Published in final edited form as:
Nat Biotechnol. 2020 February ; 38(2): 210–216. doi:10.1038/s41587-019-0363-0.

The industrial yeast *Pichia pastoris* is converted from a heterotroph into an autotroph capable of growth on CO_2

Thomas Gassler[1,2], **Michael Sauer**[1,2,3], **Brigitte Gasser**[1,2], **Michael Egermeier**[1], **Christina Troyer**[4], **Tim Causon**[4], **Stephan Hann**[4], **Diethard Mattanovich**[1,2,*], **Matthias G. Steiger**[1,2,5]

[1]Department of Biotechnology, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria

[2]Austrian Centre of Industrial Biotechnology (ACIB), Vienna, Austria

[3]CD-Laboratory for Biotechnology of Glycerol, Department of Biotechnology, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria

[4]Institute of Analytical Chemistry, Department of Chemistry, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria

[5]Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, Vienna, Austria

Abstract

The methylotrophic yeast *Pichia pastoris* is widely used in the manufacture of industrial enzymes and pharmaceuticals. Like most biotechnological production hosts, *P. pastoris* is heterotrophic and grows on organic feedstocks that have competing uses in the production of food and animal feed. In a step toward more sustainable industrial processes, we describe the conversion of *P. pastoris* into an autotroph that grows on CO_2. By addition of eight heterologous genes and deletion of three

*Corresponding author: Prof. Dr. Diethard Mattanovich, Department of Biotechnology, University of Natural Resources and Life Sciences, Vienna, Muthgasse 18, 1190 Vienna, Austria, diethard.mattanovich@boku.ac.at.

Reporting Summary
Additional information on research design is given in the Nature Life Sciences Research Reporting Summary.

Data availability statement
The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request. Genome sequencing reads of mutant strains are available at NCBI BioProject PRJNA586834.
Sequences of codon optimized genes are shown in the Supplementary Information.

Code availability statement
Mass spectrometry data was collected using Agilent Technologies MassHunter GCMS Acquisition Software B.07.06 SR1 combined with post acquisition application (Agilent SureMass). Theoretical masses were extracted from the SureMass-converted data and analyzed using the Agilent Technologies MassHunter Workstation Quantitative Analysis for TOF Build 10.0.707.0 and the Isotope Correction Toolbox Version 0.04. Genome sequence analysis including SNP calling was carried out using CLC Genomic Workbench 12.0 (QIAGEN). Microscopy pictures were processed using the Zen 2.3 lite (blue edition) software (Carl Zeiss Microscopy GmbH). Flow cytometry data was analyzed using CytExpert (version 2.3.0.84) (Beckman Coulter). No custom code was used in this study.

Author Contributions
DM conceived and initiated the project. DM, MGS, TG, MS, and BG designed the experiments. TG and ME carried out the experiments. SH, CT and TC performed mass spectrometry. TG, MGS, MS and DM analyzed the data. TG, MGS and DM wrote the manuscript. All authors read and approved the final manuscript.

Competing Financial Interests Statement
DM, TG, MGS, MS and BG are inventors of a patent application (application number PCT/EP2018/064158.) based on the results reported in this publication.

Europe PMC Funders Author Manuscripts

native genes, we engineer the peroxisomal methanol-assimilation pathway of *P. pastoris* into a CO_2 fixation pathway resembling the Calvin-Benson-Bassham cycle, the predominant natural CO_2 fixation pathway. The resulting strain can grow continuously with CO_2 as a sole carbon source at a μ_{max} of 0.008 h^{-1}. The specific growth rate was further improved to 0.018 h^{-1} by adaptive laboratory evolution. This engineered *P. pastoris* strain may promote sustainability by sequestering the greenhouse gas CO_2 and by avoiding consumption of an organic feedstock with alternative uses in food production.

Keywords

Pichia pastoris; synthetic biology; CO_2 assimilation; Calvin cycle; CBB cycle; RuBisCO

Introduction

Industrial biotech production of proteins and chemicals relies on heterotrophic host cells, which grow on feedstocks, such as glucose, sucrose or glycerol, that have competing uses in the production of food and animal feed. Biotech manufacturing would become more sustainable if it could use CO_2 as a carbon feedstock, as this would not consume organic feedstocks and would deplete atmospheric CO_2. CO_2 is converted into biomass by autotrophic organisms such as plants and algae. Channeling CO_2 into the central carbon metabolism of industrial microbes could generate autotrophic or mixotrophic organisms with the potential for increased carbon efficiency (e.g. mixed-substrate production of succinic or acetic acid from CO_2 and methanol)[1].

The Calvin-Benson-Bassham (CBB) cycle [2] is one of six naturally occurring CO_2 fixation pathways [3,4] and includes ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), which is the most abundant enzyme in the biosphere and fixes around 90% of the inorganic carbon converted into biomass [5]. Previously, metabolic engineers have endeavored to equip the heterotrophic model organisms *Escherichia coli*, *Methylobacterium extorquens* and *Saccharomyces cerevisiae* with CBB cycle genes. Expressing RuBisCO and phosphoribulokinase (PRK) genes in *E. coli* led to decreased release of CO_2 when growing on arabinose [6]. Antonosky et al. enabled “hemiautotrophic” growth of *E. coli* with a modularized carbon metabolism, utilizing pyruvate as an energy source for a heterologous CBB cycle [7]. Hemiautotrophic means that only part of the carbon assimilated originated from CO_2, with the rest derived from pyruvate. In a follow-up study the mutations that stabilized the non-native CBB cycle in *E. coli* [8] were identified. Parts of the CBB cycle have also been introduced into the yeast species *S. cerevisiae* [9–12]; ethanol yields were increased by using CO_2 as an additional electron acceptor to reoxidize NADH. More recently, the assimilatory pathway of *Methylobacterium extorquens AM1* was blocked and RuBisCO and PRK were overexpressed, resulting in a strain that can incorporate CO_2 into biomass using methanol as energy source. However, this engineered methylobacterium was unable to grow continuously on CO_2 as a sole carbon source [13]. Engineering of a synthetic autotrophic microorganism able to grow with CO_2 as the sole carbon source has not been reported to our knowledge.

Europe PMC Funders Author Manuscripts

Pichia pastoris (syn.: *Komagataella phaffii*) [14,15] is widely used to produce heterologous proteins for the biopharmaceutical [16] and enzyme markets [17]. *P. pastoris* is a chassis host for metabolic engineering [18] and a model for peroxisome biogenesis [19,20]. This industrial yeast has a methylotrophic lifestyle and uses C1 compound methanol as its sole energy and carbon source. Methanol is oxidized to formaldehyde, which then enters either a dissimilatory or assimilatory metabolic branch. In the assimilatory branch formaldehyde is converted to phosphosugars needed for biomass, whereas in the dissimilatory branch formaldehyde is oxidized to CO_2, yielding NADH. Steps of the dissimilatory branch are carried out both in the peroxisome and in the cytosol, but the assimilatory branch is localized entirely in peroxisomes [21]. During growth on methanol, peroxisomes are highly abundant in the cell and methanol utilization enzymes are highly expressed e.g. alcohol oxidase 1 (Aox1). A suite of molecular biology tools including CRISPR/Cas9 methods are available for *P. pastoris* [18,22–24] that allow entire metabolic pathways to be integrated without a need for selection markers. We report *de novo* engineering of a CBB cycle into *P. pastoris* and analysis of the capability of engineered strains to grow using CO_2 as a sole carbon source.

Results

The XuMP cycle is used as a template for the CBB cycle

The xylulose-monophosphate (XuMP) cycle for methanol assimilation of *P. pastoris* is entirely localized in peroxisomes and shares the same overall topology and many enzymes with the CBB cycle [21]. In both pathways a C1 molecule is transferred to a sugar phosphate generating a C-C bond. Methanol is oxidized by alcohol oxidase (Aox1 and Aox2 in *P. pastoris*) to formaldehyde. Subsequently, formaldehyde reacts with xylulose-5-phosphate (Xu5P) to dihydroxyacetone (DHA) and glyceraldehyde-3-phosphate (GAP) catalyzed by dihydroxyacetone synthase (Das1 and Das2 in *P. pastoris*). Xu5P is regenerated by pentose phosphate pathway (PPP) enzyme isoforms specifically enriched in peroxisomes upon cultivation on methanol [21]. In total, three mole of methanol are converted to one mole of GAP, which can be used for biomass and energy generation. Likewise, in autotrophic organisms, CO_2 is added to ribulose-1,5-bisphosphate (RuBP) to form 3-phosphoglycerate (3-PGA) in a carboxylation reaction catalyzed by RuBisCO. 3-PGA is then phosphorylated and reduced to GAP. The topology of the XuMP cycle allows its conversion into a CBB cycle by addition of six enzymatic steps (i.e. all biochemical reactions of the CBB cycle present). The overall engineering strategy is outlined in Fig. 1 (with all details in Supplementary Fig. 1). We considered a peroxisomal compartmentalization of the heterologous CBB cycle as beneficial, comparable to native C1 fixation in plant chloroplasts and in the XuMP cycle. To achieve this localization, a C-terminal peroxisomal targeting signal (PTS1) [25] was fused to each of the six heterologously expressed CBB enzymes. In a second control setup, all six heterologous genes were expressed in the cytosol without a PTS1 signal. The CBB cycle is split into 3 phases: carboxylation, reduction and regeneration [26]. For the carboxylation reaction a RuBisCO form II gene from *Thiobacillus denitrificans* (cbbM) was integrated into the genome, together with two molecular chaperone genes from *E. coli* (*groEL* and *groES*) to assist RuBisCO folding [9]. The reduction phase to yield the triose phosphates, glyceraldehyde phosphate (GAP) and dihydroxyacetone phosphate (DHAP) was achieved by integration of the genes encoding for phosphoglycerate kinase

