

Cell-free enzymatic production of glucaric acid

Kerstin Petroll

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

Supervisors:

A/Prof. Anwar Sunna

Dr. Andrew Care

Emeritus Prof. Peter Bergquist



MACQUARIE
University
SYDNEY · AUSTRALIA

Department of Molecular Sciences

Macquarie University

Sydney, Australia

November 2018

Table of contents

Table of contents	ii
Abstract	iv
Statement of originality	vii
Acknowledgments	viii
Presentations and awards	x
Poster presentations	x
Oral presentations	x
Awards	xi
List of publications	xii
List of abbreviations	xiii
Chapter 1: Tools and strategies for constructing cell-free enzyme pathways	17
1.1 Introduction	18
1.2 Manuscript 1	19
1.3 Aims and Scope of this Thesis	80
Chapter 2: Materials and methods	82
Chapter 3: Mixed-mode liquid chromatography for the rapid analysis of biocatalytic glucaric acid reaction pathways	85
3.1 Introduction	86
3.2 Contribution to Manuscript 2	87
3.3 Manuscript 2	88
Chapter 4: A novel framework for the cell-free enzymatic production of glucaric acid	120
4.1 Introduction	121
4.2 Contribution to manuscript 3	122
4.3 Manuscript 3	123
4.4 Supplementary	174
Chapter 5: Characterisation of a putative myo-inositol oxygenase from the moderately thermophilic fungus <i>Rasamsonia emersonii</i>	190
5.1 Introduction	191
5.2 Material and Methods	193
5.2.1 Reagents, vectors, strains and growth conditions	193
5.2.2 Cloning of MIOX	194

5.2.3 Identification and analysis of the putative MIOX sequence	194
5.2.4 Recombinant expression of <i>ReMIOX</i> and <i>L-ReMIOX</i>	194
5.2.5 Protein refolding	195
5.2.6 Enzyme activity measurements	196
5.3 Results and Discussion	197
5.3.1 Identification of a putative MIOX from <i>R. emersonii</i>	197
5.3.2 Recombinant expression of <i>ReMIOX</i> and <i>L-ReMIOX</i> in <i>E. coli</i>	200
5.3.3 Characterisation of MIOX	204
5.4 Conclusion	209
References	210
Chapter 6: Summary and Future Perspectives	212
Appendix A: Publication 4	224
Appendix B: Ethics Approval	238

Abstract

Cell-free biocatalysis with a number of enzymes is a fast growing and potentially high impact field in synthetic biology for the bio-manufacture of existing and novel fine or platform chemicals and biofuels. The assembly of multiple different enzymes to form synthetic pathways is a relatively new development in contrast to the single enzyme systems that have been used for decades. The cell-free approach employs enzymes outside of the cell, allowing for controllable reaction conditions and avoids metabolic repression. It prevents the diversion and loss of pathway intermediates or end-products into the cell's own metabolism. Multi-enzyme systems in a cell-free context are particularly attractive for the 'green' synthesis of high value compounds from inexpensive, simple and renewable substrates. These systems offer great versatility, allowing the investigator to 'pick, mix and test' without the need for genetic modification of the host organism and without interference from intracellular processes. Therefore, substrate conversion yields by cell-free biocatalysis can reach 100% of the theoretical value. However, major challenges for the industrial implementation of this approach includes costly enzyme preparation, enzyme stability and the dependency on expensive cofactors.

Glucaric acid (GlucA) is one of the top 12 bio-based chemicals recognised worldwide for its potential impact and application in the synthesis of greener products. GlucA can be used in the production of (biodegradable) polymers, including (hydroxylated) nylons and polyesters, offering a more sustainable and environmentally-friendly alternative to fossil fuel-derived products. The production of GlucA has been attempted mainly by chemical and microbial synthesis.

This research describes a novel cell-free and multi-enzyme biocatalytic system developed for the synthesis of GlucA. The system is composed of a synthetic six enzyme pathway designed to facilitate the synthesis of GlucA from glucose-1-phosphate (G1P) which can be derived enzymatically from various natural polymers, such as cellulose or starch, and thus represents a promising approach to utilise crude biomass for GlucA production. An integrative framework was established to achieve an economical and efficient biocatalytic process which included; i) metabolic flux analysis for system optimisation, ii) the use of thermostable enzymes for improved stability and robustness of the system, iii) immobilisation and recycling of enzymes for reduced costs and iv) a cofactor regenerating tool to reduce cofactor requirements.

To accomplish this framework, a novel analytical method was developed to engineer the GlucA production towards high titres and to monitor the pathway flux rapidly. It was based on ultra-high performance liquid chromatography (UHPLC) and refractive index detection (RID) which enabled the simultaneous analysis of GlucA and its intermediates. All selected enzymes (mostly thermostable) were fused genetically to a synthetic peptide (referred to as the “linker”) which displayed high affinity towards silica-based materials. Due to the non-invasive binding mechanism of the linker, the enzymes were co-immobilised successfully onto zeolite, a low-cost and commercially-available silica-based material. All the immobilised enzymes except the labile mouse *myo*-inositol oxygenase remained active, implying their suitability for application in a high temperature bioprocess, and allowed for their repeated use and recycling. Finally, an additional cofactor regenerating enzyme was integrated effectively into the cell-free process which maintained high cofactor levels while reducing the cofactor requirements of the pathway.

In summary, this work presents the first cell-free production of GlucA and describes a powerful framework for viable cell-free biocatalysis based on an immobilised multi-enzyme synthetic

pathway. Further work on the cell-free GlucA production is anticipated to extend the synthesis of GlucA from biomass such as cellulose and starch, and to allow increases in the overall efficiency of the system.

Statement of originality

This thesis is the result of my own work and includes nothing which 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.

Kerstin Petroll

August 2018

Acknowledgments

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

First of all, I would like to express my sincerest gratitude to my supervisors A/Prof. Anwar Sunna, Dr. Andrew Care and Emeritus Professor Peter Bergquist for the invaluable guidance and support that I have received throughout this professional and emotional PhD journey. I am extremely grateful for the opportunity to have done my PhD under their supervision as I have benefited from their vast knowledge, creativity, high research standards and optimisms, which they constantly passed on to me.

Anwar, I appreciate highly the trust that I have been given to pursue independently my own research interests while always being aided with excellent, immediate and endless support. I believe that my learning curve has been immense due to this right and fine balance of supervision and independence.

Andrew, I am highly grateful for your inspiring, creative and challenging mindset, and the strong encouragement with which you have guided me through these past and sometimes rocky three years.

I am thankful for Emeritus Professor Peter Bergquist for his constant support, critical but highly refreshing eye on my work and his persistent help to raise the standard of my writing.

I am very grateful to have worked on my PhD alongside Alex and Dominik, who have given me new perspectives, important advice and encouragement. I would like to thank Martin for his

invaluable assistance from the distance and Robert for his immense help in finding answers to nearly anything. Thanks to Dennis, Rachit, Sandra and Vinoth for the great company throughout this PhD journey and thanks to Monica, Niel, Heinrich, Alex, Tom, Xin, Hugh, Athul, Ian, Helena, Liisa, Elsa, Louise, Elizabeth, Elisabeth, Bishal, Matt, Wei, Sameera, Wisam, Wenwen, Nicole and all other people that I have worked and become friends with for the moments of support, inspiration, laughter and encouragement.

I acknowledge the international Macquarie University Research Excellence Scholarship (iMQRES) for supporting me financially over the course of this project, and the Macquarie University Postgraduate Research Fund (PGRF) funding scheme for giving me the opportunity to travel abroad to present my research and gain new experiences.

Thanks to all my friends in the distance and nearby, especially Miriam and Marlies, for their invaluable company and support through stormy times.

Finally, I am deeply grateful for the unlimited love and support from my family, regardless of what I am doing and whichever way I am going. I am deeply grateful for the love and support from Pierpaolo, for providing refreshed views on life outside of a PhD, endless understanding and the continuous encouragement to stay determined in pursuing my goals.

‘Smooth Seas Never Made A Skilled Sailor’

Franklin D. Roosevelt

Presentations and awards

Poster presentations

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

Petroll, K., Care, A., Bergquist, P.L., Sunna, A. (2017). In-vitro biocatalysis for the ‘green’ production of glucaric acid. Poster presented at the Enzyme Engineering Conference XXIV, Toulouse, France.

Petroll, K., Care, A., Bergquist, P.L., Sunna, A. (2017). In-vitro biocatalysis for the ‘green’ production of glucaric acid. Poster presented at the Synthetic Biology Australasia Conference 2017, Sydney, Australia.

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.

Petroll, K., Care, A., Bergquist, P.L., Sunna, A. (2016). Cell-free conversion of biomass to glucaric acid. Poster presented at the 17th International Biotechnology Symposium (IBS) and Exhibition in Melbourne, Australia.

Oral presentations

Petroll, K., Care, A., Bergquist, P.L., Sunna, A. (2016). Cell-free biotransformation of organic waste to the pharmaceutical compound D-glucaric acid. MQ BioFocus Research Conference 2016, Sydney, Australia.

Awards

- 2017:** **Macquarie University Postgraduate Research Fund (PGRF)**, travel grant awarded for attendance at the Enzyme Engineering XXIV conference in France and two laboratory visits in Germany. 5,000 AUD.
- 2017:** **Runner up best poster prize** at the Synthetic Biology Australasia Conference 2017, Sydney, Australia. 50 AUD
- 2017:** **People's choice award at 3 MT competition**, Faculty of Science and Engineering (FSE) Finals, Macquarie University, Australia. 150 AUD.
- 2017:** **2nd prize for 2 minutes talk and interactive discussion session** at the 6th Annual Biofocus Research Research Conference, Macquarie University, Australia. 50 AUD.
- 2015 – 2018:** **3-year international Macquarie University Research Excellence Scholarship (iMQRES)** approx. 200,000 AUD.

List of publications

This thesis includes two submitted manuscripts (manuscript 2 and 3, see Chapter 3 and 4, respectively) and two published articles (publication 1 and 4, see Chapter 1 and Appendix, respectively).

- [1] **Petroll, K.**, Kopp, D., Care, A., Bergquist, P.L., Sunna, A. Tools and strategies for constructing cell-free enzyme pathways. In print. *Biotechnology Advances*.
- [2] **Petroll, K.**, Care, A., Waterstraat, M., Bergquist, P.L. Sunna, A. Mixed-mode liquid chromatography for the rapid analysis of biocatalytic glucaric acid reaction pathways. Submitted to *Analytica Chimica Acta*.
- [3] **Petroll, K.**, Care, A., Bergquist, P.L., Sunna, A. A novel framework for the cell-free enzymatic production of glucaric acid. Submitted to *Green Chemistry*.
- [4] Care, A., **Petroll, K.**, Gibson, E. SY., Bergquist, P.L., Sunna, A. (2017). Solid-binding peptides for immobilisation of thermostable enzymes to hydrolyse biomass polysaccharides. *Biotechnology for Biofuels* 10 (1): 29. <https://doi.org/10.1186/s13068-017-0715-2>.

