

Engineering of FLS and RuMP pathway modules for synthetic methylotrophy in yeast

Vanessa Wegat

Vollständiger Abdruck der von dem TUM Campus Straubing für Biotechnologie und Nachhaltigkeit der Technischen Universität München zur Erlangung einer

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Ing Jakob Burger

Prüfende der Dissertation:

1. Prof. Dr. Volker Sieber
2. Prof. Dr. Bastian Blombach

Die Dissertation wurde am 20.04.2023 bei der Technischen Universität München eingereicht und durch den TUM Campus Straubing für Biotechnologie und Nachhaltigkeit am 12.08.2023 angenommen.

*„Wir sprechen mit Ihnen wie immer unterschiedliche Sprachen, aber die Dinge, über die wir sprechen, ändern sich dadurch nicht.“*

*Michail Bulgakow, Der Meister und Margarita*

## ACKNOWLEDGMENTS

This thesis is only a small part of my scientific journey and I would like to take this opportunity to thank all the people who have supported me during this exciting time.

First, I would like to express my sincere gratitude to my supervisor Prof. Dr. Volker Sieber, who gave me the possibility to write my dissertation in cooperation with the Fraunhofer IGB. Thank you, Volker, for the exciting topic, fruitful discussions, and the support at any time throughout the past years.

Furthermore, I would like to thank Prof. Dr. Bastian Blombach very much for taking over as my second thesis examiner and for providing the BioLector.

My special thanks also go to my supervisor at the Fraunhofer IGB, Dr. Jonathan Thomas Fabarius, whose enthusiasm, effort, and patience, added remarkably to my scientific education. I am glad that you sparked my interest in such an interesting field and that you always took the time to support me whenever I was in need.

Sincere thanks to the whole TUM CBR group and especially all the other PhD students for their kindness and encouragement during my thesis! Besides sharing expertise and giving useful advice, they also contributed to a great time outside of the lab. I want to thank Samed and Manuel in particular, who also significantly contributed to the engineering of the FLS enzyme.

To the same extent, I would like to thank all the staff at the Fraunhofer IGB in Straubing for the great atmosphere and many enjoyable moments. I want to thank Dr. Arne Roth in particular, who supported me not only in my doctoral work but also encouraged me to look at projects outside their academic relevance. I would like to thank the technicians of the biology department, Carina, Christina, Julia, Melli, Patricia and Tanja. Thank you for introducing me to the lab, your helpfulness in general and for the time spent inside and outside of work in particular. My sincere thanks also go equally to the other researchers at the Fraunhofer IGB as well as Michael H. and Michael R. for sharing their knowledge and helpful advice. I am grateful to Andreas, Dhananjai, Ferdinand, Leonardo, Melanie, Paul and Steffen for the very good cooperation, many useful conversations, joint lunch and coffee breaks, and everything beyond work. In addition, I want to thank the students I had the privilege of supervising and collaborating with. Especially to Isadora and Maurice, your drive kept me on my toes and made teaching a pleasure.

My scientific journey would have never resulted in this PhD thesis without the prior encouragement of Dr. Robert Blum, Prof. Dr. Martin Eilers, Prof. Dr. Elmar Wolf and Dr. Klaudia Marcseková. I am forever thankful for the time you put into teaching and mentoring me and giving me space to explore various projects and ideas.

Thank you as well to my fellow students in Würzburg and especially to Basti and Fabienne, who made my time at university an experience I will never forget. I am grateful for your endless help and motivation during the time we know each other. “Part of the crew, part of the ship!”

Finally, yet importantly, I want to thank the wonderful people outside of my scientific cosmos: My friends and family. You are fueling my life with your unconditional love and support. Thanks for always putting things into perspective! Lucas, you are enriching my life on a daily basis and provide the balance I need. Papa, danke, dass du mich immer herausforderst, die beste Version meiner selbst zu sein, und mich dazu bringst, den Status quo zu hinterfragen. Du bist der Grund für mein frühes Interesse an der Wissenschaft. Liebe Mama, in wissenschaftlichen Publikationen sind üblicherweise die wichtigsten Personen die Erst- und Letztautoren, welche das Projekt auf den Weg gebracht haben. So ist es auch in dieser Danksagung. Danke für deine grenzenlose Unterstützung und das Ermöglichen unzähliger Perspektiven!

## SUMMARY

Monocarbon compounds such as methanol or formate have become attractive feedstock for biotechnological processes because they can be produced sustainably from CO<sub>2</sub> and renewable energy at large scale without required fertile land. In particular, methylotrophic microorganisms can convert such carbon sources into a wide array of value-added compounds. In naturally occurring methylotrophs methanol is assimilated via the bacterial ribulose monophosphate (RuMP) or the xylulose monophosphate (XuMP) pathway in yeast. Since the metabolism of native methylotrophic organisms is yet not fully understood and the production performance of desirable metabolites often limited, synthetic methylotrophs have taken up the challenge to produce fine and bulk chemicals from C<sub>1</sub> compounds. The improved genetic accessibility and availability of genetic tools and protocols for commonly used host organisms are advantageous to engineer such tailor-made cell factories. In this regard, the engineering of non-methylotrophic platform organisms to utilize C<sub>1</sub> compounds via the implementation of heterologous methanol assimilation pathways is an attractive alternative.

This thesis aimed at enabling a *Y. lipolytica* strain to grow on methanol. The first part of this work involved the selection of a suitable host strain based on rational criteria, such as growth characteristics and methanol and formaldehyde tolerance. Based on the aforementioned metrics the *Y. lipolytica* strain DSM3286 was chosen as a host for synthetic methylotrophy. Compared to the commonly used PO1f strain, DSM3286 not only tolerates higher amounts of methanol and other relevant C<sub>1</sub> intermediates, but also achieves considerably higher biomass in minimal media.

In order to enable growth on methanol, the bacterial RuMP and the artificial FLS pathway were selected for expression in *Y. lipolytica*. The latter is based on the synthetic enzyme formolase, which converts formaldehyde directly into dihydroxyacetone. For this methanol assimilation pathway, only two enzymes need to be implemented in *Y. lipolytica*. However, as the activity of the parental formolase is very low, another aim of this dissertation was to engineer the FLS enzyme to make it more suitable for *in vivo* application. Therefore, the second part of this work focused on the engineering of the FLS enzyme. The key to success in this study was the high-throughput screening using an automatized procedure. To overcome the issues associated by applying FLS *in vivo* (i.e., high K<sub>M</sub> and a limited maximum reaction velocity), the enzyme was subjected to iterative rounds of directed evolution. A combined, 'semi-rational' approach was used to reduce the limitations of both

rational design and directed evolution. Thereby, a novel FLS variant was identified exhibiting higher expression levels and enabling better formaldehyde tolerance in *E. coli*. In addition, the engineered FLS variant L482Q/I557G depicts a  $K_M$  for the formose reaction of 23.5 mM, which is roughly a 30 % decrease compared to the parental FLS enzyme and provides increased thermostability by 5 °C.

Finally, implementation of synthetic methylotrophy in *Y. lipolytica* was tested using FLS or RuMP pathway modules in a randomized synthetic biology approach. The first step in both pathways is the methanol oxidation to formaldehyde. Therefore, three different variants were chosen and tested in a combinatorial manner. The clones that showed the best growth phenotype were isolated and characterized using a high-throughput microbioreactor and later in shake flasks. Several strains exhibited improved growth in methanol containing media supplemented with yeast extract compared to the empty vector control. These strains could be interesting starting points for further engineering and adaptive laboratory evolution to eventually enable growth on methanol as the sole carbon source.

## LIST OF PUBLICATIONS

**Wegat, V.**, Fabarius, J.T., Sieber, V. Synthetic methylotrophic yeasts for the sustainable fuel and chemical production. (2022) *Biotechnology for Biofuels and Bioproducts*, 15 (1), art. no. 113.

DOI: 10.1186/s13068-022-02210-1.

Güner, S., **Wegat, V.**, Pick, A., Sieber, V. Design of a synthetic enzyme cascade for the *in vitro* fixation of a C<sub>1</sub> carbon source to a functional C<sub>4</sub> sugar. (2021) *Green Chemistry*, 23 (17), pp. 6583-6590.

DOI: 10.1039/d1gc02226a

Fabarius, J.T., **Wegat, V.**, Roth, A., Sieber, V. Synthetic Methylotrophy in Yeasts: Towards a Circular Bioeconomy. (2021) *Trends in Biotechnology*, 39 (4), pp. 348-358.

DOI: 10.1016/j.tibtech.2020.08.008

### **Experimental contributions and planned publications**

The content of chapters 3.2 and 4.1 is partially adapted from a manuscript, intended for publication. Manuel Döring contributed to the establishment of the robot-aided high-throughput screening process and developed the pitch function (Figure S3). Samed Güner majorly contributed to the *in-silico* studies, provided the PyMol images used in Figure 13 and Figure 34 and supported the fitting of the Michaelis-Menten curve.

# CONTENTS

<b>1 INTRODUCTION</b>	<b>1</b>
<b>1.1 C<sub>1</sub> COMPOUNDS AS PROMISING MICROBIAL FEEDSTOCK</b>	<b>1</b>
<b>1.2 NATIVE METHYLOTROPHY</b>	<b>2</b>
1.2.1 OVERVIEW OF METHANOL ASSIMILATION AND DISSIMILATION PATHWAYS	3
1.2.2 APPLICATION EXAMPLES OF NATURAL METHYLOTROPHS	8
<b>1.3 SYNTHETIC METHYLOTROPHY</b>	<b>9</b>
1.3.1 PATHWAY SELECTION TO ENABLE SYNTHETIC METHYLOTROPHY	10
1.3.2 APPLICATION EXAMPLES OF SYNTHETIC METHYLOTROPHS	13
<b>1.4 THE OLEAGINOUS YEAST <i>Y. LIPOLYTICA</i></b>	<b>15</b>
1.4.1 <i>Y. LIPOLYTICA</i> AS AN INDUSTRIAL PLATFORM ORGANISM	15
1.4.2 THE ADVANCING SYNTHETIC TOOLKIT FOR ENGINEERING <i>Y. LIPOLYTICA</i>	17
1.4.3 CURRENT STATE OF SYNTHETIC C <sub>1</sub> APPROACHES IN <i>Y. LIPOLYTICA</i>	20
<b>1.5 OBJECTIVE</b>	<b>21</b>
<b>2 MATERIAL AND METHODS</b>	<b>23</b>
<b>2.1 MATERIAL</b>	<b>23</b>
2.1.1 CHEMICALS, MEDIA AND BUFFERS	23
2.1.2 OLIGONUCLEOTIDES AND PLASMIDS	29
2.1.3 MICROORGANISMS	36
2.1.4 CONSUMABLES, EQUIPMENT AND SOFTWARE	36
<b>2.2 METHODS</b>	<b>40</b>
2.2.1 MICROBIOLOGICAL METHODS	40
2.2.2 MOLECULAR METHODS	43
2.2.3 BIOCHEMICAL METHODS	49
2.2.4 ANALYTICS	53
<b>3 RESULTS</b>	<b>55</b>
<b>3.1 CHARACTERIZATION OF THE OLEAGINOUS YEAST <i>Y. LIPOLYTICA</i> AS A HOST FOR SYNTHETIC METHYLOTROPHY</b>	<b>55</b>
3.1.1 CHARACTERIZATION OF DSM3286 AND PO1F ON VARIOUS SUBSTRATES	55
3.1.2 METHANOL AND C <sub>1</sub> INTERMEDIATE TOLERANCE	56



<b>3.2 ROBOT-AIDED HIGH-THROUGHPUT ENGINEERING OF FLS ENZYME FOR <i>IN VIVO</i> SYNTHETIC ONE-CARBON METABOLISM</b>	<b>59</b>
3.2.1 EXPRESSION OF PARENTAL FLS AND ACTIVITY DETERMINATION	60
3.2.2 MODIFICATION OF THE PARENTAL FLS	62
3.2.3 DEVELOPMENT OF A ROBOT-AIDED HIGH-THROUGHPUT PROCESS	63
3.2.4 SCREENING OF MUTANT LIBRARIES AND VARIANT SELECTION	66
3.2.5 CHARACTERIZATION OF VARIANT L482Q/I557G	69
3.2.6 APPLICATION OF VARIANT L482Q/I557G IN <i>E. COLI</i> AS MEDIATOR FOR IMPROVED FALD TOLERANCE	71
<b>3.3 ENGINEERING OF <i>Y. LIPOLYTICA</i> DSM3286 FOR METHANOL ASSIMILATION</b>	<b>73</b>
3.3.1 METABOLIC PATHWAY DESIGN TO ENABLE METHANOL ASSIMILATION IN <i>Y. LIPOLYTICA</i>	73
3.3.2 OPTIMIZATION OF THE GOLDEN-GATE PROTOCOL	74
3.3.3 GOLDEN-GATE BASED ONE-POT APPROACH	75
3.3.4 ESTABLISHING A HIGH-THROUGHPUT SCREENING FOR STRAINS WITH IMPROVED METHANOL ASSIMILATION	78
3.3.5 SCREENING OF <i>Y. LIPOLYTICA</i> STRAINS WITH IMPROVED METHANOL ASSIMILATION	81
3.3.6 CHARACTERIZATION OF INITIAL SYNTHETIC METHYLOTROPHIC <i>Y. LIPOLYTICA</i> STRAINS	84
<b>4 DISCUSSION</b>	<b>87</b>
<b>4.1 CARBON ASSIMILATION REDESIGNED – THE FLS ENZYME</b>	<b>87</b>
<b>4.2 <i>YARROWIA LIPOLYTICA</i> – A NON-CONVENTIONAL HOST</b>	<b>92</b>
<b>4.3 INSTALLATION OF METHYLOTROPHIC MODULES IN <i>Y. LIPOLYTICA</i></b>	<b>95</b>
<b>4.4 SYNTHETIC METHANOL AUXOTROPHY AND POTENTIAL SOLUTIONS</b>	<b>101</b>
<b>4.5 CONCLUSION</b>	<b>104</b>
<b>4.6 FUTURE PERSPECTIVES</b>	<b>105</b>
<b>5 REFERENCES</b>	<b>107</b>
<b>6 APPENDIX</b>	<b>116</b>
<b>6.1 ADDITIONAL DATA</b>	<b>118</b>
<b>6.2 HPLC</b>	<b>127</b>
<b>6.3 GENE SEQUENCES</b>	<b>130</b>

# LIST OF TABLES

TABLE 1: OVERVIEW OF FORMALDEHYDE ASSIMILATION PATHWAYS AND THEIR CHARACTERISTICS.....	10
TABLE 2: PRODUCTION OF IMPORTANT CHEMICALS FROM CONVENTIONAL FEEDSTOCK BY <i>Y. LIPOLYTICA</i> .....	16
TABLE 3: LIST OF THE CHEMICALS USED WITHIN THIS WORK AND THEIR SUPPLIER. ....	23
TABLE 4: LIST OF THE ENZYMES USED WITHIN THIS WORK AND THEIR SUPPLIER. ....	25
TABLE 5: LIST OF THE RESTRICTION ENZYMES USED WITHIN THIS WORK, THEIR RESTRICTION SITE AND THEIR SUPPLIER.....	25
TABLE 6: LIST OF KITS USED WITHIN THIS WORK AND THEIR SUPPLIER. ....	25
TABLE 7: COMPOSITION OF USED MEDIA. ....	26
TABLE 8: COMPOSITION OF BUFFERS AND ANTIBIOTICS. ....	27
TABLE 9: OLIGONUCLEOTIDES USED IN THIS WORK.....	29
TABLE 10: PLASMIDS USED IN THIS WORK.....	31
TABLE 11: SYNTHESIZED GENES USED IN THIS WORK.....	35
TABLE 12: DIFFERENT MICROORGANISMS USED IN THIS WORK. ....	36
TABLE 13: EQUIPMENT USED WITHIN THIS WORK.....	36
TABLE 14: SOFTWARE USED IN THIS WORK. ....	38
TABLE 15: RESTRICTION REACTION. ....	44
TABLE 16: COMPOSITION AND PROGRAM FOR QUIKCHANGE PCR.....	45
TABLE 17: COMPOSITION AND PROGRAM FOR COLONY PCR.....	46
TABLE 18: FORMULATION OF SEPARATING AND STACKING GEL FOR SDS-PAGE ANALYSIS.....	51
TABLE 19: KINETIC CHARACTERIZATION OF WILD TYPE FLS AND VARIANT L482Q/I557G WITH FALD AS SUBSTRATE. ....	70
TABLE 20: MELTING POINT ( $T_M$ IN °C) FLS WILD TYPE AND VARIANT IN 100 mM PHOSPHATE BUFFER (PH 8.0) CONTAINING 1 mM $MgSO_4$ AND 0.1 mM TPP. ....	71
TABLE 21: <i>Y. LIPOLYTICA</i> STRAINS SELECTED FOR RESCREENING BASED ON THEIR HIGHER GROWTH RATE AND/OR HIGHER FINAL $OD_{600}$ COMPARED TO THE CORRESPONDING EMPTY VECTOR.....	82
TABLE 22: COMPARISON OF THE CONVENTIONAL HOST ORGANISMS <i>E. COLI</i> AND <i>S. CEREVISIAE</i> WITH <i>Y. LIPOLYTICA</i> FOR APPLICATION AS HOSTS FOR SYNTHETIC METHYLOTROPHY. ....	92
TABLE 23: SUMMARY OF PREVIOUS WORK ON THE DEVELOPMENT OF SYNTHETIC METHYLOTROPHY IN <i>Y. LIPOLYTICA</i> AND COMPARISON TO THIS WORK. ....	98

# LIST OF FIGURES

FIGURE 1: OVERVIEW OF METHANOL ASSIMILATION TO BIOMASS OR DISSIMILATION TO CO <sub>2</sub> . .....	3
FIGURE 2: METABOLIC ACCESS POINTS OF SYNTHETIC FLS, REDUCTIVE GLYCINE AND NATIVE RUMP PATHWAY INTO GLYCOLYTIC YEAST METABOLISM. ....	11
FIGURE 3: GRAPHICAL ABSTRACT OF CHOSEN METABOLIC PATHWAYS FOR SYNTHETIC METHANOL ASSIMILATION IN <i>YARROWIA LIPOLYTICA</i> . ....	22
FIGURE 4: THE CORRELATION BETWEEN DRY BIOMASS WEIGHT AND OPTICAL DENSITY FOR GROWTH OF TWO <i>Y. LIPOLYTICA</i> STRAINS. ....	56
FIGURE 5: EFFECT OF METHANOL AND C <sub>1</sub> INTERMEDIATES ON THE GROWTH BEHAVIOR OF <i>Y. LIPOLYTICA</i> DSM3286. ....	57
FIGURE 6: EFFECT OF METHANOL AND C <sub>1</sub> INTERMEDIATES ON THE GROWTH BEHAVIOR OF <i>Y. LIPOLYTICA</i> PO1F. ....	58
FIGURE 7: EFFECT OF GLYCOLALDEHYDE ON THE GROWTH BEHAVIOR OF <i>Y. LIPOLYTICA</i> DSM3286. ....	59
FIGURE 8: FORMOSE REACTION CATALYZED BY THE FORMOLASE (FLS) ENZYME. ....	60
FIGURE 9: SDS-PAGE PATTERN OF <i>E. COLI</i> BL21(DE3) pET24a::FLS-C-HIS EXPRESSION TEST AFTER 16 H OF INDUCTION AT 18 °C, AT OD <sub>600</sub> ~0.6 WITH 0.5 mM IPTG (LEFT GEL) AND INDUCTION AT ROOM TEMPERATURE AT OD <sub>600</sub> ~1 WITH 1 mM IPTG (RIGHT GEL).....	60
FIGURE 10: NADH-DEPENDENT GDH ASSAY USED TO MEASURE FLS ACTIVITY INDIRECTLY. ....	61
FIGURE 11: PRELIMINARY DATA FOR HIGH-THROUGHPUT SCREENING. ....	61
FIGURE 12: SCHEME OF THE ROBOT-AIDED HIGH-THROUGHPUT SCREENING WORKFLOW. ....	62
FIGURE 13: STRUCTURE OF FLS TO VISUALIZE RELEVANCE OF IDENTIFIED AMINO ACID RESIDUES IN CLOSE DISTANCE TO SUBSTRATE AND COFACTORS TPP AND MAGNESIUM.....	63
FIGURE 14: OPTIMIZATION EFFORTS TOWARDS A MORE RELIABLE CULTIVATION WITHIN THE SCREENING PROCESS. ....	64
FIGURE 15: PELLETS OF <i>E. COLI</i> BL21(DE3) INDICATE AN EVEN FORMATION OF BIOMASS INDEPENDENT FROM THE GENETIC BACKGROUND (E.G., FLS MUTATION). ....	65
FIGURE 16: ABSORBANCE SPECTRUM FOR TB MEDIUM. ....	66
FIGURE 17: RESULTS OF THE ROBOT-AIDED HIGH-THROUGHPUT SINGLE-MUTANT (H29, Q113, L482, I556 AND I557) SCREENING OF THE FLS ENZYME. ....	66
FIGURE 18: ROBOT-AIDED HIGH-THROUGHPUT RESCREENING OF ENGINEERED FLS CLONES USING THE FLS ENZYME ASSAY. ....	67
FIGURE 19: RESULTS OF THE ROBOT-AIDED HIGH-THROUGHPUT DOUBLE-MUTANT (L482NNK/I557NNK) SCREENING OF THE FLS ENZYME.....	68
FIGURE 20: CHARACTERIZATION OF DOUBLE MUTANT VARIANTS. ....	68

FIGURE 21: CHARACTERIZATION OF VARIANT L482Q/I557G AT LOW (13.4 mM, A), MID (67 mM, B) AND HIGH (134 mM, C) FORMALDEHYDE CONCENTRATIONS..	69
FIGURE 22: MICHAELIS MENTEN KINETIC FOR FORMOSE REACTION USING PARENTAL FLS (BLUE) OR VARIANT L482Q/I557G (GREEN).....	70
FIGURE 23: EVALUATION OF FLS VARIANT L482Q/I557G FOR <i>IN VIVO</i> APPLICATION.....	72
FIGURE 24: OVERVIEW OF GENES CHOSEN FOR SYNTHETIC METHANOL ASSIMILATION IN <i>Y. LIPOLYTICA</i> .....	73
FIGURE 25: IMPROVEMENT OF GOLDEN GATE PROTOCOL.....	74
FIGURE 26: OVERVIEW AND RESULTS OF GOLDEN GATE ASSEMBLY LEVEL 1.....	76
FIGURE 27: OVERVIEW AND RESULTS OF GOLDEN GATE ASSEMBLY LEVEL 2.....	77
FIGURE 28: METHANOL EVAPORATION IN K1 MINIMAL MEDIA AT 30 °C AND 200 RPM.....	79
FIGURE 29: OPTIMIZATION OF THE CULTIVATION CONDITIONS FOR LOWER METHANOL EVAPORATION USING THE BIOLECTOR XT.....	80
FIGURE 30: ESTABLISHING A HIGH-THROUGHPUT SCREENING FOR STRAINS WITH IMPROVED METHANOL ASSIMILATION.....	81
FIGURE 31: RESCREENING OF VARIOUS RECOMBINANT <i>Y. LIPOLYTICA</i> STRAINS ON METHANOL.....	83
FIGURE 32: CELL GROWTH OF <i>Y. LIPOLYTICA</i> STRAINS AD_N_ADH_HPS_PHI1 (A), AC_N_MDH2_FLS*2 (B) AND AC_N_MDH3_FLS3 (C) IN 25 mL OF MM WITH 15.8 g L <sup>-1</sup> METHANOL AND 0.5 g L <sup>-1</sup> YEAST EXTRACT.....	85
FIGURE 33: CELL MORPHOLOGY OF <i>Y. LIPOLYTICA</i> STRAINS AD_N_ADH_HPS_PHI 1 (A) AND AC_N_MDH3_FLS 3 (B) COMPARED TO THEIR CORRESPONDING EMPTY VECTORS.....	86
FIGURE 34: VISUALIZATION OF THE STRUCTURE OF FLS WILD TYPE AND VARIANT L482Q/I557G.....	89
FIGURE 35: THE BUILDING BLOCKS FOR ESTABLISHING SYNTHETIC METHYLOTROPHY IN YEASTS.....	103

# LIST OF ABBREVIATIONS AND ACRONYMS

## Prefixes and units

n	nano
k	kilo
μ	micro
m	milli
f	femto
°C	Degree Celsius
A	Ampere
a. u.	Arbitrary units
b	bases
d	day
Da	Dalton
g	gram
h	hour
L	litre
min	minutes
M	mol L <sup>-1</sup>
s	second
v/v	volume per volume
v/w	weight per volume

## Abbreviations

1,3BPG	1,3-Bisphosphoglycerate
6PGL	6-phosphogluconolactone
A	Adenine
aa	Amino acid
ADH	Alcohol dehydrogenase
ADP	Adenosine 5'-(trihydrogen diphosphate)

ALE	Adaptive laboratory evolution
Amp	Ampicillin
AOX	Alcohol oxidase
APS	Ammoniumpersulfate
ARP	Autonomously replicating plasmid
ATP	Adenosine 5'-(tetrahydrogen triphosphate)
<i>B. methanolicus</i>	<i>Bacillus methanolicus</i>
BLAST	Basic Local Alignment Search Tool
bp	Base pair
C	Cytosine
c	Concentration [mol L <sup>-1</sup> ]
<i>C. boidinii</i>	<i>Candida boidinii</i>
C <sub>1</sub>	One-carbon molecule
Cas9/dCas9	CRISPR associated protein 9/ deactivated Cas9
CBB	Calvin-Benson-Bassham
CCU	Carbon Capture and Utilization
CDW	Cell dry weight
CO <sub>2</sub>	Carbon dioxide
CRISPR	Clustered regularly interspaced short palindromic repeats
CSM	Complete supplement mixture
CTA	Catalase
CV	Column volume
C-terminal/terminus	Carboxy-terminal/terminus
DAK	Dihydroxyacetone kinase
DAS	Dihydroxyacetone synthase
DHA	1,3-Dihydroxypropan-2-one
DHAP	Dihydroxyacetonephosphate
DNA	Deoxyribonucleic acid
Dntp's	Deoxyribonucleoside-5'-triphosphate (dATP, dCTP, dGTP, dTTP)
DTT	Dithiothreitol

ddH <sub>2</sub> O	Bidestilled water
E4P	Erythrose 4-phosphate;
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	2,2',2'',2'''-(Ethane-1,2-diyldinitrilo)tetraacetic acid
EtOH	Ethanol
e.g.	exempli gratia
F1,6BP	Fructose 1,6-bisphosphate
FA	Formate
FALD	Formaldehyde
FDH	Formate dehydrogenase
FeSO <sub>4</sub>	Iron(II) sulfate
FLD	Formaldehyde dehydrogenase
FLS	Formolase
G	Guanine
G3P	Glyceraldehyde 3-phosphate
G6P	Glucose 6-phosphate;
GALD	Hydroxyacetaldehyde
Glu	Glucose
Gly	Glycine
GSH	Glutathione
H4F	Tetrahydrofolate
H4MPT	Tetrahydromethanopterin
H6P	Hexulose 6-phosphate
<i>H. polymorpha</i>	<i>Hansenula polymorpha</i>
H (His)	Histidine
HCl	Hydrochloric acid
HPLC	High performance liquid chromatography
HPS	3-Hexulose-6-phosphate synthase
HR	Homologues recombination
Hygro	Hygromycin B

I (Ile)	Isoleucine
i.e.	id est
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranoside
K	Nucleotides G or T (keto)
KH <sub>2</sub> PO <sub>4</sub>	Monopotassium phosphate
Kan	Kanamycin
L (Leu)	Leucine
LB	Lysogenic broth
<i>M. extorquens</i>	<i>Methylobacterium extorquens</i>
MDH	Methanol dehydrogenase
MeOH	Methanol
MgCl <sub>2</sub>	Magnesium chloride
MnCl <sub>2</sub>	Manganese chloride
MnSO <sub>4</sub>	Manganese sulfate
MTHF	5,10-methylenetetrahydrofolate
MTP	Microtiter plate
N	Nucleotides A, C, G, or T (any)
NaCl	Sodium chloride
NAD <sup>+</sup> /NADH/H <sup>+</sup>	Oxidized and reduced nicotinamide adenine dinucleotide
NADP <sup>+</sup> /NADPH <sup>+</sup>	Oxidized and reduced nicotinamide adenine dinucleotide phosphate
NaOH	Sodium hydroxide
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Ammonium sulfate
NHEJ	Non-homologous end joining
Nou	Nourseothricin
N-terminal/terminus	Amino-terminal/terminus
O <sub>2</sub>	Oxygen
OD <sub>600</sub> /OD <sub>720</sub>	Optical density at 600/720 nm
<i>P. pastoris</i>	<i>Pichia pastoris</i>
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction



pH	Negative decadic logarithm of proton concentration
PHI	6-Phosphate-3-hexuloisomerase
PGA	3-phosphoglycerate
POI	Protein of interest
PQQ	Pyrrroloquinoline quinone
Q (Gln)	Glutamine
RNA	Ribonucleic acid
rpm	Rounds per minute
RT	Room temperature
Ru5P	Ribulose 5-phosphate
RuBP	1,5-bisphosphate
RuMP	Ribulose monophosphate
S7P	Sedoheptulose 7-phosphate
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
S-FG	S-formylglutathione
S-HMG	S-hydroxymethylglutathione
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
sgRNA	Single guide RNA
SOC	Super Optimal Broth with glucose
TE	Tris-EDTA buffer
TRIS	Tris(hydroxymethyl) aminomethane
var	Variant
wt	Wild type
Xu5P	Xylulose 5-phosphate
XuMP	Xylulose monophosphate
<i>Y. lipolytica</i>	<i>Yarrowia lipolytica</i>
YPD	Yeast extract Peptone Dextrose
ZnSO <sub>4</sub>	Zinc sulfate

# 1 INTRODUCTION

## 1.1 C<sub>1</sub> compounds as promising microbial feedstock

“Encountering climate change and mitigating its impact on the environment, on our economies and on the society is the defining challenge of our time. Due to the rapid growth of the world’s population, the demand for energy is increasing dramatically every year and CO<sub>2</sub>-neutral solutions are desperately needed. The use of conventional energy sources (e.g., oil, coal and natural gas) represents by far the largest source of greenhouse gas emissions from human activities and thus contributes significantly to global warming. The depletion of fossil fuels and historical and on-going geopolitical conflicts are further reasons to commit to renewable energy sources. In this regard, the capture of CO<sub>2</sub> from an (industrial) process or even directly from the air and its subsequent utilization (Carbon Capture and Utilization, CCU) is one option to reduce industrial emissions and realizing a circular economy, provided that the energy used in capturing and converting the CO<sub>2</sub> is zero carbon. In general, CCU refers to the capture, transport and use of carbon compounds such as carbon monoxide or carbon dioxide, in which the carbon is fed into at least one further utilization cycle.”<sup>[1]</sup> The combination of CO<sub>2</sub> conversion into one-carbon (C<sub>1</sub>) compounds, such as methane, formate, or methanol and its subsequent use as a feedstock in microbial fermentation depicts one innovative option towards CCU.

Utilization of green methanol as a microbial feedstock provides several advantages: It is a water-miscible liquid, provides a high energy density and can be derived from various renewable processes such as heterogeneous chemical catalysis or electrochemical CO<sub>2</sub> reduction<sup>[2]</sup>. In addition, cultivation on methanol can support increased biomass and product yields as its combustion provides more energy ( $\Delta G^{\circ} = - 4276.6 \text{ kJ mol}^{-1}$ ) compared to

glucose oxidation ( $\Delta G^{\circ} = 2870 \text{ kJ mol}^{-1}$ )<sup>[3]</sup>. In contrast to other inexpensive plant-derived materials such as molasses, methanol is purer and its production does not compete with the use of fertile land<sup>[4]</sup>. In the past, methanol was produced by the destructive distillation (i.e., without air) of wood and was therefore known as ‘wood alcohol’<sup>[5]</sup>. Today, methanol is mainly obtained industrially from synthesis gas (syngas). Hydrogen, carbon monoxide and to some extent carbon dioxide react over a catalyst, usually a mixture of copper and zinc oxides, supported on alumina, to form methanol<sup>[6]</sup>. In contrast, “CO<sub>2</sub>-dependent methanol production is climate-friendly and independent of fossil resource usage and consequently increases the environmental benefit while reducing CO<sub>2</sub> emissions. When considering future trends in energy supply and demand, it is important to acknowledge that compelling market factors will continue to strongly influence the price of energy. Fossil fuels are not only the cause of environmental pollution and climate crisis, but also of historical and on-going conflicts and it is anticipated that future prices will increase<sup>[7]</sup>.”<sup>[1]</sup> A sustainable methanol production is therefore slowly gaining attractiveness and will eventually allow the utilization of CO<sub>2</sub> as a versatile indirect fermentation substrate. So-called native methylotrophic organisms are able to valorize methanol as their sole carbon and energy source for the production of single-cell and recombinant proteins, enzymes, and commodity chemicals. Although notable progress has been made in recent years, constraints regarding the understanding of the methylotrophic metabolism still exist. “For this reason, the development of synthetic methylotrophic cell factories based on the integration of natural or artificial methanol assimilation pathways in biotechnologically relevant microorganisms is receiving special attention.”<sup>[1]</sup>

## 1.2 Native methylotrophy

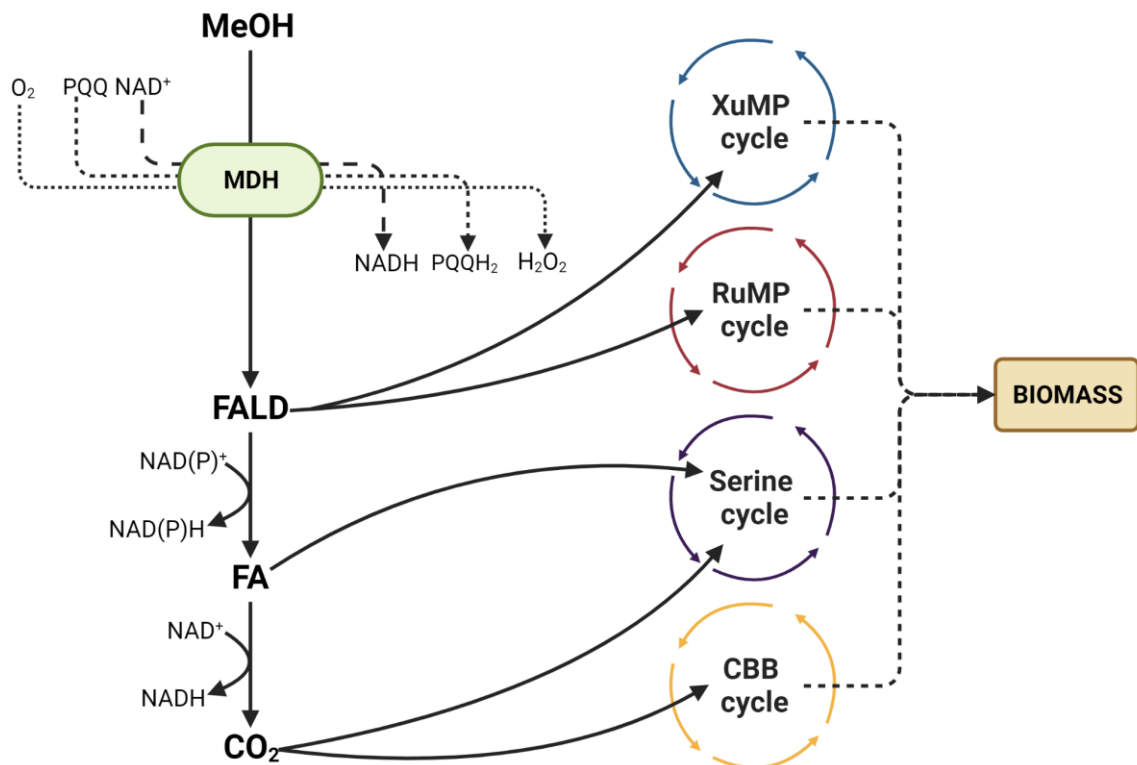
In nature, methanol occurs in many different habitats. It can be found in plants as a by-product of cell wall biosynthesis<sup>[8]</sup>, or is metabolized by microorganisms from pectin found in fruits<sup>[9]</sup>. Even though methanol is a challenging substrate for most microorganism due to its toxicity and the related toxic by-products, a divergent group of microbes have well-adapted to harness reduced monocarbon compounds for growth and energy generation. Unlike glycolytic microorganisms<sup>[10]</sup>, methylotrophic bacteria or yeasts use carbon-carbon bond formation of reduced C<sub>1</sub> substrates rather than carbon-carbon bond cleavage of sugars to produce a variety of important compounds. This ability occurs in all domains of life, ranging from aerobic and anaerobic bacteria, to yeasts and even anaerobic archaea<sup>[11]</sup>. While bacterial methylotrophs depict phylogenetic diversity, in eukaryotic methylotrophs this

metabolic trait is only found in a restricted number of yeast genera, including *Candida*, *Komagataella*, *Kuraishia*, *Ogataea*, and *Pichia*<sup>[12]</sup>. A further distinction is made between obligate methylotrophs, which are limited to C<sub>1</sub> substrates, and facultative methylotrophs, which can additionally grow on multi-carbon substrates consisting of carbon-carbon bonds<sup>[13]</sup>.

For methanol assimilation, several naturally occurring pathways have been identified, which serve as the basis of methylotrophy and enable the use of methanol for biomass generation and energy production.

### 1.2.1 Overview of methanol assimilation and dissimilation pathways

So far, four native methanol assimilation pathways have been identified (Figure 1). In bacteria the ribulose monophosphate (RuMP) pathway, the Calvin-Benson-Bassham (CBB) cycle or the serine cycle are commonly found, whereas in yeast such as *Pichia pastoris* (renamed as *Komagataella phaffii*<sup>[14]</sup>) the xylulose monophosphate (XuMP) pathway is known to support methylotrophic growth.



**Figure 1: Overview of methanol assimilation to biomass or dissimilation to CO<sub>2</sub>.** Shown are three classes of methanol assimilation enzymes (MDH) that oxidize methanol to formaldehyde, a linear cofactor-dependent formaldehyde oxidation pathway on the left and the different assimilation cycles on the right. Carbon is assimilated either directly at the level of formaldehyde or later at the level of formate or CO<sub>2</sub>. XuMP cycle, xylose monophosphate cycle RuMP cycle, ribulose

monophosphate cycle; CBB cycle, Calvin-Benson-Bassham cycle. Metabolite abbreviations: CO<sub>2</sub>: molecular carbon dioxide; FA: formate; FALD: formaldehyde; MeOH: methanol; NAD(P)<sup>+</sup>: oxidized nicotinamide adenine dinucleotide (phosphate); NAD(P)H: reduced nicotinamide adenine dinucleotide (phosphate); O<sub>2</sub>: molecular oxygen; PQQ: pyrroloquinoline quinone; PQQH<sub>2</sub>: pyrroloquinoline quinol. Created with Biorender.com.

In each of these pathways, the first step is the oxidation of methanol to formaldehyde. This reaction is carried out either via NAD- or PQQ-dependent methanol dehydrogenases, methanol oxidoreductases, or O<sub>2</sub>- dependent alcohol oxidases in yeast. Formaldehyde then branches into either assimilatory (product biomass) or dissimilatory (product CO<sub>2</sub>) metabolic pathways and plays a central role for all methylotrophic organisms. Formaldehyde “is an extremely toxic compound that non-specifically interacts with proteins and nucleic acids in all biological cells<sup>[15]</sup>. Therefore, all methylotrophic organisms must mitigate formaldehyde toxicity during growth on methanol by maintaining low intracellular formaldehyde concentrations that can be quenched either by the assimilation or dissimilation pathway<sup>[16]</sup>.”<sup>[1]</sup>

### **Methanol assimilation into biomass**

“In methylotrophic yeasts, formaldehyde is assimilated using the dihydroxyacetone (DHA) pathway, also known as the XuMP pathway<sup>[17]</sup>.”<sup>[1]</sup> “In the first step (peroxisomal), formaldehyde is condensed with xylulose 5-phosphate (Xu5P) by the peroxisomal key enzyme DAS (dihydroxyacetone synthase). Two C<sub>3</sub> compounds are formed in this reaction, namely dihydroxyacetone (DHA) and glyceraldehyde 3-phosphate (G3P) to fuel gluconeogenic reactions<sup>[18]</sup>. Subsequently, DHA and G3P are released from the peroxisomes into the cytosol<sup>[19]</sup>. Recently, it was revealed that the cytosolic localization might not occur in *P. pastoris*. It was shown by omics-level investigations of the metabolism that this yeast orchestrates all assimilation steps within the peroxisome<sup>[18]</sup>. Next, the cytosolic DHA is phosphorylated to dihydroxyacetone phosphate (DHAP) by a dihydroxyacetone kinase (DAK), the second key enzyme of the XuMP pathway<sup>[20]</sup>. The latter reaction cascade connects the C<sub>1</sub>- metabolism with the common central carbon metabolism on the level of glycolysis by formation of fructose 1,6-bisphosphate (F1,6BP) from DHAP and G3P. Subsequently, fructose 6-phosphate (F6P) is formed by dephosphorylating F1,6BP, connecting the former C<sub>1</sub>-assimilation to gluconeogenesis and the pentose phosphate pathway<sup>[18]</sup>. In particular, the F6P pool is partly harnessed for Xu5P regeneration by recruiting the non-oxidative pentose phosphate pathway branch and the associated pentose interconversion reactions. It has to be highlighted that the recruiting of a transketolase yields erythrose 4-phosphate and Xu5P from F6P and G3P. Strikingly, a distinct feature of methylotrophic pentose rearrangements is the subsequent conversion of erythrose 4-

phosphate (C<sub>4</sub>) into sedoheptulose 1,7-bisphosphate (C<sub>7</sub>) using DHAP (C<sub>3</sub>) by application of an aldolase. Finally, sedoheptulose 1,7-bisphosphate is dephosphorylated to sedoheptulose 7-phosphate, which is in turn converted by a transketolase and G3P into two Xu5P units<sup>[18].”[1]</sup>

The bacterial RuMP pathway follows a similar strategy to enable microbial growth on methanol. It is commonly found in gram-positive bacteria such as *B. methanolicus* as well as in Beta- and Gamma-proteobacteria<sup>[21]</sup>. In the RuMP pathway, formaldehyde assimilation is catalyzed by the enzyme hexulose 6-phosphate synthase (Hps), which converts formaldehyde and ribulose 5-phosphate to hexulose 6-phosphate. Next, fructose 6-phosphate is generated via isomerization of hexulose 6-phosphate by 6-phospho-3-hexuloisomerase (Phi).<sup>[13]</sup> Similar to the XuMP pathway, the RuMP pathway then relies on the regeneration of a pentose, namely ribulose 5-phosphate, which is recycled to generate glyceraldehyde G3P/DHAP every three turns. Without the tightly controlled supply of Ru5P/Xu5P, the RuMP/XuMP metabolism could not be maintained<sup>[22]</sup>.

The serine cycle occurs in methylotrophic bacteria such as *M. extorquens* AM1 and assimilates carbon in the form of methylene-tetrahydrofolate (H4F)<sup>[23]</sup>. In a first step, formaldehyde is converted to formate and further reacts with H4F to methylene-H4F via H4F-dependent enzymes<sup>[24]</sup>. These reactions use the amino acid glycine as an acceptor molecule and subsequently result in the formation of serine, the key intermediate of this pathway. Glycine is regenerated through the conversion of serine to glyoxylate and acetyl-CoA in a series of enzymatic reactions. Notably, the glycine cleavage system requires CO<sub>2</sub> and uses the produced phosphoenolpyruvate as an acceptor. Acetyl-CoA depicts the output of this cycle and is further utilized for biosynthesis of cell constituents.<sup>[25]</sup> While it is commonly acknowledged that in the XuMP and RuMP pathway “formaldehyde represents the central intermediate, some studies concluded that in methylotrophic bacteria formate is the branch point between assimilatory and dissimilatory pathways<sup>[24]</sup>. Here, the significance of the direct condensation route for methylene H4F synthesis in *M. extorquens* AM1 was assessed. It was indicated, that during laboratory growth conditions, methylene H4F is originally formed from formaldehyde via the H4MPT and H4F interconversion pathway. The latter suggests that indeed formate and not formaldehyde represents the primary metabolic branch point between assimilation and dissimilation of C<sub>1</sub> units in this microbe<sup>[24]</sup>. In turn, this additionally indicates that the spontaneous condensation of formaldehyde with H4F does not occur *in vivo*, which was confirmed recently<sup>[26].”[1]</sup>

In contrast to the other assimilation pathways, methylotrophic bacteria which use the CBB cycle assimilate carbon at the level of CO<sub>2</sub><sup>[27]</sup>. The CBB cycle is well known from plants or photosynthetic autotrophs as the Calvin Cycle<sup>[27b, 28]</sup>. In methylotrophs, methanol acts as an electron donor to generate ATP and is oxidized via formaldehyde and formate to CO<sub>2</sub>. CO<sub>2</sub> enters the CBB cycle and is carboxylated with ribulose 1,5-bisphosphate (RuBP), which was previously formed by the phosphorylation of three molecules of ribulose 5-phosphate (Ru5P). The carboxylation step is catalyzed by the enzyme ribulose bisphosphate carboxylase (RuBisCO) and results in six molecules of 3-phosphoglycerate (PGA). A kinase then phosphorylates PGA to 1,3-Bisphosphoglycerate (1,3BPG) which is converted to G3P via glyceraldehyde phosphate dehydrogenase. One G3P molecule is used for biomass generation, while the remaining molecules regenerate three molecules of Ru5P in a comparable way to the RuMP/XuMP cycle.<sup>[29]</sup> It has to be noted that CO<sub>2</sub> is assimilated in the CBB cycle in a highly ATP demanding manner<sup>[30]</sup>.

### **Methanol dissimilation into CO<sub>2</sub>**

“In order to keep the intracellular formaldehyde levels low, not only assimilation into biomass but also the dissimilation towards CO<sub>2</sub> takes place. The dissimilation is closely related to redox power generation. Particularly, the dissimilation functions as a valve to cope with toxic intracellular formaldehyde concentrations while covering NAD(P)H demand. The most frequent pathway for formaldehyde detoxification is the cytosolic thiol-dependent pathway, which employs reactive thiols as the initial formaldehyde acceptor<sup>[31]</sup>. This pathway generates redox power (i.e., NADH) and is used by methylotrophic bacteria and other non-methylotrophic organisms<sup>[32]</sup>. The produced NADH is used in cellular respiration to sustain the generation of ATP in presence of formaldehyde and supports the energy demand of the cell. Specifically, formaldehyde spontaneously reacts in the peroxisomes with glutathione (GSH) and generates S-hydroxymethylglutathione (S-HMG)<sup>[33]</sup>, which is oxidized to CO<sub>2</sub> in a subsequent cytosolic GSH-dependent oxidation cascade. In detail, the S-HMG is released from the peroxisomes into the cytosol and is oxidized to S-formylglutathione (S-FG) via a NAD<sup>+</sup>- linked and GSH-dependent formaldehyde dehydrogenase (FLD), which is shown to be essential for growth of *C. boidinii* on methanol<sup>[32]</sup>. Subsequently, S-FG is hydrolyzed via S-formylglutathione hydrolase (FGH) to formate. In the dissimilatory branch, a formate dehydrogenase (FDH) oxidizes the generated formate to CO<sub>2</sub> accompanied by NADH formation. In turn, the role of FDH is not only the formaldehyde detoxification but also retaining the redoxstate and the regulation of the glutathione level in cells. However, it was demonstrated, that FDH is not essential for growth

on methanol in *C. boidinii*. Nonetheless, as the complete genome is not yet sequenced, the existence of other FDHs cannot be excluded that supported growth during the study<sup>[32]</sup>. The latter is in contrast to the fact that FDH proteins from methylotrophic yeasts are very stable enzymes and represent about 10 to 18% of the total cellular proteins<sup>[34]</sup>.<sup>”[1]</sup> While in methylotrophic yeasts and bacteria a NAD<sup>+</sup> linked FDH is found<sup>[35]</sup>, some bacteria also exploit a cytochrome-dependent variant<sup>[36]</sup>. Surprisingly, gene mutation studies with the methylotrophic bacterium *M. extorquens* AM1, which possesses a set of three formate dehydrogenases, underlines a somewhat contrary circumstance. Here, each of the FDHs was able to sustain growth on formate, while in contrast none is required for growth on methanol or methylamine. When all FDHs are eliminated, formate appears to be consumed by an alternative route, which is not yet unraveled<sup>[37]</sup>.

Taken together, “it is still not completely understood how the efficient and dynamic distribution of formaldehyde between assimilatory and dissimilatory metabolism without toxic accumulation is conducted. However, it can be stated that compartmentalization of peroxisomal methylotrophy is highly beneficial for methylotrophic yeasts. Juxtaposed, for bacteria the formaldehyde distribution is a challenge in regard of balancing metabolic fluxes into the dissimilatory and assimilatory branch, to avoid formaldehyde accumulation.”<sup>[1]</sup>

*In vivo* flux measurements indicated that *M. extorquens* AM1 differentially distributes formaldehyde between three distinct pathway modules depending on the specific cellular methanol uptake rate<sup>[38]</sup>. In detail, the commonly assumed modules are (I) the non-enzymatic condensation reaction of formaldehyde and H4F to produce methylene-H4F, which is the C<sub>1</sub> donor for assimilation via the serine cycle. It was shown recently that this reaction is not present in *E. coli* and it is in general unlikely to serve as an effective route for formaldehyde assimilation *in vivo*<sup>[26a]</sup>. The (II) enzyme-catalyzed reaction of formaldehyde and tetrahydromethanopterin (H4MPT) resulting in methylene-H4MPT formation which is oxidized through a distinct reaction set to formate, followed by CO<sub>2</sub> dissimilation. And finally, the (III) interconversion of methylene-H4F and formate via a set of H4F-dependent reactions with subsequent assimilation of formate in the serine cycle. By combining <sup>14</sup>C and deuterium label-tracing strategies it was found that a dynamic transition from low to high formaldehyde flux occurs. With increasing formaldehyde flux, an increased distribution towards the biomass generation was shown, thereby ensuring C<sub>1</sub> related growth without formaldehyde accumulation<sup>[38]</sup>.



## 1.2.2 Application examples of natural methylotrophs

Industrial methanol fermentation dates back as early as the 1960s, when the aim was to produce microbial proteins from methanol for human and animal consumption<sup>[39]</sup>. Recent approaches have focused on engineering native methylotrophs such as *B. methanolicus*, *M. extorquens* AM1 or the yeast *P. pastoris* for the production of industrially relevant chemicals and materials. So far, the aerobic platform organism *B. methanolicus* was used to produce amino acids like L-lysine<sup>[40]</sup> and L-glutamate<sup>[41]</sup> or the food and flavor ingredient (R)-acetoin<sup>[42]</sup>. *M. extorquens* depicts another exciting methylotrophic host, which was shown to produce organic acids such as mesaconic and methylsuccinic acid<sup>[43]</sup>, dicarboxylic acid<sup>[44]</sup>, violacein<sup>[45]</sup>, and mevalonate<sup>[46]</sup> from methanol. Noteworthy, is also the production of various polyhydroxyalkanoates<sup>[47]</sup>, as these polymers are currently of interest as bio-derived and biodegradable plastics.

In terms of methylotrophic yeasts, *P. pastoris* or *H. polymorpha* have been the strains of choice. “In particular, the production of (I) human serum albumin (HSA)<sup>[48]</sup>, (II) the insulin like growth factor (IGF)<sup>[49]</sup> or (III) hepatitis B vaccines<sup>[50]</sup> was achieved. Furthermore, *P. pastoris* was engineered for the production of various protein-based polymers such as (IV) collagen<sup>[51]</sup>, (V) gelatins<sup>[52]</sup>, (VI) silk-like proteins<sup>[53]</sup> and (VII) elastin-like proteins<sup>[54]</sup>. Nevertheless, challenges like low yields, proteolytic degradation, and potential self-assembly *in vivo* may be faced when using *P. pastoris* for polymer production<sup>[55]</sup>.”<sup>[1]</sup> Advances in genomic-editing tools<sup>[56]</sup> have also led to the exploitation of *H. polymorpha* based fermentation processes. “For example, the production of various recombinant proteins such as Hepatitis E virus-like particles<sup>[57]</sup> or Ferritin (FTH1)<sup>[58]</sup> from methanol or a glycerol/methanol mixture was achieved. Moreover, several commercially available hepatitis B vaccines and other biopharmaceuticals such as hirudin, insulin and IFN $\alpha$ -2a Reiferon<sup>®</sup> are produced using *H. polymorpha*<sup>[59]</sup>.”<sup>[1]</sup>

“During the last years, there has also been an interest in methylotrophy and its application in white biotechnology as a potential silver bullet against climate change<sup>[60]</sup>. Various findings demonstrate that microorganisms play a key role in nature's carbon cycle<sup>[61]</sup> and it is therefore speculated that they can support global climate change mitigation. Selected methylotrophic microbes have the capability to utilize methane as a carbon source. Such organisms help to reduce greenhouse gas concentration in the atmosphere<sup>[62]</sup>. In addition, liquid C<sub>1</sub> substrates, sustainably produced from CO<sub>2</sub>, used for the production of bulk chemicals via fermentation can pinpoint the direction towards a cyclic bioeconomy to reduce mankind's greenhouse gas emission footprint while providing economic benefits. Already

in the early 2000s, the application of methylotrophic yeasts in the agricultural sector as biofertilizers and for the treatment of the methanol and formaldehyde containing wastewater was shown<sup>[63]</sup>.<sup>»[1]</sup>

“Furthermore, the biotechnological production of high-energy fuels by economically feasible processes has emerged as an attractive alternative to the traditional production<sup>[64]</sup>. One promising approach exploits *P. pastoris* for the production of the platform chemical and potential biofuel isopentanol.”<sup>[1]</sup> *P. pastoris* was engineered to produce 191 mg L<sup>-1</sup> of isopentanol from glucose, the highest titer recorded to date for a non-conventional yeast<sup>[65]</sup>. This represents an encouraging example that could eventually be produced from CO<sub>2</sub>-derived methanol in the future.