(*PGK1*), glyceraldehyde-3-phosphate dehydrogenase (*TDH3*) from *Ogataea polymorpha* and triosephosphate isomerase (*TPI1*) from *Ogataea parapolymorpha*. The regeneration phase was completed by the addition of *TKL1* (encoding transketolase) from *O. parapolymorpha*, needed to substitute for the peroxisomal transketolase activity of Das [27], and *PRK* (encoding phosphoribulokinase) from *Spinachia oleracea* to the five enzymes from the XuMP cycle (fructose-1,6-bisphosphatase (Fbp1), fructose-bisphosphate aldolase (Fba1-2), sedoheptulose-1,7-bisphosphatase (Shb17), ribose-5-phosphate ketolisomerase (Rki1-2), ribulose-5-phosphate 3-epimerase (Rpe1-2)) already present in *P. pastoris*. Overexpression of the *Ogataea* genes was chosen over the native genes to prevent genomic rearrangements by introduction of repetitive sequences. For each molecule of CO_2 fixed by the heterologous CBB cycle, 3 molecules of adenosine triphosphate (ATP) and 2 molecules of NADH are required. The energy is provided by methanol oxidation. The oxidation of one mole methanol yields two mole NADH which are used as reduction equivalents and for ATP production. Electrons and ATP are probably transported across the peroxisomal transport by the malate oxaloacetate shuttle and the peroxisomal adenine nucleotide transporter Ant1, respectively [28]. In order to separate the foreign fixation machinery of CO_2 from energy generation, we blocked the first steps of the methanol assimilatory pathway by deleting the genes encoding for Das1 and Das2. Furthermore, *AOX1* was deleted to reduce the formation rate of formaldehyde, which can still be formed by *AOX2* (encoding alcohol oxidase 2). These strains (listed in Supplementary Table 1) were generated in a three-step transformation procedure by a CRISPR/Cas9 mediated workflow for genome engineering in *P. pastoris* (see Fig. 2).

A peroxisomal CBB cycle enables faster growth on CO_2

Strains expressing a cytosolic (CBBc+RuBisCO) and a peroxisomal (CBBp+RuBisCO) version of the CBB cycle were obtained. In order to test the ability of these strains to grow on CO_2 as a carbon source, shake flask cultures were carried out in the presence of 5 % (v/v) CO_2 (Fig. 3). A strain harboring the same genetic setup as the CBBp+RuBisCO strain but lacking RuBisCO (CBBpΔRuBisCO) served as a negative control. This strain showed no growth over the entire cultivation phase, whereas the CBBc+RubisCO and CBBp+RuBisCO strains were able to grow (Fig. 3). Furthermore, it was observed that the strains with a peroxisomal CBB pathway (CBBp+RuBisCO) grew significantly faster increasing from 0.09 ± 0.003 to 0.89 ± 0.067 g L^{-1} CDW in 216 hours compared to the strains with a cytosolic CBB cycle, which grew from 0.09 ± 0.000 to 0.44 ± 0.060 g L^{-1} CDW within the same time frame. Further, re-integration of *AOX1* into a CBBp+RuBisCO strain did not lead to faster growth (See Supplementary Fig. 2).

The superior growth behavior of the CBBp+RuBisCO strain confirmed that it is beneficial to compartmentalize the CBB cycle by targeting it to the peroxisome and thereby replacing the native XuMP cycle. As expected the negative control (CBBpΔRuBisCO) did not grow on media containing methanol and CO_2. To further confirm that the deletion of *DAS1* and *DAS2* prevents the methanol assimilation capability of *P. pastoris*, we cultivated two controls strains, one deleted in both DAS genes and lacking PRK and RuBisCO, and the other with a peroxisomal copy of *TKL1* (Supplementary Fig. 3). Both controls did not grow on methanol, confirming that the deletion of *DAS1/DAS2* is sufficient to block the assimilation of

methanol into biomass, and that heterologous Tkl1, while sharing sequence similarity, cannot substitute for Das1/Das2 [29].

Strains expressing a CBB cycle require CO_2 for growth

In the following experiments the CO_2 incorporation characteristics of the strains expressing a peroxisomal CBB cycle (CBBp+RuBisCO) together with the negative control strain having a disrupted CBB cycle (CBBpΔRuBisCO) were tested in controlled bioreactor cultivations. After accumulation of biomass in a batch phase on glycerol, CBBp+RuBisCO and CBBpΔRuBisCO cells were fed with CO_2 and/or methanol. Thereby, two different CO_2 feeding regimes were used: in regime A, a set of both strains were first fed with methanol only and 0% (v/v) CO_2 in the inlet gas and after 135 hours the CO_2 was increased to 5 % (v/v). In regime B, a set of both strains was fed first with methanol and 5% CO_2 and switched to 0% CO_2 after 135 hours (Fig. 4). The bioreactors were operated at a high stirring- and gasflow- rate (1000 rpm and 35 sL h^{-1} respectively) to blow out CO_2 formed by methanol oxidation. With these experimental settings, the CO_2 concentration measured at the off-gas analyzer was below the detection limit of 0.1 % (v/v) also in cultures oxidizing methanol. Thus, a defined CO_2 supply of the cells was controlled by changing the CO_2 content of the inlet gas flow.

The engineered CBBp+RuBisCO strains grew with a rate of 0.033±0.005 g L^{-1} h^{-1} CDW during feeding with 5% CO_2 while it did not grow without CO_2 in the air stream. This growth – non growth behavior was the same in both feed regimes A and B, irrespective whether the CO_2 supply phase preceded or followed the CO_2-free phase. The control strain CBBpΔRuBisCO however did not grow in any of these conditions. These results demonstrate that the CBBp+RuBisCO strains take up and grow on CO_2 as a carbon source and thus have a functional CO_2 assimilation pathway. As energy source methanol can be oxidized by all strains, but with lower rates in the CBBpΔRuBisCO strain (Supplementary Fig. 4). The CBBp+RuBisCO strains only grew in the presence of both methanol and CO_2, and can restore growth on CO_2 after a prolonged period on methanol only (Regime A in Fig. 4.). Taken together these data demonstrate that the growth of the engineered strains expressing a heterologous CBB cycle in the peroxisomes of *P. pastoris* is dependent on the supply of CO_2 as carbon source.

Incorporation of CO_2 is verified by inverse ^{13}C labeling

To obtain further evidence that CO_2 is taken up and incorporated, we measured the labeling of intracellular metabolites and biomass with ^{13}C and ^{12}C by Gas Chromatography – Time of Flight Mass Spectrometry (GC-TOFMS), Liquid Chromatography (LC)-MS/MS and Elemental Analysis - Isotope Ratio Mass Spectrometry (EA-IRMS). The biomass of the engineered cells (CBBp+RuBisCO and CBBpΔRuBisCO) was first fully labeled with ^{13}C carbon by batch cultivation on minimal medium containing fully labeled ^{13}C glycerol as the sole carbon source (Fig. 5 a). After the batch phase, the ^{13}C enrichment reached 95 ± 0.5 % for the cultivation with CBBpΔRuBisCO and 97 ± 0.3 % for the CBBp+RuBisCO (Fig. 5 c). 100% ^{13}C labeling cannot be reached, due to ^{12}C carried over from the inoculum and the trace ^{12}C contamination of the utilized ^{13}C glycerol (^{13}C enrichment >99 %). After the glycerol was consumed, the feeding regime was shifted to gaseous CO_2 with a natural

isotopologue distribution (98.9 ^{12}C, 1.1 % ^{13}C) and methanol either with a natural isotopologue distribution or with ^{13}C enrichment (>99 % ^{13}C), so that the decrease of ^{13}C equals an inverse labeling with ^{12}C.

As expected, CBBp+RuBisCO strains grew under the conditions with a mean biomass formation rate of 0.040 ± 0.003 g L^{-1} h^{-1} CDW, whereas CBBpΔRuBisCO strains showed no growth but still consumed methanol. Growth of the CBBp+RuBisCO strain was accompanied with a reduction of the ^{13}C content in the biomass (Fig. 5 d). After 158 hours the ^{13}C labeling content of the CBBp+RuBisCO fed with $^{12}CO_2$ and ^{13}C methanol reached 52 ± 0.3% or 48 ± 0.2% in a technical replicate. These data matched with the predicted values based on the measurement of accumulated biomass after approximately one cell doubling. The non-growing CBBpΔRuBisCO showed no change in the ^{13}C labeling content in the final biomass over the entire cultivation period, while there was a slight reduction of the fractional label on the metabolite level probably owing to ^{12}C incorporation by carboxylation reactions. This is supported by the high inverse labeling of aspartate which is close to pyruvate carboxylase in metabolism (Supplementary Table 2). Also on $^{12}CO_2$ and ^{12}C methanol the ^{13}C labeling content of the CBBp+RuBisCO decreased according to predicted values based on the biomass measurements (all values are shown in Supplementary Table 3)). Isotopic analysis showed almost complete removal of ^{13}C carbon from the metabolite pool around the CBB cycle (see Fig. 5 e and f, Supplementary Table 4) and for free amino acids (Supplementary Table 2) already after one cell doubling.

These data show that all carbon needed for the accumulated biomass is coming from CO_2. The carbon from methanol is oxidized to CO_2, but due to the high gas-flow rate, it is efficiently removed from the bioreactor. The negative control strain (CBBpΔRuBisCO), which can oxidize methanol but cannot incorporate CO_2 into the biomass, showed consequently no significant change in the overall ^{13}C labeling pattern. Additionally, the quick turnover of ^{13}C in the key metabolites shows functionality of the integrated CBB cycle. Taken together these data confirm that a *P. pastoris* strain containing a functional CBB cycle utilizes CO_2 as a sole carbon source and methanol as energy source. *Per definitionem* the lifestyle of this strain can be described as chemoorganoautotrophic.

Continuous growth on CO_2 as sole carbon source is possible

In order to show that the chemoorganoautotrophic lifestyle of the engineered CBBp +RuBisCO can be maintained for multiple cell doublings, the strain was inoculated with a low cell density of 0.2 g L^{-1} CDW in a bioreactor and cultivated for 470 hours with a constant supply of methanol and CO_2.

