Foundational Tools for Synthetic Methylotrophy in *Saccharomyces cerevisiae*

Philip A. Kelso

A thesis submitted to Macquarie University
for the degree of Master of Research
School of Natural Sciences
25th November 2022

Author Statement

I the author of this thesis, Philip A. Kelso, state that this work has not been submitted for a higher degree to any other university or institution. All sources have been referenced appropriately.

Philip A. Kelso

25 Nov 2022

Table of Contents

Abstract

The global expansion of biomanufacturing is currently limited by the availability of sugar-based microbial feedstocks. One-carbon feedstocks, like methanol, present an enticing alternative to sugar because they can be produced from organic waste, atmospheric carbon dioxide and hydrocarbons. The development of efficient industrial microorganisms which can convert one-carbon feedstocks into valuable products is therefore a research priority. This thesis will present work from three experiments and a literature review focused on the development of tools for engineering synthetic methanol assimilation in *Saccharomyces cerevisiae,* with the intent of laying a foundation for further projects. *Experiment 1* tested multiple different methanol dehydrogenases as part of a synthetic ribulose-monophosphate cycle in *S. cerevisiae* to identify which enzymes could confer a growth advantage in the presence of methanol. *Experiment 2* demonstrated a novel genome engineering technique known as Random Assembly and Integration, to show that promoters and methylotrophy associated genes from a pre-defined library could be randomly integrated into the repeated *Ty1* retrotransposon loci in the *S. cerevisiae* genome. *Experiment 3* used a previously reported synthetic xylose utilization pathway to establish a strain of *S. cerevisiae* which can grow on xylose that will be used in future projects to engineer synthetic methanol auxotrophy.

Introduction

The global expansion of biomanufacturing is currently limited by the availability of hexose sugar which is the primary microbial feedstock. This limitation arises because sugar requires arable farmland for cultivation of sugarcane or corn, and therefore cannot support broadscale biomanufacturing without impacting the human food supply[1]. One-carbon (C1) feedstocks such as methanol have attracted intense interest as an alternative to sugar because they can be produced as by-products from other activities, and their availability is not limited by the scarcity of arable farmland [2]. For example, methanol can be derived from methane, which is generated by natural gas deposits or biogas from municipal waste. Alternatively, methanol can also be derived from synthesis gas, which can be obtained through gasification of coal or waste organic material (Figure 1)[1]. Methanol can also be produced from atmospheric CO_2, which is reduced to methanol using hydrogen generated from electrolysis [3]. Work also continues on the development of electrochemical reduction technology for the direct conversion of atmospheric CO_2 into other C1 compounds [4, 5]. As these technologies develop further, it is likely that atmospheric CO_2 will become an abundant source of methanol in the future. Further, compared to other gaseous C1 compounds (carbon monoxide, CO_2, methane etc.), methanol provides a convenient liquid feedstock for large-scale transport and industrial fermentation [6]. Therefore, the development of industrial production hosts which can utilise methanol as their sole carbon source will enable bioproduction from hydrocarbons, greenhouse gases or biomass, and reduce our dependence on arable land for the production of microbial feedstocks [7]. However, this remains a difficult task because methanol is a toxic substrate which cannot be utilised by most common industrial microorganisms.

Native methylotrophs with the capacity to utilise C1 compounds as their sole carbon source are common in nature, where they subsist on a range of different substrates such as methane, formate, methyl amine, and methanol [8 9]. However, native methylotrophs are not currently suitable for broad use in biomanufacturing due to their poor genetic characterisation and limited genetic tractability [10]. This results in limited product spectra, low yields, and slow and expensive strain engineering. Comparatively, non-methylotrophic model industrial species have benefited from decades of research and development and have extremely well characterised genetics, established synthetic biology tools and high genetic tractability [11]. This has resulted in many established high-yield bioproduction pathways and flexibility for the development of new products, as demonstrated in the two most utilised model species, *Saccharomyces cerevisiae* and *Escherichia coli* [7]. *S. cerevisiae* is now used widely to produce a broad spectrum of value-added products [12] (Figure 1). Some of these include: organic acids [13], fatty acids [14], advanced biofuels [15], pharmaceuticals [16, 17], biopharmaceuticals [18], alkaloids [19], carotenoids [20], and flavors [21].

Two tracks have emerged for developing methanol-based biomanufacturing. One track involves the engineering of native methylotrophs to improve their capacity for bioproduction to match model species. The other track involves engineering synthetic methylotrophy in established model species. Both tracks have their own merits and challenges, and it remains to be seen which will be most successful [22, 23]. In recent years, considerable progress in engineering synthetic methylotrophy in model bacteria has been made and an excellent summary of the work thus far in bacteria can be found in a review from Gregory et al. [24]. This has culminated in the development of a strain of *E. coli* which is capable of robust growth on methanol as a sole source of carbon [25]. However, there remains a significant need to develop methanol-based biomanufacturing in yeast, as yeast provides a number of advantages over bacteria for use in industry (Figure 1). Yeast has a greater tolerance for high-speed centrifugation and mixing, solvent exposure, and variations in temperature and pH, making it a more robust production host which is suitable to a broader range of industrial conditions [26]. Industrial yeast fermentation is not susceptible to bacteriophage contamination which can significantly reduce the yield of industrial bacterial fermentations [27]. Eukaryotic protein folding mechanisms and post translational modifications also mean that yeast is able to produce a broader range of complex recombinant proteins [28]. Further, yeast's eukaryotic cell structure provides additional options for metabolic engineering through organelle targeted expression of biosynthetic pathways which can benefit from organelle specific metabolic processes [29], although it is possible to localize metabolism in bacteria using bacterial microcompartments [30].

Some progress in harnessing yeast for methanol-based bioprocessing has been made in the facultative methylotroph, *Pichia pastoris* (reclassified as *Komagataella phaffii*) which has already proven to be an efficient industrial species for recombinant protein production [31]. In addition to these proteins, a handful of valuable metabolites have now been synthesised in *P. pastoris* using methanol as the sole carbon source [32]. These include: organic acids, D-lactic acid [33], malic acid [34], and 6-methylsalicylic acid [35]; the polysaccharide hyaluronic acid [36]; and the important antihypertensive compound lovastatin and its precursor monacolin [37]. Despite this progress, *P. pastoris* and other native methylotrophic yeasts remain less amenable than *S. cerevisiae* to cutting-edge synthetic biology, metabolic engineering, and systems biology approaches, and therefore have potentially narrower bioproduction capabilities [38]. Accordingly, it would be most convenient if efficient synthetic methylotrophy could be enabled in *S. cerevisiae*.

The field of engineering synthetic methylotrophy in *S. cerevisiae* is still in its infancy and only a handful of primary publications and preprints are available on the topic [39, 40, 41, 42, 43]. Some progress

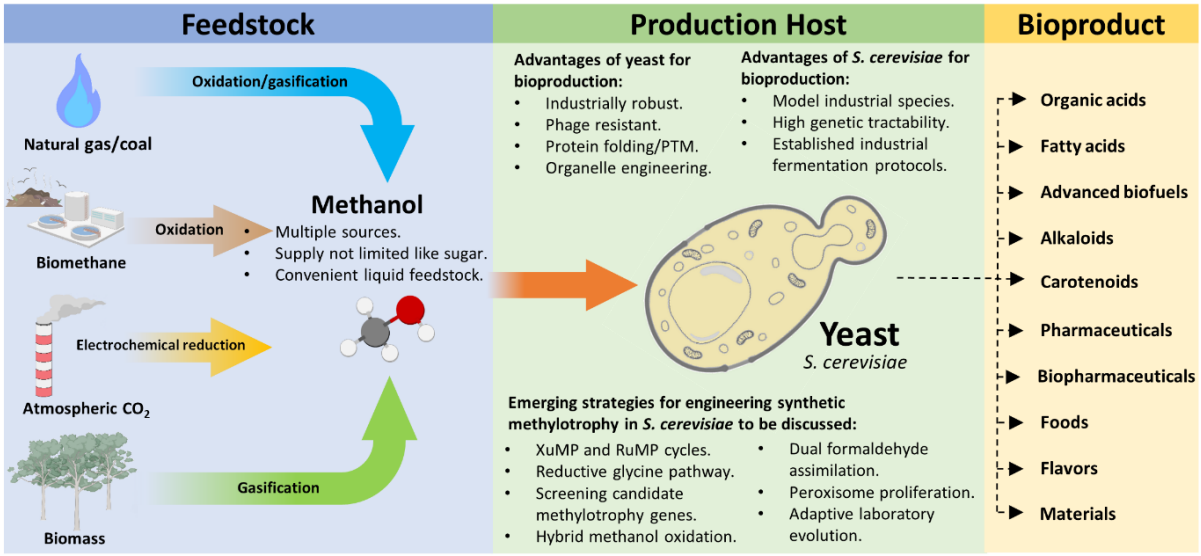

Figure 1. The future of methanol-based biomanufacturing. Methanol presents an enticing alternative to sugar-based microbial feedstocks for biomanufacturing. Unlike sugar, methanol does not require agricultural land for cultivation of sugarcane or corn and can be produced in large quantities as a by-product from other processes. Yeast presents several advantages for use in biomanufacturing and there is a pressing need to develop platform industrial yeast species for methanol-based biomanufacturing. *S. cerevisiae* is the most widely used, well developed, and versatile industrial yeast species for biomanufacturing. Therefore, it would be most convenient for industry if robust growth on methanol as the sole carbon source could be achieved *in S. cerevisiae*. However, because methanol is a toxic metabolite for most industrially relevant microbes, significant challenges remain. This image was created using Biorender.com.

towards synthetic methylotrophy in *S. cerevisiae* has been made in recent years using synthetic versions of the xylulose monophosphate (XuMP) cycle from methylotrophic yeast [39], and the ribulose monophosphate (RuMP) cycle from methylotrophic bacteria [40]. Synthetic methanol assimilation has also recently been demonstrated in the oleaginous industrial yeast, *Yarrowia lipolytica*, using a combined RuMP/XuMP synthetic methanol assimilation pathway [44]. Additionally, the core module of a synthetic reductive glycine (rGly) pathway has been engineered in *S. cerevisiae* to convert C1 formate into glycine.

This thesis will present a body of work examining the development of tools for engineering synthetic methanol assimilation in *S. cerevisiae* with the intent of laying a foundation for further projects. An initial literature review will introduce methylotrophic metabolism and discusses progress in the field of synthetic methylotrophy with a focus of how it relates to *S. cerevisiae*. The literature review will also include a discussion of the emerging strategies for engineering synthetic methylotrophy in *S. cerevisiae* and make suggestions for the design of future projects. It is worth noting that the literature review presented here has also been published as a review article in ACS Synthetic Biology [45], and has been submitted along with this thesis as an appendix. Subsequent to the literature review, work from three experiments will be presented.

Experiment 1 tested different methanol oxidation enzymes from a range of native methylotrophs as part of a synthetic RuMP cycle in *S. cerevisiae*. The first step in the RuMP cycle requires the oxidation of methanol into formaldehyde for subsequent assimilation. Therefore, it is important for

any synthetic RuMP cycle that effective methanol oxidation occurs in order to provide enough formaldehyde for sufficient carbon flux through the cycle to generate biomass. It is advantageous in metabolic engineering to test different genes from a range of organisms within a particular enzyme class to explore the potential of those enzymes with different cofactor usage, kinetics, expression levels, intracellular localisation, and solubility. So far, no published works in *S. cerevisiae* have compared methanol oxidation enzymes from different species in the same experiment, with each study only testing enzymes from a single species in isolation [39, 40, 42]. Therefore, the aim of this experiment was to test a number of methanol oxidation enzymes from different species in a synthetic RuMP cycle in *S. cerevisiae* in order to identify preferable candidates for use in future projects. The enzymes tested included nine different NAD^+ dependent methanol dehydrogenases (NAD-Mdhs) from bacterial methylotrophs and an alcohol oxidase from *P. pastoris*. The testing included serial 10-fold dilution spot assays on solid media with methanol as the sole carbon source as per Espinosa et al. [40]. Growth experiments were also conducted in liquid media with low concentration glucose and methanol. It was hypothesized that some of the enzymes tested may enable superior growth of *S. cerevisiae* in the presence of methanol and that the results of this experiment would inform the design of future synthetic methylotrophy studies in *S. cerevisiae.*

Experiment 2 tested a novel multi-locus genomic engineering technique developed by the Williams Lab, known as Random Assembly and Integration (RAIN), to introduce genes from the RuMP and XuMP cycles into *S. cerevisiae*. The technique employs a one-pot transformation method which is designed to generate expression level diversity from a pre-defined library of promoters and open reading frames (ORFs). This technique takes advantage of the innate capacity of *S. cerevisiae* for homologous recombination[46] to facilitate random assembly of promoters and ORFs from the library, and random integration of assembled expression cassettes into the *Ty1* retrotransposon, which occurs 313 times throughout the genome [47]. The idea behind employing such a technique is to negate the difficulties associated with traditional metabolic engineering approaches, where selection of genes, promoters, integration loci and copy number must be done in a rational manner and different variants engineered individually. Instead, RAIN aims to provide hundreds of variants from a single transformation experiment for subsequent high throughput screening. The primary aim of this experiment was to use the RAIN system in a one-pot transformation to generate random integrations of assembled RAIN cassettes among the transformants. If successful, it was hypothesized that RAIN would enable the random generation of hundreds of different variants containing genes from the RAIN library in a single transformation. The promoter library for this experiment contained six different promoters with varying strengths and characteristics and the

ORF library contains 18 methylotrophy associated genes from the RuMP and XuMP cycles and native pentose phosphate pathway (PPP).

For the required recombination of promoters and ORFs to occur, RAIN genetic constructs have been designed to include an identical intron sequence (*YDL191*) on the 3' end of the promoter cassettes and the 5' end of the ORF cassettes (Figure 2). This facilitates random homologous recombination of promoters and ORFs when combined in a transformation mixture. To allow for random integration of assembled RAIN cassettes into *Ty1* loci, a *Ty1* homology arm is included on the 5' end of the promoter cassettes and the 3' end of the ORF cassettes (Figure 2). To allow for selection of successful transformants a zeocin selective marker is included on the ORF cassettes. After transcription of assembled RAIN cassettes, the *YDL191* intron sequence should be cleaved out. The *YDL191* intron sequence between the promoter and ORF cassettes allows for an *in vitro* Gibson Assembly prior to transformation, and this was also tested as part of *Experiment 2*.

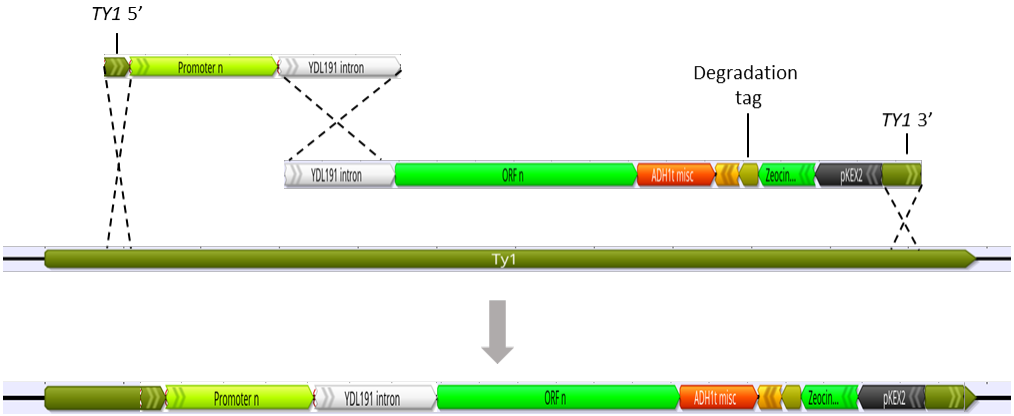

Figure 2. The RAIN genome engineering system. RAIN takes advantage of the innate capacity of *S. cerevisiae* for homologous recombination to facilitate random assembly and integration of promoters and ORFs from a pre-designed library. When promoters and ORFs are combined in a mixture and transformed in a one-pot experiment, they will randomly assemble and integrate into the *Ty1* retrotransposon, which occurs 313 times throughout the genome[47]. This generates random expression level diversity of ORFs in the library through different combinations of insertions, promoter strengths, copy number and insertion loci. After successful transformation, colonies can be screened for a desired phenotype. The homologous recombination which is required for the RAIN system to work occurs via *Ty1* homology arms and an identical intron sequence which are included on the ends of the promoter and ORF cassettes. As shown in this diagram, the ORF cassettes contain a zeocin selective marker which is under the weak KEX2 promoter with a degradation tag. This figure was extracted from Geneious Prime.

Experiment 3, engineered synthetic xylose assimilation in *S. cerevisiae* as a prelude to engineering synthetic methanol auxotrophy, also called methanol dependent growth, via a ribose-5-phosphate isomerase (*RKI1*) knockout in the PPP. Synthetic methanol auxotrophy has been employed effectively in *E. coli* as a starting point for ALE towards full synthetic methylotrophy [25]. To engineer synthetic methanol auxotrophy, cells must be grown on xylose so that central carbon metabolism relies on the PPP. Because *RKI1* is an essential gene for PPP function, knocking it out inhibits growth on xylose. PPP function and cell growth can be restored by the addition of a synthetic RuMP pathway. Importantly, the RuMP pathway requires methanol to function and cell growth becomes dependent on methanol assimilation [4]. This increases the selective pressure for

cells to assimilate methanol during ALE on xylose and methanol and increases the chances of success during ALE towards full synthetic methylotrophy.

However, *S. cerevisiae* does not possess a native capacity for growth on xylose. Therefore synthetic xylose assimilation must first be implemented if synthetic methanol auxotrophy is to be engineered. The aim of this experiment was to implement a previously reported xylose utilisation pathway from the xylose fermenting yeast *Pichia stipitis* [48], to enable growth on xylose in *S. cerevisiae*. This pathway uses aldose reductase (*XYL1*) and xylitol dehydrogenase (*XYL2*) from *P. stipitis* as well as overexpression of native xylulokinase (*XKS1*) from *S. cerevisiae*. It was hypothesized that implementation of this pathway would enable growth on xylose as previously reported by Jin et al. [48]. The xylose utilising strain established in this experiment will be used in subsequent experiments which will attempt to engineer synthetic methanol auxotrophy in *S. cerevisiae* via an *RKI1* knockout.

Literature review

This review summarises the molecular mechanisms behind methylotrophy with a focus on the RuMP and XuMP cycles, and the reductive glycine pathway, and how they may be applied in *S. cerevisiae*. Robust growth of *S. cerevisiae* using methanol as the sole carbon source is yet to be achieved. We therefore discuss different strategies to enhance methanol assimilation and metabolic cycling through the RuMP/XuMP cycles and the rGly pathway. These include: screening candidate methylotrophy genes; hybrid methanol oxidation; dual formaldehyde assimilation; peroxisome proliferation and compartmentalisation; and adaptive laboratory evolution (ALE). It is our intent that the strategies discussed in this review will aid in the design of future synthetic methylotrophy projects in *S. cerevisiae* and facilitate further progress in the field.

Methylotrophy

In nature, most methylotrophs possess at least one of four common C1 assimilation pathways: the yeast XuMP cycle; the bacterial RuMP cycle; the serine cycle; and the bacterial ribulose bisphosphate (RuBP) cycle (Figure 3) [7]. Additionally, the rGly pathway, which has so far been reported in a small number of anaerobic bacteria, is of interest to researchers studying synthetic methylotrophy [49, 50, 51]. This section will provide a brief summary of the mechanisms behind each pathway before a more detailed discussion of pathway engineering in later sections.

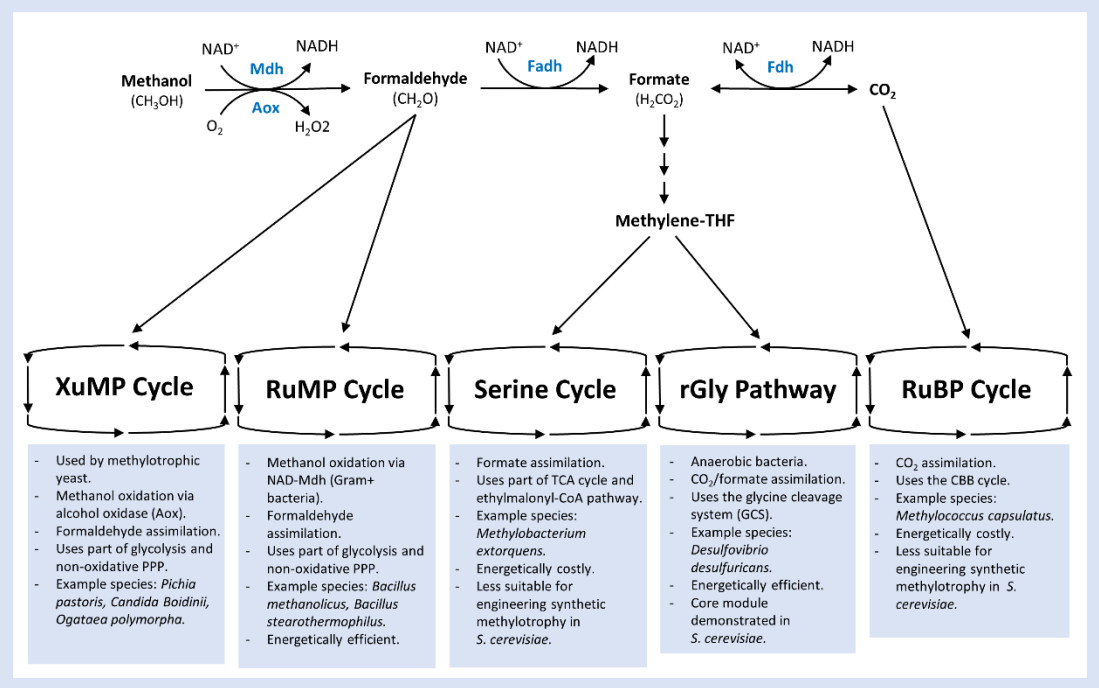

Figure 3. A general summary of methylotrophic metabolism. The four common C1 assimilation cycles in nature include the yeast XuMP cycle, bacterial RuMP cycle, serine cycle, and RuBP cycle. The rGly pathway has also recently been identified as an energetically efficient pathway of interest for synthetic methylotrophy. Both the RuMP and XuMP cycles have been used to demonstrate synthetic methanol assimilation in S. cerevisiae [39, 40]. The core module of the rGly pathway has also been demonstrated in *S. cerevisiae* [43]. The serine and RuBP cycles are less energetically favourable and appear to be less suitable for engineering synthetic methylotrophy [10]. Mdh, methanol dehydrogenase; Fadh, formaldehyde dehydrogenase; Fdh, formate dehydrogenase; Aox, alcohol oxidase; THF, tetrahydrofolate.

RuMP and XuMP cycle methylotrophy

The XuMP and RuMP cycles share a number of similarities. Both cycles involve the oxidation of methanol to formaldehyde; fixation of formaldehyde to a phosphorylated pentose acceptor molecule; cleavage of subsequent intermediates into C3 glycolytic metabolites; and rearrangement through the non-oxidative pentose phosphate pathway (PPP) to regenerate the formaldehyde acceptor (Figure 4 A, B). Despite their similarities, the RuMP and XuMP cycles differ in a number of ways. Methanol oxidation in the XuMP cycle is catalysed by an oxygen dependent alcohol oxidase (Aox) in the yeast peroxisome [52], giving formaldehyde and hydrogen peroxide (H_2O_2). In the RuMP cycle, methanol is oxidised in the bacterial cytosol via a NAD^+ dependent methanol dehydrogenase (NAD-Mdh) [53], giving formaldehyde and NADH. The major difference between the two cycles at this step is in the two different reduction products formed. Hydrogen peroxide is a highly reactive and toxic compound which must be contained within the peroxisome where it is detoxified by a catalase (Cta) into oxygen and water, a process resulting in the loss of electrons from methanol. NADH on the other hand, is a useful reducing equivalent which preserves electrons from methanol that can be used to drive other metabolic processes. Therefore, the RuMP cycle is more efficient at utilising the available electrons from methanol at this early stage of the cycle [54].

After methanol oxidation in the XuMP cycle, formaldehyde is fixed to xylulose 5-phosphate (Xu5P) by dihydroxyacetone synthase (Das), which then cleaves the complex into the important C3 sugars, glyceraldehyde 3-phosphate (G3P) and dihydroxyacetone (DHA)(Figure 4A). These C3 intermediates then enter the cytosol where DHA is phosphorylated by dihydroxyacetone kinase (Dak) into dihydroxyacetone phosphate (DHAP). It has also been reported in *P. pastoris* that DHA phosphorylation occurs in the peroxisome by a Dak paralog [55]. From here, G3P and DHAP enter glycolysis. In the RuMP cycle, formaldehyde is fixed to ribulose 5-phospahte (Ru5P) by 3-hexulose-6-phosphate synthase (Hps), giving hexulose 6-phosphate (H6P)(Figure 4B). This H6P is then isomerised into glycolytic fructose 6-phosphate (F6P) for subsequent cleavage into G3P and DHAP [56]. Although the formaldehyde acceptors differ between the two cycles, Xu5P and Ru5P are both involved in the same stage of the non-oxidative PPP. Therefore, regeneration of the formaldehyde acceptor is very similar for both cycles and occurs via transketolase (Tkt) and transaldolase (Tal) (Figure 4A, B). The three key enzymes in the XuMP cycle are Aox, Das and Dak, whereas in the RuMP cycle they are NAD-Mdh, Hps and Phi. All the other enzymes in both cycles come from glycolysis or the PPP.

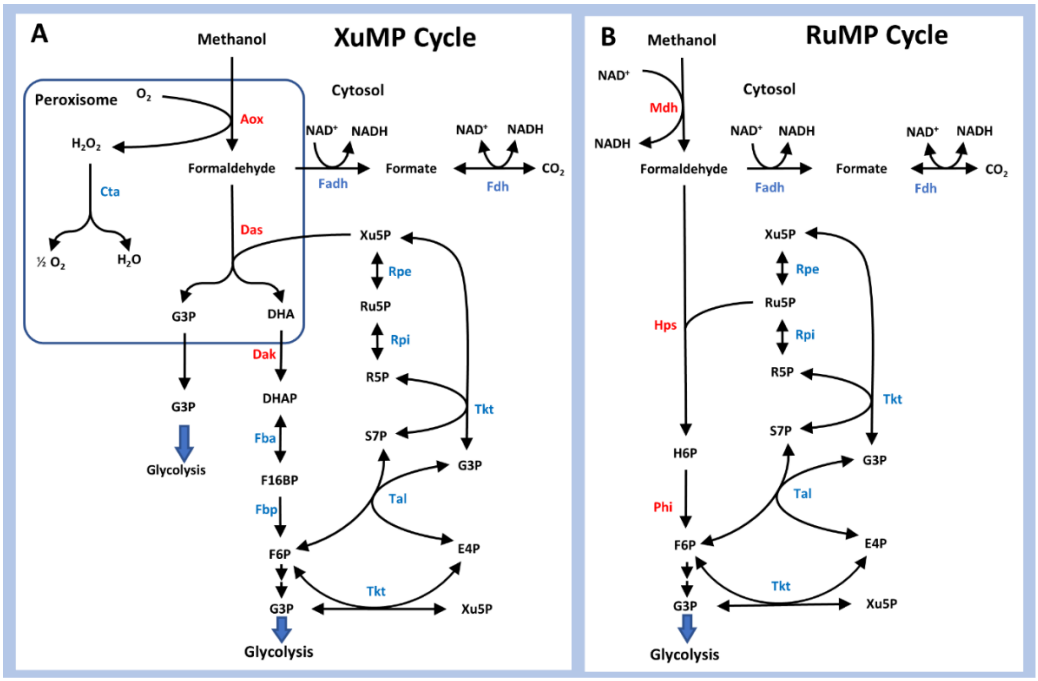

Figure 4. General scheme of the XuMP and RuMP cycles. A: Methanol metabolism via the XuMP cycle in the example species, *C. boidinii* [57]. Enzymes shown in red are specific to the XuMP cycle. **B:** Methanol metabolism via the RuMP cycle in the example species, *B. methanolicus* [58]. Enzymes shown in red are specific to the RuMP cycle. **Mdh**, methanol dehydrogenase; **Fadh**, formaldehyde dehydrogenase; **Fdh**, formate dehydrogenase; **Aox**, alcohol oxidase; **Cta**, catalase; **Das**, dihydroxyacetone synthase; **Dak**, dihydroxyacetone kinase; **Hps**, 3-hexulose-6-phosphate synthase; **Phi**, phosphohexuloisomerase; **Fba** fructose-bisphosphate aldolase; **Fbp**, fructose bisphosphatase; **Tkt**, transketolase; **Tal**, transaldolase; **Rpi**, ribose-5-phosphate isomerase; **Rpe** ribulose phosphate 3-epimerase; **G3P**, glyceraldehyde 3-phosphate; **DHA**, dihydroxyacetone; **DHAP**, dihydroxyacetone phosphate; **F16BP**, fructose-1 6-bisphposphate; **H6P**, hexulose 6-phosphate; **F6P**, fructose 6 phosphate; **E4P**, erythrose-4-phosphate; **S7P**, sedoheptulose 7-phosphate; **Xu5P**, xylulose 5-phosphate; **R5P**, ribose 5-phosphate; **Ru5P**, ribulose 5-phosphate.