List of abbreviations

ACN	acetonitrile
ANN	artificial neural networks
ATP	adenosine triphosphate
B	bound
Ctrl	control
CFPS	cell-free protein synthesis
Cys	L-cysteine
DAD	diode array detector
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EC	Enzyme Commission
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionisation
FAD	flavin adenine dinucleotide
Fe	ferrous
FID	flame ionisation detector
G°	Gibbs free energy
GA	glucuronic acid
GC	gas chromatography
GlucA	glucaric acid
GndHCl	guanidinium chloride
GST	glutathioneS-transferase
G1P	glucose-1-phosphate
G16BP	glucose-1,6-bisphosphate

G6P	glucose-6-phosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HILIC	hydrophilic interaction liquid chromatography
HPAEC	high-performance anion exchange chromatography
HPLC	high-performance liquid chromatography
I	insoluble protein fraction
IB	inclusion body
I _E	immobilisation efficiency
I _Y	immobilisation yield
IMP	inositol-1-monophosphatase (IMP)
IPS	<i>myo</i> -inositol-3-phosphate synthase
IPTG	isopropyl β -D-thiogalactoside
I _R	immobilisation recovery
I1P	1L- <i>myo</i> -inositol-1-phosphate
K _{cat}	turnover number
K _i	enzyme-inhibitor constant
K _m	Michaelis–Menten constant
L	linker
LB	Luria–Bertani broth
LC- MS	liquid chromatography - mass spectrometry
LOD	limit of detection
LOQ	limit of quantitation
M	molecular mass
MES	2-(N-morpholino)ethanesulfonic acid
min	minutes
MIOX	<i>myo</i> -inositol oxygenase

MLG	malachite green reagent
MM	mixed mode
MRM	multiple reaction monitoring mode
MS	mass spectrometry
MS/MS	tandem MS
M _w	molecular mass marker
<i>m/z</i>	mass to charge ratio
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
NMR	nuclear magnetic resonance spectroscopy
Nox	NADH oxidase
ODE	ordinary differential equations
ORF	open reading frames
RID	refractive index detector
ROS	reactive oxygen species
RSD	relative standard deviation
RT	retention time
PAD	pulsed amperometric detector
PGM	phosphoglucomutase
P _i	inorganic phosphate
pK _A	logarithmic acid dissociation constant
S	soluble protein fraction
SBA	standard binding assay
SBP	solid-binding peptide
SD	standard deviation
SDS-PAGE	sodium dodecyl sulphate - polyacrylamide gel electrophoreses
SM	start material

SRM	single reaction monitoring mode
SUMO	small ubiquitin-like modifier
T	total protein fraction
TCC	thermostatted column compartment
TFA	trifluoroacetic acid
Tris	tris(hydroxymethyl)aminomethane
U	enzyme units ($\mu\text{mol min}^{-1}$)
UB	unbound
Udh	uronate dehydrogenase
UHPLC	ultra-high-performance liquid chromatography
UV	ultraviolet
W	wash fraction
WAX	weak anion exchange chromatography
Zeo	zeolite

Chapter 1:

Tools and strategies for constructing cell-free enzyme pathways

1.1 Introduction

The construction of cell-free enzyme pathways is an exciting field of biocatalysis which evolved at the interface between synthetic biology, metabolic engineering and synthetic chemistry, and represents a promising approach for bio-manufacturing bulk and speciality chemicals and biofuels. The review contained within Chapter 1 summarises the recent achievements and current limitations of cell-free enzyme pathways. The review also proposes a range of new strategies and tools that can help to further translate the application of cell-free enzyme pathways into the bio-manufacturing industry.

The review in this chapter has been accepted for publication in the high-impact journal *Biotechnology Advances*.

1.2 Manuscript 1

Tools and strategies for constructing cell-free enzyme pathways

Kerstin Petroll^a, Dominik Kopp^a, Andrew Care^{a,b}, Peter L Bergquist^{a,b,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

*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

Keywords

Synthetic biology - *in vitro* biocatalysis – cell-free enzyme pathways - green synthesis - cell-free metabolic engineering – multi-enzyme cascades - pathway design - enzyme immobilisation - kinetic modelling - high throughput metabolite analysis.

Abstract

Single enzyme systems or engineered microbial hosts have been used for decades for various applications but the concept 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 thermodynamics and kinetics of an enzyme pathway can be very unfavourable, and the production of multiple enzymes can impose a great burden on the economics of the process. We discuss here strategies for designing synthetic pathways and highlight the requirements of substrates, enzymes and cofactor regeneration systems for improving the effectiveness and sustainability of cell-free biocatalysis. Further, we comment on methods for the immobilisation of members of a multi-enzyme pathway to enhance the economic viability of the system. Lastly, we address recent developments of integrative tools such as *in silico* pathway modelling and high throughput flux analysis with the aim of underpinning their indispensable role in the future of cell-free biocatalytic pathways and 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). 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; Swartz, 2012).

1.1 Traditional enzyme applications in academic and industrial scenarios

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., 2013; Ye et al., 2012), 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, W. et al., 2017). Similarly, the stability of enzymes can be improved through protein engineering using rational design, directed evolution or ancestral protein reconstruction (Goldsmith and Tawfik, 2017; Gumulya and Gillam, 2017; Risso et al., 2018). However, following these routes can be a tedious and a costly process. 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 instability or loss of recombinant plasmids that code for the pathway enzymes in genetically-engineered host cells can impair continuous cell-based biocatalysis (Friebs, 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.; Wilding et al., 2018; Zhang et al., 2017).

1.2 Cell-free enzyme synthesis

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 a few researchers since the 1960's (DeVries and Zubay, 1967; Nirenberg and Matthaei, 1961; Zubay, 1973), and has been demonstrated for the fermentation of sugar to ethanol by cell-free yeast extracts already in 1897 (Buchner, 1897). 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; Swartz, 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 and was subject to macromolecular degradation by native proteases and nuclease in the lysate. The Swartz laboratory embarked upon refinements of the lysate system to exploit biological processes without intact cells (summarised in (Swartz, 2006; Swartz, 2012). Recently, the principles developed have been expanded by Karim and Jewett (2016) who proposed a framework for prototyping synthetic multi-enzyme pathways and cell-free metabolic engineering (Karim and Jewett, 2016). The concept demonstrated the effective coupling of cell-

free protein synthesis and heterologous enzyme production for optimising the pathway performance in the n-butanol synthesis from glucose by employing the heterologous enzymes from *Clostridium beijinkerii*, *C. acetobutylicum* and *Treponema denticola* to supplement the native *E. coli* enzymes.

1.3 Synthetic biology and cell-free biocatalysis

Independent from the developments of cell-free enzyme synthesis, there have been many successful approaches of producing enzymes heterologously and combining purified enzymes or culture lysates into multi-enzyme biocatalytic pathways. Recent examples of synthetic cell-free pathways 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*.

Therefore, 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 coenzyme 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. A schematic overview of the tools and strategies that we will discuss in this review is given in Fig. 1.

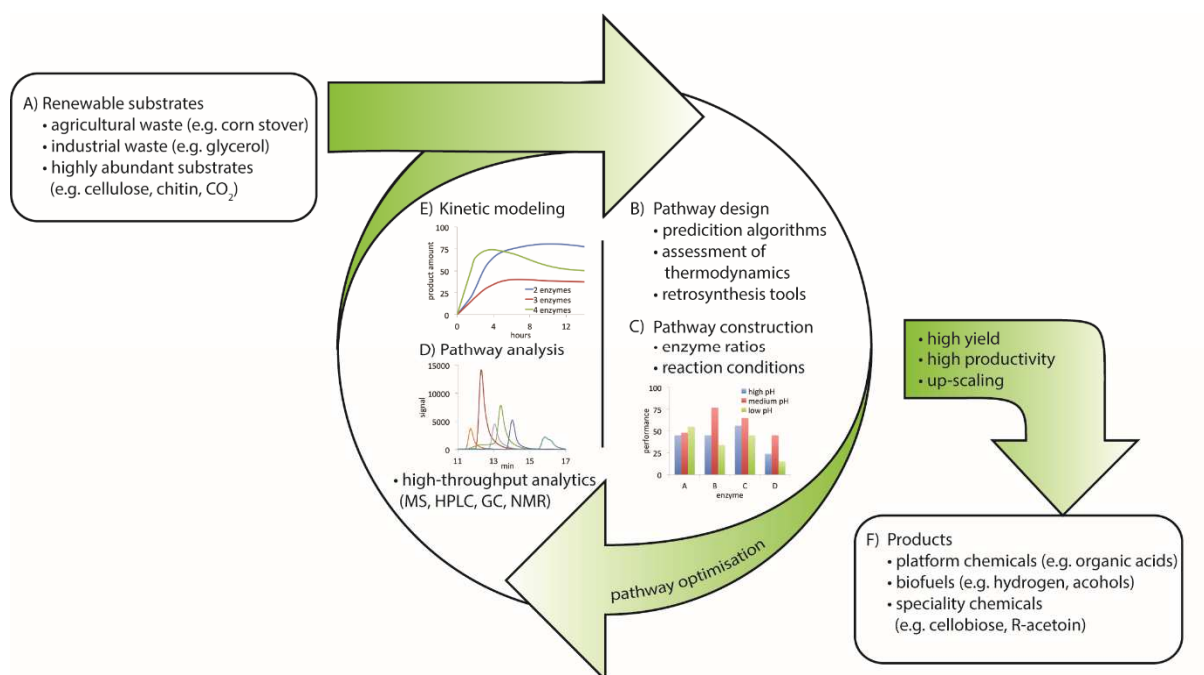


Figure 1. Diagram showing the process used to design, construct, and optimise cell-free biocatalytic pathways for the ‘green’ production of bio-based products. A) Substrates from renewable resources are selected to produce a desired product. B) A pathway is then designed based on the substrate, product and the enzymatic reactions required to achieve effective conversion. Pathway designs 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 (i.e. microorganisms), extracted and assembled into stepwise or one-pot pathway reactions. D) and E) To obtain the maximum product yield, the pathway can be optimised successively using high-throughput flux analysis and kinetic models that determine and improve the ideal enzyme ratios 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, organic acids) at a large scale.

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	Temperature (°C)	Time (h)	Product yield (% mol/mol)	Product titre (mM)	Ref
Monomeric							
CO ₂	Malate	17	30	1.5	CO ₂ fixation rate: 12 $\mu\text{M min}^{-1 b}$	0.54	(Schwande r 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 $\text{mM h}^{-1 b}$	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	(Krutsakor n 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)
	α/β -pinene				103.9	106.4	
	Fatty acids: decanoic acid, lauric acid, myristic acid, palmitic acid	30	37	4	9.3 ^d	30.04 $\text{mg L}^{-1 d}$	(Liu et al., 2017)
	Polyhydroxybutyrate	19	RT ^e	30 50	94 86	57 ^f 93.8	(Opgenort h et al., 2016). (Beer et al., 2017)
Glucuronate	α -Ketoglutarate	5	25	5	92	51.5	
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)
	<i>myo</i> -inositol	4	50 and 70	108	98	48	(Zhong et al., 2017b)

Table 1. continued

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	<i>myo</i> -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	<i>myo</i> -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, W. et al., 2017)
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)

2 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.

2.1 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), Table 1 (Wang, W. et al., 2017).

In a recent extensive report, the cell-free conversion of chitin into 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. ATP was dephosphorylated to AMP by endogenous *E. coli* enzymes that were still present after the crude lysate was heat-purified. Therefore, a polyphosphate-dependent ATP-producing and AMP-recycling pathway had to be integrated to enhance the reaction that increased the final product titre by 3.3-fold to 2.1 mM (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 Embden-Myerhof (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 *Escherichia coli* (*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 was achieved by assembling parts of the glycolytic pathway, the pentose phosphate pathway, the *Bifidobacterium* shunt and the polyhydroxybutyrate 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.

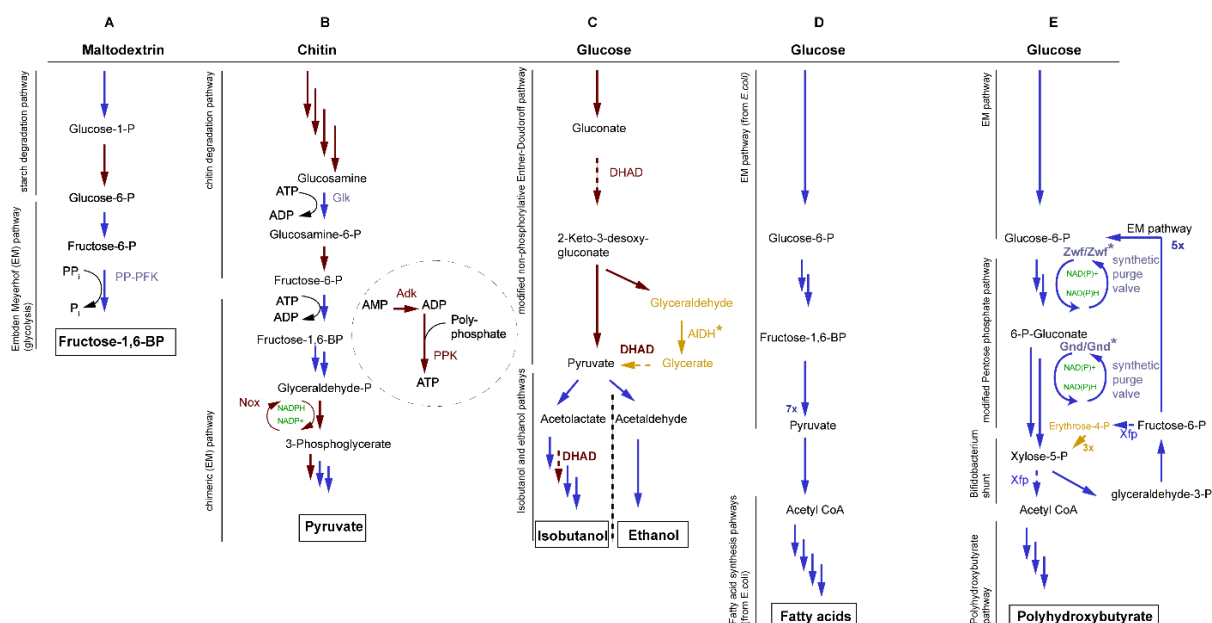


Figure 2. Schematic illustration of five recently described synthetic pathways motivated by natural metabolic pathways. Phosphate is denoted as 'P', pyrophosphate-dependent phosphofructokinase as 'PP-PFK', glucokinase as 'Glk', adenylate kinase as 'Adk', polyphosphate kinase as 'PKK' Dihydroxy acid dehydratase as 'DHAD', glyceraldehyde dehydrogenase as 'AIDH', Glucose-6-phosphate dehydrogenase as 'Zwf', 6-phosphogluconate dehydrogenase as 'Gnd' and xyululose-5-phosphate phosphoketolase as Xfp for brevity. Enzymes that were engineered are indicated with *. The number of arrows indicate the aggregate number of enzyme reactions required for the conversion of intermediates. Blue- and brown-coloured arrows indicate that the enzymes derived from bacteria and archaea, respectively, indicating from where synthetic enzyme assemblies were designed. The yellow-coloured arrows indicate 'salvaging' enzymes that were integrated for feed-back of by-products into the pathway. Dashed arrows indicate promiscuous enzymes. Cofactors that were recycled through the addition of recycling enzymes are highlighted in green. Further modifications to the natural pathways, pathway descriptions and references are described in the main text.