The growth profile is shown in Fig. 6. After induction, the cells quickly started to grow with an initial growth rate of $\mu = 0.020 \pm 0.0014$ h^{-1} in the first 20 h after inoculation. In the following the growth was reduced, but a continuous exponential growth phase between 100 up to 235 hours with a μ of 0.008 ± 0.0001 h^{-1} and a doubling time of 92 ± 1.7 h was observed reproducibly. A repetition of this experiment using a ^{13}C labelled inoculum grew with similar growth rates also after re-inoculation into fresh medium (Supplementary Fig. 5). The growth rate achieved in the prolonged exponential phase is determined as the μ_{max} of the culture under the given batch conditions. Over the entire observed 470 h of cultivation

the cells performed a mean of 4.3 cell doublings. With a ^{13}C labelled biomass as inoculum (^{13}C content above 99 %), continuous growth on CO_2 as carbon source leads to almost complete removal of cellular ^{13}C carbon (^{13}C content below 2 %) after 6.5 generations as can be seen in Supplementary Fig. 5. Furthermore, multiple re-inoculation in shake flasks demonstrated continuous growth on CO_2 over 30 generations. These data show that the rationally engineered CBBp+RuBisCO strain can grow with a chemoorganoautotrophic lifestyle without limitation. We further performed an adaptive laboratory evolution (ALE) approach using serial batch cultivations under excess of CO_2 and methanol, which led to a superior phenotype under chemoorganoautotrophic conditions. The growth rate under the relevant conditions and during the respective timeframe was increased from 0.008 to 0.018 h^{-1} for the best performing clones, which is in the range of slower growing autotrophic organisms (See Fig.6, Supplementary Fig. 6, Supplementary Table 5).

Discussion

We reengineered yeast metabolism to enable a synthetic chemoorganoautotrophic lifestyle by introducing a heterologous CBB cycle into the methylotrophic yeast *P. pastoris.* This enables the cells to utilize CO_2 as their sole carbon source, and to grow as an autotrophic organism. The heterologous CBB cycle is localized to peroxisomes by targeting the enzymes via PTS1 signals. The peroxisomal pathway enables superior growth on CO_2 compared to a cytosolic version, highlighting the importance of cellular compartmentalization, which likely facilitated functional CO_2 fixation in our set-up compared to other approaches [13]. The spatial separation can have a beneficial impact on the operation of other pathways like the pentose phosphate pathway, enabling different concentrations of key intermediates like ribose-5-phosphate or sedoheptulose-1,7-bisphosphate. In addition, a channeling effect by the close proximity of CBB cycle enzymes in the dense peroxisomal environment may improve flux rates [30,31]. *P. pastoris* may serve as a promising chassis cell to also test synthetic CO_2 fixation pathways *in vivo* such as the malonyl-CoA-oxaloacetate-glyoxylate (MOG) pathway [32] or the crotonyl–coenzyme A (CoA)/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle [33]. One needs to consider however whether the intermediate metabolites of such pathways are available in the peroxisome.

Adaptive evolution was employed to increase the maximum specific growth rate on CO_2 more than twofold, from 0.008 h^{-1} to 0.018 h^{-1}. After evolution for 30 generations single isolates were analyzed and three independent isolates (clone 5, 6 and 9) were sequenced showing in total 6, 5 or 3 mutations compared to the CBBp+RuBisCO parental strain (Supplementary Table 6). Notably, two isolates have a mutation in the heterologous *PRK* gene leading to an amino acid exchange (Ala2Gly or Thr5Ala), while the third isolate carries a mutation in *NMA1* (Thr358Ile), encoding for nicotinic acid mononucleotide adenylyltransferase. In addition clone 5 shows three mutations in rRNA loci with a frequency of 70 %. Both altered coding sequences give rise to further engineering targets. PRK catalyzes the ATP dependent phosphorylation of D-ribulose-5-phosphate and thus a key step of the CBB cycle. ATP and NADH balancing is crucial for a functional integration of a heterologous CBB cycle into the central carbon metabolism [7,8,12,34]. Nma1 is involved in maintaining high NAD levels upon growth on methanol [21,35] and mutations leading to activity changes could alter both NADH and ATP levels. In addition to its impact on ATP

consumption, an altered PRK activity would affect the availability of D-ribulose-5-phosphate for various biosynthetic routes [36].

P. pastoris has been demonstrated to be a suitable chassis for the production of enzymes and chemicals [18,37], making pathway engineering toward varied products straightforward. Implementation of CO_2 as a carbon source in industrial processes raises several considerations. Any CO_2 assimilation pathway requires energy and reducing power, provided by light or a chemical electron donor such as hydrogen, formate or methanol. While light and hydrogen require specific bioreactor equipment, a water-soluble, potentially renewable chemical energy source like methanol or formate can be directly applied in standard setups and in large-scale bioreactors[38]. More oxidized products, like organic acids, require less reducing power. Combined with a reduced co-substrate, a net CO_2 fixation can be obtained when producing organic acids such as succinic, lactic, acetic or itaconic acid from methanol and CO_2. Theoretically, up to 62 kg CO_2 are fixed when producing 100 kg succinic acid with a methanol co-feed [1]. For cost-efficient production of bulk chemicals, volumetric productivity must achieve a range of at least 2 g L^{-1} h^{-1} [39]. *P. pastoris* is well established to produce at biomass concentrations of 100 g L^{-1} or more [40]. The current best specific CO_2 assimilation rate reported here of 0.03 g g^{-1} h^{-1} (calculated on specific growth rate) would correspond to a potential volumetric productivity of about 2 g L^{-1} h^{-1} succinate, as an example. This indicates that production rates would already be close to existing processes.

We have demonstrated that conversion of a heterotroph to a synthetic autotroph is possible. Rational re-engineering of metabolism led to growth on CO_2 as the only carbon source. Adaptive evolution more than doubled growth rates, likely through adjustments in the CO_2 assimilation pathway and cofactor availability. The resulting yeast strains may form the basis of a system for producing bulk chemicals and enzymes based on CO_2 and may support mitigation of atmospheric CO_2.

Online methods

Plasmid construction

All expression cassettes and sgRNA / hCas9 plasmids were constructed by Golden Gate cloning [23,41,42]. Three native genes of *P. pastoris* (*AOX1*, *DAS1* and *DAS2*) were replaced by the coding sequences of the genes listed in Supplementary Table 7. A CRISPR/Cas9 mediated homologous recombination (HR) directed system was used to construct the strains [24]. Flanking regions needed for replacing the CDSs of *AOX1*, *DAS1* and *DAS2* were amplified from CBS7435 *wt* genomic DNA (gDNA) (Promega, Wizard® Genomic DNA Purification Kit) by PCR (NEB, Q5® High-Fidelity DNA Polymerase). The other promoters (P_{ALD4}, P_{FDH1}, P_{SHB17}, P_{PDC1} and P_{RPP1B}) and terminators (T_{IDP1}, T_{RPB1b}, T_{RPS2b}, T_{RPS3t} and $T_{RPBS17Bt}$) were prepared accordingly and derived from genomic DNA of strain CBS7435 (Supplementary Table 8). CDSs of *TDH3* and *PGK1* were amplified from gDNA of *Ogataea polymorpha* (CBS 4732) and *TKL1* and *TPI1* from gDNA of *Ogataea parapolymorpha* (CBS 11895) according to the procedure described above and a C-terminal peroxisome targeting signal 1 (PTS1) [25]encoding for SKL was added by amplification with primers carrying the additional nucleotides. The sequences encoding PRK (*Spinacia*

oleracea), RuBisCO - cbbM (*Thiobacillus denitrificans*) and the chaperones GroEL and GroES (*E. coli*) were codon optimized (GeneArt, Regensburg, Germany) (sequences are shown in the Supplementary file). Final plasmids were digested with *BpiI* (*BbsI*), (ThermoFischer Scientific, USA) to obtain linear donor DNA templates used for replacement of the three native loci. Specific single guide RNAs (sgRNAs) were designed [43] and cloned into CRIS*Pi* plasmids [24]. The genomic recognition sites for targeting the different loci with CRISPR/Cas9 were CTAGGATATCAAACTCTTCG for *AOX1*, TGGAGAATAATCGAACAAAA for *DAS1* and CGACAAACTATAAGTAGATT for *DAS2.* The final sgRNA / Cas9 plasmids were checked by enzymatic digestion and by Sanger sequencing.

Strain Construction

For cell engineering, *Komagataella phaffii (Pichia pastoris)* CBS7435 [44,45] was used as a host. Transformations were carried out as previously described [16,24,46]. Primers listed in Supplementary Table 9 were used for detection of the correct replacement events of *AOX1*, *DAS1* and *DAS2* loci and correct integration was further verified by Sanger sequencing. All strains used in this study are listed in Supplementary Table 1.

Cultivations in shake flask

Cultures were inoculated with a starting OD_{600} of 0.5 in YNB medium supplemented with 10 g L^{-1} $(NH_4)_2SO_4$. The flasks (total volume 100 mL with a working volume between 10 and 20 mL) were incubated in a CO_2 controlled incubator (5% CO_2) shaking at 180 rpm at 30°C. The methanol concentration was adjusted up to 0.5 % (v/v) after the batch phase for induction and from there on maintained at 1% (v/v). Cell growth (OD_{600} and Cell Dry Weight (CDW) measurements) and metabolite profiles (HPLC analysis) were monitored [47].

Adaptive Laboratory Evolution

The cells were evolved for 27 to 29 generations in serial batch cultivations using minimal YNB medium as described above. Each biological clone was evolved in three technical replicates. During the serial passages (2 to 3 generations each), no significant growth increase was seen from passage to passage. Because of this, the entire pool from an evolved population was regenerated on YPG medium and then characterized in comparison to the parental strains as shown in Supplementary Figure 6. If the evolved pools showed superior growth compared to the parental strains (Supplementary Figure 6 a), single colonies were isolated from this pool and further characterized (Supplementary Figure 6 b).

