anaerobic fermentation processes. This challenge is exemplified by the formation of glycerol as major by-product in yeast-based bioethanol production, which is a direct consequence of the need to reoxidize excess NADH and causes a loss of conversion efficiency. Enabling the use of CO_2 as electron acceptor for NADH oxidation in heterotrophic microorganisms would increase product yields in industrial biotechnology. A hitherto unexplored strategy to address this redox challenge is the functional expression in yeast of enzymes from autotrophs, thereby enabling the use of CO_2 as electron acceptor for NADH reoxidation. Functional expression of the Calvin cycle enzymes phosphoribulokinase (PRK) and ribulose-1,5-bisphosphate carboxylase (Rubisco) in *Saccharomyces cerevisiae* led to a 90% reduction of the by-product glycerol and a 10% increase in ethanol production. Co-expression of the *Escherichia coli* chaperones GroEL and GroES was key to successful expression of CbbM, a form-II Rubisco from the chemolithoautotrophic bacterium *Thiobacillus denitrificans* in yeast. Our results demonstrate functional expression of Rubisco in a heterotrophic eukaryote and demonstrate how incorporation of CO_2 as a co-substrate in metabolic engineering of heterotrophic industrial microorganisms can be used to improve product yields. Synthetic biology should allow for rapid insertion of this 4-gene expression cassette in industrial yeast strains to improve production, not only of 1st and 2nd generation ethanol production, but also of other renewable fuels or chemicals.

Background

The yeast *Saccharomyces cerevisiae* is not only used for the large-scale production of fuel ethanol[53], but also for industrial production of a broad and rapidly expanding range of other chemical compounds from renewable carbohydrate feedstocks[28,227]. In anaerobic, ethanol-producing cultures of *S. cerevisiae*, excess NADH generated from biosynthetic reactions, such as NAD^+-dependent oxidative decarboxylations involved in synthesis of the precursors acetyl-CoA and 2-oxoglutarate, is reoxidized by reducing part of the sugar substrate to glycerol[34]. In growing anaerobic yeast cultures, glycerol production typically accounts for 4-10% of the total sugar consumption and therefore has a significant impact on ethanol yields and process economy in both 1st and 2nd generation large-scale bioethanol production[69,154].

Using CO_2 as electron acceptor for the oxidation as NADH would be a highly attractive metabolic engineering strategy, in particular when CO_2 reduction can be coupled to the formation of the product of interest. Functional expression of the Calvin cycle enzymes phosphoribulokinase (PRK) and ribulose-1,5-bisphosphate carboxylase (Rubisco) in *S. cerevisiae* should enable the coupling of CO_2, a major product of alcoholic fermentation, to ribulose-5-phosphate, a normal intermediate of the *S. cerevisiae* pentose-phosphate pathway (Fig. 5.1). The resulting two molecules of 3-phosphoglycerate can subsequently be converted to 2 molecules each of ethanol and CO_2, with the concomitant net oxidation of 2 molecules of NADH to NAD^+ (Fig. 5.1). When ribulose-5-phosphate is formed via the oxidative pentose-phosphate pathway (Fig. 5.1), this route results in a transhydrogenase-type conversion of redox cofactors ($NADP^+ + NADH \rightarrow NADPH + NAD^+$). Since the total amount of NADPH required in biosynthesis is smaller than the amount of NADH generated[35], such a transhydrogenase-like activity cannot fully replace glycerol formation as a mechanism for reoxidizing biosynthetic NADH. However, no such constraint exists when ribulose-5-phosphate is formed from intermediates of glycolysis via the rearrangement reactions of the non-oxidative pentose-phosphate pathway (Fig. 5.1). A theoretical analysis shows that complete replacement of glycerol production with CO_2 incorporation through PRK and Rubisco can increase the ethanol yield of sugar by as much as 14% (Fig. 5.1).

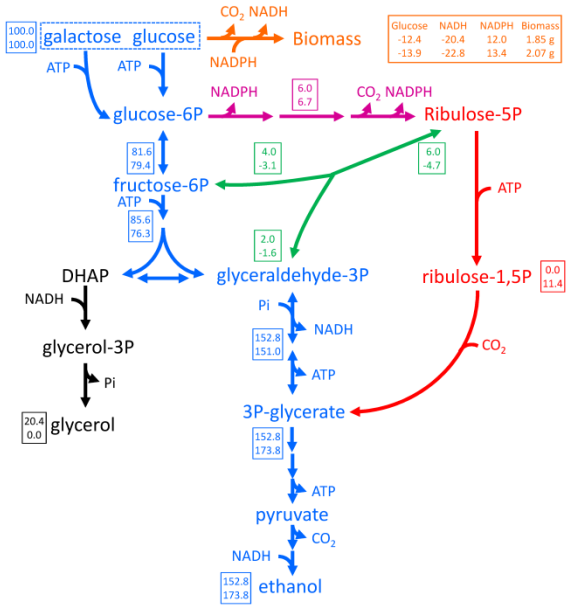

Figure 5.1: Schematic representation of central carbon metabolism and the introduced Calvin-cycle enzymes in *Saccharomyces cerevisiae*. Orange: Formation of biomass and NADH from glucose and NADPH. Stoichiometries are according to Verduyn *et al.* (1990)[35]; Blue: Redox-neutral, ATP-yielding alcoholic fermentation of glucose and galactose via the Embden-Meyerhof-Parnas glycolysis and Leloir pathways, respectively; Magenta: NADPH generation via the oxidative part of the pentose-phosphate pathway; Green: rearrangement of sugar-phosphate carbon skeletons via the non-oxidative pentose-phosphate pathway; Black: NADH oxidation by formation of glycerol through glycerol-3-phosphate dehydrogenase and glycerol-3-phosphatase; Red: heterologously expressed Calvin-cycle enzymes phosphoribulokinase and Rubisco. Numbers in boxes represents the distribution of carbon along the different pathways (in mmol) normalized for a combined glucose and galactose uptake of 100 mmol for a wild-type, glycerol-producing reference strain (top) and for a scenario in which the alternative pathways via the Calvin cycle enzymes completely replace glycerol formation as the mechanism for reoxidizing NADH formed in biosynthetic reactions (bottom). In the scenario with the Calvin cycle enzymes, ribulose-5-phosphate was assumed to be preferentially derived from the oxidative reactions of the pentose phosphate pathway. Once the generation of NADPH from these reactions matched the requirement for NADPH in biosynthesis, further ribulose-5-phosphate was derived from glycolytic intermediates via the non-oxidative pentose-phosphate pathway rearrangement reactions. The biomass yield on ATP was assumed to be identical for both scenarios.

The PRK gene from *Spinacia oleracea*[228] has previously been expressed in the yeast *Pichia pastoris*[229] and is therefore an interesting candidate for heterologous expression in *S. cerevisiae*. For Rubisco, a key enzyme in the Calvin cycle for autotrophic carbon fixation, three catalytically active forms have been described[230,231]. Prokaryotic form-II Rubisco's are encoded by single structural genes and several have been heterologously expressed in *E. coli*[232,233]. Functional expression of form-II Rubisco's in *E. coli* was shown to be strongly stimulated by the *E. coli* protein-folding chaperones GroEL and GroES[234] and expression of *Hydrogenovibrio marinus* Rubisco in *E. coli* was further stimulated by co-expression of the CbbO and CbbQ chaperones of the donor organism[235]. Very recently the structure of GroEL/GroES encapsulating Rubisco was visualized by cryo-electron microscopy[236]. Eukaryotes such as *S. cerevisiae* harbour a chaperone couple (Hsp60/Hsp10) that structurally and functionally resemble GroEL/GroES. However, these proteins are located in the mitochondria, whereas a role in Rubisco expression would require their activity in the cytosol.

In this study we investigated how to achieve functional expression of PRK and Rubisco in yeast. In view of the envisioned benefit of being encoded by single structural genes, a prokaryotic form-II Rubisco gene was expressed in *S. cerevisiae* in combination with the PRK gene from *Spinacia oleracea*. Both the promoters and coding regions for genes required for glycerol formation were left unchanged compared to the reference strain. Subsequently, the impact of the resulting CO_2 incorporation on product formation was studied, with special emphasis on the yields of ethanol and the undesired by-product glycerol.

Result and discussion

Chaperone-mediated functional expression of Rubisco in Saccharomyces cerevisiae. To study a possible requirement of heterologous chaperones for expression of Rubisco in *S. cerevisiae*, the form-II Rubisco-encoding *cbbM* gene from *T. denitrificans*[237] was codon-optimized for expression in *S. cerevisiae* and expressed from a centromeric vector, both alone and in combination with expression cassettes for the codon-optimized *E. coli groEL/groES*[238] and/or *T. denitrificans cbbO2/cbbQ2* genes[239,240]. Functional expression of *T. denitrificans* Rubisco in *S. cerevisiae*, as indicated by ribulose-1,5-bisphosphate-dependent $^{14}CO_2$ fixation by yeast cell extracts, was only observed upon co-expression of *E. coli* GroEL/GroES (Fig. 5.2). Co-expression of CbbO2/CbbQ2 did not result in a further increase of Rubisco activity (Fig. 5.2). Co-expression of bacterial chaperones has previously been shown to improve heterologous protein expression in *Pichia pastoris* and insect cells[241,242].

The positive effect of GroEL/GroES on Rubisco expression in *S. cerevisiae* demonstrates the potential value of co-expression of heterologous chaperones for metabolic pathway engineering, especially when synthetic biology is used for expression of prokaryotic enzymes in the cytosol of eukaryotes.

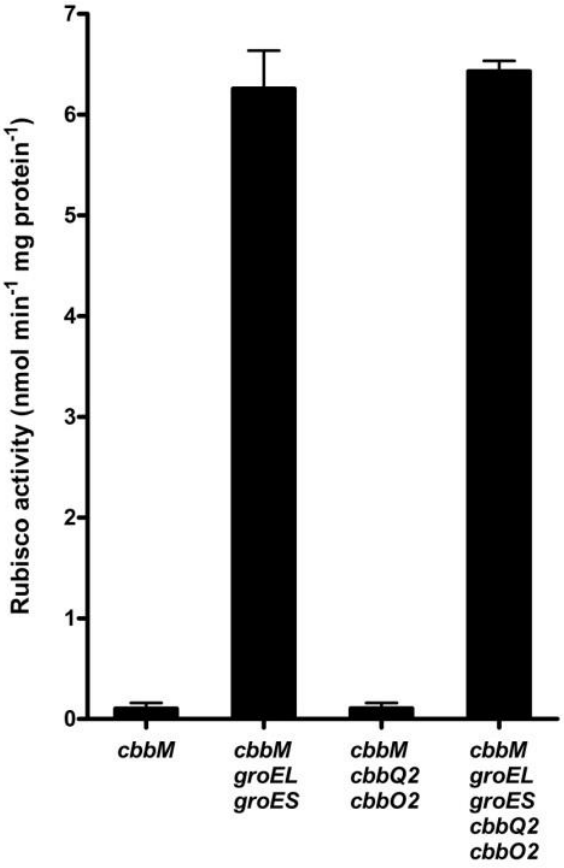

Figure 5.2 Specific ribulose-1,5-bisphosphate carboxylase (Rubisco) activity in cell extracts of *S. cerevisiae* expressing Rubisco form II CbbM from *T. denitrificans*, either alone (IMC033) or in combination the *E. coli* chaperones GroEL/GroES[238] (IMC035), the *T. denitrificans* chaperones CbbO2/CbbQ2[239] (IMC034) or all four chaperones (IMC014). Heterologously expressed genes were codon optimized for expression in yeast and expressed from a single centromeric vector. Biomass samples were taken from anaerobic batch cultures on synthetic media (pH 5.0, 30 °C), sparged with nitrogen and containing 20 g l^{-1} glucose as carbon source. Rubisco activities, measured as $^{14}CO_2$-fixation in cell extracts, in a wild-type reference strain and in *S. cerevisiae* strains expressing cbbM and cbbM-cbbQ2-cbbO2 were below the detection limit of the enzyme assay (0.2 nmol CO_2 min^{-1} mg $protein^{-1}$).

Functional expression of phosphoribulokinase in *Saccharomyces cerevisiae*. The *Spinacia oleracea* phosphoribulokinase (PRK) gene[228], which has previously been expressed in the yeast *Pichia pastoris*[229], was integrated together with *E. coli groEL/groES* and *T. denitrificans cbbO2/cbbQ2* into the *S. cerevisiae* genome at the *CAN1* locus, under control of the galactose-inducible *GAL1* promotor. This resulted in high PRK activities (approximately 15 µmol mg protein^{-1} min^{-1}) in cell extracts of *S. cerevisiae* strain IMU033 taken from carbon-limited chemostat cultures on a mixture of glucose and galactose (Table 5.1). Although relatively high background activities were measured in the reference strain without PRK (IMU032), this activity does not contribute to pathway activity (see below). We therefore assumed that the background activity observed in the reference strain was caused by an impurity in one of the chemicals used in the assay and did not reflect formation of ribulose-1,5-bisphosphate. The engineered strain IMU033, which additionally carried the centromeric expression cassette for *T. denitrificans* Rubisco, was used to quantitatively analyse the physiological impacts of the expression of Rubisco and PRK.

Table 5.1 Increased ethanol yields on sugar of an *S. cerevisiae* strain expressing phosphoribulokinase (PRK) and Rubisco. Physiological analysis of *S. cerevisiae* IMU033 expressing PRK and Rubisco and the isogenic reference strain IMU032 in anaerobic chemostat cultures, grown at a dilution rate of 0.05 h^{-1} on a synthetic medium (pH 5.0) supplemented with 12.5 g l^{-1} glucose and 12.5 g l^{-1} galactose as carbon sources. To assess the impact of CO_2 concentration, chemostat cultures were run sparged either with pure nitrogen gas or with a blend of 10% CO_2 and 90% nitrogen. Results are represented as average ± mean deviations of data from independent duplicate chemostat experiments. Data pairs labelled with the same subscripts (^{a,a}, ^{b,b}, etc.) are considered statistically different in a standard t-test ($p < 0.02$).

Yeast strain	IMU032 (reference strain)		IMU033 (expressing PRK and Rubisco)	
CO_2 in inlet gas (%)	0	10	0	10
CO2 in outlet gas (%)	0.89 ± 0.03	10.8 ± 0.0	1.02 ± 0.00	10.8 ± 0.1
Phosphoribulokinase (μmol mg protein^{-1} min^{-1})	0.58 ± 0.09	0.51 ± 0.12	14.4 ± 1.5	15.2 ± 1.0
Rubisco (nmol mg protein^{-1} min^{-1})	< 0.2*	< 0.2	4.59 ± 0.30	2.67 ± 0.28
Biomass yield on sugar (g g^{-1})	0.083 ± 0.000^{a}	0.084 ± 0.000^{b}	0.093 ± 0.001^{a}	0.095 ± 0.000^{b}
Ethanol yield on sugar (mol mol^{-1})	1.56 ± 0.03^{c}	1.56 ± 0.02^{d}	1.73 ± 0.02^{c}	1.73 ± 0.01^{d}
Glycerol yield on sugar (mol mol^{-1})	0.14 ± 0.00^{e}	0.12 ± 0.00^{f}	0.04 ± 0.00^{e,g}	0.01 ± 0.00^{f,g}

*Detection limit of enzyme activity assay.

Carbon dioxide as electron acceptor in anaerobic chemostat cultures of *Saccharomyces cerevisiae*. Quantitative physiological analysis is facilitated by the constant and highly reproducible process conditions in steady-state chemostat cultures[243,244]. Therefore, ethanol and glycerol yields of PRK- and Rubisco-expressing *S. cerevisiae* were compared to those of an isogenic reference strain in anaerobic, sugar-limited chemostats. In nitrogen-sparged cultures, the glycerol yield on sugar in the strain expressing both Calvin-cycle enzymes was 68% lower than in the reference strain, while ethanol and biomass yields on sugar were 11% and 12% higher, respectively (Table 5.1). To investigate whether the low affinity of *T. denitrificans* form-II Rubisco for CO_2 (K_{CO_2} = 0.26 mM[237]) limited its *in vivo* activity in the nitrogen-sparged cultures, additional chemostats were sparged with a 10%/90% blend of CO_2 and N_2. Indeed, this CO_2 supplementation resulted in a further decrease of the glycerol yield to a value below 10% of that of the reference strain (Table 5.1). Co-expression of Rubisco and chaperones without co-expression

of PRK (strain IMC014) did not result in decreased glycerol yield (0.13 mol mol^{-1}) compared to the reference strain IMU032 (0.12 mol mol^{-1}) in carbon-limited chemostat cultures supplemented with CO_2. This observation confirmed that expression of a heterologous phosphoribulokinase (PRK) gene is required for in vivo carbon fixation via Rubisco in yeast.

