Construction of a cell-free enzymatic pathway for the production of lactic acid from spent coffee grounds

by

Dominik Kopp

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Supervisors:

A/Prof. Anwar Sunna

Prof. Robert Willows

Department of Molecular Sciences

Macquarie University

Sydney, Australia

December 2018

Table of contents

Table of contents ii

Abstract iii

Statement of originality v

Acknowledgements vi

Presentations and awards viii

Poster presentations viii

Oral presentations viii

Outreach ix

Awards ix

List of publications x

List of abbreviations xi

Chapter 1 Alternative carbohydrate pathways - enzymes, functions and engineering **13**

1.1 Introduction 14

1.2 Manuscript 1 15

1.3 Aims and scope of this thesis 86

Chapter 2 Characterisation of a mannonate dehydratase from *Thermoplasma acidophilum* and its potential role in the catabolism of D-mannose **88**

2.1 Introduction 89

2.2 Contribution to manuscript 2 90

2.3 Manuscript 2 91

Chapter 3 Cell-free enzymatic conversion of spent coffee grounds into the platform chemical lactic acid **116**

3.1 Introduction 117

3.2 Contribution to manuscript 3 118

3.3 Manuscript 3 119

Chapter 4 Summary and future perspectives **154**

Appendix I Tools and strategies for the construction of cell-free pathways **166**

Appendix II Characterisation of the first archaeal mannonate dehydratase from *Thermoplasma acidophilum* and its potential role in the catabolism of D-mannose **184**

Appendix III Cloning, expression and purification of ManD in Saccharomyces cerevisiae **200**

Abstract

Microorganisms have been investigated extensively for their use in bio-based production of compounds for the chemical, pharmaceutical and food industry. New tools used in synthetic biology enable a fundamental rewiring of metabolic pathways within microbes. The engineering of cellular metabolisms aims to maximise production efficiencies for the sustainable, yet economically competitive manufacturing of chemicals from renewable biomass. Despite the extensive development in genetic tools, *in vivo* approaches face challenges such as interference with the host metabolism, restricted flexibility in metabolic design and limited controllability of pathway flux. Alternatively, cell-free biocatalysis has emerged as a promising technology, which is based on an almost unrestricted assembly of enzymes into synthetic and highly modular production pathways. The technology not only offers flexibility to select and combine virtually any enzyme, but also allows for rapid prototyping and testing of *de novo* assembled pathways. To date, cell-free biocatalysis has been mostly used for the production of chemicals from refined substrates such as glucose and other carbohydrates. Since the economic viability of biotechnological production processes depends heavily on substrate cost, the use of highly abundant and low-cost biomass such as waste products improves the cost-efficiency and the environmental benefit of cell-free biocatalysis.

In this work a novel cell-free enzymatic pathway was constructed for the conversion of carbohydrates obtained from spent coffee grounds into lactic acid. Lactic acid is a versatile chemical with a wide range of different applications in medical, textile and food industries and has received most of its attention for the use in form of the biodegradable polymer polylactic acid. Spent coffee grounds are a highly abundant and industrially underutilised waste compound rich in carbohydrates, especially mannose.

The catabolism of mannose in most microorganisms relies on the Embden-Meyerhof-Parnas pathway, which is high in enzyme and cofactor cost and therefore suboptimal for a synthetic cell-free approach. The new cell-free pathway described here is based on a non-phosphorylative, putative mannose metabolism from the thermophilic archaeon *Thermoplasma acidophilum*. Initial identification and characterisation of a thermostable mannonate dehydratase from this organism allowed the construction of a four-enzyme pathway for the conversion of mannose into lactic acid. All the enzymes in the pathway were derived from thermophilic organisms and were recombinantly-expressed, purified and assembled into one-pot reactions. Several reaction conditions (e.g. substrate, cofactor and enzyme concentrations) were studied and optimised

towards a conversion yield of up to 71.5 %. In order to demonstrate the feasibility of renewable substrates for cell-free biocatalysis, mannose was obtained by dilute acid hydrolysis from spent coffee grounds and converted into lactic acid using the novel pathway. The conversion from spent coffee grounds into lactic acid was further analysed via high-performance liquid chromatography and showed similar yields as those obtained from pure mannose. Complete identification of reaction intermediates and analysis of pathway flux was possible in reactions containing ^{13}C-labelled substrate via nuclear magnetic resonance spectroscopy. Identification of unspecific side reactions explained the loss of carbon and the conversion below 100% of theoretical yield. This work demonstrates the power of cell-free pathway engineering by construction of an alternative and so far undescribed carbohydrate pathway to convert waste biomass into valuable bio-based chemicals in a sustainable way.

Statement of originality

This thesis is the result of my own work and includes nothing, that is the outcome of work done in collaboration except for where specifically indicated in the text. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself. This work has not been submitted for a higher degree or diploma to any other university or institution.

Dominik Kopp

December 2018

Acknowledgements

This thesis comprises three years of full-time research conducted in the Department of Molecular Sciences at Macquarie University from December 2015 to December 2018.

First of all, I would like to thank my supervisors A/Prof. Anwar Sunna and Professor Robert Willows for the immense support during my time as a PhD candidate.
Anwar, I highly appreciate your great commitment and enthusiasm for the project and honour your style of guidance in creating space for creativity and self-development. I highly value your patience in reading and correcting the fourth PhD thesis in a single year and I am very grateful for pushing me towards new experiences, whether it was the work with public relations or applying for multiple conferences.
Robert Willows, I am very grateful for your invaluable scientific advice in many research questions, whether it concerned analytics, enzyme kinetics or biochemistry. My thesis would not have gotten this far without your help.

I also would like to thank Dr. Andrew Care who has taught me very valuable lessons scientifically and personally. Thank you for always putting up a positive and pro-active attitude and helping me to put a smile back on my face.
I also like to thank Emeritus Professor Peter L. Bergquist for his critical feedback on my project and improving my English writing skills.
Thanks also to Professor Helena Nevalainen for her support especially during the first time of my PhD and fruitful discussions in her group meetings.

I feel very thankful to my peers especially Alex for his courage to take on the PhD as the first in the group and leading the way. Also for all the inspiring conversations and being a great friend. Kerstin, for her massive support, countless discussions and strong commitment to finally master the art of high-performance liquid chromatography.

Thanks to all other amazing peers in the lab/office and your great company, Dennis, Sandra, Rachit and Vinoth. Thanks also to Monica, Alex, Tom, Xin, Niel, Heinrich, Hugh, Briardo, Roy, Athul, Ian, Liisa, Elsa, Louise, Elizabeth, Elisabeth, Bishal, Matt, Wei, Sameera, Wisam, Wenwen, Areti, Nicole and all other people that I have worked and become friends with over the three years for your support and encouragement.

I also want to thank the people in my personal life. Great thanks goes to my friend Miriam for sharing a memorable time during our studies and the encouragement to pursue this career.

Finally, my greatest thanks and love goes to my family, who has always been there for me and supported me in every way they possibly could.

Presentations and awards

Poster presentations

Kopp, D., Care, A., Bergquist, P.L., Willows, R., Sunna, A. (2018) Cell-free synthetic pathway for the conversion of spent coffee grounds into lactic acid. Poster presented at the 12th Metabolic Engineering Conference, Munich, Germany and at the European Congress on Biotechnology, Geneva, Switzerland.

Kopp, D., Care, A., Bergquist, P.L., Willows, R.D., Sunna, A. (2018) Cell-free synthetic pathway for the conversion of coffee waste into platform chemicals. Poster presented at the Bionetwork Symposium, Sydney, Australia.

Bergquist, P.L., Petroll, K., **Kopp, D.**, Care, A., Sunna, A. (2018). Synthetic biocatalytic modules for enhanced transformation of biological waste products. Poster presented at the ECI Microbial Engineering Conference, New Mexico, USA.

Kopp, D., Petroll, K.**,** Care, A., Bergquist, P.L, Sunna, A. (2017). Cell-free biocatalytic modules for biological waste conversion. Poster presented at the 7th International meeting on Synthetic Biology, Singapore.

Kopp, D., Petroll, K., Care, A., Gibson, E. SY., Bergquist, P.L., Sunna, A. (2016). Cell-free biocatalytic modules for the conversion of organic waste into platform chemicals. Poster presented at the Synthetic Biology Cutting Edge Symposium, Canberra, Australia.

Petroll, K.**, Kopp, D.**, Care, A., Bergquist, P.L., Sunna, A. (2016). Cell-free biocatalytic modules for biological waste conversion. Poster presented at 11th International Congress on Extremophiles, Kyoto, Japan.

Oral presentations

Kopp D., Care, A., Bergquist, P.L., Willows, R.D., Sunna, A. (2018). Cell-free bioconversion of spent coffee grounds into lactic acid. Rapid Fire poster presentation at the 12th Metabolic Engineering Conference in Munich, Germany.

Outreach

- Radio interviews for 2ser, ABC Hobart, Triple M Melbourne, Triple J Sydney, 3AW News, 4BC drive Brisbane
- News articles featured in/on: ABC news, New Atlas, waste management review, biofuels digest, The Lighthouse - Macquarie University, trends der zukunft, bioeconomia.com.ar, technews.tw, bastillepost

Awards

2018 **INOFEA Early Career Award for Applied Biocatalysis and Nanobiotechnology,** for the best poster presentation at the European Congress on Biotechnology in Geneva. Value: 250 Euro.

Macquarie University Postgraduate Research Fund (PGRF),
travel grant awarded for the attendance of the 12th Metabolic Engineering conference in Munich, the European Congress on Biotechnology in Geneva and several lab visits in Germany. Value: 4650 AUD.

ME-12 Junior Researcher Grant,
for the attendance of the 12th Metabolic Engineering conference in Munich. Value: 800 AUD.

Best poster at the BioNetwork Symposium at Macquarie University, Sydney
Value: 100 AUD.

2015 **Three-year international Macquarie University Research Excellence Scholarship,** (iMQRES). Approx. value: 200,000 AUD.

List of publications

This thesis includes two manuscripts, which were prepared for submission (manuscript 1 and 3, see chapter 1 and 3 respectively), one published article (manuscript 2, see chapter 2, Appendix II) and one co-authored article (manuscript 4, Appendix I).

1) **Kopp, D.**, Sunna, A. Alternative carbohydrate pathways - enzymes, functions and engineering. Prepared for submission to *Critical Reviews in Biotechnology.*

2) **Kopp, D.**; Willows, R.; Sunna, A. Characterisation of the First Archaeal Mannonate Dehydratase from Thermoplasma acidophilum and Its Potential Role in the Catabolism of D-Mannose. *Catalysts* 2019, 9, 234 (https://doi.org/10.3390/catal9030234)

3) **Kopp, D.**, Willows, R.D., Sunna, A. Cell-free enzymatic conversion for of spent coffee grounds into the platform chemical lactic acid. Prepared for submission to *Metabolic Engineering.*

4) Petroll, K., **Kopp, D.**, Care, A., Bergquist, P.L., Sunna, A. Tools and strategies for constructing cell-free enzyme pathways. Published article in *Biotechnology Advances*.

List of abbreviations

ATP	adenosine triphosphate
^{13}C	carbon-13 isotope
A	absorbance
AldT	aldohexose dehydrogenase
BCA	bicinchoninic acid assay
CFPS	cell-free protein synthesis
coA	coenzyme A
D_2O	deuterium oxide
DAD	diode array detector
DHAD	dihydroxyacid dehydratase
DNA	deoxyribonucleic acid
ED	Entner-Doudoroff
EMP	Embden-Meyerhof-Parnas
FPLC	fast protein liquid chromatography
g	gravitational force
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	histidine
HPLC	high performance liquid chromatography
IlvD/EDD	dihydroxyacid dehydratase/6-phosphogluconate dehydratase
IPTG	isopropyl β-D-1-thiogalactopyranoside
kDa	kilo Dalton
KDG	2-keto-3-deoxygluconate
KDGA	2-keto-3-deoxygluconate aldolase
K_m	Michaelis-Menten constant
LA	lactic acid
LB	lysogeny broth
LC ESI MS/MS	liquid chromatography electrospray ionization tandem mass spectrometry
LDH	L-lactate dehydroegnase
ManD	mannonate dehydratase
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
NaP	sodium phosphate
NCBI	National Center for Biotechnology Information

Ni-NTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
np-ED	non-phosphorylative Entner-Doudoroff
(n)RIU	(nano) refractive index unit
OD	optical density
PCR	polymerase chain reaction
PLA	polylactic acid
ppm	parts per million
RID	refractive index detector
RNA	ribonucleic acid
SC	synthetic complete
SC-URA	synthetic complete without uracil
SCG	spent coffee grounds
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SRS	sugar reducing standards
TBA	thiobarbituric acid
TCA	trichloracetic acid
TEV	tobacco etch virus
TMSP	3-(trimethylsilyl)propionic-2,2,3,3-d4 acid
Tris/HCl	tris(hydroxymethyl)aminomethane
Tyr	tyrosine
tRNA	transfer ribonucleic acid
U	unit
UV/Vis	ultraviolet-visible
UxuA	mannonate dehydratase (hexuronate metabolism)
V_{max}	maximal velocity
ε	molar extinction coefficient

Chapter 1

Alternative carbohydrate pathways - enzymes, functions and engineering

1.1 Introduction

The recent desire to create a bio-based economy for the production of fuels and chemical has led to the development of innovative bioprocesses using renewable resources and engineered microbial production strains. Metabolic engineering of microorganisms used for the production of chemicals builds on fundamental knowledge of metabolic pathways and their key enzymes. In the following review, the diversity of carbohydrate pathways for a multitude of different carbohydrates other than glucose is summarised. In particular key principles and enzymes of alternative pathways to the well-characterised Embden-Meyerhof-Parnas (EMP) pathway are presented. Exploring the landscape of non-phosphorylative pathways for this review eventually led to the design of the cell-free pathway for the production of lactic acid from mannose.

This chapter has been prepared as a manuscript for submission to *Critical Reviews in Biotechnology*.

1.2 Manuscript 1

Alternative carbohydrate pathways - enzymes, functions and engineering

Dominik Kopp[a] and Anwar Sunna[a,b]*

[a]Department of Molecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia

[b]Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, New South Wales 2109, Australia

*Corresponding author

Anwar Sunna

Department of Molecular Sciences

Macquarie University

North Ryde, NSW 2109 Australia

Phone: +61 2 9850 4220

Fax: +61 2 9850 8313

Email: anwar.sunna@mq.edu.au

Abstract

The Emden-Meyerhof-Parnas (EMP) pathway is regarded as the most important example of a carbohydrate pathway due to its extensive research in model organisms like *Escherichia coli*, *Saccharomyces cerevisiae* and *Bacillus subtilis*. However, these organisms do not represent the abundance of carbohydrate pathways present in all microorganisms. In fact, studies in a variety of different bacteria and archaea show that the EMP pathway is either modified from its classical example, complemented or even replaced by alternative carbohydrate pathways for example the Entner-Doudoroff (ED) pathway. Besides its prevalence in archaea, the ED pathway has been found to play important roles in soil and gut bacteria, biotechnological production strains and marine microorganisms. The ED pathway and particular its variations differ in key principles such as thermodynamics, intermediates and key enzymes from the EMP pathway. An oxidative principle, in variations of the ED pathway rather than phosphorylative strategy, found in the EMP pathway, is a reoccurring pattern in pathways for sugars such as D-galactose, D-xylose, L-arabinose, L-rhamnose and L-fucose. Increasing interest in the metabolic engineering of microorganisms for the production of sustainable chemicals from biomass builds on the fundamental knowledge of the diversity and complexity of metabolisms and enzymes. In this review we present and discuss alternative pathway strategies for *in vivo* and cell-free metabolic engineering.

Table 1 Abbreviations that are used throughout the main text

List of abbreviations	
3PG	3-phosphoglycerate
6PG	6-phosphogluconate
ABC	ATP binding cassette
CCM	central carbon metabolism
DHAD	dihydroxyacid dehydratase
DHAP	dihydroacetone phosphate
ED	Entner-Doudoroff
EDA	2-keto-3-deoxy-phosphogluconate aldolase
EDD	6-phosphogluconate dehydratase
EMP	Embden-Meyerhof-Parnas
ENO	enolase
ExuT	hexuronate transporter
F6P	fructose-6-phosphate
FBA	fructose-1,6-bisphosphate aldolase
FBP	fructose-1,6-bisphosphate
GAD	gluconate dehydratase
GAP	glyceraldehyde-3-phosphate
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GAPN	non-phosphorylative glyceraldehyde-3-phosphate dehydrogenase
GAPOR	ferredoxin dependent glyceraldehyde-3-phosphate oxidoreductase
GFOR	glucose-fructose oxidoreductase
GK	glucokinase
GLK	glycerate kinase
GNDH	gluconate-2-dehydrogenase
GNUK	gluconate kinase
IlvD/EDD	dihydroxyacid dehydratases/6-phosphogluconate dehydratases
KDF	2-keto-3-deoxy-L-fuconate
KDG	2-keto-3-deoxygluconate
KDGA	2-keto-3-deoxygluconate aldolase
KDGal	2-keto-3-deoxygalactonate
KDGK	2-keto-3-deoxygluconate kinase
KDPG	2-keto-3-deoxy-6-phosphogluconate
KDX	2-keto-3-deoxy-D-xylonate
KGUD	2-ketogluconate-6-phosphate reductase
KGUK	2-ketogluconate kinase
L-KDA	L-2-keto-3-deoxyarabinonate

L-KDR	L-2-keto-3-deoxyrhamnonate
NMR	nuclear magnetic resonance
np-ED	non-phosphorylative Entner-Doudoroff
PDH	pyruvate dehydrogenase
PEP	phosphoenolpyruvate
PFK	phosphofructokinase
PGI	phosphoglucoisomerase
PGK	phosphoglycerate kinase
PGM	phosphoglucomutase
PHA	polhydroxyalkanoates
PHB	poly(3-hydroxybutyrate)
PP	pentose phosphate
sp-ED	semi-phosphorylative Entner-Doudoroff
TCA	tricarboxylic acid
TIM	triosephospate isomerase
UxaC	uronate isomerase
UxuA	mannonate dehydratase (hexuronate metabolism)
UxuB	mannonate oxidoreductase
X5P	xylulose-5-phosphate
XAD/XylD	xylonate dehydratase
XDH	xylose dehydrogenase
xylA	α-ketoglutarate semialdehyde gene
xylB	xylose dehydrogenase gene
XylC	xylonolactonase
ZWF	zwischenferment/glucose-6-phosphate dehydrogenase

1. Introduction

Carbohydrate pathways play an essential role in energy generation and biomass formation in organisms of all domains of life. The most prominent example among carbohydrate pathways, the Embden-Meyerhof-Parnas (EMP) pathway, is considered to be present, at least in parts, in every organism (Fothergill-Gilmore and Michels, 1993). The prevalence of the EMP and the extensive research conducted mostly in eukaryotic and prokaryotic model organisms can easily lead to the misconception it is the most efficient pathway found in nature. However, studies of microorganisms from diverse environments like soil and gut bacteria, deep-branching bacteria and archaea, revealed a number of variations and alternatives to the EMP pathway (Romano and Conway, 1996; Siebers and Schönheit, 2005; Verhees et al., 2003). The EMP pathway together with the pentose phosphate (PP) pathway and the tricarboxylic acid (TCA) cycle, make up what is universally described as the central carbon metabolism (CCM) (Noor et al., 2010). However, the structure of the CCM is not fixed. Incomplete versions and variations of the CCM can be found, depending on environmental conditions and lifestyle of each organism (Peekhaus and Conway, 1998; Romano and Conway, 1996; Sudarsan et al., 2014). The Entner-Doudoroff (ED) pathway is as an alternative carbohydrate pathway to the EMP- pathway that produces pyruvate and metabolites of the lower part of glycolysis. In comparison to the EMP pathway, the ED pathway differs in thermodynamics, ATP yield, number and type of intermediate metabolites and cofactors. In a number of organisms several carbohydrate pathways coexist (Dandekar et al. 1999, Fuhrer et al. 2005, Nikel et al., 2015) even forming complex cyclic-structured metabolisms, while other organisms are devoid of single pathways such as *Zymomonas mobilis*, which lacks an EMP pathway (Sprenger, 1996). Variations of the ED pathway that do not involve any or only a limited amount of phosphorylated intermediates have been studied extensively in archaea (Bräsen et al., 2014; Romano and Conway, 1996). Recently, discoveries of the ED pathway in cyanobacteria, plants and diatoms showed that the ED pathway is more universally distributed as initially expected (Chen et al., 2016; Fabris et al., 2012). Moreover, the ED pathway and other atypical carbohydrate pathways play important roles in pathogenic gut bacteria and microorganisms commonly used in industrial biotechnology such as *Z. mobilis* and *Corynebacterium glutamicum*. Several modified versions of the ED pathway exist, predominantly represented in the kingdom of archaea. The non-phosphorylative version of the ED pathway (np-ED) is based on initial oxidation of the sugar and includes only a minimal amount of cofactors compared to the classical ED and the EMP pathways. The primary oxidation of sugars is a reoccurring principle in other oxidative pathways for the

processing of D-xylose, L-arabinose, L-rhamnose and L-fucose (Johnsen et al., 2009; Watanabe et al., 2006c; Wolf et al., 2016). These non-phosphorylative pathways are found in different species of fungi, bacteria and archaea, and they show remarkable similarities to the np-ED pathway (e.g. mechanism and key enzymes). Crucial enzymes associated with non-phosphorylative pathways include specific hexose and pentose dehydrogenases, sugar acid dehydratases and aldolases. The functional analysis of these enzymes combined with isotope-labelled carbon experiments and systems biology approaches has led to the ongoing discovery of new alternative carbohydrate pathways (Johnsen et al., 2009; Watanabe and Makino, 2009; Wichelecki et al., 2014; Wolf et al., 2016).

Over the last decades, heterotrophic microorganisms have been increasingly used in industrial biotechnology for the sustainable production of speciality and bulk chemicals, fuels and drugs. In biotechnolgical production processes, the availability and cost of carbohydrates derived from biomass play key roles to supply (a) energy for the growth of the production organism and (b) carbon for the synthesis of the product of interest. Accordingly, the identification and flux analysis of carbohydrate metabolisms is crucial for the success of metabolic engineering approaches. Engineering and optimisation of microbial production strains involves the manipulation or partial replacement of parts of the CCM. Common principles include (a) the disruption of carbon draining pathways (b) diverting carbon flux through PP and ED pathways for the generation of reduced cofactors and (c) the complete replacement of endogenous pathways with alternatives which are more suitable for the engineering purpose (Marx et al., 2003; Ng et al., 2015; Siedler et al., 2011; Tenhaef et al., 2018). Despite a comprehensive understanding of the CCM in industrial model organisms like *E. coli* and *S. cerevisiae*, other organisms, such as *Pseudomonas putida* or *Zymomonas mobilis* with less studied metabolisms have emerged as potential candidates due to advantages such as high organic acid, pH or temperature tolerance (He et al., 2014; Nikel et al., 2016; Scully and Orlygsson, 2014).

Owing to various advances in the field of synthetic biology, holistic engineering of cellular metabolisms and the construction of cellular metabolic chassis has become increasingly feasible (Aslan et al., 2017; Heider and Wendisch, 2015; Nikel et al., 2016). However, in an *in vitro* synthetic approach, enzymes from different metabolic pathways can be rapidly assembled in an unrestricted way in order to test more efficient enzymatic production routes. The field of cell-free biocatalysis provides immense opportunities for the rapid testing of *de novo* metabolic pathways assembled from

different enzymes (Petroll et al., 2018). Using this *in vitro* synthetic approach one can take advantage of alternative carbohydrate metabolisms found throughout nature in the design of novel pathways, that might not be feasible *in vivo* (Guterl et al., 2012; Honda et al., 2017; Ye et al., 2012).

In this review we provide an overview of alternative carbohydrate metabolisms in microorganisms with an emphasis on pathways which are initiated by oxidation, rather than phosphorylation. In particular, prevalence, key enzymes as well as reaction principles of the ED pathway and its variations will be discussed and compared to the EMP pathway. Additionally, an overview of non-phosphorylative pathways is provided for a variety of different carbohydrates others than glucose. Finally, we summarise and evaluate recent approaches using alternative pathways and their key enzymes for *in vivo* and *in vitro* metabolic engineering.

2. EMP and ED pathways: principles, differences and prevalence

In abstract terms, metabolic pathways can be seen as multi-objective functions optimised towards simplicity during the process of evolution (Meléndez-Hevia et al., 1997). For example, the PP pathway was found to be the simplest combinatorial solution to the problem of converting six pentoses into five hexoses (Meléndez-Hevia et al., 1994). The main purpose of the CCM in *E. coli* is (a) to provide reducing energy in the form of ATP and (b) a short route for the formation of important cellular precursors (Meléndez-Hevia et al., 1997; Noor et al., 2010). For many organisms the EMP pathway, as a fundamental part of the CCM, satisfies the role of a concise and optimised pathway balancing these two tasks without the use of any unnecessary steps (Bar-Even et al., 2012). However, a more global picture indicates that other alternative carbohydrate pathways work on different principles, which include different intermediates and a different number of enzymatic reactions. In the following, we highlight the main differences in catabolic principles between the EMP, ED and other non-phosphorylative pathways and discuss their characteristics and ubiquity in nature.

2.1. Thermodynamics

The EMP and ED pathways differ in their flux and also in their starting reactions. Despite their differences both pathways have in common that C6 sugars are converted into two C3 sugars, involving a cleavage reaction by an aldolase. However, they differ in crucial reactions steps, which lead to differences in the thermodynamic profile (Fig. 1,

Table 2). While the EMP pathway is optimised towards efficient production of ATP and symmetrical reactions in the bottom half of the pathway, the ED has poor symmetry and low ATP yield, but instead contains fewer thermodynamic unfavourable reactions (Meléndez-Hevia et al., 1997). Pathway flux is not only dependent on the overall change in Gibbs free energy, but also the order of exergonic reactions and thermodynamically unfavourable steps in the pathway. Highly exergonic reactions at the start and more endergonic reactions at the end of a pathway, promote an overall higher flux than vice versa (Meléndez-Hevia et al., 1997). In principle, the first reaction of each pathway fulfils two functions, namely driving the pathway with a highly exergonic reaction and inducing polarity of the sugar. In the EMP pathway this is achieved by phosphorylation whereas in the ED pathway this is accomplished by the oxidation of C1 on the sugar.

Table 2 The main differences of the classical EMP, ED and np-ED pathways. $\Delta_r G'^m$ represents change in Gibbs free energy at pH 7 and an ionic strength of 0.1 M for 1 mM reactant concentrations (http://equilibrator.weizmann.ac.il). Enzymes are as follows: PFK: phosphofructokinase, EDD: 6-phosphogluconate dehydratase, EDA: 2-keto-3-deoxy-phosphogluconate aldolase, GAD: gluconate dehydratase, KDGA: 2-keto-3-deoxy-gluconate aldolase, FBA: fructose-1,6-bisphosphate aldolase, TIM: triosephosphate isomerase, GAPDH: glycerladehyde-3-phosphate dehydrogenase, PGK: phosphoglycerate kinase, PGM: phosphoglucomutase ENO: enolase, GLK: glycerate kinase.

	classical EMP	classical ED	np-ED pathway
$\Delta_r G'^m$ [kJ/mol]	-63.7	-107.2	-150.8
number of reaction steps	10	9	7
mol ATP/mol glucose	2	1	0
mol NAD(P)H/mol glucose	2	2	2
key enzymes	PFK	EDD/EDA	GAD/KDGA
bottlenecks (low ΔG)	FBA, TIM, GAPDH	GAPDH, PGK, PGM, ENO	GLK, ENO

Fundamentally, the EMP pathway conserves more energy via the formation of ATP, while in the ED pathway more energy dissipates, which is reflected in the high exergonism of single enzymatic reactions (Flamholz et al., 2013). Not only has the ED pathway a greater cumulative sum in Gibbs free energy change but analysis also showed that the least favourable ED reactions are nearly twice as exergonic as the least favourable EMP reactions (Flamholz et al., 2013). In the EMP pathway, fructose-1,6-bisphosphate aldolase (FBA) and triosephosphate isomerase (TIM) are thermodynamic bottlenecks causing the cell to accumulate high amounts of substrates or enzymes in order to maintain a high forward flux in the pathway (Flamholz et al., 2013; Liang et al., 2018; Noor et al., 2014). There are other alternative carbohydrate pathways which show similar energetic and thermodynamic profiles to the ED pathway. For example, the phosphoketolase pathway is present in heterofermentative lactic acid bacteria and is a variation of the PP pathway involving a phosphoketolase enzyme

which cleaves a pentose phosphate to glyceraldehyde-3-phosphate (GAP) and acetyl-phosphate. The np-ED pathway has an even higher thermodynamic driving force compared to the classical ED pathway and is considered less optimised in the evolution towards a high production of ATP (Meléndez-Hevia et al., 1997).

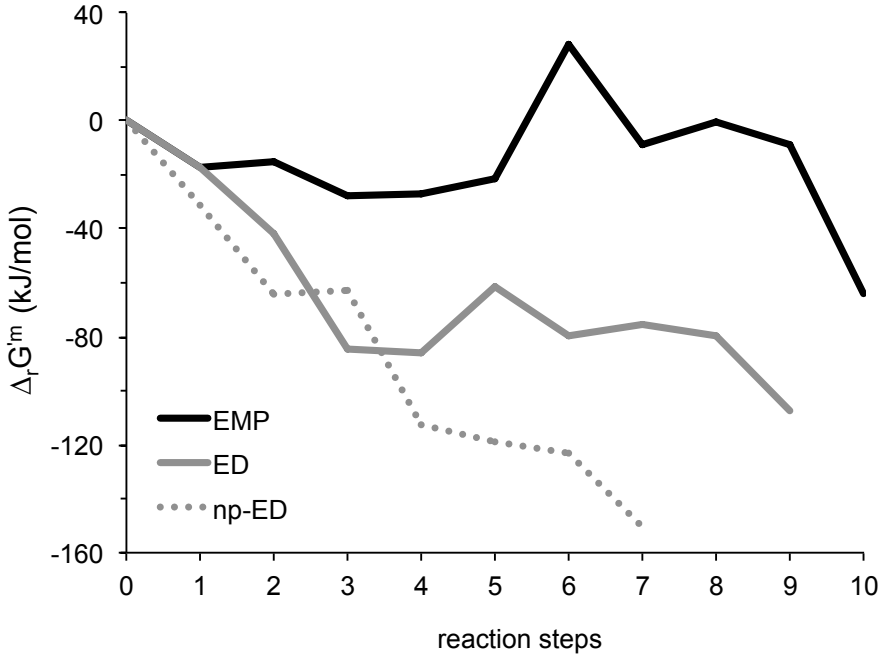


Figure 1 Comparison of the thermodynamic profiles of the classical EMP, ED and np-ED pathways at pH 7 at an ionic strength of 0.1 M (http://equilibrator.weizmann.ac.il). The lines represent the profile given 1 mM reactant concentrations ($\Delta_r G'^m$).

2.2. Energetics

Another important difference in the comparison of carbohydrate pathways is the generation of energy in form of ATP. The initial phosphorylation steps in the upper part and following dephosphorylation in the bottom part of the EMP pathway allow for the generation of high amounts of ATP, which is only accomplished in parts in the ED pathway. This implies that the EMP pathway is more optimised towards the provision of ATP than its counterparts, the PP and ED pathways. Overall, 2 mol ATP/glucose are formed in the EMP pathway, 1 mol ATP/mol glucose in the ED pathway and no net yield ATP is produced in the np-ED pathway (Table 2). Among many different transporter families such as the ATP binding cassette (ABC) transporters or the major facilitator superfamily, bacteria transport the initial substrate glucose primarily using the phosphotransferase transport system (PTS), which includes phosphorylation of the sugar (Postma et al., 1993). In eukaryotes like *S. cerevisiae* and in particular in archaea, hexose transport relies on ABC transporters, which don't involve substrate

phosphorylation but still consume one mole of ATP per hexose transported (Albers et al., 1999). This results in different possibilities for initial reactions in the np-ED, the classical ED as well as the EMP pathway. While in the np-ED glucose is oxidised to gluconate, in the classical ED and EMP pathways glucose is phosphorylated as a first step (Fig. 2A). An initial oxidation instead of a phosphorylation step implies that no ATP has to be invested in the upper part of the np-ED pathway, compared to 1 and 2 moles ATP/mol glucose in the classical ED and EMP pathways, respectively. Consequently, no phosphorylated triose is produced in the np-ED pathway until the formation of 2-phosphoglycerate, compared to 1 and 2 triosephosphates/mol glucose in the classical ED and EMP pathways, respectively. In general, the absence of initial phosphorylation comes at the expense of a lower ATP yield in the bottom part of the pathway, but instead a better thermodynamic profile can be achieved.

2.3. Modified pathways

Several microorganisms, especially fungi and archaea, utilise modifications of the EMP, the ED pathway and other sugar metabolisms (Romano and Conway, 1996; Ronimus and Morgan, 2003). A comparison of classical and modified version of ED and EMP pathways is given in Figure 2. In all studied archaea the EMP pathway is modified from the conventional type (Bräsen et al., 2014). The modifications are diverse but most striking are the difference in cofactor use for glucokinase (GK) and phosphofructokinase (PFK) and the different oxidation of the GAP intermediate. In the classical EMP pathway GK and PFK use ATP for phosphorylation, while in archaea ADP-dependent GKs and ADP or pyrophosphate-dependent PFKs are present (Fig. 2B, enzymes 17 and 20) (Kengen et al., 1994; Siebers and Klenk, 1998; Tuininga et al., 1999). In the EMP pathway, the central intermediate GAP is converted to 3-phosphoglycerate (3PG) via two enzymatic steps (a) the NAD+ dependent GAP dehydrogenase and (b) the ATP-forming phosphoglycerate kinase (PGK) (Fig. 2B enzymes 5 and 6). However, in modified archaeal pathways alternative enzymes are able to catalyse the conversion from GAP to 3PG without ATP formation (Fig. 2B, enzymes 7 and 23). Active $NAD(P)^+$ dependent non-phosphorylating GAP dehydrogenase (GAPN) have been reported in *Thermoproteus tenax* (Brunner et al., 1998), *Sulfolobus solfataricus* (Ettema et al., 2008) and *Thermococcus kodakarensis* (Matsubara et al., 2011) while a ferredoxin dependent glyceraldehyde-3-phosphate oxidoreductase (GAPOR) was found in a range of species including *Thermococcus kodakarensis* and *Pyrococcus furiosus* (Matsubara et al., 2011; Selig et al., 1997). The low redox potential of ferredoxin enables some organisms another level of energy conservation, which allows for the production of

hydrogen via ferredoxin-dependent hydrogenases. In contrast to the conversion in the classical EMP pathway via the GAPDH/PGK enzyme pair, both GAPN and GAPOR convert GAP to 3PG irreversibly. GAPDH and PGK in archaea are almost exclusively used in the gluconeogenetic direction, with exception of haloarchaea and some methanogens.

Although initially discovered in pseudomonads, the ED pathway has been found in a wide range of bacteria and archaea (Bräsen et al., 2014; Conway, 1992; Siebers and Schönheit, 2005) as well as eukaryotes and cyanobacteria (Chen et al., 2016; Conway, 1992; Fabris et al., 2012). Identification of ED and EMP pathways is usually performed through genomic screening for their characteristic enzymes. For the ED pathway these enzymes include 6-phosphogluconate dehydratase (EDD) and 2-keto-3-deoxy-6-phosphogluconate aldolase (EDA) (Fig. 2A, enzymes 2 and 3), whereas PFK and phosphoglucoisomerase (PGI) are characteristic for the operation of the classical EMP pathway (Fig. 2B, enzymes 19 and 18). However, the presence of the respective genes or even their enzyme activities is not conclusive evidence for the operation of each pathway. For example, the carbohydrate metabolism in *P. furiosus* was initially assumed to operate on the ED pathway based on the presence of almost all ED enzyme activities and the finding of a novel tungsten aldehyde oxidoreductase which was believed to be the missing link in a functional ED pathway (Mukund and Adams, 1991). However, further analysis of key metabolites by ^{13}C nuclear magnetic resonance (NMR) analyisis of cell extracts together with enzyme measurements elucidated the operation of a modified EMP pathway in *P. furiosus* (Kengen et al., 1994).

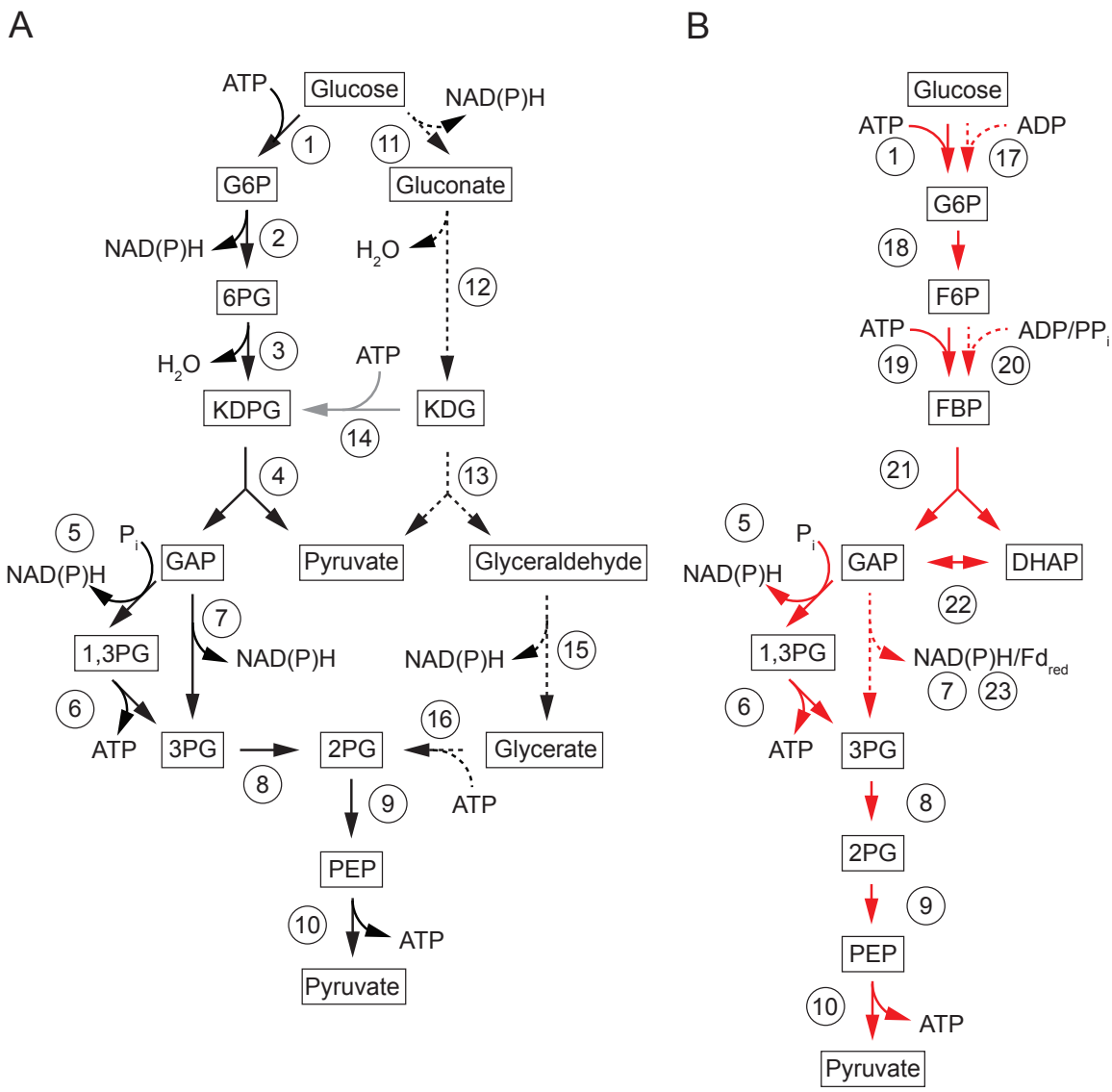


Figure 2 Overview of the ED (A) and the EMP (B) pathways and their variations adapted from (Ahmed et al., 2005) and (Siebers and Schönheit, 2005). Black dashed arrows: non-phosphorylative ED pathway, black solid arrows: classical ED pathway, grey arrow: semi-phosphorylative ED pathway, red solid arrows: classical EMP pathway, red dashed arrows: EMP pathway variations. Enzymes are indicated by numbers as follows: 1: glucokinase (GK), 2: glucose-6-phosphate dehydrogenase, 3: 6-phosphogluconate dehydratase (EDD), 4: 2-keto-3-deoxy-6-phosphogluconate aldolase (EDA), 5: glyceraldehyde-3-phosphate dehydrogenase (GAPDH), 6: phosphoglycerate kinase (PGK), 7: non-phosphorylative glyceraldehyde-3-phosphate dehydrogenase (GAPN), 8: phosphoglycerate mutase (PGM), 9: enolase (ENO), 10: pyruvate kinase (PK), 11: glucose dehydrogenase (GDH), 12: gluconate dehydratase (GAD), 13: 2-keto-3-deoxygluconate aldolase (KDGA), 14: 2-keto-3-deoxygluconate kinase (KDGK), 15: glyceraldehyde dehydrogenase, 16: glycerate kinase (GLK), 17: ADP-dependent glucokinase, 18: phosphoglucose isomerase (PGI), 19: ATP-dependent phosphofructokinase (PFK), 20: ADP-dependent phosphofructokinase, 21: fructose-1,6-bisphosphate aldolase (FBA) 22: triosephosphate isomerase (TIM), 23: glyceralaldehyde-3-phosphate oxidoreductase (GAPOR). Metabolites are as follows: G6P: glucose-6-phosphate, 6PG: 6-phosphogluconate, KDPG: 2-keto-3-deoxy-6-phosphogluconate, KDG: 2-keto-3-deoxygluconate, GAP: glyceraldehyde-3-phosphate, 1,3PG: 1,3-bisphosphoglycerate, 3PG: 3-phosphoglycerate, 2PG: 2-phosphoglycerate, PEP: phosphoenolpyruvate, F6P: fructose-6-phosphate, FBP: fructose-1,6-bisphosphate, DHAP: dihydroacetone phosphate, $Fd_{red.}$ reduced ferredoxin, PP_i: pyrophosphate, P_i: inorganic phosphate.

In addition to the classical ED pathway, modifications such as the semi- and non-phosphorylated ED pathways (sp-ED and np-ED) exist in several bacteria and archaea (Fig. 2A) (Bräsen et al., 2014; Siebers and Schönheit, 2005). When compared to the classical ED pathway, in the sp-ED and np-ED pathways, glucose is first oxidised to gluconate by a dehydrogenase instead of being phosphorylated by a kinase. This is a crucial step, since it results in an unphosphorylated pathway intermediate 2-keto-3-deoxygluconate (KDG) instead of 2-keto-3-deoxy-phosphogluconate (KDPG). In the np-ED pathway, KDG can be cleaved into pyruvate and glyceraldehyde without any phosphorylation. However, in the sp-ED pathway, the intermediate KDG is phosphorylated by a KDG kinase to yield KDPG (Fig. 2A, enzyme 14). Szymona and Duodoroff first reported the np-ED pathway in *Rhodopseudomonas spheroides* where they discovered a 'new aldolase dehydrogenase system' for the oxidation of sugars into the central metabolite KDG (Szymona and Doudoroff, 1960). The same np-ED pathway was then confirmed in other organisms including *Clostridium aceticum* (Bender and Gottschalk, 1973), *Achromobacter* (Kersters and De Ley, 1968), *Aspergillus* sp. (Elzainy et al., 1973), *Thermoplasma acidophilum* (Budgen and Danson, 1986), *S. solfataricus* (De Rosa et al., 1984), *S. acidocaldarius* (Selig et al., 1997) and *Picrophilus torridus* (Reher and Schönheit, 2006). In addition, a branched version of the ED pathway was detected in *T. tenax* and *S. solfataricus* in which the central intermediate KDG can either be converted to glyceraldehyde following the np-ED pathway or phosphorylated to KDPG following the sp-ED pathway (Ahmed et al., 2005; Kouril et al., 2013). While most bacteria contain the enzymes for the sp-ED pathway, hyperthermophilic archaea can operate on combinations of modified EMP and ED pathways (Siebers and Schönheit, 2005).

In organisms which contain multiple pathways and variations thereof, the question arises which pathways receive the most flux. A comparative study with two *Thermococcus* species, *Desulfurococcus amylolyticus T. tenax*, *S. acidocaldarius*, *P. furiosus* and *Thermotoga maritima* (Selig et al., 1997) showed the different modifications of glycolytic pathways and the flux ratio of EMP and ED pathways. *Pyrococcus furiosus*, *Thermococcus* sp. and *D. amylolyticus* were found to catabolise 100% glucose via modified EMP pathways. However, the pathway modifications were different in each organism (Kengen et al., 1994; Schäfer and Schönheit, 1992). For example, in *D. amylolyticus* GAPOR replaced the GAPDH and PGK, while in *Thermococcus* additionally ADP dependent kinases were present (Koga et al., 2000; Matsubara et al., 2011). *Thermoproteus tenax* and *T. maritima* were shown to ferment 85% glucose via an EMP pathway and 15% via an ED pathway. The crenarchaeon *T. tenax* uses several modifications in the EMP pathway, while a classical EMP pathway is

active in the eubacterium *T. maritima*. The hyperthermophilic archaeon *S. acidocaldarius* was the only organism that used the np-ED pathway exclusively. The study demonstrated how complex the modifications and combinations of both ED and EMP pathways are in only a fraction of microorganisms.

In the following section we discuss the occurrence of the ED pathway as an example for an alternative carbohydrate pathway in organisms other than archaea. We will focus on the complex CCM in *Pseudomonas*, several gut bacteria and biotechnology-relevant organisms like *Z. mobilis*. For a more comprehensive review of carbohydrate pathways present in archaea the reader is directed to available work in the literature (Bräsen et al., 2014; Siebers and Schönheit, 2005; Verhees et al., 2003).

3. Atypical ED pathways in *Pseudomonas*

The ED pathway was originally discovered in *Pseudomonas saccharophila* by tracing isotope-labelled carbon from the substrate glucose to the products pyruvate and CO_2 (Entner and Doudoroff, 1952). Pseudomonads were not only the first genus described to operate the ED pathway, but this genus also contains very atypical metabolisms, with different variations at the species level. In general, the catabolism of glucose to GAP and pyruvate in pseudomonads proceeds via the ED pathway, with 6-phosphogluconate (6PG) as a central metabolite. The EMP pathway is not operative in its classical way since pseudomonads lack a PFK, which is a characteristic enzyme for a classical operation of the EMP pathway (Lessie and Phibbs, 1984). However, other genes encoding for EMP pathway enzymes were found, that were able to perform gluconeogenesis starting from GAP (Tiwari and Campbell, 1969). In *Pseudomonas aeruginosa* there are three ways in which the central metabolite 6PG is formed from glucose, two of them are oxidative and one is phosphorylative (Del Castillo et al., 2007; Hunt and Phibbs, 1981). Under aerobic conditions glucose is oxidised to gluconate or 2-ketogluconate extracellularly by membrane bound glucose- and gluconate dehydrogenases (GDH and GNDH), respectively (Fig. 3). After transport into the cell, both compounds are phosphorylated to 6PG and 2-ketogluconate-6-phosphate (Mitchell and Dawes, 1982). However, under anaerobic conditions glucose is imported into the cell first, phosphorylated to glucose-6-phosphate (G6P), which is then further oxidised to 6PG (Hunt and Phibbs, 1983). 6PG is a central intermediate of the ED pathway in *Pseudomonas* and is converted to KDPG by EDD and subsequently cleaved by EDA into GAP and pyruvate (Fig. 3). *Pseudomonas aeruginosa*, *Pseudomonas cepacia*, and *Pseudomonas flourescens* were all shown to utilise both oxidative and phosphorylative

ways to convert glucose into 6PG, while in *P. putida* glucose is only oxidised but not phosphorylated (Nikel et al., 2015; Vicente and Canovas, 1973).

3.1. A cyclic CCM in *P. putida* KT2440

Recently, the atypical carbohydrate metabolism of *P. putida* KT2440 was investigated. Calculation of net flux analyses based on ^{13}C labelled substrate tracing was used to determine the primary processing of glucose and to elucidate the interaction of ED, PP and incomplete EMP pathways (Del Castillo et al., 2007; Nikel et al., 2015). This approach demonstrated that 90% of the carbon from glucose was oxidised to gluconate and subsequently phosphorylated to 6PG while only 10% of the glucose was directly phosphorylated by GK and then converted to 6PG (Nikel et al., 2015). Furthermore, it was shown that the central intermediate 6PG represents a metabolic node, where 91% of carbon is channelled into the ED pathway and the rest is recycled through the PP and the incomplete EMP pathways. In addition to flux analysis, knockout strains of key enzymes demonstrated the importance of each pathway. For example, knockout of the *edd* gene resulted in no growth of *P. putida* KT2440 on glucose, indicating a crucial role of the ED pathway. The relevance of the PP and EMP pathways were also investigated with knockout strain of their respective characteristic genes. Results showed that the PP pathway has a non-essential role in *P. putida* KT2440 whereas the relevance of the EMP pathway has been so far underestimated in this organism. Further, by combining the results obtained from the net flux analysis and growth of the knockout strains on glucose and succinate, it was determined that the CCM in *P. putida* operated in a cyclic fashion, which has been coined the 'EDEMP' pathway (Fig. 3) (Nikel et al., 2015).

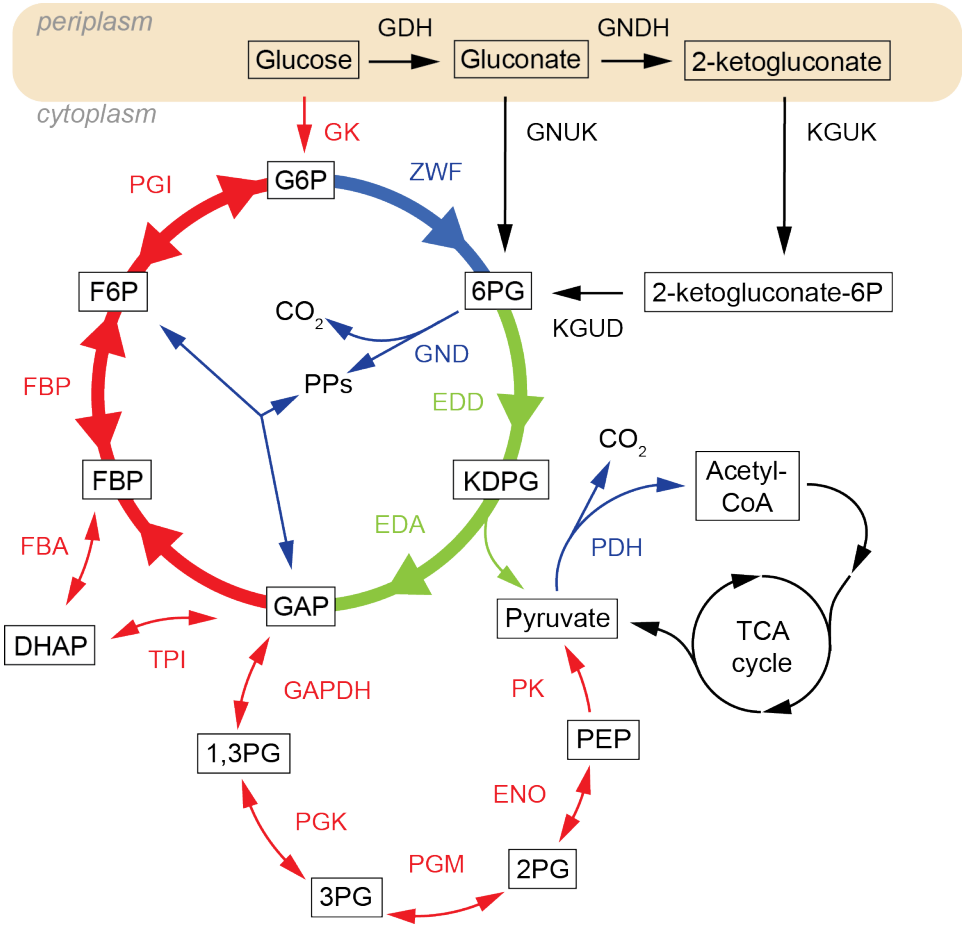


Figure 3 The carbohydrate metabolism of *P. putida* KT2240 with the main constituents of the ED (green) EMP (red) and PP (blue) pathways, called the EDEMP pathway, adapted from (Nikel et al., 2015). Metabolites are boxed and abbreviated according to Fig. 2. Enzymes are as follows: GDH: glucose dehydrogenase, GNDH: gluconate-2-dehydrogenase, GNUK: gluconate kinase, KGUK: 2-ketogluconate kinase, GK: glucokinase, PGI: glucose-6-phosphate isomerase, ZWF: zwischenferment/glucose-6-phosphate dehydrogenase, GND: 6-phosphogluconate dehydrogenase, KGUD: 2-ketogluconate-6-phosphate reductase, EDD: 6-phosphogluconate dehydratase, EDA: 2-keto-3-deoxy-6-phosphogluconate aldolase, FBA: fructose-1,6-bisphosphate aldolase, TPI: triosephosphate isomerase, FBP (enzyme): fructose-1,6-bisphosphatase, GAPDH: glyceraldehyde-3-phosphate dehydrogenase, PGK: phosphoglycerate kinase, PGM: phosphoglucomutase, ENO: enolase, PK: pyruvate kinase, PDH: pyruvate dehydroegnase. PPs: pentose phosphates.

This cyclic pathway combines parts of the ED, EMP and PP pathway for a maximal recycling of triosephosphates. The central metabolite 6PG is processed via the ED pathway to yield pyruvate and GAP. Pyruvate is used to produce acetyl-CoA while GAP can be (a) converted to pyruvate or (b) isomerised to dihydroacetone phosphate (DHAP) or c) converted to G6P through the gluconeogenetic activity of the EMP pathway enzymes. The circular metabolic route is closed by the conversion of G6P into 6PG by glucose-6-phosphate dehydrogenase (ZWF), which is part of the PP pathway (Nikel et al., 2015). The authors claimed that the cyclic route provides a pool of important cellular intermediates (DHAP, FBP and G6P) for further biomass production but also protects the cell from oxidative stress (Chavarría et al., 2013). For example, insertion of the *pfk* gene from *E. coli* into *P. putida* enables the cell to operate on the

classical EMP pathway since all other enzymes for the EMP are already present in the organism. In turn, insertion of *pfk* resulted also in a drastic change in ATP and NADPH levels, which rendered the cells more susceptible to oxidative stress from diamine and hydrogen peroxide. The change in cofactor levels was linked to the lower formation of NADPH/mole glucose in the EMP pathway in comparison to multiple NAPDH/mole glucose in the EDEMP (Nikel et al., 2015). The cyclic operation of three glycolytic pathways ED, EMP and PP has been recently discovered in *P. putida* KT2440 and it remains unclear if this metabolic strategy is widely distributed among other *Pseudomonas* strains.

Pseudomonas putida is a very attractive organism that can be used as microbial cell factory due to its high production of reducing power, ability to recycle triosephosphates and high tolerance to solvent and oxidative stress (Chavarría et al., 2013; Nikel et al., 2016; Ramos et al., 2015). Accordingly, understanding of the complete metabolic networks in this organism, especially the formation of reduced cofactors, is of great interest to the industrial biotechnology sector.

4. The ED and other carbohydrate pathways in gut bacteria

Based on the metabolic strategies utilised by *Pseudomonas* described in the previous section, it is evident that the CCM represents a unique network of pathways shaped by the environment and nutrient availability of the organism. The existence and importance of each component of the CCM (EMP, PP, TCA and ED pathways) varies within species of the same genus and upon different environmental conditions. This is especially important for gut bacteria, which have the potential to cause infections in humans. According to the nutrient-niche hypothesis postulated by Freter it is believed that ecological niches in the gastrointestinal tract facilitate the diversity of microorganisms (Freter et al., 1982). Surprisingly, still little is known about the exact composition of the carbon sources and the metabolic pathways used by gut bacteria. In organisms living on a herbivore diet the preferred mode of energy generation by gut bacteria is fermentation of carbohydrates after breakdown of polysaccharides by *Bifidobacterium* and *Ruminococcus* (Flint et al., 2008). In the intestine the carbon source availability is not only dictated by the food intake, but also bacterial cell debris and the mucus layer of the epithelium (Peekhaus and Conway, 1998). One major component of the mucus layer are mucins, a group of O-glycosylated proteins secreted by epithelial cells. Glycosylation includes sugars like L-fucose, D-galactose, sialic acids and N-acetylgalactosamine (Conway et al., 2004; Robbe et al., 2004). Besides mucins, the mucus layer consists of smaller glycoproteins, glycolipids, sugars and sugar acids

e.g. gluconate. Several of the carbohydrates present or released by the mucus are crucial for colonisation and virulence of enterohaemorrhagic *E. coli.* By disruption of different cellular functions, it has been shown, that the ED pathway is essential for the colonisation initiation of *E. coli* (Chang et al., 2004). In particular, gluconate was identified as important carbon source for colonisation initiation, while glucuronate, mannose, fucose and ribose play important roles in the maintenance of colonisation. In line with the nutrient-niche hypothesis, enterohaemorrhagic *E. coli* strains sense sugars like L-fucose, to adapt their carbon source metabolism towards the sugar the least competed for (Pacheco et al., 2012; Snider et al., 2009). In general, studying the nutrient availability and metabolism of intestinal microorganisms, especially pathogenic ones, is crucial for the understanding of infections and colonisation of microbes in humans (O'Hara and Shanahan, 2006; Patra et al., 2012; Peekhaus and Conway, 1998).

4.1. The ED pathway in *Campylobacter*

Campylobacter is a genus that is mainly known for the species *Campylobacter jejuni* and *Campylobacter coli*, the most common causes for gastroenteritis in humans. For a long time *Campylobacter* was believed to be non-glycolytic, since the organism lacks the crucial enzyme for the EMP and PP pathways, PFK and GK, respectively (Parkhill et al., 2000). Accordingly, it was assumed that *Campylobacter* acquired carbon from amino acids and other intermediates of the TCA cycle. However, since *C. jejuni* has downstream enzymes of the EMP pathway, a gluconeogenetic activity was hypothesised by Velayudhan and Kelly (Velayudhan and Kelly, 2002). *C. jejuni* possess a complete TCA cycle and is likely that this provides both, biosynthesis of necessary cellular compounds and energy. By considering that possible metabolic route for *C. jejuni,* Vorwerk et al. discovered that there are indeed species among *Campylobacter* which are able to metabolise glucose (Vorwerk et al., 2015). They discovered a genomic island in *C. coli* and *C. jejuni* subsp. *doylei* containing genes for glucose transporter and ED pathway enzymes. This so called 'plasticity region' enables a glucose metabolism by a combination of PP and np-ED pathway enzymes. Transformation of non-glycolytic *Campylobacter* strains with this plasticity region enabled them to grow on glucose as a carbon source (Vorwerk et al., 2015).
^{13}C-isotopologue profiling enables the identification of pathway metabolites and products originating from a labelled substrate (Eisenreich et al., 2010). In the case of *C. jejuni* and *C. coli* labelling of glucose revealed that glucose is not just used for energy

generation and amino acid production but also for the production of cell surface carbohydrates.

4.2. The ED pathway in *E. coli*

Escherichia coli is a main component of the intestinal micro-biome and has evolved to use a wide range of different carbon sources (Peekhaus and Conway, 1998). The central metabolism of *E. coli* features all the main carbohydrate pathways such as the EMP, ED and PP pathways. *Escherichia coli* is also able to break down amino acids and dicarboxylates via gluconeogenesis. In addition to gluconate, *E. coli* is also able to grow on D-galacturonate, D-glucuronate, L-idonate, which are imported through permeases and funnelled into the ED pathway after isomerisation and reduction/dehydratation reactions (Mandrand-Berthelot et al., 2004).

The importance of the ED pathway in *E. coli* was first indicated by colonisation experiments using *edd* and *eda* knockout mutants. For example, *eda* knockout strains *E. coli* F-18 and K-12 were not able to colonise the mouse intestine but complementation of the *eda* gene restored their ability to colonise the intestine again (Sweeney et al., 1996). However, *eda* and *edd* double knockout strains were able to colonise mice intestines but when the double knockout strain was injected into mice together with wild type parent strains the double knockout strain did not show colonisation (Sweeney et al., 1996). This experiment supports the nutrient-niche hypothesis, since the wild-type strains outcompete the knockout mutant strain.

In contrast to the essential role in *E. coli* colonisation, the ED pathway can also be complemented by other pathways upon its deletion or failed operation. For example, Waligora et al. investigated plaque formation of *Shigella flexneri*, an enteric pathogen and a main cause for diarrhea in humans (Waligora et al., 2014). Although *S. flexneri* contains a functional ED pathway and *eda* single mutants showed decreased plaque formation, double *edd* and *eda* knockout mutants showed identical plaque formation compared to the wild type when tested on epithelial cells. This indicated that despite disruption in the ED pathway, carbon catabolism can be maintained by an alternative route. The EMP pathway has shown to play a major role in *S. flexneri*. Knockout of key enzymes of the EMP pathway, PFK or pyruvate kinase (PK), results in a loss of plaque formation compared to the wild-type. Apart from the important role of the ED pathway in *E. coli and Campylobacter*, a crucial role of the ED pathway has been verified for other pathogens such as *Salmonella enterica* and *Vibrio cholerae* (Bowden et al., 2009; Diacovich et al., 2017; Patra et al., 2012).

5. The ED pathway in *Zymomonas mobilis*

Zymomonas mobilis is known for its wide applications in industrial fermentation processes, especially in the production of bioethanol. The species were originally discovered in tropical plant saps (e.g. agave) which are fermented to produce the mexican wine 'pulque' (Swings and Genetics, 1977). *Zymomonas mobilis* is often preferred over *S. cerevisiae* for the production of ethanol due to its high ethanol productivity (12%, w/v), ethanol tolerance (16%, vol/vol) and high sugar uptake and turnover (Lee et al., 2010). Furthermore, given the ability of *Z. mobilis* to utilise lignocellulosic waste in several biotechnology applications, it is of great interest how this organism catabolises its substrates. In *Z. mobilis* no EMP pathway is operative but instead the carbohydrates sucrose, glucose and fructose are metabolised exclusively via the ED pathway (Doelle et al., 1993; Fuhrer et al., 2005). Glucose and fructose are metabolised and transported into the cell by different mechanisms while sucrose is split extracellularly by a levansucrase. A glucose-fructose oxidoreductase (GFOR) so far only be found in *Z. mobilis*, catalyses the conversion of glucose to gluconolactone and fructose to sorbitol (Zachariou and Scopes, 1986). GFOR is located in the periplasm and can withstand high concentrations of sugars, especially fructose as indicated by its high K_m value. The formation of sorbitol from fructose helps the cell to counteract high osmotic stress, which it is likely to encounter naturally in dry and sugar-rich plant saps (Rehr et al., 1991; Sprenger, 1996). Apart from the formation of sorbitol, gluconolactone is cleaved by a lactonase into gluconate which is then converted to 6PG, an intermediate of the classical ED pathway. The ED pathway operates in conjunction with a pyruvate decarboxylase and an alcohol dehydrogenase, producing ethanol as a final fermentation product. A facilitated uniport system is responsible for the import of glucose, which only functions under high sugar concentration and without the need of metabolic energy (Weisser et al., 1995). Inside the cell, glucose is phosphorylated to G6P and then converted to 6PG via 6-phosphogluconolactone. Due to the exclusive operation of the ED pathway only 1 mol of ATP/mol glucose is produced in *Z. mobilis*, compared to 2 produced by the EMP pathway of *S. cerevisiae*. This often leads to the general assumption that *Z. mobilis* has to constantly cope with a low production of ATP. However, the low production of ATP/mol glucose seems to cover the cell's needs, presumably by a faster turnover of sugars and at a low protein cost of the ED pathway (Flamholz et al., 2013). The catabolic rate of glucose in *Z. mobilis* is 0.75-1 µmol glucose/mg dry weight/min (Viikari and Berry, 1988), which is three to five times faster than in yeast and accomplished by constitutive expression of enzymes involved in the fermentation process. Accordingly, based on high glucose uptake and turnover, ATP

amounts generated per time in *Z. mobilis* surpass those observed in yeast (Kalnenieks et al., 2014). The metabolism of *Z. mobilis*, including its ED pathway, has recently been subjected to systematic and modelling analyses in order to understand and fully exploit this organism's capabilities as a production strain (Kalnenieks et al., 2014).

6. Prevalence of the ED and EMP pathways

The EMP pathway is often regarded as the most efficient pathway in carbohydrate metabolism since this pathway supplies important intermediates for the cell while generating the highest ATP yield/glucose of all glycolytic pathways (Bar-Even et al., 2012; Flamholz et al., 2013). Another reason for the predominance of this pathway is that mostly model organisms (e.g., *E. coli*, *Bacillus subtilis*, *S. cerevisiae*, *Corynebacterium glutamicum)*, which metabolise glucose mostly through the EMP pathway, have been the focus of metabolic pathways and engineering research. These organisms do not necessarily represent the most abundant pathways throughout all microorganisms (Fuhrer et al., 2005). In fact there is a variety of pathways which utilise different, strategies for example the oxidative or non-phoshorylative breakdown of sugars as it is case for the np-ED pathway. This raises two questions (a) how widely distributed are alternative carbohydrate pathways and (b) is the EMP pathway the obvious choice in terms of efficiency in carbon metabolism? It is probably not possible to determine the actual percentage of all organisms operating carbohydrate pathways other than the EMP pathway, since genetic predisposition does not always account for a functional pathway. However, several approaches have been undertaken to elucidate the distribution of ED and EMP pathways. For example, Kersters and de Ley analysed 150 species belonging to 37 different bacterial genera by confirming key enzymatic activities of pathways. 24 of the 37 genera had at least one representative species, which showed ED pathway activity. It was concluded that mostly Gram-negative bacteria had a functional ED pathway whereas hardly any Gram-positive bacteria showed ED pathway activity (Kersters and De Ley, 1968).

The prevalence of the ED pathway was analysed also in representatives of the genus *Alteromonas* and *Alcaligenes,* which contain marine bacteria with non-fermentative lifestyles. All of these representatives contained active ED enzymes but none showed an active PFK, the characteristic enzyme of the EMP pathway (Baumann and Baumann, 1973). The high occurrence of the ED pathway in marine bacteria was supported by another report of 26 marine strains alongside with other *Pseudomonas* strains (Lee et al., 1977). Almost all of the tested organisms showed enzyme activities of the ED pathway, although exact numbers were not included in the report.

Cyanobacteria were long believed to metabolise hexoses via the oxidative pentose phosphate pathway (OPP) (Gómez-Baena et al., 2008). Recently, Chen et al. demonstrated that the ED pathway was responsible for the metabolism of glucose in *Synechocystis* (Chen et al., 2016). The authors reported that a mutant strain, which had the key EMP and oxidative PP pathway enzymes knocked out displayed enhanced growth under mixotrophic conditions. This suggested the presence of another operating pathway in the knockout strain. The authors were able to identify the presence of EDA, a characteristic enzyme of the ED pathway, and confirmed a functional ED pathway in *Synechocystis*. In the same study, genomic analyses revealed that most cyanobacteria lack a *pfk* but contain *eda* and most of those not containing *eda* live in freshwater. The authors hypothesised that freshwater cyanobacteria, thriving in a nutrient-rich environment, live a more heterotrophic lifestyle and therefore prefer formation of ATP via the EMP pathway. However, in nutrient-poor saltwater environments, cyanobacteria rely more on photosynthesis for ATP generation, which is why the ED pathway is the preferred metabolism, since less ATP from carbohydrates has to be produced (Chen et al., 2016). Similar to gut bacteria, these findings further support the hypothesis that environmental factors and nutrient supply dictates the prevalence/preference of a particular pathway.

Diatoms are another class of marine microorganisms with an important role in the global carbon cycle. Diatoms are responsible for fixing roughly the same amount of carbon as all rainforest on earth and play a crucial role in providing the basis of the ocean's food web (Armbrust, 2009; Field et al., 1998). Transcriptional and translational analyses of the diatom *Phaeodactylum tricornutum* revealed an upregulation of ED pathway genes *edd* and *eda* in a light-modulated manner (Fabris et al., 2012). Expression and transcript levels of ED pathway genes were downregulated in continuous light, and upregulated when switched to dark phases. Further *in vitro* assays of *P. tricornutum* verified the activity of the dehydratase and aldolase enzymes and provided further evidence of an active ED pathway in eukaryotes.

In a comparative study Fuhrer et al. investigated the glucose metabolism of phylogenetically distinct prokaryotes, with different metabolic lifestyles. Seven prokaryotes *Agrobacterium tumefaciens*, two pseudomonads, *Sinorhizobium meliloti*, *Rhodobacter sphaeroides*, *Z. mobilis*, and *Paracoccus versutus* were compared to the model organisms *E. coli* and *B. subtilis* using metabolic flux ratio analysis upon growth on glucose. In contrast to *E. coli* and *B. subtilis*, which primarily used the EMP pathway, all 7 prokaryotes metabolised glucose almost exclusively via the ED pathway. In addition, the lowest flux through the TCA cycle was encountered in *E. coli* and *B. subtilis* due to overflow metabolism and the generation of excessive amounts of

acetate.

The emergence of high-throughput genomic analysis has made possible the acquisition of large genomic data sets to study the distribution of characteristic genes is several bacterial species. For example, Flamholz et al. analysed 500 bacterial and archaeal genomes for the presence of distinctive genes from the EMP (*pfk*) and ED (*edd* and *eda*) pathways (Flamholz et al., 2013). 57% of the prokaryote genomes had the unique gene of the EMP pathway, 27% had only the genes for the ED pathway while 14% have distinctive genes for both pathways. However, this data only demonstrated the genetic predisposition for (a) possible pathway(s) and further transcriptional, proteomic and metabolic data would be required to confirm if the respective pathway(s) are being transcribed, translated and ultimately functional.

7. Important enzymes of the ED pathway/non-phosphorylative pathways

7.1. Dehydratases

7.1.1. Dehydratases in ED pathways

The initial reaction of ED pathways includes an oxidation of the sugar by an aldohexose dehydrogenase (for sp-ED and np-ED pathways) or by ZWF (classical ED pathway). Following oxidation the resulting sugar acid is dehydrated, which marks a characteristic step in the operation of all ED pathways. In the classical ED pathway of most bacteria (e.g., *Pseudomonas* spp., *E. coli* and *Z. mobilis*) this step is performed by EDD while in the archaeal np-ED pathway this step is catalysed by a gluconate dehydratase (GAD). Despite the fact that both enzymes fulfil a similar function in the pathway, EDD and GAD differ in their protein structure and characteristics. EDDs share a conserved iron-sulfur cluster and belong to the dihydroxyacid dehydratase/EDD (IlvD/EDD) family. In contrast most GADs from archaea are activated by Mg^{2+} or Mn^{2+} and belong to the enolase superfamily. In the context of their ED pathways, EDD from *E. coli*, *Z. mobilis, P. aeruginosa,* and *Helicobacter pylori* are highly conserved and share 49% residues including a 4Fe-4S cluster. The presence of the iron-sulfur cluster in EDD can hinder its use in eukaryotic production hosts, since low iron availability is limiting in eukaryotic cells (Benisch and Boles, 2014).

Several of EDDs have been studied in more detail and shown to be activated by Fe^{2+}, Mn^{2+} and Mg^{2+} as well as other reducing agents such as glutathione, cysteine and

thioglycolate (Gardner and Fridovich, 1991; Scopes and Griffiths-Smith, 1984). An overview and comparison of dehydratases is given in Table 3.

Table 3 Overview of dehydratases from IlvD/EDD, enolase and xylose isomerase-like superfamilies.
Only substrates displaying >10% of the maximal enzyme activity are shown. [1]Unclear if enzyme is part of the allocated enzyme family
glc: gluconate, xyl: xylonate, fuc: fuconate, gal: galactonate, D-man: D-mannonate, L-man: L-mannonate,D-rib: D-ribonate, iso: DL-isovalerate, dh-iso: 2,3-dihydroxyisovalerate, D-ara: D-arabinonate, gluc: glucuronate, galac: galacturonate, fuc: fuconate, L-threo: L-threonate, L-ara: L-arabinonate, L-rha: L-rhamnonate, L-lyx: L-lyxonate, L-gulo: L-gulonate, L-ido: L-idonate

enzyme	EC	abbreviation	pathway	species	active substrates	activating cofactors/additives
IlvD/EDD sf						
6-phospho-gluconate dehydratase	4.2.13	EDD	classical ED	*E. coli*	6-phosphogluconate	**Fe^{2+}**, glycerol-3-phosphate,
				Z. mobilis		**Fe^{2+}**, Mn^{2+}, glycerol, gluconate, ascorbate
				P. aerunginosa		**Fe^{2+}**, Mg^{2+}, Mn^{2+}
				P. flourescens		**Fe^{2+}**, Mg^{2+},Mn^{2+}, glutathione, cysteine
				P. putida		**Fe^{2+}**, Mg^{2+}, Mn^{2+}
gluconate dehydratase	4.2.1.39	GAD	np-ED/ sp-ED	*A. xyloxydans*	glc, xyl, fuc, gal, D-/L-man, D-rib, iso	Mg^{2+}, Mn^{2+}
dihydroxyacid dehydratase		DHAD	valine/ isoleucine synthesis	*S.solfataricus*	dh-iso, glc, xyl, D-ara, gluc, galac, D-fuc, L-threo	Mg^{2+}, Mn^{2+}, Ba^{2+}, NaCl
xylonate deydratase	4.2.1.82	XAD	pentose oxidation	*C. crescentus*	xyl	**Fe^{2+}**, Mg^{2+}
L-arabinoate dehydratase	4.2.1.25	AraC	pentose oxidation	*A. brasiliense*	L-ara, xyl	**Fe^{2+}**, Mg^{2+}
Enolase sf						
gluconate/galactonate dehydratase	4.2.1.140	GAD	np-ED/ sp-ED	*S. solfataricus*	glc, gal,	Mg^{2+}, Mn^{2+}, Ni^{2+}, Co^{2+}
				P. torridus	glc, gal	Mg^{2+}
gluconate dehydratase	4.2.1.39	GAD	np-ED/ sp-ED	*T. tenax*	glc	Co^{2+}, Mg^{2+}, Ni^{2+}, Cd^{2+} (Fe^{2+} inhibits)
				C. pasterianum[1]	glc	Fe^{2+}, Mg^{2+}, Mn^{2+}, Co^{2+}, reducing agents
				H. volcanii	glc	Mg^{2+}, Mn^{2+}
xylonate dehydratase	4.2.1.82	XAD	pentose oxidation	*H. volcanii*	xyl, glc	Mg^{2+}
				Pseudomonas sp.	xyl, glc	Fe^{2+}, Mg^{2+}
D-arabinoate dehydratase	4.2.1.5	AraD	pentose oxidation	*S. solfataricus*	D-ara	Mg^{2+}
L-rhamnonate dehydratase	4.2.1.90	RhamD	pentose oxidation	*E. coli*	L-rha, L-lyx	Mg^{2+}
				S. typhimorium	L-rha, L-lyx	Mg^{2+}
				A. vinelandii	L-rha, L-lyx, L-man	
mannonate dehydratases	4.2.1.8	ManD	diverse	*E. coli*	D-man, L-gulo	Mg^{2+}
				C. salexigens	D-man, L-gulo	Mg^{2+}
				S. enterica	D-man, L-ido	Mg^{2+}
L-fuconate dehydratase	4.2.1.68	FucD	L-fucose pathway	*X. campestris*	L-fuc, D-ara	Mg^{2+}
xylose isomerase-like sf						
mannonate dehydratase	4.2.1.8	ManD (uxuA)	hexuronate metabolism	*S. suis*	D-man	Mg^{2+}

In the np-ED pathway, which has been predominantly studied in archaea, gluconate is converted to KDG via GAD. The activity of this enzyme was observed in all cell extracts known to operate the np-ED pathway (*S. solfataricus, S. acidocaldarius, T. acidophilum, P. torridus, T. tenax, A. oxydans*) and more recently also in *Haloferax volcanii* (Sutter et al., 2016). A more detailed investigation of the properties of the GAD in archaea has been conducted for *S. solfataricus*, *P. torridus* and *T. tenax* (Ahmed et al., 2005; Kim and Lee, 2005; Lamble et al., 2004; Matsubara et al., 2014; Reher et al., 2010). GAD enzymes from *S. solfataricus* and *P. torridus* share a high protein sequence identitiy (44%) and belong to the subgroup of mandelate racemase/muconate lactonising enzymes (MR/MLE), which are part of the enolase superfamily and characteristically incorporate Mg^{2+} ions (Reher et al., 2010). The GAD from *S. solfataricus* was found to be active with gluconate and galactonate, which is in line with the previous discovery of promiscuity for the other ED pathway enzymes from the same organism such as glucose dehydrogenase and KDGA (Lamble et al., 2003).