Bioreactor Cultivations

Single colonies were picked and used for inoculation of 100 mL pre-cultures in YPD medium (28°C o/n and 180 rpm). Bioreactors were inoculated with a starting OD_{600} of 1.0 or 0.19 g L^{-1} CDW. A defined medium consisting of citrate monohydrate (2 g L^{-1}), glycerol (16 g L^{-1}), $MgSO_4$ $7H_2O$ (0.5 g L^{-1}), $(NH_4)_2HPO_4$ (12.6 g L^{-1}) KCl (0.9 g L^{-1}) and biotin (0.0004 g L^{-1}) was used. For the labeling experiment a medium containing H_3PO_4 (19.3 g L^{-1}), $CaCl_2$ $2H_2O$ (0.025 g L^{-1}), $MgSO_4$ $7H_2O$ (2.5 g L^{-1}), KOH (2.0 g L-1), NaCl (0.22 g L^{-1}), EDTA (0.6 g L^{-1}), biotin (0.0004 g L^{-1}) and fully labelled ^{13}C glycerol (8.0 g L^{-1}) was

used. Both media were supplemented with trace elements ($FeSO_4$ * 7 H_2O (0.3 g L^{-1}), $ZnCl_2$ (0.1 g L^{-1}), $CuSO_4$ * 5 H_2O (0.03 g L^{-1}), $MnSO_4$ * H_2O (0.015 g L^{-1}), $CoCl_2$ * 6 H_2O (0.004 g L^{-1}), Na_2MoO_4 * 2 H_2O (0.92 mg L^{-1}), NaI (0.036 mg L^{-1}), H_3BO_3 (0.092 mg L^{-1})). For the continuous growth fermentation YNB medium supplemented with 10 g L^{-1} $(NH_4)_2SO_4$ was used.

The bioreactor cultivations were carried out in 1.4 L DASGIP reactors (Eppendorf, Germany). Cultivation temperature was controlled at 28 °C, pH at 5.0 by addition of 12.5 % ammonium hydroxide or with 5 M NaOH and 2 M HCl in case of YNB medium. The dissolved oxygen concentration was maintained above 20 % saturation. OD_{600} and cell dry weight (CDW) measurements as well as HPLC analysis to routinely measure glycerol, glucose, methanol and citrate concentrations [47] were carried out at every sample point. Induction was done by the addition of 0.5% methanol (v/v). The CO_2 in the inlet gas was set to 1% during induction phase. After induction phase, the methanol concentration was adjusted to 1 - 1.5% (v/v) by bolus feeding and the CO_2 concentration was set to 5 % throughout the cultivations. For the labeling experiment shown in Figure 5, the gas flow rate was set to 35 sL h^{-1} and the stirrer speed to 1000 rpm until the end of the fermentation. In a second experiment (see Supplementary Figure 5), the inoculum was labeled by cultivation in YNB medium supplemented with fully labelled ^{13}C glycerol (10.0 g L^{-1}). Long term bioreactor cultivation was then carried out as described above.

Elemental Analysis - Isotope Ratio Mass Spectrometry (EA-IRMS)

For the measurement of the biomass ^{13}C content, the samples were taken during fermentation and processed at 4 °C. Volumes of cell suspension corresponding to approximately 0.5 mg of dried biomass was firstly washed with 0.1 M HCl and then twice with RO-H_2O. Until analysis, the biomass samples were stored at -20 °C. The biomass material was homogenized and weighed in tin capsules. The $^{13}C/^{12}C$ ratio was measured with an Isotope Ratio Mass Spectrometer coupled to an Elemental Analyzer (EA-IRMS). EA-IRMS measurements were carried out by Imprint Analytics, Neutal, Austria.

Derivatization and GC-TOFMS analysis of amino acids and sugar phosphates

Sampling and quenching of biomass samples for metabolome analysis were prepared as recently described [48].

Sugar phosphate isotopologue distribution analysis for the calculation of the ^{13}C labeling pattern of the metabolite backbones was carried out according to the method of Mairinger et al. [49], which is based on the gas chromatographic separation of ethoximated and trimethylsilylated metabolites and chemical ionization – time of flight mass spectrometry (GC-CI-TOFMS). Fractional labeling was calculated, with n = number of carbons, i = isotopologue and m_i = relative abundance of isotopologue, according to equation 1:

$$fractional\ labeling = \frac{\Sigma_{i=0}^{n} m_i * i}{n * \Sigma_{i=0}^{n} m_i} \quad \text{Eq 1}$$

For amino acid isotopologue distribution analysis 50 µL aliquots of the extracts were derivatized using just in time online derivatization on a sample preparation robot (MPS2, Gerstel) with N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide with 1% tert-butyldimethylchlorosilane at 85°C for 50 min. 1 µL aliquots were injected into an Agilent split/splitless injector with an Ultra Inert straight split liner in split mode (250°C, split 20:1, septum purge flow 3 mL min^{-1}). The transfer line to the MS was kept at 280 °C. The following temperature program was used: 150 °C (hold for 1 min), 4 °C min^{-1} to 236 °C, 30 °C min^{-1} to 310 °C (hold for 4 min). The electron ionization (EI) source of the instruments was operated with an electron energy of 70 eV and an emission current of 15 µA. Amino acid fragments were evaluated as described in Zamboni et al [50], extracting the theoretical mass from the SureMass-converted data within Agilent MassHunter (Version 10.0) Quantitative Analysis (extraction window is selected by the SureMass algorithm). Resulting peak areas were corrected for the H, N, O, Si and S isotopes of the derivatized molecule and the ^{13}C isotopes stemming from derivatization (not backbone C atoms) using the ICT correction toolbox [51].

In detail, gas chromatographic separation of amino acids and sugar phosphates was carried out with an Agilent 7890B gas chromatograph coupled with an Agilent Technologies 7200B Q-TOF mass spectrometer (GC-TOFMS, Agilent Technologies Inc.). The GC-TOFMS system was equipped with (1) a non-polar deactivated pre-column (3 m x 0.25 µm i.d., Phenomenex) connected to the analytical column *via* a backflush unit (purged ultimate union), (2) an Optima® 5 MS Accent (5% phenyl 95% dimethylpolysiloxane, 60 m x 250 µm x 0.25 µm) and (3) a non-polar deactivated postcolumn (3 m x 0.18 µm i.d., Phenomenex), which was connected with the analytical column *via* a second purged ultimate union T-piece (serves as restrictor column to the MS). Helium 5.0 was used as carrier gas with a constant flow of 1.1 mL min^{-1} for the pre-column, 1.2 mL min^{-1} for the analytical column and 1.4 mL min^{-1} for the post column.

Analysis for sugar phosphate isotopologue distribution analysis for the calculation of ^{13}C labeling patterns was carried out as described recently [49] with minor modifications regarding GC-TOFMS setup and data evaluation. Evaluated mass/charge ratios are summarized in Supplementary Table 10.

LC-MS/MS analysis of ribulose 1,5-bisphosphate

For measurements of ribulose 1,5-bisphosphate isotopologues, an HPIC system (Thermo Fisher Scientific IC5000+) with mass spectrometric detection (Thermo Fisher TSQ Vantage Triple Quadrupole Mass Spectrometer) was used. For each sample, 10 µL was injected onto an IonPac AG11-HC-4µm guard column (2×50 mm) followed by an IonPac AS11-HC-4µm analytical column (2×250 mm) for anion exchange chromatography separation. Effluent was suppressed using an AERS 500e 2mm electrolytic membrane suppressor with a current of 30 mA. A multistep electrolytically generated gradient program was used with an initial concentration of 10 mmol L^{-1} KOH ramped to 60 mmol L^{-1} to allow isocratic elution and a total run time of 13 min. Analytical and suppressor regenerant flow rates were 0.2 mL min^{-1}. Additional acetonitrile (flow rate of 0.1 mL min^{-1}) was mixed with effluent before entering the heated electrospray (HESI) ion source to improve ionization. Other settings were

capillary and vaporizer temperatures (350 °C), sprayer voltage (3000 V), aux gas pressure (15 arbitrary units), sheath gas pressure (40 arbitrary units) and collision gas pressure (1.5 mTorr). Multiple reaction monitoring (MRM) transitions for the $[M\text{-}H]^-$ isotopologues (*m*/*z* 309, 310, 311, 312, 313, & 314) with each yielding the fragment ions *m*/*z* 79 (50 V CE) & 97 (30 V CE) were monitored in the negative ionization mode. The ratio of the areas from the resulting MRM peaks were used to assess the relative abundances of the intact ribulose 1,5-bisphosphate isotopologues in all samples.

Genome sequencing and analysis

The yeast strains were grown in 50 mL YPD-medium over night at 30 °C under vigorous shaking, centrifuged and washed. Cell pellets were resuspended in 3 mL of SCED buffer, pH 7.5 and incubated at 37 °C for 1h after addition of 100 µL of lyticase (5 U/µL) and 20 µL of RNase A solution (100 mg/mL). 2 mL of 1% SDS were added, mixed gently and placed on ice for 5 min. After addition of 1.5 mL of 5M potassium acetate, pH 8.9, the solution was vortexed for 10 s, centrifuged at 10,000 x g for 10 min at 4°C, and the supernatant containing the genomic DNA was saved.

QIAGEN Genomic-tips (100/G) were equilibrated with 4 mL of Buffer QBT. Supernatants from the centrifugation step were vortexed for 10 s and then applied to the equilibrated genomic tips, washed twice with 7.5 mL of Buffer QC, and genomic DNA was eluted with 5 mL of prewarmed Buffer QF (50 °C). Genomic DNA was then precipitated with 3.5 mL isopropanol and collected by centrifuging at 10,000 x g for 20 min at 4°C. The pellet was washed with ethanol, air-dried and resuspended in 60 µL TE-Buffer. Genomic DNA was dissolved by shaking overnight at 25°C. Purity and concentration of the DNA was determined by agarose gel electrophoresis, and UV absorption at 260 and 280 nm.