Serine and RuBP cycles

The serine cycle differs significantly from the RuMP and XuMP cycles as it starts with the three-step conversion of formate into methylene-tetrahydrofolate (methylene-THF) (Figure 3) [59]. The key enzyme in the serine cycle is serine hydroxymethylase (Shmt) which catalyses the addition of methylene-THF to glycine giving serine [60]. Serine then undergoes a series of transamination reactions which lead to the formation of glycolytic metabolites for cell growth. Regeneration of glycine, the primary methylene-THF acceptor, is reliant on the regeneration of glyoxylate which comes from the passage of acetyl-CoA through the complicated ethylmalonyl-CoA pathway [61]. The RuBP cycle occurs in methylotrophic bacteria and uses the Calvin Benson Basham (CBB) cycle to assimilate CO_2 as an oxidation product of formate (Figure 3) [62]. Both the serine cycle and the RuBP cycle are less energetically favourable than the RuMP and XuMP cycles due to lower net ATP yields per molecule of carbon assimilated [63]. Therefore it is generally accepted that they are less suitable for engineering synthetic methanol assimilation for industrial purposes [10]. Accordingly, further discussion of the serine cycle and RuBP cycle is outside the scope of this review.

Reductive glycine pathway

The rGly pathway was first proposed in 2013 by Arren Bar-Even et al. [64] as a theoretical synthetic pathway for the efficient assimilation of formate during synthetic methylotrophy in *E. coli*. Evidence of the rGly pathway in nature has since been reported in a small number of anaerobic bacteria which reduce CO_2 to formate for assimilation [50, 51]. The core module of the rGly pathway relates to the conversion of formate to glycine, and begins with a three-step conversion of formate to methylene-THF (Figure 5) [49]. The methylene-THF is then condensed with CO_2 and NH_3 by the glycine cleavage system (GCS) to give glycine. The GCS is a four-component reversible complex which uses aminomethyl-transferase (Amt), glycine dehydrogenase (Gldc) and dihydrolypoyl dehydrogenase (Dld). Glycine can then be converted into biomass via a number of different routes. These include through the serine cycle, or directly into pyruvate via serine under aerobic conditions, or to pyruvate via acetyl-Coa under anaerobic conditions [65]. Computational flux analysis has identified the rGly pathway as an energetically efficient route for assimilating once carbon substrates, with similar net ATP yields per molecule of carbon assimilated to the RuMP pathway [65]. Therefore, the rGly pathway has gained interest as an option for engineering synthetic methylotrophy for industrial applications.

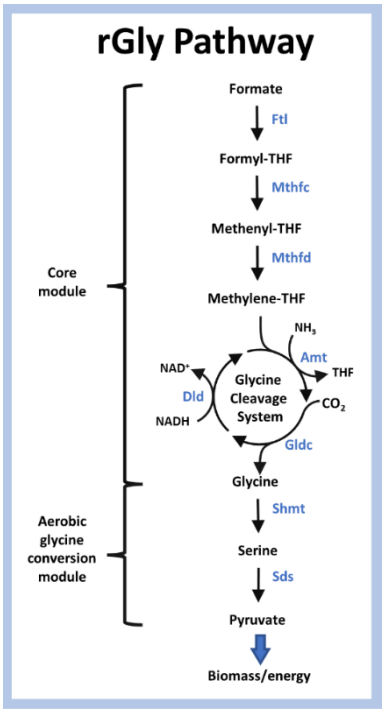

Figure 5. General scheme of the reductive glycine (rGly) pathway. The rGly pathway was first proposed as a synthetic pathway for the efficient assimilation of formate during synthetic methylotrophy [64], however evidence of the pathway in nature has since been reported in anearobic bacteria which reduce CO_2 to formate for assimilation [50, 51]. The core module of the rGly pathway relates to the conversion of formate to glycine via methylene-THF and and the glycine cleavage system. Glycine can then be assimilated into biomass via a number of different routes. This figure includes the module for the aerobic conversion of glycine to biomass. The core module of the rGly pathway has been demostrated in *S. cerevisiae* [43]. **THF,** tetrahydrofolate; **Ftl,** formate-THF ligase; **Mthfc,** methenyl-THF cyclohydrolase; **Mthfd**, methenyl-THF dehydrogenase; **Amt,** aminomethyltransferase; **Gldc**, glycine dehydrogenase; **Dld** dihydrolipoyl dehydrogenase; **Shmt**, serine hydroxymethyl transferase; **Sds**, serine deaminase.

Engineering synthetic methanol assimilation in S. cerevisiae

Engineering a synthetic XuMP cycle: Aox, Das and Dak

A synthetic XuMP cycle in *S. cerevisiae* can be engineered by the introduction of genes for Aox, Das and Dak (Figure 4A). The first study to engineer synthetic methanol assimilation in *S. cerevisiae* came from Dai et al. [39] and used genes from *P. pastoris* to construct two synthetic XuMP cycle variants which used different *DAS* paralogs, *DAS1* and *DAS2,* along with *AOX1, CTA1* and *DAK*. Neither strain was able to grow in liquid medium with methanol as the sole carbon source.

However, cultures in liquid medium with 1% methanol and 0.1% yeast extract increased growth from a starting OD_{600} of 7, by 9.2% and 11.7% for the *DAS1* and *DAS2* variants respectively when compared to a control. Methanol utilisation rates of 2.15 g/L/72 h and 2.35 g/L/72 h were also observed for the *DAS1* and *DAS2* variants respectively. Although these results are promising it is important to note that they were recorded in cultures with a starting OD_{600} of ~7 which is very high [39]. A subsequent study in *S. cerevisiae* contained in a preprint from Espinosa et al. [40] used *AOX1* and *DAS1* from *P. pastoris* to examine if these enzymes alone would be sufficient in establishing a functional XuMP cycle and methanol assimilation. The rationale was that *S. cerevisiae* already

possessed native *CTA1* and *DAK1* genes which would complete the cycle. However, growth assays on solid medium with methanol as the sole carbon source did not result in any increase in growth when compared to a control strain. This was also the case when supplemental yeast extract was provided in the medium [40]. A more recent study in *S. cerevisiae* reported in a preprint currently under review by Cell Press from Zhan et al. [42] was able to achieve methanol assimilation through a synthetic XuMP cycle via the introduction of *AOX1-DAS1-DAK* from *P. pastoris* and overexpression of the native *CTA1*. Taken together, the results from these three studies suggest that the native *DAK1* in *S. cerevisiae* is not sufficient to complete a synthetic XuMP cycle and that a full suite of heterologous *AOX1-DAS1-DAK* genes are required [39, 40, 42].

Engineering a synthetic RuMP cycle: NAD-Mdh, Hps and Phi

Using a synthetic RuMP cycle to engineer methanol assimilation in industrial microbes has so far proven most successful in bacteria [25], but it has also been shown to confer a growth advantage on methanol in *S. cerevisiae* [40]. A synthetic RuMP cycle can be engineered through the introduction of a NAD-Mdh, Hps and Phi (Figure 4B). The RuMP cycle presents a different set of advantages and challenges compared to the XuMP cycle. *In silico* modelling has identified the RuMP cycle as having the most efficient stoichiometry for methanol assimilation into biomass in *S. cerevisiae* [66]. Further, The RuMP cycle is simpler than the XuMP cycle because it does not involve transport of proteins and metabolites across the peroxisomal membrane. As discussed previously, RuMP cycle methanol oxidation generates useful reducing equivalents in the form of NADH which preserves electrons from methanol [54]. Additionally, because NAD-Mdh does not produce hydrogen peroxide, methanol oxidation in the RuMP cycle can proceed in the cytosol and is not limited to the peroxisome as in the XuMP cycle. However, cytosolic methanol oxidation is accompanied by the obvious pitfall of allowing formaldehyde, a highly reactive and toxic compound, to circulate freely around the cell. The problems associated with formaldehyde exposure and cytotoxicity through DNA-protein crosslinking have been well documented in *S. cerevisiae* [67]. Further, because Aox requires oxygen as a cofactor, XuMP cycle methylotrophy cannot function under anaerobic conditions whereas the NAD-Mdh based RuMP cycle can [68, 69].

A synthetic RuMP cycle in *S. cerevisiae* was also tested by Espinosa et al. [40] using *MDH* from *Bacillus stearothermophilus*; and *HPS* and *PHI* from *Bacillus methanolicus*. Growth assays were conducted on defined solid media with methanol as the only additional carbon source at 1%, 2% and 4%. The RuMP strain was compared to a control strain with no methanol assimilation cycle and two other strains, one with a XuMP cycle and one with a hybrid XuMP cycle. The RuMP strain showed superior growth at all methanol concentrations when compared to the control, XuMP and hybrid

XuMP strains [40]. These results suggest that the presence of a RuMP cycle does confer a growth advantage to *S. cerevisiae* when grown on methanol and that a synthetic RuMP cycle may be used to enable methanol assimilation in *S. cerevisiae*. However, it is worth noting that these growth advantages on solid media were not observed in liquid methanol media even when supplemental yeast extract was provided [40]. Interestingly, the previous study from Dai et al. [39] also attempted to engineer a synthetic RuMP cycle in *S. cerevisiae* using *MDH* from *B. methanolicus* and; Hps (*HXLA*) and Phi (*PGI*) from the non-methylotrophic *Bacillus subtilis,* but failed to observe any methanol specific growth [39]. Although *B. subtilis* possesses a RuMP cycle, it's function pertains to formaldehyde detoxification and not assimilation [70]. While not discussed by Dai et al. [39], this suggests that Hps and Phi enzymes from non-methylotrophic species should be considered with caution when designing a synthetic RuMP cycle in *S. cerevisiae*.

Engineering a reductive glycine pathway

One advantage of using the rGly pathway over the RuMP and XuMP cycles is that it is a linear pathway that does not require the regeneration of a crucial C1 acceptor molecule. Additionally, because the rGly pathway does not overlap greatly with central carbon metabolism like the RuMP and XuMP cycles, changes in the flux of the rGly pathway should have less influence on cellular metabolism [71]. Therefore, it should be easier to implement changes during strain engineering without affecting cell fitness. The utility of the rGly pathway for enabling growth on C1 substrates has recently been demonstrated in *E. coli* by Kim et al. [72]. This study employed extensive strain engineering with heterologous gene expression to construct a synthetic rGly pathway strain which was able to grow on formate as the sole carbon source with a doubling time of eight hours [72]. These results in *E. coli* present the rGly pathway as a promising candidate for use in industrial synthetic methylotrophs in the future. However, this study also demonstrates the complexities associated with engineering the rGly pathway in *E. coli* [72].

Interestingly, all the endogenous genes required for the conversion of formate to glycine through the core module of the rGly pathway are endogenous to *S. cerevisiae*. This was demonstrated by Gonzalez de la Cruz et al. [43], who implemented function of the core module of the rGly pathway in *S. cerevisiae* by the overexpression of only four native genes. As such, using *S. cerevisiae* for engineering the rGly pathway possibly presents a much simpler alternative to *E. coli.* The study began with a glycine auxotrophic strain which was generated by the deletion of serine hydroxymethyltransferase (Δ*SHM1,* Δ*SHM2)*, threonine aldolase (Δ*GLY1*) and alanine; glyoxylate aminotransferase (Δ*AGX1*) [43]. This metabolic background was used to select for strains which were only synthesising glycine from formate through a functional rGly pathway when grown in glycine

deficient media. The native genes overexpressed included *MIS1,* and the genes encoding the three GCS enzymes (Amt, Gldc and Dld). After growth in liquid media with glucose and ^{13}C formate, all the glycine in the engineered strain was shown to be ^{13}C labelled, confirming that glycine synthesis was only occurring through the rGly pathway [43]. These results are remarkable, as it is the first time that net synthesis of glycine through the rGly pathway has been demonstrated in a eukaryote. Further, they point to the great potential of the rGly pathway as a possible route towards synthetic methylotrophy in *S. cerevisiae.* However, this remains an open challenge and will require further engineering to enable efficient conversion of methanol to formate for assimilation, and efficient conversion of glycine to pyruvate for cell growth.

Review of emerging strategies to enhance synthetic methanol assimilation in S. cerevisiae

The early synthetic methylotrophy work in *S. cerevisiae* [39 40, 43] has demonstrated that the presence of a synthetic methanol assimilation cycle in *S. cerevisiae* on its own is not enough to enable robust cell growth at an industrial scale using methanol as the sole carbon source. Therefore, any strain with an engineered XuMP, RuMP or rGly cycle/pathway must be enhanced through further modification if effective methylotrophy is to be achieved.

Screening of candidate methylotrophy genes

The screening and selection of suitable methylotrophy genes for use in a desired host is a good starting point for any synthetic methylotrophy project. This is important as it is difficult to predict which enzyme homologs will function best in a given host strain. Screening can be conducted by expressing candidate enzymes in the platform host and assaying their performance. For example, Krog et al., expressed and characterised three NAD-Mdhs from two separate strains of *B. methanolicus* (MGA3 and PB1) in recombinant *E. coli* [73]. Only Mdh2 and Mdh3 from *B. methanolicus* MGA3 possessed activity in *E. coli*. Interestingly, the Mdh paralogs tested by Krog et al. differ mainly in amino acids around the active sight [73]. Although not discussed directly by Krog et al., this may explain the differences in enzyme activity which were observed.

The screening of candidate methylotrophy genes can also be applied to whole cycles. A good example of this is provided in an extensive study from Müller et al. [74], who screened a number of genes from different cycles and organisms to identify the most preferable for engineering synthetic methylotrophy in *E. coli* [74]. First, Müller et al. [74] used *in silico* modelling based on a stoichiometric genome scale model of *E. coli* [75] to identify the RuMP cycle as the best candidate. This was followed by the expression of different NAD-Mdh, Hps and Phi candidate genes from five different bacterial species [74]. Analysis of *in vivo* and *in vitro* enzyme activity identified *MDH2*, *HPS* and *PHI*

from *B. methanolicus* MGA3 as the most active in *E. coli*. Subsequent ^{13}C methanol experiments with *E. coli* containing these genes showed that ^{13}C accounted for up to 40% of the carbon in central metabolism [74].

A similar study of candidate methylotrophy genes in *S. cerevisiae* has also been reported in the previously discussed preprint from Zhan et al. [42]. This study tested sixteen strain variants with different combinations of RuMP and XuMP cycle genes. These genes included: *AOX1*, *DAS1*, *DAS2* and *DAK* from *P. pastoris*; and *MDH, MDH2, MDH3, ACT1*, *HPS* and *PHI2* from *B. methanolicus* MGA3. The variants tested included two RuMP/XuMP hybrid constructs combining *MDH-DAS1-DAK* and *AOX1-HPS-PHI*. No strains were able to grow in minimal media with methanol as the sole carbon source. However, culturing in liquid medium with 0.5% methanol and 1% yeast extract did enable increased growth in all strains when compared to the control [42]. Strains which harboured XuMP cycle *DAS1* and *DAK* along with either *AOX1* or *MDH3* showed the best growth. Further enzyme assays showed that methanol oxidation for Mdh3 was over 2-fold higher compared to the other Mdh paralogs and Aox1 was 2.3-fold higher than Mdh3. These experiments provide useful data which will aid in considerations for enzyme choice in future synthetic methylotrophy projects in *S. cerevisiae*.

Researchers continue to look beyond *B. methanolicus* for more NAD-Mdhs which could possibly be used to enhance synthetic methanol assimilation. One study identified a NAD-Mdh from *B. stearothermophilus* (*Bs*Mdh) as being preferable to the Mdhs from *B. methanolicus* (*Bm*Mdh) for use in *E. coli* [76]. This was due to *Bs*Mdh's higher methanol turnover rate to enhance the growth of synthetic methylotrophic strains when grown on methanol as the sole carbon source. Testing *Bs*Mdh in a synthetic RuMP cycle with liquid cultures of 60 mM (~0.25 %) ^{13}C methanol and 0.1% yeast extract demonstrated up to 39% ^{13}C assimilation in TCA cycle intermediates and up to 53% ^{13}C in G3P indicating that methanol assimilation was occurring. In comparison, a strain which used *Bm*Mdh2 in the same media which showed very little ^{13}C-labelling of any intracellular metabolites [76]. As previously mentioned, *Bs*Mdh was used for the construction of a synthetic RuMP cycle in *S. cerevisiae* and did increase growth on solid media with methanol as the sole carbon source [40].

The NAD-Mdhs from *B. stearothermophilus* and *B. methanolicus* are accompanied by a native nudix hydrolase activator protein (Act) which assists with methanol oxidation. This is common for many NAD-Mdhs where Act helps with cleavage of NADH after the reduction of the NAD^+ cofactor during enzyme function and reset [77, 78]. Although NAD-Mdh methanol oxidation can proceed without Act present, it has been shown that NAD-Mdh function can be considerably

improved when Act is present [74]. However, not all NAD-Mdhs use activator proteins to assist with methanol oxidation. Research in *Cupriavidus necator*, a formate utilising bacterium, identified and characterised the first NAD-Mdh that does not enhance activation from the known class of Act activator proteins, known as Mdh2 [79]. *C. necator* Mdh2 (*Cn*Mdh2) exhibited similar activity and affinity towards methanol when compared with *Bm*Mdh but did not use an Act. After ALE in *E. coli*, an isoform of *Cn*Mdh2 was characterised that exhibited a 6-fold increase in specificity for methanol compared with wild type *Cn*Mdh2 [79]. More recently, a second Act independent NAD-Mdh has been characterised in the Gram-positive *Lysinibacillus xylanilyticus*, (*Lx*Mdh) [80]. To date, no Act-independent Mdhs have been tested in *S. cerevisiae*, it is possible that they could contribute to improved methanol assimilation through a synthetic RuMP or hybrid XuMP cycle and we believe this warrants further investigation.

Hybrid methanol oxidation and dual formaldehyde assimilation

Hybrid methanol oxidation involves the mixed use of an Aox and/or Mdh during RuMP or XuMP cycle methylotrophy. This concept has been tested in *S. cerevisiae* by Zhan et al. [42] using the XuMP cycle from *P. pastoris* along with *MDH3* from *B. methanolicus* (*BmMDH3*) and its accompanying *ACT1* activator protein [42]. This established what Zhan et al. [42] called a 'double catalytic engine' to preserve electrons from methanol via the generation of NADH by *BmMDH3*. In *P. pastoris* the loss of electrons from *AOX1* based methanol oxidation is offset by an energetically efficient formaldehyde detoxification cycle which generates NADH from formaldehyde at a 2:1 molar ratio [81]. Because the formaldehyde dissimilation cycle in *S. cerevisiae* does not generate as much NADH as the one in *P. pastoris*, *BmMDH3* was used to generate NADH instead [42]. Apart from NADH generation, this hybrid methanol oxidation strategy should also provide more formaldehyde for an assimilatory cycle. After the introduction of *BmMDH3* and *ACT1* to a strain of *S. cerevisiae* which already harboured the XuMP cycle from *P. pastoris*, the resulting strain was able to increase its methanol consumption in minimal media with methanol as the sole carbon source by 155% compared to a XuMP only control strain which did not possess hybrid methanol oxidation. This hybrid methanol oxidation strategy could also work to increase methanol assimilation through a synthetic RuMP cycle where an additional Aox would complement the formaldehyde production of a NAD-Mdh and may be considered in any future RuMP cycle projects.

More recently in a paper from Wang et al. [44], the oleaginous industrial yeast *Yarrowia lipolytica* was engineered by combining the RuMP and XuMP cycles in a chimeric methanol assimilation pathway. The pathway was genomically integrated and included four copies of *MDH* from *B. stearothermophilus* and one copy each of *B. methanolicus HPS* and *PHI* for the RuMP cycle and

one copy each of *DAS1* and *DAK2* from *P. pastoris* for the XuMP cycle. Further modifications to boost formaldehyde flux were also included and will be discussed in a later section of this review. When grown in defined medium with glucose and 2% methanol, this chimeric strain was able to consume methanol at 0.42 g/L/72 h, which was significantly more than the other single RuMP or XuMP cycle variants tested [44]. This is the first study to report the simultaneous use of a XuMP and RuMP cycle in yeast. The strategy benefits from using a dual mechanism for formaldehyde assimilation via the Hps/Phi of a RuMP cycle and the Das/Dak from a XuMP cycle. It seems possible that this strategy could also be applied in *S. cerevisiae* where it may enable improved methanol assimilation over strains which contain either a single RuMP or XuMP cycle.

Peroxisome proliferation and compartmentalisation to enhance XuMP cycle methanol assimilation

In their discussion, Espinosa et al. [40] identify a lack of peroxisome proliferation as a possible limiting factor for methylotrophy via the XuMP cycle in *S. cerevisiae* [40]. In methylotrophic yeast like *P. pastoris*, peroxisome proliferation is central to methanol assimilation and occurs readily in the presence of methanol [52, 55]. Conversely, peroxisome proliferation in *S. cerevisiae* has evolved mainly in response to fatty acid oxidation [82]. Peroxisome proliferation in *S. cerevisiae* is under the control of the *PEX* gene family [83]. Deletion of *PEX30, PEX31* and *PEX32,* which all code for peroxisomal integral membrane proteins, has been shown to increase the size and number of mature peroxisomes in *S. cerevisiae* [84]. *PEX* genes are also responsible for the import of proteins across the peroxisomal membrane. This includes *PEX5* and *PEX7*, which are the two principal peroxisomal matrix protein import receptors, and are critical for peroxisome function [85]. This was tested in *S. cerevisiae* by Zhan et al. [42] by combining the expression of a XuMP cycle with the deletion of *PEX31* and *PEX32,* and the overexpression of *PEX5* before culture in liquid media with 0.5% methanol and 1% yeast extract [42]. This resulted in an observed increase in the size and number of peroxisomes.

An alternate scheme of the XuMP cycle has been reported in *P. pastoris* by Rußmayer et al. [55] where all the enzymes of the cycle, including the PPP enzymes for regeneration, are compartmentalised inside the peroxisome. This study used a multi-omics systems biology approach and concluded that regeneration of Xu5P does not occur through the cytosolic non-oxidative PPP in *P. pastoris*, but instead through a set of peroxisomal isoforms, Fba1-2p, Tal1-2p, Rpi1-2p, and Rpe1-2p. Rußmayer et al. [55] also suggest that this likely occurs in other methylotrophic yeasts. To test whether localisation of the whole XuMP cycle within the peroxisome of *S. cerevisiae* would confer advantage, Zhan et al. [42] expressed *AOX1-DAS1-DAK1* from *P. pastoris* in the peroxisome and overexpressed the native peroxisomal catalase *CTA1*. When this was combined with the

deletion of *PEX 31/32* and overexpression of *PEX5*, methanol consumption increased by 22% and biomass production increased by 25% compared to a control without these modifications [42]. This is the first study to report the manipulation of *PEX* genes to increase methanol assimilation through a synthetic XuMP cycle in *S. cerevisiae.* In our opinion, these novel strategies have merit from a rational design perspective and are worth considering for future XuMP cycle projects. This idea is further supported by the promising results from Wang et al. [44] in *Y. lipolytica*, where a chimeric XuMP/RuMP synthetic methanol assimilation pathway was engineered. Because *Y. lipolytica* is adapted to life in lipid rich environments, it possess a large number of peroxisomes for fatty acid oxidation [86]. It therefore stands to reason that modifications which increase peroxisome number and function in *S. cerevisiae* could also benefit XuMP cycle methanol assimilation.

When Zhan et al. [42] combined the above *PEX* gene modifications with peroxisomal compartmentalisation of the XuMP cycle, hybrid methanol oxidation via *BmMDH3,* and a number of other modifications, *S. cerevisiae* strain SM01-P9 was generated [42]. After culture in minimal medium with methanol as the sole carbon source this strain was able to reach a final OD_{600} of ~0.37 from a starting OD_{600} of 0.15 [42]. This reported growth on methanol as the sole carbon source in *S. cerevisiae* is interesting. However*,* as stated by Zhan et al. [42], the growth rate is too low to be industrially relevant. Furthermore, it is unclear to what extent this strain can grow on methanol after multiple rounds passages, and it is possible that it relies on reserve carbohydrates from the pre-culture. This study demonstrates just how difficult rational genome engineering of synthetic methylotrophy can be, and how even after numerous sequential modifications, an effective methylotrophic phenotype may not be established [42].

Adaptive laboratory evolution and synthetic methanol auxotrophy

Adaptive laboratory evolution can be a powerful tool for the development of a desired phenotype, including methylotrophy [87]. A full catalogue of methylotrophy associated mutations accumulated in bacteria during ALE experiments can be found in an excellent review from Wang et al. [7]. Following the discovery of a limited native capacity for methanol assimilation in *S. cerevisiae,* Espinosa et al. [41] conducted an ALE experiment to enhance growth on methanol [41]. Cultures were evolved in medium with 2% methanol and 0.1% yeast extract. After 230 generations a 44% increase in biomass was observed in the methanol cultures relative to the controls. Whole genome sequencing revealed a truncation in the *YGR076* transcription factor induced by a premature stop codon in all three methanol exposed lineages. CRISPR-Cas9 directed homologous recombination was then used to introduce a premature stop codon to *YGR076* in a parental strain. Remarkably this resulted in the same 44% increase in biomass observed in the evolved lineage, confirming that the truncation of

YGR076 was solely responsible for the increased growth [41]. This study provides a good example of how ALE and reverse genome engineering can be used to identify unknown genetic factors relating to growth on methanol *S. cerevisiae*.

Despite the benefits associated with using ALE to improve a methylotrophic phenotype, one significant downside to this strategy is that experiments must be run over a long period of time. And there is a risk that even after months or even years of ALE, a desired phenotype may still not arise. This can be addressed by using rational engineering to complement ALE and steer it in the right direction to increase the chances of success. The remainder of this section will discuss examples of this. After successful engineering of methanol assimilation in *Y. lipolytica* via a chimeric RuMP/XuMP methanol assimilation pathway, Wang et al. [44] conducted ALE to improve methanol assimilation. Engineered strains were passaged for 72 hours in YNB media with Complete Supplement Mixture (CSM) containing essential amino acids and 2% methanol. After 30 passages, a number of isolates were selected for further testing. Grown in YNB-CSM and 2% methanol, the best isolate was able to maintain a max OD_{600} of 0.4 for 72 hours from a starting OD_{600} of 0.05, which was a 54% improvement from the parental strain [44]. When the best isolates were grown in YNB without CSM and 2% methanol as the sole carbon source, starting at OD_{600} 0.06 they were able to maintain an OD_{600} of 0.05 for a long culture of >200 hours. Whereas the parental strains OD_{600} decreased during the same long culture with methanol as the sole carbon source. Post ALE, the best evolved isolate was able to consume methanol at 1.1 g/L/72 hours, a 261% increase compared to the parental strain. This project provides a very good example of how ALE can be used to enhance methylotrophy in an engineered strain of yeast. While the growth reported is not industrially relevant due to the requirement for amino acid supplementation, the fact that this evolved *Y. lipolytica* can maintain cell growth on methanol as the sole carbon source in a long culture is a significant achievement and paves the way for further advances in yeast synthetic methylotrophy [44]. This same strategy may achieve similar results if it was employed in *S. cerevisiae*.