Carbon dioxide as electron acceptor in anaerobic batch fermentations. Since industrial-scale ethanol production is routinely performed in batch fermentations[155], the impact of the expression of PRK and Rubisco was also investigated in anaerobic, CO_2-supplemented batch cultures (Fig. 5.3). Galactose was used as the carbon source for these experiments to enable efficient expression of PRK from the *GAL1* promoter. Specific growth rates of the engineered and reference strains on galactose in these anaerobic cultures were not significantly different (Fig. 5.3) and in good agreement with values reported elsewhere for this yeast strain family[245]. Apparently, expression of the Calvin cycle enzymes did not represent a major metabolic burden. Consistent with the observations in chemostat cultures, expression of the two Calvin cycle enzymes reduced glycerol formation in the batch cultures by 60% and increased the ethanol yield by 8% (Fig. 5.3 e-f).

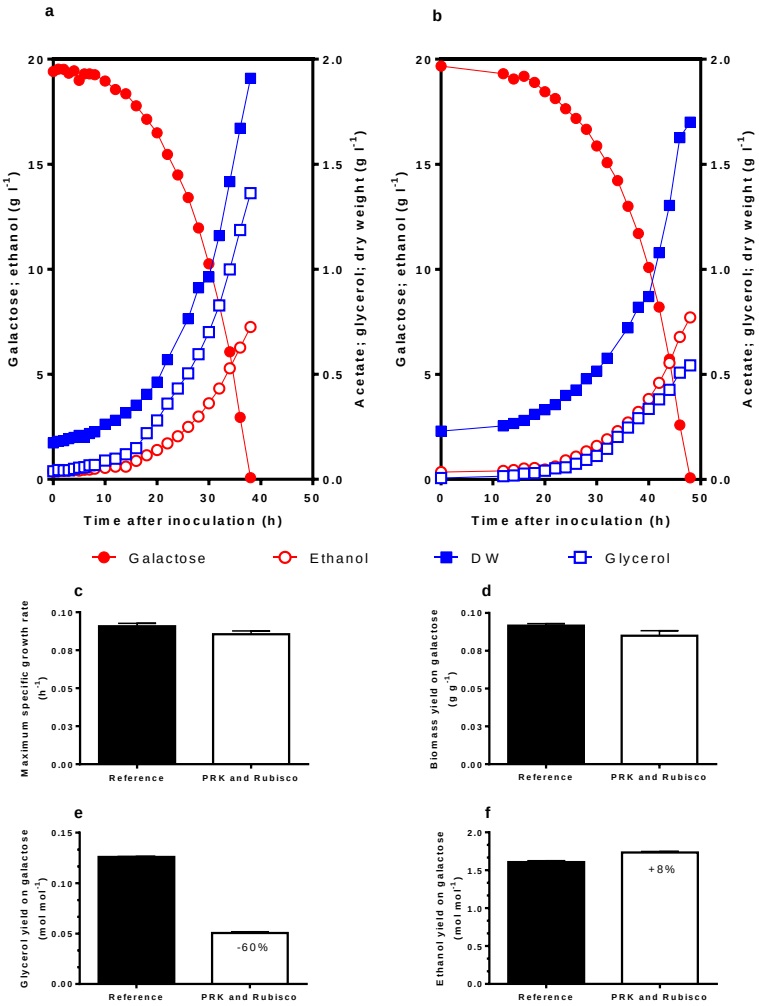

Figure 5.3. Physiological impact of expression of Calvin cycle enzymes on growth, substrate consumption and product formation in galactose-grown anaerobic batch cultures of *S. cerevisiae*. a: growth curves of isogenic reference strain *S. cerevisiae* IMU032, **b:** growth curves of *S. cerevisiae* IMU033 expressing PRK and Rubisco. Growth conditions: T = 30 °C, pH 5.0, 10% CO_2 in inlet gas. Symbols: ●, galactose; ○, ethanol; ■, Dry weight; □, glycerol. Each graph represents values for one of two independent replicate experiments, whose growth kinetic parameters differed by less than 5%. **c-f:** Calculated parameters: Maximum specific growth rate (**c**), biomass yield (**d**), glycerol yield (**e**), and ethanol yield (**f**) on galactose of the isogenic *S. cerevisiae* reference (black bars) and strain expressing PRK and Rubisco (white bars). Results are represented as average ± mean deviations of data from independent duplicate cultures. Values inside the white bars represent statistically significant differences in a standard t-test (p value < 0.02) relative to the reference strain.

Conclusion

This study provides a compelling proof of principle for the replacement of glycerol formation as the predominant redox sink in anaerobic yeast metabolism by PRK- and Rubisco-mediated incorporation of CO_2 into yeast central carbon metabolism. The loss of sugar feedstock due to glycerol production in industrial bioethanol processes has been estimated at 4% of the consumed sugar[69]. If expression of PRK and Rubisco in industrial yeast strains were to completely eliminate this loss, this could enable an additional production of 5 billion liters of ethanol from the amount of sugar used for the 2011 global ethanol production of 110 billion liters[53]. Use of CO_2 as an external electron acceptor offers important advantages over previously proposed strategies for reducing glycerol production in yeast-based bioethanol production. Optimizing the redox cofactor specificity of nitrogen assimilation in *S. cerevisiae*[69] only enables a partial reduction of glycerol production and its impact further depends on the nitrogen sources present in industrial feedstocks. Similarly, a metabolic engineering strategy that enables NADH-dependent reduction of acetic acid[154] to ethanol is dependent on the presence of acetic acid in industrial feedstocks. Further optimization of PRK and Rubisco gene expression and regulation in *S. cerevisiae* should enable the design and construction of DNA cassettes that can be easily introduced in the genomes of industrial yeast strains. Since ribulose-5-phosphate is also an intermediate in pentose metabolism by engineered *S. cerevisiae* strains[115], this approach should also be readily applicable to the yeast-based conversion of lignocellulosic hydrolysates. The observed stimulatory effect of CO_2 on the engineered strains will not hinder application of this concept in industrial bioethanol production, since large-scale processes for bioethanol production are characteristically CO_2 saturated.

Our results illustrate how metabolic engineering strategies based on the functional integration of extensively studied reactions in the central carbon metabolism of distantly related organisms enables the optimization of product yields in industrial biotechnology. Although the present study focuses on ethanol production by yeast, functional integration of autotrophic carbon-fixing enzymes in the metabolic networks of industrial microorganisms should also enable optimization of yields of other existing and novel products whose synthesis results in a net positive ATP yield.

Methods

Construction of the expression modules. Phosphoribulokinase (*PRK*) cDNA from *Spinacia oleracea* (spinach)[229] (accession number: X07654.1) was PCR-amplified

using Phusion Hot-Start Polymerase (Finnzymes, Landsmeer, The Netherlands) and the oligonucleotides XbaI_prk_FW2 and RV1_XhoI_prk (Table 5.2), and was ligated in pCR®-Blunt II-TOPO® (Life Technologies Europe BV, Bleiswijk, The Netherlands). After restriction by XbaI and XhoI, the *PRK*-containing fragment was ligated into pTEF424[246]. The *TEF1*p was later replaced by *GAL1*p from plasmid pSH47[105] by XbaI and SacI restriction/ligation, creating plasmid pUDE046 (Table 5.3).

Table 5.2 Oligonucleotides used in this study.

Number	Name	Sequence (5' to 3')	Purpose
	Primers used for cloning		
1	XbaI_prk_FW2	TGACATCTAGATGTCACAACAACAAACAATTG	Cloning of PRK into pUDE046.
2	RV1 XhoI prk	TGACATCTAGATGTCACAACAACAAACAATTG	Cloning of PRK into pUDE046.
	Primers used for *in vivo* plasmid assembly		
3	HR-cbbM-FW-65	TTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGGTACAGCTGGAGCTCAGTTTATCATTATC	Rubisco *cbbM* cassette for plasmids pUDC075, pUDC099, and pUDC100.
4	HR-cbbM-RV-65	GGAATCTGTGTAGTATGCCTGGAATGTCTGCCGTGCCATAGCCATGTATGCTGATATGTCGGTACCGGCCGCAAATTAAAG	Rubisco *cbbM* cassette for plasmids pUDC075, pUDC099, and pUDC100
5	linker-cbbO2-pRS416	ATCACTCTTACCAGGCTAGGACGACCCTACTCATGTATTGAGATCGACGAGATTTCTAGGCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGC	Linker fragment for assembly of plasmid pUDC099.
6	linker-cbbM-GroEL	GACATATCAGCATACATGGCTATGGCACGGCAGACATTCCAGGCATACTACACAGATTCCATCACTCTTACCAGGCTAGGACGACCCTACTCATGTATTGAGATCGACGAGATTTCTAGG	Linker fragment for assembly of plasmid pUDC100.
	Primers used for *in vivo* integration assembly		
7	FW pTDH3-HR-CAN1up	GTTGGATCCAGTTTTTAATCTGTCGTCAATCGAAAGTTTATTTCAGAGTTCTTCAGACTTCTTAACTCCTGTAAAAACAAAAAAAAAAAAAGGCATAGCAAGCTGGAGCTCAGTTTATC	1st cloning expression cassette linker fragment between *CAN1* upstream and *PRK* expression cassette (IMI229), and

			CAN1up-linker and *KlLEU2* expression cassette (IMI232).
8	RV linker-iHR2B	AGATATACTGCAAAGTCCGGAGCAACA GTCGTATAACTCGAGCAGCCCTCTACTT TGTTGTTGCGCTAAGAGAATGGACC	1st cloning fragment: linker fragment between CAN1up-linker and *PRK* expression cassette (IMI229).
9	RV linker-iHR6	GCTATGACCATGATTACGCCAAGCGCGC AATTAACCCTCACTAAAGGGAACAAAAG CTGGTTGCGCTAAGAGAATGGACC	1st cloning fragment: linker fragment between CAN1up-linker and *KlLEU2* expression cassette (IMI232).
10	FW pGAL1-prk HR2B	CAACAAAGTAGAGGGCTGCTCGAGTTA TACGACTGTTGCTCCGGACTTTGCAGTA TATCTGCTGGAGCTCTAGTACGGATT	2nd cloning fragment: $GAL1_p$-*PRK*-$CYC1_t$ expression cassette (IMI229) from pUDE046.
11	RV CYC1t-prk HR2	GGAATCTGTGTAGTATGCCTGGAATGT CTGCCGTGCCATAGCCATGTATGCTGAT ATGTCGTACCGGCCGCAAATTAAAG	2nd cloning fragment: $GAL1_p$-*PRK*-$CYC1_t$ expression cassette (IMI229) from pUDE046.
12	FW HR2-cbbQ2-HR3	GACATATCAGCATACATGGCTATGG	3rd cloning fragment: $PGI1_p$-*cbbQ2*-$TEF2_t$ cassette (IMI229).
13	RV HR2-cbbQ2-HR3	GGACACGCTTGACAGAATGTCAAAGG	3rd cloning fragment: $PGI1_p$-*cbbQ2*-$TEF2_t$ cassette (IMI229).
14	FW HR3-cbbO2-HR4	CGTCCGATATGATCTGATTGG	4th cloning fragment: $PGK1_p$-*cbbO2*-$ADH1_t$ cassette (IMI229).
15	RV HR3-cbbO2-HR4	CCTAGAAATCTCGTCGATCTC	4th cloning fragment: $PGK1_p$-*cbbO2*-$ADH1_t$ cassette (IMI229).
16	FW HR4-GroEL-HR5	ATCACTCTTACCAGGCTAGG	5th cloning fragment $TEF1_p$-*groEL*-$ACT1_t$ cassette (IMI229).
17	RV HR4-GroEL-HR5	CTGGACCTTAATCGTGTGCGCATCCTC	5th cloning fragment: $TEF1_p$-*groEL*-$ACT1_t$ cassette (IMI229).
18	FW HR5-GroES-HR6	CCGTATAGCTTAATAGCCAGCTTTATC	6th cloning fragment: $TPI1_p$-*groES*-$PGI1_t$ cassette (IMI229).
19	RV HR5-	GCTATGACCATGATTACGCCAAGC	6th cloning fragment:

	GroES-HR6		*TPI1*$_p$-*groES-PGI1*$_t$ cassette (IMI229).
20	FW HR6-LEU2-CAN1dwn	CCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCCTGTGAAGATCCCAGCAAAG	7th (IMI229) or 2nd (IMI232) cloning fragment: *KlLEU2* cassette from pUG73.
21	RV LEU2 HR-CAN1	AGCTCATTGATCCCTTAAACTTTCTTTTCGGTGTATGACTTATGAGGGTGAGAATGCGAAATGGCGTGGAAATGTGATCAAAGGTAATAAAACGTCATATATCCGCAGGCTAACCGGAAC	7th (IMI229) or 2nd (IMI232) cloning fragment: *KlLEU2* cassette from pUG73.
	Primers used for verification of the *in vivo* assembled constructs		
22	m-PCR-HR1-FW	GGCGATTAAGTTGGGTAACG	Diagnostic for assembly of plasmids pUDC075, pUDC099, and pUDC100.
23	m-PCR-HR1-RV	AACTGAGCTCCAGCTGTACC	Diagnostic for assembly of plasmids pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
24	m-PCR-HR2-FW	ACGCGTGTACGCATGTAAC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229
25	m-PCR-HR2-RV	GCGCGTGGCTTCCTATAATC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229
26	m-PCR-HR3-FW	GTGAATGCTGGTCGCTATAC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
27	m-PCR-HR3-RV	GTAAGCAGCAACACCTTCAG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
28	m-PCR-HR4-FW	ACCTGACCTACAGGAAAGAG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100,

			and integration in strain IMI229.
29	m-PCR-HR4-RV	TGAAGTGGTACGGCGATGC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
30	m-PCR-HR5-FW	ATAGCCACCCAAGGCATTTC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
31	m-PCR-HR5-RV	CCGCACTTTCTCCATGAGG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
32	m-PCR-HR6-FW	CGACGGTTACGGTGTTAAG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
33	m-PCR-HR6-RV	CTTCCGGCTCCTATGTTGTG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.

Rubisco form II gene *cbbM* from *T. denitrificans*[237] flanked by KpnI and SacI sites was codon optimized[247] (accession number: KC699554), synthesized at GeneArt (Life Technologies Europe BV), and ligated into pPCR-Script. The *cbbM*-containing fragment was ligated into the BamHI and SacI restricted vector pGPD_426[246] creating plasmid pBTWW002. The *cbbM* expression cassette was transferred into pRS416 using KpnI and SacI, yielding pUDC098.

Expression cassette of the specific Rubisco form II chaperones from *T. denitrificans cbbQ2* and *cbbO2*[239], and chaperones *groEL* and *groES*[238] from *E. coli* were codon optimized[247] (accession numbers: KC699555 and KC699556, respectively). The expression cassettes contained a yeast constitutive promoter and terminator, flanking the codon optimized gene. The cassette was flanked by unique 60-bp regions obtained by randomly combining bar-code sequences used in the *Saccharomyces* Genome Deletion Project[248] and an EcoRV site (GeneArt). The expression cassettes were inserted in plasmid pMK-RQ (GeneArt) using the SfiI

cloning sites yielding pUD230 (*PGI1*$_p$-*cbbQ2-TEF2*$_t$), pUD231 (*PGK1*$_p$-*cbbO2-ADH1*$_t$), pUD232 (*TEF1*$_p$-*groEL-ACT1*$_t$), and pUDE233 (*TPI1*$_p$-*groES-PGI1*$_t$) (Table 5.3). The expression cassette *TDH3*$_p$-*cbbM-CYC1*$_t$ was PCR-amplified from plasmid pBTWW002 using Phusion Hot-Start Polymerase (Finnzymes) and primers HR-cbbM-FW-65 and HR-cbbM-RV-65 in order to incorporate the 60-bp region for recombination cloning.

Table 5.3 Plasmids used in this study.