Recently, *P. pastoris* was also engineered towards CO<sub>2</sub> consumption via the Calvin–Benson–Bassham cycle<sup>[66]</sup>. “By introduction of eight heterologous genes, *P. pastoris* was converted into an autotroph capable to use CO<sub>2</sub> as its single carbon source. To separate the foreign fixation machinery of CO<sub>2</sub> from energy generation, the first steps of the XuMP pathway (AOX1, DAS1 and DAS2) were deleted. Following laboratory evolution, the engineered strain achieved a maximum growth rate of 0.018 h<sup>-1</sup><sup>[66]</sup>. Examples like these may form the basis for producing bulk- and fine-chemicals based on a sustainable CCU biotechnology and might support mitigation of atmospheric CO<sub>2</sub> in the future.”<sup>[1]</sup>

### 1.3 Synthetic methylotrophy

“Although notable progress has been made regarding the availability of genetic tools for native methylotrophic organisms, many of them are still not adequately characterized or their intrinsic capabilities to efficiently produce high value-added chemicals are limited. Besides that, a fundamental knowledge about the physiology, the genome and the metabolism is crucial for successful metabolic engineering of such microbes. Many of these aspects lack a robust basis when considering methylotrophic yeasts for engineering efficient microbial cell factories.

In consequence, industrial glycolytic yeasts or bacteria depict promising host organisms to exploit synthetic methylotrophy for efficient production of value-added products from C<sub>1</sub> substrates. Due to the long tradition of investigating such microbes, the knowledge base and available engineering tools are fundamentally broad and established to realize synthetic methylotrophy<sup>[67]</sup>. Furthermore, platform organisms like *S. cerevisiae* or the oleaginous yeast *Y. lipolytica* have the ability or were engineered to produce industrially relevant products such as bulk chemicals (e.g., monoalcohols, diols, organic acids, biopolymers) or

biofuels and precursors of biofuel molecules (e.g., alcohols, alkanes, carboxylic acids, fatty acids) with increased yield and titer. In regard of the various engineered producer strains, it is logical to switch food-related sugar substrates against methanol. This approach can enable a more sustainable, and even CO<sub>2</sub>-based, production of important chemical products via fermentation.

Especially with regard to major concerns about global climate change and increasingly difficult access to fossil fuels, synthetic methylotrophy has taken up the challenge to produce advanced biofuels and bioproducts. To this extent, expanding the substrate scope of the organism by the design and implementation of non-native carbon assimilation pathways is promising. Such an approach introduces the required enzymes and pathway modules into established industrial hosts. Subsequently, understanding and fine-tuning of redox balances, energy metabolism, carbon-fluxes as well as the transcriptional and translational regulation is mandatory for successful engineering projects to achieve beneficial efficiencies<sup>[68]</sup>.<sup>»[1]</sup>

### 1.3.1 Pathway selection to enable synthetic methylotrophy

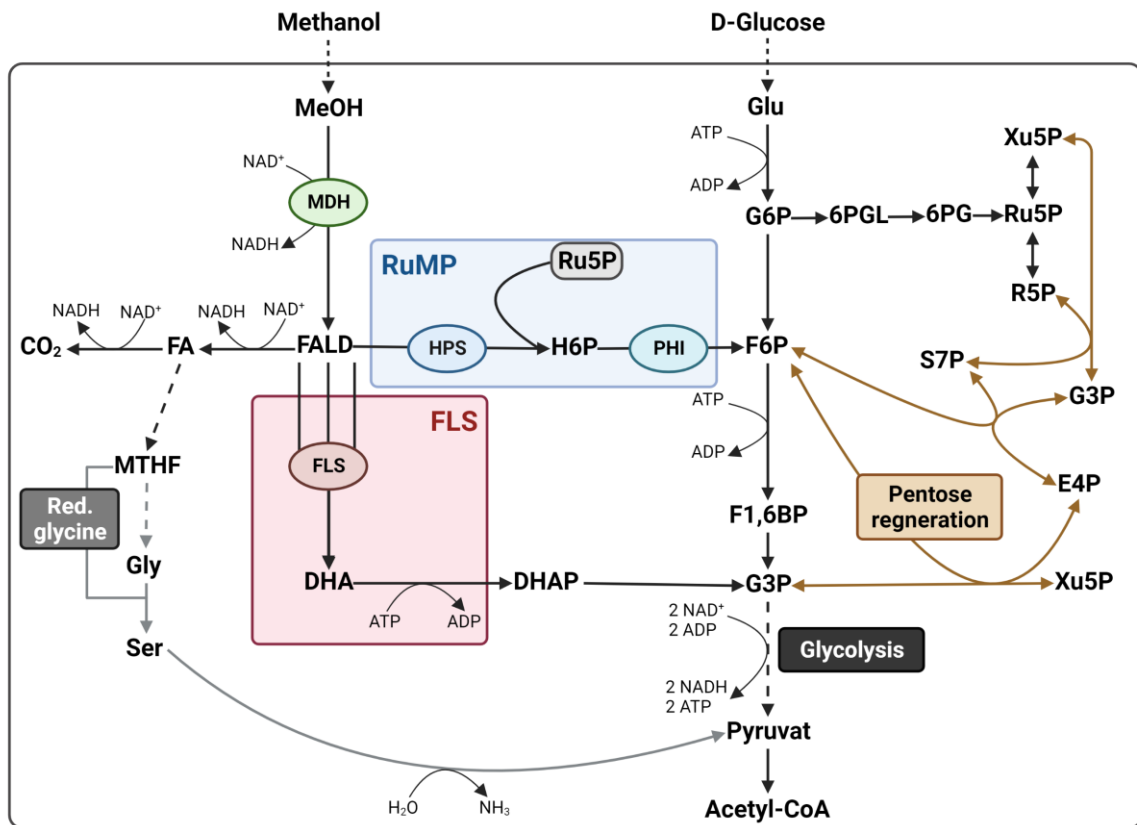
“When designing synthetic methylotrophic hosts, not only pathway kinetics and accumulation of toxic intermediates but also the stoichiometry of carbon and energy conservation have to be considered. While the serine cycle achieves the highest yield of the metabolic precursor pyruvate, it also has the highest metabolic costs in terms of ATP usage (Table 1).

**Table 1: Overview of formaldehyde assimilation pathways and their characteristics.** Adapted from<sup>[1]</sup>.

Pathway	Characteristics	Pyruvate and ATP yield	Ref.
RuMP	Cyclic assimilation pathway found in bacteria, formaldehyde enters the RuMP cycle through condensation with Ru5P	0.33 mol <sub>pyruvate</sub> mol <sub>methanol</sub> <sup>-1</sup> 0.33 mol <sub>ATP</sub> mol <sub>methanol</sub> <sup>-1</sup>	[69]
XuMP	Cyclic assimilation pathway found in yeasts, compartmentalization in the peroxisomes, formaldehyde enters the pathway through condensation with Xu5P	0.33 mol <sub>pyruvate</sub> mol <sub>methanol</sub> <sup>-1</sup> 0.66 mol <sub>ATP</sub> mol <sub>methanol</sub> <sup>-1</sup>	[70]
Serine cycle	Cyclic assimilation pathway found in bacteria, formaldehyde enters the pathway through methylene-H4F	0.5 mol <sub>pyruvate</sub> mol <sub>methanol</sub> <sup>-1</sup> -1 mol <sub>ATP</sub> mol <sub>methanol</sub> <sup>-1</sup>	[71]

The RuMP cycle and XuMP pathway yield slightly less pyruvate but, in contrast, form ATP, thus providing energy supply<sup>[72]</sup>. Regarding ATP generation, the XuMP pathway is the most promising option. Nevertheless, the corresponding AOX requires the presence of oxygen and is located in the peroxisomes, which might be a drawback depending on the used host and desired production route.”<sup>[1]</sup> This energy-intensive oxidation of methanol could also be one explanation for the higher calculated energetic efficiency for the RuMP pathway (40 – 50 %) compared to the XuMP cycle (30 – 35 %). In addition, it was shown that the RuMP cycle enables significantly higher growth rates compared to the XuMP or serine cycle.<sup>[73]</sup>

In addition to the aforementioned natural methanol assimilation pathways, there are also synthetic alternatives, such as the artificial FLS pathway or the reductive glycine pathway, among others (Figure 2).



**Figure 2: Metabolic access points of synthetic FLS, reductive glycine and native RuMP pathway into glycolytic yeast metabolism.** Highlighted is the metabolic overlap of the pathways used in this work (FLS, red and RuMP, blue) with central carbon metabolism. In addition, the reductive glycine pathway (grey) is shown. The first step of all methanol assimilation pathways is catalyzed by a methanol dehydrogenase (MDH). In the RuMP cycle formaldehyde is assimilated via pentose regeneration (brown) using the enzymes hexulose-6-phosphate synthase (HPS) and hexulose-6-phosphate isomerase (PHI). In the FLS pathway, the formolase enzyme (FLS) assimilates formaldehyde in the form of dihydroxyacetone. Metabolite abbreviations are: MeOH: methanol; Glu: D-glucose, G6P: glucose 6-phosphate; 6PGL: 6-phosphogluconolactone; Ru5P: ribulose 5-phosphate; H6P: hexulose 6-phosphate; R5P: ribose 5-phosphate; Xu5P: xylulose 5-phosphate; S7P:

sedoheptulose 7-phosphate; E4P: erythrose 4-phosphate; F6P: fructose 6-phosphate; F1,6BP: fructose 1,6-bisphosphate; G3P: glyceraldehyde 3-phosphate; FALD: formaldehyde; FA: formate; DHA: dihydroxyacetone; DHAP: dihydroxyacetonephosphate; CO<sub>2</sub>: molecular carbon dioxide; MTHF: 5,10-methylenetetrahydrofolate; Gly. glycine, Ser: serine; NAD<sup>+</sup>: oxidized nicotinamide adenine dinucleotide; NADH: reduced nicotinamide adenine dinucleotide; ADP: adenosine diphosphate; ATP: adenosine triphosphate. Partially adapted from<sup>[1]</sup>. Created with Biorender.com.

The reductive glycine pathway represents a linear route that overlaps only marginally with central metabolism, minimizing the effort required for regulatory optimization. It consists of four modules, integrated from natural occurring enzymes and forms one pyruvate molecule from two formaldehyde molecules and one CO<sub>2</sub> molecule with low energy input.<sup>[74]</sup> Only two non-native enzymes are required to establish the six steps of the reductive glycine pathway in *E. coli*. However, it faces drawbacks in terms of kinetic limitations and depending on the host organism, glycine toxicity might be a problem<sup>[75]</sup>. In addition, the toxicity of formate leading to inhibition of respiratory proteins<sup>[76]</sup> and loss of proton motive force<sup>[77]</sup> could be disadvantageous. Nonetheless, this pathway was successfully established in *E. coli* and first studies demonstrated the production of lactic acid from formate<sup>[74]</sup>.

Another elegant option to address the challenges associated with native methylotrophic pathways is the linear FLS pathway. The latter was demonstrated by Siegel et al. via the computational design of an enzyme called formolase (FLS)<sup>[78]</sup>, which can be used to realize synthetic methylotrophy. This artificial pathway is independent of oxygen and uses carbon more efficiently and with less backward flux due to the irreversibility of the FLS, therefore providing a strong driving force for methanol utilization<sup>[78-79]</sup>. Additionally, it requires no formaldehyde acceptor molecule and only consists of two enzymatic reactions: (I) Methanol oxidation to formaldehyde via an MDH and (II) Carbonylation of three formaldehyde molecules to form one dihydroxyacetone molecule via the FLS enzyme<sup>[78]</sup>. The latter reaction is generally referred to as the formose reaction, in which formaldehyde molecules combine in the presence of a base to form sugars<sup>[80]</sup>. In the case of FLS, this reaction requires two cycles of carbon-carbon coupling steps to form first glycolaldehyde (C<sub>2</sub>, GALD) and subsequently DHA (C<sub>3</sub>)<sup>[78]</sup>. To date, the low catalytic efficiency of the formose reaction (k<sub>cat</sub>/K<sub>M</sub> of 4.7 M<sup>-1</sup> s<sup>-1</sup>)<sup>[78]</sup> poses a challenge for efficient methanol utilization *in vivo*<sup>[81]</sup>. Another disadvantage of the FLS enzyme is that formaldehyde is mainly converted into the by-product GALD at low substrate concentrations. GALD is then oxidized to glyoxylate and can be further incorporated into biomass using the natural glycerate pathway. Thus, the FLS pathway via GALD leads to a lower carbon yield compared to the DHA-based counterpart and is disadvantageous for *in vivo* application. Selectivity for DHA formation even at low FALD concentrations is therefore a priority target to engineer this enzyme.<sup>[82]</sup> Introduction

of synthetic C<sub>1</sub> metabolism in *E. coli* using the FLS pathway was only successful after several rounds of adaptive laboratory evolution (ALE). Subsequent <sup>13</sup>C-labelling experiments demonstrated the incorporation of methanol into the biomass. Although the evolved strain consumed more methanol (2 g L<sup>-1</sup>) than its parent strain (1.45 g L<sup>-1</sup>), cell growth was only marginally improved.<sup>[79]</sup> The latter suggests that although more methanol was assimilated, the conversion of FALD to DHA may still be relatively low, leading to the accumulation of toxic FALD and thus hindering growth. Therefore, rapid and efficient conversion of formaldehyde to less toxic downstream metabolic intermediates is imperative, and consequently, the engineering of FLS is key to establish the methylotrophic phenotype via this non-native assimilation pathway.

### 1.3.2 Application examples of synthetic methylotrophs

Recent synthetic methylotrophy approaches focused on bacteria or the yeast *S. cerevisiae* depicting conventional host organisms that have already been used for the production of a variety of industrially relevant chemicals. And although “incorporation of labeled carbon from <sup>13</sup>C-methanol into biomass building blocks was proven in *E. coli* as well as *C. glutamicum*, growth on methanol as the sole carbon source still required yeast extract or additional sugars as energy- or carbon source<sup>[71, 83]</sup>. To overcome this limitation, a recombinant autotrophic *E. coli* strain was presented, which harnesses formate as its sole energy source for generating redox power to build up biomass completely from CO<sub>2</sub> by heterologous implementation of the Calvin cycle<sup>[84]</sup>. In detail, overexpression of FDH, Rubisco and phosphoribulokinase (PRK) enabled autotrophic growth. Initial growth experiments failed while adaptive laboratory evolution (ALE) was the key to convert the engineered strain into a fully autotrophic organism. Isotopic labeling of biomass constituents using <sup>13</sup>CO<sub>2</sub> or <sup>13</sup>C-formate, either solely or in combination proved the autotrophic growth mode. However, from an academic viewpoint these results are interesting towards true synthetic autotrophy, but far away from a real application. Another advanced example was achieved through the *de novo* design of a synthetic pathway in *E. coli* to produce acetyl-CoA from formaldehyde. It was proven that this synthetic Acetyl-CoA pathway (SACA) is the shortest, ATP-independent, carbon-conserving and oxygen-insensitive pathway for acetyl-CoA biosynthesis from a C<sub>1</sub> feedstock. The latter opens enormous possibilities for producing acetyl-CoA-derived chemicals from renewable one-carbon resources<sup>[85]</sup>. Furthermore, *E. coli* was engineered towards growth on one-carbon compounds using the reductive glycine pathway. Integration of the synthetic pathway coupled to laboratory evolution enabled growth on formate and CO<sub>2</sub> with a doubling time of ~ 8 h and growth yield of

~ 50 mg cell dry weight (CDW) g formate<sup>-1</sup>. Furthermore, growth on methanol and CO<sub>2</sub> was achieved by expressing a methanol dehydrogenase, resulting in a further increased doubling time (54 ± 5.5 h), due to the slow methanol oxidation rate<sup>[74]</sup>. This study is the first example that demonstrates true synthetic methylotrophy in a non-methylotrophic host strain.”<sup>[1]</sup>

In addition to bacteria, yeasts have enormous potential as hosts for synthetic methylotrophy and offer distinct advantages over organisms such as *E. coli* for use in industrial processes<sup>[3]</sup>. “It was shown that protein expression is superior in terms of gene expression, protein folding, and post-translational modifications of numerous eukaryotic proteins<sup>[86]</sup>. One of the most striking characteristics of yeasts is the enhanced tolerance towards acidic pH conditions<sup>[87]</sup>. Furthermore, eukaryotes are not affected by phage contamination<sup>[67]</sup>. Moreover, they possess organelles that can be harnessed for organelle directed gene expression to harness beneficial cellular functions, for example to separate formaldehyde detoxification from the cytosol<sup>[88]</sup>.”<sup>[1]</sup>

Several groups focused on installing synthetic methylotrophic modules in *S. cerevisiae*<sup>[89]</sup>. Espinosa et al. tested three different metabolic pathways to achieve growth on methanol in baker’s yeast<sup>[90]</sup>. In a first step, “the native methylotrophic yeast XuMP pathway was implemented and the expression of the associated enzymes was targeted to the peroxisomes. The latter strategy resulted in a subtle growth increase on agar plates containing YNB and 1% methanol compared to the empty vector control. In subsequent steps, engineering of a ‘hybrid’ XuMP pathway including a NAD<sup>+</sup> dependent MDH, or a bacterial RuMP pathway was conducted. Subsequently, methanol toxicity assays and <sup>13</sup>C-methanol labeling demonstrated basic functionality of the bacterial RuMP pathway. In addition, this variant seemed to be the most promising synthetic pathway, indicated by the growth profile and the increased <sup>13</sup>C-CO<sub>2</sub> production levels. Surprisingly, at higher substrate concentrations striking methanol assimilation was observed in the wild type strain. This C<sub>1</sub> assimilation was proven by <sup>13</sup>C-ethanol production from <sup>13</sup>C-methanol. The latter suggests that *S. cerevisiae* possesses native capacities for methanol oxidation and assimilation. Such findings offer new opportunities to advance microbial strain development of both, native and synthetic, one-carbon assimilation pathways in this organism.”<sup>[90]</sup>

Following the modular approach, another study demonstrated implementation of synthetic methylotrophy in *S. cerevisiae* and tested *in vivo* methanol assimilation. The strain engineering relied on genomic integration of AOX, catalase (CTA), dihydroxyacetone synthase (DAS) and dihydroxyacetone kinase (DAK) derived from *P. pastoris*. In subsequent growth experiments, the engineered strain consumed 1.04 g L<sup>-1</sup> methanol

applying shake-flask conditions with synthetic medium. The yeast produced 0.26 g L<sup>-1</sup> pyruvate and exhibited a 3.13 % improvement of biomass formation in methanol minimal medium compared to the wild type strain. Consistent with previous findings, the supplementation of yeast extract improved methanol consumption even further to 2.35 g L<sup>-1</sup> and cell growth by 11.7 %, respectively<sup>[89b]</sup>. This growth-enhancing effect of yeast extract supplementation in synthetic methylotrophy is commonly found indicating that complex media components can support synthetic methanol metabolism. Especially, biosynthesis of amino acids or vitamins and cofactors can play a key role for the observed growth dependencies<sup>[4, 18]</sup>.<sup>»[1]</sup>

“Lately, it was verified that *S. cerevisiae* has a native capacity for methylotrophy. Native methanol assimilation was confirmed through <sup>13</sup>C-tracer analysis and further improved by applying ALE. It was demonstrated that global rearrangements in central carbon metabolism and a truncation of the transcriptional regulator Ygr067cp improved growth on methanol. Nevertheless, also in this study the requirement for yeast extract in liquid methanol medium still remains a challenge.<sup>[91]</sup>”<sup>»[1]</sup>

Another interesting strain choice for synthetic methylotrophy depicts the non-conventional yeast *Y. lipolytica*. It offers “potential advantages over *S. cerevisiae* in terms of general substrate scope, metabolic pathway requirements, and physiological stress responses. It has a higher solvent tolerance in general and was shown to easily tolerate 4 % methanol as a co-substrate<sup>[68]</sup>.”<sup>»[1]</sup>

## 1.4 The oleaginous yeast *Y. lipolytica*

### 1.4.1 *Y. lipolytica* as an industrial platform organism

The oleaginous yeast *Y. lipolytica* is an obligate aerobic, ascomycetous, and non-pathogenic yeast (Biosafety Level 1) and one of the most extensively studied non-conventional yeasts<sup>[92]</sup>. It exhibits dimorphic growth and can form single cells, pseudohyphae or true hyphae<sup>[92]</sup>. The formation of these different cell forms highly depends on the growth conditions (e.g., carbon source, temperature, pH, etc.) as well as on the genetic characteristics of the respective strain<sup>[92-93]</sup>. Processes based on fermentation by the yeast *Y. lipolytica* have GRAS status ("generally recognized as safe")<sup>[94]</sup>. In addition, the biomass has already been approved by the European Food Safety Authority as a novel food and food supplement since 2019<sup>[95]</sup>. This is based on the fact that this non-conventional yeast is widespread in nature and also occurs on cheese or meat, for example<sup>[96]</sup>. Besides sugars,



*Y. lipolytica* can also use other carbon sources, such as starch, glycerol or oils and fats, for its growth<sup>[97]</sup>. In addition to tolerating extreme environments (hypersaline, heavy metal contaminated and others)<sup>[98]</sup>, this yeast species also specialized in using unusual carbon sources such as hydrocarbons<sup>[99]</sup>. *Y. lipolytica* belongs to the group of oleaginous yeasts and is characterized by a high lipid content, which accounts for more than 20 % of its biomass<sup>[100]</sup>. The majority of lipids are stored as triacylglycerides (TAGs). This physiological property makes *Y. lipolytica* particularly interesting for the production of lipid derivatives. For example, large quantities of eicosapentaenoic acid (EPA) can be produced by genetic engineering and process optimization<sup>[101]</sup>. While initial applications were still limited, the recent developments in synthetic and molecular biology (1.4.2) have enabled the production of a plethora of industrially relevant chemicals (Table 2). Other notable product examples include lipases and other hydrolytic enzymes,  $\gamma$ -decalactone and erythritol, among others<sup>[102]</sup>.

**Table 2: Production of important chemicals from conventional feedstock by *Y. lipolytica*.** Partially adapted from<sup>[1]</sup>.

Product	Procedure	Substrate	Titer	Ref.
Citric acid	Elongation of the production phase of the bioprocess with growth-decoupled citric acid production.	1.5 % glucose	~100 g L <sup>-1</sup>	[103]
Omega-3 eicosapentaenoic acid	Overexpression of the $\Delta 9/\Delta 8$ pathway (41 copies of 19 different genes) and optimization of lipid metabolism.	2 % glucose	25 % of yeast biomass	[104]
Fatty acid ethyl esters	Expression of pyruvate decarboxylase ( <i>pdh</i> ) and alcohol dehydrogenase II ( <i>adhB</i> ) from <i>Zymomonas mobilis</i> and introduction of heterologous wax ester synthases <i>ws2</i> and <i>maqu_0168</i> from disruption of competitive pathways to increase fatty acyl-CoA pool.	2 % glucose	8.2 mg L <sup>-1</sup>	[105]
Isoprene	Genome integration of <i>IspS</i> gene of <i>Pueraria montana</i> and overexpression of HMG, ERG13 and IDI.	2 % glucose	~500 $\mu$ g L <sup>-1</sup>	[106]
Limonene	Combinatorial gene (over)expression (CILS or MsLS with HMG, IDI, and tSINDPS1) and optimization of the fermentation medium.	1.5 % waste cooking oil	91.24 mg L <sup>-1</sup> (D-limonene) 83.06 mg L <sup>-1</sup> (L-limonene)	[107]

1-Decanol	Overexpression of FAR from <i>Arabidopsis thaliana</i> and FAT from <i>C. palustris</i> . Deletion of the major peroxisome assembly factor Pex10.	5 % glucose	>500 mg L <sup>-1</sup>	[108]
FAEE	Expression of WS gene from <i>Marinobacter sp.</i> and deletion of <i>PEX10</i> gene.	2 - 6 % glucose, 2 - 10 % ethanol	1.18 g L <sup>-1</sup>	[109]
FFAs	Overexpression of hybrid hFAS-EcTesA.	10 % glucose	9.67 g L <sup>-1</sup>	[110]
$\beta$ -carotene	Overexpression of $\beta$ -carotene pathway and promoter screening.	6 g h <sup>-1</sup> glucose	6.5 g L <sup>-1</sup>	[111]
Docosa-hexaenoic acid	Expression of artificial <i>pfa</i> BGC from <i>Aetherobacter fasciculatus</i> .	2.5 % glucose	350 mg L <sup>-1</sup>	[112]
Cyclopropane fatty acids	Overexpression of the <i>E. coli</i> cyclopropane fatty acid synthase gene under a hybrid promoter (hp8d) and <i>Y. lipolytica</i> LR01 gene.	7 % glucose	3.06 g L <sup>-1</sup>	[113]
Squalene	Deletion of <i>PEX10</i> and <i>URE2</i> and overexpression of <i>ERG</i> , <i>HMG</i> , and <i>DGA1</i> .	2 % glucose	502.7 mg L <sup>-1</sup>	[114]

### 1.4.2 The advancing synthetic toolkit for engineering *Y. lipolytica*

The successful transformation of a non-conventional yeast into a synthetic methylotrophic yeast depends upon industrial potential (as described in 1.4.1) and the availability of simple and efficient genetic tools. Tightly controlled gene expression is crucial in metabolic engineering, as expression levels have to be balanced to optimize metabolic flux and reduce the cost of protein synthesis. One way to achieve this is by selecting promoters according to their strength. Noteworthy, higher protein expression is not always related to stronger promoters<sup>[115]</sup>. Commonly used constitutive promoters in *Y. lipolytica* are the native pTEF<sup>[116]</sup>, which codes for translation elongation factor-1, as well as other native promoters such as pFBA1N<sup>[117]</sup>, pFBA<sup>[118]</sup>, or pGPD<sup>[119]</sup>. To expand the strength of the already existing promoters, hybrid promoters were created through the addition of tandem copies of an upstream activation sequence (UAS)<sup>[120]</sup>. In this manner, several promoters were established that provide strong, constitutive expression of a gene of interest<sup>[118, 121]</sup>.

For fermentation processes in which the growth phase has to be uncoupled from the production phase, inducible promoter systems have been developed. They add an element of temporal control by activating gene expression in response to the presence of specific inducer or repressor molecule. The first of such promoters is encoded by the native peroxisomal acyl-CoA oxidase 2 (POX2) gene<sup>[122]</sup>. The pPOX2 is inducible by fatty acids, such as oleic acid and alkanes and repressed by glucose and glycerol. While promising in the beginning, the first version of this promoter suffered from leaky expression in glucose containing media. Upon further characterization of the inducible POX2 element, repression in glucose was secured and a 48-fold induction by oleic acid achieved.<sup>[123]</sup> It was found that this inducible element is relatively versatile and can also be induced by other lipids<sup>[123]</sup>. Since then, various inducible promoters were generated including the pEYK1, which is induced by erythritol and erythrulose<sup>[124]</sup>.

Another way to adjust gene expression are genetic switches like transcription terminators (TT). In fungal systems, terminators can improve enzyme yield by altering mRNA stability and half-life, among others<sup>[125]</sup>. However, despite their importance, less effort was laid on studying terminators, which seems to be a common issue with yeast expression systems in general<sup>[126]</sup>. Usually, terminators such as the ttCyc are adapted from *S. cerevisiae* and have been shown to enhance gene expression by up to 70 % compared to common, endogenous terminators<sup>[127]</sup>. Additionally, the generation of synthetic *de novo* terminators promises to be an exciting area of research in *Y. lipolytica*.

While the mentioned features have led to an undeniable improvement in strain engineering, conventional genome engineering in *Y. lipolytica* was still a cumbersome effort. *Y. lipolytica* favors non-homologous end joining (NHEJ) over homologous repair, leading to undesirable, random integration of exogenous DNA. By deleting ku70, a DNA-binding protein responsible for double-strand break repair, the NHEJ pathway is disrupted and gene insertion can be controlled. Nevertheless, long-term stability disfavors this genotype. An important breakthrough was therefore the recently discovered CRISPR technology and its adaption for non-conventional yeasts. Schwartz et al. developed the first CRISPR/Cas9 system by expression of the active Cas9 in *Y. lipolytica*<sup>[128]</sup>. Here, the synthetic RNA polymerase III (Pol III) promoters as well as an RNA polymerase II TEF promoter using hammerhead and HDV ribozymes were used to drive sgRNA expression<sup>[128]</sup>. Furthermore, CRISPRi<sup>[129]</sup> and CRISPRa<sup>[130]</sup> systems have been engineered to express dCas9 in *Y. lipolytica* using a strong constitutive promoter as well as a SV40 nuclear localization signal. In principle, expression of Cas9/dCas9 facilitated either on plasmids or via chromosomal integration. When using

plasmids, Cas9/dCas9 is expressed by using an autonomously replicating plasmid (ARP) for the recycling of marker genes (e.g., auxotrophic markers), or via a non-ARP for transient gene expression<sup>[131]</sup>. In contrast, the integration of Cas9 into the genome allows simpler transformation and only requires a sgRNA expression cassette, thus making this system highly efficient. In addition, CRISPR-Cas9 technologies have been included in several cloning toolkits, such as YaliBricks<sup>[132]</sup> or EasyCloneYALI<sup>[133]</sup>.

The first method for assembling multiple fragments in this non-conventional yeast was presented by Gao and colleagues in 2014. The one-step integration of four genes encoding a  $\beta$ -carotene biosynthesis pathway (~11 kb, including promoter/terminator for each gene) was completed in less than one week and achieved an efficiency of up to 21 %.<sup>[134]</sup> With the aforementioned YaliBricks, a multi-part assembly kit was introduced using BioBrick standards. Using this method, the violacein pathway, consisting of five genes (~12 kb), was assembled. As a result, *Y. lipolytica* was shown to produce the bis-indole pigment, but no efficiency was reported.<sup>[132]</sup> EasyCloneYALI offers another modular cloning toolbox, consisting of standardized promoters, and plasmids, providing different selection markers on integrative vectors as well as a marker-free CRISPR-based option. Up to five vectors were integrated into target sites with an efficacy of over 80 %.<sup>[133]</sup>

An alternative are Golden Gate cloning methods, which enable rapid assemblies of modular DNA fragments into episomal vectors<sup>[135]</sup>. In 2017, this strategy has been adapted for *Y. lipolytica* and was demonstrated by assembling a synthetic carotenoid production pathway in combination with a positive selection marker. Transformation efficiencies ranged from 67 % – 90 %. The only disadvantage of this method is the number of DNA fragments that have to be processed in one reaction. For three transcription units, 12 DNA fragments have to be incorporated into the recipient vector, which leads to a higher proportion of incorrectly assembled plasmid constructs.<sup>[136]</sup> So far, researchers have mainly focused on standard ‘laboratory’ strains for which auxotrophic markers are already available. Nevertheless, wild type strains of *Y. lipolytica* bear great potential in terms of lipid, citric acid or polyol production and were shown to outperform the typical laboratory strains<sup>[137]</sup>. However, the already established tools are often dependent on genetic traits such as auxotrophic markers, making wild type engineering difficult. For this purpose, Egermeier and colleagues developed an engineering strategy independent of auxotrophic markers and therefore well suited for wild type isolates of *Y. lipolytica*<sup>[138]</sup>. The toolbox, based on the GoldenMOCS system<sup>[139]</sup>, includes vectors for the extrachromosomal expression of up to four transcription units. In combination with the provided CRISPR/Cas-9 vector, genome integration is also

possible. In addition, a set of two markers, two promoters, and two terminators provides versatility. This strategy was successfully used to overexpress glycerol kinase (GUT1) and to delete the LEU2 gene in a wild type strain, resulting in improved citric acid and erythritol production from glycerol.<sup>[138]</sup>

The mentioned advances in genetic tools greatly expand the possibilities for genome modification and metabolic engineering in this versatile host. Combined with its vast industrial potential, *Y. lipolytica* is an exciting host for implementing synthetic methylotrophy<sup>[3]</sup>.

### **1.4.3 Current state of synthetic C<sub>1</sub> approaches in *Y. lipolytica***

“Similar to *S. cerevisiae* also for the ascomycetous yeast *Y. lipolytica* a native capacity for methylotrophy in form of a non-specific alcohol dehydrogenase was proposed. A recent approach suggests using crude glycerol, which is contaminated with methanol, as a feedstock for engineered *Y. lipolytica*. In order to develop microbes, which use methanol as a co-substrate, the formaldehyde dehydrogenase (FLD) gene was identified and deleted. This prevents methanol dissimilation to CO<sub>2</sub> via formaldehyde and formate. The generated deletion strain oxidized methanol to formaldehyde without the expression of a heterologous methanol dehydrogenase. To complement the  $\Delta fld1$  strain, either HPS or DHAS were expressed and these designs enabled restoring the formaldehyde tolerance upon FLD deletion.<sup>[140]</sup>

Another approach combined metabolic engineering with ALE. By rationally constructing a chimeric assimilation pathway in *Y. lipolytica*, engineering enhanced precursor supply, and ALE, improved methanol assimilation up to 1.1 g L<sup>-1</sup> after 72 h was achieved. Here, a chimeric pathway, consisting of BsMDH, BmHPS, and BmPHI (RuMP pathway) and PpDAS1 and PpDAK2 (XuMP pathway), facilitated the most efficient methanol assimilation in *Y. lipolytica*. Furthermore, fine-tuning of methanol assimilation and enhancing formate dehydrogenation and serine pathways were exploited. In addition, upregulation of ribulose monophosphate/xylulose monophosphate (RuMP/XuMP) regeneration genes and subjecting the resulting strain to ALE were key towards improved methanol assimilation.<sup>[68]»[1]</sup>

Most recently, *Y. lipolytica* was engineered to convert methanol into succinic acid. In this strategy, rational metabolic engineering was used in combination with compartmentalization. In a first step, the XuMP pathway genes from *P. pastoris* were expressed in the peroxisomes of *Y. lipolytica*. Next, a Xu5p recycle pathway was constructed

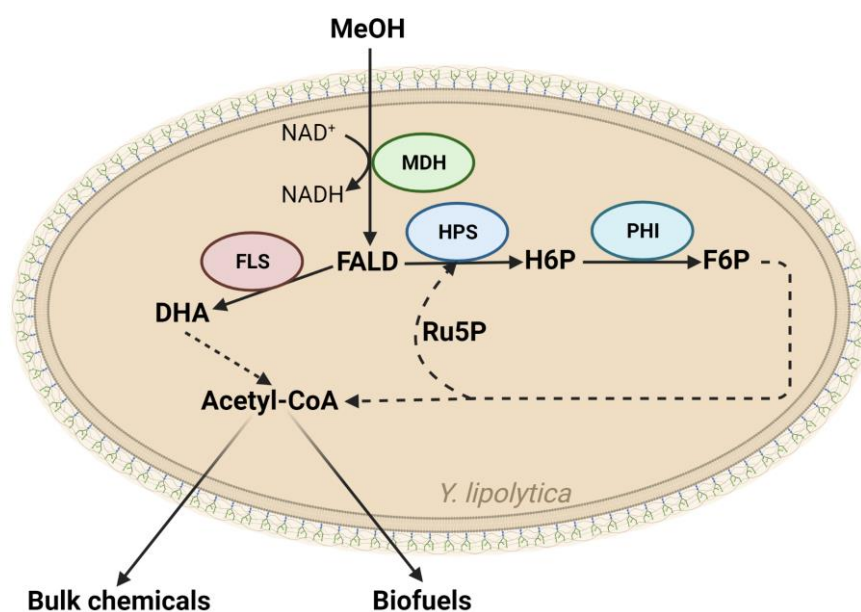
through the introduction of three genes from *Meyerozyma guilliermondii* and overexpression of enzymes from *P. pastoris* to regenerate the precursor Xu5P, resulting in enhanced cell growth and methanol utilization. After inactivation of the succinate dehydrogenase subunit 5 (sdh5), production of 0.92 g L<sup>-1</sup> succinic acid was achieved using methanol-containing media supplemented with 2 g L<sup>-1</sup> yeast extract.<sup>[141]</sup> Table 23 summarizes the current advances regarding synthetic methylotrophy in *Y. lipolytica* and compares it to this work.

## 1.5 Objective

The metabolic ability to efficiently utilize reduced C<sub>1</sub> compounds and reach high biomass yields solely on methanol, has so far been limited to natural methylotrophs. Various successful synthetic methylotrophic approaches, both in bacteria and yeast, exist but open questions and challenges remain. Aim of this thesis was to establish synthetic methylotrophy, i.e., the utilization of methanol as a carbon source, in *Y. lipolytica* and thus to produce industrially relevant products from methanol in the future. Its tolerance of harsh conditions, the lack of strict metabolic regulation and its ability to produce biofuels and bioproducts, make *Y. lipolytica* an attractive fungal host for synthetic methylotrophy. The first part of this thesis involved the selection of a suitable host strain based on rational criteria, such as methanol and formaldehyde tolerance.

In order to enable growth on methanol, pathways based on the aforementioned metrics (1.3.1) should be selected for expression in *Y. lipolytica*. Besides the reconstruction of native methylotrophic metabolic pathways such as the bacterial RuMP pathway, the establishment of new, artificial enzyme cascades provides an attractive alternative. One option uses the synthetic enzyme formolase, which converts formaldehyde directly into dihydroxyacetone. For this methanol assimilation pathway, only two enzymes need to be implemented in *Y. lipolytica*. However, as the activity of the parental formolase is very low, another aim of this dissertation was to engineer the FLS enzyme to make it more suitable for *in vivo* application. Computer-assisted modelling should be used to identify relevant positions and saturation mutagenesis performed to create a mutant FLS library. To greatly increase the number of variants which can be screened simultaneously, a robotic high-throughput screening procedure should be established and used to screen a FLS mutant library for variants with improved activity in crude lysate. *E. coli* was chosen as a well characterized host which grows quickly and is easy to transform, thus making it ideal for high-throughput screening.

If several variants show an improvement in activity compared to the parental FLS, a combinatorial approach is used to screen for synergistic effects of the single mutations by mutational combination. An improved FLS variant would be characterized by, for example, higher activity, lower  $K_M$ , better thermal stability, and/or improved expression in the host strain. Finally, the implementation of synthetic methylo-trophy in *Y. lipolytica* should be tested using FLS or RuMP pathway modules in a randomized synthetic biology approach. The first step in both pathways is the methanol oxidation to formaldehyde. Therefore, three different variants were chosen and should be tested in a combinatorial manner. The clones that showed the best growth phenotype should be isolated and characterized using a high-throughput microbioreactor followed by detailed characterization in shake flasks. An improved strain is defined by the formation of more biomass and a higher methanol consumption rate compared to the wild type strain.



**Figure 3: Graphical abstract of chosen metabolic pathways for synthetic methanol assimilation in *Yarrowia lipolytica*.** Bacterial RuMP or artificial FLS pathway should be used for methanol conversion into value-added compounds. MDH: methanol dehydrogenase; HPS: hexulose-6-phosphate synthase; PHI: hexulose-6-phosphate isomerase; FLS: formolase. Metabolite abbreviations: Acetyl-CoA: Acetyl-Coenzyme A; DHA: dihydroxyacetone; FALD: formaldehyde; H6P: hexulose 6-phosphate; MeOH: methanol;  $NAD^+$ : oxidized nicotinamide adenine dinucleotide; NADH: reduced nicotinamide adenine dinucleotide; Ru5P: ribulose 5-phosphate; R5P: ribose 5-phosphate; XU5P: xylulose 5-phosphate. Created with Biorender.com.

# 2 MATERIAL AND METHODS

## 2.1 Material

### 2.1.1 Chemicals, media and buffers

Chemicals used within this work are listed in Table 3.

**Table 3: List of the chemicals used within this work and their supplier.**

Name	Supplier
Agar Kobe I	Carl Roth GmbH, Karlsruhe, Germany
Agarose basic	AppliChem GmbH, Darmstadt, Germany
CaCl <sub>2</sub>	VWR International LLC, Radnor, USA
CaCO <sub>3</sub>	Sigma Aldrich, St. Louis, USA
Coomassie Brilliant blue G250	VWR International LLC, Radnor, USA
CoSO <sub>4</sub>	Sigma Aldrich, St. Louis, USA
CuSO <sub>4</sub>	VWR International LLC, Radnor, USA
dNTPs	Rapidozym GmbH, Berlin, Germany
DTT	VWR International LLC, Radnor, USA
Ethanol	Carl Roth GmbH, Karlsruhe, Germany
Ethidiumbromide	VWR International LLC, Radnor, USA
FeSO <sub>4</sub>	Sigma Aldrich, St. Louis, USA
Glucose	VWR International LLC, Radnor, USA
Glycerol	Carl Roth GmbH, Karlsruhe, Germany
HCl (37 %)	VWR International LLC, Radnor, USA
H <sub>2</sub> SO <sub>4</sub>	VWR International LLC, Radnor, USA
Imidazole	VWR International LLC, Radnor, USA
IPTG	Carl Roth GmbH, Karlsruhe, Germany
K <sub>2</sub> HPO <sub>4</sub>	VWR International LLC, Radnor, USA
Kanamycin sulfate	AppliChem GmbH, Darmstadt, Germany



KCl	VWR International LLC, Radnor, USA
KH <sub>2</sub> PO <sub>4</sub>	Carl Roth GmbH, Karlsruhe, Germany
Methanol	VWR International LLC, Radnor, USA
MgCl <sub>2</sub>	Carl Roth GmbH, Karlsruhe, Germany
MnCl <sub>2</sub>	Merck KGaA, Darmstadt, Germany
MnSO <sub>4</sub>	VWR International LLC, Radnor, USA
Na <sub>2</sub> HPO <sub>4</sub>	AppliChem GmbH, Darmstadt, Germany
NaCl	VWR International LLC, Radnor, USA
NAD	VWR International LLC, Radnor, USA
NADH di sodium salt	Carl Roth GmbH, Karlsruhe, Germany
NaH <sub>2</sub> PO <sub>4</sub>	VWR International LLC, Radnor, USA
SDS 20 % solution	Sigma Aldrich, St. Louis, USA
Sodium formate	Sigma Aldrich, St. Louis, USA
Thiaminechloride hydrochloride	VWR International LLC, Radnor, USA
Thiaminpyrophosphat	Sigma Aldrich, St. Louis, USA
Tris	VWR International LLC, Radnor, USA
Tris base	VWR International LLC, Radnor, USA
Tween20	VWR International LLC, Radnor, USA
Yeast extract	Carl Roth GmbH, Karlsruhe, Germany
ZnCl <sub>2</sub>	VWR International LLC, Radnor, USA
ZnSO <sub>4</sub>	Alfa Aesar, Ward Hill, USA
$\beta$ -Mercaptoethanol	Sigma Aldrich, St. Louis, USA

The different enzymes utilized in this work are summarized in Table 4.

**Table 4: List of the enzymes used within this work and their supplier.**

Name	Supplier
Antarctic Phosphatase	New England Biolabs, Ipswich, USA
DNase I	AppliChem GmbH, Darmstadt, Germany
Glycerol dehydrogenase from <i>Cellulomonas</i> sp.	Sigma Aldrich, St. Louis, USA
Lysozyme	Carl Roth GmbH, Karlsruhe, Germany
Pfu polymerase	New England Biolabs, Ipswich, USA
Phusion polymerase	New England Biolabs, Ipswich, USA
Taq polymerase	Rapidozym, Berlin, Germany
Zymolyase from <i>Arthrobacter lutes</i>	Zymo Research, Irvine, USA

The various restriction enzymes used in this study are listed in Table 5.

**Table 5: List of the restriction enzymes used within this work, their restriction site and their supplier.**

Name	Restriction site (5' to 3')	Supplier
AvrII	C   CTAGG	New England Biolabs, Ipswich, USA
Bpil (BbsI)	GAAGAC(N) <sub>2</sub>   ...	Thermo Fisher Scientific GmbH
BsaI-HF®v2	GGTCTC(N) <sub>1</sub>   ...	New England Biolabs, Ipswich, USA

The different kits used in this work are given in Table 6.

**Table 6: List of kits used within this work and their supplier.**

Name	Supplier
GeneJET Plasmid MiniPrep	Thermo Fisher Scientific GmbH
GeneJET Gel Extraction KIT	Thermo Fisher Scientific GmbH
Zymoprep Yeast Plasmid MiniPrep II	Zymo Research, Irvine, USA

The cultivation media were sterile filtered (0.2  $\mu\text{m}$ ) or autoclaved at 121 °C and 2 bar for 20 min before use. For cultivation on solid medium, 1.5 – 2 % (w/v) agar-agar was added to the liquid medium.

The composition of the used media is listed in Table 7.

**Table 7: Composition of used media.**

Medium	Components
Lysogeny Broth (LB) medium pH 7.0 adj. with HCl	5 g L <sup>-1</sup> Yeast extract 10 g L <sup>-1</sup> NaCl 10 g L <sup>-1</sup> Tryptone
SOC medium pH 7.0 adj. with HCl	5 g L <sup>-1</sup> Yeast extract 0.5 g L <sup>-1</sup> NaCl 0.3 g L <sup>-1</sup> KCl 20 g L <sup>-1</sup> Tryptone Add 10 mM MgCl <sub>2</sub> and 20 mM glucose after autoclaving
Terrific Broth (TB) medium pH 7.0 adj. with HCl	24 g L <sup>-1</sup> Yeast extract 12 g L <sup>-1</sup> Tryptone 8 mL L <sup>-1</sup> Glycerol 2.2 g L <sup>-1</sup> KH <sub>2</sub> PO <sub>4</sub> 2.2 g L <sup>-1</sup> K <sub>2</sub> HPO <sub>4</sub>
Yeast extract Peptone Dextrose (YPD) medium	20 g L <sup>-1</sup> Peptone 20 g L <sup>-1</sup> Glucose 10 g L <sup>-1</sup> Yeast extract
K1 minimal medium (Adapted from Nabou et al. <sup>[142]</sup> )	Mineral salt solution (10x, sterile filtered): 70 g L <sup>-1</sup> KH <sub>2</sub> PO <sub>4</sub> 25 g L <sup>-1</sup> Na <sub>2</sub> HPO <sub>4</sub> 15 g L <sup>-1</sup> MgSO <sub>4</sub> 5 g L <sup>-1</sup> (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> Trace elements (1000x, sterile filtered, stored at 4 °C):

	1500 mg L <sup>-1</sup> CaCl <sub>2</sub> x 2 H <sub>2</sub> O 1500 mg L <sup>-1</sup> FeCl <sub>3</sub> x 6 H <sub>2</sub> O 200 mg L <sup>-1</sup> ZnSO <sub>4</sub> x 7 H <sub>2</sub> O 600 mg L <sup>-1</sup> MnSO <sub>4</sub> x H <sub>2</sub> O 0.1 mg L <sup>-1</sup> CoCl <sub>2</sub> x 6 H <sub>2</sub> O Vitamin (1000x, sterile filtered): 300 mg L <sup>-1</sup> Thiamin-HCl C-sources: 15 g L <sup>-1</sup> Glucose or 15 g L <sup>-1</sup> Glycerol or 2 % (w/v) Methanol
--	--

The formulation of buffers and antibiotics is presented in Table 8.

**Table 8: Composition of buffers and antibiotics.**

Buffer	Components
Coomassie staining solution for SDS-PAGE	50 % (v/v) Methanol 10 % (v/v) Acetic acid 0.05 % (w/v) Coomassie Brilliant blue R250
TAE buffer pH 8.0 for agarose gel electrophoresis (50x)	242.3 g L <sup>-1</sup> Tris 16.41 g L <sup>-1</sup> EDTA 57.1 mL L <sup>-1</sup> acetic acid
Lysis Buffer	16.282 g L <sup>-1</sup> K <sub>2</sub> HPO <sub>4</sub> 887.8 mg L <sup>-1</sup> KH <sub>2</sub> PO <sub>4</sub> 246.48 mg L <sup>-1</sup> MgSO <sub>4</sub> · 7 H <sub>2</sub> O 42.4 mg L <sup>-1</sup> TPP 2.5 mg mL <sup>-1</sup> Lysozyme 0.1 mg mL <sup>-1</sup> DNase I
Buffer A	1 x PBS 246.48 mg L <sup>-1</sup> MgSO <sub>4</sub> · 7 H <sub>2</sub> O

	<p>42.4 mg L<sup>-1</sup> TPP</p> <p>1 mM β-mercaptoethanol</p> <p>10 mM Imidazole</p>
Buffer B	<p>1 x PBS</p> <p>246.48 mg L<sup>-1</sup> MgSO<sub>4</sub> · 7 H<sub>2</sub>O</p> <p>42.4 mg L<sup>-1</sup> TPP</p> <p>1 mM β-mercaptoethanol</p> <p>500 mM Imidazole</p>
Desalt Buffer	<p>16.282 g L<sup>-1</sup> K<sub>2</sub>HPO<sub>4</sub></p> <p>887.8 mg L<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub></p> <p>246.48 mg L<sup>-1</sup> MgSO<sub>4</sub> · 7 H<sub>2</sub>O</p> <p>84.9 mg L<sup>-1</sup> TPP</p>
Phosphate buffer saline (PBS) (1x)	<p>137 mM NaCl</p> <p>2.7 mM KCl</p> <p>10.1 Na<sub>2</sub>HPO<sub>4</sub></p> <p>1.76 mM KH<sub>2</sub>PO<sub>4</sub></p>
SDS running buffer pH 8.3 (10x) adj. with HCl	<p>30.29 g L<sup>-1</sup> Tris base</p> <p>144.13 g L<sup>-1</sup> Glycine</p> <p>50 mL L<sup>-1</sup> SDS solution [20 % (w/v)]</p>
Lower buffer for separation gel pH 8.8 (4x) adj. with HCl	<p>1.5 M Tris base</p> <p>40 mL L<sup>-1</sup> SDS solution [20 % (w/v)]</p>
Upper buffer for stacking gel pH 6.8 (4x) adj. with HCl	<p>0.5 M Tris base</p> <p>40 mL L<sup>-1</sup> SDS solution [20 % (w/v)]</p>
SDS sample buffer pH 6.8 (2x) adj. with HCl	<p>100 mL L<sup>-1</sup> Tris (stock solution 1.5 M)</p> <p>150 mL L<sup>-1</sup> β-mercaptoethanol</p> <p>300 mL L<sup>-1</sup> Glycine</p> <p>60 mL L<sup>-1</sup> SDS solution [20 % (w/v)]</p> <p>20 mg L<sup>-1</sup> Bromophenol blue</p>
TE Buffer	<p>10 mM Tris-Cl pH 7.4</p> <p>1 mM EDTA pH 8.0</p>

Antibiotic	Components
1000x Ampicillin stock solution	100 mg mL <sup>-1</sup> Ampicillin sodium salt in H <sub>2</sub> O
1000x Hygromycin B stock solution	300 mg mL <sup>-1</sup> Hygromycin B in H <sub>2</sub> O
1000x Kanamycin stock solution	30 mg mL <sup>-1</sup> Kanamycin sulfate in H <sub>2</sub> O
1000x Nourseothricin stock solution	400 mg mL <sup>-1</sup> Nourseothricin in H <sub>2</sub> O

## 2.1.2 Oligonucleotides and plasmids

All oligonucleotides used in this work are listed in Table 9.

**Table 9: Oligonucleotides used in this work.**

Name	Sequence 5' to 3'
pET24a_fw	GAATTGTGAGCGGATAACAATTCCC
pET24a_rev	CTTTGTTAGCAGCCGGATCTC
Ku70_sgRNA1_for	GGGTCGGCGCAGGTTGACGTGGTAATGCCACTGTAGACGGGTTTTAGAGCTAG AAATAGC
Ku70_sgRNA1_rev	GCTATTTCTAGCTCTAAAACCCGTCTACAGTGGCATTACCACGTCAACCTGCGC CGACCC
Ku70_sgRNA2_for	GGGTCGGCGCAGGTTGACGTGCATGTATAGTTATAAATTGGGTTTTAGAGCTAG AAATAGC
Ku70_sgRNA2_rev	GCTATTTCTAGCTCTAAAACCCAATTATAACTATACATCGACGTCAACCTGCGC CGACCC
Ku70_sgRNA3_for	GGGTCGGCGCAGGTTGACGTAACTCTTCATAAGGCCTTGGGTTTTAGAGCTAG AAATAGC
Ku70_sgRNA3_rev	GCTATTTCTAGCTCTAAAACCCAAGGCCTTATGAAGAGTTACGTCAACCTGCGC CGACCC
Ku70_seq_for	GGAAAATTCAGATCAGATTTGAGAGCAAAGTCCAAC
Ku70_seq_rev	TCAATTCGAACTCGTGTCTTGTGGATATCGTC
pCrispryl_AmpSeq-F	CTCATGAGCGGATACATATTTGAA
pCrispryl_M13-Rev	CAGGAAACAGCTATGACC
pCrispryl_pGEX3	CCGGGAGCTGCATGTGTCAGAGG
pCrispryl_hSpCas9-R1	CGCTCGTGCTTCTTATCCTC

pCrispryl 5' seq	GCATTTATCAGGGTTATTGTCTCATGAG
pCrispryl 3' seq	CACGAGCAGCTTGCCTATG
pBR322ori-F	GGGAAACGCCTGGTATCTTT
Kan-R	ATCGCGAGCCCATTATACC
Amp-R	ATAATACCGCGCCACATAGC
M13 Forward	TGTAACACGACGGCCAGT
FLS_H29_fw	GCCTGAAAAATGGTGTCAATMNNAATGCCATGCAGGCCAAACA
FLS_H29_rev	TGTTTGGCCTGCATGGCATTNNKATTGACACCATTTTTTCAGGC
FLS_Q113_fw	CCACCTGATCAATACCGGCMNNCAGGGTGTGGTTTCGTCA
FLS_Q113_rev	TGACGAAACCAACACCCTGNNKCGCGTATTGATCAGGTGG
FLS_L482_fw	GCCAGCTGCTGAAAGTMNNGTCCAGCCCAAGACT
FLS_L482_rev	AGTCTTGGGGCTGGACCNNKACTTTCAGCAGCTGGC
FLS_L556_fw	AAACGGGTCCATGCCAATMNGATCAGTTCCTCCGGGG
FLS_L556_rev	CCGCCGAGGAACTGATCNNKATTGGCATGGACCCGTTT
FLS_I557_fw	CGGAGGAACTGATCCTGNNKGGCATGGACCCGTTTGC
FLS_I557_rev	GCAAACGGGTCCATGCCMNNCAGGATCAGTTCCTCCG
BB2_Seq_fw	CATTAATGCAGCTGGCAC
BB2_Seq_rev	GGTTATTGTCTCATGAGCGG
BB3_YL68N_AD_fw	TGCAAGCAGCAGATTACG
BB3_YL68N_AD_rev	GCACGTCAAGACTGTCAAGG
pBR322ori-F	GGGAAACGCCTGGTATCTTT
FLS_L482Q_YaLi_fw	TGCTGGAAGTGCTGGGTCCAGCCCC
FLS_L482Q_YaLi_rev	GGGGCTGGACCCAGCACTTCCAGCA
FLS_I557G_YaLi_fw	CGAAGGGTCCATGCCGACCAGAATCAGCTCCTCA
FLS_I557G_YaLi_rev	TGAGGAGCTGATTCTGGTCGGCATGGACCCCTTCG
FLS_YaLi_seq_fw	GTATCGCCCTGGGTATTGTG
FLS_YaLi_seq_rev	GACCTTAGCGCACCAGTCAC
FLS_YaLi_seq_fw_2	CTCTGTGGAGTCCTTCTCCG

Plasmids used within this work are listed in Table 10.