The previously mentioned SM01-P9 strain of *S. cerevisiae* from Zhan et al. [42] which contained *PEX* gene modifications, peroxisomal compartmentalisation of the XuMP cycle, and hybrid methanol oxidation was subjected to ALE to improve its growth on methanol as the sole carbon source. Instead of using yeast extract as per the previous *S. cerevisiae* methanol ALE study from Espinosa et al. [41], Zhan et al. [42] used xylose, a pentose which is metabolised through the PPP. *S. cerevisiae* does not have a native capacity for xylose utilisation, so a previously reported xylose utilisation cycle [88] was first introduced. ALE began in minimal medium containing 1% methanol and 1% xylose after which xylose was slowly reduced to 0.2% [42]. After 13 months of ALE, biomass

production in xylose and methanol increased 287 % compared to the parental strain, achieving a final OD_{600} of 1.86 from a starting OD_{600} of 0.2 after 144 hours. When grown in minimal media with methanol as the sole carbon source, the evolved strain (CX01F) reached a final OD_{600} of ~0.75 from a starting OD_{600} of 0.2 after 144 hours, just under two doublings. This was an improvement on the parental strain which could only achieve 1 doubling. ^{13}C-methanol tracer analysis in CX01F revealed that 15 different amino acids had ^{13}C labelled fractions of 87% or higher. And 100% of the critical glycolytic intermediates G3P and DHAP were ^{13}C-labeled. This appeared to confirm that methanol was being metabolised [42]. The use of xylose as a co-carbon substrate instead of yeast extract is a strategy that appears to have merit for ALE of methylotrophy in *S. cerevisiae*. This is because xylose assimilation requires high carbon flux through the PPP which in turn boosts RuMP or XuMP cycling and speeds up regeneration of Ru5p and Xu5P for formaldehyde fixation. Despite these results, Zhan et al. [42] state that the reported growth of the evolved strain of *S. cerevisiae* is still too low to be industrially useful.

Despite the successful engineering of partial methylotrophy in *E. coli* [74, 76, 89, 90], *Pseudomonas putida* [91], *Corynebacterium glutamicum* [92, 93], and *S. cerevisiae* [39, 40], effective growth on methanol as the sole carbon source was not achieved in any of these studies. And the engineered strains still required supplementary carbon from low concentration yeast extract to support growth on methanol. These results show that engineered cells with the capacity for synthetic methanol assimilation will still preferentially use the supplemental carbon from yeast extract through a normally functioning central carbon metabolism. To increase the selective pressure for engineered cells to assimilate methanol, Chen et al. [25, 94] engineered synthetic methanol auxotrophy in *E. coli* by disrupting the PPP during methanol/pentose co-substrate ALE. In this strategy, the non-oxidative PPP is disrupted to prevent pentose assimilation and inhibit cell growth when cells are grown solely on xylose [7]. Cell growth on xylose can then be restored by the addition of a synthetic RuMP cycle (Mdh, Hps, Phi)(Figure 6A). Importantly, the RuMP cycle requires methanol to function, meaning that the engineered cell will not grow on xylose without methanol as a co-substrate. This strategy is what ultimately led Chen et al. [25] to successfully engineer robust synthetic methylotrophy in *E. coli*.

In their breakthrough work, Chen et al. [94] began by engineering synthetic methanol auxotrophy in *E. coli,* making growth on xylose dependent on methanol assimilation. This was achieved by disrupting the PPP through *RPIA* and *RPIB* gene knockouts which code for ribose-5-phosphate isomerase (Rpi), inhibiting the production of R5P which is essential for nucleotide biosynthesis and cell growth (Figure 6A). To restore R5P production and cell growth, a synthetic RuMP cycle was introduced which generates R5P during the rearrangement of the formaldehyde acceptor. This made

cell growth on xylose dependent on a functioning RuMP cycle and methanol [94]. Genes from the non-oxidative PPP were also introduced to improve regeneration of the Ru5P formaldehyde acceptor, including transaldolase from *Klebsiella pneumoniae* and transketolase from *Methylococcus capsulatus*. A metabolic flux model known as EMRA (Ensemble Modelling for Robustness Analysis) [95] was also used to identify that the downregulation of Phosphofructokinase (Pfk) and Glyceraldehyde 3-phosphate dehydrogenase (Gapdh) was essential for the regeneration of RuMP pathway intermediates and successful ALE [25]. This was done by deleting *PFKA* and replacing *GAPA* with *GAPC* from *E. coli* BL21. The engineered methanol auxotrophic strain was then evolved in xylose and methanol media for 20 generations before *RPIA/B* were reintroduced to improve fitness [25]. This allowed the strain to grow in minimal liquid media with just methanol and High-Def Azure (HDA) essential amino acid mixture and no other carbon source. Gradual weaning off HDA over 180 days and 21 passages lead to a strain that could use methanol as its sole carbon source. After further ALE and strain selection, the best isolate achieved optimal growth on 400mM methanol growing from OD_{600} 0.1 to 1.0 in 30 hours with a doubling time of 8 hours. This is comparable to *B. methanolicus* when grown solely on methanol [25]. This research series from Chen et al. [25, 94] is a very significant achievement and is the first example of effective synthetic methylotrophy in a platform industrial species.

It also provides a blueprint for how effective growth on methanol as the sole carbon source may be achieved in *S. cerevisiae.* More recently, synthetic methanol auxotrophy has also been engineered in *B. subtilis* via the same *RPIA/B* knockouts [96]. So far, synthetic methanol auxotrophy has only been demonstrated in bacteria using the RuMP cycle [94, 96, 97]. However, in *S. cerevisiae* it may also be possible to engineer methanol dependent growth using the XuMP cycle (Figure 6B). This could be done via the introduction of a xylose utilisation pathway, XuMP cycle and Rpi gene (*RKI1*) knockout. If this is possible, it may also mean that synthetic methanol auxotrophy could function in a cell with dual methanol assimilation capabilities as in the strain of *Y. lipolytica* which was engineered with both a XuMP and RuMP cycle [44].

Further strategies to increase formaldehyde assimilation

Careful balancing of formaldehyde flux between detoxification and assimilation cycles is important for any methylotroph [68, 71]. If too much formaldehyde is detoxified into CO_2 then less carbon will be available for assimilation through a RuMP or XuMP cycle. On the other hand, if too much formaldehyde is accumulated, DNA-protein crosslinking will occur and reduce fitness [67]. For example, the deletion of the formaldehyde dissimilation genes *ALD* for acetaldehyde dehydrogenase (Ald) and *FADH* for formaldehyde dehydrogenase (Fadh) in *C. glutamicum*, resulted in higher ^{13}C-

methanol assimilation through a synthetic RuMP cycle but also caused a toxic build-up of formaldehyde which reduced cell growth [92]. In *S. cerevisiae*, overexpression of the native Fadh gene (*SFA1*) alongside a synthetic RuMP cycle resulted in significantly more growth on solid methanol media compared to a RuMP strain without overexpression of *SFA1* [40]. However, due to the dissimilatory function of *SFA1*, it is likely that the observed growth increase on methanol was related to increased activity of the formaldehyde detoxification pathway and not an increase methanol assimilation. In this case, more carbon from formaldehyde is being lost through detoxification in CO_2 instead of being assimilated through the RuMP cycle. Therefore, it is important to consider that improved growth in the presence of methanol does not necessarily equate to improved methanol assimilation.

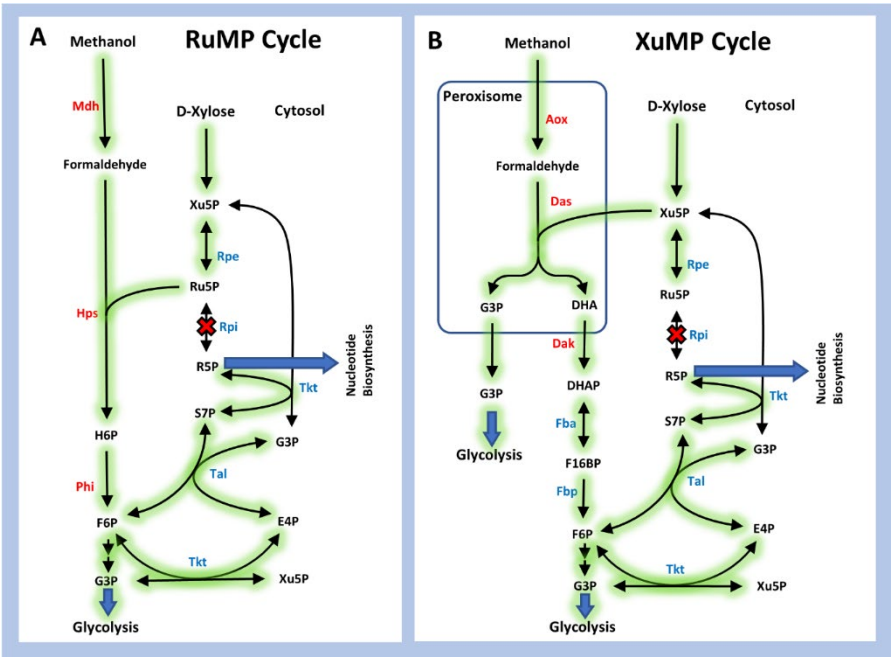

Figure 6. Possible Route for Engineering Synthetic Methanol Auxotrophy in S. cerevisiae via the RuMP and XuMP Cycles. Engineering synthetic methanol auxotrophy in *S. cerevisiae* would be a good starting point for ALE towards robust synthetic methylotrophy. **A:** Synthetic methanol auxotrophy has been demonstrated in *E. coli* using the RuMP cycle via a Rpi knockout with growth on xylose and methanol [25]. This diagram has been adapted from Chen et al. [94]. Arrows shown in green relate to methanol and xylose assimilation. Knocking out Rpi during growth on xylose prevents the production of R5P which inhibits nucleotide biosynthesis and cell growth. R5P production and cell growth can be restored through the RuMP pathway which requires methanol assimilation to function. This increases the selective pressure for cells to assimilate methanol. It may be possible that synthetic methanol auxotrophy could be engineered in *S. cerevisiae* in a similar way using a synthetic xylose assimilation pathway, RuMP cycle and Rpi knockout. **B:** It may also be possible to engineer synthetic methanol auxotrophy in *S. cerevisiae* using a synthetic xylose assimilation pathway, XuMP cycle and Rpi knockout as outlined in this figure. **Mdh,** methanol dehydrogenase; **Aox**, alcohol oxidase; **Das**, dihydroxyacetone synthase; **Dak**, dihydroxyacetone kinase; **Hps**, 3-hexulose-6-phosphate synthase; **Phi**, phosphohexuloisomerase; **Fba** fructose-bisphosphate aldolase; **Fbp**, fructose bisphosphatase; **Tkt**, transketolase; **Tal**, transaldolase; **Rpi**, ribose-5-phosphate isomerase; **Rpe** ribulose phosphate 3-epimerase; **G3P**, glyceraldehyde 3-phosphate; **DHA**, dihydroxyacetone; **DHAP**, dihydroxyacetone phosphate; **F16BP**, fructose-1 6-bisphposphate; **H6P**, hexulose 6-phosphate; **F6P**, fructose 6 phosphate; **E4P**, erythrose-4-phosphate; **S7P**, sedoheptulose 7-phosphate; **Xu5P**, xylulose 5-phosphate; **R5P**, ribose 5-phosphate; **Ru5P**, ribulose 5-phosphate.

In the evolved synthetic methylotrophic strain of *E. coli* engineered by Chen et al. [25], whole genome sequencing revealed a 4-bp insertion causing inactivation of the Fadh gene (*FRMA)*. Interestingly, inactivation of Fadh through gene knockouts or ALE mutations has accompanied improved growth

on methanol in numerous separate studies in bacteria [76, 89, 90, 94, 97]. More recently in the yeast study using *Y. lipolytica,* inactivation of the Fadh gene (*FLD1*) was also reported in the evolved strain [44]. Taken together, these results indicate that directing formaldehyde away from dissimilatory processes boosts the amount of formaldehyde available for assimilation and is beneficial for synthetic methylotrophy. This could be tested in *S. cerevisiae* by knocking out the gene for Fadh (*SFA1*) before ALE of a strain with a synthetic RuMP/XuMP cycle. Doing so would introduce a strong selective pressure for cells to improve formaldehyde assimilation in order to avoid a toxic build-up. If cells were unable to cope with a full *SFA1* knockout, then a weak-promoter expression cassette could be re-introduced to provide just enough formaldehyde detoxification for survival while still maintaining pressure for cells to improve fitness through an efficient synthetic assimilation cycle. The key aim of this strategy is to engineer a methanol assimilating cell with as little *SFA1* activity as possible, preferably none.

As more formaldehyde becomes available in an Fadh deficient cell, it must be assimilated more quickly through a RuMP or XuMP cycle to avoid DNA protein crosslinking. Apart from the strategies already discussed in this review, boosting regeneration of the formaldehyde acceptor may be useful. As previously discussed, this strategy was employed in the synthetic methylotrophic strain of *E. coli* engineered by Chen et al. [25] using exogenous transaldolase from *K. pneumoniae* and transketolase from *M. capsulatus*. In the recent work in *Y. lypolitica* from Wang et al. [44], overexpression of native regeneration enzymes *TKL1*, *PFK*, *FBA* and *RPE1* as well as heterologous expression of *GLPX* from *B. methanolicus* was employed. In doing so, they overexpressed all the genes required to convert G3P and F6P into the RuMP/XuMP cycle formaldehyde acceptors Ru5p and Xu5P [44]. This strategy could be transferred to *S. cerevisiae* where it may yield similar promising results.

Completing the reductive glycine pathway

As discussed previously, the conversion of formate to glycine through the core module of the rGly pathway has been demonstrated in *S. cerevisiae* by Gonzalez de la Cruz et al [72]. The next step towards growth on C1 substrates in *S. cerevisiae* via this strategy requires an efficient system for the aerobic conversion of glycine into biomass. The simplest and most energetically efficient route for this occurs via a two-step conversion of glycine into pyruvate via serine [64]. In this process, serine hydroxymethyltransferase (Shmt) catalyses the conversion of glycine into serine which is then converted into pyruvate by serine deaminase (Sds). In the strain of *E.coli* engineered by Kim et al. [72] to grow on formate as the sole carbon source, overexpression of native Shmt and Sds was enough to enable conversion of glycine into pyruvate for cell growth on formate. This approach was also used

in a study which engineered a synthetic rGly pathway in *C. necator,* where the two native genes for Shmt and Sds expressed on a medium strength promoter enabled growth on glycine as the sole carbon source [98]. It would be logical to test this approach in *S. cerevisiae* where overexpression of native Shmt and Sds may yield similar results.

Growth on methanol may also be possible in *S. cerevisiae* through the rGly pathway by the inclusion of enzymes for methanol oxidation. In the formate assimilating strain of *E. coli* from Kim et al. [72], growth on methanol was also tested through the rGly pathway via the addition of Mdh from *B. stearothemophilus*. The addition of Mdh meant that engineered cells were able to oxidise methanol to formaldehyde which could then be oxidised to formate via endogenous formaldehyde dehydrogenase (Fadh). This increased the doubling time considerably, from 8 hours when grown on formate to around 54 hours on methanol. This reduced growth on methanol when compared to formate was suggested to be due to the low activity of the Mdh, which meant that less carbon was available for biomass generation through the rGly pathway [72]. Despite this limited growth, these results are significant as a proof of concept to demonstrate that cell maintenance can be achieved in *E. coli* using methanol as the sole carbon source through a synthetic rGly pathway. Therefore, it stands to reason that *S. cerevisiae* might be engineered to grow on methanol through the rGly pathway in the same way. This would require the expression of a NAD-Mdh or an Aox or possibly both together utilising the hybrid methanol oxidation strategy discussed previously. Use of hybrid methanol assimilation may also alleviate the problems associated with the low activity of *Bs*Mdh encountered in the *E. coli* study [72]. Another possible bottleneck in this strategy may relate to the conversion of formaldehyde to formate by the native *S. cerevisiae* Fadh. This may be addressed by overexpressing the native Fadh from *S. cerevisiae* or expressing a heterologous Fadh from a native methylotroph which is more efficient at oxidising formaldehyde to formate, for example, Fadh from *P. pastoris*.

Literature review concluding remarks

To date, there have been four research groups which have reported work on synthetic methylotrophy in *S. cerevisiae*, with only a handful of publication and preprints available (Table 1). These efforts have provided a valuable foundation of knowledge. It has been shown that synthetic methanol assimilation in *S. cerevisiae* can be achieved through a synthetic XuMP cycle [39], a synthetic RuMP cycle and a hybrid RuMP/XuMP cycle [40, 42]. A chimeric RuMP/XuMP cycle for dual formaldehyde assimilation has also been demonstrated in *Y. lipolytica*, which is relevant to *S. cerevisiae* [44]. Additionally, the core module of the rGly pathway has been engineered in *S. cerevisiae* to convert formate to glycine [43]. From these works, it has been established that a synthetic methanol

assimilation cycle on its own is not sufficient to enable full synthetic methylotrophy with growth that is industrially relevant in *S. cerevisiae* [40]. Therefore, any engineered strain will require further modification and/or ALE to achieve this.

This review has assessed several strategies to enhance synthetic methanol assimilation which are worth considering for the future engineering of industrially relevant methylotrophy in *S. cerevisiae*. These include: careful screening and selection of appropriate methylotrophy genes; *PEX* gene modifications and peroxisomal compartmentalisation to enhance XuMP cycle methylotrophy; hybrid methanol oxidation through mixed expression of an Aox and/or NAD-Mdh; dual formaldehyde assimilation through combined expression of XuMP and RuMP cycles; increasing available formaldehyde through targeted knockout of dissimilatory enzymes; speeding up regeneration of the formaldehyde acceptor through engineering of the PPP; ALE; and engineering the rGly pathway. Despite the progress made so far [39, 40, 42, 44 43], only limited growth of *S. cerevisiae* on methanol as the sole carbon source has been achieved. This demonstrates just how difficult it is to engineer synthetic methylotrophy in *S. cerevisiae,* even after exhaustive modification and ALE.

To overcome this challenge, it may be beneficial to look towards the recent success in enabling effective synthetic methylotrophy in *E. coli* [25]. This was achieved by first engineering synthetic methanol auxotrophy, making growth on xylose dependant on a functioning RuMP cycle which requires methanol, and increasing the selective pressure on cells to assimilate methanol during co-carbon substrate ALE [94]. By imposing selection pressure for methanol assimilation, the authors were able to set the appropriate conditions for the methylotrophic phenotype to arise after being slowly weaned off xylose [25]. Importantly, this breakthrough in *E. coli* demonstrates that rational engineering can be effectively employed as a tool to enhance the power of ALE from the start of a project. It is important to note that previous experiments in *S. cerevisiae* which have employed standard modes of ALE have achieved limited success [41, 42]. Therefore, the next logical step is to implement synthetic methanol auxotrophy which will then be followed by ALE on xylose and methanol. As well as using a RuMP cycle, it may also be possible to use a XuMP cycle to engineer synthetic methanol auxotrophy in *S. cerevisiae* which would provide additional engineering options for future experiments. The successful engineering of the core module of the rGly pathway in *S. cerevisiae* also presents a possible route towards synthetic methylotrophy [43]. The next steps in this strategy would require additional modifications to allow for the efficient oxidation of methanol to formate to feed the pathway, and the efficient conversion of glycine into pyruvate for biomass generation.

Table 1. A summary of the work so far in engineering synthetic methylotrophy in yeast. Including studies in *S. cerevisiae* by four separate research groups and one group in *Y. lipolytica.*

Research group	Species	Cycle	Variant	Culture conditions/carbon source	Results/notes
Dai et al., 2017	*S. cerevisiae*	**RuMP**	***MDH*** – *B. methanolicus* (MGA3) ***HPS*** – *B. subtilis* ***PHI*** – *B. subtilis*	Minimal liquid medium with 10 g/L methanol as the sole carbon source. 72 hours, starting OD_{600} 7.5.	No cell growth. Decreased biomass. No methanol consumption.
	S. cerevisiae	**XuMP**	***AOX1*** – *P. pastoris* ***CTA1*** – *P. pastoris* ***DAS2*** – *P. pastoris* ***DAK1*** – *P. pastoris*	Minimal liquid medium with 10g/L methanol as the sole carbon source. 72 hours, starting OD_{600} 7.5.	3.31% increase in OD_{600} compared to a 5.7% decrease in the control. Methanol consumption of 1.04 g/L per 72 hours compared to none in the control.
				Minimal liquid medium with 10g/L methanol and 1g/L yeast extract. 72 hours, starting OD_{600} 7.12.	11.7% increase in OD_{600} compared to a 3.8% increase in the control. Methanol consumption of 2.35 g/L per 72 hours.
Espinosa et al., 2019b	*S. cerevisiae*	**RuMP**	***MDH*** – *B. stearothermophilus* ***HPS*** – *B. methanolicus* (MGA3) ***PHI*** – *B. methanolicus* (MGA3)	Minimal solid media with 1%, 2% and 4% methanol as the sole carbon source. 120 hours.	Growth spot assays showed a visible increase in growth at all concentrations compared to a negative control strain.
				Minimal liquid medium with 0.1% yeast extract and 2% methanol. 120 hours.	No significant difference in growth compared to negative control.
			MDH – *B. stearothermophilus* ***HPS*** – *B. methanolicus* (MGA3) ***PHI*** – *B. methanolicus* (MGA3) Overexpression: ***TAL1***	Minimal solid media with 1%, 2% and 4% methanol. 120 hours.	Growth spot assays showed a visible increase in growth at all methanol concentrations when compared to negative control. Superior growth compared to a RuMP strain without *TAL1* overexpression. Superior growth to RuMP strain with *TKL1* overexpression.
			MDH – *B. stearothermophilus* ***HPS*** – *B. methanolicus* (MGA3) ***PHI*** – *B. methanolicus* (MGA3) Overexpression: ***TKL1***	Minimal solid media with 1%, 2% and 4% methanol.	Growth spot assays showed a visible increase in growth at all methanol concentrations when compared to negative control and had superior growth compared to a RuMP strain without *TKL1* overexpression.
		XuMP	***AOX1*** – *P. pastoris* ***DAS1*** – *P. pastoris*	Minimal solid media with 1%, 2% and 4% methanol.	Marginal increase in visible growth at 1% methanol but no difference at 2% or 4% methanol when compared to negative control. Inferior growth to all RuMP strains.
		Hybrid RuMP/XuMP	***MDH*** – *B. stearothermophilus* ***DAS1*** – *P. pastoris*	Minimal solid media with 1%, 2% and 4% methanol.	Growth spot assays showed a visible increase in growth at all concentrations compared to a negative control strain.
Zhan et al., 2021, (pre-print)	*S. cerevisiae*	**XuMP**	***AOX1*** – *P. pastoris* ***DAS1*** – *P. pastoris* ***DAK2*** – *P. pastoris* ***CDC19*** – *P. pastoris* ***MDH3*** – *B. methanolicus* (MGA3) ***PEX31Δ*** ***PEX32Δ*** Overexpression: ***PEX5*** ***MDH2*** (malate dehydrogenase 2) ***MDH3*** (malate dehydrogenase 3) ***FBA1/2*** 13 months of ALE in 0.2% xylose and 1% methanol.	Minimal liquid medium with 0.2% xylose and 1% methanol. 144 hours, starting OD_{600} 0.2	Increase in biomass to final OD_{600} 1.86. Over three doublings. Compared to the parental strain which could only achieve final OD_{600} of 0.48.
				Minimal liquid medium with 1% ^{13}C- methanol as the sole carbon source. 144 hours, starting OD_{600} 0.2	Increase in biomass to final OD_{600} ~0.75. Just under two doublings. 100% of G3P and DHAP labeled with ^{13}C. Compared to the parental strain which could only reach a final OD_{600} of 0.37, just under one doubling.
Wang et al., 2021	*Y. lipolytica*	**Chimeric RuMP/XuMP**	***MDH*** – *B. stearothermophilus* (4 copies) ***HPS*** – *B. methanolicus* (MGA3) ***PHI*** – *B. methanolicus* (MGA3) ***AOX1*** – *P. pastoris* ***DAS1*** – *P. pastoris* ***DAK2*** – *P. pastoris* ***GLPXC*** – *B. methanolicus* (MGA3) ***FLD1Δ*** Overexpression: ***TKL1*** ***FBA*** ***PFK*** ***RPE1*** 3 months of ALE on Complete Supplement Mixture (CSM) essential amino acid supplement and 2% methanol.	Minimal liquid media with CSM and 2% methanol. 72 hours, starting OD_{600} 0.01.	Increase in biomass to final OD_{600} to 0.25. Just over 4 doublings. Compared to parental strain which could not maintain and cell growth.
				Minimal liquid media with 2% methanol as the sole carbon source. 200 hours, starting OD_{600} 0.06.	Able to maintain OD_{600} of 0.05 for 200 hours. Compared to a parental strain which could not maintain and growth.

Methods

Experiment 1

Cloning

This experiment was conducted in *S. cerevisiae* CEN.PK2-1C (*URA3Δ, HIS3Δ, LEU2Δ* and *TRP1Δ*) [99]. Cloning started with a pRS416 shuttle vector that contained a synthetic RuMP cycle which was constructed by Espinosa et al [40] (*pTEF1-BmHPS-ADH1t-pTEF2-BmPHI-tPHO5-pTDH3-BsMDH-tCYC1t*-pRS416) (Figure 7, Supplementary Table 3). The pRS416 vector contains a *URA3* marker for minus uracil (-URA) selection in in CEN.PK2-1C and an *AMPR* marker for ampicillin selection in *E. coli*.

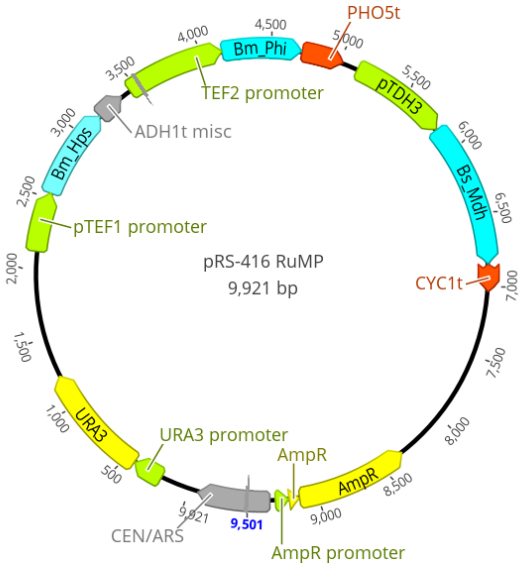

Figure 7. Plasmid map of the starting vector used for cloning. This pRS416 vector containing a synthetic RuMP cycle was constructed by Espinosa et al. [40] and used as a starting point for cloning in this study. RuMP cycle genes are shown in blue and include *BsMDH*, *BmHPS* and *BmPHI*. Gene blocks for the Mdhs tested were cloned in to replace *BsMDH*. Extracted from Geneious Prime.

A search of the literature identified ten enzymes of interest which could be tested to oxidise methanol as part of the RuMP cycle. These included NAD-Mdhs from *B. methanolicus* strains MGA3 and PB1[58, 74], *B. stearothermophilus*[76], *C. necator*[79] and *L. xylanilyticus*[80], as well as *AOX1* from *P. pastoris*[100] (Table 2). Nucleotide sequences for the methanol oxidation genes tested were codon optimised for *S. cerevisiae* in Geneious Prime and ordered as Gene blocks from Twist Bioscience in San Francisco (Supplementary table 1).

Table 2. List of RuMP pathway Variants which were tested in this study. The majority of the methanol oxidation genes tested in this experiment encoded NAD-Mdhs from different native methylotrophic bacteria. *AOX1* from *P. pastoris* was also tested. Each of the test genes were included in a synthetic RuMP cycle which used *HPS* and *PHI* from *B. methanolicus* MGA3.