Besides the archaeal dehydratases, GADs from *Clostridium pasteurianum* and *Achromobacter xylosoxidans* have been investigated more thoroughly (Bender and Gottschalk, 1973; Kim and Lee, 2008). The dehydratase from *A. xylosoxidans* was characterised closely in regards to its biochemical characteristics and evolutionary relationship. Surprisingly, the enzyme shows high substrate promiscuity in particular with D-xylonate, and is phylogenetically more related to the bacterial IlvD/EDD enzymes instead of the GADs of archaea (Kim and Lee, 2008).

7.1.2. Dehydratases of the Enolase superfamily

Besides the importance of EDD and GAD in ED pathways, sugar acid dehydratases have a wide range of different functions in other metabolic pathways. Most studies focus on metabolic routes starting from glucose or galactose as a substrate and neglect alternative substrates that can be used by organisms. The investigation of substrate specificities of enzymes like dehydratases or dehydrogenases are often indicators for so far undiscovered carbohydrate pathways (Johnsen and Schönheit, 2004). In particular the study of single enzymes in superfamilies has helped to elucidate enzymes with identical mechanisms but different physiological roles (Rakus et al., 2008; Wichelecki et al., 2014b; Yew et al., 2007). Many dehydratases involved in the np-ED pathway in archaea are members of the enolase

superfamily, which encompasses enzymes with the same bidomain structure and a conserved mechanism. In addition to dehydratases involved in the catabolism of hexoses in the ED pathway, also dehydratases from the pentose degradation pathways belong to the enolase superfamily e.g. xylonate dehydratase from *Haloferax volcanii* (Johnsen et al., 2009) and D-arabinonate dehydratase from *S. solfataricus* (Brouns et al., 2006). All enzymes of this superfamily share structural similarities including a N-terminal capping domain, which dictates substrate specificity and a TIM-barrel domain for acid/base catalysis. All members of the superfamily share the same partial reaction in which a base in the active site abstracting the alpha proton of a carboxylic acid. Subsequently an enolate anion intermediate is formed and stabilised by a Mg^{2+} ion (Babbitt and Gerlt, 1997; Gerlt et al., 2005). The enzymes of this group share conserved ligands which are essential for Mg^{2+} binding. However, since the basic/acidic catalysts are not conserved across all members, the superfamily can be divided into four subgroups according to the position and identity of the catalytic residues. The four subgroups, enolase, 3-methylaspartate ammonia lyases, muconate lactonising enzymes (MLE) and mandelate racemases (MR) are classified by the reactions their members are able to catalyse and their conserved structural characteristics.
Most of the dehydratases are part of the MR/MLE subgroups, which are characterised by the presence of a histidine base for the abstraction of the alpha proton. Among these subgroups several dehydratases have been identified including D-glucarate/L-idarate dehydratase, D-altronate/D-mannonate dehydratase, D-galactonate dehydratase, D-gluconate dehydratase and L-rhamnonate dehydratase (Gulick et al., 2001; Kim and Lee, 2008; Lamble et al., 2004; Rakus et al., 2008, 2007). Further investigation of the MR/MLE subgroups, in particular annotated mannonate dehydratases, revealed that substrate specificity differed greatly from that expected from gene annotation (Wichelecki et al., 2014b). For example, 43 members of the mannonate dehydratase subgroup were recombinantly expressed to determine their enzyme activities. Based on their substrate specificity and catalytic efficiency, all tested enzymes could be functionally divided into three groups (a) enzymes with high efficiency on mannonate (b) enzymes with low efficiency on mannonate and/or gluconate (c) enzymes with no activity on either substrate. High activity mannonate dehydratases were found in organisms that lack the *uxuA* gene which encodes for an additional mannonate dehydratase and is known to be involved in the degradation of hexuronates (Rothe et al., 2013). The authors hypothesised that in those organisms the role of the high-efficiency mannonate dehydratases was to compensate for

the missing UxuA enzyme in the hexuronate pathway. In contrast, organisms with low-efficiency ManDs mostly contained an *uxuA* gene, which raised the question about the physiological function of low-efficiency mannonate dehydratases in the corresponding organisms. Although further studies found no physiological roles for the low-efficiency dehydratases, they suggested that these enzyme could mark putative starting points for the metabolism of other sugar acids (Wichelecki et al., 2014a, Wichelecki et al., 2014b)

In *Chromohalobacter salexigens* DSM3043 and *E. coli* CFT073, the low-efficiency mannonate dehydratase revealed activity towards L-gulonate, while in *Salmonella enterica* subsp. *enterica* the enzyme showed activity towards L-idonate. However, only in *C. salexigens* the dehydratase was assumed to be a part of a physiological relevant metabolism, whereas the dehydratase in *E. coli* and *S. enterica* could not be assigned to a metabolic pathway (Wichelecki et al., 2014b).

The enolase superfamily is an example of structurally similar enzymes showing diversity in their physiological function (Gerlt et al., 2005). On a mechanistic level the diversification in function is due to variation in loops and other binding domains but not the catalytic domain. This is similar to the family of amidohydrolases in which diverse loop structures determine the substrate specificity of the enzymes but do not necessarily interact with the substrate themselves. For example, among a group of 64 amidohydrolases only 18 enzymes showed an interaction of the substrate with one of the loops (Seibert and Raushel, 2005). Consequently, sequence similarity in catalytic sites does not necessarily correlate with similarity in physiological function and therefore needs to be confirmed by enzyme activity assays.

7.1.3. Dehydratases in the hexuronate metabolism

Glucose has been one of the main substrates in the investigation of many pathways (e.g. the ED pathway) and their enzymes. However, as seen with gut bacteria (section 4), microrganisms have access to a wide range of substrates beyond glucose. For example, hexuronic acids like glucuronate and galacturonate can serve as carbon sources in the metabolism of different organisms including *Dickeya dadantii* (formerly known as *Erwinia chrysanthemi*), *E. coli*, *B. subtilis*, *B. stearothermophilus* and *Aspergillus niger* (Hugouvieux-Cotte-Pattat and Robert-Baudouy, 1987; Kuivanen et al., 2016; Mekjian et al., 1999; Portalier et al., 1980; Shulami et al., 1999). Glucuronate is a component of plant cell walls present in the form of glucuronoxylan, which can be broken down by fungi and

bacteria (Reis et al., 1994; Shulami et al., 1999). In the intestine, glucuronate and galacturonate are present in the mucus and can serve as a carbon source for gut bacteria such as *E. coli* (Chang et al., 2004; Fabich et al., 2008). There are 4 different known pathways for the degradation of glucuronate, with 3 of them involving a dehydratase. The most ubiquitous pathway present in *E. coli*, *B. subtilis* and *D. dadantii* features the genes *uxuA*, *uxuB*, *uxaC* and *exuT*, *kdgA* and *kdgK* which code for enzymes and transporters in the catabolism of hexuronates (Fig. 4). Glucuronate is transported into the cell via the transporter ExuT and then converted into the central metabolites pyruvate and GAP by the action of five different enzymes (Peekhaus and Conway, 1998). First, glucuronate is isomerised into fructuronate via an uronate isomerase (UxaC). A mannonate oxidoreductase (UxuB) converts fructuronate to mannonate, which is then converted into KDG by the action of a mannonate dehydratase (UxuA). KDG is also present as central intermediate of the sp-ED and np-ED pathways where it can be phosphorylated by KDG kinase (KDGK) and cleaved into pyruvate and GAP by KDG aldolase (KDGA). The dehydratase involved here has a similar function as the GAD in the archaeal np-ED pathway, since it is the crucial enzyme for the formation of the central intermediate KDG. However, dehydratases of the hexuronate cluster *uxuA* are mostly part of the xylose isomerase-like superfamily and show only very limited protein sequence identity to dehydratases of the enolase or IlvD/EDD superfamilies.

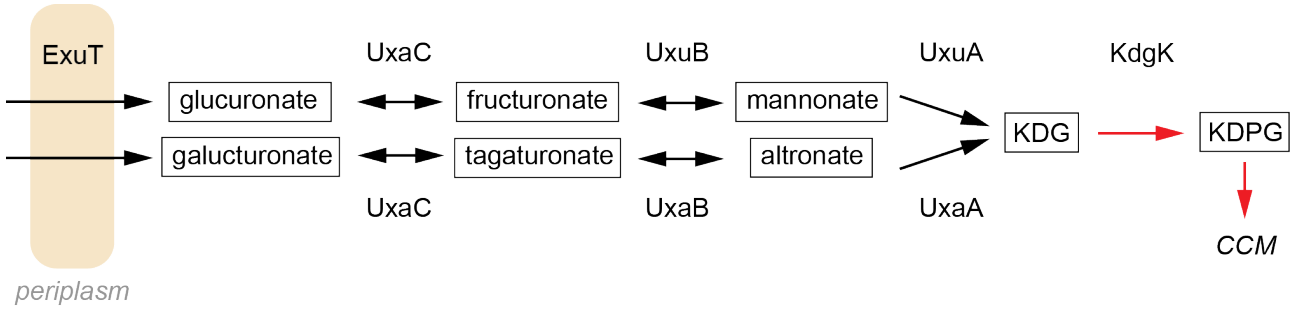


Figure 4 Metabolism of hexuronates in *E. coli.* Red arrows indicate analogous route to the ED pathway. ExuT: hexuronate transporter, UxaC: hexuronate isomerase, UxuB: mannonate oxidoreductase, UxaB: altronate oxidoreductase, UxuA: mannonate dehydratase, UxaA: altronate dehydratase, KdgK: 2-keto-3-deoxygluconate kinase, KDG: 2-keto-3-deoxygluconate, KDPG:, 2-keto-3-deoxy-6-phosphogluconate, CCM: central carbon metabolism.

7.1.4. Dehydratases of the IlvD/EDD superfamily

The IlvD/EDD superfamily comprises EDDs of the classical ED pathway as well as dehydratases involved in the synthesis of the branched amino acids, leucine, isoleucine and valine, in eukaryotes. Despite the difference in their physiological roles, the two enzyme groups IlvD and EDD are structurally similar (Egan et al., 1992). Members of the IlvD/EDD superfamily have been shown to contain Fe-S clusters in different coordination states. Stable [2Fe-2S] clusters have been found in a number of IlvD/EDD superfamily representatives in spinach (Flint and Emptage, 1988), *Rhizobium leguminosarum* (Rahman et al., 2017) and *Caulobacte*r *crescentus* (Rahman et al., 2018) while oxygen-labile [4Fe-4S] cluster have been identified in *Z. mobilis* and *E. coli* (Flint et al., 1993; Rodriguez et al., 1996). Other dehydratases from the IlvD/EDD superfamily were identified in the oxidative pentose pathways of *Azospirillum brasilense, Burkholderia cepacia* and *C. crescentus* (Stephens et al., 2007; Watanabe et al., 2006c). The crystal structure of the D-xylonate dehydratase from *C. crescentus* was recently solved and revealed a tetrameric structure containing a [2Fe-2S] cluster including a Mg^{2+} ion in the monomer-monomer interface, which is similar to the previously studied L-arabinonate dehydratase from *R. leguminosarum* (Rahman et al., 2018, 2017). The D-xylonate dehydratase from *C. crescentus* shares 31% sequence identity with the EDD from *E. coli*, indicating an evolutionary relationship between the oxidative D-xylose pathway and the ED pathway (Stephens et al., 2007). Similar to the enzymes from the enolase superfamiliy, enzymes from the IlvD/EDD superfamiliy often display substrate promiscuity. For example, the dihydroxyacid dehydratase (DHAD) from *S. solfataricus*, showed highest activity towards its natural substrate 2,3-dihydroxyisovalerate while maintaining more than 40% of its activity with D-gluconate, D-xylonate and D-arabonate and 10% of activity with D-glucuronate, D-galacturonate, D-fuconate and L-threonate (Kim and Lee, 2006). The specific activity of the *S. solfataricus* DHAD heterologously-expressed in *E. coli* was determined at 47 U/mg with 2,3-dihydroxyisovalerate as substrate (Kim and Lee, 2006). Carsten et al. expressed the same enzyme and measured a specific activity of 0.32 U/mg on 2,3-dihydroxyisovalerate using high-performance liquid chromatography (Carsten et al., 2015). Based on the results reported by Carsten et al, the DHAD appears to have a relatively low catalytic efficiency. However, DHAD remains an enzyme of industrial interest

in particular because of its versatility and promiscuous activity. More specifically, DHAD has been used in several cell-free biocatalytic systems for the production of acetoin from glycerol, the production of L-lactic acid from glucose and the production of isobutanol and ethanol from glucose (Gao et al., 2015; Xie et al., 2018; Guterl et al., 2012). In general, enzyme promiscuity can be used as a starting point for the engineering of enzyme towards new properties and higher efficiencies (Bornscheuer et al. 2004, Khanal et al. 2015).

7.2. Aldolases

Carbon-carbon bond formation is considered one of the most important and challenging steps in organic synthetic chemistry. Fundamental cellular processes such as the formation of carbohydrates, amino- and α-hydroxy acids are dependent on aldolases. In chemical synthesis, the production of stereo- and regio specific compounds is of great importance, but virtually impossible to be achieved with non-biochemical methods. Instead, the use of chemo-, regio- and stereoselective enzymes provides more control over the formation of enantiopure compounds. Accordingly, aldolases have been used for the synthesis of carbohydrates, sugar analogs and enzyme inhibitors (Clapés et al., 2010; Hecquet et al., 2009; Wong and Whitesides, 1983). Besides their importance for applications in organic chemistry, aldolases play a crucial role in many naturally occurring carbohydrate pathways. For example, in both the EMP and the ED pathways, a C6 sugar is converted into two C3 compounds by an aldolase. Aldolases can be divided into two different classes. Class I aldolases have a conserved lysine residue in the active site which enables the abstraction of a proton to form a Schiff base intermediate. This class of aldolases work independently of bivalent metal ions. In contrast, class II aldolases are known to be dependent on cofactors like Zn^{2+}, Fe^{2+} or Co^{2+} ions for the deprotonation step.

Most KDG- and KDPG aldolases from the *Archaea* domain have been shown to either be promiscuous towards different 2-keto-3-deoxy sugar acids or lack stereoselectivity (Reher et al., 2010; Theodossis et al., 2004; Wolf et al., 2016). Since aldolases are found at crucial points in metabolic routes, their substrate preference can be a key indicator for the existence of an endogenous metabolic pathway. For example, the KDG aldolase from *S. solfataricus* was shown to lack stereoselectivity since 2-keto-3-deoxygalactonate (KDGal) as well as KDG are cleaved into pyruvate and D-glyceraldehdye. This was first observed in

the reverse reaction, the addition of D-glyceraldehyde with pyruvate, which yielded KDG and KDGal in equal amounts (Lamble et al., 2003). Since the first enzyme of the np-ED pathway in *S. solfataricus*, glucose dehydrogenase, showed promiscuity towards galactose, the observed promiscuity of KDG aldolase supported the fact that both sugars, glucose and galactose, can be metabolised in the np-ED pathway of *S. solfataricus* (Lamble et al., 2003). In contrast to *S. solfataricus*, the thermophilic anaerobe *T. tenax* is known to metabolise glucose through a modified EMP pathway, the sp-ED and the np-ED pathways. In the sp-ED pathway, phosphorylation occurs at the level of the 2-keto-3-deoxy sugar acid, which is why the KDG aldolase from *T. tenax* cleaves both KDPG and KDG (Siebers et al., 2004). The operation of the sp-ED pathway was long believed to be exclusive to haloarchaea, whereas the np-ED pathway was considered to be present in (hyper-)thermophilic archaea. However, comparative genetic analysis of both thermophilic archaea and haloarchaea revealed that the ED pathway cluster is conserved in *T. tenax*, *S. solfataricus*, *S. tokodaii* and *Halobacterium sp. NRC*. ^{14}C-tracer analysis on carbohydrate metabolism and biochemical analysis of recombinantly expressed KDG aldolases from *T. tenax* and *S. solfataricus*, revealed that both enzymes show promiscuity towards unphosphorylated and phosphorylated compounds. Therefore both organisms are able to operate on sp-ED and np-ED pathways (Ahmed et al., 2005).

The crystal structures of KDG aldolases from *S. solfataricus* (Theodossis et al., 2004), *S. acidocaldarius* (Wolterink-van Loo et al., 2007), and *T. tenax* (Pauluhn et al., 2008) were studied in further detail and found to show similar flexibility towards their aldehyde acceptors. All studied enzymes belong to the *N*-acetylneuraminate lyase superfamily and depend on the formation of a Schiff base intermediate in the active site usually containing a lysine residue. All three enzymes processed KDG and KDGal, which is in line with the pathway promiscuity. However, there is no physiological reason for the aldolase promiscuity in *T. tenax* since the organism is not able to grow on galactose (Siebers et al., 2004). Promiscuity of enzymes is often interpreted as a starting point for the divergence of function and the development of novel metabolic routes in the course of evolution (Khersonsky et al., 2006). It seems likely that the substrate flexibility in the aldolase of *T. tenax* is just a relic of a former evolutionary step towards the development of other metabolisms, e.g. a galactose dependent np-ED pathway. In summary, engineering of aldolases can be performed to increase their value in industrial applications. For example the activity at low temperatures for a hyperthermophilic aldolase for synthesis reactions at

moderate temperatures has been improved (Van Der Oost et al., 2009). In industry, many synthesis reactions rely on stereoselectivity of the biocatalyst, which is not a common feature for aldolases from most archaea. However, promiscuous enzymes display a good starting point for directed evolution or rational design (Copley et al. 2015, Hammer et. al, 2017, Wang et al. 2017). For example, the promiscuous KDPG aldolase from *S. solfataricus* was modified via site-directed mutagenesis to improve its diastereoselectivity (Royer et al., 2010). In the reaction of aldol addition, the wild type enzyme produces almost equimolar amounts of KDG and KDGal. Via an informed change of certain amino acid residues, a mutant library of 30 different variants could be created. Two mutants were identified that, despite having lower catalytic efficiencies, achieved high diastereoselectivity of either 93% KDG or 88% KDGal.

8. Alternative metabolisms of sugars other than glucose

The variety of carbohydrate pathways in organisms is dictated by their environmental conditions. Glucose has become the first choice as carbon source in molecular- and microbiology practice but also in industrial biotechnological processes. However, most microorganisms originate from an environment that display a wider substrate spectrum than the one found in laboratory-based cultivations. Lignocellulose which consist of cellulose, hemicellulose and lignin is the most abundant non-edible biomass on earth. The average lignocellulosic biomass contains about 20-40% hemicellulose, 30-50% cellulose and 15-25% lignin (Kenney et al., 2013). In particular, hemicelluloses contain a variety of anhydrides of hexoses and pentoses such as xylan, arabinan, glucan, galactan, mannan and uronic acids. Corn stover is the most abundant lignocellulosic waste product in the U.S. and contains on average 37.4% cellulose, 21.1% xylan, 18% lignin, 2.9% arabinan, 2% galactan and 1.6% mannan (Aden et al., 2002). Although this composition varies with the source of the biomass, the second most abundant component after cellulose is xylan, which makes up 15-30% in annual plants, 20-25% in hardwoods, and 7-12% in softwoods. In order to optimise biotechnological fermentation processes towards economic viability, readily available biomass (e.g. waste-streams) can be exploited to reduce the cost of substrate. Metabolic engineering in production strains such as *E. coli*, *S. cerevisiae*, *C. glutamicum* is performed to exploit abundant biomass and achieve high yields, productivities and titers. However, the design capacity in rational metabolic engineering

approaches is inspired by fundamental knowledge and discovery of carbohydrate pathways. Fungi, bacteria and archaea harbour a vast variety of metabolic pathways to catabolise the sugars present in lignocellulosic biomass, offering opportunities for the design of industrial biotechnological processes. The following section highlights unusual pathways for sugars such as D-xylose, L-arabinose, L-rhamnose and L-fucose and the main enzymes involved in those pathways, which might find use in the field of metabolic engineering.

8.1. Pentoses

8.1.1. D-xylose metabolism

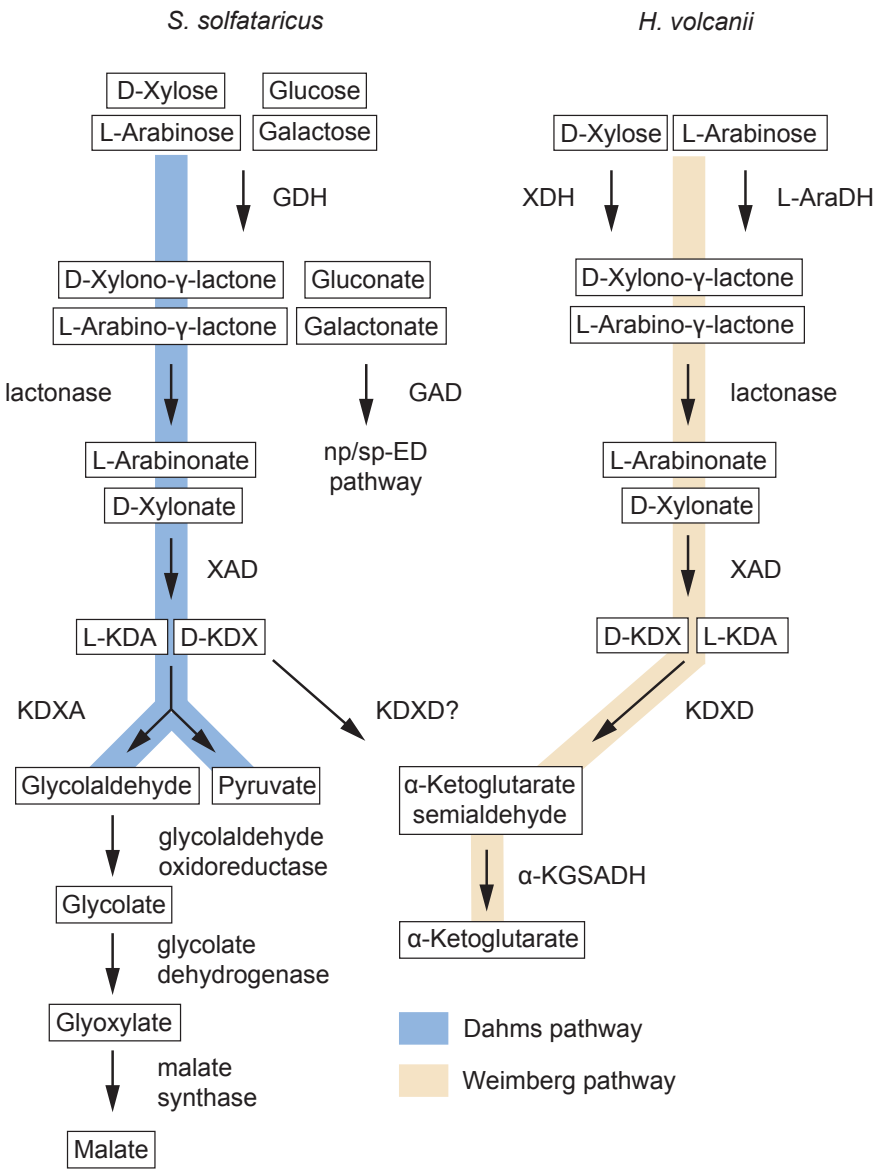
Hydrolysis of lignocellulosic biomass yields glucose and xylose as the most abundant monosaccharides available for bioconversion (Kobayashi and Fukuoka, 2013). Many microorganisms naturally use xylose as a growth substrate and several bacteria, yeast and fungi are fermenting xylose to ethanol (McMillan, 1993). The main applied research efforts in this area have been focused on strain improvement by incorporating other xylose metabolic pathways, reducing the amount of xylitol production and solving redox imbalances by modulating gene expression of the non-oxidative PP pathway (Moysés et al., 2016). In order to make use of most of the carbon in a lignocellulosic feedstock, simultaneous fermentation of glucose and xylose has been investigated increasingly in the last years (Oreb et al., 2012). Four different xylose degradation pathways have been described in microorganisms so far. Most bacteria use the xylose isomerase pathway converting xylose into xylulose-5-phosphate (X5P), which enters the CCM via the PP pathway. In a first step, D-xylose is isomerised to D-xylulose by xylulose isomerase, followed by phosphorylation to X5P by xylulokinase. The oxo-reductive pathway is present in fungi and some yeast and similarly to the XI pathway also yields X5P via phosphorylation of xylulose. However, in contrast to a single isomerisation, D-xylopyranose is reduced to xylitiol followed by an oxidation to xylulose, which is then phosphorylated. The Weimberg and Dahms pathways, which in contrast do not rely on phosphorylation, metabolise several pentose sugars including L-arabinose and D-xylose to intermediates that are used in the CCM e.g. the TCA cycle (Fig. 5). Instead of a conversion to the PP pathway intermediate (X5P), xylose is oxidised to the central TCA intermediate α-

ketoglutarate in the Weimberg pathway and to pyruvate and glycolaldehyde in the Dahms pathway (Dahms, 1974; Weimberg, 1961). The Weimberg and Dahms pathways share the initial oxidation of xylose to xylono-gamma-lactone and after spontaneous or enzymatic hydrolysis of the lactone, xylonate is converted to 2-keto-3-deoxy-D-xylonate (KDX) by a xylonate dehydratase. The processing of the 2-keto-3-deoxy-sugar acid differs in the Dahms and Weimberg pathways. In the Dahms pathway a KDX aldolase cleaves the C6 sugar into glycolaldehyde and pyruvate while in the Weimberg pathway a KDX dehydratase produces α-ketoglutarate semialdehyde which is then converted to the TCA intermediate α-ketoglutarate.

Only 40 years after the discovery of the Weimberg and Dahms pathways in *Pseudomonas* (Weimberg, 1961), genes and enzymes of the oxidative pathways were first identified for the metabolism of L-arabinose in *A. brasilense* (Watanabe et al., 2006a, 2006c) and of D-xylose in *C. crescentus* (Stephens et al., 2007). Stephens et al. identified the xylose operon by screening a transposon mutagenesis library, in which two differently annotated dehydrogenase genes, *xylA* and *xylB,* were identified. Recombinant expression in *E. coli* verified that the dehydrogenase activity was attributed to the *xylB* gene, while *xylA* encoded a α-ketoglutarate semialdehyde dehydrogenase, which generated α-ketoglutarate, the final product of the pathway. Other genes in the operon encoded for the enzymes xylonolactonase and the xylonate dehdyratase. The operon from *C. crescentus* is among the most used operons to convert production strains like *C. glutamicum* or *S. cerevisiae* into xylose utilising microbes (Radek et al., 2014; Salusjärvi et al., 2017; Wasserstrom et al., 2018). In particular the lactonase has been used in *S. cerevisiae* for the production of xylonate from xylose (Nygård et al., 2014; Toivari et al., 2012).

The xylose metabolism in archaea has remained unknown until a xylose-specific dehydrogenase in *Haloarcula marismortui* suggested an oxidative metabolism of the pentose sugar similar to that found in other non-phosphoryative pathways in archaea (Johnsen and Schönheit, 2004). The essential genes for a phosphorylative pathway via xylose isomerase and xylulose kinase are missing in *H. volcanii*, supporting the fact that the organism is only operating an oxidative xylose pathway. The identification of the complete oxidative xylose pathway in archaea was elucidated in *H. volcanii* by a combination of growth experiments on ^{13}C labelled xylose, DNA microarray analysis and enzyme activity assays (Johnsen et al., 2009). The metabolism of the archaeon was found to resemble the Weimberg pathway, which starts with the oxidation by a D-xylose

dehydrogenase, followed by the conversion to α-ketoglutarate by a xylonate dehydratase and two dehydration steps (Fig. 5). As part of the study, the genes for the oxidative pathway were identified and their corresponding enzymes, xylose dehydrogenase (XDH) and xylonate dehydratase (XAD) were characterised. In contrast to the dehydratase from *C. crescentus*, which was found to be a member of the IlvD/EDD superfamily, the xylonate dehydratase from *Haloferax volcanii* belonged to the enolase superfamily. Xylose pathways in both, *H. volcanii* and *S. solfataricus* share promiscuity with the pentose L-arabinose, which is further discussed in the following section and presented in Fig. 5.



GDH: glucose dehydrogenase, XDH: D-xylose dehydrogenase, L-AraDH, L-arabinose dehydrogenase, GAD: gluconate dehydratase, XAD: xylonate dehydratase KDGA: 2-keto-3-deoxygluconate aldolase, L-KDA: 2-keto-3-deoxy-L-arabinonate, D-KDX: 2-keto-3-deoxy-D-xylonate, KDXD: KDX dehydratase, α-KGSADH: α-ketoglutarate semialdehyde dehydrogenase

Figure 5 Overview of the Weimberg (brown) and Dahms (blue) pathways and promiscuous pathways for D-xylose and L-arabinose in *S. solfataricus* and *H. volcanii* GDH: glucose dehydrogenase, XDH: D-xylose dehydrogenase, L-AraDH, L-arabinose dehydrogenase, GAD: gluconate dehydratase, XAD: xylonate dehydratase, D-KDX: 2-keto-3-deoxy-D-xylonate, L-KDA: 2-keto-3-deoxy-L-arabinonate KDXA: KDX aldolase, KDXD: KDX dehydratase, α-KGSADH: α-ketoglutarate semialdehyde dehydrogenase.

Although the xylose mechanism discovered in *H. volcanii* shows a similar pattern to non-phosphorylative pathways like the np-ED, it is unique since the substrate is oxidised to α-ketoglutarate instead of pyruvate and glyceraldehyde.

8.1.2. L-arabinose metabolism

Together with xylose, arabinose is part of many types of lignocellulosic biomass. The degradation of the natural isomer L-arabinose in most bacteria works via an isomerase-dependent pathway (Van de Werken et al., 2008). However, in some bacteria and archaea, L-arabinose can be metabolised via the Weimberg, Dahms or pathway variations. The first genetic analysis of an oxidative L-arabinose pathway was performed in *A. brasiliense,* followed by studies in the archaeal species, *S. solfataricus*, *S. acidocaldarius* and *H. volcanii* (Brouns et al., 2006; Johnsen et al., 2013; Nunn et al., 2010; Watanabe et al., 2006c). The identification of a L-arabinose dehydrogenase indicated the existence of an oxidative pathway, which could be confirmed by the identification of the missing enzymes L-arabinonate dehydratase, L-2-keto-3-deoxyarabinonate (L-KDA) dehydratase, and L-arabinonolactonase including the last enzyme α-ketoglutaric acid semialdehyde dehydrogenase (Watanabe et al., 2006a, 2006b). Similar to the xylonate metabolism in *H. volcanii*, α-ketoglutarate is the final product of the pathway in *A. brasiliense* entering the TCA cycle. Unlike the xylonate dehydratase from *H. volcanii*, the N-terminal amino acid sequence of L-arabinonate dehydratase showed greater similarity to proteins belonging to the IlvD/EDD superfamily than those of the enolase superfamily and thus an evolutionary relationship between the L-arabinose pathway in *A. brasiliense* and the ED pathway and its variations has been suggested (Watanabe et al., 2006c).

The L-arabinose metabolism in the *Archaea* kingdom has been studied in *S. solfataricus*, *S. acidocaldarius* and *H. volcanii.* In large parts the archaeal pathways resemble the oxidative pathway found in *A. brasiliense*, with the difference of a higher level of substrate promiscuity and a few modifications in *Sulfolobus* (Fig. 5) (Brouns et al., 2006; Johnsen et al., 2013). Based on the activity of the glucose dehydrogenase with D-xylose and L-

arabinose, Nunn et al. suggested that glucose, galactose, L-arabinose and D-xylose are oxidised by the same dehydrogenase in *Sulfolobus* (Lamble et al., 2003; Nunn et al., 2010). However, the following dehydratation of the pentose sugar acids is performed by a C5 specific dehydratase which converts D-xylonate and L-arabonate to KDX and L-KDA, respectively. The end products of the D-xylose/L-arabinose pathway are glycolaldehyde and pyruvate. Pyruvate enters the CCM via the TCA, while two more enzymes convert glycolaldehyde into acetyl-CoA. Glycolaldehyde is further converted to malate via glycolaldehyde oxidoreductase and glycolate dehydrogenase (Nunn et al., 2010).

In contrast to the high level of promiscuity observed for C5 and C6 sugars in *S. solfataricus,* L-arabinose and D-xylose pathways are only partially promiscuous in *H. volcanii*. Here, different dehydrogenases are present for the oxidation of the pentoses and hexoses (Johnsen et al., 2013).

8.1.3. L-rhamnose metabolism

L-rhamnose is found in a variety of glycosides, mostly as a structural polysaccharide. In plants, L-rhamnose is part of seed cell walls as constituent of the polysaccharide rhamnogalacturonan (Kim et al., 2012). The mechanism of L-rhamnose utilisation has been studied in a number of organisms which can grow on L-rhamnose as a sole carbon source including *E. coli*, *T. acidophilum, Pichia stipitis, Pichia pastoris* and *A. niger* (Baldomà and Aguilar, 1988; Khosravi et al., 2017; Kim et al., 2012; Liu et al., 2016; Twerdochlib et al., 1994). At least three main pathways for the degradation of L-rhamnose in microorganisms have been described so far. The phosphorylative pathway, which is present in most bacteria, was discovered together with the catabolism of L-fucose in *E. coli* (Boronat and Aguilar, 1981). The respective pathway cluster contains genes encoding for an isomerase, a kinase and an aldolase. These enzymes convert L-rhamnose to L-lactaldehyde and DHAP. DHAP readily enters the glycolysis, whereas the fate of lactaldehyde is determined by the availability of oxygen. Under aerobic conditions L-lactaldehdye is converted to L-lactate, which can be further processed to pyruvate. However, under anaerobic conditions lactaldehyde is converted into the fermentation end product 1,2-propanediol, which is exported from the cell.

A non-phosphorylative L-rhamnose pathway, resembling the np-ED pathway for glucose in numerous enzymatic steps, has been described in eukaryotic organisms (Fig. 6). First

evidence for the non-phosphorylative rhamnose pathway was found in the yeast-like fungus *Pullularia pullulans* in which the enzymes L-rhamnonate dehydratase and 2-keto-3-deoxy-L-rhamnonate (L-KDR) aldolase were shown to convert rhamnose into L-lactaldehyde and pyruvate (Rigo et al., 1985). A similar pattern was observed in the fungi *P. stipitis* and *Debaryomyces hansenii,* (Twerdochlib et al., 1994). Based on phylogenetic relationships of genes involved in ED and other alternative sugar pathways, Watanabe et al. identified the L-rhamnose pathway cluster in the fungi *P. stipitis* and *D. hansenii* and found similar clusters also in bacteria such as *Azotobacter vinelandii* (Watanabe et al., 2008). The cluster consisted of the four-characteristic enzyme: L-rhamnose dehydrogenase, L-rhamnono-1,4-lactonase, L-rhamnonate dehydratase and L-KDR aldolase. They also found that Cluster of Orthologous Groups (COGs) for enzymes of different sugar pathways varied, indicating that the pathways do not originate from a single ancestor. However, the total number of the COG is fairly small for all enzymes of known alternative sugar pathways. This suggests that different pathways evolved from a combination of a small number of ancestral enzymes. Each enzyme of the L-rhamnose pathway was further characterised and high levels of substrate promiscuity has been observed. Among 16 different sugars tested the L-rhamnose dehydrogenases from *P. stipitis*, *D. hansenii* and *A. vinelandii* showed activity with L-rhamnose, L-lyxose, L-mannose and L-fucose, indicating that C2, C3 and C4 epimers can be oxidised by the dehydrogenase. In contrast, the L-rhamnono-γ-lactonase showed highly specific activity towards rhamnono-γ-lactone among 11 different sugar acids. The rhamnonate dehydratase was found to be similar to L-tartrate/galactarate dehydratase from *Salmonella typhimorium* and L-fuconate dehydratase from *Xanthomonas campestris.* Similar to the dehydrogenase, the L-rhamnonate dehydratase showed high promiscuity and displayed high activity towards L-lyxonate and L-rhamnonate while retaining activity towards L-mannonate and L-fuconate (Watanabe et al., 2008).

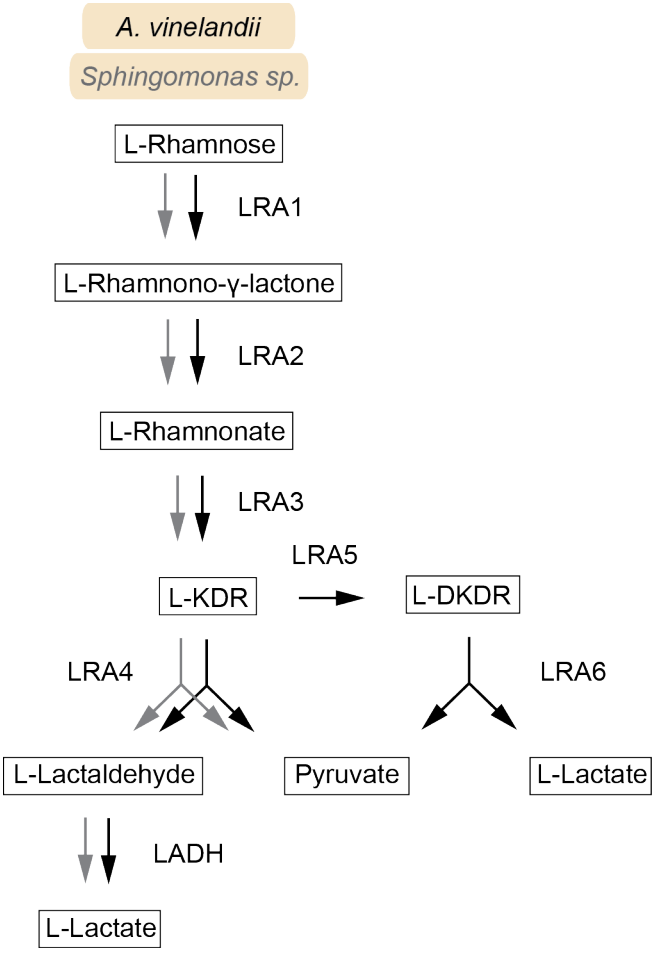


Figure 6 Overview of L-rhamnose metabolism in *A. vinelandii* (black lines) and *Sphingomonas* sp. (grey lines). LRA1, L-rhamnose-1-dehydrogenase, LRA2, L-rhamnono-lactonase, LRA3: L-rhamnonate dehydratase, LRA4: L-2-keto-3-deoxyrhamnonate aldolase, LRA5: L-2-keto-3-deoxyrhamnonate dehydrogenase, LRA6: L-2,4-diketo-3-deoxyrhamnonate hydrolase, LADH: L-lactaldehyde dehydrogenase, L-KDR: L-2-keto-3-deoxyrhamnonate, L-DKDR: L-2,4-diketo-3-deoxyrhamnonate

Other studies investigating the L-rhamnose pathway in *Sphingomonas* revealed two more genes (LRA5 and LRA6) in addition to the cluster found in *A. vinelandii*. The enzymes encoded by these genes resulted in a modification of the previously described non-phosphorylative L-rhamnonate pathway (Fig. 6 grey arrows). In species of *Sphingomonas* the intermediate product L-2-keto-3-deoxyrhamnonate can be converted via a dehydrogenase and a hydrolyase into pyruvate and L-lactate (Watanabe and Makino, 2009). Since the enzymes displayed strong similarity to already known dehydrogenase and dehydratases from L-arabinose and L-rhamnose pathways and the pathway seemed to be promiscuous, it was hypothesised that these enzymes evolved via gene duplication.

The gene cluster of L-rhamnose metabolism was compared in 32 fungal species. The cluster first described in *P. stipitis* is only present in an identical form in closely related species. Interestingly, the aldolase (LRA4), is not found as part of the cluster in many species. In contrast a transcription factor TRC1 is found to be a conserved part of the cluster in many species (Koivistoinen et al., 2012)

The study of promoters in the L-rhamnose cluster of *P. pastoris* helped to identify strong rhamnose-inducible promoters for expression of target proteins (Liu et al., 2016). Commonly used promoters for heterologous expression in *P. pastoris* are methanol inducible promoters, which bear disadvantages due to the hazardous and flammable nature of the inducer methanol. The induction with L-rhamnose overcomes these drawbacks and therefore displays a preferred option for the production of therapeutic and food-grade products.

8.1.4. L-fucose metabolism

L-fucose is a common component of N- and O-linked glycans and glycolipids present in mammals. The heoxose is highly abundant as part of glycans covering epithelial cells in the intestine, making this sugar an interesting target for medical research (Pacheco et al., 2012). L-fucose is also present in cell walls of plants, fungi and bacteria as part of their exopolysaccharide structure. Different species of seaweed contain fucodian, a sulfated polysaccharide structure consisting of L-fucose. (Vanhooren and Vandamme, 1999). Several species of gut bacteria attach to fucosylated glycans in the intestine (Hugdahl et al., 1988) and release fucosidases that cleave L-fucose monomers which then can be imported into the cell and used as a carbon source (Hooper et al., 1999). Although *E. coli* does not express fucosidases, an L-fucose metabolism still exists. In *E. coli*, L-fucose is metabolised similarly to L-rhamnose and converted to L-lactataldehyde via a phosphorylative pathway. This phosphorylative pathway for L-fucose has been verified in *E. coli* and *Bacteroides* (Cocks et al., 1974; Hooper et al., 1999). Similarly to hexoses in the glycolysis, an isomerase, a kinase and an aldolase are fundamental parts of the pathway (Baldomà and Aguilar, 1988). L-lactaldehyde is processed further depending on the presence of oxygen as described earlier (Cocks et al., 1974).

Comparative microarray studies identified a genomic island in *C. jejuni* which exhibited a six- to nine-fold upregulation upon growth on L-fucose (Stahl et al., 2011). However, only a gene encoding for a fucose transporter (*fucP*) and an aldolase (*fucA*) showed homology to known genes in the phosphorylative pathway for L-fucose in *E. coli*. Particularly, the absence of a gene encoding for a L-fucolose kinase (*fucK*) suggested the presence of a different pathway for L-fucose catabolism in *C. jejuni*. The existence of an alternative non-phosphorylative L-fucose pathway is supported by the discovery of a L-fuconate dehydratase in *X. campestris* (Yew et al., 2006). The enzyme was found to have promiscuous activity towards L-galactonate, D-arabinonate, D-altronate, L-talonate, and D-ribonate. Together with the identification of L-fucose pathway cluster a non-phosphorylative processing of L-fucose was suggested, which proceeds via similar pattern to previously mentioned L-rhamnose pathway in *Sphingomonas*. However, the 2-keto-3-deoxy sugar is supposed to be oxidised to 2,4-diketo-3-deoxy-L-fuconate, which can be cleaved into pyruvate and L-lactate.

Finally, a complete non-phosphorylative L-fucose pathway could be solved in a systematic approach in *S. solfataricus* (Wolf et al., 2016). The study compared metabolic intermediates, transcripts and phenotypes of cells grown on L-fucose to cells grown on D-glucose. Additionally, genes were identified and enzymes tested for activity. Loci previously identified to play a role in the metabolism of L-arabinose displayed up-regulated transcript levels in cells grown on L-fucose. Enzyme assays verified that L-fucose was metabolised via a promiscuous pathway, which also catabolises L-arabinose.

In summary, three different fucose degradation pathways have been described so far. The archaeal one, identified in *S. solfataricus*, the oxidative pathway in *X. campestris*, and the phosphorylative pathway in other bacteria.

8.1.5. D-mannose metabolism

Mannose occurs mostly in form of a homo- or heteropolymer found in the hemicellulose component of plants such as ivory nut, fenugreek, guar gums and coffee beans. Mannose is also highly abundant in the N-linked glycosylation of proteins (Sharma et al., 2014). In most organisms, D-mannose is phosphorylated to mannose-6-phosphate by a hexokinase and can be further processed depending on cellular needs. If mannose is used as a growth substrate and directed towards the EMP pathway, mannose-6-phosphate is isomerised to

fructose-6-phosphate by a phosphomannoisomerase. In eukaryotic cells, mannose can also be directed towards glycosylation via conversion by a phosphomannomutase to form α-D-mannose-1-phosphate. Since mannose is used less frequently to culture strains in microbiological practice, the research on mannose catabolism in microorganisms is not as extensive as for more common sugars like the ones previously discussed. Many bacteria and archaea contain hydrolytic enzymes to break down carbohydrate polymers containing mannose. α-mannosidases from bacteria including *Bacillus* spp. (Maruyama et al., 1994), *Thermotoga maritima* (Duffaud et al., 1997; Nakajima et al., 2006) and few archaea (Angelov et al., 2006) were investigated for their processing of glycosylated proteins in eukaryotic cells (Numao et al., 2000). In contrast β-mannanases and endo-1,4-β-mannosidases have been studied in a range of microorganisms for their hydrolytic activity on polymers containing 1,4-β-mannan such as those found in hemicellulose (Duffaud et al., 1997; Gibbs et al., 1996; Sunna, 2010; Sunna et al., 2000). However, no non-phosphorylative version similar to the np-ED, Weimberg or Dahms pathways has been identified so far for mannose. In *S. solfataricus*, RNAseq of L-fucose and D-glucose grown cells revealed a down-regulation of several genes that could be involved in a mannose metabolic pathway. For example, an α-mannosidase, a putative endo-1,4-β-mannosidase, an additional sugar dehydrogenase and a glyceraldehyde oxidoreductase showed lower transcription and expression levels (Wolf et al., 2016). Since *S. solfataricus* and other archaea, which operate on the np-ED pathway already contain KDGA, a promiscuous pathway with mannose as a substrate could be possible with the presence of only two enzymes. For a complete non-phosphorylative conversion starting from mannose, only a mannose-specific aldohexose dehydrogenase and a mannonate dehydratase are needed. In *S. solfataricus*, glucose dehydrogenase and gluconate dehydratase show activity with numerous substrates, but not with D-mannose and D-mannonate, respectively (Kim and Lee, 2005; Lamble et al., 2003). In addition, although highly promiscuous, dihydroxyacid dehydratase (DHAD) has been shown to work with a variety of different sugar acids, but not with D-mannonate (Kim and Lee, 2006).

However, dehydration of mannonate occurs in other sugar pathways, which could compensate missing enzymes in an oxidative mannose pathway. One example is the hexuronate metabolism, which has been studied in several bacteria and was described earlier in section 7.1.3. In the catabolism of hexuronates, a mannonate dehydratase (UxuA) converts mannonate to KDG, which is equivalent to the main intermediate of the np-ED

pathway. In mutant strains of *Rhodospseudomonas spheroides,* the compounds mannonate, KDG and traces of gluconic acid accumulate during growth on mannose, suggesting that a non-phosphorylative oxidation occurs (Szymona and Doudoroff, 1960). These results suggest the presence of a so far unknown alternative mannose metabolism.

9. Alternative carbohydrate pathways for metabolic engineering

The EMP pathway investigated in *E. coli* is considered a to be evolutionary optimised towards the simultaneous production of essential cellular intermediates and generation of energy (Bar-Even et al., 2012; Meléndez-Hevia et al., 1997; Noor et al., 2010). However, the objective in metabolic engineering, a highly efficient production of a desired compound, differs quite substantially from an evolutionary optimisation process which is mainly driven towards the maintenance of cellular energy and supply of cellular intermediates. That is why in the context of metabolic engineering the CCM is often 'trimmed' and carbon is directed towards product formation. Most importantly, several principles have been used to tweak metabolisms towards higher product yield, productivity and titer. High product yields can only be ensured if most of the carbon provided as substrate ends up as product. However, carbon loss usually occurs in the absence of a final electron acceptor (e.g. oxygen) and even under aerobic conditions (Radek et al., 2014). This can be observed in many common production strains which channel excess of reducing power towards undesired by-products, such as ethanol formation in *S. cerevisiae* and acetate production in *E. coli*. Reduced cofactors such as the availability of NAD(P)H are important for the generation of many compounds (e.g. isobutanol) in engineered strains (Blombach et al., 2011; Chemler et al., 2010). In addition, thermodynamics and pathway flux are important aspects to consider when high productivities need to be met. In most metabolic engineering approaches, a higher ATP supply in the cell is preferred since it promotes biomass formation and biosynthesis of metabolites (Hara and Kondo, 2015). However, high ATP generation in a pathway is often associated with a lower thermodynamic driving force and higher enzyme cost (Flamholz et al., 2013). Hence, metabolisms with high flux through the EMP pathway providing the cell with an excess of energy in form of ATP and certain intermediates might not be the most desirable option especially in cases where a high flux is desirable.

In contrast, diverting flux through pathways operating on a (semi-) oxidative principle rather than a phosphorylative principle can change the thermodynamic profile, NAD(P)H generation, lower enzyme amounts and reduce carbon loss (Flamholz et al., 2013; Ng et al., 2015). Since metabolic engineering has become a wide field of study, we focus on approaches in which especially elements of the (np-)ED and oxidative xylose pathways were used to overcome challenges and to develop improved production strains.

9.1. *In vivo* metabolic engineering approaches

9.1.2. NADPH production

The environmentally sustainable production of fuels has been a major focus of the biotechnology industry for many years. Bioethanol produced from lignocellulosic feedstock has been established as a biofuel and is currently produced on a commercial scale (Zabed et al., 2017). Among different types of biofuel, butanol features several advantages over ethanol such as a higher energy value, lower hygroscopy and the potential to be used for other applications besides fuel production. The *in vivo* isobutanol production has been demonstrated in various production hosts including *B. subtilis, S. cerevisiae, C. glutamicum and E. coli* by overexpressing and exploiting the valine biosynthesis pathway (Atsumi et al., 2008; Chen et al., 2011; Li et al., 2012; Smith et al., 2010). Recent engineering and systematic approaches of the keto-acid pathway lead to the notion that for high-yield production it was crucial to (a) replace acetolactate synthase and ketoisovalerate synthase with more efficient alternative enzymes and (b) to ensure NADPH availability since the key enzymes alcohol dehydrogenase and keto-acid reductoisomerase are dependent on reduced cofactor. The importance of balancing the cofactor level was demonstrated in the overexpression of the NADPH balancing pyridine nucleotide transhydrogenase (PntAB) in isobutanol producing *E. coli* and *C. glutamicum* strains (Blombach et al., 2011; Weckbecker and Hummel, 2004). PntAB reduces $NADP^+$ by oxidising NADH using a proton gradient over the cytoplasmic membrane (Jackson, 2003). Hence, high NADPH levels are ensured by change of flux through NADH generating reactions.

A study of glycolysis in an isobutanol producing *B. subtilis* strain showed that upon increase of glucose levels, more G6P and fructose-6-phosphate (F6P) accumulated compared to downstream metabolites such as FBP, GAP, 3PG and 2-phosphoglycerate.

This indicated a bottleneck in the conversion of F6P to FBP, which marks a fundamental step in glycolysis and for the downstream production of isobutanol. An *E. coli* strain was engineered to circumvent this problem, which contained an ED pathway from *Z. mobilis.* The introduction of ED pathway required less enzymes and produced more NADPH, which was crucial for a NADPH-dependent production of isobutanol. Compared to the original *E. coli* isobutanol production strain, the new strain containing the *eda* and *edd* genes from *Z. mobilis* showed an increase of 56.8%, 47.4% and 88.1% in titer, yield and production rate, respectively (Liang et al., 2018). The study provides an example for the construction of a production pathway in which the generation of reduced cofactors plays a major factor in product formation. The availability of NADPH can also significantly improve the productivity and yield of other production pathways relying on NADPH, like those for polyphenols, leucocyanidin, (+)-catechin (Chemler et al., 2010) and xylitol (Chin and Cirino, 2011).

In most cellular metabolism, reduced cofactor is not primarily produced through the EMP pathway but mainly by enzymes of the PP pathway and TCA cycle. Yields in production strains could be improved by either overexpressing NADPH producing enzymes or by diverting flux from the EM to the PP pathway by deletion of phosphoglucose isomerase (*pgi)* (Marx et al., 2003) or phosphofructokinases (*pfk)* genes (Siedler et al., 2011). For example, in *E. coli* the coproduction of hydrogen and ethanol from glucose was improved by diverting the flow through the PP pathway, since only 2 mol NADH can be produced from 1 mol glucose via glycolysis. In contrast, an average of 3.67 mol NAD(P)H were produced from 1 mol glucose when glucose was metabolised via the PP pathway (Sundara Sekar et al., 2017).

Diverting flux through the PP pathway by knockout of *pfk* or *pgi* is an established method to increase the NADPH level in the cell, which result in higher product formation for production pathways depending on high amounts of reduced cofactor. However, in doing so, a substantial amount of the carbon dissipates in form of CO_2 via the PP pathway. To overcome this problem, Ng et al. directed flux through the ED pathway, which recycles NADPH without carbon loss (Ng et al., 2015). Since engineering the endogenous ED pathway of *E. coli* is hindered by its strong and multi-leveled regulation, the complete ED pathway from *Z. mobilis* was instead reconstructed in *E. coli*. Two operons consisting of 5 genes from *Z. mobilis* were constructed using the 'Operon Calculator' software. The first operon consisted of genes encoding for the enzymes ZWF and PGI and the second of the

genes for the main pathway enzymes EDD, EDA and the 6-phosphogluconolactonase. The initial production strain containing a complete ED pathway exhibited already a 4.8-fold improvement of NADPH regeneration over the parent *E. coli* strain. Different enzyme expression levels were tested by genome mutagenesis, and yielded a variant with 25-fold increase in NADPH levels compared to the parent strain. The enhanced NADPH levels were tested in a terpenoid production strain and resulted in a 97% higher product yield. The authors suggested that, in principle, the ED operon could function as a 'drop-in module' which can be applied to any production strain, that benefits from higher NADPH levels (Ng. et al. 2015)

The incorporation of the ED pathway was also shown to be beneficial for the production of poly(3-hydroxybutyrate) (PHB) in *E. coli.* PHB is one of the most common examples for polyhydroxyalkanoates (PHAs), a group of polyesters that provides a more sustainable alternative to fossil fuel-derived plastics. PHAs can be produced via fermentation and offer a range of advantageous properties like thermoplasticity and biodegradability (Gao et al., 2011). Crucial to a high production of PHBs in most bacteria is the availability of acetyl-CoA and reduced cofactor NAD(P)H. PHB synthesis can be established in *E. coli* via recombinant expression of PHB synthesis genes. Proteome analysis of PHB-producing recombinant *E. coli* strains showed that enzymes of the glycolytic and ED pathways were present in increased levels, suggesting that a high glycolytic flux through these enzymes is important for biosynthesis of PHB (Han et al., 2001).

Zhang et al established a recombinant *E. coli* strain producing PHB via a L-serine pathway (Zhang et al., 2014). However, additional overexpression of the ED pathway genes only resulted in increased growth rates but not higher PHB yields. Only a combination of the expression of a) L-serine biosynthesis genes b) ED pathway genes and c) pyruvate dehydrogenase complex (PDH) significantly increased PHB levels, demonstrating that an additional glycolytic bypass via pyruvate can further increase flux towards the end product. As a comparison, the resulting strain with an engineered serine deaminase pathway, ED pathway and overexpressed PDH complex achieved 81.1% PHB content. In previous strains increasing NADPH by channeling flux through PP pathway (e.g. by overexpressing a transaldolase), achieved only 52.3% PHB content (Song et al., 2006).

9.1.3. Changing ATP supply

In addition to recycling NAD(P)+, ATP production and availability plays an important role in metabolic engineering. ATP production is essential to maintain cell growth, but is also dependent on the organism's environment and its nutrient availability. It was hypothesised earlier that cyanobacteria in nutrient-depleted environments depend less on ATP synthesis via the EMP pathway, but instead rely on ED pathway and photosynthetic ATP production (Chen et al., 2016). Similarly, organisms with a high glycolytic flux (e.g. *Z. mobilis*) do not rely on high ATP production via the EMP, but instead use the ED pathway. In metabolically engineered strains, ATP has to be maintained and often compensated by and additional ATP producing reaction. However, an excess production of ATP in a production strain causes an increase in biomass at the expense of target product formation (De Kok et al., 2012). This can be observed when ethanol yields in yeast strains fermenting maltose are compared to those fermenting glucose. The symport of maltose is accomplished by proton motive force over the membrane, which is maintained by an H^+-ATPase causing a lower net ATP yield in the cell in comparison to cells grown on glucose. The lower ATP concentration in maltose grown yeast strains results in a 25% decrease in biomass, but a 9% increase in ethanol production over glucose-grown cells (Weusthuis et al., 1993).

Since the ED pathway and its variations yield less ATP/mole glucose compared to the EMP pathway, substituting the EMP pathway with the ED pathway could potentially increase ethanol production in yeast strains fermenting glucose. This was attempted by Benisch and Boles by inserting a complete ED pathway in a *S. cerevisiae* strain with a knocked-out EMP and PP pathway (Benisch and Boles, 2014). Disruption of the EMP and PP pathways was accomplished by the deletion of the key enzyme PFK and 6-phosphogluconate dehydrogenase (GND), while establishing an ED pathway was attempted by expression of the *eda* and *edd* genes. However, the heterologous expression of the ED pathway enzymes turned out to be a challenging task and none of the three tested EDAs from *E. coli*, *Z. mobilis* and *P. putida* showed sufficient expression levels. Only after codon optimisation, the *E. coli* dehydratase could be expressed in detectable amounts. Yet, the enzyme displayed only very low activity (2.7 mU/mg cell extract) in transformed *S. cerevisiae*. Since the EDA from *E. coli* is a [4Fe-4S] cluster-containing enzyme, it was suggested that the low availability of essential metal ions in the yeast's cytosol restricted a sufficient loading of the enzyme with FeS resulting in low activity. However, further attempts to increase the activity of the enzyme by directing the enzyme folding into the more iron-loaded mitochondria or increasing the bioavailability of iron in the yeast cell were

unsuccessful. This example shows that in theory only two enzymes (EDD and EDA) are required to redirect flux from the EMP pathway into a metabolism that ferments glucose primarily via the ED pathway. However, the *in vivo* approach is hampered by the incompatibility of archaeal genes and yeast as an expression host.

9.1.4. Oxidative xylose pathway in vivo metabolic engineering

On an industrial level, xylose fermentation is mainly used for the production of ethanol or xylitiol. However, in the last decade xylonate has gained attention due to its potential applications in food, chemical and pharmaceutical industries (Toivari et al., 2012). Xylonate production via the oxidative route displays a good example for how oxidative pathways can serve as a 'low in enzyme-cost' conversion of a biomass into a valuable product. In the oxidative Weimberg and Dahms pathways (discussed in sections 8.1.1 - 8.1.3) xylose is oxidised by a dehydrogenase to xylonolactone. Lactone hydrolysis to yield xylonate occurs spontaneously or is accomplished at a faster rate by a lactonase, potentially yielding the desired product xylonate in only one or two enzymatic steps.

Microbial oxidation of xylose represents an alternative to conventional chemical oxidation with palladium catalysts and conversion with purified glucose oxidases, which show promiscuous activity towards xylose. Pure enzymes can be immobilised and produce organic acids at a high purity level. However, the product of oxidation is acidic which often results in decreased activity or denaturation of pure enzymes (Ramachandran et al., 2006). While metal catalyst and free enzyme oxidation are highly efficient on pure xylose, oxidation by organisms such as *Gluconobacter oxydans* can be easily performed in mixed biomass hydrolysates. *Gluconobacter oxydans* and *Pseudomonas fragi* naturally metabolise xylose in an oxidative fashion and have been exploited for the production of xylonate (Buchert et al., 1988). Although these strains showed a conversion rate of 90%, a further increase in productivity suffers from inhibitory compounds in the hydrolysates. Common cell factories such as *E. coli* and *S. cerevisiae* have been transformed with genes enabling an efficient oxidative xylose metabolism for the production of xylonate (Cao et al., 2013). Growth on xylose via an oxidative pathway has already been achieved in *P. putida.* Interestingly, only the expression of a xylonate dehydratase was indispensable for growth on D-xylose (Meijnen et al., 2009). The biomass yield could be further increased by more than 10% by additional expression of the remaining genes of the xylose operon from *C. crescentus*, including genes for KDX dehydratase and α-ketoglutaric semialdehyde

dehydrogenase. The same xylose operon was used to engineer a more carbon-efficient *C. glutamicum* production strain that grows on xylose (Radek et al., 2014). *Corynebacterium glutamicum* is used as production strain for amino acids at an industrial scale and establishing an oxidative xylose pathway is advantageous in several ways. For example, the end product of the Weimberg pathway is α-ketoglutarate, an important intermediate for amino acid production, while carbon loss can be lowered by catabolism of xylose via the Weimberg pathway. The *C. glutamicum* strain harbouring the oxidative xylose pathway showed several distinct growth stages. During consumption of xylose, xylonate levels increased transiently, but decreased again in the final growth stage when substrate became limiting again, suggesting that the dehydration of xylonate to KDX is the rate-limiting step in the engineered Weimberg pathway. Besides the accumulation of xylonate, xylitol was found as a byproduct of endogenous reactions during growth on xylose, which was further identified to act as growth inhibitor when converted to X5P (Radek et al., 2016). This demonstrates that an oxidative xylose metabolism is a good choice for the construction of a xylose-utilising production strain, since cytotoxic effect of X5P and other intermediates of the phosphorylative xylose pathway are not accumulating.

Similar challenges are observed for the production of xylonate in *S. cerevisiae.* Spontaneous hydrolysis of xylonate occurs only slowly, which is why enzymatic hydrolysis by a xylonolactonase leads to higher initial production of xylonate. However, the cleavage of xylonolactone has been shown to decrease the pH to a level which is in turn lowering the vitality of cells after a high initial production rate of xylonate (Nygård et al., 2014).

In another approach, a complete oxidative xylose pathway has been used for the production of ethylene glycol or glycolic acid in *S. cerevisiae* (Salusjärvi et al., 2017). The selected dehydratase from *C. crescentus* belonged to the IlvD/EDD enzyme family, which is known to be dependent on a FeS cluster. As already seen in the expression of EDD for the reconstitution of the ED pathway in *S. cerevisiae,* proper folding and loading of an FeS-dependent enzyme in eukaryotes can be troublesome (Benisch and Boles, 2014). However, the activity of the xylonate dehydratase from *C. crescentus*, which is also an iron-sulfur protein, could be increased 27-fold by targeting its expression to the mitochondria in a strain lacking an iron regulon repressor (Δfra2) (Salusjärvi et al., 2017).

A promising strategy to achieve the decoupling of growth from product formation in one organism is to establish orthogonal pathways for each purpose with a minimised cross talk.

A cellular chassis for *E. coli* has been used to implement a *de novo* pathway for the production of 1,4-butanediol from xylose (Liu and Lu, 2015). 1,4-butanediol is one of many important platform chemicals used in the chemical industry for the production of plastic, fibres and pharmaceuticals. The pathway is based on the oxidative xylose metabolism as found in *C. crescentus* and is mostly independent from the cellular metabolism of *E. coli*. At first a cellular chassis was created by knockout of the D-xylose isomerase and KDX aldolase genes in order to disable the host from naturally growing on xylose. Then a *de novo* pathway was constructed based on XDH and XAD from *C. crescentus*, KDX dehydratase from *H. volcanii*, and keto-acid decarboxylases from *Lactobacillus lactis* and *P. putida*. Crosstalk between the xylose pathway for the production of 1,4-butanediol and CCM of the cell could be minimised by disrupting the D-xylose isomerase gene (*xylA*) which normally allows xylose to enter the PP pathway of the cell. In total, the *de novo* xylose pathway featured three oxydoreductions, which were dependent on one mole of reduced cofactor from the cellular chassis. The dependence of cofactors from the cell's metabolism indicated that the production pathway is not entirely independent from the cell metabolism. However, the implementation of the *de novo* xylose pathway resulted in lower levels of acetic acid production and no fermentation to ethanol, which indicated the independence of the production pathway.

For a comprehensive summary of metabolic engineering of the xylose oxidative pathway, we refer the reader to the recently published review by Valdehuesa et al. (Valdehuesa et al., 2018).

9.2. Cell-free engineering approaches

Metabolic engineering has developed into a rapidly advancing field with an increasing number of microorganisms to be exploited as microbial factories. Although synthetic biology approaches and computational tools allow the design and construction of more complex and sophisticated networks *in vivo*, a comprehensive fully modular and predictive system similar to those found in synthetic chemistry has not been achieved yet. However, in a cell-free metabolic engineering approach, many limitations which are encountered in living cells can be overcome by an almost unrestricted combination of enzymes in a single or sequential reaction(s) (Dudley et al., 2015). Enzymatic pathways can be assembled from crude lysates of recombinant cells, purified enzymes, immobilised and/or commercially

available enzymes (Dudley et al., 2016; Gao et al., 2015; Hold et al., 2016; You and Zhang, 2012). In all cases, the *in vitro* approach facilitates quick testing of a pathway by changing its enzymes, substrates, cofactors or reaction conditions. In addition, in combination with a high-throughput analytical method, the *in vitro* approach enables fast pathway prototyping and testing (Hold et al., 2016). Although cell-free metabolic engineering has been proven to be suitable for testing design and feasibility of conversions, it is still facing challenges when scaling up production processes. The main challenges that have to be overcome in order to increase the attractivity for cell-free metabolic engineering for larger scale productions are efficient soluble production of enzymes, their performance, stability and potential resuability in repeated production cycles. In the future, immobilisation of whole cell-free pathways could lower the cost of enzymes and in turn decrease the overall cost for cell-free production of platform chemicals (Petroll et al. 2018). High-throughput screening tools combined with the low price of gene synthesis can be used to identify yet undiscovered enzymes with better characteristics. Additionally, enzyme engineering can help to improve specificity, kinetics and robustness of each catalyst.

For the construction of cell-free pathways, all cofactors that are normally produced by the endogenous metabolism of the cell have to be either provided by enzymes or supplied externally (Welch and Scopes, 1985). This challenge can be overcome by choosing alternative enzymes to circumvent the use of cofactors or by balancing cofactors via implementation of additional reactions or complete cofactor salvaging systems (Beer et al., 2017; Honda et al., 2016; Nowak et al., 2015; Opgenorth et al., 2014). As mentioned earlier, one of the main purposes of sugar utilisation in the EMP pathway is the generation of reducing power in the form of ATP (Flamholz et al., 2013). However, in a minimal cell-free environment, ATP is not required to sustain energy and the generation of different cofactors should be kept at a minimum. In the conventional EMP pathway, 2 molecules of ATP and 2 molecules of NAD(P)H are formed by converting FBP into 3PG via GAPDH and the phosphoglycerate kinase (PGK). However, in variations of the EMP pathway found in archaea (section 2.3), this step is accomplished by a non-phosphorylative glycerlaldehyde-3-phosphate dehydrogenase (GAPN) (Brunner et al., 1998; Ettema et al., 2008; Matsubara et al., 2011).

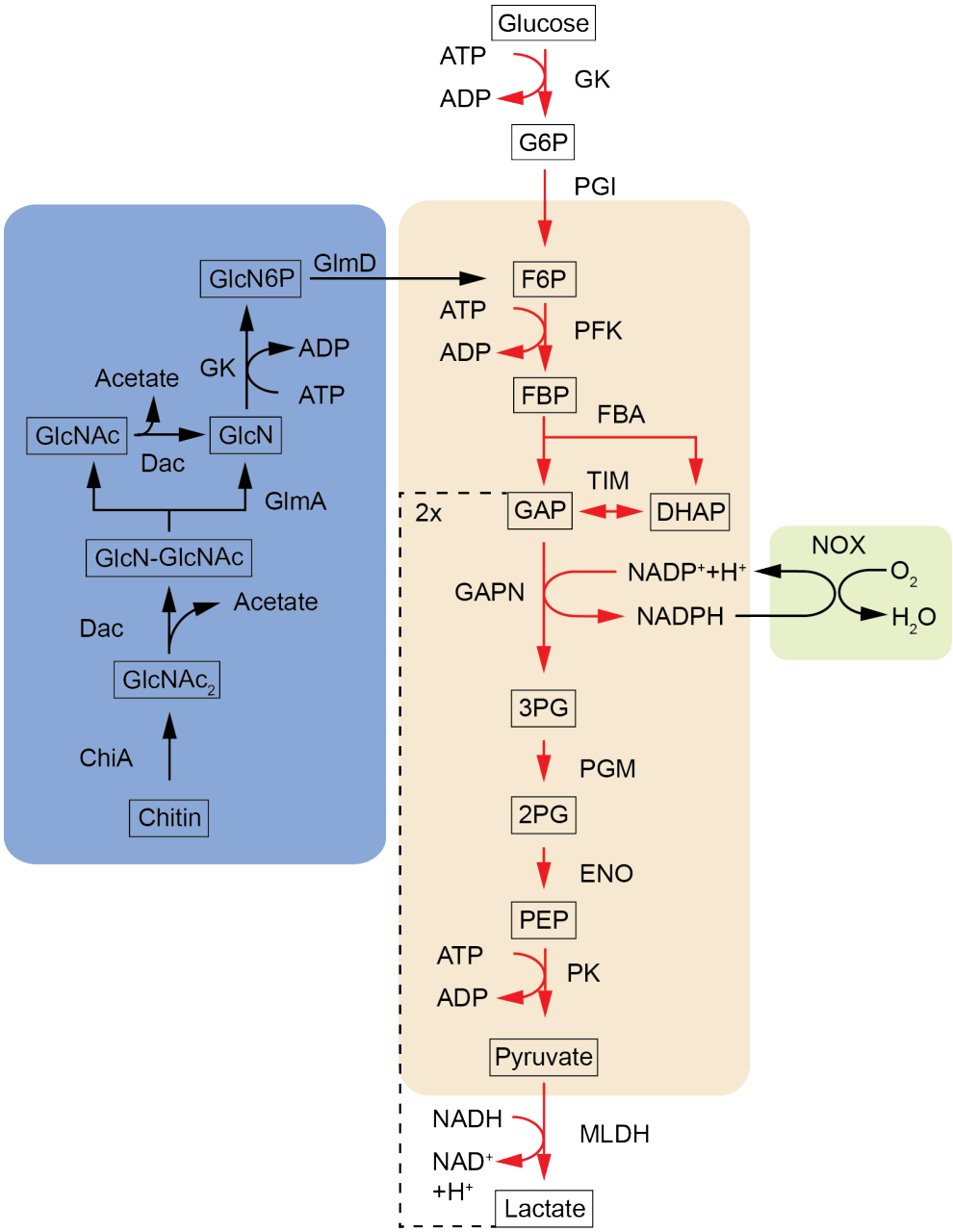


Figure 7 Examples of cofactor balanced cell-free glycolytic pathways adapted from Ye et al. 2012 and Honda et al. 2017. A glycolytic core module (brown) including GAPN instead of a GAPDH/PGK combination was used for the ATP-balanced conversion of glucose to lactose (red arrows). A combination with a chitin degradation module (blue) and a cofactor recycling module (green) enabled the conversion of colloidal chitin to pyruvate (black arrows). Glycolytic enzymes and pathway intermediates are as indicated in Fig. 2B. Other enzymes are as follows: ChiA: chitinase, Dac: diacetylchitobiose deacetylase, GK: glucokinase, GlmD: glucosamine-6-phosphate deaminase, NOX: NADH oxidase, malate/lactate dehydrogenase. Metabolites are: GlcN: glucosamine, GlcNAc: N-acetylglucosamine, GlcN6P: glucosamine-6-phosphate.

Ye et al. developed a chimeric non-ATP forming version of the EMP pathway by exchanging the GAPDH/PGK enzyme pair with GAPN and thereby circumventing the production of excess ATP (Fig. 7, red arrows) (Ye et al., 2012). Not only did the use of

GAPN solve the problem of ATP imbalance but it also reduced the amount of thermolabile triosephosphates. Since the whole pathway is based on thermophilic enzymes from *Thermus thermophilus, T. kodakarensis* and *Pyrococcus horikoshii*, an operating temperature of 50°C could be achieved. In order to test the feasibility of the pathway a malate/lactate dehydrogenase from *T. thermophilus* was introduced, which resulted in the production of lactate from glucose with a turnover of 31 ATP. Although enzyme activities were stable over 10 h in the reaction, decomposition of NAD^+ and NADH was observed. In particular GAPN, which has a high *K*m for NAD^+, showed a decrease in activity upon decomposition of the cofactor leading to the accumulation and eventual break down of the heat labile intermediates GAP and DHAP. However, a 100% yield of lactate could be reached when small amounts of NADH were supplemented during the reaction.
The scope of this chimeric pathway was further extended by using colloidal chitin instead of glucose as substrate for the production of pyruvate (Fig. 7, blue and green modules) (Honda et al., 2017). Chitin is a major component of the exoskeleton of crustaceans and fungal cell walls and is among the most abundant organic compounds on earth (Gooday, 1990). The polymer consists of a chain of β-1,4-linked-N-acetylglucosamine subunits that can be hydrolysed by chitinases yielding oligosaccharides of glucosamine. The previously designed chimeric pathway was connected to a chitin degradation pathway from *T. kodakarensis,* which consisted of 5 more enzymes (chitinase, diacetylchitobiose deacetylase, β-D-glucosaminidase, glucokinase and glucose-6-phosphate deaminase). Although the whole pathway was cofactor balanced, degradation of ATP could still be observed. Since only semi-purified enzymes were used for the pathway, it was suggested that endogenous ATPdegrading enzymes were active during the reaction. Addition of two more enzymes, adenylate kinase and polyphosphate kinase, regenerated ATP from polyphosphate and AMP, which resulted in an increase of the titer from 0.62 mM to 2.1 mM (Honda et al., 2017).

A commonly encountered challenge in the conversion of biomass into platform chemicals is the incompatibility of thermophilic enzymes used for hydrolysis and mesophilic growth temperatures of production strains such as *S. cerevisiae*, *C. glutamicum* or *E. coli* (Mutturi and Lide, 2013). The chitin degradation pathway overcomes the problem by combining the two elements at similar temperatures, since enzymes for hydrolysis and production of a chemical are derived from a thermophilic background and combined in an unrestricted cell-free environment.