2.2 *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. 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. 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, L. et al., 2017). 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 efficiency (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. 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. As a consequence, computationally-guided *de novo* pathway design can 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).

2.3 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 CO₂, providing a promising alternative to the six CO₂ 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 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 CO₂ 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 theoretically perceived as i) certain intermediates generated within the pathway might inhibit other pathway enzymes, ii) intermediates are metabolised to undesired by-products by promiscuous enzymes, or iii) the overall enzyme kinetics are insufficient to support an effective pathway flux. Sub-optimal pathway performances then need to be improved by changing the input settings or by undergoing the traditional “design – build – test and learn” cycle routinely proposed for synthetic biology.

3 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), more than 5 billion m³ tons of plant biomass waste is generated annually from agriculture (<http://web.unep.org/gpwm/what-we-do/waste-agricultural-biomass>) 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 an 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 (2015) 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 more than 10 purified enzymes. The productivity of the pathway achieved was $54 \text{ mmol H}_2 \text{ L}^{-1} \text{ h}^{-1}$ 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. Saccharolytic enzymes usually operate at temperatures above 45°C , whereas fermentative organisms (e.g. *E. coli*, *S. cerevisiae*) can't sustain these high temperatures. The use of wild type or engineered thermophilic strains for fermentation still suffers from low yields, which is why a cell-free enzymatic pathway is a promising solution to overcome the challenges of combined saccharification and fermentation. Accordingly, Honda et al. (2017) reported a one-pot bioconversion of chitin to pyruvate using a cell-free pathway consisting of twelve thermophilic enzymes recombinantly-expressed in *E. coli* (table 1, figure 2 B, (Honda et al., 2017).

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 the price of glycerol has plummeted due to increased biodiesel production, 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 (table 1), (Gao et al., 2015). Pyruvate itself is a versatile C3 platform chemical that can be transformed further into the C4 platform chemical, acetoin. Acetoin exists as (3R)-acetoin and (3S)-acetoin and both isomers are widely used 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 obtain a more enantiopure acetoin, Gao et al. extended the synthetic cell-free pathway to convert pyruvate selectively into (3R)-acetoin by introducing a α -acetolactate synthase and α -acetolactate decarboxylase (table 1), (Gao et al., 2015). An 85.5% of the theoretical yield from glycerol to (3R)-acetoin (stereoisomeric purity of 95.4%) conversion 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 natural *Lactobacillus*-aided fermentation process (Hull et al., 1996). In the United States alone, more than 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. (2017) 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 and 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). Recently, alternative pathways have been discovered for carbohydrate metabolism which rely on previously unknown enzymes (Watanabe et al., 2008; Wolf et al., 2016). These new pathways 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.

4 The key players in biocatalysis – the enzymes

4.1 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; Karim and Jewett, 2018; Swartz, 2006; Swartz, 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 and catalytic efficiency (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. $> 80\text{ }^{\circ}\text{C}$ and denature irreversibly only at temperatures above their temperature optimum (Daniel et al., 2008). The rationale of using thermostable enzymes stems from i) 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) and ii) 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 iii) (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*, *Thermatoga maritima*) and hyperthermophilic archaea (e.g., *Archaeoglobus 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; Zhong et al., 2017b). The reactions were conducted at temperatures $> 50\text{ }^{\circ}\text{C}$ 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).

4.2 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; Daniel and Danson, 2013; Daniel et al., 2009; Eysenhard 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). Accordingly, they propose that successful screening for high(er) temperature enzymes will need to take into account both thermal stability and activity (via T_{eq}). The Equilibrium 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; Daniel and Danson, 2013). Instead, successful screening for high(er) temperature enzymes will need to account for both thermal

stability and activity, 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).

5 Cofactors 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. Cell-free cofactor recycling can be achieved by integrating cofactor regeneration pathways into the biocatalytic process to minimise cofactor dependency. 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 O₂ is reduced to H₂O 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). 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 (Opgeorth et al., 2016; van der Donk and Zhao,

2003; Zhang and Hess, 2017). An overview on requirements on cofactor regeneration pathways for cost-effective biocatalysis was given by (Zhao and van der Donk, 2003).

In cases where the biocatalytic process is conducted at temperatures $> 50\text{ }^{\circ}\text{C}$, heat-labile coenzymes like NAD(P)H and ATP degrade rapidly and make any regeneration process superfluous. For example, at $40\text{ }^{\circ}\text{C}$ and pH 6, the half-life of NADPH and NADH was shown to be 1 and 7 h, respectively (Wu et al., 1986). At $95\text{ }^{\circ}\text{C}$ and pH 7, the half-life of ATP and NADH was 1 h and less than five minutes, respectively, while NADPH was completely degraded (Wiegel and Michael, 2002), (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 (Honda et al., 2016) 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 $60\text{ }^{\circ}\text{C}$ as compared to a no salvage system where NAD^{+} was degraded to 1 mM. Another solution is to circumvent the dependency on expensive and heat-labile cofactors by designing alternative pathways that avoid the use of coenzymes completely, for example, through the assembly of ATP-independent metabolic pathways (Fig. 2A and 2C) or by replacing or engineering NADPH-dependent enzymes with NADH-dependent analogues (Chen et al., 2016; Huang et al., 2016)

As another alternative, synthetic NAD-based biomimetics have been developed that retain the enzyme-recognition moieties. The advantage of mimetics 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; Nowak et al., 2017a). However, enzymes may not always accept these biomimetics and may need to be adapted initially through laborious enzyme

engineering (Nowak et al., 2017b). For example, the cofactor binding site of a glucose dehydrogenase from *Sulfolobus solfataricus* (SsGDH) was engineered by computational design to regenerate a NADH-biomimetic with higher catalytic efficiency (Nowak et al., 2017b). 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 unfavourable coenzyme 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.

6 Co-immobilisation of multi-enzyme pathways

6.1 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) (Liu et al., 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).

6.2 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).

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 enzyme-gel interaction is based on physical adsorption between amino acids present on the enzyme's outer surface and functional groups present on the matrix. 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). Unlike simple physical adsorption, the weak interactions of enzyme and matrix are stabilised through entrapment and prevented from enzyme leaching that can occur under harsher reaction conditions, such as high flow rates, high ionic strength, variations in pH or temperature, as the binding is weak and reversible. 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., 2016; Liang et al., 2015; 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) 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 (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 (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).

Affinity tags. 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). They have been used to immobilise proteins in various biotechnological applications, including enzymes for biocatalysis (Care et al., 2015, 2016). For example, a SBP that binds strongly to inexpensive silica-containing materials was fused genetically to three different thermostable hemicellulases (Care et al., 2017; Sunna et al., 2013). 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. It is particularly significant that the natural zeolite used by Care et al (Care et al., 2017) is a bulk product and natural zeolites, especially clinoptilolite, have a variety of applications mainly due to their abundance and ion-

exchange properties. In the water and wastewater treatment industries, clinoptilolites have been introduced successfully for the removal of ammonium, the most common cation found in wastewaters. They are used widely as a catalyst to separate and purify gases in the petrochemical industry and as a component of fertilizer in agriculture and as animal feed additives. Despite the extensive variety of applications for natural zeolite, to our knowledge, they have not been used as an immobilisation matrix or carrier for proteins or other biomolecules.

In summary, 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.

Table 2. Table summarising the various techniques used for the co-immobilisation of multiple enzymes or biocatalytic pathways (composed of three or more enzymes). 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).

N° of enzymes	Biocatalyst	Application	Immobilisation	Support matrix	Recovered activity (%)	Relative activity ^a (%)	Re-usability ^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 cohesin-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)

Table 2. continued.

3	Bgl	Individual	Affinity	Zeolite	NA	99	50 ^{e,i} / 65 ^{h,i}	(Care et al., 2017b)
	Man	enzyme	binding via			56	25 ^{e,h} / 40 ^{h,i}	
	Xyn	reactions ^g	silica-binding-SBP/cross-linking			41	30 ^{e,h} / 60 ^{h,i}	
4	GOx	GOx and	Physical	Zinc/aden	94 ^d	260 in	NA	(Liang et al., 2016)
	HRP	HRP: two-	entrapment	osine	100 ^d	the		
	Lip	enzyme			77 ^d	glucose		
	Amyl	system for		5'-mono-phosphate (AMP) hydrogel	90 ^d	detection system		
		glucose detection. Lip and α -amylase: Individual enzyme reactions ^f						
5	GOX	Individual	Physical	Polyacryli	91 ^j	NA	NA	(Zore et al., 2017)
	AP	enzyme	entrapment/	c acid	970 ^j			
	LDH	reactions ^g	cross-linking	(PAA)	40 ^j			
	HRP			nanogel	54 ^j			
	Lip				450 ^j			
7	GalK	Cell-free	Affinity	Nickel-	140 ^k	142 ^l	NA	(Liu et al., 2002)
	GALT	production of	binding via	functional	96 ^k			
	UDPGP	UDP-	hexa-histidine	ised	71 ^k			
	PPA	galactose	tag	agarose	100 ^k			
	UMPK	from		beads	110 ^k			
	NDK	galactose			53 ^k			
	PPK				125 ^k			
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)

^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⁻¹ d⁻¹).

7 *In silico* kinetic modelling for improving pathway yields

7.1 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 modeled *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. (1987) used kinetic models of the reaction to calculate the optimum ratio of the enzymes, the optimum coenzyme concentration and the optimum concentration of the prochiral reaction intermediate, phenylpyruvate (Schmidt et al. 1987). 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 and applied mathematical models for each reaction step and for the complete system. 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. (2015) 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 $54 \text{ mmol } H_2 \cdot L^{-1} \cdot h^{-1}$.

Multi-enzyme systems for the industrial production of pharmaceutical and fine chemicals often rely on expensive cofactors that are one of the main 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. (2009) 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.

7.2 Simulation and optimisation

Describing and predicting *in vivo* metabolisms by models derived from single enzyme reactions is often impossible since in most cases, the kinetic parameters in multi enzyme pathways are very different (Teusink et al., 2000). However, simulations based on cell-free biocatalysis may provide valuable information for optimising multi-enzyme pathways *in vivo*. For example, a kinetic model that was based on kinetic data from the cell-free production of lactones (Milker et al., 2017) provided insight into bottlenecks encountered during the *in vivo* production. Specifically, a high discrepancy between the theoretical model and *in vivo* results revealed that the fastest enzyme of the pathway became rate limiting due to unfavorable cofactor conditions. 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).

7.3 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). In contrast, the Ensemble Modeling for Robustness analysis (EMRA) was shown to correctly predict the likelihood of system instability to occur within a cell-free enzyme pathway without prior knowledge of enzyme parameters (Theisen et al., 2016). Importantly, the EMRA investigates intrinsic system instability and its incapability to reach a steady-state upon drastical metabolic changes. This is

the case in batch-fed systems or when pathways constitute regulatory metabolic branch points. EMRA differs from sensitivity analysis where optimal conditions are identified by gradually altering the system conditions, which might not always be relevant for complex and highly regulated multi-enzyme pathways. Likewise, models were also 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 recognized one enzyme to as being responsible for maximum hydrogen yield (Ardao and Zeng, 2013; Ye et al., 2009).

7.4 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 more details in 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.

8 High-throughput metabolite analysis for system optimisation

8.1 Evaluation of methods

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 thermodynamics and kinetics, or due to the formation of unexpected by-products through promiscuous enzyme action. These understanding of these factors is crucial for 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.