Library preparation and sequencing was done at the Vienna Biocenter Core Facilty for Next Generation Sequencing (Vienna, Austria) on an Illumina MiSeq (PE300) sequencing platform. Sequence analysis was carried out using CLC Genomic Workbench 12.0 (QIAGEN, Aarhus). Sequences were mapped against the *K. phaffii* strain CBS7435 [45] and variants were detected using the fixed ploidy variant detection tool setting ploidy to 1. Detected variants of the evolved strains were filtered against the detected variants of the unevolved parental strains with a threshold frequency setting of 65 %.

Determination of viability, cell size and morphology

Cell morphology was routinely checked using a Zeiss Axio Observer.Z1/7 microscope equipped with a LCI Plan-Neofluar 63x/1.3 water-immersion objective. Cell morphology was analyzed in bright-field mode. Pictures were processed with the Zen 2.3 lite (blue edition) software (Carl Zeiss Microscopy GmbH).

For cell size and viability measurements respective cultures were washed and diluted in PBS (KH_2PO_4 0.24 g L^{-1}, Na_2HPO_4 2 H_2O 1.8 g L^{-1}, KCl 0.2 g L^{-1}, NaCl 8 g L^{-1}) followed by brief sonicatation. For viability measurements, cells were stained in the presence of 2mM propidium iodide (PI) before fluorescence (638 nm laser), forward and sideward scatter were measured at CytoFlex S flow cytometer (Beckman Coulter). Ethanol fixed cells were used as negative control. The size distribution of the cell population and the mean cell size was

Europe PMC Funders Author Manuscripts

determined by calibration using a polystyrene particle size standard kit (SPHERO Flow Cytometry Size Standard Kit, Spherotech). Flow cytometry data was analyzed using CytExpert (version 2.3.0.84).

Statistical Methods

All analyses were done using Microsoft Excel or the CytExpert (version 2.3.0.84, Beckman Coulter) software. Data are presented as individual values with error bars representing *s.d.* or *s.e.m.* as indicated in the figure legends.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work has been supported by the Federal Ministry for Digital and Economic Affairs (bmwd), the Federal Ministry for Transport, Innovation and Technology (bmvit), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol, Government of Lower Austria and ZIT - Technology Agency of the City of Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG, granted to DM and MGS. We also thank the Austrian Science Fund for support of DM and TG (FWF W1224—Doctoral Program on Biomolecular Technology of Proteins—BioToP). The funding agencies had no influence on the conduct of this research. EQ BOKU VIBT GmbH is acknowledged for providing fermentation and mass spectrometry equipment. We kindly acknowledge the technical support from Sarah Heinzl and Philipp Tondl. We also kindly thank Hannes Rußmayer for help during labeling experiments.

References

1. Steiger MG, Mattanovich D, Sauer M. Microbial organic acid production as carbon dioxide sink. FEMS Microbiol Lett. 2017; 364
2. Bassham JA, et al. The Path of Carbon in Photosynthesis. XXI. The Cyclic Regeneration of Carbon Dioxide Acceptor 1. J Am Chem Soc. 1954; 76:1760–1770.
3. Berg IA. Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. Appl Environ Microbiol. 2011; 77:1925–36. [PubMed: 21216907]
4. Claassens NJ, Sousa DZ, dos Santos VAPM, de Vos WM, van der Oost J. Harnessing the power of microbial autotrophy. Nat Rev Microbiol. 2016; 14:692–706. [PubMed: 27665719]
5. Erb TJ, Zarzycki J. A short history of RubisCO: the rise and fall (?) of Nature's predominant CO_2 fixing enzyme. Curr Opin Biotechnol. 2018; 49:100–107. [PubMed: 28843191]
6. Zhuang Z-Y, Li S-Y. Rubisco-based engineered *Escherichia coli* for *in situ* carbon dioxide recycling. Bioresour Technol. 2013; 150:79–88. [PubMed: 24152790]
7. Antonovsky N, et al. Sugar Synthesis from CO_2 in *Escherichia coli*. Cell. 2016; 166:115–125. [PubMed: 27345370]
8. Herz E, et al. The genetic basis for the adaptation of *E. coli* to sugar synthesis from CO_2. Nat Commun. 2017; 8
9. Guadalupe-Medina V, et al. Carbon dioxide fixation by Calvin-Cycle enzymes improves ethanol yield in yeast. Biotechnol Biofuels. 2013; 6:125. [PubMed: 23987569]
10. Li Y-J, et al. Engineered yeast with a CO_2-fixation pathway to improve the bio-ethanol production from xylose-mixed sugars. Sci Rep. 2017; 7
11. Xia P-F, et al. Recycling Carbon Dioxide during Xylose Fermentation by Engineered *Saccharomyces cerevisiae*. ACS Synth Biol. 2017; 6:276–283. [PubMed: 27744692]
12. Papapetridis I, et al. Optimizing anaerobic growth rate and fermentation kinetics in *Saccharomyces cerevisiae* strains expressing Calvin-cycle enzymes for improved ethanol yield. Biotechnol Biofuels. 2018; 11:17. [PubMed: 29416562]

13. Schada von Borzyskowski L, et al. An engineered Calvin-Benson-Bassham cycle for carbon dioxide fixation in *Methylobacterium extorquens* AM1. Metab Eng. 2018; 47:423–433. [PubMed: 29625224]
14. Phaff HJ, Miller MW, Shifrine M. The taxonomy of yeasts isolated from *Drosophila* in the Yosemite region of California. Antonie Van Leeuwenhoek. 1956; 22:145–161. [PubMed: 13340701]
15. Kurtzman CP. Description of *Komagataella phaffii* sp. nov. and the transfer of *Pichia pseudopastoris* to the methylotrophic yeast genus *Komagataella*. Int J Syst Evol Microbiol. 2005; 55:973–976. [PubMed: 15774694]
16. Gasser B, et al. Pichia pastoris: protein production host and model organism for biomedical research. Future Microbiol. 2013; 8:191–208. [PubMed: 23374125]
17. Liu L, et al. How to achieve high-level expression of microbial enzymes. Bioengineered. 2013; 4:212–223. [PubMed: 23686280]
18. Peña DA, Gasser B, Zanghellini J, Steiger MG, Mattanovich D. Metabolic engineering of *Pichia pastoris*. Metab Eng. 2018; 50:2–15. [PubMed: 29704654]
19. Agrawal G, Shang HH, Xia Z-J, Subramani S. Functional regions of the peroxin Pex19 necessary for peroxisome biogenesis. J Biol Chem. 2017; 292:11547–11560. [PubMed: 28526747]
20. Ma C, Agrawal G, Subramani S. Peroxisome assembly: matrix and membrane protein biogenesis. J Cell Biol. 2011; 193:7–16. [PubMed: 21464226]
21. Rußmayer H, et al. Systems-level organization of yeast methylotrophic lifestyle. BMC Biol. 2015; 13:80. [PubMed: 26400155]
22. Weninger A, Hatzl A-M, Schmid C, Vogl T, Glieder A. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. J Biotechnol. 2016; 235:139–149. [PubMed: 27015975]
23. Prielhofer R, et al. Golden*Pi*CS: a Golden Gate-derived modular cloning system for applied synthetic biology in the yeast *Pichia pastoris*. BMC Syst Biol. 2017; 11:123. [PubMed: 29221460]
24. Gassler T, Heistinger L, Mattanovich D, Gasser B, Prielhofer R. CRISPR/Cas9-Mediated Homology-Directed Genome Editing in *Pichia pastoris*. Methods Mol Biol. 2019; 1923:211–225. [PubMed: 30737742]
25. Lee PC. Peroxisome targeting of lycopene pathway enzymes in *Pichia pastoris*. Methods Mol Biol. 2012; 898:161–9. [PubMed: 22711124]
26. Raines CA, Lloyd JC, Dyer TA. New insights into the structure and function of sedoheptulose-1,7-bisphosphatase; an important but neglected Calvin cycle enzyme. J Exp Bot. 1999; 50:1–8.
27. Waites MJ, Quayle DJR. Dihydroxyacetone Synthase: a Special Transketolase for Formaldehyde Fixation from the Methylotrophic Yeast *Candida boidinii* CBS 5777. Journal of General Microbiology. 1983; 129:935–944.
28. Visser WF, van Roermund CWT, Ijlst L, Waterham HR, Wanders RJA. Metabolite transport across the peroxisomal membrane. Biochem J. 2007; 401:365–75. [PubMed: 17173541]
29. de Koning W, Gleeson MAG, Harder W, Dijkhuizen L. Regulation of methanol metabolism in the yeast *Hansenula polymorpha*. Arch Microbiol. 1987; 147:375–382.
30. Siu K-H, et al. Synthetic scaffolds for pathway enhancement. Curr Opin Biotechnol. 2015; 36:98–106. [PubMed: 26322735]
31. Hammer SK, Avalos JL. Harnessing yeast organelles for metabolic engineering. Nat Chem Biol. 2017; 13:823–832. [PubMed: 28853733]
32. Bar-Even A, Noor E, Lewis NE, Milo R. Design and analysis of synthetic carbon fixation pathways. Proc Natl Acad Sci. 2010; 107:8889–8894. [PubMed: 20410460]
33. Schwander T, Schada von Borzyskowski L, Burgener S, Cortina NS, Erb TJ. A synthetic pathway for the fixation of carbon dioxide *in vitro*. Science. 2016; 354:900–904. [PubMed: 27856910]
34. Barenholz U, et al. Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points. Elife. 2017; 6:e20667. [PubMed: 28169831]
35. Anderson RM, et al. Manipulation of a nuclear NAD+ salvage pathway delays aging without altering steady-state NAD+ levels. J Biol Chem. 2002; 277:18881–90. [PubMed: 11884393]