Implementation of complete pathways in microbial hosts can be restricted by the challenge of expressing the desired genes *in vivo* in a functional manner. Particularly with pathways from different backgrounds (e.g. archaea), proper enzyme folding and low enzymatic activity inside the host may ultimately restrict the functional operation of a new pathway (Benisch and Boles, 2014; Wasserstrom et al., 2018). Construction of pathways in a cell-free environment allows for more flexibility and engineering freedom. Especially low enzyme activities can be compensated by addition of higher enzyme amounts in the cell-free process. For examples, the advantageous features of the np-ED pathway were used in the design of a cell-free pathway for the production of ethanol and isobutanol from glucose (Guterl et al., 2012). The reaction cascade established by Guterl et al. consisted of two main parts. In the first half of the conversion one mole glucose was oxidised to two moles of pyruvate via a cofactor balanced pathway based on enzymes of the np-ED pathway from *S. solfataricus S. acidocaldarius* and *T. acidophilum*. In the second half of the conversion, two molecules of pyruvate were converted either into one molecule isobutanol or two molecules ethanol depending on the enzymes added. In the natural np-ED pathway found in *S. solfataricus* and *T. acidophilum*, the key intermediate KDG is split into glyceraldehyde and pyruvate. However, in the modified cell-free system, two more enzymes were introduced to recycle glyceraldehyde back to pyruvate. In a conventional glycolytic pathway, 13 enzymes, NADH and defined concentrations of ATP are necessary to convert glucose to ethanol. In contrast, the oxidative pathway design based on the np-ED pathway converted glucose to ethanol via 6 enzymes and a single cofactor (Guterl et al., 2012; Welch and Scopes, 1985). A similar strategy and enzymes were used to produce lactate from glucose via pyruvate (Xie et al., 2018). The cofactor-balanced pathway consists of 5 enzymes, which was optimised towards maximum yield by adjustment of enzyme loadings, buffer strength and the amount of added cofactor. Using the cell-free pathway, L-lactate was produced from pyruvate at ~90% yield.

The reconstruction of the np-ED pathway for the production of ethanol, isobutanol and lactate demonstrates the suitability of oxidative pathways in particularly for a cell-free engineering strategy. The need for a minimal amount of cofactors and reduced amount of enzymes, compared to phosphorylative pathways such as the EMP pathway, makes oxidative pathways a suitable template for the construction of cell-free conversions. The variety of oxidative non-phosphorylative pathways is not limited to glucose but also other sugars that are highly abundant in biomass, such as xylose, arabinose, mannose and

galactose can be metabolised in a similar way. So far, it has not been demonstrated that the oxidative pathways for these sugars represent a feasible alternative for the production of platform chemicals through a cell-free conversion.

Cell-free pathway engineering is developing towards a truly synthetic approach for metabolic pathway design. In the design space, the only limitations are set by stoichiometry, thermodynamics and influence of the intermediates and cofactors brought into or produced within the system. Therefore, cell-free approaches enable the implementation and combination of pathways found in nature, pathways with novel enzymes and pathways with novel chemistries (Erb et al., 2017).

10. Conclusion

In this overview we highlighted the diversity and abundance of carbohydrate pathways present in microorganisms with an emphasis on the ED pathway. Numerous studies with non-model organisms from different phylogenetic backgrounds suggested that the ED pathway plays an important role in phylogenetically distant organisms (Chen et al., 2016; Fuhrer et al., 2005). Despite a lower energetic yield in terms of ATP formation, the non-phosphorylative strategy present in the np-ED, Weimberg and Dahms pathways seems to be beneficial for the metabolism of different sugars. Sugar dehydrogenases, sugar acid dehydratases and aldolases are key enzymes in those pathway and may give indications to hidden carbohydrate metabolisms (Lamble et al., 2003; Theodossis et al., 2004; Wichelecki et al., 2014b; Wolf et al., 2016; Wolterink-van Loo et al., 2007). As a result of the continuous increase in sequencing power and a decrease in cost, more genetic information is being constantly discovered which require more powerful computational tools for the identification of genes and their hypothetical proteins. Over the last decade the extensive use of computational prediction decreased the accuracy in annotations especially for large enzyme superfamilies (Schnoes et al., 2009). Manual curation, specifically for enzymes and their substrate specificity, is necessary to identify hidden functions and metabolisms. As seen in discoveries for previous pathways, comparative metabolomics and transcriptomic studies can aid in the discovery of new pathways and uncover complex metabolic networks (Wolf et al., 2016). In the intersection between synthetic biology and metabolic engineering, modularity and standardisation of parts

including carbohydrate modules is highly desirable, especially with a shift towards higher complexity in pathways. The design, construction and testing of synthetic pathways in a cell-free manner enables rapid prototyping of novel synthetic pathways or verification of hypothetical natural ones.

Acknoledgements

DK is supported by an international Macquarie University Research Excellence Scholarship (iMQRES).

References

Aden, A., Ruth, M., Ibsen, K., Jechura, J., Neeves, K., Sheehan, J., Wallace, B., Montague, L., Slayton, A., Lukas, J., 2002. Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. No. NREL/TP-510-32438. Natl. Renew. energy lab golden co.

Ahmed, H., Ettema, T.J.G., Tjaden, B., Geerling, A.C.M., van der Oost, J., Siebers, B., 2005. The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: a re-evaluation. Biochem. J. 390, 529–540.

Albers, S.-V., Elferink, M.G.L., Charlebois, R.L., Sensen, C.W., Driessen, A.J.M., Konings, W.N., 1999. Glucose transport in the extremely thermoacidophilic *Sulfolobus solfataricus* involves a high-affinity membrane-integrated binding protein. J. Bacteriol. Bacteriol. 181, 4285–4291.

Angelov, A., Putyrski, M., Liebl, W., 2006. Molecular and biochemical characterisation of α-glucosidase and α-mannosidase and their clustered genes from the thermoacidophilic archaeon *Picrophilus torridus*. J. Bacteriol. 188, 7123–7131.

Armbrust, E.V., 2009. The life of diatoms in the world's oceans. Nature 459, 185–192.

Aslan, S., Noor, E., Bar-Even, A., 2017. Holistic bioengineering: rewiring central metabolism for enhanced bioproduction. Biochem J 474, 3935–3950.

Atsumi, S., Hanai, T., Liao, J.C., 2008. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. Nature 451, 86–89.

Babbitt, P.C., Gerlt, J., 1997. Understanding enzyme superfamilies - Chemistry as the fundamental determinant in the evolution of new catalytic activities. J. Biol. Chem. 272, 30591–30594.

Baldomà, L., Aguilar, J., 1988. Metabolism of L-fucose and L-rhamnose in *Escherichia coli*: aerobic-anaerobic regulation of L-lactaldehyde dissimilation. J. Bacteriol. 170, 416–421.

Bar-Even, A., Flamholz, A., Noor, E., Milo, R., 2012. Rethinking glycolysis: on the biochemical logic of metabolic pathways. Nat. Chem. Biol. 8, 509.

Baumann, L., Baumann, P., 1973. Enzymes of glucose catabolism in cell-free extracts of non-fermentative marine eubacteria. Can J Microbiol 19, 302–304.

Beer, B., Pick, A., Sieber, V., 2017. *In vitro* metabolic engineering for the production of α-ketoglutarate. Metab. Eng. 40, 5–13.

Bender, R., Gottschalk, G., 1973. Purification and properties of d-gluconate dehydratase from *Clostridium pasteurianum*. Eur. J. Biochem. 40, 309–321.

Benisch, F., Boles, E., 2014. The bacterial Entner-Doudoroff pathway does not replace glycolysis in *Saccharomyces cerevisiae* due to the lack of activity of iron-sulfur cluster enzyme 6-phosphogluconate dehydratase. J. Biotechnol. 171, 45–55.

Blombach, B., Riester, T., Wieschalka, S., Ziert, C., Youn, J.W., Wendisch, V.F., Eikmanns, B.J., 2011. *Corynebacterium glutamicum* tailored for efficient isobutanol production. Appl. Environ. Microbiol. 77, 3300–3310.

Bornscheuer, U.T., Kazlauskas, R.J., 2004. Minireviews Reaction Specificity of Enzymes Catalytic Promiscuity in Biocatalysis : Using Old Enzymes to Form New Bonds and Follow New Pathways. Angew. Chemie - Int. Ed. 43, 6032–6040.

Boronat, A., Aguilar, J., 1981. Metabolism of L-fucose and L-rhamnose in *Escherichia coli*: Differences in induction of propanediol oxidoreductase. J. Bacteriol. 147, 181–185.

Bowden, S.D., Rowley, G., Hinton, J.C.D., Thompson, A., 2009. Glucose and glycolysis are required for the successful infection of macrophages and mice by *Salmonella enterica* serovar typhimurium. Infect. Immun. 77, 3117–3126.

Bräsen, C., Esser, D., Rauch, B., Siebers, B., 2014. Carbohydrate metabolism in archaea: current insights into unusual enzymes and pathways and their regulation. Microbiol. Mol. Biol. Rev. 78, 89–175.

Brouns, S.J.J., Walther, J., Snijders, A.P.L., Van De Werken, H.J.G., Willemen, H.L.D.M., Worm, P., De Vos, M.G.J., Andersson, A., Lundgren, M., Mazon, H.F.M., Van Den Heuvel, R.H.H., Nilsson, P., Salmon, L., De Vos, W.M., Wright, P.C., Bernander, R., Van Der Oost, J., 2006. Identification of the missing links in prokaryotic pentose oxidation pathways: Evidence for enzyme recruitment. J. Biol. Chem. 281, 27378–27388.

Brunner, N.A., Brinkmann, H., Siebers, B., Hensel, R., 1998. NAD+ - dependent glyceraldehyde-3-phosphate dehydrogenase from *Thermoproteus tenax*. Jour. Biol. Chem. 273, 6149–6156.

Buchert, J., Puls, J., Poutanen, K., 1988. Comparison of *Pseudomonas fragi* and *Gluconobacter oxydans* for production of xylonic acid from hemicellulose hydrolyzates. Appl. Microbiol. Biotechnol. 28, 367–372.

Budgen, N., Danson, M.J., 1986. Metabolism of glucose via a modified Entner-Doudoroff pathway in the thermoacidophilic archaebacterium *Thermoplasma acidophilum*. FEBS Lett. 196, 207–210.

Cao, Y., Xian, M., Zou, H., Zhang, H., 2013. Metabolic engineering of *Escherichia coli* for the production of xylonate. PLoS One 8, e67305.

Carsten, J.M., Schmidt, A., Sieber, V., 2015. Characterisation of recombinantly expressed dihydroxy-acid dehydratase from *Sulfobus solfataricus* - A key enzyme for the conversion of carbohydrates into chemicals. J. Biotechnol. 211, 31–41.

Chang, D.-E., Smalley, D.J., Tucker, D.L., Leatham, M.P., Norris, W.E., Stevenson, S.J., Anderson, A.B., Grissom, J.E., Laux, D.C., Cohen, P.S., Conway, T., 2004. Carbon nutrition of *Escherichia coli* in the mouse intestine. Proc. Natl. Acad. Sci. U. S. A. 101, 7427–32.

Chavarría, M., Nikel, P.I., Pérez-Pantoja, D., De Lorenzo, V., 2013. The Entner–Doudoroff pathway empowers *Pseudomonas putida* KT2440 with a high tolerance to oxidative stress. Environ. Microbiol. 15, 1772–1785.

Chemler, J.A., Fowler, Z.L., McHugh, K.P., Koffas, M.A.G., 2010. Improving NADPH availability for natural product biosynthesis in *Escherichia coli* by metabolic engineering. Metab. Eng. 12, 96–104.

Chen, X., Nielsen, K.F., Borodina, I., Kielland-Brandt, M.C., Karhumaa, K., 2011. Increased isobutanol production in *Saccharomyces cerevisiae* by overexpression of genes in valine metabolism. Biotechnol. Biofuels 4, 1–12.

Chen, X., Schreiber, K., Appel, J., Makowka, A., Fähnrich, B., Roettger, M., Hajirezaei, M.R., Sönnichsen, F.D., Schönheit, P., Martin, W.F., Gutekunst, K., 2016. The Entner–Doudoroff pathway is an overlooked glycolytic route in cyanobacteria and plants. Proc. Natl. Acad. Sci. 113, 5441–5446.

Chin, J.W., Cirino, P.C., 2011. Improved NADPH supply for xylitol production by engineered *Escherichia coli*with glycolytic mutations. Biotechnol. Prog. 27, 333–341.

Clapés, P., Fessner, W.D., Sprenger, G.A., Samland, A.K., 2010. Recent progress in stereoselective synthesis with aldolases. Curr. Opin. Chem. Biol. 14, 154–167.

Cocks, G.T., Aguilar, T., Lin, E.C., 1974. Evolution of L-1, 2-propanediol catabolism in *Escherichia coli* by recruitment of enzymes for L-fucose and L-lactate metabolism. J. Bacteriol. 118, 83–88.

Conway, T., 1992. The Entner-Doudoroff pathway: history, physiology and molecular biology. FEMS Microbiol. Lett. 103, 1–28.

Conway, T., Krogfelt, K.A., Cohen, P.S., 2004. The Life of commensal *Escherichia coli* in the mammalian ntestine. EcoSal Plus 1, 1–16.

Copley, S.D., 2015. An evolutionary biochemist's perspective on promiscuity. Trends Biochem. Sci. 40, 72–78.

Dahms, S.A., 1974. 3-Deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. Biochem. Biophys. Res. Commun. 60, 1433–1439.

Dandekar, T., Schuster, S., Snel, B., Huynen, M., Bork, P., 1999. Pathway alignment: application to the comparative analysis of glycolytic enzymes.. Biochem. J. 343, 115–124.

De Kok, S., Kozak, B.U., Pronk, J.T., Van Maris, A.J.A., 2012. Energy coupling in *Saccharomyces cerevisiae*: Selected opportunities for metabolic engineering. FEMS Yeast Res. 12, 387–397

De Rosa, M., Gambacorta, A., Nicolaus, B., Giardina, P., Poerio, E., Buonocore, V., 1984. Glucose metabolism in the extreme thermoacidophilic archaebacterium *Sulfolobus solfataricus*. Biochem. J 224, 407–414.

Del Castillo, T., Ramos, J.L., Rodríguez-Herva, J.J., Fuhrer, T., Sauer, U., Duque, E., 2007. Convergent peripheral pathways catalyze initial glucose catabolism in *Pseudomonas putida*: genomic and flux analysis. J. Bacteriol. 189, 5142–5152.

Diacovich, L., Lorenzi, L., Tomassetti, M., Méresse, S., Gramajo, H., 2017. The infectious intracellular lifestyle of *Salmonella enterica* relies on the adaptation to nutritional conditions within the *Salmonella*-containing vacuole. Virulence 8, 975–992.

Doelle, H.W., Kirk, L., Crittenden, R., Toh, H., Doelle, M.B., 1993. *Zymomonas Mobilis*-science and industrial application. Crit. Rev. Biotechnol. 13, 57–98.

Dudley, Q.M., Anderson, K.C., Jewett, M.C., 2016. Cell-Free Mixing of *Escherichia coli* crude extracts to prototype and rationally engineer high-titer mevalonate synthesis. ACS Synth. Biol. 5, 1578–1588.

Dudley, Q.M., Karim, A.S., Jewett, M.C., 2015. Cell-free metabolic engineering: Biomanufacturing beyond the cell. Biotechnol. J. 10, 69–82.

Duffaud, G.D., McCutchen, C.M., Leduc, P., Parker, K.N., Kelly, R.M., 1997. Purification and characterisation of extremely thermostable β- mannanase, β-mannosidase, and α-galactosidase from the hyperthermophilic eubacterium *Thermotoga neapolitana* 5068. Appl. Environ. Microbiol. 63, 169–177.

Egan, S.E., Fliege, R., Tong, S., Egan, S.E., Fliege, R., Tong, S., Shibata, A., Wolf, R.E., Conway, T., 1992. Molecular characterisation of the Entner-Doudoroff pathway in *Escherichia coli*: sequence analysis and localization of promoters for the *edd-eda* operon. J Bacteriol 174, 4638–4646.

Eisenreich, W., Dandekar, T., Heesemann, J., Goebel, W., 2010. Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. Nat. Rev. Microbiol. 8, 401–412.

Elzainy, T.A., Hassan, M.M., Allam, A.M., 1973. New pathway for nonphosphorylated degradation of gluconate by *Aspergillus niger* 114, 457–459.

Entner, N., Doudoroff, M., 1952. Glucose and gluconic acid oxidation of *Pseudomonas saccharophila*. J. Biol. Chem. 196, 853–862.

Erb, T.J., Jones, P.R., Bar-Even, A., 2017. Synthetic metabolism: metabolic engineering meets enzyme design. Curr. Opin. Chem. Biol. 37, 56–62.

Ettema, T.J.G., Ahmed, H., Geerling, A.C.M., Van Der Oost, J., Siebers, B., 2008. The non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN) of *Sulfolobus solfataricus*: A key-enzyme of the semi-phosphorylative branch of the Entner-Doudoroff pathway. Extremophiles 12, 75–88.

Fabich, A.J., Jones, S.A., Chowdhury, F.Z., Cernosek, A., Anderson, A., Smalley, D., McHargue, J.W., Hightower, G.A., Smith, J.T., Autieri, S.M., Leatham, M.P., Lins, J.J., Allen, R.L., Laux, D.C., Cohen, P.S., Conway, T., 2008. Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. Infect. Immun. 76, 1143–1152.

Fabris, M., Matthijs, M., Rombauts, S., Vyverman, W., Goossens, A., Baart, G.J.E., 2012. The metabolic blueprint of *Phaeodactylum tricornutum* reveals a eukaryotic Entner-Doudoroff glycolytic pathway. Plant J. 70, 1004–1014.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., Falkowski, P., 1998. Primary production of the biosphere: Integrating terrestrial and oceanic components. Science (80-.). 281, 237–240.

Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W., Milo, R., 2013. Glycolytic strategy as a tradeoff between energy yield and protein cost. Proc. Natl. Acad. Sci. 110, 10039–10044.

Flint, D.H., Emptage, M.H., 1988. Dihydroxy acid dehydratase from spinach contains a [2Fe-2S] cluster. J. Biol. Chem. 263, 3558–3564.

Flint, D.H., Emptage, M.H., Finnegan, M.G., Fu, W., Johnson, M.K., 1993. The role and properties of the iron-sulfur cluster in *Escherichia coli* dihydroxy-acid dehydratase. J. Biol. Chem. 268, 14732–14742.

Flint, H.J., Bayer, E.A., Rincon, M.T., Lamed, R., White, B.A., 2008. Polysaccharide utilization by gut bacteria: Potential for new insights from genomic analysis. Nat. Rev. Microbiol. 6, 121–131.

Fothergill-Gilmore, L.A., Michels, P.A.M., 1993. Evolution of glycolysis. Prog. Biophys. Mol. Biol. 59, 105–235.

Freter, R., Brickner, H., Fekete, J., Vickerman, M.M., Carey, K.E., Carey, E., 1982. Survival and implantation of *Escherichia coli* in the intestinal tract survival and implantation of *Escherichia coli* in the intestinal tract. Infect. Immun. 39, 686–703.

Fuhrer, T., Fischer, E., Sauer, U., 2005. Experimental identification and quantification of glucose metabolism in seven bacterial species. J. Bacteriol. 187, 1581–1590.

Gao, C., Li, Z., Zhang, L., Wang, C., Li, K., Ma, C., Xu, P., 2015. An artificial enzymatic reaction cascade for a cell-free bio-system based on glycerol. Green Chem. 17, 804–807.

Gao, X., Chen, J.C., Wu, Q., Chen, G.Q., 2011. Polyhydroxyalkanoates as a source of chemicals, polymers, and biofuels. Curr. Opin. Biotechnol. 22, 768–774.

Gagoski, D., Shi, Z., Keld, L., Vickers, C.E., Mahler, S., Speight, R., Johnston, W.A., Alexandrov, K., 2017. Cell-free pipeline for discovery of thermotolerant xylanases and *endo*-1,4-β-glucanases. J. Biotechnol. 259, 191–198.

Gardner, P.R., Fridovich, I., 1991. Superoxide sensitivity of the *Escherichia coli* 6-phosphogluconate dehydratase. J. Biol. Chem. 266, 1478–1483.

Gerlt, J.A., Babbitt, P.C., Rayment, I., 2005. Divergent evolution in the enolase superfamily: The interplay of mechanism and specificity. Arch. Biochem. Biophys. 433, 59–70.

Gibbs, M.D., Elinder, A.U., Reeves, R.A., Bergquist, P.L., 1996. Sequencing, cloning and expression of a β-1,4-mannanase gene, *man*A, from the extremely thermophilic anaerobic bacterium, *Caldicellulosiruptor* Rt8B.4. FEMS Microbiol. Lett. 141, 37–43.

Gómez-Baena, G., López-Lozano, A., Gil-Martínez, J., Lucena, J.M., Diez, J., Candau, P., García-Fernández, J.M., 2008. Glucose uptake and its effect on gene expression in *Prochlorococcus*. PLoS One 3, 1–11.

Gooday, G.W., 1990. The ecology of chitin degradation. Springer, Boston, MA.

Gulick, A.M., Hubbard, B.K., Gerlt, J.A., Rayment, I., 2001. Evolution of enzymatic activities

in the enolase superfamily: Identification of the general acid catalyst in the active site of D-glucarate dehydratase from *Escherichia coli*. Biochemistry 40, 10054–10062.

Guterl, J.K., Garbe, D., Carsten, J., Steffler, F., Sommer, B., Reiße, S., Philipp, A., Haack, M., Rühmann, B., Koltermann, A., Kettling, U., Brück, T., Sieber, V., 2012. Cell-free metabolic engineering: production of chemicals by minimized reaction cascades. ChemSusChem 5, 2165–2172.

Hammer, S.C., Knight, A.M., Arnold, F.H., 2017. Design and evolution of enzymes for non-natural chemistry. Curr. Opin. Green Sustain. Chem. 7, 23–30.

Han, M., Yoon, S.S., Lee, S.Y., 2001. Proteome analysis of metabolically engineered *Escherichia coli* producing Poly(3-Hydroxybutyrate). J. Bacteriol. 183, 301–308.

Hara, K.Y., Kondo, A., 2015. ATP regulation in bioproduction. Microb. Cell Fact. 14, 1–9.

He, M.X., Wu, B., Qin, H., Ruan, Z.Y., Tan, F.R., Wang, J.L., Shui, Z.X., Dai, L.C., Zhu, Q.L., Pan, K., Tang, X.Y., Wang, W.G., Hu, Q.C., 2014. *Zymomonas mobilis*: A novel platform for future biorefineries. Biotechnol. Biofuels 7, 101.

Hecquet, L., Hélaine, V., Charmantray, F., Lemaire, M., 2009. Enzymes Catalyzing C-C Bond Formation for the Synthesis of Monosaccharide Analogs, Modern Biocatalysis: Stereoselective and Environmentally Friendly Reactions.

Heider, S.A.E., Wendisch, V.F., 2015. Engineering microbial cell factories: metabolic engineering of *Corynebacterium glutamicum* with a focus on non-natural products. Biotechnol. J. 10, 1170–1184.

Hold, C., Billerbeck, S., Panke, S., 2016. Forward design of a complex enzyme cascade reaction. Nat. Commun. 7, 1–8.

Honda, K., Hara, N., Cheng, M., Nakamura, A., Mandai, K., Okano, K., Ohtake, H., 2016. *In vitro* metabolic engineering for the salvage synthesis of NAD+. Metab. Eng. 35, 114–120.

Honda, K., Kimura, K., Ninh, P.H., Taniguchi, H., Okano, K., Ohtake, H., 2017. *In vitro* bioconversion of chitin to pyruvate with thermophilic enzymes. J. Biosci. Bioeng. 124, 296–301.

Hooper, L. V., Xu, J., Falk, P.G., Midtvedt, T., Gordon, J.I., 1999. A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. Proc. Natl. Acad. Sci. 96, 9833–9838.

Hugdahl, M.B., Beery, J.T., Doyle, M.P., 1988. Chemotactic behavior of *Campylobacter jejuni*. Infect. Immun. 56, 1560–1566.

Hugouvieux-Cotte-Pattat, N., Robert-Baudouy, J., 1987. Hexuronate catabolism in *Erwinia Chrysanthemi*. J. Bacteriol. 169, 1223–1231.

Hunt, J.C., Phibbs, P. V., 1983. Regulation of alternate peripheral pathways of glucose catabolism during aerobic and anaerobic growth of *Pseudomonas aeruginosa*. J. Bacteriol. 154, 793–802.

Hunt, J.C., Phibbs, P.V. V., 1981. Failure of *Pseudomonas aeruginosa* to form membrane-associated glucose dehydrogenase activity during anaerobic growth with nitrate. Biochem. Biophys. Res. Commun. 102, 1393–1399.

Jackson, J.B., 2003. Proton translocation by transhydrogenase. FEBS Lett. 545, 18–24.

Johnsen, U., Dambeck, M., Zaiss, H., Fuhrer, T., Soppa, J., Sauer, U., Schönheit, P., 2009. D-xylose degradation pathway in the halophilic archaeon *Haloferax volcanii*. J. Biol. Chem. jbc-M109.

Johnsen, U., Schönheit, P., 2004. Novel xylose dehydrogenase in the halophilic archaeon *Haloarcula marismortui*. J. Bacteriol. 186, 6198–6207.

Johnsen, U., Sutter, J.M., Zaiß, H., Schönheit, P., 2013. L-Arabinose degradation pathway in the haloarchaeon *Haloferax volcanii* involves a novel type of l-arabinose

dehydrogenase. Extremophiles 17, 897–909.

Kalnenieks, U., Pentjuss, A., Rutkis, R., Stalidzans, E., Fell, D.A., 2014. Modeling of *Zymomonas mobilis* central metabolism for novel metabolic engineering strategies. Front. Microbiol. 5, 42.

Kengen, S.W.M., De Bok, F.A.M., Van Loo, N.D., Dijkema, C., Stams, A.J.M., De Vos, W.M., 1994. Evidence for the operation of a novel Embden-Meyerhof pathway that involves ADP-dependent kinases during sugar fermentation by *Pyrococcus furiosus*. J. Biol. Chem. 269, 17537–17541.

Kenney, K.L., Smith, W.A., Gresham, G.L., Westover, T.L., 2013. Understanding biomass feedstock variability. Biofuels 4, 111–127.

Kersters, K., De Ley, J., 1968. The occurrence of the Entner-Doudoroff pathway in bacteria. Antonie Van Leeuwenhoek 34, 393–408.

Khersonsky, O., Roodveldt, C., Tawfik, D.S., 2006. Enzyme promiscuity: evolutionary and mechanistic aspects. Curr. Opin. Chem. Biol. 10, 498–508.

Khanal, A., Mcloughlin, S.Y., Kershner, J.P., Copley, S.D., 2014. Differential Effects of a Mutation on the Normal and Promiscuous Activities of Orthologs : Implications for Natural and Directed Evolution. Mol. Biol. Evol. 32, 100–108.

Khosravi, C., Kun, R.S., Visser, J., Aguilar-Pontes, M.V., De Vries, R.P., Battaglia, E., 2017. In vivo functional analysis of L-rhamnose metabolic pathway in *Aspergillus niger*: A tool to identify the potential inducer of RhaR. BMC Microbiol. 17, 1–12.

Kim, S., Lee, S.B., 2008. Identification and characterisation of the bacterial d-gluconate dehydratase in *Achromobacter xylosoxidans*. Biotechnol. Bioprocess Eng. 13, 436–444.

Kim, S., Lee, S.B., 2006. Catalytic promiscuity in dihydroxy-acid dehydratase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. J. Biochem. 139, 591–596.

Kim, S., Lee, S.B., 2005. Identification and characterisation of *Sulfolobus solfataricus* D-gluconate dehydratase: a key enzyme in the non-phosphorylated Entner-Doudoroff pathway. Biochem. J. 387, 271–80.

Kim, S.M., Paek, K.H., Lee, S.B., 2012. Characterisation of NADP+ - specific L-rhamnose dehydrogenase from the thermoacidophilic Archaeon *Thermoplasma acidophilum*. Extremophiles 16, 447–454.

Kobayashi, H., Fukuoka, A., 2013. Synthesis and utilisation of sugar compounds derived from lignocellulosic biomass. Green Chem. 15, 1740–1763.

Koga, S., Takahashi, M., Sakasegawa, S., Shimizu, S., Ohshinia, T., 2000. Biochemical characterisation, cloning, and sequencing of ADP - dependent (AMP-Forming) glucokinase from two hyperthermophilic archaea, *Pyrococcus furiosus* and *Thermococcus litoralis*. J. Biochem. 128, 1079–1085.

Koivistoinen, O.M., Arvas, M., Headman, J.R., Andberg, M., Penttilä, M., Jeffries, T.W., Richard, P., 2012. Characterisation of the gene cluster for L-rhamnose catabolism in the yeast *Scheffersomyces* (*Pichia*) *stipitis*. Gene 492, 177–185.

Kouril, T., Wieloch, P., Reimann, J., Wagner, M., Zaparty, M., Albers, S.V., Schomburg, D., Ruoff, P., Siebers, B., 2013. Unraveling the function of the two Entner-Doudoroff branches in the thermoacidophilic Crenarchaeon *Sulfolobus solfataricus* P2. FEBS J. 280, 1126–1138.

Kuivanen, J., Sugai-Guérios, M.H., Arvas, M., Richard, P., 2016. A novel pathway for fungal D-glucuronate catabolism contains an L-idonate forming 2-keto-L-gulonate reductase. Sci. Rep. 6, 1–9.

Lamble, H.J., Heyer, N.I., Bull, S.D., Hough, D.W., Danson, M.J., 2003. Metabolic pathway promiscuity in the Archaeon *Sulfolobus solfataricus* revealed by studies on glucose

dehydrogenase and 2-keto-3-deoxygluconate aldolase. J. Biol. Chem. 278, 34066–34072.

Lamble, H.J., Milburn, C.C., Taylor, G.L., Hough, D.W., Danson, M.J., 2004. Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in *Sulfolobus solfataricus*. FEBS Lett. 576, 133–136.

Lee, K.Y., Park, J.M., Kim, T.Y., Yun, H., Lee, S.Y., 2010. The genome-scale metabolic network analysis of *Zymomonas mobilis* ZM4 explains physiological features and suggests ethanol and succinic acid production strategies. Microb. Cell Fact. 9, 1–12.

Lee, J. V, Gibson, D.M., Shewan, J.M., 1977. A numerical taxonomic study of some Pseudomonas-like marine bacteria. J. Gen. Microbiol. 98, 439–451.

Lessie, T.G., Phibbs, P.V., 1984. Alternative pathways of carbohydrate utilization in Pseudomonads. Annu.Rev.Microbiol. 38, 359–387.

Li, S., Huang, D., Li, Y., Wen, J., Jia, X., 2012. Rational improvement of the engineered isobutanol-producing *Bacillus subtilis* by elementary mode analysis. Microb. Cell Fact. 11, 1–12.

Liang, S., Chen, H., Liu, J., Wen, J., 2018. Rational design of a synthetic Entner – Doudoroff pathway for enhancing glucose transformation to isobutanol in *Escherichia coli*. J. Ind. Microbiol. Biotechnol. 45, 187–199.

Liu, B., Zhang, Y., Zhang, X., Yan, C., Zhang, Y., Xu, X., Zhang, W., 2016. Discovery of a rhamnose utilization pathway and rhamnose-inducible promoters in *Pichia pastoris*. Sci. Rep. 6, 1–8.

Liu, H., Lu, T., 2015. Autonomous production of 1,4-butanediol via a de novo biosynthesis pathway in engineered *Escherichia coli*. Metab. Eng. 29, 135–141.

Mandrand-Berthelot, M.-A., Condemine, G., Hugouvieux-Cotte-Pattat, N., 2004. Catabolism of hexuronides, hexuronates, aldonates, and aldarates. EcoSal Plus 1, 1–21.

Maruyama, Y., Nakajima, T., Ichishima, E., 1994. A 1,2-α-D-mannosidase from a *Bacillus* sp.: purification, characterisation, and mode of action. Carbohydr. Res. 251, 89–98.

Marx, A., Hans, S., Möckel, B., Bathe, B., De Graaf, A.A., 2003. Metabolic phenotype of phosphoglucose isomerase mutants of *Corynebacterium glutamicum*. J. Biotechnol. 104, 185–197.

Matsubara, K., Köhling, R., Schönenberger, B., Kouril, T., Esser, D., Bräsen, C., Siebers, B., Wohlgemuth, R., 2014. One-step synthesis of 2-keto-3-deoxy-D-gluconate by biocatalytic dehydration of D-gluconate. J. Biotechnol. 191, 69–77.

Matsubara, K., Yokooji, Y., Atomi, H., Imanaka, T., 2011. Biochemical and genetic characterisation of the three metabolic routes in *Thermococcus kodakarensis* linking glyceraldehyde 3-phosphate and 3-phosphoglycerate. Mol. Microbiol. 81, 1300–1312.

McMillan, J.D., 1993. Xylose Fermentation to Ethanol: A Review, NREL/TP-421-4944.

Meijnen, J.P., De Winde, J.H., Ruijssenaars, H.J., 2009. Establishment of oxidative D-xylose metabolism in *Pseudomonas putida* S12. Appl. Environ. Microbiol. 75, 2784–2791.

Mekjian, K.R., Bryan, E.M., Beall, B.W., Moran, C.P., 1999. Regulation of hexuronate utilization in *Bacillus subtilis*. J. Bacteriol. 181, 426–433.

Meléndez-Hevia, E., Waddell, T.G., Heinrich, R., Monter, F., 1997. Theoretical approaches to the evolutionary optimization of glycolysis chemical analysis. Eur. J. Biochem. 244, 527–543.

Meléndez-Hevia, E., Waddell, T.G., Montero, F., 1994. Optimization of metabolism: the evolution of metabolic pathways toward simplicity through the game of the pentose phosphate cycle. J. theor. biol. 166, 201–220.

Mitchell, C., Dawes, E., 1982. The role of oxygen in the regulation of glucose metabolism , transport and the tricarboxylic acid cycle in *Pseudomonas aeruginosa*. J. Gen. Microbiol. 128, 49–59.

Moysés, D.N., Reis, V.C.B., de Almeida, J.R.M., de Moraes, L.M.P., Torres, F.A.G., 2016. Xylose fermentation by *Saccharomyces cerevisiae*: Challenges and prospects. Int. J. Mol. Sci. 17, 1–18.

Mukund, S., Adams, M.W.W., 1991. The novel tungsten-iron-sulfur protein of the hyperthermophilic archaebacterium, *Pyrococcus furiosus*, is an aldehyde ferredoxin oxidoreductase. J. Biol. Chem. 266, 14208–14216.

Mutturi, S., Lide, G., 2013. Effect of temperature on simultaneous saccharification and fermentation of pretreated spruce and arundo. Ind. Eng. Chem. Res. 52, 1244–1251.

Nakajima, M., Fushinobu, S., Imamura, H., Shoun, H., Wakagi, T., 2006. Crystallization and preliminary X-ray analysis of cytosolic α-mannosidase from *Thermotoga maritima*. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 62, 104–105.

Ng, C.Y., Farasat, I., Maranas, C.D., Salis, H.M., 2015. Rational design of a synthetic Entner-Doudoroff pathway for improved and controllable NADPH regeneration. Metab. Eng. 29, 86–96.

Nikel, P.I., Chavarría, M., Danchin, A., de Lorenzo, V., 2016. From dirt to industrial applications: *Pseudomonas putida* as a synthetic biology chassis for hosting harsh biochemical reactions. Curr. Opin. Chem. Biol. 34, 20–29.

Nikel, P.I., Chavarría, M., Fuhrer, T., Sauer, U., De Lorenzo, V., 2015. *Pseudomonas putida* KT2440 strain metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and pentose phosphate pathways. J. Biol. Chem. jbc-M115.

Noor, E., Bar-Even, A., Flamholz, A., Reznik, E., Liebermeister, W., Milo, R., 2014. Pathway thermodynamics highlights kinetic obstacles in central metabolism. PLoS Comput. Biol. 10.

Noor, E., Eden, E., Milo, R., Alon, U., 2010. Central carbon metabolism as a minimal biochemical walk between precursors for biomass and energy. Mol. Cell 39, 809–820.

Nowak, C., Beer, B., Pick, A., Roth, T., Lommes, P., Sieber, V., 2015. A water-forming NADH oxidase from *Lactobacillus pentosus* suitable for the regeneration of synthetic biomimetic cofactors. Front. Microbiol. 6, 1–9.

Numao, S., He, S., Evjen, G., Howard, S., Tollersrud, O.K., Withers, S.G., 2000. Identification of Asp197 as the catalytic nucleophile in the family 38 α-mannosidase from bovine kidney lysosomes. FEBS Lett. 484, 175–178.

Nunn, C.E.M., Johnsen, U., Schönheit, P., Fuhrer, T., Sauer, U., Hough, D.W., Danson, M.J., 2010. Metabolism of pentose sugars in the hyperthermophilic archaea Sulfolobus solfataricus and *Sulfolobus acidocaldarius*. J. Biol. Chem. 285, 33701–33709.

Nygård, Y., Maaheimo, H., Mojzita, D., Toivari, M., Wiebe, M., Resnekov, O., Gustavo Pesce, C., Ruohonen, L., Penttilä, M., 2014. Single cell and in vivo analyses elucidate the effect of xylC lactonase during production of D-xylonate in *Saccharomyces cerevisiae*. Metab. Eng. 25, 238–247.

O'Hara, A.M., Shanahan, F., 2006. The gut flora as a forgotten organ. EMBO Rep. 7, 688–693.

Opgenorth, P.H., Korman, T.P., Bowie, J.U., 2014. A synthetic biochemistry molecular purge valve module that maintains redox balance. Nat. Commun. 5, 1–8.

Oreb, M., Dietz, H., Farwick, A., Boles, E., 2012. Novel strategies to improve co-fermentation of pentoses with d-glucose by recombinant yeast strains in lignocellulosic hydrolysates. Bioengineered 3, 347–351.

Pacheco, A.R., Munera, D., Waldor, M.K., Sperandio, V., Ritchie, J.M., 2012. Fucose sensing regulates bacterial intestinal colonization. Nature 492, 113–117.

Parkhill, J., Wren, B.W., Mungall, K., Ketley, J.M., Churcher, C., Basham, D., Chillingworth, T., Davies, R.M., Feltwell, T., Holroyd, S., Jagels, K., Karlyshev, A. V, Moule, S., Pallen, M.J., Penn, C.W., Quail, M.A., Rajandream, M.-A., Rutherford, K.M., van Vliet, A.H.M., Whitehead, S., Barrell, B.G., 2000. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. Nature 403, 665-668.

Patra, T., Koley, H., Ramamurthy, T., Ghose, A.C., Nandy, R.K., 2012. The Entner-Doudoroff pathway is obligatory for gluconate utilization and contributes to the pathogenicity of *Vibrio cholerae*. J. Bacteriol. 194, 3377–3385.

Pauluhn, A., Ahmed, H., Lorentzen, E., Buchinger, S., Schomburg, D., Siebers, B., Pohl, E., 2008. Crystal structure and stereochemical studies of KD(P)G aldolase from *Thermoproteus tenax*. Proteins Struct. Funct. Genet. 72, 35–43.

Peekhaus, N., Conway, T., 1998. What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. J. Bacteriol. 180, 3495–3502.

Petroll, K., Kopp, D., Care, A., Bergquist, P.L., Sunna, A., 2018. Tools and strategies for constructing cell-free enzyme pathways. Biotechnol. Adv. in press.

Portalier, R., Robert Baudouy, J., Stoeber, F., 1980. Regulation of *Escherichia coli* K-12 hexuronate system genes: exu regulon. J. Bacteriol. 143, 1095–1107.

Postma, P.W.W., Lengeler, J.W.W., Jacobson, G.R.R., 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. 57, 543–594.

Radek, A., Krumbach, K., Gätgens, J., Wendisch, V.F., Wiechert, W., Bott, M., Noack, S., Marienhagen, J., 2014. Engineering of *Corynebacterium glutamicum* for minimized carbon loss during utilization of D-xylose containing substrates. J. Biotechnol. 192, 156–160.

Radek, A., Müller, M.F., Gätgens, J., Eggeling, L., Krumbach, K., Marienhagen, J., Noack, S., 2016. Formation of xylitol and xylitol-5-phosphate and its impact on growth of D-xylose-utilizing *Corynebacterium glutamicum* strains. J. Biotechnol. 231, 160–166.

Rahman, M.M., Andberg, M., Koivula, A., Rouvinen, J., Hakulinen, N., 2018. The crystal structure of D-xylonate dehydratase reveals functional features of enzymes from the Ilv/ED dehydratase family. Sci. Rep. 8, 23–25.

Rahman, M.M., Andberg, M., Thangaraj, S.K., Parkkinen, T., Penttilä, M., Jänis, J., Koivula, A., Rouvinen, J., Hakulinen, N., 2017. The crystal structure of a bacterial L-arabinonate dehydratase contains a [2Fe-2S] cluster. ACS Chem. Biol. 12, 1919–1927.

Rakus, J.F., Fedorov, A.A., Fedorov, E. V., Glasner, M.E., Hubbard, B.K., Delli, J.D., Babbitt, P.C., Almo, S.C., Gerlt, J.A., 2008. Evolution of enzymatic activities in the enolase superfamily: L-rhamnonate dehydratase. Biochemistry 47, 9944–9954.

Rakus, J.F., Fedorov, A.A., Fedorov, E. V, Glasner, M.E., Vick, J.E., Babbitt, P.C., Almo, S.C., Gerlt, J.A., 2007. Evolution of enzymatic activities in the enolase superfamily: D-mannonate dehydratase from *Novosphingobium aromaticivorans*. Biochemistry 46, 12896–12908.

Ramachandran, S., Fontanille, P., Pandey, A., Larroche, C., 2006. Gluconic Acid: properties, applications and microbial production. Food Technol. Biotechnol. 44, 185–195.

Ramos, J.L., Cuenca, M.S., Molina-Santiago, C., Segura, A., Duque, E., Gómez-Garciá, M.R., Udaondo, Z., Roca, A., 2015. Mechanisms of solvent resistance mediated by

interplay of cellular factors in *Pseudomonas putida*. FEMS Microbiol. Rev.

Reher, M., Fuhrer, T., Bott, M., Schönheit, P., 2010. The nonphosphorylative Entner-Doudoroff pathway in the thermoacidophilic euryarchaeon *Picrophilus torridus* involves a novel 2-Keto-3-deoxygluconate- specific aldolase. J. Bacteriol. 192, 964–974.

Reher, M., Schönheit, P., 2006. Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner-Doudoroff pathway, constitute a novel enzyme family within the aldehyde d. FEBS Lett. 580, 1198–1204.

Rehr, B., Wilhelm, C., Sahm, H., 1991. Production of sorbitol and gluconic acid by permeabilized cells of *Zymomonas mobilis*. Appl. Microbiol. Biotechnol. 35, 144–148.

Reis, D., Vian, B., Roland, J.C., 1994. Cellulose-glucuronoxylans and plant cell wall structure. Micron 25, 171–187.

Rigo, L.U., Marechal, L.R., Vieira, M.M., Viega, L.A., 1985. Oxidative pathway for L-rhamnose degradation in *Pullularia pullulans*. Candian J. Microbiol. 31, 817-.

Robbe, C., Capon, C., Coddeville, B., Michalski, J.-C., 2004. Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract. Biochem J 384, 307–316.

Rodriguez, M., Wedd, A.G., Scopes, R.K., 1996. 6-phosphogluconate dehydratase from *Zymomonas mobilis*: an iron-sulfur-manganese enzyme. Biochem. Mol. Biol. Int. 38, 783—789.

Romano, A.H., Conway, T., 1996. Evolution of carbohydrate metabolic pathways. Res. Microbiol. 147, 448–455.

Ronimus, R.S., Morgan, H.W., 2003. Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. Archaea 1, 199–221.

Rothe, M., Alpert, C., Loh, G., Blaut, M., 2013. Novel insights into *E. coli*'s hexuronate metabolism: KduI facilitates the conversion of galacturonate and glucuronate under osmotic stress conditions. PLoS One 8, e56906.

Royer, S.F., Haslett, L., Crennell, S.J., Hough, D.W., Danson, M.J., Bull, S.D., 2010. Structurally informed site-directed mutagenesis of a stereochemically promiscuous aldolase to afford stereochemically complementary biocatalysts. J. Am. Chem. Soc. 132, 11753–11758.

Salusjärvi, L., Toivari, M., Vehkomäki, M.L., Koivistoinen, O., Mojzita, D., Niemelä, K., Penttilä, M., Ruohonen, L., 2017. Production of ethylene glycol or glycolic acid from D-xylose in *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. 101, 8151–8163.

Schäfer, T., Schönheit, P., 1992. Maltose fermentation to acetate, CO_2 and H_2 in the anaerobic hyperthermophilic archaeon *Pyrococcus furiosus*: evidence for the operation of a novel sugar fermentation pathway. Arch. Microbiol. 158, 188–202.

Schnoes, A.M., Brown, S.D., Dodevski, I., Babbitt, P.C., 2009. Annotation error in public databases: misannotation of molecular function in enzyme superfamilies. PLoS Comput. Biol. 5, e1000605.

Scopes, R.K., Griffiths-Smith, K., 1984. Use of differential dye-ligand chromatography with affinity elution for enzyme purification: 6-phosphogluconate dehydratase from *Zymomonas mobilis*. Anal. Biochem. 136, 530–534.

Scully, S.M., Orlygsson, J., 2014. Recent advances in second generation ethanol production by thermophilic bacteria. Energies 8, 1–30.

Seibert, C.M., Raushel, F.M., 2005. Structural and catalytic diversity within the amidohydrolase superfamily. Biochemistry 44, 6383–6391.

Selig, M., Xavier, K.B., Santos, H., Schönheit, P., 1997. Comparative analysis of Embden-

Meyerhof and Entner-Doudoroff glycolytic pathways in hyperthermophilic archaea and the bacterium *Thermotoga*. Arch. Microbiol. 167, 217–232.

Sharma, V., Ichikawa, M., Freeze, H.H., 2014. Mannose metabolism: more than meets the eye. Biochem BIophys Res Commun 453, 220–228.

Shulami, S., Gat, O., Sonenshein, A.L., Shoham, Y., 1999. The glucuronic acid utilization gene cluster from *Bacillus stearothermophilus* T-6. J. Bacteriol. 181, 3695–3704.

Siebers, B., Klenk, H., 1998. PPi -dependent phosphofructokinase from *Thermoproteus tenax*, an archaeal descendant of an ancient line in phosphofructokinase evolution. J. Bacteriol. 180, 2137–2143.

Siebers, B., Schönheit, P., 2005. Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. Curr. Opin. Microbiol. 8, 695–705.

Siebers, B., Tjaden, B., Michalke, K., Do, C., Ahmed, H., Zaparty, M., Gordon, P., Sensen, C.W., Zibat, A., Klenk, H., Schuster, S.C., Hensel, R., 2004. Reconstruction of the central carbohydrate metabolism of *Thermoproteus tenax* by use of genomic and biochemical data 186, 2179–2194.

Siedler, S., Bringer, S., Bott, M., 2011. Increased NADPH availability in *Escherichia coli*: Improvement of the product per glucose ratio in reductive whole-cell biotransformation. Appl. Microbiol. Biotechnol. 92, 929–937.

Smith, K.M., Cho, K.M., Liao, J.C., 2010. Engineering *Corynebacterium glutamicum* for isobutanol production. Appl. Microbiol. Biotechnol. 87, 1045–1055.

Snider, T.A., Fabich, A.J., Conway, T., Clinkenbeard, K.D., 2009. *E. coli* O157:H7 catabolism of intestinal mucin-derived carbohydrates and colonization. Vet. Microbiol. 136, 150–154.

Song, B.-G., Kim, T.-K., Jung, Y., Lee, Y.-H., 2006. Modulation of talA gene in pentose phosphate pathway for overproduction of poly-beta-hydroxybutyrate in transformant *Escherichia coli* harboring phbCAB operon. J. Biosci. Bioeng. 102, 237–240.

Sprenger, G.A., 1996. Carbohydrate metabolism in *Zymomonas mobilis*: a catabolic highway with some scenic routes. FEMS Microbiol. Lett. 145, 301–307.

Stahl, M., Friis, L.M., Nothaft, H., Liu, X., Li, J., Szymanski, C.M., Stintzi, A., 2011. L-Fucose utilization provides *Campylobacter jejuni* with a competitive advantage. Proc. Natl. Acad. Sci. 108, 7194–7199.

Stephens, C., Christen, B., Fuchs, T., Sundaram, V., Watanabe, K., Jenal, U., 2007. Genetic analysis of a novel pathway for D-xylose metabolism in *Caulobacter crescentus*. J. Bacteriol. 189, 2181–2185.

Sudarsan, S., Dethlefsen, S., Blank, L.M., Siemann-Herzberg, M., Schmid, A., 2014. The functional structure of central carbon metabolism in *Pseudomonas putida* KT2440. Appl. Environ. Microbiol. 80, 5292–5303.

Sundara Sekar, B., Seol, E., Park, S., 2017. Co-production of hydrogen and ethanol from glucose in *Escherichia coli* by activation of pentose-phosphate pathway through deletion of phosphoglucose isomerase (*pgi*) and overexpression of glucose-6-phosphate dehydrogenase (*zwf*) and 6-p. Biotechnol. Biofuels 10, 1–12.

Sunna, A., 2010. Modular organisation and functional analysis of dissected modular β-mannanase CsMan26 from *Caldicellulosiruptor* Rt8B.4. Appl. Microbiol. Biotechnol. 86, 189–200.

Sunna, A., Gibbs, M.D., Chin, C.W.J., Nelson, P.J., Bergquist, P.L., 2000. A gene encoding a novel multidomain β-1,4-mannanase from *Caldibacillus cellulovorans* and action of the recombinant enzyme on kraft pulp. Appl. Environ. Microbiol. 66, 664–670.

Sutter, J.M., Tästensen, J.B., Johnsen, U., Soppa, J., Schönheit, P., 2016. Key enzymes of the semiphosphorylative Entner-Doudoroff Pathway in the haloarchaeon *Haloferax*

volcanii: characterisation of glucose dehydrogenase, gluconate dehydratase, and 2-keto-3-deoxy-6-phosphogluconate aldolase. J. Bacteriol. 198, 2251–2262.

Sweeney, N.J., Laux, D.C., Cohen, P.S., 1996. *Escherichia coli* F-18 and *E. coli* K-12 *eda* mutants do not colonize the streptomycin-treated mouse large intestine. Infect. Immun. 64, 3504–3511.

Swings, J., Genetics, M., 1977. The Biology of *Zymomonas*. Bacteriol. Rev. 41, 1–46.

Szymona, M., Doudoroff, M., 1960. Carbohydrate metabolism in *Rhodopseudomonas spheroides*. J. Gen. Microbiol. 22, 167–183.

Tenhaef, N., Brüsseler, C., Radek, A., Hilmes, R., Unrean, P., Marienhagen, J., Noack, S., 2018. Production of D-xylonic acid using a non-recombinant *Corynebacterium glutamicum* strain. Bioresour. Technol. 268, 332–339.

Theodossis, A., Walden, H., Westwick, E.J., Connaris, H., Lamble, H.J., Hough, D.W., Danson, M.J., Taylor, G.L., 2004. The structural basis for substrate promiscuity in 2-keto-3-deoxygluconate aldolase from the Entner-Doudoroff pathway in *Sulfolobus solfataricus*. J. Biol. Chem. 279, 43886–4892.

Tiwari, N.P., Campbell, J.J.R., 1969. Enzymatic control of the metabolic activity of *Pseudomonas aeruginosa* grown in glucose and succinate media. Biochim. Biophys. Acta 192, 395–401.

Toivari, M., Nygård, Y., Kumpula, E.P., Vehkomäki, M.L., Benčina, M., Valkonen, M., Maaheimo, H., Andberg, M., Koivula, A., Ruohonen, L., Penttilä, M., Wiebe, M.G., 2012. Metabolic engineering of *Saccharomyces cerevisiae* for bioconversion of D-xylose to D-xylonate. Metab. Eng. 14, 427–436.

Tuininga, J.E., Verhees, H.C., Van der Oost, J., Kengen, W.M.S., Stams, A.J.M., Vos, W.M. De, 1999. Molecular and biochemical characterisation of the ADP-dependent phosphofructokinase from the hyperthermophilic archaeon *Pyrococcus furiosus*. J. Biol. Chem. 274, 21023–21028.

Twerdochlib, A.L., Pedrosa, F., Funayama, S., 1994. L-Rharnnose metabolism in *Pichia stipitis* and *Debaryomyces polymolphus*. Candian J. Microbiol. 40, 896–902.

Valdehuesa, K.N.G., Ramos, K.R.M., Nisola, G.M., Bañares, A.B., Cabulong, R.B., Lee, W.K., Liu, H., Chung, W.J., 2018. Everyone loves an underdog: metabolic engineering of the xylose oxidative pathway in recombinant microorganisms. Appl. Microbiol. Biotechnol. 102, 7703–7716.

Van de Werken, H.J.G., Brouns, S.J.J., Van der Oost, J., 2008. Pentose metabolism in archaea, Archaea: New Models for Prokaryotic Biology. Norfolk, UK: Caister Academic Press.

Van Der Oost, J., Wolterink-Van Loo, S., Siemerink, M.A.J., Perrakis, G., Kaper, T., Kengen, S.W.M., 2009. Improving low-temperature activity of *Sulfolobus acidocaldarius* 2-keto-3-deoxygluconate aldolase. Archaea 2, 233–239.

Vanhooren, P.T., Vandamme, E.J., 1999. L-Fucose: occurrence, physiological role, chemical, enzymatic and microbial synthesis. J. Chem. Technol. Biotechnol. 74, 479–497.

Velayudhan, J., Kelly, D.J., 2002. Analysis of gluconeogenic and anaplerotic enzymes in *Campylobacter jejuni*: An essential role for phosphoenolpyruvate carboxykinase. Microbiology 148, 685–694.

Verhees, C.H., Kengen, S.W.M., Tuininga, J.E., Schut, G.J., Adams, M.W.W., De Vos, W.M., Van Der Oost, J., 2003. The unique features of glycolytic pathways in Archaea. Biochem. J 375, 231–246.

Vicente, M., Canovas, J.L., 1973. Glucolysis in *Pseudomonas putida*: physiological role of alternative routes from the analysis of defective mutants. J. Bacteriol. 116, 908–914.

Viikari, L., Berry, D.R., 1988. Carbohydrate metabolism in *Zymomonas*. Crit. Rev. Biotechnol. 7, 237–261.

Vorwerk, H., Huber, C., Mohr, J., Bunk, B., Bhuju, S., Wensel, O., Spröer, C., Fruth, A., Flieger, A., Schmidt-Hohagen, K., Schomburg, D., Eisenreich, W., Hofreuter, D., 2015. A transferable plasticity region in *Campylobacter coli* allows isolates of an otherwise non-glycolytic food-borne pathogen to catabolize glucose. Mol. Microbiol. 98, 809–830.

Waligora, E.A., Fisher, C.R., Hanovice, N.J., Rodou, A., Wyckoff, E.E., Payne, S.M., 2014. Role of intracellular carbon metabolism pathways in *Shigella flexneri* virulence. Infect. Immun. 82, 2746–2755.

Wasserstrom, L., Portugal-Nunes, D., Almqvist, H., Sandström, A.G., Lidén, G., Gorwa-Grauslund, M.F., 2018. Exploring D-xylose oxidation in *Saccharomyces cerevisiae* through the Weimberg pathway. AMB Express 8, 33.

Wang, J., Li, G., Reetz, M.T., 2017. Enzymatic site-selectivity enabled by structure-guided directed evolution. Chem. Commun. 53, 3916–3928.

Watanabe, S., Kodak, T., Makino, K., 2006a. Cloning, expression, and characterisation of bacterial L-arabinose 1-dehydrogenase involved in an alternative pathway of L-arabinose metabolism. J. Biol. Chem. 281, 2612–2623.

Watanabe, S., Kodaki, T., Makino, K., 2006b. A novel α-ketoglutaric semialdehyde dehydrogenase: Evolutionary insight into an alternative pathway of bacterial L-arabinose metabolism. J. Biol. Chem. 281, 28876–28888.

Watanabe, S., Makino, K., 2009. Novel modified version of nonphosphorylated sugar metabolism - An alternative L-rhamnose pathway of *Sphingomonas* sp. FEBS J. 276, 1554–1567.

Watanabe, S., Saimura, M., Makino, K., 2008. Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. J. Biol. Chem. 283, 20372–20382.

Watanabe, S., Shimada, N., Tajima, K., Kodaki, T., Makino, K., 2006c. Identification and characterisation of L-arabonate dehydratase, L-2-Keto-3-deoxyarabonate dehydratase, and L-arabinolactonase involved in an alternative pathway of L-arabinose metabolism: Novel evolutionary insight into sugar metabolism. J. Biol. Chem. 281, 33521–33536.

Weckbecker, A., Hummel, W., 2004. Improved synthesis of chiral alcohols with *Escherichia coli* cells co-expressing pyridine nucleotide transhydrogenase, NADP+-dependent alcohol dehydrogenase and NAD+-dependent formate dehydrogenase. Biotechnol. Lett. 26, 1739–1744.

Weimberg, R., 1961. Pentose oxidation by *Pseudomonas fragi*. J. Biol. Chem. 236, 629–635.

Weisser, P., Kramer, R., Sahm, H., Sprenger, G.A., 1995. Functional expression of the glucose transporter of *Zymomonas mobilis* leads to restoration of glucose and fructose uptake in *Escherichia coli* mutants and provides evidence for its facilitator action. J. Bacteriol. 177, 3351–3354.

Welch, P., Scopes, R.K., 1985. Studies on cell-free metabolism: Ethanol production by a yeast glycolytic system reconstituted from purified enzymes. J. Biotechnol. 2, 257–273.

Weusthuis, R.A., Adams, H., Scheffers, W.A., Van Dijken, J.P., 1993. Energetics and kinetics of maltose transport in *Saccharomyces cerevisiae*: a continuous culture study. Appl. Environ. Microbiol. 59, 3102–3109.

Wichelecki, D.J., Balthazor, B.M., Chau, A.C., Vetting, M.W., Fedorov, A.A., Fedorov, E. V., Lukk, T., Patskovsky, Y. V., Stead, M.B., Hillerich, B.S., Seidel, R.D., Almo, S.C.,

Gerlt, J.A., 2014. Discovery of function in the enolase superfamily: D-mannonate and D-gluconate dehydratases in the D-mannonate dehydratase subgroup. Biochemistry 53, 2722–2731.

Wichelecki, D.J., Alyxa, J., Vendiola, F., Jones, A.M., Al-obaidi, N., Almo, S.C., Gerlt, J.A., 2014b. Investigating the physiological roles of low-efficiency D-mannonate and D-gluconate dehydratases in the enolase superfamily: pathways for the catabolism of L-gulonate and L-idonate. Biochemistry 53, 5692–5699.

Wolf, J., Stark, H., Fafenrot, K., Albersmeier, A., Pham, T.K., Müller, K.B., Meyer, B.H., Hoffmann, L., Shen, L., Albaum, S.P., Kouril, T., Schmidt-Hohagen, K., Neumann-Schaal, M., Bräsen, C., Kalinowski, J., Wright, P.C., Albers, S.V., Schomburg, D., Siebers, B., 2016. A systems biology approach reveals major metabolic changes in the thermoacidophilic archaeon *Sulfolobus solfataricus* in response to the carbon source L-fucose versus D-glucose. Mol. Microbiol. 102, 882–908.

Wolterink-van Loo, S., van Eerde, A., Siemerink, M. a J., Akerboom, J., Dijkstra, B.W., van der Oost, J., 2007. Biochemical and structural exploration of the catalytic capacity of *Sulfolobus* KDG aldolases. Biochem. J. 403, 421–30.

Wong, C., Whitesides, M., 1983. Chemical and enzymatic syntheses of 6-deoxyhexoses. Conversion to 2,5-dimethyl-4-hydroxy-2,3-dihydrofuran-3-one (Furaneol) and analogues. J. Org. Chem. 48, 3493–3497.

Xie, L., Wei, X., Zhou, X., Meng, D., Zhou, R., Zhang, Y.H.P.J., Xu, S., You, C., 2018. Conversion of D-glucose to L-lactate via pyruvate by an optimized cell-free enzymatic biosystem containing minimized reactions. Synth. Syst. Biotechnol. 1–7.

Ye, X., Honda, K., Sakai, T., Okano, K., Omasa, T., Hirota, R., Kuroda, A., Ohtake, H., 2012. Synthetic metabolic engineering-a novel, simple technology for designing a chimeric metabolic pathway. Microb. Cell Fact. 11, 120.

Yew, W.S., Fedorov, A.A., Fedorov, E. V., Almo, S.C., Gerlt, J.A., 2007. Evolution of enzymatic activities in the enolase superfamily: L-talarate/galactarate dehydratase from *Salmonella typhimurium* LT2. Biochemistry 46, 9564–9577.

Yew, W.S., Fedorov, A.A., Fedorov, E. V., Rakus, J.F., Pierce, R.W., Almo, S.C., Gerlt, J.A., 2006. Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from *Xanthomonas campestris*. Biochemistry 45, 14582–14597.

You, C., Zhang, Y.-H.P., 2012. Self-assembly of synthetic metabolons through synthetic protein scaffolds: one-step purification, co-immobilization, and substrate channeling. ACS Synth. Biol. 2, 102–110.

Zabed, H., Sahu, J.N., Suely, A., Boyce, A.N., Faruq, G., 2017. Bioethanol production from renewable sources: Current perspectives and technological progress. Renew. Sustain. Energy Rev. 71, 475–501.

Zachariou, M., Scopes, R.K., 1986. Glucose-fructose oxidoreductase, a new enzyme isolated from *Zymomonas mobilis* that is responsible for sorbitol production. J. Bacteriol. 167, 863–869.

Zhang, Y., Lin, Z., Liu, Q., Li, Y., Wang, Z., Ma, H., Chen, T., Zhao, X., 2014. Engineering of serine-deamination pathway, Entner-Doudoroff pathway and pyruvate dehydrogenase complex to improve poly(3-hydroxybutyrate) production in *Escherichia coli*. Microb. Cell Fact. 13, 1–11.

1.3 Aims and scope of this thesis

The overall aim of this thesis was to construct an enzymatic pathway, allowing a conversion of mannose, the most abundant carbohydrate in spent coffee grounds, to the versatile platform chemical lactic acid. An optimal cell-free pathway design features a high thermodynamic driving force, only depends on a low number of enzymes and cofactors and ideally functions at elevated temperatures.

In order to realise the comprehensive aim, specific steps of this study included:

I. To select and discover enzymes which are able to convert mannose to lactic acid at a low protein and cofactor cost.

II. To characterise new enzymes, that are essential for the conversion.

III. To express and purify enzymes, study their kinetics and characteristics

IV. To assemble the pathway and assess its functionality with commercially available mannose using various analytical tools

V. Hydrolyse readily available spent coffee grounds, to obtain mannose and proof operation of the cell-free pathway from waste biomass.

The main work of this thesis is divided in to three chapters (**chapters 1-3**), which are either published or prepared for submission to peer-reviewed journals.
Chapters include the review about carbohydrate catabolisms, highlighting priniciples and enzymes of non-phosphorylative pathways, which were applied for enzyme discovery and the construction of the cell-free pathway (manuscript 1, **chapter 1**).
The second manuscript (**chapter 2**) describes the finding and characterisation of a mannonate dehydratase and its potential role in a mannose metabolism, which was essential for the construction of the cell-free pathway presented in **chapter 3.** The third manuscript describes the construction and functionality of the cell-free pathway and its application to convert mannose obtained from spent coffee grounds into lactic acid

(**chapter 3**). A summary and perspectives on future research including the optimisation of the pathway is given in **chapter 4.** Additional information about renewable substrates in biocatalysis and the use of kinetic modelling was contributed to the published review found in **Appendix I** (sections 5 and 9). The construction of a yeast strain expressing a mannonate dehydratase, which was has not been investigated further, is described in **Appendix III**.

Chapter 2

Characterisation of a mannonate dehydratase from *Thermoplasma acidophilum* and its potential role in the catabolism of D-mannose

2.1 Introduction

The previous chapter highlighted the presence of oxidative pathways for a multitude of different carbohydrates. However, an oxidative pathway starting from mannose has not been identified. First indication for such as pathway was given by the finding of a mannose-specific dehydrogenase in *Thermoplasma acidophilum* (Nishiya et al., 2014). This inspired me to search for a gene encoding for a mannonate dehydratase in the same organism. A functional mannonate dehydratase would enable the second step of an oxidative mannose pathway *in vitro* and *in vivo.* In the following chapter a mannonate dehydratase from *T. acidophilum* (ManD) was identified, expressed, purified and functionally characterised. Interestingly, the gene annotation of ManD indicates a role in the metabolism of hexuronates. However, the genomic neighbourhood of the ManD gene and mannonate dehydratase genes from closely related species indicate a potential role in a oxidative metabolism of mannose. Therefore, it is demonstrated here that a mannose oxidation via the previously identified mannose-specific dehydrogenase is able to produce the substrate for ManD and its subsequent conversion of mannonate to 2-keto-3-deoxy-gluconate (KDG). Additionally we found that archaeal mannonate dehydratases share a very high amino acid sequence identity and are highly activated by the bivalent cations Co^{2+} and Ni^{2+}.

During the revision of the original thesis, the manuscript of chapter 2 was accepted and published. However, in the uploaded corrected thesis, chapter 2 has not been replaced by the published article in order to address all of the examiners comments and to keep references and terminology compatible within the chapters. A copy of the final published article can be found in Appendix II of the corrected thesis.

2.2 Contribution to manuscript 2

The concept of this publication was developed together with my main supervisor Anwar Sunna. The experiments were designed, performed and data was analysed by me. Advise on the experimental design was given by my supervisors Anwar Sunna and Robert Willows. The manuscript was written by me and was reviewed by all authors. The contributions of each author are given in Table 2.1

Table 2.1 Author contributions for manuscript 2

	Dominik Kopp	Robert Willows	Anwar Sunna
Conception	●		●
Experiment design	●	●	●
Data Collection	●		
Data Analysis	●		
Manuscript	●	●	●

2.3 Manuscript 2

Characterisation of a mannonate dehydratase from *Thermoplasma acidophilum* and its potential role in the catabolism of D-mannose

Dominik Kopp[a], Robert D. Willows[a,b] and Anwar Sunna[a,b]*

[a] Department of Molecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia

[b] Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, New South Wales 2109, Australia

*Corresponding author

Anwar Sunna

Department of Molecular Sciences

Macquarie University

North Ryde, NSW 2109 Australia

Phone: +61 2 9850 4220

Fax: +61 2 9850 8313

Email: anwar.sunna@mq.edu.au

Abstract

Mannonate dehydratases catalyse the dehydration reaction from mannonate to 2-keto-3-deoxygluconate as part of the hexuronic acid metabolism in bacteria. Previous biochemical and physiological studies mostly focused on examples from bacteria or members of the enolase superfamily. Here, we identified and characterised the first archaeal mannonate dehydratase (ManD) from the thermoacidophilic archaeon *Thermoplasma acidophilum* that does not belong to the enolase superfamily. The enzyme shows high specificity towards D-mannonate and its lactone, D-mannono-1,4-lactone. The activity of most bacterial mannonate dehydratases (e.g. in *Escherichia coli*) is dependent on Fe^{2+}, Mn^{2+} and Mg^{2+}, while the archaeal mannonate dehydratase studied shows equally high activity with Co^{2+} or Ni^{2+}. The gene encoding for ManD is located adjacent to a previously studied mannose-specific aldohexose dehydrogenase (AldT) in the genome of *T. acidophilum*. Using nuclear magnetic resonance (NMR) spectroscopy, we demonstrated that the mannose-specific AldT produces the substrates for ManD, demonstrating the possibility for an oxidative metabolism of mannose in *T. acidophilum*. Among previously studied mannonate dehydratases, ManD shows closest homology to enzymes belonging to the xylose isomerase-like superfamily. Genetic analysis revealed that closely related mannonate dehydratases among archaea, are not located in a hexuronate gene cluster like in bacteria, but next to putative aldohexose dehydrogenases, implying a different physiological role of mannonate dehydratases in those archaeal species.

Keywords

mannonate dehydratase, mannose metabolism, *Thermoplasma acidophilum*, mannono-1,4-lactone, bivalent cations, 2-keto-3-deoxygluconate

1 Introduction

Mannonate dehydratases (EC 4.2.1.8) catalyse the conversion of mannonate to 2-keto-3-deoxygluconate (KDG) and have been studied as part of the hexuronic acid gene cluster in several organisms such as *Escherichia coli*, *Bacillus stearothermophilus*, *Bacillus subtilis* and *Erwinia chrysanthemi* (Hugouvieux-Cotte-Pattat and Robert-Baudouy, 1987; Mekjian et al., 1999; Portalier et al., 1980; Robert-Baudouy et al., 1982). The hexuronate gene cluster encodes enzymes involved in the metabolism of glucuronate and galacturonate (Mandrand-Berthelot et al., 2004). Galacturonate is a common sugar acid present in pectin a constituent of plant cell walls and can serve as the only carbon source for growth of some bacteria (Lawford and Rousseau, 1997; Mekjian et al., 1999). Furthermore, glucuronate is present in the intestine of mammals, providing a carbon source for anaerobic gut bacteria such as *E. coli* (Peekhaus and Conway, 1998). As part of the glucuronate/galacturonate metabolism, mannonate dehydratase converts mannonate to KDG. In *E. coli,* and some species of *Erwinia,* KDG is phosphorylated and cleaved into pyruvate and 3-phosphoglycerate entering the central carbon metabolism (Mandrand-Berthelot et al., 2004) Mannonate dehydratases are represented in different enzyme families, such as the xylose isomerase-like superfamily or the enolase superfamily.

Mannonate dehydratases encoded in the bacterial hexuronate gene clusters belong to the xylose isomerase-like superfamily. Crystal structures have been solved for xylose isomerase-like mannonate dehydratases from *Streptococcus suis, E. coli K12* and *Enterococcus faecalis* in native form and in complex with Mn^{2+} ions (Qiu et al., 2012; Zhang et al., 2009). It has been shown that His311 and Tyr325 in the binding pocket were crucial for the activity of the dehydratase in Gram-positive bacteria such as *S. suis*. The basic His311 extracts the C2-proton, while Tyr325 is necessary for the abstraction of the OH group at carbon 3 of the substrate. Members of the enolase superfamily show a conserved structure and reaction mechanism, but differ in their physiological functions. Within their conserved barrel structure, mannonate dehydratases of this family share conserved ligand binding sites for Mg^{2+}, which is essential for the stabilisation of an enediolate intermediate (Wichelecki et al., 2014). Among mannonate dehydratases of the enolase superfamily, some representatives with diverse functions have been found which are involved in catabolism of sugar acids other than mannonate (Wichelecki et al., 2014b). Crystal structures of mannonate dehydratases belonging to the enolase superfamily have been solved from

Chromohalobacter salexigens and *Novosphingobium aromaticivorans* (Rakus et al., 2007; Wichelecki et al., 2014). However, archaeal mannonate dehydratases have not been investigated so far.

Here, we used genomic analysis of archaea to identify a functional mannonate dehydratase from *T. acidophilum* (ManD) that showed a high amino acid identity to mannonate dehydratases from the xylose isomerase-like superfamily. In the genome of *T. aciodphilum,* ManD is located adjacent to an aldohexose dehydrogenase (AldT) which was shown to catalyse the oxidation of mannose to mannonate with high specificity (Nishiya et al., 2014; Yasutake et al., 2007). However, the physiological significance for the oxidation and the oxidation products were not investigated further. In principle, mannonate produced by AldT could be further converted by ManD into the central carbon metabolite KDG and enable generation of energy and biomass from mannose.

In this work we characterised a recombinantly-expressed and purified ManD from the *T. acidophilum*. ManD was found to be dependent on bivalent cations and showed higher activity with mannonate compared to D-mannono-1,4-lactone in physiological buffer. Genomic analysis of other thermoacidophilic archaea showed the location of mannonate dehydratases outside of the known hexuronate gene cluster but adjacent to putative aldohexose dehydrogenases, suggesting the possibility of a different physiological function for the xylose isomerase-like mannonate dehdyratases in archaea.

2 Materials and Methods

2.1 Cloning, expression and purification of enzymes

All plasmids were constructed by amplification of genes from *T. acidophilum* DSM 1728 genomic DNA (DSMZ, Braunschweig, Germany) using PCR, restriction and ligation into vectors according to standard protocols (Sambrook et al., 2001). Primer pairs and restriction enzymes used for the construction of each expression plasmid are summarised in Table 1. *E. coli* strain α-select (Bioline, Alexandria, AUS) was used in all initial cloning procedures. The expression plasmids pProEX HTa-Ta0753, pETDuet-1-Ta0085m, pETDuet-1-Ta0752 and pETDuet-1-Ta0756 were then used to transform BL21-CodonPlus (DE3)-RIL competent cells (Agilent Technologies, Santa Clara, USA).

Plasmid pETDuet-1-AldT was used to transform BL21 (DE3) cells (NEB, Ipswitch, USA). BL21-CodonPlus (DE3)-RIL strains containing the respective expression plasmid were grown in 500 ml Luria–Bertani medium (LB) supplemented with carbenicillin (100 µg/ml) and chloramphenicol (35 µg/ml) at 37°C to an OD_{600} of 0.6 before induction was performed with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 20°C for 16-18 h. The BL21 (DE3) cells carrying the expression plasmid pETDuet-1-AldT were grown in 500 ml LB medium containing carbenicillin (100 µg/ml) to an OD_{600} of 0.5 at 37°C before induction was performed with 0.4 mM IPTG at 37°C for 4 h. All cultures were harvested by centrifugation at 5000 x *g* for 20 min, resuspended in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7 and lysed by three passages through a French pressure cell. Soluble fractions of the lysates were obtained by centrifugation at 15000 x *g* for 20 min. His-tagged enzymes were purified by subjecting the soluble extracts to a 5 ml nickel-nitrilotriacetic acid (Ni-NTA) affinity column (GE Healthcare) equilibrated with buffer containing 50 mM sodium phosphate (NaP) 300 mM NaCl and 20 mM imidazole. The proteins were eluted isocratically with a final concentration of 400 mM imidazole. The buffer of the eluates was exchanged to 50 mM HEPES pH 7 by 3 rounds of centrifugation in Amicon Ultra centrifugal filters (Millipore) with a 30 kDa molecular mass cut-off.

In order to obtain a ManD without a N-terminal His-tag, the purified ManD obtained from the pProEX HTa-Ta0753 expression was incubated with a TEV Protease (Promega, Madison, USA). Digestion was performed according to the manufacturer's protocol with addition of 10% glycerol (v/v, final concentration) and incubation for 1.5 h at 30°C followed by incubation at 4°C overnight. Removal of the His-tag and the TEV Protease was achieved by a second Ni-NTA purification using a 1ml ZetaSep Ni-NTA affinity column (emp Biotech, Berlin, Germany). The flow-through contained the pure protein and cleavage of the His-tag was verified by SDS-polyacrylamide gel electrophoresis (PAGE).

Table 1 Oligonucleotides used in this study. Engineered restriction sites are in lowercase. RE: restricton enzyme.