**Table 10: Plasmids used in this work.**

Plasmid	Description	Source
pET24a	Bacterial vector with a kanamycin resistance marker for inducible expression of T7-tagged proteins.	Novagen
pET24a_FLS_cHis	KanR, expression of FLS wild type with C-terminal his tag	Received from the chair of Chemistry of Biogenic Resources, TU Munich <sup>[143]</sup>
pET24a_FLS_L482Q_I557G_cHis	KanR, expression of FLS variant L482Q/I557G with C-terminal his tag	This work
BB1_L_23_syn_BsaI	KanR, linker to construct BB1 from PCR product with FS2 and FS3	Addgene #89915
BB1_12_YlpTEF	KanR, promoter pTEF of <i>Y. lipolytica</i> for GoldenMOCS	Addgene #117819
BB1_12_YlpGPD	KanR, promoter pGPD of <i>Y. lipolytica</i> for GoldenMOCS	Addgene #117820
BB1_34_YlCyc1TT	KanR, transcription terminator Cyc1TT of <i>Y. lipolytica</i> for GoldenMOCS	Addgene #117821
BB1_34_YlMig1TT	KanR, transcription terminator Mig1TT of <i>Y. lipolytica</i> for GoldenMOCS	Addgene #117822
BB2_L_AB_syn_BbsI	AmpR, linker to construct BB2 from BB1(FS1 - FS4) with FSA and FSB	Addgene #89917
BB2_L_BC_syn_BbsI	AmpR, linker to construct BB2 from BB1(FS1 - FS4) with FSB and FSC	Addgene #89918
BB2_L_CD_syn_BbsI	AmpR, linker to construct BB2 from BB1(FS1 - FS4) with FSC and FSD	Addgene #89919
pMEG_BB3_YL18H_AB	HygR, GoldenMOCS backbone for 1 TU in <i>Y. lipolytica</i>	Addgene #117827



pMEG_BB3_YL18H_AC	HygR, GoldenMOCS backbone for 2 TU in <i>Y. lipolytica</i>	Addgene #117828
pMEG_BB3_YL68N_AB	NrsR, GoldenMOCS backbone for 1 TU in <i>Y. lipolytica</i>	Addgene #117829
pMEG_BB3_YL68N_AC	NrsR, GoldenMOCS backbone for 2 TU in <i>Y. lipolytica</i>	Addgene #117830
pMEG_BB3_YL68N_AD	NrsR, GoldenMOCS backbone for 3 TU in <i>Y. lipolytica</i>	Addgene #117831
pMEG_BB3_YL68N_AE	NrsR, GoldenMOCS backbone for 4 TU in <i>Y. lipolytica</i>	Addgene #117832
BB1_ADH	KanR, Encodes alcohol dehydrogenase (ADH) from <i>B. stearothersophilus</i> for GoldenMOCS	This work
BB1_MDH2	KanR, Encodes methanol dehydrogenase 2 (MDH2) from <i>B. methanolicus</i> MGA3 for GoldenMOCS	This work
BB1_MDH3	KanR, Encodes methanol dehydrogenase 3 (MDH3) from <i>B. methanolicus</i> MGA3 for GoldenMOCS	This work
BB1_HPS	KanR, Encodes 3-hexulose-6-phosphate synthase (HPS) from <i>B. methanolicus</i> for GoldenMOCS	This work
BB1_PHI	KanR, Encodes 6-phosphate-3-hexuloisomerase (PHI) from <i>B. methanolicus</i> for GoldenMOCS	This work
BB1_FLS	KanR, Encodes Formolase (FLS) for GoldenMOCS	This work
BB2_L_AB_YlpTEF_ADH_YMig1TT	AmpR, Transcription unit providing the ADH gene under the control pTEF and Mig1TT of <i>Y. lipolytica</i>	This work

BB2_L_AB_YlpTEF_ADH_YlCyc1TT	AmpR, Transcription unit providing the ADH gene under the control pTEF and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpGPD_ADH_YlMig1TT	AmpR, Transcription unit providing the ADH gene under the control pGPD and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpGPD_ADH_YlCyc1TT	AmpR, Transcription unit providing the ADH gene under the control pGPD and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpTEF_MDH2_YlMig1TT	AmpR, Transcription unit providing the MDH2 gene under the control pTEF and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpTEF_MDH2_YlCyc1TT	AmpR, Transcription unit providing the MDH2 gene under the control pTEF and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpGPD_MDH2_YlMig1TT	AmpR, Transcription unit providing the MDH2 gene under the control pGPD and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpGPD_MDH2_YlCyc1TT	AmpR, Transcription unit providing the MDH2 gene under the control pGPD and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpTEF_MDH3_YlMig1TT	AmpR, Transcription unit providing the MDH3 gene under the control pTEF and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpTEF_MDH3_YlCyc1TT	AmpR, Transcription unit providing the MDH3 gene under the control pTEF and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpGPD_MDH3_YlMig1TT	AmpR, Transcription unit providing the MDH3 gene under the control pGPD and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_AB_YlpGPD_MDH3_YlCyc1TT	AmpR, Transcription unit providing the MDH3 gene under the control pGPD and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpTEF_HPS_YlMig1TT	AmpR, Transcription unit providing the HPS gene under the control pTEF and Mig1TT of <i>Y. lipolytica</i>	This work

BB2_L_BC_YlpTEF_HPS_YlCyc1TT	AmpR, Transcription unit providing the HPS gene under the control pTEF and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpGPD_HPS_YlMig1TT	AmpR, Transcription unit providing the HPS gene under the control pGPD and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpGPD_HPS_YlCyc1TT	AmpR, Transcription unit providing the HPS gene under the control pGPD and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpTEF_FLS_YlMig1TT	AmpR, Transcription unit providing the FLS wt gene under the control pTEF and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpTEF_FLS_YlCyc1TT	AmpR, Transcription unit providing the HPS gene under the control pTEF and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpGPD_FLS_YlMig1TT	AmpR, Transcription unit providing the FLS wt gene under the control pGPD and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpGPD_FLS_YlCyc1TT	AmpR, Transcription unit providing the FLS wt gene under the control pGPD and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpTEF_FLS*_YlMig1TT	AmpR, Transcription unit providing the FLS L482Q/I557G gene under the control pTEF and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpTEF_FLS*_YlCyc1TT	AmpR, Transcription unit providing the FLS L482Q/I557G gene under the control pTEF and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpGPD_FLS*_YlMig1TT	AmpR, Transcription unit providing the FLS L482Q/I557G gene under the control pGPD and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_BC_YlpGPD_FLS*_YlCyc1TT	AmpR, Transcription unit providing the FLS L482Q/I557G gene under	This work

	the control pGPD and Cyc11TT of <i>Y. lipolytica</i>	
BB2_L_CD_YlpTEF_PHI_YIMig1TT	AmpR, Transcription unit providing the PHI gene under the control pTEF and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_CD_YlpTEF_PHI_YICyc1TT	AmpR, Transcription unit providing the PHI gene under the control pTEF and Cyc11TT of <i>Y. lipolytica</i>	This work
BB2_L_CD_YlpGPD_PHI_YIMig1TT	AmpR, Transcription unit providing the PHI gene under the control pGPD and Mig1TT of <i>Y. lipolytica</i>	This work
BB2_L_CD_YlpGPD_PHI_YICyc1TT	AmpR, Transcription unit providing the PHI gene under the control pGPD and Cyc11TT of <i>Y. lipolytica</i>	This work

All synthesized genes used in this work are listed in Table 11.

**Table 11: Synthesized genes used in this work.**

<b>gBlock</b>	<b>Description</b>	<b>Source</b>
BsAdh	Alcohol dehydrogenase from <i>B. stearrowthermophilus</i> with FS2 and 3 and BsaI restriction sites, codon optimized for <i>Y. lipolytica</i>	BioCat GmbH
BmMdh2	NAD <sup>+</sup> -dependent methanol dehydrogenase 2 from <i>B. methanolicus</i> MGA3 with FS2 and 3 and BsaI restriction sites, codon optimized for <i>Y. lipolytica</i>	BioCat GmbH
BmMdh3	NAD <sup>+</sup> -dependent methanol dehydrogenase 2 from <i>B. methanolicus</i> MGA3 with FS2 and 3 and BsaI restriction sites, codon optimized for <i>Y. lipolytica</i>	BioCat GmbH
BmHps	3-Hexulose-6-phosphate synthase from <i>B. methanolicus</i> with FS2 and 3 and BsaI restriction sites, codon optimized for <i>Y. lipolytica</i>	BioCat GmbH
BmPhi	6-phosphate-3-hexuloisomerase from <i>B. methanolicus</i> with FS2 and 3 and BsaI restriction sites, codon optimized for <i>Y. lipolytica</i>	BioCat GmbH
Fls-cHis	Formolase with C-terminal his tag and FS2 and 3 and BsaI restriction sites, codon optimized for <i>Y. lipolytica</i>	BioCat GmbH

### 2.1.3 Microorganisms

The strains used within this work (Table 12) are part of the strain collection of the Fraunhofer Institute for Interfacial Engineering and Biotechnology IGB, Bio-, electro- and chemocatalysis BioCat, Straubing.

**Table 12: Different microorganisms used in this work.**

Microorganism	Strain	Genotype	Source
<i>E. coli</i>	DH10B	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL nupG /pMON14272 / pMON7124	Invitrogen
<i>E. coli</i>	BL21(DE3)	F- ompT gal dcm lon hsdSB(rB- mB-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	Novagen
<i>Y. lipolytica</i>	DSM3286	Wild type	DSMZ
<i>Y. lipolytica</i>	PO1f (ATCC MYA-2613)	MatA, leu2-270, ura3-302, xpr2-322, axp-2	ATCC

### 2.1.4 Consumables, equipment and software

Reaction tubes, micro well plates, inoculating loops, falcons and other disposable plastic items were purchased from the companies Beckman Coulter, Eppendorf, Greiner Bio-One, Roth, Sarstedt, and VWR, among others. Laboratory glass bottles were used from Schott AG. Shake flasks were purchased from VWR. Table 13 lists the most commonly used devices within this work.

**Table 13: Equipment used within this work.**

Equipment	Type and Source
ÄKTA™ system	ÄKTA™ purifier, GE Healthcare Europe GmbH, Freiburg ÄKTA™ pure, GE Healthcare Europe GmbH, Freiburg
Centrifuges	Avanti J-E Centrifuge, Beckmann Coulter GmbH, Brea, USA Micro Star 17, VWR International LLC, Radnor, USA Multifuge X3R. Thermo Fisher Scientific GmbH, Waltham, Massachusetts, USA

BioLector	BioLector® I, m2p-labs GmbH, Baesweiler BioLector® XT Microbioreactor, Beckmann Coulter GmbH, Brea, USA
Electrophoresis system	Mini-Sub Cell GT System, Bio-Rad Laboratories, Inc., Hercules, California USA
Electroporator	ECM™ 630 Exponential Decay Wave Electroporator, BTX, Holliston, Massachusetts, USA
Freezer	GGU 1500, Liebherr-International Deutschland GmbH, Biberach an der Riß 6382, GFL Gesellschaft für Labortechnik mbH, Burgwedel Revco CxF, Thermo Fisher Scientific GmbH, Waltham, Massachusetts, USA
Heating block	ThermoMixer® C, Eppendorf AG, Hamburg Thermomixer Comfort, Eppendorf AG, Hamburg
HPLC system	LC-20AD, Shimadzu Corporation, Kyoto, Japan
Incubator and shaker	Innova40, Innova42 & Innova44, Eppendorf AG, Hamburg TiMix 5 control, Edmund Bühler GmbH, Hechingen Heratherm™, Thermo Fisher Scientific GmbH, Waltham, Massachusetts, USA
Microscope	BX53M, Olympus Corporation, Shinjuku, Tokio, Japan
Magnetic stirrer	MR Hei-Tec, Heidolph Instruments, Schwabach
NanoDrop	ND-Lite, Thermo Fischer Scientific GmbH, Waltham, Massachusetts, USA
PCR thermal cycler	peqSTAR 2X, VWR International LLC, Radnor, USA
pH meter	pPhenomenal, VWR International LLC, Radnor, USA
Photometers	Infinite M200, Tecan Group Ltd., Männedorf, Schweiz FLUOstar Omega, BMG Labtech, Ortenberg UV-1800, Shimadzu Corporation, Kyoto, Japan
Refrigerator	FKEX 1800, Liebherr-International Deutschland GmbH, Biberach an der Riß
Sonifier	SONOPULS HD 2200/ MS 73, BANDELIN electronic GmbH & Co. KG, Berlin
Scales	Cubis MSA balance, Sartorius AG, Göttingen Explorer Precision EX2202/E, OHAUS Corporation, Parsippany, New Jersey, United States
SDS-PAGE system	Agarose gel chamber GH100 Series, Biostep GmbH, Jahnsdorf

Sterile workbench	Safe 2020, Thermo Fisher Scientific GmbH, Waltham, Massachusetts, USA
UV fluorescent table	Gel iX20 Imager, Intas Science Imaging Instruments GmbH, Göttingen
Vortex mixer	Analogue vortex mixer, VWR International LLC, Radnor, USA
Water bath	EH-27, JULABO GmbH, Seelbach (Schutter)
Water purification system	Milli Q ultra pure system, Merck KGaA, Darmstadt

Table 14 lists the software used within this work.

**Table 14: Software used in this work.**

Software	Source
AutoDock 4	Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009) ( <a href="https://autodock.scripps.edu/">https://autodock.scripps.edu/</a> )
Basic Local Alignment Search Tool (BLAST)	National Centre for Biotechnology Information ( <a href="http://blast.ncbi.nlm.nih.gov/">http://blast.ncbi.nlm.nih.gov/</a> )
Biolection 5	m2p-labs, Baesweiler, Germany
BioRender	BioRender, Toronto, Ontario, Canada ( <a href="https://www.biorender.com/">https://www.biorender.com/</a> )
Braunschweig Enzyme Database (BRENDA)	Technische Universität Braunschweig, Institut für Biochemie und Biotechnologie ( <a href="http://brenda-enzymes.info/">http://brenda-enzymes.info/</a> )
ChemDraw 22	PerkinElmer, Inc, Waltham, Massachusetts, USA
Clone Manager 9	Scientific & Educational Software, Cary, USA
HPLC Analysis Editor	Shimadzu Corporation, Kyoto, Japan
HPLC Postrun Analysis	Shimadzu Corporation, Kyoto, Japan
Intas	Intas Science Imaging Instruments GmbH, Göttingen, Germany
KEGG data base	Kanehisa Laboratories, Kyoto, Japan ( <a href="https://www.genome.jp/kegg/pathway.html">https://www.genome.jp/kegg/pathway.html</a> )
NEBiocalculator	New England Biolabs, Ipswich, Massachusetts, USA ( <a href="https://nebiocalculator.neb.com/">https://nebiocalculator.neb.com/</a> )
Omega Data-Analysis 210R3	BMG Labtech, Ortenberg
Origin 2022b	OriginLab Corporation, Northampton, Massachusetts, USA

Primer3 Input	Whitehead Institute for Biomedical Research, Cambridge Massachusetts, USA  ( <a href="http://primer3.ut.ee/">http://primer3.ut.ee/</a> )
PyMol 2.4.1	Schrodinger, LLC, New York, USA
Tm Calculator	New England Biolabs, Ipswich, Massachusetts, USA  ( <a href="https://tmcalculator.neb.com/">https://tmcalculator.neb.com/</a> )  Agilent, Santa Clara, California, USA  ( <a href="https://www.agilent.com/store/primerDesignProgram.jsp">https://www.agilent.com/store/primerDesignProgram.jsp</a> )
YASARA	YASARA Biosciences GmbH, Wien, Austria



## 2.2 Methods

### 2.2.1 Microbiological methods

#### 2.2.1.1 Cultivation and storage of microorganisms

*E. coli* was cultivated in liquid or on solid medium at 37 °C. Solid cultivation was performed overnight on LB medium containing 1.5 % (w/v) agar-agar and the corresponding antibiotics. For liquid cultivation, a single colony was used to inoculate 10 mL TB or LB medium containing the appropriate antibiotics. The preculture was incubated overnight at 37 °C and 150 rpm and inoculated into main cultures with a volume up to 250 mL in shaking flasks. The main cultures were further cultivated at 37 °C and 135 rpm. For the liquid cultivation of *E. coli*, baffled flasks were used.

*Y. lipolytica* strains were grown in full medium with glucose (YPD) or minimal medium (K1)<sup>[142]</sup> with the appropriate carbon sources and supplements. Liquid cultivation was performed in unbaffled flasks at 30 °C and 200 rpm. For solid cultivation, YPD or K1 medium was supplemented with 2 % (w/v) agar-agar. Selection was achieved by plating on K1 agar plates supplemented with Nourseothricin at 400 µg mL<sup>-1</sup> or Hygromycin B at 300 µg mL<sup>-1</sup>. Cells were cultivated at 30 °C for 2 – 4 days.

For long-term storing of microorganisms 750 µL of a cell suspension were combined with 750 µL of 60 % (v/v) glycerol and stored at – 80 °C.

#### 2.2.1.2 Preparation of competent cells for transformation

##### Preparation of chemically competent *E. coli*

A preculture of *E. coli* was incubated overnight at 37 °C and 150 rpm in LB-medium. Subsequently, using the grown preculture, 250 mL of SOB medium was inoculated with an OD<sub>600</sub> of 0.1 and grown to an OD<sub>600</sub> of 0.6 – 0.8 at 20 °C and 200 rpm. Next, the cells were kept on ice for 10 min and then harvested by centrifugation (10 min, 2,500 g, 4 °C). The pellet was resuspended in 80 mL of cold TB buffer and incubated on ice for additional 10 min. Afterwards the cells were harvested again and the pellet resuspended gently in 20 mL TB buffer containing 7 % DMSO. The cell suspension was ultimately incubated on ice for 10 min and subsequently divided into 500 µL aliquots, which were frozen in liquid nitrogen and stored at – 80 °C.

### **Preparation of electrocompetent *E. coli***

*E. coli* was incubated overnight in LB-low salt medium at 37 °C and 150 rpm. The next morning, 400 mL of LB-low salt medium were inoculated using the preculture with an OD<sub>600</sub> of 0.2. The main culture was grown to an OD<sub>600</sub> of 0.5 – 0.6 at 37 °C and 150 rpm. After incubation on ice for 30 min, the cultures were centrifuged for 15 min at 4,000 rpm and 4 °C. The resulting pellets were resuspended in 50 mL sterile H<sub>2</sub>O and centrifuged again for 15 min at 4,000 rpm and 4 °C. Subsequently, the pellets were reconstituted in 50 mL of 10 % (v/v) glycerol and re-centrifuged using the same parameters. After repeating the centrifugation step once more, the resulting cell pellet was resuspended in 10 % (v/v) glycerol to result in an OD<sub>600</sub> of 50. Finally, 70 µL aliquots of the cells were shock-frozen in liquid nitrogen and stored at – 80 °C.

### **Preparation of chemically competent *Y. lipolytica***

Preparation of chemically competent *Y. lipolytica* and transformation was performed according to the lithium acetate method established by Barth and Gaillardin, 1996<sup>[96]</sup>. In brief, a 10 mL preculture (YPD, pH 4.0) of the recipient strain was incubated overnight at 30 °C, 200 rpm. The preculture was used to inoculate 10 mL main culture (YPD, pH 4.0) with an OD<sub>600</sub> of 0.05. The cells were allowed to grow to an OD<sub>600</sub> of approximately 9 – 12 and then harvested by centrifugation (5 min, 4,000 g, RT). After washing twice in TE buffer and discarding the supernatant, the pellet was diluted in 0.1 M lithium acetate (pH 6.0) to an OD<sub>600</sub> of 5 and incubated for 1 h at 30 °C and 90 rpm. After a final centrifugation step (5 min, 4,000 g, RT), the pellet was resuspended in 1/10 volume of 0.1 M lithium acetate (pH 6.0), resulting in a final OD<sub>600</sub> of 50. The cells were stored at 4 °C and used within one week.

### **Preparation of electrocompetent *Y. lipolytica***

Preparation of competent *Y. lipolytica* and transformation was performed according to Markham et al., 2018<sup>[144]</sup>. A 10 mL preculture (YPD) of the recipient strain was incubated overnight at 30 °C, 200 rpm and used to inoculate 10 mL main culture (YPD) at an OD<sub>600</sub> of 0.05. The main culture was transferred to a 50 mL Falcon tube on the next morning and centrifuged for 5 min, 500 g at RT. After discarding the supernatant, the cell pellets were washed twice with 50 mL of 1 M sorbitol. Finally, the cells were resuspended in 1 M sorbitol to result in a final OD<sub>600</sub> of 30. 200 µL aliquots of these competent cells were either used directly for transformation or stored at – 80 °C for later use.

### 2.2.1.3 Determination of the cell dry weight and optical density

To determine the cell dry weight (CDW) from shake flask experiments, 400 mL cell culture was centrifuged (10 min, 8,000 g, RT) in 40 mL portions in falcon tubes, which were pre-dried (48 h, 80 °C) and pre-weighted. The resulting cell pellets were dried at 80 °C for 72 h. After cooling for 30 min at RT, the falcons were weighted again and the dry mass calculated. The derived masses were used to correlate the OD<sub>600</sub> signal to biomass concentration.

The determination of optical density as a measure of biomass was conducted at a wavelength of 600 nm in 1.5 mL disposable plastic cuvettes with a layer thickness of 1 cm. If the OD<sub>600</sub> exceeded 1.5, the sample was diluted 1/10 or 1/20 in medium, respectively.

### 2.2.1.4 Growth curves of *Y. lipolytica*

Initially, 10 mL of K1 minimal medium supplemented with 15 g L<sup>-1</sup> glucose and the corresponding antibiotics were inoculated with *Y. lipolytica* empty vector control or variants and incubated at 30 °C and 200 rpm for 48 h. These precultures were washed twice and used to inoculate 25 mL of K1 minimal medium supplemented with 2 % methanol, 0.5 g L<sup>-1</sup> yeast extract and the corresponding antibiotic with an OD<sub>600</sub> of 0.05. Cultivation was performed at 30 °C and 200 rpm using independent triplicates. After the lag phase, the OD<sub>600</sub> values of the individual cultures were measured every 1.5 h. To quantify methanol uptake, every 1.5 h a sample of 1 mL culture broth was centrifuged (3 min, 13,000 rpm, RT) and the supernatant was stored at – 20 °C until further analysis using HPLC (2.2.4.1). The specific growth rate  $\mu$  was calculated from an exponential fit of optical density as a function of time during exponential growth.

### 2.2.1.5 Cultivation of *Y. lipolytica* in a high-throughput microbioreactor

For obtaining high-throughput growth profiles of *Y. lipolytica* empty vector control or variants, 1 mL microtiter plate (MTP, FlowerPlate<sup>®</sup>, Beckman Coulter, Brea, United States of America) cultures were incubated at 30 °C and 1,000 rpm in the BioLector<sup>®</sup> XT (Beckman Coulter, Brea, United States of America). In detail, precultures with 10 mL of K1 minimal medium containing 15 g L<sup>-1</sup> glucose were inoculated using a single colony grown on K1 agar plates. Precultures were incubated overnight in a shaker at 30 °C and 200 rpm and washed twice with K1 minimal medium supplemented with 2 % methanol prior to inoculation. For MTP cultures, 5 mL of K1 minimal medium was supplemented with 2 % methanol, 2 g L<sup>-1</sup> yeast extract and the corresponding antibiotic and inoculated with an OD<sub>600</sub> of 0.05. Subsequently, 1 mL of each main culture was transferred to the MTP plate wells. Cultivation was carried out in randomly distributed triplicates. If not indicated otherwise, the MTPs were

sealed with sterile, gas-permeable sealing foil with evaporation reduction layer (Beckman Coulter, Brea, United States), allowing ventilation of the wells at reduced evaporation rates.

#### 2.2.1.6 Formaldehyde toxicity of *E. coli*

A 10 mL preculture of *E. coli* BL21(DE3) was incubated overnight at 37 °C, 150 rpm in TB (Terrific broth) medium supplemented with 50 µg mL<sup>-1</sup> kanamycin. This preculture was used to inoculate 25 mL of a second preculture (TB medium supplemented with 50 µg mL<sup>-1</sup> kanamycin, 250 mL baffled flask) with an OD<sub>600</sub> of 0.05 and grown at 37 °C and 150 rpm. When an OD<sub>600</sub> of 1 was reached, 1 mM IPTG was added to induce protein expression. Main cultures were then performed using the BioLector<sup>®</sup> XT (Beckman Coulter, Brea, United States of America). The prior to incubation induced medium was transferred to the flower plate (MTP-48-BOH, Lot. 1202, m2p Labs, Germany), equipped with optodes for online monitoring of DO and pH value. All cultivations were performed in triplicates at 25 °C, a shaking frequency of 1,000 rpm, a shaking diameter of 3 mm and a filling volume of 1 mL per well. MTPs were sealed with sterile, non-woven, gas-permeable sealing foil (Beckman Coulter, Brea, United States). Following 1 h of FLS expression, 3 mM formaldehyde were added manually to all wells, to analyze the effects of the mutant FLS on the tolerance of *E. coli* against formaldehyde. Empty vector (pET24a) was used as a negative control.

For the determination of wet biomass, 200 mL of TB supplemented with 50 µg mL<sup>-1</sup> kanamycin was inoculated with an OD<sub>600</sub> of 0.05 and cultivated at 37 °C, 150 rpm. As described previously, the temperature was lowered to 25 °C when OD<sub>600</sub> reached 1, and 1 mM IPTG was added. After 1 h of cultivation at 25 °C and 150 rpm, 3 mM formaldehyde was added. Cultivation was continued for further 24 h and *E. coli* cell pellets were harvested by centrifugation for 10 min at 8,000 g, at 4 °C. Supernatant was discarded thoroughly, and wet biomass was weighted. The determination of formaldehyde toxicity was conducted in three individual biological replicates.

## 2.2.2 Molecular methods

### 2.2.2.1 Plasmid isolation

For plasmid isolation from *E. coli* a 10 mL preculture in LB medium containing the appropriate antibiotic was inoculated with a single clone bearing the desired plasmid. After incubation at 37 °C and 150 rpm overnight, 2 mL of the culture was harvested by centrifugation (2 min, 13,300 g). The supernatant was removed and the cell pellet was treated

according to the manufacturers' manual of the GeneJET Plasmid MiniPrep kit. For DNA elution 30  $\mu\text{L}$  of pre-warmed (10 min, 65 °C) sterile MilliQ water was used.

For plasmid isolation from *Y. lipolytica* a 10 mL preculture in K1 or YPD medium including the corresponding antibiotic was inoculated with a single colony and incubated at 30 °C and 200 rpm overnight. On the next day, 2 mL of the culture was centrifuged (3,000 g, 2 min) and the supernatant discarded. The pellet was treated according to the manufacturers' manual of the Zymoprep Yeast Plasmid MiniPrep II kit. After the addition of solution 3, the tubes were centrifuged at 13,300 g for 10 min instead of the recommended 3 min. DNA elution was performed in two steps with 10  $\mu\text{L}$  of sterile, prewarmed (65 °C) MilliQ water.

DNA concentration was measured using UV/Vis spectroscopy (2.2.2.9).

#### 2.2.2.2 Restriction digestion

Sequence specific restriction digest was performed to cleave the DNA of the backbone plasmid or the amplified PCR-inserts. Furthermore, this technique was exploited to check for positive clones after plasmid isolation (2.2.2.1). The incubation conditions such as time, temperature, and buffer for restriction enzymes were adjusted according to the manufacturer's instructions. The restriction digest was prepared as shown in Table 15 and mixed by carefully pipetting up and down.

**Table 15: Restriction reaction.**

Component	Volume
DNA (1 $\mu\text{g}/\mu\text{L}$ )	10 $\mu\text{L}$
10x reaction buffer	5 $\mu\text{L}$
Restriction enzyme 1	2 $\mu\text{L}$
Restriction enzyme 2	2 $\mu\text{L}$
ddH <sub>2</sub> O	31 $\mu\text{L}$
Total	50 $\mu\text{L}$

If plasmids were used for subsequent ligation, dephosphorylation was performed by adding antarctic phosphatase according to the manufacturers' manual. The fragments obtained from the restriction digestion were examined using agarose gel electrophoresis (2.2.2.6) and, if required, isolated from the gel (2.2.2.6).

### 2.2.2.3 Ligation

After restriction digestion and gel purification of inserts and backbone, the two double-stranded DNA fragments were ligated. For ligation, the quantities of vector and insert DNA were determined considering the number of base pairs (bp). Typically, one bp corresponds to a molecular mass of  $6.6 \times 10^{-4} \text{ ng fmol}^{-1}$ . The quotient of vector and insert DNA was established between 1:3 (100 fmol vector: 300 fmol insert DNA) and 1:5 (100 fmol vector: 500 fmol insert DNA) and the respective volumes were calculated according Equation 1.

$$[\text{Eq. 1}]: \quad c_{\text{DNA}} [\text{fmol } \mu\text{L}^{-1}] = \frac{c (\text{DNA} / \text{ng } \mu\text{L}^{-1})}{6,6 \cdot 10^{-4} \text{ ng fmol}^{-1} \times \text{base pairs}}$$

As a control, both insert and vector were ligated without the other.

### 2.2.2.4 Polymerase Chain Reaction

#### QuikChange (QC) PCR

QC PCR was used to replace various amino acids of the FLS enzyme. The composition of the master mix and the program of the QC PCR are shown in Table 16. After the PCR was completed, 10 U DpnI were added to the PCR mix and incubated for 4 h at 37 °C. Next, 2  $\mu\text{L}$  were used to transform electrocompetent *E. coli* BL21(DE3) (2.2.2.7). The annealing temperature ( $T_A$ ) was calculated using the QuikChange Primer Design program from Agilent.

**Table 16: Composition and program for QuikChange PCR.**

Composition		Program			
Pfu HF Buffer (10x)	2.5 $\mu\text{L}$		Time	T [°C]	Cycles
Template (50 ng/ $\mu\text{L}$ )	1 $\mu\text{L}$	Initial denaturation	2 min	95	1 x
dNTPs (10 mM)	0.5 $\mu\text{L}$	Denaturation	30 s	95	16 x
Forward-Primer (10 $\mu\text{M}$ )	0.5 $\mu\text{L}$	Annealing	1 min	$T_A$	
Reverse-Primer (10 $\mu\text{M}$ )	0.5 $\mu\text{L}$	Extension	2 min/kb	68	
Pfu polymerase	0.5 $\mu\text{L}$	Final extension	10 min	72	1 x
ddH <sub>2</sub> O	19.5 $\mu\text{L}$				
total	25 $\mu\text{L}$				

## Colony PCR

Colony PCR with Taq polymerase was used to screen for positive clones. Single colonies were resuspended in PCR tubes containing 20  $\mu\text{L}$  of cPCR master mix. The composition of the master mix and the program of the cPCR are shown in Table 17. The annealing temperature ( $T_A$ ) was defined at 5  $^\circ\text{C}$  beneath the melting point of the lower primer.

**Table 17: Composition and program for colony PCR.**

Composition		Program			
Template	~ 10 ng		Time	T [ $^\circ\text{C}$ ]	Cycles
Primer forward (10 $\mu\text{M}$ )	0.1 $\mu\text{L}$	Initial denaturation	5 min	95	1 x
Primer reverse (10 $\mu\text{M}$ )	0.1 $\mu\text{L}$	Denaturation	30 s	95	30 x
dNTPs (10 mM)	0.5 $\mu\text{L}$	Annealing	30 s	$T_A$	
Buffer (5x)	4 $\mu\text{L}$	Extension	1 min/kb	72	
Taq polymerase (5 U $\mu\text{L}^{-1}$ )	0.2 $\mu\text{L}$	Final extension	10 min	72	1 x
ddH <sub>2</sub> O	ad 20 $\mu\text{L}$				
total	20 $\mu\text{L}$				

Afterwards, the fragments were analyzed by agarose gel electrophoresis (2.2.2.6).

### 2.2.2.5 Golden Gate cloning

A detailed outline of the basic principles of GoldenMOCS can be found in Sarkari et al.<sup>[139]</sup>. In general, the episomal Golden-MOCS plasmids used in this study were ordered from Addgene (Table 10) and the experimental design adapted from the manufacturer's manual (Egermeier, Sauer, & Marx, 2019<sup>[138]</sup>). The inserts MDH2, MDH3, ADH and FLS were codon optimized for *Y. lipolytica* and synthesized by BioCat GmbH (Heidelberg, Germany, Table 11) as gBlocks. Golden Gate assembly was performed in a one-pot reaction using a type IIS restriction endonuclease (20 U BbsI or 20 U BsaI), 100 U T4 ligase, 2 mM ATP and CutSmart buffer. For assembly of first stage plasmids (BB1) 10 fmol of backbone and 20 fmol of insert were used. For later stages (BB2 and BB3), the donor plasmids, containing the inserts, were used in equimolar ratios at a final concentration of 10 fmol. The assembly underwent incubation for 45 cycles of restriction and ligation (37  $^\circ\text{C}$ , 1 min and 16  $^\circ\text{C}$ , 2.5 min), and then used to chemically transform *E. coli* DH10B (2.2.2.7). Colony PCR was used to confirm the presence and the correct size of the plasmid in the transformed strains. Stage 3 plasmids (BB3) were used to chemically transform *E. coli* DH10B, isolated as

described in chapter 2.2.2.7 and then used to transform *Y. lipolytica* using the lithiumacetate method established by Barth and Gaillardin 1996<sup>[96]</sup> (2.2.2.8).

#### 2.2.2.6 Agarose gel electrophoresis

To separate DNA from RNA and proteins, gel electrophoresis was used. According to the expected fragment size, 0.8 – 2 % agarose gels were prepared. For DNA fragments < 500 bp 2 % agarose, for fragments between 500 – 4,000 bp 1 % agarose and for fragments > 4,000 bp 0.8 % agarose was used. The required amount of agarose was diluted in 200 mL 1 x TAE buffer and boiled until complete solubilization.

For analytical purposes up to 5 µL DNA samples were mixed with 1 µL of loading dye (6x) and loaded onto the gel. For preparative electrophoresis the entire sample was loaded after mixing with loading dye (6x). 7 µL of 100 bp DNA ladder (#N3231) or 1 kb DNA ladder (#N3232) from New England Biolabs were used as a marker. The gel run was performed at 110 V for 45 min to 1.5 h. For analysis, an UV fluorescent table (Intas Gel iX20 Imager documentation system) was used to detect the ethidium bromide which binds DNA by intercalating between base pairs. For this purpose, the agarose gels were dyed with the ethidium bromide staining solution at RT for 10 minutes. Prior to analysis, the gel was destained in TAE buffer for at least 10 min at RT to minimize the background.

For the purification of DNA fragments, a gel slice containing the DNA fragment was excised and prepared according to the manual of the GeneJET Gel Extraction Kit. DNA was eluted from the column using 30 µL sterile MilliQ water.

#### 2.2.2.7 Transformation of competent *E. coli*

For transformation via heat shock, 100 µL of chemically competent cells were thawed on ice and up to 2 µL of vector DNA or 10 µL of Golden Gate mixture were added, mixed carefully and incubated on ice for 30 min. Afterwards a heat shock was performed in a water bath at 42 °C for 90 s. The tubes were then placed on ice for 1 min and 1 mL of SOC medium was added. Next, the cells were incubated for 1 h at 37 °C and 150 rpm. Finally, 100 – 200 µL cell suspension was plated on selective agar plates and incubated overnight at 37 °C.

For electroporation, 100 µL of competent cells were defrosted on ice and 1 µL of vector DNA was added and mixed gently. The cells were pipetted into an electroporation cuvette with a 1 mm gap (MBP) and electroporation was performed using 50 µF, 125 Ω, and 1380 V for BL21(DE3) or 25 µF, 200 Ω, and 2000 V for DH10B. After electroporation, cells were quenched with 1 mL of pre-warmed SOC medium and incubated for 1 h at 37 °C and



150 rpm. Subsequently, 100  $\mu\text{L}$  cell suspension was plated on selective agar plates and incubated overnight at 37 °C.

#### 2.2.2.8 Transformation of competent *Y. lipolytica*

For transformation of *Y. lipolytica* using heat shock, 5 – 10  $\mu\text{g}$  of DNA was mixed with 5  $\mu\text{L}$  of carrier DNA (salmon sperm DNA, 5  $\text{mg mL}^{-1}$  in 50 mM TE buffer, pH 8.0) and 100  $\mu\text{L}$  of competent cells and incubated for 15 min at 30 °C. After addition of 0.7 mL of 30 % PEG 8000 (dissolved in 0.1 M lithium acetate, pH 6.0), the cells were shaken for 1 h at 30 °C and 90 rpm. Next, a heat shock was performed at 39 °C for 10 min. Cells were then resuspended in 1.2 mL of 0.1 M lithium acetate pH 6.0. Thereafter, 200  $\mu\text{L}$  were plated on K1 agar plates with the appropriate supplements and antibiotics. Plates were incubated at 30 °C. Depending on the plasmid and strain used, the transformants were visible as colonies after three to five days.

For transformation of *Y. lipolytica* using electroporation, 200  $\mu\text{L}$  of competent cells were mixed with up to 5  $\mu\text{g}$  of linearized DNA and transferred to an electroporation cuvette with 1 mm gap (MBP). Cells were electroporated at 25  $\mu\text{F}$ , 200  $\Omega$ , and 2700 V. Next, cells were quenched with 600  $\mu\text{L}$  YPD two times, and transferred to a 15 mL Falcon tube. Cells were recovered at 30 °C, 90 rpm for one hour. After recovery, cells were centrifuged at 500 g for 5 min and washed with 1 mL of 1 M sorbitol. Cell pellets were then resuspended in 1 mL of 1 M sorbitol. Finally, 200  $\mu\text{L}$  were plated on K1 selective agar plates containing the desired carbon source and the corresponding antibiotic. Plates were incubated for two to five days at 30 °C.

#### 2.2.2.9 UV/Vis spectroscopy

For the determination of DNA, the Nano Drop ND-Lite spectrometer (PeqLab) was used. It measures the absorbance at 260 nm, whereas an absorbance of 1 equal 50  $\text{ng } \mu\text{L}^{-1}$  of dsDNA. The DNA concentration is calculated according to Equation 2 considering the dilution factor:

$$[\text{Eq. 2}] c_{\text{dsDNA}} [\text{ng } \mu\text{L}^{-1}] = A_{260\text{nm}} \times 50 \text{ ng } \mu\text{L}^{-1} \times df$$

$A_{260\text{nm}}$ : absorbance at 260 nm

$df$ : dilution factor

The NanoDrop furthermore calculates the purity of the nucleic acids by determining the ratio between the absorption maximum of proteins at 280 nm and the absorption maximum of nucleic acids at 260 nm. The quotient for DNA should not exceed 1.8.

To determine the concentration of proteins, the UV-1800 spectrophotometer (Shimadzu) was used. The protein concentration was measured at 280 nm using a quartz cuvette (Ultra-Micro Cell QS LP 10 mm, Hellma Analytics) and calculated based on the *Lambert-Beer* law, taking into account the dilution factor (Equation 3).

$$[\text{Eq. 3}]: \quad A_{280\text{nm}} = \varepsilon \times c_p \times d \times d_f$$

$A_{280\text{nm}}$ : absorbance at 280 nm

$\varepsilon$ : molar attenuation coefficient [ $\text{L mol}^{-1} \text{cm}^{-1}$ ]

$c_p$ : protein concentration

$d$ : thickness of the cuvette [cm]

$d_f$ : dilution factor

The molar attenuation coefficients for each protein were computed using the ProtParam tool available from <https://web.expasy.org>.

#### 2.2.2.10 DNA sequencing

The sequencing of the constructed plasmid to verify the DNA sequence was performed at Eurofins MWG GmbH or Microsynth Seqlab GmbH. The primers used for sequencing are given in Table 9.

### 2.2.3 Biochemical methods

#### 2.2.3.1 Protein expression

Parental FLS and variants were expressed using TB medium and IPTG. A single colony of *E. coli* BL21(DE3) carrying plasmids with the gene of interest was inoculated into 10 mL TB medium supplemented with  $50 \mu\text{g mL}^{-1}$  kanamycin and grown overnight in a 100 mL baffled flask containing at  $37^\circ\text{C}$  and 150 rpm. On the next morning, the preculture was used to inoculate 200 mL main culture (TB medium supplemented with  $50 \mu\text{g mL}^{-1}$  kanamycin, 1 L baffled flask) with an  $\text{OD}_{600}$  of 0.05. The culture was grown at  $37^\circ\text{C}$  and 135 rpm. When an  $\text{OD}_{600}$  of 0.8 – 1 was reached, 1 mM IPTG was added and the temperature was lowered to  $25^\circ\text{C}$ . After approximately 20 h, the cells were pelleted by centrifugation at 8,000 g for 10 min and  $4^\circ\text{C}$ . Supernatant was discarded and the cell pellets used for cell disruption.

#### 2.2.3.2 Lysis and cell disruption by sonication

All cell disruptions in this work were done via sonication using the HD 2200 sonotrode from Bandelin. First, the cell pellet (10 % w/v) was dissolved in lysis buffer containing  $2.5 \text{ mg mL}^{-1}$  lysozyme and  $0.1 \text{ mg mL}^{-1}$  DNase. After incubation on ice for 1 h, the cell suspension was sonicated. Sonication was performed five times for 1 min each (0.5 s cycle,

55 %) on ice. Between each sonication, the cells were incubated on ice for 1 min. After sonication, the cell suspension was centrifuged (30 min, 30,000 g, 4 °C) and the supernatant transferred to a Falcon tube. The obtained cell lysates were either used for enzyme assays or subjected to protein purification (2.2.3.3).

#### 2.2.3.3 Protein purification by affinity chromatography

Parental FLS and variants carried a hexahistidine tag at the C-terminus and were purified via Immobilized Metal Affinity Chromatography (IMAC). Purification was performed using an Äkta protein purification system. The crude cell lysate was loaded onto a 1 mL HisTrap<sup>TM</sup> FF column (HisTrap HP) using a flow rate of 3.5 mL min<sup>-1</sup>. The column was then washed with buffer A for 30 column volumes (CV) with a 3.5 mL min<sup>-1</sup> flow rate. By washing with 20 CV of 10 % buffer B with a flow rate of 3.5 mL min<sup>-1</sup>, weakly bound substances such as other *E. coli* proteins were eluted from the column. To elute the protein of interest, the concentration of buffer B was increased to 60 % for 20 CV. A higher concentration of imidazole breaks the interaction between the nickel ions and the hexahistidine tag, releasing the tagged protein from the Ni-column. In the final step of the gradient, 5 CV of 100 % elution buffer were used to remove all solidly bound substances still remaining on the column. Finally, a single fraction containing the protein of interest was collected and the column was equilibrated with buffer A using a flow rate of 3.5 mL min<sup>-1</sup> for 20 CV.

After protein purification, the buffer was exchanged to remove imidazole and salt in the eluted protein. For this purpose, the PD-10 Desalting column was equilibrated with 10 CV of protein buffer and then loaded with up to 2.5 mL of the purified protein solution. The flow-through was discarded and the protein eluted using 3.5 mL of protein buffer.

Protein purity was analyzed using SDS-PAGE (2.2.3.4).

#### 2.2.3.4 SDS-PAGE according to Laemmli<sup>[145]</sup>

SDS-PAGE was used for the separation of proteins. Sodium dodecyl sulfate (SDS) denatures proteins which are then separated according to their molecular weight. First, 100 µL of protein sample were mixed with 100 µL of SDS sample buffer (6x) and boiled for 5 min at 95 °C. After heat-denaturation, approximately 10 – 30 µg of each sample were separated on a 12 % separating gel, topped with a 5 % stacking gel. The components of both gels are shown in Table 18. In addition, 7 µL of a molecular weight marker (Unstained Protein Standard, Broad Range (10 – 200 kDa) from NEB) were used as a size standard. Gel electrophoresis was performed at 25 mA for 45 – 60 min. Subsequently, the gels were stained

with Coomassie staining solution at RT for 20 min. To reduce the background, the gels were boiled twice in ddH<sub>2</sub>O and analyzed using the Intas Gel iX20 Imager documentation system.

**Table 18: Formulation of separating and stacking gel for SDS-PAGE analysis.**

Component	Separating gel (12 %)	Stacking gel (5 %)
Acrylamide-, bis-acrylamide stock solution [30 % (w/v), 37.5:1]	3 mL	0.65 mL
ddH <sub>2</sub> O	4.29 mL	3 mL
Buffer (4x, separating and stacking gel)	2.5 mL	1.25 mL
SDS [10 % (w/v)]	100 µL	50 µL
Ammonium persulfate [10 % (w/v)]	100 µL	50 µL
Tetramethylethylenediamine	10 µL	5 µL
Total	10 mL	5 mL

#### 2.2.3.5 3D structure modelling and *in silico* engineering

In order to find suitable target positions, docking studies using the crystal structure of FLS (PDB ID: 4QQ8) were performed. The 2D structures of ligands (TPP, FALD & DHA) were created using ChemDraw Professional™ (Version 22.0), then transformed into 3D structures (.pdb) and their energy minimized using Chem3D™ (Version 17.0). Docking experiments were executed with YASARA, utilizing AutoDock VINA for docking the ligand to a receptor. The FLS receptor was prepared according to the YASARA standard protocol. PyMOL (Version 2.4.1) was used for *in silico* analysis. FLS variants were generated accordingly, amino acid residues were swapped to the desired amino acid, and the resulting protein energy minimized.

#### 2.2.3.6 Directed evolution and FLS mutant screening

To overcome the challenges associated with the application of the FLS enzyme *in vivo*, mutant libraries were designed and tested in a robotic-aided high-throughput screening. Positions H29, Q113, L482, I556 and I557 were identified as hot spots upon *in silico* analysis and docking studies using YASARA AutoDock (2.2.3.5) and subjected to site-saturation mutagenesis (2.2.2.4). The subsequent library screening was carried out in deep-well-plates (96DWP). For the liquid precultures, each well was filled with 1.2 mL TB media supplemented with 50 µg mL<sup>-1</sup> kanamycin and inoculated with single colonies using an automatic colony picker (CP7200, Norgren Systems, United Kingdom). Three wells were

inoculated with empty vector controls, four with wild type FLS and another four wells served as sterile controls. After incubation for 14.75 h at 37 °C and 1,000 rpm, OD<sub>710</sub> was measured using a robotic pipetting system (Tecan Freedom Evo 200, Switzerland). To equalize the cell concentration, the cell cultures were diluted individually according to their OD<sub>710</sub> values using an empirical developed pitch function (Equation 4).

[Eq. 4]:

$$y = 87.678x^6 - 325.59x^5 + 484.09x^4 - 361.5x^3 + 142.19x^2 - 28.814x + 3.6536$$

The precultures were used to inoculate the main culture plates (96DWP) prefilled with 1 mL medium and incubated for 3.5 h at 37 °C and 1,000 rpm. In addition, 100 µL of the preculture were mixed with 100 µL 60 % glycerol in a microtiter plate and frozen at – 80 °C and served as retention samples. For protein expression, 108 µL of isopropyl-β-D-thiogalactoside (IPTG) solution was added to each main culture well at a final concentration of 1 mM. Plates were incubated at 25 °C and 1,000 rpm for 20 h and then centrifuged (15 min, 3,000 g, 4 °C). The supernatant was discarded and pellets were resuspended in 500 µL lysis buffer and incubated at 37 °C and 1000 rpm for 1.5 h. After centrifugation (15 min, 3,000 g, 4 °C), 300 µL clear cell lysate was transferred to a microtiter plate via the pipetting robot and analyzed using the GDH assay (2.2.3.7).

#### 2.2.3.7 GDH assay and kinetic constant determination

DHA production from FALD was measured spectrophotometrically at 340 nm. In a coupled enzyme assay<sup>[78]</sup>, FALD was converted to DHA, which in turn was reduced by an NADH-dependent glycerol dehydrogenase (GDH). The assay was performed with either supernatant containing FLS wild type or variants or purified enzyme. In detail, 100 µL of supernatant or 100 µL of purified protein (300 µg mL<sup>-1</sup>) were combined with 100 µL assay mix (100 mM phosphate buffer, pH 8.0 supplemented with 1 mM MgSO<sub>4</sub>, 1.6 mM NADH, 268 mM FALD, 0.1 mM TPP and 100 µg mL<sup>-1</sup> glycerol dehydrogenase). NADH concentrations were assayed over a 1-h period at 30 °C and 340 nm.

To validate the robustness of the screening, the Z-factor<sup>[146]</sup> was determined for the GDH assay using FLS wild type (positive control) lysate or pET24a empty vector (negative control) lysate according to Equation 5.

$$[\text{Eq. 5}]: \quad Z = 1 - \frac{3(\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$

$\sigma_p$ : standard deviation of positive control

$\sigma_n$ : standard deviation of negative control

$\mu_p$ : mean value of positive control

$\mu_n$ : mean value of negative control

Kinetic constants were measured over a 1-h period using the previously described GDH assay. 100  $\mu\text{L}$  of purified protein (300  $\mu\text{g mL}^{-1}$ ) was combined with 100  $\mu\text{L}$  assay mix (100 mM phosphate buffer, pH 8.0 supplemented with 1 mM  $\text{MgSO}_4$ , 1.6 mM NADH, 0.1 mM TPP and 100  $\mu\text{g mL}^{-1}$  glycerol dehydrogenase). Substrate concentration ranged from 1.25 to 125 mM. Kinetic constants for FLS wild type or variant L482Q/I557G using FALD as a substrate were determined by employing nonlinear regression to fit the rate of product production as a function of substrate concentration using the Michaelis–Menten equation (Equation 6).

$$\text{[Eq. 6]: } y = \frac{v_{max} \times X}{K_M + X}$$

$v_{max}$ : maximum reaction rate

$K_M$ : Michaelis–Menten constant

#### 2.2.3.8 Thermal shift assay (TSA)

Melting temperature ( $T_m$ ) of parental FLS and variants was determined indirectly by following their interaction with SYPRO™ orange. This dye fluoresces when binding to exposed hydrophobic regions of the enzyme during unfolding. TSA measurements were performed in triplicates using the CFX96 Touch Real-Time PCR detection system (Biorad, USA). 21  $\mu\text{L}$  of protein buffer (100 mM phosphate buffer, pH 8.0; 1 mM  $\text{MgSO}_4$ ; 0.1 mM TPP) were mixed with 2  $\mu\text{L}$  of diluted SYPRO™ orange (5000x in DMSO) and 2  $\mu\text{L}$  of purified enzyme (1  $\text{mg mL}^{-1}$ ). The assay was performed using a temperature gradient from 5 – 100 °C with an applied heating rate of 0.5 °C/5 sec.

### 2.2.4 Analytics

#### 2.2.4.1 High-performance liquid chromatography (HPLC)

Separation and quantification of metabolite concentrations was performed using an HPLC system from Shimadzu (LC-20AD pumps, SIL-20AC auto sampler, CTO-20AC oven, CBM-20A controller) via a Rezex™ ROA-Organic Acid H+ (8 %) column (300 mm x 7.8 mm) and an isocratic flow using 0.005  $\text{mol L}^{-1}$   $\text{H}_2\text{SO}_4$  as an eluent at a flow rate of 0.5  $\text{mL min}^{-1}$  and 65 °C for 27 min. For metabolites FALD and DHA, the method was changed to 45 °C. Prior to HPLC analysis, all samples were centrifuged for 3 min at 13,800 rpm. Subsequently, 10  $\mu\text{L}$  of standard or sample were applied to the system. Glucose,

methanol, FALD and glycerol were detected with RI, DHA and GALD were detected with UV detector. Qualification and quantification of metabolite concentrations was conducted using external standards (6.2). The resulting chromatograms were interpreted using the LabSolution post-run analysis software.

# 3 RESULTS

## 3.1 Characterization of the oleaginous yeast *Y. lipolytica* as a host for synthetic methylotrophy

A *Y. lipolytica* strain to be selected as a host for synthetic methylotrophy must fulfill various criteria. It should be able to produce an industrially relevant product, it should grow fast in minimal to rich growth media, and it must be safe to handle. In addition, it would ideally already tolerate increased concentrations of methanol, and the intermediates FALD, GALD and DHA.