Name	Relevant genotype	Source/reference
pFL451	*AOX1*$_p$-prk (Spinach)-*AOX1*$_t$ (*HIL2*$_p$-D2 *HIS4* Amp centromeric)	Brandes *et al.* (1996)[229].
pCR®-Blunt II-TOPO	*bla*	Life Technologies Europe BV.
pTEF424_TEF	*TRP1* 2μ *bla*	Mumberg *et al.* (1995)[246].
pSH47	*URA3 CEN6 ARS4 GAL1*$_p$-*cre-CYC1*$_t$ *bla*	Güldener *et al.* (1996)[105].
pUD0E46	*TRP1* 2μ *GAL1*$_p$-*prk-CYC1*$_t$ *bla*	This study.
pPCR-Script	*bla*	Life Technologies Europe BV.
pGPD_426	*URA3* 2μ *bla*	Mumberg *et al.* (1995)[246].
pRS416	*URA3 CEN6 ARS4 bla*	Mumberg *et al.* (1995)[246].
pBTWW002	*URA3* 2μ *TDH3*$_p$-*cbbM-CYC1*$_t$ *bla*	This study.
pUDC098	*URA3 CEN6 ARS4 TDH3*$_p$-*cbbM-CYC1*$_t$ *bla*	This study.
pMK-RQ	*nptII*	Life Technologies Europe BV.
pUD230	*PGI1*$_p$-*cbbQ2-TEF2*$_t$ *nptII*	Life Technologies Europe BV.
pUD231	*PGK1*$_p$-*cbbO2-ADH1*$_t$ *nptII*	Life Technologies Europe BV.
pUD232	*TEF1*$_p$-*groEL-ACT1*$_t$ *nptII*	Life Technologies Europe BV.
pUD233	*TPI1*$_p$-*groES-PGI1*$_t$ *nptII*	Life Technologies Europe BV.
pUDC075	*URA3 CEN6 ARS4 TDH3*$_p$-*cbbM-CYC1*$_t$;*PGI1*$_p$-*cbbQ2-TEF2*$_t$;*PGK1*$_p$-*cbbO2-ADH1*$_t$;*TEF1*$_p$-*groEL-ACT1*$_t$;*TPI1*$_p$-*groES-PGI1*$_t$ *bla*	This study.
pUDC099	*URA3 CEN6 ARS4 TDH3*$_p$-*cbbM-CYC1*$_t$;*PGI1*$_p$-*cbbQ2-TEF2*$_t$;*PGK1*$_p$-*cbbO2-ADH1*$_t$ *bla*	This study.
pUDC100	*URA3 CEN6 ARS4 TDH3*$_p$-*cbbM-CYC1*$_t$;*TEF1*$_p$-*groEL-ACT1*$_t$;*TPI1*$_p$-*groES-PGI1*$_t$ *bla*	This study.

Strain construction, isolation and maintenance. All *Saccharomyces cerevisiae* strains used (Table 5.4) belong to the CEN.PK family[158,159]. All strains were grown in 2% w/v glucose synthetic media[35] supplemented with 150 mg l^{-1} uracil

when required[106] until they reached end exponential phase, then sterile glycerol was added up to *ca.* 30% v/v and aliquot of 1 ml were stocked -80 °C.

The strain IMC014 that co-expressed the Rubisco form II *cbbM* and the four chaperones *cbbQ2*, *cbbO2*, *groEL*, and *groES* was constructed using a previously published *in vivo* transformation associated recombination[249]. 200 fmol of each expression cassette were pooled with 100 fmol of the KpnI/SacI linearized pRS416 backbone in a final volume of 50 µl and transformed in CEN.PK 113-5D using the lithium acetate protocol[250] (Fig. 5.4.a). Cells were selected on synthetic medium. Correct assembly of the fragment of pUDC075 was confirmed by performing multiplex PCR on transformant colonies using primers enabling amplification over the regions used for homologous recombination (Table 5.2) and by restriction analysis after retransformation of the isolated plasmid in *E. coli* DH5α. pUDC075 was sequenced by Next Gen Seq Illumina (100-bp reads paired-end, 50Mb) and assembled with Velvet[251]. The assembled sequence did not contain mutations in any of the assembled expression cassettes. The strains IMC034 and IMC035 that expressed *cbbM/cbbQ2/cbbO2* and *cbbM/groEL/groES* respectively were constructed using the same *in vivo* assembly method with the following modification. To construct plasmids pUDC099 and pUDC100, 120 bp cbbO2-pRS416 linker and cbbM-GroEL linker were used to close the assembly respectively (Table 5.2), 100 fmol of each of complementary 120 bp oligonucleotides were added to the transformation. The strain IMC033 that only expressed the *cbbM* gene was constructed by transforming CEN.PK113-5D with pUDC098.

To construct the strain IMU033 that co-expressed *PRK*, *cbbM*, *cbbQ2*, *cbbO2*, *groEL*, *groES*, the intermediate strain IMI229 was constructed by integrating *PRK*, the four chaperones and *KlLEU2*[224] at the *CAN1* locus by *in vivo* homologous integration in CEN.PK102-3A (Fig. 5.4.b). The expression cassettes were PCR amplified using Phusion Hot-Start Polymerase (Finnzymes), the corresponding oligonucleotides and DNA templates (Table 5.2). Finally, the strain IMI229 was transformed with pUDC100 that carries the Rubisco form II *cbbM* and the two *E. coli* chaperones *groEL* and *groES*.

Strain IMI232 was constructed by transforming CEN.PK102-3A with the *KlLEU2* cassette. IMI232 was finally transformed with the plasmid p426_GPD to restore prototrophy resulting in the reference strain IMU032.

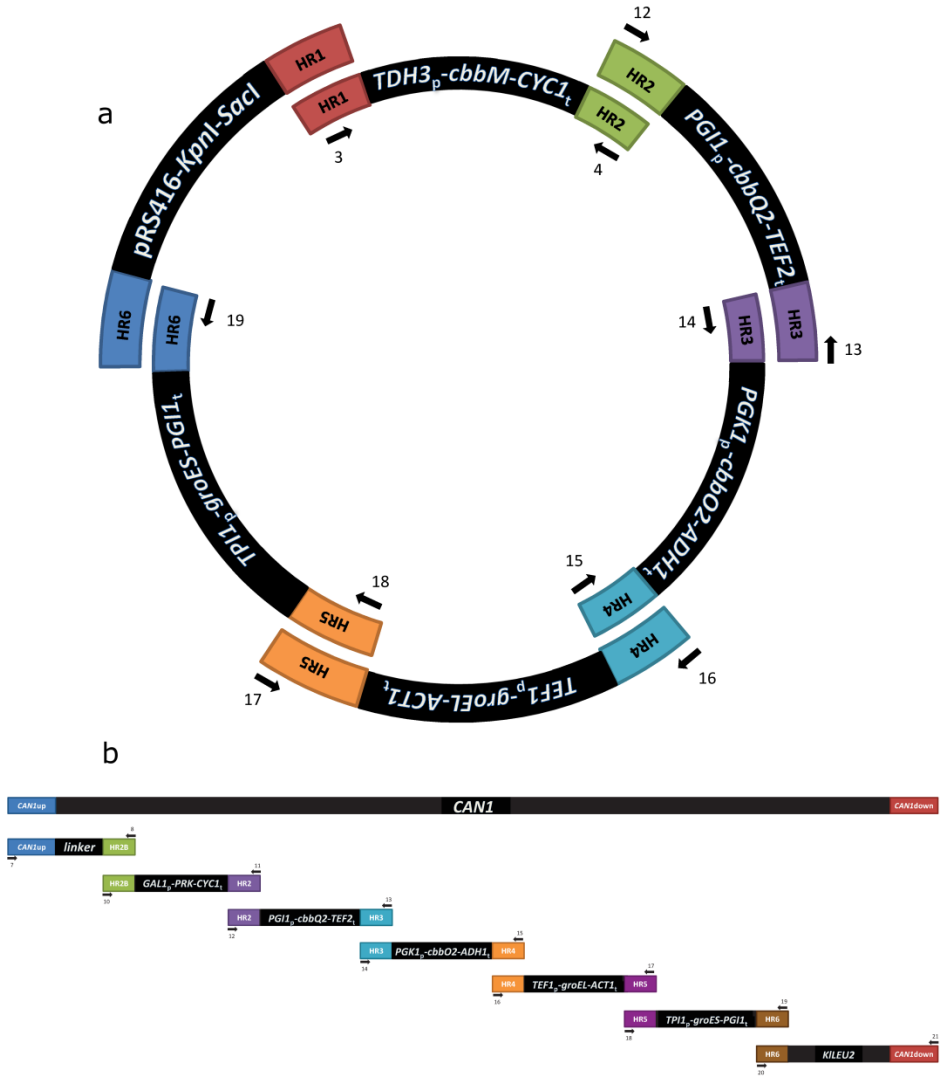

Figure 5.4 Synthetic biology constructs for the heterologous expression of Rubisco and PRK in *S. cerevisiae*. (a) *In vivo* assembly of Rubisco expression plasmid pUDC075, and **(b)** *in vivo* assembly and integration of PRK and chaperone proteins in *CAN1* locus of *Saccharomyces cerevisiae* strain IMI229. Each fragment represents a different expression cassette or plasmid backbone. All fragments used in assembly experiments were flanked by 60-bp sequences used for *in vivo* recombination, either enabling the assembly of plasmids or the integration assembled constructs into the *S. cerevisiae* genome. Arrows and numbers indicate primers used in the construction of the cassette.

Table 5.4 *Saccharomyces cerevisiae* strains used in this study.

Strain	Relevant genotype	Source/reference
CEN.PK113-5D	*MAT***a** *ura3-52*	Euroscarf.
CEN.PK102-3A	*MAT***a** *ura3-52 leu2-3, 112*	Euroscarf.
IMC014	*MAT***a** *ura3-52* pUDC075 (*CEN6 ARS4 URA3 TDH3*$_p$*-cbbM-CYC1*$_t$ *PGI1*$_p$*-cbbQ2-TEF2*$_t$ *PGK1*$_p$*-cbbO2-ADH1*$_t$ *TEF1*$_p$*-groEL-ACT1*$_t$ *TPI1*$_p$*-groES-PGI1*$_t$)	This study.
IMC033	*MAT***a** *ura3-52* pUDC098 (*CEN6 ARS4 URA3 TDH3*$_p$*-cbbM-CYC1*$_t$)	This study.
IMC034	*MAT***a** *ura3-52* pUDC099 (*CEN6 ARS4 URA3 TDH3*$_p$*-cbbM-CYC1*$_t$ *PGI1*$_p$*-cbbQ2-TEF2*$_t$ *PGK1*$_p$*-cbbO2-ADH1*$_t$ *cbbO2-pRS416 linker*)	This study.
IMC035	*MAT***a** *ura3-52* pUDC100 (*CEN6 ARS4 URA3 TEF1*$_p$*-groEL-ACT1*$_t$ *TPI1*$_p$*-groES-PGI1*$_t$ *cbbM-GroEL linker*)	This study.
IMI229	*MAT***a** *ura3-52 leu2-3, 112 can1*Δ*::GAL1*$_p$*-prk-CYC1*$_t$*, PGI1*$_p$*-cbbQ2-TEF2*$_t$*, PGK1*$_p$*-cbbO2-ADH1*$_t$*, TEF1*$_p$*-groEL-ACT1*$_t$*, TPI1*$_p$*-groES-PGI1*$_t$ *KILEU2*	This study.
IMI232	*MAT***a** *ura3-52 leu2-3, 112 can1*Δ*::KILEU2*	This study.
IMU032	IMI232 p426_GPD (2μ *URA3*)	This study.
IMU033	IMI229 pUDC100 (*CEN6 ARS4 URA3 TEF1*$_p$*-groEL-ACT1*$_t$ *TPI1*$_p$*-groES-PGI1*$_t$ *cbbM-groEL* linker)	This study.

Experimental set-up of chemostat and batch experiments. Anaerobic chemostat cultivation was performed essentially as described[252] but with 12.5 g l^{-1} glucose and 12.5 g l^{-1} galactose as the carbon source and where indicated, a mixture of 10% CO_2/90% N_2 replaced pure nitrogen as the sparging gas. Residual glucose and galactose concentrations were determined after rapid quenching[253] using commercial enzymatic assays for glucose (Boehringer, Mannheim, Germany) and D-galactose (Megazyme, Bray, Ireland). Anaerobic bioreactor batch cultures were grown essentially as described[154], but with 20 g l^{-1} galactose and a sparging gas consisting of 10% CO_2 and 90% N_2. Biomass and metabolite concentrations in batch and chemostat and batch cultures were determined as described by Guadalupe *et al.* (2010)[154]. In calculations of ethanol fluxes and yields, ethanol evaporation was corrected based on a first-order evaporation rate constant of 0.008 h^{-1} in the bioreactor set-ups and under the conditions used in this study[154,252].

Enzyme assays for phosphoribulokinase and Rubisco. Cell extracts for analysis of phosphoribulokinase (PRK) activity were prepared as described previously[254]. PRK activity was measured at 30 °C by a coupled spectrophotometric assay[255]. Reaction rates were proportional to the amounts of cell extract added. Protein concentrations were determined by the Lowry method[256] using bovine serum albumin as a standard.

Cell extracts for Rubisco activity assays were prepared as described[254], with two modifications: Tris-HCl (1 mM, pH 8.2) containing 20 mM $MgCl_2 \cdot 6H_2O$, 5 mM of DTT, 5 mM $NaHCO_3$ was used as sonication buffer and Tris-HCl (100 mM, pH 8.2), 20 mM $MgCl_2 \cdot 6H_2O$ and 5 mM of DTT as freezing buffer. Rubisco activity was determined by measuring $^{14}CO_2$-fixation (PerkinElmer, Groningen, The Netherlands) as described[257] and measuring radioactive counts in a TRI-CARB® 2700TR Series liquid scintillation counter (PerkinElmer, Groningen, The Netherlands), using Ultima Gold™ scintillation cocktail (PerkinElmer, Groningen, The Netherlands). Protein concentrations were determined by the Lowry method[256] using standard solutions of bovine serum albumin dissolved in 50 mM Tris-HCl (pH 8.2).

Acknowledgements

The PhD project of VGM was carried out within the research programme of the Kluyver Centre for Genomics of Industrial Fermentation, which is part of The Netherlands Genomics Initiative / Netherlands Organization for Scientific Research. We thank Mirjam Fuchs, Nakul Barfa, Peter de Bruijn, Astrid van der Meer and Daniel de Vries for experimental assistance and Niels Kuijpers, Tim Vos and Frank Koopman for advice on molecular biology and Gijs Kuenen for critically reading the manuscript.

Reference list

1. Camill, P. *Global Change: An Overview*. http://www.nature.com/scitable/knowledge/library/global-change-an-overview-13255365#. Access date: 1-11-2012
2. Hall, C., Tharakan, P., Hallock, J., Cleveland, C., & Jefferson, M. Hydrocarbons and the evolution of human culture. *Nature* **426**, 318-322 (2003).
3. Soetaert, W. & Vandamme, E. The impact of industrial biotechnology. *Biotechnol. J.* **1**, 756-769 (2006).
4. Kuik, O. J., Lima, M. B., & Gupta, J. Energy security in a developing world. *Wiley Interdiscip. Rev. Clim. Change* **2**, 627-634 (2011).
5. International Energy Agency *World Energy Outlook 2010* (International Energy Agency, OECD, Paris, 2010).
6. Langeveld, J. W. A. & Sanders, J. P. M. General Introduction in *The biobased economy: biofuels, materials and chemicals in the post-oil era* (Earthscan, London, 2010).
7. Dale, B. E. Energy consumption, wealth, and biofuels: helping human beings achieve their potential. *Biofuels, Bioprod. Bioref.* **6**, 1-3 (2012).
8. Jacobson, M. Z. Review of solutions to global warming, air pollution, and energy security. *Energy Environ. Sci.* **2**, 148-173 (2009).
9. Lewis, N. S. & Nocera, D. G. Powering the planet: Chemical challenges in solar energy utilization. *Proc. Natl. Acad. Sci.* **103**, 15729-15735 (2006).
10. Louwrier, A. Industrial products - the return to carbohydrate-based industries. *Biotechnol. Appl. Biochem.* **27**, 1-8 (1998).
11. Dale, B. E. "Greening" the chemical industry: research and development priorities for biobased industrial products. *J. Chem. Technol. Biotechnol.* **78**, 1093-1103 (2003).
12. Jordan, N. *et al.* Sustainable development of the agricultural bio-economy. *Science* **316**, 1570-1571 (2007).
13. Ragauskas, A. J. *et al.* The path forward for biofuels and biomaterials. *Science* **311**, 484-489 (2006).
14. Karsenti, E. Self-organization in cell biology: a brief history. *Nat. Rev. Mol. Cell Biol.* **9**, 255-262 (2008).
15. Otero, J. M., Panagiotou, G., & Olsson, L. Fueling industrial biotechnology growth with bioethanol in *Biofuels* (Springer Berlin Heidelberg, 2007).
16. McGovern, P. E. *et al.* Fermented beverages of pre- and proto-historic China. *Proc. Natl. Acad. Sci.* **101**, 17593-17598 (2004).
17. Swiegers, J. H., Bartowsky, E. J., Henschke, P. A., & Pretorius, I. S. Yeast and bacterial modulation of wine aroma and flavour. *Aust. J. Grape Wine Res.* **11**, 139-173 (2005).
18. Wegman, M. A., Janssen, M. H., van Rantwijk, F., & Sheldon, R. A. Towards biocatalytic synthesis of β-lactam antibiotics. *Adv. Synth. Catal.* **343**, 559-576 (2001).
19. Schoemaker, H. E., Mink, D., & Wubbolts, M. G. Dispelling the myths-biocatalysis in industrial synthesis. *Science* **299**, 1694-1697 (2003).
20. van Maris, A. *et al.* Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Antonie van Leeuwenhoek* **90**, 391-418 (2006).
21. Papini, M., Salazar, M., & Nielsen, J. Systems biology of industrial microorganisms in *Biosystems Engineering I* (Springer Berlin Heidelberg, 2010).
22. Gavrilescu, M. & Chisti, Y. Biotechnology — a sustainable alternative for chemical industry. *Biotech. Adv.* **23**, 471-499 (2005).
23. Mustacchi, R., Hohmann, S., & Nielsen, J. Yeast systems biology to unravel the network of life. *Yeast* **23**, 227-238 (2006).
24. Porter, J. R. Antony van Leeuwenhoek: tercentenary of his discovery of bacteria. *Bacteriol. Rev.* **40**, 260-269 (1976).