Experiment 1: RuMP cycle variants and controls tested in *S. cerevisiae* CEN.PK2-1C
B. methanolicus MGA3 *MDH-BmHPS-BmPHI*-pRS416

B. methanolicus MGA3 *MDH2-BmHPS-BmPHI*-pRS416
B. methanolicus MGA3 *MDH3-BmHPS-BmPHI*-pRS416
B. methanolicus PB1 *MDH-BmHPS-BmPHI*-pRS416
B. methanolicus PB1 *MDH1-BmHPS-BmPHI*-pRS416
B. methanolicus PB1 *MDH2-BmHPS-BmPHI*-pRS416
B. stearothemophilus MDH-BmHPS-BmPHI-pRS416
C. necator MDH2-BmHPS-BmPHI-pRS416
L. xylanilyticus MDH-BmHPS-BmPHI-pRS416
P. pastoris AOX-BmHPS-BmPHI-pRS416
BmHPS-BmPHI (RuMP cycle minus Mdh control)
pRS-416 (empty vector control)

The pRS416 RuMP cycle vector was linearised via PCR using primers which excluded the ORF for *BsMDH* (PK0003, PK0004)(Figure 8A). Gene blocks were amplified with primers that had 5' extensions encoding 20 bp homologous overlaps for insertion into the backbone via Gibson assembly (Figure 8B, Supplementary table 2). Primers were ordered from Integrated DNA

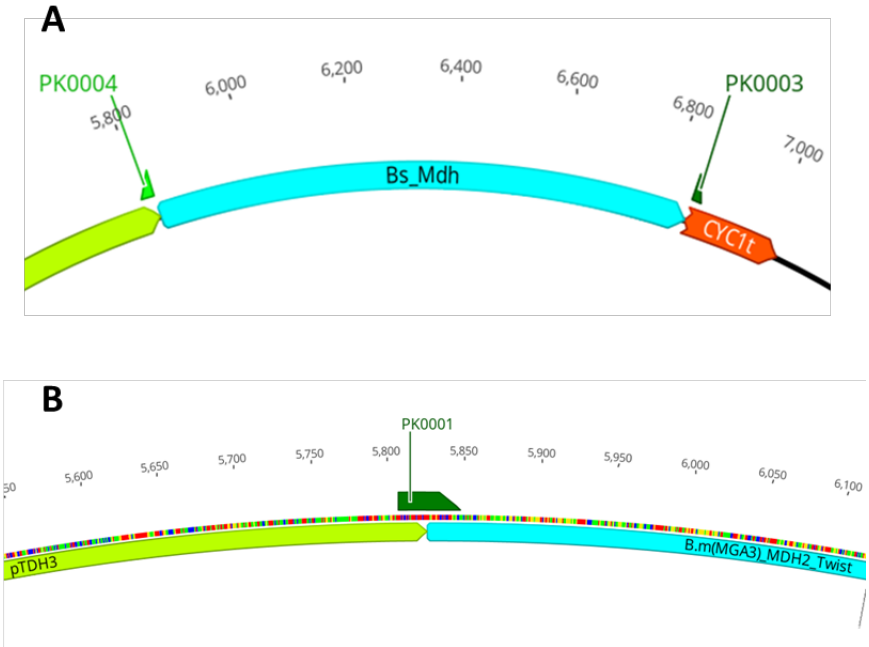

Figure 8. Primer design for cloning of methanol oxidation enzymes into pRS-416 RuMP. A. Cloning began by first linearising the pRS-416 RuMP vector with PK0004 and PK0003 primers to give a backbone that excluded the *BsMDH* ORF. **B.** Gene blocks were amplified concurrently using primers which were unique to each insert that had a 20 bp overlap with the ends of the backbone for subsequent Gibson Assembly of inserts and backbone (Supplementary table 2). Extracted from Geneious Prime

Technologies (IDT) in Singapore. All PCR reactions were conducted using an Eppendorf Vapo-protect 96 well thermocycler. Reactions contained 10 µL of 5x HF reaction buffer from New England Biolabs (NEB), 1 µL of 10 mM DNTPs from NEB, 1 µL of each primer, 0.5 µL of Phusion

High-fidelity DNA Polymerase (Phusion) from NEB and 35.5 µL of sterile MilliQ water for a final reaction volume of 50 µL. PCR products were digested with 1 µL of *Dpn*I restriction enzyme (NEB) per 50 µL at 37 °C for one hour before PCR clean-up with a QIAquick PCR Purification Kit from Qiagen as per the manufacturer's instructions.

Gel electrophoresis and imaging of PCR products was used to confirm correct band sizes for each of the inserts and backbone using 1% agarose, 1x Tris Acetate, EDTA (TAE) gels run at 100 mVs for one hour (Figure 9).

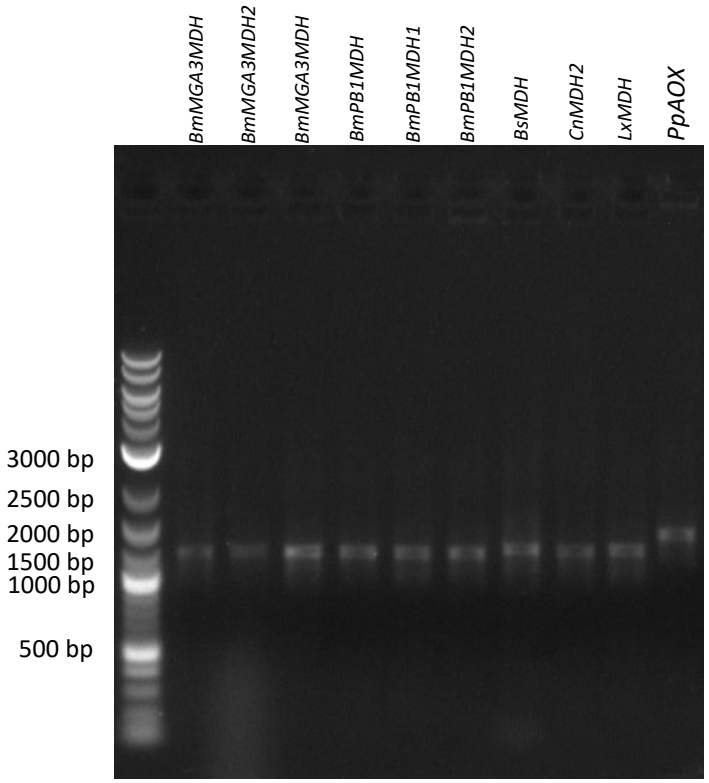

Figure 9. Amplified methanol oxidation gene blocks. PCR products were cleaned up using a Quiagen PCR clean up kit then run on a 1x TAE, 1% agarose gel to confirm correct band sizes for each insert to be used for cloning.

Gibson assembly was done using NEBuilder HiFi DNA Assembly Master Mix (NEB). Inserts and backbone were Gibson assembled in reactions containing a total of 0.02 pmlos of assembly DNA at a ratio of 2:1 insert to backbone, 10 µL of HiFi DNA Assembly Master Mix and sterile MilliQ water to a final volume of 20 µL. Gibson reactions were incubated in a thermocycler at 50° C for 3 hours. Gibson assembled vectors were then transformed in *E. coli.* 50 µL aliquots of chemically competent *E. coli* DH10B cells were taken out of -80° C storage and thawed on ice for 30 minutes. 5 µL of each Gibson assembly product was added to a tube of thawed cells and flicked gently five times to mix. Mixtures were kept on ice for 30 minutes then heat shocked in a 42° C water bath for 45 seconds. Mixtures were then placed back on ice for 2 minutes before 200 µL of 1x Luria Broth (LB) (Sigma Aldrich) growth medium was added to each tube and outgrown in a shaking incubator

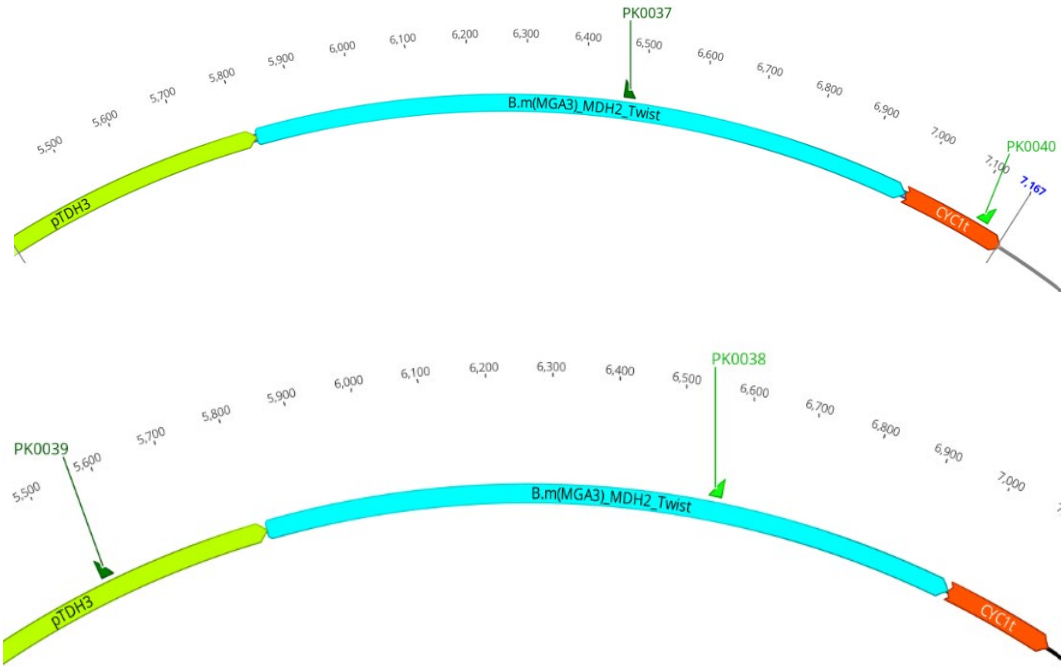

Figure 10. Primer design for diagnostic *E. coli* colony PCR. Primer pairs were designed that spanned the junction of each insert to confirm that correct insertion had taken place during Gibson assembly. This gave two unique PCR products for each insert which could be visualised and confirmed in 1% agarose gels during electrophoresis.

at 37° C, 240 rpm for 1 hour. After outgrowth, the total volume of the transformation mixtures were plated out onto selective LB, 100 µg/mL ampicillin plates and incubated at 37° C overnight. Transformants were picked randomly by hand and subject to diagnostic colony PCR. Two primer sets were designed for each insert spanning the 5' and 3' insert junctions (Figure 10, Supplementary table 2). A small mass of cells was picked off the plate with a pipette tip and resuspended in 50 µL of sterile MilliQ water in 150 µL PCR tubes. Cell suspensions were incubated in a thermocycler at 95° C for 10 minutes to lyse the cells. Lysates were spun at 3000 x g at 4° C for 10 minutes. 1 µL of supernatant was used as the template DNA for the PCR reactions. PCRs were conducted using the same protocol as outlined for the gene blocks and backbone however were scaled down to 10 µL. Diagnostic PCR products were run out in 1% agarose 1x TAE buffer gels at 100 mVs for 1 hour. Successful assemblies were indicated by two bands of unique size for each variant which spanned the junctions of each insert. Successful *E. coli* transformants were grown overnight in 1x LB, 100 µg/mL ampicillin liquid media. A Quiagen plasmid miniprep extraction was conducted to extract each of the newly constructed RuMP cycle variants for yeast transformation. Glycerol stocks were also taken for later use by adding 40% glycerol in a 1:1 ratio to cells and stored at -80° C.

Yeast plasmid transformations

500 µL of overnight CEN.PK2-1C culture in 1x Yeast-extract Peptone Dextrose (YPD), 2% glucose medium, was aliquoted into a 1.5 mL microcentrifuge tube for each plasmid to be transformed. Cells were spun at 13,000 x g and supernatant removed. The transformation mixture was then added to each tube and contained 48 µL polyethylene glycol 3350 50% w/v (Sigma Aldrich), 7 µL 1 M

lithium acetate (Sigma Aldrich), 10 µL boiled single stranded carrier DNA from herring sperm (Sigma Aldrich), and 1 µg in solution of each RuMP cycle variant in its respective tube. Pellets were resuspended in the transformation mixture and tubes were incubated in a 42° C water bath for one hour. After incubation, tubes were spun at 4000 x g for 30 seconds and the supernatant containing the transformation mixture was removed. Cells were then resuspended in 500 µL sterile MilliQ water before 50 µL of each cells suspension was plated out onto 1x Yeast Nitrogen Base without amino acids (YNB), 2% glucose, 2% agarose, -URA plates and placed in a 30 °C incubator for two days then stored at 4° C.

Spot assays

To test whether the different methanol oxidation enzymes could confer differential growth on solid media with methanol as the sole carbon source, a series of spot assay were conducted. Three biological replicates (individual colonies) of each variant were grown in an overnight culture of 1x YNB, 2% glucose, -URA. Precultures were then inoculated in 1x YNB, 2% glucose, -URA at OD_{600} 0.5 and grown at 30° C 200 rpm for 3 hours to exponential phase. Cells were spun and washed twice in sterile water to remove residual glucose from the preculture. Cell suspensions in water at OD_{600} 1 were made for each variant from which serial tenfold dilutions were made down to 10^{-4}. 3 µL of each dilution of each replicate was spotted out on rectangular Plus Plates (Singer Instruments) which contained either 1x YNB, -URA with no additional carbon source; 1x YNB, 2% glucose, -URA; 1x YNB, 2% methanol, -URA; and 1x YNB, 4% methanol, -URA. Plates were incubated at 30° C for five days and imaged in a Phenobooth (Singer Instruments).

Liquid growth experiments

The same strains from the spot assays were grown in a preculture of 1x YNB, 0.25% glucose, 2% methanol, -URA for 48 hours at 30° C 800 rpm in a Greiner Bio CellStar 96 well clear flat bottom plate vessel in an Infors MultiTron Pro shaking incubator with a 3mm orbital. This was followed by growth tests in a Greiner Bio CellStar 96 well clear flat bottom plate where 100 µL of the same medium as the preculture was inoculated from the precultures to a starting OD_{600} of 0.015. 8 replicates of each variant were included which fit conveniently onto a single 96 well plate. Tests were run in an Agilent Synergy H1 Biotek microplate reader for 45 hours. The plate was incubated at 30° C and shaken in double orbital continuous shake mode at 800 rpm for the duration. Optical densities were measured at 600 nm every 20 minutes throughout. To prevent evaporation of wells during the experiment, 150 µL of sterile MilliQ water was added to each void space in between the wells of the plate and the edge of the plate was wrapped four times in Parafilm.

Data analysis

Growth curves of the 8 replicants for each variant were averaged. Mean maximum OD_{600} values were then recorded for each variant. Differences in mean maximum OD_{600} were first tested with a one-way ANOVA in the GraphPad Prism software package. This was followed by a Dunnett's multiple comparison of all the test variants against the RuMP -Mdh control. The Dunnett's test gives and adjusted P value to reduce the probability of making a type-1 error during analysis. The significance limit was $P < 0.05$.

Experiment 2

RAIN library design

The promoter library contained five different promoters, each with different strengths and characteristics. These included: p*ADH2* which is glucose repressed/alcohol inducible; p*CYC1* which is weak and low glucose inducible; p*PDA1* which is weak and constitutive; p*SSA1* which is alcohol inducible; and p*SSB1* which is also alcohol inducible[101] (Table 3). The ORF library contained a variety of genes from the RuMP and XuMP cycles from different organisms and from the PPP of *S. cerevisiae*. These included *DAS1, DAS2 and DAK* from *P. pastoris*, *HPS and PHI* from *Mycobacterium gastri, HPS* from *Methylococcaceae buryatense* and *HPS* and *PHI* from *B. methanolicus* (Table 4). Native *S. cerevisiae* genes from the PPP included *TKL1, TKL2, GND1, GND2, ZWF1, TAL1, NQM1*, *RKI1, SOL3,* and *SOL4*.

The RAIN promoter cassettes included a consensus *Ty1* homology arm on the 5' end of the cassette followed by the promoter sequence followed by the *YDL191* intron sequence on the 3' end.
The ORF cassettes included the *YDL191* intron on the 5' end followed by the ORF sequence followed by the *ADH1* terminator. Downstream of the ORFs a zeocin selective marker cassette (*ZEO*) was included which was comprised of the *KEX2* promoter, *ZEO* ORF, 50 base pair degradation tag and *URA3* terminator. The *ZEO* marker was designed with the weak *KEX2* promoter and CL1 degradation tag [102]. This was done so that selected transformants would possess a higher number of integrations because a single copy of the low expression *ZEO* markers would not be sufficient to enable growth on selective plates with higher concentrations of zeocin. The RAIN library was ordered from GeneWiz in China and came encoded on pRS413 plasmids (Supplementary table 3).

Table 3. RAIN promoter and ORF library. Sequences for each construct from the RAIN library are available in the supplementary materials (Supplementary table 3).

RAIN library			
Promotor library			
Promoter	**Strength/characteristics**		
p*ADH2*	Glucose repressed/ ethanol induced		
p*CYC1*	Low glucose inducible/ weak		
p*PDA1*	Constitutive /weak		
p*SSB1*	Chaperone promoter/ ethanol induced		
p*SSA1*	Chaperone promoter/ethanol induced		
p*TEF1*	Translational elongation factor/ strong		
ORF library			
Species	**Gene**	**Protein**	**Function**
P. pastoris	*DAS2*	Dihydroxyacetone synthase 2	Formaldehyde Assimilation
P. pastoris	*DAS1*	Dihydroxyacetone synthase 1	
P. pastoris	*DAK*	Dihydroxyacetone kinase	
M. gastri	*PHI*	6-phosphate-3-hexuloisomerase	
M. gastri	*HPS*	3-hexulose-6-phosphate synthase	
M. buryatense	*HPS*	3-hexulose-6-phosphate synthase	
B. methanolicus	*PHI*	6-phosphate-3-hexuloisomerase	
B. methanolicus	*HPS*	3-hexulose-6-phosphate synthase	
S. cerevisiae	*TKL2*	Transketolase 2	Pentose phosphate pathway
S. cerevisiae	*TKL1*	Transketolase 1	
S. cerevisiae	*GND1*	6-phosphogluconate dehydrogenase, decarboxylating 1	
S. cerevisiae	*ZWF1*	Glucose-6-phosphate 1-dehydrogenase	
S. cerevisiae	*GND1*	6-phosphogluconate dehydrogenase, decarboxylating 2	
S. cerevisiae	*TAL1*	Transaldolase	
S. cerevisiae	*NQM1*	Transaldolase NQM1	
S. cerevisiae	*RKI1*	Ribose-5-phosphate isomerase	
S. cerevisiae	*SOL3*	6-phosphogluconolactonase 3	
S. cerevisiae	*SOL4*	6-phosphogluconolactonase 4	

CRISPR design

To facilitate better integration of RAIN cassettes, CRISPR *CAS9* with a gRNA targeting the *Ty1* sites was included in the transformations. CRISPR vector pWS171 was ordered from the lab of Tom Ellis via Addgene (Addgene plasmid # 90518; http://n2t.net/addgene:90518; RRID: Addgene_90518). pWS171 is a two part vector system that requires co-transformation of a *CAS9* containing backbone and gRNA containing fragment. After co-transformation the *CAS9* backbone

and gRNA fragment assemble into a functional vector via yeast homologous recombination and the CRISPR *CAS9* system becomes active. The pWS171 vector expresses *CAS9* from a strong p*PGK1* promoter. However, for this project p*PGK1* was replaced with a weak p*PDA1* promoter. Cloning of the gRNA and p*PDA1* were completed by a lab colleague working on a similar project engineering multiple integration of RAIN cassettes into *Ty1* (F. Harrison 2022, unpublished). PCR and clean-up of the backbone and gRNA fragment were completed as per the protocol outlined in *Experiment 1* of this thesis.

Preparation and transformation of the RAIN library

Promoter and ORF cassettes were amplified in 50 µL PCR reactions *Dpn*I digested and cleaned up as per the protocol in *Experiment 1* (Supplementary table 2). One µg of DNA solution for each of the promoter cassettes was combined into a stock promoter mix. This was also done for the ORF cassettes. This produced stock mixtures of the promoters and ORFs in roughly equal proportions of each cassette from the libraries. Equal amounts of the promoter and ORF mixtures were Gibson assembled as per the protocol outlined in *Experiment 1* (Figure 11). A strain of *S. cerevisiae* BY4741 *SGS1*Δ (*HIS3*Δ, *LEU2*Δ, *MET15*Δ, *URA3*Δ) was used for the transformations. The rationale behind using a *SGS1* knockout strain will be outlined in the results and discussions section.

Overnight cultures of BY4741 *SGS1Δ* were used to inoculate precultures in baffled shake flasks with 25 mL 1x YPD at a starting OD_{600} of 0.125. Precultures were then grown at 30 °C, 200 rpm until cells reached an OD_{600} of 0.5, or two doublings. The total volume of the precultures were then poured into sterile 50 mL Falcon tubes and spun at 4000 x g at room temperature for 10 minutes. Supernatants were removed, and pellets resuspended in 5 mL of 0.1 M lithium acetate and spun again at 4000 x g for 10 minutes at room temperature. Supernatants were removed again and pellets were resuspended in 100 µL of 0.1 M LiOAc and the total volume transferred into 1.5 mL microcentrifuge tubes. Transformation DNA was then added and cells were incubated at room temperature for 30 minutes. Transformation DNA included 45 femtomoles each of the CRISPR backbone and gRNA fragment, then either 500 ng each of the promotor and ORF mixtures, or 1 µg of the Gibson assembled promoter/ORF mixture (Figure 11). This gave two different transformation mixtures, those which contained unassembled promoter and ORF cassettes were designated *Yeast Assembly*. Those which contained Gibson assembled promoter and ORF mixture were designated *Gibson Assembly*. Another transformation mixture which contained no RAIN DNA was also included as a negative control.

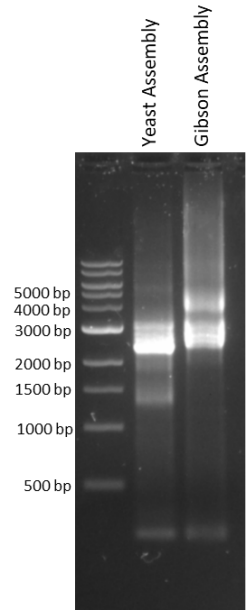

Figure 11. Gibson Assembly of RAIN promoter and ORF libraries. The left lane contains equal amounts of the RAIN promoter and ORF libraries in an unassembled mixture. The right lane contains the same RAIN library mixture after *in vitro* Gibson Assembly. These formed the basis of the two transformation mixtures that were used in *experiment 2*. Cells transformed with the unassembled promoter and ORF libraries were designated *Yeast Assembly*. Cells that were transformed with the Gibson assembled promoter and ORF libraries were designated *Gibson Assembly*.

After 30 minutes of incubation in 100 µL of 0.1 M LiOAc, the remainder of the transformation mixtures were added. This included 600 µL of 50% w/v PEG 3350, 100 µL DMSO, 90 µL 1 M lithium acetate, 10 µL single stranded carrier DNA from boiled herring sperm (NEB) and Sterile MilliQ water to bring the final volume of each transformation mixture to 1 mL prior to a further 30 minutes at room temperature. Tubes were then heat shocked in a 42 °C water bath for 14 minutes, mixing gently by inversion halfway through. After heat shock, tubes were incubated at room temperature for 10 minutes then spun at 4000 x g for 2 minutes and supernatants were discarded. Pellets were resuspended gently in 250 µL of 5 mM calcium chloride and allowed to sit at room temperature for a further 10 minutes. Tubes were spun again at 4000 x g, supernatants removed, and pellets resuspended gently in 600 µL of 1x YNB, 2% glucose and outgrown on a shaking heat block for 3 hours at 30 °C 800rpm. One hundred µL of each outgrown transformation were then plated onto 1x YPD, 2% glucose plates with zeocin concentrations of 60, 150 and 250 µg/mL and placed in a 30° C incubator for 2 days, imaged and stored at 4 °C.

Yeast colony PCR

Colony PCR was conducted on zeocin transformants to estimate the percentage of transformants that had successful integrations into any *Ty1* loci. These reactions used a forward primer on the 5' *Ty1* homology arm of the promoter cassettes (FH0026) and a reverse primer on the *ZEO* marker of the ORF cassettes (PK0072). This would amplify any assembled RAIN cassettes and give different band sizes according to whichever promoter and ORF pair that were included. A small mass of cells

was picked off the transformation plates with a pipette tip and resuspended in 50 µL of 20 mM sodium hydroxide and incubated in a thermocycler at 95 °C for 5 minutes. Lysates were spun at 4000 x g for 10 minutes. PCR reactions contained 5 µL of OneTaq master mix (NEB), 1 µL each of the forward and reverse primers, 1 µL of supernatant from the cell lysates and 2 µL of sterile MilliQ water for a final volume of 10 µL. A touchdown PCR protocol was used starting with an annealing temperature of 68 °C, decreasing by 1 °C per cycle for 15 cycles, with the annealing temperature held for 30 seconds and five minute extension times at 68 °C, followed by a further 30 cycles with the a constant annealing temperature of 50 °C and extension the same as during the touch-down phase. Both touch-down and constant phases had denaturation temperatures of 95 °C for 30 s. PCR products were run in 1% agarose 1x TAE gels at 100 mVs for 1 hour.

Experiment 3

The xylose utilisation pathway used in this study (*pTDH3-XYL1-ADH1t-pPGK1-XYL2-CYC1t-pTEF1-XKS1-PHO5t*) was ordered from Genewiz (China) and came encoded on a pRS415 vector which harbours the *LEU2* gene for minus leucin (-LEU) amino acid selection (Figure 12, Supplementary table 3). The use of *XYL1* and *XYL2* from *P. stipitis* and overexpression of native *XKS1* has been reported previously by Jin et al. to enable growth of *S. cerevisiae* on xylose [48]. The vector was transformed into CEN.PK2-1C using the same protocol outlined for plasmid transformation in *Experiment 1*. Transformants were plated out onto 1x YNB 2% glucose -LEU plates to select for the plasmid. To confirm that transformed cells were able to assimilate xylose, ten individual colonies were picked from the transformation plate and grown in plastic 50 mL Falcon tubes in 10 mL 1x YNB -LEU with 2% xylose as the sole carbon source at 30^{o} C 200 RPM for five days. After five days a visual inspection indicated which strains were able to grow on xylose, once it was established that the vector did enable growth on xylose a growth curve experiment was conducted to examine how strains would grow on varying concentrations of xylose. Three xylose utilising strains from the previous passage and three wild type controls were grown overnight in 1x YNB 2% glucose with the requisite amino acids. Each replicate was then grown in a preculture with the same media at a starting OD_{600} of 0.5. After 3 hours of preculture, cells were washed twice with sterile water to remove any residual glucose. Test cultures in 50 mL plastic Falcon tubes with 10 mL of 1x YNB, requisite amino acids and either 0.5%, 1%, 1.5% and 2% xylose were inoculated at a starting OD_{600} of 0.05 and grown at 30 °C 200 RPM. Test cultures were grown for six days and OD_{600}s were measured every 24 hours in a spectrophotometer. Three biological replicates of CEN.PK2-1C harbouring an empty pRS415 vector were also grown in 2% xylose as a negative control.

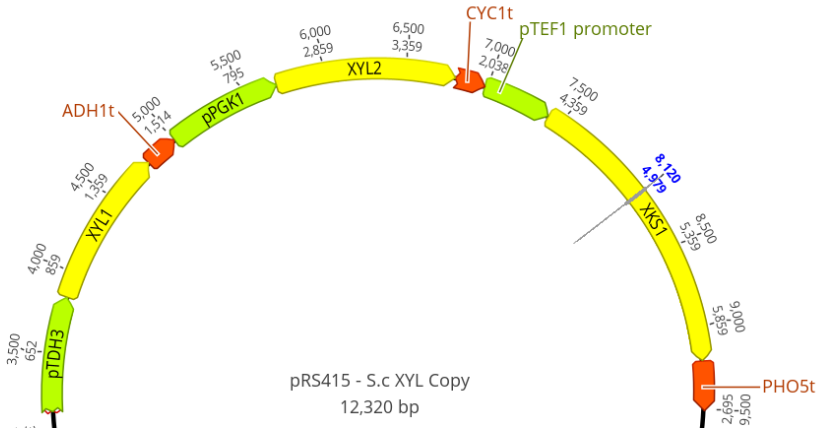

Figure 12. Design of the heterologous xylose utilisation pathway used in this study. The xylose utilisation pathway was encoded on a pRS415 vector and included *XYL1* and *XYL2* from *P. stipitis* and *XKS1* from *S. cerevisiae* on the strong TEF1 promoter. This figure was extracted from Geneious Prime.