Expression plasmid	amplified gene locus	oligonucleotide used	oligonucleotide sequence (5' –> 3')	RE used
pETDuet-1-AldT	Ta0754	AldT-fw	TGATggatccCATGTTCAGCGATCTAAGGGA	BamHI
	Ta0754	AldT-rev	ATTATgaattcCTCATTCTGGCGTGCTTATGG	EcoRI
pProEX HTa-Ta0753	Ta0753	ManD_woATG_fw	TGATggatccTAAGCTCAGAATAGCCGAGATC	BamHI
	Ta0753	ManD_HindIII_rev	TGATaagcttTCAGGAAATCTGTTTATGATTTG	HindIII
pETDuet-1-Ta0085m	Ta0085m	Ta0085_BamHI_fw	TGATggatccTGGAGAAAAATCGCGTGAGA	BamHI
	Ta0085m	Ta0085_SalI_rev	TGATgtcgacTTATTTTTGTCGGTATGGCTTG	SalI
pETDuet-1-Ta0752	Ta0752	Ta_rspA_BamHI_fw	TGATggatccTAGCAAGATAAAATCGCTCGAAA	BamHI
	Ta0752	Ta_rspA_HindIII_rev	ATTATaagcttTCATTCGAAGTCATATCTTATCC	HindIII
pETDuet-1-Ta0756	Ta0756	Ta0756_SalI_fw	TGATgtcgacTTTCTGAATTCGTCGAAGTCTTT	SalI
	Ta0756	Ta0756_NotI_rev	ATTATggccgcTCAGGAGAATCCTCTCTTCTC	NotI

2.2 Protein analysis

Purified proteins were separated by sodium dodecyl sulfate-polyacrylamide gelelectrophoresis (SDS-PAGE) using 4-15% Tris-glycine polyacrylamide gels (Bio-Rad Laboratories, Hercules, USA) and visualised with Coomassie Brilliant Blue. For the identification of ManD the single protein band at 38 kDa was excised from the gel and analysed by liquid chromatography electrospray ionization tandem mass spectrometry (LC ESI MS/MS) at the Australian Proteome Analysis Facility (APAF, Macquarie University) as described elsewhere (Atack et al., 2015). Protein identification was performed using the mascot software (Matrixscience, USA). The native molecular mass of the protein was determined by size exclusion chromatography on a ÄKTA pure FPLC system using a Superdex Increase 200 10/300GL column (GE Healthcare) equilibrated with 10 mM NaP buffer, 140 mM NaCl pH 7. A calibration curve was obtained with molecular mass markers thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa) and ovalbumin (43 kDa).

2.3 NMR analysis

D-Mannono-1,4-lactone (>95%) was purchased from TCI Chemicals Co. (Japan). For NMR spectra of the sugar acid in different substrate forms, 1 M D-Mannono-1,4-lactone solutions were prepared in 1 M NaOH, 1 M HCl and in ultrapure water. All solutions

were incubated for 1 h at room temperature, before each solution was diluted to 100 mM using ultrapure water. Then, 300 µl of each 100 mM solution (equivalent to 5 mg sugar acid) was freeze-dried and resuspended in 500 µl 0.1 M sodium phosphate (NaP) buffer pH 7. After transfer of each sample to 5 mm NMR tubes 10% D_2O and 40 µM 3-(trimethylsilyl)propionic-2,2,3,3-d4 acid (TMSP) chemical shift standard were added. 1D decoupled ^{13}C NMR spectra of non-isotopically labelled samples were acquired on a 500 MHz Bruker Avance III HD NMR equipped with a BBFO probe at 50°C (323 K), using power-gated proton decoupling with a 90° pulse with 512 scans and a 3 s relaxation delay between scans.

Oxidation of D-[1,6-$^{13}C_2$]mannose was performed in reactions containing 5 mM of the substrate, 5 mM NAD^+, 20 mM NaP buffer at pH 7 and 13.5 µg purified AldT. Prior to the addition of purified enzyme and after 30 min of reaction at 50°C, 1D power-gated proton decoupled ^{13}C NMR spectra were acquired using a 90° pulse with 4 scans and a 3 s relaxation delay between scans.

Visualisation of all spectra was performed with iNMR 6.0 (http://www.inmr.net). After Fourier transformation and automatic phase correction, an exponential visual weighting factor of 1.5 and a smoothing factor of 10 were applied to all spectra.

2.4 Enzyme activity

D-mannono-1,4-lactone was either used after preparation in 50 mM HEPES at pH 7 (indicated as 'lactone in buffer') or after hydrolysis to its free acid form (indicated as 'mannonate' or 'free sugar acid'). Hydrolysis was performed according to Lamble et al. by preparing 1 M stock solutions of D-mannono-1,4-lactone in 1 M NaOH and incubation at room temperature for 1 h before dilutions were prepared in 50 mM HEPES pH 7 (Lamble et al., 2004). For temperature optimum, pH optimum and thermostability, activity was quantified using the thiobarbituric acid (TBA) assay, according to a modification by Buchanan et al. (Buchanan et al., 1999; Skoza and Mohos, 1976). Unless stated otherwise, reactions (60 µl) contained 0.5 - 1.5 µg pure enzyme, 10 mM mannonate, 50 mM HEPES pH 7 (adjusted at 55°C) and 1 mM $CoSO_4$. Reactions were stopped by incubation on ice or if needed by addition of 6 µl 12.5% trichloracetic acid (TCA). 50 µl of the reaction mixture was oxidised using 125 µl of 25 mM periodic acid in 0.25 M H_2SO_4 at room temperature for 20 min. The oxidation was stopped by addition of 250 µl of 2% (w/v) sodium arsenite in 0.5 M HCl. Finally, 1 ml of 0.3% TBA in water was added to the samples and the reaction mixture was boiled for 10 min. The

formation of 2-keto-3-deoxy sugar acid was determined by reading A_{549nm} using 96-well microtiter plates in a Spectrostar Nano plate reader (BMG Labtech, Ortenberg, Germany) and quantified using the molar extinction coefficient of 67.8 x 10^3 M^{-1} x cm^{-1}.
The effect of pH on enzyme activity was analysed in reactions with 1 µg enzyme and 120 mM universal buffer (Britton and Robinson, 1931) instead of HEPES. Reactions were incubated for 2 h at 55°C before they were analysed with the TBA Assay. In reactions with D-mannono-1,4-lactone in buffer, pH decreased during the course of the experiment, which was accounted for in the analysis by measurement of the pH after the reaction. The optimal temperature for enzyme activity was determined in reactions containing 1 µg enzyme. Reactions were performed for 2 h at different temperatures before being analysed with the TBA Assay. For thermostability assays, 120 mM universal buffer was used and the purified enzyme (1 µg) was incubated at temperatures between 55°C and 95°C in the absence of substrate. Samples were removed at different times and the residual activity was measured in reactions with 10 mM mannonate for 45 min at 55°C.

The semicarbazide assay was used according to Wichelecki et al. (Wichelecki et al., 2014) to determine the effect of metal ions, chelators and reducing agents, and the acquisition of Michalis-Menten kinetics (K_m, V_{max}). Each reaction (60 µl) was incubated with 240 µl semicarbazide reagent at room temperature for 1 h. A_{250nm} was read in a UV transparent microtiter plate (Thermo Scientific) using a Spectrostar Nano plate reader. Product formation was quantified using a standard curve of 2-keto-3-deoxygluconate (Sigma-Aldrich) prepared in the same assay buffer.
For the effect of metals, chelators and reducing agents, pure enzyme (0.6 µg) was incubated with 1 mM of each additive in 50 mM HEPES pH 7 for 1 h on ice. Enzyme activity was then measured in reactions containing 10 mM D-mannono-1,4-lactone in 50 mM HEPES pH 7, incubated for 1 h at 55°C and then analysed with the semicarbazide assay. Kinetic data for the determination of K_m and V_{max} were performed in triplicate reactions (60 µl), containing 1 µg purified ManD, 50 mM HEPES pH 7, 1 mM $CoSO_4$ and different substrate concentrations (1-50 mM). Linear increase of reaction product was assured over 1 h reaction time. Reactions were performed for 45 min at 55°C before product formation was determined using the semicarbazide assay.
Substrate specificity was analysed using 10 different sugar acids (Fig. 8B). Reactions contained 0.5 µg enzyme, 50 mM HEPES pH 7, 1 mM $CoSO_4$ and 10 mM of each sugar

acid. Reactions were incubated for 16 h at 55°C before they were analysed using the semicarbazide assay.

2.5 High performance liquid chromatography (HPLC) analysis

Reactions catalysed by ManD and standards of KDG and mannonate were analysed using an Agilent 1290 HPLC system (Agilent Technologies, Santa Clara, USA) connected to a refractive index detector (RID) G1362A (Agilent Technologies). Samples were analysed on an organic acid column (Agilent HiPlex H^+) with 10 mM H_2SO_4 as a mobile phase at a flow rate of 0.6 ml/min. The column was heated to 80°C and the RID was set to 55°C. Reactions were stopped by addition of 7 µl 12.5% TCA to a 60 µl reaction. After short centrifugation, 10 µl of the supernatant was injected into the HPLC for analysis.

2.6 Statistical analysis

The programs Excel (Microsoft Corporation, USA) and Prism 6 (GraphPad software, La Jolla, USA) were used for statistical analysis of ManD kinetic data. An unpaired two-tailored t-test was used to analyse the difference in V_{max} and K_m for both substrates.

3 Results and Discussion

3.1 Screening for a functional mannonate dehydratase in *T. acidophilum*

Based on the finding of the mannose-specific AldT, we screened the genomic neighbourhood of its gene locus (Ta0754) for enzymes which could display activity towards mannonate. In the proximity of Ta0754 several genes are located that encode for hypothetical dehydratases, suggesting their potential to convert mannonate to KDG (Table 2). Gene Ta0753 is located adjacent to the previously described mannose-specific AldT and is annotated as *uxuA* D-mannonate dehydratase indicating its role in the hexuronate metabolism in most bacteria. We also investigated genes annotated as L-rhamnonate dehydratase (Ta0752) and galactonate dehydratase (Ta0756), since a common characteristic for dehydratases from the enolase superfamily is their high level of promiscuity (Lamble et al., 2004; Yew et al., 2006). Ta0752 and Ta0756 are annonated as member of the enolase superfamily (subgroup MR/MLE), which contains

enzymes with conserved mechanisms and structure but yet unknown substrate specificities and physiological functions (Gerlt et al., 2005; Wichelecki et al., 2014). Candidate Ta0085m was investigated as it has been assumed to be a gluconate dehydratase in the non-phosphorylating Entner-Doudoroff (np-ED) pathway (Sun et al. 2010). Some gluconate dehydratases of the np-ED pathway have been shown to be promiscuous towards several substrates, such as the gluconate dehydratase in *Sulfolobus solfataricus* (Kim and Lee, 2005; Lamble et al., 2004).

All four candidates (Ta0085m, Ta0753, Ta0752, Ta0756) were genetically fused to His-tags and expressed in *E. coli* containing additional tRNAs for rare codons. However, only the expression of Ta0753 yielded a soluble and functional enzyme whereas expression of the other candidates did not yield enough soluble protein to be purified via their His-tag (Fig. S1). We identified the protein product of the Ta0753 gene as a functional mannonate dehydratase from *T. acidophilum* (from here on referred to as ManD) and compared its properties and genomic context to other previously studied mannonate dehydratases.

Table 2 Overview of genes potentially encoding for mannonate dehydratases. MR/MLE: Mandelate racemase/Muconate lactonising enzyme. The gene product of Ta0753 is referred to as ManD in the following and was further investigated as part of this study.

old locus tag	genebank identifier/ new locus tag	KEGG annotation	protein family	soluble expression
Ta0085m	CAC11233.1/ TA_RS00440	MR/MLE family protein	MR/MLE	no
Ta0752	CAC11886.1/ TA_RS03865	L-rhamnonate dehydratase	MR/MLE	no
Ta0753	CAC11887.1/ TA_RS03870	*uxuA* D-mannonate dehydratase	uxuA family	yes
Ta0756	CAC11889.1/ TA_RS03885	galactonate dehydratase	MR/MLE	no

Among previously characterised mannonate dehydratases, ManD shares highest sequence identity with enzymes in the xylose isomerase-like superfamily such as the mannonate dehydratase from *S. suis* (31.5% sequence identity) (Zhang et al., 2009). Key amino acid residues for substrate binding (His311 and Tyr325) and binding of Mn^{2+} (Asp310 and Cys237) in the structure of *S. suis* are conserved in the amino acid sequence of ManD. However, a much higher amino acid sequence identity (62.9 %) is

shared between ManD and mannonate dehydratases from closely related archaeal species, *Ferroplasma acidarmanus* and *Ferroplasma acidophilum*.

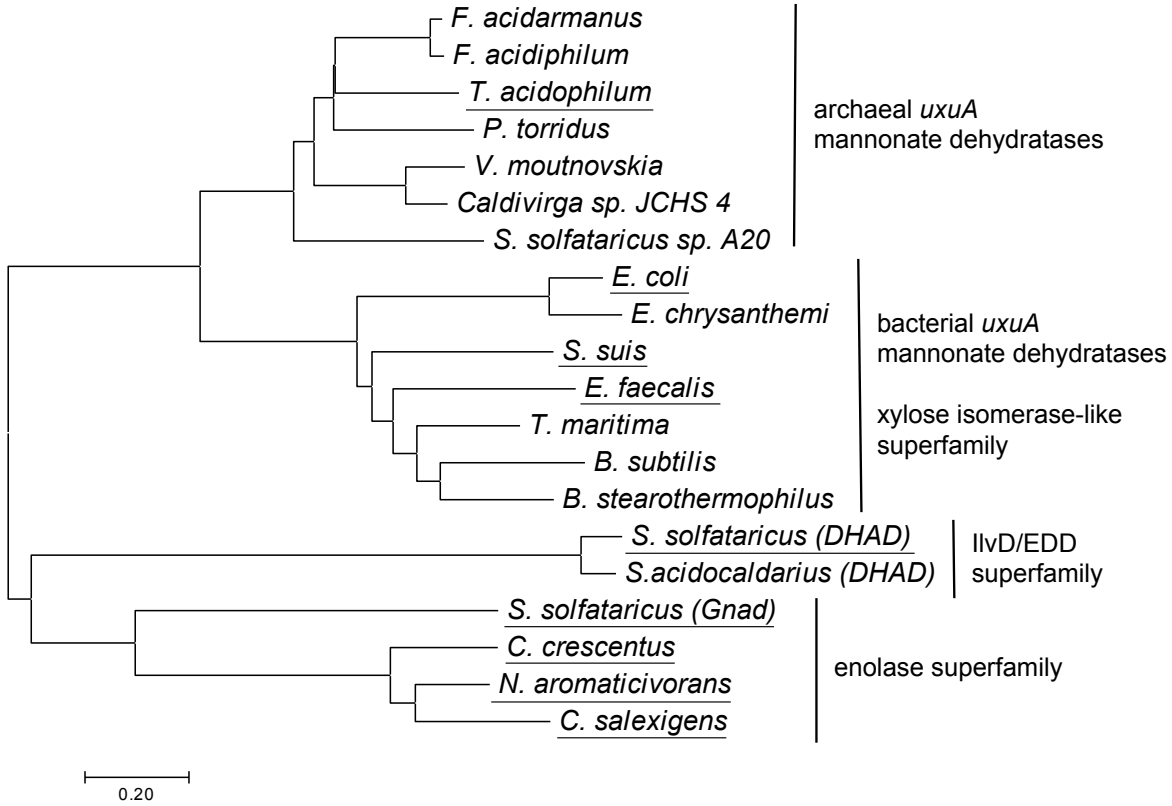


Figure 1 Phylogenetic relationship of different putative and confirmed dehydratases belonging to different enzyme families. Characterised enzymes are underlined. Evolutionary analyses were conducted in MEGA 7. 20 protein sequences were aligned using MUSCLE. The phylogenetic tree was inferred using the neighbour-joining method. The scale bar indicates an evolutionary distance of 0.20 nucleotide per position in the sequence. Protein sequences were retrieved from the Uniprot database and their entry codes are as follows. *Ferroplasma acidarmanus* - S0AL33, *Ferroplasma acidophilum* - A0A1V0N416, *Thermoplasma acidophilum* - Q9HK52, *Picrophilus torridus* - Q6L2R9, *Vulcanisaeta moutnovskia* - F0QYL3, *Caldivirga* sp. JCHS 4 - A0A101XEY7, *Sulfolobus solfataricus* sp. A20 - A0A1C8ZTN0, *Escherichia coli* K12 - P24215, *Erwinia chrysanthemi/Dickeya dadantii* - E0SEP1, *Streptococcus suis* - A0A142UME2, *Enterococcus faecalis* - Q82ZC9, *Thermotoga maritima* - Q9WXS4, *Bacillus subtilis* - O34346, *Bacillus stearothermophilus* - A0A087LHB3, *Sulfolobus solfataricus* DHAD - Q97UB2, *Sulfolbus acidocaldarius* DHAD Q4J860, *Sulfolobus solfataricus* Gnad - Q97U27, *Caulobacter crescentus* - Q9AAR4, *Novosphingobium aromaticivorans* - A4XF23, *Chromohalobacter salexigens* - Q1QT89.

Phylogenetic analysis showed that mannonate dehydratases from *T. acidophilum*, *F. acidarmanus* and *F. acidophilum* are closer related to bacterial mannonate dehydratases from the xylose isomerase-like superfamily, than to bacterial or archaeal mannonate dehydratases of the enolase superfamily (Fig. 1). The only two archaeal mannonate dehydratases characterised so far are from *N. aromaticivorans* and *C. salexigens*, and have been identified as members of the enolase superfamiliy (Rakus et al., 2007; Wichelecki et al., 2014). Mannonate dehydratases of the enolase superfamily have a substantially different structure from those in the xylose isomerase-like

superfamily and, therefore, are only distantly related. Despite fulfilling a similar function, the dihydroxyacid dehydratase from *S. solfataricus,* which belongs to the IlvD/EDD superfamily, is rather unrelated to the xylose isomerase-like and enolase superfamilies. For bacterial xylose isomerase-like mannonate dehydratases, a physiological function of the enzyme in the catabolism of hexuronates has been demonstrated (Ashwell, 1962). Although ManD is annotated as *uxuA* mannonate dehydratase, which implies a role in the metabolism of hexuronates, it is not present in a classical hexuronate gene cluster known from bacteria such as the one described in *B. stearothermophilus* (Fig. 2) (Shulami et al., 1999). In contrast, the gene encoding for ManD is located adjacent to AldT in the genome of *T. acidophilum*. Similarly, the two closely related mannonate dehydratases in *F. acidarmanus* and *F. acidophilum* are also located adjacent to aldohexose dehydrogenases. A different physiological role for these mannonate dehydratases can therefore be proposed. In the following, we functionally characterised ManD and show that AldT is able to produce the substrate needed for a subsequent conversion mediated by ManD.

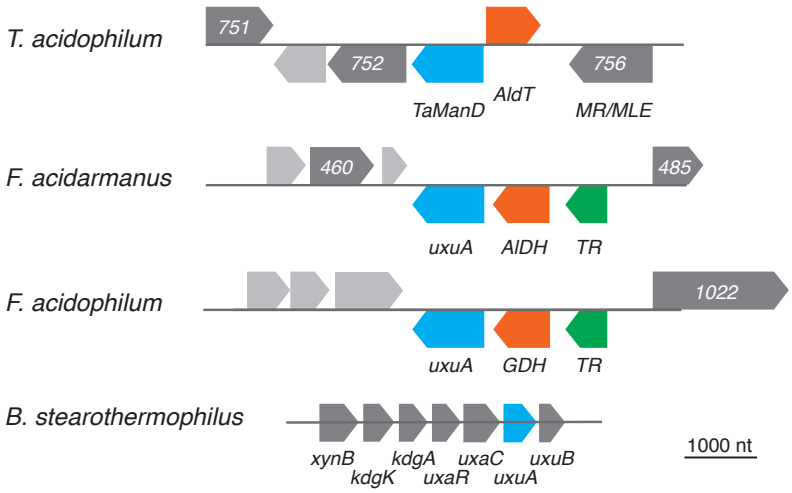


Figure 2 Schematic maps of *uxuA* mannonate dehydratases and their genomic neighbourhoods in *T. acidophilum*, *F. acidarmanus, F. acidophilum* and *B. stearothermophilus*. Blue/ManD/uxuA: uxuA D-mannonate dehydratases, orange/AldT/AlDH: aldohexose dehydrogenases, green/TR: transcriptional regulators, light grey: annotated as hypothetical proteins. 750: FAA hydrolase family protein, 752: L-rhamnonate dehydratase, 460: transposase, 485: acetyl-CoA synthetase, 1022: acetate-CoA ligase, *xynB*: β-xylosidase, *kdgA:* 2-keto-3-deooxy-6-phosphogluconate aldolase, *kdgK*: 2-keto-3-deoxygluconate kinase, *uxaR:* regulatory protein, *uxaC*: hexuronate isomerase, *uxuB*: mannonate oxidoreductase.

3.2 Substrate conformation for ManD activity

Before acquiring enzyme kinetics and studying the characteristics of ManD, we investigated which substrate conformation is encountered by ManD under physiological conditions. In aqueous solutions free sugar acids coexist with their lactone in an equilibrium, which is defined by the stability of the lactone, the temperature and the pH of the solution. Many lactones hydrolyse spontaneously in water, although several lactones including D-arabinonolactone (Palleroni and Duodoroff, 1957), L-rhamnonolactone (Watanabe et al., 2008) and D-xylonolactone (Stephens et al., 2007) have been reported to be hydrolysed by lactonases. Mannonate is able to form a lactone by covalent bond formation between carbon 1 and carbon 5 (δ-lactone) or carbon 1 and carbon 4 (γ-lactone) (Fig. 3A). The equilibrium of the two different envelope forms has been studied by NMR spectroscopy and it was found that the γ-lactone is strongly favoured over the δ-lactone (Wałaszek and Horton, 1982). However, these studies do not describe equilibria between lactone and free acid form. In a cellular environment mannonate can either be produced by a mannonate oxidoreductase (*uxuB*) as part of the catabolism of hexuronic acids or by an aldohexose dehydrogenase such as AldT in a hypothetical oxidative catabolism of mannose. In order to study the conformation of the substrate for ManD in a physiological environment, we first acquired decoupled ^{13}C NMR spectra for D-mannono-1,4-lactone in physiological buffer at pH 7, in its lactone form (incubated with HCl) and after hydrolysis to the free sugar acid (incubated with NaOH) (Fig. 3B). The spectra obtained indicated that the lactone and the free sugar can be distinguished by their chemical shifts. In the free sugar acid form, carbon 6 yields a chemical shift at 63.51 ppm, whereas in D-mannono-1,4-lactone carbon 6 displays a chemical shift at 63.02 ppm. D-mannono-1,4-lactone in buffer shows both chemical shifts, indicating that D-mannono-1,4-lactone and D-mannonate are present in an equilibrium at pH 7.

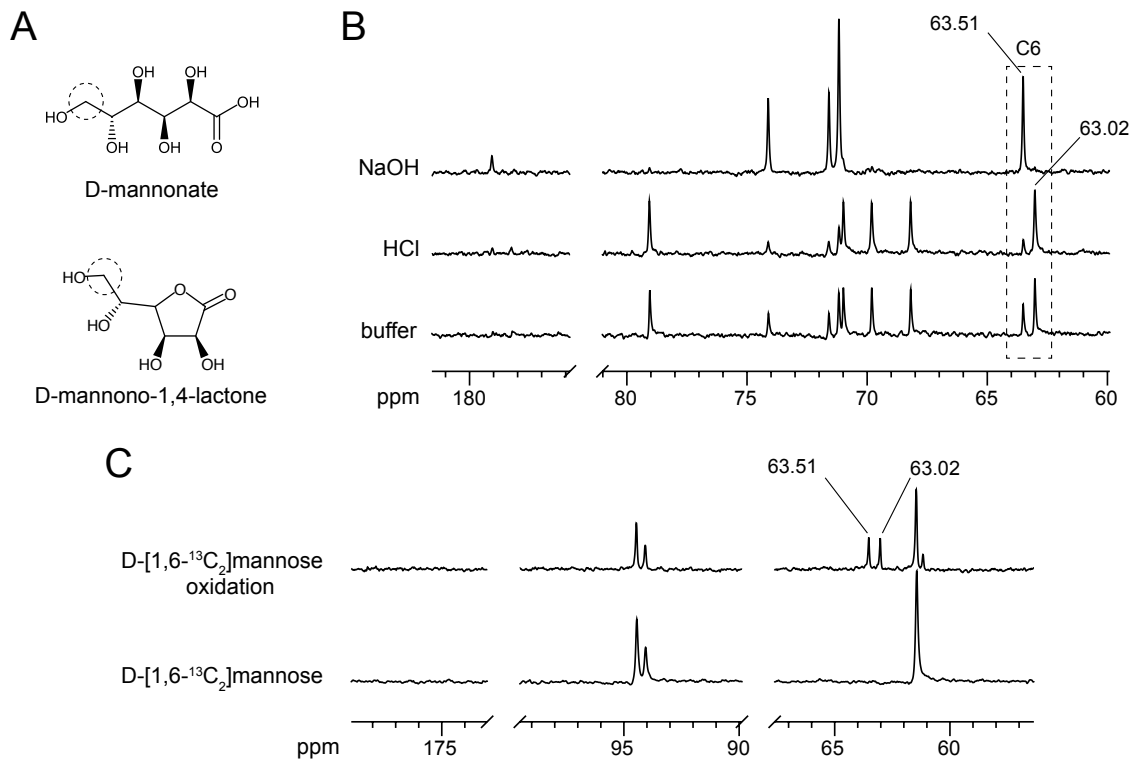


Figure 3 Analysis of D-mannono-1,4-lactone under different conditions and products of AldT-mediated oxidation of isotope labelled mannose**. A)** Chemical structures of the free sugar acid, D-mannonate and its lactone, D-mannono-1,4-lactone. Carbon 6, which was used to differentiate the compounds in NMR spectra, is circled for both structures. **B)** Decoupled 1D ^{13}C NMR spectra of D-mannono-1,4-lactone incubated in 1 M NaOH, 1M HCl and in water with subsequent dilution in 0.1 M NaP buffer ph 7 (buffer). Chemical shift for carbon 6 is shown by a dashed rectangle. Characteristic chemical shifts for the lactone (63.02 ppm) and for the free sugar acid (63.51 ppm) are indicated. **C)** Decoupled 1D ^{13}C NMR spectra before (bottom) and after (top) oxidation of mannose with AldT. Chemical shifts of carbon 6 for lactone and free sugar acid are indicated.

Next, we oxidised isotope-labelled D-[1,6-$^{13}C_2$]mannose using recombinantly-expressed AldT and acquired 1D ^{13}C NMR spectra of the reaction product (Fig. 3C). The chemical shifts for carbon 6 of lactone and D-mannonate were used to analyse the reaction product of the mannose oxidation with AldT. The spectrum after oxidation of D-[1,6-$^{13}C_2$]mannose proved that AldT activity is able to provide an equilibrium of D-mannono-1,4-lactone and D-mannonate for the subsequent dehydration by ManD.

3.3 Expression and purification of ManD

For the further biochemical characterisation, ManD was expressed together with a tabacco etch virus (TEV) sequence, which allowed the proteolytic cleavage of the His-tag after the first metal affinity purification (Fig. 4). In order to study the effect of bivalent metal cofactors on enzyme activity, cleavage of the His-tag from the recombinant enzyme was necessary since cations (Ni^{2+}, Mg^{2+}) might influence enzyme activity assays of a His-tagged enzyme. After His-tag cleavage, the recombinant enzyme showed a single band on SDS-PAGE with an apparent mass of 38 kDa (Fig. 4). Peptide mass fingerprinting was performed with LC ESI MS/MS and NCBI database searches using the mascot software confirmed the identity of the purified protein as *uxuA* mannonate dehydratase from *T. acidophilum*. In addition, the apparent native molecular mass of the protein was estimated from size exclusion chromatography to be 225 kDa, indicating a hexameric structure of the protein.

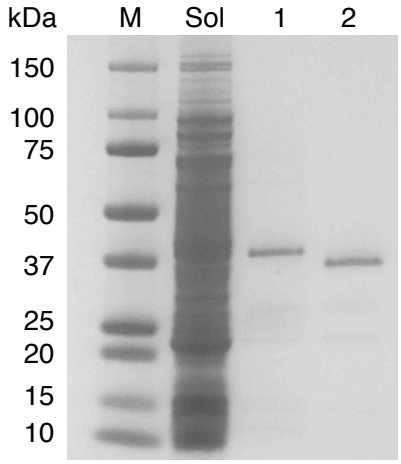


Figure 4 Purification of the Ta0753 gene product (ManD) after expression in *E. coli.* M: Bio-rad Precision Plus Protein Standard, Sol: soluble cell extract, 1: Enzyme after first His-tag purification, expected size 42 kDa, 2: Enzyme after TEV cleavage and second His-tag purification, expected size 38 kDa.

3.4 Enzyme characterisation

Initial tests showed that ManD lost almost its complete activity after purification. This observation suggested that essential cofactors for the activity of ManD were lost during the purification process. Accordingly, the effect of several additives was tested for their influence on the activity of ManD with D-mannono-1,4-lactone. ManD showed strongly enhanced activity in the presence of Mn^{2+}, Ni^{2+}, Mg^{2+}, Ca^{2+} and Co^{2+}, but no increase in activity with Fe^{2+}, Zn^{2+} was observed, and activity decreased with Cu^{2+} (Fig. 5).

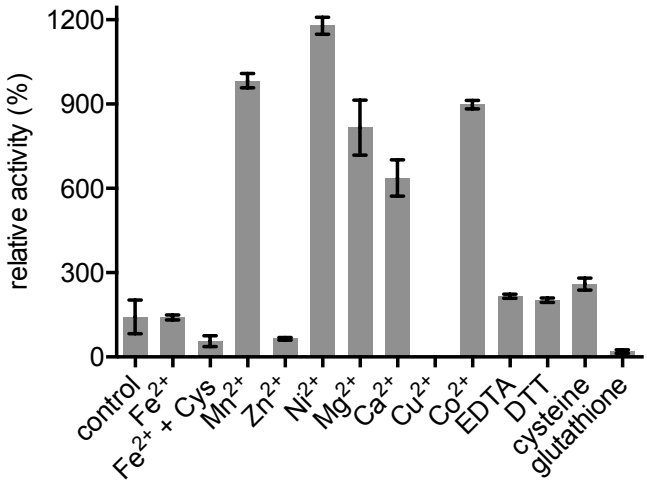


Figure 5 Effect of metal ions, chelating and reducing agents on the activity of ManD. Purified enzyme (0.6 µg) was pre-incubated for 1 h with 1 mM of each additive in 50 mM HEPES pH 7. Activity was then determined in reactions containing 10 mM D-mannono-1,4-lactone in 50 mM HEPES pH 7 incubated for 1 h at 55°C before analysis with the semicarbazide assay. Fe^{2+} + Cys: Fe^{2+} was prepared with 1 mM cysteine. Error bars represent standard deviation of mean values for duplicate experiments.

None of the bacterial mannonate dehydratases described so far have been shown to have an equally strong activation with Co^{2+} and Ni^{2+} as observed with ManD. Early investigations of bacterial mannonate dehydratase focussed on *E. coli* and showed different strengths in activation after incubation with bivalent cations (Dreyer, 1987; Robert-Baudouy et al., 1982; Robert-baudouy et al., 1975; Robert-Baudouy and Stoeber, 1973). In summary, all studies showed that Fe^{2+} and Mn^{2+} in combination with reducing agents resulted in highest enzyme activity (90-100% of maximum enzyme activity). Incubation with other bivalent cations resulted in less activity compared to the enzyme's maximum activity (Co^{2+}: 27-80%, Ni^{2+}: 5-40% and Zn^{2+}: 5-25 %, depending on the study).

More recently, crystal structures of xylose isomerase-like mannonate dehydratases from *S. suis* and *E. coli* have been solved and revealed the presence of primarily Mn^{2+} and lower amounts of Mg^{2+} and Zn^{2+} in binding sites of the enzymes (Qiu et al., 2012; Zhang et al., 2009). Although ManD shares high amino acid sequence identity with mannonate dehydrates from the xylose isomerase superfamily, its high activation by Ni^{2+} and Co^{2+} is not observed by previous studies of bacterial members of the same superfamily. The effect of metal ions was studied in terms of its activation. However, metal ion that act as cofactors can also have a stabilising effect, which was demonstrated in other studies. Amino acid residues essential for the binding of Mn^{2+} in ManD from *S. suis* are conserved in the protein seqeunce of ManD from *T. acidophilum*.

Therefore, one could assume that the different metal ions bind to the same amino acid residues.

T. acidophilum was originally isolated from a hot and acidic environment and accordingly this archaeon displays optimal temperature and pH for growth at 59°C and pH 1-2, respectively (Darland et al., 1970). The purified ManD was active between 35 and 70°C, with an optimal temperature for activity at 65°C (Fig. 6A). Thermostability of the ManD enzyme in the absence of substrate was studied at temperatures between 55°C and 95°C, and activity decreases rapidly once incubated at temperatures higher than 50°C (Fig. 6B) The enzyme retains its full activity at 55°C for at least an hour, whereas complete inactivation was observed at 75°C, 85°C and 95°C within 90 min, 60 min, and 15 min, respectively. The difference in the relatively high optimal temperature of the enzyme to the comparably low thermostability indicated that the enzyme is more prone to inactivation in the absence of substrate. Previous characterisations from this organism have shown that enzymes usually obtain an optimal temperature of around 55°C (Jung and Lee, 2005; Kim et al., 2012; Reher and Schönheit, 2006; Yasutake et al., 2007). In contrast to the moderately thermostable ManD, enzymes isolated from the same organism, like the glyceraldehyde dehydrogenase, showed nearly unaltered activity after incubation at 60°C over 2 h (Jung and Lee, 2006; Reher and Schönheit, 2006).

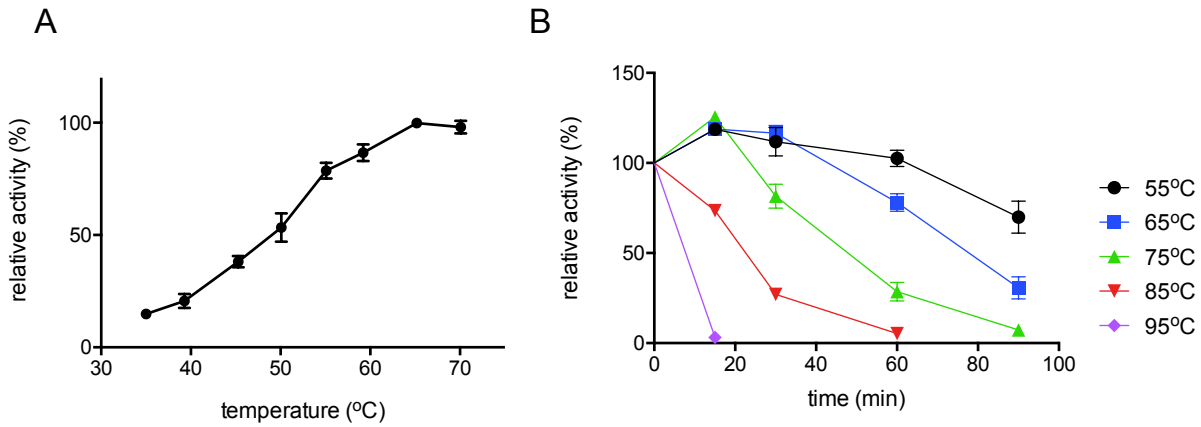


Figure 6 A) Effect of temperature on purified ManD enzyme activity. **B)** Thermostability of purified ManD enzyme incubated over 90 min at various temperatures. Values are mean of duplicate experiments. Error bars represent standard deviation of mean values.

Following the identification of the crucial metal cofactors required for ManD activity and determination of its optimal reaction temperature, we acquired enzyme kinetics data for ManD. Considering the fact that under physiological conditions an equilibrium of both lactone and free sugar acids exists, the kinetic data was acquired in reactions with

substrate provided in these two different forms. The substrates, were either hydrolysed with NaOH to obtain free sugar acid, D-mannonate, or the lactone was prepared in buffer, equalling an equilibrium of lactone and free sugar acid. ManD displayed a higher activity in reactions with D-mannonate compared to reactions with D-mannono-1,4-lactone prepared in buffer at pH 7 (Fig. 7A). The maximal velocity of the reaction, V_{max}, with D-mannonate as substrate was higher (2.39 ± 0.11 U/mg) compared to reactions with the lactone (1.90 ± 0.06 U/mg) prepared in buffer (p-value 0.06). The affinity of the enzyme did not differ significantly for the two forms of the substrate (K_m for mannonate: 5.37 ± 0.9 mM, K_m for lactone in buffer: 4.90 ± 0.5 mM).

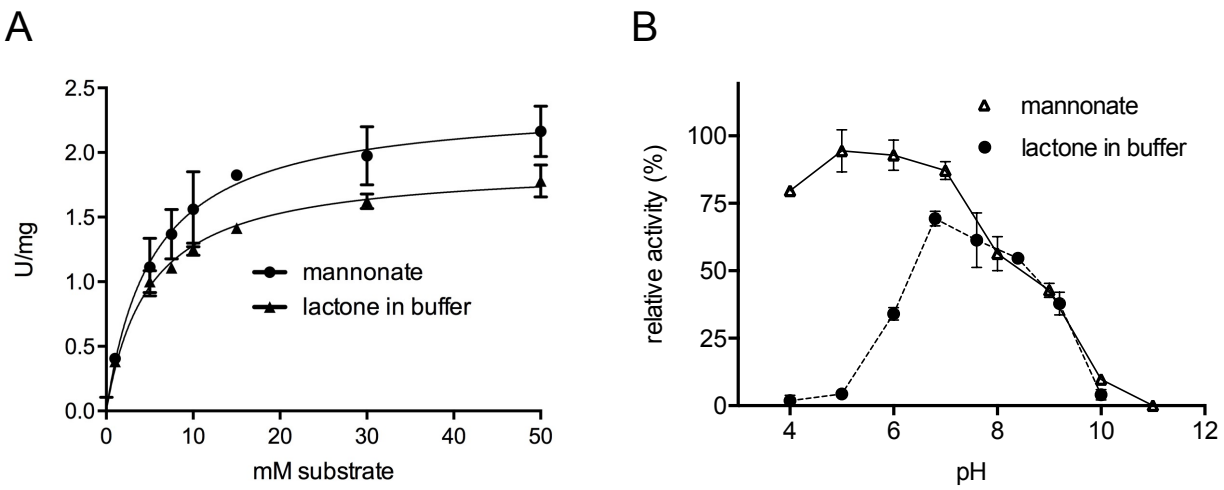


Figure 7 A) Comparison of enzyme kinetics of ManD with the substrate D-mannono-1,4-lactone prepared in buffer (lactone in buffer) and with D-mannonate (prepared by hydrolysis with NaOH). **B)** pH optima of reactions with the substrate in both forms. Activity values are expressed as relative activity (%) normalised to the highest overall activity observed in the assay. Values are mean of duplicate experiments. Error bars represent standard deviation of mean values. For V_{max} and K_m values with different substrate conformations a statistical analysis was perfromed using an unpaired t-test in Prism 6 (Graphpad software).

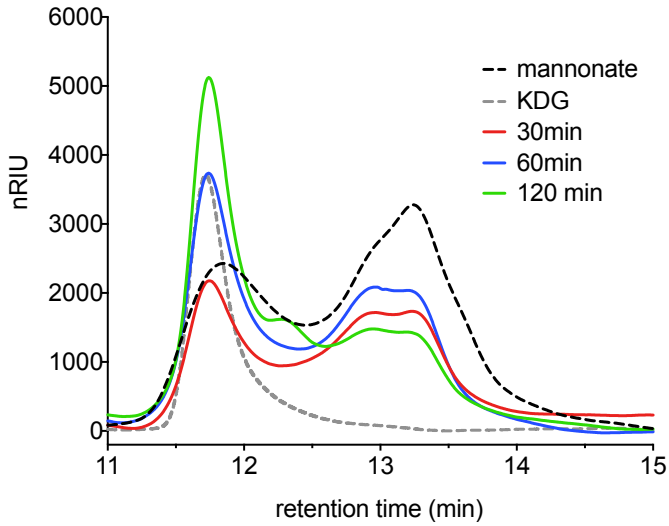
The pH optimum of ManD was dependent on the form of the substrate. ManD displayed highest activity between pH 5 and 7 with mannonate, while with D-mannono-1,4-lactone in buffer, the enzyme displayed overall lower activities with an apparent optimum at pH 7 (Fig. 7B). Unlike previously studied enzymes from *T. acidophilum* (e.g., AldT), ManD did not retain maximum activity above pH 7 with both substrates (Nishiya et al., 2014). Based on the observations from experiments with different forms of substrate, we assume that D-mannonate is either the only form or at least more accessible to ManD than its lactone. At low pH, ManD shows low activity with D-mannono-1,4-lactone, supposedly because the lactone does not hydrolyse to the free acid. In contrast, if D-mannono-1,4-lactone was hydrolysed to D-mannonate and used as substrate at low pH,

high enzyme activity should be observed. Similarly, the difference in V_{max} for D-mannonate and D-mannono-1,4-lactone in buffer can be explained by slow hydrolysis of the lactone to the free acid. Reactions with hydrolysed D-mannonate occur faster compared to the lactone in buffer, since more substrate is readily available.

3.5 Product identification and substrate specificity

In order to identify the product of the ManD reaction with mannonate, reactions containing the substrate and Co^{2+} were analysed at several time points by HPLC (Fig. 8A). Although a complete separation of the product KDG and mannonate could not be achieved using our organic acid column, an increase of KDG in measurements was observed. In order to test for substrate promiscuity, 10 different sugar acids (Fig. 8B) were incubated with the purified ManD enzyme and tested for the formation of 2-keto-3-deoxy analogues using the semicarbazide assay. D-mannonate and D-mannono-1,4-lactone were the only substrates that showed a positive reaction after incubation for 16 h (data not shown). This is in contrast to other dehydratases, including members of the enolase superfamily and the IlvD/EDD superfamily, such as the L-fuconate dehydratase from *Xanthomonas campestris* (Yew et al., 2006) or the dihydroxyacid dehydratase from *S. solfataricus* (Carsten et al., 2015; Kim and Lee, 2006), which has been shown to be active with a multitude of different sugar acids.

A



B

tested sugar acids	relative activity
D-mannonate	100%
D-mannono-1,4-lactone	79%
D-gluconic acid sodium salt	n.d.
D-galactono-1,4-lactone	n.d.
D-glucuronic acid	n.d.
D-galacturonic acid sodium salt	n.d.
L-gulonic acid-γ-lactone	n.d.
D-arabinoic acid sodium salt	n.d.
D-fuconic acid lithium salt	n.d.
L-fuconic acid sodium salt	n.d.
L-rhamnonic acid lithium salt	n.d.

Figure 8 A) Reactions with ManD and 10 mM mannonate analysed by HPLC equipped with a RID detector. Standards: 10 mM mannonate (dashed black line), 2.5 mM KDG (dashed grey line). Reactions: stopped after 30 min (red line), stopped after 60 min (green line), stopped after 120 min (blue line). nRIU: normalised refractive index units **B)** List of sugar acids that were tested as substrates for ManD. Activity is expressed relative (%) to the maximal activity obtained in the assay. n.d.: not detected. Lactones except for D-mannono-1,4-lactone were hydrolysed by

incubation in 1 M NaOH for 1 h at room temperature and then diluted in 50 mM HEPES pH 7 to obtain free sugar acids. All sugar acids were tested at a concentration of 10 mM in duplicate experiments.

4 Conclusion

In the present study, we characterised a functional mannonate dehydratase from the archaeon *T. acidophilum* (ManD) by recombinant expression in *E. coli*. The gene encoding for the mannonate dehydratase was found adjacent to a previously described aldohexose dehydrogenase (AldT) gene in the genome of *T. acidophilum* (Nishiya et al., 2014). Previously, it has been shown that the AldT from *T. acidophilum* selectively oxidises mannose, but the physiological function behind that oxidation has not been investigated (Fushinobu et al., 2011; Yasutake et al., 2007). Using NMR spectroscopy, we were able to show that mannonate and mannono-1,4-lactone is produced via oxidation of mannose by AldT. Kinetic assays confirmed that ManD was able to convert both the sugar acid and its lactone to the central intermediate KDG at neutral pH, without the help of a further lactonase. This resembles the second step of many oxidative pathways studied in archaea, including those for sugars like glucose, galactose, rhamnose, arabinose and xylose (Bräsen et al., 2014). Interestingly, ManD and mannonate dehydratase in *F. acidarmanus* and *F. acidophilum* share high amino acid sequence identity. Although the gene annotations for those mannonate dehydratases indicate a role in the hexuronate metbaolism they are all located adjacent to (putative) aldohexose dehydrogenases. It remains to be seen whether a non-phosphorylative pathway starting from mannose, similar to the np-ED pathway from glucose and galactose, exists in thermophilic archaea like *T. acidophilum*, *F. acidarmanus* and *F. acidophilum*.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

DK is supported by an international Macquarie University Research Excellence Scholarship (iMQRES). We thank the Australian Proteome Analysis Facility (Macquarie University, Sydney, Australia) for their support in performing LC ESI MS/MS experiments and Dr. Nicole Cordina from the Macquarie University NMR facility from the Department of Molecular Sciences for her help in performing NMR studies.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

References

Ashwell, G., 1962. [21] Enzymes of glucuronic and galacturonic acid metabolism in bacteria, in: Methods in Enzymology. Elsevier, pp. 190–208.

Atack, J.M., Srikhanta, Y.N., Fox, K.L., Jurcisek, J.A., Brockman, K.L., Clark, T.A., Boitano, M., Power, P.M., Jen, F.E.C., McEwan, A.G., Grimmond, S.M., Smith, A.L., Barenkamp, S.J., Korlach, J., Bakaletz, L.O., Jennings, M.P., 2015. A biphasic epigenetic switch controls immunoevasion, virulence and niche adaptation in non-typeable Haemophilus influenzae. Nat. Commun. 6, 1–12.

Bräsen, C., Esser, D., Rauch, B., Siebers, B., 2014. Carbohydrate metabolism in archaea: current insights into unusual enzymes and pathways and their regulation. Microbiol. Mol. Biol. Rev. 78, 89–175.

Britton, H.T.S., Robinson, R.A., 1931. CXCVIII.—Universal buffer solutions and the dissociation constant of veronal. J. Chem. Soc. 1456–1462.

Buchanan, C.L., Connaris, H., Danson, M.J., Reeve, C.D., Hough, D.W., 1999. An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. Biochem. J 343, 563–570.

Carsten, J.M., Schmidt, A., Sieber, V., 2015. Characterisation of recombinantly expressed dihydroxy-acid dehydratase from *Sulfobus solfataricus* - A key enzyme for the conversion of carbohydrates into chemicals. J. Biotechnol. 211, 31–41.

Darland, G., Brock, T.D., Samsonoff, W., Conti, S.F., 1970. A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile. Science (80-.). 170, 1416–1418.

Dreyer, J., 1987. The role of iron in the activation of mannonic and altronic acid hydratases, two Fe-requiring hydro-lyases. Eur. J. Biochem 166, 623–630.

Fushinobu, S., Nishimasu, H., Hattori, D., Song, H.-J., Wakagi, T., 2011. Structural basis for the bifunctionality of fructose-1,6-bisphosphate aldolase/phosphatase. Nature 478, 538–541.

Gerlt, J.A., Babbitt, P.C., Rayment, I., 2005. Divergent evolution in the enolase superfamily: The interplay of mechanism and specificity. Arch. Biochem. Biophys. 433, 59–70.

Hugouvieux-Cotte-Pattat, N., Robert-Baudouy, J., 1987. Hexuronate catabolism in *Erwinia Chrysanthemi*. J. Bacteriol. 169, 1223–1231.

Jung, J.H., Lee, S.B., 2006. Identification and characterisation of *Thermoplasma acidophilum* glyceraldehyde dehydrogenase: a new class of NADP+-specific aldehyde dehydrogenase. Biochem. J. 397, 131–138.

Jung, J.H., Lee, S.B., 2005. Identification and characterisation of *Thermoplasma acidophilum* 2-keto-3-deoxy-D-gluconate kinase: A new class of sugar kinases. Biotechnol. Bioprocess Eng. 10, 535–539.

Kim, S., Lee, S.B., 2006. Catalytic promiscuity in dihydroxy-acid dehydratase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. J. Biochem. 139, 591–596.

Kim, S., Lee, S.B., 2005. Identification and characterisation of *Sulfolobus solfataricus* D-gluconate dehydratase: a key enzyme in the non-phosphorylated Entner-Doudoroff pathway. Biochem. J. 387, 271–80.

Kim, S.M., Paek, K.H., Lee, S.B., 2012. Characterisation of NADP+-specific L-rhamnose dehydrogenase from the thermoacidophilic Archaeon *Thermoplasma acidophilum*. Extremophiles 16, 447–454.

Lamble, H.J., Milburn, C.C., Taylor, G.L., Hough, D.W., Danson, M.J., 2004. Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in *Sulfolobus solfataricus*. FEBS Lett. 576, 133–136.

Lawford, H.G., Rousseau, J.D., 1997. Fermentation of biomass-derived glucuronic acid by *pet* expressing recombinants of *E. coli* B. Appl. Biochem. Biotechnol. 63–65, 221–241.

Mandrand-Berthelot, M.-A., Condemine, G., Hugouvieux-Cotte-Pattat, N., 2004.

Catabolism of hexuronides, hexuronates, aldonates, and aldarates. EcoSal Plus 1, 1–21.

Mekjian, K.R., Bryan, E.M., Beall, B.W., Moran, C.P., 1999. Regulation of hexuronate utilization in *Bacillus subtilis*. J. Bacteriol. 181, 426–433.

Nishiya, Y., Tamura, N., Tamura, T., 2014. Analysis of bacterial glucose dehydrogenase homologs from thermoacidophilic archaeon *Thermoplasma acidophilum*: finding and characterisation of aldohexose dehydrogenase. Biosci. Biotechnol. Biochem. 68, 2451–2456.

Palleroni, N.J., Duodoroff, M., 1957. Metabolism of Carbohydrate by *Pseudomonas Saccharophila* III. Oxidation of Arabinose. J. Bacteriol. 74, 180–185.

Peekhaus, N., Conway, T., 1998. What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. J. Bacteriol. 180, 3495–3502.

Portalier, R., Robert Baudouy, J., Stoeber, F., 1980. Regulation of *Escherichia coli* K-12 hexuronate system genes: exu regulon. J. Bacteriol. 143, 1095–1107.

Qiu, X., Tao, Y., Zhu, Y., Yuan, Y., Zhang, Y., Liu, H., Gao, Y., Teng, M., Niu, L., 2012. Structural insights into decreased enzymatic activity induced by an insert sequence in mannonate dehydratase from Gram negative bacterium. J. Struct. Biol. 180, 327–334.

Rakus, J.F., Fedorov, A.A., Fedorov, E. V, Glasner, M.E., Vick, J.E., Babbitt, P.C., Almo, S.C., Gerlt, J.A., 2007. Evolution of enzymatic activities in the enolase superfamily: D-mannonate dehydratase from *Novosphingobium aromaticivorans*. Biochemistry 46, 12896–12908.

Reher, M., Schönheit, P., 2006. Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner-Doudoroff pathway, constitute a novel enzyme family within the aldehyde d. FEBS Lett. 580, 1198–1204.

Robert-Baudouy, J., Jimeno-Abendano, J., Stoeber, F., 1982. D-Mannonate and D-altronate dehydratases of *Escherichia coli* K12. Methods Enzymol. 90, 288–294.

Robert-baudouy, J.M., Jimeno-Abendano, J., Stoeber, F.R., 1975. Individualité des hydrolyases mannonique et altronique chez *E. coli* K-12. Biochimie 57, 1–8.

Robert-Baudouy, J.M., Stoeber, F.R., 1973. Purification et propriétés de la D-mannonate hydrolyase d'*Escherichia coli*. Biochim. Biophys. Acta -Enzymology 309, 473–485.

Sambrook, J., Russell, D.W., Russell, D.W., 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.

Shulami, S., Gat, O., Sonenshein, A.L., Shoham, Y., 1999. The glucuronic acid utilization gene cluster from *Bacillus stearothermophilus* T-6. J. Bacteriol. 181, 3695–3704.

Skoza, L., Mohos, S., 1976. Stable thiobarbituric acid chromophore with dimethyl sulphoxide. Application to sialic acid assay in analytical de-O-acetylation. Biochem. J. 159, 457–62.

Stephens, C., Christen, B., Fuchs, T., Sundaram, V., Watanabe, K., Jenal, U., 2007. Genetic analysis of a novel pathway for D-xylose metabolism in *Caulobacter crescentus*. J. Bacteriol. 189, 2181–2185.

Sun, N., Pan, C., Nickell, S., Mann, M., Baumeister, W., Nagy, I., 2010. Quantitative proteome and transcriptome analysis of the archaeon *Thermoplasma acidophilum* cultured under aerobic and anaerobic conditions. J. Proteome Res. 9, 4839–4850.

Wałaszek, Z., Horton, D., 1982. Conformational studies on aldonolactones by NMR spectroscopy. Conformations of d-glucono-, D-mannono, D-gulono-, D-galactono-1,4-lactone in solution. Carbohydr. Res. 105, 131–143.

Watanabe, S., Saimura, M., Makino, K., 2008. Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. J.

Biol. Chem. 283, 20372–20382.

Wichelecki, D.J., Balthazor, B.M., Chau, A.C., Vetting, M.W., Fedorov, A.A., Fedorov, E. V., Lukk, T., Patskovsky, Y. V., Stead, M.B., Hillerich, B.S., Seidel, R.D., Almo, S.C., Gerlt, J.A., 2014. Discovery of function in the enolase superfamily: D-mannonate and D-gluconate dehydratases in the D-mannonate dehydratase subgroup. Biochemistry 53, 2722–2731.

Wichelecki, D.J., Alyxa, J., Vendiola, F., Jones, A.M., Al-obaidi, N., Almo, S.C., Gerlt, J.A., 2014. Investigating the physiological roles of low-efficiency D-mannonate and D-gluconate dehydratases in the enolase superfamily: pathways for the catabolism of L-gulonate and L-idonate. Biochemistry 53, 5692–5699.

Yasutake, Y., Nishiya, Y., Tamura, N., Tamura, T., 2007. Structural Insights into unique substrate selectivity of *Thermoplasma acidophilum* D-aldohexose dehydrogenase. J. Mol. Biol. 367, 1034–1046.

Yew, W.S., Fedorov, A.A., Fedorov, E. V., Rakus, J.F., Pierce, R.W., Almo, S.C., Gerlt, J.A., 2006. Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from *Xanthomonas campestris*. Biochemistry 45, 14582–14597.

Zhang, Q., Gao, F., Peng, H., Cheng, H., Liu, Y., Tang, J., Thompson, J., Wei, G., Zhang, J., Du, Y., Yan, J., Gao, G.F., 2009. Crystal structures of *Streptococcus suis* mannonate dehydratase (ManD) and its complex with substrate: Genetic and biochemical evidence for a catalytic mechanism. J. Bacteriol. 191, 5832–5837.

Supplementary information

Characterisation of a mannonate dehydratase from *Thermoplasma acidophilum* and its potential role in the catabolism of D-mannose

Dominik Kopp, Robert D. Willows and Anwar Sunna

[a]Department of Molecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia

[b]Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, New South Wales 2109, Australia

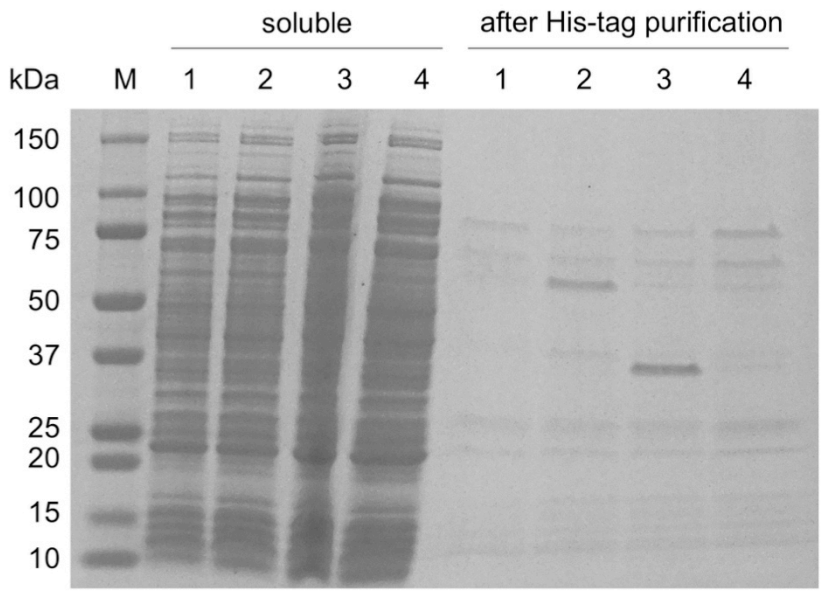


Figure S1 SDS-PAGE showing soluble expression and purification of gene products potentially encoding for a mannonate dehydratases in *T. acidophilum*. 1: Ta0085m 2: Ta0752 3: Ta0753 4: Ta0756. Expected molecular masses (kDa) for 1, 2, 3, 4 are 44.8, 43.9, 42 and 47.6. M: Bio-Rad Precision Plus Protein Standard.

Chapter 3

Cell-free enzymatic conversion of spent coffee grounds into the platform chemical lactic acid

3.1 Introduction

The previous chapter described the finding of a mannonate dehydratase and its potential role in an *in vivo* metabolism. The following chapter demonstrates how this enzyme is used to construct a cell-free synthetic pathway to convert mannose, the most abundant sugar in spent coffee grounds, into L-lactic acid. The synthetic pathway assembled and presented in this work represents a proof-of-concept and has not been optimised yet for industrial yields and titers at a large scale. The production of L-lactic acid is usually achieved with acid tolerant *S. cerevisiae* strains or *Lactobacilli* by fermentation of carbohydrate-rich crops. More sustainable substrates such as waste streams have the potential to lower the price of bio-production of chemicals while achieving a better economical footprint (see contribution to Appendix I, section 5). Spent coffee grounds are an excellent example, as they are an industrially underutilised waste substrate, which is highly abundant and rich in carbohydrates. The construction of cell-free biocatalysis enables a rapid prototyping of synthetic pathways by circumventing numerous obstacles encountered in *in vivo* processes. In the following work, the concept of cell-free biocatalysis is applied to test the hypothetical oxidative mannose pathway and to demonstrate the conversion of waste substrates into platform chemicals.

This chapter has been prepared as a manuscript for submission to *Metabolic Engineering*.

3.2 Contribution to manuscript 3

The concept of this publication was developed together with my main supervisor Anwar Sunna. The experiments were designed, performed by me. Data analysis was conducted by me with help of Robert Willows. Advice on the experimental design was given by my supervisors Anwar Sunna and Robert Willows. The manuscript was written by me and was reviewed by all authors. The contributions of each author are given in Table 3.1

Table 3.1 Author contributions for manuscript 3

	Dominik Kopp	Robert Willows	Anwar Sunna
Conception	●		●
Experiment design	●	●	●
Data Collection	●		
Data Analysis	●	●	
Manuscript	●	●	●

3.3 Manuscript 3

Cell-free enzymatic conversion of spent coffee grounds into the platform chemical lactic acid

Dominik Kopp[a], Robert D. Willows[a,b] and Anwar Sunna[a,b]*

[a]Department of Molecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia

[b]Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, New South Wales 2109, Australia

*Corresponding author

Anwar Sunna

Department of Molecular Sciences

Macquarie University

North Ryde, NSW 2109 Australia

Phone: +61 2 9850 4220

Fax: +61 2 9850 8313

Email: anwar.sunna@mq.edu.au

Abstract

Spent coffee grounds (SCG) are an underutilised waste resource rich in the polysaccharide galactomannan. Here, we use a cell-free approach for the construction of a novel enzymatic pathway for the conversion of mannose, the most abundant sugar in SCG, into L-lactic acid. The pathway is based on a putative oxidative mannose metabolism from *Thermoplasma acidophilum* and uses a previously characterised mannonate dehydratase to convert mannose into lactic acid via 4 enzymatic reactions. The cofactor-balanced pathway consists of an aldohexose dehydrogenase, a mannonate dehydratase, a 2-keto-3-deoxygluconate aldolase and a lactate dehydrogenase, and uses NADH as a single cofactor. Single enzymes were first studied under reaction conditions and then tested in sequential and one-pot pathway reactions. Adjustment of enzyme loadings, substrate and cofactor concentrations, and buffer ionic strength allowed the identification of crucial reaction parameters and bottlenecks. Moreover, reactions with isotope labelled mannose enabled the monitoring of pathway intermediates and revealed a reverse flux in the conversion process. Finally, 4.4 ± 0.1 mM lactic acid was produced from 14.57 ± 0.7 mM SCG-derived mannose. The conversion efficiency of the pathway could be improved in the future by addition of a glyceraldehyde recycling process and enzyme engineering.

Keywords

enzyme biocatalysis, mannose, cell-free, spent coffee grounds, lactic acid, NMR flux analysis

1 Introduction

The biotechnological production of versatile chemicals in a sustainable way depends on the availability on low-cost biomass and production systems capable to compete with established base chemicals derived from fossil fuels. Lactic acid (LA) is a versatile chemical with a wide range of different applications in medical, textile and food industries and has received most of its attention for the production of the biodegradable polymer polylactic acid (PLA) (Castro-Aguirre et al., 2016; Dusselier et al., 2013). The biotechnological production of LA via fermentation of sugars by yeast is favoured over chemical synthesis, since an optically pure product can be obtained. Mainly glucose obtained from carbohydrate-rich feedstock such as sugar cane, sugar beet, corn or sweet potato is used in fermentative processes (Vink et al., 2003). However, high production yields by conventionally used lactic acid bacteria is limited by several factors including high nutrient requirements and low pH stability of production strains. Accordingly, this has led to a growing interest in using metabolically engineered strains with improved properties for LA production (Chen et al., 2011; Sauer et al., 2010). Despite the great development of genetic tools for metabolic engineering *in vivo*, some physiological challenges and limitation in yields remain. As an alternative, cell-free biocatalysis has the potential to overcome a variety of challenges encountered during *in vivo* metabolic engineering (Dudley et al., 2015; Morgado et al., 2016, Petroll et al. 2018). In cell-free biocatalysis, purified enzymes or cell lysates are assembled into synthetic pathways, which can be rapidly tested and easily adjusted. Common obstacles of *in vivo* biosynthesis of chemicals like product toxicity, limited cellular uptake and secretion or genetic compatibility with the production host can be overcome in a cell-free environment. To date, this cell-free approach has been successfully applied to produce a range of different bulk and speciality chemicals including iso- and n-butanol, ethanol, hydrogen, α−ketoglutarate, mevalonate and dihydroacetone phosphate (Beer et al., 2017; Dudley et al., 2016; Guterl et al., 2012; Hold et al., 2016; Ye et al., 2009). Besides the immense diversity in enzyme combinations, mostly refined substrates like glucose have been used as starting substrates for these pathways. The drive towards a lower ecological footprint, further reduction of costs and avoidance of competition with food resources increased the efforts towards non-edible or wasted biomass. More recently, alternative feedstocks to hydrolysates from lignocellulosic biomass, such as chitin, corn steep water or glycerol have been implemented in cell-free bioconversions for the production of chemicals and fuels (Gao et al., 2015; Honda et al., 2017; Li et al., 2017).

Spent coffee grounds (SCG) represent an abundant and so far industrially underexploited waste resource. Yearly, more than 6 million tons of SCG are generated worldwide, most of which end up in landfills (Tokimoto et al., 2005). Approaches to reutilise some of this waste have been undertaken to produce drop-in biofuels via oil extraction (Caetano et al., 2012; Kondamudi et al., 2008), as a source for bioactive compounds and antioxidants (Ramalakshmi et al., 2009), for the extraction of immunostimulatory oligosacharides (Takao et al., 2006) and the extraction of phenolic compounds (Zuorro and Lavecchia, 2012). Various fungal strains have been cultivated on SCG for the production of ethanol and the extraction of polyphenolic compounds (Machado et al., 2012; Mussatto et al., 2012). Several different microbial species have been employed for the production of polyhydroxyalkanoates and carotenoids from oil and sugars in SCG (Obruca et al., 2015). Recently, acid-hydrolysed SCG was used to cultivate several strains of lactic acid bacteria. However, fermentation with microbial strains, can be impaired by toxic compounds such as polyphenols, tannins and caffeine present in SCG and results in low yields. Despite the fact that on average 50% of SCG is composed of carbohydrates, only little research has been undertaken to exploit this component of SCG (Campos-Vega et al., 2015). Several carbohydrate analysis have shown that mannose is the most abundant sugar present in SCG (Ballesteros et al., 2014; Jooste et al., 2013; Mussatto et al., 2011). Mannose can be used by many organisms as a growth substrate. For example, mannose catabolism in industrial production strains like *Saccharomyces cerevisiae* follows the classical Embden-Meyerhof-Parnas (EMP) pathway and comprises at least 10 enzymatic reactions, 2 different cofactors, heat labile triosephosphates and unfavourable isomerisation reactions (Fig. 1B). In contrast, non-phosphorylative or oxidative pathways such as the non-phosphorylative Entner-Doudoroff (np-ED) pathway do not use ATP, rely on a lower enzyme cost and display a more thermodynamically driven design which result in a high flux compared to the EMP pathway (Flamholz et al., 2013; Meléndez-Hevia et al., 1997). Despite the discovery of oxidative pathways for many sugars, such as glucose, galactose, xylose, L-rhamnose and L-arabinose, a mannose-specific pathway has not been described yet.

Here, we report the design of a functional non-phosphorylative synthetic pathway to convert mannose obtained from SCG into the platform chemical LA (Fig. 1A). Key enzymes of the pathway include a newly identified mannonate dehydratase and a 2-keto-3-deoxygluconate (KDG) aldolase from the thermoacidophilic archaeon

Thermoplasma acidophilum. In addition we demonstrate that a so far unknown, non-phosphorylative catabolism of mannose is possible in the archaeon *T. acidophilum*.

2 Material and Methods

2.1 Reagents and Chemicals

All chemicals were purchased from Sigma Aldrich (St. Louis, USA) unless otherwise stated. D-[1,6-$^{13}C_2$]mannose was obtained from Omicron Biochemicals, Inc. (South Bend, USA), D-mannono-1,4-lactone was obtained from TCI Chemicals (Tokyo, Japan) and was either prepared in water or hydrolysed according to Lamble et al. (Lamble et al., 2004) to obtain the free sugar acid D-mannonate. Briefly, for hydrolysis, 1 M stock solutions of D-mannono-1,4-lactone were prepared in 1 M NaOH and incubated at room temperature for 1 h before dilutions were prepared in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid buffer (HEPES) pH 7.

2.2 Bacterial strains and plasmids

All DNA preparations, manipulations and digestions were carried out according to Sambrook (Sambrook et al., 2001). DNA fragments encoding the different proteins were amplified by polymerase chain reaction (PCR) using the primer pairs described in Table S1. The *Escherichia coli* strains used in this study were *E. coli* α-select cells (Bioline, Alexandria, AUS), *E. coli* BL21(DE3) cells (NEB, Ipswitch, USA) and *E. coli* BL21 CodonPlus-(DE3)RIL (Agilent Technologies, Santa Clara, USA). The plasmids used were pETDuet-1 (Novagen, Madison, USA) and pProEX Htc (Invitrogen, Carlsbad, USA). All *E. coli* strains were grown at 37°C in Luria-Bertani medium (LB) supplemented with 100 µg/ml carbenicillin except for strain BL21-CodonPlus (DE3)-RIL which was supplemented with 100 µg/ml carbenicillin and 50 µg/ml chloramphenicol (final concentrations). To generate the expression plasmids pETDuet-1-AldT (encoding the aldohexose dehydrogenase, AldT) and pProEx Htc-Ta1157 (encoding the 2-keto-3-deoxygluconate aldolase, KDGA), the genes Ta0754 and Ta1157 were amplified directly from *Thermoplasma acidophilum* DSM 1728 (DSMZ, Braunschweig, Germany) genomic DNA. The amplified Ta0754 DNA fragment was digested with BamHI/EcoRI and ligated into similarly cut pETDuet-1 vector. The Ta1157 DNA fragment was digested with EcoRI/HindIII and ligated into similarly cut pProEX Htc vector. To generate the expression plasmid pETDuet-1-BsLDH (encoding the lactate

dehydrogenase from *Bacillus stearothermophilus,* LDH*),* the *ldh* gene sequence was synthesised as a gBlock from IDT DNA Technologies (Singapore). The DNA sequence was amplified by PCR, digested with BamHI/HindIII and ligated into similarly cut pETDuet-1 vector. The expression plasmids were used to transform chemically competent *E. coli* α-select (Bioline).

The expression plasmid for ManD, pProEx Hta-Ta0753 was constructed previously and is described elsewhere (Kopp et al. 2018, submitted).

2.3 Production and purification of recombinant proteins

Expression of AldT and LDH was performed in *E. coli* BL21 (DE3) cells, while ManD and KDGA were expressed in BL21-CodonPlus (DE3)-RIL cells. All strains were grown at 37°C to an OD_{600} of 0.4 - 0.7, before expression was induced by addition of 0.4 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). All expression were performed at 20°C for at least 18 h except for the expression of LDH, which was performed at 37°C for 4 h. Cells were harvested by centrifugation at 5000 x *g* for 20 min at 4 ^{o}C. Harvested cells were resuspended in 50 mM HEPES buffer (pH 7) and ruptured by three passages through a French pressure cell. Soluble fraction of each lysate was obtained by centrifugation at 15000 x *g* for 20 min at 4°C. The final soluble fractions containing the His-tagged proteins were loaded onto a 5 ml Ni-NTA affinity column (GE Healthcare) previously equilibrated with buffer containing 50 mM NaH_2PO_4, 300 mM NaCl and 20 mM imidazole, pH 8. The column was washed extensively with the same buffer containing 40 mM imidazole before enzymes were eluted isocratically with a final concentration of 400 mM imidazole. Fractions containing the desired enzyme were selected and pooled based on the purification chromatogram acquired at 280 nm. The buffer of eluted enzyme solutions was exchanged to 50 mM HEPES pH 7 using an Amicon Ultra centrifugal filter unit (30-kDa cutoff, Millipore). Purity and size of the enzymes were analysed by sodium dodecyl sulfate-polyacrylamide gelectrophoresis (SDS-PAGE) followed by Coomassie staining. Protein concentrations were determined using a Pierce BCA Protein assay Kit (Thermo Fisher Scientific, Waltham, USA). All purified enzyme preparations were stored at -80°C.

2.4 Enzyme assays

All photometrical assays were performed in a Jasco V-760 UV/Vis spectrophotometer equipped with a heating device PAC-743R set to 50°C. For assays measuring NADH changes, the enzyme activity (AldT and LDH) was calculated using a molar extinction coefficient of ε_{340} = 6.22 mM^{-1} cm^{-1} (Haid et al. 1975). One unit of enzyme activity was defined as the amount of enzyme necessary to convert 1 µmol substrate per minute. For all enzyme assays, reactions were performed in triplicates and all data were plotted and analysed using Prism 6.0 (GraphPad software, La Jolla, CA).

The activity of AldT was assayed photometrically by measuring the increase in NADH at 340 nm. Reactions (800 µl) contained 1 - 200 mM D-mannose and 5 mM NAD^+ in 20 mM sodium phosphate (NaP) buffer pH 7. Reaction mixtures were preheated for 5 min before enzyme was added to start the reaction. KDGA activity was assayed in the direction of aldol addition. Reactions (100 µl) were performed in 0.2 ml PCR tubes in a waterbath at 50°C and contained 0.5 - 40 mM pyruvate and D-glyceraldehyde in 20 mM NaP buffer pH 7. Reactions were incubated for 15 min before 10 µl of 12.5% (w/v) tricholoroacetic acid (TCA) was added to stop the reaction. Formation of 2-keto-deoxy sugar acid was determined by a modified thiobarbituric acid (TBA) assay (Buchanan et al., 1999; Skoza and Mohos, 1976). 50 µl of each reaction was oxidised by addition of 125 µl 25 mM periodic acid in 0.25 M H_2SO_4 and incubated at room temperature for 20 min. The oxidation was stopped by addition of 250 µl of 2% (w/v) sodium arsenite in 0.5 M HCl followed by addition of 1 ml of 0.3% TBA before samples were boiled for 10 min. The formation of 2-keto-3-deoxy sugar acid was measured by reading the absorbance at 549 nm in a Spectrostar Nano platereader (BMG Labtech, Ortenberg, Germany) and quantification using a molar extinction coefficient of 67.8 mM^{-1} cm^{-1}. The activity of LDH was assayed photometrically by measuring the decrease in NADH at 340 nm. For the determination of K_m and V_{max} values, reactions (500µl) contained pyruvate (0 - 50 mM) and 0.2 mM NADH in 20 mM NaP buffer pH 7. Reaction mixtures were preheated at 50°C for 200 s before reactions were started by the addition of NADH.

2.5 Sequential reactions

Sequential reactions were performed with a final concentration of 20 mM mannose, 20 mM NAD^+, 0.1 mM $CoSO_4$ in 50 mM HEPES pH 7. Reactions were started with the addition of AldT (1.44 mU). ManD (0.063 mU) was added after 1.5 h, KDGA (4.59 mU) after 16.5 h and LDH (52.2 mU) with 10 mM NADH were added after 17.5 h. The

complete reaction (AldT+ManD+KDGA+LDH) was terminated at 18.5 h. 60µl of each reaction was stopped with 6 µl 12.5% (w/v) TCA and then subjected to HPLC analysis.

2.6 One-pot reactions

Reactions with varying enzyme units were performed in 100 µl volume. Each reaction contained all 4 enyzmes (AldT, ManD, KDGA add LDH), 5 mM mannose, 5 mM NAD^+ and 0.1 mM $CoSO_4$ in 100 mM NaP buffer pH 7. Reactions with varying mannose and NAD^+ concentrations contained 0.2 U/ml AldT, 0.1 U/ml ManD, 0.4 U/ml KDGA, 11.6 U/ml LDH and 0.1 mM $CoSO_4$ in 100 mM NaP buffer pH 7. One-pot reactions were incubated at 50°C for 18 h on a Thermomixer C without shaking. After incubation, 35 µl of 1.9 M $HClO_4$ was added to 65 µl of each sample to stop the reaction and precipitate proteins. This was followed by addition of 13 µl 5 M KOH to neutralise the reaction before samples were centrifuged and subjected to HPLC analysis.

2.7 Thermostability of enzymes

All enzymes were incubated at 50°C for 1 h in the absence of substrate but at the same loading (U/ml) as used in the one-pot reactions (0.2 U/ml AldT, 0.1 U/ml ManD, 0.4 U/ml KDGA, 11.6 U/ml LDH). Samples were removed at different times and the residual enzyme activity was measured by spectrophotometric, colorimetric or HPLC analysis. All reactions were performed at least in duplicate.

The residual activity of AldT was assayed in reactions containing 10 mM mannose and NAD^+ in 20 mM NaP buffer pH 7. Reactions were stopped by the addition of 10 µl of 12.5% TCA. Samples were then subjected to HPLC.

The residual activity of KDGA was tested in the aldol cleavage direction, by quantification of pyruvate and glyceraldehyde. Reactions (100 µl) contained 5 mM 2-keto-3-deoxygluconate (KDG) in 20 mM NaP buffer pH 7. Reactions were stopped by the addition of 10 µl of 12.5% TCA. Pyruvate formation was measured by subjecting 10 µl of each sample to HPLC analysis.

The residual activity of ManD was assayed in 60 µl reactions containing 10 mM hydrolysed mannonate and 1 mM $CoSO_4$ in 20 mM NaP buffer pH 7. The reactions were stopped by addition of 6 µl 12.5% TCA. 50 µl of each sample were analysed by the TBA method as described under the KDGA activity assay.