### 3.1.1 Characterization of DSM3286 and PO1f on various substrates

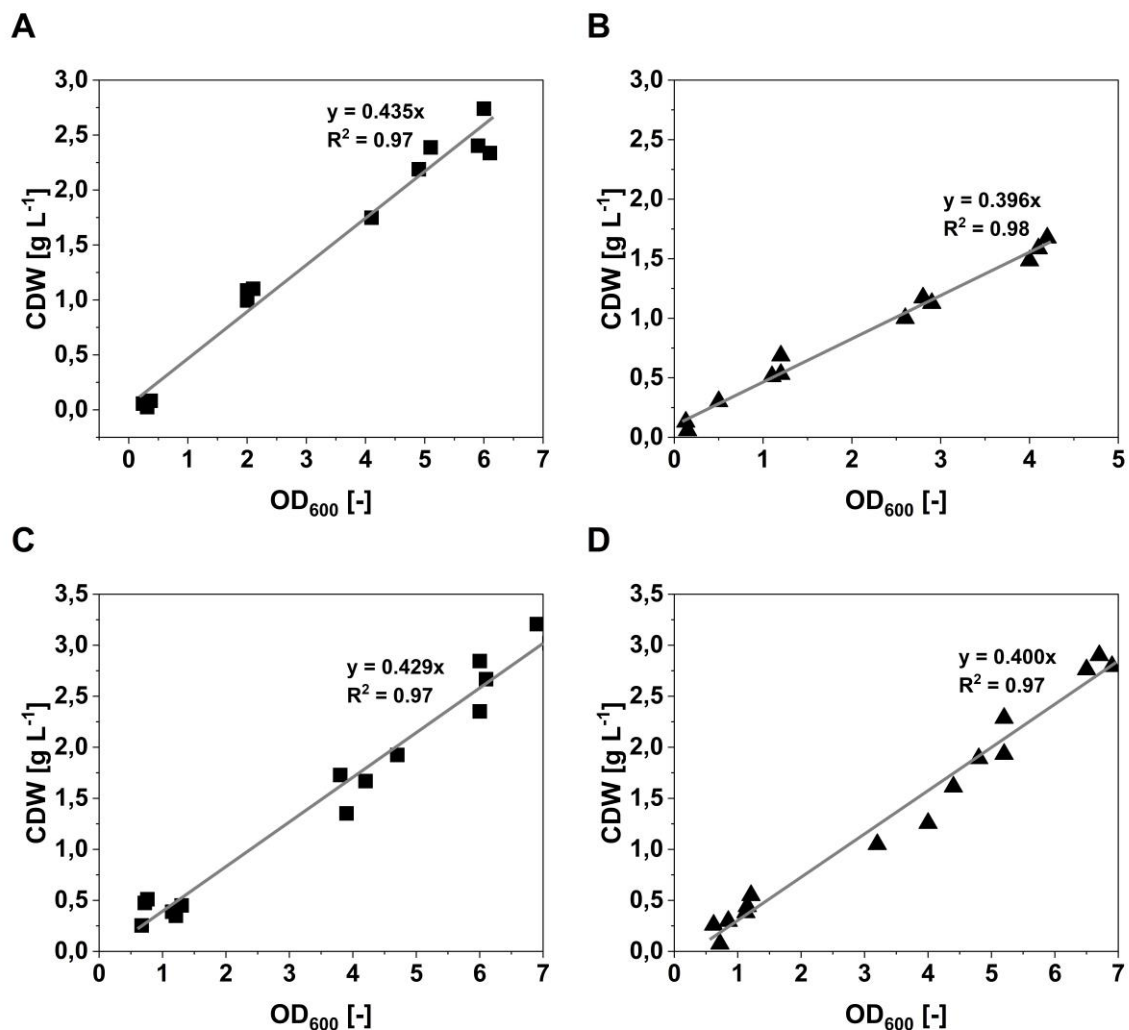
Two promising strains were evaluated as hosts for synthetic methylotrophy based on rational criteria. The most frequently used host is the PO1f strain, which is easy to manipulate due to a leucine and uracil auxotrophy and the ability to use sucrose as the sole carbon source<sup>[120]</sup>. DSM3286 on the other hand depicts an attractive wild type candidate with its ability to produce single cell oil highly suitable for quality biodiesel production in a cost-effective manner<sup>[142]</sup>.

The initial step was to determine the relationship between biomass cell dry weight (CDW) and optical density (OD, measured at 600 nm) to evaluate the performance of both strains in shake flasks using simple carbon sources. Figure 4 illustrates the OD-CDW correlation, and allows the interpolation of absorbance values to dry weight using either glucose (Figure 4A, B) or glycerol (Figure 4C, D) as the sole carbon source. On glucose, DSM3286 reached significantly higher OD and CDW values than PO1f (0.396 g L<sup>-1</sup> CDW per 1 OD<sub>600</sub>), with a linear relationship of  $y = 0.435x$  and with a positive correlation of 0.97. Using glycerol as carbon source, both strains exhibited higher OD values compared to growth on glucose. Nevertheless, strain DSM3286 reached again higher CDW values compared to PO1f. When grown on glycerol, a factor of 0.429 or 0.400 g L<sup>-1</sup> CDW per 1 OD<sub>600</sub> unit was obtained for *Y. lipolytica* DSM3286 or *Y. lipolytica* PO1f, respectively.

This is in accordance with growth curves obtained from culturing both *Y. lipolytica* strains on various substrates in the BioLector I (Figure S1). DSM3286 reaches higher OD values compared to PO1f at all conditions and also tolerates small amounts of methanol (7.9 g L<sup>-1</sup>)



or formate ( $1.5 \text{ g L}^{-1}$ ) when combined with glucose in minimal media. Even though PO1f depicts slightly higher growth rates (Table S1), only marginal biomass formation can be observed in K1 minimal media (MM). For both strains, growth on methanol or formate as the single carbon source was hard to detect, therefore the associated growth rates could not be reliably calculated.

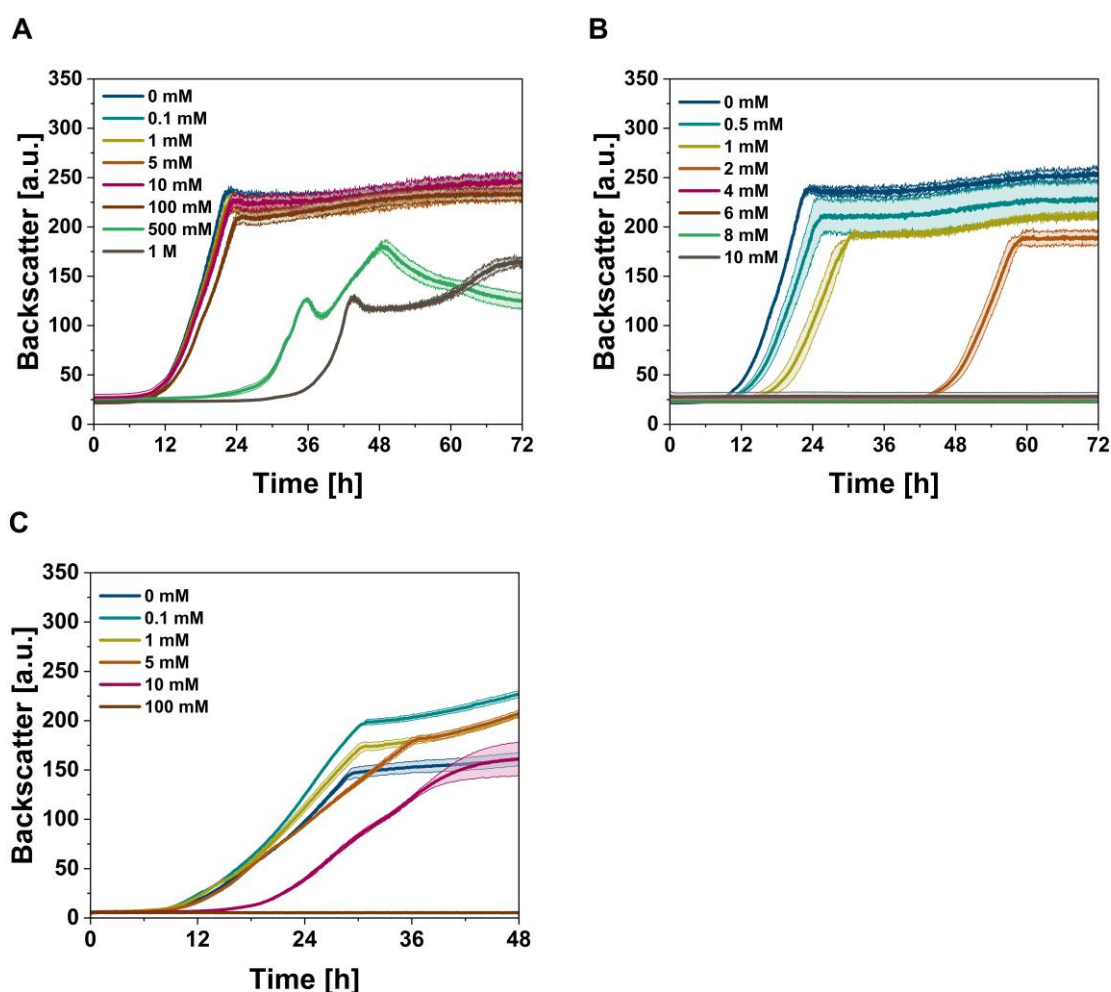


**Figure 4: The correlation between dry biomass weight and optical density for growth of two *Y. lipolytica* strains.** Shown are the OD-CDW correlations of strains DSM3283 (squares) and PO1f (triangles) using 2 L shake flasks and a culture volume of 400 mL. Cells were grown in minimal medium using  $15 \text{ g L}^{-1}$  glucose (A and B) or  $15 \text{ g L}^{-1}$  glycerol (C and D) as carbon sources, respectively. Cultivation was conducted at  $30 \text{ }^\circ\text{C}$  and 200 rpm in three independent replicates.

### 3.1.2 Methanol and C<sub>1</sub> intermediate tolerance

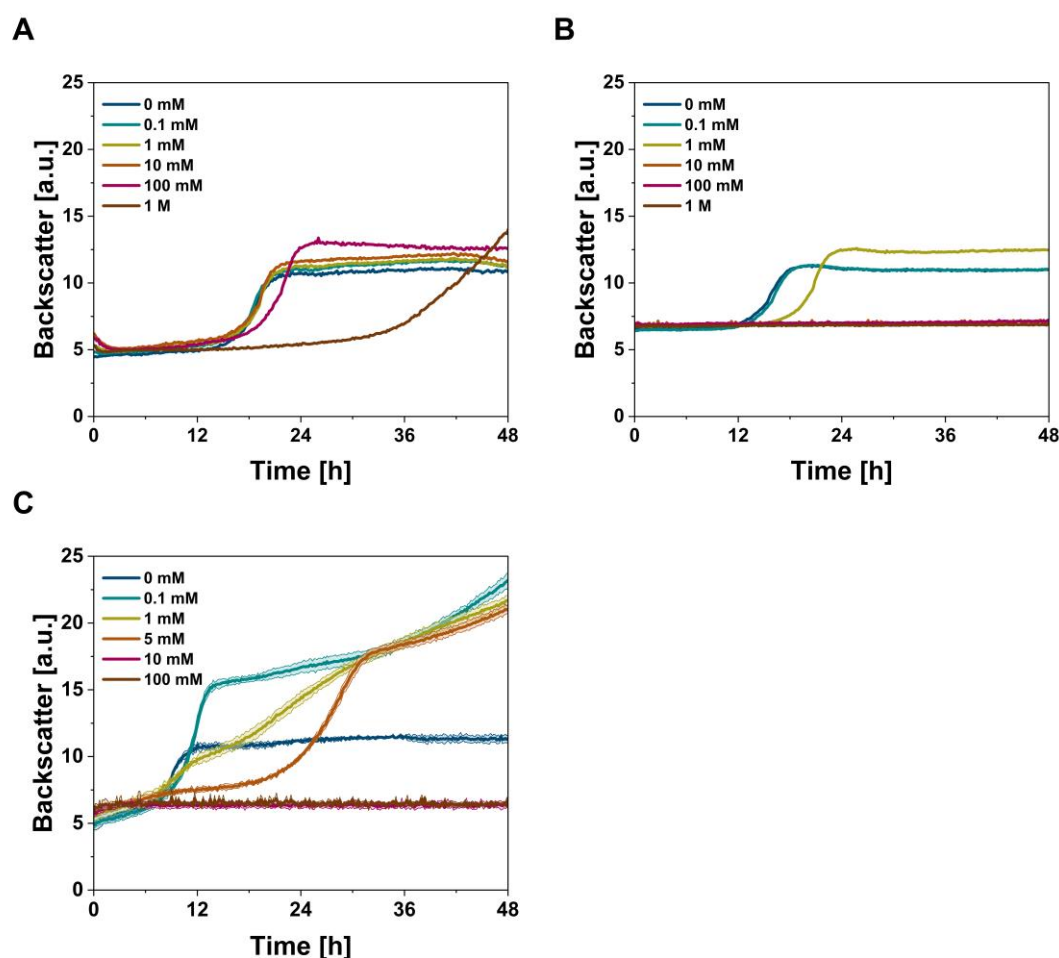
Next, the effect of different amounts of methanol, and the associated intermediates FALD and DHA on the growth of wild type *Y. lipolytica* DSM3286 was tested in microtiter-scale cultivation (i.e., 1 mL). Cells were grown in K1 MM supplemented with  $15 \text{ g L}^{-1}$  glucose in the presence of different concentrations of methanol or the previously named intermediates

(Figure 5). In these conditions, the addition of 100 mM methanol caused a slight reduction of the maximal growth rate, while higher methanol concentrations led to a prolonged lag-phase and decreased biomass formation. In addition, cell growth in two phases was observed at high (500 mM and above) methanol concentrations (Figure 5A). FALD (Figure 5B) is highly toxic for *Y. lipolytica* DSM3286 and is only tolerated up to concentrations of 2 mM. While the addition of FALD leads to a lower biomass formation in general, addition of small amounts of DHA (up to 5 mM) appear to have a positive growth effect. Nevertheless, supplementation of 10 mM DHA is characterized by a prolonged lag-phase and a lower growth rate, whereas 100 mM DHA abolishes growth altogether (Figure 5C).



**Figure 5: Effect of methanol and C<sub>1</sub> intermediates on the growth behavior of *Y. lipolytica* DSM3286.** Shown is the effect of methanol (A), formaldehyde (B), and dihydroxyacetone (C). Cells were cultivated in liquid minimal medium supplemented with 15 g L<sup>-1</sup> glucose. Cultivations were performed in 48-well plates, in 1 mL culture volume at 30 °C and 1000 rpm. Biomass gain was measured using the BioLector I or the BioLector XT. Shown are the mean values of three biological replicates. Error bars indicate standard deviations. In (B) 4 mM (pink), 6 mM (brown) and 8 mM (green) lines are hidden under 10 mM (black) line.

The widely used *Y. lipolytica* strain PO1F was also tested for its methanol, FALD and DHA tolerance (Figure 6), but growth on K1 medium is impaired in this strain so that high biomass growth was not achieved. A positive effect on growth was observed up to a concentration of 100 mM methanol (Figure 6A). Similar to strain DSM3286, FALD (Figure 6B) is highly toxic for *Y. lipolytica* PO1f and is tolerated only at very low concentrations (maximum 1 mM). In addition, a concentration of 1 mM FALD leads to a prolonged lag phase and marginal higher OD values. Supplementation of up to 5 mM of DHA improves biomass production, whereas 10 mM DHA diminishes growth altogether (Figure 6C).

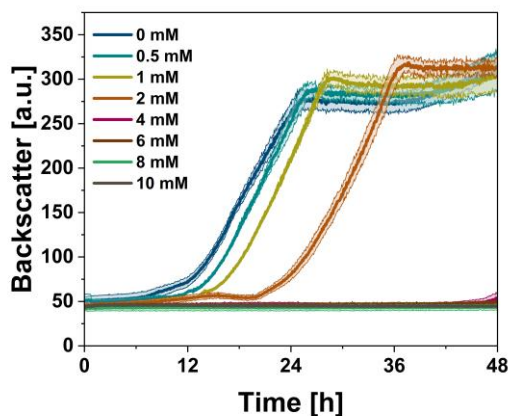


**Figure 6: Effect of methanol and C<sub>1</sub> intermediates on the growth behavior of *Y. lipolytica* PO1f.** Shown is the effect of methanol (A), formaldehyde (B), and dihydroxyacetone (C). Cells were cultivated in liquid minimal medium supplemented with 15 g L<sup>-1</sup> glucose. Cultivations were performed in 48-well plates, in 1 mL culture volume at 30 °C and 1000 rpm. Biomass gain was measured using the BioLector I. Shown are the mean values of two (A, B) or three (C) biological replicates. Error bars indicate standard deviations (C).

Based on these findings, strain DSM3286 was chosen as a host for synthetic methylotrophy (see also chapter 3.3). Compared to *Y. lipolytica* PO1f, DSM3286 not only tolerates higher

amounts of methanol and other relevant intermediates, but also achieves considerably higher biomass in minimal media.

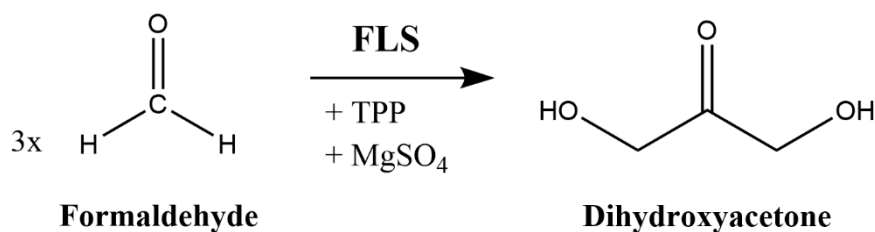
One obstacle when using the FLS pathway *in vivo* is the formation of the side product GALD. Therefore, strain DSM3286 was additionally tested for its tolerance towards the intermediate of the FLS pathway, GALD. Figure 7 depicts the effect of various GALD concentrations on the growth behavior of *Y. lipolytica* DSM3286. Similar to FALD, GALD is highly toxic for the cells and concentrations exceeding 2 mM cease cell growth.



**Figure 7: Effect of glycolaldehyde on the growth behavior of *Y. lipolytica* DSM3286.** Cells were cultivated in liquid minimal medium supplemented with 15 g L<sup>-1</sup> glucose. Cultivations were performed in 48-well plates, in 1 mL culture volume at 30 °C and 1000 rpm. Biomass gain was measured using the BioLector XT. Shown are the mean values of three biological replicates. Error bars indicate standard deviations. 4 mM (pink), 6 mM (brown) and 8 mM (green) lines are hidden under 10 mM (black) line.

### 3.2 Robot-aided high-throughput engineering of FLS enzyme for *in vivo* synthetic one-carbon metabolism

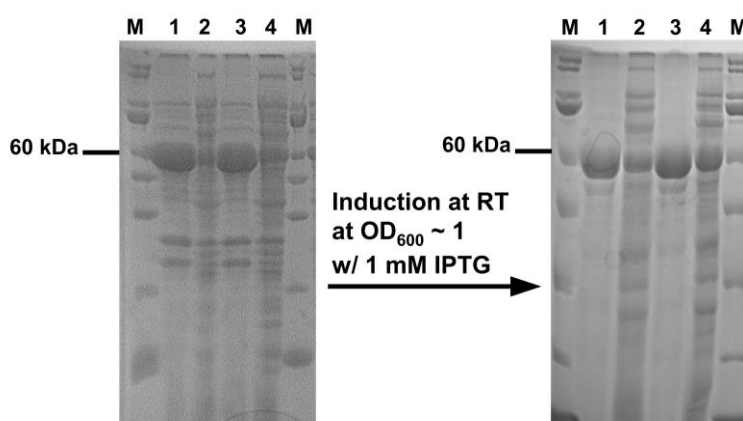
The next part of the work relates to the engineering of the FLS enzyme that catalyzes the formation of DHA out of three FALD molecules (Figure 8). To date, the high  $K_M$  and the limited maximum reaction velocity of FLS pose a challenge for the efficient methanol utilization in synthetic methylotrophy. For that reason, the FLS enzyme should be engineered towards advantageous properties for *in vivo* application. The resulting variant could be also interesting for realizing synthetic methylotrophy based on the FLS pathway in *Y. lipolytica*.



**Figure 8: Formose reaction catalyzed by the formolase (FLS) enzyme.** Formation of dihydroxyacetone via the carboligation of three formaldehyde molecules. FLS enzyme relies on cofactors magnesium and thiaminpyrophosphate (TPP).

### 3.2.1 Expression of parental FLS and activity determination

The native enzyme along with two his-tag variants (N- and C-terminal) were cloned into the pET24a expression vector at the Chair of Chemistry of Biogenic Resources, TU Munich and the successful incorporation of the gene was verified by sequencing (2.2.2.10). Expression tests using *E. coli* BL21(DE3) showed visible expression although most enzyme was located in the pellet and not in the lysate as shown by SDS-PAGE (Figure 9, left). When changing the parameters for FLS induction (switch to room temperature, induction at  $OD_{600} \sim 1$  with 1 mM IPTG) higher levels of soluble expression of the C-terminal his-tag variant were achieved (Figure 9, right).



**Figure 9: SDS-PAGE pattern of *E. coli* BL21(DE3) pET24a::FLS-C-his expression test after 16 h of induction at 18 °C, at  $OD_{600} \sim 0.6$  with 0.5 mM IPTG (left gel) and induction at room temperature at  $OD_{600} \sim 1$  with 1 mM IPTG (right gel).** M: protein standard; lane 1: FLS-containing pellet; lane 2: FLS-containing supernatant; lane 3: FLS-containing pellet; lane 4: FLS-containing supernatant. Experiment was performed in duplicates.

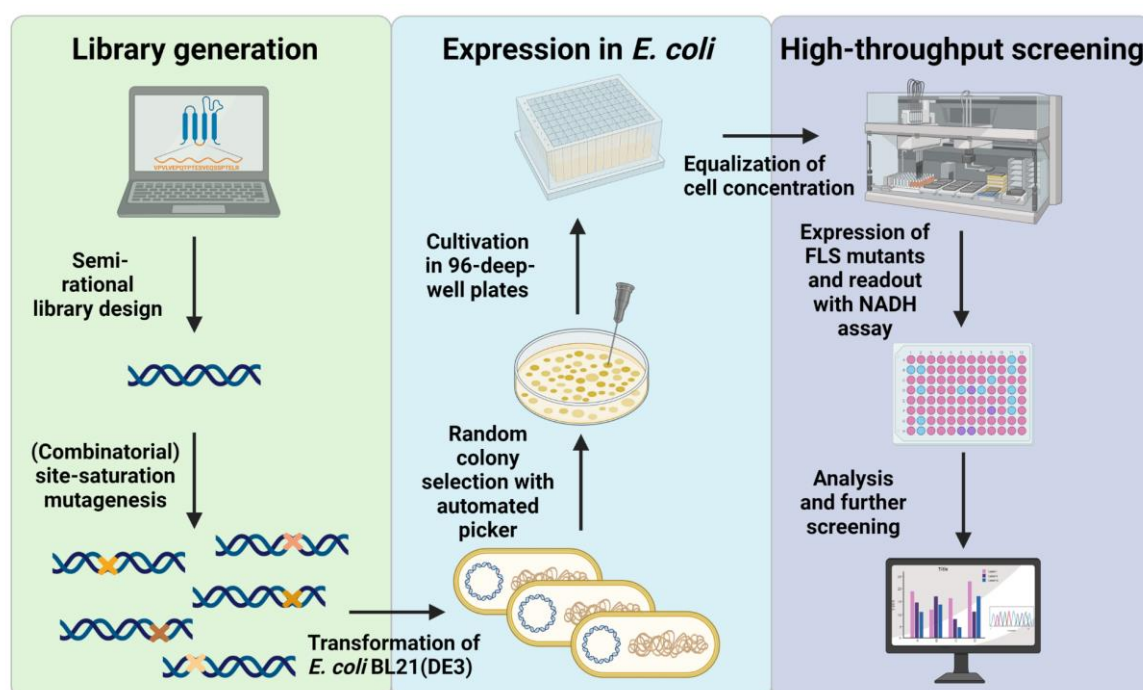
Enzyme engineering relies on mutant libraries, which are screened for variants with beneficial traits such as higher activity, improved stability, modified substrate specificity, or greater resilience to temperature or pH fluctuations. This requires a readout to evaluate the catalytic activity of variants in comparison to the parental enzyme. Usually, an assay linked to a biochemical readout, such as a change in fluorescence or absorbance through either



Furthermore, preliminary experiments showed that the empty vector lysate also was active when incubated with NADH and GDH. However, the addition of FALD (even at lower concentrations) silences this effect (Figure 11B), thus making a screening using the crude cell extract possible.

### 3.2.2 Modification of the parental FLS

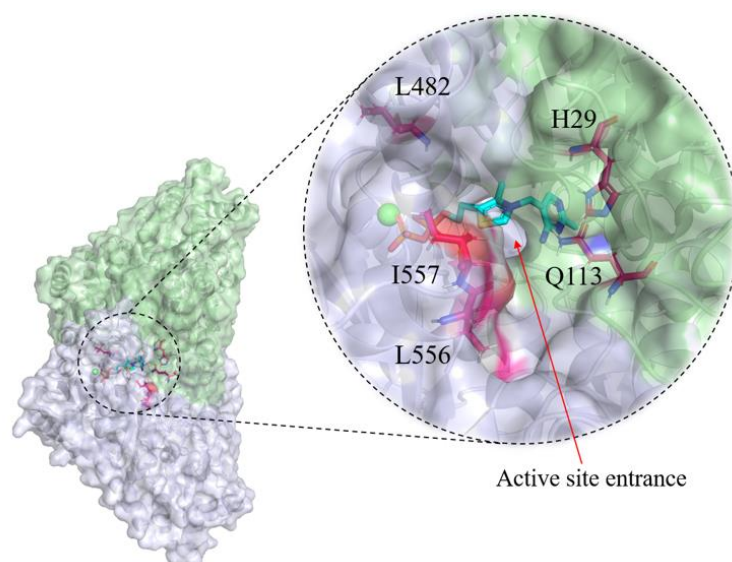
Next, semi-rational libraries were designed to be tested in a robot-assisted screening assay (Figure 12). In brief, a mutant library was used to transform *E. coli* BL21(DE3) and isolated as bacterial colonies on agar plates. To screen the performance variation of enzyme variants occurring within the established libraries, randomly selected clone colonies were picked by an automated colony picker and subsequently grown in 96-deep-well plates. The resulting preculture was used to inoculate the main culture on the next day. Protein expression was induced via the addition of IPTG at an OD<sub>600</sub> of 1. After harvesting the biomass, the cells were disrupted and crude lysate was used to assay FLS activity.



**Figure 12: Scheme of the robot-aided high-throughput screening workflow.** First, a FLS mutant library was generated through semi-rational design and site-saturation mutagenesis by using QC PCR. The obtained mutants were used to transform *E. coli* BL21(DE3) which were then cultivated in 96-deep-well plates. After equalization of cell concentration, FLS variants were induced with IPTG and the lysate used in an NADH-dependent coupled enzyme assay. Variants were ranked according their lysate activity. Figure created with BioRender.com.

Upon *in silico* analysis and docking studies using YASARA AutoDock, positions H29, Q113, L482, I556 and I557 were chosen for site-saturation mutagenesis. To identify relevant amino acid positions for site-saturation mutagenesis, the C2 atom of the thiazole ring of the

cofactor TPP was selected and all amino acid residues in the range of 12 angstrom were examined. This position was chosen as the C2 of the thiazole ring depicts the carbanion within the ylide and thus plays a pivotal role in the first step of the formose reaction<sup>[82]</sup>. In detail, the substrate FALD reversibly forms an enamine intermediate with the thiazoliumylide, which can then perform one round of carboligation with a second FALD molecule to form the C<sub>2</sub> intermediate GALD. GALD can either leave and reenter the active site, or undergo a further ligation with a third FALD molecule to form the C<sub>3</sub> product DHA.<sup>[82]</sup> Positions that were already investigated in the initial work of Siegel et al. were not considered. The remaining positions are all located in the active site pocket and/or the substrate tunnel pocket and should have an influence on the formose reaction. Figure 13 depicts the structure of the FLS enzyme harboring the cofactors (TPP and magnesium) and illustrates the positions of the amino acid residues identified as hot spots.



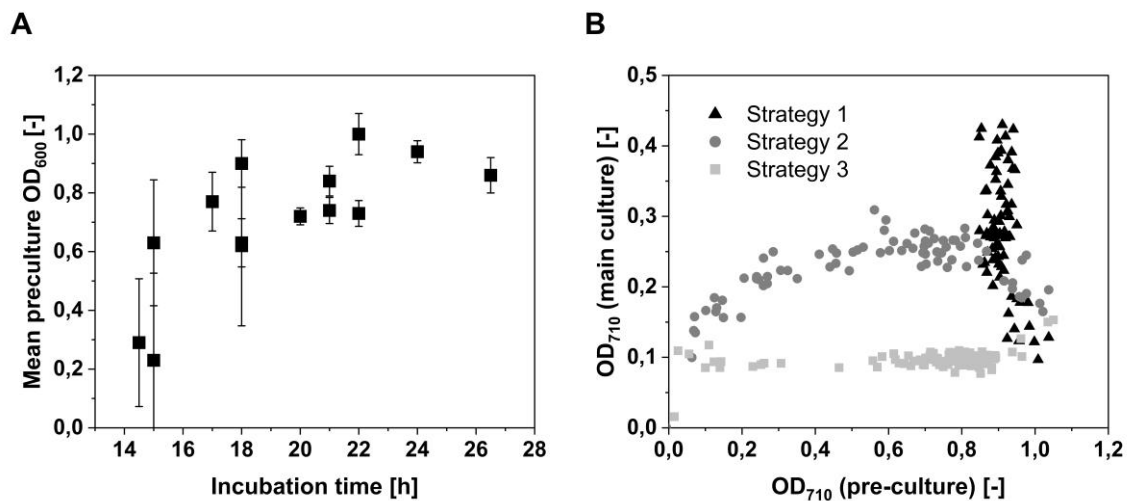
**Figure 13: Structure of FLS to visualize relevance of identified amino acid residues in close distance to substrate and cofactors TPP and Magnesium.** The active site is composed of two subunits (blue and green) while the oligomer structure is a tetramer formed from two homodimers. Magnesium is depicted in green, the structure of TPP is shown in turquoise/blue and C2 of the thiazole ring is marked yellow. Positions mutated in this study (H29, Q113, L482, I556 and I557) are highlighted. Rendered from PDB 4QQ8<sup>[78]</sup>.

### 3.2.3 Development of a robot-aided high-throughput process

To achieve a reliable screening process, it was necessary to adjust the automated inoculation procedure. When cultivated in 96-deep-well plates, the FLS wild type and variants showed different growth rates, which led to a time-dependent deviation of OD<sub>600</sub> in the exponential phase of parallel precultures (Figure 14A). This effect furthermore led to OD<sub>600</sub> deviations in the main cultures, making a simultaneous protein induction at similar OD<sub>600</sub> values impossible. Consequently, the latter issue influences heavily the formolase yield in the



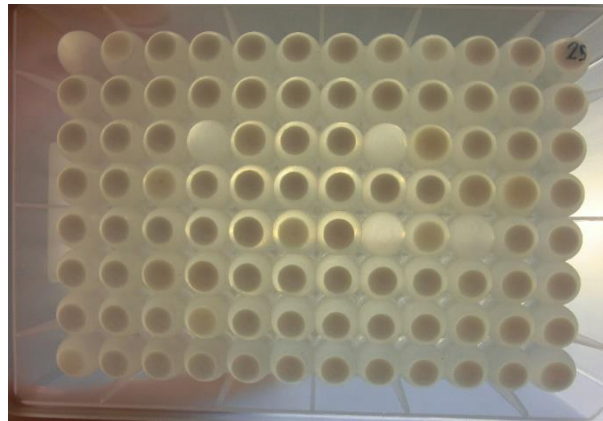
lysate. Ideally, parallel cultures would show similar growth behavior, so that all robot-aided steps (i.e., inoculation of main culture, induction of protein expression) could be carried out simultaneously.



**Figure 14: Optimization efforts towards a more reliable cultivation within the screening process.** (A) OD values of *E. coli* cultures, inoculated by automated colony picker and incubated in 1.2 mL TB at 37 °C, 1000 rpm on a shaker. 60  $\mu$ L aliquots were mixed with 270  $\mu$ L water and measured in a 96-well-microtiterplate. Each data point represents mean OD value, calculated from 85 measurements of parallel cell cultures grown on 96-well-deep-well-plates. Error bars represent standard deviation from mean OD. (B) Dependence of main culture OD (measured in early log-phase, diluted 1/10) on preculture OD (diluted 1/10) with different inoculation strategies. Strategy 1 (black triangles) depicts OD values of main cultures, inoculated with the same amount of cells from precultures grown for 20 h into stationary phase. Strategy 2 (dark grey circles) depicts OD values of a main cultures, which was inoculated with different volumes of 14.75 h precultures, depending on measured preculture OD and diluted in proportional manner. Strategy 3 (light grey squares) represents OD values of main cultures, which were inoculated from precultures grown for 14.75 h, based on the measured OD and taking into account an empirically derived pitch function that corrects the distribution of the varying precultures to achieve similar growth behavior.

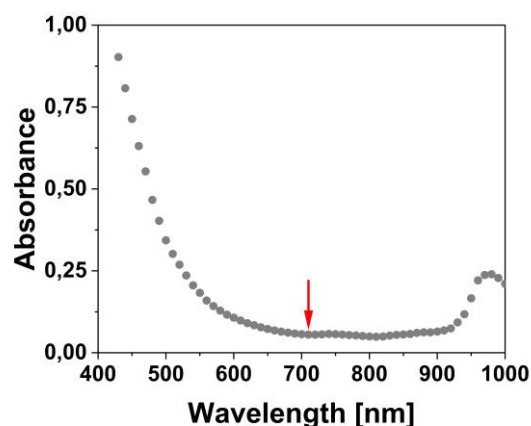
In total, three different strategies were tested to achieve similar main culture OD values and to facilitate concurrent induction (Figure 14B). In strategy 1, the preculture incubation time was extended to 20 h to ensure that all precultures reached stationary phase with nearly similar OD values (about 6 % standard deviation). However, application of this strategy resulted in high deviations (about 28 % standard deviation) of the main culture OD values. The second approach comprised a fully automatic procedure to equalize the cell concentrations in the exponential growth phase by individual dilution depending on the specific preculture OD values. First, preculture plates were incubated for 14.75 h, just before first cultures reached their stationary phase. Target OD was then divided by net OD values, resulting in individual dilution factors used for the inoculation of the main culture (the detailed procedure is summarized in Figure S2). Contrary to expectations, this proportional dilution procedure alone (e.g., a doubled OD in the preculture was compensated by a doubled dilution factor) did not lead to a satisfactory equalization of main culture OD (Figure 14B,

strategy 2), impeding a concurrent protein induction. As strategy 2 indicated that the deviation of the main culture OD depends on the different growth behavior of the precultures and shows a characteristic, recurring shift, a pitch function was introduced (2.2.3.6, Eq. 4). This function estimates how much one distribution must be shifted to match another one. The pitch function was used to correct the shift by calculating the required inoculation volume necessary for the main culture to grow evenly (Figure S3). After multiplying the target ODs in the described dilution procedure with correction factors resulting from the empirical pitch function, finally similar main culture OD values were achieved at the same time and simultaneous induction was possible (Figure 14B, strategy 3). In addition, this optimized procedure resulted in comparable cell pellet sizes, indicating comparable amounts of biomass (Figure 15).



**Figure 15: Pellets of *E. coli* BL21(DE3) indicate an even formation of biomass independent from the genetic background (e.g., FLS mutation).** Pellets were obtained prior to cell lysis using the optimized protocol including the pitch function.

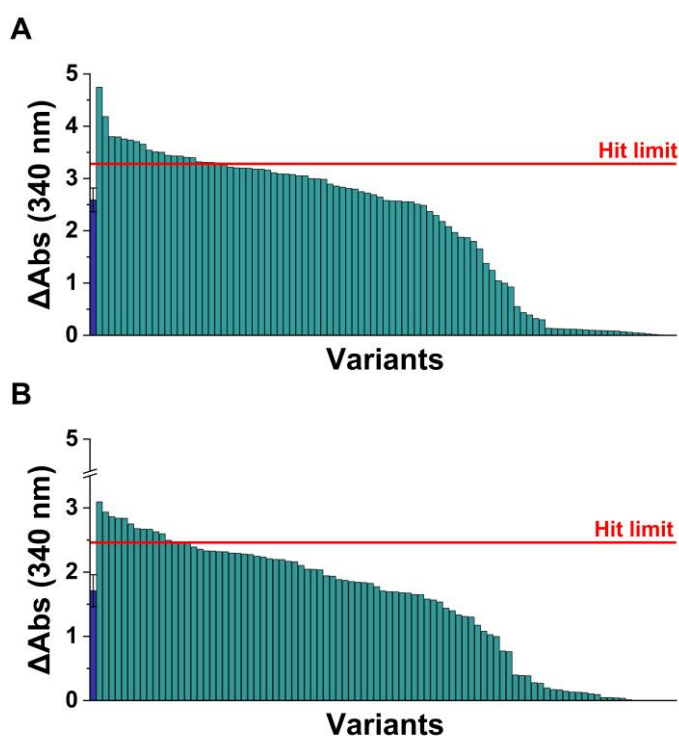
Since the performance of the equalization process depends fundamentally on the OD measurements, high precision at this step was crucial. This was achieved by increasing the filling volume and therefore the path length up to the maximum well capacity. In addition, OD was measured at 710 nm instead of the usual wavelength of 600 nm. Here, the ratio of cell culture extinction to media blank absorption was maximal (Figure 16).



**Figure 16: Absorbance spectrum for TB medium.** Red arrow indicates absorbance at 710 nm with minimal background noise. Absorption scan was performed in triplicates.

### 3.2.4 Screening of mutant libraries and variant selection

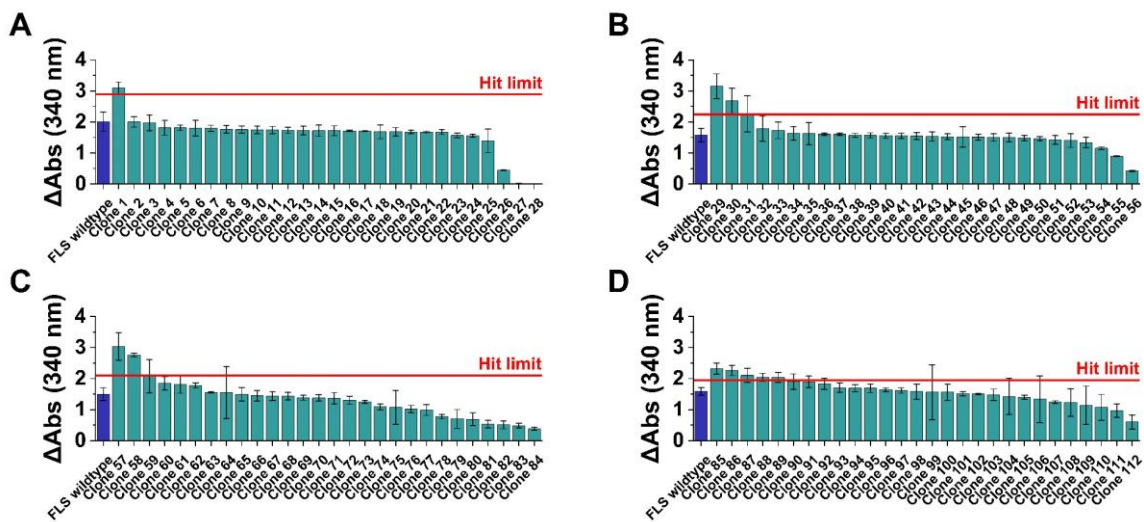
The optimized high-throughput approach including the previously described pitch function was then used for the engineering of the formolase enzyme. Figure 17 depicts exemplarily the measured activity for the mutant library in comparison to the wild type FLS. Shown is the change of absorbance ( $\Delta\text{Abs}$ ) of both wild type FLS (blue) and variants (green) using the NADH-dependent GDH assay at 340 nm. To evaluate only relevant variants, the hit limit (indicated in red) was set at  $3 \times \text{S.D.}$  above the FLS control mean.



**Figure 17: Results of the robot-aided high-throughput single-mutant (H29, Q113, L482, I556 and I557) screening of the FLS enzyme.** Activity of FLS wild type (blue bar at the left end of the

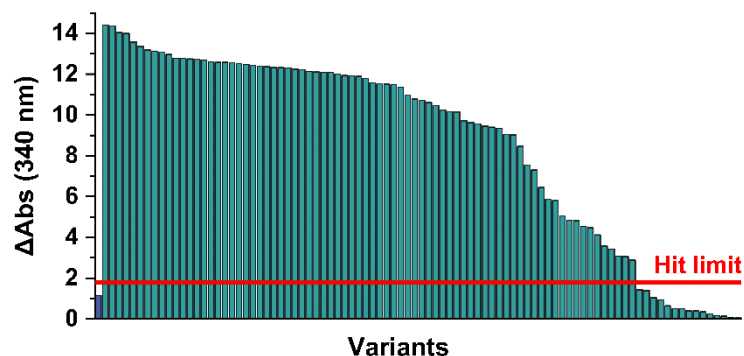
x axis) and mutant variants (green) was determined in a NADH-dependent enzyme assay. Shown are two exemplary plates (**A**, **B**). In total five plates were screened, resulting in 93 % expected coverage of mutational space per library. Hit limit (red) was set at  $3 \times$  S.D. away from the FLS wild type mean. In total five plates (one for each mutation) were screened. Over 100 clones were identified exceeding the predefined hit limit.

To verify, whether these variants contributed to higher lysate activity compared to the wild type FLS, they were subjected to a rescreen using the same method. To minimize plate effects on cell growth and protein expression, variants were expressed in triplicates and randomly distributed in the 96-deep-well plates. The change of absorbance, which correlates with lysate enzyme activity, is shown in Figure 18. Eleven variants ( $\sim 10\%$  of rescreened variants) exceeded the hit limit ( $3 \times$  S.D. above the FLS wild type mean). Sanger sequencing revealed that the most promising variants have a mutation at position L482 or I557.



**Figure 18: Robot-aided high-throughput rescreening of engineered FLS clones using the FLS enzyme assay.** Over 100 clones of single mutant libraries were subjected to a second screen (rescreen). Shown are four exemplary plates (**A-D**). Delta absorbance ( $\Delta$ Abs) of FLS wild type (blue) and variants (green) was determined in an NADH-dependent enzyme assay at 340 nm. Error bars represent standard deviation (S.D.),  $n = 3$ . Hit limit was set at  $3 \times$  S.D. from the FLS wild type controls.

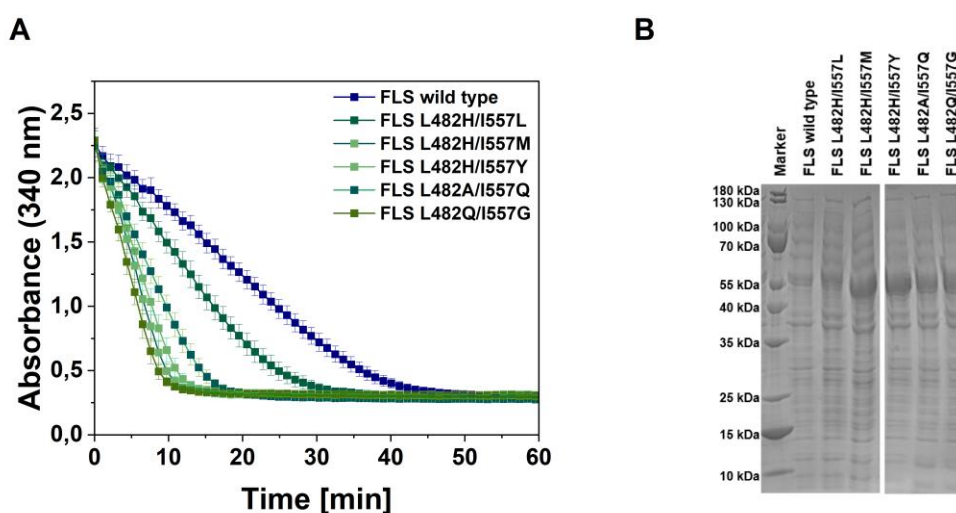
To identify whether beneficial combinatorial effects can be achieved by mutating both positions (L482 and I557) it was necessary to construct a double-mutant library. Therefore, the five most promising variants (L482H, L482A, L482R, I557G and I557T) were chosen and altered individually at the other position (I557 and L482, respectively) with the NNK motif. The resulting double mutant library was then analyzed using the robot-aided high-throughput screening platform. Figure 19 depicts the results of the double mutant screening process exemplarily for one analyzed plate.



**Figure 19: Results of the robot-aided high-throughput double-mutant (L482NNK/I557NNK) screening of the FLS enzyme.** Activity of FLS wild type (blue bar at the left end of the x axis) and double-mutant variants (green) was determined in an NADH-dependent enzyme assay. In total five plates were screened, resulting in 93 % expected coverage of protein space per library. Hit limit (red) was set at  $3 \times$  S.D. above the FLS wild type mean.

More than 80 % of the double mutant lysates performed strikingly better in the GDH assay than the wild type lysate. The combination of L482NNK with I557NNK synergistically lead to an improvement of up to 10-fold compared to the wild type. This represents a considerably higher improvement in activity than has been observed for any previous FLS single mutant (Figure 17), whose lysates showed only up to a 2-fold improvement.

Next, the lysates of the best variants were analyzed individually as triplicates using shake flasks. Figure 20A depicts the NADH-dependent coupled enzyme assay for five exemplarily variants (green) at 134 mM FALD and compares it to the FLS wild type (blue). All tested double mutant lysates are more active than the wild type control. Variant L482Q/I557G showed the biggest improvement in this setup and was chosen for further examination. In addition, SDS-PAGE analysis revealed that all double mutants showed better expression patterns in comparison to the wild type (Figure 20B).



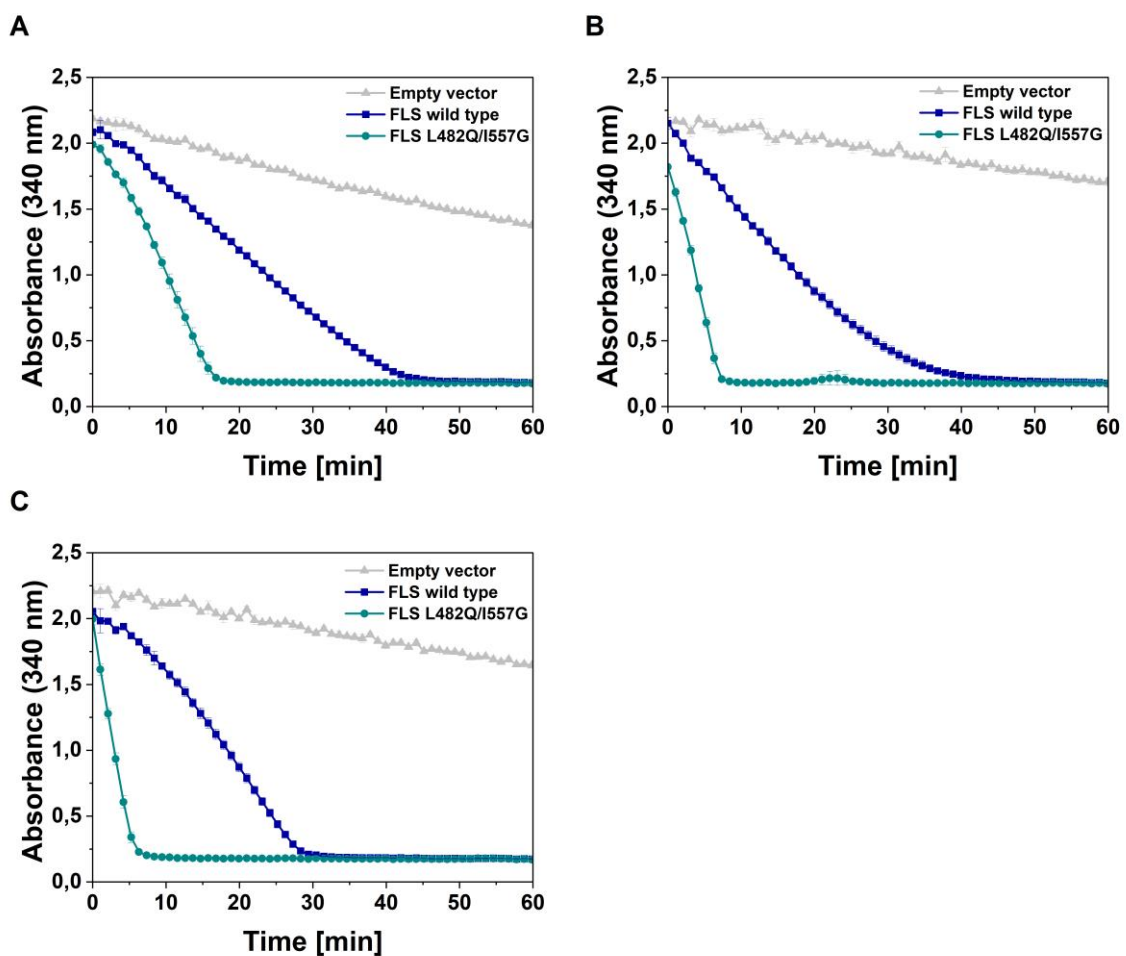
**Figure 20: Characterization of double mutant variants.** (A) NADH-dependent coupled enzyme assay at high (134 mM) formaldehyde levels for FLS wild type (blue) lysate and various double

mutant (green) lysates. Change of absorbance at 340 nm corresponds to decreasing NADH concentrations. Error bars represent standard deviation (S.D.) of three independent replicates. **(B)** SDS-PAGE of lysates containing FLS wild type or double mutant variants. M: protein standard.

### 3.2.5 Characterization of variant L482Q/I557G

To evaluate whether the variant L482Q/I557G is suitable for *in vivo* application, various parameters were determined and compared to the wild type FLS.

First, the lysate activity of the variant (green) was tested at low (13.4 mM), mid (67 mM) and high (134 mM) FALD concentrations (Figure 21A-C) and compared to the wild type (blue). The lysate of the empty vector pET24a (grey) served as a negative control.

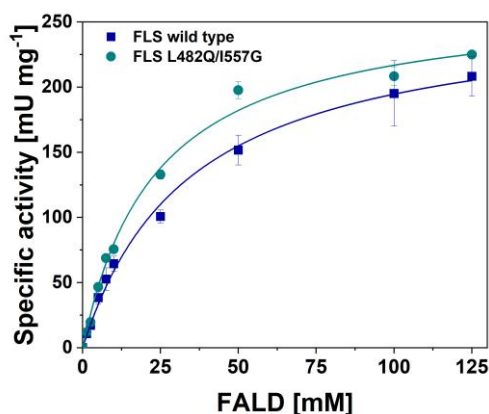


**Figure 21: Characterization of variant L482Q/I557G at low (13.4 mM, A), mid (67 mM, B) and high (134 mM, C) formaldehyde concentrations.** NADH-dependent coupled enzyme assay was performed with lysate containing parental FLS (blue), variant L482Q/I557G (green) or pET24a (empty vector) as negative control (grey). Change of absorbance at 340 nm corresponds to decreasing NADH concentrations. Experiment was performed using three biological replicates. Error bars represent standard deviation.

To ensure comparable results, total protein concentrations of the lysates were determined and diluted accordingly, so that equal amounts of total protein were used for all samples. It

was shown that the lysate of variant L482Q/I557G is not only more active at high but also at lower substrate concentrations compared to the wild type FLS. The latter property is essential for using this variant *in vivo* as desirably high selectivity towards DHA formation is only observed at high FALD concentrations.

Next, the his-tagged enzymes were purified (Figure S4) and used for activity tests to determine kinetic parameters. To ensure reproducibility, both enzymes were purified three times from biomass obtained from single individual clone colonies and the assays were performed in duplicates. A substrate range between 1.25 mM and 125 mM FALD was investigated. The  $K_M$  and  $v_{max}$  values were obtained using nonlinear regression fitting of rate of product production as a function of substrate concentration using the Michaelis–Menten equation (Figure 22). FLS variant L482Q/I557G depicts a  $K_M$  for the formose reaction of 23.5 mM, which is roughly a 30 % decrease compared to the parental FLS enzyme (Table 19).



**Figure 22: Michaelis Menten kinetic for formose reaction using parental FLS (blue) or variant L482Q/I557G (green).** Activities were assayed using the NADH-dependent GDH assay at 340 nm and 30 °C. The steady state portion of the monitored reaction was used to determine the specific activity at each substrate concentration. Each enzyme was purified in three biological replicates. Assay was performed in duplicates. Error bars represent standard deviation of six samples.

**Table 19: Kinetic characterization of wild type FLS and variant L482Q/I557G with FALD as substrate.**

Enzyme	$v_{max}$ (mU mg <sup>-1</sup> )	$K_M$ (mM)
FLS wild type	262 ± 12	34.4 ± 4.1
FLS L482Q/I557G	268 ± 6	23.5 ± 1.5

In addition to the improved kinetic parameters, the thermal stability of the enzymes is another crucial property to consider when realizing synthetic methylotrophy based on FLS. To determine their protein melting points, purified wild type or variant FLS were subjected to a thermal shift assay (Figure S5). By systematically increasing the temperature of a buffered solution and simultaneously monitoring fluorescence emission, it is possible to observe the thermally induced denaturation of the protein of interest. The change in melting temperature results in a "thermal shift" that quantifies the stabilization of the protein under different buffer or additive conditions<sup>[147]</sup>. Table 20 compares the derived melting points for wild type FLS and variant.

**Table 20: Melting point ( $T_m$  in °C) FLS wild type and variant in 100 mM phosphate buffer (pH 8.0) containing 1 mM MgSO<sub>4</sub> and 0.1 mM TPP.**

Enzyme	$T_m$ [°C]
FLS wild type	45.8 ± 0.2
FLS L482Q/I557G	50.8 ± 0.2

For the variant L482Q/I557G a 5 °C higher melting point compared to the parental FLS was measured, signifying an increase of thermal stability of this variant.

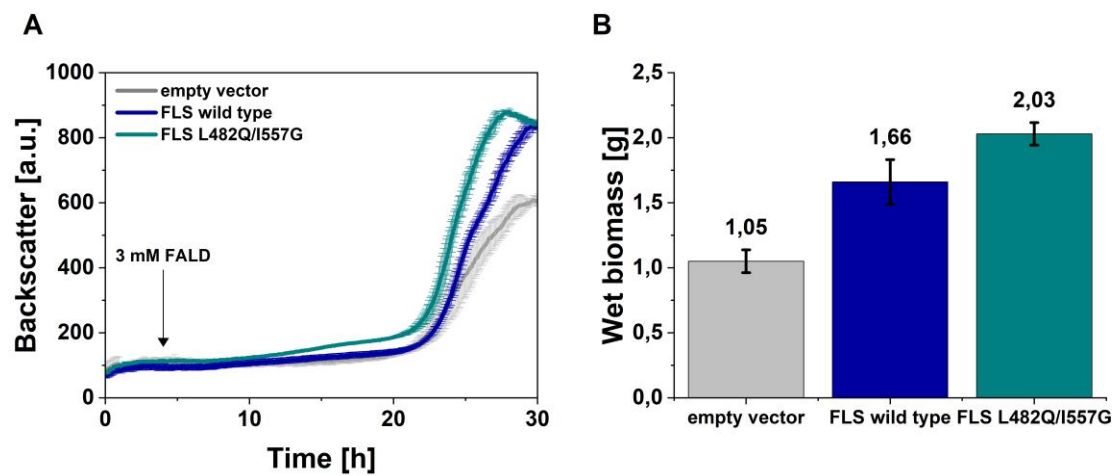
### 3.2.6 Application of variant L482Q/I557G in *E. coli* as mediator for improved FALD tolerance

One challenge to overcome when realizing synthetic methylotrophy *in vivo* is the toxicity regarding the intermediate FALD.