Results and discussion

Experiment 1. Testing of different methanol oxidation enzymes as part of a synthetic RuMP cycle

Enzymes from the RuMP and XuMP cycles of different donor organism have been shown to exhibit varying performance when expressed in a heterologous host [42, 73, 74]. Therefore, a good starting point for any synthetic methylotrophy project is to test the performance of a number of different enzyme candidates in the host of choice if no previous reports are available in the literature. Krog et al. expressed the six NAD-Mdh paralogous from the *B.methanolicus* MGA3 and PB1 strains in *E.coli* and found that only *MDH2* and *MDH3* from the MGA3 strain possessed significant activity in *E. coli*[73]. Müller et al. also screened a number of RuMP cycle gene candidates in *E. coli* and found that the best performing combination of genes was *MDH2, HPS* and *PHI* from *B. methanolicus* MGA3[74].

In *S.cerevisiae*, Zhan et al., have tested a number of RuMP and XuMP cycle gene candidates including genes for methanol oxidation, including *AOX1* from *P.pastoris;* and *MDH, MDH2* and *MDH3* from *B. methanolicus* MGA3 [42]. Their results indicated that *AOX1* and *MDH3* were the most active in *S. cerevisiae*. The *MDH* gene from *B. stearothermophilus* has also been tested in a synthetic RuMP cycle in *S. cerevisiae* by Espinosa et al. [40]. However, no published work on *S. cerevisiae* has made a direct comparison of NAD-Mdhs from different bacterial species in the same experiment, and it remains unclear which methanol oxidation enzymes are preferable for use in a synthetic RuMP cycle in *S. cerevisiae*. Therefore, this study has attempted to clarify this question by taking a growth testing based approach to identify preferable methanol oxidation enzymes for use in a synthetic RuMP cycle in *S. cerevisiae*.

Spot assays on solid media did not provide any clarity on the differential performance of any of the enzymes tested (Figure 13). This was made evident by the fact that none of the strains were able to outgrow the RuMP cycle minus Mdh control strain in more than one replicate at 2% and 4% methanol. Initially, the lack of any definitive results from these experiments was surprising given that in similar growth spot trials, Espinosa et al. were able to show that a strain harbouring the same RuMP cycle vector with *B.sMDH* used in this study was able to outgrow the empty vector control [40]. However, upon further consideration of the results from both studies, the results from the present study make more sense. Because Espinosa et al. were testing whole pathways and not individual methanol oxidation enzymes, they did not engineer a RuMP cycle minus Mdh control as was done in the present study.

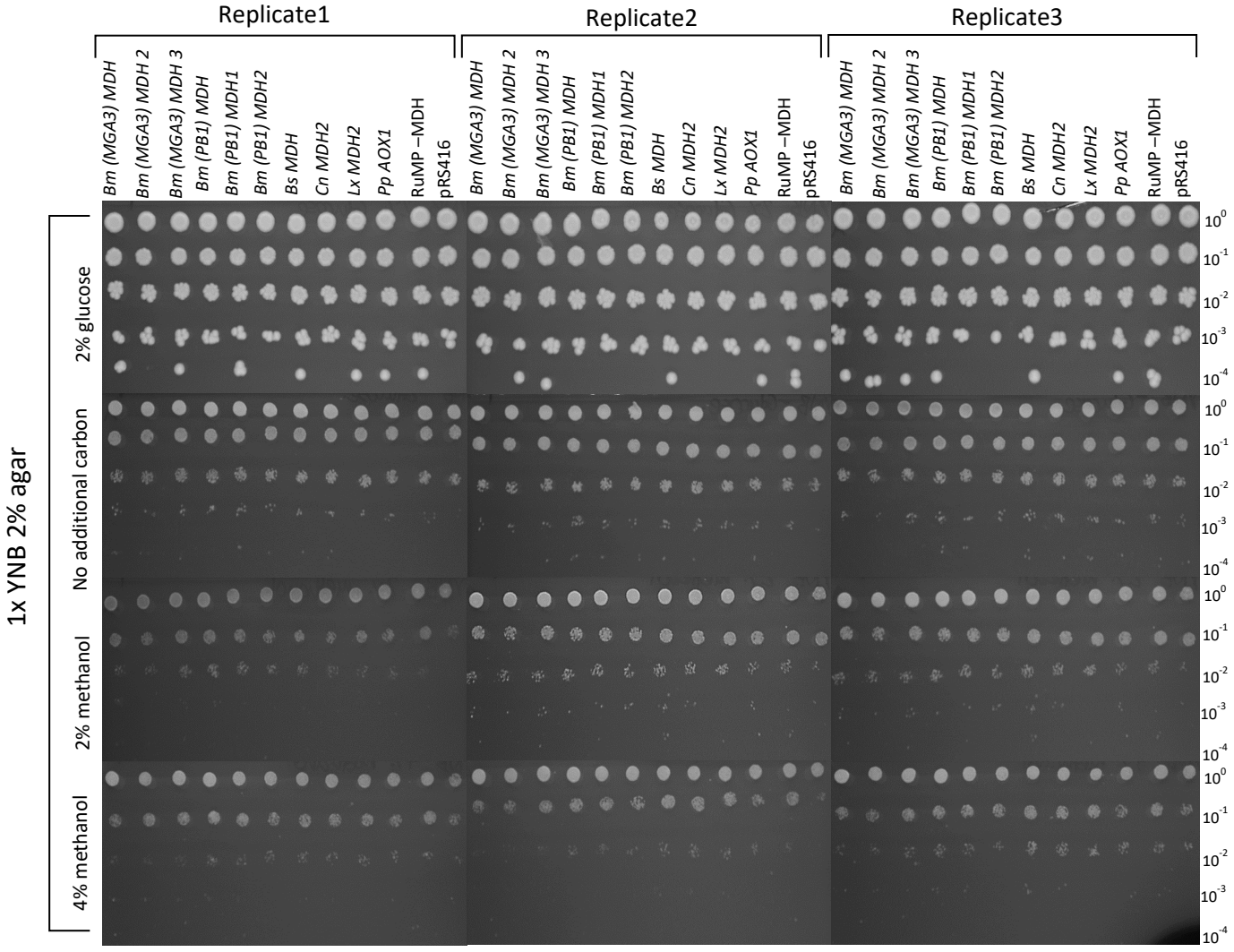

Figure 13. Spot assays of RuMP/MDH variants on solid media with methanol as the only additional carbon source. Three biological replicates of each RuMP/MDH variant were grown overnight, then passaged into precultures (starting OD600 0.5) and grown to exponential phase. Cells were washed and resuspended in water at OD600 1. Serial tenfold dilutions down to 10-4 were done for each variant and 3 µL of each dilution spotted on to solid media.

In contrast to the spot assays on solid media (Figure 13), growth trials in liquid medium (1xYNB, .25% glucose, 2% methanol) were able to provide more definitive results (Figure 14). This liquid medium was designed to limit the potential for biomass formation from glucose, while providing

initial metabolic activity to initiate or supplement any potential methylotrophic metabolism. An initial examination of the mean growth curves for each variant indicated that *Bm*MGA3*MDH3* and *Bm*PB1*MDH2* had grown more than both the RuMP –Mdh and empty vector controls (Figure 14). Mean maximum OD_{600} was used as the measure of strain performance across all variants. A one way ANOVA done in GraphPad Prism confirmed that a significant difference did exist among mean maximum ODs ($F = 33.2$, $R^2 = 0.815$, $P = <0.0001$).

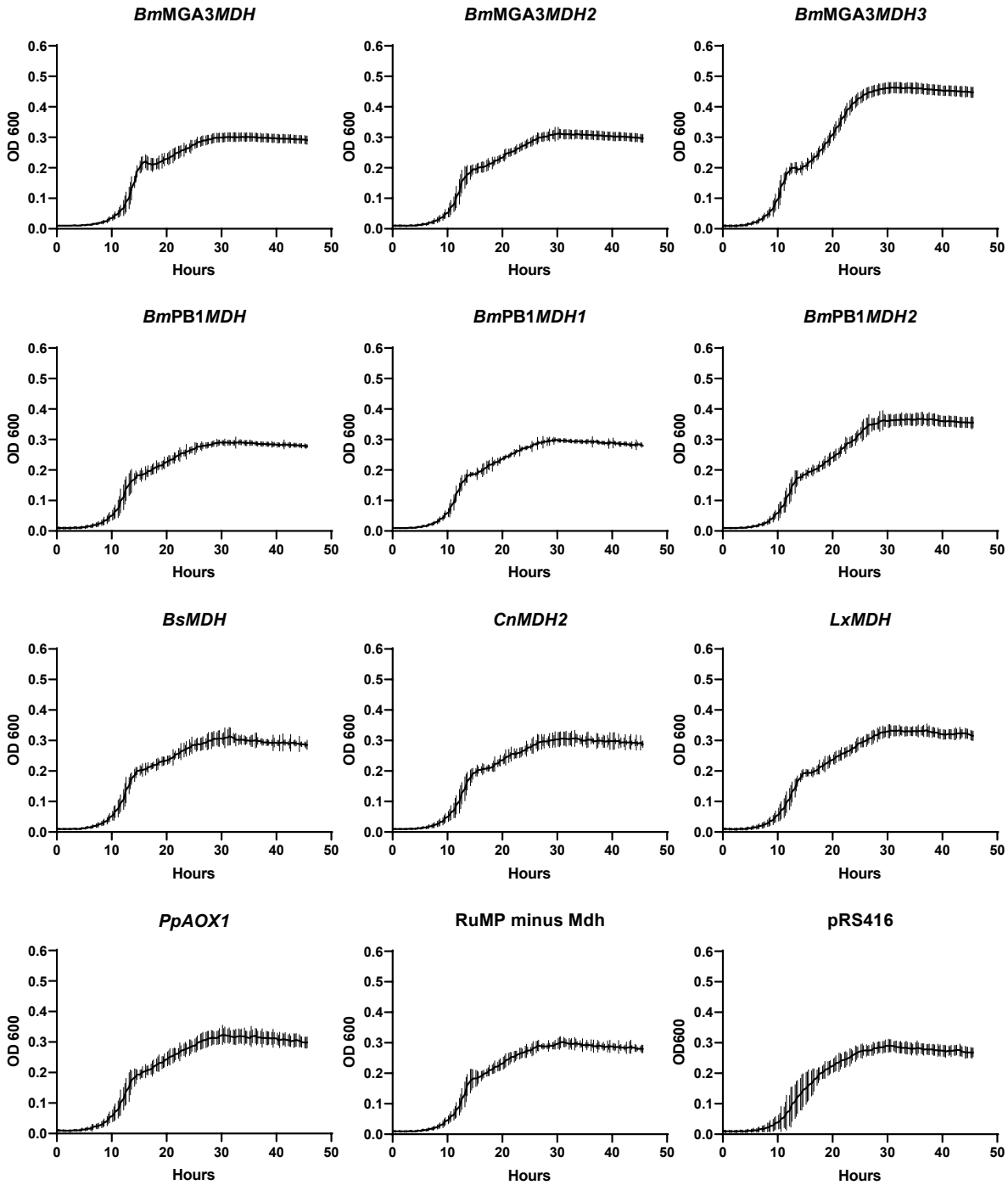

Figure 14. Growth curves of RuMP/MDH variants in 1xYNB, 0.25% glucose, 2% methanol. Strains were grown in 1xYNB, 0.25% glucose, 2% methanol for 48 hours before being passaged into the same media at a starting OD_{600} of 0.015. Variants were grown and growth curves recorded in an Agilent Synergy H1 micro plate reader at 30° C 800 rpm for 45 hours making OD_{600} readings every 20 minutes. Error bars shown are 95% confidence intervals, with lines representing the mean of 8 replicates.

This was followed by a Dunnett's multiple comparison test, comparing mean maximum OD_{600} of each variant to the RuMP -Mdh control. Of the variants tested, two were significantly higher than the RuMP -Mdh control, these were *Bm*MGA3*MDH3* and *Bm*PB1*MDH2* (Table 4). The RuMP *Bm*MGA3*MDH3* variant was the highest performer, reaching a mean maximum OD_{600} of 0.461 which represented a 52.6% increase compared to the RuMP -Mdh control. The second highest performer was the *Bm*PB1*MDH2* with a mean maximum OD_{600} of 0.367, representing a 21.6% increase compared to the RuMP -Mdh control. All other variants, including the empty vector control did not record a mean maximum OD_{600} that was significantly different from the RuMP -Mdh control (Table 4).

Table 4. Results from growth trials of RuMP/MDH variants in 1x YNB, 0.25% glucose, 2% methanol. Comparison of mean maximum OD_{600} of each test variant to the RuMP -Mdh control variant. Significance of differences in mean maximum OD_{600} was tested with a Dunnett's multiple comparison test in the GraphPad Prism software package.

RuMP cycle variant	Starting OD_{600}	Mean max OD_{600} ($\bar{y}$)	95% CI	SD (σ)	Sample size (n)	Difference vs. RuMP minus MDH ($\bar{y}_1 - \bar{y}_2$)	% Difference vs. RuMP minus Mdh	Adjusted P value	Significantly different ($\bar{y}_1 \neq \bar{y}_2$)?
*Bm*MGA3*MDH*	0.015	0.300	± 0.0119	0.0171	8	-0.00200	-0.550%	0.999	No
*Bm*MGA3*MD2*	0.015	0.310	± 0.0127	0.0184	8	0.00800	+2.73%	0.992	No
*Bm*MGA3*MDH3*	0.015	0.461	± 0.0133	0.0192	8	0.159	+52.6%	<0.0001	Yes
*Bm*PB1*MDH*	0.015	0.292	± 0.00693	0.0100	8	-0.0100	-3.39%	0.974	No
*Bm*PB1*MDH1*	0.015	0.294	± 0.0060	0.00870	8	-0.00800	-2.75%	0.992	No
*Bm*PB1*MDH2*	0.015	0.367	± 0.0186	0.0268	8	0.0650	+21.6%	<0.0001	Yes
BsMDH	0.015	0.310	± 0.0218	0.0315	8	0.00800	+2.71%	0.993	No
CnMDH2	0.015	0.307	± 0.0208	0.0301	8	0.00500	+1.50%	0.999	No
LxMDH	0.015	0.333	± 0.0156	0.0225	8	0.0310	+10.2%	0.079	No
PpAOX1	0.015	0.323	± 0.0265	0.0382	8	0.0210	+7.02%	0.414	No
pRS416 control	0.015	0.290	± 0.0176	0.0238	7	-0.0120	-3.83%	0.955	No
RuMP minus Mdh control ($\bar{y}_2$)	0.015	0.302	± 0.0130	0.0188	8	N/A	N/A	N/A	N/A

The results from this study validate and extend those from Zhan et al., who identified *Bm*MGA3*MDH3* as being the most active of the three Mdh paralogs from *B. methanolicus* MGA3 when expressed in *S. cerevisiae* [42]. However, the results of the *AOX1* strain in this study did correspond with Zhan et al., who found that *AOX1* provided superior methanol oxidation over the Mdhs tested[42]. This discrepancy could be examined further with follow up experiments that relate specifically to enzyme activities. These could include metabolomics based experiments with growth on ^{13}C methanol as per the work from Espinosa et al. [40], where fractions of ^{13}C labelled formaldehyde could be compared between variants. An invitro enzyme assay using a Nash reagent

to measure formaldehyde in cell free extracts as per Müller et al.[58], may also provide more information as to methanol oxidation enzyme activity in *S. cerevisiae*. Activity of the different NAD-Mdhs could also be examined further using an alcohol dehydrogenase activity assay kit that measures NADH concentrations. Such a study may benefit from testing in *ADH2*Δ to reduce background NADH production. This idea is supported by a study from Espinosa et al. which showed that *ADH2* was upregulated in the presence of methanol [40], and a separate study from the same group on solid media which showed that growth of *ADH2*Δ strains was severely inhibited in the presence of methanol[6]. Issues relating to background NADH activity could also be mitigated by including a poly His tag on expressed NAD-Mdhs for purification prior to *in vitro* enzyme assays, something which was not considered at the outset of this study. It is worth noting that NADH based assays cannot be used for *PpAOX1* because it does not produce NADH.

The results from this study also correspond with the liquid growth trials from Espinosa et al.[40], where the same RuMP cycle with *BsMDH* was used in a 0.1% yeast extract and 2% methanol medium. As in the present study, Espinosa et al. did not observe any increase in the growth of the RuMP/*BsMDH* test strain compared to an empty vector control. As discussed in the literature review, *BsMDH* has been shown to have a lower K_m and higher V_{max} for methanol than Mdhs from *B. methanolicus* when expressed in *E. coli* [76]. However this does not correspond with the liquid growth results from *Espinosa* et al.[40], or the present study. *BsMDH* was included in the combined RuMP/XuMP cycle employed by Wang et al. [44] in *Y. lypolitica*, achieving good results. However, this was in a strain of *Y. lypolitica* which included four copies of *BsMDH*, suggesting that this enzyme may be less active when expressed in yeast. Similarly, *CnMDH2* has been shown to be more active that other NAD-Mdhs when expressed in *E.coli* but did not confer any growth advantageous in this study[79]. On the other hand, *LxMDH* has been shown to be more active than other NAD-Mdhs in *E. coli* and did confer a growth advantage in this study. Taken together, these mixed results highlight the importance of screening multiple candidate methylotrophy genes in the host of choice at the beginning of a synthetic methylotrophy project if sufficient published data is not available.

Dai et al. also tested a RuMP cycle in *S. cerevisiae* using *MDH, HPS* and *PHI* from *B. methanolicus* MGA3, however did not report any data from growth trials and only state that the RuMP cycle used was not able to assimilate methanol, before proceeding on with XuMP pathway testing [39]. It is difficult to compare the XuMP cycle results from Dai et al., to the results from this study because they used a very high starting OD_{600} of 7, 1% methanol and 0.1% yeast extract. However, the results

of the present study could suggest that Espinosa et al. and Dai et al. did not observe favourable growth in the RuMP pathway strains tested due to the low activity of the NAD-Mdhs used (*Bm*MGA3*MDH* and *BsMDH*) in *S. cerevisiae*. If this was the case then, insufficient formaldehyde would be available for the RuMP cycle and cells could not gain a growth advantage from the methanol present.

Based on the findings from this study it can be said that *Bm*MGA3*MDH3* and *Bm*PB1*MDH2* are likely to confer significant growth advantages when used in conjunction with a synthetic RuMP cycle (Figure 15). Therefore, the use of these two NAD-Mdhs should be given close consideration by researchers during the design of future synthetic methylotrophy projects in *S. cerevisiae*. It is important to note that these results cannot be used to make any inferences about how these different enzymes may improve synthetic methanol assimilation through a RuMP cycle. This could be examined further in a series of follow up metabolomics experiments using ^{13}C methanol and mass spectrometry to measure the formation of ^{13}C labelled G3P and DHAP. This study also presents the first reported use of *CnMDH* and *LxMDH* for the design of a synthetic methanol assimilation pathway in *S. cerevisiae*. A useful follow up experiment which would complement the data reported here, would be to include genes for the corresponding ACT activator proteins of the *Bm*Mdhs. The inclusion of ACTs has been shown to increase the *in vitro* activity of *Bm*Mdhs and it is possible that they may further increase the growth advantages observed in *Bm*MGA3*MDH3* and *Bm*PB1*MDH2* or improve the growth of strains harbouring other ACT activated Mdhs [73, 77].

In this study the strains were passaged in the test medium (1x YNB, 0.25% glucose, 2% methanol) for 48 hours prior to the commencement of growth trials. In previous iterations of this experiment test cultures were inoculated with washed cells taken directly from overnight cultures done in 1x YNB, 2% glucose with no methanol. As a result, variations in growth between strains was barely detectable. However, after a 48 hour passage in the test media with 0.25% glucose and 2% methanol, growth differences in the test cultures become much more pronounced as per the results reported in this thesis. It is likely that a 48 hour passage in the presence of methanol gives cells time to physiologically adapt to allow for methanol consumption. Without this important first passage in methanol, any heterologous methanol assimilation pathway may not function effectively. This was not done for the growth spot trials on solid media due to time constraints, and could explain why these results were inconclusive. Therefore, it is recommended that for growth trials in future synthetic methylotrophy projects, test strains are passaged in the methanol containing medium at least once before the commencement of testing. This experiment has also demonstrated that 0.25%

glucose and 2% methanol are appropriate substrate concentrations for liquid media growth trials in *S. cerevisiae* to observe variable results in synthetic methylotrophy projects.

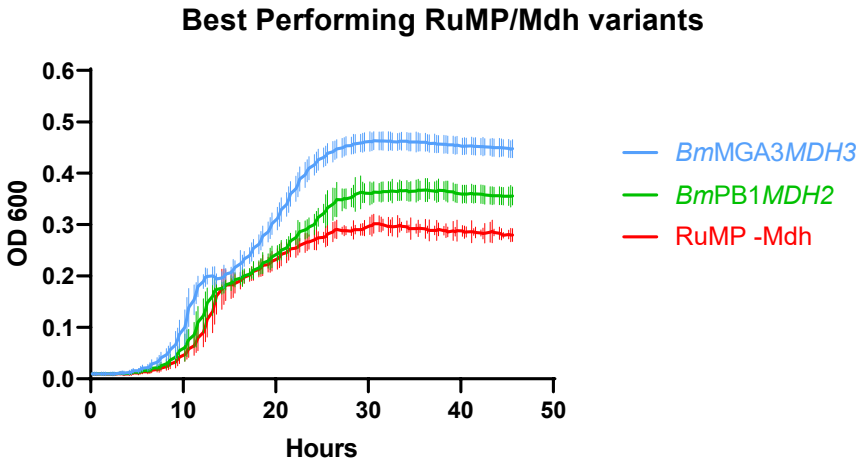

Figure 15. Growth curves of two best performing RuMP/Mdh variants. Of the ten RuMP/Mdh variants tested, *Bm*MGA3*MDH3* and *Bm*PB1*MDH2* were the only variants which displayed significantly improved growth compared to the RuMP -Mdh control in 1x YNB, 0.25% glucose, 2% methanol medium. It is recommended that these three NAD-Mdhs are considered closely by researchers for the design of future methylotrophy projects. Error bars shown are 95% confidence intervals.

Experiment 2. Random Assembly and Integration (RAIN) to engineer synthetic methanol assimilation in S. cerevisiae

As discussed in the literature review, efforts to engineer synthetic methanol assimilation in yeast so far have tested a variety of enzymes from different pathways and organisms. Despite the number of variants tested, effective growth of *S. cerevisiae* on methanol as the sole carbon source has not been achieved. This demonstrates the difficulties associated with the rational engineering strategies employed in the field so far. Mostly, the synthetic methanol assimilation strategies that have partially worked in yeast so far have involved the testing of only one copy of each enzyme in an engineered pathway [39, 40, 42]. However, Wang et al. have used four copies of *BsMDH* in their chimeric RuMP/XuMP strain of *Y. lypolitica* [44]. Some of the studies in yeast have also tested overexpression of different enzymes from glycolysis and the non-oxidative PPP to enhanced regeneration of the formaldehyde acceptor of the RuMP and XuMP pathways [40, 44]. Overexpression of the native formaldehyde detoxification enzyme FadH has also been examined in *S. cerevisiae* [40]. From this we can see that testing every combination of every gene from all the pathways and at different copy numbers and promoter strengths is not achievable for most research labs. The RAIN system demonstrated here attempts to overcome these difficulties by generating expression level diversity of promoters and ORFs from the library in a one pot transformation that yields hundreds to thousands of semi-rationally designed pathway variants for further screening. The transformation protocol used yielded good numbers of transformants selected on 60, 150 and 250 µg/mL of zeocin

for both the Yeast Assembly and Gibson Assembly transformations with no growth in any of the negative control plates (Figure 16).

An initial colony PCR was used to examine what percentage of transformants had a successful RAIN assembly and integration (Figure 17). The PCR reaction used a forward primer on the *Ty1* homology arm of the promoter cassettes (FH0026) and a reverse primer on the *ZEO* marker of the ORF cassettes (PK0072). Using primers that straddled the promoter and ORF cassettes in this way meant that amplification could not occur unless successful assembly and integration into the genome had occurred. Different band sizes indicated that assembly and integration of different promoters and ORFs from the RAIN library. This PCR was able to show that a large proportion of both the Yeast Assembly and Gibson Assembly 60 µg/mL zeocin transformants had successful amplification of an assembled RAIN cassette (Figure 17). Successful amplifications were identified by the presence of at least one clear band in the 2.5 to 4kb range, which brackets the size range of possible PCR products for this reaction. It was also encouraging to see that a few colonies displayed multiple bands within that size range. Of the 42 Yeast Assembly colonies tested, 35 had successful amplification of an assembled RAIN cassette or 83%. Of the 42 Gibson Assembly colonies tested, 31 had successful amplification of an assembled RAIN cassette or 73%. It is assessed that such a PCR can be used as a reliable diagnostic to initially assess if a RAIN transformation has resulted in successful *Ty1* integrations.

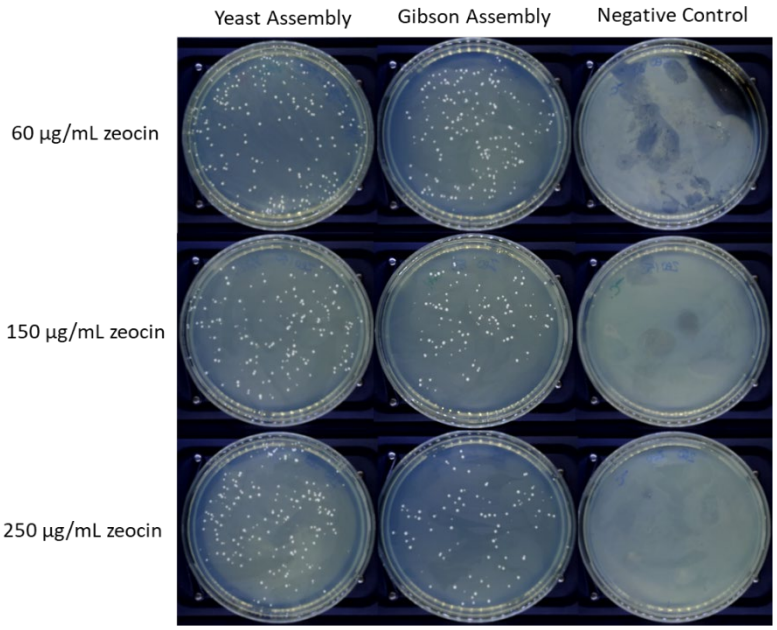

Figure 16. RAIN transformation plates

Transformations of strain BY4741 *SGS1*Δ with the RAIN promoter and ORF libraries selected on 1x YPD solid media with 60, 150 and 250 µg/mL zeocin. Yeast Assembly colonies were transformed with unassembled promoter and ORF cassettes from the RAIN library. Gibson Assembly colonies were transformed with RAIN promoter and ORF cassettes which had first undergone an *in vitro* Gibson Assembly. Each plate received 100 µL of undiluted cells post outgrowth. These plates are representative of two separate transformations for both the Yeast and Gibson Assemblies.