8.2 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 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 subsequent 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 oxidising 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 in sample preparation, the use of HPLC or GC provides high robustness (i.e. less variation between samples measurements caused by the sample preparation) to the analysis 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 at 200 nm) to mM (DAD at 190 nm or RID) and poor selectivity since they cannot characterise analytes by unique chemical or structural features other than their chromatographic behavior as their retention time. 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. 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 30-60 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.

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. Single dot (•), poor performance; two dots (••), moderate performance; three dots (•••), high performance.

	Sample throughput	Costs (per run, purchase, maintenance)	Selectivity	Sensitivity	Robustness	Method development/validation	Data analysis	Multiple metabolite analysis ^a
Enzymatic assays	••	•••	••	•••	•	••	•••	•
Colorimetric assays	••	•••	•	•••	••	•••	•••	•
HPLC	•	••						
UHPLC	••	••						
DAD (≤ 200 nm)		••	•	•	•••	••	••	•
DAD (≥ 200 nm)		••	••	••	•••	••	••	•
PAD		••	••	••	•••	••	••	•
RID		••	•	•	••	••	••	•
GC-FID	•	••	••	••	••	••	••	•
LC-MS/GC-MS	••	•	•••	•••	•••	•	••	•••
MS ^b	•••	•	••	••	•	•	•	••
MS/MS ^b	•••	•	•••	•••	••	•	••	•••
NMR	••	•	••	•	•••	•••	•	••

^a e.g. > 10 metabolites

^b Direct injection

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, G6P,

glyceraldehyde-3-phosphate and fructose-1,6-bisphosphate were not separated and cannot be detected under these conditions (Krutsakorn et al., 2013; Opgenorth et al., 2016; Wang, W. et al., 2017; 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; Ye et al., 2013; Schwander et al., 2016; Dudley et al., 2016; Korman et al., 2017; Guterl et al., 2012). Despite their widespread use, these assays are laborious 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, 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 detection of coenzymes becomes impractical when they are recycled *in situ* or involved in high temperature processes in which they may be susceptible to degradation or when cofactor turnover does not correspond directly with substrate turnover (Zhang and Hess, 2017).

8.3 State of the art metabolite analysis

As a result of these issues, the use of highly selective and sensitive (nM - μ M) 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 coenzymes 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 75000 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 sec) 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 coenzyme supply (Hold et al., 2016).

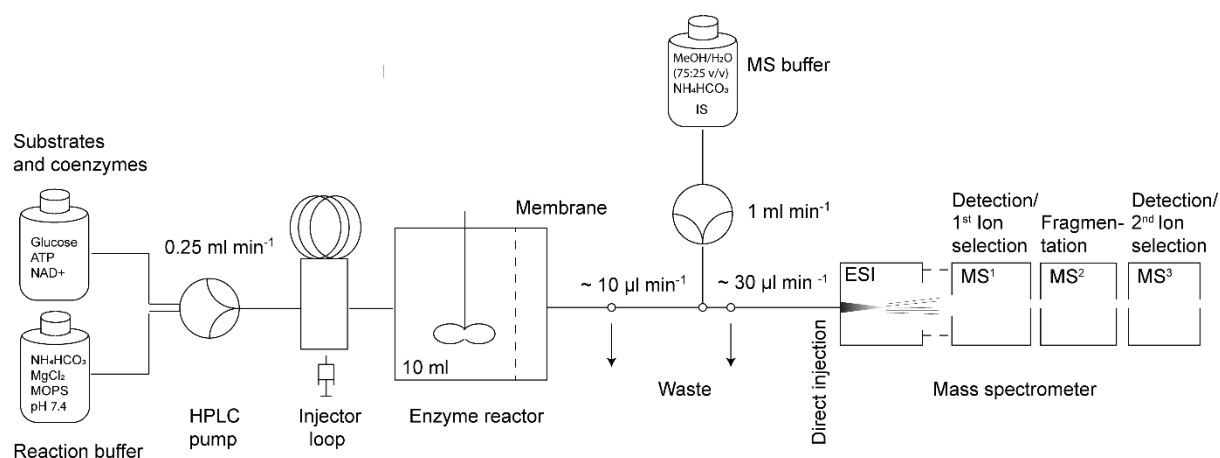


Figure 3. Schematic illustration of the experimental set-up for the real-time metabolite analysis of a synthetic enzyme pathway using electron spray ionisation (ESI)–triple quadrupole mass spectrometry, adapted from (Hold et al., 2016). Reaction substrates and coenzymes were mixed with the reaction buffer 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 enzymes. The effluent was diluted with MS-compatible buffer and internal standard (IS) before it was injected into the MS system.

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 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

(2PG) (Hold et al., 2016). 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). 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 second acquisition time) (Kautz et al., 2005). However, the lower sensitivity, longer analysis time, challenges in interpreting spectra from complex reaction mixtures and similarly high costs when compared to MS has limited the widespread application of NMR in the analysis of biocatalytic pathways (Emwas et al., 2013). Another powerful method is Raman spectroscopy (RS) which usually suffers from low sensitivity (Bumrah and Sharma, 2016). However, improved sensitivities (in the lower mM range) have been achieved by employing UV resonance Raman (UVR) 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 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 metabolite analysis of biocatalytic processes.

9 Conclusions and outlook

Cell-free multi-enzyme biocatalysis is expanding its scope steadily, 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, with 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 discovery and characterisation of more effective and stable biocatalytic enzyme pathways 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 computing power and machine learning may 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 might ensure that maximum performance is achieved. Ultimately, a vital crosstalk 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) and the Australian Research Council (CE140100003).

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.(US).
- Alonso, A., Marsal, S., Julià, A., 2015. Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. *Frontiers in Bioengineering and Biotechnology* 3, 23.
- Alter, T., Blank, L., Ebert, B., 2018. Genetic Optimization Algorithm for Metabolic Engineering Revisited. *Metabolites* 8(2), 33.
- 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. *Chemical Engineering Science* 87, 183-193.
- Bar-Even, A., Flamholz, A., Noor, E., Milo, R., 2012. Rethinking glycolysis: on the biochemical logic of metabolic pathways. *Nature Chemical Biology* 8, 509.
- Bar-Even, A., Milo, R., Noor, E., Tawfik, D.S., 2015. The Moderately Efficient Enzyme: Futile Encounters and Enzyme Floppiness. *Biochemistry* 54(32), 4969-4977.
- Bar-Even, A., Noor, E., Lewis, N.E., Milo, R., 2010. Design and analysis of synthetic carbon fixation pathways. *Proceedings of the National Academy of Sciences* 107(19), 8889-8894.
- Barbosa, O., Ortiz, C., Berenguer-Murcia, Á., Torres, R., Rodrigues, R.C., Fernandez-Lafuente, R., 2015. Strategies for the one-step immobilisation–purification of enzymes as industrial biocatalysts. *Biotechnology Advances* 33(5), 435-456.
- Beer, B., Pick, A., Sieber, V., 2017. *In vitro* metabolic engineering for the production of alpha-ketoglutarate. *Metabolic Engineering* 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. *Journal of Materials Chemistry B* 3(26), 5232-5240.
- Bornscheuer, U.T., 2016. Biocatalysis: Successfully Crossing Boundaries. *Angewandte Chemie International Edition* 55(14), 4372-4373.
- Bornscheuer, U.T., 2018. The fourth wave of biocatalysis is approaching. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376(2110).
- 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.
- Buchner, E., 1897. Alkoholische Gärung ohne Hefezellen. *Berichte der deutschen chemischen Gesellschaft* 30(1), 1110-1113.

- Bujara, M., Schumperli, M., Pellaux, R., Heinemann, M., Panke, S., 2011. Optimization of a blueprint for *in vitro* glycolysis by metabolic real-time analysis. *Nature Chemical Biology* 7(5), 271-277.
- Bumbrah, G.S., Sharma, R.M., 2016. Raman spectroscopy – Basic principle, instrumentation and selected applications for the characterization of drugs of abuse. *Egyptian Journal of Forensic Sciences* 6(3), 209-215.
- Care, A., Bergquist, P.L., Sunna, A., 2015. Solid-binding peptides: smart tools for nanobiotechnology. *Trends in Biotechnology* 33(5), 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. *Biotechnology for Biofuels* 10(1), 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. *Nature Biotechnology* 32(10), 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 Catalysis* 6(8), 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*. *Scientific Reports* 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. *Microbial Cell Factories* 16(1), 8.
- Cherubini, F., 2010. The biorefinery concept: Using biomass instead of oil for producing energy and chemicals. *Energy Conversion and Management* 51(7), 1412-1421.
- Chong, S., 2014. Overview of cell-free protein synthesis: historic landmarks, commercial systems, and expanding applications. *Current Protocols in Molecular Biology* 108, 16.30.11-11.
- Cira, L.A., Huerta, S., Hall, G.M., Shirai, K., 2002. Pilot scale lactic acid fermentation of shrimp wastes for chitin recovery. *Process Biochemistry* 37(12), 1359-1366.

- Corey, E.J., 1988. Robert Robinson Lecture. Retrosynthetic thinking-essentials and examples. *Chemical Society Reviews* 17(0), 111-133.
- Daniel, R.M., Cowan, D.A., 2000. Biomolecular stability and life at high temperatures. *Cellular and molecular life sciences: CMLS* 57(2), 250-264.
- Daniel, R.M., Danson, M.J., 2010. A new understanding of how temperature affects the catalytic activity of enzymes. *Trends in Biochemical Sciences* 35(10), 584-591.
- Daniel, R.M., Danson, M.J., 2013. Temperature and the catalytic activity of enzymes: A fresh understanding. *FEBS Letters* 587(17), 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. *The Biochemical Journal* 425(2), 353-360.
- de Souza, R., Miranda, L.S.M., Bornscheuer, U.T., 2017. A Retrosynthesis Approach for Biocatalysis in Organic Synthesis. *Chemistry (Weinheim an der Bergstrasse, Germany)* 23(50), 12040-12063.
- Delépine, B., Duigou, T., Carbonell, P., Faulon, J.-L., 2017. RetroPath2.0: A retrosynthesis workflow for metabolic engineers. *bioRxiv. Metabolic engineering* 45, 58-170.
- 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). *Journal of the American Chemical Society* 125(38), 11476-11477.
- DeVries, J.K., Zubay, G., 1967. DNA-directed peptide synthesis. II. The synthesis of the alpha-fragment of the enzyme beta-galactosidase. *Proceedings of the National Academy of Sciences of the United States of America* 57(4), 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 Synthetic Biology* 5(12), 1578-1588.
- Eisenthal, R., Peterson, M.E., Daniel, R.M., Danson, M.J., 2006. The thermal behaviour of enzyme activity: implications for biotechnology. *Trends in Biotechnology* 24(7), 289-292.
- Elias, M., Wieczorek, G., Rosenne, S., Tawfik, D.S., 2014. The universality of enzymatic rate-temperature dependency. *Trends in Biochemical Sciences* 39(1), 1-7.
- Emwas, A.-H., Salek, R., Griffin, J., Merzaban, J., 2013. NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations. *An Official Journal of the Metabolomics Society* 9(5), 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. *Biotechnology Journal* 9(11), 1446-1457.
- Findrik, Z., Vasic'-Racki, D., Lutz, S., Daussmann, T., Wandrey, C., 2005. Kinetic modeling of acetophenone reduction catalysed by alcohol dehydrogenase from *Thermoanaerobacter* sp. *Biotechnology Letters* 27(15), 1087-1095.
- Findrik, Z., Vasic-Racki, D., 2007. Biotransformation of D-methionine into L-methionine in the cascade of four enzymes. *Biotechnology and Bioengineering* 98(5), 956-967.
- 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. *Analytical Chemistry* 89(22), 12527-12532.
- Fontaine, N., Grondin-Perez, B., Cadet, F., Offmann, B., 2015. Modeling of a Cell-Free Synthetic System for Biohydrogen Production. *Journal of Computer Science & Systems Biology* 8(3).
- 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 Catalysis* 7(1), 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. *Applied and environmental microbiology*. AEM-00550.
- 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 Chemistry* 17(2), 804-807.
- 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. *Bioorganic & Medicinal Chemistry Letters* 22(5), 1931-1935.
- Geueke, B., Riebel, B., Hummel, W., 2003. NADH oxidase from *Lactobacillus brevis*: a new catalyst for the regeneration of NAD. *Enzyme and Microbial Technology* 32(2), 205-211.
- Goldsmith, M., Tawfik, D.S., 2017. Enzyme engineering: reaching the maximal catalytic efficiency peak. *Current Opinion in Structural Biology* 47, 140-150.
- Gooday, G.W., 1990. The Ecology of Chitin Degradation, in: Marshall, K.C. (Ed.) *Advances in Microbial Ecology*. 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. *Pacific Symposium on Biocomputing*. Pacific Symposium on Biocomputing, 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. *Biotechnology Journal* 12(7), 1700030-n/a.
- 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. *The Biochemical Journal* 474(1), 1-19.
- Guo, W., Sheng, J., Feng, X., 2017. Mini-review: *In vitro* Metabolic Engineering for Biomanufacturing of High-value Products. *Computational and Structural Biotechnology Journal* 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(11), 2165-2172.
- Hadadi, N., Hatzimanikatis, V., 2015. Design of computational retrobiosynthesis tools for the design of *de novo* synthetic pathways. *Current Opinion in Chemical Biology* 28(Supplement C), 99-104.
- Hogan, M.C., Woodley, J.M., 2000. Modelling of two enzyme reactions in a linked cofactor recycle system for chiral lactone synthesis. *Chemical Engineering Science* 55(11), 2001-2008.
- Hold, C., Billerbeck, S., Panke, S., 2016. Forward design of a complex enzyme cascade reaction. *Nature Communications* 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(.). *Metabolic Engineering* 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. *Journal of Bioscience and Bioengineering* 124(3), 296-301.
- Honig, M., Sondermann, P., Turner, N.J., Carreira, E.M., 2017. Enantioselective Chemo- and Biocatalysis: Partners in Retrosynthesis. *Angewandte Chemie (International ed. in English)* 56(31), 8942-8973.
- Huang, C.J., Lin, H., Yang, X., 2012. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *Journal of Industrial Microbiology & Biotechnology* 39(3), 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(+). *Scientific Reports* 6, 32644.