Remaining activity of LDH was assayed in reactions (100 µl) with 5 mM pyruvate and NADH. Reactions were stopped by the addition of 10 µl 12.5% TCA to the reaction. LA

formation was measured by subjecting 10 µl of each sample to HPLC analysis. For determination of half-lives of AldT and ManD, data points were fitted with Prism 6 (GraphPad software) to a non-linear regression with a plateau constrain equal to 0 (Fig. S3)

2.8 Hydrolysis of spent coffee grounds

SCG was obtained from a university coffee shop and was dried at 60°C for one week before storage in an air-tight container at 4°C. SCG monosaccharides were extracted by dilute acid hydrolysis according to the standard National Renewable Energy Laboratory (NREL) laboratory analytical procedure described by Sluiter et al. (Sluiter et al., 2011). A set of standard reducing sugars (SRS) consisting of 30 mg of D-xylose, D-glucose, D-galactose, L-arabinose and D-mannose was subjected to the same hydrolysis procedure as SCG in order to correct for any losses of monosaccharides during the procedure. In this process 1.5 ml of 72% (v/v) H_2SO_4 was added to 150 mg SCG and incubated at 50°C for 1 h under constant stirring in 100 ml pressure tubes. The solution was then diluted to 4% H_2SO_4 by addition of ultrapure water and autoclaved for 45 min at 121°C. 25 ml of the solution was neutralised by addition of $CaCO_3$ until pH 6 was reached and filtered through 0.22 µm sterile filter. The final monosaccharide content was analysed by HPLC.

2.9 HPLC Analysis

LA yields were determined using an Agilent 1290 HPLC System (Agilent Technologies, Santa Clara, USA) connected to a diode array detector (DAD) (G4212A) and a refractive index detector (RID) G1362A (Agilent Technologies) heated to 55°C. Chromatographic separation of LA was achieved on a Bio-Rad Aminex HPX-87H organic acid column at 65°C with 5 mM H_2SO_4 used as a mobile phase. Standards of L-lactic acid treated equal to the reactions were used for quantification (Fig. 6A). The flow rate was adjusted to 0.6 ml/min. Unless specified otherwise, 10 µl sample was injected for each analysis.

Sequential pathway reactions, separation of pathway intermediates, residual KDGA activity and residual LDH activity were analysed on the HPLC systems equipped with an Agilent Hi-Plex H^+ organic acid column connected to the RID. Standard curves for pyruvate and L-lactic acid were used to determine product formation for residual KDGA

and LDH activity, respectively. 10 mM H_2SO4 was used as a mobile phase at a flow rate of 0.6 ml/min. The column temperature was set to 80°C and the RID heated to 55°C.

Analysis of monosaccharide obtained from SCG dilute acid hydrolysis, residual activity of AldT and quantification of remaining mannose in time course reactions was achieved on the HPLC system equipped with a Hi-Plex Pb column (Agilent) connected to a RID. Ultrapure water was used as a mobile phase at a flow rate of 0.6 ml/min. The column temperature was set to 85°C and the RID heated to 55°C. The mannose concentration in the hydrolysate was calculated using a standard curve prepared with D-mannose standards (Fig. S6B) and corrected for changes observed with the SRS.

2.10 NMR analysis

Nuclear magnetic resonance spectroscopy (NMR) measurements of one-pot reactions were carried out at 50°C (323K) on a 500 MHz Bruker Avance III HD NMR equipped with a BBFO probe. 20% D_2O and 0.4% 10 mM 3-(trimethylsilyl)propionic-2,2,3,3-d4 acid (TMSP) chemical shift standard were added to an enzyme mixture containing 0.2 U/ml AldT, 0.1 U/ml ManD, 0.4 U/ml KDGA, 11.6 U/ml LDH and a substrate solution containing 10 mM D-[1,6-$^{13}C_2$]mannose, 10 mM NAD^+ and 100 mM NaP buffer pH 7.5. Time point zero of the time course reaction was acquired with 500 µl of substrate solution in the absence of enzymes. One-pot reactions were started by combining enzyme mixture and substrate solution. Subsequently reactions were transferred to 5 mm NMR tubes, and 1D power-gated ^{1}H decoupled ^{13}C NMR spectra were acquired with 16 scans at a pulse angle of 90° with a relaxation delay of 3s between scans.
Standards for mannonate and mannono-1,4-lactone were prepared by dissolving D-mannono-1,4-lactone in 1 M NaOH or 1 M HCl with incubation for 1 h which was followed by dilution in water to 100 mM. Then, 300 µl of each 100 mM solution (equivalent to 5 mg sugar acid) was freeze-dried and resuspended in 500 µl 100 mM NaP buffer pH 7. After transfer of each sample to 5 mm NMR tubes, 10% D_2O and 40 µM TMSP chemical shift standard was added. 5 mg of D-glyceraldehyde, pyruvate and L-lactic acid and 1 mg of KDG were each resuspended in 100 mM NaP buffer pH 7.5 to obtain 1D decoupled ^{13}C NMR spectra of non-isotopically labelled samples for comparison to reaction intermediates. Spectra of intermediates were acquired using power-gated proton decoupling at a 90° pulse with 512 scans and a 3 s relaxation delay

between scans. Since the chemical shift changed with pH, L- lactic acid standard (5 mg each) had to buffered to reach pH 7.5.

In the labelled time course reaction, C6 compounds (mannose, mannono-1,4-lactone, mannonate and KDG) contained two labelled carbons (carbon 1 and carbon 6). Due to a combination of proton nuclear Overhauser effect (NOE) and shorter relaxtion times the carbons with lower chemical shift have higher signal intensities in the decoupled ^{13}C NMR. Accordingly, only lower chemical shifts were used for integration (Table S2). Visualisation of all spectra was performed with iNMR 6.0. After Fourier transformation with an exponential weighting factor of 1.5 was performed followed by and automatic phase correction and a smoothing factor of 10 were applied to all spectra.

3 Results and Discussion

3.1 Pathway design and enzymes selection

We designed a cell-free enzymatic pathway for the bioconversion of LA from mannose (Fig. 1A). The first 3 reactions of the assembled pathway are catalysed by enzymes from the archaeon *T. acidophilum.* In the first reaction step of the pathway, mannose is oxidised via a previously described aldohexose dehydrogenase (AldT) to mannonolactone using NAD^+ (Nishiya et al., 2014). Mannonolactone hydrolyses spontaneously at neutral pH to mannonate, which is then converted by the mannonate dehydratase (ManD) to 2-keto-3-deoxygluconate (KDG). A KDG aldolase (KDGA) subsequently cleaves KDG into pyruvate and D-glyceraldehyde. Pyruvate is then oxidised by a L-lactate dehydrogenase (LDH) from *B. stearothermophilus* with the use of NADH produced in the first oxidation reaction of the pathway. The turnover of NADH is balanced since one mole NADH is produced per mole mannose being oxidised, and one mole NADH is oxidised per mole pyruvate being reduced in the last reaction. Since KDGA converts KDG to equal ratios (50% each) of pyruvate and D-glyceraldehyde, and only pyruvate is converted to LA, the current design of the pathway only allows for 50% of the carbon in the mannose to be converted into LA.

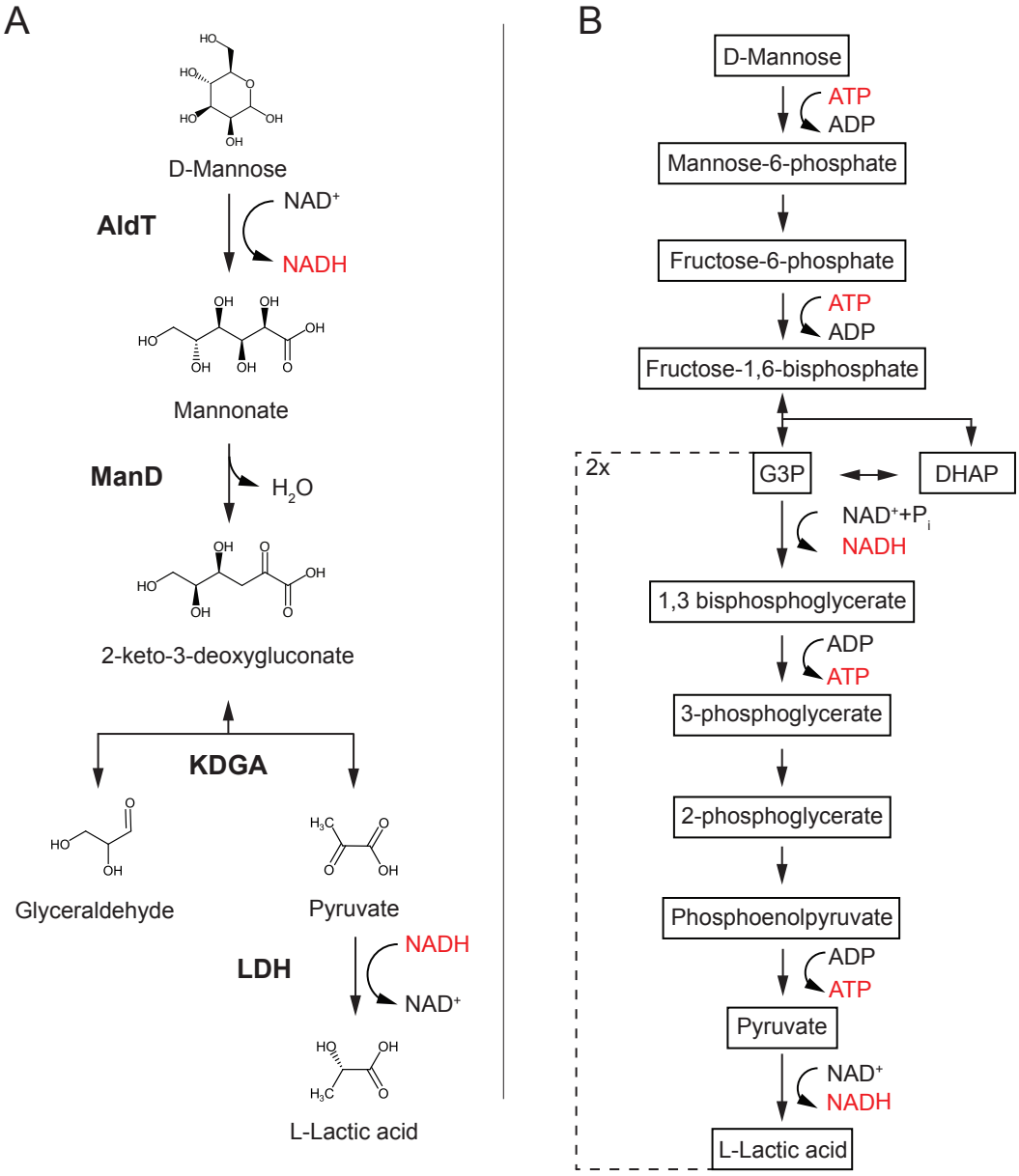


Figure 1 Synthetic pathway for the cell-free conversion of mannose into LA (A) compared to EMP-based pathway found in most production hosts such as *S. cerevisiae* (B). Enzymes are in bold, reduced cofactor in red. AldT: aldohexose dehydrogenase, ManD: mannonate dehydratase, KDGA: 2-keto-3-deoxygluconate aldolase, G3P: glyceraldehyde-3-phosphate, DHAP: dihydroacetone phosphate.

The enzymes used in this pathway were selected according to their temperature and pH compatibility. AldT has been described previously to be stable above pH 7 with an optimum activity at pH 10, while retaining full activity after at least 15 min incubation at 60°C (Nishiya et al., 2014). Recently, a putative mannonate dehydratase from *T. acidophilum* (ManD) was recombinantly expressed in *E. coli,* purified and fully characterised (Kopp et al. 2018, submitted). ManD shows specific activity towards mannonate and mannono-1,4-lactone at a pH 7 with a strong decrease in activity above

neutral pH. The enzyme showed maximal activity at 65°C and remained stable for at least 1 h at 55°C.

A KDG aldolase in *T. acidophilum* has not been identified before, but indication for the corresponding genes were given based on homology to a KDG aldolase from *P. torridus* (Reher et al., 2010). Accordingly, the gene locus Ta1157 was recombinantly-produced in *E. coli* and the purified enzyme was shown to successfully cleave KDG to pyruvate and glyceraldehyde. Since this enzyme is also derived from the same organism as AldT and ManD, we assumed a similar thermostability and pH range for activity. In the final step of the reaction cascade, the NADH-dependent L-lactate dehydrogenase (LDH) from *B. stearothermophilus* was chosen because of its thermostability (fully active after 20 min at 70°C) at a neutral pH (Murphey et al., 1967). This LDH has been successfully expressed in *E. coli* and the purified enzyme displayed high activity (241 +/- 30 U/mg) with a reasonably low K_m for pyruvate (1.48 ± 0.05 mM) (Halliwell et al., 2001). All single enzyme experiments and the final pathway assemblies were performed with purified enzymes in order to exclude potential side reactions (e.g. cofactor consuming reactions) present in *E. coli* expression lysates (Fig. S1).

3.2 Single enzymes and reaction conditions

The kinetic data available for AldT and LDH were acquired previously at 25°C and in different buffer systems (Halliwell et al., 2001; Nishiya et al., 2014). However, the cell-free pathway presented here is intended to function at 50°C. We chose NaP buffer at neutral pH during the selected operating temperature because it faces low temperature dependency and does not interfere with ^{13}C NMR. Accordingly, in order to normalise the kinetic data for each enzyme of the pathway, all kinetic assays were performed at 50°C in the same buffer system and a pH of 7 (Tab. 1, Fig. S2). Under these conditions, the V_{max} of AldT was only 5.2 U/mg, which is 4.6-fold lower than that previously reported for AldT (Nishiya et al., 2014). We determined the K_m value of 9.73 ± 0.78 mM for D-mannose with NAD^+, which is lower than the suggested previously (D-mannose with NAD^+ of 25.0 ± 1.8 mM) (Nishiya et al., 2014). Since in both cases enzymes were heterologously-expressed and His-tag purified, the difference is attributed to a different pH and temperature of the assays. ManD was investigated as part of a separate study towards two different substrates, since the oxidation product of mannose is present in an equilibrium of free sugar acid (D-mannonate) and its lactone (D-mannono-1,4-lactone) (Kopp et al. 2018, submitted). The enzyme displayed low V_{max} values for both

substrates (2.39 and 1.9 U/mg for D-mannonate and D-mannono-lactone in buffer, respectively), with a relatively high K_m values (5.37 and 4.90 mM for D-mannonate and D-mannono-lactone in buffer, respectively). KDGA from *T. acidophilum* was analysed in the direction of aldol addition with pyruvate and glyceraldehyde as substrates. The enzyme exhibited a high V_{max} of 17.04 ± 1.03 mM which is in a similar range as compared to previously characterised KDG aldolases from different *Sulfolobus* spp. (V_{max} of *S. solfataricus* for pyruvate: 15.7 ± 0.3 U/mg) (Buchanan et al., 1999). The K_m of the KDGA from *T. acidophilum* for both substrates combined (8.07 ± 1.29 mM) was high but in a comparable range to values reported for the KDG aldolase from *S. solfataricus* (for D,L-glyceraldehyde: 5.2 ± 0.1 mM) (Buchanan et al., 1999). LDH displayed very high specific activity (194.3 ± 11.8 U/mg) for pyruvate under our assay conditions which was comparable to those reported previously (241 ± 30 U/mg, Halliwell et al., 2001). However, the K_m of 10.95 ± 1.67 mM for pyruvate was found to be substantially different from that previously reported (1.65 ± 0.12 mM, Halliwell et al., 2001).

Table 1 K_m and V_{max} values of selected pathway enzymes. Values for AldT, KDGA, LDH were obtained in 20 mM NaP buffer pH 7 at 50°C. Values are means of triplicate experiments and standard deviation (±) is indicated. Kinetic data for ManD was previously obtained at 50°C in 50 mM HEPES pH 7 (Kopp et al. 2018, submitted). n.d.: not determined.

enzyme	substrate	K_m [mM]	V_{max} [U/mg]	k_{cat} [s^{-1}]	k_{cat}/K_m [mM^{-1} s^{-1}]	pH opt
AldT	mannose	9.73 ± 0.78	5.20 ± 0.10	2.51	0.26	10[1]
ManD	mannonate	5.37 ± 0.90	2.39 ± 0.11	1.67	0.30	7[2]
	mannono-1,4-lactone	4.90 ± 0.53	1.90 ± 0.06	1.33	0.26	
KDGA	pyruvate + glyceraldehyde	8.07 ± 1.29	17.04 ± 1.03	10.51	1.30	n.d.
LDH	pyruvate	10.95 ± 1.67	194.3 ± 11.8	116.58	10.65	5-7[3]

[1] (Nishiya et al., 2014), [2] (Kopp et al. 2018, submitted) [3](Clarke et al., 1988)

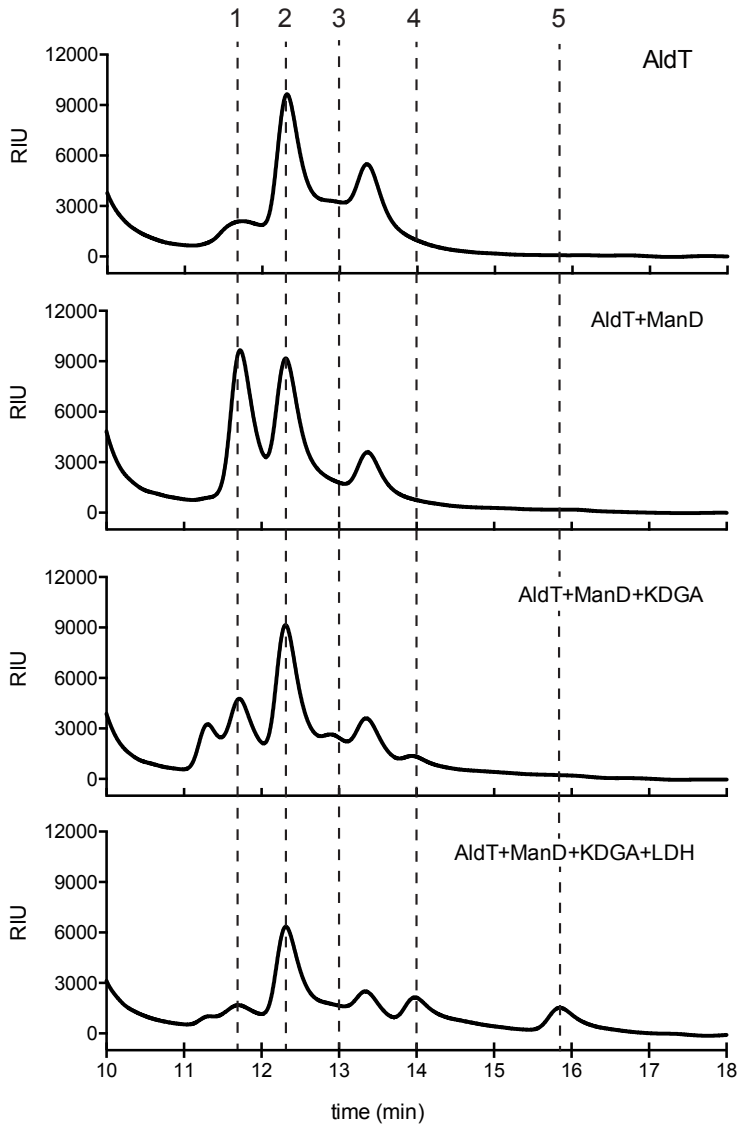


Figure 2 Exemplary sequential enzymatic pathway reactions demonstrating the activity of each enzyme and product formation after addition of all enzymes. Reactions contained a final concentration of 20 mM mannose and NAD^+ in NaP buffer (50mM) and 0.1 mM $CoSO_4$. Enzyme loadings were: AldT: 0.16 U/ml, ManD: 0.07 U/ml, KDGA 0.51 U/ml, 5.8 U/ml LDH. Retention times of standards for pathway intermediates 1: KDG (11.69 min), 2: Mannose (12.29 min), 3: Pyruvate (13 min), 4: Glyceraldehyde (14 min), 5: L-Lactic acid (15.9 min). RIU: refractive index units. For separation of standards see Fig. S4.

3.3 Sequential reactions and enzyme validation

Proof-of-concept biotransformation was conducted initially in sequential reactions with D-mannose and equal amount of NAD^+. Each enzyme of the pathway was added sequentially and individual product accumulation was monitored by HPLC (Fig. 2). Besides a clearly separated and quantifiable LA peak, an incomplete separation of all other pathway intermediates only allowed a qualitative analysis (Fig. S4). The first

reaction was performed by AldT with mannose and NAD^+, yielding mannonate in different conformations represented by peaks at 11.7 min and 13.3 min in the chromatogram. The subsequent addition of ManD yielded KDG while remaining amounts of mannose were observed. The addition of KDGA cleaved the C6 sugar into two C3 compounds, glyceraldehyde and pyruvate. LA formation was observed only after addition of the last enzyme (LDH) but not in any of the preceding reactions. After addition of all enzymes, mostly mannose and smaller amounts of mannonate, glyceraldehyde and KDG, including several unspecific products, were still visible. The initial yield of the proof-of-concept reaction was 18.9 ± 6.3% with 3.78 ± 1.26 mM LA produced from 20 mM mannose.

3.4 Validation and optimisation of one-pot reactions

Next, we performed one-pot reactions under varying parameters to identify the most important factors influencing the efficiency of the process. In order to investigate the effect of enzyme composition, the loading of each individual enzyme was reduced to 10% while keeping the other enzymes of the pathway at a constant loading (Table 2). Lower loads of AldT and LDH resulted in comparably low LA yields of 45.5 ± 3.6 % and 45.5 ± 1.1 %, respectively. The most drastic effect was obtained when the ManD loading was reduced, which resulted in a LA yield of only 20.2 ± 0.7 %. This suggests that the ManD loading appeared to be the most critical in the one-pot reaction. No significant difference was observed between lower KDGA loading and the standard loading, indicating that a 10-fold reduction in KDGA could be used without compromising the final LA yield. The standard enzyme loading of 0.2 U/ml AldT, 0.1 U/ml ManD, 0.4 U/ml KDGA and 11.6 U/ml resulted in a LA yield of 71.5 ± 4.1 % from 5 mM mannose and NAD^+. This enzyme composition was used for following one-pot reactions.

Table 2 Variation of each individual enzyme loading and its effect on LA yield. Red cells indicate decreased enzyme loading (10% of standard enzyme loading). Reactions were performed at least in duplicate and contained 5 mM mannose and NAD^+, 0.1 mM $CoSO_4$ in 0.1 M NaP buffer. Standard deviation of the mean (±) is indicated.

Enzyme	Enzyme amount [U/ml]				standard enzyme loading [U/ml]
AldT	0.02	0.2	0.2	0.2	0.2
ManD	0.1	0.01	0.1	0.1	0.1
KDGA	0.4	0.4	0.04	0.4	0.4
LDH	11.6	11.6	11.6	1.6	11.6
Yield (%)	**45.5 ± 3.6**	**20.2 ± 0.7**	**68.6 ± 0.04**	**45.5 ± 1.1**	**71.5 ± 4.1**
Titer (mM LA)	**2.3 ± 0.18**	**1 ± 0.03**	**3.4 ± 0.002**	**2.3 ± 0.05**	**3.6 ± 0.21**

Besides enzyme concentration, process duration and operational temperature directly influence enzymatic stability. An efficient process aims not only at a high yield but also high titres after a process-specific period of time. Half-life times for enzymes found in the literature are often acquired at higher temperatures but at assay times much shorter than those used in their actual operational conditions. In addition, thermostability depends very much on the concentration of the particular enzyme during the thermostability assay, which in many cases differ from the actual concentration used in the process. Therefore, in order to assess the operational stability of each enzyme in the pathway, enzymes were assayed at their standard enzyme loading, in the same buffer and at 50°C over a 20 h period (Fig. S3). Under these conditions LDH and KDGA remained fully active over the entire assay period. AldT displayed a half-life of 9.8 h while retaining 30% of its initial activity after 20 h. The half-life of ManD was only 1.4 h and no activity could be detected after 6 h.

3.5 Substrate and cofactor concentration

The initial substrate concentrations of mannose and in particular the cofactor NAD^+ play important roles in the economic viability of the pathway, since substrates and cofactors, beside enzymes, are major cost drivers in a cell-free reaction process. *In vivo*, endogenous cofactors are regenerated continuously, e.g. through coupled side reactions. However, *in vitro*, cofactors are supplied externally into a cell-free biocatalytic system and may considerably affect its cost-effectiveness. The cell-free pathway assembled here consisted of purified enzymes and, therefore, lacked the endogenous

cofactors required to start and drive the first reaction of the pathway. Accordingly, NAD^+ has to be supplied at the beginning of the pathway in sufficient amounts to catalyse the initial oxidation of mannose to mannonate until it is recycled again by LDH at the end of the pathway. However, since the pathway was assembled with thermostable enzymes and designed to operate at 50°C, cofactor thermal stability becomes an important factor to consider. Incubation of NAD^+ at 60°C in HEPES buffer pH 8 over a 6 h period has been shown to result in 50% decomposition of the cofactor into ADP-ribose and nicotinamide (Honda et al., 2016). Krutsakorn et al. found that NAD+ and NADH levels decreased to 50% during a reaction time of 4 h in 50 mM HEPES pH 7 and supplemented additional cofactor to achieve maximum conversion (Krutsakorn et al., 2013). In another study, no degradation was observed in 50 mM HEPES pH 7 at 60°C for 1 h (Ye et al., 2012). Besides its thermal instability, NAD^+ has also been shown to inhibit LDH from *Thermus thermophilus* almost completely above a concentration of 0.5 mM (Ye et al., 2012).

In order to study the effects of cofactors on the pathway performance, the production of LA was compared at different mannose concentrations in combination with varying NAD^+ concentrations (Fig. 3A). For reactions with varying NADH concentrations (e.g. 1.25, 2.5 and 5 mM NAD^+) at an equal mannose concentration (e.g. 5 mM mannose), no major difference in LA yield was observed. This suggested that cofactor decomposition was not affecting LA yields at this stage. The same was true for reactions with higher mannose concentrations (10 and 25 mM mannose), indicating that NAD^+ was efficiently recycled to NADH. However, independent of the cofactor concentration, an increase in mannose concentration caused a decrease in LA yield (Fig. 3A). We investigated whether the decrease in yield at high substrate concentrations was caused by insufficient buffer capacity, since a high production of organic acids from sugar acidifies the reaction. The ionic strength of the buffer was increased from 100 to 400 mM in reactions with 25 and 50 mM mannose at equal NAD^+ concentrations (Fig. 3B). Enhancing the ionic strength increased the LA yield from 2.6 ± 0.3% to 20.5 ± 0.5% in reactions with 50 mM mannose while the LA yield slightly decreased in reactions with 25 mM mannose. This suggested that only at very high mannose concentrations (50 mM) acidification of the solution was limiting the production of LA. For the decrease in LA yield in reactions with up to 25 mM mannose, there must have been another reason. Although no substrate inhibition was observed in single enzyme kinetic studies (Fig. S2), it is possible that pathway intermediates other than the natural substrate of each enzyme act as inhibitor.

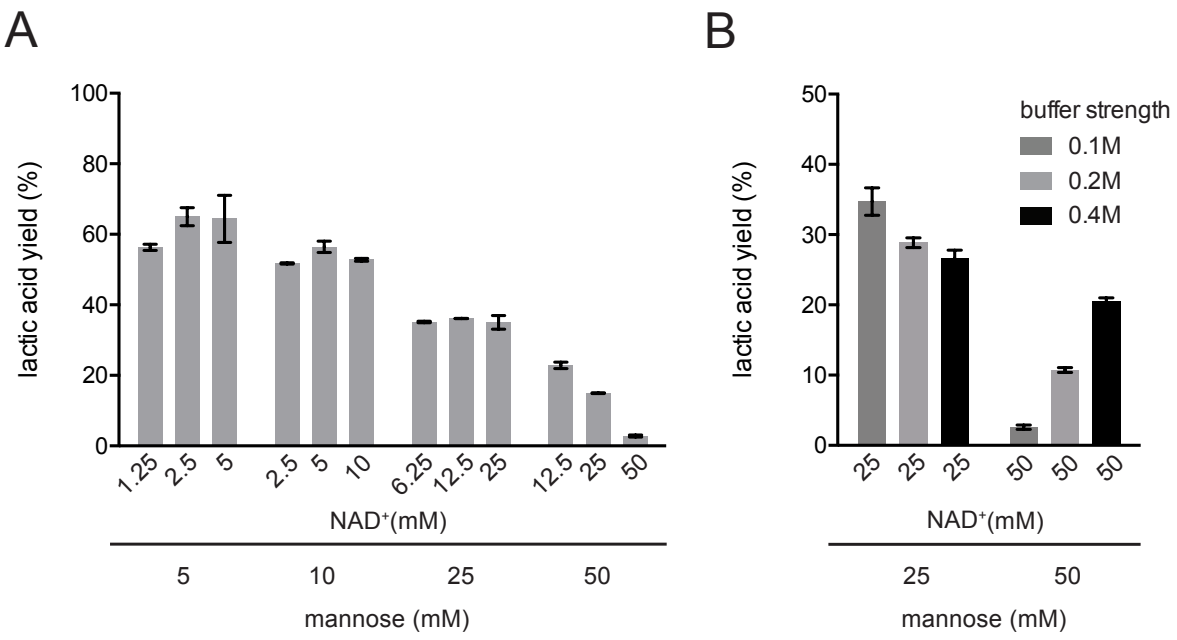


Figure 3 A) Effect of different mannose and cofactor concentrations on LA yield. Reactions were performed in 0.1 M sodium phosphate buffer pH 7.5 with standard enzymes loading **B)** Effect of ionic strength on LA yield in reactions with high substrate concentrations. Reactions were performed in duplicate with 25 and 50 mM mannose, equimolar NAD^+ concentration and with standard enzymes loading. Values display mean of duplicate experiments. Error bars represent the standard deviation.

3.6 LA production from SCG

The feasibility of the cell-free pathway for the conversion of biomass to LA was tested with sugars obtained from SCG. In order to substitute purposely-refined sugar as substrate, we obtained mannose directly from SCG galactomannan. SCG were dried, subjected to a standard dilute acid hydrolysis procedure and analysed by HPLC (Fig. 4A) (Sluiter et al., 2011). The mannose content of 100 mg SCG was 28.7 ± 1.5 mg. However, due to harsh hydrolysis conditions 19.5 % of mannose was lost, leaving 24 ± 1.26 mg/100 mg SCG available for subsequent conversion. The mannose content is comparable to previous obtained values using similar dilute acid methods (Mussatto et al., 2011). The hydrolysate was neutralised, concentrated and used as substrate for the cell-free enzymatic production of LA. The conversion with mannose obtained from SCG was compared to reactions with pure mannose by monitoring the decrease in mannose and the accumulation of LA over time (Fig. 4B).

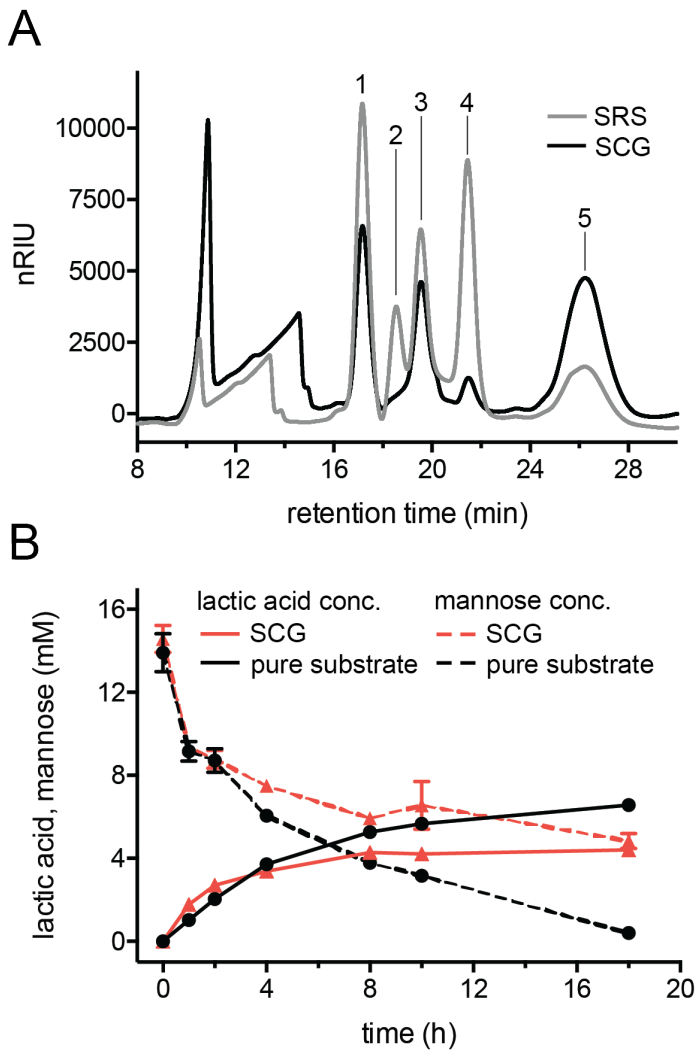


Figure 4 A) Analysis of monosaccharides after dilute acid hydrolysis of SCG. Samples were neutralised with $CaCO_3$ and concentrated 10-fold before subjected to HPLC-RID analysis. SRS: sugar reducing standards (grey line), SCG: spent coffee grounds hydrolysate (black line). Retention times of standards are indicated. 1: D-glucose 2: D-xylose 3: D-galactose 4: L-arabinose 5: D-mannose. **B)** Production of LA from mannose obtained from SCG (red lines) compared to production from pure mannose (black lines). Decrease of mannose (dashed lines) and increase of LA (solid lines) were monitored by HPLC-RID analysis and quantified by mannose and L-lactic acid standards (Fig. S6). Reactions contained 14.3 mM mannose, 3.75 mM NAD^+, 0.1 mM $CoSO_4$, and standard enzyme loading in 0.1 M NaP buffer pH 7. Reactions were performed in triplicate at $50^{O}C$. Error bars represent standard deviation.

HPLC-RID analysis of the reactions demonstrated that LA was produced successfully from SCG hydrolysate (Fig. 4B). In reactions with pure mannose as substrate, 6.6 ± 0.1 mM LA could be produced from 13.9 ± 0.9 mM mannose (47.4 ± 3.9% yield), whereas in reactions with mannose obtained from SCG, 4.4 ± 0.1 mM LA was produced from 14.57 ± 0.7 mM mannose (30.3 ± 1.5% yield). Based on the concentration of mannose remaining compared to the amount of product formed, it was evident that a significant amount of mannose was not converted to LA in reactions from SCG. In contrast, mannose was almost completely oxidised in reactions with pure mannose. Taking into account that 50% of the carbon is not converted to LA but is accumulating as

glyceraldehyde, 7 mM mannose are lost on average in in the reactions with pure mannose and 9.7 mM mannose in the reactions with SCG. Further investigation has to be performed to identify inhibitory compounds present in the SCG hydrolysate, which limit an optimal pathway perormance as demonstrated with pure D-mannose. Dilute acid hydrolysis combined with heat can result in high amounts of toxic substances such 5- hydromethylfurfural (5HMF) and levulinic acid, which inhibit microbial growth (Obruca et al. 2014). Wehter these compounds are also responsible for lower enzyme activity has to be tested.

3.7 Real-time tracking of the cell-free pathway using ^{13}C NMR

Monitoring of intermediate metabolites in cell-free pathways is a crucial tool in understanding flux and identifying bottlenecks. Since the highest yield of the pathway was 71.5% with 5 mM mannose and NAD^+, the rest of the carbon was probably lost in incomplete or unspecific side reactions. Using several HPLC setups and conditions did not result in a clear separation of all pathway metabolites (Fig. S4). This has been shown previously to be a challenging task for similarly polar organic acids (Xie et al., 2018). To address this challenge, we performed reactions with ^{13}C labelled substrate in combination with nuclear magnetic resonance (NMR) spectroscopy to trace every carbon during the reaction in the system. Reactions with D-[1,6-$^{13}C_2$]mannose were performed at 50°C over a time period of 16 h, while decoupled 1D ^{13}C-NMR spectra were acquired every 6.5 min (Fig. 5A). Chemical shifts appearing during the experiment, were allocated to pathway intermediates by comparison to 1D ^{13}C-NMR spectra of unlabelled standards of reaction intermediates (Table S2, Fig. S5). Peak areas for each chemical shift were integrated for all acquired spectra over the time of the reaction. The integrals for each chemical shift represent the relative amount of each intermediate in relation to the maximal integral occurring during the reaction (Fig. 5B). Absolute intensity between peaks cannot be accurately used, as this would require longer acquisition times and a different pulse program to prevent ^{1}H NOE enhancement of the ^{13}C signal, which would also require longer acquisition times. However, the compromise of short acquisition times and use of relative maximum intensity still allows monitoring of the flux for each intermediate at a high time resolution.

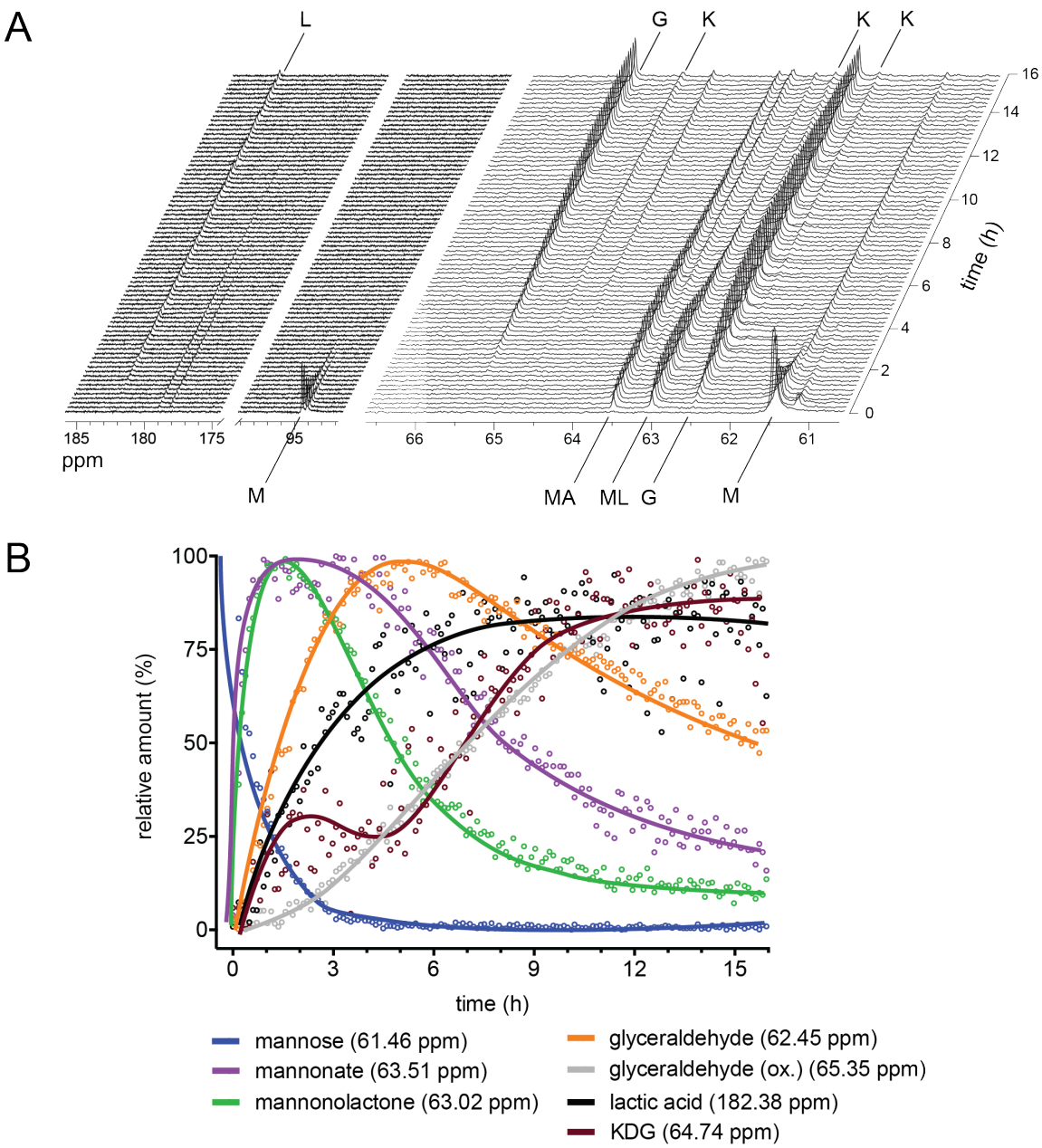


Figure 5 A) Real-time ^{13}C NMR analysis of one-pot reaction performed with 10 mM D-[1,6-$^{13}C_2$]mannose and 10 mM NAD^+ for 16 h. M: mannose, MA: mannonate, ML: mannonolactone K: KDG, G: glyceraldehyde L: L-lactic acid, glyceraldehyde (ox.): oxidisation product of glyceraldehyde. **B)** Relative amounts of the intermediates derived from integration of areas for characteristic chemical shifts for each pathway intermediate. Only integrals for the lower chemical shift of C6 compounds were analysed. Lines show approximations and were manually fitted.

D-[1,6-$^{13}C_2$]mannose was identified in ^{13}C NMR by the doublet at 94.09/94.46 ppm representing the α- and β-anomers of D-mannose and a single peak at 61.46 ppm representing carbon 6 (Fig. 5A). Both peaks decreased in their area very rapidly and mannose was fully consumed after 3 h (Fig. 5B). The reaction products of the initial oxidation of mannose, mannonate and D-mannono-1,4-lactone, was identified by a chemical shift difference at its carbon 6. Both oxidation products appeared almost simultaneously, but the level of D-mannono-1,4-lactone decreased earlier in

comparison to the free sugar acid, D-mannonate. Most probably this was due to the hydrolysis of the lactone into its free acid form. About 20-25% of mannonate was still not converted after 16 h of incubation, which contributed to the loss of carbon in the cell-free conversion to LA. The next intermediate in the reaction, KDG, was represented by three chemical shifts, which were observed in the standard and in the one-pot reaction (Table S2). The most pronounced shift occurred at 64.74 ppm, while two smaller peaks were visible at 62.24 ppm and 62.78 ppm. After a sharp increase and decrease during the first 3 h, KDG levels continued to accumulate linearly throughout the rest of the reaction. A possible reason for the accumulation of KDG might be related to the low positive change in Gibbs energy (+6.2 kJ/mol) of the aldolase cleavage reaction. Although LDH is highly active and converts pyruvate rapidly, it has a lower affinity to pyruvate than KDGA (Tab. 1). Therefore, an aldol addition reaction of glyceraldehyde and pyruvate to KDG is likely to occur. Additionally, we observed a lack in diastereoselectivity of the aldolase (data not shown) similar to that observed with other aldolases such as the one from *S. acidocaldarius*, which yields not only KDG but also 2-keto-3-deoxygalactonate (KDGal) (Theodossis et al., 2004). KDGal potentially caused more unallocated chemical shifts in the NMR analysis. Chemical shifts for pyruvate did not appear in the time-course reaction, since LDH was expected to convert its substrate rapidly into LA. Interestingly, two different chemical shifts were assigned to glyceraldehyde, which follow different kinetics. The integral of the chemical shift at 62.45 ppm reached a maximum after ~6h reaction time and then continued to decrease until the end of the reaction. In contrast, the integral of chemical shift at 65.35 ppm increased almost in a linear fashion, which is indicating an unspecific oxidation of glyceraldehyde over time. The final product LA accumulated right from the start of the reaction. The strongest increase in product levels was observed already in the first 1-2 h of the reaction, eventually reaching a steady state after approximately 9 h, which is in line with the observations made from HPLC analysis (Fig. 4B).

The Maillard reaction has been described as a non-enzymatic browning reaction between amino- and carbonyl groups under heat (Maillard, 1912) that can affect the productivity of enzymatic pathways (Cheng et al., 2018). In our pathway, mannose was incubated with proteins and phosphate-based buffers which are known to promote the Maillard reaction. Based on this, one would expect that that mannose degraded and enzymes were inactivated over the time of the reaction (Bell, 1997). However, we were able to exclude this effect as no rise in unspecific chemical shifts could be observed in a

control reaction with mannose and bovine serum albumin (BSA) incubated under the standard conditions and temperature used with the cell-free pathway.

The pathway presented here does not fully exploit the carbohydrate substrate, since only pyruvate is converted to LA and glyceraldehyde remains unused. Guterl et al. demonstrated a recycling process for glyceraldehyde by implementation of an engineered aldehyde dehydrogenase and a promiscuous dihydroxyacid dehydratase (DHAD) (Guterl et al., 2012; Steffler et al., 2013). The natural promiscuity of the enzyme could be exploited to convert 100% carbon towards KDG while keeping the NAD^+/NADH ratio balanced. A similar recycling process could be combined with the pathway presented in this work to enable a conversion of the complete carbon in mannose and avoid the accumulation of glyceraldehyde.

4 Conclusion

We successfully demonstrated a proof of concept for a cell-free synthetic enzymatic pathway for the conversion of mannose, a major constituent of SCG, into the platform chemical LA. The pathway is free of heat-labile phosphorylated compounds, requires only NADH and has fewer enzymatic reaction steps compared to the EMP pathway. In comparison to a cell-based implementation of the pathway, the cell-free approach overcomes challenging aspects such as insufficient enzyme activity *in vivo.* This can be caused by insufficient iron loading, encountered upon heterologous expression of enzymes from an archaeal background in *S. cerevisiae* (Benisch and Boles, 2014). Another advantage is that the performance of the pathway can be rapidly assessed and potential bottlenecks isolated in a shorter time. Although the pathway presented here can be considered synthetic, its operation is in principle possible *in vivo* and resembles similar oxidative metabolisms such as the np-ED, the Weimberg or Dahms pathways (Budgen and Danson, 1986; Dahms, 1974; Weimberg, 1961). Growth experiments with *T. acidophilum* on mannose in conjunction with transcription analysis could be used to verify this hypothesis.

Conflict of Interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

DK is supported by an international Macquarie University Research Excellence Scholarship (iMQRES).

Acknoledgments

We thank Dr. Nicole Cordina from the Macquarie University NMR facility from the Department of Molecular Sciences for her help in performing NMR studies.

Supplementary information

Supplementary data associated with this article can be found in the online version.

References

Ballesteros, L.F., Teixeira, J.A., Mussatto, S.I., 2014. Chemical, functional, and structural properties of spent coffee grounds and coffee silverskin. Food Bioprocess Technol. 7, 3493–3503.

Beer, B., Pick, A., Sieber, V., 2017. *In vitro* metabolic engineering for the production of α-ketoglutarate. Metab. Eng. 40, 5–13.

Bell, L.N., 1997. Maillard reaction as influenced by buffer type and concentration. Food Chem. 59, 143–147.

Benisch, F., Boles, E., 2014. The bacterial Entner-Doudoroff pathway does not replace glycolysis in *Saccharomyces cerevisiae* due to the lack of activity of iron-sulfur cluster enzyme 6-phosphogluconate dehydratase. J. Biotechnol. 171, 45–55.

Buchanan, C.L., Connaris, H., Danson, M.J., Reeve, C.D., Hough, D.W., 1999. An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. Biochem. J 343, 563–570.

Budgen, N., Danson, M.J., 1986. Metabolism of glucose via a modified Entner-Doudoroff pathway in the thermoacidophilic archaebacterium *Thermoplasma acidophilum*. FEBS Lett. 196, 207–210.

Caetano, N.S., Silva, V.F.M., Mata, T.M., 2012. Valorization of coffee grounds for biodiesel production. Chem. Eng. Trans. 26, 267–272.

Campos-Vega, R., Loarca-Piña, G., Vergara-Castañeda, H. a., Oomah, B.D., 2015. Spent coffee grounds: a review on current research and future prospects. Trends Food Sci. Technol. 45, 24–36.

Castro-Aguirre, E., Iñiguez-Franco, F., Samsudin, H., Fang, X., Auras, R., 2016. Poly(lactic acid)-Mass production, processing, industrial applications, and end of life. Adv. Drug Deliv. Rev. 107, 333–366.

Chen, X., Nielsen, K.F., Borodina, I., Kielland-Brandt, M.C., Karhumaa, K., 2011. Increased isobutanol production in *Saccharomyces cerevisiae* by overexpression of genes in valine metabolism. Biotechnol. Biofuels 4, 1–12.

Cheng, K., Zheng, W., Chen, H., Job Zhang, Y.-H.P., 2018. Upgrade of wood sugar D-xylose to a value-added nutraceutical by *in vitro* metabolic engineering. Metab. Eng. 52, 1–8.

Clarke, A.R., Wilks, H.M., Barstow, D.A., Atkinson, T., Chia, W.N., Holbrook, J.J., 1988. An investigation of the contribution made by the carboxylate group of an active site histidine-aspartate couple to binding and catalysis in lactate dehydrogenase. Biochemistry 27, 1617–1622.

Dahms, S.A., 1974. 3-Deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. Biochem. Biophys. Res. Commun. 60, 1433–1439.

Dudley, Q.M., Anderson, K.C., Jewett, M.C., 2016. Cell-Free Mixing of *Escherichia coli* crude extracts to prototype and rationally engineer high-titer mevalonate synthesis. ACS Synth. Biol. 5, 1578–1588.

Dudley, Q.M., Karim, A.S., Jewett, M.C., 2015. Cell-free metabolic engineering: Biomanufacturing beyond the cell. Biotechnol. J. 10, 6982.

Dusselier, M., Van Wouwe, P., Dewaele, A., Makshina, E., Sels, B.F., 2013. Lactic acid as a platform chemical in the biobased economy: The role of chemocatalysis. Energy Environ. Sci. 6, 1415–1442

Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W., Milo, R., 2013. Glycolytic strategy as a tradeoff between energy yield and protein cost. Proc. Natl. Acad. Sci. 110, 10039–10044.

Gao, C., Li, Z., Zhang, L., Wang, C., Li, K., Ma, C., Xu, P., 2015. An artificial enzymatic reaction cascade for a cell-free bio-system based on glycerol. Green Chem. 17, 804–807.

Guterl, J.K., Garbe, D., Carsten, J., Steffler, F., Sommer, B., Reiße, S., Philipp, A., Haack, M., Rühmann, B., Koltermann, A., Kettling, U., Brück, T., Sieber, V., 2012. Cell-free metabolic engineering: production of chemicals by minimized reaction cascades. ChemSusChem 5, 2165–2172.

Halliwell, C.M., Morgan, G., Ou, C., Cass, A.E.G., 2001. Introduction of a (poly-)histidine Tag in L-lactate dehydrogenase produces a mixture of active and inactive molecules. Anal. Biochem. 295, 257–261.

Hald, E., Lehmann, P., Ziegenhorn, J., 1975. Molar Absorptivities of β-NADH and β-NAD at 260 nm. Clin. Chem. 21, 884–887.

Hold, C., Billerbeck, S., Panke, S., 2016. Forward design of a complex enzyme cascade reaction. Nat. Commun. 7, 1–8.

Honda, K., Hara, N., Cheng, M., Nakamura, A., Mandai, K., Okano, K., Ohtake, H., 2016. *In vitro* metabolic engineering for the salvage synthesis of NAD+. Metab. Eng. 35, 114–120.

Honda, K., Kimura, K., Ninh, P.H., Taniguchi, H., Okano, K., Ohtake, H., 2017. *In vitro* bioconversion of chitin to pyruvate with thermophilic enzymes. J. Biosci. Bioeng. 124, 296–301.

Jooste, T., García-Aparicio, M.P., Brienzo, M., Van Zyl, W.H., Görgens, J.F., 2013. Enzymatic hydrolysis of spent coffee ground. Appl. Biochem. Biotechnol. 169, 2248–2262.

Kondamudi, N., Mohapatra, K.S., Misra, M., 2008. Spent coffee grounds as a versatile source of green energy. J. Agric. Food Chem. 56, 11757–11760.

Kopp, D., Willows, R.D., Sunna, A., 2018. Characterisation of a mannonate dehydratase from *Thermoplasma acidophilum* and its potential role in the catabolism of D-mannose. Sci. Rep. sumbitted.

Krutsakorn, B., Honda, K., Ye, X., Imagawa, T., Bei, X., Okano, K., Ohtake, H., 2013. *In vitro* production of n-butanol from glucose. Metab. Eng. 20, 84–91.

Lamble, H.J., Milburn, C.C., Taylor, G.L., Hough, D.W., Danson, M.J., 2004. Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in *Sulfolobus solfataricus*. FEBS Lett. 576, 133–136.

Li, Z., Zhang, M., Jiang, T., Sheng, B., Ma, C., Xu, P., Gao, C., 2017. Enzymatic cascades for efficient biotransformation of racemic lactate derived from corn steep water. ACS Sustain. Chem. Eng. 5, 3456–3464.

Machado, E.M.S., Rodriguez-Jasso, R.M., Teixeira, J.A., Mussatto, S.I., 2012. Growth of fungal strains on coffee industry residues with removal of polyphenolic compounds. Biochem. Eng. J. 60, 87–90.

Maillard, L.C., 1912. Action des acides amines sur les sucres; formation des melanoidines par voie methodique. Comptes R. Acad. Sci.(Paris) 154, 66–68.

Meléndez-Hevia, E., Waddell, T.G., Heinrich, R., Monter, F., 1997. Theoretical approaches to the evolutionary optimization of glycolysis chemical analysis. Eur. J. Biochem. 244, 527–543.

Morgado, G., Gerngross, D., Roberts, T.M., Panke, S., 2016. Synthetic biology for cell-free biosynthesis: fundamentals of designing novel in vitro multi-enzyme reaction networks. Synth. Biol. Eng. 20, 117–146.

Murphey, W.H., Barnaby, C., Lin, F.J., Kaplan, N.O., 1967. Malate dehydrogenases II. Purification and properties of *Bacillus subtilis*, *Bacillus stearothermophilus*, and *Escherichia coli* malate dehydrogenases. J. Biol. Chem. 242, 1548–1559.

Mussatto, S.I., Carneiro, L.M., Silva, J.P.A., Roberto, I.C., Teixeira, J.A., 2011. A study on chemical constituents and sugars extraction from spent coffee grounds. Carbohydr. Polym. 83, 368–374.

Mussatto, S.I., Machado, E.M.S., Carneiro, L.M., Teixeira, J.A., 2012. Sugars metabolism and ethanol production by different yeast strains from coffee industry wastes hydrolysates. Appl. Energy 92, 763–768.

Nishiya, Y., Tamura, N., Tamura, T., 2014. Analysis of bacterial glucose dehydrogenase homologs from thermoacidophilic archaeon *Thermoplasma acidophilum*: finding and characterisation of aldohexose dehydrogenase. Biosci. Biotechnol. Biochem. 68, 2451–2456.

Obruca, S., Benesova, P., Petrik, S., Oborna, J., Prikryl, R., Marova, I., 2014. Production of polyhydroxyalkanoates using hydrolysate of spent coffee grounds. Process Biochem. 49, 1409–1414.

Obruca, S., Benesova, P., Kucera, D., Petrik, S., Marova, I., 2015. Biotechnological conversion of spent coffee grounds into polyhydroxyalkanoates and carotenoids. N. Biotechnol. 32, 569–574.

Ramalakshmi, K., Mohan, L.J., Takano-ishikawa, Y., Goto, M., 2009. Bioactivities of low-grade green coffee and spent coffee in different in vitro model systems. Food Chem. 115, 79–85.

Reher, M., Fuhrer, T., Bott, M., Schönheit, P., 2010. The nonphosphorylative Entner-Doudoroff pathway in the thermoacidophilic euryarchaeon *Picrophilus torridus* involves a novel 2-Keto-3-deoxygluconate- specific aldolase. J. Bacteriol. 192, 964–974.

Sambrook, J., Russell, D.W., Russell, D.W., 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.

Sauer, M., Porro, D., Mattanovich, D., Branduardi, P., 2010. 16 years research on lactic acid production with yeast – ready for the market? Biotechnol. Genet. Eng. Rev. 27, 229–256.

Skoza, L., Mohos, S., 1976. Stable thiobarbituric acid chromophore with dimethyl sulphoxide. Application to sialic acid assay in analytical de-O-acetylation. Biochem. J. 159, 457–62.

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., Crocker, D., 2011. Determination of structural carbohydrates and lignin in biomass. Lab. Anlaytical Proced. 1617, 1–16.

Steffler, F., Guterl, J.K., Sieber, V., 2013. Improvement of thermostable aldehyde dehydrogenase by directed evolution for application in synthetic cascade biomanufacturing. Enzyme Microb. Technol. 53, 307–314.

Takao, I., Fujii, S., Ishii, A., Han, L., Kumao, T., Ozaki, K., Asakawa, A., 2006. Effects of mannooligosaccharides from coffee mannan on fat storage in mice fed a high fat diet. J. Heal. Sci. 52, 333–337.

Theodossis, A., Walden, H., Westwick, E.J., Connaris, H., Lamble, H.J., Hough, D.W., Danson, M.J., Taylor, G.L., 2004. The structural basis for substrate promiscuity in 2-keto-3-deoxygluconate aldolase from the Entner-Doudoroff pathway in *Sulfolobus solfataricus*. J. Biol. Chem. 279, 43886–4892.

Tokimoto, T., Kawasaki, N., Nakamura, T., Akutagawa, J., Tanada, S., 2005. Removal of lead ions in drinking water by coffee grounds as vegetable biomass. J. Colloid Interface Sci. 281, 56–61.

Vink, E.T.H., Rábago, K.R., Glassner, D.A., Gruber, P.R., 2003. Applications of life cycle assessment to NatureWorks^{TM} polylactide (PLA) production. Polym. Degrad. Stab. 80, 403–419.

Weimberg, R., 1961. Pentose oxidation by *Pseudomonas fragi*. J. Biol. Chem. 236, 629–35.

Xie, L., Wei, X., Zhou, X., Meng, D., Zhou, R., Zhang, Y.H.P.J., Xu, S., You, C., 2018. Conversion of D-glucose to L-lactate via pyruvate by an optimized cell-free enzymatic biosystem containing minimized reactions. Synth. Syst. Biotechnol. 1–7.

Ye, X., Honda, K., Sakai, T., Okano, K., Omasa, T., Hirota, R., Kuroda, A., Ohtake, H., 2012. Synthetic metabolic engineering-a novel, simple technology for designing a chimeric metabolic pathway. Microb. Cell Fact. 11, 120.

Ye, X., Wang, Y., Hopkins, R.C., Adams, M.W.W., Evans, B.R., Mielenz, J.R., Zhang,

Y.H.P., 2009. Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. ChemSusChem 2, 149–152.
Zuorro, A., Lavecchia, R., 2012. Spent coffee grounds as a valuable source of phenolic compounds and bioenergy. J. Clean. Prod. 34, 49–56.

Supplementary data

Cell-free enzymatic conversion of spent coffee grounds into the platform chemical lactic acid

Dominik Kopp[a], Robert D. Willows[a,b] and Anwar Sunna[a,b]*

[a]Department of Molecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia

[b]Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, New South Wales 2109, Australia

Table S1 Oligonucleotides used in the study. All oligonucleotides were synthesised by IDT DNA Technologies (Singapore). For oligonucleotide sequences, restriction sites are written in lowercase.

Oligonucleotide name	Gene amplified	Oligonucleotide sequence (5' −> 3')
AldT-fw	aldohexose dehydrogenase	TGATggatccCATGTTCAG CGATCTAAGGGA
AldT-rev	aldohexose dehydrogenase	ATTATgaattcCTCATTCTG GCGTGCTTATGG
Ta1157_EcoRI_fw	KDG aldolase	ATATTgaattcCTACAAGG GTATCGTAACGCC
Ta1157_HindIII_rev	KDG aldolase	TGATaagcttATTGCTTCTT CGGGATTTTG
BsLDH_BamHI_fw	L-lactate dehydrogenase	ATTATggatccTAAAAACA ACGGTGGAGCCC
BsLDH_HindIIII_rev	L-lactate dehydrogenase	TGATATaagcttTCATCGC GTAAAAGCACGG

Table S2 Comparison of predicted chemical shifts, acquired chemical shifts of unlabelled standards and chemical shifts appearing in pathway reaction with D-[1,6-$^{13}C_2$]mannose and 10 mM NAD^+. Standards were obtained in the same buffer as the reaction took place as described in Material and Methods. n.d.: not detected. Prediction was performed with ChemDraw Professional 16.0 software.

	prediction		pathway reaction		standards	
	C1	C6	C1	C6	C1	C6
mannose	98.5	62.2	94.463 (α) 94.087 (β)	61.459	94.463 (α) 94.087 (β)	61.435
mannonate	176.2	64.4	179.002	63.518	n.d.	63.509
mannonolactone	175.3	64.6	178.168	63.006	n.d.	63.024
KDG	63.8	162.8	62.24 62.784 64.736	-	62.228 62.782 64.727	n.d.
pyruvate	-	162.8	-	170.595	-	-
glyceraldehyde	59.2	-	62.447 65.345	-	62.490 65.340	-
L-lactic acid	-	183	182.367	-	-	182.548

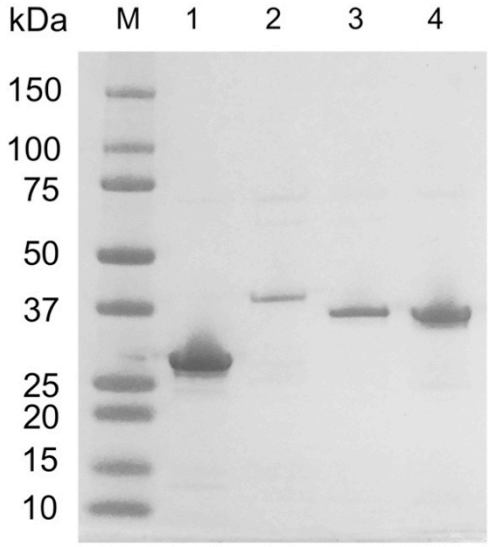


Figure S1 SDS-PAGE of the purified enzymes for the cell-free conversion of mannose to LA. AldT (1) and LDH (4) were expressed in *E.coli* BL21 cells. ManD (2) and KDGA (3) were expressed in BL21-CodonPlus (DE3)-RIL cells. Molecular masses of AldT, ManD, KDGA and LDH were 29 kDa, 42 kDa, 37 kDa and 36 kDa respectively. All enzymes were purified via their N-terminal His-tag using metal-affinity chromatography. M: Bio-Rad Precision Plus Protein Standard.

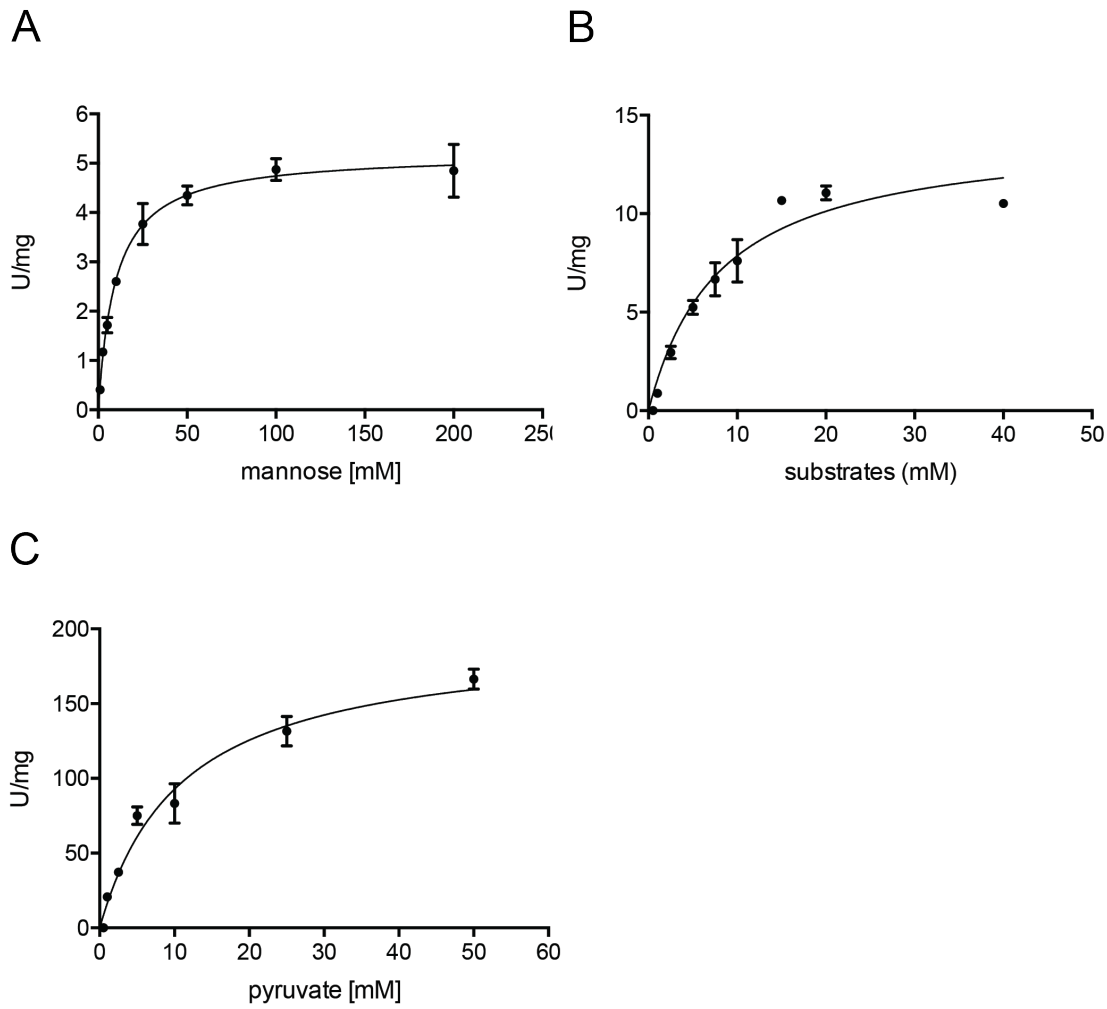


Figure S2 Kinetic data of purified enzymes. Values are mean of triplicate experiments. Error bars represent standard deviation. All data was fitted to Michaelis-Menten kinetics for one substrate. A) AldT for the substrate D-mannose B) KDGA measured in the direction of aldol addition with pyruvate and glyceraldehyde as substrates C) LDH for the substrate pyruvate. Curve fitting was performed with Prism 6 (GraphPad software).

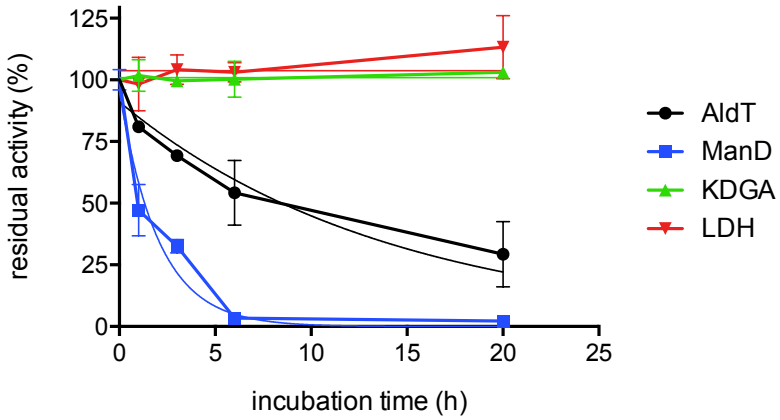


Figure S3 Thermostability of AldT (black line), ManD (blue line), KDGA (green line) and LDH (red line) at their standard enzyme loading at 50°C over the duration of the process time. Residual activity is expressed relative to activity without any incubation at 50°C.

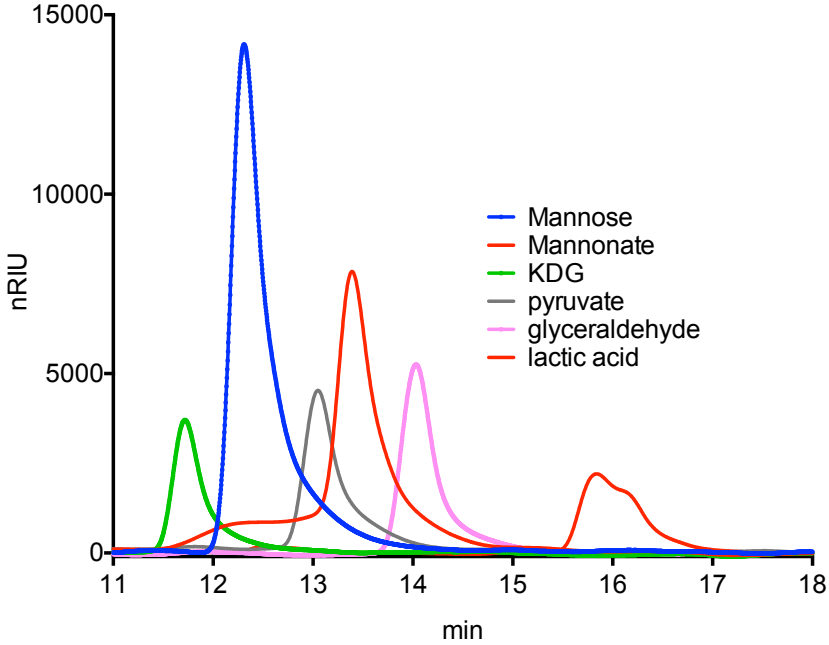


Figure S4 Separation of pathway intermediates on organic acid column. 10 µl of 10 mM standards except from KDG (2.5 mM) were subjected to Hiplex H^+ column connected to a RID (see Material and Methods for further details).

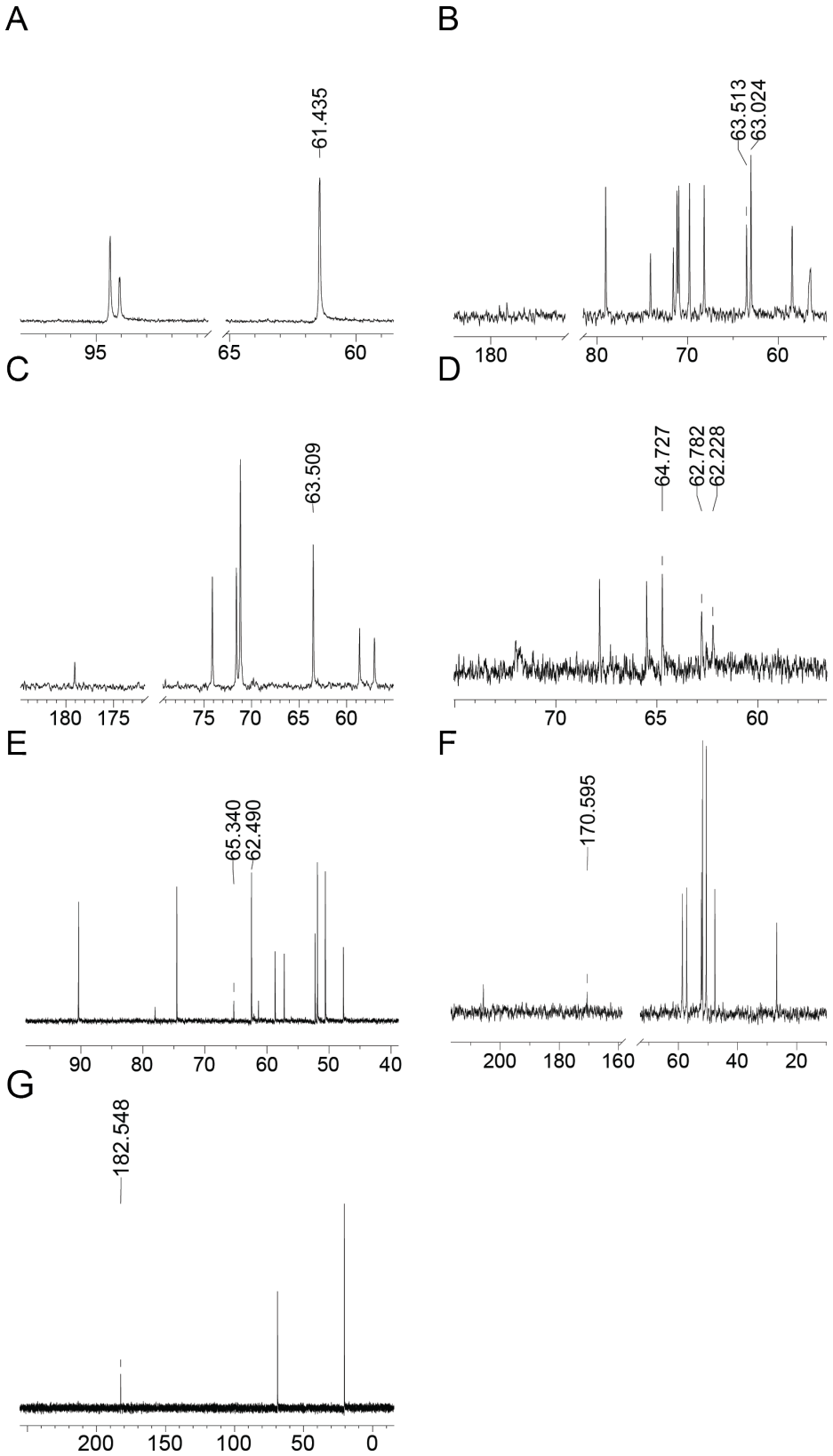


Figure S5 Decoupled 1D ^{13}C NMR spectra of standards. All values are expressed in ppm. All standards were acquired in 0.1 M NaP buffer pH 7.5 at 55°C. A: D-[1,6-$^{13}C_2$]mannose, B: mannonate, C: mannono-1,4-lactone, D: KDG, E: glyceraldehyde, F: pyruvate, G: L-lactic acid. Labelled peaks were found in real-time analysis of pathway reactions, except for peak at 170.595 ppm for pyruvate.

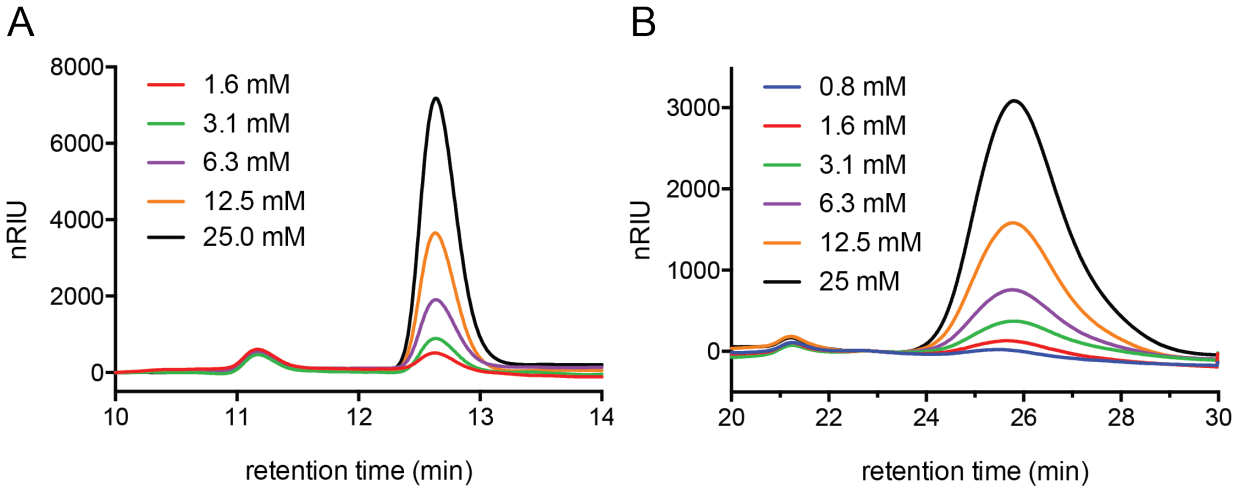


Figure S6 Standards used for the quantification of L-lactic acid (A) and mannose (B) in time course experiments for the conversion of mannose obtained from SCG and pure mannose into LA. Standards for LA were analysed on a Bio-Rad Aminex HPX-87H column and standards for mannose on an Agilent HiPlex Pb column (see Material and Methods for further details).