25. Polsinelli, M., Romano, P., Suzzi, G., & Mortimer, R. Multiple strains of *Saccharomyces cerevisiae* on a single grape vine. *Lett. Appl. Microbiol.* **23**, 110-114 (1996).
26. Stefanini, I. *et al.* Role of social wasps in *Saccharomyces cerevisiae* ecology and evolution. *Proc. Natl. Acad. Sci.*(2012).
27. Pasteur, L. Mémoire sur la fermentation appelée lactique (Extrait par l'auteur). *C. R. Seances Acad. Sci.* **45**, 913-916 (1857).
28. Hong, K. K. & Nielsen, J. Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cell. Mol. Life Sci.* **69**, 2671-2690 (2012).
29. Kim, I. K., Roldão, A., Siewers, V., & Nielsen, J. A systems-level approach for metabolic engineering of yeast cell factories. *FEMS Yeast Res.* **12**, 228-248 (2012).
30. Nielsen, J. & Jewett, M. C. Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **8**, 122-131 (2008).
31. Goffeau, A. *et al.* Life with 6000 genes. *Science* **274**, 546-567 (1996).
32. Weusthuis, R. A., Lamot, I., van der Oost, J., & Sanders, J. P. M. Microbial production of bulk chemicals: development of anaerobic processes. *Trends Biotechnol.* **29**, 153-158 (2011).
33. Barnett, J. A. A history of research on yeasts 5: the fermentation pathway. *Yeast* **20**, 509-543 (2003).
34. van Dijken, J. P. & Scheffers, W. A. Redox balances in the metabolism of sugars by yeasts. *FEMS Microbiol. Rev.* **32**, 199-224 (1986).
35. Verduyn, C., Postma, E., Scheffers, W. A., & van Dijken, J. P. Physiology of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. *J. Gen. Microbiol.* **136**, 395-403 (1990).
36. Bakker, B. M. *et al.* Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* **25**, 15-37 (2001).
37. Förster, J., Famili, I., Fu, P., Palsson, B. Ø., & Nielsen, J. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* **13**, 244-253 (2003).
38. Pronk, J. T., Steensma, H. Y., & van Dijken, J. P. Pyruvate metabolism in *Saccharomyces cerevisiae*. *Yeast* **12**, 1607-1633 (1996).
39. Stuart, R. A. Supercomplex organization of the oxidative phosphorylation enzymes in yeast mitochondria. *J. Bioenerg. Biomembr.* **40**, 411-417 (2008).
40. Kane, P. M., Yamashiro, C. T., & Stevens, T. H. Biochemical characterization of the yeast vacuolar H^+-ATPase. *J. Biol. Chem.* **264**, 19236-19244 (1989).
41. van Urk, H., Postma, E., Scheffers, W. A., & van Dijken, J. P. Glucose transport in crabtree-positive and crabtree-negative yeasts. *J. Gen. Microbiol.* **135**, 2399-2406 (1989).
42. Postma, E., Verduyn, C., Scheffers, W. A., & van Dijken, J. P. Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **55**, 468-477 (1989).
43. Hensing, M. C. M., Rouwenhorst, R. J., Heijnen, J. J., Dijken, J. P., & Pronk, J. T. Physiological and technological aspects of large-scale heterologous-protein production with yeasts. *Antonie van Leeuwenhoek* **67**, 261-279 (1995).
44. Albers, E., Larsson, C., Lidén, G., Niklasson, C., & Gustafsson, L. Influence of the nitrogen source on *Saccharomyces cerevisiae* anaerobic growth and product formation. *Appl. Environ. Microbiol.* **62**, 3187-3195 (1996).
45. Athenstaedt, K. & Daum, G. Phosphatidic acid, a key intermediate in lipid metabolism. *Eur. J. Biochem.* **266**, 1-16 (1999).
46. Albertyn, J., Hohmann, S., Thevelein, J. M., & Prior, B. A. *GPD1*, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in *Saccharomyces cerevisiae*, and its expression is regulated by the high-osmolarity glycerol response pathway. *Mol. Cell. Biol.* **14**, 4135-4144 (1994).
47. Larsson, K., Ansell, R., Eriksson, P., & Adler, L. A gene encoding sn-glycerol 3-phosphate dehydrogenase (NAD^+) complements an osmosensitive mutant of *Saccharomyces cerevisiae*. *Mol. Microbiol.* **10**, 1101-1111 (1993).
48. Eriksson, P., André, L., Ansell, R., Blomberg, A., & Adler, L. Cloning and characterization of *GPD2*, a second gene encoding sn-glycerol 3-phosphate dehydrogenase (NAD^+) in *Saccharomyces cerevisiae*, and its comparison with *GPD1*. *Mol. Microbiol.* **17**, 95-107 (1995).

49. Ansell, R., Granath, K., Hohmann, S., Thevelein, J. M., & Adler, L. The two isoenzymes for yeast NAD+-dependent glycerol 3-phosphate dehydrogenase encoded by *GPD1* and *GPD2* have distinct roles in osmoadaptation and redox regulation. *EMBO J.* **16**, 2179-2187 (1997).
50. Norbeck, J., Påhlman, A. K., Akhtar, N., Blomberg, A., & Adler, L. Purification and characterization of two isoenzymes of DL-glycerol-3-phosphatase from *Saccharomyces cerevisiae* - identification of the corresponding *GPP1* and *GPP2* genes and evidence for osmotic regulation of Gpp2p expression by the osmosensing mitogen-activated protein kinase signal transduction pathway. *J. Biol. Chem.* **271**, 13875-13881 (1996).
51. Solomon, B. D., Barnes, J. R., & Halvorsen, K. E. Grain and cellulosic ethanol: History, economics, and energy policy. *Biomass Bioenergy* **31**, 416-425 (2007).
52. Kovarik, B. *Henry Ford, Charles Kettering and the Fuel of the Future*. http://www.environmentalhistory.org/billkovarik/research/henry-ford-charles-kettering-and-the-fuel-of-the-future/. Access date: 1-11-2012
53. Renewable Fuels Association. *World Fuel Ethanol Production*. http://ethanolrfa.org/pages/World-Fuel-Ethanol-Production. Access date: 25-11-2012
54. Tenenbaum, D. J. Food vs. fuel diversion of crops could cause more hunger. *Environ. Health Perspect.* **116**, (2008).
55. Hahn-Hägerdal, B., Galbe, M., Gorwa-Grauslund, M. F., Lidén, G., & Zacchi, G. Bio-ethanol — the fuel of tomorrow from the residues of today. *Trends Biotechnol.* **24**, 549-556 (2006).
56. Singh, A., Nigam, P. S., & Murphy, J. D. Renewable fuels from algae: An answer to debatable land based fuels. *Bioresour. Technol.* **102**, 10-16 (2011).
57. Lynd, L. R., Zyl, W. H., McBride, J. E., & Laser, M. Consolidated bioprocessing of cellulosic biomass: an update. *Curr. Opin. Biotechnol.* **16**, 577-583 (2005).
58. Mussatto, S. I. *et al.* Technological trends, global market, and challenges of bio-ethanol production. *Biotechnol. Adv.* **28**, 817-830 (2010).
59. Argueso, J. L. *et al.* Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. *Genome Res.* **19**, 2258-2270 (2009).
60. Basso, L. C., De Amorim, H. V., De Oliveira, A. J., & Lopes, M. L. Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res.* **8**, 1155-1163 (2008).
61. Thomas, K. C., Hynes, S. H., & Ingledew, W. M. Practical and theoretical considerations in the production of high concentrations of alcohol by fermentation. *Process Biochem.* **31**, 321-331 (1996).
62. Zhao, X. Q. & Bai, F. W. Mechanisms of yeast stress tolerance and its manipulation for efficient fuel ethanol production. *J. Biotechnol.* **144**, 23-30 (2009).
63. Palmqvist, E., Grage, H., Meinander, N. Q., & Hahn-Hägerdal, B. Main and interaction effects of acetic acid, furfural, and p-hydroxybenzoic acid on growth and ethanol productivity of yeasts. *Biotechnol. Bioeng.* **63**, 46-55 (1999).
64. Taherzadeh, M. J., Niklasson, C., & Lidén, G. Acetic acid—friend or foe in anaerobic batch conversion of glucose to ethanol by *Saccharomyces cerevisiae*? *Chem. Eng. Sci.* **52**, 2653-2659 (1997).
65. Pampulha, M. E. & Loureiro-Dias, M. C. Combined effect of acetic acid, pH and ethanol on intracellular pH of fermenting yeast. *Appl. Microbiol. Biotechnol.* **31**, 547-550 (1989).
66. Boender, L. G. M., de Hulster, E. A. F., van Maris, A. J. A., Daran-Lapujade, P. A. S., & Pronk, J. T. Quantitative physiology of *Saccharomyces cerevisiae* at near-zero specific growth rates. *Appl. Environ. Microbiol.* **75**, 5607-5614 (2009).
67. Genon, G., Gianetto, A., & Specchia, V. Production of ethanol with *Saccharomyces cerevisiae* in a continuous reactor Note III: an experimental, kinetic study with suspended, recycled biomass. *Chem. Eng. Commun.* **23**, 245-258 (1983).
68. van Loosdrecht, M. C. M. & Henze, M. Maintenance, endogeneous respiration, lysis, decay and predation. *Water Sci. Technol.* **39**, 107-117 (1999).
69. Nissen, T. L., Kielland-Brandt, M. C., Nielsen, J., & Villadsen, J. Optimization of ethanol production in *Saccharomyces cerevisia*e by metabolic engineering of the ammonium assimilation. *Metab. Eng.* **2**, 69-77 (2000).
70. Grosz, R. & Stephanopoulos, G. Physiological, biochemical, and mathematical studies of micro-aerobic continuous ethanol fermentation by *Saccharomyces cerevisiae*. I: Hysteresis,

oscillations, and maximum specific ethanol productivities in chemostat culture. *Biotechnol. Bioeng.* **36**, 1006-1019 (1990).
71. Kuriyama, H. & Kobayashi, H. Effects of oxygen supply on yeast growth and metabolism in continuous fermentation. *J. Ferment. Bioeng.* **75**, 364-367 (1993).
72. Ryu, D. D. Y., Kim, Y. J., & Kim, J. H. Effect of air supplement on the performance of continuous ethanol fermentation system. *Biotechnol. Bioeng.* **26**, 12-16 (1984).
73. Weusthuis, R. A., Visser, W., Pronk, J. T., Scheffers, W. A., & van Dijken, J. P. Effects of oxygen limitation on sugar metabolism in yeasts: a continuous-culture study of the Kluyver effect. *Microbiology* **140**, 703-715 (1994).
74. Alfenore, S. *et al.* Aeration strategy: a need for very high ethanol performance in *Saccharomyces cerevisiae* fed-batch process. *Appl. Microbiol. Biotechnol.* **63**, 537-542 (2004).
75. Thomas, K. C. & Ingledew, W. M. Fuel alcohol production: effects of free amino nitrogen on fermentation of very-high-gravity wheat mashes. *Appl. Environ. Microbiol.* **56**, 2046-2050 (1990).
76. Blomberg, A. & Adler, L. Roles of glycerol and glycerol-3-phosphate dehydrogenase (NAD^+) in acquired osmotolerance of *Saccharomyces cerevisiae*. *J. Bacteriol.* **171**, 1087-1092 (1989).
77. Maeda, T., Wurgler-Murphy, S. M., & Saito, H. A two-component system that regulates an osmosensing MAP kinase cascade in yeast. *Nature* **369**, 242-245 (1994).
78. Nevoigt, E. & Stahl, U. Osmoregulation and glycerol metabolism in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* **21**, 231-241 (1997).
79. Hohmann, S. Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol. Mol. Biol. R.* **66**, 300-372 (2002).
80. Palmqvist, E. Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresour. Technol.* **74**, 25-33 (2000).
81. Bellissimi, E., van Dijken, J. P., Pronk, J. T., & van Maris, A. J. A. Effects of acetic acid on the kinetics of xylose fermentation by an engineered, xylose-isomerase-based *Saccharomyces cerevisiae* strain. *FEMS Yeast Res.* **9**, 358-364 (2009).
82. Mussatto, S. I. & Roberto, I. C. Alternatives for detoxification of diluted-acid lignocellulosic hydrolyzates for use in fermentative processes: a review. *Bioresour. Technol.* **93**, 1-10 (2004).
83. Almeida, J. R. M. *et al.* Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. *J. Chem. Technol. Biotechnol.* **82**, 340-349 (2007).
84. Rudolf, A., Galbe, M., & Lidén, G. Controlled fed-batch fermentations of dilute-acid hydrolysate in pilot development unit scale. *Appl. Biochem. Biotechnol.* **114**, 601-617 (2004).
85. Taherzadeh, M. J., Niklasson, C., & Lidén, G. Conversion of dilute-acid hydrolyzates of spruce and birch to ethanol by fed-batch fermentation. *Bioresour. Technol.* **69**, 59-66 (1999).
86. Bailey, J. E. Toward a science of metabolic engineering. *Science* **252**, 1668-1675 (1991).
87. Nielsen, J. Metabolic engineering. *Appl. Microbiol. Biotechnol.* **55**, 263-283 (2001).
88. Ogata, H. *et al.* KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nuc. Acids Res.* **27**, 29-34 (1999).
89. Kuyper, M. *et al.* Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Res.* **5**, 399-409 (2005).
90. Wisselink, H. W. *et al.* Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose. *Appl. Environ. Microbiol.* **73**, 4881-4891 (2007).
91. Dequin, S. & Barre, P. Mixed lactic acid-alcoholic fermentation by *Saccharomyes cerevisiae* expressing the *Lactobacillus casei* L(+)-LDH. *Nat. Biotech.* **12**, 173-177 (1994).
92. Gorsich, S. W. *et al.* Tolerance to furfural-induced stress is associated with pentose phosphate pathway genes *ZWF1*, *GND1*, *RPE1*, and *TKL1* in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **71**, 339-349 (2006).
93. Gombert, A. K. *et al.* Functional characterization of the oxaloacetase encoding gene and elimination of oxalate formation in the β-lactam producer *Penicillium chrysogenum*. *Fungal Genetics and Biology* **48**, 831-839 (2011).
94. Wisselink, H. W., Toirkens, M. J., Wu, Q., Pronk, J. T., & van Maris, A. J. A. Novel evolutionary engineering approach for accelerated utilization of glucose, xylose, and arabinose mixtures by engineered *Saccharomyces cerevisiae* strains. *Appl. Environ. Microbiol.* **75**, 907-914 (2009).