- Hull, S.R., Yang, B.Y., Venzke, D., Kulhavy, K., Montgomery, R., 1996. Composition of corn steep water during steeping. *Journal of Agricultural and Food Chemistry* 44(7), 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 Letters* 581(3), 413-420.
- Jewett, M.C., Calhoun, K.A., Voloshin, A., Wu, J.J., Swartz, J.R., 2008. An integrated cell-free metabolic platform for protein production and synthetic biology. *Molecular Systems Biology* 4, 220.
- Jia, F., Narasimhan, B., Mallapragada, S., 2014. Materials-based strategies for multi-enzyme immobilisation and co-localization: A review. *Biotechnology and Bioengineering* 111(2), 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. *Environmental Progress* 26(4), 338-348.
- Karim, A.S., Jewett, M.C., 2016. A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. *Metabolic Engineering* 36, 116-126.
- Karim, A.S., Jewett, M.C., 2018. *Cell-Free Synthetic Biology for Pathway Prototyping, Methods in Enzymology*. Academic Press.
- Kautz, R.A., Karger, B.L., Goetzinger, W.K., 2005. High-throughput microcoil NMR of compound libraries using zero-dispersion segmented flow analysis. *Journal of Combinatorial Chemistry* 7(1), 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(1), 111-127.
- Korman, T.P., Opgenorth, P.H., Bowie, J.U., 2017. A synthetic biochemistry platform for cell free production of monoterpenes from glucose. *Nature Communications* 8, 15526.
- 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 Science* 23(5), 576-585.
- 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. *Chemical Society Reviews* 43(8), 2587-2627.
- Kruger, A., Schafers, C., Schroder, C., Antranikian, G., 2018. Towards a sustainable biobased industry - Highlighting the impact of extremophiles. *New Biotechnology* 40(Pt A), 144-153.
- Krutsakorn, B., Honda, K., Ye, X., Imagawa, T., Bei, X., Okano, K., Ohtake, H., 2013. *In vitro* production of n-butanol from glucose. *Metabolic Engineering* 20(Supplement C), 84-91.

- Kuchler, A., Yoshimoto, M., Luginbuhl, S., Mavelli, F., Walde, P., 2016. Enzymatic reactions in confined environments. *Nature Nanotechnology* 11(5), 409-420.
- Kumar, A., Wang, L., Ng, C.Y., Maranas, C.D., 2018. Pathway design using *de novo* steps through uncharted biochemical spaces. *Nature Communications* 9(1), 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. *Biotechnology Advances* 28(4), 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 Sustainable Chemistry & Engineering* 5(4), 3456-3464.
- Liang, H., Jiang, S., Yuan, Q., Li, G., Wang, F., Zhang, Z., Liu, J., 2016. Co-immobilisation of multiple enzymes by metal coordinated nucleotide hydrogel nanofibers: improved stability and an enzyme cascade for glucose detection. *Nanoscale* 8(11), 6071-6078.
- Liang, K., Ricco, R., Doherty, C.M., Styles, M.J., Bell, S., Kirby, N., Mudie, S., Haylock, D., Hill, A.J., Doonan, C.J., Falcato, P., 2015. Biomimetic mineralization of metal-organic frameworks as protective coatings for biomacromolecules. *Nature Communications* 6, 7240.
- Liese, A., Hilterhaus, L., 2013. Evaluation of immobilized enzymes for industrial applications. *Chemical Society Reviews* 42(15), 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. *Journal of Molecular Catalysis B: Enzymatic* 4(1), 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. *Biotechnology Advances* 35(8), 950-970.
- 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. *Chemical communications (Cambridge, England)* 49(36), 3766-3768.
- 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(4), 348-355.
- 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. Liu, Z., Zhang, Y., Jia, X., Hu, M., Deng, Z., Xu, Y., & Liu, T. (2017). *ACS synthetic biology*, 6(4), 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. *International Journal of Hydrogen Energy* 40(30), 9113-9124.

- Luo, J., Meyer, A.S., Mateiu, R.V., Pinelo, M., 2015. Cascade catalysis in membranes with enzyme immobilisation for multi-enzymatic conversion of CO₂ to methanol. *New Biotechnology* 32(3), 319-327.
- Lynd, L.R., Wyman, C.E., Gerngross, T.U., 1999. Biocommodity Engineering. *Biotechnology Progress* 15(5), 777-793.
- Martín del Campo, J.S., Rollin, J., Myung, S., Chun, Y., Chandrayan, S., Patiño, R., Adams, M.W.W., Zhang, Y.H.P., 2013. High-Yield Production of Dihydrogen from Xylose by Using a Synthetic Enzyme Cascade in a Cell-Free System. *Angewandte Chemie International Edition* 52(17), 4587-4590.
- McCall, J., 2005. Genetic algorithms for modelling and optimisation. *Journal of Computational and Applied Mathematics* 184(1), 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(17), 3420-3427.
- Myung, S., Zhang, Y.H.P., 2013. Non-Complexed Four Cascade Enzyme Mixture: Simple Purification and Synergetic Co-stabilization. *PLOS ONE* 8(4), 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 Letters* 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. *Biotechnology and Bioengineering* 112(1), 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. *Proceedings of the National Academy of Sciences* 47(10), 1588-1602.
- 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. *Frontiers in Microbiology* 6(957).
- Nowak, C., Pick, A., Csepei, L.-I., Sieber, V., 2017a. Characterization of Biomimetic Cofactors According to Stability, Redox Potentials, and Enzymatic Conversion by NADH Oxidase from *Lactobacillus pentosus*. *ChemBioChem* 18(19), 1944-1949.
- Nowak, C., Pick, A., Lommes, P., Sieber, V., 2017b. Enzymatic Reduction of Nicotinamide Biomimetic Cofactors Using an Engineered Glucose Dehydrogenase: Providing a Regeneration System for Artificial Cofactors. *ACS Catalysis* 7(8), 5202-5208.
- Opgenorth, P.H., Korman, T.P., Bowie, J.U., 2014. A synthetic biochemistry molecular purge valve module that maintains redox balance. *Nature Communication* 5, 4113.
- Opgenorth, P.H., Korman, T.P., Bowie, J.U., 2016. A synthetic biochemistry module for production of bio-based chemicals from glucose. *Nature Chemical Biology* 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 immobilisation and its application in l-xylulose production. *RSC Advances* 7(6), 3488-3494.
- Paul, C.E., Hollmann, F., 2016. A survey of synthetic nicotinamide cofactors in enzymatic processes. *Applied Microbiology and Biotechnology* 100, 4773-4778.
- Peterson, M.E., Eienthal, R., Danson, M.J., Spence, A., Daniel, R.M., 2004. A New Intrinsic Thermal Parameter for Enzymes Reveals True Temperature Optima. *Journal of Biological Chemistry* 279(20), 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 Chemistry* 19(2), 372-375.
- Pleissner, D., Lin, C.S.K., 2013. Valorisation of food waste in biotechnological processes. *Sustainable Chemical Processes* 1(1), 21.
- Poppe, L., Vertessy, B.G., 2018. The Fourth Wave of Biocatalysis Emerges- The 13 th International Symposium on Biocatalysis and Biotransformations. *Chembiochem* 19(4), 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 (Oxford, England)* 33(21), 3454-3460.
- Quin, M.B., Wallin, K.K., Zhang, G., Schmidt-Dannert, C., 2017. Spatial organization of multi-enzyme biocatalytic cascades. *Organic & Biomolecular Chemistry* 15(20), 4260-4271.
- Randek, J., Mandenius, C.-F., 2017. On-line soft sensing in upstream bioprocessing. *Critical Reviews in Biotechnology*, 1-16.
- Reetz, M.T., 2013. Biocatalysis in Organic Chemistry and Biotechnology: Past, Present, and Future. *Journal of the American Chemical Society* 135(34), 12480-12496.
- Rios-Solis, 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 catalysed synthesis of chiral amino-alcohols. *Biochemical Engineering Journal* 73, 38-48.
- Risso, V.A., Sanchez-Ruiz, J.M., Ozkan, S.B., 2018. Biotechnological and protein-engineering implications of ancestral protein resurrection. *Current Opinion in Structural Biology* 51, 106-115.
- Rocha-Martín, J., Rivas, B.d.l., Muñoz, R., Guisán, J.M., López-Gallego, F., 2012. Rational Co-Immobilisation of Bi-Enzyme Cascades on Porous Supports and their Applications in Bio-Redox Reactions with *In situ* Recycling of Soluble Cofactors. *ChemCatChem* 4(9), 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 coutilisation and kinetic modeling. *Proceedings of the National Academy of Sciences* 112(16), 4964-4969.
- Rosenblum, G., Cooperman, B.S., 2014. Engine out of the chassis: cell-free protein synthesis and its uses. *FEBS Lett* 588(2), 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. *Nature Catalysis* 1(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 (New York, N.Y.)* 329(5989), 305-309.
- Scherkus, C., Schmidt, S., Bornscheuer, U.T., Groger, H., Kara, S., Liese, A., 2017. Kinetic insights into ϵ -caprolactone synthesis: Improvement of an enzymatic cascade reaction. *Biotechnology and Bioengineering* 114(6), 1215-1221.
- Schmidt-Dannert, C., Lopez-Gallego, F., 2016. A roadmap for biocatalysis - functional and spatial orchestration of enzyme cascades. *Microbial Biotechnology* 9(5), 601-609.
- Schmidt, E., Vasić-Rački, Đ., Wandrey, C., 1987. Enzymatic production of l-phenylalanine from the racemic mixture of d,l-phenyllactate. *Applied Microbiology and Biotechnology* 26(1), 42-48.
- 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 (New York, N.Y.)* 354(6314), 900-904.
- Secundo, F., 2013. Conformational changes of enzymes upon immobilisation. *Chemical Society Reviews* 42(15), 6250-6261.
- Sheldon, R.A., Brady, D., 2018. The limits to biocatalysis: pushing the envelope. *Chemical Communications* 54(48), 6088-6104.
- Sheldon, R.A., Pereira, P.C., 2017. Biocatalysis engineering: the big picture. *Chemical Society Reviews*.
- Sheldon, R.A., van Pelt, S., 2013. Enzyme immobilisation in biocatalysis: why, what and how. *Chemical Society Reviews* 42(15), 6223-6235.
- Shi, J., Wu, Y., Zhang, S., Tian, Y., Yang, D., Jiang, Z., 2018. Bioinspired construction of multi-enzyme catalytic systems. *Chemical Society Reviews* 47(12), 4295-4313.
- Sokic-Lazic, D., Minteer, S.D., 2008. Citric acid cycle biomimic on a carbon electrode. *Biosensors and Bioelectronics* 24(4), 939-944.