Chapter 4

Summary and future perspectives

The work presented in this PhD project established and validated a novel enzymatic conversion of a highly abundant waste product into a commonly used platform chemical. Through the implementation of a cell-free approach, it was possible to construct a conversion route based on a hypothetical mannose metabolism from the archaeon *T. acidophilum*. In contrast to the EMP pathway, which is commonly found in industrial organisms, the synthetic pathway is based on an oxidative principle. This allowed the assembly of simplified conversion pathway consisting of only 4 instead of 11 enzymes, no ATP requirement and a balancing of NADH cofactor. Several key points can be concluded from the successful proof of concept work presented here: (a) ManD marks the missing link between the oxidation of mannose and the conversion of 2-keto-3-deoxygluconate (KDG) into pyruvate and, therefore, enabled an oxidative mannose pathway; (b) hydrolysed spent coffee grounds (SCG) can be enzymatically converted to lactic acid (LA) at similar yield efficiencies as with pure mannose; (c) real-time NMR analysis is an effective tool for the tracking and verification of intermediates in a cell-free pathway and allows the identification of accumulation and/or unwanted side products.

Chapter 1 highlighted the different principles and reaction steps in oxidative and phosphorylative pathways using as example the ED and EMP pathways. In particular, we discussed aspects of oxidative pathways such as a superior thermodynamic profile, low enzyme cost and different cofactor usage, which are features that can be used to an advantage in metabolic engineering approaches. Special attention was given to key enzymes such as dehydratases and aldolases, two types of enzymes that were used in the subsequent construction of the cell-free pathway for the conversion of SCG to LA.

Chapter 2 summarised the identification of a so far uncharacterised mannonate dehydratases from *T. acidophilum*. The conformation of the sugar acid substrate for a dehydratase is often neglected in characterisation studies. For ManD, the presence of lactone and free sugar acid form determined its maximal activity. Elucidation of the enzyme's unique cofactor dependency raised new questions about the structure and cofactor binding mechanism of this enzyme, since none of the bacterial mannonate dehydratases described so far showed an equally strong activation with Co^{2+} and Ni^{2+}. Additional phylogenetic analysis suggested a possible different physiological role for UxuA mannonate dehydratases in archaea in a possible oxidative mannose pathway. This hypothesis could be investigated further by analysis of dehydrogenases and

dehydrates of closely related *Ferroplasma acidarmanus* and *Ferroplasma acidophilum* or with transcriptional analysis of mannose grown cells. Time limitations and prioritisation of the cell-free pathway design did not allow us to pursue additional physiological studies to answer this question and this should be part of future studies. However, the functionality and concerted raction of all three *T. acidophilum* enzymes, AldT, ManD and KDGA, strongly suggest the possibility of functional mannose pathway in some archaea similar to the np-ED pathway

In addition to the identified gene for ManD, several other genes encoding for potentially thermostable mannonate dehydratases were explored at first. Judging by annotation and previous proteomic studies, the gene locus Ta0085m is supposed to encode for a gluconate/galactonate dehydratase, the crucial enzyme of the np-ED pathway (Kim and Lee, 2005; Sun et al., 2010). Although *T. acidophilum* was one of the first organisms in which the np-ED pathway was found, key enzymes, namely the gluconate dehydratase and KDG aldolase were not identified at the time (Budgen and Danson, 1986). Expression of Ta0085m and other genes were attempted during this work, but hindered by low solubility of the recombinantly expressed protein.
Recombinant expression using *E. coli* BL21 strains, which have plasmids containing rare tRNAs has previously been proven to be a successful tool for the expression of enzymes from thermophilic origin (Kim et al., 2012; Nishiya et al., 2014; Steffler et al., 2013). However, this expression system was not successful for the expression of Ta0085m and soluble yields of ManD were quite low. A range of different approaches were attempted to increase the solubility of the recombinant proteins, especially ManD. Although they were mostly unsuccessful, ManD could be sucessfully expressed in *S. cerevisiae*, which is described in more detail in Appendix III.
Future apporaches to overcome low solubility in recombinant protein production could include heterologous expression in different expression systems. Several archaeal expression systems have been established and are being optimised towards higher expression yields using different inducible promoters (Albers et al., 2006; Contursi et al., 2003; Wagner et al., 2014). However, to the best of our knowledge, there are so far no successful reports of genes from *T. acidophilum* that have been expressed using archaeal expression systems.

In Chapter 3 the newly identified and characterised ManD in combination with three other enzymes was used to assemble a novel synthetic pathway for the conversion of

mannose into LA. The pathway design, assembly and operation was demonstrated in sequential cascade-like and one-pot reactions, in which pure mannose was successfully converted to LA. Furthermore, the applicability of the synthetic pathway in the conversion of sugars obtained from organic waste was demonstrated successfully using mannose obtained directly from dilute acid hydrolysis of SCG. This proof of concept represents the first report on the enzymatic conversion of mannose obtained from SCG into LA. Pathway optimisation and major bottlenecks were identified using variations of enzyme loading and substrate concentration, different ionic strength of the buffer, enzyme thermostability tests and flux analysis using real-time ^{13}C NMR analysis. This PhD project was conducted according to Macquarie University 3-years PhD candidature. Due to this time constrain many interesting research questions were not pursued or investigated further. In the following section, several future lines of research are presented that could help optimise the pathway and answer some of the questions sparked throughout this work.

The mannose conversion pathway presented here can be further optimised towards a 100% theoretical yield. Glyceraldehyde is the major by-product of the pathway, which is presumably also driving a reverse reaction of KDGA resulting in a steady increase of KDG over time. Elimination of these carbon losses requires a conversion of glyceraldehyde to pyruvate and its further conversion into LA. Guterl et al. presented a complete conversion of glucose to pyruvate via a np-ED pathway, including a recycling of glyceraldehyde to pyruvate, by implementing an engineered glyceraldehyde dehydrogenase and a promiscuous dihydroxyacid dehydratase (DHAD) (Guterl et al., 2012). The engineered glyceraldehyde dehydrogenase oxidises glyceraldehyde to glycerate under reduction of NAD^+ to NADH (Steffler et al., 2013). In its multi-functional role, DHAD converts gluconate to KDG but also glycerate to pyruvate (Carsten et al., 2015). The amount of NAD^+ produced is used in the conversion of pyruvate to either isobutanol or ethanol and renders the whole pathway cofactor-balanced. An implementation of the recycling module presented by Guterl et al. into the pathway developed in this work is feasible and presented in Figure 1. In principle, the recycling of glyceraldehyde enables a doubling of the present yield with the addition of only two enzymes.

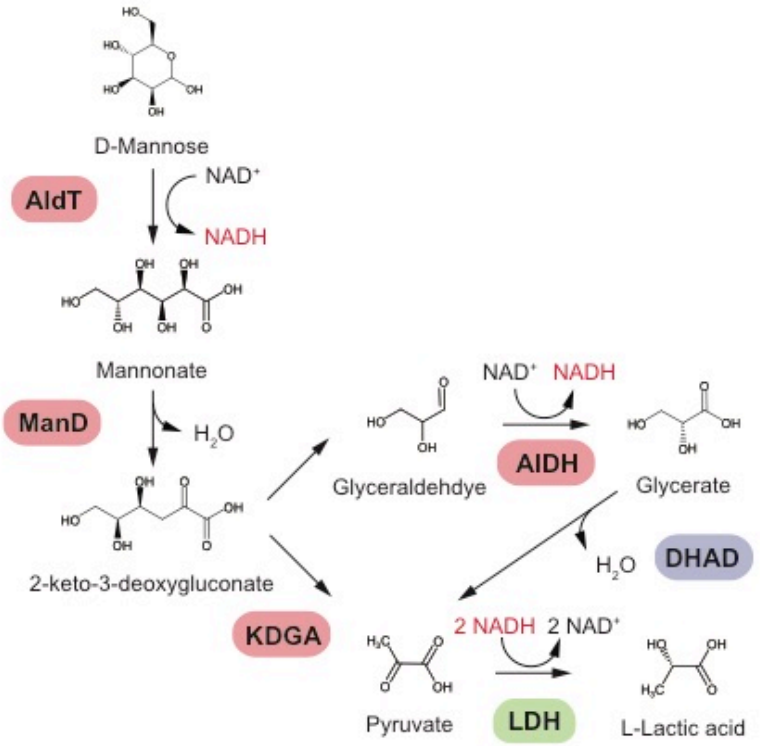


Figure 1 Extension of the pathway presented in this work to incorporate AIDH and DHAD enzymes to recycle glyceraldehyde to pyruvate as demonstrated by Guterl et al. (Guterl et al., 2012). AldT: aldohexose dehydrogenase, ManD: mannonate dehydratase, KDGA: 2-keto-3-deoxygluconate aldolase, AIDH: aldehyde dehydrogenase, DHAD: dihydroxyacid dehydratase, LDH: lactate dehydrogenase. Enzymes in red derived from *T. acidophilum*, in green: from *B. stearothermophilus*, in purple: from *S. solfataricus*.

A striking observation made in one-pot reactions was that an increase in the original reaction substrate (mannose) correlated with a decrease in the final product (LA) yield. Enzymes can be inhibited often by their own substrate, but none of the enzyme kinetics in our pathway indicated such an effect. Instead, product inhibition by a pathway intermediate could be responsible for this decreased in LA yield. This effect has been observed in other cell-free pathways, e.g. for the production of α-ketoglutarate, in which NADH showed an inhibitory effect on uronate dehydrogenase and KDG dehydratase (Beer et al., 2017). The inhibitory effect can be addressed by either avoiding accumulation of certain intermediates or by choosing homologous enzymes which are not affected by the inhibitory intermediate (Beer et al., 2017; Krutsakorn et al., 2013; Ye et al., 2012). Accumulation of intermediates can be avoided by adjusting enzyme amounts to improve forward flux, timely controlled feed-in of substrates and enzymes, or removal of products (Myung et al., 2014; Ye et al., 2012).

In our pathway, high levels of glyceraldehyde were present and NMR analysis also indicated an oxidation of glycerlaldehyde. Further experiments using NMR with intermediates in combination with different enzymes, could reveal unspecific reactions or oxidation products of intermediates.

The use of different enzyme concentrations and single enzyme kinetics identified ManD as the most limiting enzyme in the mannose conversion pathway. In general, ManD displayed low catalytic activity, low thermostability and low soluble yields upon expression. In future studies two possible approaches could be pursued to eliminate this weak link in the conversion process. The most straightforward approach would be the screening for superior ManD variants from other organisms, while a more complex possibility would be an improvement of the existing enzyme using different engineering techniques. Wichelecki et al. performed a screening of a large set of mannonate dehydratases from mostly bacterial backgrounds with the aim to uncover unusual substrate specificities among the members of the enolase superfamily (Wichelecki et al., 2014). 42 enzymes were expressed, purified and tested individually for their activity. Enzymes with different levels of catalytic efficiencies were identified. However enzymes from thermophilic backgrounds were not included in the study (Wichelecki et al., 2014).
Recently, Gagoski et al. presented an alternative approach to the classical laborious process of screening for functional enzymes. The combination of low-cost gene synthesis and cell-free protein synthesis (CFPS) allowed for a streamlined method that identified 30 putative thermostable endo-1,4-β-glucanases and xylanases (Gagoski et al., 2017). Research on the translation machinery in archaea could assist in the development and modulate archaeal extracts that can be used for CFPS from thermophilic origins (Becker et al., 2012). First expression yields with cell-free extracts from *Thermococcus kodakarensis* were quite low, but could be improved to the range of 100 µg/ml within 30 min after modulating cell extract components and the use of a mutant strain, with a disrupted heat shock regulator gene (Endoh et al., 2007, 2006).
A second possibility to improve the efficiency of the dehydration step in the pathway would be the improvement of the existing ManD by enzyme engineering. In general, enzyme engineering can be divided into three strategies: (a) directed evolution and combinatorial approaches; (b) rational design and (c) data-driven approaches. All engineering techniques have delivered their own success stories. Directed evolution for the improvement of thermotolerance has been used extensively, e.g. for fungal cellobiohydrolases by random mutagenesis and recombination of beneficial mutations

(Denard et al., 2015; Wu and Arnold, 2013). Again, the increased use of directed evolution approaches for enzyme improvement requires a development of capable high-throughput enzyme production and powerful screening techniques in order to exploit the full potential of large genetic libraries. In a rational engineering approach, beneficial mutations in flexible loops of an *E. coli* transketolase were identified using different approaches (Yu et al., 2017). Screening of over 40 mutant versions led to the identification of three variants, which showed superior performance over the wild-type enzyme. A combination of mutations was found to have a tripled half-life at 60°C and a 5-fold increase in specific activity at 65°C. More recently, the reconstruction of ancestral enzyme versions was shown to deliver enzyme variants with improved temperature and solvent tolerance (Gumulya et al., 2018). The improvement was demonstrated on animal cytochromes P450 and bacterial ketol-acid reductoisomerases, with a half-life of one hour at 66°C for the ancestral version of cytochrome P450 and a 3.5-fold increase in specific activity for the engineered *E. coli* ketol-acid reductoisomerase.

An important measure to improve the economic viability of the cell-free pathway is to lower the cost of enzyme by reducing its amount, production and purification cost. The lowest possible amount of enzyme to be used without sacrificing product yield can be identified by experimentally testing different enzyme loads. However, this task quickly becomes too laborious when the number of enzymes increase in the pathway. Kinetic modelling has been used to address this problem and to identify pathway bottlenecks, optimal enzyme ratios, and minimal enzyme loadings (Rollin et al., 2015; Ye et al., 2009; Zhong et al., 2017). Each reaction can be described as ordinary differential equation by incorporating empirical or literature kinetic parameters for each enzyme in the pathway. Solving these equations with algorithms (e.g. genetic algorithms) towards set goals such as high yield and/or low enzyme loads enables the identification of the minimal enzyme amounts that are required to meet a certain yield (Rollin et al., 2015). Kinetic modelling is also a powerful tool to define the right reaction parameters of multi-enzyme systems and can become an even more powerful tool if combined with the acquisition of reliable and high-density data from pathway reactions (Hold et al., 2016). A more thorough discussion about the use and limits of kinetic modelling can be found in the contribution to the review by Petroll et al., 2018 (Appendix I, section 9).

Another approach to reduce enzyme cost involves the use of crude cell lysates instead of purified enzymes (Bujara and Panke, 2012; Dudley et al., 2016). However, while

crude lysates seem to be cheaper at first, purified enzymes can potentially be reused and diminish the influence of unspecific side reactions caused by enzyme activity and endogenous enzymes present in crude lysates. Immobilisation of enzymes could help improve operation performance by enhancing enzyme stability, enabling reusability while at the same time simplifying the purification process. An overview of immobilisation techniques is given in Petroll et al. 2018. Recently, a facile immobilisation technique mediated by solid-binding peptides has been developed for enzymes used in biocatalysis (Care et al., 2017, 2016). This technology allows the direct immobilisation of enzymes onto low-cost industrial inorganic matrices like silica or zeolite. However, different immobilisation techniques would have to be considered, since genetic fusion of binding peptides or other modifications may result in insoluble expression or loss in enzyme activity. Immobilisation of all pathway enzymes would enable reuse of the enzymes and thus, help the cell-free process to compete with cell-based production with egards to long-term operation and high product titres.

The pathway in this work was validated using a readily available waste product as substrate into a platform chemical. Most cell-free processes use refined substrates, but more recently there is a trend towards the use of renewable resources as substrates for the production of chemicals (Cheng et al., 2018; Gao et al., 2015; Honda et al., 2017; Rollin et al., 2015). A more detailed consideration on the topic of renewable substrates in cell-free biocatalysis can be found in the contribution to the recently published review Petroll et al. 2018 (Appendix I, section 5).

In order to use lignocellulose as a substrate for bioconversion, polysaccharide structures are usually hydrolysed into simpler sugars. We used a standard dilute acid hydrolysis protocol to extract mannose sugar from SCG. Although dilute acid hydrolysis has emerged as an efficient way to hydrolyse lignocellulosic biomass, the technique has several drawbacks such as the release of potential toxic compounds for cells and enzymes and the relatively high cost, especially when compared to other physicochemical pre-treatment methods (Aden et al., 2002). Alternatively, biological hydrolysis methods could be integrated to achieve a more environmentally friendly process. Several fungal species have been shown to use SCG as growth substrate and effectively hydrolyse galactomannan (Machado et al., 2012). Although species of *Penicillium* and *Aspergillus* released up to 10 mg reducing sugars/g SCG, it is to question how much of the released mannose from SCG is consumed by the fungi itself, and whether fungal pretreatment and subsequent enzymatic conversion can be

combined. Potential secretion of fungal proteases into the hydrolysate could impair the stability of enzymes and lower the efficiency of a subsequent conversion process.
The conversion yield of our pathway with the SCG hydrolysate was in average 17.2 % lower compared to yields obtained from conversions with pure mannose. The difference in yield is likely to be caused by inhibitory compounds of acid-hydrolysed SCG, such as furfural, hydroxymethylfurfural, acetic acid and phenolic compounds (Mussatto et al., 2012). In order to match the yields of conversion with pure mannose identification of inhibitory compounds and their avoidance without a huge increase in process costs is needed. Polyphenolic compounds have a wide range of different applications in the food and pharmaceutical industries and can be extracted before fermentation or enzymatic conversion processes (Machado et al., 2012).

The pathway presented in this work was designed to specifically target mannose, but the pool of carbohydrates present in SCG could be further exploited. After mannose, galactose is usually found as the second most abundant sugar in SCG (Ballesteros et al., 2014; Jooste et al., 2013; Mussatto et al., 2011). Enzymes characterised in a variety of archaea including *S. solfataricus* would allow the extension of the cell-free reaction system by oxidation of glucose, galactose and arabinose. Namely the glucose dehydrogenase and gluconate dehydratase have been extensively studied and showed sufficient promiscuous activity towards galactose and galactonate, respectively (Kim and Lee, 2005; Lamble et al., 2004, 2003). In order to reduce enzyme costs, broadening the substrate promiscuity by engineering the aldohexose dehydrogenase and the already promiscuous DHAD could also enable conversion of all sugars found in SCG hydrolysates.

References

Aden, A., Ruth, M., Ibsen, K., Jechura, J., Neeves, K., Sheehan, J., Wallace, B., Montague, L., Slayton, A., Lukas, J., 2002. Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. No. NREL/TP-510-32438. Natl. Renew. energy lab golden co.

Albers, S., Jonuscheit, M., Dinkelaker, S., Urich, T., Kletzin, A., Tampe, R., Driessen, A.J.M., 2006. Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Appl. Environ. Microbiol. 72, 102–111.

Ballesteros, L.F., Teixeira, J.A., Mussatto, S.I., 2014. Chemical, functional, and structural properties of spent coffee grounds and coffee silverskin. Food Bioprocess Technol. 7, 3493–3503.

Becker, T., Franckenberg, S., Wickles, S., Shoemaker, C.J., Anger, A.M., Armache, J.P., Sieber, H., Ungewickell, C., Berninghausen, O., Daberkow, I., Karcher, A., Thomm, M., Hopfner, K.P., Green, R., Beckmann, R., 2012. Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. Nature 482, 501–506.

Beer, B., Pick, A., Sieber, V., 2017. *In vitro* metabolic engineering for the production of α-ketoglutarate. Metab. Eng. 40, 5–13.

Budgen, N., Danson, M.J., 1986. Metabolism of glucose via a modified Entner-Doudoroff pathway in the thermoacidophilic archaebacterium *Thermoplasma acidophilum*. FEBS Lett. 196, 207–210.

Bujara, M., Panke, S., 2012. In silico assessment of cell-free systems. Biotechnol. Bioeng. 109, 2620–2629.

Care, A., Bergquist, P.L., Sunna, A., 2016. Solid-binding peptides: immobilisation strategies for extremophile biocatalysis in biotechnology, in: Biotechnology of Extremophiles: Advances and Challenges. Springer International Publishing, pp. 637–674.

Care, A., Petroll, K., Gibson, E.S.Y., Bergquist, P.L., Sunna, A., 2017. Solid-binding peptides for immobilisation of thermostable enzymes to hydrolyse biomass polysaccharides. Biotechnol. Biofuels 10, 1–16.

Carsten, J.M., Schmidt, A., Sieber, V., 2015. Characterisation of recombinantly expressed dihydroxy-acid dehydratase from *Sulfobus solfataricus* - A key enzyme for the conversion of carbohydrates into chemicals. J. Biotechnol. 211, 31–41.

Cheng, K., Zheng, W., Chen, H., Job Zhang, Y.-H.P., 2018. Upgrade of wood sugar D-xylose to a value-added nutraceutical by *in vitro* metabolic engineering. Metab. Eng. 52, 1–8.

Contursi, P., Cannio, R., Prato, S., Fiorentino, G., Rossi, M., Bartolucci, S., 2003. Development of a genetic system for hyperthermophilic Archaea : expression of a moderate thermophilic bacterial alcohol dehydrogenase gene in *Sulfolobus solfataricus*. FEMS Microbiol. Lett. 218, 115–120.

Denard, C.A., Ren, H., Zhao, H., 2015. Improving and repurposing biocatalysts via directed evolution. Curr. Opin. Chem. Biol.

Dudley, Q.M., Anderson, K.C., Jewett, M.C., 2016. Cell-Free Mixing of *Escherichia coli* crude extracts to prototype and rationally engineer high-titer mevalonate synthesis. ACS Synth. Biol. 5, 1578–1588.

Endoh, T., Kanai, T., Imanaka, T., 2007. A highly productive system for cell-free protein synthesis using a lysate of the hyperthermophilic archaeon, *Thermococcus kodakaraensis*. Appl. Microbiol. Biotechnol. 74, 1153–1161.

Endoh, T., Kanai, T., Sato, Y.T., Liu, D. V., Yoshikawa, K., Atomi, H., Imanaka, T.,

2006. Cell-free protein synthesis at high temperatures using the lysate of a hyperthermophile. J. Biotechnol. 126, 186–195.

Gagoski, D., Shi, Z., Keld, L., Vickers, C.E., Mahler, S., Speight, R., Johnston, W.A., Alexandrov, K., 2017. Cell-free pipeline for discovery of thermotolerant xylanases and *endo*-1,4-β-glucanases. J. Biotechnol. 259, 191–198.

Gao, C., Li, Z., Zhang, L., Wang, C., Li, K., Ma, C., Xu, P., 2015. An artificial enzymatic reaction cascade for a cell-free bio-system based on glycerol. Green Chem. 17, 804–807.

Gumulya, Y., Baek, J.-M., Wun, S.-J., Thomson, R.E.S., Harris, K.L., Hunter, D.J.B., Behrendorff, J.B.Y.H., Kulig, J., Zheng, S., Wu, X., Wu, B., Stok, J.E., De Voss, J.J., Schenk, G., Jurva, U., Andersson, S., Isin, E.M., Bodén, M., Guddat, L., Gillam, E.M.J., 2018. Engineering highly functional thermostable proteins using ancestral sequence reconstruction. Nat. Catal. 1, 878–888.

Guterl, J.K., Garbe, D., Carsten, J., Steffler, F., Sommer, B., Reiße, S., Philipp, A., Haack, M., Rühmann, B., Koltermann, A., Kettling, U., Brück, T., Sieber, V., 2012. Cell-free metabolic engineering: production of chemicals by minimized reaction cascades. ChemSusChem 5, 2165–2172.

Hold, C., Billerbeck, S., Panke, S., 2016. Forward design of a complex enzyme cascade reaction. Nat. Commun. 7, 1–8.

Honda, K., Kimura, K., Ninh, P.H., Taniguchi, H., Okano, K., Ohtake, H., 2017. *In vitro* bioconversion of chitin to pyruvate with thermophilic enzymes. J. Biosci. Bioeng. 124, 296–301.

Jooste, T., García-Aparicio, M.P., Brienzo, M., Van Zyl, W.H., Görgens, J.F., 2013. Enzymatic hydrolysis of spent coffee ground. Appl. Biochem. Biotechnol. 169, 2248–2262.

Kim, S., Lee, S.B., 2005. Identification and characterisation of *Sulfolobus solfataricus* D-gluconate dehydratase: a key enzyme in the non-phosphorylated Entner-Doudoroff pathway. Biochem. J. 387, 271–80.

Kim, S.M., Paek, K.H., Lee, S.B., 2012. Characterisation of NADP+-specific L-rhamnose dehydrogenase from the thermoacidophilic Archaeon *Thermoplasma acidophilum*. Extremophiles 16, 447–454.

Krutsakorn, B., Honda, K., Ye, X., Imagawa, T., Bei, X., Okano, K., Ohtake, H., 2013. *In vitro* production of n-butanol from glucose. Metab. Eng. 20, 84–91.

Lamble, H.J., Heyer, N.I., Bull, S.D., Hough, D.W., Danson, M.J., 2003. Metabolic pathway promiscuity in the archaeon *Sulfolobus solfataricus* revealed by studies on glucose dehydrogenase and 2-keto-3-deoxygluconate aldolase. J. Biol. Chem. 278, 34066–34072.

Lamble, H.J., Milburn, C.C., Taylor, G.L., Hough, D.W., Danson, M.J., 2004. Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in *Sulfolobus solfataricus*. FEBS Lett. 576, 133–136.

Machado, E.M.S., Rodriguez-Jasso, R.M., Teixeira, J.A., Mussatto, S.I., 2012. Growth of fungal strains on coffee industry residues with removal of polyphenolic compounds. Biochem. Eng. J. 60, 87–90.

Mussatto, S.I., Carneiro, L.M., Silva, J.P.A., Roberto, I.C., Teixeira, J.A., 2011. A study on chemical constituents and sugars extraction from spent coffee grounds. Carbohydr. Polym. 83, 368–374.

Mussatto, S.I., Machado, E.M.S., Carneiro, L.M., Teixeira, J.A., 2012. Sugars metabolism and ethanol production by different yeast strains from coffee industry wastes hydrolysates. Appl. Energy 92, 763–768.

Myung, S., Rollin, J., You, C., Sun, F., Chandrayan, S., Adams, M.W.W., Zhang, Y.P., 2014. *in vitro* metabolic engineering of hydrogen production at theoretical yield from sucrose. Metab. Eng. 24, 70–77.

Nishiya, Y., Tamura, N., Tamura, T., 2014. Analysis of bacterial glucose dehydrogenase

homologs from thermoacidophilic archaeon *Thermoplasma acidophilum*: finding and characterisation of aldohexose dehydrogenase. Biosci. Biotechnol. Biochem. 68, 2451–2456.

Petroll, K., Kopp, D., Care, A., Bergquist, P.L., Sunna, A., 2018. Tools and strategies for constructing cell-free enzyme pathways. Biotechnol. Adv. in press.

Pick, A., Beer, B., Hemmi, R., Momma, R., Schmid, J., Miyamoto, K., Sieber, V., 2016. Identification and characterisation of two new 5-keto-4-deoxy-D-glucarate dehydratases/decarboxylases. BMC Biotechnol. 16, 1–10.

Rollin, J.A., Martin del Campo, J., Myung, S., Sun, F., You, C., Bakovic, A., Castro, R., Chandrayan, S.K., Wu, C.-H., Adams, M.W.W., Senger, R.S., Zhang, Y.-H.P., 2015. High-yield hydrogen production from biomass by *in vitro* metabolic engineering: mixed sugars coutilization and kinetic modeling. Proc. Natl. Acad. Sci. 112, 4964–4969.

Steffler, F., Guterl, J.K., Sieber, V., 2013. Improvement of thermostable aldehyde dehydrogenase by directed evolution for application in synthetic cascade biomanufacturing. Enzyme Microb. Technol. 53, 307–314.

Sun, N., Pan, C., Nickell, S., Mann, M., Baumeister, W., 2010. Quantitative proteome and transcriptome analysis of the archaeon *Thermoplasma acidophilum* cultured under aerobic and anaerobic conditions. J. Proteome Res. 9, 4839–4850.

Wagner, M., Wagner, A., Kort, J.C., Ghosh, A., Rauch, B., Siebers, B., Albers, S.-V., 2014. Investigation of the *malE* promoter and *malR*, a positive regulator of the maltose regulon , for an Improved expression system in *Sulfolobus acidocaldarius*. Appl. Environ. Microbiol. 80, 1072–1081.

Wichelecki, D.J., Balthazor, B.M., Chau, A.C., Vetting, M.W., Fedorov, A.A., Fedorov, E. V., Lukk, T., Patskovsky, Y. V., Stead, M.B., Hillerich, B.S., Seidel, R.D., Almo, S.C., Gerlt, J.A., 2014. Discovery of function in the enolase superfamily: D-mannonate and D-gluconate dehydratases in the D-mannonate dehydratase subgroup. Biochemistry 53, 2722–2731.

Wu, I., Arnold, F.H., 2013. Engineered thermostable fungal Cel6A and Cel7A cellobiohydrolases hydrolyze cellulose efficiently at elevated temperatures. Biotechnol. Bioeng. 110, 1874–1883.

Ye, X., Honda, K., Sakai, T., Okano, K., Omasa, T., Hirota, R., Kuroda, A., Ohtake, H., 2012. Synthetic metabolic engineering-a novel, simple technology for designing a chimeric metabolic pathway. Microb. Cell Fact. 11, 120.

Ye, X., Wang, Y., Hopkins, R.C., Adams, M.W.W., Evans, B.R., Mielenz, J.R., Zhang, Y.H.P., 2009. Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. ChemSusChem 2, 149–152.

Yu, H., Yan, Y., Zhang, C., Dalby, P.A., 2017. Two strategies to engineer flexible loops for improved enzyme thermostability. Sci. Rep. 7, 1–15.

Zhong, C., Wei, P., Zhang, Y.H.P., 2017. A kinetic model of one-pot rapid biotransformation of cellobiose from sucrose catalyzed by three thermophilic enzymes. Chem. Eng. Sci. 161, 159–166.

Appendix I

ARTICLE IN PRESS

Biotechnology Advances xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Tools and strategies for constructing cell-free enzyme pathways

Kerstin Petroll[a], Dominik Kopp[a], Andrew Care[a,b], Peter L. Bergquist[a,c], Anwar Sunna[a,b,*]

[a] Department of Molecular Sciences, Macquarie University, Sydney, Australia
[b] Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, Australia
[c] Department of Molecular Medicine & Pathology, University of Auckland, Auckland, New Zealand

ARTICLE INFO

Keywords:
Synthetic biology
In vitro biocatalysis
Cell-free enzyme pathways
Multi-enzyme cascades
Pathway design
Enzyme immobilisation
Kinetic modelling
Metabolite analysis

ABSTRACT

Single enzyme systems or engineered microbial hosts have been used for decades but the notion of assembling multiple enzymes into cell-free synthetic pathways is a relatively new development. The extensive possibilities that stem from this synthetic concept makes it a fast growing and potentially high impact field for biomanufacturing fine and platform chemicals, pharmaceuticals and biofuels. However, the translation of individual single enzymatic reactions into cell-free multi-enzyme pathways is not trivial. In reality, the kinetics of an enzyme pathway can be very inadequate and the production of multiple enzymes can impose a great burden on the economics of the process. We examine here strategies for designing synthetic pathways and draw attention to the requirements of substrates, enzymes and cofactor regeneration systems for improving the effectiveness and sustainability of cell-free biocatalysis. In addition, we comment on methods for the immobilisation of members of a multi-enzyme pathway to enhance the viability of the system. Finally, we focus on the recent development of integrative tools such as *in silico* pathway modelling and high throughput flux analysis with the aim of reinforcing their indispensable role in the future of cell-free biocatalytic pathways for biomanufacturing.

1. Introduction

Enzyme technology has become a prominent field in synthetic biology and biomanufacturing during the 20th and 21st centuries (Bornscheuer et al., 2012; Reetz, 2013; Sheldon and Brady, 2018). Advances in molecular biological, genetical and computational tools have promoted the discovery and design of new enzymes, the improvement of known enzymes and facilitated enzyme production along with reduced production costs. These achievements have caused the classical concept of biocatalysis, the use and study of single enzyme reactions, to be expanded to the use of multiple enzymes that are combined rationally into natural and non-natural synthetic biocatalytic pathways (Bornscheuer, 2018; Poppe and Vértessy, 2018). The field of enzyme technology and biocatalysis has become a primary focus for the 'green' organic synthesis of chemicals due to the increased demand for sustainable and environmentally-friendly biomanufacturing (Bornscheuer et al., 2012; France et al., 2017; Sperl and Sieber, 2018; Swartz, 2006, 2012).

Enzymes are outstanding catalysts because of their high specificity and stereo-, regio- and chemo-selectivity, which can translate into high product yields and low generation of undesirable side products (France et al., 2017). However, due to their biological origins, many enzymes are tightly regulated allosterically and often suffer from substrate or product inhibition at concentrations higher than in their intracellular environment. Often, they are inherently unstable under industrial conditions, making them both short-lived and economically non-viable. Industry has relied heavily on chemical catalysts as a consequence since they are cheaper, more robust and can be evaluated rapidly. Nevertheless, when high stereochemistry, mild reaction conditions and environmentally-friendly operations are desired, the application of enzymes is preferred over conventional chemical catalysts (Bornscheuer, 2016; Keasling et al., 2012) particularly in the case of drug synthesis where there is a demand for high optical purity. As a result, single enzyme reactions increasingly complement chemical catalysis for enantioselective steps. The potential and limitations of this challenging field was reviewed recently (Rudroff et al., 2018).

Several strategies have been developed for the introduction of enzymes into industrial applications and to address the challenges outlined. For example, to circumvent enzyme inhibition from high substrate or product concentrations, substrates can be continuously fed into the reactor (Ye et al., 2012 , 2013), products can be removed *in-situ* or enzymes can be engineered to accept higher substrate or product concentrations (Bornscheuer et al., 2012; Desantis et al., 2003; Gruber et al., 2017; Savile et al., 2010; Wang et al., 2017b). Similarly, the stability of enzymes can be improved through protein engineering using rational design, directed evolution or ancestral protein reconstruction

* Corresponding author at: Department of Molecular Sciences, Macquarie University, North Ryde, NSW 2109, Australia.
E-mail address: anwar.sunna@mq.edu.au (A. Sunna).

https://doi.org/10.1016/j.biotechadv.2018.11.007
Received 3 September 2018; Received in revised form 22 October 2018; Accepted 20 November 2018
0734-9750/ © 2018 Elsevier Inc. All rights reserved.

Please cite this article as: Petroll, K., Biotechnology Advances, https://doi.org/10.1016/j.biotechadv.2018.11.007

(Goldsmith and Tawfik, 2017; Gumulya and Gillam, 2017; Risso et al., 2018). Where possible, it may be feasible to replace less stable enzymes with thermostable counterparts which usually are intrinsically more robust (see later).

Industrial biomanufacturing using single enzyme reactions commonly has been performed within living cells (*in vivo*), where enzymes are produced continuously and cell walls or internal compartments protect the enzymes from degradation (Bornscheuer et al., 2012; France et al., 2017). However, operating multi-enzyme pathways *in vivo* is challenging and often results in low yields due to the metabolic complexity of the host where substrates and products may be diverted into other cellular metabolic pathways or are toxic to the cells (Guo et al., 2017). The genetic instability or loss of recombinant plasmids that code for the pathway enzymes in genetically-engineered host cells can impair continuous cell-based biocatalysis (Friehs, 2004; Huang et al., 2012). As a consequence, multi-enzyme pathways have been assembled increasingly in cell-free systems allowing the investigator to 'pick, mix and test' an enzyme system without the need for complex genetic modifications of the host organism and without interference from intracellular processes. Accordingly, this approach provides high flexibility and control for optimising and debugging the system (Cherubini, 2010; Schmidt-Dannert and Lopez-Gallego, 2016; Taniguchi et al., 2017; Wilding et al., 2018; Zhang et al., 2017).

2. Cell-free enzyme synthesis and applications

The use of multiple enzymes in a cell-free context is not a new concept and has been practised in the form of conventional cell-free protein synthesis (CFPS) by researchers since the 1960's (Devries and Zubay, 1967; Nirenberg and Matthaei, 1961; Zubay, 1973). However, cell-free biocatalysis has become a critical procedure recently with extensive possibilities when linked to synthetic biology. Cell-free expression of proteins occurs without the presence of cells but all of the components necessary for RNA translation *in vitro* are present. The history and advances in the field have been subjected to many reviews and the interested reader is referred to them and they will not be discussed here (for example, (Chong, 2014; Jewett et al., 2008; Karim and Jewett, 2016; Rosenblum and Cooperman, 2014; Swartz, 2006, 2012; Whittaker, 2013; Yang et al., 2012). Generally, CFPS is not considered to be suitable for large-scale protein production but the work of Swartz and colleagues and the company Sutro Biopharma (www.sutrobio.com) have shown that high productivity could be achieved at the hundreds of litres scale with selected proteins. The original system was simply a cell lysate with the debris removed but was problematic as it soon ran out of energy (*e.g.* ATP) and amino acids, and was subjected to macromolecular degradation by native proteases in the lysate. The Swartz laboratory embarked upon refinements of the lysate system to exploit biological processes without intact cells. This approach overcame problems such as toxicity, transport limitation, cell membrane effects and provided alternative energy sources (summarised in Swartz, 2006, 2012). Later a system was developed by Shimizu et al. (2005) which used 32 purified factors for initiation, elongation and termination of protein synthesis and allowed the production and purification of enzymatically-active proteins. The system was called PURE (Protein Synthesis Using Recombinant Elements) and is available commercially as PURExpress (New England Biolabs). This system is suitable for small-scale laboratory experiments but the cost of the components ruled it out for large scale applications.

3. Synthetic biology and cell-free biocatalysis

Recently, the principles of cell-free protein synthesis have been expanded by Karim and Jewett (2016). The authors proposed a framework for biosynthetic pathway development using cell-free protein synthesis driven metabolic engineering (CFPS-ME). CFPS-ME enabled the rapid optimisation of a 17-enzyme pathway for the production of n-butanol from glucose using heterologous enzymes from *Clostridium beijinkerii*, *Cs. acetobutylicum* and *Treponema denticola* to supplement the native *E. coli* enzymes (Karim and Jewett, 2016). This study demonstrated that the cell-free protein expression translated well into the *in vivo* system and that CFPS-ME could be used to prototype biocatalytic pathways for their reconstitution *in vivo* more rapidly than with conventional metabolic engineering.

Recent examples of other synthetic cell-free pathways based on purified enzymes or culture lysates that utilise between 3 and 30 enzymes are given in Table 1. Most of these pathways start from simple and renewable precursors such as glucose or starch and have proved to be very effective as they reached nearly 100% of the theoretical conversion yield. Many of the intermediates or products are central carbon metabolites (*e.g.* pyruvate), which are likely to be diverted into a cell's own metabolism rather than the synthetic pathway if the reaction was performed *in vivo*. It should be noted that the process temperature is often higher than 37 °C and many products at the concentrations produced would be toxic to cells such as terpenes (Korman et al., 2017), and isobutanol (Guterl et al., 2012) and would have been unworkable if carried out *in vivo*.

Cell-free biocatalysis is an attractive partner for synthetic biology in that many of the restraints exercised by cells are eliminated and the concept of cell-free metabolic engineering becomes a reality as well as allowing more rapid screening of protein pathways. It must be stressed that it is vital to understand the parameters involved and that the setting up of functional biocatalytic pathways using genes from different organisms is not a trivial exercise. We outline some of the more important issues and discuss the latest advances, current limitations and future potential of cell-free multi-enzyme biocatalysis with particular focus on synthetic pathway design strategies, renewable substrates available, enzyme requirements, and cofactors regeneration systems. In addition, we comment on techniques to immobilise whole enzyme pathways to allow repeated biocatalyst use with consequent reduction in costs. Unlike other recent reviews in this area, we emphasise integrative tools like *in silico* kinetic modelling and high-throughput flux analysis to promote the potential of cell-free synthetic enzyme pathways. This review is not intended to be comprehensive and the interested reader is referred to other recent excellent reviews on multi-enzyme biocatalysis (Sperl and Sieber, 2018), (Sheldon and Brady, 2018).

4. Pathway design strategies

Traditionally, the design of cell-free biocatalytic pathways has been based on existing metabolic pathways that were translated into cell-free biocatalytic processes. Modern synthetic biological techniques now allow the design of synthetic pathways in which different natural and/or *in silico*-predicted pathways or their components are selected, integrated, modified, and assembled into a customised pathway (Fig. 1).

4.1. Synthetic pathways derived from natural metabolic pathways

Synthetic enzyme pathways that are modified versions or combinations of existing metabolic pathways often are capable of producing similar or higher yields than their natural counterparts but with reduced enzyme demands and cofactor components, thus representing a more economically-viable approach to biocatalysis. For example, starch can be converted into the glycolytic high-energy intermediate fructose-1-6-bisphosphate (FBP) in a completely adenosine triphosphate (ATP)-independent and less expensive manner by employing an α-glucan phosphorylase. In this way, starch was first digested to glucose-1-phosphate (G1P), consuming inorganic phosphate instead of ATP and the G1P was then converted to fructose-1-phosphate (F1P). By introducing a pyrophosphate-dependent kinase, F1P was phosphorylated to fructose-1-6-bisphosphate (FBP) consuming pyrophosphate instead of ATP (Fig. 2A and Table 1) (Wang et al., 2017b).

In a recent extensive report, the cell-free conversion of chitin into

ARTICLE IN PRESS

Table 1
Recently described cell-free enzyme pathways for the production of biofuels or high value chemicals from inexpensive and renewable substrates (categorised as monomeric, dimeric and polymeric substrates).

Substrate	Product	Enzymes[a]	Reaction temperature (°C)	Reaction time (h)	Production yield (% mol/mol)	Production titre (mM)	References
Monomeric							
CO_2	Malate (CO_2 fixation)	17	30	1.5	CO_2 fixation rate: 12 μM min^{-1}[b]	0.54	(Schwander et al., 2016)
Glycerol	Pyruvate	3	50	24	93.2	9.3	(Gao et al., 2015)
	R-Acetoin	5	50	24	85.5	4.4	
Glucose	Hydrogen	14	37	NA	4.8 mM h^{-1} [b]	NA[c]	(Lu et al., 2015)
	Ethanol	6	50	19	53	28.7	(Guterl et al., 2012)
	Isobutanol	9		23	57	10.3	
	n-Butanol	16	50	7	82	3.5	(Krutsakorn et al., 2013)
	L-Lactate	10	50	10	100	12	(Ye et al., 2012)
	L-Lactate	10	70	3	100	2.2	(Ninh et al., 2015)
	L-Lactate	5	50	8	90	48.2	(Xie et al., 2018)
	Malate	10	50	4	60	2.8	(Ye et al., 2013)
	Mevalonate	3	37	20	46.9	119	(Dudley et al., 2016)
	Limonene	27	25	168	88.3	91.7	(Korman et al., 2017)
	a/β-pinene				103.9	106.4	
	Fatty acids: decanoic, lauric, myristicand palmitic acids	30	37	4	9.3[d]	30.04 mg L^{-1} [d]	(Liu et al., 2017)
	Polyhydroxybutyrate	19	RT[e]	30	94	57[f]	(Opgenorth et al., 2016)
				50	86	93.8	
Glucuronate	α-Ketoglutarate	5	25	5	92	51.5	(Beer et al., 2017)
Xylose	Hydrogen	13	50	30	96	NA	(Martín del Campo et al., 2013)
Dimeric							
Lactose (Racemic)	Pyruvate	2–3	30	7	97.9	20.3	(Li et al., 2017)
	Acetoin			5	92.7	9.8	
	Acetaldehyde			4	66.9	13.6	
Cellobiose	Hydrogen	13	32	150	93.3	NA	(Ye et al., 2009)
Sucrose	Cellobiose	3	60	10	62.3	62.3	(Zhong et al., 2017a)
	Myo-inositol	4	50 and 70	108	98	48	(Zhong et al., 2017b)
Polymeric							
Cellopentaose	Hydrogen	14	32	150	68	NA	(Ye et al., 2009)
Corn stover [g]	Hydrogen	14	40	78	100	2.3 mM·h^{-1} [b]	(Rollin et al., 2015)
Amylose	*Myo*-inositol	4	90	6	96	81	(Fujisawa et al., 2017)
Soluble starch		5	90	4	73	58	
Potato		5	90	6	139 and 146	57 and 61	
Maltodextrin	*Myo*-inositol	6	70	48	98	54.9	(You et al., 2017)
		5	70	48	76	527.3	
Maltodextrin	Fructose-1-6-bisphosphate	4	70	7	22.5	125	(Wang et al., 2017b)
Chitin	Pyruvate	14	70	5	42	2.1	(Honda et al., 2017)

[a] Number of enzymes in the pathway.
[b] Expressed as pathway productivity.
[c] NA; not available.
[d] Sum of all fatty acids.
[e] RT; room temperature.
[f] Monomer equivalents.
[g] Contains about 3% cellulose and 21% xylan (Aden et al., 2002).

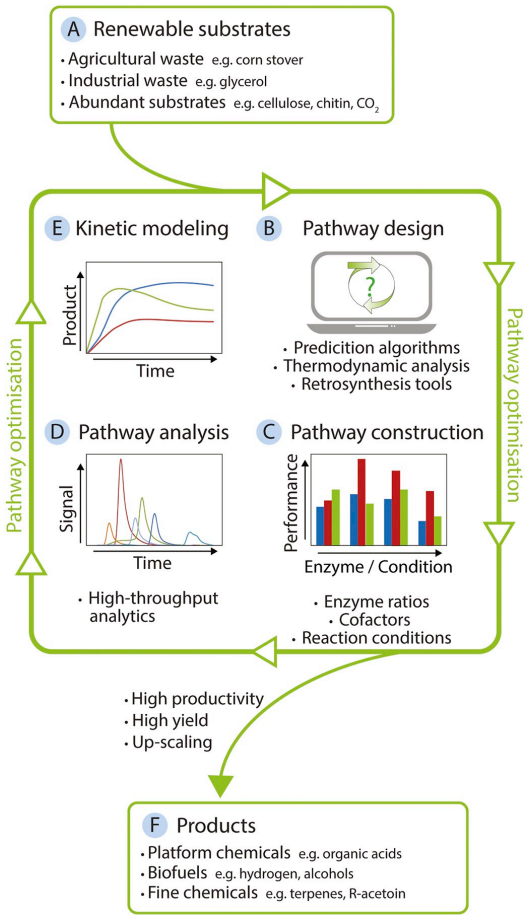


Fig. 1. Diagram showing the process used to design, construct, analyse and optimise cell-free biocatalytic pathways for the 'green' production of biobased products. A) Substrates from renewable resources are selected to produce a desired product. B) A pathway is then designed based on the substrate, product and the enzymatic reactions required to achieve effective conversion. Pathways design can be derived or optimised *in silico*, for example by using computational 'retrosynthesis' based algorithms that can determine the shortest route, lowest number of reactions and thermodynamic viability. C) The required enzymes are produced in recombinant hosts (*e.g.*, *E. coli*), extracted and assembled into stepwise or one-pot pathway reactions. D) and E) To obtain the maximum product yield, the pathway can be optimised further using high-throughput flux analysis and kinetic models that determine and improve the ideal enzyme ratios, cofactors and reaction conditions. This approach also can help to test and identify enzymes within a pathway that limit high product yields. Accordingly, limiting enzymes can be improved by protein engineering or exchanged with alternative enzymes. F) Iterative cycles of design, construction, analysis and modelling increase yields and productivity in order to produce valuable bulk or speciality chemicals (*e.g.*, fuels and organic acids) at a large scale.

the metabolite, pyruvate, was accomplished by integrating various components of three metabolic pathways based on enzymes derived from hyperthermophilic archaea and bacteria (Fig. 2B). The pathways included: (i) a chitin degradation pathway; (ii) the classical bacterial/eukaryotic Embden-Myerhof (EM) glycolytic pathway; and (iii) the archaeal modified EM pathway (Honda et al., 2017). A glucosamine kinase that originally linked the chitin degradation pathway with the glycolytic pathway was not well characterised yet and therefore, a promiscuous, well known glucose kinase was used instead. Additionally, merging the classical and the archaeal-modified EM pathway into a 'chimeric EM pathway' allowed *in situ* ATP recycling and reduced enzyme usage. After an initial production of 0.62 mM pyruvate in 5 h, the system was improved further by incorporation of a thermophilic adenylate kinase and a thermophilic polyphosphate kinase, which reduced ATP degradation by the host ATP-degrading enzyme(s). A 3.3-fold increase in the final product titre to 2.1 mM was obtained with the optimised pathway (encircled, Fig. 2B).

In another elegant example, the cell-free conversion of glucose to isobutanol and ethanol was established through the assembly of the archaeal non-phosphorylative Entner-Doudoroff (ED) pathway and the isobutanol and ethanol pathways (Fig. 2C; Guterl et al. (2012)). The original pathways were optimised for cofactor and enzyme usage and pathway yields. First, the selection of a dihydroxy acid dehydratase (DHAD) with substrate promiscuity towards three intermediates of the pathways reduced the total amount of enzyme needed. DHAD promiscuity also enabled the by-product glycerate to be feed-back into the pathway ('salvaged') and recycled as pyruvate without the use of additional ATP, unlike the natural non-phosphorylative ED pathway. Further, a highly specific glyceraldehyde dehydrogenase (AlDH) from *Thermoplasma acidophilum* was selected for the oxidation of glyceraldehyde to glycerate to prevent the oxidation of acetaldehyde or isobutyraldehyde produced downstream into undesirable intermediates. The activity of AlDH with NAD^+ was improved further by directed evolution-guided enzyme engineering (Steffler et al., 2013).

The cell-free production of fatty acids from glucose was achieved through a synthetic reconstruction of the classical EM glycolytic pathway and the fatty acid pathway (Fig. 2D) (Liu et al., 2017). These authors demonstrated that the synthetic assembly of enzymes from both *E. coli* and *Enterococcus faecalis* and the replacement of the NAD^+-dependent glyceraldehyde 3-phosphate dehydrogenase (GAPDH) from *E. coli* with a $NADP^+$-dependent GAPDH from *Clostridium acetobutylicum* increased NADPH availability and dramatically improved fatty acid production.

The cell-free conversion of glucose to the bioplastic polyhydroxybutyrate (PHB) was achieved by assembling parts of the glycolytic pathway, the pentose phosphate pathway, the *Bifidobacterium* shunt and the PHB pathway (Fig. 2E; (Opgenorth et al., 2016). The key to a high pathway flux was the integration of two synthetic purge valve systems which balanced the NADPH concentration *in situ* and the addition of a salvaging pathway to recycle the erythrose-4-phosphate which was produced from fructose-6-phosphate due to the activity of a promiscuous enzyme that was required downstream. The purge valve systems consisted of a wild type NADPH-dependent enzyme, its engineered NADH-dependent analogue and a NADH oxidase (Nox) (Opgenorth et al., 2014). When NADPH is oversupplied due to flux imbalances, wild type enzymes are inhibited while the engineered enzymes produce the substrate for the ensuing reaction, reducing NAD^+ instead of $NADP^+$. The reduced NADH equivalent is then regenerated to NAD^+ by Nox and available for subsequent reactions. The NADPH oversupply eventually is reversed through downstream reactions and the purge valve systems switches back to the normal situation with the use of the wildtype NADPH-dependent enzyme.

4.2. De novo synthetic enzyme pathways

Pathways may be designed completely *de novo* in cases where natural or nature-derived synthetic pathways still fail to allow synthesis of the desired target compound in an efficient manner. This *in silico* approach is constrained mainly by the availability of enzymes, their kinetic properties and the reactions they catalyse, the thermodynamics of the process and as well, by inadequate or incorrect annotation of sequences with no proof of specific enzyme activity as claimed.

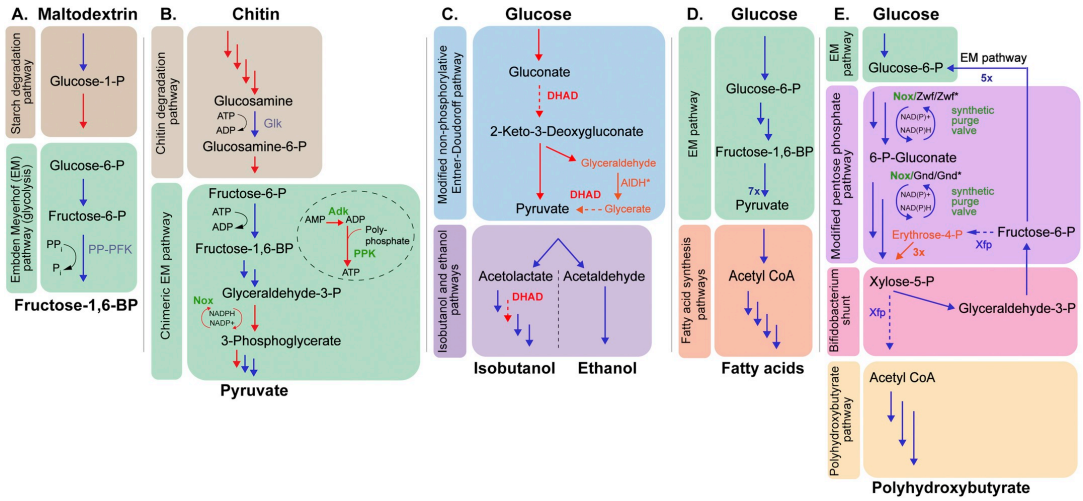


Fig. 2. Schematic illustration of five recently described synthetic pathways based on natural metabolic pathways (A–E). P, phosphate; PP-PFK, pyrophosphate-dependent phosphofructokinase; Glk, glucokinase; Adk, adenylate kinase; PKK, polyphosphate kinase; Nox, NADH oxidase; DHAD, dihydroxy acid dehydratase; AlDH, glyceraldehyde dehydrogenase; Zwf, glucose-6-phosphate dehydrogenase; Gnd, 6-phosphogluconate dehydrogenase and Xfp, xyululose-5-phosphate phosphoketolase. Blue- and red-coloured arrows indicate that the enzymes derived from bacteria or archaea, respectively. Dashed arrows indicate promiscuous enzymes. The orange-coloured arrows indicate 'salvaging' enzymes that were integrated for feed-back of undesired by-products (also highlighted in orange) into the pathway. Enzymes that were added for recycling of cofactors are highlighted in green. Enzymes that were engineered are indicated with *. Further modifications to the natural pathways, pathway descriptions and references are described in the main text. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Retrosynthetic analysis that traditionally was introduced for designing *de novo* synthetic pathways in organic chemistry now has been adopted in synthetic biocatalysis in order to guide *de novo* enzyme pathway design effectively (de Souza et al., 2017; Green and Turner, 2016; Honig et al., 2017). In general, retrosynthetic analysis describes the process whereby the target molecule is disconnected systematically into a 'simpler' precursor until an easily-accessible and low cost starting material is identified (Corey, 1988). This same concept applies to biocatalysis ('biocatalytic retrosynthesis') where a pathway is suggested for the interconversion of a target product into available start materials (*e.g.* supplied by the cell's metabolism) based on enzyme-catalysed reactions and feasible thermodynamics (Turner and O'Reilly, 2013). However, with the increasing diversity of enzyme-catalysed reactions as a result of the identification, design or engineering of novel enzymes, this approach is difficult to master intuitively. Accordingly, computational algorithms have been developed that encode enzymatic reactions into reaction rules and build reaction networks *in silico* based on retrosynthetic analysis (Delépine et al., 2017; Hadadi and Hatzimanikatis, 2015; Kumar et al., 2018). Reaction rules define the change of atoms in the substrates and products and are usually computed automatically using metabolic databases such as KEGG (Goto et al., 1997) or the Enzyme Commission (EC) nomenclature (Wang et al., 2017a). The retrosynthetic networks then are evaluated or ranked based on criteria such as enzyme availability, compatibility of selected enzymes (similar pH and temperature optima), compounds' toxicities, thermodynamics and maximum theoretical yields (Delépine et al., 2017; Feher et al., 2014; Kumar et al., 2018).

This computational approach of 'biocatalytic retrosynthesis' can reveal extensive possibilities in synthetic biocatalysis by offering novel synthetic routes based on existing or hypothetical enzymes (Turner and O'Reilly, 2013). The integration of hypothetical enzymes is beneficial when a product, starting material or intermediate displays similarity but not homology to an enzyme's native substrate. The development of hypothetical enzymes is readily accessible due to the pioneering work of Frances Arnold in directed evolution to engineer enzymes with catalytic mechanisms 'new to nature' (Hammer et al., 2017; Renata et al., 2015). As a consequence, computationally-guided *de novo* pathway design can now serve as a valuable starting point for engineering enzymes with distinctive functions or specificities (Hadadi and Hatzimanikatis, 2015). Current challenges associated with algorithm-based pathway design include the need to define reaction rules for enzyme promiscuity, multimolecular reactions (*e.g.* transaminases), novel (synthetic) compounds and reliance on possibly incomplete or incorrect annotation of enzyme reactions in databases which can limit the scope of this approach significantly and may require elaborate refinements (Delépine et al., 2017; Kumar et al., 2018; Poux et al., 2017).

4.3. An example of biocatalytic retrosynthesis

Recently, Erb and co-workers employed 'biocatalytic retrosynthesis' to design a novel cell-free synthetic cycle for the continuous fixation of CO_2, providing a promising alternative to the six CO_2 fixation pathways found in nature (Schwander et al., 2016). The pathway was created *in silico* using a variant of constraint-based modelling that defined the shortest carbon fixation cycle for a given set of well characterised and more efficient, carboxylating enzymes than those used in natural carbon fixation pathways (Bar-Even et al., 2010; Schwander et al., 2016). Reaction rules were based on the standard reaction classes defined by the EC standard nomenclature and were not restricted to the use of hypothetical enzymes. The pathways were ranked according to their potential to generate a complete pathway cycle, to yield a valuable output from the assimilated carbon (for example, malate), for regeneration of the substrate through the cycle, for favourable thermodynamics and low ATP usage. As a result, seven different carbon fixation cycles were formulated of which one cycle was composed of enzymes that had been characterised previously. The first tests failed to fulfil a complete carbon fixation cycle. Consequently, multiple pathway corrections were implemented, such as rational (active-site)

engineering of enzymes for enhanced specificity, the addition of enzymes for cofactor recycling, detoxification of reactive oxygen species (ROS) and re-feeding of dead-end metabolites into the cycle. The final synthetic pathway was composed of a network of 17 enzymes derived from 9 different organisms from all domains of life and was capable of assimilating CO_2 at a rate of 5 nmol/min/mg core cycle proteins. Notably, this cell-free synthetic pathway was five times more efficient than the most common natural carbon fixation pathways *in vivo*. This example demonstrates that *in silico* computational design of *de novo* enzyme pathways can advance the capacity to exploit significantly a complex biological system for innovative applications. However, it also showed that cell-free multi-enzyme biocatalysis is more complex and unpredictable than previously expected and needs to be compensated for by either improving the input settings or by undergoing the traditional "design – build – test and learn" cycle routinely proposed for synthetic biology.

5. Renewable substrates in biocatalysis

The most common substrates used in multi-enzyme biocatalysis are those supplied from refinery (glucose) or costly fermentation processes (glucose-6-phosphate) (Dudley et al., 2016; Guterl et al., 2012; Krutsakorn et al., 2013; Liu et al., 2017; Woodward et al., 2000; Ye et al., 2012; Zhu and Zhang, 2017). As the result of a continuous growth in world food consumption coupled with an increasing demand for sustainable production of bio-commodities, the use of renewable waste represents a more attractive source of substrates than those that are purposely refined (Koutinas et al., 2014). Potential sources include agricultural, industrial and municipal (food) wastes as their main components include biopolymers such as cellulose, hemicellulose, lignin and starch, all of which have been identified as having considerable potential as substrates for biofuel production (Pleissner and Lin, 2013). According to the United Nations Environment Programme (UNEP), > 5 billion m^3 tons of plant biomass waste is generated annually from agriculture (UNEP, 2014) and the average agricultural lignocellulosic biomass contains about 50–60% cellulose/hemicellulose and 15–25% lignin (Kenney et al., 2013). The most abundant agricultural feedstock in the United States is corn stover which is the main residue after corn harvest and contains about 3% cellulose and 21% xylan (Aden et al., 2002).

The availability and price of biomass as substrate is crucial for the economic viability and sustainability of a large-scale (bio)chemical production. For example, in the production of bioethanol from corn or sugarcane, the price of the feedstock constitutes approximately 40 to 75% of the final price of the commodity (Lynd et al., 1999). Therefore, careful substrate selection is critical and the introduction of alternative inexpensive substrates like biomass derived from organic municipal waste may improve the economic viability of these processes. Cell-free biocatalysis has substantial potential as an alternative approach for the production of bio-commodities from organic wastes. The major biocatalytic challenge for the utilisation of this biomass is its chemical complexity and the ability to channel the major degradation products into a multi-enzyme system without compromising enzyme activities. This problem has been addressed successfully by the many strategies that have been implemented for the supply of inexpensive substrates for various pathways. For example, Rollin et al. used a combination of acid hydrolysis and enzymatic degradation (cellulase/hemicellulase mixture) to convert corn stover into its monomers, xylose and glucose (Rollin et al., 2015). These sugars served as substrates for phosphorylation using polyphosphate glucokinase and hydrogen generation using a cell-free synthetic pathway composed of > 10 purified enzymes. The productivity of the pathway achieved was 54 mmol H_2 L^{-1} h^{-1} through optimisation of the pathway's reaction temperature, substrate, and enzyme concentrations.

Chitin is the second most abundant organic compound in nature and is primarily present in cell walls of fungi, exoskeletons from arthropods and scales of fish (Gooday, 1990). Approximately 6 to 8 million tons of waste crab, shrimp and lobster shells are produced yearly worldwide (Yan and Chen, 2015). There have been various studies describing the utilisation of chitin in fermentations (Cira et al., 2002) but a persistent challenge remains in the combination of saccharolytic enzymes to break down the polymer with the addition of fermentative organisms to further process the glucose. Recently, this problem was approached by Honda et al. who reported a one-pot bioconversion of colloidal chitin. to pyruvate using a cell-free pathway consisting of twelve thermophilic enzymes that had been expressed recombinantly in *E. coli* (Honda et al., 2017) (see Table 1 and Fig. 2B).

Glycerol is the main by-product from the bioethanol and biodiesel production and constitutes approximately 10% of the total biodiesel generated (Johnson and Taconi, 2007). The biodiesel industry was projected to reach 37 billion gallons by 2016 (Lam et al., 2010) resulting in 68% of the world glycerol produced from this industry (Vivek et al., 2017). Since then, the price of glycerol has plummeted due to increased biodiesel production and researchers have explored alternative applications for the use of glycerol as substrate in biocatalysis. Gao et al. designed a simplified synthetic cell-free pathway consisting of only three enzymes (alditol oxidase, dihydroxy acid dehydratase, and catalase) to produce pyruvate from glycerol without the addition of any cofactors (Gao et al., 2015) (Table 1). Pyruvate itself is a versatile C3 platform chemical that can be transformed further into the C4 platform chemical, acetoin. Acetoin exists as (3*R*)-acetoin and (3*S*)-acetoin and both isomers are used widely in several industries (Xiao and Lu, 2014; Yang et al., 2017). Acetoin produced through microbial fermentation usually results in a mixture of both isomers which has less value than enantiopure acetoin (Xiao and Lu, 2014). To resolve this effect, Gao et al. developed the synthetic cell-free pathway further to convert pyruvate selectively into (3*R*)-acetoin by introducing a α-acetolactate synthase and α-acetolactate decarboxylase (Gao et al., 2015) (Table 1). An 85.5% of the theoretical yield from glycerol to (3*R*)-acetoin conversion (stereoisomeric purity of 95.4%) was obtained when all five enzymes were combined with glycerol in a one-pot reaction.

Corn steep water is a valuable byproduct of the corn milling process. Racemic lactate (20–50 g/L) is produced naturally during the corn steeping process through the action of a *Lactobacillus*-aided fermentation process (Hull et al., 1996). In the United States alone, > 500,000 tons of racemic lactate can be separated from corn steep water (Li et al., 2017). The main impediment for the utilisation of this lactate in several lactate-based industries is its lack of optical purity. Li et al. designed different cell-free multi-enzyme pathways to convert the racemic lactate from corn steep water into the platform chemicals acetaldehyde, pyruvate and acetoin (Li et al., 2017). The main enzymes of the pathway were D-lactate oxidase, L-lactate oxidase, pyruvate decarboxylase and catalase. The pathway was capable of producing optically pure L- or D-lactate by incorporating stereo selective oxidases.

The examples given here represent only a small fraction of possible waste streams and process byproducts which could be used as potential substrates for cell-free bio-manufacturing (for an overview see Table 1). A wide range of pathways for carbohydrates other than glucose exist predominantly in archaea and fungi. Similar to the non-phosphorylative ED pathway, other sugars present in biomass such as xylose, rhamnose and arabinose can be metabolised in an oxidative manner. More recently, pathways have been identified in *Pichia stipitis* and *Debaryomyces hansenii* for L-rhamnose and in *Sulfolobus solfataricus* for L-fucose and D-arabinose (Watanabe et al., 2008; Wolf et al., 2016). These alternative pathways differ substantially from classical glycolytic routes such as the EM pathway in their amount and type of enzymes, thermodynamic profile and their use of phosphorylated cofactors. Continuing fundamental research in the field of carbohydrate metabolisms may provide more combinatorial opportunities for the assembly of synthetic cell-free pathways with less common carbohydrates as substrates. However, it is important to stress that for industrial production, cost drivers other than raw material must be considered, such as

transportation and pretreatment costs of the waste stream.

6. The key players in biocatalysis – the enzymes

6.1. Advantages of thermostable enzymes

Cell-free biocatalysis offers great process flexibility, allowing enzymes from a diverse range of natural (and sometimes non-natural) sources to be combined readily and rapidly tested for their compatibility and efficiency (Guo et al., 2017; Karim and Jewett, 2016, 2018; Swartz, 2006, 2010). Thus, the choice of reactions that can be performed under cell-free conditions is expanding constantly as enzyme discovery and protein engineering methods continue to advance. However, limitations still exist, as cell-free processes remain completely empirical since an enzyme's performance within non-natural reaction parameters and its interactions with foreign intermediates is still difficult to predict (Sheldon and Pereira, 2017). When constructing a cell-free pathway, enzymes preferably should exhibit high catalytic efficiency (low K_m and high k_{cat}), high stability, high substrate specificity and selectivity, low substrate and product inhibition profiles, similar operational conditions (temperature and pH) and preferably, accept inexpensive NADH or inorganic phosphate instead of NADPH or ATP.

As these properties are not common, the frequent situation that arises is the quest for improvements in the efficiency, stability, or specificity of selected enzymes. Either directed evolution or computational design (rational engineering) or a combination of both usually are proposed as solutions, with neither approach being trivial. Tawfik and collaborators have outlined many of the problems associated with enzyme improvement in several reviews, for example, the problem of the extensive sequence space occupied by an enzyme and the achievement of maximal catalytic efficiency (Goldsmith and Tawfik, 2017) selectivity (Bar-Even et al., 2015; Tawfik, 2014) and rate-temperature dependency (Elias et al., 2014). Among the many attributes of enzyme performance, one that is highly desirable for many cell-free processes is that enzymes should be thermostable, especially in commercial applications where such a trait would aid re-use. Thermostable enzymes display higher robustness towards heat and organic solvents than their mesophilic counterparts, operate optimally at high temperatures *e.g.* > 80 °C and denature irreversibly only at temperatures above their temperature optimum (Daniel et al., 2008). The rationale of using thermostable enzymes stems from the fact that thermostable enzymes can be heat-purified conveniently and inexpensively from crude cell lysates when expressed heterologously in a mesophilic host (*e.g. E. coli*) (Cheng et al., 2015; Guterl et al., 2012; You et al., 2017). For example, Ninh et al. (2015) co-expressed nine thermostable enzymes from one single operon in *E. coli* for the production of pyruvate from glucose and heat-purified all enzymes by simple heat treatment of the crude extract at 70 °C for 30 min. Another potential advantage of using thermostable enzymes include; i) at high temperatures, the reaction equilibrium favours product formation due to increased mass transfer, lower viscosity of the reaction fluid and improved reactant solubility and ii) (untested) claims of lower risk of common microbial contaminants in the reaction vessels. Accordingly, several cell-free pathways have been assembled recently with enzymes derived from extremely thermophilic bacteria (*e.g.*, *Thermus thermophilus, Thermotoga maritima*) and hyperthermophilic archaea (*e.g.*, *Archaeglobus fulgidus, Sulfolobus solfataricus*) (Fujisawa et al., 2017; Honda et al., 2017; Krutsakorn et al., 2013; Ninh et al., 2015; You et al., 2017; Zhong et al., 2017a, b). The reactions were conducted at temperatures > 50 °C at which the enzymes retain activity over several hours and achieve nearly 100% production of the theoretical conversion yields at high substrate and product concentrations (Table 1).

6.2. Classical and a new interpretation of enzyme kinetics

The quest for improvement in thermal stability can be pursued either through enzyme discovery *e.g.* guided by metagenomics on samples taken from geothermal environments or through enzyme engineering when thermostable analogues are lacking (Krüger et al., 2018). However, the proposal of a new model for the effect of temperature on enzyme activity may have a considerable influence on the understanding of enzymes in areas ranging from biodiversity screening, enzyme engineering and enzyme production in a bioreactor (Daniel and Danson, 2010, 2013; Daniel et al., 2009; Eisenthal et al., 2006; Peterson et al., 2004). The effect of temperature on enzyme activity has been interpreted traditionally as increased temperature gives increased activity - but results in a loss of activity as a result of enzyme denaturation (the Classical Model). The model of Daniel and Danson and collaborators introduces an inactive non-denatured form of the enzyme as an intermediate (the Equilibrium Model). Their results support the notion that the active site is more flexible than the whole protein and loss of activity occurs before denaturation. The Equilibrium Model introduces a new thermal parameter, T_{eq}, where the concentrations of the active and inactive enzyme are equal and allows prediction of the temperature at which the enzyme activity is maximal (T_{opt}, (Peterson et al., 2004)). The model implies that in enzyme engineering studies using rational or directed mutational procedures, an increase in thermostability will not necessarily lead to enhanced activity at high temperatures (Daniel and Danson, 2010, 2013). Instead, successful screening for high(er) temperature enzymes will need to account for both thermal stability and activity (*via* T_{eq}), unlike the Classical Model which suggests both activity and resistance to thermal denaturation can be achieved by selecting for thermal activity alone. Acceptance of these considerations will have an effect also on enzyme reactions in a reactor to give a product, as the output at given times and temperatures will vary between the two *e.g.* Classical and Equilibrium models (Eisenthal et al., 2006).

7. Cofactor regeneration systems for cost-effective biocatalysis

Cofactors such as nicotinamide adenine dinucleotide (phosphate) NAD(P)H/NAD(P)$^+$ or adenine tri- or diphosphate (ATP and ADP) are organic compounds that are integral in many redox and phosphoryl-dependent enzymatic reactions. *In vivo*, cofactors are regenerated continuously through downstream or coupled side reactions. *In vitro*, cofactors are expensive and their external addition into a cell-free biocatalytic system reduces its cost-effectiveness.

7.1. Cofactor regeneration

Cell-free cofactors recycling can be achieved by integrating cofactor regeneration pathways into the biocatalytic process to minimise cofactor dependency and costs. A common example is NADH which is accumulated often in the reduced form when the substrates of a pathway are oxidised (Bar-Even et al., 2012). The oxidised form (NAD$^+$) can be recovered by adding a water-forming NADH oxidase to maintain a continuous pathway flux, where O_2 is reduced to H_2O and NADH is oxidised to NAD$^+$ in this reaction (Beer et al., 2017; Gao et al., 2012; Geueke et al., 2003; Nowak et al., 2015). The reduced form NADH can be recovered from NAD$^+$ and inexpensive substrates such as formate or glucose using a formate or glucose dehydrogenase (Van der Donk and Zhao, 2003). ATP regeneration can be achieved economically by employing polyphosphate kinases which accept inexpensive polyphosphate as phosphate donor for the synthesis of ATP from ADP or AMP (Nocek et al., 2008; Schwander et al., 2016).

Another advantage of *in situ* recycling additional to the economics consideration is that the accumulation of either the reduced or oxidised form of the cofactors often is inhibitory to the enzyme and the reaction can be driven in the direction favoured by continuous removal of either form of cofactor (Van der Donk and Zhao, 2003; Zhang and Hess, 2017). For example, the synthetic purge valve system developed by (Opgenorth et al., 2014) and described in Section 2.1, effectively

eliminated the build-up of NADPH in the synthetic PHB pathway and increased the overall pathway flux two-fold (Opgenorth et al., 2016). A general overview on cofactor regeneration pathway requirements for cost-effective biocatalysis was provided by (Zhao and van der Donk, 2003). For ATP regeneration systems the interested reader is referred to (Andexer and Richter, 2015).

7.2. Cofactors in high-temperature enzyme system

In cases where the biocatalytic process is conducted at temperatures > 50 °C, heat-labile cofactors like NAD(P)H and ATP degrade rapidly and make any regeneration process superfluous. For example, at 40 °C and pH 6, the half-life of NADPH and NADH was shown to be 1 and 7 h, respectively (Wu et al., 1986). At 95 °C and pH 7, the half-life of ATP and NADH was 1 h and less than 5 min, respectively, while NADPH was completely degraded (Wiegel and Michael, 2002). However, as these compounds are used by organisms that grow at > 100 °C there must be some compensatory mechanisms utilised by cells (Daniel and Cowan, 2000). The loss of cofactors can be compensated for by continuous supplementation into the cell-free system, a step that greatly increases the costs (Huang et al., 2016; Ye et al., 2012). Honda and co-workers addressed this problem by introducing a cell-free "NAD salvaging pathway" composed of eight recombinant enzymes which were capable of maintaining a constant NAD^+ concentration of 4 mM for 15 h at 60 °C as compared to a no salvage system where NAD^+ was degraded to 1 mM (Honda et al., 2016). Another solution is to circumvent the dependency on expensive and heat-labile cofactors by designing alternative pathways that avoid the use of cofactors completely, for example, through the assembly of ATP-independent metabolic pathways (Fig. 2A and C) or by replacing or engineering NADPH-dependent enzymes with NADH-dependent analogues (Chen et al., 2016; Huang et al., 2016).

7.3. Biomimetics to replace cofactors

As another alternative, synthetic NAD-based biomimetics have been developed that retain the enzyme-recognition moieties. The advantage of biomimetics is that they can be altered to be more stable and less expensive than their natural counterparts. For example, NADH oxidase characterised from *Lactobacillus pentosus* has been reported to accept and regenerate various biomimetics (Nowak et al., 2015, 2017). However, enzymes may not always accept these biomimetics and may need to be adapted initially through laborious enzyme engineering (Nowak et al., 2017). For example, the cofactor binding site of a glucose dehydrogenase from *Sulfolobus solfataricus* (SsGDH) was engineered by computational design to regenerate a NADH-biomimetic more efficiently than the wildtype enzyme (Nowak et al., 2017). The application of synthetic NAD-based biomimetics was reviewed recently (Paul and Hollmann, 2016) and a more expanded discussion is recommended to the interested reader.