95. Edwards, J. S. & Palsson, B. O. Systems Properties of the *Haemophilus influenzae Rd* Metabolic Genotype. *J. Biol. Chem.* **274**, 17410-17416 (1999).
96. Bro, C., Regenberg, B., Förster, J., & Nielsen, J. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab. Eng.* **8**, 102-111 (2006).
97. Henry, C. S., Broadbelt, L. J., & Hatzimanikatis, V. Thermodynamics-Based Metabolic Flux Analysis. *Biophys. J.* **92**, 1792-1805 (2007).
98. Henry, C. S., Broadbelt, L. J., & Hatzimanikatis, V. Discovery and analysis of novel metabolic pathways for the biosynthesis of industrial chemicals: 3-hydroxypropanoate. *Biotechnol. Bioeng.* **106**, 462-473 (2010).
99. Griffith, F. The significance of *pneumococcal* types. *J. Hyg.* **27**, 113-159 (1928).
100. Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* **171**, 737-738 (1953).
101. Avery, O. T., MacLeod, C. M., & McCarty, M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types induction of transformation by a desoxyribonucleic acid fraction isolated from *pneumococcus* type III. *J. Exp. Med.* **79**, 137-158 (1944).
102. Saiki, R. K. *et al.* Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350-1354 (1985).
103. Cohen, S. N., Chang, A. C. Y., Boyer, H. W., & Helling, R. B. Construction of biologically functional bacterial plasmids *in vitro*. *Proc. Natl. Acad. Sci.* **70**, 3240-3244 (1973).
104. Smith, H. O. & Welcox, K. W. A Restriction enzyme from *Hemophilus influenzae*: I. Purification and general properties. *J. Mol. Biol.* **51**, 379-391 (1970).
105. Güldener, U., Heck, S., Fiedler, T., Beinhauer, J., & Hegemann, J. H. A new efficient gene disruption cassette for repeated use in budding yeast. *Nuc. Acids Res.* **24**, 2519-2524 (1996).
106. Pronk, J. T. Auxotrophic yeast strains in fundamental and applied research. *Appl. Environ. Microbiol.* **68**, 2095-2100 (2002).
107. Dani, G. M. & Zakian, V. A. Mitotic and meiotic stability of linear plasmids in yeast. *Proc. Natl. Acad. Sci.* **80**, 3406-3410 (1983).
108. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., & Stahl, F. W. The double-strand-break repair model for recombination. *Cell* **33**, 25-35 (1983).
109. Gibson, D. G. *et al.* Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215-1220 (2008).
110. Keasling, J. D. Synthetic biology for synthetic chemistry. *ACS Chem. Biol.* **3**, 64-76 (2008).
111. Sauer, U. Evolutionary engineering of industrially important microbial phenotypes. *Adv. Biochem. Eng. Biotechnol.* **73**, 129-169 (2001).
112. Oud, B., van Maris, A. J. A., Daran, J. M., & Pronk, J. T. Genome-wide analytical approaches for reverse metabolic engineering of industrially relevant phenotypes in yeast. *FEMS Yeast Res.* **12**, 183-196 (2012).
113. Lee, S. K., Chou, H., Ham, T. S., Lee, T. S., & Keasling, J. D. Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr. Opin. Biotechnol.* **19**, 556-563 (2008).
114. Nevoigt, E. Progress in Metabolic Engineering of *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. R.* **72**, 379-412 (2008).
115. Hahn-Hägerdal, B., Karhumaa, K., Fonseca, C., Spencer-Martins, I., & Gorwa-Grauslund, M. Towards industrial pentose-fermenting yeast strains. *Appl. Microbiol. Biotechnol.* **74**, 937-953 (2007).
116. Zaldivar, J., Nielsen, J., & Olsson, L. Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. *Appl Microbiol Biotechnol* **56**, 17-34 (2001).
117. Björkqvist, S., Ansell, R., Adler, L., & Lidén, G. Physiological response to anaerobicity of glycerol-3-phosphate dehydrogenase mutants of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **63**, 128-132 (1997).
118. Shen, B., Hohmann, S., Jensen, R. G., & Bohnert, H. Roles of sugar alcohols in osmotic stress adaptation. Replacement of glycerol by mannitol and sorbitol in yeast. *Plant Physiol.* **121**, 45-52 (1999).

119. Jain, V., Divol, B., Prior, B., & Bauer, F. Elimination of glycerol and replacement with alternative products in ethanol fermentation by *Saccharomyces cerevisiae*. *J. Ind. Microbiol. Biotechnol.* **38**, 1427-1435 (2011).
120. Nevoigt, E. *et al.* Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **72**, 5266-5273 (2006).
121. Hubmann, G., Guillouet, S., & Nevoigt, E. Gpd1 and Gpd2 fine-tuning for sustainable reduction of glycerol formation in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **77**, 5857-5867 (2011).
122. Pagliardini, J. *et al.* Quantitative evaluation of yeast's requirement for glycerol formation in very high ethanol performance fed-batch process. *Microb. Cell Fact.* **9**, 36 (2010).
123. Heux, S., Cachon, R., & Dequin, S. Cofactor engineering in *Saccharomyces cerevisiae*: Expression of a H2O-forming NADH oxidase and impact on redox metabolism. *Metab. Eng.* **8**, 303-314 (2006).
124. Hou, J., Lages, N. F., Oldiges, M., & Vemuri, G. N. Metabolic impact of redox cofactor perturbations in *Saccharomyces cerevisiae*. *Metab. Eng.* **11**, 253-261 (2009).
125. Lopez de Felipe, F., Kleerebezem, M., de Vos, W. M., & Hugenholtz, J. Cofactor Engineering: a Novel Approach to Metabolic Engineering in *Lactococcus lactis* by Controlled Expression of NADH Oxidase. *J. Bacteriol.* **180**, 3804-3808 (1998).
126. Guo, Z. p., Zhang, L., Ding, Z. y., & Shi, G. y. Minimization of glycerol synthesis in industrial ethanol yeast without influencing its fermentation performance. *Metab. Eng.* **13**, 49-59 (2011).
127. Anderlund, M. *et al.* 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**, 2333-2340 (1999).

128. Nissen, T. L., Hamann, C. W., Kielland-Brandt, M. C., Nielsen, J., & Villadsen, J. Anaerobic and aerobic batch cultivations of *Saccharomyces cerevisiae* mutants impaired in glycerol synthesis. *Yeast* **16**, 463-474 (2000).
129. Nissen, T. L., Anderlund, M., Nielsen, J., Villadsen, J., & Kielland-Brandt, M. C. Expression of a cytoplasmic transhydrogenase in *Saccharomyces cerevisiae* results in formation of 2-oxoglutarate due to depletion of the NADPH pool. *Yeast* **18**, 19-32 (2001).
130. Hou, J., Vemuri, G. N., Bao, X., & Olsson, L. Impact of overexpressing NADH kinase on glucose and xylose metabolism in recombinant xylose-utilizing *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **82**, 909-919 (2009).
131. Vemuri, G. N., Eiteman, M. A., McEwen, J. E., Olsson, L., & Nielsen, J. Increasing NADH oxidation reduces overflow metabolism in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **104**, 2402-2407 (2007).
132. Geertman, J. M., van Dijken, J. P., & Pronk, J. T. Engineering NADH metabolism in *Saccharomyces cerevisiae*: formate as an electron donor for glycerol production by anaerobic, glucose-limited chemostat cultures. *FEMS Yeast Res.* **6**, 1193-1203 (2006).
133. Hou, J., Scalcinati, G., Oldiges, M., & Vemuri, G. N. Metabolic impact of increased NADH availability in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **76**, 851-859 (2010).
134. Nissen, T. L., Hamann, C. W., Kielland-Brandt, M. C., Nielsen, J., & Villadsen, J. Anaerobic and aerobic batch cultivations of *Saccharomyces cerevisiae* mutants impaired in glycerol synthesis. *Yeast* **16**, 463-474 (2000).
135. Blomberg, A. & Adler, L. Physiology of osmotolerance in fungi. *Adv. Microb. Physiol.* **33**, 145-212 (1992).
136. Scheffers, W. A. Stimulation of fermentation in yeasts by acetoin and oxygen. *Nature* **210**, 533-534 (1966).
137. van den Berg, M. A. *et al.* The two acetyl-coenzyme A synthetases of *Saccharomyces cerevisiae* differ with respect to kinetic properties and transcriptional regulation. *J. Biol. Chem.* **271**, 28953-28959 (1996).
138. Ciriacy, M. Genetics of alcohol-dehydrogenase in *Saccharomyces cerevisiae* .1. Isolation and genetic-analysis of *adh* mutants. *Mutat. Res.* **29**, 315-326 (1975).

139. Ferrández, A., García, J. L., & Díaz, E. Genetic characterization and expression in heterologous hosts of the 3- (3-hydroxyphenyl)propionate catabolic pathway of *Escherichia coli* K-12. *J. Bacteriol.* **179**, 2573-2581 (1997).
140. van Dijken, J. P. *et al.* An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb. Technol.* **26**, 706-714 (2000).
141. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541-1553 (1999).
142. Gietz, R. D. & Akio, S. New yeast-*Escherichia coli* shuttle vectors constructed with *in vitro* mutagenized yeast genes lacking six-base pair restriction sites. *Gene* **74**, 527-534 (1988).
143. Burke,D., Dawson,D., & Stearns,T. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory Press Plainview, NY,2000).
144. Abbott, D. A., van den Brink, J., Minneboo, I. M. K., Pronk, J. T., & van Maris, A. J. A. Anaerobic homolactate fermentation with *Saccharomyces cerevisiae* results in depletion of ATP and impaired metabolic activity. *FEMS Yeast Res.* **9**, 349-357 (2009).
145. Abbott, D. A. *et al.* Catalase overexpression reduces lactic acid-induced oxidative stress in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **75**, 2320-2325 (2009).
146. Albertyn, J., Vantonder, A., & Prior, B. A. Purification and characterization of glycerol-3-phosphate dehydrogenase of *Saccharomyces cerevisiae*. *FEBS Lett.* **308**, 130-132 (1992).
147. Taherzadeh, M. J., Lidén, G., Gustafsson, L., & Niklasson, C. The effects of pantothenate deficiency and acetate addition on anaerobic batch fermentation of glucose by *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **46**, 176-182 (1996).
148. Horváth, I. S., Taherzadeh, M. J., Niklasson, C., & Lidén, G. Effects of furfural on anaerobic continuous cultivation of *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **75**, 540-549 (2001).
149. Pronk, J. T. *et al.* Energetic aspects of glucose-metabolism in a pyruvate-dehydrogenase-negative mutant of *Saccharomyces cerevisiae*. *Microbiology* **140**, 601-610 (1994).
150. Sonderegger, M., Schümperli, M., & Sauer, U. Metabolic engineering of a phosphoketolase pathway for pentose catabolism in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **70**, 2892-2897 (2004).
151. de Jong-Gubbels, P., Luttik, M. A. H., Steensma, H. Y., van Dijken, J. P., & Pronk, J. T. Overproduction of acetyl-coenzyme A synthetase isoenzymes in respiring *Saccharomyces cerevisiae* cells does not reduce acetate production after exposure to glucose excess. *FEMS Microbiol. Lett.* **165**, 15-20 (1998).
152. Sá-Correia, I., dos Santos, S. C., Teixeira, M. C., Cabrito, T. R., & Mira, N. P. Drug: H^+ antiporters in chemical stress response in yeast. *Trends Microbiol.* **17**, 22-31 (2009).
153. Athenstaedt, K., Weys, S., Paltauf, F., & Daum, G. Redundant systems of phosphatidic acid biosynthesis via acylation of glycerol-3-phosphate or dihydroxyacetone phosphate in the yeast *Saccharomyces cerevisiae*. *J. Bacteriol.* **181**, 1458-1463 (1999).
154. Guadalupe Medina, V., Almering, M. J. H., van Maris, A. J. A., & Pronk, J. T. Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. *Appl. Environ. Microbiol.* **76**, 190-195 (2010).
155. Bai, F. W., Anderson, W. A., & Moo-Young, M. Ethanol fermentation technologies from sugar and starch feedstocks. *Biotechnol. Adv.* **26**, 89-105 (2008).
156. Jones, A. M., Thomas, K. C., & Ingledew, W. M. Ethanolic fermentation of blackstrap molasses and sugar cane juice using very high gravity technology. *J. Agric. Food Chem.* **42**, 1242-1246 (1994).
157. Laluce, C. Current aspects of fuel ethanol production in Brazil. *Crit. Rev. Biotechnol.* **11**, 149-161 (1991).
158. Entian, K. D. & Kötter, P. Yeast genetic strain and plasmid collections in *Methods in Microbiology: Yeast Gene Analysis Second Edition* (Academic Press, Amsterdam, 2007).
159. Nijkamp, J. *et al. De novo* sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN.PK113-7D, a model for modern industrial biotechnology. *Microb. Cell Fact.* **11**, 36 (2012).
160. Herskowitz, I. & Jensen, R. E. Putting the *HO* gene to work: Practical uses for mating-type switching in *Methods in Enzymology: Guide to Yeast Genetics and Molecular Biology* (Academic Press, Amsterdam, 1991).

161. Sugawara, N. & Haber, J. Monitoring DNA recombination initiated by HO endonuclease in *DNA Repair Protocols* (Humana Press, New York, 2012).
162. Bahalul, M., Kaneti, G., & Kashi, Y. Ether-zymolyase ascospore isolation procedure: an efficient protocol for ascospores isolation in *Saccharomyces cerevisiae* yeast. *Yeast* **27**, 999-1003 (2010).
163. Swinnen, S., Thevelein, J. M., & Nevoigt, E. Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **12**, 215-227 (2012).
164. Verduyn, C., Postma, E., Scheffers, W. A., & van Dijken, J. P. Energetics of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. *J. Gen. Microbiol.* **136**, 405-412 (1990).
165. Rudolf, A., Alkasrawi, M., Zacchi, G., & Lidén, G. A comparison between batch and fed-batch simultaneous saccharification and fermentation of steam pretreated spruce. *Enzyme Microb. Technol.* **37**, 195-204 (2005).
166. Taherzadeh, M. J., Millati, R., & Niklasson, C. Continuous cultivation of dilute-acid hydrolysates to ethanol by immobilized *Saccharomyces cerevisiae*. *Appl. Biochem. Biotechnol.* **95**, 45-57 (2001).
167. Racenis, P. V. *et al.* The acyl dihydroxyacetone phosphate pathway enzymes for glycerolipid biosynthesis are present in the yeast *Saccharomyces cerevisiae*. *J. Bacteriol.* **174**, 5702-5710 (1992).
168. Athenstaedt, K. & Daum, G. 1-Acyldihydroxyacetone-phosphate reductase (Ayr1p) of the yeast *Saccharomyces cerevisiae* encoded by the open reading frame YIL124w is a major component of lipid particles. *J. Biol. Chem.* **275**, 235-240 (2000).
169. Hounsa, C. G., Brandt, E. V., Thevelein, J., Hohmann, S., & Prior, B. A. Role of trehalose in survival of *Saccharomyces cerevisiae* under osmotic stress. *Microbiology* **144**, 671-680 (1998).
170. Mahmud, S. A. *et al.* Effect of trehalose accumulation on response to saline stress in *Saccharomyces cerevisiae*. *Yeast* **26**, 17-30 (2009).
171. Takagi, H. Proline as a stress protectant in yeast: physiological functions, metabolic regulations, and biotechnological applications. *Appl. Microbiol. Biotechnol.* **81**, 211-223 (2008).
172. Koppram, R., Albers, E., & Olsson, L. Evolutionary engineering strategies to enhance tolerance of xylose utilizing recombinant yeast to inhibitors derived from spruce biomass. *Biotechnol. Biofuels* **5**, 32 (2012).
173. Kuyper, M. *et al.* Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting *Saccharomyces cerevisiae* strain. *FEMS Yeast Res.* **5**, 925-934 (2005).
174. Zelle, R. M., Harrison, J. C., Pronk, J. T., & van Maris, A. J. A. Anaplerotic role for cytosolic malic enzyme in engineered *Saccharomyces cerevisiae* strains. *Appl. Environ. Microbiol.* **77**, 732-738 (2011).
175. de Kok, S. *et al.* Laboratory evolution of new lactate transporter genes in a *jen1Δ* mutant of *Saccharomyces cerevisiae* and their identification as *ADY2* alleles by whole-genome resequencing and transcriptome analysis. *FEMS Yeast Res.* **12**, 359-374 (2012).
176. Adamo, G. M., Brocca, S., Passolunghi, S., Salvato, B., & Lotti, M. Laboratory evolution of copper tolerant yeast strains. *Microb. Cell Fact.* **11**, 1 (2012).
177. Wright, J. *et al.* Batch and continuous culture-based selection strategies for acetic acid tolerance in xylose-fermenting *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **11**, 299-306 (2011).
178. Portnoy, V. A., Bezdan, D., & Zengler, K. Adaptive laboratory evolution - harnessing the power of biology for metabolic engineering. *Curr. Opin. Biotechnol.* **22**, 590-594 (2011).
179. Wagner, A. Neutralism and selectionism: a network-based reconciliation. *Nat. Rev. Genet.* **9**, 965-974 (2008).
180. Elena, S. F. & Lenski, R. E. Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* **4**, 457-469 (2003).
181. Loh, E., Salk, J. J., & Loeb, L. A. Optimization of DNA polymerase mutation rates during bacterial evolution. *Proc. Natl. Acad. Sci.* **107**, 1154-1159 (2010).
182. Stanek, M. T., Cooper, T. F., & Lenski, R. E. Identification and dynamics of a beneficial mutation in a long-term evolution experiment with *Escherichia coli*. *BMC Evol. Biol.* **9**, 302 (2009).
183. Kvitek, D. J. & Sherlock, G. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genet.* **7**, e1002056 (2011).