36. Rouzeau C, et al. Adaptive response of yeast cells to triggered toxicity of phosphoribulokinase. Res Microbiol. 2018; 169:335–342. [PubMed: 29964131]
37. Yang Z, Zhang Z. Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: A review. Biotechnol Adv. 2018; 36:182–195. [PubMed: 29129652]
38. Olah GA. Towards Oil Independence Through Renewable Methanol Chemistry. Angew Chemie Int Ed. 2013; 52:104–107.
39. Werpy T, Petersen G. Value Added Chemicals from Biomass: Volume I - Results of Screening for Potential Candidates from Sugars and Synthesis Gas. US Department of Energy. 2004; doi: 10.2172/15008859
40. Mattanovich, D, Jungo, C, Wenger, J, Dabros, M, Maurer, M. Yeast Suspension CultureIndustrial Scale Suspension Culture of Living Cells. Wiley-Blackwell; 2014. 94–129.
41. Engler C, Gruetzner R, Kandzia R, Marillonnet S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS One. 2009; 4:e5553. [PubMed: 19436741]
42. Sarkari P, Marx H, Blumhoff ML, Mattanovich D, Steiger MG. An efficient tool for metabolic pathway construction and gene integration for *Aspergillus niger*. Bioresour Technol. 2017; 245:1327–1333. [PubMed: 28533066]
43. Gao Y, Zhao Y. Self-processing of ribozyme-flanked RNAs into guide RNAs *in vitro* and *in vivo* for CRISPR-mediated genome editing. J Integr Plant Biol. 2014; 56:343–349. [PubMed: 24373158]
44. Küberl A, et al. High-quality genome sequence of *Pichia pastoris* CBS7435. J Biotechnol. 2011; 154:312–320. [PubMed: 21575661]
45. Valli M, et al. Curation of the genome annotation of *Pichia pastoris* (*Komagataella phaffii*) CBS7435 from gene level to protein function. FEMS Yeast Res. 2016; 16:fow051. [PubMed: 27388471]
46. Cregg, JM, Russell, KA. Pichia Protocols. Vol. 103. Humana Press; 1998. 27–40.
47. Blumhoff ML, Steiger MG, Mattanovich D, Sauer M. Targeting enzymes to the right compartment: Metabolic engineering for itaconic acid production by *Aspergillus niger*. Metab Eng. 2013; 19:26–32. [PubMed: 23727192]
48. Rußmayer H, et al. Metabolomics sampling of *Pichia pastoris* revisited: rapid filtration prevents metabolite loss during quenching. FEMS Yeast Res. 2015; 15
49. Mairinger T, et al. Gas Chromatography-Quadrupole Time-of-Flight Mass Spectrometry-Based Determination of Isotopologue and Tandem Mass Isotopomer Fractions of Primary Metabolites for ^{13}C-Metabolic Flux Analysis. Anal Chem. 2015; 87:11792–11802. [PubMed: 26513365]
50. Zamboni N, Fendt S-M, Rühl M, Sauer U. ^{13}C-based metabolic flux analysis. Nat Protoc. 2009; 4:878–892. [PubMed: 19478804]
51. Jungreuthmayer C, Neubauer S, Mairinger T, Zanghellini J, Hann S. *ICT*: isotope correction toolbox. Bioinformatics. 2015; 32

Europe PMC Funders Author Manuscripts

Editors summary

A yeast species used to produce proteins and chemicals is engineered to grow solely on the greenhouse gas CO_2.

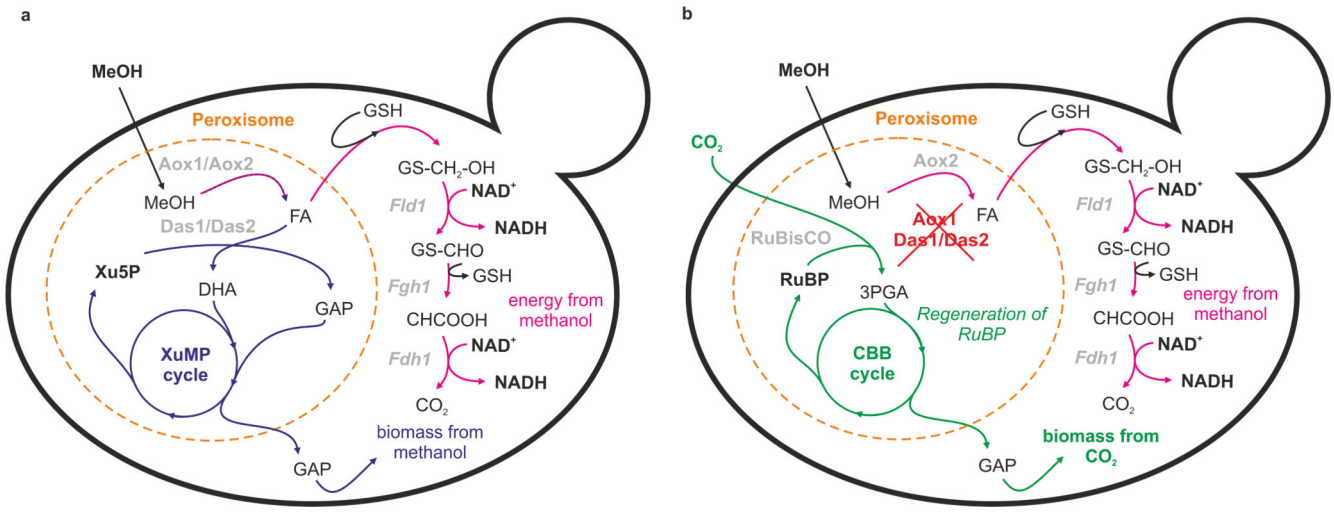

Figure 1. Scheme for engineering chemoorganoautotrophy in *Pichia pastoris*.

(**a**) **Wild type *P. pastoris*** uses methanol (MeOH) for biomass generation in the assimilatory branch of the methanol utilization by fixation of formaldehyde (FA) to xylulose-5-phosphate (Xu5P) which is regenerated in the xylulose monophosphate (XuMP) cycle (purple pathway) and as an energy source in the dissimilatory branch (pink pathway) by oxidation to carbon dioxide (CO_2) under formation of NADH. (**b**) In an **engineered strain** methanol assimilation is blocked by deletion of dihydroxyacetone synthase (*DAS1* and *DAS2*) and a Calvin-Benson-Bassham (CBB) cycle is integrated. Additionally the alcohol oxidase 1 gene (*AOX1*) was deleted. RuBisCO carboxylates ribulose-1,5-bisphosphate (RuBP) which is regenerated in the CBB cycle; abbreviations: alcohol oxidase (Aox), formaldehyde dehydrogenase (Fld1), S-formylglutathione hydrolase (Fgh1), formate dehydrogenase (Fdh1), ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), dihydroxyacetone (DHA), glyceraldehyde-3-phosphate (GAP), 3-phosphoglycerate (3PGA), glutathione (GSH), ribulose-1,5-bisphosphate (RuBP); A detailed depiction of the two pathways including all involved enzymes is given in Supplementary Fig. 1.

Europe PMC Funders Author Manuscripts

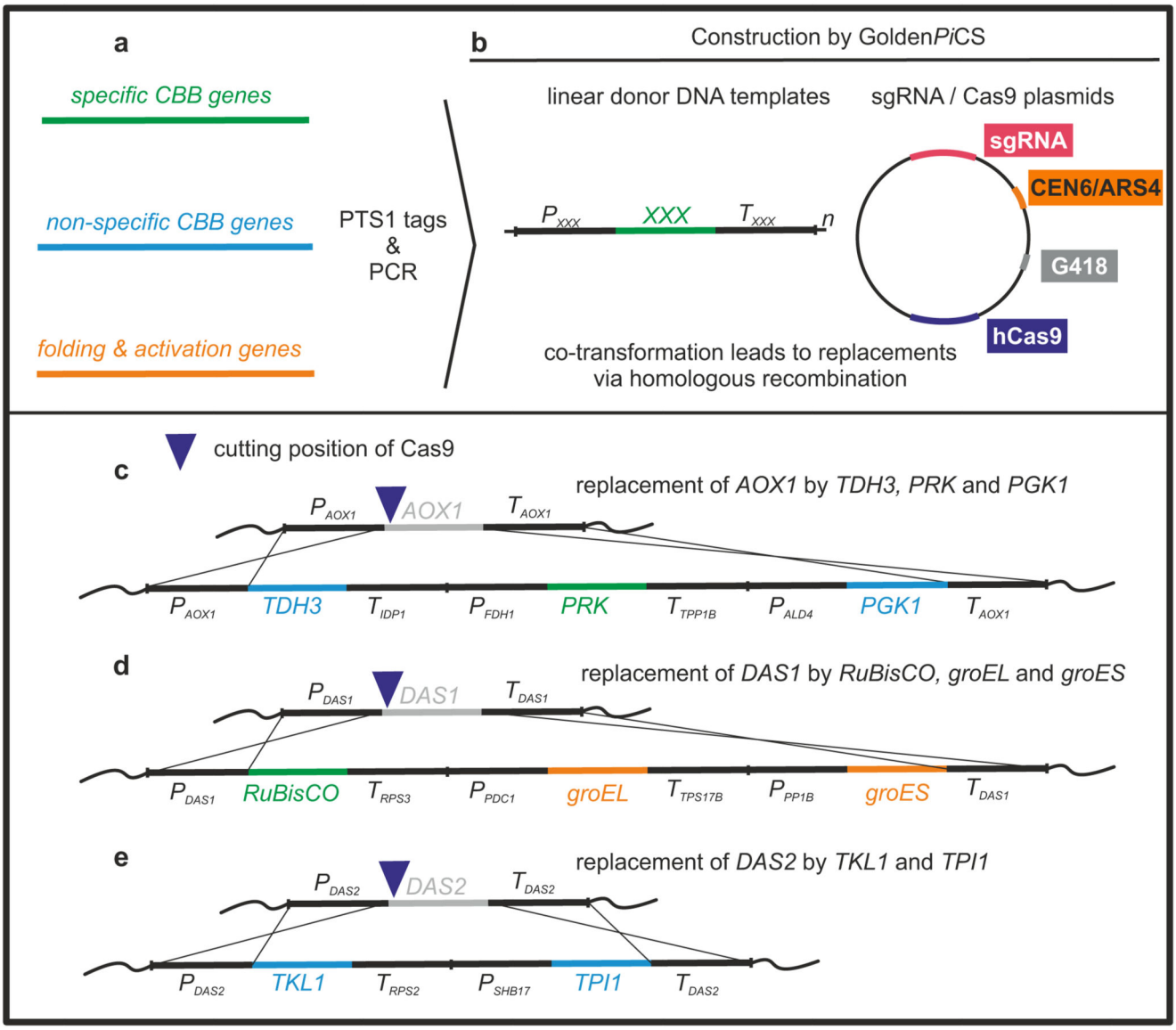

Figure 2. Engineering scheme for generation of *P. pastoris* strains with a heterologous CBB cycle.