- Sperl, J.M., Sieber, V., 2018. Multienzyme Cascade Reactions—Status and Recent Advances. *ACS Catalysis* 8(3), 2385-2396.
- Steffler, F., Guterl, J.K., Sieber, V., 2013. Improvement of thermostable aldehyde dehydrogenase by directed evolution for application in Synthetic Cascade Biomanufacturing. *Enzyme and Microbial Technology* 53(5), 307-314.
- Sunna, A., Chi, F., Bergquist, P.L., 2013. A linker peptide with high affinity towards silica-containing materials. *New Biotechnology* 30(5), 485-492.
- Swartz, J., 2006. Developing cell-free biology for industrial applications. *Journal of Industrial Microbiology & Biotechnology* 33(7), 476-485.
- Swartz, J.R., 2012. Transforming biochemical engineering with cell-free biology. *AIChE Journal* 58(1), 5-13.
- Swartz, M., 2010. HPLC detectors: a brief review. *Journal of Liquid Chromatography & Related Technologies* 33(9-12), 1130-1150.
- Taniguchi, H., Okano, K., Honda, K., Modules for *in vitro* metabolic engineering: Pathway assembly for bio-based production of value-added chemicals. *Synthetic and Systems Biotechnology* 2(2), 65-74.
- Tawfik, D.S., 2014. Accuracy-rate tradeoffs: how do enzymes meet demands of selectivity and catalytic efficiency? *Current Opinion in Chemical Biology* 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. *European Journal of Biochemistry* 267(17), 5313-5329.
- Theisen, M.K., Lafontaine Rivera, J.G., Liao, J.C., 2016. Stability of Ensemble Models Predicts Productivity of Enzymatic Systems. *PLOS Computational Biology*. 12(3), e1004800.
- Torres, P., Batista-Viera, F., 2017. Immobilized Trienzymatic System with Enhanced Stabilization for the Biotransformation of Lactose. *Molecules (Basel, Switzerland)* 22(2).
- van der Donk, W.A., Zhao, H., 2003. Recent developments in pyridine nucleotide regeneration. *Current Opinion in Biotechnology* 14(4), 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. *Biotechnology and Bioengineering* 102(5), 1475-1482.
- Vasic-Racki, D., Kragl, U., Liese, A., 2003. Benefits of enzyme kinetics modelling. *Chemical and Biochemical Engineering Quarterly* 17(1), 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. *Bioresource Technology* 239, 507-517.

- Wang, L., Dash, S., Ng, C.Y., Maranas, C.D., 2017. A review of computational tools for design and reconstruction of metabolic pathways. *Synthetic and Systems Biotechnology* 2(4), 243-252.
- Wang, W., Liu, M., You, C., Li, Z., Zhang, Y.-H.P., 2017. ATP-free biosynthesis of a high-energy phosphate metabolite fructose 1,6-diphosphate by *in vitro* metabolic engineering. *Metabolic Engineering* 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. *The Journal of Biological Chemistry* 283(29), 20372-20382.
- Wheeldon, I., Minter, S.D., Banta, S., Barton, S.C., Atanassov, P., Sigman, M., 2016. Substrate channelling as an approach to cascade reactions. *Nature Chemistry* 8(4), 299-309.
- Whittaker, J.W., 2013. Cell-free protein synthesis: the state of the art. *Biotechnology Letters* 35(2), 143-152.
- Wiegel, J., Michael, A.W., 2002. *Thermophiles: The Keys to the Molecular Evolution and the Origin of Life?* CRC Press.
- Wilding, K.M., Schinn, S.-M., Long, E.A., Bundy, B.C., 2018. The emerging impact of cell-free chemical biosynthesis. *Current Opinion in Biotechnology* 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. *Molecular Microbiology* 102(5), 882-908.
- Woodward, J., Orr, M., Cordray, K., Greenbaum, E., 2000. Biotechnology: enzymatic production of biohydrogen. *Nature* 405(6790), 1014-1015.
- Wu, J.T., Wu, L.H., Knight, J.A., 1986. Stability of NADPH: effect of various factors on the kinetics of degradation. *Clinical Chemistry* 32(2), 314-319.
- Wu, X., Yang, C., Ge, J., 2017. Green synthesis of enzyme/metal-organic framework composites with high stability in protein denaturing solvents. *Bioresources and Bioprocessing* 4(1), 24.
- Xiao, Z., Lu, J.R., 2014. Strategies for enhancing fermentative production of acetoin: A review. *Biotechnology Advances* 32(2), 492-503.
- 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. *Synthetic and Systems Biotechnology*.
- Yan, N., Chen, X., 2015. Sustainability: Don't waste seafood waste. *Nature* 524(7564), 155-157.

- 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. *Critical Reviews in Biotechnology* 37(8), 990-1005.
- Yang, W.C., Patel, K.G., Wong, H.E., Swartz, J.R., 2012. Simplifying and streamlining *Escherichia coli*-based cell-free protein synthesis. *Biotechnology Progress* 28(2), 413-420.
- Ye, X., Honda, K., Morimoto, Y., Okano, K., Ohtake, H., 2013. Direct conversion of glucose to malate by synthetic metabolic engineering. *Journal of Biotechnology* 164.
- 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. *Microbial Cell Factories* 11(1), 120.
- 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 catalysed by enzyme cocktails. *ChemSusChem* 2(2), 149-152.
- You, C., Myung, S., Zhang, Y.H.P., 2012. Facilitated Substrate Channeling in a Self-Assembled Trifunctional Enzyme Complex. *Angewandte Chemie International Edition* 51(35), 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. *Biotechnology and Bioengineering* 114(8), 1855-1864.
- You, C., Zhang, Y.H.P., 2013. Self-Assembly of Synthetic Metabolons through Synthetic Protein Scaffolds: One-Step Purification, Co-immobilisation, and Substrate Channeling. *ACS Cynthetic Biology* 2(2), 102-110.
- Zampieri, M., Sekar, K., Zamboni, N., Sauer, U., 2017. Frontiers of high-throughput metabolomics. *Current Opinion in Chemical Biology* 36, 15-23.
- Zdarta, J., Meyer, A., Jesionowski, T., Pinelo, M., 2018. A General Overview of Support Materials for Enzyme Immobilisation: Characteristics, Properties, Practical Utility. *Catalysts* 8(2), 92.
- Zhang, Y., Hess, H., 2017. Toward Rational Design of High-efficiency Enzyme Cascades. *ACS Catalysis* 7(9), 6018-6027.
- Zhang, Y., Tsitkov, S., Hess, H., 2016. Proximity does not contribute to activity enhancement in the glucose oxidase–horseradish peroxidase cascade. *Nature Communications* 7, 13982.
- Zhang, Y.P., Sun, J., Ma, Y., 2017. Biomanufacturing: history and perspective. *Journal of Industrial Microbiology & Biotechnology* 44(4-5), 773-784.
- Zhao, H., van der Donk, W.A., 2003. Regeneration of cofactors for use in biocatalysis. *Current Opinion in Biotechnology* 14(6), 583-589.

Zhong, C., Wei, P., Zhang, Y.-H.P., 2017a. A kinetic model of one-pot rapid biotransformation of cellobiose from sucrose catalysed by three thermophilic enzymes. *Chemical Engineering Science* 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 Catalysis* 7(9), 5992-5999.

Zhu, Z., Zhang, Y.H.P., 2017. *In vitro* metabolic engineering of bioelectricity generation by the complete oxidation of glucose. *Metabolic Engineering* 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 Advances* 7(47), 29563-29574.

Zubay, G., 1973. *In vitro* synthesis of protein in microbial systems. *Annual Review of Genetics* 7, 267-287.

1.3 Aims and Scope of this Thesis

The overarching aim of this thesis was to establish the cell-free production of the platform chemical GlucA. GlucA has been produced previously only by chemical or metabolic synthesis but those processes are environmental unfriendly or lack in efficiency. The rationale for this aim was the high degree of control that can be exerted on cell-free biocatalytic systems and the great flexibility by which such systems can be manipulated and optimised. In this way, their efficiency and sustainability may be extended beyond the limits of chemical or microbial synthesis and may promote next generation bio-manufacturing based on cell-free biocatalysis.

To realise the comprehensive aim as stated, the specific objectives of this study included:

- i) To design a synthetic pathway for the cell-free enzymatic production of GlucA from inexpensive starting materials.
- ii) To develop an analytical method for the simultaneous analysis of GlucA and all pathway intermediates to allow metabolic flux analysis and optimisation of the pathway.
- iii) To employ thermostable enzymes for high system robustness.
- iv) To immobilise all enzymes of the pathway onto a low-cost and industrially-compatible silica-based material for recycling enzymes and to reduce enzyme production costs.
- v) To establish a cofactor recycling system to achieve reduced cofactor requirements for the process and minimise operational costs.

The main work presented in this thesis has been summarised in 5 different chapters. Three of these chapters are presented in manuscript formats which are either submitted or ready for submission to peer-reviewed scientific journals. Chapters for publication include a review article which establishes the context and the current understanding of the research being

reported in this thesis (publication 1, **Chapter 1**). The second manuscript (**Chapter 3**) describes the novel analytical method developed for simultaneous analysis of GlucA and its intermediates. The third manuscript for publication (**Chapter 4**) covers results from the cell-free production of GlucA, including pathway design, enzyme immobilisation and cofactor recycling.

Chapter 2 lists the materials and methods used in this work with referral to the chapters where a detailed description of each method can be found. Additional work that resulted from investigation of the characterisation of a putative, thermostable MIOX enzyme is described in **Chapter 5**, which is not presented for publication as the results are preliminary. A concluding summary and future perspectives are provided in **Chapter 6**. Preliminary work for this thesis which resulted in one publication is provided in the **Appendix**.

Chapter 2:

Materials and methods

Materials and methods employed in this work are summarised in Table 2.1. The detailed descriptions for each method can be found in the referred chapter.

Table 2.1. Materials and methods employed in this work.

Materials and methods	Refer to Chapter
UHPLC-RID method development	3
UHPLC-RID method validation	3
Synthetic GlucA pathway design	4
Thermodynamic pathway calculations	4
Rational enzyme selection for cell-free GlucA production	4
Bacterial strains, culture conditions and plasmid construction	3, 4, 5
Production and partial purification of recombinant enzymes	3, 4, 5
Enzyme activity assays and enzyme characterisation	3, 4, 5
Substrate inhibition test	4
Enzyme stability assays	4
Immobilisation of single enzymes	4
Co-immobilisation of enzymes and recycling studies	4
Cell-free enzymatic synthesis of GlucA from G1P	4
Sample analysis with UHPLC-RID	3, 4

Table 2.1. Continued.

LC-MS analysis	4
NMR spectroscopy	4
Database search	5
Multiple sequence alignment	5
Biochemical characterisation of a putative, thermostable MIOX	5
Protein refolding	5

Chapter 3:

Mixed-mode liquid chromatography for the rapid analysis of biocatalytic glucaric acid reaction pathways

3.1 Introduction

Metabolic flux analysis is a critical tool in the realisation and optimisation of cell-free multi-enzyme pathways. Monitoring the concentrations of substrate, intermediates and final product allows users to rapidly identify pathway bottlenecks and adjust specific parameters (e.g. enzyme, substrate and cofactor concentrations) to optimise the metabolic flux towards high conversion yields and product titres. However, current analytical approaches employed for the analysis of GlucA and intermediates rely on a combination of enzymatic, colorimetric and chromatographic techniques, which are laborious, time consuming and impractical for rapid and effective metabolic flux analysis. Therefore, a novel analytical method was developed, which is simple, rapid, accurate and precise. The new method was used to monitor the cell-free production of glucaric acid in a synthetic enzymatic pathway.

This manuscript has been submitted to *Analytica Chimica Acta* for publication as a peer-reviewed paper.

3.2 Contribution to Manuscript 2

The concept of this publication was developed in partnership with my supervisors Anwar Sunna and Andrew Care. The experiments were designed, conducted and analysed by me. Advice on the experimental design was given by Martin Waterstraat. The manuscript was prepared by me and was reviewed by all co-authors before submission to the journal. The contributions of each author are detailed in Table 3.1.

Table 3.1 Author contribution summary for Publication 2.

	Kerstin Petroll	Martin Waterstraat	Andrew Care	Peter Bergquist	Anwar Sunna
Conception	•		•		•
Experiment Design	•	•			
Data Collection	•				
Data Analysis	•				
Manuscript	•	•	•	•	•

Keywords

Mixed mode ultra-high performance liquid chromatography; hydrophilic interaction chromatography; weak anion exchange chromatography; biocatalysis; bio-based glucaric acid.