The message here is that importance of unfavorable cofactors dependency should not be underestimated since it can compromise the economics and effectiveness of cell-free biocatalysis and accordingly, needs to be addressed early in the pathway-construction process.

8. Co-immobilisation of multi-enzyme pathways

8.1. Challenges and benefits

The immobilisation of enzymes onto solid matrices can greatly enhance the operational performance and cost-effectiveness of biocatalytic processes. Immobilised enzymes are typically more stable under extreme physical (heat) and chemical (organic solvents) conditions compared to their free, non-immobilised counterparts (Liang et al., 2016; Planchestainer et al., 2017; Torres and Batista-Viera, 2017). An advantage is that immobilised enzymes can be removed from a reaction mixture with ease, enabling enzyme recycling (reuse), simplified product purification and lower production costs (Table 2) (Liese and Hilterhaus, 2013; Secundo, 2013; Sheldon and van Pelt, 2013). Enzyme immobilisation may be employed to partition enzymes into separate compartments to prevent the formation of undesirable side-reactions resulting from enzymes that are promiscuous for multiple substrates in the pathway, or to prevent substrate/product inhibition caused by metabolites that are present elsewhere in the pathway (Gruber et al., 2017). The major obstacle for the widespread application of enzyme immobilisation is the partial or complete loss of an enzyme's activity caused by conformational changes that occur upon contact with the solid matrix (Sheldon and van Pelt, 2013). Finding the optimal immobilisation strategy for an enzyme usually requires laborious trial-and-error experiments and becomes even more challenging when multiple enzymes must be attached to the same carrier (France et al., 2017). Despite these limitations, numerous studies have shown that the catalytic efficiencies of pathways were improved when multiple enzymes were co-immobilised within close proximity (Table 2) (Liang et al., 2016; Patel et al., 2017; Rocha-Martín et al., 2012; You et al., 2012; You and Zhang, 2013). The original assumption to explain this phenomenon was that the enhanced catalytic activity resulted from substrate channelling, the intramolecular guidance of a substrate from the active site of one enzyme to the other when enzymes are placed < 10 nm away from each other (Castellana et al., 2014; Kuchler et al., 2016; Wheeldon et al., 2016). As a result, elaborate strategies have been developed to target and immobilise enzymes to artificial DNA or protein scaffolds in a highly organised manner that mimics substrate channelling of multi-enzyme complexes *in vivo,* and the interested reader is referred to two comprehensive recent reviews on this topic (Quin et al., 2017; Shi et al., 2018). However, the original assumption that substrate channelling was the primary cause for the enhanced catalytic activity *in vitro* has been disputed and claimed to occur due to enhanced enzyme activities as solely as a result of the immobilisation or the microenvironment of the scaffold (Zhang and Hess, 2017; Zhang et al., 2016). As well, these authors argued that it was not *active* substrate channelling but substrate *confinement* (the prevention of intermediates escaping into the bulk medium) that was the primary cause of enhanced production rates when enzymes were in close proximity (Zhang and Hess, 2017). This argument is supported further by Chado and coworkers who proposed that the controlled localisation of enzymes onto predefined scaffolds provides negligible improvements on the overall pathway productivity as compared to random co-immobilisation (Chado et al., 2016). *In silico*-derived and experimentally-derived data revealed that simply altering the enzyme stoichiometry of a pathway in favour of a rate-limiting enzyme(s) can improve product yields significantly (Begum et al., 2015; Chado et al., 2016; Fujisawa et al., 2017; Krutsakorn et al., 2013; Liu et al., 2017; Opgenorth et al., 2016; Rollin et al., 2015; Zhang and Hess, 2017). While enzyme amounts can be adjusted easily in solution, this action may not be the case for immobilised enzymes. First, to achieve optimal stoichiometry of an enzyme pathway bound to a matrix it is necessary to control the immobilisation yield (the amount bound) of each enzyme. Secondly, it needs to be ensured that the matrix has enough capacity to bind substantial quantities of enzyme. Therefore, matrices should offer large surface areas or pores such as micro- and mesoporous matrices (*e.g.*, zeolites, amorphous silica, porous carbons or nanomaterials) and the interested reader is referred to a review published recently on this topic that summarises work in this field (Zdarta et al., 2018).

8.2. Co-immobilisation strategies

Most of the examples reported of enzyme co-immobilisation onto a uniform matrix involve a maximum of two enzymes and these strategies are reviewed elsewhere (Jia et al., 2014). Multi-enzyme pathways require the co-immobilisation of three or more enzymes, which is more difficult and thus rarely documented. We discuss here two strategies,

physical entrapment and affinity-based immobilisation, that are simple, readily utilisable in low-cost industrial applications, can be performed under mild conditions and are effective for the simultaneous co-immobilisation of multiple enzymes. In addition, we present a comprehensive account of co-immobilisation strategies and have limited the selection to those that involve three or more enzymes (Table 2).

8.2.1. Physical entrapment

Enzymes in cell-free biocatalysis are frequently entrapped *in-situ* in artificially-generated gel/matrices (polyacrylamide, alginate, silica sol-gel, metal-organic gels, Table 2). The immobilisation conditions are very mild and are expected to preserve the protein's native structure and activity and thus have proved to be successful for a variety of different enzymes (Jesionowski, Zdarta et al. 2014, Table 2). In many cases, the weak interactions of enzyme and matrix are stabilised through the entrapment and may prevent enzyme leaching that can occur under harsh reaction conditions, such as high flow rates, high ionic strength, variations in pH or temperature, as the binding is weak and reversible. However, additional covalent attachment might be required in situations when the physical restraints are too weak to prevent enzyme leakage under operational conditions (Sheldon and van Pelt, 2013). Physically-entrapped enzymes were shown to be protected from inactivation through heat, pH changes, protease attacks or toxic compounds such as organic solvents which are crucial considerations for industrial biocatalysis (Liang et al., 2015, 2016; Patel et al., 2017). For example, three different enzymes, cytochrome *c*, horseradish peroxidase and *Candida antarctica* lipase B were entrapped individually in a metal-organic framework (ZIF-8) and they exhibited a higher thermostability and tolerance towards organic solvents as compared to the free enzymes (Wu et al., 2017). However, synthetic matrices (*e.g.* sol-gels, silica gels, nanoparticles based gels) with high mineral content and nanometre-sized pores may affect substrate access and diffusion rates and thus influence the catalytic efficiency of an enzyme pathway (Table 2), (Begum et al., 2015; Liang et al., 2016). Sokic-Lazic and Minteer compensated for this effect by modifying a Nafion® polymer to be composed of enlarged micellar pores for the functional co-entrapment of 13 different enzymes (Table 2) (Sokic-Lazic and Minteer, 2008). This approach succeeded in preventing limited substrate access/diffusion and resulted in a methanol- oxidizing biofuel cell that yielded 1.5 times higher power density than calculated theoretically from that of the single enzymes system. Similarly, the use of polyacrylamide gels, metal-organic frameworks or hydrogels for immobilisation of various enzymes has resulted in a complete recovery or even enhancement of enzymatic activities, suggesting that these gels preserve the enzymes' native activity and are accessible by enzyme substrates (Liang et al., 2016; Patel et al., 2017; Zore et al., 2017).

8.2.2. Affinity-based immobilisation

Affinity-based enzyme immobilisation is another growing trend in cell-free biocatalysis as it is a mild and versatile approach (Table 2). Here, an enzyme is modified genetically with an affinity protein or peptide tag positioned away from the active site to prevent structural and functional impairment that mediates the strong attachment of the tagged enzyme to a matrix that displays the complementary affinity ligand (Barbosa et al., 2015). Accordingly, affinity-based enzyme immobilisation can be used simultaneously for protein purification (Planchestainer et al., 2017; You and Zhang, 2013). An affinity tag used frequently is the polyhistidine-tag which binds selectively to functionalised matrices that display divalent cations *e.g.* nitrilotriacetic acid (Ni-NTA) (Liu et al., 2002; Planchestainer et al., 2017; Rocha-Martín et al., 2012). However, functionalising matrices is expensive, making many affinity-tags unsuitable for large-scale industrial biocatalysis.

A promising alternative to conventional affinity tags are solid-binding peptides (SBPs), which exhibit high binding affinity to a diverse range of solid materials (Care et al., 2015; Sunna et al., 2013). They have been used to immobilise proteins in various biotechnological applications, including enzymes for biocatalysis (Care et al., 2015, 2016, 2017). For example, a SBP that binds strongly to inexpensive silica-containing materials was fused genetically to three different thermostable hemicellulases (Table 2) (Care et al., 2016, 2017). The SBP-tagged enzymes were co-immobilised onto a low-cost zeolite and shown to hydrolyse various hemicellulose substrates at 80 °C. The zeolite-immobilised enzymes maintained between 25 and 50% of their initial activity after 12 reaction cycles at 80 °C, demonstrating the stable interaction of the SBP with the matrix and their potential application in industrial biocatalytic processes. The approach was enhanced by covalently crosslinking the zeolite-bound enzymes using glutaraldehyde, resulting in 40–65% residual activity remaining after 12 reaction cycles at 80 °C (Table 2). Alternatively, enzymes can be tagged with dockerin domains which bind with high affinity to its corresponding cohesin protein domain. A scaffold of three different cohesin protein domains was displayed on genetically modified yeast cells which facilitated the immobilisation of a three-enzyme pathway for the cell-free conversion of methanol to CO_2 (Table 2) (Liu et al., 2013). Similarly, three different dockerin-containing enzymes were assembled onto a synthetic scaffold which was expressed recombinantly in *E. coli* and modified to contain the corresponding cohesin domains and a cellulose-binding-module. The scaffold mediated the one-step purification of the three enzymes from *E. coli* cell lysates and co-immobilisation of all three enzymes onto cellulose (Table 2) (You and Zhang, 2013).

As shown in Table 2, most of the co-immobilised enzyme systems yield a higher activity (either total product titre or productivity, *e.g.*, titre over time) than that achieved using free enzymes. This effect may be attributed to substrate refinement and thus higher local substrate concentrations and higher pathway fluxes (see section 6.1), or could be a result from the protection and stabilisation of enzymes from inactivation upon immobilisation. There are only two reports where the reusability of co-immobilised enzyme pathways was examined (Table 2). In both cases, the co-immobilised enzymes were easily recovered, reintroduced into the reaction and reused for 10–12 rounds, thus, allowing for cost-effective reuse of the biocatalysts.

Physical entrapment or affinity binding are promising strategies for co-immobilisation of whole enzyme pathways. They facilitate substrate confinement and improve pathway productivity, while the increase in enzyme stability and re-usability can reduce process costs directly.

9. *In silico* kinetic modelling for improving pathway yields

9.1. Introduction to kinetic modelling

Kinetic modelling is a promising tool for the *in silico* simulation, control and optimisation of biological processes (Vasic-Racki et al., 2003). It also can be a powerful tool for pathway assembly and the optimisation of cell-free systems where biocatalysis is more readily defined and simplified than for *in vivo* metabolic pathways. No cellular processes are present in a cell-free environment that interfere with the desired reactions and changes in concentration of intermediates can be attributed accurately to the enzymatic reactions. In most approaches, kinetic modelling can help to adjust the initial enzyme, substrate and cofactor loadings (Rollin et al., 2015; Ye et al., 2009; Zhong et al., 2017a) and determination of the optimal parameters of a reaction (Hold et al., 2016; Zhong et al., 2017a). Initially, enzyme systems were modelled *in silico* to identify ideal enzyme ratios and potentially inhibitory substrates. For example, studies of the kinetic parameters of a fungal carbonyl reductase-catalysed enantioselective reduction of 2-octanone to (*S*)-2-octanol resulted in a 9-fold increase in turnover numbers of even poorly soluble ketones (Liese et al., 1998). The production of L-phenylalanine from the racemate D, L-phenyllactate is a two-step reaction involving the enzymes D- and L-hydroxyisocaproate dehydrogenase and L-phenylalanine dehydrogenase. Schmidt et al. used kinetic models of the reaction to calculate the optimum ratio of the

ARTICLE IN PRESS

Table 2
Table summarising the various techniques used for the co-immobilisation of multiple enzymes or biocatalytic pathways (composed of three or more enzymes).

N°of enzymes	Biocatalyst	Application	Immobilisation technique	Support matrix	Recovered activity (%)	Relative activity[a](%)	Reusability[b]	Ref
3	FDH FalDH ADH	Cell-free production of methanol from CO_2	Physical entrapment	Cellulose filter membrane	NA	100	100[c]	(Luo et al., 2015)
3	ADH FalDH FDH	Enzymatic biofuel cell: methanol oxidation	Affinity binding *via* cohesion-dockerin pair	Synthetic protein scaffold displayed on yeast surface	NA	510	NA	(Liu et al., 2013)
3	β-Gal AI XI	Cell-free production of fructose and tagatose from lactose	Physical adsorption/covalent cross-linking	Epoxy-activated acrylic supports (Eupergit C and C 250 L)	70-98[d]	138 (fructose) 80 (tagatose)	NA	(Torres and Batista-Viera, 2017)
3	Cel GlK G6PD	Cell-free production of 6-phospho-gluconolactone from cellobiose	Affinity binding *via* hexa-arginine tag/ encapsulation	Silica	75[e]/58[f] 90[e]/85[f] 80[e]/45[f]	80	NA	(Begum et al., 2015)
3	TIM Ald FBP	Cell-free production of fructose-6-phosphate from glyceraldehyde-3-phosphate	Affinity binding *via* cohesin-dockerin pair	Synthetic protein scaffold attached to cellulose	NA	1000	NA	(You and Zhang, 2013)
3	Bgl Man Xyn	Individual enzyme reactions[g]	Affinity binding *via* silica-binding-SBP/ cross-linking	Zeolite	NA	99 56 41	50[e,i]/65[h,i] 25[e,h]/40[h,i] 30[e,h]/60[h,i]	(Care et al., 2017b)
4	GOx HRP Lip Amyl	GOx and HRP: two-enzyme system for glucose detection. Lip and α-amylase: Individual enzyme reactions[f]	Physical entrapment	Zinc/adenosine 5′-mono-phosphate (AMP) hydrogel	94[d] 100[d] 77[d] 90[d]	260 in the glucose detection system	NA	(Liang et al., 2016)
5	GOX AP LDH HRP Lip	Individual enzyme reactions[g]	Physical entrapment/cross-linking	Polyacrylic acid (PAA) nanogel	91[j] 970[j] 40[j] 54[j] 450[j]	NA	NA	(Zore et al., 2017)
7	GalK GALT UDPGP PPA UMPK NDK PPK	Cell-free production of UDP-galactose from galactose	Affinity binding *via* hexa-histidine tag	Nickel-functionalised agarose beads	140[k] 96[k] 71[k] 100[k] 110[k] 53[k] 125[k]	142[l]	NA	(Liu et al., 2002)

(*continued on next page*)

Table 2 (*continued*)

N°of enzymes	Biocatalyst	Application	Immobilisation technique	Support matrix	Recovered activity (%)	Relative activity[a] (%)	Reusability[b]	Ref
13	Enzymes from methanol metabolism and citric acid cycle	Enzymatic biofuel cell: ethanol oxidation	Physical entrapment	Modified Nafion®	NA	145	NA	(Sokic-Lazic and Minteer, 2008)

Abbreviations: FDH (formate dehydrogenase); FalDH (formaldehyde dehydrogenase); ADH (alcohol dehydrogenase); β-Gal (β-galactosidase); AI (L-arabinose (D-galactose) isomerase); XI (D-xylose (D-glucose) isomerase); Cel (cellulase); GlK (glucose kinase); G6PD (glucose-6-phosphate dehydrogenase); TPI (triosephosphate isomerase); Ald (aldolase); FBP (fructose 1,6-bisphosphatase); Bgl (β-glucosidase); Man (β-mannanase); Xyn (β-xylanase); GOx (glucose oxidase); HRP (horseradish peroxidase); Lip (lipase); Amyl (α-amylase); AP (acid phosphatase); LDH (lactate dehydrogenase); GalK (galactokinase); GALT (galactose-1-phosphate uridyl transferase); UDPGP (UDP-glucose pyrophosphorylase); PPA (inorganic pyrophosphatase); UMPK (uridylate kinases); NDK (nucleoside diphosphate kinase); PPK (polyphosphate kinase); NA (not available).

[a] Relative activity represents product titre (%) or pathway productivity (%) relative to the values obtained from a free enzyme pathway.
[b] Reusability is expressed as residual activity (%) after n cycles.
[c] After 10 cycles, each for 30 min at room temperature.
[d] Retained activity from single immobilised enzymes.
[e] Enzymes were individually immobilised through affinity binding.
[f] Enzymes were first immobilised through affinity binding and subsequently encapsulated.
[g] Individual reactions indicate that the co-immobilised enzymes catalyse different reactions and are not part of a multi-enzyme pathway.
[h] Enzymes were individually immobilised through affinity binding and subsequent cross-linking.
[i] After 12 cycles each for 10 min at 80 °C.
[j] Recovered activities from enzymes immobilised according to scheme 2 in respective reference.
[k] Calculated from V_{max}.
[l] Derived from pathway productivity (g L d^{-1}).

enzymes, the optimum cofactors concentration and the optimum concentration of the prochiral reaction intermediate, phenylpyruvate (Schmidt et al., 1987). In order to establish the model, exact kinetic parameters of all enzymes for forward and reverse reactions had to be determined. The parameters enabled the formulation of mass-balance equations which can be solved using a Runge-Kutta integration. Solving the mass balance equations enabled estimation of the effect of changing enzyme and substrate amounts, resulting in a prediction of an optimal enzyme ratio in the system. The kinetic optimisation of the process resulted in the enzymatic production of L-phenylalanine from the racemic mixture of D, L-phenyllactate with a 43% conversion.

More recently, a four-enzyme system for the conversion of D-methionine into L-methionine was described with a mathematical model (Findrik and Vasic-Racki, 2007). They analysed the kinetics of all enzymes involved in the conversion process that were used to generate mass-balance equations and applied mathematical models for each reaction step and for the complete system. The model allowed the optimisation of initial reaction conditions to lower the amount of enzymes and cofactors used for a complete conversion. Validation of the mathematical models in batch reactor experiments resulted in a 100% conversion of D-methionine to L-methionine. Most kinetic models for multi-enzymatic systems can be described by ordinary differential equations (ODE) which are either based on mass balance equations of known parameters like K_m or k_{cat} (Rollin et al., 2015; Ye et al., 2009) or on experimental data of metabolic fluxes in the reaction system (Hold et al., 2016). ODE describes outputs like yield and volumetric productivity and different algorithms can be used to solve these equations. One type is exemplified by genetic algorithms, which are a solution-search and optimisation method inspired by the principles of natural selection and developed initially to solve evolutionary relationships. It iteratively evolves a set of solutions to given problems and is now being exploited to solve multi-objective functions (Alter et al., 2018; McCall, 2005). For example, Rollin et al. described the high-yield production of H_2 using a cell-free enzymatic system for the conversion of glucose and xylose from plant biomass (Rollin et al., 2015). They used a genetic algorithm in combination with a global sensitivity analysis to identify the enzymes which had the greatest impact on reaction rate and yield. The model was used to optimise enzyme loadings and reaction temperatures and resulted in a 67-fold increase in H_2 productivity to 54 mmol $H_2 \cdot L^{-1} \cdot h^{-1}$. Computational modelling using algorithms has been primarily developed for *in vivo* metabolic engineering. However, *in vitro* systems are highly suited for modelling approaches, since a closed and precise design facilitates an easier and robust simulation. In theory, challenges encountered in genome-scale models for the metabolic modelling of production hosts can be overcome by the modelling of *in vitro* systems. For example, in a framework for building dynamic cell-free metabolic models, Wayman et al. found that complex allosteric regulations in cell-free systems can be described by simple reaction rules (Wayman et al., 2015). When these rules were integrated into network models, regulatory patterns such as product inhibition could be predicted for simulated cell-free reactions. In general, the use of high computing power and algorithms for kinetic modelling becomes more important with the increase in complexity of the cell-free reaction system and a higher amount of acquired kinetic data (see Section 7.4).

Multi-enzyme systems for the industrial production of pharmaceutical and fine chemicals often rely on expensive cofactors that are one of the cost drivers in industrial biocatalysis (see cofactors section). Mathematical models describing cofactor-dependent enzyme systems have been introduced to lower production costs. In one example, a model based on differential rate equations was used to investigate the recycling of NADPH in lactone synthesis (Hogan and Woodley, 2000). The model facilitated the optimisation of enzyme proportions and cofactor concentrations to achieve improved reaction rates. In addition, the model helped to predict cofactor stability as function of pH. Similarly, Van Hecke et al. developed a kinetic model to describe the bi-enzymatic oxidation of lactose to lactobionic acid in the presence of a

redox mediator (Van Hecke et al., 2009). Lactose was converted to lactobionic acid by the enzyme cellobiose dehydrogenase with the use of a 2,2′-azino-bis (3-ethylbenzthiazoline-6-sulphonic acid) as redox mediator which was continuously recycled by a laccase enzyme with the supply of oxygen as final electron acceptor. Implementation of the kinetic model allowed investigation of the productivity of the process by simultaneously varying enzyme and redox mediator concentrations. Accordingly, the systems' behavior could be predicted for different cofactor and enzyme concentrations at a given oxygenation rate with the model predictions in close agreement with the experimental data obtained in mini-reactors.

9.2. Simulation and optimisation

Describing and predicting *in vivo* metabolism by models derived from cell-free experiments is often difficult since in most cases, the cell-free conditions are so different to the conditions present in the cell. So far, models based on cell-free enzyme reactions and kinetic parameters from the literature only partially matched *in vivo* metabolic pathways (Teusink et al., 2000). However, cell-free approaches may provide novel information for *in vivo* biosynthesis. A multi-enzyme system with three purified glycolytic enzymes from *E. coli* was simulated and matched empirical data. The model was created from experimental data on single enzyme reactions and could simulate the outcome of a cell-free multi-enzyme system without further fine tuning of the kinetic parameters (Ishii et al., 2007). Although the simulation was not compared to actual *in vivo* data, the multi-enzyme experiments were simulated with physiologically-meaningful conditions. In another example, it was possible to model a redox process for the synthesis of lactones (Milker et al., 2017). In order to create a kinetic model that estimates changes in the *in vivo* system, kinetic parameters were acquired in cell-free reactions and used in non-linear fittings for *in vivo* experimental data. Despite the fast kinetic observed, the final enzyme of the cascade could be identified by the model as the rate limiting step due to the low concentration of its active form. Further, cell-free experiments with an excess of cofactors provided evidence that their low supply *in vivo* drastically limited product formation.

Besides finding the ideal enzyme as well as cofactor and substrate concentrations, computational models of cell-free biocatalysis can speed up the process of finding the ideal reactor design and reaction conditions, for example, the reaction temperature (Ardao and Zeng, 2013; Findrik et al., 2005; Rollin et al., 2015; Vasic-Racki et al., 2003). A comparative kinetic model simulating the production of ε-caprolactone *via* a three-enzyme pathway showed that a fed-batch reactor enabled a 28% faster conversion than normal batch synthesis. In addition, the *in silico* model describing the production modes yielded > 90% accuracy for ε-caprolactone production (Scherkus et al., 2017). An *in silico* kinetic model suggested that running a hyperthermophilic hydrogenase in a second reactor separate from the rest of the pathway would lead to a 2-fold increase in the production of hydrogen (Ardao and Zeng, 2013) and the introduction of a second pathway that consumes an inhibitory compound of the hydrogen production pathway was predicted to increase the productivity 8-fold.

One crucial aspect in kinetic modelling is to identify the most affected parameters of a system through a sensitivity analysis for rapid and effective process control. In a traditional sensitivity analysis, one parameter is changed whereas all the other parameters remain fixed. Combined with *in silico* simulations, a sensitivity analysis helped to identify bottlenecks in the production of chiral alcohols and running the analysis identified a high K_m value cell-free enzyme with a single substrate as the key effector in a multi-enzyme reaction (Rios-Solis et al., 2013).

9.3. Limitations of current in silico kinetic models

The kinetic models may not fit the dynamics of multi-enzyme systems as most are based on data from the literature or single experiments under defined conditions (Myung and Zhang, 2013; Zhong et al., 2017b). For example, a model for the production of cellobiose from sucrose initially was created by using kinetic parameters from the literature but did not match experimental results. Only after the model was supplemented with parameters acquired under the appropriate reaction conditions could the kinetic simulation predict optimal enzyme ratios and then facilitated 62% cellobiose conversion (Zhong et al., 2017b). Alternatively, novel models can be created by establishing artificial neural networks (ANN) in systems where the kinetic parameters of the enzyme reactions are unknown (Fontaine et al., 2015). ANN's consist of multiple computational units which act like neurons in a biological network, receiving several input functions from other neurons. Input and output functions are described through experimental data (*e.g.*, initial reaction conditions and conversion rates) which is used to simulate an optimised output. This strategy requires sufficient experimental data to be able to derive a valid model. For example, an ANN was developed for cell-free hydrogen production using empirical data available in the literature (Fontaine et al., 2015). The model highlighted the claim that different enzymes influence the hydrogen output whereas previous models mostly recognised one enzyme to as being responsible for maximum hydrogen yield (Ardao and Zeng, 2013; Ye et al., 2009).

9.4. The way forward with kinetic modelling

The final adjustment of these models must be based on truly empirical data although the development and accuracy of kinetic models is progressing steadily. High throughput measurement of different experimental setups and their feedback into the kinetic models should be incorporated in the process design to obtain accurate prediction models. This strategy was used efficiently in the cell-free production of the platform chemical dihydroxyacetone phosphate (DHAP) (Hold et al., 2016). High throughput, real time measurements were used to acquire the kinetic parameters from 22 different experiments which were employed to adjust the kinetic model and to determine the conditions required for optimal product yield and cofactor concentration (see Section 8.3).

Mathematical models can be powerful tools for the optimisation of production pathways. However, there are influences such as protein-protein interactions, local substrate concentrations or molecular crowding effects which are hard to model (Ishii et al., 2007). Experimental data acquisition and curation is as important as the further development of more sophisticated algorithms for the kinetic optimisation of cell-free bioprocesses.

10. High-throughput metabolite analysis for system optimisation

The main pitfalls for high product yields by a multi-enzyme pathway are bottlenecks in the metabolic flux such as the accumulation of reaction intermediates through enzyme inhibition or unfavorable kinetics and the formation of unexpected by-products through promiscuous enzyme action. All these factors are crucial for understanding, developing, and optimising systems in synthetic biology and biocatalysis (Hold et al., 2016; Nguyen et al., 2012; Swartz, 2012). In this section, we summarise current strategies used to monitor the metabolic flux in multi-enzyme biocatalysis and propose state-of-the-art technologies that are available to overcome these limitations. Advantages and limitations of these technologies are summarised in Table 3.

10.1. Examples of product measurements in enzyme pathways

Currently, there are few examples in which all metabolites of a multi-enzyme pathway (*i.e.* comprising > 6 metabolites) have been measured simultaneously. This barrier is due mainly to a lack of analytical techniques that can target the chemical diversity of the

Table 3
The advantages and limitations of enzymatic assays, colorimetric assays, chromatographic approaches, mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR) as analytical tools for the simultaneous metabolite analysis of cell-free biocatalytic processes. (U)HPLC, (ultra-) high performance liquid chromatography; GC, gas chromatography; DAD, diode array detector (at wavelengths below 200 or above 200 nm); PAD, pulsed amperometric detector; RID, refractive index detector; FID, flame ionisation detector. (•), poor performance; (••), moderate performance; (•••), high performance.

Analytical tool	Sample throughput	Costs (run, purchase, maintenance)	Selectivity	Sensitivity	Robustness	Method development/ validation	Data interpretation	Flux analysis[a]
Enzymatic	••	•••	••	•••	•	••	•••	•
Colorimetric	••	•••	•	•••	••	•	•••	•
HPLC	•	••						
UHPLC	••	••						
DAD (≤200 nm)		••	•	••	•••	••	••	•
DAD (≥200 nm)		••	•••	••	•••	••	••	••
PAD		••	••	••	•••	••	••	•
RID		••	•	•	••	••	••	•
GC-FID	•	••	••	••	••	••	••	•
LC-MS/GC–MS	••	•	•••	•••	••	•	•••	•••
MS[b]	•••	•	••	••	•	•	•	••
MS/MS[b]	•••	•	•••	•••	••	•	••	•••
^{1}H NMR	••	•	••	•	•••	•••	•	••
^{13}C-NMR[c]	••	•	••	••	•••	•••	••	••

[a] Involves the simultaneous analysis of multiple metabolites, *e.g.* > 10 metabolites.
[b] Direct injection.
[c] Performed with ^{13}C-labelled substrates.

metabolites potentially generated within a multi-enzyme pathway (Lin et al., 2017). Therefore, pathways are often 'monitored' by measuring only their final product, *e.g. myo*-inositol (Fujisawa et al. (2017), hydrogen (Lu et al., 2015), and isoprene (Cheng et al., 2017; Korman et al., 2014) or by analysing the product and a few easily-detectable metabolites, such as sugars (glucose, fructose), alcohols (butanol, glycerol, myo-inositol) and organic acids (pyruvate, glycerate, malate, lactate, gluconate (Gao et al., 2015; Honda et al., 2017; Krutsakorn et al., 2013; Ye et al., 2013; Zhong et al., 2017b). Analysis is accomplished mainly by using high-performance liquid chromatography (HPLC) for liquid metabolites or gas chromatography (GC) for volatile metabolites to separate the reaction mixture and subsequently, for detection of single metabolites with on-line coupled detection systems. Common detectors used with HPLC are the diode array detector (DAD) which detects UV–visible absorbent compounds, the pulsed amperometric detector (PAD) which allows detection of oxidizing or reducing compounds and refractive index detector (RID) which detects everything with a refractive index different from that of the mobile phase and when using GC, the flame ionisation detector (FID). The advantage of LC or GC is that the sample preparation is simple (unless compounds require derivatisation to be volatile for detection by GC) and only includes the removal of proteins, cell debris and an exchange of buffers. Because of this simplicity, the use of HPLC or GC provides high robustness to the detection system and consequently, higher reproducibility with minimal interference from the sample's other components (Table 3).

A major limitation of these detectors is their medium to low sensitivity, for example, μM (PAD or DAD) to mM (RID) and poor selectivity since they cannot characterise analytes by unique chemical or structural features other than their chromatographic behavior (*e.g.*, retention time with RID), or certain functional groups (*e.g.*, DAD and PAD). Another possible problem is that several metabolites co-elute and cannot be differentiated due to the diverse nature of metabolites and other non-target compounds present in a sample which renders this approach less suitable for flux analysis (Table 3). Regardless of these issues, it is common practice to confirm the identity of a product based solely on whether it exhibits the same retention time as its commercial standard without taking into account that unexpected by-products with the same retention time may have been formed unexpectedly in the pathway. Another limitation of HPLC or GC is their low sample throughput as the separation time in HPLC or GC is usually between 30 and 60 min, or 5–10 min when using ultra-high performance liquid chromatography (UHPLC), resulting in a maximum throughput of 24 to 300 samples per day.

Most of the HPLC methods used in cell-free biocatalysis employ ion exclusion, weak cation exchange, weak anion exchange or hydrophilic interaction liquid chromatography (HILIC) columns that are able to separate small polar molecules such as sugars, organic acids and alcohols. However, highly polar or charged metabolites, such as NADH, ATP, glucose-6-phosphate, glyceraldehyde-3-phosphate and fructose-1,6-bisphosphate were not separated and detected under these conditions (Krutsakorn et al., 2013; Opgenorth et al., 2016; Wang et al., 2017a; Ye et al., 2013). Their detection is performed commonly using enzymatic or colorimetric assays as the concentration of these metabolites can give important clues about the dynamics of the enzyme system (Krutsakorn et al., 2013; Gao et al., 2015; Ye et al., 2012, 2013; Schwander et al., 2016; Dudley et al., 2016; Korman et al., 2017; Guterl et al., 2012). Despite their widespread use, these assays are time-consuming and moderately selective as enzymes can be promiscuous and colouring agents often detect functional groups rather than single molecules. These considerations mean that these assays are not capable of detecting multiple metabolites simultaneously and thus unsuitable for flux analysis, they are influenced by the sample composition such as metals, pH, lipids, salts and other metabolites and the assay conditions and consequently are not very robust (Table 3). Furthermore, the enzymatic detection of cofactors becomes impractical when they are recycled *in situ* or when cofactor turnover does not correspond directly with substrate turnover (Zhang and Hess, 2017).

10.2. Current approaches to simultaneous and high-throughput metabolite analysis

As a result of these issues, the use of highly selective and sensitive (nM - μM) detectors such as mass spectrometry (MS) becomes particularly attractive for avoiding these limitations (Table 3). In MS, the analytes are first ionised and transferred into the gas phase, for example by electrospray ionisation (ESI), then separated according to their mass to charge ratios (m/z) and detected *via* a secondary electron multiplier. Consequentially, several hundred compounds can be analysed simultaneously, including cofactors and phosphorylated substrates (Zampieri et al., 2017). Tandem MS (MS/MS) provides the additional option of selecting single or multiple ions, fragmenting them and detecting all fragmented ions in a second MS assay (single or multiple reaction monitoring mode (SRM, MRM); Table 3, Fig. 3. Ion

fragmentation patterns are metabolite-specific and provide additional structural information on the targeted molecules while allowing detection of targeted compounds with higher sensitivity and selectivity as compared to MS.

These substantial advantages associated with MRM-MS have been used to analyse samples without LC or GC –based sample separation prior to MS. Samples were directly injected into the MS system which resulted a in dramatic increase of the sample throughput (*e.g.* > 1 sample/s and in total 75,000 samples per day (Zampieri et al., 2017). In another elegant example, MRM-MS was employed for the real-time analysis of a synthetic enzyme pathway for the cell-free production of DHAP from glucose (Bujara et al., 2011; Hold et al., 2016). Samples were extracted continuously from a stirred tank reactor, filtered, diluted with MS-compatible buffers and introduced into the detector *via* direct injection (Fig. 3). High-density data acquisition (every 8 s) was used systematically to assess optimal enzyme concentrations needed by an *E.coli* cell extract to direct the glycolysis flux towards DHAP (Bujara et al., 2011). The results served as an initial template for the design of a synthetic operon to be used in *E. coli* that was optimised rapidly to express the enzymes in previously established optimal ratios in *E. coli* through this real-time metabolite analysis. After iterative cycles of design-build-test, the operon produced a 2.5-times higher DHAP production (2.5 mM) in the cell-free extract than when enzymes were added externally. The same real-time MRM-MS set-up was then applied later to advance this complex pathway into the cell-free production of glyceraldehyde-3-phosphate (0.25 mM) using 10 purified enzymes from various organisms and a minimalised cofactor supply (Hold et al., 2016).

A major limitation of direct injection/MS without preceding chromatography is that the enzyme samples or cell lysates often contain large amounts of salts and other metabolites. These compounds can suppress ionisation and lead to inaccurate quantification in addition to a continuous contamination of the MS system. To identify and compensate for ion suppression, elaborate validation schemes (for example, by spiking with isotopically-labelled internal standards into the sample) needed to be integrated into the process. This action adds time and costs to the overall method development and sample analysis. Another limitation of MS concerns the analysis of isomers or isobars, which are ions with the same nominal m/z ratios and in the case of tandem MS, generate similar fragmentation patterns and thus they cannot be discriminated from each other (for example glucose-6-phosphate and fructose-6-phosphate, glyceraldehyde-3-phosphate and dihydroxyacetone phosphate or 3-phosphoglycerate and 2-phosphoglycerate (Hold et al., 2016)). However, the chromatographic separation of these isobaric metabolites prior to MS is essential for such samples if their measurement is important. Finally, the purchase and maintenance of a mass spectrometer is very costly. Despite the considerable advantages of MS, the cost factor is a limitation in its widespread use in synthetic biology laboratories.

Other methods besides MS can be applied for high-throughput metabolite analysis including nuclear magnetic resonance spectroscopy (NMR) which is a selective, simple to perform, robust and a non-destructive analysis method for liquid samples (Alonso et al., 2015). ^{1}H NMR-based metabolomics has been applied in the characterisation of various biological samples such as blood/urine and yeast media (Alonso et al., 2015) and has been used in high-throughput applications (*e.g.* 1.5 min/sample with a 16 s acquisition time) (Kautz et al., 2005). However, the isotope ^{1}H is naturally highly abundant (99.9%) and thus interference of the pathway flux measurements from enzymes, buffers and endogenous metabolites in the cell-free extracts greatly reduce the sensitivity of this approach. In addition, difficulties in interpreting spectra from complex reaction mixtures, longer analysis time, and similarly high costs when compared to MS has limited the widespread application of ^{1}H NMR in the flux analysis of biocatalytic pathways (Table 3) (Emwas et al., 2013). Alternatively, the use of ^{13}C-labelled substrates in combination with ^{13}C NMR spectroscopy is a valuable yet underutilised tool to monitor pathway fluxes and to identify unambiguously the conversion of substrate to product throughout a multi-enzyme pathway. As the natural abundance of ^{13}C is only 1.1%, the use of ^{13}C-enriched substrates reduce the background noise, increase the sensitivity and simplify spectra analysis (Table 3) (Klein and Heinzle, 2012). Therefore, ^{13}C-based flux analysis has been applied readily for *in vivo* metabolic engineering (McAtee et al., 2015). Disadvantages of this approach are the higher costs associated with the substrate labelling procedure and different enzyme kinetics that may be observed with the heavy substrate when compared to the natural substrate due to isotope

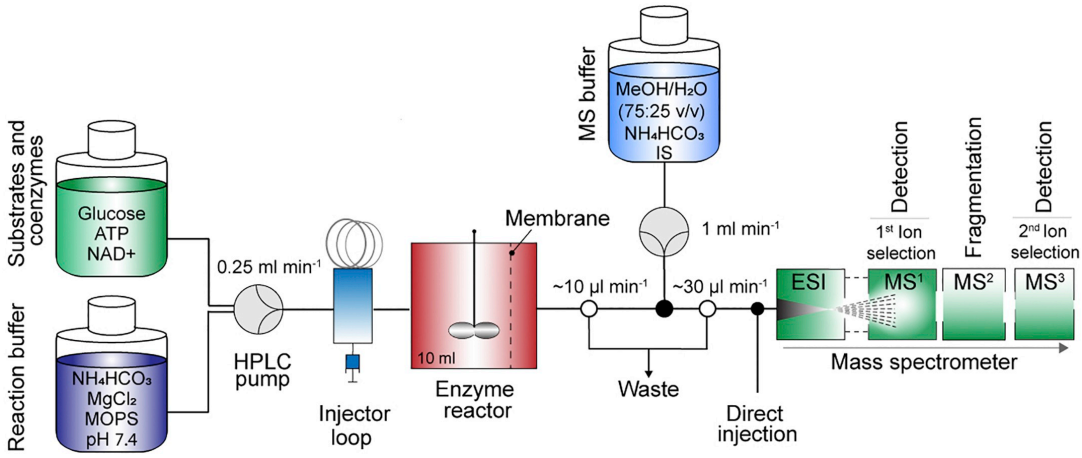


Fig. 3. Schematic illustration of the experimental set-up for the real-time metabolite analysis of a synthetic enzyme pathway using electron spray ionisation (ESI)–triple quadrupole mass spectrometry, adapted from Hold et al. (2016). Reaction substrates and cofactors were mixed with the reaction buffer using an HPLC pump system, and the mixture was fed continuously into the enzyme reactor. Enzymes were injected into the influent and incubated with the substrates in a continuously stirred tank reactor (600 rpm) maintained at 30 °C. From the reactor, samples were collected continuously and passed through a cellulose membrane to retain the enzymes. The effluent was diluted with MS-compatible buffer and internal standard (IS) before it was injected into the MS system.

ARTICLE IN PRESS

effects (Cleland, 2005).

Another method is Raman spectroscopy (RS) which usually suffers from low sensitivity (Bumbrah and Sharma, 2016). However, improved sensitivities (in the lower mM range) have been achieved by employing UV resonance Raman (UVRR) spectroscopy for detection of aromatic metabolites in a whole-cell enzyme pathway in real-time manner (Fisk et al., 2017).

Currently, LC/GC–MS or direct injection-MRM-MS are readily available for metabolite analysis and offer valuable strategies to ensure sufficient throughput and data quality that is required to match the rapid developments in the design and build phase of synthetic biology. Further advances are expected in the field of spectroscopic techniques and they likely offer a range of new possibilities for flux analysis of biocatalytic processes.

11. Conclusions and outlook

These are exciting times for cell-free multi-enzyme biocatalysis and progress is being made at considerable pace, whether it is for the production of biofuels and bio-commodities, the synthesis of chiral building blocks, or for replacing single steps in classical organic synthesis routes (Bornscheuer, 2018; Rudroff et al., 2018; Sheldon and Brady, 2018). The design of more sophisticated enzyme pathways is a challenging task, yet crucial to the success of biocatalysis in academic and industrial settings. We envisage that progress in biocatalytic retrosynthesis and computational pathway design will accelerate the development of novel pathways and further expand the repertoire of enzyme-driven synthesis routes, for the ultimate goal of replacing uneconomical or yet non-viable routes.

Enzymes have largely overcome their traditional reputation as fragile and ineffectual catalysts with the aid of advances in enzyme engineering and immobilisation. Accordingly, we speculate that the immobilisation of whole enzyme pathways and their recycling is not only achievable but will have a substantial impact on the uses of multi-enzyme biocatalysis in industrial applications. Advances in high-throughput and *in silico* screening of enzyme libraries and metagenomic data will expedite the characterisation of more effective and stable biocatalytic enzyme pathways. However, this requires more knowledge about enzyme systems and the metabolic flux in form of high density data that are fed back into the models to improve their accuracy and predictability. This data could be provided by high-throughput metabolite analysis in a real-time manner. The advantages of MS-based metabolite analysis in terms of throughput, sample preparation, accuracy and versatility are obvious. Nevertheless, many cell-free biocatalytic processes rely on low-throughput detection systems mainly because it is a laborious and intricate process to develop and validate methods for the analysis of multiple metabolites and because MS systems and their maintenance are very costly. Computational tools that recognise structural features of the analytes to derive an effective and integrated analysis strategy may provide guidance in developing suitable analytical approaches. Nevertheless, it has become clear that we need to reinforce the importance of analytical chemistry as an aid to synthetic biology and request appropriate infrastructure and funding schemes to equip laboratories with suitable analytical facilities.

In summary, we anticipate that future achievements in pathway design, enzyme discovery, engineering and immobilisation, will benefit substantially the application of multi-enzyme biocatalysis for industrial manufacturing and extend the portfolio of renewable and sustainable biomanufacturing processes. The increasing availability of increased computing power and machine learning will provide a platform for more stochastic tools and algorithms and reinforce bioinformatics as a central part of cell-free biocatalysis design and application. The implementation of high-throughput, automated system analysis will ensure that maximum performance is achieved. Ultimately, a vital cross-talk of mathematical models and experimental data in real time in form of smart soft sensors could be undertaken to achieve a fully automated process control and optimisation (Randek and Mandenius, 2017).

Acknowledgements

KP and DK are supported by an International Macquarie University Research Excellence Scholarship(iMQRES). AC is supported by a Cancer Institute New South Wales Early Career Fellowship(Project Number: ECF171114).

References

Aden, A., Ruth, M., Ibsen, K., Jechura, J., Neeves, K., Sheehan, J., Wallace, B., Montague, L., Slayton, A., Lukas, J., 2002. Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover. National Renewable Energy Lab., Golden, CO, USA.
Alonso, A., Marsal, S., Julià, A., 2015. Analytical methods in untargeted metabolomics: state of the art in 2015. Front. Bioeng. Biotechnol. 3, 23.
Alter, T., Blank, L., Ebert, B., 2018. Genetic optimization algorithm for metabolic engineering revisited. Metabolites 8, 33.
Andexer, J.N., Richter, M., 2015. Emerging enzymes for ATP regeneration in biocatalytic processes. Chembiochem 16, 380–386.
Ardao, I., Zeng, A.-P., 2013. *In silico* evaluation of a complex multi-enzymatic system using one-pot and modular approaches: application to the high-yield production of hydrogen from a synthetic metabolic pathway. Chem. Eng. Sci. 87, 183–193.
Barbosa, O., Ortiz, C., Berenguer-Murcia, Á., Torres, R., Rodrigues, R.C., Fernandez-Lafuente, R., 2015. Strategies for the one-step immobilization–purification of enzymes as industrial biocatalysts. Biotechnol. Adv. 33, 435–456.
Bar-Even, A., Noor, E., Lewis, N.E., Milo, R., 2010. Design and analysis of synthetic carbon fixation pathways. Proc. Natl. Acad. Sci. U. S. A. 107, 8889–8894.
Bar-Even, A., Flamholz, A., Noor, E., Milo, R., 2012. Rethinking glycolysis: on the biochemical logic of metabolic pathways. Nat. Chem. Biol. 8, 509.
Bar-Even, A., Milo, R., Noor, E., Tawfik, D.S., 2015. The moderately efficient enzyme: futile encounters and enzyme floppiness. Biochemistry 54, 4969–4977.
Beer, B., Pick, A., Sieber, V., 2017. *In vitro* metabolic engineering for the production of alpha-ketoglutarate. Metab. Eng. 40, 5–13.
Begum, G., Goodwin, W.B., Deglee, B.M., Sandhage, K.H., Kröger, N., 2015. Compartmentalisation of enzymes for cascade reactions through biomimetic layer-by-layer mineralization. J. Mater. Chem. B 3, 5232–5240.
Bornscheuer, U.T., 2016. Biocatalysis: successfully crossing boundaries. Angew. Chem. Int. Ed. 55, 4372–4373.
Bornscheuer, U.T., 2018. The fourth wave of biocatalysis is approaching. Philos. Trans. Royal Soc. A. 376 (20170063).
Bornscheuer, U.T., Huisman, G.W., Kazlauskas, R.J., Lutz, S., Moore, J.C., Robins, K., 2012. Engineering the third wave of biocatalysis. Nature 485, 185.
Bujara, M., Schumperli, M., Pellaux, R., Heinemann, M., Panke, S., 2011. Optimization of a blueprint for *in vitro* glycolysis by metabolic real-time analysis. Nat. Chem. Biol. 7, 271–277.
Bumbrah, G.S., Sharma, R.M., 2016. Raman spectroscopy – basic principle, instrumentation and selected applications for the characterization of drugs of abuse. Egypt. J. Forensic Sci. 6, 209–215.
Care, A., Bergquist, P.L., Sunna, A., 2015. Solid-binding peptides: smart tools for nanobiotechnology. Trends Biotechnol. 33, 259–268.
Care, A., Bergquist, P.L., Sunna, A., 2016. Solid-binding peptides: immobilisation strategies for extremophile biocatalysis in biotechnology. In: Rampelotto, P.H. (Ed.), Biotechnology of Extremophiles: Advances and Challenges. Springer International Publishing, Cham, pp. 637–674.
Care, A., Petroll, K., Gibson, E.S.Y., Bergquist, P.L., Sunna, A., 2017. Solid-binding peptides for immobilisation of thermostable enzymes to hydrolyse biomass polysaccharides. Biotechnol. Biofuels. 10, 29.
Castellana, M., Wilson, M.Z., Xu, Y., Joshi, P., Cristea, I.M., Rabinowitz, J.D., Gitai, Z., Wingreen, N.S., 2014. Enzyme clustering accelerates processing of intermediates through metabolic channeling. Nat. Biotechnol. 32, 1011–1018.
Chado, G.R., Stoykovich, M.P., Kaar, J.L., 2016. Role of dimension and spatial arrangement on the activity of biocatalytic cascade reactions on scaffolds. ACS Catal. 6, 5161–5169.
Chen, H., Zhu, Z., Huang, R., Zhang, Y.P., 2016. Coenzyme engineering of a hyperthermophilic 6-phosphogluconate dehydrogenase from NADP+ to NAD+ with its application to biobatteries. Sci. Rep. 6, 36311.
Cheng, K., Zhang, F., Sun, F., Chen, H., Percival Zhang, Y.H., 2015. Doubling power output of starch biobattery treated by the most thermostable isoamylase from an Archaeon Sulfolobus tokodaii. Sci. Rep. 5, 13184.
Cheng, T., Liu, H., Zou, H., Chen, N., Shi, M., Xie, C., Zhao, G., Xian, M., 2017. Enzymatic process optimization for the *in vitro* production of isoprene from mevalonate. Microb. Cell Factories 16, 8.
Cherubini, F., 2010. The biorefinery concept: using biomass instead of oil for producing energy and chemicals. Energy Convers. Manag. 51, 1412–1421.
Chong, S., 2014. Overview of cell-free protein synthesis: historic landmarks, commercial systems, and expanding applications. Curr. Protoc. Mol. Biol. 108, 16–30.
Cira, L.A., Huerta, S., Hall, G.M., Shirai, K., 2002. Pilot scale lactic acid fermentation of shrimp wastes for chitin recovery. Process Biochem. 37, 1359–1366.
Cleland, W.W., 2005. The use of isotope effects to determine enzyme mechanisms. Arch. Biochem. Biophys. 433, 2–12.

Corey, E.J., 1988. Robert Robinson lecture. Retrosynthetic thinking-essentials and examples. Chem. Soc. Rev. 17, 111–133.
Daniel, R.M., Cowan, D.A., 2000. Biomolecular stability and life at high temperatures. Cell. Mol. Life Sci. 57, 250–264.
Daniel, R.M., Danson, M.J., 2010. A new understanding of how temperature affects the catalytic activity of enzymes. Trends Biochem. Sci. 35, 584–591.
Daniel, R.M., Danson, M.J., 2013. Temperature and the catalytic activity of enzymes: a fresh understanding. FEBS Lett. 587, 2738–2743.
Daniel, R.M., Danson, M.J., Hough, D.W., Lee, C.K., Peterson, M.E., Cowan, D.A., 2008. Enzyme stability and activity at high temperatures. Protein Adaptation in Extremophiles. 1–34.
Daniel, R.M., Peterson, M.E., Danson, M.J., Price, N.C., Kelly, S.M., Monk, C.R., Weinberg, C.S., Oudshoorn, M.L., Lee, C.K., 2009. The molecular basis of the effect of temperature on enzyme activity. Biochem. J. 425, 353–360.
Delépine, B., Duigou, T., Carbonell, P., Faulon, J.-L., 2017. RetroPath2.0: a retrosynthesis workflow for metabolic engineers. Metab. Eng. 45, 158–170.
Desantis, G., Wong, K., Farwell, B., Chatman, K., Zhu, Z., Tomlinson, G., Huang, H., Tan, X., Bibbs, L., Chen, P., Kretz, K., Burk, M.J., 2003. Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM). J. Am. Chem. Soc. 125, 11476–11477.
Devries, J.K., Zubay, G., 1967. DNA-directed peptide synthesis. II. The synthesis of the alpha-fragment of the enzyme beta-galactosidase. Proc. Natl. Acad. Sci. U. S. A. 57, 1010–1012.
Dudley, Q.M., Anderson, K.C., Jewett, M.C., 2016. Cell-free mixing of *Escherichia coli* crude extracts to prototype and rationally engineer high-titer mevalonate synthesis. ACS Synth. Biol. 5, 1578–1588.
Eisenthal, R., Peterson, M.E., Daniel, R.M., Danson, M.J., 2006. The thermal behaviour of enzyme activity: implications for biotechnology. Trends Biotechnol. 24, 289–292.
Elias, M., Wieczorek, G., Rosenne, S., Tawfik, D.S., 2014. The universality of enzymatic rate-temperature dependency. Trends Biochem. Sci. 39, 1–7.
Emwas, A.-H., Salek, R., Griffin, J., Merzaban, J., 2013. NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations. Metabolomics 9, 1048–1072.
Feher, T., Planson, A.G., Carbonell, P., Fernandez-Castane, A., Grigoras, I., Dariy, E., Perret, A., Faulon, J.L., 2014. Validation of RetroPath, a computer-aided design tool for metabolic pathway engineering. Biotechnol. J. 9, 1446–1457.
Findrik, Z., Vasic-Racki, D., 2007. Biotransformation of D-methionine into L-methionine in the cascade of four enzymes. Biotechnol. Bioeng. 98, 956–967.
Findrik, Z., Vasic'-Racki, D., Lutz, S., Daussmann, T., Wandrey, C., 2005. Kinetic modeling of acetophenone reduction catalyzed by alcohol dehydrogenase from *Thermoanaerobacter* sp. Biotechnol. Lett. 27, 1087–1095.
Fisk, H., Xu, Y., Westley, C., Turner, N.J., Micklefield, J., Goodacre, R., 2017. From multistep enzyme monitoring to whole-cell biotransformations: development of real-time ultraviolet resonance Raman spectroscopy. Anal. Chem. 89, 12527–12532.
Fontaine, N., Grondin-Perez, B., Cadet, F., Offmann, B., 2015. Modeling of a cell-free synthetic system for biohydrogen production. Omics International 8, 132–139.
France, S.P., Hepworth, L.J., Turner, N.J., Flitsch, S.L., 2017. Constructing biocatalytic cascades: *In vitro* and *in vivo* approaches to de novo multi-enzyme pathways. ACS Catal. 7, 710–724.
Friehs, K., 2004. Plasmid copy number and plasmid stability. In: Scheper, T. (Ed.), New Trends and Developments in Biochemical Engineering. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 47–82.
Fujisawa, T., Fujinaga, S., Atomi, H., 2017. An in vitro enzyme system for the production of *myo*-inositol from starch. Appl. Environ. Microbiol. 83 (e00550-00517).
Gao, H., Tiwari, M.K., Kang, Y.C., Lee, J.-K., 2012. Characterization of H2O-forming NADH oxidase from *Streptococcus pyogenes* and its application in l-rare sugar production. Bioorg. Med. Chem. Lett. 22, 1931–1935.
Gao, C., Li, Z., Zhang, L., Wang, C., Li, K., Ma, C., Xu, P., 2015. An artificial enzymatic reaction cascade for a cell-free bio-system based on glycerol. Green Chem. 17, 804–807.
Geueke, B., Riebel, B., Hummel, W., 2003. NADH oxidase from *Lactobacillus brevis*: a new catalyst for the regeneration of NAD. Enzym. Microb. Technol. 32, 205–211.
Goldsmith, M., Tawfik, D.S., 2017. Enzyme engineering: reaching the maximal catalytic efficiency peak. Curr. Opin. Struct. Biol. 47, 140–150.
Gooday, G.W., 1990. The ecology of chitin degradation. In: Marshall, K.C. (Ed.), Adv. Microb. Ecol. Springer US, Boston, MA, pp. 387–430.
Goto, S., Bono, H., Ogata, H., Fujibuchi, W., Nishioka, T., Sato, K., Kanehisa, M., 1997. Organizing and computing metabolic pathway data in terms of binary relations. Pac. Symp. Biocomput. 175–186.
Green, A.P., Turner, N.J., 2016. Biocatalytic retrosynthesis: redesigning synthetic routes to high-value chemicals. Perspectives in Science. 9, 42–48.
Gruber, P., Marques, M.P.C., O'Sullivan, B., Baganz, F., Wohlgemuth, R., Szita, N., 2017. Conscious coupling: the challenges and opportunities of cascading enzymatic microreactors. Biotechnol. J. 12, 1700030.
Gumulya, Y., Gillam, E.M., 2017. Exploring the past and the future of protein evolution with ancestral sequence reconstruction: the 'retro' approach to protein engineering. Biochem. J. 474, 1–19.
Guo, W., Sheng, J., Feng, X., 2017. Mini-review: *in vitro* metabolic engineering for biomanufacturing of high-value products. Comput. Struct. Biotechnol. J. 15, 161–167.
Guterl, J.K., Garbe, D., Carsten, J., Steffler, F., Sommer, B., Reisse, S., Philipp, A., Haack, M., Ruhmann, B., Koltermann, A., Kettling, U., Bruck, T., Sieber, V., 2012. Cell-free metabolic engineering: production of chemicals by minimized reaction cascades. ChemSusChem 5, 2165–2172.
Hadadi, N., Hatzimanikatis, V., 2015. Design of computational retrobiosynthesis tools for the design of *de novo* synthetic pathways. Curr. Opin. Chem. Biol. 28, 99–104 Supplement C.
Hammer, S.C., Knight, A.M., Arnold, F.H., 2017. Design and evolution of enzymes for non-natural chemistry. Curr. Opin. Green Sustain. Chem. 7, 23–30.
Hogan, M.C., Woodley, J.M., 2000. Modelling of two enzyme reactions in a linked cofactor recycle system for chiral lactone synthesis. Chem. Eng. Sci. 55, 2001–2008.
Hold, C., Billerbeck, S., Panke, S., 2016. Forward design of a complex enzyme cascade reaction. Nat. Commun. 7, 12971.
Honda, K., Hara, N., Cheng, M., Nakamura, A., Mandai, K., Okano, K., Ohtake, H., 2016. *In vitro* metabolic engineering for the salvage synthesis of NAD. Metab. Eng. 35, 114–120.
Honda, K., Kimura, K., Ninh, P.H., Taniguchi, H., Okano, K., Ohtake, H., 2017. *In vitro* bioconversion of chitin to pyruvate with thermophilic enzymes. J. Biosci. Bioeng. 124, 296–301.
Honig, M., Sondermann, P., Turner, N.J., Carreira, E.M., 2017. Enantioselective chemo- and biocatalysis: partners in retrosynthesis. Angew. Chem. Int. Ed. Engl. 56, 8942–8973.
Huang, C.J., Lin, H., Yang, X., 2012. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. J. Ind. Microbiol. Biotechnol. 39, 383–399.
Huang, R., Chen, H., Zhong, C., Kim, J.E., Zhang, Y.-H.P., 2016. High-throughput screening of coenzyme preference change of thermophilic 6-phosphogluconate dehydrogenase from NADP(+) to NAD(+). Sci. Rep. 6, 32644.
Hull, S.R., Yang, B.Y., Venzke, D., Kulhavy, K., Montgomery, R., 1996. Composition of corn steep water during steeping. J. Agric. Food Chem. 44, 1857–1863.
Ishii, N., Suga, Y., Hagiya, A., Watanabe, H., Mori, H., Yoshino, M., Tomita, M., 2007. Dynamic simulation of an *in vitro* multi-enzyme system. FEBS Lett. 581, 413–420.
Jewett, M.C., Calhoun, K.A., Voloshin, A., Wuu, J.J., Swartz, J.R., 2008. An integrated cell-free metabolic platform for protein production and synthetic biology. Mol. Syst. Biol. 4, 220.
Jia, F., Narasimhan, B., Mallapragada, S., 2014. Materials-based strategies for multi-enzyme immobilization and co-localization: a review. Biotechnol. Bioeng. 111, 209–222.
Johnson, D.T., Taconi, K.A., 2007. The glycerin glut: Options for the value-added conversion of crude glycerol resulting from biodiesel production. Environ. Prog. 26, 338–348.
Karim, A.S., Jewett, M.C., 2016. A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. Metab. Eng. 36, 116–126.
Karim, A.S., Jewett, M.C., 2018. Cell-free synthetic biology for pathway prototyping. Methods Enzymol. 608, 31–57.
Kautz, R.A., Karger, B.L., Goetzinger, W.K., 2005. High-throughput microcoil NMR of compound libraries using zero-dispersion segmented flow analysis. J. Comb. Chem. 7, 14–20.
Keasling, J.D., Mendoza, A., Baran, P.S., 2012. A constructive debate. Nature 492, 188.
Kenney, K.L., Smith, W.A., Gresham, G.L., Westover, T.L., 2013. Understanding biomass feedstock variability. Biofuels. 4, 111–127.
Klein, S., Heinzle, E., 2012. Isotope labeling experiments in metabolomics and fluxomics. Wiley Interdiscip. Rev. Syst. Biol. Med. 4, 261–272.
Korman, T.P., Sahachartsiri, B., Li, D., Vinokur, J.M., Eisenberg, D., Bowie, J.U., 2014. A synthetic biochemistry system for the *in vitro* production of isoprene from glycolysis intermediates. Protein Sci. 23, 576–585.
Korman, T.P., Opgenorth, P.H., Bowie, J.U., 2017. A synthetic biochemistry platform for cell free production of monoterpenes from glucose. Nat. Commun. 8, 15526.
Koutinas, A.A., Vlysidis, A., Pleissner, D., Kopsahelis, N., Lopez Garcia, I., Kookos, I.K., Papanikolaou, S., Kwan, T.H., Lin, C.S., 2014. Valorization of industrial waste and by-product streams via fermentation for the production of chemicals and biopolymers. Chem. Soc. Rev. 43, 2587–2627.
Krüger, A., Schäfers, C., Schröder, C., Antranikian, G., 2018. Towards a sustainable bio-based industry – Highlighting the impact of extremophiles. New Biotech. 40, 1–184.
Krutsakorn, B., Honda, K., Ye, X., Imagawa, T., Bei, X., Okano, K., Ohtake, H., 2013. *In vitro* production of n-butanol from glucose. Metab. Eng. 20 (Supplement C), 84–91.
Kuchler, A., Yoshimoto, M., Luginbuhl, S., Mavelli, F., Walde, P., 2016. Enzymatic reactions in confined environments. Nat. Nano. 11, 409–420.
Kumar, A., Wang, L., Ng, C.Y., Maranas, C.D., 2018. Pathway design using *de novo* steps through uncharted biochemical spaces. Nat. Commun. 9, 184.
Lam, M.K., Lee, K.T., Mohamed, A.R., 2010. Homogeneous, heterogeneous and enzymatic catalysis for transesterification of high free fatty acid oil (waste cooking oil) to biodiesel: a review. Biotechnol. Adv. 28, 500–518.
Li, Z., Zhang, M., Jiang, T., Sheng, B., Ma, C., Xu, P., Gao, C., 2017. Enzymatic cascades for efficient biotransformation of racemic lactate derived from corn steep water. ACS Sustain. Chem. Eng. 5, 3456–3464.
Liang, K., Ricco, R., Doherty, C.M., Styles, M.J., Bell, S., Kirby, N., Mudie, S., Haylock, D., Hill, A.J., Doonan, C.J., Falcaro, P., 2015. Biomimetic mineralization of metal-organic frameworks as protective coatings for biomacromolecules. Nat. Commun. 6, 7240.
Liang, H., Jiang, S., Yuan, Q., Li, G., Wang, F., Zhang, Z., Liu, J., 2016. Co-immobilization of multiple enzymes by metal coordinated nucleotide hydrogel nanofibers: improved stability and an enzyme cascade for glucose detection. Nanoscale 8, 6071–6078.
Liese, A., Hilterhaus, L., 2013. Evaluation of immobilized enzymes for industrial applications. Chem. Soc. Rev. 42, 6236–6249.
Liese, A., Zelinski, T., Kula, M.R., Kierkels, H., Karutz, M., Kragl, U., Wandrey, C., 1998. A novel reactor concept for the enzymatic reduction of poorly soluble ketones. J. Mol. Catal. B Enzym. 4, 91–99.
Lin, J.L., Wagner, J.M., Alper, H.S., 2017. Enabling tools for high-throughput detection of metabolites: metabolic engineering and directed evolution applications. Biotechnol. Adv. 35, 950–970.
Liu, Z., Zhang, J., Chen, X., Wang, P.G., 2002. Combined biosynthetic pathway for *de novo* production of UDP-galactose: catalysis with multiple enzymes immobilized on

agarose beads. Chembiochem 3, 348–355.
Liu, F., Banta, S., Chen, W., 2013. Functional assembly of a multi-enzyme methanol oxidation cascade on a surface-displayed trifunctional scaffold for enhanced NADH production. Chem. Commun (Camb). 49, 3766–3768.
Liu, Z., Zhang, Y., Jia, X., Hu, M., Deng, Z., Xu, Y., Liu, T., 2017. *In vitro* reconstitution and optimization of the entire pathway to convert glucose into fatty acid. ACS Synth. Biol. 6, 701–709.
Lu, F., Smith, P.R., Mehta, K., Swartz, J.R., 2015. Development of a synthetic pathway to convert glucose to hydrogen using cell free extracts. Int. J. Hydrog. Energy 40, 9113–9124.
Lynd, L.R., Wyman, C.E., Gerngross, T.U., 1999. Biocommodity engineering. Biotechnol. Prog. 15, 777–793.
McAtee, A.G., Jazmin, L.J., Young, J.D., 2015. Application of isotope labeling experiments and 13C flux analysis to enable rational pathway engineering. Curr. Opin. Biotechnol. 36 (C), 50–56.
McCall, J., 2005. Genetic algorithms for modelling and optimisation. J. Comput. Appl. Math. 184, 205–222.
Milker, S., Fink, M.J., Oberleitner, N., Ressmann, A.K., Bornscheuer, U.T., Mihovilovic, M.D., Rudroff, F., 2017. Kinetic modeling of an enzymatic redox cascade *in vivo* reveals bottlenecks caused by cofactors. ChemCatChem 9, 3420–3427.
Myung, S., Zhang, Y.H.P., 2013. Non-complexed four cascade enzyme mixture: simple purification and synergetic co-stabilization. PLoS One 8, e61500.
Nguyen, Q.-T., Merlo, M.E., Medema, M.H., Jankevics, A., Breitling, R., Takano, E., 2012. Metabolomics methods for the synthetic biology of secondary metabolism. FEBS Lett. 586 (15), 2177–2183.
Ninh, P.H., Honda, K., Sakai, T., Okano, K., Ohtake, H., 2015. Assembly and multiple gene expression of thermophilic enzymes in *Escherichia coli* for *in vitro* metabolic engineering. Biotechnol. Bioeng. 112, 189–196.
Nirenberg, M.W., Matthaei, J.H., 1961. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. Proc. Natl. Acad. Sci. U. S. A. 47, 1588–1602.
Nocek, B., Kochinyan, S., Proudfoot, M., Brown, G., Evdokimova, E., Osipiuk, J., Edwards, A.M., Savchenko, A., Joachimiak, A., Yakunin, A.F., 2008. Polyphosphate-dependent synthesis of ATP and ADP by the family-2 polyphosphate kinases in bacteria. Proc. Natl. Acad. Sci. U. S. A. 105, 17730–17735.
Nowak, C., Beer, B., Pick, A., Roth, T., Lommes, P., Sieber, V., 2015. A water-forming NADH oxidase from *Lactobacillus pentosus* suitable for the regeneration of synthetic biomimetic cofactors. Front. Microbiol. 6, 957.
Nowak, C., Pick, A., Lommes, P., Sieber, V., 2017. Enzymatic reduction of nicotinamide biomimetic cofactors using an engineered glucose dehydrogenase: providing a regeneration system for artificial cofactors. ACS Catal. 7, 5202–5208.
Opgenorth, P.H., Korman, T.P., Bowie, J.U., 2014. A synthetic biochemistry molecular purge valve module that maintains redox balance. Nat. Commun. 5, 4113.
Opgenorth, P.H., Korman, T.P., Bowie, J.U., 2016. A synthetic biochemistry module for production of bio-based chemicals from glucose. Nat. Chem. Biol. 12 (6), 393–395.
Patel, S.K.S., Otari, S.V., Chan Kang, Y., Lee, J.-K., 2017. Protein-inorganic hybrid system for efficient his-tagged enzymes immobilization and its application in l-xylulose production. RSC Adv. 7 (6), 3488–3494.
Paul, C.E., Hollmann, F., 2016. A survey of synthetic nicotinamide cofactors in enzymatic processes. Appl. Microbiol. Biotechnol. 100, 4773–4778.
Peterson, M.E., Eisenthal, R., Danson, M.J., Spence, A., Daniel, R.M., 2004. A new intrinsic thermal parameter for enzymes reveals true temperature optima. J. Biol. Chem. 279, 20717–20722.
Planchestainer, M., Contente, M.L., Cassidy, J., Molinari, F., Tamborini, L., Paradisi, F., 2017. Continuous flow biocatalysis: production and in-line purification of amines by immobilised transaminase from *Halomonas elongata*. Green Chem. 19, 372–375.
Pleissner, D., Lin, C.S.K., 2013. Valorisation of food waste in biotechnological processes. Sustain. Chem. Process. 1, 21.
Poppe, L., Vértessy, B.G., 2018. The fourth wave of biocatalysis emerges—the 13th international symposium on biocatalysis and biotransformations. Chembiochem 19, 284–287.
Poux, S., Arighi, C.N., Magrane, M., Bateman, A., Wei, C.-H., Lu, Z., Boutet, E., Bye-A-Jee, H., Famiglietti, M.L., Roechert, B., UniProt Consortium, T., 2017. On expert curation and scalability: UniProtKB/Swiss-Prot as a case study. Bioinformatics 33, 3454–3460.
Quin, M.B., Wallin, K.K., Zhang, G., Schmidt-Dannert, C., 2017. Spatial organization of multi-enzyme biocatalytic cascades. Org. Biomol. Chem. 15, 4260–4271.
Randek, J., Mandenius, C.-F., 2017. On-line soft sensing in upstream bioprocessing. Crit. Rev. Biotechnol. 1–16.
Reetz, M.T., 2013. Biocatalysis in organic chemistry and biotechnology: past, present, and future. J. Am. Chem. Soc. 135, 12480–12496.
Renata, H., Wang, Z.J., Arnold, F.H., 2015. Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. Angew. Chem. Int. Ed. Engl. 54, 3351–3367.
Ríos-Solís, L., Bayir, N., Halim, M., Du, C., Ward, J.M., Baganz, F., Lye, G.J., 2013. Non-linear kinetic modelling of reversible bioconversions: Application to the transaminase catalyzed synthesis of chiral amino-alcohols. Biochem. Eng. J. 73, 38–48.
Risso, V.A., Sanchez-Ruiz, J.M., Ozkan, S.B., 2018. Biotechnological and protein-engineering implications of ancestral protein resurrection. Curr. Opin. Struct. Biol. 51, 106–115.
Rocha-Martín, J., Rivas, B.d.l., Muñoz, R., Guisán, J.M., López-Gallego, F., 2012. Rational co-immobilization of bi-enzyme cascades on porous supports and their applications in bio-redox reactions with *in situ* recycling of soluble cofactors. ChemCatChem 4, 1279–1288.
Rollin, J.A., Martin Del Campo, J., Myung, S., Sun, F., You, C., Bakovic, A., Castro, R., Chandrayan, S.K., Wu, C.-H., Adams, M.W.W., Senger, R.S., Zhang, Y.-H.P., 2015. High-yield hydrogen production from biomass by *in vitro* metabolic engineering: mixed sugars coutilization and kinetic modeling. Proc. Natl. Acad. Sci. U. S. A. 112, 4964–4969.
Rosenblum, G., Cooperman, B.S., 2014. Engine out of the chassis: cell-free protein synthesis and its uses. FEBS Lett. 588, 261–268.
Rudroff, F., Mihovilovic, M.D., Gröger, H., Snajdrova, R., Iding, H., Bornscheuer, U.T., 2018. Opportunities and challenges for combining chemo- and biocatalysis. Nat. Catal. 1, 12–22.
Savile, C.K., Janey, J.M., Mundorff, E.C., Moore, J.C., Tam, S., Jarvis, W.R., Colbeck, J.C., Krebber, A., Fleitz, F.J., Brands, J., Devine, P.N., Huisman, G.W., Hughes, G.J., 2010. Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. Science 329, 305–309.
Scherkus, C., Schmidt, S., Bornscheuer, U.T., Groger, H., Kara, S., Liese, A., 2017. Kinetic insights into -caprolactone synthesis: Improvement of an enzymatic cascade reaction. Biotechnol. Bioeng. 114, 1215–1221.
Schmidt, E., Vasić-Rački, Đ., Wandrey, C., 1987. Enzymatic production of l-phenylalanine from the racemic mixture of D,L-phenyllactate. Appl. Microbiol. Biotechnol. 26, 42–48.
Schmidt-Dannert, C., Lopez-Gallego, F., 2016. A roadmap for biocatalysis - functional and spatial orchestration of enzyme cascades. Microb. Biotechnol. 9, 601–609.
Schwander, T., Schada Von Borzyskowski, L., Burgener, S., Cortina, N.S., Erb, T.J., 2016. A synthetic pathway for the fixation of carbon dioxide *in vitro*. Science 354, 900–904.
Secundo, F., 2013. Conformational changes of enzymes upon immobilisation. Chem. Soc. Rev. 42, 6250–6261.
Sheldon, R.A., Brady, D., 2018. The limits to biocatalysis: pushing the envelope. Chem. Commun. 54, 6088–6104.
Sheldon, R.A., Pereira, P.C., 2017. Biocatalysis engineering: the big picture. Chem. Soc. Rev. 46, 2678–2691.
Sheldon, R.A., van Pelt, S., 2013. Enzyme immobilisation in biocatalysis: why, what and how. Chem. Soc. Rev. 42, 6223–6235.
Shi, J., Wu, Y., Zhang, S., Tian, Y., Yang, D., Jiang, Z., 2018. Bioinspired construction of multi-enzyme catalytic systems. Chem. Soc. Rev. 47, 4295–4313.
Shimizu, Y., Kanamori, T., Ueda, T., 2005. Protein synthesis by pure translation systems. Methods 36, 299–304.
Sokic-Lazic, D., Minteer, S.D., 2008. Citric acid cycle biomimic on a carbon electrode. Biosens. Bioelectron. 24, 939–944.
de Souza, R., Miranda, L.S.M., Bornscheuer, U.T., 2017. A retrosynthesis approach for biocatalysis in organic synthesis. Chemistry 23, 12040–12063.
Sperl, J.M., Sieber, V., 2018. Multienzyme cascade reactions—status and recent advances. ACS Catal. 8, 2385–2396.
Steffler, F., Guterl, J.K., Sieber, V., 2013. Improvement of thermostable aldehyde dehydrogenase by directed evolution for application in synthetic cascade biomanufacturing. Enzym. Microb. Technol. 53, 307–314.
Sunna, A., Chi, F., Bergquist, P.L., 2013. A linker peptide with high affinity towards silica-containing materials. New Biotechnol. 30, 485–492.
Swartz, J., 2006. Developing cell-free biology for industrial applications. J. Ind. Microbiol. Biotechnol. 33, 476–485.
Swartz, M., 2010. HPLC detectors: a brief review. J. Liq. Chromatogr. Relat. Technol. 33, 1130–1150.
Swartz, J.R., 2012. Transforming biochemical engineering with cell-free biology. AICHE J. 58, 5–13.
Taniguchi, H., Okano, K., Honda, K., 2017. Modules for *in vitro* metabolic engineering: Pathway assembly for bio-based production of value-added chemicals. Synth. Syst. Biotechnol. 2, 65–74.
Tawfik, D.S., 2014. Accuracy-rate tradeoffs: how do enzymes meet demands of selectivity and catalytic efficiency? Curr. Opin. Chem. Biol. 21, 73–80.
Teusink, B., Passarge, J., Reijenga, C.A., Esgalhado, E., van der Weijden, C.C., Schepper, M., Walsh, M.C., Bakker, B.M., van Dam, K., Westerhoff, H.V., Snoep, J.L., 2000. Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. Eur. J. Biochem. 267, 5313–5329.
Torres, P., Batista-Viera, F., 2017. Immobilized trienzymatic system with enhanced stabilization for the biotransformation of lactose. Molecules 22, E284.
Turner, N.J., O'Reilly, E., 2013. Biocatalytic retrosynthesis. Nat. Chem. Biol. 9, 285–288 (UNEP. 2014. United Nations Environment Programme 2014 Annual Report).
Van der Donk, W.A., Zhao, H., 2003. Recent developments in pyridine nucleotide regeneration. Curr. Opin. Biotechnol. 14, 421–426.
Van Hecke, W., Bhagwat, A., Ludwig, R., Dewulf, J., Haltrich, D., Van Langenhove, H., 2009. Kinetic modeling of a bi-enzymatic system for efficient conversion of lactose to lactobionic acid. Biotechnol. Bioeng. 102, 1475–1482.
Vasic-Racki, D., Kragl, U., Liese, A., 2003. Benefits of enzyme kinetics modelling. Chem. Biochem. Eng. Q. 17, 7–18.
Vivek, N., Sindhu, R., Madhavan, A., Anju, A.J., Castro, E., Faraco, V., Pandey, A., Binod, P., 2017. Recent advances in the production of value added chemicals and lipids utilizing biodiesel industry generated crude glycerol as a substrate–Metabolic aspects, challenges and possibilities: an overview. Bioresour. Technol. 239, 507–517.
Wang, L., Dash, S., Ng, C.Y., Maranas, C.D., 2017a. A review of computational tools for design and reconstruction of metabolic pathways. Synth. Syst. Biotechnol. 2, 243–252.
Wang, W., Liu, M., You, C., Li, Z., Zhang, Y.-H.P., 2017b. ATP-free biosynthesis of a high-energy phosphate metabolite fructose 1,6-diphosphate by *in vitro* metabolic engineering. Metab. Eng. 42, 168–174.
Watanabe, S., Saimura, M., Makino, K., 2008. Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. J. Biol. Chem. 283, 20372–20382.
Wayman, J., Sagar, A., Varner, J., 2015. Dynamic modeling of cell-free biochemical networks using effective kinetic models. Processes 3, 138.
Wheeldon, I., Minteer, S.D., Banta, S., Barton, S.C., Atanassov, P., Sigman, M., 2016.