It is known that *E. coli* has a minimal inhibitory concentration (MIC) of FALD in the range of 5 mM<sup>[148]</sup>. A similar sensitivity was confirmed for the empty vector control when incubated in TB medium supplemented with FALD concentrations ranging from 0 – 5 mM (Figure S6). If FLS supports FALD detoxification into DHA *in vivo*, the host cells should be able to escape the toxic effects occurring at 3 mM FALD leading to a decrease of inhibition. To evaluate this predicted behavior, *E. coli* BL21(DE3) cells expressing empty vector control, FLS wild type or variant L482Q/I557G were grown in TB medium. After FLS expression for 1 h via the addition of IPTG, 3 mM FALD was added. Figure 23A shows the growth curves of wild type (blue) and variant FLS (green) at 30 °C and 1000 rpm, measured using the BioLector XT. Empty vector pET24a (grey) served as a negative control. It was observed that a prolonged lag phase until 20 h after FALD addition impedes cell growth. However, cells expressing FLS variant L482Q/I557G escaped the lag phase earlier and



reached consequently higher OD values with an increased growth rate in comparison to the wild type and empty vector control.



**Figure 23: Evaluation of FLS variant L482Q/I557G for *in vivo* application.** (A) Growth curves of *E. coli* cells expressing wild type (blue) or variant FLS (green) in 3 mM formaldehyde. Cells were induced with IPTG and grown in TB. After 4 h, 3 mM formaldehyde was added. Empty vector (grey) provides negative control without FLS gene. Each point represents the mean culture density of three independent cultures. (B) Determination of wet biomass of empty vector control (grey), FLS wild type (blue) and FLS L482Q/I557G (green) measured subsequently after growth in TB medium containing 3 mM formaldehyde for 24 h. At an OD<sub>600</sub> of 1, IPTG was used to induce FLS expression, and after 1 h, 3 mM formaldehyde was added. Experiment was performed using three biological replicates. Error bars represent standard deviation.

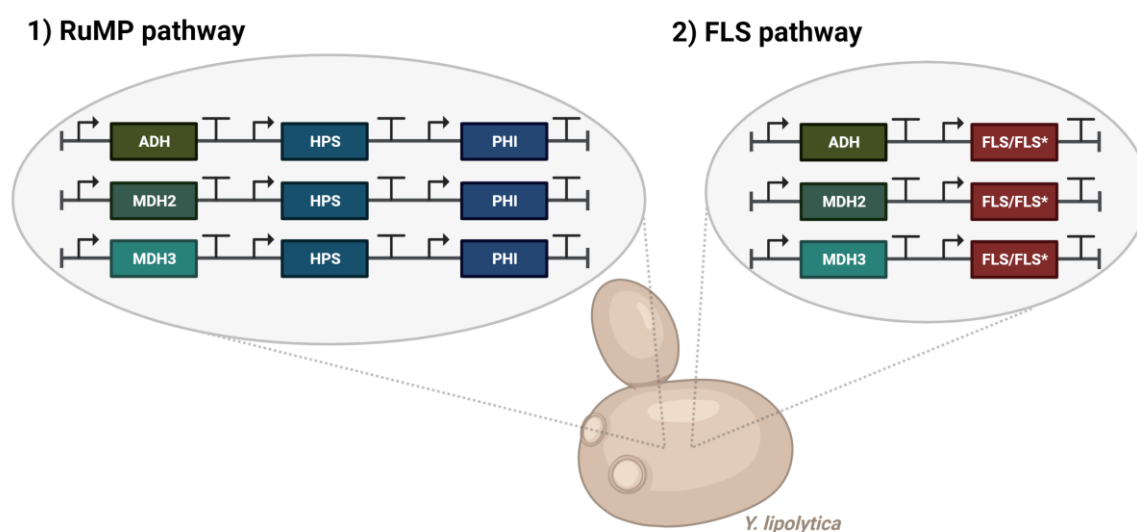
To confirm that the increased FALD tolerance is related to improved conversion of FALD to DHA, and exclude that a change in cell morphology upon the presence of stressors is responsible for the observed improved growth, wet biomass was determined in a similar setting. In detail, wet biomass concentration was determined by calculating the weight of *E. coli* BL21(DE3) expressing empty vector control, FLS wild type or variant L482Q/I557G. The cells were grown in TB medium and 1 h after induction 3 mM FALD was added. After 24 h the cells were harvested and wet biomass was weighted after deduction of the supernatant. Figure 23B depicts the measured wet biomass of all three strains upon treatment with 3 mM FALD. The variant L482Q/I557G (green) exhibited the highest formation of wet biomass compared to both wild type FLS (blue) and empty vector control (grey). Cells expressing variant L482Q/I557G form more than 20 % more biomass than cells expressing the wild type FLS. These results are in line with the observed BioLector XT cultivations. Therefore, it is concluded that the FLS variant L482Q/I557G provides improved properties for *in vivo* application, e.g., synthetic methylotrophy based on FLS (see also 4.1).

### 3.3 Engineering of *Y. lipolytica* DSM3286 for methanol assimilation

To achieve methanol utilization in *Y. lipolytica*, two steps are required. First, methanol is oxidized to the toxic intermediate FALD and second FALD is assimilated into the central carbon metabolism (Figure 2). In this work, two pathways should be tested in *Y. lipolytica*, the well-studied bacterial RuMP and the artificial FLS pathway. Both pathways share the first step, which is catalyzed by an NADH-dependent methanol dehydrogenase. Therefore, three different enzyme variants were chosen and should be tested in a combinatorial manner.

#### 3.3.1 Metabolic pathway design to enable methanol assimilation in *Y. lipolytica*

For the methanol oxidation step, MDH2 and MDH3 from *Bacillus methanolicus* MGA3 (BmMDH) were selected, based on their kinetic parameters and their performance in synthetic methylotrophic *E. coli*<sup>[149]</sup>. Furthermore, an unspecific alcohol dehydrogenase (ADH) from *Bacillus stearothermophilus* was chosen due to its low  $K_M$  for methanol<sup>[150]</sup>. For the FALD assimilation via the RuMP pathway, HPS and PHI from *B. methanolicus* (BmHPS and BmPHI) were selected providing superior kinetics and higher activities compared to homologs<sup>[150]</sup>. FALD assimilation using the FLS pathway is realized using the parental FLS or the previously mentioned variant L482Q/I557G, shown to have advantageous properties in *E. coli*. Including the mutated FLS variant, a total of nine metabolic pathways were individually introduced and tested in *Y. lipolytica* (Figure 24).



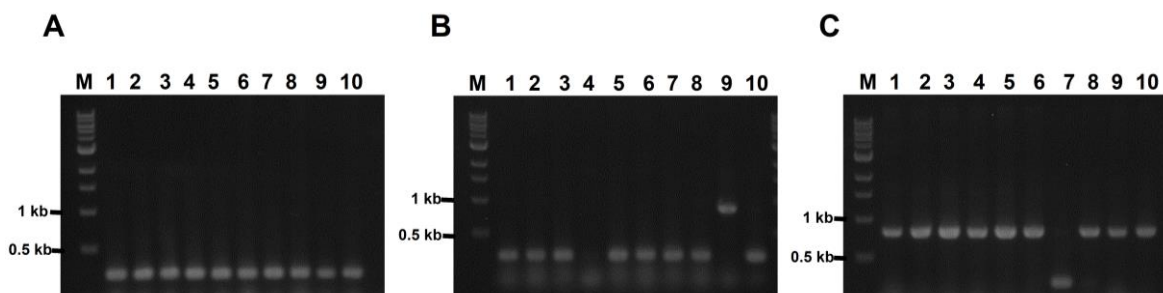
**Figure 24: Overview of genes chosen for synthetic methanol assimilation in *Y. lipolytica*.** The bacterial RuMP pathway (left side) consists of three genes, whereas the artificial FLS pathway (right side) needs the expression of two genes. The first step of both routes is the oxidation of methanol to formaldehyde. This step is carried out by either an unspecific alcohol dehydrogenase (ADH) or a

methanol dehydrogenase (MDH2 and MDH3). Formaldehyde is then assimilated using either 3-hexulose-6-phosphate synthase (HPS) and 6-phospho-3-hexuloisomerase (PHI) or the formolase (FLS) enzyme or the mutated variant FLS L482Q/I557G (FLS\*). This combination results in nine different pathways, which were engineered in a synthetic biology approach and individually introduced and tested in *Y. lipolytica*. Figure created with Biorender.com.

To facilitate rapid testing of the mentioned pathways including a combination of promoters and terminators with varying strength, an episomal Golden Gate kit was selected and used in a one-pot approach<sup>[138]</sup>.

### 3.3.2 Optimization of the Golden-Gate protocol

For this purpose, it was necessary to optimize the Golden Gate method because standard protocols for construction of the stage 1 plasmid failed (Figure 25A). In the original Golden Gate protocol 10 fmol backbone and 20 fmol insert were used and the assembly was incubated for 45 cycles of restriction and ligation (37 °C, 1 min and 16 °C, 2.5 min). The resulting assembly was then used to transform *E. coli*, isolated and digested using a restriction enzyme and loaded onto an agarose gel. As this resulted in no visible band at the expected size (~ 0.8 kbps in the case of insert PHI), different parameters of the Golden Gate assembly were altered. First, donor plasmid:insert ratio of 1:2 (molar) was used at a final concentration of 75 ng. This resulted in 10 % (1 out of 10) of the assembled plasmid being the correct size as seen in Figure 25B. To achieve a higher efficiency, needed for the later stages of the cloning, in which more fragments are ligated into one vector, a second optimization step was introduced. In detail, after the 45 cycles of restriction and ligation, a 5 min 60 °C step (final digestion) and a 5 min 80 °C step (heat inactivation) was added for background reduction. Figure 25C shows the results of the final protocol exemplarily with PHI as an insert (expected size ~ 0.8 kbps). Nine out of ten assemblies showed the correct band size, thus improving the protocol tremendously.



**Figure 25: Improvement of Golden Gate protocol.** (A) In the initial protocol, 10 fmol backbone and 20 fmol insert were used. (B) Next, a plasmid:insert ratio of 1:2 (molar) at a final concentration of 75 ng was exploited. (C) Last, a 5 min 60 °C step and a 5 min 80 °C step was introduced. For each gel (1.5 % agarose) the size of the expected products was ~ 0.8 kbps. M: DNA standard, lane 1 – 10: different clones of the Golden Gate assembly.

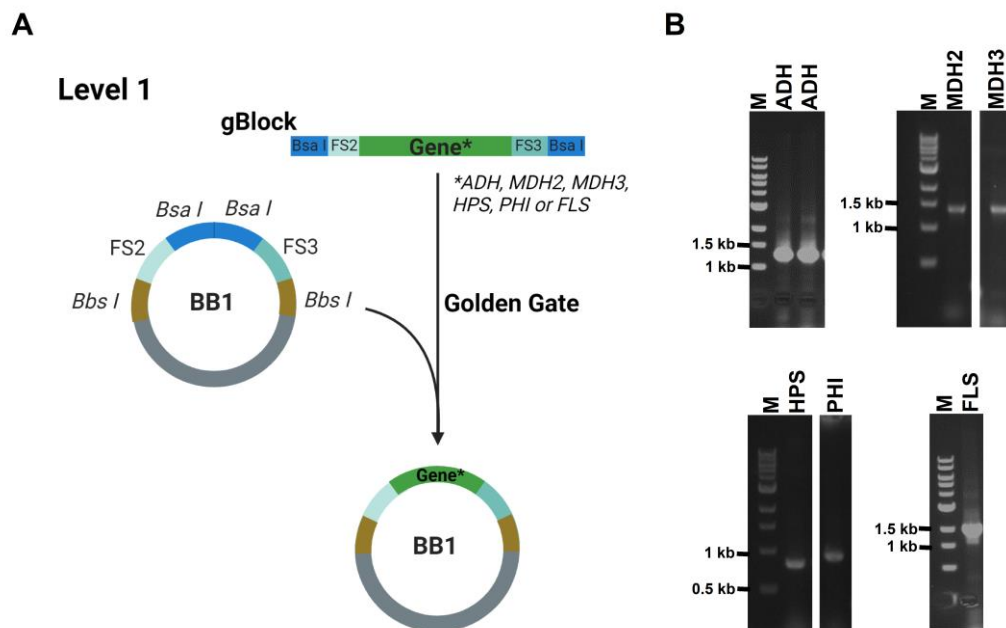
The Golden Gate kit provides extrachromosomal plasmids with two dominant markers, Nourseothricin and HygromycinB, and is therefore suitable for wild type *Y. lipolytica* strains such as DSM3286. Initially, the idea was to express the RuMP pathway on a plasmid providing Nourseothricin resistance, while expressing the FLS pathway on a backbone containing the gene for HygromycinB resistance. This would allow the expression of both plasmids in one cell and thus being convenient for testing combinatorial synergies of both pathways. However, when testing the HygromycinB resistance of *Y. lipolytica* on commonly used media, PO1f and DSM3286 were both sensitive to the antibiotic-supplemented YPD media but somewhat insensitive when streaked on HygromycinB containing K1 minimal media (Figure S7). Therefore, the genes of the FLS pathway should be expressed on both vector systems. Methanol could then serve as a second marker for growth in K1 minimal media and if this approach fails, the library based on Nourseothricin backbone serves as a backup.

### 3.3.3 Golden-Gate based one-pot approach

The Golden Gate-related approach used in this work provides a combinatorial approach to shuffle the genes of interest under control of YlpTEF- or pGPD-promoter and the transcription terminators YICyc1TT or YIMig1TT, respectively<sup>[138]</sup>. This cloning system is based on three levels depicting the individual plasmid backbone 1 to 3 (BB1–BB3). The first level (BB1) is designed for exchangeable DNA modules and every plasmid carries one of the specific functional units (promoter, gene and terminator). The second level assembles one individual gene with a promoter and a terminator on backbone 2 in a cloning step using the part plasmids from level 1. These functional units depict a defined expression cassette of the combinatorial space. In the last cloning step, the functional units are assembled into the extrachromosomal backbone BB3 harboring a resistance marker for either Nourseothricin or HygromycinB and an autonomous replication sequence (CenArs68) for the episomal expression in *Y. lipolytica*.<sup>[138]</sup>

The genes of interest (ADH, MDH2, MDH3, HPS, PHI and FLS) were codon optimized for *Y. lipolytica*, ordered as gBlocks providing the BsaI restriction sites and fusion sites 2 and 3 and cloned into BB1\_L\_23\_syn\_BsaI using BsaI resulting in BB1\_ADH, BB1\_MDH2, BB1\_MDH3, BB1\_HPS, BB1\_PHI and BB1\_FLS (Figure 26A). Intrinsic BsaI or BbsI sites were eliminated prior to gBlock synthesis by exchanging the last base pair in a codon without altering the amino acid. Chemically competent *E. coli* DH10B were transformed with 10  $\mu$ L of each Golden Gate mix and on average 70 % of clones analyzed showed the correct band

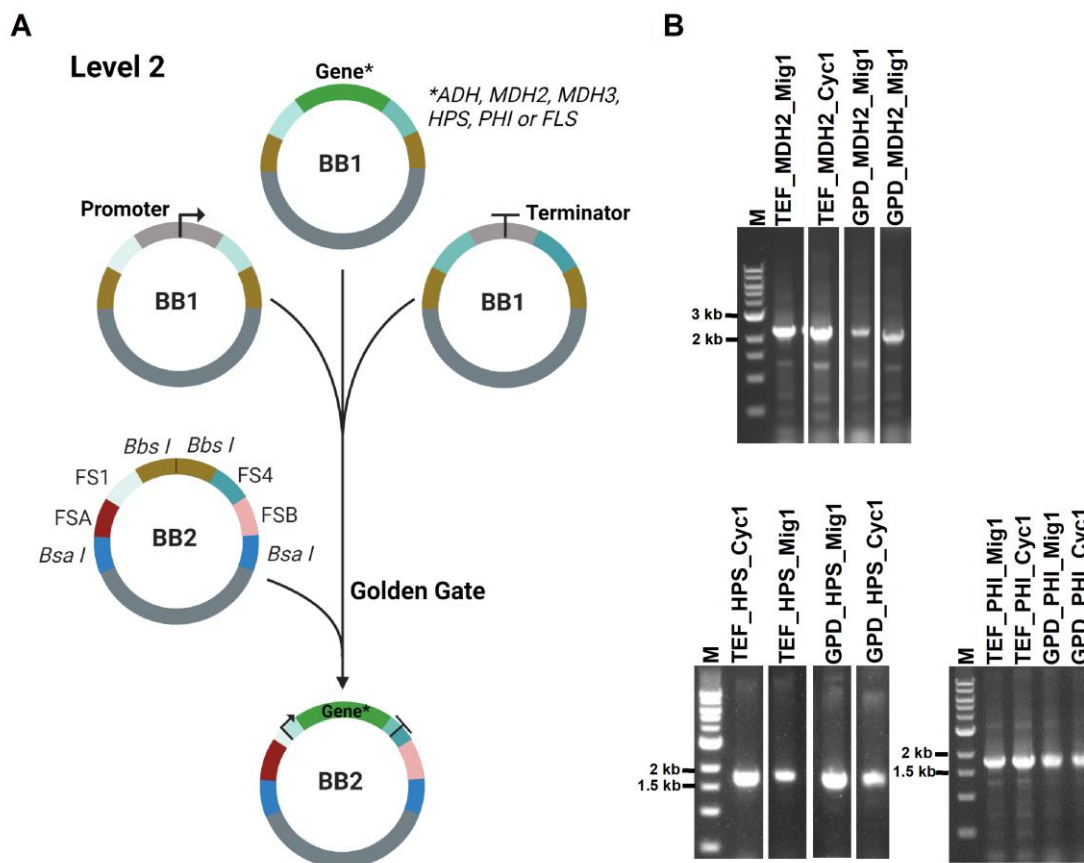
size after amplification by colony PCR. Figure 26B shows the desired size of the PCR product for each gene exemplarily. For ADH, two clones are shown, as some clones exhibited more than one band. Sanger sequencing confirmed the correct insertion.



**Figure 26: Overview and results of Golden Gate assembly level 1.** (A) The gene of interest (ADH, MDH2, MDH3, HPS, PHI or FLS) was cloned in a Golden Gate assembly as a gBlock into BB1\_L\_23\_syn\_BsaI (BB1) using restriction enzyme BsaI. Created with Biorender.com, adapted from Egermeier et al.<sup>[138]</sup>. (B) Colony PCR confirmed the correct insertion event. Shown are 1.5 % agarose gels. Size of the expected PCR products: ADH ~ 1.2 kbps; MDH2/MDH3 ~ 1.2 kbps; HPS ~ 0.7 kbps; PHI ~ 0.8 kbps; FLS ~ 1.7 kbps. M: DNA standard.

Next, each gene was assembled with a promoter and a terminator and cloned into backbone 2 (Figure 27A). The combination of six genes with one of two available promoters and terminators yields four possible expression cassette designs per gene, resulting in 24 different level 2 plasmids. Eight clones per plasmid were analyzed using colony PCR. Figure 27B shows exemplarily the four possible promoter-gene-terminator combinations for the three genes of the RuMP pathway, namely MDH2, HPS, and PHI.

To test the FLS variant L482Q/I557G also in *Y. lipolytica*, the four level 2 FLS wild type plasmids (namely BB2\_TEF\_FLS\_Mig1, BB2\_TEF\_FLS\_Cyc1, BB2\_GPD\_FLS\_Mig1, and BB2\_GPD\_FLS\_Cyc1) were each mutated using QuikChange site-directed mutagenesis (2.2.2.4).



**Figure 27: Overview and results of Golden Gate assembly level 2.** (A) The arrangement of the fusion sites FS1 to FS4 ensures the proper order of promoter, gene of interest and transcription terminator. For each gene, four promoter/terminator combinations are possible. Resulting in 24 plasmids (6 genes \* 4 combinations) for level 2. Created with Biorender.com, adapted from Egermeier et al.<sup>[138]</sup>. (B) Colony PCR confirmed the correct assembly of the expression cassette. Shown are 1 % agarose gels exemplarily for MDH2, HPS and PHI. Size of the expected PCR products: Promoter-MDH2-Terminator ~ 2.5 kbps; Promoter-HPS-Terminator ~ 1.9 kbps; Promoter-PHI-Terminator ~ 1.8 kbps. M: DNA standard.

After conforming the identity of the resulting 28 plasmids, the complete RuMP or FLS/FLS\* pathway was randomly assembled using backbone 3 (BB3) and the generated level 2 plasmids used to transform *E. coli* DHB10. In principle, the use of *E. coli* as intermediate host would not have been necessary, and one could have transformed *Y. lipolytica* with the plasmids right away, but this plasmid shuffle provides a certain security.

Next, *Y. lipolytica* DSM3286 was transformed with the obtained plasmid library and after incubation for 4 days on selective agar plates (K1 + 2 % MeOH + 400  $\mu\text{g mL}^{-1}$  NTC/ 300  $\mu\text{g mL}^{-1}$  HygroB, 30 °C) colonies were selected for further screening. The average transformation efficiency was around  $1 \cdot 10^3$  colony forming units per  $\mu\text{g}$  of plasmid DNA.

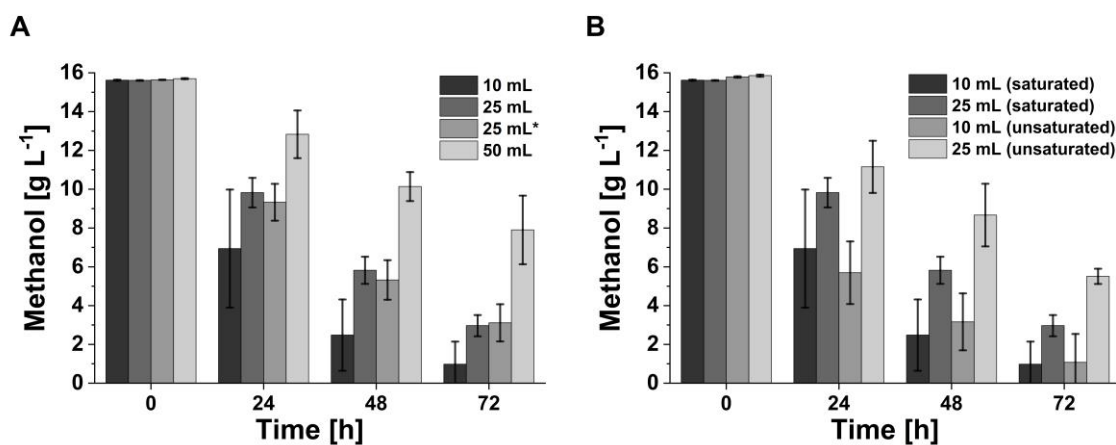
### 3.3.4 Establishing a high-throughput screening for strains with improved methanol assimilation

To find *Y. lipolytica* strains with improved growth on methanol, a screening procedure was established.

Since methanol is known to evaporate readily already at room temperature, caused by its low boiling point (64.7 °C<sup>[151]</sup>), the commonly used cultivation methods were evaluated regarding methanol evaporation. Ideally, substrate loss through evaporation is minimized to ensure effective methanol utilization when screening for variants with improved methanol assimilation properties.

To assess the amount of methanol evaporated under various conditions, K1 minimal media was supplemented with 15.8 g L<sup>-1</sup> methanol and incubated at 30 °C and 200 rpm. Subsequently, samples were taken at different time points and the amount of methanol was determined by HPLC. Figure 28 depicts the remaining amount of methanol after 24, 36 and 72 h using different shake flasks and comparing two shaking incubators. Each shake flask was filled with 10 % of its total volume. In addition to standard Erlenmeyer shake flasks, 250 mL baffled shake flasks equipped with a 0.2 µm PTFE membrane screw cap (medium grey, \*) were tested. Methanol incubated using 10 mL volume (light grey) depicted by far the highest amount of methanol evaporation. Using a volume of 50 mL (light grey) depicted a high standard deviation, but showed the least methanol evaporation overall. Using 25 mL in shake flasks (dark grey) or flasks equipped with membrane screw caps (medium grey) resulted in a steady but non-negligible methanol evaporation.

One solution against evaporation loss is to specify a humidification system for the incubator shaker. The shaker used in Figure 28A was equipped with a water bath but due to its design leads to a higher air exchange. Therefore, it was compared to a second, smaller shaker with an unsaturated atmosphere (Figure 28B). Surprisingly more methanol remained in the unsaturated atmosphere using 25 mL culture volume (light grey) compared to same volume incubated in a shaker equipped with a water bath (dark grey).



**Figure 28: Methanol evaporation in K1 minimal media at 30 °C and 200 rpm.** K1 minimal media was supplemented with 15.8 g L<sup>-1</sup> methanol and the filling volume was 10 % of total volume. **(A)** Comparison of 10 mL (black), 25 mL (dark grey), 25 mL (\*equipped with membrane screw cap, baffled, medium grey) and 50 mL culture volume in un baffled shake flasks using a shaker with a humidification system. **(B)** Comparison of a standard incubation shaker with a shaker using a water bath. Culture volumes of 10 mL (black) and 25 mL (dark grey) were incubated in a saturated environment and compared to the same flasks (10 mL volume, medium grey; 25 mL volume, light grey) in an unsaturated environment. Shown are the mean values of three independent replicates. Error bars indicate standard deviations.

Since the antibiotic Nourseothricin used in later screening experiments is relatively expensive, reducing the culture volume would also reduce costs. Ultimately, it was decided to use a filling volume of 25 mL in standard Erlenmeyer flasks and the smaller shaker for the cultivation of *Y. lipolytica* on methanol as a compromise between methanol evaporation and cultivation costs. These preliminary data underline the importance of an evaporation control and exhibit how error-prone a screening for synthetic methylotrophic *Y. lipolytica* would be.

Another option for cultivation on methanol instead of commonly used shake flasks depicts 48-well plates and the usage of a high-throughput microreactor like the BioLector XT, which enables the real-time evaluation of biomass. Here again, the amount of methanol evaporation was measured prior to the actual screening. Figure 29A depicts the amount of evaporated methanol in g L<sup>-1</sup> after 48 h of incubation at 30 °C and 1,000 rpm using a flow rate of 30 mL min<sup>-1</sup>. Methanol evaporation differs greatly depending on the position of the well. Especially wells at the right bottom (colored in red) suffer from high methanol evaporation (up to 6 g L<sup>-1</sup>). After optimizing the cultivation in the BioLector XT by reducing the flow rate to 10 mL min<sup>-1</sup> and the usage of a sealing foil with less permeability, a lower and more consistent evaporation rate throughout the plate was achieved (Figure 29B).



**A**

	1	2	3	4	5	6	7	8
A	1,16	1,64	1,66	1,98	2,33	2,38	2,46	1,60
B	1,42	1,87	2,10	2,54	2,81	3,05	3,23	2,19
C	1,66	2,20	2,56	2,43	2,79	3,30	3,93	2,23
D	1,43	2,55	2,48	3,66	3,39	4,17	5,06	2,66
E	1,67	2,41	3,52	4,03	4,83	5,13	6,39	3,37
F	1,89	2,73	3,09	2,85	3,99	4,77	4,31	2,96

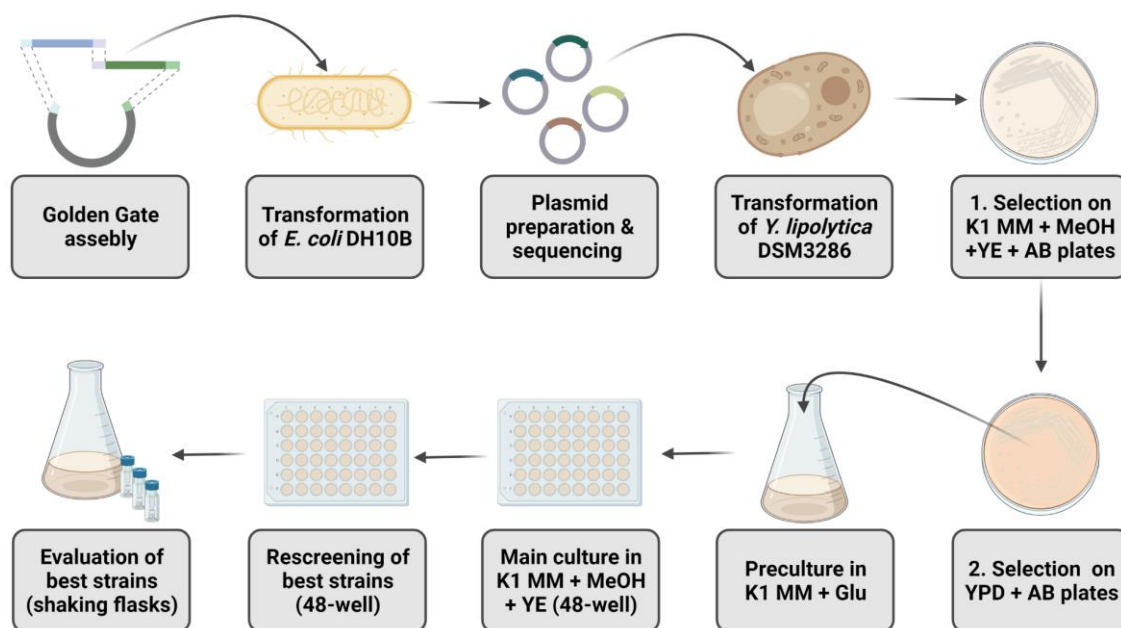
**B**

	1	2	3	4	5	6	7	8
A	1,47	0,96	0,94	1,29	1,46	1,45	1,36	1,33
B	0,97	1,34	1,36	1,31	1,33	1,54	1,69	1,45
C	0,94	0,85	1,14	1,30	1,72	1,78	1,98	1,55
D	1,09	1,78	1,36	1,64	2,05	1,98	2,20	2,37
E	1,62	1,78	1,77	1,69	2,18	2,13	2,00	1,80
F	1,56	1,89	2,03	1,98	1,90	2,17	1,75	1,77

**Figure 29: Optimization of the cultivation conditions for lower methanol evaporation using the BioLector XT.** (A) Amount of evaporated methanol in g L<sup>-1</sup> per well after 24 h using a permeable sealing foil and a flow rate of 30 mL min<sup>-1</sup>. (B) Amount of evaporated methanol in g L<sup>-1</sup> per well after 48 h using a gas-permeable sealing foil with evaporation reduction layer and a flow rate of 10 mL min<sup>-1</sup>. Both experiments were performed at 30 °C and 1000 rpm with a filling volume of 1 mL/well.

To keep screening efforts to a minimum and minimize the influence of methanol evaporation, the BioLector XT was chosen for high-throughput screening for strains with improved methanol assimilation. Figure 30 summarizes the established screening procedure used to screen for *Y. lipolytica* strains with improved methanol assimilation. After assembling the RuMP and FLS pathway on an episomal plasmid, using Golden Gate, the resulting one-pot mixture was used to transform *E. coli* DH10B. DNA preparation was performed and *Y. lipolytica* transformed with the purified plasmid library. In addition, random single *E. coli* clones were used to confirm the BB3 plasmids. *Y. lipolytica* strains expressing variants of the library were streaked on minimal media agar plates containing methanol for a first selection of strains. The corresponding empty vectors serve as a negative control. For each pathway, five clones were picked and transferred to YPD plates containing the corresponding antibiotic. This ensures that only uncontaminated single colonies are used for further characterization and prevents antibiotic resistance (see also Figure S7). Next, each clone was used to inoculate a preculture of 10 mL K1 + 15 g L<sup>-1</sup> glucose. After 36 h this preculture was washed twice with methanol containing media and used for inoculation of the 48-well flowerplates in triplicates. Cultivation on methanol was conducted using the BioLector XT and biomass formation was measured online. To only characterize significant hits, a second screen (rescreen) was performed using the same method and additionally measuring pH and O<sub>2</sub> as a second layer of valuable parameters. The best hits were then characterized in shake flasks to have a more precise determination of biomass. In addition,

HPLC samples were taken to quantify the methanol uptake of each strain compared to the empty vector controls. Furthermore, a media control without inoculation was conducted to account for the influence of the evaporated methanol.



**Figure 30: Establishing a high-throughput screening for strains with improved methanol assimilation.** First, RuMP and FLS pathway variants are assembled using the previously described Golden Gate protocol. After transformation of *E. coli* and preparation of the library, *Y. lipolytica* is transformed. In addition, random *E. coli* clones are used for sequencing the third level of the Golden Gate assembly. A first selection of possible synthetic methylotrophic strains is performed on minimal media agar plates containing  $15.8 \text{ g L}^{-1}$  methanol,  $0.5 \text{ g L}^{-1}$  yeast extract (YE) and the corresponding antibiotic. Interesting clones are then transferred onto YPD plates, which serve as a control. Single colonies are used to inoculate a preculture in glucose ( $15 \text{ g L}^{-1}$ ) containing minimal media. After 36 h, this culture is used to inoculate the main culture (K1 +  $15.8 \text{ g L}^{-1}$  methanol,  $2 \text{ g L}^{-1}$  yeast extract) and biomass formation on methanol is measured using the BioLector XT. The preculture is also used to generate backups in the form of glycerol stocks. The best strains are re-evaluated in a 48-well format in triplicates. Last, the superior strains are characterized in shake flasks and methanol consumption is measured via HPLC. Strains containing the empty vectors serve as negative control.

### 3.3.5 Screening of *Y. lipolytica* strains with improved methanol assimilation

To evaluate the strains expressing variations of the bacterial RuMP or the artificial FLS pathway, *Y. lipolytica* harboring the respective genes were cultivated in 48-well flowerplates as described in (2.2.1.5, and Figure 30). In the first screening, a total of 58 clones were screened. Since only marginal biomass formation on methanol was expected and lower OD values are associated with high background noise using the BioLector XT, the  $\text{OD}_{600}$  was also measured manually at the end of each cultivation (Figure S8). In addition,  $2 \text{ g L}^{-1}$  of yeast extract was added to the medium (instead of the usual  $0.5 \text{ g L}^{-1}$ ) to ensure that the cells form enough biomass to be able to record it with the BioLector XT. To ensure, that cells will

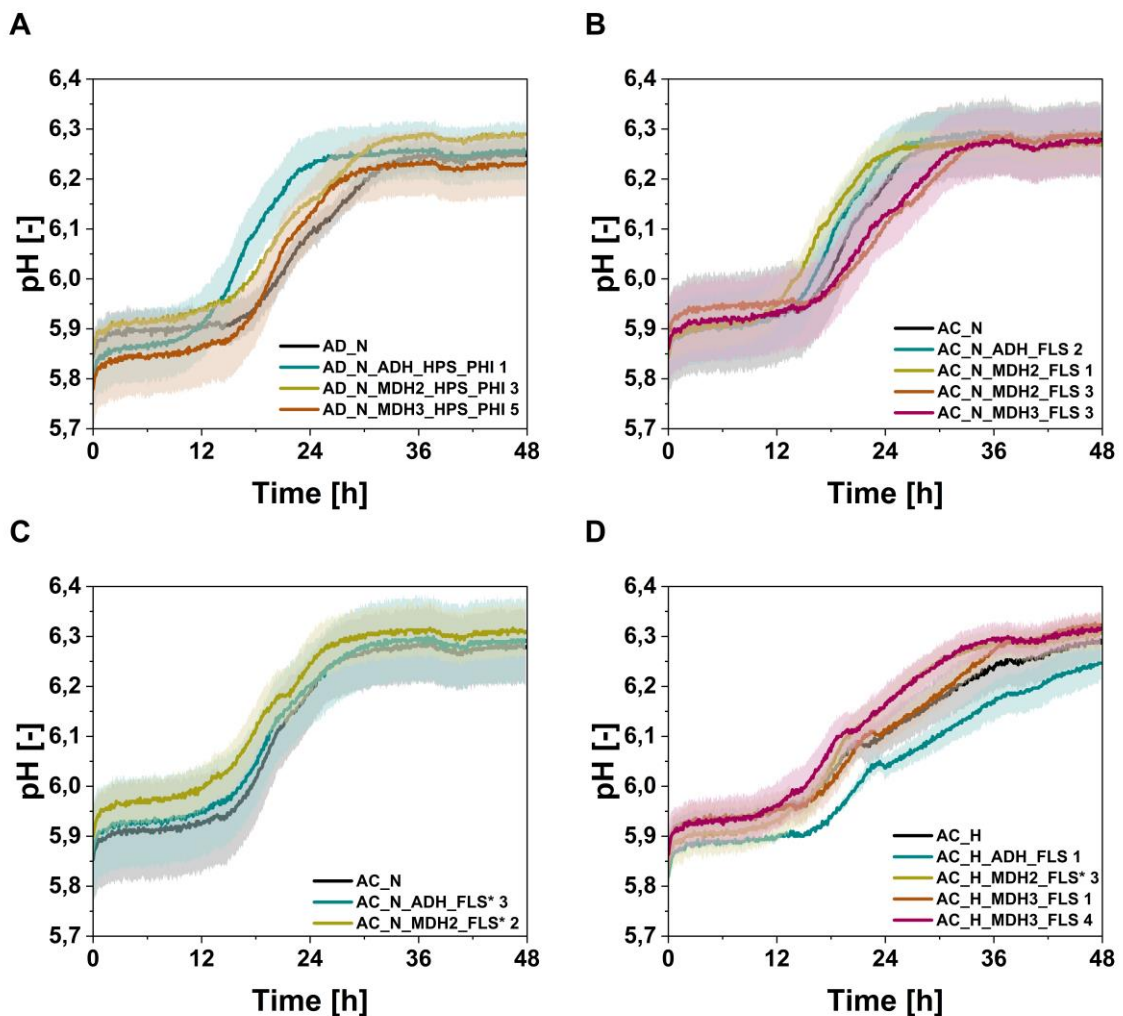
not grow on residual glucose, viable cells from precultures were washed twice in methanol containing MM prior to inoculating the main culture. The hit limit was set at  $1 \times$  S.D. above the empty vector control mean. Figure S9 depicts exemplarily various strains of the first screening rounds and compares their growth behavior to their corresponding empty vector control.

Different to expectations, the strains with the best growth behavior were not always also the strains with the highest final OD. To prevent that advantageous strains were overlooked due to a non-robust screening, strains that fulfilled both criteria, as well as only one of the two (higher growth rate or higher final OD compared to the corresponding empty vector control), were selected for rescreening. Table 21 sums up all strains selected for rescreening based on the set criteria. Strains expressing the FLS variant L482Q/I557G are abbreviated as FLS\*.

**Table 21: *Y. lipolytica* strains selected for rescreening based on their higher growth rate and/or higher final OD<sub>600</sub> compared to the corresponding empty vector.** Brackets indicate only marginal improvement compared to the empty vector.

Strain	Higher growth rate	Higher final OD <sub>600</sub>
AD_N_ADH_HPS_PHI clone 1	✓	
AD_N_MDH2_HPS_PHI clone 3		✓
AD_N_MDH3_HPS_PHI clone 5	✓	✓
AC_N_ADH_FLS clone 2	✓	
AC_N_MDH2_FLS clone 1	✓	
AC_N_MDH2_FLS clone 3		✓
AC_N_MDH3_FLS clone 2	(✓)	
AC_N_MDH3_FLS clone 3		✓
AC_N_ADH_FLS L482Q/I557G clone 1	✓	
AC_N_ADH_FLS L482Q/I557G clone 3	✓	
AC_N_MDH2_FLS L482Q/I557G clone 2	✓	
AC_H_ADH_FLS clone 1	✓	
AC_H_MDH3_FLS clone 1	✓	
AC_H_MDH3_FLS clone 4		(✓)
AC_H_MDH2_FLS L482Q/I557G clone 3		(✓)

Fulfilling these criteria, 15 strains were selected for a rescreen. In addition to biomass formation, in the rescreen pH and DO were also recorded. Especially a change in pH is relevant, as it is related to biomass formation and therefore serves as an indirect measurement for growth. This is particularly interesting because it avoids the high background noise that interferes with biomass measurement. In addition, the final OD<sub>600</sub> was determined manually. Figure 31 depicts the growth-related change in pH of the rescreened *Y. lipolytica* strains in comparison to the corresponding empty vector control.



**Figure 31: Rescreening of various recombinant *Y. lipolytica* strains on methanol.** Shown is the pH of *Y. lipolytica* DSM3286 cells expressing the RuMP pathway (A), the FLS pathway using Nourseothricin vector (B), the FLS L482Q/I557G (abbreviated FLS\*) pathway using the Nourseothricin vector (C), or both FLS pathway (wild type and variant) using the HygromycinB vector (D). Comparison of strains expressing methanol assimilation pathways (colored) to their corresponding empty vector (black). Cells were cultivated in liquid minimal medium supplemented with 15.8 g L<sup>-1</sup> methanol and 2 g L<sup>-1</sup> yeast extract. Shown are the mean values of three biological replicates. Error bars indicate standard deviations.

Cells expressing the RuMP pathway (namely AD\_N\_ADH\_HPS\_PHI, green and AD\_N\_MDH2\_HPS\_PHI, yellow) change the pH faster than the empty vector control

black), suggesting a shorter lag phase when cultivated on minimal media supplemented with methanol (Figure 31A). In addition, some strains based on the FLS pathway perform better on methanol than their corresponding empty vector control. Especially the strains AC\_N\_MDH2\_FLS 1 (Figure 31B, yellow) and AC\_N\_ADH\_FLS 2 (Figure 31B, green) as well as AC\_H\_MDH3\_FLS 4 (Figure 31D, pink) and AC\_H\_MDH2\_FLS\* 3 (Figure 31D, yellow, behind pink curve) depict interesting choices for further characterization. Again, not all observed growth curves/changes in pH correlate with the measured final OD<sub>600</sub>. Thus, none of the FLS\* strains based on the Nourseothricin backbone showed a significantly faster change in pH compared to the control, but a higher final OD<sub>600</sub> (Figure S10).

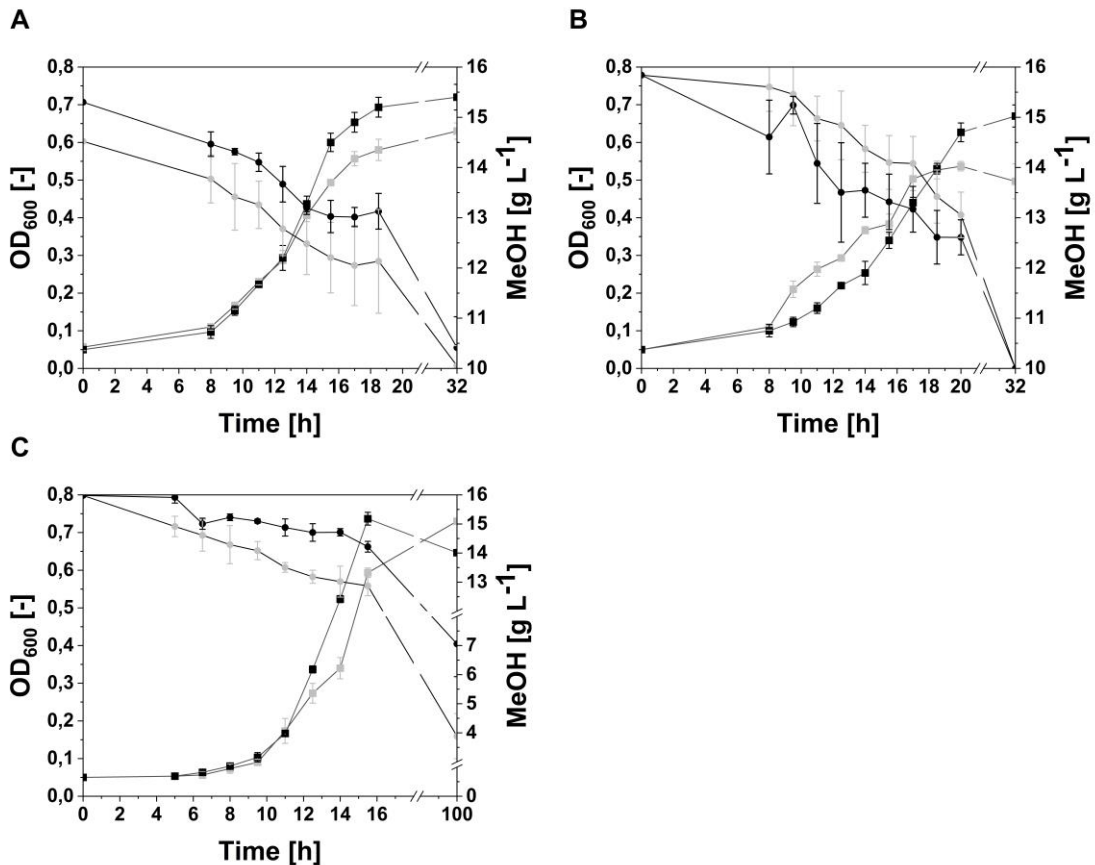
Based on their growth behavior/ change in pH and the measured final OD<sub>600</sub> (Figure S10), seven strains were chosen for further characterization.

### **3.3.6 Characterization of initial synthetic methylotrophic *Y. lipolytica* strains**

Next, methanol assimilation of engineered strains either carrying a variant of the RuMP pathway or of the FLS pathway were evaluated. Each strain was compared to the corresponding empty vector control, which further served as a calibration for methanol evaporation during cultivation. When grown in MM containing 15.8 g L<sup>-1</sup> methanol and 0.5 g L<sup>-1</sup> yeast extract, the strain AD\_N\_ADH\_HPS\_PHI1, was observed to reach the highest OD<sub>600</sub> values (Figure 32A). Furthermore, HPLC data indicated that this strain consumed the most methanol with a rate of 0.016 mg h<sup>-1</sup>. When these experiments were repeated (Figure S11A), strain AD\_N\_ADH\_HPS\_PHI1 showed again improved growth in methanol containing media compared to the empty vector control, but the calculated methanol consumption rate was considerably lower (0.004 mg h<sup>-1</sup>). It has to be noted that methanol evaporation is still an issue and rather high standard deviation for methanol consumption were observed, making a precise measurement difficult. Furthermore, the measurements are very close to the detection limit of HPLC, indicating that the strain expressing the RuMP methanol assimilation pathway might consume methanol, albeit in minimal quantities. Nevertheless, significantly higher OD values were reached after 15.5 h (unpaired two-tailed t-test, P = 0.0007), 17 h (P = 0.0124) and 18.5 h (P = 0.0109) for the cells expressing the RuMP pathway compared to the empty vector control.

For the evaluated FLS strains, only strains AC\_N\_MDH3\_FLS3 and AC\_N\_MDH2\_FLS\*2 showed slightly higher biomass formation when cultivated on methanol and yeast extract

compared to the empty vector control (Figure 32B, C). However, both strains did not seem to consume more methanol than the empty vector control. When repeating the experiment, AC\_N\_MDH3\_FLS3 showed only higher final OD<sub>600</sub> values but not a higher growth rate compared to the empty vector control (Figure S11C). Also, for AC\_N\_MDH2\_FLS\*2 an improved growth on methanol could not be confirmed (Figure S11B).

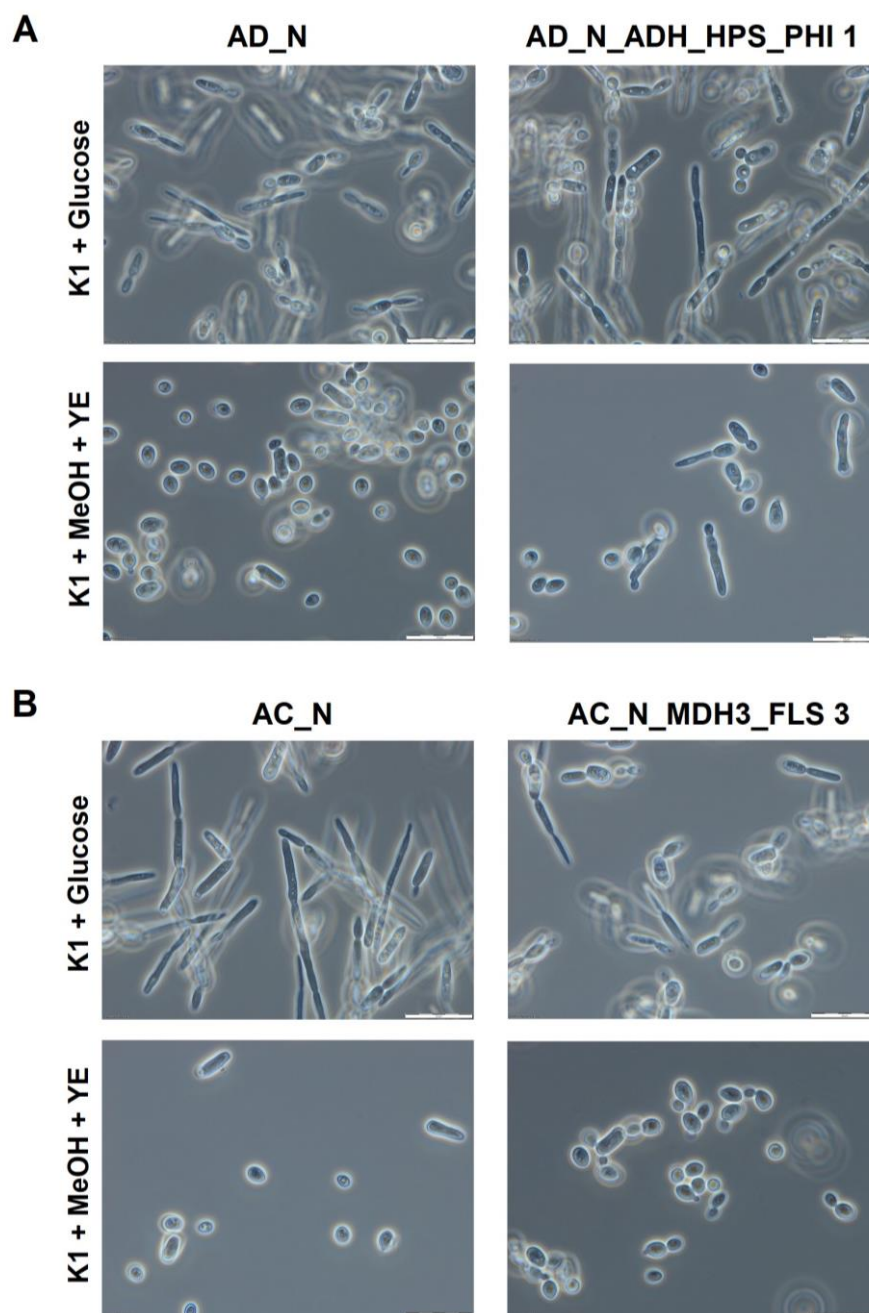


**Figure 32: Cell growth of *Y. lipolytica* strains AD\_N\_ADH\_HPS\_PHI1 (A), AC\_N\_MDH2\_FLS\*2 (B) and AC\_N\_MDH3\_FLS3 (C) in 25 mL of MM with 15.8 g L<sup>-1</sup> methanol and 0.5 g L<sup>-1</sup> yeast extract.** Engineered strains are shown in black, whereas the corresponding empty vector control (light grey) was used as a reference value for native methanol consumption and methanol evaporation during cultivation. Squares indicate OD<sub>600</sub> value, whereas circles indicate measured methanol concentration. Shown are the mean values of three independent replicates. Error bars indicate standard deviations.

The other evaluated strains, namely AC\_N\_ADH\_FLS2, AC\_N\_MDH2\_FLS1, AC\_H\_MDH2\_FLS\*3 and AC\_H\_MDH3\_FLS4, neither showed an improved growth nor a higher methanol consumption rate compared to their corresponding controls (Figure S12).

Furthermore, it was shown that cells grown on glucose (and in the absence of methanol) form pseudohyphae, while cells cultivated on methanol retained the yeast morphology. Remarkably, for strain AD\_N\_ADH\_HPS\_PHI 1 not only improved cell growth was observed, but cells also partially formed pseudohyphae when grown on methanol, whereas

the control depicted the yeast shape (Figure 33A). This behavior was observed exclusively for strain AD\_N\_ADH\_HPS\_PHI 1, but not for strain AC\_N\_MDH3\_FLS 3 (Figure 33B).



**Figure 33: Cell morphology of *Y. lipolytica* strains AD\_N\_ADH\_HPS\_PHI 1 (A) and AC\_N\_MDH3\_FLS 3 (B) compared to their corresponding empty vectors.** Cells were grown for 36 h ins K1 minimal media containing either 15 g L<sup>-1</sup> glucose or 15.8 g L<sup>-1</sup> methanol and 0.5 g L<sup>-1</sup> yeast extract. Cultivation was performed at 30 °C and 200 rpm. Images were taken at 100 × magnifications using a light microscope under phase contrast mode and dispersion oil. Bars represent 20 μm.

# 4 DISCUSSION

## 4.1 Carbon assimilation redesigned – the FLS enzyme

The installation of synthetic methylotrophy (i.e., native methylotrophic pathway modules) into non-methylotrophic host organisms, such as bacteria or yeasts, has proven to be a difficult task due to many unforeseen limitations<sup>[79, 152]</sup>. Obstacles include enzyme inefficiency, an unfavourable chemical driving force or unanticipated inherent complexity, among others<sup>[153]</sup>. A promising approach, besides the utilization of the native routes, is the establishment of novel synthetic pathways that circumvent the obstacles by using tailored enzymes and metabolic shunts. In this regard, it is recognized that linear metabolic pathways with almost no overlaps with the native central metabolism of the host are advantageous<sup>[78, 154]</sup>. Such pathways enable a more streamlined carbon entry via the connecting metabolite into the central metabolism (Figure 2). However, it is also known that overlaps of pathways lead to an increased degree of flexibility of the metabolic network<sup>[155]</sup>. In consequence, the linear FLS pathway is limited to the tight regulation of the entry metabolite FALD. Therefore, it is crucial to enable rapid and selective conversion of FALD by FLS to support the *in vivo* application of this enzyme in non-methylotrophic host organisms (i.e., *E. coli* or *Y. lipolytica*).