Abstract

Glucaric acid (GlucA) has been identified as one of the top 10 potential bio-based chemicals for replacement of oil-based chemicals. Several synthetic enzyme pathways have been engineered in bacteria and yeast to produce GlucA from glucose and *myo*-inositol. However, the yields and titres achieved with these systems remain too low for the requirements of a bio-based GlucA industry. A major limitation for the optimisation of GlucA production via synthetic enzymatic pathways are the laborious analytical procedures required to detect the final product (GlucA) and pathway intermediates. We have developed a novel method for the simple and simultaneous analysis of GlucA and pathway intermediates to address this limitation using mixed mode (MM) HILIC and weak anion exchange chromatography (WAX), referred to as MM HILIC/WAX, coupled with RID. Isocratic mobile phase conditions and the sample solvent were optimised for the separation of GlucA, glucose-1-phosphate (G1P), glucose-6-phosphate (G6P), inositol-1-phosphate (I1P), *myo*-inositol and glucuronic acid (GA). The method showed good repeatability, precision and excellent accuracy with detection and quantitation limits (LOD and LOQ) of 1.5-2 and 577 mM, respectively. The method developed was used for monitoring the enzymatic synthesis of the final step in the GlucA pathway, and showed that GlucA was produced from GA with near 100% conversion and a titre of 9.2 g L⁻¹.

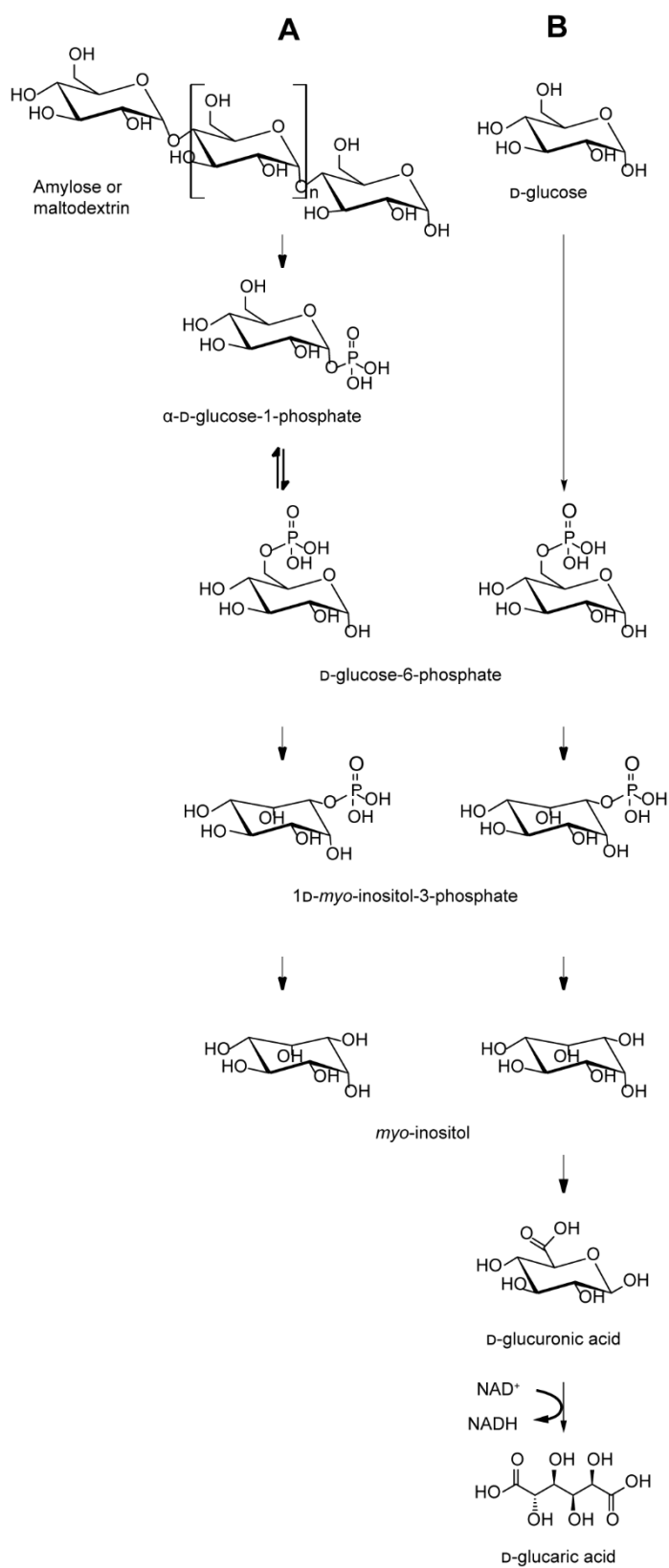
1 Introduction

Glucaric acid (GlucA) is a glucose-derived compound with a high market value due to its widespread applications as a renewable bio-detergent [1], anti-corrosive [2], a starting material for bio-based polymers (e.g. Nylon 6.6 or hydroxylated nylon) [3, 4], as an agent against cancer [5] or an agent to prevent diarrhoea that accompanies cancer treatment [6]. The production of bio-based GlucA from renewable resources like glucose, sucrose or *myo*-inositol has become more economical with the development of novel biocatalytic pathways that are able to catalyse the conversion of these sugars to GlucA in fewer steps than the natural pathway. The synthetic biocatalytic pathways were derived from combinations of components of natural pathways present in eukaryotic and bacteria/archaea and have been genetically-engineered into several microorganisms such as *Escherichia coli*, *Saccharomyces cerevisiae* and *Pichia pastoris* to enable the metabolic *in vivo* production of GlucA (Fig. 1B). Accordingly, engineered *E. coli* systems have been reported to produce 0.012 - 2.5 g L⁻¹ GlucA from glucose [7-10] 1.42 g L⁻¹ GlucA from sucrose [11] and 4.85 g L⁻¹ GlucA from *myo*-inositol [12]. GlucA production in engineered yeast has been reported for *S. cerevisiae* with 0.98 g L⁻¹ from glucose [13], and 6 g L⁻¹ from *myo*-inositol [14], while *P. pastoris* was able to produce 6.61 g L⁻¹ from a mixture of glucose and *myo*-inositol [15]. Recently, two cell-free systems were developed for the biocatalytic production of the GlucA precursor *myo*-inositol from starch by modifying the upper part of the GlucA pathway described above [16, 17] (Fig 1A). Combining enzymes from these new and previous pathways would be a promising approach for producing GlucA from crude, unrefined biomass and would be less expensive and more sustainable. Despite these recent advances in the production of bio-based GlucA, the maximum yield reported for GlucA was only 0.25 g per g of glucose [8] and 0.56 g per g of *myo*-inositol [14], which is too inefficient for cost-effective GlucA production on an industrial scale.

There are several potential bottlenecks within a synthetic biocatalytic pathway that may contribute to low metabolic fluxes. These include the accumulation of intermediates due to unfavourable thermodynamics and enzyme kinetics, thermal-/pH-induced degradation of intermediates or enzyme promiscuity towards intermediates that can result in their removal from the pathway. Comprehensive understanding of the dynamics of an enzyme pathway requires the analysis of all reactants, intermediates and products. Only then accurate conclusions can be drawn that are necessary for the effective elimination of pathway bottlenecks [18-22]. Common intermediates in the GlucA and *myo*-inositol pathways are α -D-glucose-1-phosphate (G1P), D-glucose-6-phosphate (G6P), D-*myo*-Inositol-3-phosphate (I1P), *myo*-inositol and D-glucuronic acid (GA) (Fig 1). No method has been described so far for the simultaneous detection of GlucA and its intermediates. Therefore, these compounds are either detected laboriously by the combination of multiple enzymatic, colorimetric or HPLC-based assays or not detected at all, which has hindered the development and optimisation of high-yielding bio-based GlucA systems [7, 11, 15-17].

We describe here a new and simple method for simultaneously quantitating GlucA and its intermediates; G1P, G6P and I1P, *myo*-inositol and GA using UHPLC with isocratic elution and online-coupled RID. The method is rapid (less than 10 min), simple (minimal sample preparation required), highly accurate and precise with LOD and LOQs in the lower mM range. Consequently, this novel approach provides a valuable tool for the assessment and optimisation of GlucA production in metabolically-engineered microorganisms and cell-free systems.

Figure 1. Synthetic enzyme pathways for the production of (A) *myo*-inositol from either amylose or maltodextrin [16, 17], or (B) GlucA from glucose [7].



2 Experimental

2.1 Chemicals

All chemicals and reagents were purchased from Sigma-Aldrich (Sigma-Aldrich, Castle Hill, Australia) unless stated otherwise. Luria-broth (LB) media was purchased from Invitrogen (Invitrogen, Carlsbad, USA) and Isopropyl β -D-thiogalactoside (IPTG) was purchased from Promega (Promega, Corporation, Madison, USA). Single stock solutions of α -D-Glucose-1-phosphate (G1P) disodium salt hydrate, D-glucose-6-phosphate (G6P) sodium salt, D-*myo*-Inositol-3-phosphate (IIP) (sodium salt) (Cayman Chemical, Michigan, USA), *myo*-inositol and D-glucuronic acid (GA) were prepared at a concentration of 200 mM in deionised water (18 M Ω) from a Milli-Q system (Millipore, Bedford, USA). Stock solutions of D-glucaric acid (GlucA) potassium salt were prepared at a concentration of 150 mM in deionised water. All solutions were filtered through a 0.45 μ m filter (Sartorius AG, Goettingen, Germany) and stored in portions at -30 °C. Standard working solutions were either prepared freshly by appropriate dilutions or from stock solutions that were stored at -30 °C and used for a maximum of three freeze-thawing cycles.

2.2 UHPLC-RID analysis

UHPLC was performed on a 1290 Infinity LC System (Agilent Technologies, Santa Clara, USA) equipped with a binary gradient pump, an autosampler, a degasser and a temperature-controlled column compartment (TCC). Detection was carried out using an RID (1260 model, Agilent Technologies). Samples (4 μ l) were injected onto a GlycanPac AXH-1 column (Thermo Fisher Scientific, Waltham, USA, 2.1 x 150 mm, 1.9 μ m particle diameter) after passing an online-coupled 0.2 μ m particle filter. The mobile phase of the method developed consisted of acetonitrile (ACN):ammonium formate (70 mM, pH 3.5, 80:20, v/v). The pH of

the buffer was adjusted with formic acid to 3.5 before mixing with ACN. The flowrate was set to 0.3 ml min⁻¹ and the TCC was maintained at 25 °C.

2.3 Method development

The selection of the appropriate column was based on an initial assessment of established analytical techniques and columns for the target compounds. To improve the separation, the composition of the mobile phase, temperature and flow rate were optimised by maintaining a single parameter and changing conditions one at a time. The separation of all compounds was assessed by using a mixture of all standard compounds (referred to as standard mixture in the following text) at concentrations of 5 or 10 mM in ACN:ammonium formate (50 mM, pH 3.5, 50:50, v/v), unless stated otherwise. After optimisation of the mobile phase, the composition of the sample solvent was amended to improve analyte separation and maximise signal intensity. Peaks were identified from the retention time (RT) assessed for single standards.

2.4 Method validation

To evaluate the sensitivity and linearity of the method, standard mixtures at 5, 10, 20, 30, and 40 mM were prepared in triplicate. Sample solvent was 50 mM ammonium formate, 50 mM dehydrate 4-(2-hydroxyethylhydroxyethyl) piperazine-1-ethanesulfonic acid (HEPES) buffer (pH 8) or 50 mM (2-(*N*-morpholino) ethanesulfonic acid) (MES) buffer (pH 8) with 30% vol. ACN and 0.02% vol. trifluoroacetic acid (TFA). Linear regression was used to calculate the calibration functions. The LOD and LOQ were determined as the concentrations of standard compounds which resulted in a signal-to-noise ratio of approximately 3:1 and 10:1, respectively [23]. To evaluate the repeatability and precision of the method, the standard compounds at concentrations of 10 and 30 mM were analysed in triplicate, and RSD of the peak areas were

calculated. The repeatability of the RT was calculated as RSD from six individual injections of each standard compound at a concentration of 10 mM.