ARTICLE IN PRESS

Substrate channelling as an approach to cascade reactions. Nat. Chem. 8, 299–309.
Whittaker, J.W., 2013. Cell-free protein synthesis: the state of the art. Biotechnol. Lett. 35, 143–152.
Wiegel, J., Michael, A.W., 2002. Thermophiles: The Keys to the Molecular Evolution and the Origin of Life? CRC Press, Protein Science.
Wilding, K.M., Schinn, S.-M., Long, E.A., Bundy, B.C., 2018. The emerging impact of cell-free chemical biosynthesis. Curr. Opin. Biotechnol. 53, 115–121.
Wolf, J., Stark, H., Fafenrot, K., Albersmeier, A., Pham, T.K., Muller, K.B., Meyer, B.H., Hoffmann, L., Shen, L., Albaum, S.P., Kouril, T., Schmidt-Hohagen, K., Neumann-Schaal, M., Brasen, C., Kalinowski, J., Wright, P.C., Albers, S.V., Schomburg, D., Siebers, B., 2016. A systems biology approach reveals major metabolic changes in the thermoacidophilic archaeon *Sulfolobus solfataricus* in response to the carbon source L-fucose versus D-glucose. Mol. Microbiol. 102, 882–908.
Woodward, J., Orr, M., Cordray, K., Greenbaum, E., 2000. Biotechnology: enzymatic production of biohydrogen. Nature 405, 1014–1015.
Wu, J.T., Wu, L.H., Knight, J.A., 1986. Stability of NADPH: effect of various factors on the kinetics of degradation. Clin. Chem. 32, 314–319.
Wu, X., Yang, C., Ge, J., 2017. Green synthesis of enzyme/metal-organic framework composites with high stability in protein denaturing solvents. Bioresour. Bioprocess. 4, 24.
Xiao, Z., Lu, J.R., 2014. Strategies for enhancing fermentative production of acetoin: a review. Biotechnol. Adv. 32, 492–503.
Yan, N., Chen, X., 2015. Sustainability: don't waste seafood waste. Nature 524, 155–157.
Yang, W.C., Patel, K.G., Wong, H.E., Swartz, J.R., 2012. Simplifying and streamlining *Escherichia coli*-based cell-free protein synthesis. Biotechnol. Prog. 28, 413–420.
Yang, T., Rao, Z., Zhang, X., Xu, M., Xu, Z., Yang, S.-T., 2017. Metabolic engineering strategies for acetoin and 2,3-butanediol production: advances and prospects. Crit. Rev. Biotechnol. 37, 990–1005.
Ye, X., Wang, Y., Hopkins, R.C., Adams, M.W., Evans, B.R., Mielenz, J.R., Zhang, Y.H., 2009. Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. ChemSusChem 2, 149–152.
Ye, X., Honda, K., Sakai, T., Okano, K., Omasa, T., Hirota, R., Kuroda, A., Ohtake, H., 2012. Synthetic metabolic engineering-a novel, simple technology for designing a chimeric metabolic pathway. Microb. Cell Factories 11, 120.
Ye, X., Honda, K., Morimoto, Y., Okano, K., Ohtake, H., 2013. Direct conversion of glucose to malate by synthetic metabolic engineering. J. Biotechnol. 164, 34–40.
You, C., Zhang, Y.H.P., 2013. Self-assembly of synthetic metabolons through synthetic protein scaffolds: one-step purification, co-immobilization, and substrate channeling. ACS Synth. Biol. 2, 102–110.
You, C., Myung, S., Zhang, Y.H.P., 2012. Facilitated substrate channeling in a self-assembled trifunctional enzyme complex. Angew. Chem. Int. Ed. Engl. 51, 8787–8790.
You, C., Shi, T., Li, Y., Han, P., Zhou, X., Zhang, Y.P., 2017. An in vitro synthetic biology platform for the industrial biomanufacturing of *myo*-inositol from starch. Biotechnol. Bioeng. 114, 1855–1864.
Zampieri, M., Sekar, K., Zamboni, N., Sauer, U., 2017. Frontiers of high-throughput metabolomics. Curr. Opin. Chem. Biol. 36, 15–23.
Zdarta, J., Meyer, A., Jesionowski, T., Pinelo, M., 2018. A general overview of support materials for enzyme immobilization: characteristics, properties, practical utility. Catalysts 8, 92.
Zhang, Y., Hess, H., 2017. Toward rational design of high-efficiency enzyme cascades. ACS Catal. 7, 6018–6027.
Zhang, Y., Tsitkov, S., Hess, H., 2016. Proximity does not contribute to activity enhancement in the glucose oxidase–horseradish peroxidase cascade. Nat. Commun. 7, 13982.
Zhang, Y.P., Sun, J., Ma, Y., 2017. Biomanufacturing: history and perspective. J. Ind. Microbiol. Biotechnol. 44, 773–784.
Zhao, H., van der Donk, W.A., 2003. Regeneration of cofactors for use in biocatalysis. Curr. Opin. Biotechnol. 14, 583–589.
Zhong, C., Wei, P., Zhang, Y.-H.P., 2017a. A kinetic model of one-pot rapid biotransformation of cellobiose from sucrose catalyzed by three thermophilic enzymes. Chem. Eng. Sci. 161, 159–166.
Zhong, C., You, C., Wei, P., Zhang, Y.-H.P., 2017b. Thermal cycling cascade biocatalysis of *myo*-inositol synthesis from sucrose. ACS Catal. 7, 5992–5999.
Zhu, Z., Zhang, Y.H.P., 2017. In vitro metabolic engineering of bioelectricity generation by the complete oxidation of glucose. Metab. Eng. 39, 110–116.
Zore, O.V., Pande, P., Okifo, O., Basu, A.K., Kasi, R.M., Kumar, C.V., 2017. Nanoarmoring: strategies for preparation of multi-catalytic enzyme polymer conjugates and enhancement of high temperature biocatalysis. RSC Adv. 7, 29563–29574.
Zubay, G., 1973. *In vitro* synthesis of protein in microbial systems. Annu. Rev. Genet. 7, 267–287.

Appendix II

catalysts

MDPI

Article

Characterisation of the First Archaeal Mannonate Dehydratase from *Thermoplasma acidophilum* and Its Potential Role in the Catabolism of D-Mannose

Dominik Kopp [1], Robert Willows [1,2] and Anwar Sunna [1,2,*]

[1] Department of Molecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia; dominik.kopp@hdr.mq.edu.au (D.K.); robert.willows@mq.edu.au (R.W.)
[2] Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, New South Wales 2109, Australia
* Correspondence: anwar.sunna@mq.edu.au; Tel.: +61-2-9850-4220

Received: 13 February 2019; Accepted: 24 February 2019; Published: 3 March 2019

check for updates

Abstract: Mannonate dehydratases catalyse the dehydration reaction from mannonate to 2-keto-3-deoxygluconate as part of the hexuronic acid metabolism in bacteria. Bacterial mannonate dehydratases present in this gene cluster usually belong to the xylose isomerase-like superfamily, which have been the focus of structural, biochemical and physiological studies. Mannonate dehydratases from archaea have not been studied in detail. Here, we identified and characterised the first archaeal mannonate dehydratase (TaManD) from the thermoacidophilic archaeon *Thermoplasma acidophilum*. The recombinant TaManD enzyme was optimally active at 65 °C and showed high specificity towards D-mannonate and its lactone, D-mannono-1,4-lactone. The gene encoding for TaManD is located adjacent to a previously studied mannose-specific aldohexose dehydrogenase (AldT) in the genome of *T. acidophilum*. Using nuclear magnetic resonance (NMR) spectroscopy, we showed that the mannose-specific AldT produces the substrates for TaManD, demonstrating the possibility for an oxidative metabolism of mannose in *T. acidophilum*. Among previously studied mannonate dehydratases, TaManD showed closest homology to enzymes belonging to the xylose isomerase-like superfamily. Genetic analysis revealed that closely related mannonate dehydratases among archaea are not located in a hexuronate gene cluster like in bacteria, but next to putative aldohexose dehydrogenases, implying a different physiological role of mannonate dehydratases in those archaeal species.

Keywords: mannonate dehydratase; mannose metabolism; *Thermoplasma acidophilum*; mannono-1,4-lactone; 2-keto-3-deoxygluconate; aldohexose dehydrogenase

1. Introduction

Mannonate dehydratases (EC 4.2.1.8) catalyse the conversion of mannonate to 2-keto-3-deoxygluconate (KDG) and have been studied as part of the hexuronic acid gene cluster in several organisms, such as *Escherichia coli*, *Bacillus stearothermophilus*, *Bacillus subtilis* and *Erwinia chrysanthemi* [1–4]. The hexuronate gene cluster encodes enzymes involved in the metabolism of glucuronate and galacturonate [5]. Glucuronate is a common sugar acid present in glucuronoxylan, a constituent of plant cell walls which can serve as the only carbon source for growth of some bacteria [3,6,7]. Glucuronate is also present in the mucus layer of mammals, providing a carbon source for anaerobic gut bacteria, such as *E. coli* [8,9]. As part of the hexuronate metabolism, mannonate dehydratase converts mannonate to KDG (Figure 1A). In *E. coli*, and some species of *Erwinia*, KDG is phosphorylated and cleaved into pyruvate and 3-phosphoglycerate, which can be further metabolised in the tricarboxylic acid cycle, or the Entner-Doudoroff pathway [4,5].

Mannonate dehydratases are represented in different enzyme families, such as the xylose isomerase-like superfamily or the enolase superfamily. Mannonate dehydratases encoded in the bacterial hexuronate gene clusters usually belong to the xylose isomerase-like superfamily. Crystal structures have been solved for xylose isomerase-like mannonate dehydratases from *Streptococcus suis*, *E. coli* K12 and *Enterococcus faecalis* in native form and in complex with Mn^{2+} ions [10,11]. His311 and Tyr325 in the binding pocket were identified as crucial for the activity of the mannonate dehydratases from Gram-positive bacteria, such as *S. suis*. However, in Gram-negative bacteria (e.g., *E. coli* K12), an additional inserted sequence in the binding pocket rendered the dehydratase less active [11]. Members of the enolase superfamily show a conserved structure and reaction mechanism, but differ in their physiological functions. Within their conserved barrel structure, mannonate dehydratases of this family share conserved ligand-binding sites for Mg^{2+}, which are essential for the stabilisation of the enediolate intermediate [12]. Among mannonate dehydratases of the enolase superfamily, some representatives with diverse functions have been found, which are involved in the catabolism of sugar acids other than glucuronate or galacturonate [13]. Several crystal structures have been solved for mannonate dehydratases from the enolase superfamily, including enzymes from *Chromohalobacter salexigens* and *Novosphingobium aromaticivorans* [12,14]. However, to the best of our knowledge, no archaeal mannonate dehydratase has been investigated so far.

Here, we used genomic analysis to identify the first functional mannonate dehydratase from the thermoacidophilic archaeon *Thermoplasma acidophilum* (TaManD). The archaeal mannonate dehydratase was recombinantly-expressed in *E. coli*, purified and functionally characterised. TaManD showed high amino acid sequence identity to bacterial mannonate dehydratases from the xylose isomerase-like superfamily. In the genome of *T. acidophilum*, the gene encoding for TaManD is located adjacent to an aldohexose dehydrogenase (AldT), which has been shown previously to catalyse the oxidation of mannose to mannonate with high specificity [15,16]. However, the physiological significance of the oxidation and its products were not investigated further. We identified the products of an AldT-mediated oxidation of D-mannose using NMR spectroscopy and confirmed that TaManD is able to convert the products to KDG. This demonstrates that in principle, a mannose metabolism based on AldT and ManD is possible in *T. acidophilum* (Figure 1B).

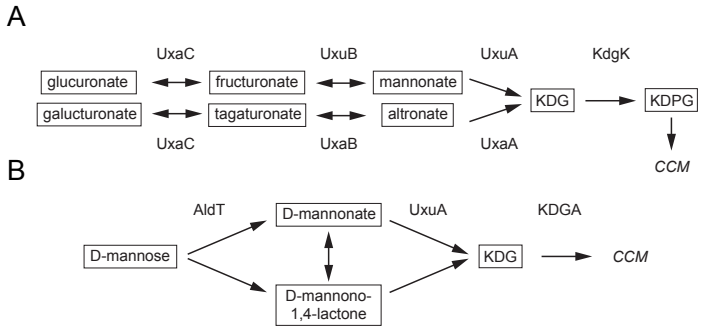


Figure 1. Hexuronate metabolism in *E. coli* and possible mannose catabolism in *T. acidophilum*. (**A**) Role of *uxuA* mannnonate dehdyratase in dissimilation of hexuronates in *E. coli* adapted from Peekhaus and Conway [8]. (**B**) Role of *uxuA* mannonate dehydratase in a possible mannose metabolism in *T. acidophilum*. UxaC: hexuronate isomerase, UxuB: mannonate oxidoreductase, UxaB: altronate oxidoreductase, UxuA: mannonate dehydratase, UxaA: altronate dehydratase, KdgK: 2-keto-3-deoxygluconate kinase, KDG: 2-keto-3-deoxygluconate, KDPG: 2-keto-3-deoxy-6-phosphogluconate, AldT: aldose dehydrogenase, KDGA: 2-keto-3-deoxygluconate aldolase, CCM: central carbon metabolism.

2. Results and Discussion

2.1. Screening for a Functional Mannonate Dehydratase in T. acidophilum

Based on the previously identified mannose-specific AldT, we searched the genomic neighbourhood of its gene locus (Ta0754) for enzymes which could display activity towards mannonate. In the proximity of Ta0754, several genes are located that encode for hypothetical dehydratases, suggesting their potential to convert mannonate to KDG (Table S1). We identified the protein product of the Ta0753 gene as a functional mannonate dehydratase from *T. acidophilum* by heterologous expression in *E. coli*, and compared its properties and genomic context to other previously studied mannonate dehydratases.

Among previously characterised mannonate dehydratases, TaManD shares highest protein sequence identity (31.5%) with enzymes in the xylose isomerase-like superfamily, such as the mannonate dehydratase from *S. suis* [10]. Key amino acid residues for substrate binding (His311 and Tyr325) and binding of the cofactor Mn^{2+} (Asp310, Cys237, His199 and His266) in *S. suis* are conserved in the amino acid sequence of TaManD. A much higher amino acid sequence identity (62.9%) is shared between TaManD and putative mannonate dehydratases from closely related archaeal species, *Ferroplasma acidarmanus* and *Ferroplasma acidiphilum*, suggesting an archaeal clade of mannonate dehydratases.

Phylogenetic analysis showed that mannonate dehydratases from *T. acidophilum*, *F. acidarmanus* and *F. acidiphilum* are more closely related to bacterial mannonate dehydratases from the xylose isomerase-like superfamily than to bacterial or archaeal mannonate dehydratases of the enolase superfamily (Figure 2). Mannonate dehydratases of the enolase superfamily have a substantially different structure from those in the xylose isomerase-like superfamily, and therefore, are only distantly related to TaManD. Despite fulfilling a similar function, the dihydroxy-acid dehydratase from *S. solfataricus*, which belongs to the IlvD/EDD superfamily, is rather unrelated to the xylose isomerase-like and enolase superfamilies [17].

For bacterial xylose isomerase-like mannonate dehydratases, a physiological function of the enzyme in the catabolism of hexuronates has been demonstrated [18,19]. Although TaManD is annotated as *uxuA* mannonate dehydratase, which implies a role in the metabolism of hexuronates, it is not present in a classical hexuronate gene cluster known from bacteria such as *E. coli*, *B. subtilis* or *B. stearothermophilus* (Figure 3) [2,3]. In contrast, the gene encoding for TaManD is located adjacent to AldT in the genome of *T. acidophilum*. Similarly, the two closely related putative mannonate dehydratases in *F. acidarmanus* and *F. acidiphilum* are also located adjacent to putative aldohexose dehydrogenases. Therefore, a different physiological role can be proposed for these mannonate dehydratases. In the following, we functionally characterise TaManD and show that AldT is able to produce the substrate needed for a subsequent conversion to KDG mediated by TaManD.

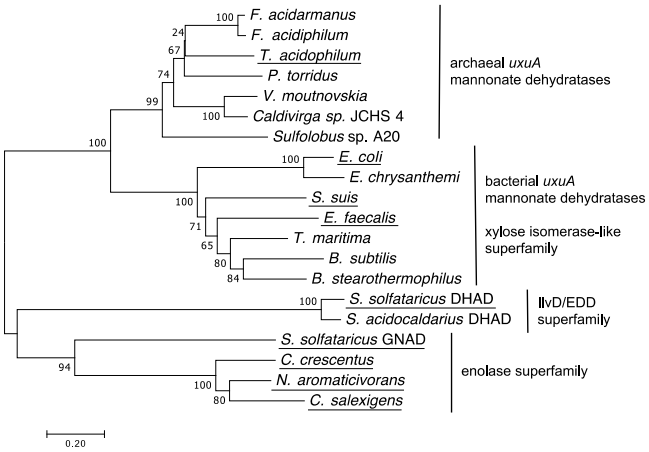


Figure 2. Phylogenetic relationship of different putative and confirmed dehydratases belonging to different enzyme families. Characterised enzymes are underlined. Evolutionary analyses were conducted in MEGA 7. 20 protein sequences were aligned using MUSCLE. The phylogenetic tree was inferred using the neighbour-joining method. The scale bar indicates an evolutionary distance of 0.20 nucleotide per position in the sequence. The number next to the nodes represent bootstrap confidence values estimated from 500 replicates. Protein sequences were retrieved from the Uniprot database and their entry codes are as follows. *Ferroplasma acidarmanus*: S0AL33, *Ferroplasma acidiphilum*: A0A1V0N416, *Thermoplasma acidophilum*: Q9HK52, *Picrophilus torridus*: Q6L2R9, *Vulcanisaeta moutnovskia*: F0QYL3, *Caldivirga* sp. JCHS 4: A0A101XEY7, *Sulfolobus* sp. A20: A0A1C8ZTN0, *Escherichia coli* K12: P24215, *Erwinia chrysanthemi*/*Dickeya dadantii*: E0SEP1, *Streptococcus suis*: A0A142UME2, *Enterococcus faecalis*: Q82ZC9, *Thermotoga maritima*: Q9WXS4, *Bacillus subtilis*: O34346, *Bacillus stearothermophilus*: A0A087LHB3, *Sulfolobus solfataricus* DHAD: Q97UB2, *Sulfolobus acidocaldarius* DHAD: Q4J860, *Sulfolobus solfataricus* GNAD: Q97U27, *Caulobacter crescentus*/*Caulobacter vibroides*: Q9AAR4, *Novosphingobium aromaticivorans*: A4XF23, *Chromohalobacter salexigens*: Q1QT89.

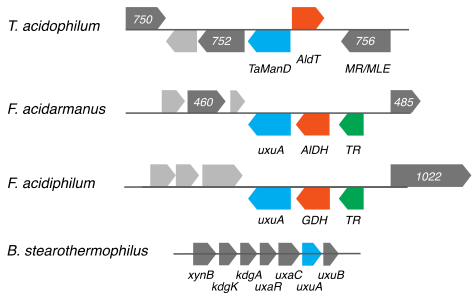


Figure 3. Schematic maps of uxuA mannonate dehydratases and their genomic neighbourhoods in *T. acidophilum*, *F. acidarmanus*, *F. acidiphilum* and *B. stearothermophilus*. Genes are annotated in the NCBI database as follows: In blue, *TaManD*/*uxuA*: *T. acidophilum* mannonate dehydratase/uxuA D-mannonate dehydratases; in orange, *AldT*/*AlDH*/*GDH*: aldohexose dehydrogenase/aldose dehydrogenase/glucose-1-dehydrogenase; in green, TR: transcriptional regulator; in light grey, annotated as hypothetical protein. Note: 750: FAA hydrolase family protein, 752: L-rhamnonate dehydratase, 756: mandelate racemase/muconate lactonizing enzyme family protein, 460: transposase, 485: acetyl-coenzyme A synthetase, 1022: acetate-CoA ligase, *xynB*: β-xylosidase, *kdgK*: 2-keto-3-deoxygluconate kinase, *kdgA*: 2-keto-3-deoxy-6-phosphogluconate aldolase, *uxaR*: regulatory protein, *uxaC*: hexuronate isomerase, *uxuB*: mannonate oxidoreductase.

2.2. Substrate Conformation for TaManD Activity

Before acquiring enzyme kinetics and studying the characteristics of TaManD, we investigated which substrate conformation is encountered by TaManD under physiological conditions. In aqueous solutions, free sugar acids coexist with their lactone in an equilibrium, which is defined by the stability of the lactone, the temperature and the pH of the solution. Many lactones hydrolyse spontaneously in water, although several lactones, including D-arabinonolactone [20], L-rhamnonolactone [21] and D-xylonolactone [22], have been reported to be hydrolysed by lactonases. Mannonate is able to form a lactone by covalent bond formation between carbon 1 and carbon 5 (δ-lactone) or carbon 1 and carbon 4 (γ-lactone) (Figure 4A). The equilibrium of the two different envelope forms has been studied by NMR spectroscopy and it was found that the γ-lactone is strongly favoured over the δ-lactone [23]. However, these studies do not describe equilibria between lactone and free acid form in a physiological buffer. In a cellular environment mannonate can either be produced by a mannonate oxidoreductase (UxuB) as part of the catabolism of hexuronic acids or by an aldohexose dehydrogenase, such as AldT, in a hypothetical oxidative mannose catabolism (Figure 1). In order to study the conformation of the substrate for TaManD in a physiological environment, we first acquired decoupled ^{13}C NMR spectra for D-mannono-1,4-lactone in physiological buffer at pH 7, in its lactone form (incubated with HCl) and after hydrolysis to the free sugar acid (incubated with NaOH) (Figure 4B). The spectra obtained indicated that the lactone and the free sugar acid can be distinguished by their chemical shifts. In the free sugar acid form, carbon 6 yields a chemical shift at 63.51 ppm, whereas in D-mannono-1,4-lactone, carbon 6 displays a chemical shift at 63.02 ppm. D-mannono-1,4-lactone in buffer shows both chemical shifts, indicating that D-mannono-1,4-lactone and D-mannonate are present in an equilibrium at pH 7.

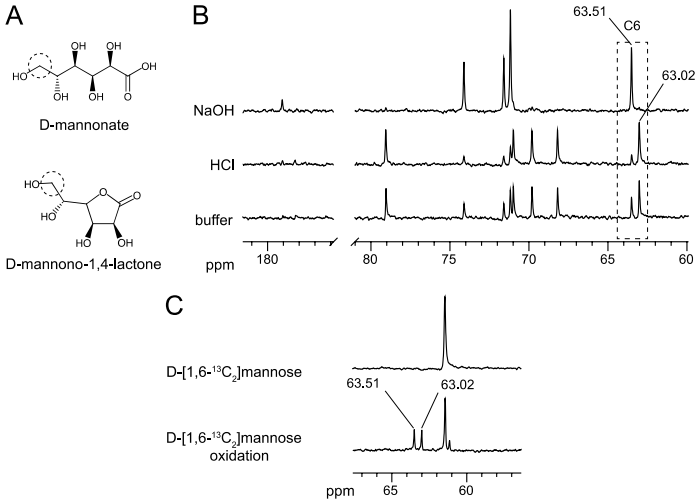


Figure 4. Analysis of D-mannono-1,4-lactone under different conditions and products of AldT-mediated oxidation of isotope labelled D-mannose. (**A**) Chemical structures of the free sugar acid, D-mannonate, and its lactone, D-mannono-1,4-lactone. Carbon 6, which was used to differentiate the two different substrate forms in NMR spectra, is circled for both structures. (**B**) Decoupled 1D ^{13}C NMR spectra of D-mannono-1,4-lactone after incubation in 1 M NaOH, 1 M HCl and in 0.1 M sodium phosphate (NaP) buffer pH 7. Chemical shift for carbon 6 is shown by a dashed rectangle. Characteristic chemical shifts for the lactone (63.02 ppm) and for the free sugar acid (63.51 ppm) are indicated. (**C**) Decoupled 1D ^{13}C NMR spectra before (**top**) and after (**bottom**) oxidation of labelled D-mannose with AldT. Chemical shifts of carbon 6 for lactone and free sugar acid are indicated.

Next, we oxidised isotope-labelled D-[1,6-$^{13}C_2$]mannose using recombinantly-expressed AldT and acquired 1D ^{13}C NMR spectra of the reaction product. The chemical shifts for carbon 6 of mannono-1,4-lactone and D-mannonate were used to analyse the reaction products of the mannose oxidation with AldT, as these were close together but well resolved and will have similar relaxation rates in both forms. The spectrum after partial AldT oxidation of D-[1,6-$^{13}C_2$]mannose (Figure 4C) shows that equal quantities of D-mannono-1,4-lactone and D-mannonate are produced for the subsequent dehydration by TaManD.

2.3. Expression and Purification of TaManD

For the further biochemical characterisation, TaManD was expressed together with a tobacco etch virus (TEV) sequence, which allowed the proteolytic cleavage of the His-tag after the first metal affinity purification. In order to study the effect of bivalent metal cofactors on enzyme activity, cleavage of the His-tag from the recombinant enzyme was necessary, since cations (Ni^{2+}, Mg^{2+}) might influence enzyme activity assays of a His-tagged enzyme. After His-tag cleavage, the recombinant enzyme showed a single band on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) with an apparent mass of 38 kDa (Figure S1). Peptide mass fingerprinting was performed with liquid chromatography-electrospray ionisation-tandem mass spectrometry (LC ESI MS/MS) and the National Center for Biotechnology Information (NCBI) database searches using the mascot software confirmed the identity of the purified protein as uxuA mannonate dehydratase from *T. acidophilum*. In addition, the apparent native molecular mass of the protein was estimated from size exclusion chromatography to be 225 kDa, indicating a hexameric structure of the protein.

2.4. Enzyme Characterisation

Initial tests showed that TaManD lost almost its complete activity after purification, suggesting that essential cofactors for activity were removed during the purification process. Accordingly, the effect of several additives was tested for their influence on the activity of TaManD with D-mannono-1,4-lactone. TaManD showed strongly enhanced activity in the presence of β-mercaptoethanol and the metal ions Mn^{2+}, Ni^{2+}, Mg^{2+}, Ca^{2+} and Co^{2+} (Figure 5). The highest activation (100%) was observed in the presence of β-mercaptoethanol, which was significantly higher than the activating bivalent cation Ni^{2+} (71.5% of maximum enzyme activity, p-value: 0.0022). Mn^{2+}, Mg^{2+} and Co^{2+} also activated TaManD and resulted in 50–60% increase in maximum activity when compared to the TaManD control (without any additive). No significant difference in activity was observed in the presence of Fe^{2+}, Zn^{2+}, Cu^{2+}, EDTA, DTT and glutathione.

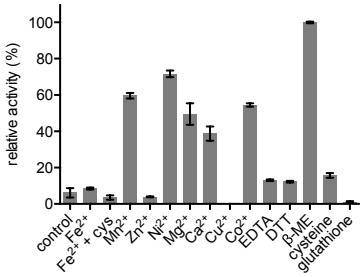


Figure 5. Effect of metal ions, chelating and reducing agents on the activity of TaManD. Purified enzyme (0.6 μg) was pre-incubated for 1 h with 1 mM of each additive in 50 mM HEPES pH 7. Activity was then determined in reactions containing 10 mM D-mannono-1,4-lactone in 50 mM HEPES pH 7 incubated for 1 h at 55 °C before analysis with the semicarbazide assay. Activity is expressed in relation to the maximum enzyme activity. Fe^{2+} + cys: Fe^{2+} was prepared with 1 mM cysteine. DTT: dithioerythritol, EDTA: ethylendiaminetetraacetate, β-ME: β-mercaptoethanol.

TaManD was activated by similar bivalent ions and reducing factor, as previously studied bacterial mannonate dehydratases from the xylose isomerase-like superfamily. However, none of the bacterial mannonate dehydratases described so far have been shown to have an equally strong activation with both metal ions, Co^{2+} and Ni^{2+}, as observed with TaManD. Early investigations of mannonate dehydratases focused on *E. coli* and showed different strengths in activation after incubation with bivalent cations [1,24–26]. In summary, most studies showed that β-mercaptoethanol, Fe^{2+} and Mn^{2+} resulted in highest enzyme activity (90–100% of maximum enzyme activity). Incubation with other bivalent cations resulted in less activity compared to the enzyme's maximum activity (Co^{2+}: 27–80%, Ni^{2+}: 5–40% and Zn^{2+}: 5–25%, depending on the study). More recently, crystal structures of xylose isomerase-like mannonate dehydratases from *S. suis* and *E. coli* have been solved and revealed the presence of primarily Mn^{2+} and lower amounts of Mg^{2+}, Ni^{2+} and Zn^{2+} in their binding sites [10,11].

T. acidophilum was originally isolated from a hot and acidic environment and accordingly this archaeon displays optimal temperature and pH for growth at 59 °C and pH 1–2, respectively [27]. The purified TaManD was active between 35 °C and 70 °C, with an optimal temperature for activity at 65 °C (Figure 6A). Thermostability of the enzyme in the absence of substrate was studied at temperatures between 55 °C and 95 °C (Figure 6B). The enzyme retains its full activity at 55 °C for at least an hour, whereas complete inactivation was observed at 75 °C, 85 °C and 95 °C within 90 min, 60 min and 15 min, respectively. The difference in the relatively high optimal temperature of the enzyme to the comparably low thermostability indicated that the enzyme is more prone to inactivation in the absence of substrate. Other enzymes obtained from this organism have been shown to display optimal temperatures in the range of 55 °C to 70 °C [15,28–30].

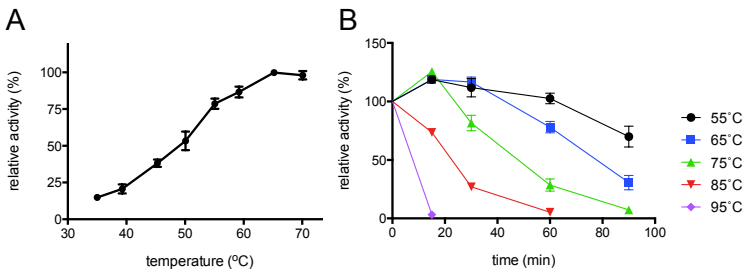


Figure 6. Effect of temperature on TaManD activity. (**A**) Optimum temperature of TaManD. (**B**) Thermostability of purified TaManD enzyme (1 µg) incubated over 90 min at various temperatures.

Following the identification of the crucial metal cofactors required for TaManD activity and determination of its optimal reaction temperature, we acquired enzyme kinetic data for the purified enzyme. Considering the fact that under physiological conditions an equilibrium of both lactone and free sugar acid exists, the kinetic data was acquired in reactions with substrate provided in these two different forms. The substrate was either hydrolysed with NaOH to obtain free sugar acid, D-mannonate, or the lactone was prepared in buffer, equaling an equilibrium of lactone and free sugar acid. TaManD displayed a higher activity in reactions with D-mannonate compared to reactions with D-mannono-1,4-lactone prepared in buffer at pH 7 (Table 1). The maximal velocity of the reaction, V_{max}, with D-mannonate as substrate was slightly higher (24–27%) compared to reactions with the lactone prepared in buffer (p-value: 0.06), whereas no difference in the affinity between the two different forms of the substrate could be observed (Table 1, Figure S2).

The pH optimum of TaManD was dependent on the form of the substrate. TaManD displayed highest activity between pH 5 and 7 with mannonate, while with D-mannono-1,4-lactone in buffer, the enzyme displayed overall lower activities with an apparent optimum at pH 7 (Figure 7). Unlike previously studied enzymes from *T. acidophilum* (e.g., AldT), TaManD did not retain maximum activity above pH 7 with both substrates [15].

Table 1. TaManD kinetic data with D-mannonate (prepared by hydrolysis with NaOH) and with D-mannono-1,4-lactone prepared in buffer (pH 7). Non-linear fitting was performed (Figure S2).

Substrate	K_m (mM)	V_{max} (U/mg)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
D-mannonate	5.37 ± 0.90	2.39 ± 0.11	1.64	0.30
D-mannono-1,4-lactone in buffer	4.90 ± 0.53	1.90 ± 0.06	1.33	0.26

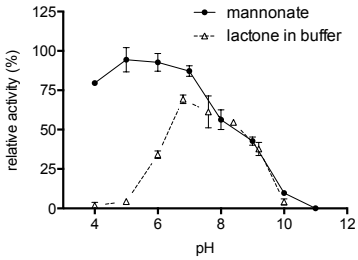


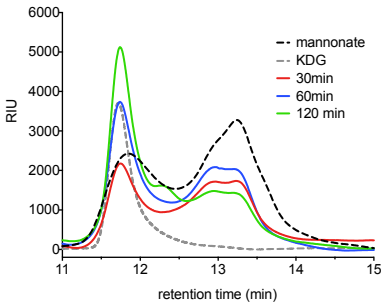
Figure 7. The pH optima of reactions with D-mannono-1,4-lactone prepared in buffer pH 7 (lactone in buffer) and with D-mannonate (prepared by hydrolysis with NaOH). Activity values are expressed as relative activity (%) normalised to the highest overall activity observed in the assay.

Based on the observations from experiments with different forms of substrate, we assume that D-mannonate is either the only form, or at least more accessible to TaManD than its lactone. At low pH, TaManD shows low activity with D-mannono-1,4-lactone, supposedly because the lactone does not hydrolyse to the free acid. In contrast, if D-mannono-1,4-lactone was hydrolysed to D-mannonate and used as substrate at low pH, high enzyme activity should be observed. Similarly, the difference in V_{max} for D-mannonate and D-mannono-1,4-lactone in buffer can be explained by slow hydrolysis of the lactone to the free acid. Reactions with hydrolysed D-mannonate occur faster compared to the lactone in buffer, since more substrate is readily available.

2.5. Product Identification and Substrate Specificity

In order to identify the product of the TaManD reaction with mannonate, reactions containing the substrate and Co^{2+} were analysed at several time points by high-performance liquid chromatography (HPLC) (Figure 8A). Although a complete separation of the product KDG and mannonate could not be achieved on different organic acid columns, an increase of KDG was observed. In order to test for substrate promiscuity, 11 different sugar acids (Figure 8B) were incubated with purified TaManD and tested for the formation of 2-keto-3-deoxy analogues using the semicarbazide assay. D-mannonate and D-mannono-1,4-lactone were the only substrates that showed a positive reaction after incubation for 16 h. This is in contrast to other dehydratases, including members of the enolase superfamily and the IlvD/EDD superfamily, such as the L-fuconate dehydratase from *Xanthomonas campestris* [31] or the dihydroxy-acid dehydratase from *Sulfolobus solfataricus* [17,32], which have been shown to be active with a multitude of different sugar acids.

A



B

tested sugar acids	relative activity
D-mannonate	100%
D-mannono-1,4-lactone	79%
D-gluconic acid sodium salt	n.d.
D-galactono-1,4-lactone	n.d.
D-glucuronic acid	n.d.
D-galacturonic acid sodium salt	n.d.
L-gulonic acid-γ-lactone	n.d.
D-arabinoic acid sodium salt	n.d.
D-fuconic acid lithium salt	n.d.
L-fuconic acid sodium salt	n.d.
L-rhamnonic acid lithium salt	n.d.

Figure 8. Substrate specificity and product analysis of reactions with TaManD. (**A**) Reactions with TaManD and 10 mM mannonate analysed by HPLC equipped with a refractive index detector. Standards: 10 mM mannonate (dashed black line), 2.5 mM KDG (dashed grey line). Reactions: stopped after 30 min (red line), stopped after 60 min (blue line), stopped after 120 min (green line). RIU: refractive index units. (**B**) List of sugar acids that were tested as substrates for TaManD. Activity is expressed relative (%) to the maximal activity obtained in the assay; n.d.: no signal detected after 16 h of incubation. Lactones except for D-mannono-1,4-lactone were hydrolysed by incubation in 1 M NaOH for 1 h at room temperature and then diluted in 50 mM HEPES pH 7 to obtain free sugar acids. All sugar acids were tested at a concentration of 10 mM in duplicate experiments.

3. Materials and Methods

3.1. Cloning, Expression and Purification of Enzymes

All plasmids were constructed by amplification of genes from *T. acidophilum* DSM 1728 genomic DNA (DSMZ, Braunschweig, Germany) using PCR, restriction and ligation into vectors according to standard protocols [33]. Primer pairs and restriction enzymes used for the construction of each expression plasmid are summarised in Table S2. *E. coli* strain α-select (Bioline, Sydney, Australia) was used in all initial cloning procedures. The expression plasmid pProEX HTa-Ta0753 was then used to transform BL21-CodonPlus (DE3)-RIL competent cells (Agilent Technologies, Santa Clara, CA, USA). Plasmid pETDuet-1-AldT was used to transform BL21 (DE3) cells (NEB, Ipswitch, Burlington, MA, USA). BL21-CodonPlus (DE3)-RIL strains containing the respective expression plasmid were grown in 500 mL Luria–Bertani medium (LB) supplemented with carbenicillin (100 μg/ml) and chloramphenicol (35 μg/ml) at 37 °C to an OD_{600} of 0.6 before induction was performed with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 20 °C for 16–18 h. The BL21 (DE3) cells carrying the expression plasmid pETDuet-1-AldT were grown in 500 mL LB medium containing carbenicillin (100 μg/ml) to an OD_{600} of 0.5 at 37 °C before induction was performed with 0.4 mM IPTG at 37 °C for 4 h. All cultures were harvested by centrifugation at 5000× *g* for 20 min, resuspended in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7 and lysed by three passages through a French pressure cell. Soluble fractions of the lysates were obtained by centrifugation at 15,000× *g* for 20 min. His-tagged enzymes were purified by subjecting the soluble extracts to a 5 mL nickel-nitrilotriacetic acid (Ni-NTA) affinity column (GE Healthcare) equilibrated with buffer containing 50 mM sodium phosphate (NaP), 300 mM NaCl and 20 mM imidazole. The proteins were eluted isocratically with a final concentration of 400 mM imidazole. The buffer of the eluates was exchanged to 50 mM HEPES pH 7 by three rounds of centrifugation in Amicon Ultra centrifugal filters (Millipore) with a 30 kDa molecular weight cut-off.

In order to obtain TaManD without a N-terminal His-tag, the enzyme obtained after pProEX HTa-Ta0753 expression and first His-tag purification was incubated with a TEV Protease (Promega,

Madison, WI, USA). Digestion was performed according to the manufacturer's protocol with addition of 10% glycerol (v/v, final concentration) and incubation for 1.5 h at 30 °C followed by incubation at 4 °C overnight. Removal of the His-tag and the TEV Protease was achieved by a second Ni-NTA purification using a 1 mL ZetaSep Ni-NTA affinity column (emp Biotech, Berlin, Germany). The flow-through contained the pure protein and cleavage of the His-tag was verified by SDS-PAGE.

3.2. Protein Analysis

Purified proteins were separated by SDS-PAGE using 4–15% Tris-glycine polyacrylamide gels (Bio-Rad Laboratories, Hercules, USA) and visualised with Coomassie Brilliant Blue. For the identification of TaManD, the single protein band at 38 kDa was excised from the gel and analysed by LC ESI MS/MS at the Australian Proteome Analysis Facility (APAF, Macquarie University), as described elsewhere [34]. Protein identification was performed using the mascot software (2.4.1, Matrixscience, Boston, MA, USA). The native molecular mass of the recombinant protein was determined by size exclusion chromatography on an ÄKTA pure FPLC system using a Superdex Increase 200 10/300 GL column (GE Healthcare) equilibrated with 10 mM NaP buffer, 140 mM NaCl, pH 7. A calibration curve was obtained with molecular weight markers thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa) and ovalbumin (43 kDa).

3.3. Nuclear Magentic Resonance (NMR) Analysis

D-Mannono-1,4-lactone (>95%) was purchased from TCI Chemicals Co. (Tokyo, Japan). For NMR spectra of the sugar acid in different substrate forms, 1 M D-Mannono-1,4-lactone solutions were prepared in 1 M NaOH, 1 M HCl and in ultrapure water. All solutions were incubated for 1 h at room temperature, before each solution was diluted to 100 mM using ultrapure water. Then, 300 µL of each 100 mM solution (equivalent to 5 mg sugar acid) was freeze-dried and resuspended in 500 µL 0.1 M NaP buffer, pH 7. The pH of all three samples was evaluated prior to NMR analysis. The substrate incubated in HCl displayed pH 6, while the substrate in water or NaOH displayed pH 7. After transfer of each sample to 5 mm NMR tubes, 10% D_2O and 40 µM 3-(trimethylsilyl)propionic-2,2,3,3-d4 acid (TMSP) chemical shift standard was added. The 1D decoupled ^{13}C NMR spectra of non-isotopically labelled samples were acquired on a 500 MHz Bruker Avance III HD NMR equipped with a BBFO probe at 50 °C (323 K), using power-gated proton decoupling with a 90 ° pulse with 512 scans and a 3 s relaxation delay between scans. Oxidation of D-[1,6-$^{13}C_2$]mannose was performed in reactions containing 5 mM of the substrate, 5 mM NAD^+, 20 mM NaP buffer at pH 7 and 13.5 µg purified AldT. Prior to the addition of purified enzyme and after 30 min of reaction at 50 °C, 1D power-gated proton decoupled ^{13}C NMR spectra were acquired using a 90 ° pulse with 4 scans and a 3 s relaxation delay between scans. Visualisation of all spectra was performed with iNMR 6.0 (http://www.inmr.net). After Fourier transformation and automatic phase correction, an exponential visual weighting factor of 1.5 and a smoothing factor of 10 were applied to all spectra.

3.4. Enzyme Activity

D-mannono-1,4-lactone was either used after preparation in 50 mM HEPES at pH 7 (indicated as "lactone in buffer") or after hydrolysis to its free acid form (indicated as "mannonate" or "free sugar acid"). Hydrolysis was performed according to Lamble et al. by preparing 1 M stock solutions of D-mannono-1,4-lactone in 1 M NaOH and incubation at room temperature for 1 h before dilutions were prepared in 50 mM HEPES pH 7 [35]. For temperature optimum, pH optimum and thermostability, activity was quantified using the thiobarbituric acid (TBA) assay, according to a modification by Buchanan et al. [36,37]. Unless stated otherwise, reactions (60 µL) contained 0.5–1.5 µg pure enzyme, 10 mM mannonate, 50 mM HEPES pH 7 (adjusted at 55 °C) and 1 mM $CoSO_4$. Reactions were stopped by incubation on ice or if needed by addition of 6 µL 12.5% trichloracetic acid (TCA). Next, 50 µL of the reaction mixture was oxidised using 125 µL of 25 mM periodic acid in 0.25 M H_2SO_4 at room temperature for 20 min. The oxidation was stopped by addition of 250 µL of 2% (w/v) sodium arsenite

in 0.5 M HCl. Finally, 1 mL of 0.3% TBA in water was added to the samples and the reaction mixture was boiled for 10 min. The formation of 2-keto-3-deoxy sugar acid was determined by reading A_{549nm} using 96-well microtiter plates in a Spectrostar Nano plate reader (BMG Labtech, Ortenberg, Germany) and quantified using the molar extinction coefficient of 67.8×10^3 $M^{-1} \times cm^{-1}$.

The effect of pH on enzyme activity was analysed in duplicate reactions with 1 µg enzyme and 120 mM universal buffer [38]. Reactions were incubated for 2 h at 55 °C before they were analysed with the TBA Assay. In reactions with D-mannono-1,4-lactone in buffer, pH decreased during the course of the experiment, which was accounted for in the analysis by measurement of the pH after the reaction. The optimal temperature for enzyme activity was determined in duplicate reactions containing 1 µg enzyme. Duplicate reactions were performed for 1 h at different temperatures before being analysed with the TBA Assay. For thermostability assays, 120 mM universal buffer was used and purified enzyme (1 µg) was incubated in duplicates at temperatures between 55 °C and 95 °C in the absence of substrate. Samples were removed at different times and the residual activity was measured in reactions with 10 mM mannonate for 45 min at 55 °C.

The semicarbazide assay was used according to Wichelecki et al. [12] to determine the effect of metal ions, chelators and reducing agents, and the acquisition of Michalis-Menten kinetics (K_m, V_{max}). Each reaction (60 µL) was incubated with 240 µL semicarbazide reagent at room temperature for 1 h. A_{250nm} was read in a UV transparent microtiter plate (Thermo Scientific, Waltham, MA, USA) using a Spectrostar Nano plate reader. Product formation was quantified using a standard curve of 2-keto-3-deoxygluconate (Sigma-Aldrich, St. Louis, MO, USA) prepared in the same assay buffer. K_m, V_{max} values were estimated using non-linear fitting in Prism 6 (6.0c, GraphPad software, San Diego, CA, USA). For the effect of metals, chelators and reducing agents, pure enzyme (0.6 µg) was incubated with 1 mM of each additive in 50 mM HEPES pH 7 for 1 h at room temperature. Activity was measured in duplicate reactions containing 10 mM D-mannono-1,4-lactone in 50 mM HEPES pH 7, incubated for 1 h at 55 °C and then analysed with the semicarbazide assay. Kinetic data for the determination of K_m and V_{max} were performed in duplicate reactions (60 µL), containing 1 µg purified TaManD, 50 mM HEPES pH 7, 1 mM $CoSO_4$ and different substrate concentrations (1–50 mM). Linear increase of reaction product was assured over 1 h reaction time. Reactions were performed for 45 min at 55 °C before product formation was determined using the semicarbazide assay and the KDG standard curve. Substrate specificity was analysed using 11 different sugar acids (Figure 8B). Reactions were performed in duplicate and contained 0.5 µg enzyme, 50 mM HEPES pH 7, 1 mM $CoSO_4$ and 10 mM of each sugar acid. Reactions were incubated for 16 h at 55 °C before they were analysed using the semicarbazide assay.

3.5. High Performance Liquid Chromatography (HPLC) Analysis

Reactions catalysed by TaManD and standards of mannonate and KDG were analysed using an Agilent 1290 HPLC system connected to a refractive index detector (RID) G1362A (Agilent Technologies, Santa Clara, CA, USA). Samples were analysed on an organic acid column (Agilent HiPlex H^+) with 10 mM H_2SO_4 as a mobile phase at a flow rate of 0.6 ml/min. The column was heated to 80 °C and the RID was set to 55 °C. Reactions were stopped by addition of 7 µL 12.5% TCA to a 60 µL reaction. After short centrifugation, 10 µL of the supernatant was used for HPLC analysis.

3.6. Statistical Analysis

Statistical analysis was performed for enzyme kinetics and effect of different metals and additives on enzyme activity, by two-tailed unpaired t-tests using Prism 6.

4. Conclusions

In this study, we present the first purification and characterisation of a functional archaeal mannonate dehydratase. The gene encoding for the mannonate dehydratase was found adjacent to a previously described aldohexose dehydrogenase (AldT) gene in the genome of *T. acidophilum* [15].

Previously, it has been shown that AldT from *T. acidophilum* selectively oxidises mannose, but the physiological function behind that oxidation was not investigated. Using NMR spectroscopy, we were able to show that mannonate and mannono-1,4-lactone are produced via oxidation of D-mannose by AldT. Kinetic assays confirmed that TaManD was able to convert both the sugar acid and its lactone to the central intermediate KDG at neutral pH, without further help of a lactonase. This resembles the second step of many oxidative pathways studied in archaea, including those for sugars like glucose, galactose, rhamnose, arabinose and xylose [39]. The amino acid sequence of TaManD and those of the putative mannonate dehydratases from *F. acidarmanus* and *F. acidiphilum* share high amino acid sequence identity. Although the gene annotations for those mannonate dehydratases indicate a role in the hexuronate metabolism (*uxuA*), they are all located adjacent to (putative) aldohexose dehydrogenases. It remains to be seen whether a non-phosphorylative pathway starting from mannose, similar to the non-phosphorylative Entner-Doudoroff pathway from glucose and galactose, exists in thermophilic archaea like *T. acidophilum*, *F. acidarmanus* and *F. acidiphilum*.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4344/9/3/234/s1, Figure S1: Purification of the Ta0753 gene product (TaManD) after expression in *E. coli*, Figure S2: TaManD kinetic data Table S1: Overview of genes, potentially encoding for mannonate dehydratases located in the genomic neighbourhood of AldT (Ta0754), Table S2: Oligonucleotides used in this study.

Author Contributions: Conceptualization, D.K. and A.S.; methodology, experimental design, D.K., R.W., and A.S.; data collection, D.K.; data analysis, D.K.; writing—original draft preparation, D.K. and A.S.; writing—review and editing, D.K., R.W., and A.S.; supervision, A.S.; project administration, A.S.

Funding: D.K. is supported by an international Macquarie University Research Excellence Scholarship (iMQRES).

Acknowledgments: We thank the Australian Proteome Analysis Facility (Macquarie University, Sydney, Australia) for their support in performing LC ESI MS/MS experiments and Nicole Cordina from the Macquarie University NMR facility from the Department of Molecular Sciences for her help in performing NMR studies.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Robert-Baudouy, J.; Jimeno-Abendano, J.; Stoeber, F. D-Mannonate and D-altronate dehydratases of *Escherichia coli* K12. *Methods Enzymol.* **1982**, *90*, 288–294. [PubMed]
2. Shulami, S.; Gat, O.; Sonenshein, A.L.; Shoham, Y. The glucuronic acid utilization gene cluster from *Bacillus stearothermophilus* T-6. *J. Bacteriol.* **1999**, *181*, 3695–3704. [PubMed]
3. Mekjian, K.R.; Bryan, E.M.; Beall, B.W.; Moran, C.P. Regulation of hexuronate utilization in *Bacillus subtilis*. *J. Bacteriol.* **1999**, *181*, 426–433. [PubMed]
4. Hugouvieux-Cotte-Pattat, N.; Robert-Baudouy, J. Hexuronate catabolism in *Erwinia Chrysanthemi*. *J. Bacteriol.* **1987**, *169*, 1223–1231. [CrossRef] [PubMed]
5. Mandrand-Berthelot, M.-A.; Condemine, G.; Hugouvieux-Cotte-Pattat, N. Catabolism of hexuronides, hexuronates, aldonates, and aldarates. *EcoSal Plus* **2004**, *1*, 1–21. [CrossRef] [PubMed]
6. Reis, D.; Vian, B.; Roland, J.C. Cellulose-glucuronoxylans and plant cell wall structure. *Micron* **1994**, *25*, 171–187. [CrossRef]
7. Lawford, H.G.; Rousseau, J.D. Fermentation of biomass-derived glucuronic acid by *pet* expressing recombinants of *E. coli* B. *Appl. Biochem. Biotechnol.* **1997**, *63–65*, 221–241. [CrossRef]
8. Peekhaus, N.; Conway, T. What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. *J. Bacteriol.* **1998**, *180*, 3495–3502. [PubMed]
9. Chang, D.-E.; Smalley, D.J.; Tucker, D.L.; Leatham, M.P.; Norris, W.E.; Stevenson, S.J.; Anderson, A.B.; Grissom, J.E.; Laux, D.C.; Cohen, P.S.; et al. Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7427–7432. [CrossRef] [PubMed]
10. Zhang, Q.; Gao, F.; Peng, H.; Cheng, H.; Liu, Y.; Tang, J.; Thompson, J.; Wei, G.; Zhang, J.; Du, Y.; et al. Crystal structures of *Streptococcus suis* mannonate dehydratase (ManD) and its complex with substrate: Genetic and biochemical evidence for a catalytic mechanism. *J. Bacteriol.* **2009**, *191*, 5832–5837. [CrossRef] [PubMed]

11. Qiu, X.; Tao, Y.; Zhu, Y.; Yuan, Y.; Zhang, Y.; Liu, H.; Gao, Y.; Teng, M.; Niu, L. Structural insights into decreased enzymatic activity induced by an insert sequence in mannonate dehydratase from Gram negative bacterium. *J. Struct. Biol.* **2012**, *180*, 327–334. [CrossRef] [PubMed]
12. Wichelecki, D.J.; Balthazor, B.M.; Chau, A.C.; Vetting, M.W.; Fedorov, A.A.; Fedorov, E.V.; Lukk, T.; Patskovsky, Y.V.; Stead, M.B.; Hillerich, B.S.; et al. Discovery of function in the enolase superfamily: D-mannonate and D-gluconate dehydratases in the D-mannonate dehydratase subgroup. *Biochemistry* **2014**, *53*, 2722–2731. [CrossRef] [PubMed]
13. Wichelecki, D.J.; Alyxa, J.; Vendiola, F.; Jones, A.M.; Al-obaidi, N.; Almo, S.C.; Gerlt, J.A. Investigating the physiological roles of low-efficiency D-mannonate and D-gluconate dehydratases in the enolase superfamily: pathways for the catabolism of L-gulonate and L-idonate. *Biochemistry* **2014**, *53*, 5692–5699. [CrossRef] [PubMed]
14. Rakus, J.F.; Fedorov, A.A.; Fedorov, E.V.; Glasner, M.E.; Vick, J.E.; Babbitt, P.C.; Almo, S.C.; Gerlt, J.A. Evolution of enzymatic activities in the enolase superfamily: D-mannonate dehydratase from *Novosphingobium aromaticivorans*. *Biochemistry* **2007**, *46*, 12896–12908. [CrossRef] [PubMed]
15. Nishiya, Y.; Tamura, N.; Tamura, T. Analysis of bacterial glucose dehydrogenase homologs from thermoacidophilic archaeon *Thermoplasma acidophilum*: finding and characterization of aldohexose dehydrogenase. *Biosci. Biotechnol. Biochem.* **2014**, *68*, 2451–2456. [CrossRef] [PubMed]
16. Yasutake, Y.; Nishiya, Y.; Tamura, N.; Tamura, T. Structural Insights into unique substrate selectivity of *Thermoplasma acidophilum* D-aldohexose dehydrogenase. *J. Mol. Biol.* **2007**, *367*, 1034–1046. [CrossRef] [PubMed]
17. Kim, S.; Lee, S.B. Catalytic promiscuity in dihydroxy-acid dehydratase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *J. Biochem.* **2006**, *139*, 591–596. [CrossRef] [PubMed]
18. Ashwell, G. [21] Enzymes of glucuronic and galacturonic acid metabolism in bacteria. In *Methods in enzymology*; Academic Press: Cambridge, MA, USA, 1962; Volume 5, pp. 190–208, ISBN 0076-6879.
19. Portalier, R.; Robert-Baudouy, J.; Stoeber, F. Regulation of *Escherichia coli* K-12 hexuronate system genes: *exu* regulon. *J. Bacteriol.* **1980**, *143*, 1095–1107. [PubMed]
20. Palleroni, N.J.; Duodoroff, M. Metabolism of Carbohydrate by *Pseudomonas Saccharophila* III. Oxidation of Arabinose. *J. Bacteriol.* **1957**, *74*, 180–185. [PubMed]
21. Watanabe, S.; Saimura, M.; Makino, K. Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. *J. Biol. Chem.* **2008**, *283*, 20372–20382. [CrossRef] [PubMed]
22. Stephens, C.; Christen, B.; Fuchs, T.; Sundaram, V.; Watanabe, K.; Jenal, U. Genetic analysis of a novel pathway for D-xylose metabolism in *Caulobacter crescentus*. *J. Bacteriol.* **2007**, *189*, 2181–2185. [CrossRef] [PubMed]
23. Wałaszek, Z.; Horton, D. Conformational studies on aldonolactones by NMR spectroscopy. Conformations of d-glucono-, D-mannono, D-gulono-, D-galactono-1,4-lactone in solution. *Carbohydr. Res.* **1982**, *105*, 131–143. [CrossRef]
24. Dreyer, J. The role of iron in the activation of mannonic and altronic acid hydratases, two Fe-requiring hydro-lyases. *Eur. J. Biochem* **1987**, *166*, 623–630. [CrossRef] [PubMed]
25. Robert-Baudouy, J.M.; Jimeno-Abendano, J.; Stoeber, F.R. Individualité des hydrolyases mannonique et altronique chez *E. coli* K-12. *Biochimie* **1975**, *57*, 1–8. [CrossRef]
26. Robert-Baudouy, J.M.; Stoeber, F.R. Purification et propriétés de la D-mannonate hydrolyase *d'Escherichia coli*. *Biochim. Biophys. Acta -Enzymology* **1973**, *309*, 473–485. [CrossRef]
27. Darland, G.; Brock, T.D.; Samsonoff, W.; Conti, S.F. A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile. *Science* **1970**, *170*, 1416–1418. [CrossRef] [PubMed]
28. Jung, J.H.; Lee, S.B. Identification and characterization of *Thermoplasma acidophilum* 2-keto-3-deoxy-D-gluconate kinase: A new class of sugar kinases. *Biotechnol. Bioprocess Eng.* **2005**, *10*, 535–539. [CrossRef]
29. Kim, S.M.; Paek, K.H.; Lee, S.B. Characterization of $NADP^+$-specific L-rhamnose dehydrogenase from the thermoacidophilic Archaeon *Thermoplasma acidophilum*. *Extremophiles* **2012**, *16*, 447–454. [CrossRef] [PubMed]

30. Reher, M.; Schönheit, P. Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner-Doudoroff pathway, constitute a novel enzyme family within the aldehyde dehydrogenase superfamily. *FEBS Lett.* **2006**, *580*, 1198–1204. [PubMed]
31. Yew, W.S.; Fedorov, A.A.; Fedorov, E.V.; Rakus, J.F.; Pierce, R.W.; Almo, S.C.; Gerlt, J.A. Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from *Xanthomonas campestris*. *Biochemistry* **2006**, *45*, 14582–14597. [CrossRef] [PubMed]
32. Carsten, J.M.; Schmidt, A.; Sieber, V. Characterization of recombinantly expressed dihydroxy-acid dehydratase from *Sulfobus solfataricus*—A key enzyme for the conversion of carbohydrates into chemicals. *J. Biotechnol.* **2015**, *211*, 31–41. [CrossRef] [PubMed]
33. Sambrook, J.; Russell, D.W.; Russell, D.W. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2001; Volume 49, ISBN 978-1-936113-42-2.
34. Atack, J.M.; Srikhanta, Y.N.; Fox, K.L.; Jurcisek, J.A.; Brockman, K.L.; Clark, T.A.; Boitano, M.; Power, P.M.; Jen, F.E.C.; McEwan, A.G.; et al. A biphasic epigenetic switch controls immunoevasion, virulence and niche adaptation in non-typeable Haemophilus influenzae. *Nat. Commun.* **2015**, *6*, 1–12. [CrossRef] [PubMed]
35. Lamble, H.J.; Milburn, C.C.; Taylor, G.L.; Hough, D.W.; Danson, M.J. Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in *Sulfolobus solfataricus*. *FEBS Lett.* **2004**, *576*, 133–136. [CrossRef] [PubMed]
36. Buchanan, C.L.; Connaris, H.; Danson, M.J.; Reeve, C.D.; Hough, D.W. An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. *Biochem. J.* **1999**, *343*, 563–570. [CrossRef] [PubMed]
37. Skoza, L.; Mohos, S. Stable thiobarbituric acid chromophore with dimethyl sulphoxide. Application to sialic acid assay in analytical de-O-acetylation. *Biochem. J.* **1976**, *159*, 457–462. [CrossRef] [PubMed]
38. Britton, H.T.S.; Robinson, R.A. CXCVIII.—Universal buffer solutions and the dissociation constant of veronal. *J. Chem. Soc.* **1931**, 1456–1462. [CrossRef]
39. Bräsen, C.; Esser, D.; Rauch, B.; Siebers, B. Carbohydrate metabolism in archaea: current insights into unusual enzymes and pathways and their regulation. *Microbiol. Mol. Biol. Rev.* **2014**, *78*, 89–175. [CrossRef] [PubMed]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Supplementary Materials: Characterisation of the first archaeal mannonate dehydratase from *Thermoplasma acidophilum* and its potential role in the catabolism of D-mannose

Dominik Kopp, Robert Willows, Anwar Sunna*

***anwar.sunna@mq.edu.au**

Table S1. Overview of genes, potentially encoding for mannonate dehydratases located in the genomic neighbourhood of AldT (Ta0754). The gene product of Ta0753 is referred to as TaManD in the following and was further investigated as part of this study. MR/MLE, mandelate racemase/muconate lactonising enzyme.

old locus tag	new locus tag	NCBI gene annotation	protein family
Ta0752	TA_RS03865	L-rhamnonate dehydratase	MR/MLE enolase superfamily
Ta0753	TA_RS03870	D-mannonate dehydratase	uxuA xylose isomerase-like superfamily
Ta0756	TA_RS03885	mandelate racemase/muconate lactonzing enzyme family protein	MR/MLE enolase superfamily

Table S2. Oligonucleotides used in this study. Engineered restriction sites are displayed in lowercase. RE, restricton enzyme.

Expression plasmid	gene locus	oligonucleotide used	oligonucleotide sequence (5' –> 3')	RE used
pETDuet-1-AldT	Ta0754	AldT-fw	TGATggatccCATGTTCAGCGATCTAAGGGA	BamHI
	Ta0754	AldT-rev	ATTATgaattcCTCATTCTGGCGTGCTTATGG	EcoRI
pProEX HTa-Ta0753	Ta0753	ManD_woATG_fw	TGATggatccTAAGCTCAGAATAGCCGAGATC	BamHI
	Ta0753	ManD_HindIII_rev	TGATaagcttTCAGGAAATCTGTTTATGATTTG	HindIII

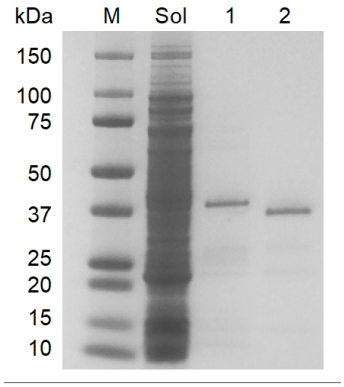


Figure S1. Purification of the Ta0753 gene product (TaManD) after expression in *E. coli*. M, Bio-rad Precision Plus Protein Standard; Sol, soluble cell extract; 1, enzyme after first His-tag purification (expected size 42 kDa); 2, enzyme after TEV cleavage and second His-tag purification (expected size 38 kDa).

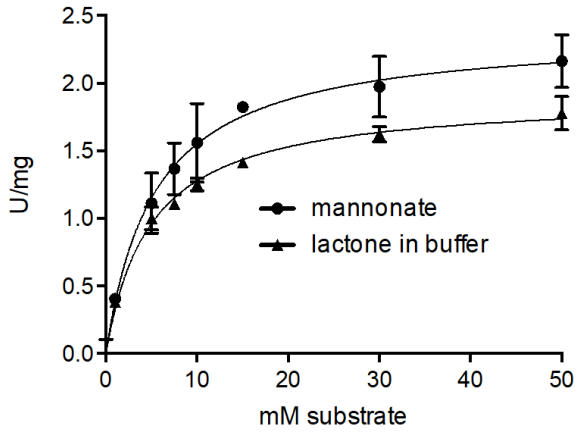


Figure S2 TaManD kinetic data. Kinetic measurements with D-mannonate (prepared by hydrolysis with NaOH) and with D-mannono-1,4-lactone prepared in buffer (pH 7). Non-linear fitting was applied using Prism 6 (GraphPad software). Error bars indicate standard error of mean values.

Appendix III

Cloning, expression and purification of ManD in *Saccharomyces cerevisiae*

Introduction

The process of developing, testing and up-scaling of cell-free enzymatic pathways depends on enzyme performance and available enzyme amounts. The use of the mannonate dehydratase from *Thermoplasma acidophilum* (ManD) for a potential up-scaling of the cell-free conversion process is hampered by its low solubility in *Escherichia coli*. Various methods were tried to increase the soluble expression of ManD, including genetic fusion with solubility tags, codon optimisation, autoinduction, heat-inducible promoters, different *E. coli* expression cell lines and expression in *Saccharomyces cerevisiae*. None of the methods showed an increase in soluble yields compared to the method used in chapter 2. However, the expression in yeast produced enough soluble protein that could be purified via immobilised metal affinity chromatography. The recombinant enzyme was found inactive in its purified form and in soluble and insoluble lysates of the expression strain. However, at the stage of the project when these experiments were performed, it was not clear yet that ManD is dependent on essential bivalent cations for activity. The construction of the ManD expressing *S. cerevisiae* strain is documented and results of enzyme expression and purification are presented.

Methods

Strains and media

For the expression and assembly of plasmids via homologous recombination the yeast strain CEN.PK2.1C (*MAT*a *ura3-52 trp1-289 leu2-3,112 his3*-Δ1 *MAL2-8*^{C} *SUC2*) was used. The strain was kindly provided by Dr. H. Kroukamp (Macquarie University). Media used for growth of *S. cerevisiae* are listed in Table 2. For non-selective cultivation, cells were grown in synthetic complete (SC) medium. For selection of plasmids containing a *URA* marker, cells were propagated in SC-URA medium. Plasmids were propagated and maintained in *E. coli* α-select cells (Bioline) grown on lysogeny broth (LB) medium

supplemented with carbenicillin (50 µg/ml).

Table 2 Media used for growth of *S.cerevisiae*. All chemicals were purchased from Sigma Aldrich (St. Louis, USA)

synthetic complete (SC)		synthetic complete without uracil (SC-URA)	
yeast extract	2%	Difco™ yeast nitrogen base without amino acids	5.34 g/l
peptone	2%	yeast synthetic drop-out medium supplements without uracil	1.92 g/l
glucose	2%	glucose	2%

Plasmid construction and yeast transformation

All DNA preparations, manipulations and digestions were carried out according to Sambrook and Russell (Sambrook et al., 2001). The gene Ta0753 was amplified using the primer pairs ManD-fw and ManD-woSTOP-His-rev (Table 3) directly from *Thermoplasma acidophilum* DSM 1728 (DSMZ, Braunschweig, Germany) genomic DNA. The amplified Ta0753 DNA was digested with BamHI and XhoI and ligated into similarly cut pET22b vector (Novagen). Plasmid pET22b-ManD-His was used to create the ManD-His fragment using primer pairs Y1 and Y4. Plasmid pRS426 containing a phosphoglycerate kinase promoter (PGK1p) sequence, the secretion signal XYNSEC and a cytochrome c terminator (CYC1t) sequence was used as a PCR template to generate fragment 2 and 3 using the primer pairs 2-micron-fw/2-micron-rev and ura-p-fw/ura-p-rev (Table 4). Yeast cell were then transformed with fragments 1, 2 and ManD-His using the lithium acetate method as described elsewhere (Gietz and Schiestl, 2007). Transformants were selected for growth on SC-URA plates and the presence of the gene was verified by sequencing.

Table 3 Oligonucleotides used for the construction of the pET22b-ManD-His plasmid. Restriction sites are indicated in lowercase.

plasmid	gene/ fragment	oligonucleotides used	oligonucleotide sequence (5' −> 3')	RE used
pET22b-ManD-HIs	Ta0753	ManD-fw	TGATggatccTATGAAGCTCAGAATAGCCGAG	BamHI
	Ta0753	ManD-woSTOP-His-rev	TATctcgagGGAAATCTGTTTATGATTTGACAAAG	XhoI

Table 3 Oligonucleotides used for construction of plasmid pRS426-PGK1p-ManD-His-CYC1t. Overlapping sequences used for homologous recombination are underlined.

fragment	template	oligonucleotides used	oligonucleotide sequence (5' −> 3')
1	pRS426-PGK1p- CYC1t	2-micron fw	ATATTTGCTAGCGGCGCGCCTTTTCCTTTGT
		2-micron rev	TATATGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGG
2	pRS426-PGK1p-xynsec-CYC1t	ura-p-fw	TATACTAAGAAACCATTATTATCATGACATTAACC
		ura-p-rev	TATAGACGATGTCTTAATTAAATTTGTTGTAAAAAGTAGATAATTACTTCC
Man-His	ManD-His	Y1	CTTTTTACAACAAATTTAATTAAGACATCGTCGGATCCTATGAAGCTCAGAA
		Y4	CATGATATCGACAAAGGAAAAGGCGCGCCGCTAGCAGCAGCCGGATCTCAGTG

Protein expression and purification

The yeast strain containing the expression plasmid (pRS426-PGK1p-ManD-His-CYC1t) was cultivated in 250 ml SC-URA at 30°C. A sample used as an expression control was prepared by centrifugation of 1 ml culture at OD_{600} 0.3. Cells were grown for expression up to OD_{600} of 1.45. The culture was harvested by centrifugation at 5000 x *g* for 10 min, washed once with ultrapure water and then resuspended in native lysis buffer (50mM NaH_2PO4, 300mM NaCl, 10mM imidazole, pH8) supplemented with complete protease inhibitor cocktail (Roche Diagnostics) and 100 U /ml DNAse I (Sigma Aldrich). Cells were homogenised by three passages through a French pressure cell and soluble fractions of the lysates were obtained by centrifugation at 15000 x *g* for 30 min. The remaining pellet was used for analysis of insoluble fraction. Supernatant was used for analysis of the soluble fraction. His-tagged enzymes were purified by subjecting the soluble fraction to Ni-NTA Fast start column (Qiagen, Hilden, Germany) equilibrated with native wash buffer (50mM NaH_2PO4, 300mM NaCl, 20mM imidazole, pH8). The His-tagged enzyme was eluted with a final concentration of 250 mM imidazole. Samples of the elution were taken for protein analysis and enzyme activity tests. Purified enzyme

was dialysed against 1 L 50 mM Tris/HCl pH 7.5 in two passages. Dialysed enzyme was used for enzyme activity tests.

Protein analysis

All samples taken during expression and purification were separated by SDS-PAGE using 4-12% Bis-Tris polyacrylamide gels (Invitrogen, Carlsbad, USA) and visualised with Coomassie Brilliant Blue.

Enzyme activity test

Purified enzyme before and after dialysis as well as soluble and insoluble lysates used tested for enzyme activity. Activity was measured using the thiobarbituric acid (TBA) assay according to a modification by Buchanan et al. (Buchanan et al., 1999; Skoza and Mohos, 1976). 100 µl reactions contained 10 µl of lysate or purified enzyme, 10 mM D-mannono-1,4-lactone and 120 mM Universal buffer pH 7 (Britton and Robinson, 1931). Reactions were incubated for 1 h at 55°C and were stopped by incubation on ice. 50 µl of the reaction mixture was oxidised using 125 µl of 25 mM periodic acid in 0.25 M H_2SO_4 at room temperature for 20 min. The oxidation was stopped by addition of 250 µl of 2% (w/v) sodium arsenite in 0.5 M HCl. 1 ml of 0.3% TBA in water was added to the samples and the reaction mixture was boiled for 10 min. The formation of 2-keto-3-deoxy sugar acid was analysed by reading A_{549nm} using 96-well microtiter plates in a Spectrostar Nano plate reader (BMG Labtech, Ortenberg, Germany). 1 µM 2-keto-3-deoxygluconate was used as positive control for the TBA Assay.

Results

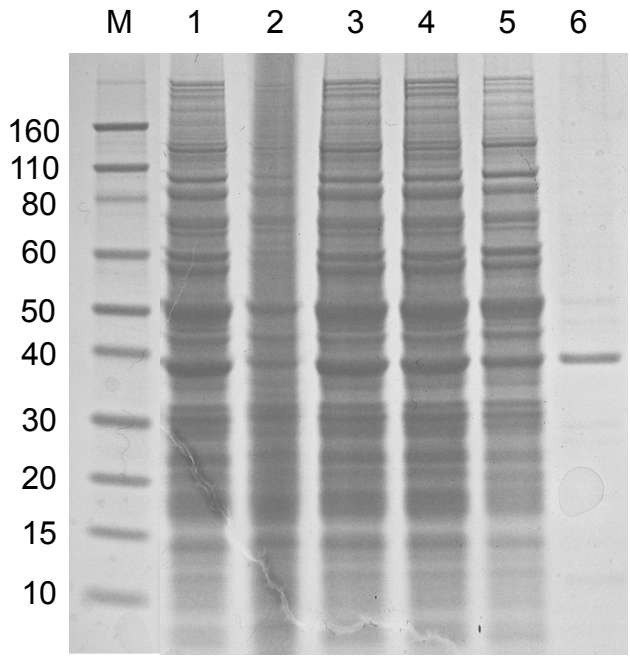


Figure 1 SDS-PAGE showing soluble expression and purification of ManD expressed in *S. cerevisiae* cells. M: Marker Novex Sharp Pre-Stained Protein Standard, 1: soluble lysate, 2: insoluble lysate, 3: lysate subjected to Ni-NTA chromatography, 4: flow-through of Ni-NTA chromatography, 5: wash of Ni-NTA chromatography, 6: elution of purified enzyme. Expected size of ManD-His: 38 kDa.

A recombinant ManD was produced in *S. cerevisiae*, further purified by metal affinity chromatography (Fig. 1) and tested for enzyme activity. In reaction with D-mannono-1,4-lactone, the purified enzyme, soluble and insoluble fractions of lysed cells did not show any activity. However, at a later stage of this research project, bivalent cations were found to be essential for the activity of ManD (chapter 2). Re-assessing mannonate dehydratase activity with purified enzyme supplemented with Co^{2+} or Ni^{2+} and in intact yeast cells is subject to future research. Since *S. cerevisiae* is normally dependent on the EMP pathway it would be of interest to test if its mannose metabolism could be switched to an oxidative pathway using ManD and the synthetic pathway presented in chapter 3.

References

Britton, H.T.S., Robinson, R.A., 1931. CXCVIII.—Universal buffer solutions and the dissociation constant of veronal. J. Chem. Soc. 1456–1462.

Buchanan, C.L., Connaris, H., Danson, M.J., Reeve, C.D., Hough, D.W., 1999. An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. Biochem. J 343, 563–570.

Gietz, D.R., Schiestl, R., 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. 2, 31.

Skoza, L., Mohos, S., 1976. Stable thiobarbituric acid chromophore with dimethyl sulphoxide. Application to sialic acid assay in analytical de-O-acetylation. Biochem. J. 159, 457–62.