184. Kao, K. C. & Sherlock, G. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nat. Genet.* **40**, 1499-1504 (2008).
185. Conrad, T. M., Lewis, N. E., & Palsson, B. O. Microbial laboratory evolution in the era of genome-scale science. *Mol. Syst. Biol.* **7**, 509 (2011).
186. Bailey, J. E. *et al.* Inverse metabolic engineering: A strategy for directed genetic engineering of useful phenotypes. *Biotechnol. Bioeng.* **52**, 109-121 (1996).
187. Bro, C. & Nielsen, J. Impact of 'ome' analyses on inverse metabolic engineering. *Metab. Eng.* **6**, 204-211 (2004).
188. Teusink, B., Bachmann, H., & Molenaar, D. Systems biology of lactic acid bacteria: a critical review. *Microb. Cell Fact.* **10**, S11 (2011).
189. Ratcliff, W. C., Denison, R. F., Borrello, M., & Travisano, M. Experimental evolution of multicellularity. *Proc. Natl. Acad. Sci.* **109**, 1595-1600 (2012).
190. Bonner, J. T. The origins of multicellularity. *Integr. Biol. Issues News Rev.* **1**, 27-36 (1998).
191. Knoll, A. H. The multiple origins of complex multicellularity. *Annu. Rev. Earth Planet Sci.* **39**, 217-239 (2011).
192. Rokas, A. The origins of multicellularity and the early history of the genetic toolkit for animal development. *Annu. Rev. Genet.* **42**, 235-251 (2008).
193. Linden, T., Peetre, J., & Hahnhagerdal, B. Isolation and characterization of acetic acid-tolerant galactose-fermenting strains of *Saccharomyces cerevisiae* from a spent sulfite liquor fermentation plant. *Appl. Environ. Microbiol.* **58**, 1661-1669 (1992).
194. Tao, W. T. Multicellularity of a unicellular organism in response to DNA replication stress. *Res. Microbiol.* **160**, 87-88 (2009).
195. Koschwanez, J. H., Foster, K. R., & Murray, A. W. Sucrose utilization in budding yeast as a model for the origin of undifferentiated multicellularity. *PLoS Biol.* **9**, e1001122 (2011).
196. Boraas, M. E., Seale, D. B., & Boxhorn, J. E. Phagotrophy by a flagellate selects for colonial prey: A possible origin of multicellularity. *Evol. Ecol.* **12**, 153-164 (1998).
197. Hong, K. K., Vongsangnak, W., Vemuri, G. N., & Nielsen, J. Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *Proc. Natl. Acad. Sci.* **108**, 12179-12184 (2011).
198. Soares, E. V. Flocculation in *Saccharomyces cerevisiae*: a review. *J. Appl. Microbiol.* **110**, 1-18 (2011).
199. Stratford, M. Evidence for two mechanisms of flocculation in *Saccharomyces cerevisiae*. *Yeast* **5**, S441-S445 (1989).
200. Nijkamp, J. F. *et al. De novo* detection of copy number variation by co-assembly. *Bioinformatics* **28**, 3195-3202 (2012).
201. Voth, W. P., Olsen, A. E., Sbia, M., Freedman, K. H., & Stillman, D. J. *ACE2*, *CBK1*, and *BUD4* in budding and cell separation. *Eukaryot. Cell* **4**, 1018-1028 (2005).
202. Saputo, S., Chabrier-Rosello, Y., Luca, F. C., Kumar, A., & Krysan, D. J. The RAM network in pathogenic fungi. *Eukaryot. Cell* **11**, 708-717 (2012).
203. Doolin, M. T., Johnson, A. L., Johnston, L. H., & Butler, G. Overlapping and distinct roles of the duplicated yeast transcription factors Ace2p and Swi5p. *Mol. Microbiol.* **40**, 422-432 (2001).
204. Sbia, M. *et al.* Regulation of the yeast Ace2 transcription factor during the cell cycle. *J. Biol. Chem.* **283**, 11135-11145 (2008).
205. Kuranda, M. J. & Robbins, P. W. Chitinase is required for cell separation during growth of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **266**, 19758-19767 (1991).
206. Herth, W. & Schnepf, E. The fluorochrome, calcofluor white, binds oriented to structural polysaccharide fibrils. *Protoplasma* **105**, 129-133 (1980).
207. Chen, G. B., Rubinstein, B., & Li, R. Whole chromosome aneuploidy: Big mutations drive adaptation by phenotypic leap. *Bioessays* **34**, 893-900 (2012).
208. Koszul, R., Caburet, S., Dujon, B., & Fischer, G. Eucaryotic genome evolution through the spontaneous duplication of large chromosomal segments. *EMBO J.* **23**, 234-243 (2004).
209. Yona, A. H. *et al.* Chromosomal duplication is a transient evolutionary solution to stress. *Proc. Natl. Acad. Sci.* **109**, 21010-21015 (2012).
210. Wolfe, K. H. & Shields, D. C. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708-713 (1997).

211. Kamran, M. *et al.* Inactivation of transcription factor gene *ACE2* in the fungal pathogen *Candida glabrata* results in hypervirulence. *Eukaryot. Cell* **3**, 546-552 (2004).
212. MacCallum, D. M. *et al.* Different consequences of *ACE2* and *SW15* gene disruptions for virulence of pathogenic and nonpathogenic yeasts. *Infect. Immun.* **74**, 5244-5248 (2006).
213. Ma, X. *et al.* Mutation hot spots in yeast caused by long-range clustering of homopolymeric sequences. *Cell Rep.* **1**, 36-42 (2012).
214. Dechering, K. J., Konings, R. N. H., Cuelenaere, K., & Leunissen, J. A. M. Distinct frequency-distributions of homopolymeric DNA tracts in different genomes. *Nuc. Acids Res.* **26**, 4056-4062 (1998).
215. Verduyn, C., Postma, E., Scheffers, W. A., & Vandijken, J. P. Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* **8**, 501-517 (1992).
216. Daran-Lapujade, P. *et al.* An atypical *PMR2* locus is responsible for hypersensitivity to sodium and lithium cations in the laboratory strain *Saccharomyces cerevisiae* CEN.PK113-7D. *FEMS Yeast Res.* **9**, 789-792 (2009).
217. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
218. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
219. Ewing, B. & Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**, 186-194 (1998).
220. Porro, D., Brambilla, L., & Alberghina, L. Glucose metabolism and cell size in continuous cultures of *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* **229**, 165-171 (2003).
221. Daran-Lapujade, P., Daran, J. M., van Maris, A. J. A., de Winde, J. H., & Pronk, J. T. Chemostat-based micro-array analysis in baker's yeast. *Adv. Microb. Physiol.* **54**, 257-311 (2009).
222. Schmitt, M. E., Brown, T. A., & Trumpower, B. L. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nuc. Acids Res.* **18**, 3091-3092 (1990).
223. Gietz, D. R. & Woods, R. A. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method in *Methods in Enzymology, Guide to Yeast Genetics and Molecular and Cell Biology - Part B* (Academic Press, 2002).
224. Güldener, U., Heinisch, J., Koehler, G. J., Voss, D., & Hegemann, J. H. A second set of *loxP* marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nuc. Acids Res.* **30**, e23 (2002).
225. Engel, S. R. & Cherry, J. M. The new modern era of yeast genomics: community sequencing and the resulting annotation of multiple *Saccharomyces cerevisiae* strains at the *Saccharomyces* Genome Database. *Database* **2013**, (2013).
226. Mulder, N. & Apweiler, R. InterPro and InterProScan: tools for protein sequence classification and comparison in *Comparative Genomics* (Humana Press, 2007).
227. Nielsen, J., Larsson, C., van Maris, A., & Pronk, J. Metabolic engineering of yeast for production of fuels and chemicals. *Curr. Opin. Biotechnol.* **24**, 398-404 (2013).
228. Milanez, S. & Mural, R. J. Cloning and sequencing of cDNA encoding the mature form of phosphoribulokinase from spinach. *Gene* **66**, 55-63 (1988).
229. Brandes, H. K., Hartman, F. C., Lu, T. Y., & Larimer, F. W. Efficient expression of the gene for spinach phosphoribulokinase in *Pichia pastoris* and utilization of the recombinant enzyme to explore the role of regulatory cysteinyl residues by site-directed mutagenesis. *J. Biol. Chem.* **271**, 6490-6496 (1996).
230. Shively, J. M., van Keulen, G., & Meijer, W. G. Something from almost nothing: Carbon Dioxide Fixation in Chemoautotrophs. *Annu. Rev. Microbiol.* **52**, 191-230 (1998).
231. Tabita, F. R., Satagopan, S., Hanson, T. E., Kreel, N. E., & Scott, S. S. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J. Exp. Bot.* **59**, 1515-1524 (2008).
232. Quivey, J. & Robert Tabita, F. Cloning and expression in *Escherichia coli* of the form II ribulose 1,5-bisphosphate carboxylase/ oxygenase gene from *Rhodopseudomonas sphaeroides*. *Gene* **31**, 91-101 (1984).

233. Somerville, C. & Somerville, S. Cloning and expression of the *Rhodospirillum rubrum* ribulose bisphosphate carboxylase gene in *Escherichia coli*. *Mol. Gen. Genet.* **193**, 214-219 (1984).
234. Goloubinoff, P., Gatenby, A. A., & Lorimer, G. H. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. *Nature* **337**, 44-47 (1989).
235. Hayashi, N. R., Arai, H., Kodama, T., & Igarashi, Y. The *cbbQ* genes, located downstream of the form I and form II RubisCO genes, affect the activity of both RubisCOs. *Biochem. Biophys. Res. Commun.* **265**, 177-183 (1999).
236. Chen, D. H. *et al.* Visualizing GroEL/ES in the act of encapsulating a folding protein. *Cell* **153**, 1354-1365 (2013).
237. Hernández, J. M., Baker, S. H., Lorbach, S. C., Shively, J. M., & Tabita, F. R. Deduced amino acid sequence, functional expression, and unique enzymatic properties of the form I and form II ribulose bisphosphate carboxylase/oxygenase from the chemoautotrophic bacterium *Thiobacillus denitrificans*. *J. Bacteriol.* **178**, 347-356 (1996).
238. Durfee, T. *et al.* The complete genome sequence of *Escherichia coli* DH10B: Insights into the biology of a laboratory workhorse. *J. Bacteriol.* **190**, 2597-2606 (2008).
239. Beller, H. R. *et al.* The genome sequence of the obligately chemolithoautotrophic, facultatively anaerobic bacterium *Thiobacillus denitrificans*. *J. Bacteriol.* **188**, 1473-1488 (2006).
240. Hayashi, N. R., Arai, H., Kodama, T., & Igarashi, Y. The novel genes, *cbbQ* and *cbbO*, located downstream from the RubisCO genes of *Pseudomonas hydrogenothermophila*, affect the conformational states and activity of RubisCO. *Biochem. Biophys. Res. Commun.* **241**, 565-569 (1997).
241. Jariyachawalid, K., Laowanapiban, P., Meevootisom, V., & Wiyakrutta, S. Effective enhancement of *Pseudomonas stutzeri* D-phenylglycine aminotransferase functional expression in *Pichia pastoris* by co-expressing *Escherichia coli* GroEL-GroES. *Microb. Cell Fact.* **11**, 47 (2012).
242. Martínez-Alonso, M. et al. Rehosting of bacterial chaperones for high-quality protein production. *Appl. Environ. Microbiol.* **75**, 7850-7854 (2009).
243. Daran-Lapujade, P. *et al.* Role of transcriptional regulation in controlling fluxes in central carbon metabolism of *Saccharomyces cerevisiae*: a chemostat culture study. *J. Biol. Chem.* **279**, 9125-9138 (2004).
244. Harder, W., Kuenen, J. G., & Matin, A. Microbial selection in continuous culture. *J. Appl. Microbiol.* **43**, 1-24 (1977).
245. Garcia Sanchez, R., Hahn-Hägerdal, B., & Gorwa-Grauslund, M. *PGM2* overexpression improves anaerobic galactose fermentation in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **9**, 40 (2010).
246. Mumberg, D., Müler, R., & Funk, M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**, 119-122 (1995).
247. Grote, A. *et al.* JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nuc. Acids Res.* **33**, W526-W531 (2005).
248. Winzeler, E. A. *et al.* Functional characterization of the *Saccharomyces cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901-906 (1999).
249. Kuijpers, N. *et al.* A versatile, efficient strategy for assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60 bp synthetic recombination sequences. *Microb. Cell Fact.* **12**, 47 (2013).
250. Gietz, R. D. & Woods, R. Yeast transformation by the LiAc/SS carrier DNA/PEG method in *Yeast Protocol* (Humana Press, 2006).
251. Zerbino, D. & Birney, E. Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821-829 (2008).
252. Basso, T. O. *et al.* Engineering topology and kinetics of sucrose metabolism in *Saccharomyces cerevisiae* for improved ethanol yield. *Metab. Eng.* **13**, 694-703 (2011).
253. Mashego, M. R., van Gulik, W. M., Vinke, J. L., & Heijnen, J. J. Critical evaluation of sampling techniques for residual glucose determination in carbon-limited chemostat culture of *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **83**, 395-399 (2003).

254. Luttik, M. A. H. et al. The *Saccharomyces cerevisiae ICL2* gene encodes a mitochondrial 2-methylisocitrate lyase involved in propionyl-coenzyme A metabolism. *J. Bacteriol.* **182**, 7007-7013 (2000).
255. MacElroy, R. D., Mack, H. M., & Johnson, E. J. Properties of phosphoribulokinase from *Thiobacillus neapolitanus*. *J. Bacteriol.* **112**, 532-538 (1972).
256. Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265-275 (1951).
257. Beudeker, R. F., Cannon, G. C., Keunen, J. G., & Shively, J. M. Relations between d-ribulose-1,5-bisphosphate carboxylase, carboxysomes and CO_2 fixing capacity in the obligate chemolithotroph *Thiobacillus neapolitanus* grown under different limitations in the chemostat. *Arch. Microbiol.* **124**, 185-189 (1980).

Summary

of the PhD thesis
"Novel strategies for engineering redox metabolism in *Saccharomyces cerevisiae*"

In its search to decrease the environmental impact of the production of materials and food, and for other socio-economic reasons, mankind has recently taken the first steps into a paradigm shift from a petrochemical-based society to a new, sustainable and to a significant extent bio-based society. In this new scenario, biomass derived from agriculture and forestry industry and/or its residues are used for the generation of chemicals and materials that are currently derived from oil. To produce these chemicals, microorganisms and/or enzyme-catalyzed reactions can be used to convert the carbohydrates present in the biomass into a wide variety of products. Industrial biotechnology studies these conversion processes and aims to improve them. Nowadays, the single largest process in industrial biotechnology, in terms of product volume, is the production of ethanol from sugars using baker's yeast (*Saccharomyces cerevisiae*). The study of this model organism in terms of physiology, metabolism and genetics has contributed to significant improvements in industrial biotechnology. However, to compete with the efficiency and economics of petrochemical industry, also for products other than ethanol, further major breakthroughs are necessary. To achieve these advances, a knowledge-based redesign, followed by implementation of genetic changes via state-of-the-art molecular biology tools is used to improve cellular activities of microbial cell factories. This discipline in applied science is called Metabolic Engineering.

An important constraint in the design of metabolic engineering strategies is the balancing of reactions in metabolism that involve the transfer of electrons by the redox cofactor couples NADH/NAD^+ and NADPH/$NADP^+$. For instance, redox cofactor balancing plays a central role in the formation of glycerol as the major by-product of alcoholic fermentation by bakers' yeast. Under aerobic conditions, oxygen is used as the final electron acceptor. However, under the anaerobic conditions that are required for cost effective production of ethanol, the yeast *S. cerevisiae* produces glycerol to reoxidize the "excess"-NADH derived from biosynthetic reactions. This PhD thesis aimed to explore new strategies to increase the flexibility of the (anaerobic) redox metabolism of *S. cerevisiae* for the production

of fuels and chemicals. After a general introduction to industrial biotechnology and *S. cerevisiae*, **Chapter 1** describes the concept of metabolic engineering and how it has been applied in the past to reduce the formation of glycerol formation in alcoholic fermentation.