(**a-b**) **Plasmid construction**: (**a**) Genes for the CBB cycle including CBB – specific (green – source are autotroph organisms), non-specific genes (blue – from *Ogataea* strains) and chaperones (from *Escherichia coli*) were amplified by PCR (*TDH3*, *PGK1*, *TPI1* and *TKL1*), or synthesized (*RuBisCO*, *PRK*, g*roEL* and g*roES*, codon optimized). The C-terminal PTS1 tag encoding for –SKL was added by PCR if needed. The entire gene names, source organisms and protein identifiers are summarized in Supplementary Table 7. (**b**) Generation of linear donor DNA templates and single guide RNA (sgRNA) / Cas9 plasmids; (**c – e**) **Homologous recombination (HR) mediated replacement of the native sequences** by co – transformation of a linear donor DNA with a sgRNA / Cas9 plasmid; replacement of (**c**) *AOX1* by an expression cassette for *TDH3*, *PRK* and *PGK1*, (**d**) *DAS1* by *RuBisCO*, *groEL* and *groES* and (**e**) *DAS2* by *TKL1* and *TPI1* was done consecutively in three transformation steps and is drawn schematically. The strains lacking one or more

heterologous gene were constructed accordingly with altered donor DNA fragments. All strains constructed by this workflow are listed in Supplementary Table 1.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

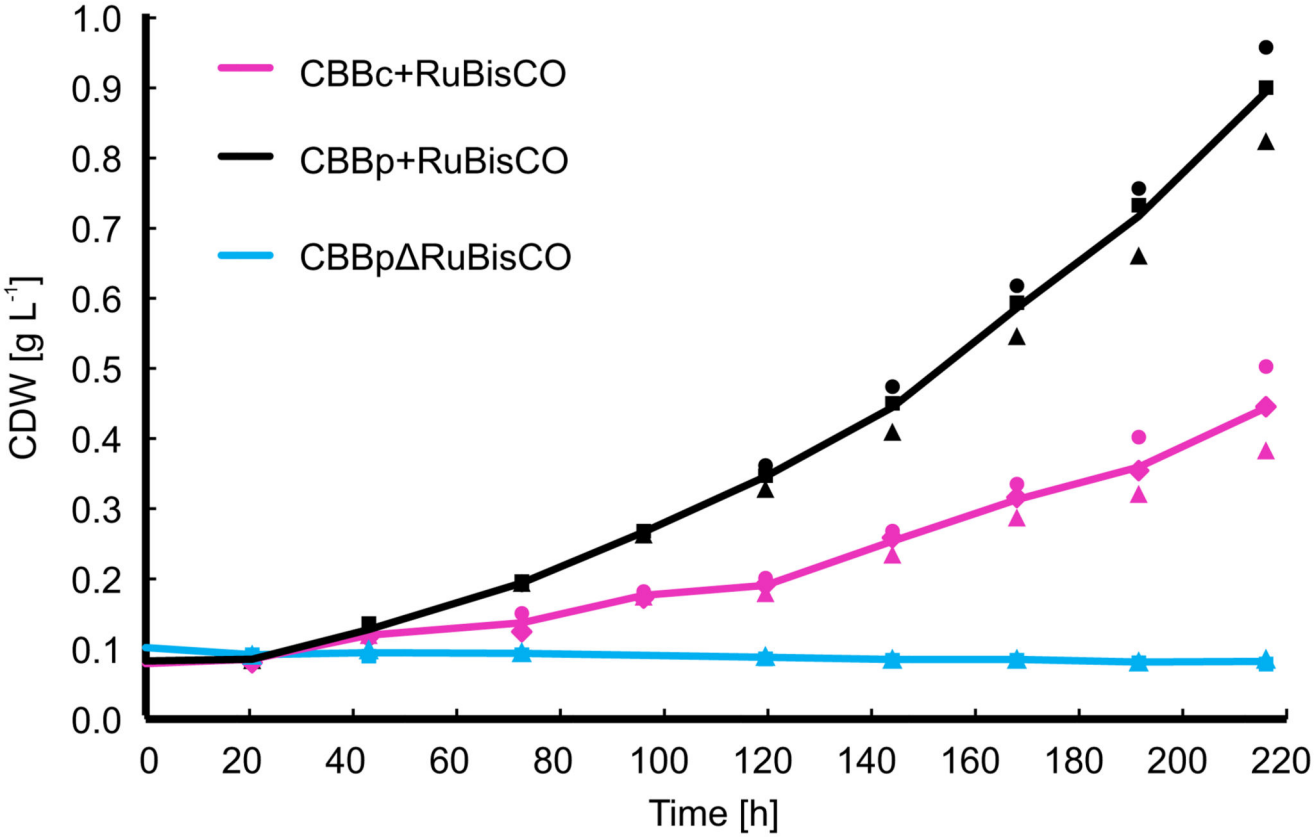

Figure 3. Peroxisomal targeting of the CBB pathway leads to increased growth.

Cultivation of strains with cytosolic (CBBc+RuBisCO) and peroxisomal (CBBp+RuBisCO) pathway expression; Cell Dry Weight (CDW) values are calculated from OD_{600} measurements (correlation: 1 OD_{600} unit = 0.191 g L^{-1}CDW), for CBBpΔRuBisCO the mean of two individual cultivations including single values is shown (blue line plus signs), for CBBp+RuBisCO (black line) and CBBc+RuBisCO (pink line) the means of three biological replicates (independent transformations) with single values (black and pink signs) is shown.

Europe PMC Funders Author Manuscripts

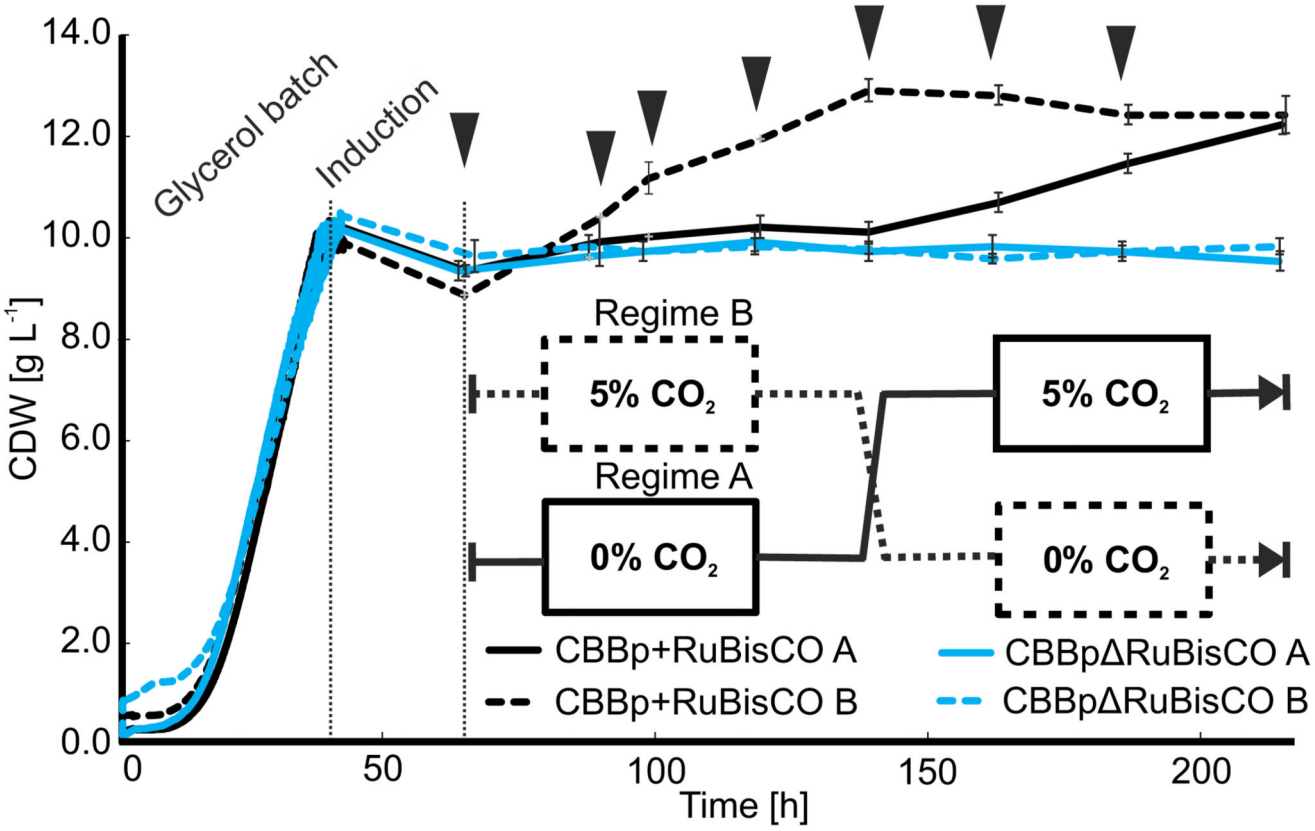

Figure 4. Growth of engineered CBBp+RuBisCO strain depends on the supply of CO_2 as a carbon source.