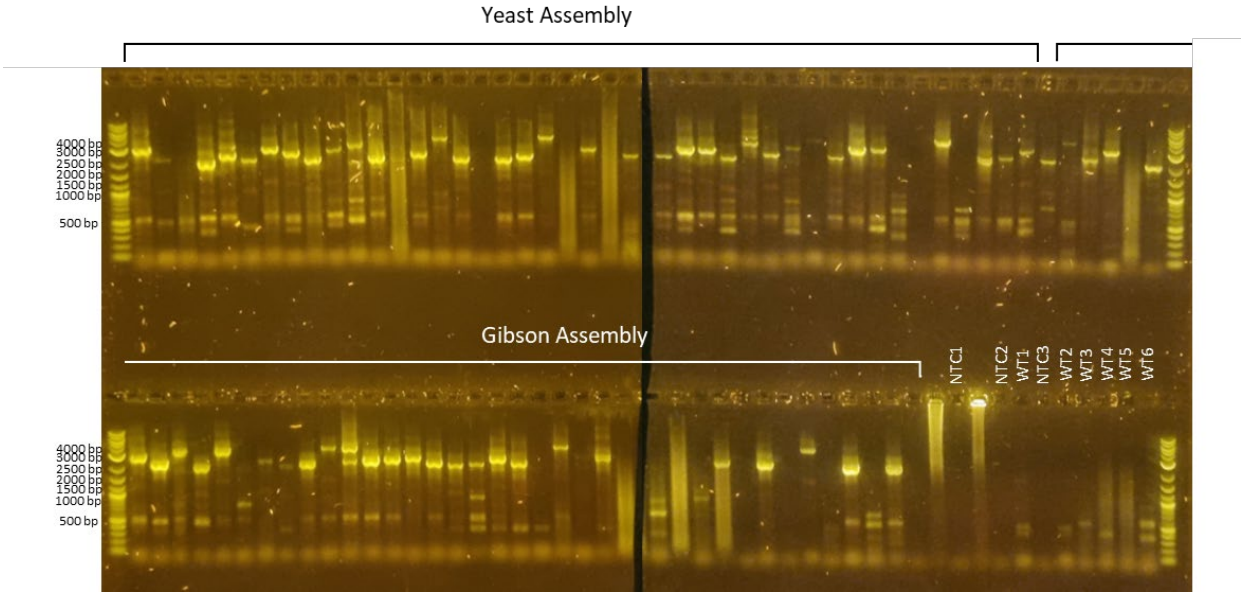

Figure 17. Initial colony PCR of RAIN transformants selected on 1xYPD 60 µg/mL zeocin plates. This PCR reaction used a forward primer on the *Ty1* homology arm of the promoter cassettes (FH0026) and a revers primer on the *ZEO* marker of the ORF cassettes (PK0072). Therefore, amplification could not occur unless successful assembly and integration into the genome had occurred. Different band sizes indicate the assembly and integration of different promoters and ORFs from the RAIN library. This image is a composite of two photos taken of the same gel cut in half as it was too wide to fit on the lightbox. No template controls (NTCs) and wild type (WT) controls are shown as marked.

The random variation in band sizes between colonies indicated that different promoters and ORFs from the library had assembled and integrated. To investigate whether transformants selected at higher concentrations of zeocin may have more integrations, a follow up colony PCR with samples of transformants selected at 60, 150, and 250 µg/mL of zeocin was run. While there did appear to be more thick bright bands from the colonies selected at 150 and 250 µg/mL than those selected on 60 µg/mL, these results are in no way quantitative and cannot be used to draw any conclusions about the effect of selective zeocin concentration on number of RAIN integrations. Therefore, the logical next step for this project would be to proceed with qPCR using a primer set which amplifies a portion of the *ZEO* ORF to provide quantitative data on the effect that selective zeocin concentration has on RAIN integration copy number. This would require the engineering of a single *ZEO* copy reference strain for comparisons to be made.

This project used a weak p*KEX2* promoter and degradation tag on the *ZEO* marker to allow for tuneable integration of RAIN cassettes at different selective concentrations of zeocin. This concept has been well demonstrated before. Parekh et al. used the weakly expressed bacterial *NEO* kanamycin resistance gene to δ integrate up to 30 copies of bovine pancreatic trypsin inhibitor (*BPTI*) at *TY2* δ sites [103]. Semkiv et al. used an attenuated version of the *KANMX* geneticin resistance gene to integrate up to 10 copies of alkaline phosphatase (*PHO8*). However, earlier iterations of this RAIN experiment which relied on non-CRISPR-mediated δ integration yielded

very few transformants at 60 µg/mL of zeocin and none at higher concentrations. To improve integration efficiency CRISPR was used. Importantly, *CAS9* was put under a weak promoter (p*PDA1*) to avoid excessive cutting of *Ty1* loci throughout the whole genome which had been shown previously, *in vitro*, to reduce cell fitness and transformation efficiency (F. Harrison 2022, unpublished). The transformation protocol was also changed, from initially using the yeast high-efficiency transformation method reported by Gietz et al. [104], to an adapted version of the super high-efficiency yeast transformation method published by William Shaw from the Tom Ellis lab on benchling.com [105]. After these steps had been optimised the results improved to what has been reported here. It was surprising that transformants were able to grow at 250 µg/mL given that the *ZEO* marker is under such a weak promoter and also attenuated with a degradation tag. This also suggests that multiple integrations have occurred. It would have been useful to repeat this experiment and select transformants at even higher concentrations of zeocin to identify the upper limit of how many integrations are possible with this system, however time did not permit. qPCR of strains selected across the full possible spectrum of zeocin couple with growth screening in methanol containing media could provide a clearer blueprint of how this system can be tuned to the optimum range of insertions for maximum fitness benefits. Apart from engineering synthetic methylotrophy, it is also clear to see how this system could be applied to the engineering of any metabolic pathway of interest. With this in mind it would be interesting to design a second ORF library with genes of the reductive glycine pathway in an attempt to engineer formate assimilation in *S. cerevisiae* [43].

Because *ADH2* has been shown to oxidise methanol in *S. cerevisiae* [40] the RAIN ORF library was designed without enzymes for methanol oxidation. However, based off the results from *Experiment 1* it is recommended that *Bm*MGA3*MDH3*, *Bm*PB1*MDH2* and *LxMDH* be added to the ORF library. It may also be beneficial to include *AOX1* from *P. pastoris* based off the results from Zhan et al., who were able to demonstrate the benefits of dual methanol oxidation [42]. The results from this experiment have demonstrated that the RAIN genomic engineering system can facilitate random assembly and integration of promoters and ORFs from the library in a single one pot transformation. To further investigate if this protocol can improve growth in the presence of methanol it is recommended that these transformants be subject to the same growth trials in *Experiment 1*.

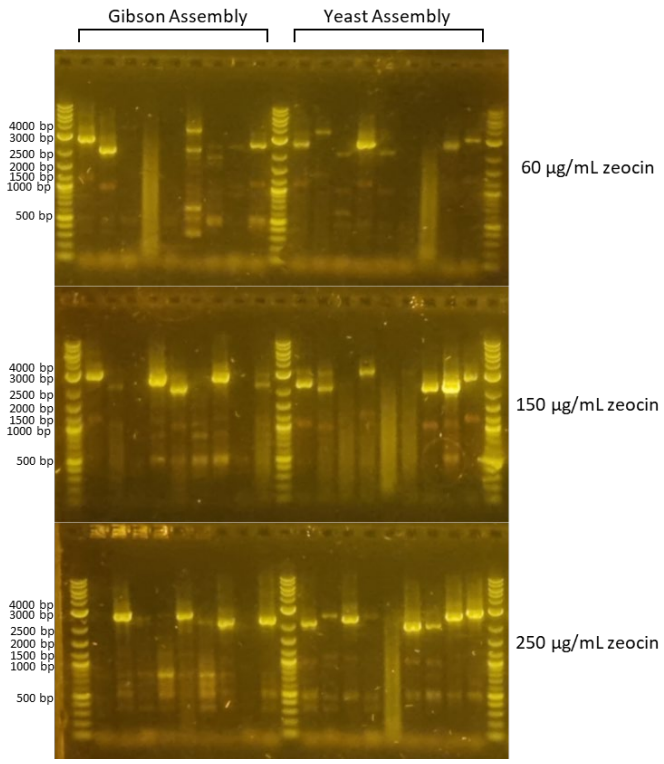

Figure 18. **PCR comparison of RAIN transformants selected on 1x YPD 60, 150 and 150 µg/mL zeocin.** NTCs and WT controls were imaged on a separate gel which is not shown here. This image is a composite of three separate gels.

Experiment 3. Engineering synthetic xylose utilisation in S. cerevisiae

This experiment set out to establish synthetic xylose utilisation in *S. cerevisiae* as a prelude to engineering synthetic methanol auxotrophy. Methanol auxotrophy has been shown to be a good starting point in ALE towards full synthetic methylotrophy [94]. In order to implement methanol auxotrophy, cells must be grown on xylose so that central metabolism occurs through the PPP. This allows for inhibition of cell growth on xylose by knocking out the gene encoding Rpi. Growth on xylose can then be restored with a synthetic RuMP cycle, which importantly, requires methanol assimilation to function. This forces cells to assimilate some methanol at the beginning of ALE and increases the chance of successfully evolving synthetic methylotrophy[25]. The xylose utilisation pathway used in this experiment (*PsXYL1, PsXYL2* and *ScXKS1*) did enable growth on xylose as the sole carbon source in *S. cerevisiae* as expected (Figure 19). Strains grown on 2% xylose where able to reach a maximum OD_{600} of 6.1. Strains grown on 1.5% xylose reached a maximum OD_{600} of 5.3. In 1% xylose, strains reached a maximum OD_{600} of 5.5 and in 0.5% xylose strains grew to a maximum OD_{600} of 2.9. Interestingly there did not appear to be a major difference in final OD_{600} between strains grown in 1.5% or 1% xylose. However, strains grown in 1.5% xylose did appear to grow slightly faster than those grown in 1% xylose, reaching an OD_{600} of 1.3 on day three compared to 0.8 in the 1% xylose strains (Figure 19). Strains grown on 1.5% xylose also appeared to enter exponential phase and subsequent stationary phase around 12 hours earlier than those grown in 1%

xylose.

It is difficult to compare the results from this study with any from the early literature regarding xylose utilisation in engineered *S. cerevisiae* [48, 106, 107, 108, 109]. This is because the earlier work on xylose fermentation in *S. cerevisiae* was focused on using *S. cerevisiae* for the industrial conversion of xylose to ethanol. Therefore, studies often used xylose as a co-substrate with glucose or yeast extract [106, 107]. In studies where engineered *S. cerevisiae* was grown on xylose as the sole carbon source, no growth curves are presented, as these were fermentation studies conducted in bioreactors that were concerned with rates of xylose utilisation and ethanol production [48, 110].

More recently, synthetic xylose utilisation has been engineered in yeast by two separate groups working on synthetic methylotrophy, these are the two previously discussed works of Zhan et al.[42] in *S. cerevisiae* and Wang et al.[44] in *Y. lipolytica*. In both studies, synthetic xylose utilisation was engineered for use in xylose/methanol co-substrate ALE [42, 44]. However, no specific data has been reported on the growth rate of strains on xylose as the sole carbon source. In the work from Zhan et al., ALE of their engineered strain of *S. cerevisiae* started on 1% xylose and 1% methanol before gradual reduction in xylose to a minimum of 0.2% over 13 months[42]. It is unclear how long Zhan et al.'s passages were carried out, as it is not reported. During the engineering of *Y. lypolitica* by Wang et al., ALE was started at 0.75% xylose and 2% methanol with a gradual reduction to 0.25% xylose with 72 hour passages [44]. Based off the results from this study, and when considering the strategies employed by Zhan et al.[42] and Wang et al.[44], 1% xylose appears to be an appropriate

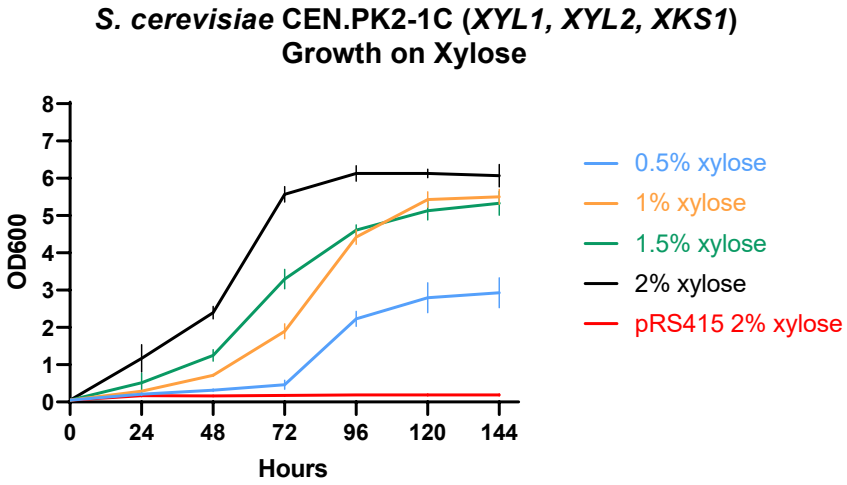

Figure 19. Growth of xylose utilising *S. cerevisiae* with differing concentrations of xylose as the sole carbon source. CEN.PK2-1C was transformed with a plasmid harbouring the heterologous xylose utilisation pathway (*PsXYL1-PsXYL2-ScXKS*-pRS415*)* and grown for six days in defined liquid media containing differing concentrations of xylose as the sole carbon source. Optical density (600 nm) of each strain was measured in a spectrophotometer every 24 hours. Values plotted are the mean of 3 biological replicates ± one standard deviation. An empty pRS415 vector strain was included as a control.

starting concentration for xylose/methanol ALE. This is because 1% xylose appears to be the lowest starting concentration that allows for reliable passaging over 72 hours whereas 0.5% xylose is not sufficient (Figure 19). However, it is important to note that neither of the efforts from Zhan et al.[42] and Wang et al.[44] were successful in evolving effective growth on methanol as the sole carbon source, highlighting the difficulties associated with standard modes of ALE for evolving synthetic methylotrophy. Therefore, as discussed in the literature review, it may be beneficial to first implement synthetic methanol auxotrophy before ALE to improve the chances of success.

Engineering synthetic methanol auxotrophy in *S. cerevisiae* is not as straightforward as it is in bacteria [96]. For synthetic methanol auxotrophy to work on xylose and methanol, the gene encoding Rpi must be knocked out, something which is easily done in bacteria [94]. However, in *S. cerevisiae,* the gene encoding Rpi, *RKI1*, is an essential gene and *RKI1* deficient cells are not viable [111]. This presents a major obstacle to engineering synthetic methanol auxotrophy in *S. cerevisiae* and is presumably why it has not yet been achieved. Interestingly, Wang et al. tried to knock out *RKI1* in *Y. lypolitica* to implement synthetic methanol auxotrophy but were unable to do so [44]. This suggests that *RKI1* is also an essential gene in other yeasts. It is still unclear why *RKI1* deficient *S. cerevisiae* are not viable and it is suggested that Rpi must have another essential function in yeast apart from its role in the non-oxidative PPP [111].

Interestingly, it has been reported that a single nucleotide polymorphism (SNP) in *RKI1* (566 G>A), resulting in an amino acid substitution (Arg189> Lys), can inactivate Rpi whilst cells remain viable[111]. This results in cells becoming auxotrophic for the essential B vitamin, pyridoxine (vitamin B12), as *RKI1* is responsible for production of R5P and Ru5P which are the precursors for pyridoxine. When supplemented with pyridoxine RKI1$^{Arg189>Lys}$ cells are viable whereas with a full gene deletion will not grow at all. In other words, when a non-functional version of *RKI1* is transcribed, the cell loses Rpi activity but remains viable, presumably because the inactive Rpi is able to fulfill any other essential role it plays in the cell. This presents a possible solution to engineering synthetic methanol auxotrophy in *S. cerevisiae*. If the *RKI1*$^{Arg189>Lys}$ SNP can be introduced alongside the synthetic xylose utilisation pathway used in this study and a synthetic RuMP/XuMP cycle, it may be possible to implement methanol dependant growth on xylose and methanol. Whether or not this would work is unknown and requires further investigation. Successful engineering of synthetic methanol auxotrophy in *S. cerevisiae* would be an excellent starting point for ALE of full synthetic methylotrophy.

Conclusion

Enabling the widescale use of methanol as a feedstock for industrial fermentation has the potential to unlock vast reserves of waste carbon and facilitate a global expansion of biotechnology. *S. cerevisiae* is well placed as the industrial eukaryote of choice to maximise the conversion of methanol into useful products. However, as demonstrated in this thesis, significant challenges remain on the path towards methanol based biomanufacturing in *S. cerevisiae*. The literature review provided an introduction to methylotrophic metabolism and outlined the current state of the field in engineering synthetic methylotrophy, paying particular attention to the work in yeast. A number of suggestions were also made as to how further progress could be made. The key recommendation from the literature review is that engineering synthetic methanol auxotrophy in *S. cerevisiae* should be a high priority. This can then be followed by ALE towards possible full synthetic methylotrophy using xylose and methanol as co-substrates. This strategy has born great success in enabling full synthetic methylotrophy in *E. coli* [25] and is sound from an evolutionary biology perspective, where the selective pressure to assimilate methanol is increased by making growth reliant on it.

Experiment 1 tested 10 different methanol oxidation enzymes from different bacterial methylotrophs and one yeast as part of a synthetic RuMP cycle. From growth trials in liquid medium with 0.25% glucose and 2% methanol, two of the enzymes tested were shown to confer a significant growth advantage of 52.6% compared to a negative control strain. These were, *Bm*MGA3*MDH3* and *Bm*PB1*MDH2*. From these results it is recommended that these two NAD-Mdhs be given close consideration for use in future synthetic methylotrophy projects in *S. cerevisiae*. This study also demonstrates the utility of testing different candidate methylotrophy genes at the outset of a synthetic methylotrophy project if sufficient published data is not available for the host of choice.

Experiment 2 demonstrated that the novel RAIN genome engineering technique can facilitate random assembly and integration of promoters and ORFs from a predefined library into the *Ty1* retrotransposon, generating hundreds of variants in a single transformation. The protocol reported here has been shown to yield good numbers of transformants with high rates of successful integrations. In order to gain quantitative data on how the concentration of selective zeocin influences RAIN cassette integration copy number, it is recommended that further qPCR experiments are conducted. Transformants from this study should also undergo growth trials in 0.25% glucose, 2% methanol liquid media as per *Experiment 1* to examine whether RAIN can confer growth advantages in engineered cells in the presence of methanol. The RAIN library in its current design does not contain enzymes for methanol oxidation. Based off the results from *Experiment 1* it is recommended that *Bm*MGA3*MDH3* and *Bm*PB1*MDH2* are included in the RAIN

library to provide sufficient formaldehyde for any functioning methanol assimilation pathway which arises. The RAIN library may also benefit from the inclusion of *AOX*1 from *P. pastoris* based of the results from Zhan et al. [42].

Experiment 3 used a previously reported synthetic xylose utilisation pathway[106] to establish a strain of *S. cerevisiae* which could grow on xylose. This was done as a prelude to engineering synthetic methanol auxotrophy via inactivation of *RKI1* and addition of a synthetic RuMP cycle. As discussed previously, the implementation of synthetic methanol auxotrophy provides an advantageous starting point for ALE towards full synthetic methylotrophy and has been employed with great success in *E. coli*[25]. It is assessed that it may yield similar success in *S. cerevisiae* if the challenges associated with engineering synthetic methanol auxotrophy can be overcome. Namely the fact that *RKI1* deficient *S. cerevisiae* are not viable. It may be possible to negate this by introducing an amino acid substitution in *RKI1* which has been reported to inactivate the gene product but allow cells to remain viable[111]. Therefore, this should be the next step to advance this experiment.

Continued progress in the field of synthetic methylotrophy has made robust growth of *S. cerevisiae* on methanol a realistic proposition. Use of one carbon feedstocks will be essential for the expansion of a sustainable bioeconomy in the future. The work presented in this thesis has laid a foundation from which further research should be undertaken. Despite the challenges which remain, a number of credible paths forward have been identified which may bare success in the future.

Abbreviations

General: **ALE**, adaptive laboratory evolution; **C1**, one-carbon; **C3**, three-carbon; **CRISPR**: clustered regularly interspaced short palindromic repeats; **CSM**, Complete Supplement Mixture (essential amino acids); **HDA**, High-def Azure essential amino acid mixture; **OD_{600}**, optical density at 600 nano metres; **YNB**, yeast nitrogen base.

Metabolic pathways: **CBB cycle**, Calvin Benson Basham cycle; **PPP**, pentose phosphate pathway; **rGly pathway**, reductive glycine pathway; **RuBP cycle**, ribulose bisphosphate cycle; **RuMP cycle**, ribulose monophosphate cycle; **TCA cycle**, citric acid cycle; **XuMP cycle**, xylulose monophosphate cycle.

Metabolites: **ATP**, adenosine triphosphate; **DHA**, dihydroxyacetone; **DHAP**, dihydroxyacetone phosphate; **E4P**, erythrose-4-phosphate; **F6P**, fructose 6-phosphate; **G3P**, glyceraldehyde 3-phosphate; **H6P**, hexulose 6-phosphate; **NAD(H)**, nicotinamide adenine dinucleotide; **R5P**, ribose-5-phosphate; **Ru5P**, ribulose 5-phosphate; **S7P**, sedoheptulose 7-phosphate; **THF**, tetrahydrofolate; **Xu5P**, xylulose 5-phosphate.

Enzymes/proteins:
Act, nudix hydrolase activator protein; **Ald**, acetaldehyde dehydrogenase; **Amt**, aminomethyl-transferase **Aox**, alcohol oxidase; **Cta**, catalase; **Dak**, dihydroxyacetone kinase; **Das**, dihydroxyacetone synthase; **Dld**, dihydrolypoyl dehydrogenase; **Fadh**, formaldehyde dehydrogenase; **Fba**, fructose-bisphosphate aldolase; **Fbp**, fructose bisphosphatase; **Fdh**, formate dehydrogenase; **GCS**, glycine cleavage system; **Gldc**, glycine dehydrogenase; **Hps**, 3-hexulose-6-phosphate synthase; **Mdh**, methanol dehydrogenase; **NAD-Mdh**, NAD^+ dependent methanol dehydrogenase; **PQQ-Mdh**, pyrroloquinoline quinone dependent methanol dehydrogenase; **Rpe**, ribulose phosphate 3-epimerase; **Rpi**, ribose-5-phosphate isomerase; **Sds**, serine deaminase; **Shmt**, serine hydroxymethylase; **Tal**, transaldolase; **Tkt**, transketolase.

References

(1) Clomburg, J. M.; Crumbley, A. M.; Gonzalez, R. Industrial biomanufacturing: The future of chemical production. *Science* **2017**, *355* (6320).
(2) Tuyishime, P.; Sinumvayo, J. P. Novel outlook in engineering synthetic methylotrophs and formatotrophs: a course for advancing C1-based chemicals production. *World Journal of Microbiology and Biotechnology* **2020**, *36* (8), 1-16.
(3) Bos, M.; Kersten, S.; Brilman, D. Wind power to methanol: Renewable methanol production using electricity, electrolysis of water and CO2 air capture. *Applied energy* **2020**, *264*, 114672.
(4) Chen, C.; Kotyk, J. F. K.; Sheehan, S. W. Progress toward commercial application of electrochemical carbon dioxide reduction. *Chem* **2018**, *4* (11), 2571-2586.
(5) Liu, Y.; Li, F.; Zhang, X.; Ji, X. Recent progress on electrochemical reduction of CO2 to methanol. *Current Opinion in Green and Sustainable Chemistry* **2020**, *23*, 10-17.
(6) Espinosa, M. I.; Williams, T. C.; Pretorius, I. S.; Paulsen, I. T. Benchmarking two Saccharomyces cerevisiae laboratory strains for growth and transcriptional response to methanol. *Synthetic and systems biotechnology* **2019**, *4* (4), 180-188.
(7) Wang, Y.; Fan, L.; Tuyishime, P.; Zheng, P.; Sun, J. Synthetic methylotrophy: a practical solution for methanol-based biomanufacturing. *Trends in biotechnology* **2020**, *38* (6), 650-666.
(8) Chistoserdova, L.; Kalyuzhnaya, M. G.; Lidstrom, M. E. The expanding world of methylotrophic metabolism. *Annual review of microbiology* **2009**, *63*, 477-499.
(9) Mosin, O.; Ignatov, I. Metabolism and Physiology of Methylotrophic Microorganisms. *Journal of Medicine, Physiology and Biophysics* **2014**, *6*, 61-84.
(10) Whitaker, W. B.; Sandoval, N. R.; Bennett, R. K.; Fast, A. G.; Papoutsakis, E. T. Synthetic methylotrophy: engineering the production of biofuels and chemicals based on the biology of aerobic methanol utilization. *Current opinion in biotechnology* **2015**, *33*, 165-175.
(11) Lian, J.; Mishra, S.; Zhao, H. Recent advances in metabolic engineering of Saccharomyces cerevisiae: new tools and their applications. *Metabolic engineering* **2018**, *50*, 85-108.
(12) Li, M.; Borodina, I. Application of synthetic biology for production of chemicals in yeast Saccharomyces cerevisiae. *FEMS Yeast Research* **2015**, *15* (1), 1-12.
(13) Borodina, I.; Kildegaard, K. R.; Jensen, N. B.; Blicher, T. H.; Maury, J.; Sherstyk, S.; Schneider, K.; Lamosa, P.; Herrgård, M. J.; Rosenstand, I. Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in Saccharomyces cerevisiae via β-alanine. *Metabolic engineering* **2015**, *27*, 57-64.
(14) Yazawa, H.; Iwahashi, H.; Kamisaka, Y.; Kimura, K.; Uemura, H. Production of polyunsaturated fatty acids in yeast Saccharomyces cerevisiae and its relation to alkaline pH tolerance. *Yeast* **2009**, *26* (3), 167-184.
(15) Buijs, N. A.; Siewers, V.; Nielsen, J. Advanced biofuel production by the yeast Saccharomyces cerevisiae. *Current opinion in chemical biology* **2013**, *17* (3), 480-488.
(16) Galanie, S.; Thodey, K.; Trenchard, I. J.; Interrante, M. F.; Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science* **2015**, *349* (6252), 1095-1100.
(17) Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M.; Tai, A.; Main, A.; Eng, D. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **2013**, *496* (7446), 528-532.
(18) Wang, G.; Huang, M.; Nielsen, J. Exploring the potential of Saccharomyces cerevisiae for biopharmaceutical protein production. *Current opinion in biotechnology* **2017**, *48*, 77-84.
(19) Li, Y.; Li, S.; Thodey, K.; Trenchard, I.; Cravens, A.; Smolke, C. D. Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proceedings of the National Academy of Sciences* **2018**, *115* (17), E3922-E3931.
(20) Verwaal, R.; Wang, J.; Meijnen, J.-P.; Visser, H.; Sandmann, G.; van den Berg, J. A.; van Ooyen, A. J. High-level production of beta-carotene in Saccharomyces cerevisiae by successive transformation with carotenogenic genes from Xanthophyllomyces dendrorhous. *Applied and environmental microbiology* **2007**, *73* (13), 4342-4350.
(21) Tippmann, S.; Chen, Y.; Siewers, V.; Nielsen, J. From flavors and pharmaceuticals to advanced biofuels: production of isoprenoids in Saccharomyces cerevisiae. *Biotechnology journal* **2013**, *8* (12), 1435-1444.