The ability of yeast cells to use glycerol as a redox sink can be eliminated by the double deletion of the genes encoding for glycerol-3-phosphate dehydrogenase (*GPD1* and *GPD2* in *S. cerevisiae*). In order to grow anaerobically, such mutants depend on the availability of an external electron acceptor, such as for instance externally added acetoin. However, addition of such a compound is too expensive to use on industrial scale. Of course, this is not a problem when the external electron acceptor is already present in the industrial feedstock. Acetic acid is present in small quantities in first generation feedstocks for ethanol production and in larger amounts as product from hydrolysis processes in lignocellulosic biomass (second generation feedstocks). Since, acetic acid is more oxidized than ethanol, its NADH-dependent reduction to ethanol could theoretically obviate the need for glycerol production in anaerobic cultures of *S. cerevisiae*. In practice, this does not occur because acetic acid is almost fully dissociated at the near-neutral pH inside yeast cells. Other microorganisms can use acetic acid as electron acceptor. These microorganisms use a linear pathway that first activates acetic acid to acetyl-CoA in a reaction that costs two ATP equivalents. In two subsequent reactions coupled to NADH oxidation, acetyl-CoA is first reduced to acetaldehyde, which is then reduced to ethanol. Only the reaction that reduces acetyl-CoA towards acetaldehyde, catalyzed by an (acetylating) acetaldehyde dehydrogenase, is not naturally present in *S. cerevisiae*. In **Chapter 2**, the replacement of glycerol formation as redox sink by acetic acid as external electron acceptor was studied in anaerobic cultures of a *gpd1*Δ *gpd2*Δ *S. cerevisiae* expressing the *Escherichia coli mhpF* gene encoding an (acetylating) acetaldehyde dehydrogenase. Growth of the constructed strain (*gpd1*Δ *gpd2*Δ *mhpF*) at a maximum specific growth rate of 0.14 h^{-1} was dependent on the presence of acetic acid. Under these conditions, the strain did not produce glycerol and showed a 13% higher ethanol yield on glucose than the isogenic reference strain (*GPD1 GPD2*). These results are a proof of concept for this metabolic engineering strategy, where glycerol formation was replaced by the removal of, otherwise inhibitory, acetic acid from lignocellulosic hydrolysates, leading to a significant increase in ethanol yield on glucose.

Besides its role in redox-cofactor balancing, glycerol is also the main compatible solute of *S. cerevisiae*, accumulating inside yeast cells when they face high extracellular osmotic pressure. This type of stress is especially important in first generation alcoholic fermentation, where a high initial sugar concentration is present

at the start of the production process. In **Chapter 3**, the tolerance to high osmotic pressure of a strain lacking functional glycerol-3-phosphate dehydrogenases (Gpd$^-$) and able to use acetate as electron acceptor was studied. Based on these findings, evolutionary engineering for anaerobic growth at high sugar concentrations (1 M glucose) was used to obtain an osmotolerant Gpd$^-$ *S. cerevisiae* strain expressing an (acetylating) acetaldehyde dehydrogenase. After the desired phenotype was obtained, single colonies were isolated and characterized under relevant conditions. An isolated evolved Gpd$^-$ strain grew anaerobically at 1 M glucose at a maximum specific growth rate of 0.12 h^{-1} in the presence of acetic acid (2 g l^{-1}). Surprisingly, formation of glycerol was observed again towards the end of the fermentation, albeit at much lower concentrations than in the Gpd$^+$ reference strain under identical conditions. Moreover, the evolved strain exhibited an apparent higher ethanol yield on glucose than the reference strain, reaching a value of 1.84 mol mol^{-1} (92% of the theoretical ethanol yield on glucose), when cultured in the presence of 3 g l^{-1} of acetic acid and 1 M glucose. Genetic analysis of the evolved strain revealed that this evolved phenotype was the consequence of one dominant chromosomal mutation, and one mutation in the plasmid-borne *mhpF* gene for anaerobic growth.

In industrial biotechnology, the use of evolutionary engineering coupled to reverse metabolic engineering is a powerful tool in the development of strains with desired phenotypes and in the transfer of these characteristics to different strains. However, during laboratory or industrial evolution, also undesired phenotypes are observed, which are equally important to comprehend: only by understanding their molecular basis, such phenotypes can be removed from production strains. Cell flocculation and/or aggregation were frequently observed during laboratory evolution in sequential batch cultures under a wide variety of selective culture conditions. Although flocculation can be a desired phenotype in beer brewing, where it may facilitate separation of the yeast, it is not desirable during evolutionary engineering, where it diverts the selective pressure away from the target of interest and furthermore complicates the generation of single cell lines. This phenomenon, which was also observed in some of the evolutionary engineering lines with Gpd$^-$ *S. cerevisiae* strains that were performed in the context of this thesis, was further studied in **Chapter 4** with a strain obtained from laboratory evolution for faster consumption of glucose-galactose mixtures. Reverse metabolic engineering of 'multicellular' strains of *S. cerevisiae* obtained from this study revealed that genome duplication and deregulation of the cell cycle were key elements in the development of a multicellular/agglomeration phenotype. Whole genome sequencing of two single colony isolates from independent laboratory evolutions showed that multicellular phenotype resulted from different point mutations in *ACE2* gene, a

key transcriptional regulator in the separation between bud and mother cell in yeast. Moreover, both final evolved strains became diploids, whereas the original parental strain was haploid. The multicellular phenotype was reverted by introduction of a functional copy of the original *ACE2* allele in the evolved strains. Introduction of the mutant allele and doubling the genome size by mating in the parent strain led to the same multicellular, fast-sedimenting phenotype that was observed in the evolved strains. These results do not only shed light on the mutations that underlie the evolution of multicellular yeast strains, but can also be applied to induce or eliminate cell aggregation in industrial strains.

Major breakthroughs in industrial biotechnology benefit from rapidly developing techniques in synthetic biology. These advantages allow scientists to use bolder and more creative metabolic engineering strategies. An unexplored metabolic engineering strategy in redox metabolism is the use of carbon dioxide as electron acceptor via enzymes from autotrophic microorganisms. Carbon dioxide is a by-product of yeast fermentation and therefore abundantly present in industrial fermentation processes. The Calvin cycle, present in different autotrophic organisms, uses ATP and NADPH for carbon dioxide fixation and provides metabolic building blocks required in biosynthesis. Phosphoribulokinase and ribulose-1,5-bisphosphate carboxylase/oxidase (Rubisco) are the two key enzymes of the Calvin cycle and have been involved in the fixation of the majority of the organic carbon available in nature. Using state-of-the-art synthetic biology techniques, we functionally expressed both enzymes in a *S. cerevisiae* strain and studied its physiology in anaerobic carbon-limited chemostat cultures at a dilution rate of 0.05 h^{-1} (**Chapter 5**). Functional expression of a single-subunit Rubisco from the chemolithoautotrophic bacterium *Thiobacillus denitrificans* required the co-expression of chaperones proteins GroEL and GroES from *E. coli*, and resulted in the first functional expression of Rubisco in a heterotrophic eukaryotic microorganism. Anaerobic chemostat cultures sparged with pure nitrogen gas showed a 68% lower glycerol yield on sugars, and 11% and 12% higher ethanol and biomass yields on sugars, respectively, than an isogenic reference strain. Increasing the concentration of dissolved carbon dioxide by purging the fermenters with a gas mixture of 10% v/v CO_2 and 90% v/v nitrogen resulted in a further decrease of the glycerol yield on sugars to less than 10% of that observed in the reference strain. To study the performance of the engineered strain using a fermentation setup more relevant to industrial application, anaerobic batch cultures were run on 20 g l^{-1} galactose in laboratory fermenters sparged with a 10%-90 % mixture of CO_2 and nitrogen. Under these conditions, the constructed strain showed no differences in growth kinetics when compared to its isogenic reference, while its glycerol yield on

galactose decreased by 60% and its ethanol yield on galactose increased by 8% relative to the reference strain. This strategy not only demonstrates the potential of using carbon dioxide as electron acceptor in the metabolic engineering of yeasts and other microorganisms, but also illustrates how co-expression of chaperone proteins can aid the functional expression of bacterial proteins in yeast cytosol.

To conclude, this thesis presents two novel metabolic engineering strategies (Chapter 2 & 5) that increased the flexibility of redox metabolism in anaerobic fermentation by implementing novel redox sinks in baker's yeast. Both strategies have high potential to substantially contribute to optimizing product formation by *S. cerevisiae* in anaerobic industrial biotechnological production processes. In such processes even a small increase in product yield on substrate, given the large volumes of production, can result in large economic benefits without the introduction of new process steps. Even though novel strategies can show high potential for industrial application, an integrative view on process optimization should also consider potential negative effects on strain robustness. Evolutionary engineering (Chapter 3) and reverse metabolic engineering (Chapter 4) of desired and undesired phenotypes, provide additional powerful tools in the implementation of metabolic engineering strategies. The chapters of this thesis represent small but significant steps in the vigorous research needed to enable wide-spread, economically viable and sustainable production of transport fuels and chemicals in a bio-based society.

Samenvatting

van het proefschrift "Nieuwe strategieën voor gerichte verandering van de redoxstofwisseling in *Saccharomyces cerevisiae*"

Tegenwoordig probeert men steeds meer de invloed te verminderen die de productie van materialen en voedsel heeft op het milieu, maar ook vanwege andere sociaaleconomische redenen vindt er een verschuiving plaats van een samenleving gebaseerd op petrochemie naar een nieuwe, duurzame en voor een groot gedeelte biologisch gebaseerde samenleving. Hiertoe worden de grondstoffen zoals aardolie, dat nu vooral gebruikt wordt voor de productie van chemicaliën en materialen, vervangen door biomassa en residuen uit de landbouw- en bosbouwindustrie. De koolwaterstoffen uit deze zogenaamde 'groene grondstoffen' kunnen met behulp van enzymen of micro-organismen omgezet worden naar een breed scala van producten. Het vakgebied 'industriële biotechnologie' spitst zich toe op het onderzoeken en vervolgens verbeteren van deze conversies. Vandaag de dag is, qua tonnage, de conversie van suikers naar ethanol door bakkersgist (*Saccharomyces cerevisiae*) het grootste industrieel-biotechnologische proces. De studie van dit modelorganisme op het gebied van fysiologie, metabolisme en genetica heeft enorm bijgedragen aan de optimalisatie van biotechnologische processen. Echter, om financieel en op efficiëntie te kunnen concurreren met de petrochemische industrie, zowel voor ethanol als andere producten, zijn er grote doorbraken nodig. Om deze vooruitgangen te boeken is een kennis-gebaseerd herontwerp, gevolgd door de implementatie van genetische veranderingen met behulp van "state-of-the-art" moleculair biologische technieken vereist om de cellulaire activiteiten van microbiologische celfabriekjes te verbeteren. Deze discipline in de toegepaste wetenschap wordt "metabolic engineering" genoemd.

Een belangrijk punt in het ontwerp van (heterologe) metabole routes is het balanceren van reacties waarbij voor elektronenoverdracht de redoxcofactoren NADH/NAD^+ of NADPH/$NADP^+$ nodig zijn. Om deze redoxevenwichten in stand te houden produceert *S. cerevisiae* tijdens anaërobe groei het bijproduct glycerol. Bij aërobe groei is zuurstof de laatste elektronenacceptor. Echter, onder de anaërobe condities die gebruikt worden voor de industriële productie van alcohol, produceert bakkersgist glycerol om de overmaat van NADH, gevormd bij anabole reacties, te re-oxideren en zodoende de NADH/NAD^+ homeostase te handhaven.

In dit proefschrift werd gezocht naar andere mogelijkheden voor de re-oxidatie van NADH met een focus op de productie van groene moleculen en brandstoffen. In **Hoofdstuk 1** wordt, in een algemene inleiding over industriële biotechnologie en *S. cerevisiae*, geschetst wat "metabolic engineering" inhoudt en hoe het toegepast kan worden om glycerolproductie tijden de alcoholische fermentatie te verminderen.

De productie van de elektronenacceptor glycerol kan in bakkersgist geëlimineerd worden door beide genen die coderen voor glycerol-3-fosfaat dehydrogenase (*GPD1* en *GPD2*) te verwijderen. Voor anaërobe groei hebben zulke mutanten dan wel de toevoeging van een andere elektronenacceptor nodig, bijvoorbeeld acetoïne. Op industriële schaal brengt dit echter hoge kosten met zich mee en is daarmee niet wenselijk. Dit probleem vervalt echter wanneer in de grondstof chemische verbindingen aanwezig zijn die als elektronenacceptor zouden kunnen dienen. Een kandidaat is azijnzuur, dat in kleine hoeveelheden aanwezig is in eerste-generatie grondstoffen voor ethanolproductie en in ruime mate in gehydrolyseerde lignocellulose-houdende biomassa (tweede- generatie grondstoffen). Omdat acetaat meer geoxideerd is dan ethanol, zou, in theorie, de NADH-afhankelijke reductie ervan naar ethanol de noodzaak van glycerolproductie in anaërobe *S. cerevisiae*-culturen kunnen wegnemen. Helaas heeft in gist het enzym aceetaldehyde dehydrogenase niet acetaat maar azijnzuur als substraat. Omdat, onder fysiologische condities, azijnzuur vrijwel alleen in gedissocieerde vorm voorkomt vindt de reductie van acetaat naar aceetaldehyde via deze route niet plaats. Er zijn echter wel andere micro-organismen die acetaat gebruiken als elektronenacceptor. In deze organismen wordt, door een ATP-afhankelijke reactie, eerst acetaat geactiveerd naar acetyl-CoA. Vervolgens wordt in twee reacties, die gekoppeld zijn aan NADH- oxidatie, acetyl-CoA gereduceerd tot aceetaldehyde door acetylerend aceetaldehyde dehydrogenases (A-ALD), waarna aceetaldehyde gereduceerd wordt tot ethanol door alcohol dehydrogenase. In *S. cerevisiae* ontbreekt alleen het acetylerend acetaldehyde dehydrogenase. In **Hoofdstuk 2** wordt beschreven hoe anaërobe glycerolproductie vervangen kan worden door de introductie van een A-ALD uit *E. coli* (*mhpF*) in een *gpd1*⊿ *gpd2*⊿ stam en het toevoegen van azijnzuur. De anaërobe groei van deze stam (*gpd1*⊿ *gpd2*⊿ *mhpF*) met een maximale specifieke groeisnelheid van 0,14 uur^{-1} was strikt afhankelijk van toegevoegd azijnzuur. Onder deze condities produceerde de stam geen glycerol en de ethanolopbrengst op glucose was 13% hoger dan de referentie stam (*GPD1 GPD2*). Deze strategie, waarbij glycerolproductie vervangen wordt door acetaatconsumptie, is een belangrijke innovatie omdat er zowel toxisch acetaat uit de lignocellulose-hydrolysaten wordt opgeruimd als meer ethanol geproduceerd wordt.

Als tweede belangrijke rol is glycerol ook een belangrijk "compatible solute", waarvan ophoping in de cel de tolerantie voor hoge extracellulaire osmotische druk vergroot. Dit is extra belangrijk voor groei op de eerste-generatie grondstoffen, waar in het begin van alcoholische fermentatie hoge suikerconcentraties aanwezig zijn. In **Hoofdstuk 3** werd gekeken naar de tolerantie van een stam zonder functionele glycerol-3-fosfaat dehydrogenases (Gpd$^-$) voor hoge osmotische druk met acetaat als elektronenacceptor. Vervolgens werd, om een osmotolerante stam te krijgen, evolutionaire engineering toegepast op een Gpd$^-$ stam, met een acetylerend aceetaldehyde dehydrogenase, bij hoge suikerconcentraties (1 M glucose), onder anaërobe condities. Na het verkrijgen van het gewenste fenotype werden enkele kolonies geïsoleerd, die verder gekarakteriseerd werden onder relevante condities. Eén van deze Gpd$^-$ mutanten groeide anëeroob bij 1 M glucose en 2 g l^{-1} azijnzuur met een maximale specifieke groeisnelheid van 0,12 uur^{-1}. Verrassend genoeg maakte deze stam aan het einde van de fermentatie toch glycerol aan, zij het met veel lagere concentraties dan de Gpd$^+$ referentie stam onder gelijke condities. Daarnaast had de geëvolueerde stam een hogere schijnbare ethanolopbrengst op glucose dan de referentiestam, met waardes tot 1,84 mol mol^{-1} (92% van het theoretisch maximum), bij kweken met 3 g l^{-1} azijnzuur en 1 M glucose. Genetische analyse wees uit dat dit geëvolueerde fenotype het resultaat was van een dominante chromosomale mutatie en één mutatie op het plasmide met het *mhpF* gen.