The parental FLS is characterized by a high  $K_M$  and limited maximum reaction rate<sup>[78]</sup>, and should therefore be improved by enzyme engineering. To advance the FALD conversion by FLS, a robot-aided high-throughput screening had to be developed. Although smaller libraries and simple screening methods can be sufficient for enzyme engineering projects<sup>[156]</sup>, there are many examples where larger libraries are required<sup>[157]</sup>. Especially for enzymes that are difficult to engineer, a systematic high-throughput screening significantly increases the probability of obtaining the desired properties and reduces the required time and costs. In order to screen larger libraries with sufficient coverage, a protein engineering platform must be exploited that allows rapid screening by using genotype-to-phenotype correlation as a readout required for directed evolution<sup>[158]</sup>.

In a first step, the screening was optimized towards robustness and reproducibility. This was achieved through developing a fully automated procedure and the application of a pitch function to achieve similar OD values at the time of protein induction in main culture



regimes. The advantages of this method are not only the even formation of biomass but also a lower error rate. Another possibility would have been to use autoinduction medium, which is commonly used for the induction of any lac promoter-driven, IPTG-inducible protein expression in recombinant *E. coli*<sup>[159]</sup>. However, it is very difficult to fine-tune the level of expression using this media. As seen in Figure 9, the concentration of IPTG, the temperature and the OD<sub>600</sub> value at which to start the expression play a major role in FLS expression. Furthermore, Sobrado et al. observed partial proteolytic processing of a C-terminal membrane anchor when using autoinduction but not with IPTG induction<sup>[160]</sup>. Therefore, no autoinduction medium was used in this case and instead an automated approach using a pitch function was developed to achieve similar protein expression of wild type and variants.

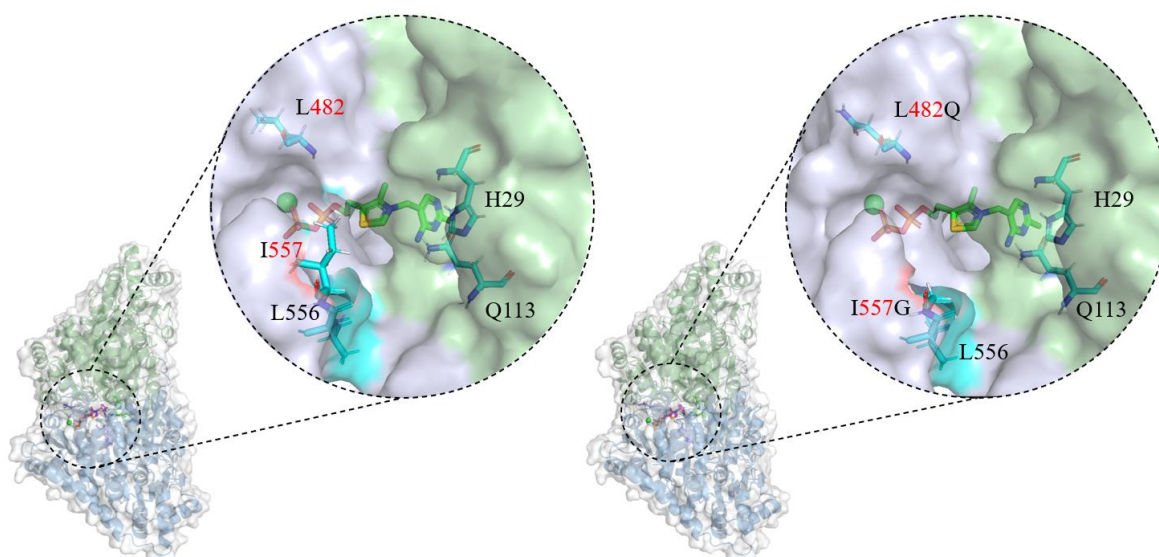
Then, the optimized high-throughput method was used to screen different semi-rational libraries with mutations at positions H29, Q113, L482, I556 and I557. After rescreening more than 100 clones, only 11 variants remained, even though a strict hit limit ( $3 \times$  S.D. above the FLS wild type mean) was applied. One reason for this could be the rather low Z-factor ( $Z = 0.4$ ), which was determined to judge whether the GDH assay is precise enough for use in a full-scale, high-throughput screen. Commonly, a Z-factor  $> 0.5$  is considered as a good assay, whereas a Z-factor  $< 0.5$  is interpreted as a marginal assay<sup>[146]</sup>. However, it has been shown that requiring  $Z > 0.5$  is likely to prevent many potentially helpful screens from being carried out at all and that assays with  $Z < 0.5$  can almost always find useful enzymes without producing too many false positives if an appropriate threshold is chosen<sup>[161]</sup>. This has also proven to be true in the case of the FLS screening.

As the best variants revealed mutations at position L482 and I557 upon sequencing, a double mutant library was constructed to identify potential beneficial combinatorial effects of mutating both positions at once. In contrast to random mutagenesis, focusing on specific amino acid positions drastically reduces library sizes, while taking evolutionary variability and mechanistic features into account when considering amino acid identity can lead to libraries with higher functional content<sup>[162]</sup>. If a library is altered at two positions (instead of only one) with the NNK motif, the mutations result in 1024 different sequences, but only 400 unique variants on the amino acid level. In order to achieve about 95 % library coverage, screening of at least 3068 ( $3 \times 1024$ ) clones would be required<sup>[163]</sup>. To keep laborious screening efforts to a minimum, the five most promising variants (three with a mutation at position L482, two with a mutation at position I557) were chosen and altered at the other position (I557 and L482, respectively) with the NNK motif. This combinatorial approach allows the screening of only one plate per double mutation (instead of the 3068 clones that

would otherwise be required) to still achieve a sufficient coverage of ~ 93 %. By combining L482NNK with I557NNK, more than 80 % of the screened double mutant lysates performed strikingly better (up to 10-fold) in the assay compared to the wild type lysate. These results demonstrate that advantageous FLS mutants obtained by saturation mutagenesis can be retained and significantly improved by recombining them with other beneficial positions in a semi-rational approach.

The importance of the amino acid replacement at position L482 and I557 is suggested by the dramatic increase in activity of the identified double mutants. This is in line with findings of Zhang et al. who also identified position L482 as crucial for enzyme activity<sup>[164]</sup>. By exchanging leucine for serine at residue 482, a significant activity increase was achieved in the conversion of acetaldehyde to acetoin via FLS. Computational analysis revealed that FLS L482S retained more hydrogen bond contacts compared to the wild type. In addition, it was shown that the mutation L482S leads to stronger contacts of active site residue W480 with the substrate acetaldehyde than the wild type.<sup>[164]</sup>

The combination of the positions L482 and I557 within the amino acid chain, the associated change in 3D structure and the specific exchange for beneficial amino acids supports increased activity. In detail, the active site of FLS is formed by two subunits (Figure 34, shown in blue and green) with position I557 being located at the entrance of the active site (red).



**Figure 34: Visualization of the structure of FLS wild type and variant L482Q/I557G.** Shown are the cofactors TPP (pink, blue) and magnesium (light green). The active site is composed of two subunits (blue and green, transparent). Position I557 (red) is located at the entrance of the active site. Rendered from PDB 4QQ8<sup>[78]</sup>.

It is known that the cofactor TPP and a magnesium ion are commonly located at the interface of the subunits of carbonylating enzymes, implying a common catalytic mechanism for this enzyme type<sup>[165]</sup>. Unlike other carbonylating enzymes such as VvSucA, which form a positively charged surface at the entrance of the active site, FLS wild type depicts a relatively hydrophobic surface, beneficial for its original substrate, the hydrophobic benzaldehyde<sup>[165b]</sup>. In regard of the fact that the initial reaction rate likely depends on the affinity between FALD and the surfaces of the active site, a change from a non-polar, hydrophobic residue (leucine) to a polar, neutral residue (glutamine) at position 482 positively influences the affinity towards the hydrophilic substrate FALD. At position 557, which is directly located at the entrance of the active site, a replacement of isoleucine to glycine seems to be beneficial for substrate (FALD) entry and product (DHA) release. Especially, branched, stiff, nonpolar side chains provide hydrophobic surfaces by their residues that limit internal flexibility and are known for their stabilizing properties. Even though both amino acids display aliphatic properties, it is speculated that the replacement from bulkier isoleucine ( $131.18 \text{ g mol}^{-1}$ ) to smaller glycine ( $75.07 \text{ g mol}^{-1}$ ) leads to a decrease in hydrophobicity and provides more space for substrate entrance (i.e., faster entrance into active center) due to a more compact side chain<sup>[166]</sup>.

Besides a lower  $K_M$ , the double mutant L482Q/I557G is expressed at higher levels compared to the parental FLS (Figure 20B). This is especially interesting as previous studies showed that the low protein expression level of FLS was the main challenge, even when a highly effective promoter was used<sup>[167]</sup>. Targeting protein expression through mutations is not an uncommon approach. Although rational strategies can sometimes lead to success<sup>[168]</sup>, high-throughput approaches are often required when it is not completely understood why a protein is not expressed well. Using combinatorial mutagenesis and selection for improved signal sequences that positively affect the expression level, mutants with up to 5.5-fold-increased  $\beta$ -lactamase expression level were isolated in *E. coli*<sup>[169]</sup>. Similarly, Sarkar et al. used directed evolution in conjunction with random mutagenesis followed by a high-throughput screening to achieve a 10-fold increase in functional expression of the neurotensin receptor in *E. coli*. In addition, a slight increase in thermostability was accomplished.<sup>[170]</sup>

In general, it is difficult to identify mutations that stabilize proteins, as substitutions often result in a less stable protein<sup>[171]</sup>. However, the FLS variant L482Q/I557G showed a  $5 \text{ }^\circ\text{C}$  higher melting point compared to the parental FLS (Table 20), signifying an increase of thermal stability of this variant. A more stable enzyme is not only helpful in terms of

expression, purification, formulation and storage<sup>[172]</sup>, but also important regarding *in vivo* application as stability *in vitro* and *in vivo* is often determined by similar principles<sup>[171]</sup>.

Taken together, the here developed FLS variant L482Q/I557G depicted not only a lower  $K_M$  and a higher level of expression, but also excels through improved thermal stability. Moreover, it is concluded that this variant provides better properties for *in vivo* application, e.g., synthetic methylotrophy based on FLS. This is evident from *in vivo* experiments in FALD-containing media, where cells expressing the double mutated enzyme, or also to a minor extent the wild type, emerged from the lag phase faster than empty vector cells and also reached higher biomass concentrations corresponding to an increased biomass yield.

Although previous studies demonstrated the usability of the linear FLS pathway in bioconversion of methanol<sup>[79]</sup>, it has to be tested whether the herein generated variants represent better choices for realizing synthetic methylotrophy in *E. coli*. While using artificial enzymes for methanol assimilation remains cumbersome, the FLS variant L482Q/I557G represents a first step towards this ambitious goal. Future work should focus on <sup>13</sup>C-metabolic flux analysis to quantify the uptake of FALD into biomass via the (mutated) FLS. The application of *in vivo* <sup>13</sup>C labeling could serve as a basis for a better understanding of synthetic methylotrophy in general and the challenges associated with the artificial FLS pathway in particular. The development of completely functional synthetic methylotrophs could effectively enable a methanol-driven production of dozens of industrially relevant products in the future.

Likewise, the recent developments in next-generation sequencing and high-throughput screening offer enormous opportunities for enzyme engineering with regard to different targets. Such approaches got more valuable and efficient in biotechnology in recent years<sup>[158]</sup>. For example, the robot-aided process described here has already been used to engineer FLS towards a broader product profile by us. First, FLS variants with enhanced two-carbon activity were screened upon saturation mutagenesis of position H29 and Q113 and a variant with high activity for GALD formation but reduced activity for DHA was found. Furthermore, a variant (H29I/Q113S) catalyzing the formation of the C<sub>4</sub> product erythrulose out of GALD was discovered. When using both variants in an enzyme cascade, the production of erythrulose from FALD was demonstrated.<sup>[143]</sup> In the future, enzyme engineering projects will not only benefit from robotic platforms but also more and more from artificial intelligence (AI). Combining directed evolution with AI and machine learning might minimize the number of trial-and-error experiments and dramatically accelerate and improve success in the search for better enzymes. Despite being a relatively new field of

study, several enzyme engineering projects have already benefited greatly from machine learning<sup>[173]</sup>, which might revolutionize enzyme engineering henceforward.

## 4.2 *Yarrowia lipolytica* – a non-conventional host

Up until recently, the field of synthetic methylotrophy has focused on using model systems such as *E. coli* and *S. cerevisiae* as non-native methylotrophic hosts. Even though in-depth research over the last decades has led to a profound understanding of their metabolism and the development of powerful genetic tools, they are not always the most favorable strain choice. The oleaginous yeast *Y. lipolytica*, on the other hand, is another promising host to install methylotrophic modules for the assimilation of C<sub>1</sub> substrates such as methanol. In general, it has many superior properties over the mentioned model organisms, including high flux through the citric acid cycle and through acyl-CoA intermediates, as well as the ability to utilize a broad range of substrates. In addition, its inherently high tolerance towards chemicals and harsh conditions provides further advantages when constructing synthetic methylotrophy based on *Y. lipolytica*. Table 22 summarizes the benign properties of this non-conventional yeast and compares it to *E. coli* and *S. cerevisiae*.

**Table 22: Comparison of the conventional host organisms *E. coli* and *S. cerevisiae* with *Y. lipolytica* for application as hosts for synthetic methylotrophy.** Partially adapted from<sup>[3]</sup>.

Properties	<i>E. coli</i>	<i>S. cerevisiae</i>	<i>Y. lipolytica</i>	Ref.
Substrates				
Conventional substrates (glucose, molasses etc.)	+	+	+	[174]
Non-conventional substrates (oil, hydrocarbons etc.)			+	[174b, 175]
Metabolism				
Aerobic	+	+	+	[175-176]
Anaerobic	+	+	-	[176]
Genetic accessibility				
HR	+	+	-	[175, 177]
CRISPR/Cas9	+		+	[128, 178]
CRISPRi	+		+	[129, 179]

(Episomal) toolkits	+	+	+	[138, 180]
Protein expression				
PTM	-	+	+	[86b]
Chaperones	+	+	+	[86b]
Titer	0	+	+	[86b]
Robustness				
Stress tolerance	-	+	+	[181]
pH tolerance	-	+	+	[175, 182]
Solvent tolerance	0	0	+	[183]
Compartmentalization	-	+	+	[184]

First, the robustness of *Y. lipolytica* towards toxic substrates as a requirement for synthetic methylotrophy was evaluated. Specifically, two *Y. lipolytica* strains were characterized for their tolerance towards methanol and the intermediates associated with its assimilation. Cells were grown in K1 medium supplemented with 15 g L<sup>-1</sup> glucose in the presence of different concentrations of methanol. Addition of 100 mM methanol reduced the growth of DSM3286 slightly. Also, higher methanol concentrations (0.5 – 1 M) were tolerated, but cell growth in two phases was observed and lead to a prolonged lag-phase (Figure 5). For the strain PO1f, a small beneficial effect was detected when supplemented with 0.1 M methanol, but in general, lower OD values were reached (Figure 6).

Vartiainen et al. also tested the impact of methanol on the growth of two *Y. lipolytica* wild type strains<sup>[140]</sup>. Strains W29 and VTT C-00365 were grown on 0.2 M glycerol in the presence of various concentrations of up to 1 M methanol. For strain VTT C-00365, the addition of 0.1 M or more methanol resulted in a notable reduction in the maximum growth rate, while growth of W29 strain was already so severely compromised under these conditions that growth rates could not be determined. When adding 1 M methanol, the growth rate of *Y. lipolytica* VTT C-00365 was reduced to 33 % of the maximal growth rate without methanol.<sup>[140]</sup>

In contrast, another study observed that the supplementation of up to 4 % methanol ( $\cong \sim 1$  M) had beneficial effects of the growth of *Y. lipolytica* strain W29 when grown on

minimal media and 20 g L<sup>-1</sup> glucose. Yet, HPLC analysis showed no change in methanol concentration for any of the conditions tested, implying that the enhanced growth was not due to native methanol assimilation.<sup>[68]</sup> This is contradictory to the results of Vartiainen et al., which suggest a native methanol capacity of *Y. lipolytica* in the form of inherent alcohol dehydrogenases<sup>[140]</sup>. However, it should be noted that both studies used different C-sources (glycerol and glucose) for their experiments, which can lead to different conclusions. It is known that *Y. lipolytica* usually prefers glycerol and that its uptake and metabolism differ considerably from that of glucose<sup>[137a, 185]</sup>. This is in line with experiments performed in this work (Figure 4, Figure S1), where it was shown that when cultivated on glycerol strains DSM3286 and PO1f both exhibited higher OD values compared to growth on glucose. It would be interesting to test whether methanol supplementation on glycerol also has a beneficial effect on the strains studied in this work. Methanol is a frequent contaminant in raw glycerol and could act as a valuable co-substrate in fermentation. In addition, the effect of methanol evaporation was observed in this work as a parameter that has to be taken into account. The latter can strongly influence the obtained results leading to misinterpretations of measured methanol concentrations.

When methanol is assimilated into biomass, the first intermediate is the toxic compound FALD. As expected, FALD was much more growth inhibiting for both examined *Y. lipolytica* strains than methanol (Figure 5, Figure 6). At concentrations exceeding 2 mM, cell growth of DSM3286 ceased. Other studies came to similar conclusions and reported that up to a concentration of 1 mM FALD was tolerated rather well, but higher concentrations were highly inhibiting<sup>[68, 140]</sup>. It was shown that *Y. lipolytica* possesses formaldehyde-degrading enzymes<sup>[68]</sup> and that the deletion of formaldehyde dehydrogenase gene *fld1* resulted in reduced growth when supplemented with FALD concentrations as low as 0.1 mM<sup>[140]</sup>. Studies on baker's yeast suggest that while membrane structure is a major target of methanol toxicity, FALD affects mostly genes related to protein synthesis<sup>[186]</sup>. Furthermore, it was shown that *S. cerevisiae* mutant strains deficient in multiple DNA repair pathways such as HR and postreplication repair were extremely sensitive to FALD, indicating that FALD toxicity is related to various forms of DNA damage in yeast<sup>[187]</sup>. Even though no such studies were conducted for *Y. lipolytica* so far, it is very likely that FALD toxicity in this yeast is based on similar principles.

For the intermediate GALD of the FLS pathway, a similar toxicity as for FALD was found for *Y. lipolytica* DSM3286 (Figure 7). At GALD concentrations exceeding 2 mM, cell growth was completely abolished. Thus, *Y. lipolytica* shows a slightly higher tolerance

towards GALD than *S. cerevisiae*, for which concentrations of 1 mM were already sufficient to inhibit fermentation<sup>[188]</sup>. Bioethanol can be produced by degradation of lignocellulose with pressurized hot water and further used for fermentation by *S. cerevisiae*<sup>[189]</sup>. However, it was shown that the by-product GALD is the key inhibitor of bioethanol fermentation. A genome-wide screen demonstrated that genes encoding alcohol dehydrogenase, methylglyoxal reductase, and the ubiquitin ligase complex play a major role in GALD sensitivity.<sup>[188]</sup> Even though the mechanism of GALD toxicity for *Y. lipolytica* remains elusive, similar patterns may be responsible.

Last, the influence of DHA on the growth of *Y. lipolytica* was investigated (Figure 5, Figure 6). Concentrations of up to 5 mM enhanced the growth of both strains, whereas 10 mM reduced the growth of DSM3286 and 1 M impaired it completely. For PO1f 10 mM DHA was already sufficient to abolish growth completely (Figure 6C). It is known that glycerol can be metabolized in two ways, one of which converts glycerol into DHA<sup>[185a]</sup>. It is therefore not surprising that at least small amounts of DHA contribute to better growth of *Y. lipolytica*.

Taken together, *Y. lipolytica* in general, and the wild type strain DSM3286 in particular are exciting hosts for synthetic methylotrophy. However, tight metabolic regulation is imperative to reduce FALD, GALD and to some extent DHA toxicity during growth on methanol. It is worth noting that not only the host but also the choice of the specific strain is important for methylotrophic engineering projects. This is suggested not only by results attained from this thesis, but also by the aforementioned examples, where different strains were examined before proceeding with the installation of methylotrophic modules in the most suitable strain.

### **4.3 Installation of methylotrophic modules in *Y. lipolytica***

To convert *Y. lipolytica* into a synthetic methylotroph, the well-studied bacterial RuMP as well as the artificial FLS pathways were chosen. Synthetic biology approaches can empower synthetic methylotrophy by pathway prototyping and a mix-and-match library design of genetic components (promoters, terminators, coding sequences, transcriptional factors, among others.) In this regard, a one-pot assembly for diverse combinations of DNA parts was chosen and used to transform *Y. lipolytica* DSM3286. The resulting library was then tested regarding growth on methanol.

For the RuMP pathway, HPS and PHI from *B. methanolicus* were chosen due to their superior kinetics (i.e., low  $K_M$ ) and combined with either ADH (from *B. stearrowthermophilus*) MDH2 or MDH3 (both from *B. methanolicus* MGA3). Testing various Mdhs *in vivo* was



necessary, as it is difficult to predict which heterologous enzymes will work best in a particular host strain. The use of a synthetic RuMP cycle to engineer methanol assimilation in industrial microbes has so far proven to be successful, particularly in bacteria<sup>[190]</sup>, but has also been shown to confer a growth advantage on methanol in baker's yeast<sup>[89a]</sup>. It offers various advantages but also challenges compared to the XuMP cycle (see also 1.3.1). Here, in total 15 clones expressing a version of the RuMP pathway were evaluated regarding their capacity to grow on methanol. In the end, only three strains were rescreened, each expressing a different Mdh. While all clones depicted better growth on methanol compared to the empty vector control, strain ADH\_HPS\_PHI 1 excelled by depicting a higher growth rate than the other strains tested (Figure S9). This was confirmed in shake flasks experiment, performed in MM with 2 % MeOH and 0.5 g L<sup>-1</sup> YE in triplicates. Twice, strain ADH\_HPS\_PHI 1 demonstrated improved growth on methanol compared to the empty vector control (Figure 32, Figure S11). Nevertheless, only a marginal methanol consumption was measured and growth solely on methanol was barely detectable.

Several hypotheses have been put forward concerning possible bottlenecks in methanol assimilation. Specifically, the concentration of Mdhs is a constraint and the inadequate kinetic and thermodynamic properties of methanol oxidation by NAD-dependent Mdhs are widely recognized<sup>[150, 191]</sup>. The Adh from *B. stearothersophilus* provides very good properties for *in vivo* applications, e.g., a lower K<sub>M</sub>. In general, NADH-dependent Mdhs show several magnitude lower affinity and reaction rates for methanol oxidation (K<sub>M</sub> ~ 170 mM, V<sub>max</sub> = 0.015 units mg<sup>-1</sup>) in comparison to PQQ-dependent MDHs<sup>[149]</sup>. This rather low substrate affinity seems to limit the methanol assimilation flux, while a high NADH/NAD<sup>+</sup> ratio negatively affects the Gibbs free energy of methanol oxidation<sup>[192]</sup>. The used Adh from *B. stearothersophilus* (in literature often named as BsMdh) shares only about 20 % amino acid identity with the Mdh from *B. methanolicus*<sup>[150]</sup> and was shown to be advantageous for use in *E. coli* compared to the Mdhs from *B. methanolicus*. While a synthetic RuMP cycle in combination with the BsMdh showed up to 39 % <sup>13</sup>C assimilation from labelled methanol in the TCA cycle, BmMdh2 demonstrated only marginal <sup>13</sup>C-labeling of any intracellular metabolites.<sup>[150]</sup> The same BsMdh was also used for methanol assimilation in baker's yeast and lead to increased growth on methanol containing solid media<sup>[89a]</sup>. This is in line with experiments performed in this work, where the most beneficial RuMP strain also carried a version of the BsMdh. While screening and selection of suitable Mdhs was a good starting point, the precise activity of the tested Mdhs in *Y. lipolytica*

remains elusive. Assaying their performance to determine their *in vivo* enzyme activities would be a reasonable next step.

So far, it is also not yet known which promoter-terminator combinations have prevailed, as plasmid purification from *Y. lipolytica* remains cumbersome. In general, isolating an episomal plasmid from yeast is a nontrivial task for several reasons<sup>[193]</sup>. First, the herein used plasmids are large (up to ~ 8 kb) and have a low copy number resulting in a low plasmid DNA yield. Since the isolation procedure breaks the yeast chromosomes, plasmid DNA is heavily contaminated by genomic DNA. In addition, yeasts are able to replicate more than one plasmid at a time and unwanted random integration of exogenous DNA into the yeast genome might occur.<sup>[193]</sup> For this reason, several plasmid isolation procedures were tested and optimized but without success. Plasmid purification from *Y. lipolytica* and the subsequent utilization of *E. coli* as a shuttle has also been unsuccessful so far. The next step would be to perform several colony PCRs to identify the promoter-terminator combinations for each gene and each strain. Additionally, sequencing of gDNA could reveal whether some strains have integrated parts of the plasmids (including the antibiotic resistance) into their genome. If successful, the prevailed promoter-terminator combinations could give valuable insights for further metabolic engineering of synthetic methylotrophic *Y. lipolytica* strains. Furthermore, the expansion of the Golden Gate library would be another conceivable approach to achieve improved methanol assimilation.

Besides the RuMP pathway, also the linear FLS pathway was expressed and assayed in a similar manner. Utilization of this artificial pathway is independent of pentose regeneration and flux changes have less influence on cellular metabolism as it does not overlap with central carbon metabolism<sup>[78, 154]</sup>. Combining both wild type FLS or variant L482Q/I557G with the mentioned Mdhs, a total of 43 clones were screened. Most tested strains showed only marginal improvement when grown on methanol compared to the empty vector control (Figure S12). When examining six strains in shake flasks, only AC\_N\_MDH3\_FLS 3 and AC\_N\_MDH2\_FLS\* 3 showed slightly higher biomass formation but no consumption of methanol was detectable. It can be assumed from the available kinetic data (see also 3.2.5), that the FLS is rate-limiting. A major challenge common to all methanol assimilation pathways is the toxicity of FALD, limiting cell growth at very low concentrations. Although the FLS variant L482Q/I557G depicts a higher affinity towards FALD compared to the parental FLS, its  $K_M$  is still above a multiple of the concentration that cells can withstand. Therefore, it is vital to keep the concentration of FALD low which is difficult to achieve and could further lead to kinetic limitations for enzymes assimilating FALD into biomass. This

is especially true for the formolase reaction, which needs not one but three FALD molecules to generate DHA<sup>[78]</sup>. In addition, FALD assimilation competes with efficient formaldehyde detoxification mechanisms. In order to achieve higher carbon assimilation through FLS, these mechanisms could be deleted. However, the deletion could in turn lead to an even higher sensitivity to the accumulation of FALD, as it has already been shown elsewhere<sup>[140]</sup>. Contrary to expectations, none of the screened strains expressing the mutated FLS variant performed strikingly better than cells expressing the non-mutated version. Even though the parental FLS was engineered towards a lower  $K_M$ , variant L482Q/I557G does not necessarily have to be advantageous also in *Y. lipolytica*. Furthermore, enzyme expression levels are very strain-specific and a better enzyme expression in *E. coli* does not automatically lead to better expression in yeasts. It would be interesting to perform the FALD toxicity experiment (Figure 23) with both parental and variant FLS also in *Y. lipolytica* to see whether similar observations are made.

Ultimately, the findings from this chapter are congruent to studies conducted by Wang et al. using the FLS enzyme in *E. coli*. It was shown that the expression of the FLS pathway did not lead to improved growth on methanol. Only several rounds of ALE finally lead to success in the form methanol assimilation into biomass.<sup>[79]</sup> In summary, there are numerous challenges to overcome in order to achieve methanol assimilation based on the formolase pathway. Converting methanol to DHA in *Y. lipolytica* would be a first step to confirm the utility of FLS and allow improvements through further engineering. Then other sections of the pathway can be added and optimized until it is fully assembled.

The field of synthetic methylotrophy in *Y. lipolytica*, similar to that in *S. cerevisiae*<sup>[194]</sup>, is still in its infancy. Besides this work, so far only three research groups have reported work on methanol assimilation in *Y. lipolytica* (Table 23) and only one work is available regarding formate assimilation<sup>[195]</sup>.

**Table 23: Summary of previous work on the development of synthetic methylotrophy in *Y. lipolytica* and comparison to this work.**

Strain	Pathway	Genes	Procedure and outcome	Ref.
VTT C-00365	RuMP w/o MDH	HPS – <i>B. methanolicus</i> PHI – <i>B. methanolicus</i> DAS1 – <i>P. pastoris</i> FLD1Δ	Identification and deletion of the FLD gene. Expression of HPS or DHAS restored the formaldehyde tolerance of the deletion strain.  Expression of HPS and PHI enzymes and cultivation in bioreactors using 0.5 M methanol and yeast extract or glycerol as a co-substrate. Formaldehyde production was	[140]

			detected but methanol conversion to biomass was not observed. Growth curves were not shown.	
ST6512	Chimeric RuMP/XuMP	MDH – <i>B. stearothersophilus</i> HPS – <i>B. methanolicus</i> PHI – <i>B. methanolicus</i> AOX1 – <i>P. pastoris</i> DAS1 – <i>P. pastoris</i> DAK2 – <i>P. pastoris</i> GLPXC – <i>B. methanolicus</i>  FLD1Δ  Overexpression of TKL1 FBA PFK RPE1	Expression of a chimeric pathway, followed by ALE using CSM and 2% methanol.  Growth in minimal media, supplemented with CSM and 2% methanol. Cultivation for 72 h with start OD <sub>600</sub> of 0.01. Increase of final OD <sub>600</sub> to 0.25 (just above 4 doublings) compared to parental strain, which could not maintain cell growth.	[68]
PO1f	XuMP	AOX1 – <i>P. pastoris</i> DAS1 – <i>P. pastoris</i> DAK2 – <i>P. pastoris</i> GLPXC – <i>P. pastoris</i> XYR – <i>M. guilliermondii</i> XDH – <i>M. guilliermondii</i> XYK – <i>M. guilliermondii</i>  Overexpression of Fba2p – <i>P. pastoris</i> Tal2p – <i>P. pastoris</i> Fbp1p – <i>P. pastoris</i> hsp70 – native	Peroxisomal expression of the XuMP pathway and construction of the xylose utilization pathway to provide the precursor of Xu5P.  Cultivation in SD medium supplemented with 1 g L <sup>-1</sup> of yeast extract, w/w/o 10 g L <sup>-1</sup> methanol led to a final OD <sub>600</sub> of 0.96 at 72 h.  Through rewiring the Xu5P recycle pathway and overexpression of hsp70, a final OD <sub>600</sub> of 10.5 and a methanol consumption of 5.2 g L <sup>-1</sup> was reached after 72 h when using 20 g L <sup>-1</sup> methanol and 20 g L <sup>-1</sup> xylose as co-substrates.	[141]
DSM3286	RuMP or FLS	ADH – <i>B. stearothersophilus</i> / MDH2 – <i>B. methanolicus</i> / MDH3 – <i>B. methanolicus</i> / HPS – <i>B. methanolicus</i> PHI – <i>B. methanolicus</i>  ADH – <i>B. stearothersophilus</i> / MDH2 – <i>B. methanolicus</i> / MDH3 – <i>B. methanolicus</i> / FLS/ FLS L482Q/I557G – synthetic	Episomal expression of combinations of the RuMP or the FLS pathway.  Growth in minimal media, supplemented with 0.5 g L <sup>-1</sup> YE and 2% methanol. Cultivation for 32 h/100 h with start OD <sub>600</sub> of 0.05. Increase to a final OD <sub>600</sub> to 0.7 (nearly 4 doublings) compared to empty vector control, which depicted lower OD <sub>600</sub> values.	This work

Independent from the used methanol assimilation pathway, only marginal biomass yields were achieved. While Wang et al. reached more than four doublings using a chimeric RuMP/XuMP, Zhang and colleagues reported a final OD<sub>600</sub> of 0.96 after 72 h using a peroxisomal XuMP cycle. As all groups used different cultivation conditions, it is hard to

rank the different strategies according to their performance. However, several factors are worth highlighting when comparing the different approaches and will therefore be discussed in more detail. It is remarkable that both strategies using either the chimeric RuMP/XuMP or solely the XuMP pathway depicted the highest OD<sub>600</sub> value after 72 h.<sup>[68, 141]</sup> This is in contrast to growth curves recorded in this work where the strains depending on the pathway and the particular cultivation method reached stationary phase after 20 to 36 h. As this was also seen with the empty vector controls, one possible explanation is the used host DSM3286 whose genotype and growth behavior differs considerably from the strains used in other publications. This is consistent with initial experiments with strains DSM3286 and PO1f, where the latter was shown to have a higher sensitivity to methanol and associated by-products in minimal medium (Figure 6). Hence, it would be valuable to establish the XuMP pathway also in DSM3286 to determine whether this would result in a shorter lag-phase/higher growth rate.

Wang and colleagues also reported an effect of methanol on cell morphology<sup>[68]</sup>. While the cells grown in the absence of methanol maintained pseudohyphae, cells subjected to methanol depicted the typical yeast morphology<sup>[68]</sup>. Similar observations were made in this work for both empty vector controls and strains expressing a version of the FLS pathway. Remarkably, the ADH\_HPS\_PHI 1 strain partially also showed pseudohyphae formation when grown on methanol and yeast extract (Figure 33). In general, *Y. lipolytica* undergoes a transition from yeast-to-hyphal growth in response to environmental conditions, such as, pH, temperature, oxygenation, and carbon and nitrogen concentration<sup>[92, 102]</sup>. Although the precise mechanisms of dimorphic transition in *Y. lipolytica* remain elusive, it is suggested that changes to hyphal growth are related to stress response through multiple kinase pathways<sup>[196]</sup>. It is speculated that the improved growth of strain ADH\_HPS\_PHI 1 on methanol leads to a so far unknown cellular stress response, which could result in the observed morphological changes. While *Y. lipolytica* is a promising industrial workhorse, in many industrial processes' hyphae formation is rather obstructive. It was shown that the deletion of MHY1 abolished hyphae formation with only minor negative side effects and a weak positive effect on lipid accumulation<sup>[197]</sup>. Having the overall goal in mind to produce lipids from renewable methanol via synthetic methylotrophic *Y. lipolytica* in the future, MHY1 is an initial promising target to address improved lipid accumulation while simultaneously reduce hyphae formation.

## 4.4 Synthetic methanol auxotrophy and potential solutions

Overall, only a handful of strategies have been reported so far on the topic of synthetic methylotrophy in *Y. lipolytica*. Similar to *S. cerevisiae*, early work in this oleaginous yeast demonstrated that the expression of a synthetic methanol assimilation cycle alone is not sufficient to enable robust cell growth using methanol as the sole carbon and energy source and requires co-substrates, such as yeast extract or complete supplement mixtures (CSM).

Recent findings suggest that the link to the “pentose phosphate pathway is essential for synthetic methylotrophy on the one hand for pentose regeneration of RuMP and XuMP pathway and on the other hand for the synthesis of complex biomass precursors or vitamins and cofactors<sup>[18]</sup>.”<sup>[1]</sup> “Beside the oxidative pentose phosphate pathway branch, yielding mainly NADPH for assimilatory reactions, the non-oxidative pentose phosphate pathway branch has also a central role in methanol assimilation during methylotrophy<sup>[18]</sup>. In particular, the pentose rearrangement reactions are of importance to regenerate the metabolites Xu5P and Ru5P<sup>[198]</sup>. Specifically, both molecules are used as acceptors for formaldehyde assimilation in XuMP and RuMP cycle, respectively<sup>[199]</sup>. Therefore, the depletion of these metabolites has vast influence on the pathway efficiencies. In consequence, constant replenishment of the pentose pools has to be ensured by appropriate metabolic flux distribution to drive the individual assimilation cycles of XuMP and RuMP.”<sup>[1]</sup>

“In addition to pentose regeneration, the importance of the non-oxidative pentose phosphate branch is even deeper intertwined with the metabolism. It was shown that methanol utilization in *P. pastoris* is associated with the overproduction of vitamins and cofactors<sup>[18]</sup>. These molecules are required for the recruited enzymes. Here, the synthesis of flavin adenine mononucleotide or riboflavin are examples, which require pentoses as precursor metabolites<sup>[18]</sup>. Such findings can explain why almost all synthetic methylotrophs require small amounts of complex media components for growth on sole methanol as the carbon source. In this regard the specific role of yeast extract and associated compounds that stimulate cell growth during synthetic methylotrophy is not fully known to date. Yeast extract is a complex hydrolysate of yeast biomass, which provides carbon, sulfur, trace nutrients, vitamin B complex and other important growth factors<sup>[200]</sup>.”<sup>[1]</sup>

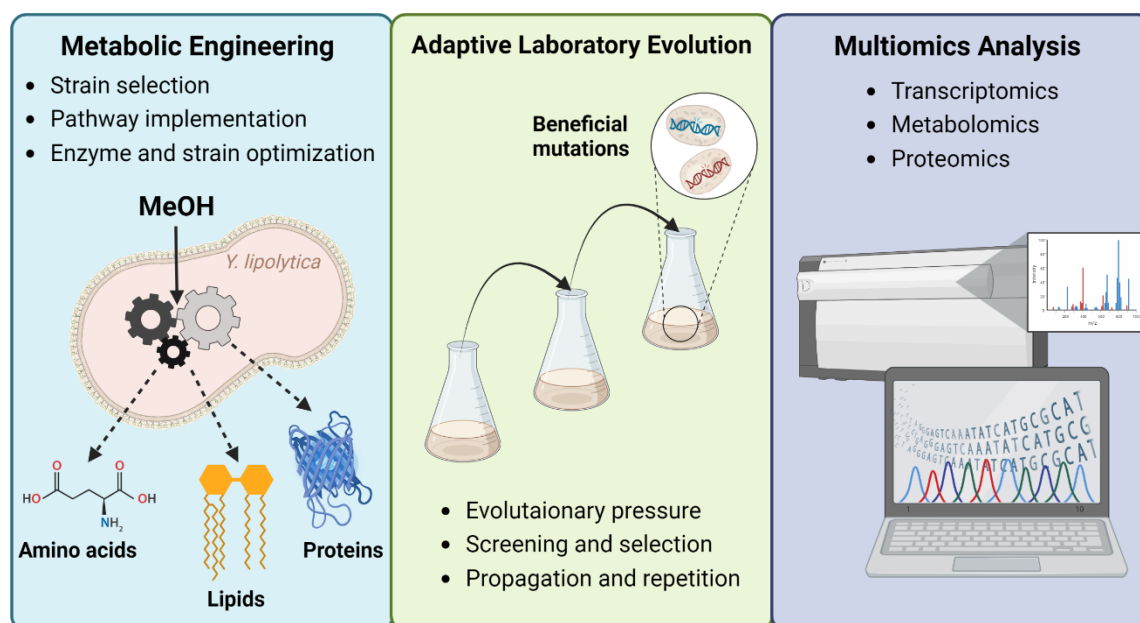
Despite the successful development of partial methylotrophy in *Pseudomonas putida*<sup>[201]</sup>, *C. glutamicum*<sup>[83, 202]</sup>, or *S. cerevisiae*<sup>[89a, 89b]</sup>, no true synthetic methylotrophy solely on methanol was achieved in any of these studies. A milestone was achieved using *E. coli*<sup>[16]</sup>.

<sup>203]</sup> that could also serve as a blueprint for synthetic methylotrophy in yeast. “To release the growth-dependence of yeast extract, a synthetic methylotrophic *E. coli* strain was optimized in the absence of yeast extract in a laboratory evolution approach<sup>[203]</sup>. Initial depletion of yeast extract led to reduced growth. Strikingly, after nine passages, an increased optical density was reached. Interestingly, this biomass formation outcompeted even the unevolved strain using yeast extract supplementation. To understand the underlying principle mechanism, genome sequencing of the evolved strains resolved associated mutations in genes encoding glutathione-dependent formaldehyde oxidation (*frmA*), NAD(H) homeostasis/ biosynthesis (*nadR*), phosphopentomutase (*deoB*), and gluconate metabolism (*gntR*)<sup>[203]</sup>. The identified mutations in *deoB* induced a genetic loss of function. This is remarkable since the associated enzyme represents a branch point in the RuMP cycle. It catalyzes the transfer of a phosphate group between the C<sub>1</sub> and C<sub>5</sub> carbon atoms of ribose or deoxyribose, respectively<sup>[204]</sup>. Even though the yeast extract supplementation is not yet fully understood, these data indicate again influence on the pentose phosphate pathway level. Due to the complex interconnection of the pentose regeneration and the synthesis of vitamins and cofactors (or even other unknown aspects) a combination of metabolic engineering and systems-level analysis can support successful installation of synthetic methylotrophy in yeast.”<sup>[1]</sup>

Another strategy couples growth on xylose with methanol assimilation<sup>[16]</sup>. First, the pentose phosphate pathway was disrupted through deletion of *RPIA/B* and installing methanol-utilizing genes (*MEDH*, *HPS*, and *PHI*) resulting in cell growth dependent on xylose and methanol<sup>[16]</sup>. In later work, *PFKA* was deleted and *GAPA* was replaced with *GAPC*, followed by ALE for 20 generations<sup>[152]</sup>. Next, *RPIA/B* was reintroduced for improved fitness of *E. coli* BL21 and cultivations were performed in minimal media supplemented with methanol and High-Def Azure (HDA) essential amino acid mixture without any other carbon sources. Gradual deprivation of HDA over 180 days resulted in a strain that could use methanol as its sole feedstock. After further laboratory evolution, the best isolate achieved growth on 400 mM methanol and reached a final OD<sub>600</sub> of 1.0 in 30 h with a doubling time of 8 h.<sup>[152]</sup> In addition, synthetic methanol auxotrophy was recently also generated in *B. subtilis* by the same *RPIA/B* knockouts<sup>[205]</sup>, suggesting a good starting point for engineering projects using the RuMP pathway. Even though similar strategies for yeast are under evaluation, so far, no growth solely on methanol has been reported.

Taken together, the “most recent findings suggest, that at this stage, ALE plays a more important role than rational metabolic engineering in constructing synthetic methylotrophy.

By combining both strategies, exciting advances for using C<sub>1</sub> compounds as a feedstock for synthetic methylotrophic eukaryotes can be reached (Figure 35). In conjunction with next-generation sequencing and omics-technologies, ALE can reveal relationships between genotypes and phenotypes, as well as the molecular mechanisms underlying the desired complex phenotypes. However, in order for ALE to be successful in generating strains with improved C<sub>1</sub>- utilization, the substrate-of-interest should be coupled with cellular growth or survival.”<sup>[1]</sup>



**Figure 35: The building blocks for establishing synthetic methylotrophy in yeasts.** The combination of metabolic engineering and adaptive laboratory evolution leads to exciting advances in the utilization of C<sub>1</sub> compounds (e.g., methanol) for the sustainable production of industrially relevant chemicals. Evolved strains are characterized for fitness improvements and integrated multiomics analysis helps to identify the most important changes on systems level, thus providing a deeper understanding of methylotrophy in general. Figure adapted from<sup>[1]</sup> and created with Biorender.com.

Also in this work, the improvement of the best strains by means of ALE was considered. However, due to problems with sterility, the experiment was cancelled at an early stage (data not shown). One obstacle was that the strains used for ALE carried the methanol assimilation pathway on an episomal plasmid instead of genomic integration. It cannot be ruled out that the cells had already lost the plasmid after a few generations and thus lost resistance towards the used antibiotics, allowing other, resistant microorganisms to take over. The use of a plasmid in ALE is also disadvantageous in terms of mutations preferentially occurring on the plasmid (e.g., higher copy number, stronger promoter) instead of resulting in the desired genomic mutations. Therefore, the methylotrophic modules should first be integrated into the genome of *Y. lipolytica*, preferably coupled to survival, and then ALE should be carried



out in a continuous cultivation. Using a chemostat for cultivation can be beneficial when nutritional components are limited<sup>[206]</sup>. In a continuous approach, cell growth and environmental conditions are kept constant, making it easier to select cells with higher growth rates. Altogether, a “combination of advanced metabolic engineering, *in silico* modeling, and automation to maximize evolutionary efficiency should be considered. Finally, the subsequent omics-analysis of the evolved strains can lead to new insights into the mode of action and further genetic targets to improve efficiency of synthetic methylotrophy in eukaryotes even more<sup>[207]</sup>.”<sup>[1]</sup>

## 4.5 Conclusion

In conclusion, several strategies for the installation of methylotrophy in *Y. lipolytica*, or in yeast in general, exist, but many obstacles remain. In this work, the bacterial RuMP pathway and the artificial FLS pathway were introduced into *Y. lipolytica* using a synthetic biology approach. Prior, the parental FLS was engineered towards better properties for *in vivo* application. For that, *E. coli* was chosen as the host organism because it is easy to manipulate and exhibits rapid growth, which is advantageous for high-throughput screening methods. In a semi-rational combinatorial approach, FLS was engineered towards a lower  $K_M$  and higher expression levels in *E. coli*. In addition, variant L482Q/I557G depicted improved thermal stability and *E. coli* cells expressing the double mutated enzyme emerged from the lag phase faster and reached higher biomass concentrations when grown in FALD containing media.

Next, variants of the RuMP cycle and the linear FLS pathway (including the L482Q/I557G mutant) were expressed in *Y. lipolytica*. While several strains depicted improved growth on methanol compared to the control, there is still a long way to achieve true synthetic methylotrophy solely on methanol. In particular, dependencies on supplements such as yeast extract and the associated pentose regeneration need to be understood in more depth to find solutions for methylotrophic auxotrophy. Since the genetically engineered *Y. lipolytica* strains presented in this work are capable of growing on methanol when supplemented with yeast extract, it is hypothesized that genomic integration of the methylotrophic metabolic pathways used in combination with ALE could likely result in a strain capable of growing on methanol as the sole carbon source. The DMS3286 strain used in this thesis is furthermore naturally capable of producing high yields of lipids from sugars<sup>[142]</sup>. It would be exciting to produce these and other industrially relevant products from CO<sub>2</sub>-derived methanol in the future. It was shown that the deletion of peroxisomal biogenesis factor 10 (PEX10) leads to

an increase in lipid content in various *Y. lipolytica* engineering projects<sup>[208]</sup>. By combining methanol assimilation with PEX10 deletion, among others, the non-conventional yeast *Y. lipolytica* could turn into an efficient platform for the sustainable production of lipids in the future.

Collectively, the chapters included in this dissertation demonstrate that high-throughput approaches and the screening of large libraries depict one fruitful way to deal with current biotechnological questions. The development of synthetic methylotrophy in yeast remains a challenge, but one worth taking on. For this purpose, yeasts have excellent potential due to their tolerance towards harsh conditions and their ability to separate toxic intermediates in organelles. These traits can be used to develop superior synthetic methylotrophs capable of recycling CO<sub>2</sub>-derived methanol, and thus contributing to a future circular economy. With the recent developments in artificial intelligence and machine learning techniques, the used methods could be further improved in terms of robustness, rapidity and quality of the results.

## 4.6 Future perspectives

“Due to the depletion of fossil fuels and concerns about environmental pollution, there is an urgent need to develop sustainable and climate neutral products and chemicals. In this regard, the application of CO<sub>2</sub>-derived C<sub>1</sub> feedstock received great attention. One reason is that the key feedstock CO<sub>2</sub>, is virtually unlimited<sup>[209]</sup>. Therefore, start-ups and established companies strived into the field and research efforts are on-going. The valorization of gaseous C<sub>1</sub> substrates might be one possibility in the solution space to face climate change and obtain sustainable commercial production processes for bulk and fine-chemicals as well as biofuels. In addition, CCU approaches to fix CO<sub>2</sub> into methanol or formate followed by fermentation is promising for future directions into a sustainable and cyclic bioeconomy.

The nature of methylotrophs empowers these microbes to utilize renewable derived C<sub>1</sub> feedstock, depicting them as attractive biotechnological platform strains for industrial strain development. In particular, from the viewpoint of the bioprocess, these strains provide key features to establish sustainable bioprocesses. Nevertheless, challenges remain and limit their broad use on large and commercially feasible scales so far, in terms of applying gaseous or liquid substrates. In particular, remaining challenges are low conversion and growth rates using gaseous substrates, low biomass yields and a lack of reliable genetic engineering tools when considering native methylotrophs. From the viewpoint of synthetic methylotrophy, implementation of functional genetic methylotrophic modules in established industrial host organism were so far introduced but are still limited. To date, the literature indicates that the

installation of metabolic regeneration cycles such as supporting carbon re-entry towards the pentose phosphate pathway from C<sub>1</sub> fueled central carbon metabolism is a crucial target for synthetic methylotrophy. Moreover, the recent engineering of industrially relevant microbes, such as *E. coli* or *S. cerevisiae*, towards utilization of methanol or formate as the sole carbon source succeeded. Despite this, the future for both approaches, native and synthetic methylotrophy, seems promising, as the tools and technologies are now emerging to push the frontier towards efficient C<sub>1</sub>- utilization in a modern bioeconomy.”<sup>[1]</sup>

## 5 REFERENCES

- [1] V. Wegat, J. T. Fabarius, V. Sieber, *Biotechnol Biofuels Bioprod* **2022**, *15*, 113.
- [2] aQ. I. Roode-Gutzmer, D. Kaiser, M. Bertau, *ChemBioEng Reviews* **2019**, *6*, 209-236; bT. J. Deka, A. I. Osman, D. C. Baruah, D. W. Rooney, *Environmental Chemistry Letters* **2022**, *20*, 3525-3554.
- [3] J. T. Fabarius, V. Wegat, A. Roth, V. Sieber, *Trends Biotechnol* **2021**, *39*, 348-358.
- [4] W. Zhang, M. Song, Q. Yang, Z. Dai, S. Zhang, F. Xin, W. Dong, J. Ma, M. Jiang, *Biotechnol Biofuels* **2018**, *11*, 260.
- [5] L. F. Distillation, *Industrial & Engineering Chemistry* **1926**, *18*, 929-930.
- [6] L. A. Pellegrini, G. Soave, S. Gamba, S. Langè, *Applied Energy* **2011**, *88*, 4891-4897.
- [7] J. D. Colgan, *International Security* **2013**, *38*, 147-180.
- [8] R. Fall, A. A. Benson, *Trends in Plant Science* **1996**, *1*, 296-301.
- [9] B. Schink, J. G. Zeikus, *Current Microbiology* **1980**, *4*, 387-389.
- [10] J. M. Peregrin-Alvarez, S. Tsoka, C. A. Ouzounis, *Genome Res* **2003**, *13*, 422-427.
- [11] H. Yurimoto, K. Shiraishi, Y. Sakai, *Microorganisms* **2021**, *9*.
- [12] aC. P. Kurtzman, *Int J Syst Evol Microbiol* **2005**, *55*, 973-976; bG. Péter, D. Dlačny, J. Tornai-Lehoczki, C. P. Kurtzman, *Antonie Van Leeuwenhoek* **2005**, *88*, 241-247; cS.-O. Suh, M. Blackwell, C. P. Kurtzman, M.-A. Lachance, *Mycologia* **2006**, *98*, 1006-1017; dS. Limtong, N. Srisuk, W. Yongmanitchai, H. Yurimoto, T. Nakase, *Int J Syst Evol Microbiol* **2008**, *58*, 302-307.
- [13] C. Anthony, *Sci Prog* **1975**, *62*, 167-206.
- [14] C. P. Kurtzman, *Journal of Industrial Microbiology and Biotechnology* **2009**, *36*, 1435-1435.
- [15] aN. Kato, M. Kano, Y. Tani, K. Ogata, *Agricultural and Biological Chemistry* **1974**, *38*, 111-116; bM. Y. Feldman, *Prog Nucleic Acid Res Mol Biol* **1973**, *13*, 1-49.
- [16] C.-T. Chen, F. Y. H. Chen, I. W. Bogorad, T.-Y. Wu, R. Zhang, A. S. Lee, J. C. Liao, *Metabolic Engineering* **2018**, *49*, 257-266.
- [17] J. P. Van Dijken, W. Harder, A. J. Beardsmore, J. R. Quayle, *FEMS Microbiology Letters* **1978**, *4*, 97-102.
- [18] H. Rußmayer, M. Buchetics, C. Gruber, M. Valli, K. Grillitsch, G. Modarres, R. Guerrasio, K. Klavins, S. Neubauer, H. Drexler, M. Steiger, C. Troyer, A. Al Chalabi, G. Krebiehl, D. Sonntag, G. Zellnig, G. Daum, A. B. Graf, F. Altmann, G. Koellensperger, S. Hann, M. Sauer, D. Mattanovich, B. Gasser, *BMC Biology* **2015**, *13*, 80.
- [19] H. Yurimoto, M. Oku, Y. Sakai, *Int J Microbiol* **2011**, *2011*, 101298.
- [20] N. Kato, H. Yoshikawa, K. Tanaka, M. Shimao, C. Sakazawa, *Archives of Microbiology* **1988**, *150*, 155-159.
- [21] M. E. Lidstrom, in *The Prokaryotes: Volume 2: Ecophysiology and Biochemistry* (Eds.: M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, E. Stackebrandt), Springer New York, New York, NY, **2006**, pp. 618-634.
- [22] T. Brautaset, M. Ø. Jakobsen, M. C. Flickinger, S. Valla, T. E. Ellingsen, *J Bacteriol* **2004**, *186*, 1229-1238.
- [23] L. V. Chistoserdova, M. E. Lidstrom, *J Bacteriol* **1994**, *176*, 7398-7404.
- [24] G. J. Crowther, G. Kosály, M. E. Lidstrom, *J Bacteriol* **2008**, *190*, 5057-5062.
- [25] H. Smejkalová, T. J. Erb, G. Fuchs, *PLoS One* **2010**, *5*.