(22) Zhang, M.; Yuan, X.-J.; Zhang, C.; Zhu, L.-P.; Mo, X.-H.; Chen, W.-J.; Yang, S. Bioconversion of methanol into value-added chemicals in native and synthetic methylotrophs. *Current Issues in Molecular Biology* **2019**, *33*, 225-236.
(23) Antoniewicz, M. R. Synthetic methylotrophy: Strategies to assimilate methanol for growth and chemicals production. *Current opinion in biotechnology* **2019**, *59*, 165-174.
(24) Gregory, G. J.; Bennett, R. K.; Papoutsakis, E. T. Recent advances toward the bioconversion of methane and methanol in synthetic methylotrophs. *Metabolic Engineering* **2021**, *71*, 99-116.
(25) Chen, F. Y.-H.; Jung, H.-W.; Tsuei, C.-Y.; Liao, J. C. Converting Escherichia coli to a synthetic methylotroph growing solely on methanol. *Cell* **2020**, *182* (4), 933-946. e914.
(26) Albers, E.; Larsson, C. A comparison of stress tolerance in YPD and industrial lignocellulose-based medium among industrial and laboratory yeast strains. *Journal of Industrial Microbiology and Biotechnology* **2009**, *36* (8), 1085-1091.
(27) Wêgrzyn, G. Bacteriophage contamination: is there a simple method to reduce its deleterious effects in laboratory cultures and biotechnological factories? *Journal of Applied Genetics* **2004**, *45* (1), 111-120.
(28) Vieira Gomes, A. M.; Souza Carmo, T.; Silva Carvalho, L.; Mendonça Bahia, F.; Parachin, N. S. Comparison of yeasts as hosts for recombinant protein production. *Microorganisms* **2018**, *6* (2), 38.
(29) Hammer, S. K.; Avalos, J. L. Harnessing yeast organelles for metabolic engineering. *Nature chemical biology* **2017**, *13* (8), 823-832.
(30) Wagner, H. J.; Capitain, C. C.; Richter, K.; Nessling, M.; Mampel, J. Engineering bacterial microcompartments with heterologous enzyme cargos. *Engineering in Life Sciences* **2017**, *17* (1), 36-46.
(31) Looser, V.; Bruhlmann, B.; Bumbak, F.; Stenger, C.; Costa, M.; Camattari, A.; Fotiadis, D.; Kovar, K. Cultivation strategies to enhance productivity of Pichia pastoris: a review. *Biotechnology advances* **2015**, *33* (6), 1177-1193.
(32) Gao, J.; Jiang, L.; Lian, J. Development of synthetic biology tools to engineer Pichia pastoris as a chassis for the production of natural products. *Synthetic and systems biotechnology* **2021**, *6* (2), 110-119.
(33) Yamada, R.; Ogura, K.; Kimoto, Y.; Ogino, H. Toward the construction of a technology platform for chemicals production from methanol: D-lactic acid production from methanol by an engineered yeast Pichia pastoris. *World Journal of Microbiology and Biotechnology* **2019**, *35* (2), 37.
(34) Guo, F.; Dai, Z.; Peng, W.; Zhang, S.; Zhou, J.; Ma, J.; Dong, W.; Xin, F.; Zhang, W.; Jiang, M. Metabolic engineering of Pichia pastoris for malic acid production from methanol. *Biotechnology and Bioengineering* **2021**, *118* (1), 357-371.
(35) Gao, L.; Cai, M.; Shen, W.; Xiao, S.; Zhou, X.; Zhang, Y. Engineered fungal polyketide biosynthesis in Pichia pastoris: a potential excellent host for polyketide production. *Microbial cell factories* **2013**, *12* (1), 1-14.
(36) Jeong, E.; Shim, W. Y.; Kim, J. H. Metabolic engineering of Pichia pastoris for production of hyaluronic acid with high molecular weight. *Journal of biotechnology* **2014**, *185*, 28-36.
(37) Liu, Y.; Tu, X.; Xu, Q.; Bai, C.; Kong, C.; Liu, Q.; Yu, J.; Peng, Q.; Zhou, X.; Zhang, Y. Engineered monoculture and co-culture of methylotrophic yeast for de novo production of monacolin J and lovastatin from methanol. *Metabolic engineering* **2018**, *45*, 189-199.
(38) Fabarius, J. T.; Wegat, V.; Roth, A.; Sieber, V. Synthetic Methylotrophy in Yeasts: Towards a Circular Bioeconomy. *Trends in Biotechnology* **2020**, *39* (4), 348-358.
(39) Dai, Z.; Gu, H.; Zhang, S.; Xin, F.; Zhang, W.; Dong, W.; Ma, J.; Jia, H.; Jiang, M. Metabolic construction strategies for direct methanol utilization in Saccharomyces cerevisiae. *Bioresource technology* **2017**, *245*, 1407-1412.
(40) Espinosa, M. I.; Gonzalez-Garcia, R. A.; Valgepea, K.; Plan, M.; Scott, C.; Pretorius, I. S.; Marcellin, E.; Paulsen, I. T.; Williams, T. Engineering and Evolution of Methanol Assimilation in Saccharomyces cerevisiae. *bioRxiv* **2020**, 717942.
(41) Espinosa, M. I.; Gonzalez-Garcia, R. A.; Valgepea, K.; Plan, M. R.; Scott, C.; Pretorius, I. S.; Marcellin, E.; Paulsen, I. T.; Williams, T. C. Adaptive laboratory evolution of native methanol assimilation in Saccharomyces cerevisiae. *Nature communications* **2020**, *11* (1), 1-12.
(42) Zhan, C.; Li, X.; Baidoo, E. E.; Yang, Y.; Sun, Y.; Wang, S.; Wang, Y.; Wang, G.; Nielsen, J.; Keasling, J. D. The Glyoxylate-Serine Pathway Enables Conversion of Saccharomyces cerevisiae to a Synthetic Methylotroph. *Saccharomyces cerevisiae* **2021**.

(43) Gonzalez de la Cruz, J.; Machens, F.; Messerschmidt, K.; Bar-Even, A. Core catalysis of the reductive glycine pathway demonstrated in yeast. *ACS synthetic biology* **2019**, *8* (5), 911-917.
(44) Wang, G.; Olofsson-Dolk, M.; Hansson, F. G.; Donati, S.; Li, X.; Chang, H.; Cheng, J.; Dahlin, J.; Borodina, I. Engineering Yeast Yarrowia lipolytica for Methanol Assimilation. *ACS Synthetic Biology* **2021**, *10* (12), 3537-3550.
(45) Kelso, P. A.; Chow, L. K.; Carpenter, A. C.; Paulsen, I. T.; Williams, T. C. Toward Methanol-Based Biomanufacturing: Emerging Strategies for Engineering Synthetic Methylotrophy in Saccharomyces cerevisiae. *ACS Synthetic Biology* **2022**, *11* (8), 2548-2563.
(46) Gibson, D. G.; Benders, G. A.; Axelrod, K. C.; Zaveri, J.; Algire, M. A.; Moodie, M.; Montague, M. G.; Venter, J. C.; Smith, H. O.; Hutchison, C. A. One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic Mycoplasma genitalium genome. *Proceedings of the National Academy of Sciences* **2008**, *105* (51), 20404-20409.
(47) Curcio, M. J.; Lutz, S.; Lesage, P. The Ty1 LTR-retrotransposon of budding yeast, Saccharomyces cerevisiae. *Microbiology spectrum* **2015**, *3* (2), 1.
(48) Jin, Y.-S.; Laplaza, J. M.; Jeffries, T. W. Saccharomyces cerevisiae engineered for xylose metabolism exhibits a respiratory response. *Applied and environmental microbiology* **2004**, *70* (11), 6816-6825.
(49) Claassens, N. J. Reductive Glycine Pathway: A Versatile Route for One-Carbon Biotech. *Trends in Biotechnology* **2021**.
(50) Figueroa, I. A.; Barnum, T. P.; Somasekhar, P. Y.; Carlström, C. I.; Engelbrektson, A. L.; Coates, J. D. Metagenomics-guided analysis of microbial chemolithoautotrophic phosphite oxidation yields evidence of a seventh natural CO2 fixation pathway. *Proceedings of the National Academy of Sciences* **2018**, *115* (1), E92-E101.
(51) Sánchez-Andrea, I.; Guedes, I. A.; Hornung, B.; Boeren, S.; Lawson, C. E.; Sousa, D. Z.; Bar-Even, A.; Claassens, N. J.; Stams, A. J. The reductive glycine pathway allows autotrophic growth of Desulfovibrio desulfuricans. *Nature communications* **2020**, *11* (1), 1-12.
(52) van der Klei, I. J.; Yurimoto, H.; Sakai, Y.; Veenhuis, M. The significance of peroxisomes in methanol metabolism in methylotrophic yeast. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **2006**, *1763* (12), 1453-1462.
(53) Heggeset, T. M.; Krog, A.; Balzer, S.; Wentzel, A.; Ellingsen, T. E.; Brautaset, T. Genome sequence of thermotolerant Bacillus methanolicus: features and regulation related to methylotrophy and production of L-lysine and L-glutamate from methanol. *Applied and environmental microbiology* **2012**, *78* (15), 5170-5181.
(54) Zhu, T.; Zhao, T.; Bankefa, O. E.; Li, Y. Engineering unnatural methylotrophic cell factories for methanol-based biomanufacturing: challenges and opportunities. *Biotechnology advances* **2020**, *39*, 107467.
(55) Rußmayer, H.; Buchetics, M.; Gruber, C.; Valli, M.; Grillitsch, K.; Modarres, G.; Guerrasio, R.; Klavins, K.; Neubauer, S.; Drexler, H. Systems-level organization of yeast methylotrophic lifestyle. *BMC biology* **2015**, *13* (1), 1-25.
(56) Yurimoto, H.; Kato, N.; Sakai, Y. Genomic organization and biochemistry of the ribulose monophosphate pathway and its application in biotechnology. *Applied microbiology and biotechnology* **2009**, *84* (3), 407-416.
(57) Yurimoto, H.; Kato, N.; Sakai, Y. Assimilation, dissimilation, and detoxification of formaldehyde, a central metabolic intermediate of methylotrophic metabolism. *Chemical Record (New York, NY)* **2005**, *5* (6), 367-375.
(58) Müller, J. E.; Heggeset, T. M.; Wendisch, V. F.; Vorholt, J. A.; Brautaset, T. Methylotrophy in the thermophilic Bacillus methanolicus, basic insights and application for commodity production from methanol. *Applied microbiology and biotechnology* **2015**, *99* (2), 535-551.
(59) Crowther, G. J.; Kosály, G.; Lidstrom, M. E. Formate as the main branch point for methylotrophic metabolism in Methylobacterium extorquens AM1. *Journal of bacteriology* **2008**, *190* (14), 5057-5062.
(60) Pfeifenschneider, J.; Brautaset, T.; Wendisch, V. F. Methanol as carbon substrate in the bio-economy: Metabolic engineering of aerobic methylotrophic bacteria for production of value-added chemicals. *Biofuels, Bioproducts and Biorefining* **2017**, *11* (4), 719-731.
(61) Šmejkalová, H.; Erb, T.; Fuchs, G. Methanol Assimilation in Methylobacterium extorquens AM1: Demonstration of All Enzymes and Their Regulation *PLoS One* **2010**, *5* (10), e13001.

(62) Keltjens, J. T.; Pol, A.; Reimann, J.; den Camp, H. J. O. PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. *Applied microbiology and biotechnology* **2014**, *98* (14), 6163-6183.
(63) Anthony, C. The biochemistry of methylotrophs. **1982**.
(64) Bar-Even, A.; Noor, E.; Flamholz, A.; Milo, R. Design and analysis of metabolic pathways supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2013**, *1827* (8-9), 1039-1047.
(65) Cotton, C. A.; Claassens, N. J.; Benito-Vaquerizo, S.; Bar-Even, A. Renewable methanol and formate as microbial feedstocks. *Current opinion in biotechnology* **2020**, *62*, 168-180.
(66) Comer, A. D.; Long, M. R.; Reed, J. L.; Pfleger, B. F. Flux balance analysis indicates that methane is the lowest cost feedstock for microbial cell factories. *Metabolic Engineering Communications* **2017**, *5*, 26-33.
(67) de Graaf, B.; Clore, A.; McCullough, A. K. Cellular pathways for DNA repair and damage tolerance of formaldehyde-induced DNA-protein crosslinks. *DNA repair* **2009**, *8* (10), 1207-1214.
(68) Zhang, W.; Zhang, T.; Wu, S.; Wu, M.; Xin, F.; Dong, W.; Ma, J.; Zhang, M.; Jiang, M. Guidance for engineering of synthetic methylotrophy based on methanol metabolism in methylotrophy. *RSC advances* **2017**, *7* (7), 4083-4091.
(69) Zhang, W.; Zhang, T.; Song, M.; Dai, Z.; Zhang, S.; Xin, F.; Dong, W.; Ma, J.; Jiang, M. Metabolic engineering of Escherichia coli for high yield production of succinic acid driven by methanol. *ACS Synthetic Biology* **2018**, *7* (12), 2803-2811.
(70) Yasueda, H.; Kawahara, Y.; Sugimoto, S.-i. Bacillus subtilis yckG and yckF encode two key enzymes of the ribulose monophosphate pathway used by methylotrophs, and yckH is required for their expression. *Journal of bacteriology* **1999**, *181* (23), 7154-7160.
(71) Klein, V. J.; Irla, M.; Gil López, M.; Brautaset, T.; Fernandes Brito, L. Unravelling Formaldehyde Metabolism in Bacteria: Road towards Synthetic Methylotrophy. *Microorganisms* **2022**, *10* (2), 220.
(72) Kim, S.; Lindner, S. N.; Aslan, S.; Yishai, O.; Wenk, S.; Schann, K.; Bar-Even, A. Growth of E. coli on formate and methanol via the reductive glycine pathway. *Nature chemical biology* **2020**, *16* (5), 538-545.
(73) Krog, A.; Heggeset, T. M.; Müller, J. E.; Kupper, C. E.; Schneider, O.; Vorholt, J. A.; Ellingsen, T. E.; Brautaset, T. Methylotrophic Bacillus methanolicus encodes two chromosomal and one plasmid born NAD+ dependent methanol dehydrogenase paralogs with different catalytic and biochemical properties. *PloS one* **2013**, *8* (3), e59188.
(74) Müller, J. E.; Meyer, F.; Litsanov, B.; Kiefer, P.; Potthoff, E.; Heux, S.; Quax, W.; Wendisch, V. F.; Brautaset, T.; Portais, J.-C. Engineering Escherichia coli for methanol conversion. *Metabolic Engineering* **2015**, *28*, 190-201.
(75) Feist, A. M.; Henry, C. S.; Reed, J. L.; Krummenacker, M.; Joyce, A. R.; Karp, P. D.; Broadbelt, L. J.; Hatzimanikatis, V.; Palsson, B. Ø. A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular systems biology* **2007**, *3* (1), 121.
(76) Whitaker, W. B.; Jones, J. A.; Bennett, R. K.; Gonzalez, J. E.; Vernacchio, V. R.; Collins, S. M.; Palmer, M. A.; Schmidt, S.; Antoniewicz, M. R.; Koffas, M. A. Engineering the biological conversion of methanol to specialty chemicals in Escherichia coli. *Metabolic engineering* **2017**, *39*, 49-59.
(77) Ochsner, A. M.; Müller, J. E.; Mora, C. A.; Vorholt, J. A. In vitro activation of NAD-dependent alcohol dehydrogenases by Nudix hydrolases is more widespread than assumed. *FEBS letters* **2014**, *588* (17), 2993-2999.
(78) Ochsner, A. M.; Müller, J. E.; Mora, C. A.; Vorholt, J. A. In vitro activation of NAD-dependent alcohol dehydrogenases by Nudix hydrolases and attempts at in vivo investigation. *Characterization of novel methanol dehydrogenases and identification of the gene set required for methylotrophic growth* **2017**, *588*, 17.
(79) Wu, T.-Y.; Chen, C.-T.; Liu, J. T.-J.; Bogorad, I. W.; Damoiseaux, R.; Liao, J. C. Characterization and evolution of an activator-independent methanol dehydrogenase from Cupriavidus necator N-1. *Applied microbiology and biotechnology* **2016**, *100* (11), 4969-4983.
(80) Lee, J.-Y.; Park, S.-H.; Oh, S.-H.; Lee, J.-J.; Kwon, K. K.; Kim, S.-J.; Choi, M.; Rha, E.; Lee, H.; Lee, D.-H. Discovery and biochemical characterization of a methanol dehydrogenase from Lysinibacillus xylanilyticus. *Frontiers in bioengineering and biotechnology* **2020**, *8*, 1-10.

(81) Jordà, J.; Jouhten, P.; Cámara, E.; Maaheimo, H.; Albiol, J.; Ferrer, P. Metabolic flux profiling of recombinant protein secreting Pichia pastoris growing on glucose: methanol mixtures. *Microbial cell factories* **2012**, *11* (1), 1-14.
(82) Hiltunen, J. K.; Mursula, A. M.; Rottensteiner, H.; Wierenga, R. K.; Kastaniotis, A. J.; Gurvitz, A. The biochemistry of peroxisomal β-oxidation in the yeast Saccharomyces cerevisiae. *FEMS microbiology reviews* **2003**, *27* (1), 35-64.
(83) Yan, M.; Rayapuram, N.; Subramani, S. The control of peroxisome number and size during division and proliferation. *Current opinion in cell biology* **2005**, *17* (4), 376-383.
(84) Vizeacoumar, F. J.; Torres-Guzman, J. C.; Bouard, D.; Aitchison, J. D.; Rachubinski, R. A. Pex30p, Pex31p, and Pex32p form a family of peroxisomal integral membrane proteins regulating peroxisome size and number in Saccharomyces cerevisiae. *Molecular biology of the cell* **2004**, *15* (2), 665-677.
(85) Fodor, K.; Wolf, J.; Reglinski, K.; Passon, D.; Lou, Y.; Schliebs, W.; Erdmann, R.; Wilmanns, M. Ligand-induced compaction of the PEX5 receptor-binding cavity impacts protein import efficiency into peroxisomes. *Traffic (Copenhagen, Denmark)* **2014**, *16* (1), 85-98.
(86) Nicaud, J. M. Yarrowia lipolytica. *Yeast* **2012**, *29* (10), 409-418.
(87) Sandberg, T. E.; Salazar, M. J.; Weng, L. L.; Palsson, B. O.; Feist, A. M. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. **2019**.
(88) Watanabe, S.; Kodaki, T.; Makino, K. Complete reversal of coenzyme specificity of xylitol dehydrogenase and increase of thermostability by the introduction of structural zinc. *Journal of Biological Chemistry* **2005**, *280* (11), 10340-10349.
(89) Bennett, R. K.; Gonzalez, J. E.; Whitaker, W. B.; Antoniewicz, M. R.; Papoutsakis, E. T. Expression of heterologous non-oxidative pentose phosphate pathway from Bacillus methanolicus and phosphoglucose isomerase deletion improves methanol assimilation and metabolite production by a synthetic Escherichia coli methylotroph. *Metabolic engineering* **2018**, *45*, 75-85.
(90) Gonzalez, J. E.; Bennett, R. K.; Papoutsakis, E. T.; Antoniewicz, M. R. Methanol assimilation in Escherichia coli is improved by co-utilization of threonine and deletion of leucine-responsive regulatory protein. *Metabolic engineering* **2018**, *45*, 67-74.
(91) Koopman, F. W.; De Winde, J. H.; Ruijssenaars, H. J. C 1 compounds as auxiliary substrate for engineered Pseudomonas putida S12. *Applied microbiology and biotechnology* **2009**, *83* (4), 705-713.
(92) Witthoff, S.; Schmitz, K.; Niedenführ, S.; Nöh, K.; Noack, S.; Bott, M.; Marienhagen, J. Metabolic engineering of Corynebacterium glutamicum for methanol metabolism. *Applied and environmental microbiology* **2015**, *81* (6), 2215-2225.
(93) Leßmeier, L.; Pfeifenschneider, J.; Carnicer, M.; Heux, S.; Portais, J.-C.; Wendisch, V. F. Production of carbon-13-labeled cadaverine by engineered Corynebacterium glutamicum using carbon-13-labeled methanol as co-substrate. *Applied microbiology and biotechnology* **2015**, *99* (23), 10163-10176.
(94) Chen, C.-T.; Chen, F. Y.-H.; Bogorad, I. W.; Wu, T.-Y.; Zhang, R.; Lee, A. S.; Liao, J. C. Synthetic methanol auxotrophy of Escherichia coli for methanol-dependent growth and production. *Metabolic engineering* **2018**, *49*, 257-266.
(95) Lee, Y.; Rivera, J. G. L.; Liao, J. C. Ensemble Modeling for Robustness Analysis in engineering non-native metabolic pathways. *Metabolic engineering* **2014**, *25*, 63-71.
(96) Gao, B.; Zhao, N.; Deng, J.; Gu, Y.; Jia, S.; Hou, Y.; Lv, X.; Liu, L. Constructing a methanol-dependent Bacillus subtilis by engineering the methanol metabolism. *Journal of Biotechnology* **2022**, *343*, 128-137.
(97) Meyer, F.; Keller, P.; Hartl, J.; Gröninger, O. G.; Kiefer, P.; Vorholt, J. A. Methanol-essential Growth of Escherichia Coli. *Nature communications* **2018**, *9* (1), 1508.
(98) Claassens, N. J.; Bordanaba-Florit, G.; Cotton, C. A.; De Maria, A.; Finger-Bou, M.; Friedeheim, L.; Giner-Laguarda, N.; Munar-Palmer, M.; Newell, W.; Scarinci, G. Replacing the Calvin cycle with the reductive glycine pathway in Cupriavidus necator. *Metabolic Engineering* **2020**, *62*, 30-41.
(99) Entian, K.-D.; Kötter, P. 25 yeast genetic strain and plasmid collections. *Methods in microbiology* **2007**, *36*, 629-666.
(100) Koch, C.; Neumann, P.; Valerius, O.; Feussner, I.; Ficner, R. Crystal Structure of Alcohol Oxidase from Pichia pastoris. *PLoS ONE* **2016**, *11* (2).

(101) Peng, B.; Williams, T. C.; Henry, M.; Nielsen, L. K.; Vickers, C. E. Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: a comparison of yeast promoter activities. *Microbial cell factories* **2015**, *14* (1), 1-11.
(102) Gilon, T.; Chomsky, O.; Kulka, R. G. Degradation signals for ubiquitin system proteolysis in Saccharomyces cerevisiae. *The EMBO journal* **1998**, *17* (10), 2759-2766.
(103) Parekh, R.; Shaw, M.; Wittrup, K. An integrating vector for tunable, high copy, stable integration into the dispersed Ty delta sites of Saccharomyces cerevisiae. *Biotechnology progress* **1996**.
(104) Gietz, R. D.; Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature protocols* **2007**, *2* (1), 31-34.
(105) Shaw, W. *Yeast Transformation - Super-High Efficiency*. 2016. https://benchling.com/protocols/hYSdel7a/yeast-transformation-super-high-efficiency (accessed.
(106) Jin, Y.-S.; Lee, T.-H.; Choi, Y.-D.; Ryu, Y.-W.; Seo, J.-H. Conversion of Xylose to Ethanol by Recombinant Saccharomyces cerevisiae Containing Genes for Xylose Reductase and Xylose Reductase and xylitol Dehydrogenase from Pichia stipitis. *Journal of microbiology and biotechnology* **2000**, *10* (4), 564-567.
(107) Ho, N. W.; Chen, Z.; Brainard, A. P. Genetically engineered Saccharomyces yeast capable of effective cofermentation of glucose and xylose. *Applied and environmental microbiology* **1998**, *64* (5), 1852-1859.
(108) Kötter, P.; Ciriacy, M. Xylose fermentation by Saccharomyces cerevisiae. *Applied microbiology and biotechnology* **1993**, *38* (6), 776-783.
(109) Tantirungkij, M.; Izuishi, T.; Seki, T.; Yoshida, T. Fed-batch fermentation of xylose by a fast-growing mutant of xylose-assimilating recombinant Saccharomyces cerevisiae. *Applied microbiology and biotechnology* **1994**, *41* (1), 8-12.
(110) Jin, Y.-S.; Ni, H.; Laplaza, J. M.; Jeffries, T. W. Optimal growth and ethanol production from xylose by recombinant Saccharomyces cerevisiae require moderate D-xylulokinase activity. *Applied and environmental Microbiology* **2003**, *69* (1), 495-503.
(111) Kondo, H.; Nakamura, Y.; Dong, Y.-X.; Nikawa, J.-i.; Sueda, S. Pyridoxine biosynthesis in yeast: participation of ribose 5-phosphate ketol-isomerase. *Biochemical Journal* **2004**, *379* (1), 65-70.

Supplementary data

Supplementary table 1. Gene block sequences used in experiment 1.