In industriële biotechnologie is evolutionaire engineering in combinatie met reverse engineering een zeer krachtig hulpmiddel voor het verkrijgen van gewenste fenotypes en het overzetten van gewenste eigenschappen naar andere stammen. Soms worden er in industriële- en labevolutie echter ook ongewenste fenotypes gevormd. Het is evenzeer belangrijk ook deze fenotypes te onderzoeken: alleen door de moleculaire basis te begrijpen kunnen productiestammen ontdaan worden van deze ongewenste eigenschappen. Celflocculatie en/of -aggregatie is een veelvoorkomend fenomeen in laboratorium-evolutie-experimenten waarvoor sequentiële batchculturen gebruikt worden. Dit fenomeen is waargenomen onder verschillende selectieve kweekcondities. Hoewel flocculatie een gewenst fenotype kan zijn in de bierindustrie waar het de scheiding van het bier en de gist vereenvoudigt, is dit niet een gewenst fenotype tijdens evolutionaire engineering, waar het de selectiedruk voor het gewenste fenotype verstoort en het verkrijgen van individuele cellijnen vermoeilijkt. Dit fenomeen, wat ook werd waargenomen in enkele evolutionaire engineeringlijnen van Gpd$^-$ *S. cerevisiae* stammen in de context van dit proefschrift, is verder bestudeerd in **Hoofdstuk 4** met een stam verkregen door labevolutie voor snellere consumptie van glucose-galactose mengsels. Reverse

metabole engineering van de geëvolueerde 'multi-cellulaire' *S. cerevisiae* stammen, verkregen uit deze studie, liet zien dat genoomduplicatie en deregulatie van de celcyclus belangrijk waren voor de ontwikkeling van een multi-cellulair/agglomererend fenotype. Het bepalen van de DNA-volgorde van het hele genoom van twee individuele kolonies uit twee onafhankelijke evolutielijnen wees uit dat het multi-cellulaire fenotype het resultaat was van verschillende puntmutaties in het *ACE2* gen, een belangrijke transcriptie-regulator voor het scheiden van de knop en de moedercel in een delende gist. Verder waren beide geëvolueerde stammen diploïd geworden, hoewel de moederstam haploïd was. Het multi-cellulaire fenotype van de geëvolueerde stammen kon ongedaan gemaakt worden door een functioneel kopie van het originele *ACE2* gen tot expressie te brengen. De introductie van een gemuteerd *ACE2* allel met een door kruising verkregen genoomduplicatie in de moederstam leidde tot hetzelfde multi-cellulaire, snel sedimenterende fenotype als de geëvolueerde stammen. Deze resultaten verklaren niet alleen welke mutaties ten gronde liggen aan de evolutie van multi-cellulaire stammen, maar kunnen ook toegepast worden om in industriële stammen celaggregatie te induceren of te elimineren.

Belangrijke doorbraken in industriële biotechnologie hebben veel te danken aan de snelle ontwikkeling van technieken in de synthetische biologie. Hierdoor kunnen wetenschappers gedurfdere en creatievere metabolic engineering-strategieën toepassen. Een onverkende metabole engineeringstrategie in het redoxmetabolisme is het gebruik van koolstofdioxide als elektronenacceptor met behulp van enzymen uit autotrofe micro-organismen. Koolstofdioxide is een bijproduct van gistfermentatie en dus overvloedig aanwezig in industriële fermentatieprocessen. De Calvincyclus, aanwezig in verschillende autotrofe organismen, gebruikt ATP en NADPH voor koolstoffixatie en levert metabole bouwstenen die nodig zijn voor biosynthese. Fosforibulokinase en ribulose-1,5-bifosfaatcarboxylase/oxidase (Rubisco) zijn de twee belangrijkste enzymen van de Calvincyclus en daarmee verantwoordelijk voor een groot deel van de natuurlijke fixatie van koolstofdioxide. Met behulp van state-of-the-art synthetische biologische technieken werden beide enzymen functioneel tot expressie gebracht in een *S. cerevisiae*-stam. Vervolgens werd de fysiologie bestudeerd in anaërobe, koolstof-gelimiteerde chemostaatculturen met een verdunningssnelheid van 0,05 uur^{-1} (**Hoofdstuk 5**). Functionele expressie van een single-subunit Rubisco van de chemolithoautotrofe bacterie *Thiobacillus denitrificans* vereiste de co-expressie van de chaperone eiwitten GroEL en GroES van *E. coli*, resulterend in de allereerste functionele expressie van Rubisco in een heterotroof eukaryoot micro-organisme. Anaërobe chemostaatcultures van deze giststam waar puur stikstofgas doorheen werd geleid hadden een 68% lagere

glycerolopbrengst op suikers en, respectievelijk, 11% en 12% hogere ethanol- en biomassaopbrengsten op glucose in vergelijking tot een isogene referentiestam. De concentratie van opgelost koolstofdioxide in de cultures werd verhoogd door er een gasmengsel van 10% v/v CO_2 en 90% v/v stikstof door te leiden. Hierdoor werd de glycerolopbrengst nog lager, zelfs minder dan 10% dan in cultures van de referentiestam. Om de eigenschappen van de Rubisco-giststam verder te onderzoeken in een fermentatiesysteem met relevantie voor industriële toepassingen, werden er anaërobe batchculturen gekweekt op 20 g l^{-1} galactose in labfermentors, waardoor een mengsel van 10%/90% van CO_2 en stikstof werd geleid. Onder deze condities vertoonde de stam geen verschillen in groeikinetiek in vergelijking tot de isogene referentiestam, echter de glycerolopbrengst op galactose was 60% lager en de ethanolopbrengst was 8% hoger dan de referentiestam. Deze strategie laat niet alleen de potentie zien van het gebruik van koolstofdioxide als elektronenacceptor in de metabolic engineering van gisten en andere micro-organismes, maar het illustreert ook hoe de co-expressie van chaperone-eiwitten kan helpen bij de functionele expressie van bacteriële eiwitten in het cytosol van gist.

Tot slot: dit proefschrift laat twee nieuwe metabolic engineering strategieën zien (Hoofdstuk 2 & 5) waarmee de flexibiliteit van het redoxmetabolisme in anaërobe fermentatieprocessen vergroot kan worden door de implementatie van nieuwe redox 'sinks' in het metabolisme van bakkersgist. Beide strategieën hebben veel potentie om substantieel bij te dragen aan de optimalisatie van productvorming door *S. cerevisiae* in anaërobe industriële biotechnologische productieprocessen. In zulke processen kan, gezien de grote productievolumes, zelfs een kleine verhoging van de productopbrengst een groot economisch voordeel te weeg brengen, dit zonder nieuwe processtappen toe te hoeven voegen. Hoewel de nieuwe strategieën grote potentie kunnen hebben voor industriële toepassingen, moeten in een geïntegreerde analyse voor procesoptimalisatie ook de potentiële negatieve effecten op stamrobuustheid meegenomen worden. Evolutionaire engineering (Hoofdstuk 3) en reverse metabolic engineering (Hoofdstuk 4) van gewenste en ongewenste fenotypes leveren krachtige hulpmiddelen voor de implementatie van metabole engineeringstrategieën. De hoofdstukken van dit proefschrift representeren kleine, maar significante stappen in het grondige onderzoek wat nodig is om wijdverbreid, economisch haalbare en duurzame productie van chemicaliën en transportbrandstoffen in een biologisch gebaseerde samenleving mogelijk te maken.

Curriculum Vitae

Víctor Gabriel Guadalupe Medina was born on the 23rd of August 1982 in Rancagua, Chile. After his primary and secondary education at the Instituto O'Higgins in Rancagua, Víctor started his Civil Engineering studies at the Catholic University of Chile (UC Chile, Santiago) in March 2001. Based on his fascination for manufacturing processes, Victor chose the subject Industrial Engineering with a specialization in Bioprocess Engineering in his 2nd year of studies. During a 3-month research internship at Fluxome A/S and at the Centre for Microbial Biotechnology from the Danish Technical University in 2006, Victor discovered his passion for scientific research. In March 2007, Victor started his MSc studies in Biotechnology at UC Chile. In his MSc thesis project, he studied the production of aromatic compounds by commercial wine yeast strains of *Saccharomyces cerevisiae* under the supervision of Francisco Pizarro and prof. Eduardo Agosin. After obtaining his MSc degree in Biotechnology in August 2008, Víctor picked up his back pack and moved to Europe to start his PhD research at the Industrial Microbiology section of the Department of Biotechnology of the Delft University of Technology in September 2008. Under the supervision of Jack Pronk and Ton van Maris, Victor explored new strategies for metabolic engineering of redox metabolism in *S. cerevisiae*. The results of this research are compiled in this thesis: "Novel strategies for engineering redox metabolism in *Saccharomyces cerevisiae*". In February 2013, Víctor started the next step in his scientific career in 'The Verstrepen Lab' for Systems Biology at Katholieke Universiteit Leuven in Belgium to study feral yeast strains using high-throughput techniques and genome sequencing.

List of Publications

Molina A.M., **Guadalupe V.**, Varela C., Swiegers J.H., Pretorius I.S., Agosin E. 2009. Differential synthesis of fermentative aroma compounds of two related commercial wine yeast strains. *Food Chemistry* **117**: 189-195.

Guadalupe Medina V.G., Almering M.J.H., van Maris A.J.A., Pronk J.T. 2010. Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. *Applied and Environmental Microbiology* **76**:190-195.

Guadalupe-Medina V., Metz B., Oud B., van der Graaf C.M., Mans R., Pronk J.T., van Maris A.J.A. 2013. Evolutionary engineering of glycerol-3-phosphate dehydrogenase-negative acetate-reducing *Saccharomyces cerevisiae* enables high ethanol yields at one-molar glucose concentrations. *Microbial Biotechnology*, early view article.

Oud B., **Guadalupe-Medina V.**, Nijkamp J.F., de Ridder D., Daran J.M., Pronk J.T., van Maris A.J.A. 2013. Genome duplication and mutations in *ACE2* cause multicellular, fast-sedimenting phenotypes in evolved *Saccharomyces cerevisiae*. Submitted for publication.

Guadalupe-Medina V., Wisselink H.W., Luttik M.A.H., de Hulster E., Daran J. M., Pronk J.T. and van Maris A.J.A. 2013. Carbon dioxide fixation by Calvin-cycle enzymes improves ethanol yield in yeast. *Biotechnology for Biofuels*, **6**:125.

Guadalupe Medina V. G., Almering M. J. H., van Maris A. J. A., and Pronk J.T. 2009. Fermentative glycerol-free ethanol production. International Publication Number WO2011/010923 A1.

Acknowledgements

It is fun to write my thesis acknowledgements at the end of my 4th year living in Europe, looking back at so many experiences and memories of people that have helped me to grow professionally and personally. I will do my best to remember everyone, but please forgive if I forget to mention someone.

My first big 'hartelijk bedankt' goes to Jack. Almost 6 years have passed since I received his answer to my e-mail asking for an opportunity to do a PhD in his lab. A few months later, I still remember his enthusiastic answer after I accepted his offer to join the IMB group. Moreover, I am very thankful to him for the life-changing chance to be part of all "joint exciting scientific adventures" we had during these last 4 years, and for sharing his great enthusiasm for science. The second 'van harte bedankt' goes to Ton. In my opinion, Jack's enthusiasm has a perfect match with Ton's excellent critical thinking. I have learned a lot of this valuable quality from Ton, and I thank him for all his support and tutoring. My PhD project was part of program 1 of the Kluyver Centre of Genomics of Industrial Fermentations, which was supported by The Netherlands Genomics Initiative. I want to thank both these institutions for their financial support and the chance this has given me to advance my scientific career.

Many thanks go also to the rest of the permanent staff of the IMB group. Big thanks Jean-Marc and Marinka for helping me learn the basics of molecular biology and for their contribution to my thesis. A big muchas gracias to Marijke for her contribution to my thesis, her total support in my IRI-adventures, and other quests in the enzymatic world. Erik, the 2nd fermentation master of my life, I thank him very much for his contribution, the constant support and for sharing his wisdom with me. To Pascale and Marcel, I am grateful for your help, comments, and enjoyable conversations on science and other matters.

Ook mijn hartelijke dank voor de MSD dames: Apilena, Astrid en Jannie. Bedankt voor jullie geduld met laat steleriseren, in het bijzonder met al die 20 L glucose flessen en natuurlijk voor jouw hulp in Nederlands leren en oefenen. Thanks also to all the rest of the support staff in the Biotechnology Department, especially Jos and Sjaak and the workshop guys: Arno, Marko, and Marcel. En bijzonder bedankt ook voor Mevrouw Naima, het was altijd een blij-moment op de dag om met jou praten en te zeggen mucho trabajo poco dinero.

Very big thanks goes to my 4 students: Mirjam, Frank, Robert, and Rudi. It was a pleasure to work with all of them and helping them to develop their professional skills. I am very thankful for the opportunity they gave me to learn how

to motivate people and for their important contributions to my thesis. Also thanks to all the many students that I met in the last 4 years, who are also important to the great environment in the Industrial Microbiology Section, always bringing fresh air to the labs.

Dear (ex-)colleagues, I would like to mention you all, but I prefer to avoid the risk of forgetting one of you. I prefer to say to you all that it was a privileged to work with you, that I have learned a lot from all of you and that I am thankful for your help in my development in these 4 years. A very special thanks to two molecular ladies: Barbara C. (AKA 12006), another big teacher in the molecular world, but most important a very dear friend; and Lizanne, the movie-fan/afternoon-choco-partner kiddo that taught me a few molecular tricks and shared fun coffee-break talks. Big thanks go to Jeremiah and Huong, from which I learned a lot about life and I missed having them around a lot after they left. Also big thanks to Rintze, it was fun and a honor to be your paranymph. Very big thanks go to my own paranymphs: Harmen and Barbara K., for their support and help in the thesis post-processing and party. Ein sehr groβes Danke schön to Benjamin for all the fun and mutual learning shared in the glycerol-project. It was great to see during the last year all the advances that are bringing the concept much closer to industrial reality. Also big thanks go to Bart for all the great science and life conversations, and especially for sharing the awesome world of reverse metabolic engineering, sequence analysis, and flocculation. A final special thanks for the good times goes to my office-mates in the fish tank: Joost, Lucie, Rintze, and Irina, and in room .150: Gabriele, Filipa, Gijs, Dick, Bianca, Ruben, Nick, and though short but not less pleasant, Reno and Emilio. I would like to thank especially Gijs for his enthusiasm and support on the Rubisco project, it was a pleasure to talk science with him and an honor to share an office with him.

I would also like to thank Fundacion Andes, Jochen Förster, Prof. Eduardo Agosin from UC Chile, and Prof. Jens Nielsen from Chalmers University, for their support on my 1st science adventure in Denmark, where I discovered my passion for research.

Now, being a Spanish speaking person, I am sorry, but there are people that I need to thank in my own language.

Muchas gracias a todos los amigos hispano-parlantes con quienes siempre pude compartir una cerveza en las tardes de Delft para transformar Holanda en Chile, aunque haya sido por un par de horas. En especial para Manuel y Carlos, con quienes los asados de verano, pocos pero buenos, serán definitivamente necesario repetirlos. Para mi apañero Fernando y Mariëlle, hartelijke bedankt por todo el apoyo, la amistad y las ricas cenas.

También un saludo especial va para mis amigos en Chile (en especial Luciano, Martin, JotaA, Álvaro, y Nico), gracias a quienes las visitas a la Madre Patria fueron cortas pero entretenidas. Muchas gracias también a Pancho Pizarro y Prof. Eduardo Agosin por enseñarme mis primeros pasos en la investigación.

Finalmente quiero agradecer de todo mi corazón a mi familia, a quienes he extrañado mucho en estos 4 años. Mis queridos sobrinos: Cata, Jose, Maxi, Conita y Anto, me da una pena enorme no estar cerca para verlos crecer, pero espero con ganas conocer los adultos que pronto serán, de los cuáles estaré muy orgulloso. A mis hermanos, a mi Tialy y a mis padres, muchas gracias por el apoyo y el cariño de toda una vida, todo lo que soy se lo debo a ustedes. Los quiero mucho, y aunque no sé cuando será el tiempo para estar cerca, tengan presente que siempre los tengo en mis pensamientos y en mi corazón.

My last words are for Sarah. Guapa, words are not enough but thank you very much for helping me grow in so many aspects, and for all your love and support during the last 2 years and a half. I am very proud of you, especially because of your courage, and I am happy to have you by my side to start a new life in Leuven.