- [26] aH. He, E. Noor, P. A. Ramos-Parra, L. E. García-Valencia, J. A. Patterson, R. I. Díaz de la Garza, A. D. Hanson, A. Bar-Even, *Metabolites* **2020**, *10*, 65; bR. Peyraud, K. Schneider, P. Kiefer, S. Massou, J. A. Vorholt, J.-C. Portais, *BMC Systems Biology* **2011**, *5*, 189.
- [27] aR. S. Hanson, T. E. Hanson, *Microbiological Reviews* **1996**, *60*, 439-471; bC. A. Raines, *Photosynthesis Research* **2003**, *75*, 1-10.
- [28] M.-L. Champigny, E. Bismuth, *Physiologia Plantarum* **1976**, *36*, 95-100.
- [29] J. A. Bassham, A. A. Benson, L. D. Kay, A. Z. Harris, A. T. Wilson, M. Calvin, *Journal of the American Chemical Society* **1954**, *76*, 1760-1770.
- [30] M. Mangiapia, K. Scott, *FEMS Microbiol Lett* **2016**, *363*.
- [31] N. H. Chen, K. Y. Djoko, F. J. Veyrier, A. G. McEwan, *Front Microbiol* **2016**, *7*, 257.
- [32] B. Lee, H. Yurimoto, Y. Sakai, N. Kato, *Microbiology* **2002**, *148*, 2697-2704.
- [33] R. P. Mason, J. K. M. Sanders, A. Crawford, B. K. Hunter, *Biochemistry* **1986**, *25*, 4504-4507.
- [34] O. Negruta, O. Csutak, I. Stoica, E. Rusu, T. Vassu, **2010**, *15*, 5369-5375.
- [35] aS. J. Allen, J. J. Holbrook, *Gene* **1995**, *162*, 99-104; bY. Asano, T. Sekigawa, H. Inukai, A. Nakazawa, *J Bacteriol* **1988**, *170*, 3189-3193.
- [36] M. Jormakka, B. Byrne, S. Iwata, *Curr Opin Struct Biol* **2003**, *13*, 418-423.
- [37] L. Chistoserdova, M. Laukel, J. C. Portais, J. A. Vorholt, M. E. Lidstrom, *J Bacteriol* **2004**, *186*, 22-28.
- [38] C. J. Marx, S. J. Van Dien, M. E. Lidstrom, *PLoS Biology* **2005**, *3*, e16.
- [39] S. Upadhyaya, S. Tiwari, N. Arora, D. P. Singh, *Microbes and Environmental Management* **2016**, 259-279.
- [40] T. Brautaset, M. D. Williams, R. D. Dillingham, C. Kaufmann, A. Bennaars, E. Crabbe, M. C. Flickinger, *Appl Environ Microbiol* **2003**, *69*, 3986-3995.
- [41] V. F. Wendisch, **2018**.
- [42] E. B. Drejer, D. T. C. Chan, C. Haupka, V. F. Wendisch, T. Brautaset, M. Irla, *Green Chemistry* **2020**, *22*, 788-802.
- [43] F. Sonntag, M. Buchhaupt, J. Schrader, *Appl Microbiol Biotechnol* **2014**, *98*, 4533-4544.
- [44] L. Pöschel, E. Gehr, M. Buchhaupt, *Applied Microbiology and Biotechnology* **2022**, *106*, 6713-6731.
- [45] H. T. Q. Le, D. H. A. Mai, J.-G. Na, E. Y. Lee, *Metabolic Engineering* **2022**, *72*, 150-160.
- [46] W. F. Liang, L. Y. Cui, J. Y. Cui, K. W. Yu, S. Yang, T. M. Wang, C. G. Guan, C. Zhang, X. H. Xing, *Metab Eng* **2017**, *39*, 159-168.
- [47] aD. Bourque, Y. Pomerleau, D. Groleau, *Applied microbiology and biotechnology* **1995**, *44*, 367-376; bZ. B. Mokhtari-Hosseini, E. Vasheghani-Farahani, A. Heidarzadeh-Vazifekhoran, S. A. Shojaosadati, R. Karimzadeh, K. K. Darani, *Bioresource technology* **2009**, *100*, 2436-2443.
- [48] K. Kobayashi, S. Kuwae, T. Ohya, T. Ohda, M. Ohyama, H. Ohi, K. Tomomitsu, T. Ohmura, *Journal of Bioscience and Bioengineering* **2000**, *89*, 55-61.
- [49] Z. Chen, H. Chen, X. Wang, X. Ma, B. Huang, *Protein Expression and Purification* **2007**, *52*, 239-248.
- [50] H. Lünsdorf, C. Gurramkonda, A. Adnan, N. Khanna, U. Rinas, *Microb Cell Fact* **2011**, *10*, 48-48.
- [51] aA. Vuorela, J. Myllyharju, R. Nissi, T. Pihlajaniemi, K. I. Kivirikko, *EMBO J* **1997**, *16*, 6702-6712; bM. Nokelainen, H. Tu, A. Vuorela, H. Notbohm, K. I. Kivirikko, J. Myllyharju, *Yeast* **2001**, *18*, 797-806.

- [52] M. W. T. Werten, T. J. van den Bosch, R. D. Wind, H. Mooibroek, F. A. de Wolf, *Yeast* **1999**, *15*, 1087-1096.
- [53] aM. W. T. Werten, A. P. H. A. Moers, T. Vong, H. Zuilhof, J. C. M. van Hest, F. A. de Wolf, *Biomacromolecules* **2008**, *9*, 1705-1711; bW. A. Gaines, W. R. Marcotte, Jr., *AATCC Rev* **2011**, *11*, 75-79.
- [54] aR. E. Sallach, V. P. Conticello, E. L. Chaikof, *Biotechnol Prog* **2009**, *25*, 1810-1818; bR. Schipperus, R. L. Teeuwen, M. W. Werten, G. Eggink, F. A. de Wolf, *Appl Microbiol Biotechnol* **2009**, *85*, 293-301.
- [55] M. W. T. Werten, G. Eggink, M. A. Cohen Stuart, F. A. de Wolf, *Biotechnol Adv* **2019**, *37*, 642-666.
- [56] aK. Wefelmeier, B. E. Ebert, L. M. Blank, S. Schmitz, *Front Bioeng Biotechnol* **2022**, *10*, 876316; bJ. H. Sohn, E. S. Choi, C. H. Kim, M. O. Agaphonov, M. D. Ter-Avanesyan, J. S. Rhee, S. K. Rhee, *J Bacteriol* **1996**, *178*, 4420-4428.
- [57] C. Su, L. Li, Z. Jin, X. Han, P. Zhao, L. Wang, C. Jiang, Y. Wang, W. Wang, D. Xu, N. Zhu, *Sheng Wu Gong Cheng Xue Bao* **2017**, *33*, 653-663.
- [58] E. Eilert, C. P. Hollenberg, M. Piontek, M. Suckow, *J Biotechnol* **2012**, *159*, 172-176.
- [59] J. H. C. Manfrão-Netto, A. M. V. Gomes, N. S. Parachin, *Front Bioeng Biotechnol* **2019**, *7*, 94.
- [60] M. Kumar, R. Saxena, R. S. Tomar, P. K. Rai, D. Paul, in *Microbes for Climate Resilient Agriculture* **2018**, pp. 149-164.
- [61] L. Schada von Borzyskowski, F. Severi, K. Krüger, L. Hermann, A. Gilardet, F. Sippel, B. Pommerenke, P. Claus, N. S. Cortina, T. Glatter, S. Zauner, J. Zarzycki, B. M. Fuchs, E. Bremer, U. G. Maier, R. I. Amann, T. J. Erb, *Nature* **2019**, *575*, 500-504.
- [62] G. Borrel, D. Jézéquel, C. Biderre-Petit, N. Morel-Desrosiers, J.-P. Morel, P. Peyret, G. Fonty, A.-C. Lehours, *Research in Microbiology* **2011**, *162*, 832-847.
- [63] aP. Kaszycki, H. Koloczek, *Microbiol Res* **2000**, *154*, 289-296; bP. Kaszycki, M. Tyszka, P. Malec, H. Kołoczek, *Biodegradation* **2001**, *12*, 169-177.
- [64] Y. Kung, W. Runguphan, J. D. Keasling, *ACS Synthetic Biology* **2012**, *1*, 498-513.
- [65] W. Siripong, C. Angela, S. Tanapongpipat, W. Runguphan, *Enzyme and Microbial Technology* **2020**, *138*, 109557.
- [66] T. Gassler, M. Sauer, B. Gasser, M. Egermeier, C. Troyer, T. Causon, S. Hann, D. Mattanovich, M. G. Steiger, *Nature biotechnology* **2020**, *38*, 210-216.
- [67] S. Ostergaard, L. Olsson, J. Nielsen, *Microbiol Mol Biol Rev* **2000**, *64*, 34-50.
- [68] G. Wang, M. Olofsson-Dolk, F. G. Hansson, S. Donati, X. Li, H. Chang, J. Cheng, J. Dahlin, I. Borodina, *ACS Synthetic Biology* **2021**, *10*, 3537-3550.
- [69] J. E. N. Müller, T. M. B. Heggeset, V. F. Wendisch, J. A. Vorholt, T. Brautaset, *Applied Microbiology and Biotechnology* **2015**, *99*, 535-551.
- [70] F. S. Hartner, A. Glieder, *Microb Cell Fact* **2006**, *5*, 39.
- [71] W. B. Whitaker, N. R. Sandoval, R. K. Bennett, A. G. Fast, E. T. Papoutsakis, *Current Opinion in Biotechnology* **2015**, *33*, 165-175.
- [72] C. J. R. Frazão, T. Walther, *Chemie Ingenieur Technik* **2020**, *92*, 1680-1699.
- [73] C. A. R. Cotton, N. J. Claassens, S. Benito-Vaquero, A. Bar-Even, *Current Opinion in Biotechnology* **2020**, *62*, 168-180.
- [74] S. Kim, S. N. Lindner, S. Aslan, O. Yishai, S. Wenk, K. Schann, A. Bar-Even, *Nature Chemical Biology* **2020**, *16*, 538-545.
- [75] M. Eisenhut, H. Bauwe, M. Hagemann, *FEMS Microbiol Lett* **2007**, *277*, 232-237.
- [76] P. Nicholls, *Biochemical and Biophysical Research Communications* **1975**, *67*, 610-616.
- [77] T. Warnecke, R. T. Gill, *Microb Cell Fact* **2005**, *4*, 25.

- [78] J. B. Siegel, A. L. Smith, S. Poust, A. J. Wargacki, A. Bar-Even, C. Louw, B. W. Shen, C. B. Eiben, H. M. Tran, E. Noor, J. L. Gallaher, J. Bale, Y. Yoshikuni, M. H. Gelb, J. D. Keasling, B. L. Stoddard, M. E. Lidstrom, D. Baker, *Proceedings of the National Academy of Sciences* **2015**, *112*, 3704-3709.
- [79] X. Wang, Y. Wang, J. Liu, Q. Li, Z. Zhang, P. Zheng, F. Lu, J. Sun, *Bioresources and Bioprocessing* **2017**, *4*, 41.
- [80] R. Breslow, *Tetrahedron Letters* **1959**, *1*, 22-26.
- [81] A. Bar-Even, *Biochemistry* **2016**, *55*, 3851-3863.
- [82] S. Poust, J. Piety, A. Bar-Even, C. Louw, D. Baker, J. D. Keasling, J. B. Siegel, *ChemBioChem* **2015**, *16*, 1950-1954.
- [83] S. Witthoff, K. Schmitz, S. Niedenführ, K. Nöh, S. Noack, M. Bott, J. Marienhagen, *Appl Environ Microbiol* **2015**, *81*, 2215-2225.
- [84] S. Gleizer, R. Ben-Nissan, Y. M. Bar-On, N. Antonovsky, E. Noor, Y. Zohar, G. Jona, E. Krieger, M. Shamshoum, A. Bar-Even, R. Milo, *Cell* **2019**, *179*, 1255-1263.e1212.
- [85] X. Lu, Y. Liu, Y. Yang, S. Wang, Q. Wang, X. Wang, Z. Yan, J. Cheng, C. Liu, X. Yang, H. Luo, S. Yang, J. Gou, L. Ye, L. Lu, Z. Zhang, Y. Guo, Y. Nie, J. Lin, S. Li, C. Tian, T. Cai, B. Zhuo, H. Ma, W. Wang, Y. Ma, Y. Liu, Y. Li, H. Jiang, *Nat Commun* **2019**, *10*, 1378.
- [86] aG. L. Rosano, E. A. Ceccarelli, *Frontiers in Microbiology* **2014**, *5*; bA. M. Vieira Gomes, T. Souza Carmo, L. Silva Carvalho, F. Mendonça Bahia, N. S. Parachin, *Microorganisms* **2018**, *6*.
- [87] M. Palma, J. F. Guerreiro, I. Sá-Correia, *Frontiers in Microbiology* **2018**, *9*.
- [88] W. C. DeLoache, Z. N. Russ, J. E. Dueber, *Nature communications* **2016**, *7*, 11152-11152.
- [89] aM. I. Espinosa, R. A. Gonzalez-Garcia, K. Valgepea, M. Plan, C. Scott, I. S. Pretorius, E. Marcellin, I. T. Paulsen, T. C. Williams, *bioRxiv* **2020**, 717942; bZ. Dai, H. Gu, S. Zhang, F. Xin, W. Zhang, W. Dong, J. Ma, H. Jia, M. Jiang, *Bioresour Technol* **2017**, *245*, 1407-1412; cC. Zhan, X. Li, Y. Yang, Y. Sun, S. Wang, Y. Wang, G. Wang, J. Nielsen, J. Keasling, Y. Chen, Z. Bai, *SSRN Electronic Journal* **2021**.
- [90] M. I. Espinosa, K. Valgepea, R. A. Gonzalez-Garcia, C. Scott, I. S. Pretorius, E. Marcellin, I. T. Paulsen, T. C. Williams, *bioRxiv* **2019**, 717942.
- [91] M. I. Espinosa, R. A. Gonzalez-Garcia, K. Valgepea, M. R. Plan, C. Scott, I. S. Pretorius, E. Marcellin, I. T. Paulsen, T. C. Williams, *Nature Communications* **2020**, *11*, 5564.
- [92] G. Barth, C. Gaillardin, *FEMS Microbiol Rev* **1997**, *19*, 219-237.
- [93] F. M. Pérez-Campo, A. Domínguez, *Curr Microbiol* **2001**, *43*, 429-433.
- [94] M. Groenewald, T. Boekhout, C. Neuvéglise, C. Gaillardin, P. W. van Dijck, M. Wyss, *Crit Rev Microbiol* **2014**, *40*, 187-206.
- [95] D. Turck, J. Castenmiller, S. de Henauw, K. I. Hirsch-Ernst, J. Kearney, A. Maciuk, I. Mangelsdorf, H. J. McArdle, A. Naska, C. Pelaez, K. Pentieva, A. Siani, F. Thies, S. Tsabouri, M. Vinceti, F. Cubadda, K. H. Engel, T. Frenzel, M. Heinonen, R. Marchelli, M. Neuhäuser-Berthold, A. Pötting, M. Poulsen, Y. Sanz, J. R. Schlatter, H. van Loveren, R. Ackerl, H. K. Knutsen, *Efsa j* **2019**, *17*, e05594.
- [96] G. Barth, C. Gaillardin, in *Nonconventional Yeasts in Biotechnology: A Handbook* (Ed.: K. Wolf), Springer Berlin Heidelberg, Berlin, Heidelberg, **1996**, pp. 313-388.
- [97] aP. Fickers, P.-H. Benetti, Y. Waché, A. Marty, S. Mauersberger, M. S. Smit, J.-M. Nicaud, *FEMS Yeast Research* **2005**, *5*, 527-543; bM. A. Coelho, P. Amaral, I. Belo, *Curr. Res. Technol. Educ. Top. Appl. Microbiol. Microb. Biotechnol.* **2010**, *2*.
- [98] D. Mamaev, R. Zvyagilskaya, *FEMS Yeast Res* **2021**, *21*.

- [99] C. L. Flores, C. Gancedo, *Eukaryot Cell* **2005**, *4*, 356-364.
- [100] J.-M. Nicaud, *Yeast* **2012**, *29*, 409-418.
- [101] D. Xie, E. N. Jackson, Q. Zhu, *Applied Microbiology and Biotechnology* **2015**, *99*, 1599-1610.
- [102] B. Zieniuk, A. Fabiszewska, *World J Microbiol Biotechnol* **2018**, *35*, 10.
- [103] L. Moeller, M. Grünberg, A. Zehnsdorf, B. Strehlitz, T. Bley, *Engineering in Life Sciences* **2010**, *10*, 311-320.
- [104] D. Xie, E. Miller, P. Sharpe, E. Jackson, Q. Zhu, *Biotechnology and Bioengineering* **2017**, *114*, 798-812.
- [105] T.-K. Ng, A.-Q. Yu, H. Ling, N. K. Pratomo Juwono, W. J. Choi, S. S. J. Leong, M. W. Chang, *Journal of Bioscience and Bioengineering* **2020**, *129*, 31-40.
- [106] K. M. Shaikh, A. A. Odaneth, *Biotechnology Progress* **2021**, *37*, e3201.
- [107] S. Li, L. Rong, S. Wang, S. Liu, Z. Lu, L. Miao, B. Zhao, C. Zhang, D. Xiao, K. Pushpanathan, A. Wong, A. Yu, *Chemical Engineering Science* **2022**, *249*, 117342.
- [108] C. D. Rutter, C. V. Rao, *Metabolic Engineering* **2016**, *38*, 139-147.
- [109] Q. Gao, X. Cao, Y.-Y. Huang, J.-L. Yang, J. Chen, L.-J. Wei, Q. Hua, *ACS Synthetic Biology* **2018**, *7*, 1371-1380.
- [110] P. Xu, K. Qiao, W. S. Ahn, G. Stephanopoulos, *Proceedings of the National Academy of Sciences* **2016**, *113*, 10848-10853.
- [111] M. Larroude, E. Celinska, A. Back, S. Thomas, J. M. Nicaud, R. Ledesma-Amaro, *Biotechnol Bioeng* **2018**, *115*, 464-472.
- [112] K. Gemperlein, D. Dietrich, M. Kohlstedt, G. Zipf, H. S. Bernauer, C. Wittmann, S. C. Wenzel, R. Müller, *Nature Communications* **2019**, *10*, 4055.
- [113] N. Imatoukene, A. Back, M. Nonus, B. Thomasset, T. Rossignol, J. M. Nicaud, *J Ind Microbiol Biotechnol* **2020**, *47*, 403-412.
- [114] L.-J. Wei, X. Cao, J.-J. Liu, S. Kwak, Y.-S. Jin, W. Wang, Q. Hua, *Appl Environ Microbiol* **2021**, *87*, e00481-00421.
- [115] R. Dulermo, F. Brunel, T. Dulermo, R. Ledesma-Amaro, J. Vion, M. Trassaert, S. Thomas, J.-M. Nicaud, C. Leplat, *Microb Cell Fact* **2017**, *16*, 31.
- [116] S. Müller, T. Sandal, P. Kamp-Hansen, H. Dalbøge, *Yeast* **1998**, *14*, 1267-1283.
- [117] S. P. Hong, J. Seip, D. Walters-Pollak, R. Rupert, R. Jackson, Z. Xue, Q. Zhu, *Yeast* **2012**, *29*, 59-72.
- [118] J. Blazeck, L. Liu, H. Redden, H. Alper, *Appl Environ Microbiol* **2011**, *77*, 7905-7914.
- [119] D. e. al., **2006**.
- [120] C. Madzak, B. Tréton, S. Blanchin-Roland, *J Mol Microbiol Biotechnol* **2000**, *2*, 207-216.
- [121] X. Xiong, S. Chen, *ACS Synthetic Biology* **2020**, *9*, 2208-2213.
- [122] G. Pignède, H.-J. Wang, F. Fudalej, M. Seman, C. Gaillardin, J.-M. Nicaud, *Appl Environ Microbiol* **2000**, *66*, 3283-3289.
- [123] M. Shabbir Hussain, I. Wheeldon, M. A. Blenner, *Biotechnol J* **2017**, *12*.
- [124] M. Trassaert, M. Vandermies, F. Carly, O. Denies, S. Thomas, P. Fickers, J.-M. Nicaud, *Microb Cell Fact* **2017**, *16*, 141.
- [125] J. V. Geisberg, Z. Moqtaderi, X. Fan, F. Oszolak, K. Struhl, *Cell* **2014**, *156*, 812-824.
- [126] M. Larroude, T. Rossignol, J. M. Nicaud, R. Ledesma-Amaro, *Biotechnol Adv* **2018**, *36*, 2150-2164.
- [127] K. A. Curran, N. J. Morse, K. A. Markham, A. M. Wagman, A. Gupta, H. S. Alper, *ACS Synth Biol* **2015**, *4*, 824-832.
- [128] C. M. Schwartz, M. S. Hussain, M. Blenner, I. Wheeldon, *ACS Synthetic Biology* **2016**, *5*, 356-359.



- [129] C. Schwartz, K. Frogue, A. Ramesh, J. Misa, I. Wheeldon, *Biotechnol Bioeng* **2017**, *114*, 2896-2906.
- [130] C. Schwartz, N. Curtis, A. K. Löbs, I. Wheeldon, *Biotechnol J* **2018**, *13*, e1700584.
- [131] T.-Q. Shi, H. Huang, E. J. Kerkhoven, X.-J. Ji, *Applied Microbiology and Biotechnology* **2018**, *102*, 9541-9548.
- [132] L. Wong, J. Engel, E. Jin, B. Holdridge, P. Xu, *Metabolic Engineering Communications* **2017**, *5*, 68-77.
- [133] C. Holkenbrink, M. I. Dam, K. R. Kildegaard, J. Beder, J. Dahlin, D. Doménech Belda, I. Borodina, *Biotechnol J* **2018**, *13*, e1700543.
- [134] S. Gao, L. Han, L. Zhu, M. Ge, S. Yang, Y. Jiang, D. Chen, *Biotechnol Lett* **2014**, *36*, 2523-2528.
- [135] C. Engler, R. Kandzia, S. Marillonnet, *PLoS One* **2008**, *3*, e3647.
- [136] E. Celińska, R. Ledesma-Amaro, M. Larroude, T. Rossignol, C. Pauthenier, J. M. Nicaud, *Microb Biotechnol* **2017**, *10*, 450-455.
- [137] aM. Egermeier, H. Russmayer, M. Sauer, H. Marx, *Front Microbiol* **2017**, *8*, 49; bE. Carsanba, S. Papanikolaou, P. Fickers, H. Erten, *Yeast* **2019**, *36*, 319-327; cJ. Quarterman, P. J. Slininger, C. P. Kurtzman, S. R. Thompson, B. S. Dien, *Appl Microbiol Biotechnol* **2017**, *101*, 3319-3334.
- [138] M. Egermeier, M. Sauer, H. Marx, *FEMS Microbiology Letters* **2019**, 366.
- [139] P. Sarkari, H. Marx, M. L. Blumhoff, D. Mattanovich, M. Sauer, M. G. Steiger, *Bioresource technology* **2017**, *245*, 1327-1333.
- [140] E. Vartiainen, P. Blomberg, M. Ilmén, M. Andberg, M. Toivari, M. Penttilä, *Fungal Biology and Biotechnology* **2019**, *6*, 27.
- [141] S. Zhang, F. Guo, Q. Yang, Y. Jiang, S. Yang, J. Ma, F. Xin, T. Hasunuma, A. Kondo, W. Zhang, M. Jiang, *Green Chemistry* **2023**, *25*, 183-195.
- [142] K. Nambou, C. Zhao, L. Wei, J. Chen, T. Imanaka, Q. Hua, *Bioresource Technology* **2014**, *173*, 324-333.
- [143] S. Güner, V. Wegat, A. Pick, V. Sieber, *Green Chemistry* **2021**, *23*, 6583-6590.
- [144] K. A. Markham, S. Vazquez, H. S. Alper, *FEMS Yeast Res* **2018**, *18*.
- [145] U. K. Laemmli, *Nature* **1970**, *227*, 680-685.
- [146] J. H. Zhang, T. D. Chung, K. R. Oldenburg, *J Biomol Screen* **1999**, *4*, 67-73.
- [147] K. Huynh, C. L. Partch, *Curr Protoc Protein Sci* **2015**, *79*, 28.29.21-28.29.14.
- [148] P. G. Mazzola, A. F. Jozala, L. C. d. L. Novaes, P. Moriel, T. C. V. Penna, *Brazilian Journal of Pharmaceutical Sciences* **2009**, *45*, 241-248.
- [149] A. Krog, T. M. Heggeset, J. E. Müller, C. E. Kupper, O. Schneider, J. A. Vorholt, T. E. Ellingsen, T. Brautaset, *PLoS One* **2013**, *8*, e59188.
- [150] W. B. Whitaker, J. A. Jones, R. K. Bennett, J. E. Gonzalez, V. R. Vernacchio, S. M. Collins, M. A. Palmer, S. Schmidt, M. R. Antoniewicz, M. A. Koffas, *Metabolic engineering* **2017**, *39*, 49-59.
- [151] K. Kobayashi, in *Encyclopedia of Astrobiology* (Eds.: R. Amils, M. Gargaud, J. Cernicharo Quintanilla, H. J. Cleaves, W. M. Irvine, D. Pinti, M. Viso), Springer Berlin Heidelberg, Berlin, Heidelberg, **2014**, pp. 1-1.
- [152] F. Y. Chen, H. W. Jung, C. Y. Tsuei, J. C. Liao, *Cell* **2020**, *182*, 933-946.e914.
- [153] A. Bar-Even, E. Noor, N. E. Lewis, R. Milo, *Proceedings of the National Academy of Sciences* **2010**, *107*, 8889-8894.
- [154] O. Yishai, L. Goldbach, H. Tenenboim, S. N. Lindner, A. Bar-Even, *ACS Synthetic Biology* **2017**, *6*, 1722-1731.
- [155] P. Schneider, S. Klamt, *Bioinformatics* **2019**, *35*, 3063-3072.
- [156] Y. Nov, *PLoS One* **2013**, *8*, e68069.
- [157] C. M. Miton, N. Tokuriki, *Protein Science* **2016**, *25*, 1260-1272.

- [158] C. K. Longwell, L. Labanieh, J. R. Cochran, *Current Opinion in Biotechnology* **2017**, *48*, 196-202.
- [159] B. G. Fox, P. G. Blommel, *Curr Protoc Protein Sci* **2009**, *Chapter 5*, 5.23.21-25.23.18.
- [160] P. Sobrado, M. A. Goren, D. James, C. K. Amundson, B. G. Fox, *Protein Expr Purif* **2008**, *58*, 229-241.
- [161] H. Bar, A. Zweifach, *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **2020**, *25*, 1000-1008.
- [162] S. Lutz, *Curr Opin Biotechnol* **2010**, *21*, 734-743.
- [163] S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z.-G. Zhang, D. J. Opperman, M. T. Reetz, J. P. Acevedo, *ACS Synthetic Biology* **2013**, *2*, 83-92.
- [164] L. Zhang, R. Singh, S. D. Z. Guo, J. Li, F. Chen, Y. He, X. Guan, Y. C. Kang, J.-K. Lee, *Green Chemistry* **2018**, *20*, 230-242.
- [165] aA. S. Demir, P. Ayhan, S. B. Sopaci, *CLEAN – Soil, Air, Water* **2007**, *35*, 406-412; bP.-W. Seo, H.-J. Jo, I. Y. Hwang, H.-Y. Jeong, J.-H. Kim, J.-W. Kim, E. Y. Lee, J.-B. Park, J.-S. Kim, *Catalysis Science & Technology* **2020**, *10*, 79-85.
- [166] T. Miyata, S. Miyazawa, T. Yasunaga, *Journal of Molecular Evolution* **1979**, *12*, 219-236.
- [167] K. Peng, D. Guo, Q. Lou, X. Lu, J. Cheng, J. Qiao, L. Lu, T. Cai, Y. Liu, H. Jiang, *ACS Synthetic Biology* **2020**, *9*, 2902-2908.
- [168] aN. Nasoohi, K. Khajeh, M. Mohammadian, B. Ranjbar, *International journal of biological macromolecules* **2013**, *60*, 56-61; bS. R. SELVARAJ, A. N. SCHELLER, H. Z. MIAO, R. J. KAUFMAN, S. W. Pipe, *Journal of thrombosis and haemostasis* **2012**, *10*, 107-115.
- [169] T. M. B. Heggeset, V. Kucharova, I. Nærdal, S. Valla, H. Sletta, T. E. Ellingsen, T. Brautaset, *Appl Environ Microbiol* **2013**, *79*, 559-568.
- [170] C. A. Sarkar, I. Dodevski, M. Kenig, S. Dudli, A. Mohr, E. Hermans, A. Plückthun, *Proceedings of the National Academy of Sciences* **2008**, *105*, 14808-14813.
- [171] L. Foit, G. J. Morgan, M. J. Kern, L. R. Steimer, A. A. von Hacht, J. Titchmarsh, S. L. Warriner, S. E. Radford, J. C. Bardwell, *Mol Cell* **2009**, *36*, 861-871.
- [172] R. Chapman, M. H. Stenzel, *Journal of the American Chemical Society* **2019**, *141*, 2754-2769.
- [173] aA. H.-W. Yeh, C. Norn, Y. Kipnis, D. Tischer, S. J. Pellock, D. Evans, P. Ma, G. R. Lee, J. Z. Zhang, I. Anishchenko, B. Coventry, L. Cao, J. Dauparas, S. Halabiya, M. DeWitt, L. Carter, K. N. Houk, D. Baker, *Nature* **2023**, *614*, 774-780; bW. P. Russ, M. Figliuzzi, C. Stocker, P. Barrat-Charlaix, M. Socolich, P. Kast, D. Hilvert, R. Monasson, S. Cocco, M. Weigt, R. Ranganathan, *Science* **2020**, *369*, 440-445; cW. S. Mak, X. Wang, R. Arenas, Y. Cui, S. Bertolani, W. Q. Deng, I. Tagkopoulos, D. K. Wilson, J. B. Siegel, *Biochemistry* **2020**, *59*, 3834-3843.
- [174] aK. Rumbold, H. J. van Buijsen, V. M. Gray, J. W. van Groenestijn, K. M. Overkamp, R. S. Slomp, M. J. van der Werf, P. J. Punt, *Bioeng Bugs* **2010**, *1*, 359-366; bM. Spagnuolo, M. Shabbir Hussain, L. Gambill, M. Blenner, *Front Microbiol* **2018**, *9*, 1077.
- [175] Y.-K. Park, R. Ledesma-Amaro, *Trends in Biotechnology* **2023**, *41*, 242-254.
- [176] aN. A. Yasid, M. D. Rolfe, J. Green, M. P. Williamson, *Royal Society Open Science* **2016**, *3*, 160187; bM. Krantz, B. Nordlander, H. Valadi, M. Johansson, L. Gustafsson, S. Hohmann, *Eukaryot Cell* **2004**, *3*, 1381-1390.
- [177] S. C. Kowalczykowski, D. A. Dixon, A. K. Eggleston, S. D. Lauder, W. M. Rehrauer, *Microbiol Rev* **1994**, *58*, 401-465.

- [178] aY. Li, Z. Lin, C. Huang, Y. Zhang, Z. Wang, Y.-j. Tang, T. Chen, X. Zhao, *Metabolic Engineering* **2015**, *31*, 13-21; bJ. E. DiCarlo, J. E. Norville, P. Mali, X. Rios, J. Aach, G. M. Church, *Nucleic acids research* **2013**, *41*, 4336-4343.
- [179] R. W. Bradley, *Biotechnology Reports* **2021**, *32*, e00680.
- [180] aY. Tong, T. S. Jørgensen, C. M. Whitford, T. Weber, S. Y. Lee, *Nat Commun* **2021**, *12*, 5206; bM. E. Lee, W. C. DeLoache, B. Cervantes, J. E. Dueber, *ACS Synthetic Biology* **2015**, *4*, 975-986.
- [181] A. Zakrzewska, G. van Eikenhorst, J. E. Burggraaff, D. J. Vis, H. Hoefsloot, D. Delneri, S. G. Oliver, S. Brul, G. J. Smits, *Mol Biol Cell* **2011**, *22*, 4435-4446.
- [182] A. Peña, N. S. Sánchez, H. Álvarez, M. Calahorra, J. Ramírez, *FEMS Yeast Research* **2015**, *15*.
- [183] C. Walker, S. Ryu, C. T. Trinh, *Metab Eng* **2019**, *54*, 83-95.
- [184] A. A. Sibirny, *FEMS Yeast Res* **2016**, *16*.
- [185] aM. Workman, P. Holt, J. Thykaer, *AMB Express* **2013**, *3*, 58; bA. M. Erian, M. Egermeier, H. Marx, M. Sauer, *Yeast* **2022**, *39*, 323-336.
- [186] D. Yasokawa, S. Murata, Y. Iwahashi, E. Kitagawa, R. Nakagawa, T. Hashido, H. Iwahashi, *Applied Biochemistry and Biotechnology* **2010**, *160*, 1685-1698.
- [187] M. North, B. D. Gaytán, C. Romero, V. Y. De La Rosa, A. Loguinov, M. T. Smith, L. Zhang, C. D. Vulpe, *Frontiers in Genetics* **2016**, *7*.
- [188] L. N. Jayakody, N. Hayashi, H. Kitagaki, *Biotechnol Lett* **2011**, *33*, 285-292.
- [189] J. T. Cunha, P. O. Soares, S. L. Baptista, C. E. Costa, L. Domingues, *Bioengineered* **2020**, *11*, 883-903.
- [190] P. Keller, E. Noor, F. Meyer, M. A. Reiter, S. Anastassov, P. Kiefer, J. A. Vorholt, *Nature Communications* **2020**, *11*, 5403.
- [191] B. M. Woolston, J. R. King, M. Reiter, B. Van Hove, G. Stephanopoulos, *Nature communications* **2018**, *9*, 2387.
- [192] Y. Wang, L. Fan, P. Tuyishime, P. Zheng, J. Sun, *Trends in Biotechnology* **2020**, *38*, 650-666.
- [193] A. Byrd, R. St-Arnaud, in *Two-Hybrid Systems: Methods and Protocols* (Ed.: P. N. MacDonald), Humana Press, Totowa, NJ, **2001**, pp. 107-119.
- [194] P. A. Kelso, L. K. M. Chow, A. C. Carpenter, I. T. Paulsen, T. C. Williams, *ACS Synth Biol* **2022**, *11*, 2548-2563.
- [195] W. A. van Winden, R. Mans, S. Breestraat, R. A. J. Verlinden, Á. Mielgo-Gómez, E. A. F. de Hulster, H. M. C. J. de Bruijn, H. J. Noorman, *Biotechnology and Bioengineering* **2022**, 1-10.
- [196] K. R. Pomraning, E. L. Bredeweg, E. J. Kerkhoven, K. Barry, S. Haridas, H. Hundley, K. LaButti, A. Lipzen, M. Yan, J. K. Magnuson, B. A. Simmons, I. V. Grigoriev, J. Nielsen, S. E. Baker, *mSphere* **2018**, *3*.
- [197] O. Konzock, J. Norbeck, *PLoS One* **2020**, *15*, e0231161.
- [198] T. Zhu, T. Zhao, O. E. Bankefa, Y. Li, *Biotechnology Advances* **2020**, *39*, 107467.
- [199] H. Yurimoto, N. Kato, Y. Sakai, *Chem Rec* **2005**, *5*, 367-375.
- [200] O. Zarei, S. Dastmalchi, M. Hamzeh-Mivehroud, *Iran J Pharm Res* **2016**, *15*, 907-913.
- [201] F. W. Koopman, J. H. de Winde, H. J. Ruijssenaars, *Appl Microbiol Biotechnol* **2009**, *83*, 705-713.
- [202] L. Leßmeier, J. Pfeifenschneider, M. Carnicer, S. Heux, J. C. Portais, V. F. Wendisch, *Appl Microbiol Biotechnol* **2015**, *99*, 10163-10176.
- [203] F. Meyer, P. Keller, J. Hartl, O. G. Gröninger, P. Kiefer, J. A. Vorholt, *Nature Communications* **2018**, *9*, 1508.
- [204] H. O. Kammen, R. Koo, *J Biol Chem* **1969**, *244*, 4888-4893.

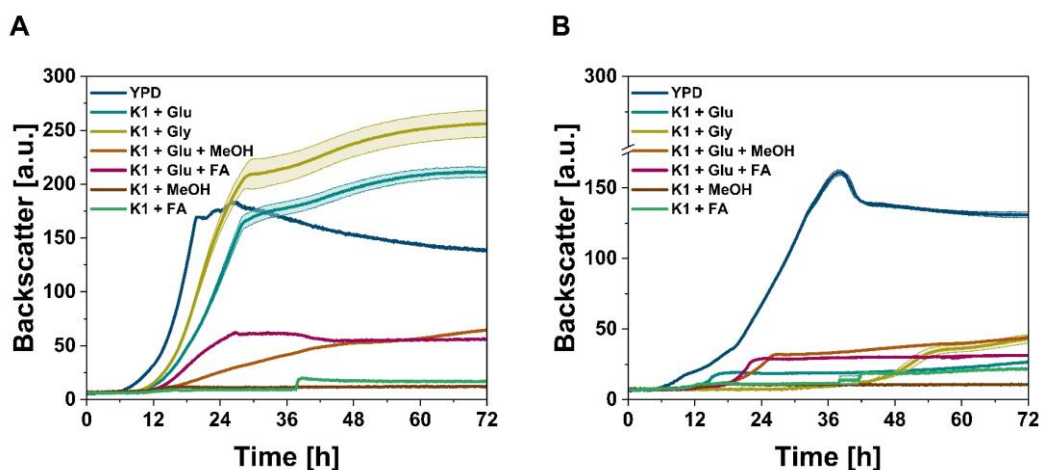
- [205] B. Gao, N. Zhao, J. Deng, Y. Gu, S. Jia, Y. Hou, X. Lv, L. Liu, *J Biotechnol* **2022**, *343*, 128-137.
- [206] M. Dragosits, D. Mattanovich, *Microb Cell Fact* **2013**, *12*, 64.
- [207] Y. Wu, A. Jameel, X.-H. Xing, C. Zhang, *Trends in Biotechnology* **2022**, *40*, 38-59.
- [208] S.-Y. Zeng, H.-H. Liu, T.-Q. Shi, P. Song, L.-J. Ren, H. Huang, X.-J. Ji, *European Journal of Lipid Science and Technology* **2018**, *120*, 1700352.
- [209] S. E. Manahan, *Fundamentals of environmental chemistry*, Lewis Publishers, Boca Raton [Fla.], **1993**.

## 6 APPENDIX

FIGURE S1: GROWTH CURVE OF <i>Y. LIPOLYTICA</i> DSM3286 (A) AND PO1F (B) ON VARIOUS SUBSTRATES. ....	118
FIGURE S2: SCHEMATIC OVERVIEW OF ROBOT-AIDED HIGH-THROUGHPUT ENGINEERING USING A TECAN PLATFORM. ....	119
FIGURE S3: DEVELOPMENT OF A SHIFT- AND PITCHFUNCTION IN ORDER TO MINIMIZE THE INFLUENCE OF PRECULTURE OD. ....	120
FIGURE S4: SDS-PAGE ANALYSIS OF THE PURIFICATION OF FLS WILDTYPE AND VARIANT L482Q/I557G EXPRESSED IN <i>E. COLI</i> BL21(DE3) FOR 16 H AT 25 °C USING 1.0 mM IPTG. ....	121
FIGURE S5: THERMAL DENATURATION OF FLS WILD TYPE (A, BLUE LINE) AND VARIANT (B, GREEN LINE) AT PH 8.0. ....	121
FIGURE S6: GROWTH CURVE OF <i>E. COLI</i> BL21(DE3) EXPRESSING THE EMPTY VECTOR pET24A CULTIVATED IN TB MEDIUM SUPPLEMENTED WITH VARIOUS FALD CONCENTRATIONS. ....	122
FIGURE S7: GROWTH OF <i>Y. LIPOLYTICA</i> STRAINS PO1F AND DSM3286 ON HYGROMYCINB CONTAINING AGAR PLATES. ....	122
FIGURE S8: FINAL OD <sub>600</sub> OF SCREENED <i>Y. LIPOLYTICA</i> STRAINS ON METHANOL. ....	123
FIGURE S9: SCREENING OF VARIOUS <i>Y. LIPOLYTICA</i> STRAINS ON METHANOL. ....	124
FIGURE S10: FINAL OD <sub>600</sub> OF RESCREENED <i>Y. LIPOLYTICA</i> STRAINS ON METHANOL. ....	125
FIGURE S11: CELL GROWTH OF <i>Y. LIPOLYTICA</i> STRAINS AD_N_ADH_HPS_PHI1 (A), AC_N_MDH2_FLS*2 (B) AND AC_N_MDH3_FLS3 (C) IN 25 mL OF MM WITH 2% METHANOL AND 0.5 G L <sup>-1</sup> YEAST EXTRACT.. ....	125
FIGURE S12: CELL GROWTH OF <i>Y. LIPOLYTICA</i> STRAINS AC_H_MDH2_FLS*3 (A), AC_H_MDH3_FLS4 (B), AC_N_ADH_FLS2 (C) AND AC_N_MDH2_FLS1 (D) IN 25 mL OF MM WITH 2% METHANOL AND 0.5 G L <sup>-1</sup> YEAST EXTRACT. ....	126

FIGURE S13: HPLC CHROMATOGRAM OF STANDARDS CONTAINING GLUCOSE (RT 12.4 MIN), D-LACTATE (RT 16.6 MIN), FORMATE (RT 17.7 MIN) AND METHANOL (RT 23.5 MIN) MEASURED WITH RI DETECTOR..	127
FIGURE S 14: HPLC CHROMATOGRAM OF 5 mM - 100 mM STANDARDS OF FORMALDEHYDE DETECTED WITH RI DETECTOR. FALD DEPICTED A RETENTION TIME OF 17.2 MIN.....	127
FIGURE S 15: HPLC CHROMATOGRAM OF 5 mM - 100 mM STANDARDS OF GLYCOLALDEHYDE DETECTED WITH RI DETECTOR. GALD DEPICTED A RETENTION TIME OF 15.8 MIN. ....	127
FIGURE S 16: HPLC CHROMATOGRAM OF 5 mM - 100 mM STANDARDS OF DIHYDROXYACETONE DETECTED WITH UV DETECTOR. DHA DEPICTED A RETENTION TIME OF 18.8 MIN. ....	128
FIGURE S 17: MEASURED CALIBRATION CURVE OF GLUCOSE IN A RANGE OF 0 – 20 G L <sup>-1</sup> . ....	128
FIGURE S 18: MEASURED CALIBRATION CURVE OF D-LACTATE IN A RANGE OF 0 – 5 G L <sup>-1</sup> . ....	128
FIGURE S 19: MEASURED CALIBRATION CURVE OF FORMATE IN A RANGE OF 0 – 5 G L <sup>-1</sup> .....	129
FIGURE S 20: MEASURED CALIBRATION CURVE OF METHANOL IN A RANGE OF 0 – 20 G L <sup>-1</sup> . ....	129
FIGURE S21: MEASURED CALIBRATION CURVE OF METHANOL IN A RANGE OF 0 – 15 G L <sup>-1</sup> . ....	130
TABLE S1: CALCULATED GROWTH RATE FOR <i>Y. LIPOLYTICA</i> STRAIN DSM3286 AND PO1F ON VARIOUS SUBSTRATES, CULTIVATION PERFORMED IN THE BIOLECTOR I.....	118

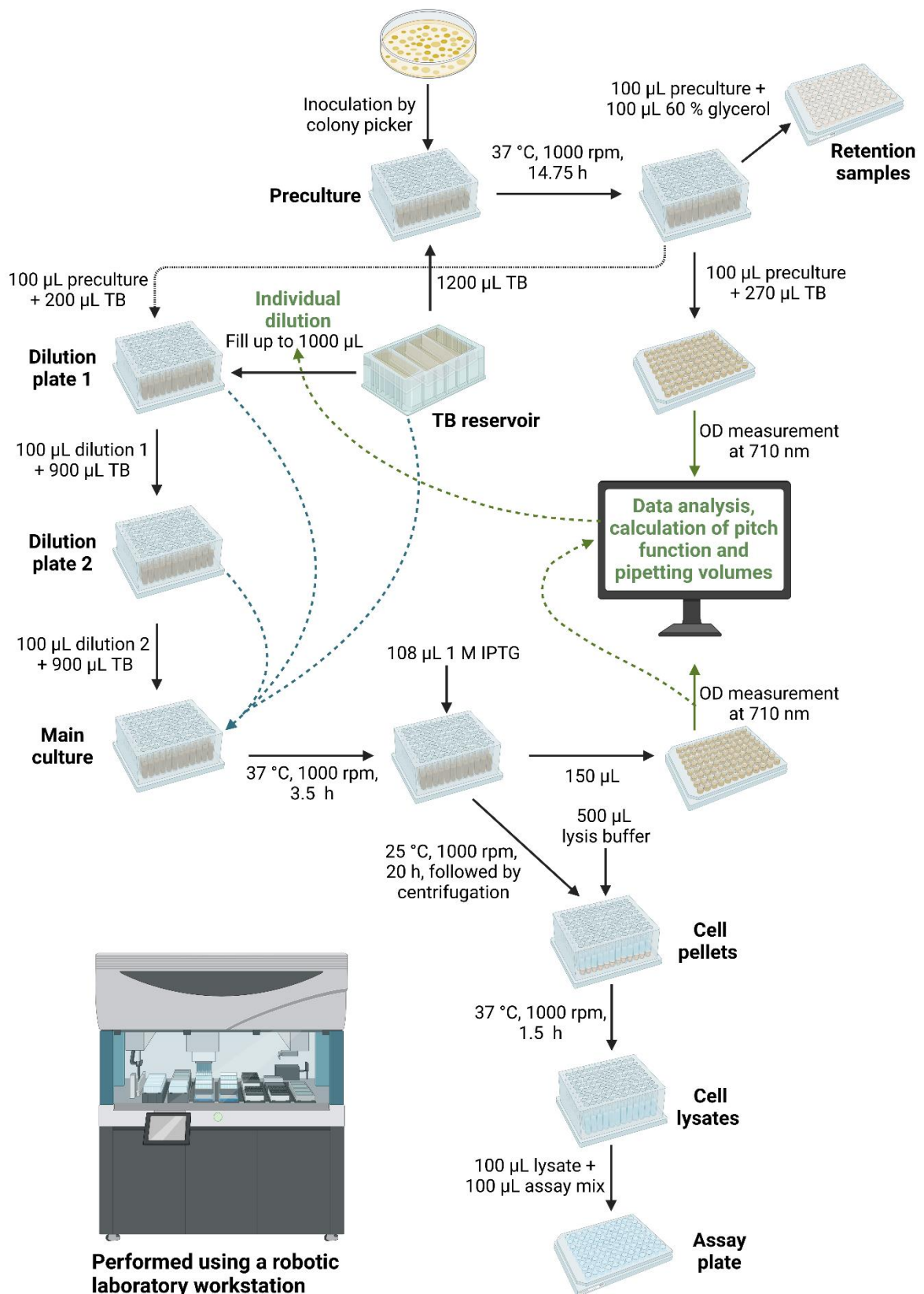
## 6.1 Additional Data



**Figure S1: Growth curve of *Y. lipolytica* DSM3286 (A) and PO1f (B) on various substrates.** Cells were cultivated in YPD, or minimal medium K1 supplemented with either 15 g L<sup>-1</sup> glucose, 15 g L<sup>-1</sup> glycerol, 7.9 g L<sup>-1</sup> methanol, or 1.5 g L<sup>-1</sup> formate as the single carbon source or in combination with glucose. Shown are the mean values of three biological replicates. Error bars represent standard deviations.

**Table S1: Calculated growth rate for *Y. lipolytica* strain DSM3286 and PO1f on various substrates, cultivation performed in the BioLector I.**

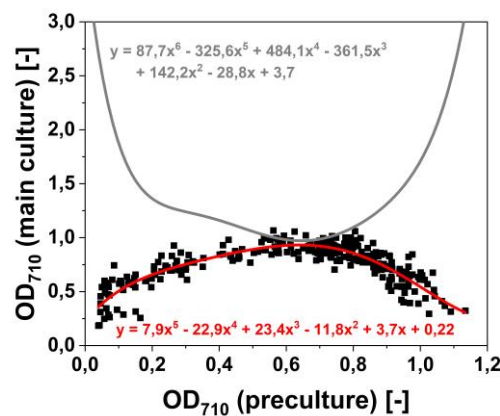
	Growth rate $\mu$ [h <sup>-1</sup> ] DSM3286	Growth rate $\mu$ [h <sup>-1</sup> ] PO1f
YPD	0.30	0.26
K1 + Glu	0.27	0.18
K1 + Gly	0.26	0.16
K1 + Glu + MeOH	0.13	0.15
K1 + Glu + FA	0.20	0.20
K1 + MeOH	0.04	0.04



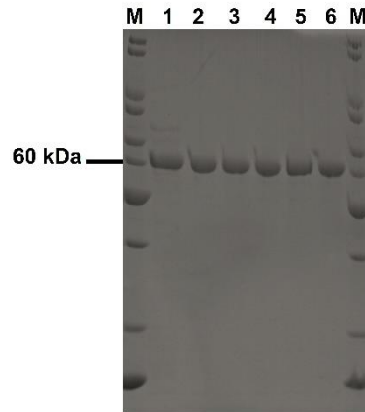
**Figure S2: Schematic overview of robot-aided high-throughput engineering using a Tecan platform.** For liquid preculture, 96-Well-Deep-Well-plates (96DWP) were filled with 1200 µL TB medium per well. Single colonies were picked by an automatic colony picker and preculture plates were incubated for 14.75 h at 37 °C and 1000 rpm. By use of the MCA 96, 90 µL were mixed with



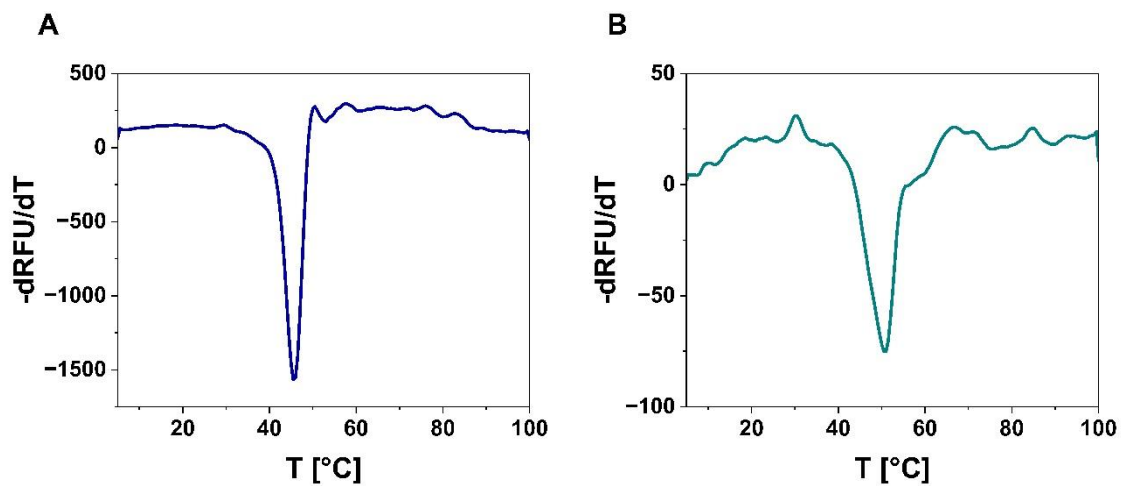
270  $\mu\text{L}$  TB and OD710 was measured. The OD-values were exported to a spreadsheet calculation program. Averaged blank OD was subtracted and target OD was divided by net OD values, resulting in individual dilution factors. 100  $\mu\text{L}$  aliquots of preculture were added to 200  $\mu\text{L}$  TB in a dilution plate I by MCA 96. To this plate, individual volumes of TB were added for further dilution by 8-Channel-LiHa using 1 mL pipette tips, based on calculation. After mixing, 100  $\mu\text{L}$  aliquots of this dilution were transferred into dilution plate II and mixed with 900  $\mu\text{L}$  TB for further tenfold dilution. Next, 100  $\mu\text{L}$  of these dilutions were mixed with 700  $\mu\text{L}$  TB in the main culture plate. For some cultures showing very low preculture OD, the dilution procedure in dilution plate I was out of range, limited by the maximal well capacity. Here, individual direct transfers from dilution plate I or dilution plate II to the main culture plate were executed by the 8-Channel-LiHa for complementation. Finally, the wells of the main culture plate were filled up to 1 mL with fresh TB. In addition, 100  $\mu\text{L}$  preculture were mixed with 100  $\mu\text{L}$  60 % Glycerol and stored at  $-80\text{ }^{\circ}\text{C}$ . Main cultures were incubated for 3.5 h at  $37\text{ }^{\circ}\text{C}$  and 1000 rpm. Next, OD710 was measured and 108  $\mu\text{L}$  IPTG solution was added for induction of formolase expression. Formolase was expressed at  $25\text{ }^{\circ}\text{C}$ , 1000 rpm for 20 h. Black lines indicate volume transfer, blue dashed lines represent individual volumes, and green dashed lines represent data transfer. Figure created with Biorender.com.



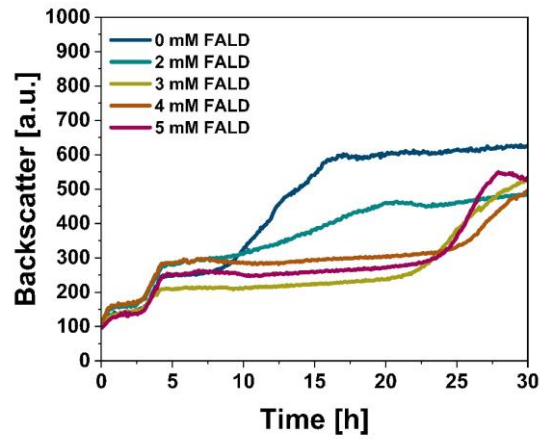
**Figure S3: Development of a shift- and pitchfunction in order to minimize the influence of preculture OD.** Black squares represent normalized OD values of four main culture plates, inoculated with different volumes of 14,75 h-precultures, depending on measured preculture OD and diluted in a proportional manner. Grey line represents shift function with a polynomial fit. Red line represents pitchfunction for shift-correction.