Biomass formation in engineered CBBp+RuBisCO (solid (Regime A) and dashed (Regime B) black lines) is shown compared to the control strains, which lack RuBisCO, GroEL and GroES (CBBpΔRuBisCO solid (A) and dashed (B) blue lines. Cells were cultivated in batch phase on 16.0 g L^{-1} glycerol to a CDW of ~ 10 g L^{-1} and then induced with 0.5 % methanol (v/v) and 1 % CO_2, afterwards methanol concentration was adjusted to 1 % (v/v) (arrows). After induction only CBBp+RuBisCO B and CBBpΔRuBisCO B were co-fed with 5 % CO_2. After 3 days and occurrence of pronounced growth, the CO_2 supply was set to 0 % for CBBp+RuBisCO B and CBBpΔRuBisCO B and increased to 5 % for CBBp+RuBisCO A and CBBpΔRuBisCO A, so that two independent replicates with and without CO_2 supply were achieved. Cell Dry Weight (CDW) values were calculated from OD_{600} measurements (correlation: 1 OD_{600} unit = 0.191 g L^{-1} CDW, error bars indicate *s.e.m.*).

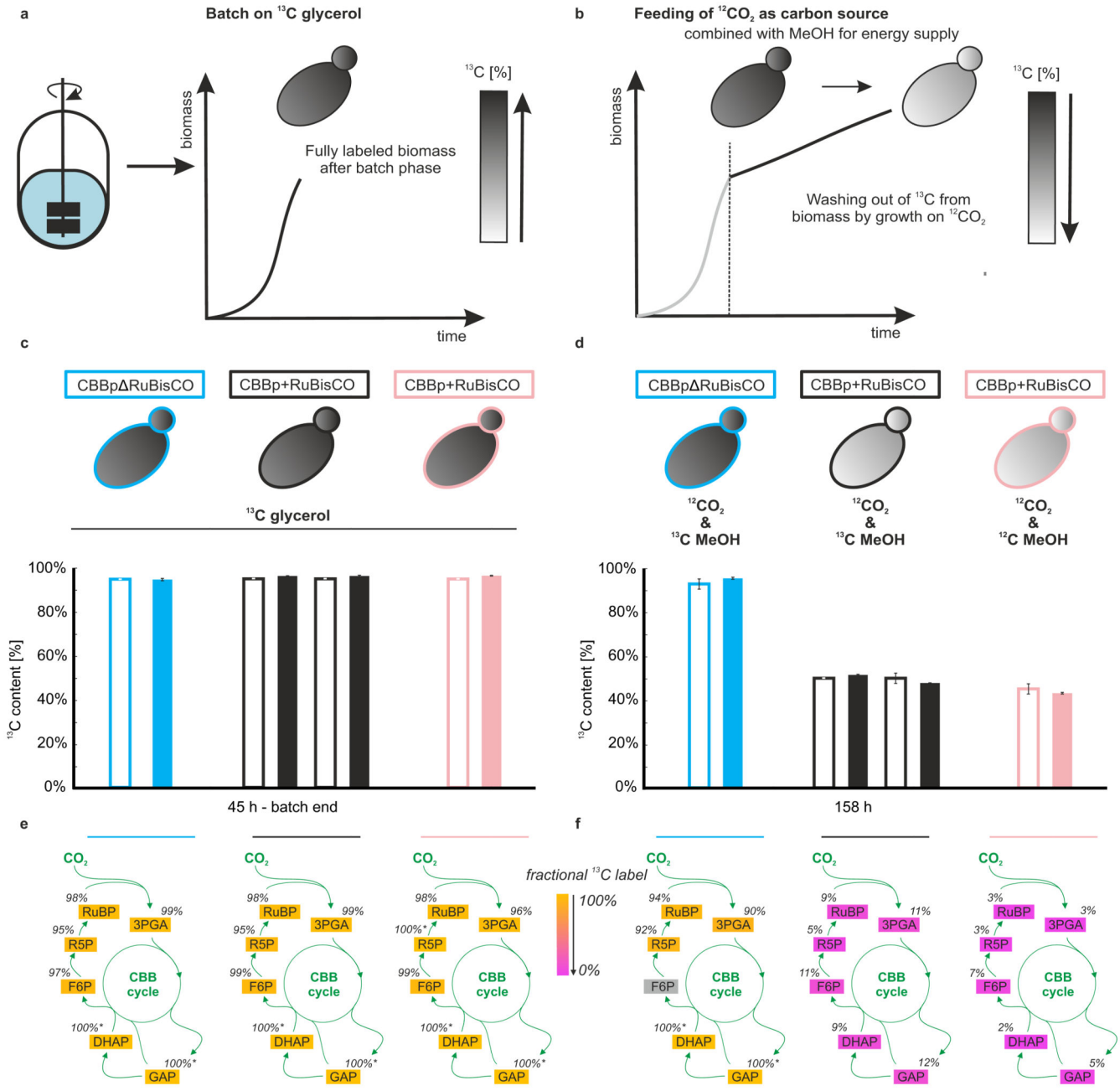

Figure 5. Labeling of engineered *P. pastoris* cells with ^{13}C glycerol and ^{12}C carbon dioxide.

(**a**) CBBpΔRuBisCO and CBBp+RuBisCO (three technical replicates) cells were enriched for ^{13}C by cultivation on fully labeled ^{13}C glycerol as the sole carbon source. White-to-black color key indicates the ^{13}C [%] of the biomass resulting from each condition; (**b**) After the batch phase the cells were fed with methanol (0.5 -1.0 % (v/v) and 5 % $^{12}CO_2$ which led to growth in CBBp+RuBisCO and no growth for the CBBpΔRuBisCO. ^{13}C methanol was used for two reactors containing CBBp+RuBisCO (black) and one reactor CBBpΔRuBisCO (blue), while ^{12}C methanol was used for one reactor with CBBp+RuBisCO (pink). Incorporation of $^{12}CO_2$ leads to a decrease of the total ^{13}C content in the biomass. (**c** – **d**)

Bar charts show total ^{13}C content after the batch phase (45 h) (**c**) and after 158 h of cultivation (**d**), open bars show calculated values and filled bars show measured values obtained by EA-IRMS (error bars indicate *s.e.m.*); (**e** and **f**) isotopic analysis of key metabolites from the CBB cycle by GC-MS and LC-MS, Shown is the fractional labeling calculated from the isotopologue distribution of each metabolite in [%], For values marked with (*) only the M+3 or M+5 with 3 or 5 incorporated ^{13}C atoms was detectable resulting in 100% fractional label (see Supplementary Information). Instrumental and Biological precision calculations are included in Supplementary Table 4).

Europe PMC Funders Author Manuscripts

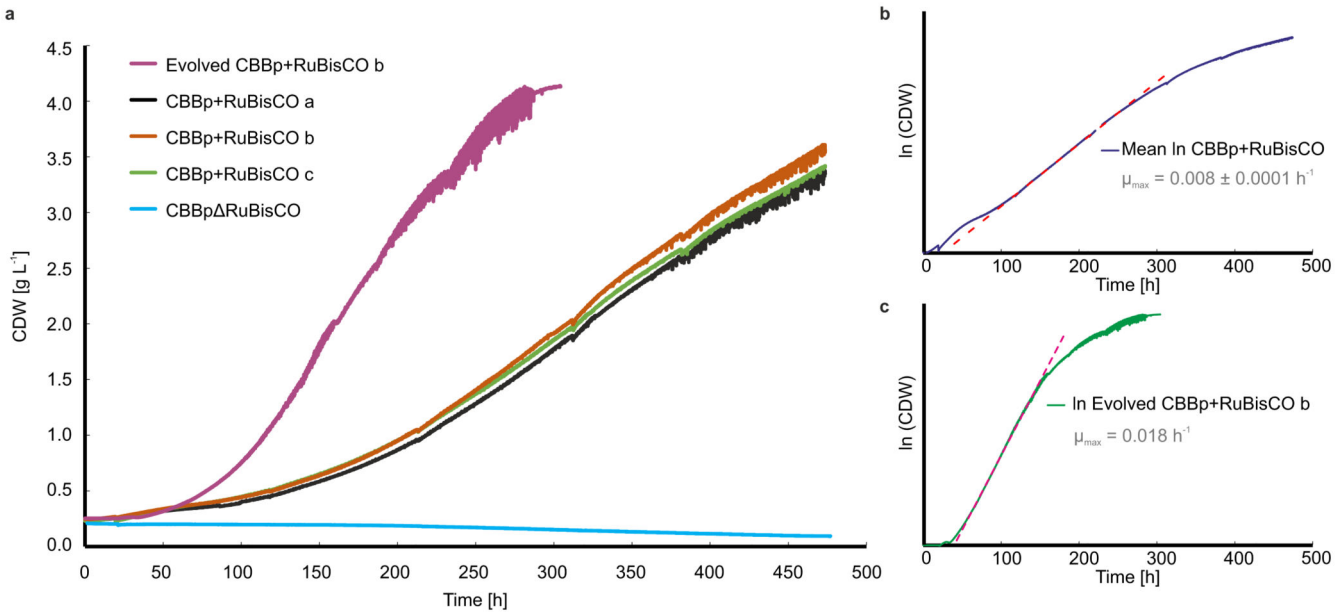

Figure 6. Bioreactor cultivation of CBBp+RuBisCO and CBBpΔRuBisCO strains inoculated at low cell density and grown in the presence of CO_2.

(**a**) During cultivation, cells were fed with 5 % CO_2 and methanol. For CBBp+RuBisCO CDW values of three biological replicates (individual transformations) are shown. For CBBpΔRuBisCO and the best evolved CBBp+RuBisCO b clone, values for a single cultivation each are shown. (**b** and **c**) Logarithmic CDW of original clones (n=3) and evolved CBBp+RuBisCO single isolate (red dotted lines show linear fitting).

Europe PMC Funders Author Manuscripts