Gene block name	Gene block sequence
*Bm*MGA3*MDH*	ATGACCACCAACTTTTTCATACCACCTGCCTCTGTCATTGGTAGAGGGGCCGTAAAAGAAGTCGGTACTAGATT GAAACAAATTGGCGCAAAAAAAGCCTTGATCGTGACAGATGCTTTTTTACACAGCACAGGCTTGTCTGAAGAA GTTGCAAAGAACATTAGAGAAGCTGGGGTTGACGTTGCAATATTTCCGAAGGCACAACCGGATCCGGCTGACA CTCAAGTGCATGAAGGAGTAGATGTGTTTAAACAAGAGAATTGTGACTCACTGGTATCTATTGGTGGAGGTTC CTCCCATGATACAGCAAAAGCCATCGGGTTAGTCGCCGCAAACGGTGGTAGAATAAATGATTACCAAGGGGTT AATAGTGTGGAAAAAACCCGTCGTCCCGGTTGTGGCTATCACCACGACGGCCGGTACTGGATCAGAAACTACCT CCTTAGCTGTCATTACCGACTCCGCCAGGAAAGTAAAAATGCCAGTGATAGATGAGAAAATCACACCAACTGT CGCAATTGTCGATCCGGAATTGATGGTTAAGAAGCCAGCCGGGTTGACGATAGCAACTGGGATGGATGCTTTA AGCCATGCCATTGAGGCATATGTTGCAAAAGGTGCTACCCCAGTAACCGATGCCTTTGCAATTCAGGCCATGA AGCTTATCAATGAATACTTGCCAAAAGCTGTTGCAAATGGTGAAGATATAGAAGCTCGTGAAAAGATGGCTTA TGCTCAATACATGGCAGGCGTCGCTTTTAATAACGGAGGATTAGGACTAGTGCATTCAATCTCGCATCAGGTTG GCGGAGTTTATAAGTTACAACATGGTATCTGTAATTCGGTTAACATGCCTCATGTTTGCGCCTTCAACCTGATC GCTAAAACAGAGAGATTTGCACATATAGCTGAATTATTGGGTGAGAACGTTGCTGGTTTGTCGACTGCAGCTG CCGCTGAAAGAGCCATTGTAGCTCTGGAAAGAATCAACAAGAGTTTTGGAATACCATCTGGGTATGCAGAGAT GGGTGTTAAGGAAGAAGATATAGAGTTATTAGCAAAGAATGCATACGAAGACGTTTGCACTCAATCCAATCCT AGGGTGCCAACGGTTCAAGACATTGCACAAATAATCAAGAATGCGATGTAA
*Bm*MGA3*MDH2*	ATGAAGAACACACAAAGTGCCTTCTATATGCCTAGTGTGAATTTGTTTGGTGCCGGCAGTGTCAATGAGGTGG GAACGAGATTAGCGGGATTGGGTGTGAAAAAAGCCCTTTTAGTTACTGATGCTGGGTTGCATAGCCTGGGTCTT TCAGAAAAAATAGCTGGTATAATTAGAGAAGCTGGGGTGGAAGTGGCTATTTTTCCAAAGGCTGAACCTAATC CAACAGATAAAAATGTTGCGGAAGGTCTTGAGGCTTACAATGCAGAAAATTGTGACTCGATTGTTACTCTAGG CGGTGGATCCAGCCACGATGCTGGTAAAGCTATAGCTTTAGTTGCGGCAAATGGAGGTACAATACATGACTAC GAAGGGGTGGATGTCAGTAAAAAGCCAATGGTTCCACTTATTGCCATTAATACTACCGCTGGAACTGGTAGCG AACTTACAAAATTCACTATCATTACTGATACAGAAAGAAAAGTAAAAATGGCAATAGTCGATAAACACGTTAC ACCCACATTAAGTATAAATGACCCGGAATTAATGGTTGGTATGCCGCCGTCCTTGACCGCTGCCACTGGCCTTG ATGCATTGACGCATGCTATTGAGGCTTATGTATCCACGGGTGCAACTCCAATAACAGATGCCCTGGCAATTCAG GCCATTAAGATCATATCTAAGTATTTACCCAGAGCAGTTGCAAATGGTAAGGATATTGAGGCTAGAGAACAGA TGGCCTTTGCACAAAGCTTAGCTGGGATGGCCTTCAATAATGCAGGTTTGGGATATGTGCATGCTATTGCTCAC CAACTAGGTGGCTTTTACAACTTCCCGCATGGAGTATGTAACGCCATCTTGTTGCCTCACGTATGTCGTTTTAAC CTGATAAGCAAGGTAGAAAGGTATGCGGAAATTGCAGCATTCCTTGGAGAAAACGTTGACGGTTTGAGCACCT ATGAAGCGGCAGAGAAGGCCATAAAGGCTATTGAAAGGATGGCTAGAGACTTAAATATTCCCAAAGGCTTTA AGGAATTGGGTGCAAAAGAGGAGGACATAGAGACATTGGCAAAGAATGCAATGAATGACGCGTGTGCTTTGA CGAATCCCAGAAAACCTAAACTGGAAGAAGTGATACAAATTATAAAAAACGCTATGTGA
*Bm*MGA3*MDH3*	ATGACTAATACACAATCAGCGTTCTTTATGCCCTCAGTAAATTTATTTGGTGCCGGTAGTGTAAACGAAGTAGG TACCAGACTAGCCGATTTAGGTGTTAAGAAGGCTTTGCTTGTAACAGATGCCGGCTTACACGGCCTAGGCTTAT CTGAGAAGATCTCATCTATAATCAGAGCTGCCGGTGTTGAAGTTTCCATTTTTCCAAAAGCTGAACCTAATCCT ACAGACAAAAACGTAGCTGAAGGTTTGGAAGCCTACAATGCTGAGAATTGCGATAGCATAGTTACCCTAGGCG GTGGGTCCTCTCACGATGCTGGAAAAGCCATTGCCTTAGTCGCTGCTAATGGAGGCAAAATCCACGATTATGA AGGTGTTGACGTGAGCAAAGAACCTATGGTCCCACTTATTGCTATTAACACTACGGCCGGCACAGGATCAGAA CTAACCAAGTTTACTATTATCACTGATACAGAAAGAAAAGTAAAAATGGCTATTGTAGACAAACATGTAACCC CTACATTATCAATTAATGATCCGGAGCTTATGGTTGGTATGCCTCCATCTCTAACCGCGGCTACTGGTCTAGAT GCATTAACACATGCAATAGAAGCATATGTCTCTACAGGAGCAACGCCAATAACAGATGCCTTAGCCATCCAAG CAATTAAAATCATCTCTAAGTATTTACCACGTGCTGTCGCCAACGGAAAGGATATTGAGGCGCGTGAGCAAAT GGCCTTCGCTCAGAGTTTAGCTGGTATGGCATTTAATAACGCTGGATTAGGTTATGTACATGCCATTGCTCATC AATTGGGGGGTTTTTATAATTTTCCACATGGTGTTTGTAATGCTGTTCTTCTACCATACGTTTGTAGGTTTAACT TGATCTCTAAAGTCGAAAGATATGCCGAAATTGCAGCTTTCCTAGGTGAAAACGTCGACGGTTTGTCCACATAC GATGCAGCTGAAAAGGCTATTAAAGCTATTGAAAGAATGGCTAAAGATTTGAACATCCCCAAAGGTTTTAAAG AACTAGGCGCAAAAGAAGAAGACATAGAGACTCTAGCCAAAAATGCAATGAAAGATGCTTGTGCTTTGACGA ATCCTAGGAAACCTAAACTGGAAGAAGTAATTCAAATTATTAAGAATGCTATGTGA
*Bm*PB1*MDH*	ATGACCCAAAGAAACTTCTTTATTCCACCCGCTTCGGTAATTGGTAGGGGTGCAGTGAAAGAGGTCGGAACCA GATTAAAGCAAATAGGTGCTACCAAAGCCTTAATTGTCACAGACGCTTTCTTACATGGCACTGGGTTATCAGAA GAAGTGGCGAAGAACATTAGAGAGGCTGGCTTAGATGCCGTCATTTTTCCAAAAGCTCAACCTGATCCGGCTG ACACGCAAGTCCACGAGGGAGTTGATATATTTAAACAAGAAAAATGTGATGCTTTAGTGTCGATTGGTGGTGG TAGTAGTCATGACACAGCTAAGGCGATTGGCTTGGTAGCCGCTAACGGCGGAAGAATTAACGATTATCAAGGC GTCAACTCCGTCGAAAAACCCGTGGTCCCAGTTGTGGCAATAACTACTACTGCTGGAACGGGATCGGAAACAA CGTCATTAGCGGTTATTACCGACTCTGCTCGTAAGGTAAAGATGCCGGTAATTGACGAAAAAATTACCCCAAC CGTTGCCATAGTGGACCCAGAGCTAATGGTGAAAAAACCGGCGGGCCTAACGATTGCAACAGGTATGGATGCC TTATCTCATGCAATCGAGGCCTATGTTGCTAAAAGAGCCACTCCAGTAACTGATGCATTCGCTATTCAGGCTAT GAAATTGATTAACGAGTACTTGCCGAGAGCAGTGGCTAATGGTGAAGACATTGAAGCTAGGGAAGCGATGGC CTATGCACAATATATGGCAGGTGTAGCATTTAATAACGGAGGATTGGGCCTGGTACACTCTATTTCACACCAAG TCGGAGGGGTTTATAAACTTCAACATGGTATCTGTAACTCGGTTAATATGCCACATGTCTGTCAATTTAACTTA ATAGCTAGGACTGAGAGATTTGCTCATATAGCTGAATTGTTGGGTGAAAATGTCAGTGGTTTATCGACCGCTAG TGCTGCAGAAAGAGCTATAGTGGCACTACAGAGATACAATAAGAATTTTGGAATCCCATCCGGATATGCCGAG ATGGGTGTAAAAGAGGAAGATATTGAGCTTCTGGCGAATAATGCCTATCAGGATGTGTGTACATTAGACAATC CTAGAGTCCCAACTGTACAAGATATTGCTCAAATCATCAAGAATGCTTTGTAA
*Bm*PB1*MDH1*	ATGACAAAGACAAAGTTTTTTATTCCTTCCTCAACAGTATTCGGTCGTGGCGCTGTCAAAGAAGTTGGTGCTAG ACTTAAGGCAATAGGCGCAACAAAAGCACTAATTGTAACTGATGCTTTTTTACATTCAACTGGACTGAGTGAG GAGGTTGCCAAGAACATACGTGAGGCTGGATTGGATGTCGTAATTTTCCCTAAGGCTCAACCGGATCCAGCGG ATACGCAGGTTCACGAAGGTGTTGAAGTGTTCAAGCAAGAAAAATGTGATGCCCTAGTCTCTATCGGTGGTGG CTCCAGTCACGATACGGCGAAAGGTATCGGTTTAGTTGCCGCCAACGGCGGCCGTATAAATGATTATCAAGGG GTCAACTCAGTCGAGAAACAAGTAGTACCACAAATTGCAATCACTACGACAGCCGGTACGGGTAGTGAAACCA CATCTTTGGCCGTCATTACCGACAGTGCCAGGAAAGTCAAAATGCCCGTAATAGATGAAAAAATTACCCCAAC GGTCGCTATTGTAGACCCTGAACTAATGGTAAAAAAACCTGCTGGATTAACAATAGCAACGGGCATGGATGCT

	TTGAGTCATGCAATTGAAGCTTATGTTGCTAAGAGAGCAACTCCTGTTACAGATGCATTCGCTATTCAAGCCAT GAAGCTAATCAACGAATATTTGCCTAAGGCAGTCGCTAACGGGGAGGATATTGAAGCTAGAGAAGCGATGGCT TACGCACAGTACATGGCTGGCGTAGCTTTTAATAATGGCGGTTTGGGGCTGGTGCACAGCATTAGTCATCAAGT GGGAGGTGTGTACAAACTGCAGCACGGTATCTGCAATAGCGTTGTCATGCCTCATGTTTGTCAGTTTAATTTGA TTGCAAGAACCGAAAGATTTGCACATATTGCAGAACTTTTAGGCGAGAACGTTTCTGGTTTAAGTACCGCTTCT GCTGCCGAAAGAACTATAGCTGCGTTGGAAAGATACAACAGAAACTTCGGCATTCCCAGTGGTTACAAAGCCA TGGGGGTGAAAGAAGAAGACATAGAATTATTAGCTAATAATGCTATGCAAGATGTGTGCACTCTAGATAACCC TCGTGTTCCAACGGTTCAAGACATACAACAAATTATTAAGAATGCACTATAA
*Bm*PB1*MDH2*	ATGACAAATACTCAATCAATATTTTACATCCCTTCAGTAAACCTATTTGGACCGGGGTCTGTGAACGAAGTAGG TACTAGATTGGCAGGTCTAGGTGTTAAGAAAGCATTATTAGTCACTGACGCTGGTTTGCATGGACTAGGGTTGA GCGAGAAGATTGCGTCGATTATTAGGGAAGCGGGTGTTGAGGTGCTTATCTTTCCCAAAGCCGAACCAAACCC AACTGATAAAAATGTGGCAGAGGGATTAGAGGTGTACAACGCTGAAAACTGTGATTCAATCGTCACCTTAGGG GGTGGTTCGAGCCACGATGCAGGTAAGGGCATTGCCTTAGTGGCTGCCAATGGTGGTACAATTTACGACTACG AAGGAGTTGACAAGTCGAAGAAGCCAATGGTCCCATTAATAGCAATTAACACTACAGCTGGTACTGGTTCTGA GTTGACTAGATTTACAATAATAACCGACACAGAGAGAAAAGTGAAGATGGCCATCGTTGATAAACATGTAACC CCCACACTAAGTATTAACGATCCTGAGTTGATGGTCGGCATGCCACCAAGTCTAACTGCAGCAACTGGTCTGG ATGCCCTAACACACGCTATTGAAGCATACGTCTCGACGGCTGCTACACCAATTACTGATGCTCTAGCCATACAA GCTATTAAGATCATTTCTAAGTACTTACCTAGGGCTTTCGCCAATGGCAAGGACATGGAAGCTAGGGAGCAGA TGGCATTCGCTCAAAGCTTAGCGGGTATGGCATTTAATAATGCTTCCCTAGGGTACGTACATGCAATAGCTCAT CAATTTGGTGGCTTCTACAATTTCCCTCACGGTGTTTGTAACGCTATACTTCTGCCTCATGTATGTAGATTCAAC TTGATTTCCAAAGTTGAAAGATTTGCGGAAATTGCCGCGCTATTAGGCGAAAACGTTGCTGGATTATCAACAA GAGAAGCGGCGGAAAAGGGGATCAAAGCTATTGAGCGTATGGCTAAAGACTTGAATATCCCAAGAGGTTTTA AAGAACTTGGGGCTAAGGAAGAAGATATTGTTACATTAGCCGAAAACGCCATGAAGGATGCTACAGCGCTGA CTAACCCTAGAAAACCTAAACTAGAAGAGGTTATTCAGATAATAAAGAATGCGATGTAA
CnMDH2	ATGACTCATCTAAACATAGCTAACAGAGTCGATTCATTCTTTATCCCTTGCGTAACTCTTTTCGGGCCTGGGTGC GCAAGAGAGACTGGAGCTAGAGCTCGTTCACTGGGTGCTAGGAAAGCTTTAATTGTTACCGATGCCGGACTTC ACAAAATGGGCTTGTCAGAAGTCGTAGCTGGCCACATAAGAGAAGCTGGATTGCAAGCGGTGATCTTTCCGGG TGCCGAACCAAACCCAACTGATGTAAATGTGCACGACGGTGTTAAGTTGTTCGAAAGAGAGGAGTGTGACTTC ATCGTGAGCTTAGGGGGTGGTTCCTCGCACGATTGTGCTAAGGGTATAGGGTTGGTGACAGCAGGTGGTGGTC ATATTAGGGATTATGAAGGTATTGACAAATCTACAGTCCCTATGACTCCTTTAATTTCGATCAATACTACTGCA GGAACTGCAGCAGAGATGACTAGATTCTGCATAATCACCAATTCGAGCAATCATGTGAAAATGGCCATAGTCG ATTGGAGATGTACGCCTCTTATCGCCATTGATGATCCATCTTTGATGGTCGCTATGCCACCTGCGTTAACTGCCG CAACTGGTATGGACGCTCTAACTCATGCAATTGAGGCATATGTGTCAACAGCTGCCACCCCCATTACTGATGCG TGTGCGGAGAAAGCTATTGTCCTAATTGCTGAATGGTTGCCAAAAGCTGTAGCAAATGGAGATTCTATGGAAG CGAGAGCCGCAATGTGCTATGCTCAGTACCTTGCCGGTATGGCTTTTAACAACGCCAGTTTAGGCTATGTACAT GCTATGGCTCATCAATTGGGTGGTTTTTATAATTTGCCACACGGGGTTTGTAATGCAATACTGCTGCCACATGT CTCCGAATTCAACTTAATAGCAGCCCCTGAGAGGTATGCCAGAATTGCAGAATTGCTGGGTGAAAATATTGGC GGCCTTTCGGCTCATGATGCCGCTAAGGCTGCAGTTTCGGCCATAAGAACATTGAGTACCTCTATTGGGATACC CGCTGGACTAGCTGGTTTGGGAGTGAAGGCAGACGACCATGAAGTCATGGCAAGCAATGCTCAGAAAGACGC TTGCATGTTAACCAATCCAAGGAAGGCTACTCTTGCTCAGGTAATGGCGATCTTCGCGGCTGCAATGTAA
LxADH	ATGAGTGACGTGTTGAAACAGTTCGTGATGCCGAAGACAAATTTGTTTGGTCCTGGAGCGATACAGGAAGTAG GCACTAGATTGAATGACTTAGAAGTTAAAAAGACCTTGATTGTGACTGATGAAGGCTTGCACAAACTAGGTCT TTCCGAACAGATAGCTAACATTATTACTGCTGCAGGAATTGACGTCGCTATCTTTCCAAAAGCTGAACCTAATC CGACTGATCAAAACATTGAAGATGGTATTGCCGTTTATCATGCTGAAAACTGTGATTCCATCGTATCGTTGGGT GGCGGTAGCGCACACGACGCTGCAAAAGGTATTGGCTTGATTGCCTCAAACGGCGGTAGAATCCACGATTATG AAGGCGTAGATAAGTCTCAGAATCCGTTGGTACCCCTGATCGCTATCAATACTACGGCAGGTACAGCCTCTGA AATGACGAGATTTACAATAATAACAGATACTGCACGTAAGGTAAAAATGGCCATTGTTGATAAACACGTGACC CCATTATTGAGCATCAACGATCCGGAATTAATGATTGGTTTACCACCAGCCCTTACTGCTGCGACAGGTCTTGA TGCATTGACACATGCCATCGAATCCTTTGTTTCCACTAATGCTACCCCCATTACGGATGCATGTGCTGAGAAAG TTTTACAATTGGTGCCTGAATATCTACCGAGAGCATACGCAAACGGCGCCGACTTAGAAGCGAGGGAACAAAT GGTATACGCACAATTTTTGGCTGGTATGGCTTTTAATAACGCCTCTTTGGGCTATGTCCATGCAATTGCTCATCA ATTGGGGGGTTTTTATAATCTGCCTCATGGAGTTTGTAATGCTATCCTTTTACCACACGTATGTCGTTTCAATCT TACTGCTAGAACCGAAAGGTTCGCAAGAATCGCCGAGTTGTTAGGTGAAAATGTGGAAGCGTTATCGAAGAGA GATGCTGCTGAGAAGGCTATTGTTGCCATTGAGAATCTTTCTAGAGATCTTAACATACCCTCTGGCTTCAGAGA ATTGGGGGCAAAAGATGAGGATATCGAAATATTGGCTAAAAATGCTATGCTTGATGTATGTGCAGCCACTAAT CCAAGAAAGGCTACTTTGGAGGAAATTAAGCAAATCATCACAAACGCTATGGGCCCTGTCGCTAAAAAAGAG GAGAGCCTAGAAGCAGTCGCATTGTCATAA
PpAOX1	ATGGCTATTCCCGAAGAGTTTGACATTTTAGTACTTGGCGGTGGATCATCCGGTTCTTGCATAGCCGGTAGACT GGCAAATTTAGATCATAGTTTGAAAGTTGGTTTAATCGAAGCTGGTGAGAATAATTTAAATAACCCATGGGTTT ATCTGCCAGGCATATATCCAAGGAATATGAAGTTAGATTCCAAAACCGCGTCTTTCTATACCTCTAACCCAAGT CCACATTTGAATGGAAGGAGAGCCATTGTACCCTGCGCAAATGTTTTAGGTGGTGGCAGCTCCATCAACTTTAT GATGTATACTAGAGGCTCTGCTAGCGACTATGATGACTTCCAAGCAGAGGGTTGGAAAACCAAAGACCTATTG CCGCTAATGAAGAAAACGGAAACCTATCAAAGGGCTTGCAATAATCCAGATATCCACGGTTTCGAAGGACCAA TTAAAGTGTCCTTTGGAAATTACACATACCCAGTTTGTCAGGACTTCTTGAGAGCCTCAGAATCGCAGGGTATA CCCTATGTTGATGACCTAGAAGATCTGGTGACTGCACATGGTGCTGAACATTGGTTGAAATGGATAAATCGTG ATACAGGTCGTAGATCCGACTCCGCTCATGCTTTCGTCCATTCAACGATGAGGAACCATGATAACCTTTATTTG ATATGTAATACCAAAGTCGACAAAATTATAGTCGAAGACGGTAGAGCTGCGGCTGTAAGGACTGTACCAAGCA AGCCACTAAACCCAAAGAAGCCTTCTCACAAGATTTACAGGGCTAGAAAACAGATAGTTTTATCCTGTGGTAC TATTAGTTCACCATTGGTTCTTCAAAGATCAGGATTTGGTGATCCCATCAAATTGAGGGCCGCGGGTGTAAAGC CATTGGTCAACTTGCCAGGTGTCGGAAGAAATTTCCAAGATCATTATTGCTTTTTTAGTCCATACCGTATTAAG CCACAGTATGAATCATTTGATGATTTTGTTAGAGGTGATGCAGAAATACAGAAAAGAGTATTTGATCAGTGGT ACGCTAACGGTACAGGTCCTCTTGCCACAAACGGGATTGAAGCGGGCGTAAAGATTAGACCAACACCCGAAG AATTGTCACAAATGGACGAGAGTTTCCAAGAAGGTTACAGAGAGTATTTCGAAGATAAGCCAGATAAACCAGT AATGCATTATAGTATAATCGCTGGCTTCTTCGGTGATCATACGAAAATACCTCCGGGTAAGTATATGACGATGT TCCATTTTTTAGAGTATCCTTTCTCTAGAGGCAGTATTCATATTACATCACCAGATCCATACGCTGCGCCAGACT TCGATCCAGGATTTATGAATGATGAAAGAGACATGGCTCCAATGGTTTGGGCTTACAAAAAAAGCAGAGAAAC GGCAAGAAGAATGGATCACTTCGCCGGAGAAGTGACATCACATCACCCCTTGTTCCCCTATTCTTCAGAAGCG AGAGCACTAGAAATGGATCTGGAAACTTCAAATGCTTATGGTGGTCCGCTTAATTTGTCTGCTGGTTTGGCACA

	CGGATCTTGGACTCAACCACTAAAAAAACCTACTGCAAAAAATGAAGGGCATGTTACCTCAAATCAAGTGGAG CTGCATCCCGATATCGAATATGACGAAGAAGACGACAAAGCTATTGAGAACTACATTAGAGAGCATACAGAA ACAACTTGGCATTGCTTAGGTACTTGTTCAATCGGGCCAAGAGAAGGATCTAAGATCGTTAAGTGGGGAGGCG TTTTGGACCACAGAAGTAATGTCTACGGAGTTAAAGGTCTAAAAGTGGGTGACTTATCAGTATGCCCAGACAA TGTAGGTTGCAACACTTACACCACTGCTTTACTTATCGGTGAAAAAACTGCCACTTTGGTTGGCGAGGATTTGG GCTATAGCGGAGAAGCATTGGATATGACCGTACCACAATTCAAATTAGGCACCTACGAAAAAACTGGGTTAGC AAGGTTCTGA

Supplementary table 2. Catalogue of primers used for work in this thesis.

Primer pair		Target
Forward	Reverse	
PK0007: ACACACATAAACAAACAAAATGACCACCAACTTTTTCATACCAC	PK0008: TTTTCGGTTAGAGCGGATTTACATCGCATTCTTGATTATTTGTGC	*Bm*MGA3*MDH*
PK0004: ACACACATAAACAAACAAAATGAAGAACACACAAAGTGCCT	PK0002: TTTTCGGTTAGAGCGGATTCACATAGCGTTTTTTATAATTTGT	*Bm*MGA3*MDH2*
PK0005: ACACACATAAACAAACAAAATGACTAATACACAATCAGCGTTCT	PK0006: TTTTCGGTTAGAGCGGATTCACATAGCATTCTTAATAATTTGA	*Bm*MGA3*MDH3*
PK0013: ACACACATAAACAAACAAAATGACCCAAAGAAACTTCTT	PK0014: TTTTCGGTTAGAGCGGATTTACAAAGCATTCTTGATGATTTGAGC	*Bm*PB1*MDH*
PK0009: ACACACATAAACAAACAAAATGACAAAGACAAAGTTTTTTATTCCTTC	PK0010: TTTTCGGTTAGAGCGGATTTATAGTGCATTCTTAATAATTTGTTGT	*Bm*PB1*MDH1*
PK0011: ACACACATAAACAAACAAAATGACAAATACTCAATCAAT	PK0012: TTTTCGGTTAGAGCGGATTTACATCGCATTCTTTATTATCTGA	*Bm*PB1*MDH2*
PK0015: ACACACATAAACAAACAAAATGACTCATCTAAACATAGCTAACAGA	PK0016: TTTTCGGTTAGAGCGGATTTACATTGCAGCCGCGAAGA	*CnMDH2*
PK0017: ACACACATAAACAAACAAAATGAGTGACGTGTTGAAACAGT	PK0018: TTTTCGGTTAGAGCGGATTTATGACAATGCGACTGCTTCT	*LxADH*
PK0019: ACACACATAAACAAACAAAATGGCTATTCCCGAAGAGTTTGA	PK0020: TTTTCGGTTAGAGCGGATTCAGAACCTTGCTAACCCAGT	*PpAOX1*
PK0039: CACAAGGCAATTGACCCACG	PK0044: GTAACTTATAAACTCCGCCAACC	5' junction of *Bm*MGA3*MDH* insert for RuMP vector
PK0037: TGCCCTGGCAATTCAGGC	PK0040: TAATGTTACATGCGTACACGCG	3' junction of *Bm*MGA3*MDH* insert for RuMP vector

PK0039: CACAAGGCAATTGACCCACG	PK0038: AGCTAAGCTTTGTGCAAAGGC	5' junction of *Bm*MGA3*MDH2* insert for RuMP vector
PK0037: TGCCCTGGCAATTCAGGC	PK0040: TAATGTTACATGCGTACACGCG	3' junction of *Bm*MGA3*MDH2* insert for RuMP vector
PK0039: CACAAGGCAATTGACCCACG	PK0042: ATACCAGCTAAACTCTGAGCG	5' junction of *Bm*MGA3*MDH3* insert for RuMP vector
PK0041: ATATGTCTCTACAGGAGCAACG	PK0040: TAATGTTACATGCGTACACGCG	3' junction of *Bm*MGA3*MDH3* insert for RuMP vector
PK0039: CACAAGGCAATTGACCCACG	PK0050: CAGACATGTGGCATATTAACCG	5' junction of *Bm*PB1*MDH* insert for RuMP vector
PK0049: GGCCTATGTTGCTAAAAGAGC	PK0040: TAATGTTACATGCGTACACGCG	3' junction of *Bm*PB1*MDH* insert for RuMP vector
PK0039: CACAAGGCAATTGACCCACG	PK0046: GTTCTCGCCTAAAAGTTCTGC	5' junction of *Bm*PB1*MDH1* insert for RuMP vector
PK0045: CTGGCGTAGCTTTTAATAATGGC	PK0040: TAATGTTACATGCGTACACGCG	3' junction of *Bm*PB1*MDH1* insert for RuMP vector
PK0039: CACAAGGCAATTGACCCACG	PK0048: CTTTAGCCATACGCTCAATAGC	5' junction of *Bm*PB1*MDH2* insert for RuMP vector
PK0047: CCTCACGGTGTTTGTAACGC	PK0040: TAATGTTACATGCGTACACGCG	3' junction of *Bm*PB1*MDH2* insert for RuMP vector
PK0039: CACAAGGCAATTGACCCACG	PK0052: CAGCAATTCTGCAATTCTGGC	5' junction of *CnMDH2* insert for RuMP vector
PK0051: CCACACGGGGTTTGTAATGC	PK0040: TAATGTTACATGCGTACACGCG	3' junction of *CnMDH2* insert for RuMP vector
PK0039: CACAAGGCAATTGACCCACG	PK0054: GGCAACAATAGCCTTCTCAGC	5' junction of *LxADH* insert for RuMP vector
PK0053: CCGAAAGGTTCGCAAGAATCG	PK0040: TAATGTTACATGCGTACACGCG	3' junction of *LxADH* insert for RuMP vector
PK0039: CACAAGGCAATTGACCCACG	PK0056: CTGGTTTATCTGGCTTATCTTCG	5' junction of *PpAOX1* insert for RuMP vector
PK0055: CGAAGAATTGTCACAAATGGACG	PK0040: TAATGTTACATGCGTACACGCG	3' junction of *PpAOX1* insert for RuMP vector
PK001: ACACACATAAACAAAC	PK0002: TTTTCGGTTAGAGCGGA	Rump vector backbone

AAAATGAAGAACACAC AAAGTGCCT	TTCACATAGCGTTTTTTA TAATTTGT	
FH0026: TGTTGGAATAAAAATC AACTATCATCTACTAA	FH0027: CTATTAATAGGATAAAA AATGAAATAATATTCCA AATTTTTA	RAIN promotor cassettes
intronF: ATGGTATGTTTGAGAT GAACAAAATAATAAAG	Ty1-2RL: CGGCCGCAGATCTTGAG AAA	RAIN ORF cassettes
FH0026: TGTTGGAATAAAAATC AACTATCATCTACTAA	PK0072: CGAGATCGGTGAACAGC C	RAIN diagnostic primers straddling promotor and ORF cassettes

Supplementary table 3. Catalogue of genetic constructs used for each experiment.
This table contains a summary of the genetic constructs used for each experiment. Full access to GenBank files for each construct is available at the following link: https://mqoutlook-my.sharepoint.com/:f:/g/personal/philip_kelso_hdr_mq_edu_au/EmkIGlPt8ehOgLYDKj4rt-QBebL4D7zLsEFOZnl5UUP-iQ?e=bczasz**.** Access through this link expires on the 20th February 2023. For subsequent access at a later date please contact the author directly via email: philip.kelso@HDR.mq.edu.au**.**

Experiment number	Files
Experiment 1	pRS-416, empty vector.gb RuMP cycle pRS416 -MDH control (BmHPS _ BmPHI).gb RuMP cycle pRS416, BmMGA3MDH _ (BmHPS_BmPHI).gb RuMP cycle pRS416, BmMGA3MDH2 _ (BmHPS_BmPHI).gb RuMP cycle pRS416, BmMGA3MDH3 _ (BmHPS_BmPHI).gb RuMP cycle pRS416, BmPB1MDH _ (BmHPS_BmPHI).gb RuMP cycle pRS416, BmPB1MDH1 _ (BmHPS_BmPHI).gb RuMP cycle pRS416, BmPB1MDH2 _ (BmHPS_BmPHI).gb RuMP cycle pRS416, BsMDH _ (BmHPS_BmPhi).gb RuMP cycle pRS416, CnMDH2 _ (BmHPS_BmPHI).gb RuMP cycle pRS416, LxADH _ (BmHPS_BmPHI).gb RuMP cycle pRS416, PpAOX1 _ (BmHPS_BmPHI).gb
Experiment 2	RAIN ORF S. cerevisiae SOL4.gb RAIN ORF S. cerevisiae TAL1.gb RAIN ORF S. cerevisiae TKL1.gb RAIN ORF S. cerevisiae TKL2.gb RAIN ORF S. cerevisiae ZWF1.gb RAIN ORF S.cerevisiae GND1.gb RAIN promoter cassettes.gb RAIN promoter pADH2.gb RAIN promoter pCYC1.gb RAIN promoter pPDA1.gb RAIN promoter pSSA1.gb RAIN promoter pSSB1.gb RAIN promoter pTEF1.gb RAIN ORF B. methanolicus HPS.gb RAIN ORF B. methanolicus PHI.gb RAIN ORF M. buryatense HPS.gb RAIN ORF M. gastri HPS.gb

	RAIN ORF M. gastri HPS_PHI fusion.gb RAIN ORF M. gastri PHI.gb RAIN ORF P. pastoris DAK.gb RAIN ORF P. pastoris DAS1.gb RAIN ORF P. pastoris DAS2.gb RAIN ORF S. cerevisiae GND2.gb RAIN ORF S. cerevisiae NQM1.gb RAIN ORF S. cerevisiae RKI1.gb RAIN ORF S. cerevisiae RPE1.gb RAIN ORF S. cerevisiae SOL3.gb
Experiment 3	pRS415 empty vector.gb Xylose utilisiation parthway pRS415, PsXYL1 _ PsXYL2 _ ScXKS1.gb