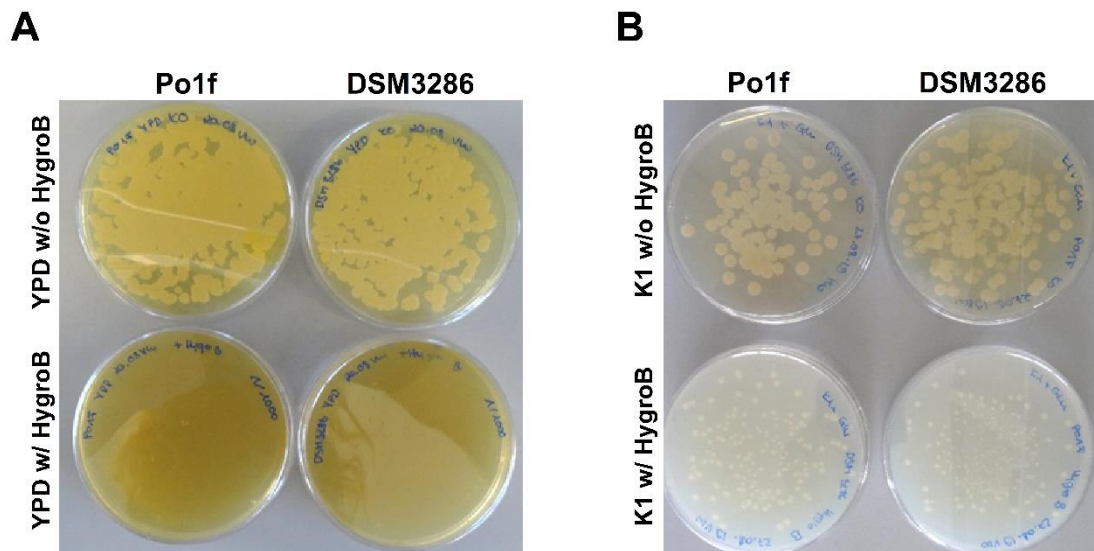
**Figure S4: SDS-PAGE analysis of the purification of FLS wildtype and variant L482Q/I557G expressed in *E. coli* BL21(DE3) for 16 h at 25 °C using 1.0 mM IPTG.** M: protein standard; lane 1-3: Purified FLS wild type after buffer exchange as used in the GDH assay; lane 4-6: Purified FLS variant L482Q/I557G after buffer exchange as used in the GDH assay. Expression and purification was performed with three biological replicates.



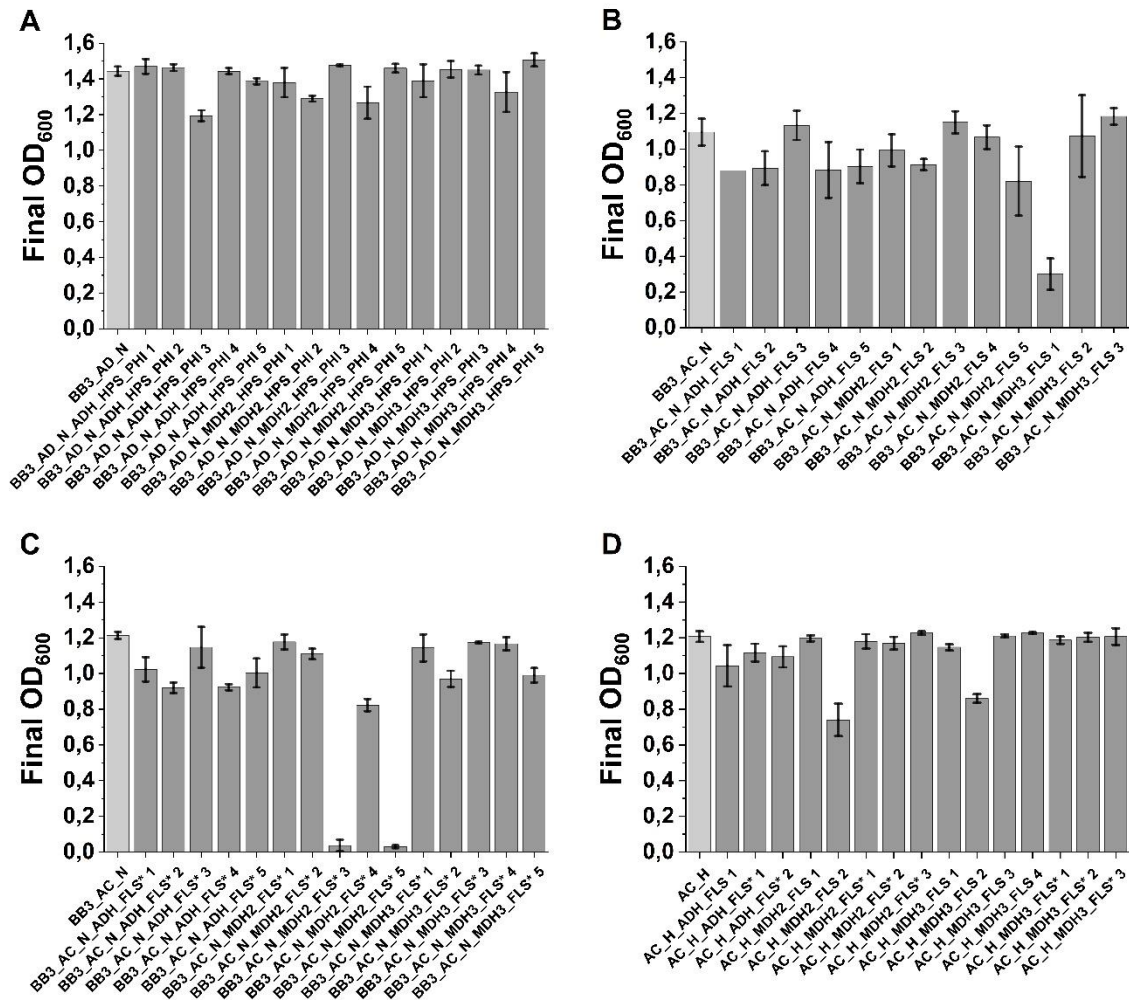
**Figure S5: Thermal denaturation of FLS wild type (A, blue line) and variant (B, green line) at pH 8.0.** The protein concentration is 1 mg mL<sup>-1</sup>. Shown is the mean of three replicates.



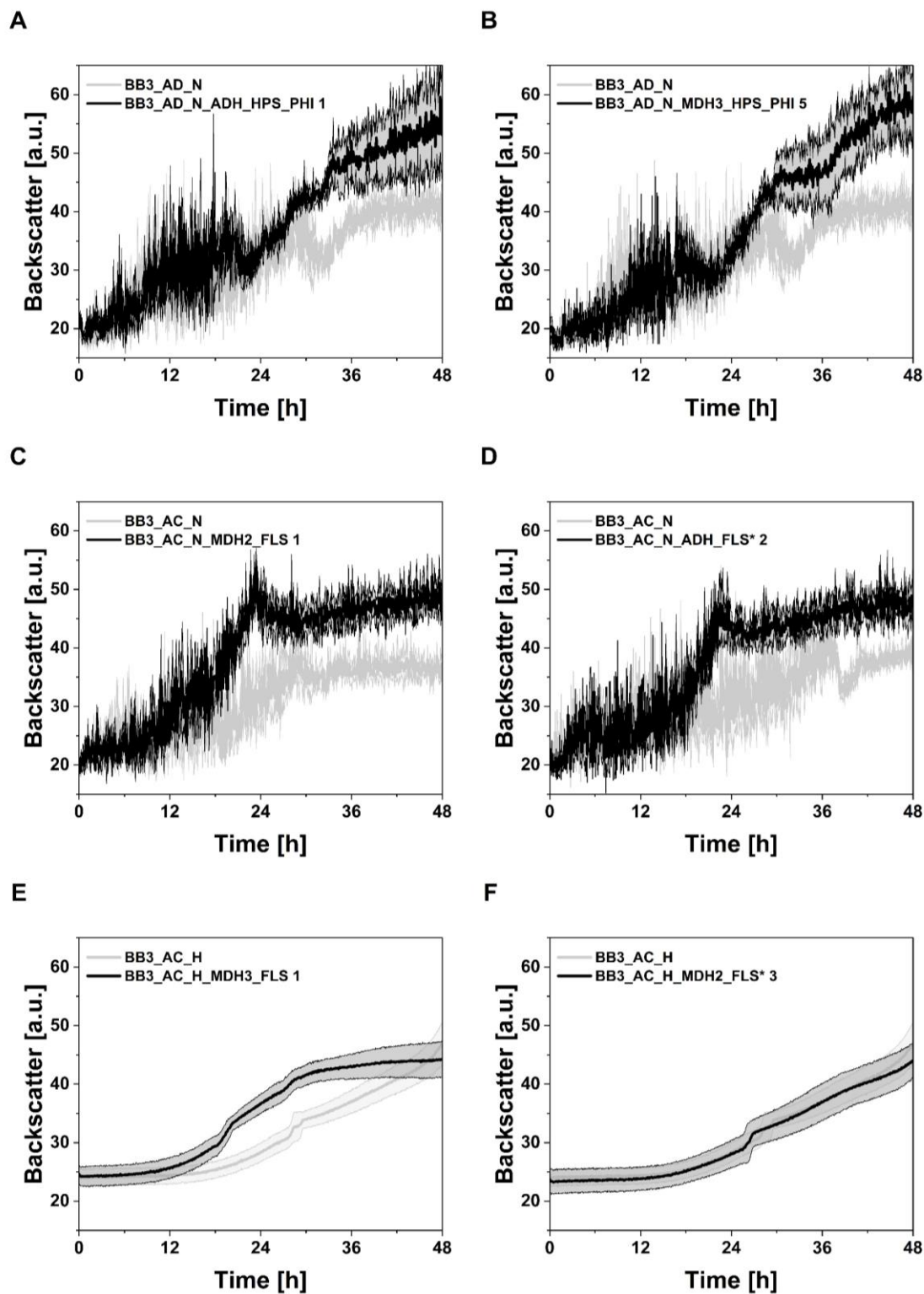
**Figure S6: Growth curve of *E. coli* BL21(DE3) expressing the empty vector pET24a cultivated in TB medium supplemented with various FALD concentrations.** Shown are the mean values of two replicates.



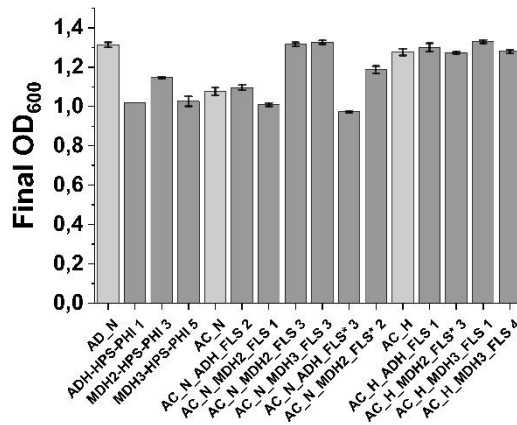
**Figure S7: Growth of *Y. lipolytica* strains PO1f and DSM3286 on HygromycinB containing agar plates.** Strains are able to bypass HygromycinB resistance on glucose ( $15 \text{ g L}^{-1}$ ) supplemented K1 minimal media plates (A), but not on YPD plates (B).



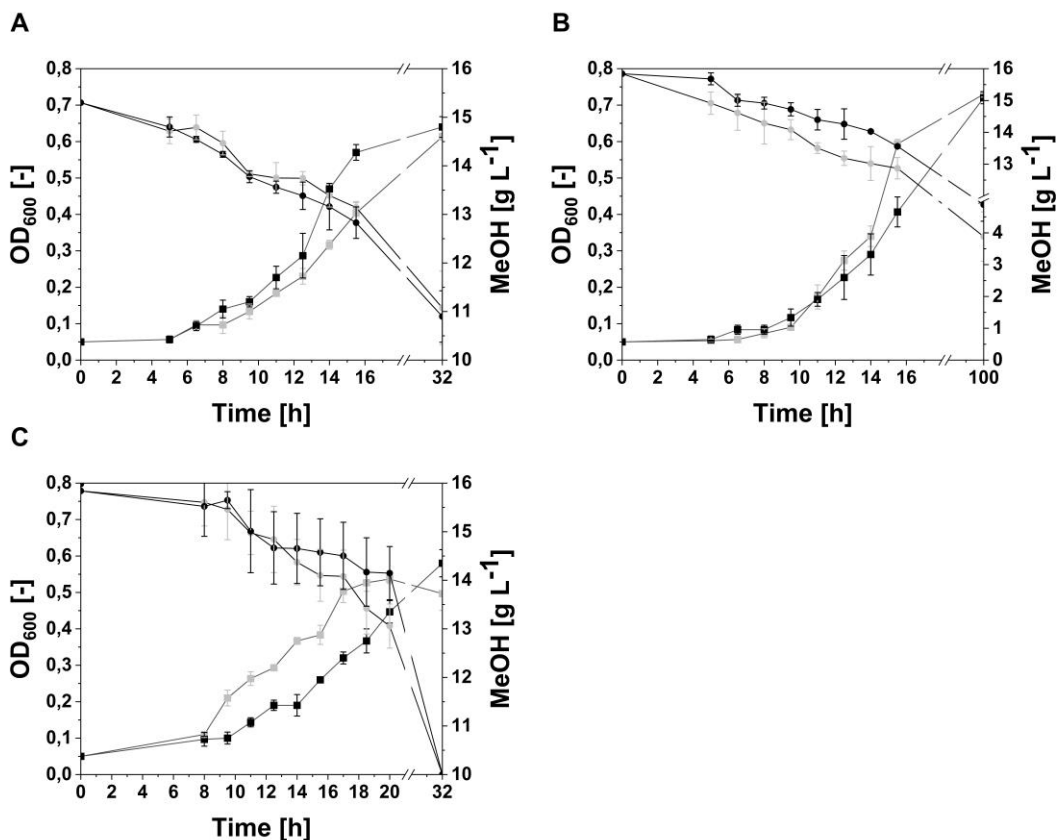
**Figure S8: Final OD<sub>600</sub> of screened *Y. lipolytica* strains on methanol.** Shown are *Y. lipolytica* DSM3286 cells expressing the RuMP pathway (A), the FLS pathway using Nourseothricin vector (B), the FLS L482Q/I557G pathway using the Nourseothricin vector (C), or both FLS pathway (wild type and variant) using the HygromycinB vector (D). Comparison of strains expressing methanol assimilation pathways (dark grey) to their corresponding empty vector (light grey). Cells were cultivated in liquid minimal medium supplemented with 15.8 g L<sup>-1</sup> methanol and 2 g L<sup>-1</sup> yeast extract. Shown are the mean values of three biological replicates. Error bars indicate standard deviations.



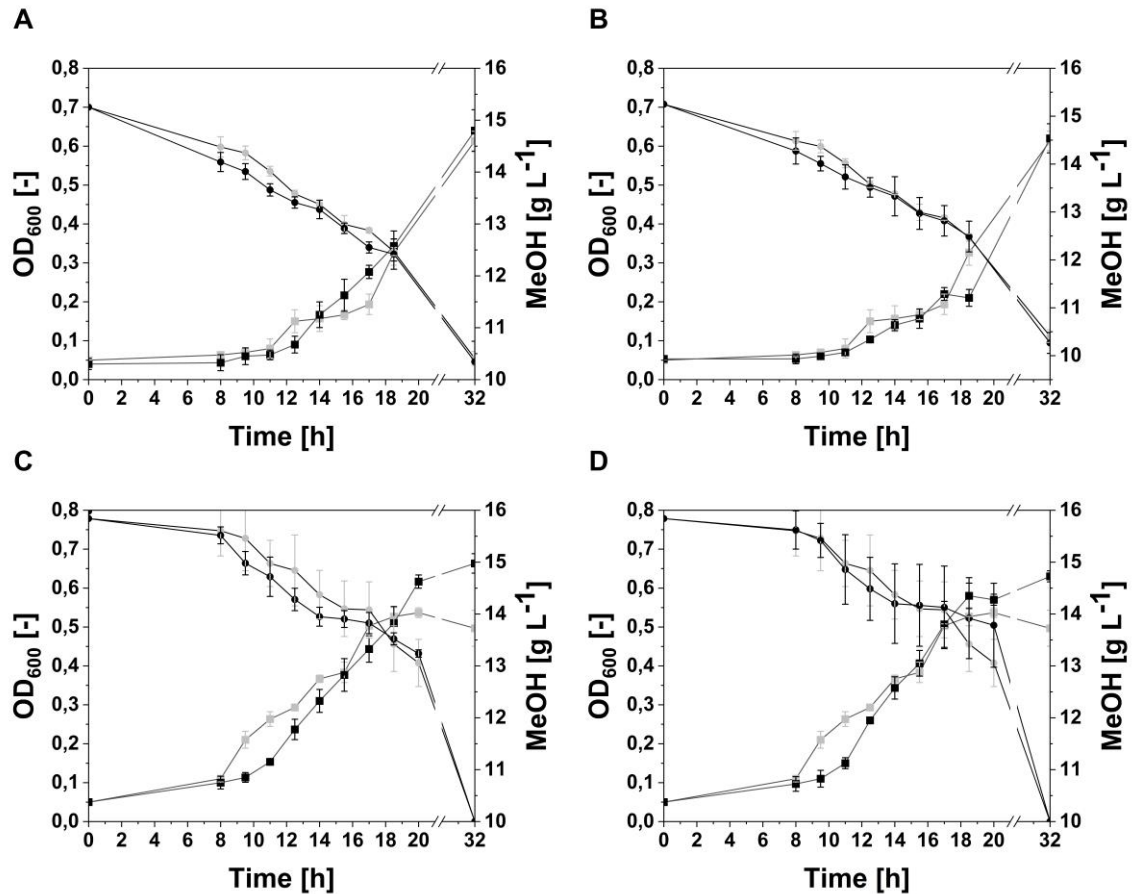
**Figure S9: Screening of various *Y. lipolytica* strains on methanol.** Shown are *Y. lipolytica* DSM3286 cells expressing the RuMP pathway (A, B), the FLS pathway (C, E) or the FLS pathway using the FLS variant L482Q/I557G (D, F). Comparison of strains expressing methanol assimilation pathways (black) to their corresponding empty vector (light grey). Cells were cultivated in liquid minimal medium supplemented with 15.8 g L<sup>-1</sup> methanol and 2 g L<sup>-1</sup> yeast extract. In total, 58 strains were screened. Shown are the mean values of three biological replicates. Error bars indicate standard deviations.



**Figure S10: Final OD<sub>600</sub> of rescreened *Y. lipolytica* strains on methanol.** Shown are *Y. lipolytica* DSM3286 cells expressing the RuMP pathway or the FLS pathway based on wild type or variant (\*) FLS. Comparison of strains expressing methanol assimilation pathways (dark grey) to their corresponding empty vector (light grey). Cells were cultivated in liquid minimal medium supplemented with 15.8 g L<sup>-1</sup> methanol and 2 g L<sup>-1</sup> yeast extract. Shown are the mean values of three biological replicates. Error bars indicate standard deviations.



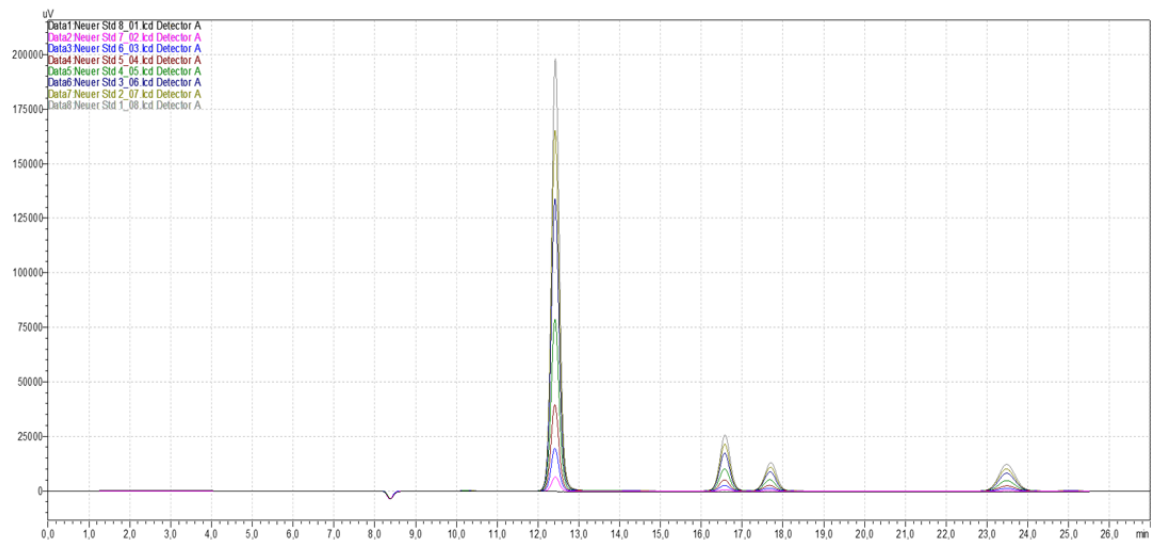
**Figure S11: Cell growth of *Y. lipolytica* strains AD\_N\_ADH\_HPS\_PHI1 (A), AC\_N\_MDH2\_FLS\*2 (B) and AC\_N\_MDH3\_FLS3 (C) in 25 mL of MM with 2% methanol and 0.5 g L<sup>-1</sup> yeast extract.** Engineered strains are shown in black, whereas the corresponding empty vector control (light grey) was used as a benchmark for the natural methanol consumption and methanol evaporation during the strain cultivation. Squares indicate OD<sub>600</sub> value, whereas circles indicate measured methanol concentration. Shown are the mean values of three independent replicates. Error bars indicate standard deviations.



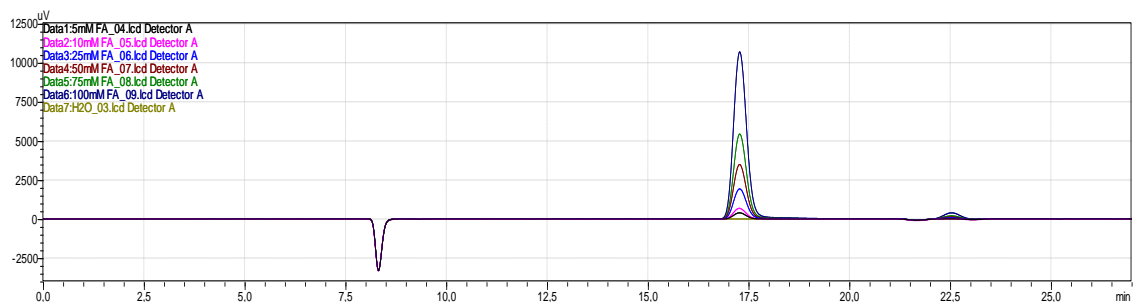
**Figure S12: Cell growth of *Y. lipolytica* strains AC\_H\_MDH2\_FLS\*3 (A), AC\_H\_MDH3\_FLS4 (B), AC\_N\_ADH\_FLS2 (C) and AC\_N\_MDH2\_FLS1 (D) in 25 mL of MM with 2% methanol and 0.5 g L<sup>-1</sup> yeast extract.** Engineered strains are shown in black, whereas the corresponding empty vector control (light grey) was used as a benchmark for the natural methanol consumption and methanol evaporation during the strain cultivation. Squares indicate OD<sub>600</sub> value, whereas circles indicate measured methanol concentration. Shown are the mean values of three independent replicates. Error bars indicate standard deviations.

## 6.2 HPLC

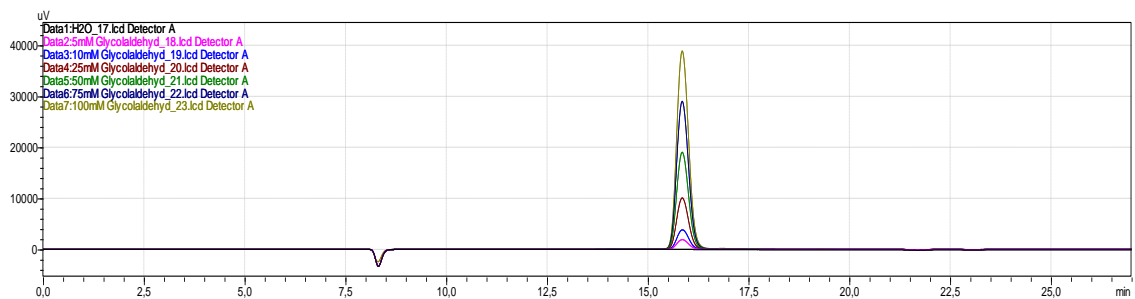
### Chromatograms of HPLC standards



**Figure S13: HPLC chromatogram of standards containing Glucose (RT 12.4 min), D-Lactate (RT 16.6 min), Formate (RT 17.7 min) and Methanol (RT 23.5 min) measured with RI detector.**

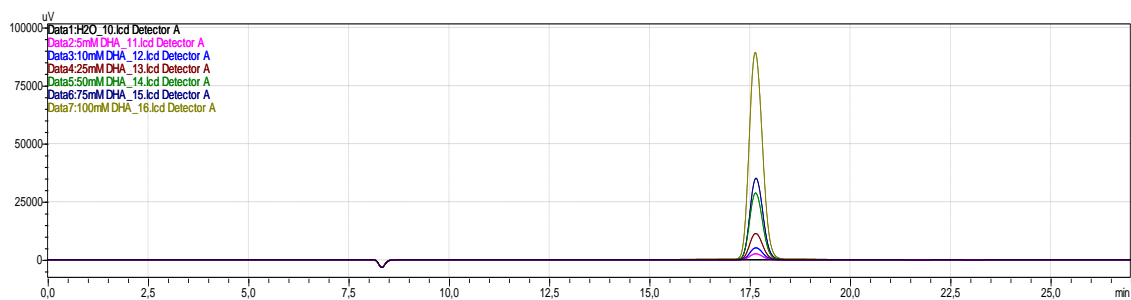


**Figure S 14: HPLC chromatogram of 5 mM - 100 mM standards of formaldehyde detected with RI detector. FALD depicted a retention time of 17.2 min.**



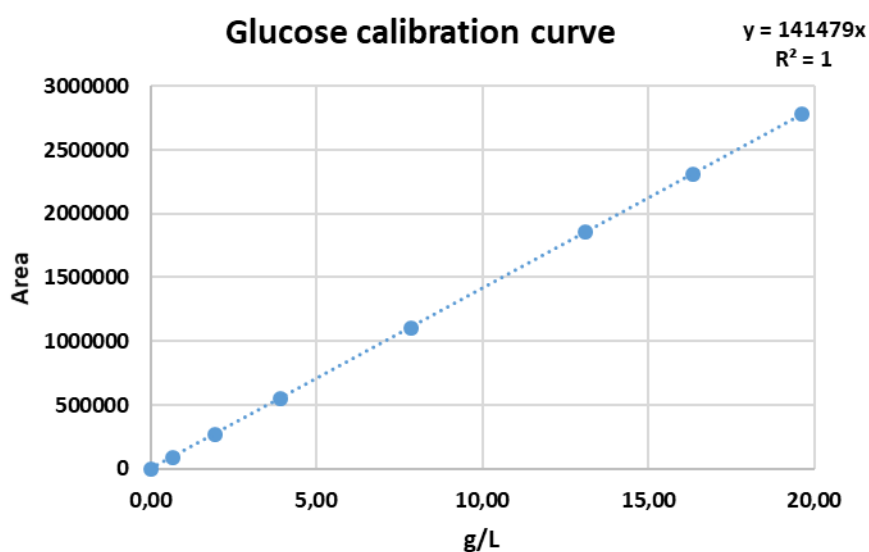
**Figure S 15: HPLC chromatogram of 5 mM - 100 mM standards of glycolaldehyde detected with RI detector. GALD depicted a retention time of 15.8 min.**



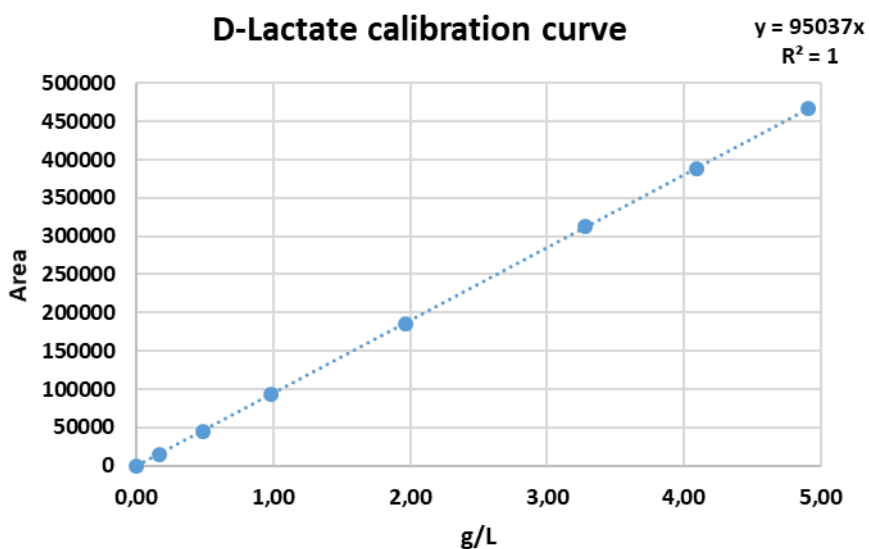


**Figure S 16: HPLC chromatogram of 5 mM - 100 mM standards of dihydroxyacetone detected with UV detector. DHA depicted a retention time of 18.8 min.**

### Calibration of HPLC standards



**Figure S 17: Measured calibration curve of glucose in a range of 0 – 20 g L<sup>-1</sup>.**



**Figure S 18: Measured calibration curve of D-lactate in a range of 0 – 5 g L<sup>-1</sup>.**

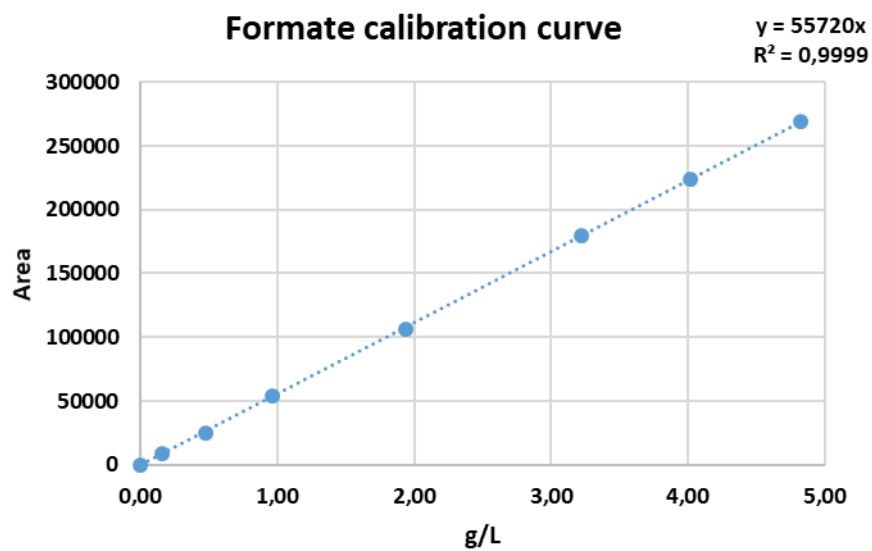


Figure S 19: Measured calibration curve of formate in a range of 0 – 5 g L<sup>-1</sup>.

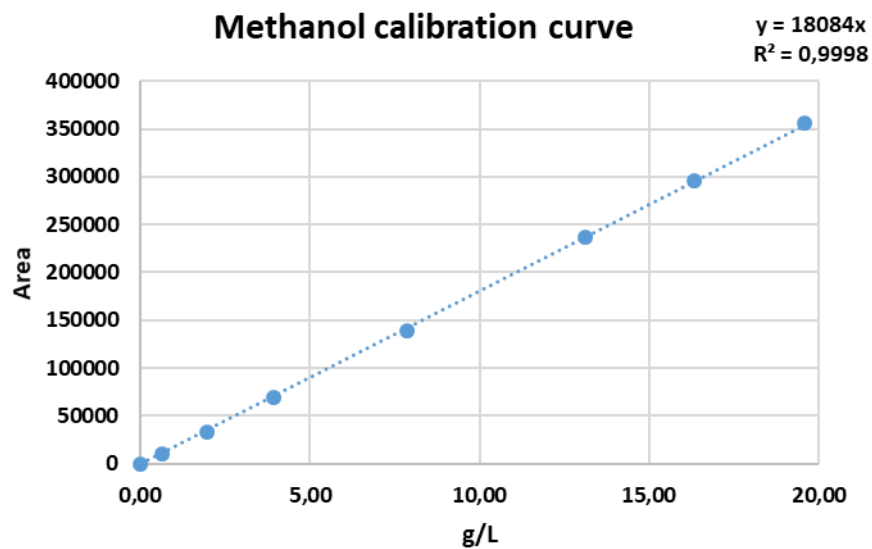


Figure S 20: Measured calibration curve of methanol in a range of 0 – 20 g L<sup>-1</sup>.

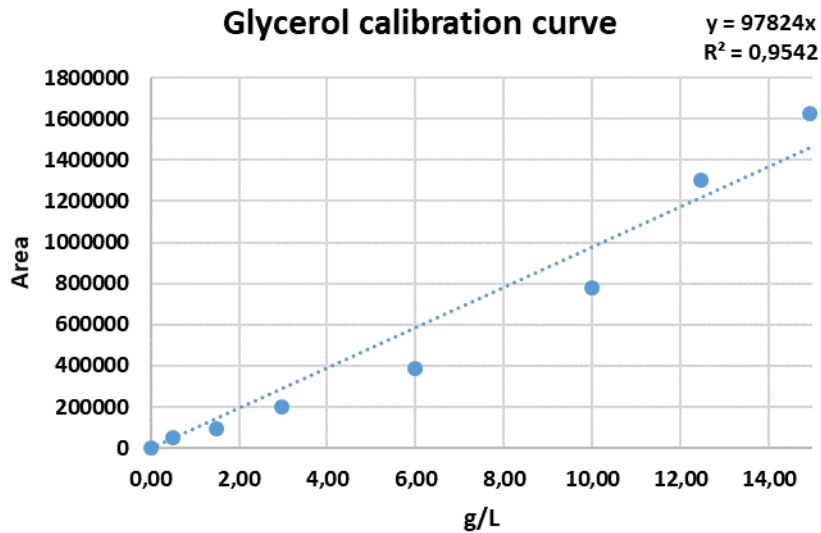


Figure S21: Measured calibration curve of methanol in a range of 0 – 15 g L<sup>-1</sup>.

### 6.3 Gene sequences

Sequencing results for double mutant FLS variants:

	1370	1380	1390	1400	1410	1420	1430	1440
FLS wild type	*	*	*	*	*	*	*	*
FLS L482H/I557L	ccattggcgaatttcgacacccctggtacgcaaacagctgccgctgattgtaattatcatgaacaaccagctcttggggctgg							
FLS L482H/I557M	ccattggcgaatttcgacacccctggtacgcaaacagctgccgctgattgtaattatcatgaacaaccagctcttggggctgg							
FLS L482H/I557Y	ccattggcgaatttcgacacccctggtacgcaaacagctgccgctgattgtaattatcatgaacaaccagctcttggggctgg							
FLS L482A/I557Q	ccattggcgaatttcgacacccctggtacgcaaacagctgccgctgattgtaattatcatgaacaaccagctcttggggctgg							
FLS L482Q/I557G	ccattggcgaatttcgacacccctggtacgcaaacagctgccgctgattgtaattatcatgaacaaccagctcttggggctgg							
	1450	1460	1470	1480	1490	1500	1510	1520
FLS wild type	*	*	*	*	*	*	*	*
FLS L482H/I557L	accctgcactttcagcagctggccgttggtcctaaccgtgtcaccggcaccgcctggaaaatggttcctatcacggcgt							
FLS L482H/I557M	accctgcactttcagcagctggccgttggtcctaaccgtgtcaccggcaccgcctggaaaatggttcctatcacggcgt							
FLS L482H/I557Y	accctgcactttcagcagctggccgttggtcctaaccgtgtcaccggcaccgcctggaaaatggttcctatcacggcgt							
FLS L482A/I557Q	accctgcactttcagcagctggccgttggtcctaaccgtgtcaccggcaccgcctggaaaatggttcctatcacggcgt							
FLS L482Q/I557G	accctgcactttcagcagctggccgttggtcctaaccgtgtcaccggcaccgcctggaaaatggttcctatcacggcgt							
	1530	1540	1550	1560	1570	1580	1590	1600
FLS wild type	*	*	*	*	*	*	*	*
FLS L482H/I557L	tgctgcgccattcgggtgctgatggttaccacgtcgactctgtcgagagcttcagcgcgctctggctcaggcactggcac							
FLS L482H/I557M	tgctgcgccattcgggtgctgatggttaccacgtcgactctgtcgagagcttcagcgcgctctggctcaggcactggcac							
FLS L482H/I557Y	tgctgcgccattcgggtgctgatggttaccacgtcgactctgtcgagagcttcagcgcgctctggctcaggcactggcac							
FLS L482A/I557Q	tgctgcgccattcgggtgctgatggttaccacgtcgactctgtcgagagcttcagcgcgctctggctcaggcactggcac							
FLS L482Q/I557G	tgctgcgccattcgggtgctgatggttaccacgtcgactctgtcgagagcttcagcgcgctctggctcaggcactggcac							
	1610	1620	1630	1640	1650	1660	1670	1680
FLS wild type	*	*	*	*	*	*	*	*
FLS L482H/I557L	acaaccgcccggcctgcatacaagttgctgtggccctggaccocgatcccgcggaggaaactgatcctgattggcatggac							
FLS L482H/I557M	acaaccgcccggcctgcatacaagttgctgtggccctggaccocgatcccgcggaggaaactgatcctgattggcatggac							
FLS L482H/I557Y	acaaccgcccggcctgcatacaagttgctgtggccctggaccocgatcccgcggaggaaactgatcctgattggcatggac							
FLS L482A/I557Q	acaaccgcccggcctgcatacaagttgctgtggccctggaccocgatcccgcggaggaaactgatcctgattggcatggac							
FLS L482Q/I557G	acaaccgcccggcctgcatacaagttgctgtggccctggaccocgatcccgcggaggaaactgatcctgattggcatggac							

Sequencing results for double mutant FLS translated into protein sequence:

	430	440	450	460	470	480	490
	*	*	*	*	*	*	*
FLS wt	mgvfgftalgaqvadleagrtilvtgdgsvgysigefdtlvrkqlpliviimnnqswgwtlhfqqlavg						
FLS L482H/I557L	mgvfgftalgaqvadleagrtilvtgdgsvgysigefdtlvrkqlpliviimnnqswgwtlhfqqlavg						
FLS L482H/I557M	mgvfgftalgaqvadleagrtilvtgdgsvgysigefdtlvrkqlpliviimnnqswgwtlhfqqlavg						
FLS L482H/I557Y	mgvfgftalgaqvadleagrtilvtgdgsvgysigefdtlvrkqlpliviimnnqswgwtlhfqqlavg						
FLS L482A/I557Q	mgvfgftalgaqvadleagrtilvtgdgsvgysigefdtlvrkqlpliviimnnqswgwtahfqqqlavg						
FLS L482Q/I557G	mgvfgftalgaqvadleagrtilvtgdgsvgysigefdtlvrkqlpliviimnnqswgwtqhfqqqlavg						
	500	510	520	530	540	550	560
	*	*	*	*	*	*	*
FLS wt	pnrvtgtrlengsyhgvaafgadyhvdsvsfsaalaqalahnrpacinvaldpippeeliligmd						
FLS L482H/I557L	pnrvtgtrlengsyhgvaafgadyhvdsvsfsaalaqalahnrpacinvaldpippeeliligmd						
FLS L482H/I557M	pnrvtgtrlengsyhgvaafgadyhvdsvsfsaalaqalahnrpacinvaldpippeeliligmd						
FLS L482H/I557Y	pnrvtgtrlengsyhgvaafgadyhvdsvsfsaalaqalahnrpacinvaldpippeeliligmd						
FLS L482A/I557Q	pnrvtgtrlengsyhgvaafgadyhvdsvsfsaalaqalahnrpacinvaldpippeeliligmd						
FLS L482Q/I557G	pnrvtgtrlengsyhgvaafgadyhvdsvsfsaalaqalahnrpacinvaldpippeeliligmd						

MDH2:

GACAGGTCTCCCATGAAGAACACCCAGTCTGCCTTCTACATGCCCTCTGTGAACCTGTT  
CGGCGCCGGTTCGTC AACGAGGTGGGTACCCGACTGGCCGGCCTGGGAGTCAAGAA  
GGCCCTGCTGGTACCGACGCCGGCCTGCATTCCCTGGGCCTGTCCGAGAAGATCGCC  
GGTATTATTCGAGAGGCCGGTGTGCGAGGTGCCATCTTCCCCAAGGCCGAGCCCAACC  
CCACCGACAAGAACGTGCGCCGAGGGCCTGGAGGCCTACAACGCCGAGA ACTGTGACT  
CTATTGTCACCCTGGGCGGCGGTTCTCCACGACGCTGGTAAGGCCATTGCCCTGGTC  
GCCGCCAACGGTGGTACCATCCACGACTACGAGGGTGTGGACGTCTCCAAGAAGCCC  
ATGGTCCCCCTGATCGCCATCAACACCACCGCCGGCACC GGTTCCGAGCTGACCAAGT  
TCACCATTATCACCGACACCGAGCGAAAGGTCAAGATGGCCATCGTGGACAAGCACG  
TGACCCCCACCCTGTCTATTAACGACCCCGAGCTGATGGTGGCATGCCCCCTCCCTG  
ACCGCCGCTACTGGTCTGGACGCCCTGACCCACGCCATCGAGGCCTACGTGTCCACCG  
GCGCCACCCCTATTACCGACGCCCTGGCCATTCAGGCCATTAAGATTATCTCCAAGTA  
CCTGCCCCGAGCCGTCGCCAACGGTAAGGACATTGAGGCCCAGAGCAGATGGCCTTC  
GCCCAGTCCCTGGCCGGCATGGCCTTCAACAACGCCGGTCTGGGCTACGTCCACGCCA  
TTGCCACCAGCTGGGCGGCTTCTACAACCTTCCCCACGGTGTGTGTAACGCCATCCTG  
CTGCCCCACGTGTGCCGATTCAACCTGATCTCTAAGGTGGAGCGATACGCCGAGATCG  
CCGCTTCCCTGGGTGAGAACGTGGACGGCCTGTCTACCTACGAGGCCGCGGAGAAGGC  
CATTAAAGGCCATTGAGCGAATGGCCCGAGACTTGAACATTCCCAAGGGCTTCAAGGAG  
CTGGGTGCCAAGGAGGAGGACATTGAGACTCTGGCCAAGAACGCCATGAACGACGCC  
TGCGCCCTGACCAACCCCGAAAGCCCAAGCTGGAGGAGGTCATTCAGATCATTAAGA  
ACGCCATGTAAGCTTGGAGACCGATC

MDH3:

GACAGGTCTCCCATGACCAACACCCAGTCTGCCTTCTTCATGCCCTCCGTGAACCTGTT  
CGGTGCCGGTCTGTCAACGAGGTGGGCACCCGACTGGCCGACCTGGGTGTCAAGAA  
GGCCCTGCTGGTGACCGACGCCGGCCTGCATGGTCTGGGTCTGTCCGAGAAGATCTCT  
TCCATTATTCGAGCCGCCGGTGTGCGAGGTGAGCATCTTCCCCAAGGCCGAGCCCAACC  
CCACCGACAAGAACGTGGCCGAGGGCCTGGAGGCCTACAACGCCGAGA ACTGTGACT  
CCATTGTCACCCTGGGTGGTGGTTCCTCTCACGACGCCGGTAAGGCCATCGCCCTGGTC  
GCCGCTAACGGCGGAAAGATCCACGACTACGAGGGCGTGCACGTCTCTAAGGAGCCC  
ATGGTGGCCCTGATCGCCATCAACACCACCGCCGGCACC GGTTCCGAGCTGACCAAGT  
TCACCATTATTACCGACACCGAGCGAAAGGTGAAGATGGCCATCGTCGACAAGCACGT  
GACCCCCACCCTGTCTATTAACGACCCCGAGCTGATGGTGGGTATGCCCCCTCTCTGA  
CCGCCGCCACCGACTTGACGCCCTGACCCACGCCATCGAGGCCTACGTGTCTACCGG

CGCCACCCCATCACCGACGCCCTTGCTATTCAGGCCATCAAGATCATTAGCAAGTAC  
CTGCCCCGAGCCGTGGCCAACGGCAAGGACATCGAGGCCCGAGAGCAGATGGCCTTC  
GCCAGTCTCTGGCCGGCATGGCCTTCAACAACGCCGGCCTGGGCTACGTCCACGCCA  
TCGCTCACAGCTGGGCGGCTTCTACAACCTCCCCACGGTGTCTGTAACGCCGTCTG  
CTGCCCTACGTGTGTCGATTCAACCTGATTAGCAAGGTCGAGCGATAACGCCGAGATCG  
CCGCTTCTGGGTGAGAACGTGGACGGTCTGTCCACCTACGACGCCGCCGAGAAGGC  
CATTAAGGCCATCGAGCGAATGGCCAAGGACCTGAACATTCCCAAGGGCTTCAAGGA  
GCTGGGTGCCAAGGAGGAGGACATCGAGACTCTGGCCAAGAACGCCATGAAGGACGC  
CTGCGCCCTGACCAACCCCCGAAAGCCCAAGCTGGAGGAGGTCATCCAGATTATCAAG  
AACGCCATGTAAGCTTGGAGACCGATC

HPS:

GACAGGTCTCCCATGGAGCTGCAGCTGGCCCTGGACCTGGTGAACATCGAAGAAGCC  
AAGCAGGTCGTCGCCGAGGTGCAGGAATACGTTGACATCGTCGAGATCGGCACCCCC  
GTGATCAAGATTTGGGGACTGCAGGCCGTGAAGGCCGTCAAGGACGCATTCCCCCACC  
TCCAGGTGCTGGCCGACATGAAAACCATGGACGCCGCTGCCTACGAGGTGGCTAAGG  
CCGCCGAGCACGGAGCCGACATCGTGACCATCCTGGCCGCTGCCGAGGACGTGTCTAT  
CAAGGGCGCCGTCGAAGAGGCTAAGAAGCTGGGCAAGAAGATCCTGGTGGACATGAT  
CGCCGTGAAGAACCTGGAGGAACGAGCCAAGCAGGTTGACGAGATGGGCGTGGACTA  
CATCTGCGTCCACGCCGGCTACGACCTCCAGGCCGTGGCAAGAACCCCTGGACGAC  
CTGAAGCGAATCAAGGCCGTGGTGAAGAACGCCAAGACCGCCATTGCCGGCGGAATC  
AAGCTGGAGACTCTGCCCGAGGTGATCAAGGCCGAGCCCGACCTGGTGTATCGTGGGC  
GGAGGAATCGCCAACCAGACCGACAAGAAGGCTGCCGCCGAGAAGATCAACAAGCTG  
GTGAAGCAGGGCCTGTAAGCTTGGAGACCGATC

PHI:

GACAGGTCTCCCATGATCTCGATGCTGACCACCGAGTTTCTGGCCGAGATCGTGAAGG  
AGCTGAACCTTTCTGTCAACCAGATCGCCGACGAGGAGGCCGAGGCCCTGGTCAACGG  
CATCCTGCAGTCTAAGAAGGTGTTCTGTCGCTGGCGCCGGACGATCTGGCTTCATGGCC  
AAGTCTTTCGCCATGCGAATGATGCACATGGGCATCGACGCCTACGTGGTGGGCGAGA  
CTGTGACCCCAACTACGAGAAGGAGGACATCCTGATCATCGGCTCTGGCTCTGGCGA  
GACTAAGTCTCTGGTGTCTATGGCCCAGAAGGCCAAGTCTATCGGCGGCACCATTGCC  
GCCGTGACCATCAACCCCGAGTCTACCATCGGACAGCTGGCCGACATTGTGATCAAGA  
TGCCCGGCTCTCCAAGGACAAGTCTGAGGCCCGAGAGACTATCCAGCCCATGGGATC  
TCTGTTTCGAGCAGACCCTGCTGCTGTTCTACGACGCCGTGATCCTGCGATTATGGAGA  
AGAAGGGCCTGGACACCAAGACCATGTACGGCCGACACGCCAACCTGGAGTAGGCTT  
GGAGACCGATC

ADH:

GACAGGTCTCCCATGAAGGCCGCCGTGCTGGAGCAGTTCAAGGAGCCCCTGAAGATTA  
AGGAGGTGGAGAAGCCCACCATCTCTTACGGTGAGGTCCTGGTGCGAATCAAGGCCTG  
TGGTGTCTGTCACACCGACCTGCACGCCGCCACGGCGATTGGCCTGTGAAGCCCAAG  
CTGCCCCGATTCCCGGTCACGAGGGCGTGGGCATTGTTCGAGGAGGTCGGTCCCGGTG  
TCACCCACCTGAAGGTGGGCGACCGAGTGGGCATTCCCTGGCTGTACTCTGCCTGCGG  
TCACTGCGACTACTGCCTGTCTGGCCAGGAGACTCTGTGCGAGCACCAGAAGAACGCC  
GGTACTCCGTGGACGGCGGTTACGCCGAGTACTGTGCGAGCCGCCGCCGACTACGTG  
TGAAGATCCCCGACAACCTGTCCTTCGAGGAGGCCGCCCCATCTTCTGCGCCGGAGT  
CACCACCTACAAGGCCCTGAAGGTCACCGGTGCCAAGCCCAGGTGAGTGGGTGGCCATC  
TACGGCATTGGCGGCCTGGGCCACGTGGCCGTTTCAAGTACGCCAAGGCCATGGGCCTGA  
ACGTCGTGGCCGTCGACATTGGTGCAGAGAAGCTGGAGCTGGCCAAGGAGCTGGGTG  
CCGACCTGGTCTGTAACCCCTGAAGGAGGACGCCGCCAAGTTCATGAAGGAGAAGG  
TGGGCGGCGTGCACGCCGCCGTTGTCACTGCTGTGTCCAAGCCCGCCTTCCAGTCTGCC  
TACAACCTCCATTCGACGAGGCGGTGCCTGCGTCTGGTCCGGTCTGCCTCCCGAGGAGA  
TGCCCATTCATCTTCGACACCGTCTGAACGGCATCAAGATTATTGGCTCCATTGTC  
GGCACCCGAAAGGACCTGCAGGAGGCCCTGCAGTTCGCCGCCGAGGGTAAGGTCAAG

ACCATCATTGAGGTCCAGCCCCTGGAGAAGATTAACGAGGTGTTGACCGAATGCTGA  
AGGGTCAGATTAACGGCCGAGTCGTCTGACCCTGGAGGACAAGTAAGCTTGGAGAC  
CGATC

FLS-oHis:

GACAGGTCTCCCATGGCCATGATCACCGGTGGCGAGCTGGTGGTGCGAACCCTGATCA  
AGGCCGGCGTGGAGCACCTGTTCCGGTCTGCACGGTATCCACATTGACACCATCTTCCA  
GGCCTGCCTGGACCACGACGTCCCCATTATCGACACCCGACACGAGGCCGCCGCCGA  
CATGCTGCTGAGGGATACGCCCCGAGCCGGTGCCAAGCTGGGCGTGGCTCTGGTGACCG  
CCGGTGGAGGCTTCACCAACGCCGTACCCCCATCGCCAACGCCCGAACCAGCCGAAC  
CCCCGTGCTGTTCCCTGACCGGTTCCGGTGGCCCTGCGAGACGACGAGACTAACACCCTG  
CAGGCCGGCATCGACCAGGTGGCCATGGCCGCTCCCATTACCAAGTGGGCCACCCGAG  
TGATGGCCACCGAGCACATCCCCGACTGGTCATGCAGGCCATCCGAGCCGCCCTGTG  
CGCCCCCTCGAGGTCCTGTTCTGCTGGACCTGCCCTGGGACATCCTGATGAACCAGATC  
GACGAGGACTCTGTGATCATTCCCGACCTGGTGCTGTCTGCCACGGCGCTCACCCCG  
ACCCTGCTGATCTGGACCAGGCCCTGGCCCTGCTGCGAAAGGCCGAGCGACCCGTGAT  
CGTCTGGGTTCTGAGGCCTCCCGAACCGCCCGAAAGACCGCCCTGTCTGCCTTCGTG  
GCCGCCACCGGTGTCCCCGTGTTGCTGACTACGAGGGTCTGTCCATGCTGTCTGGTCT  
GCCCGACGCCATGCGAGGCGGCCTTGTGCAGAACCTGTACTCCTTCGCCAAGGCCGAC  
GCCGCCCCCGATCTGGTTCTGATGCTGGGCGCCCGATTCCGGTCTGAACACCGGTACG  
GTTCCGGTCAGCTGATCCCCACTCTGCCAGGTGATCCAGGTCGACCCCGACGCTG  
TGAGCTGGGCCGACTGCAGGGTATCGCCCTGGGTATTGTGGCCGACGTCGGCGGCACC  
ATTGAGGCCCTGGCCCAGGCTACCGCCAGGATGCTGCCTGGCCCCGACCGAGGTGACT  
GGTGCGCTAAGGTCACCGACCTGGCCCAGGAGCGATACGCCTCCATCGCCGCCAAGTC  
TAGCTCTGAGCACGCCCTGCACCCCTTCCACGCCTCTCAGGTGATCGCCAAGCACGTG  
GACGCCGGTGTACCGTGGTGGCCGACGGTGGTCTGACCTACCTGTGGCTGTCTGAGG  
TCATGTCTCGAGTCAAGCCCGGTGGCTTCCCTGTGCCACGGCTACCTGAACTCTATGGGT  
GTGGGCTTCGGCACCGCCCTGGGTGCTCAGGTGGCCGATCTGGAGGCCGGTCGACGAA  
CCATTCTGGTCACCGGCGACGGCTCCGTCGGTTACTCCATTGGCGAGTTCGACACCCTG  
GTCCGAAAGCAGCTGCCCCTGATCGTCATTATTATGAACAACCAGTCTTGGGGCTGGA  
CCCTGCACTTCCAGCAGCTGGCCGTGGGTCCCAACCGAGTCACCGGCACCCGACTGGA  
GAACGGTTCCTACCACGGTGTGGCCGCCGCCTTCGGCGCTGATGGATAACACGTGGAC  
TCTGTGGAGTCCTTCTCCGCCGCCCTGGCCCAGGCCCTTGCTCATAACCGACCCGCCTG  
CATTACGTGGCCGTCGCCCTGGACCCCATCCCCCTGAGGAGCTGATTCTGATTGGC  
ATGGACCCCTTCGCCGGCTCTACCGAGAACCTGTACTTCCAGTCTGGCGCCTAAGCTTG  
GAGACCGATC