The accuracy and precision of the method was determined by spiking bacterial cell lysates with the target compounds. *E. coli* BL21(DE3) cells (New England Biolabs, Ipswich, USA) were grown in LB medium supplemented with carbenicillin (50 µg/ml) at 37 °C with shaking (250 rpm). After reaching an A_{600} between 0.7 and 0.9, cells were harvested by centrifugation (4000 rcf, 20 min, 4 °C) and stored at -30 °C. For cell lysis, cells were resuspended in ice-cold lysis buffer (50 mM HEPES, pH 8.0 or 50 mM MES, pH 6 with a cells (g): lysis buffer (ml) ratio of 1:4 (m/v)) and disrupted by three passages through a French pressure cell at 10,000 Psi with further incubation with 25 U Benzonase for 20 min at 4 °C. Cellular debris was removed by centrifugation (12,000 rcf, 30 min, 4 °C) and the final crude lysate was filtered successively through 0.45 µm and 0.2 µm sterile filters (Sartorius AG) followed by storage at -30 °C. Standard mixtures of *myo*-inositol, GA, G1P, G6P and GlucA were spiked into the lysates in three independent replicates at final concentrations of 10 mM and 20 mM, while IIP was spiked into the lysates at only a final concentration of 10 mM. UHPLC-RID confirmed that these compounds were under the detectable level in the original lysates and thus are suitable 'empty' matrices. 20 µL of the spiked bacterial cell lysates were ultrafiltered through a 10 kDa cut-off filter (ACROPREP ADV.350µL 10K, Pall Corporation, New York, USA) using a multi-plate vacuum manifold (Pall Corporation) to remove bacterial proteins. ACN was added to the filtrates to a final content of 30% vol. prior to UHPLC analysis. TFA was added to all samples to a final concentration of 0.02% vol.

2.5 Sample preparation

The UHPLC-RID-method developed was used to analyse the biocatalytic conversion of GA to GlucA by uronate dehydrogenase (Udh) from *Fulvimarina pelagi* (HTCC2506) in bacterial cell lysates. For the heterologous production of Udh, the DNA sequence encoding the Udh gene (protein sequence GenBank™ Q0FXI7) was codon optimised for expression in *E. coli* and synthesised as gBlocks® Gene Fragments (Integrated DNA Technologies, Illinois, USA). The *udh* gene synthesised was introduced into the expression vector pET22b (Novagen, Merck Millipore) and transformed into *E. coli* BL21 (DE3) cells. They were grown in LB medium supplemented with 50 mg/ml carbenicillin at 37 °C with shaking (250 rpm) until the A_{600} was approximately 0.6. Recombinant protein synthesis was induced by the addition of 0.4 mM isopropyl β -D-thiogalactoside (IPTG) and cultivation of cells at 16 °C for 21h. After harvesting the cells by centrifugation (4000 rfc, 30 min, 4 °C), crude cell lysates were obtained as described in section 2.4. Enzyme assays were performed by incubating the Udh-containing crude lysate at 40 °C with 40 mM GA and 40 mM NAD^+ in 50 mM HEPES buffer, pH 8. As a negative control, enzyme assays were performed with crude lysates from *E. coli* that did not express Udh. Enzymes were inactivated after incubation at different time intervals (0, 1, and 2 h) by boiling for 15 min. After cooling on ice and the addition of TFA to a final concentration of 0.02% vol., the samples were ultrafiltered through a 10 kDa cut off filter plate, mixed with ACN to a final concentration of 30% vol. and subjected to UHPLC-RID analysis as described in section 2.4.

3 Results and discussion

3.1 Method development

3.1.1 Selection of a suitable column and detector

Column

We performed an extensive review of the established analytical techniques present within the literature to determine the optimal chromatographic mode for the separation of GlucA and its precursor compounds. As shown in both Fig. 1 and Table 1, GlucA and its precursors represent a combination of neutral and multiple anionic compounds that have a variety of functional groups but very similar structures, which makes their chromatographic separation highly problematic. Specifically, *myo*-inositol is a glucose-derived sugar alcohol (neutral), while all other compounds are anionic. GlucA is the dicarboxylic acid of glucose whereas GA is the monocarboxylic acid of glucose and G1P and G6P are constitutional isomers of a glucose monophosphate ester. I1P is a sugar alcohol phosphate ester and a conformational isomer of G1P. Ion exclusion chromatography, which is often applied to separate neutral and weakly dissociated compounds such as organic acids, was shown to retain *myo*-inositol, GlucA and GA only, whereas G1P, G6P and I1P co-eluted near the void volume (data not shown). However, the ion exclusion mechanisms were not sufficient to separate the structurally-similar compounds GlucA and GA (data not shown). Therefore, when analysing the biocatalytic pathway described above, GlucA frequently was separated from GA by a boronic acid affinity gel prior to ion exclusion chromatography [7, 15-17, 24]. High-performance anion exchange chromatography (HPAEC) often is applied to separate multiple anionic compounds but it is known to retain neutral compounds poorly. For example, HPAEC has been employed to either separate GlucA from GA or G1P from G6P and other sugar phosphates [25-29], but there are

no existing reports in which any of those compounds have been separated in combination with a neutral compound. In contrast to the strong anion exchange mechanism in HPAEC, WAX is more suitable for the separation of neutral and anionic compounds. Accordingly, WAX has been shown to separate *myo*-inositol from various anionic sugar phosphates, but the conformational isomers of G1P and I1P were unable to be resolved under these conditions [30]. HILIC has been established as an effective technique to separate highly polar compounds, but may be inadequate for the separation of multiple acidic compounds [31, 32]. For example, it has been shown that under HILIC conditions, *myo*-inositol was separated from various sugars but co-eluted with G6P [33]. Similarly, the constitutional isomers G1P and G6P were not resolved under HILIC conditions [34]. However, when HILIC was combined with ion-pairing reagents in the mobile phase, sugar phosphates, including G1P and G6P and neutral sugars such as sucrose and lactose were separated efficiently [35]. When HILIC was combined with WAX on a MM WAX/reversed phase (RP) column, G1P, G6P and various other sugar phosphates were resolved [36]. Therefore, we anticipated that a MM column, which features both HILIC and WAX retention mechanisms, would be most suitable for the separation of the chosen analytes since our target compounds represented a combination of neutral and multiple anionic compounds. Accordingly, we selected the silica-based GlycanPac AXH-1 column for development and optimisation of the chromatographic separation.

Table 1. List of target compounds for the method developed and their molecular masses (M), logarithmic acid dissociation constants (pK_A), and physiological charge states.

Compound	Abbreviation	M	pK _{A1} ; pK _{A2}	Physiological charge
<i>Myo</i> -inositol	Ino	180.16	12.3	0
D-glucuronic acid	GA	194.14	3.2	-1
α -D-glucose-1-phosphate	G1P	260.14	1.73; 6.21 ^a	-1
D-glucose-6-phosphate	G6P	260.14	1.86; 6.36 ^a	-1
D- <i>myo</i> -inositol-3-phosphate	I1P	260.14	1.91; 6.41 ^a	-1
D-glucaric acid	GlucA	210.14	2.83; 3.5 ^b	-2

^a pK_{A1} and pK_{A2} values were adapted from [36].

^b pK_{A2} was estimated according to [37].

Detector

We compared the suitability of multiple UHPLC detectors in the development of a method that was applicable for all of our compounds and was simple, robust and accessible with minimal technical requirements. Typically, carbohydrates, sugar phosphates and uronic acids can be separated by ion exchange chromatography like HPAEC and detected by pulsed amperometric detection (PAD) [38]. This highly sensitive detection method relies on the electrooxidation of oxidisable compounds (in principle - also reduction of reducible compounds) such as aldehyde groups of sugars in solutions with high pH. However, GlucA (which is already oxidised) will not be detected under the same conditions as the other target compounds. Furthermore, sugar acids (GA and GlucA) can be analysed by UV detection at 205 nm or 210 nm [39, 40], but this detector is unsuitable for sugar phosphates and *myo*-inositol which absorb poorly at these wavelengths. Instead, many other compounds (e.g. compounds with a keto group or conjugated

π -electron systems) absorb UV light at 205 or 210 nm which results in UV detection being highly unselective at these wavelengths [41]. While MS is suitable for the analysis of all these compounds due to its high sensitivity and selectivity, the required set-up and validation processes are lengthy and the acquisition and maintenance of the instrument is expensive. Therefore, the best detector system for our purpose was identified to be RID, which is highly economical and reliable, and is able to detect all compounds with a refractive index that is different from the mobile phase and as such, is universal. The low sensitivity of RID has been reported as one of the main drawbacks of this detection system [41]. However, the aim of enzyme pathways is to achieve high-yield and high-titre production and hence, the concentrations of the target product(s) and intermediates are often in the mM (or g L⁻¹) range and therefore within the detection range of RID. Further drawbacks of RID were its incompatibility with gradient elution, since the refractive index of the eluent changes with the mobile phase composition, and as well as its incapability to detect analytes selectively [41]. Therefore, methods based on RID require the use of isocratic elution and baseline separation of all analytes and matrix-derived substances.

3.1.2 Optimisation of the mobile phase composition

Generally, MM HILIC/WAX columns allow the sensitive tuning of the analyte's retention through variations of the organic content, pH and ionic strength of the mobile phase [42, 43]. Accordingly, the separation of our target compounds was optimised by successively changing these parameters (Fig. 2), as well as adjusting the flow rate and column temperature.

Organic modifier content

Preliminary experiments indicated that the organic modifier content had the largest effect on the retention of the target compounds while the pH and ionic strength were effective to optimise the separation of the acidic compounds. The organic modifier content plays a major role in

HILIC since the primary separation mechanism was postulated to be a partitioning of the analyte between a water layer that solvates a polar stationary phase (e.g. silica) and a bulk organic mobile phase in dependence on their polarity [31, 32, 44]. Consequently, a mobile phase with high organic modifier content causes polar analytes to partition into the aqueous layer and to be retained by the stationary phase [32]. We observed that when the aqueous buffer was maintained at 70 mM ammonium formate, pH 3.5, the most suitable retention and best separation of all acidic compounds was at 80% vol. ACN (Fig. 2A). This finding is in agreement with previous reports, where under HILIC conditions the water excess on the silica matrix was maximal at an ACN content of 80% vol. [45]. Increasing (e.g. > 80%) or decreasing (e.g. < 80%) the ACN content results in more ACN being adsorbed onto the silica surface and multilayers of ACN and H₂O being formed on the surface which consequently impairs the retention of highly polar compounds [45]. The order of elution under these conditions was 1) *myo*-inositol; 2) GA; 3) G1P; 4) G6P; 5) I1P and 6) GlucA which correlates directly with the compounds' (decreasing) pK_A values and thus their degree of ionisation and hydrophilicities (Table 1). At pH 3.5 in the aqueous phase, the sugar acids and sugar phosphates exist largely in their dissociated carboxyl and phosphate forms, resulting in electrostatic interactions between the anions and the stationary phase. The observation that all sugar phosphates were almost baseline separated, although they exhibit very similar pK_A values and thus very similar ionisation degrees, suggested that ion exchange mechanisms are highly effective alongside HILIC. Similarly, it has been reported previously that the separation efficiency of G1P and G6P under HILIC/WAX was the highest at 80% vol. ACN [36]. *Myo*-inositol with a pK_A of 12.3 remains neutral under these chromatographic conditions and therefore, its retention is attributed solely to HILIC, demonstrating further that the column effectively supports both ion exchange and HILIC mechanisms at 80% vol. ACN.

When the ACN content was reduced below 80% vol. the retention of all compounds decreased and this effect was most notable for the acidic compounds (Fig. 2A). Similar results were

reported by others and were attributed to a decrease of water solvating the stationary phase (as was also described above), and which results in a loss of the analytes' interaction with the stationary phase [31, 46]. When the organic modifier content was increased to 90% vol. ACN, the RT of *myo*-inositol increased to 7 min. A similar chromatographic behaviour of *myo*-inositol was observed in a previous HILIC study [33]. This result indicated that in MM HILIC/WAX, HILIC prevails over ion exchange mechanisms only at very high solvents. Likewise, the acidic compounds were retained more at 90% vol. ACN than at lower ACN proportions but co-eluted (data not shown), which supports the assumption that in MM, ion exchange becomes ineffective in presence of a high organic solvent concentration.

Eluent pH

The pH conditions were optimised by maintaining the ACN content at 80 % vol. and the ammonium formate concentration at 70 mM, while changing the pH of the aqueous phase gradually from 3 to 4.5 (Fig. 2B). The retention of neutral *myo*-inositol and weak acidic GA was barely subjective to pH changes, while the stronger acidic sugar phosphates and GlucA were retained less when the pH dropped from 4.5 to 3.5. This loss in retention might be attributed to a suppression of the compounds' dissociation and, accordingly, a decreased electrostatic interaction with the stationary phase. Interestingly, when the eluent pH was further reduced to 3.0, all acidic compounds were retained more than at pH 3.5 analytes (Fig. 2B). Although the exact column chemistry of the GlycanPac AXH-1 is not disclosed, most HILIC and WAX columns contain secondary amine groups [43, 44, 47]. Thus, the observed increase of the analytes' retention at pH 3 may be explained with an increase in the positive charge of the stationary phase and a higher anion exchange capacity [48]. However, the increased peak tailing at pH 3 is indicative for undesired, secondary interactions between the compounds and the stationary phase [43]. In conclusion, the highest separation efficiency was achieved at pH

3.5 with GA and G1P being baseline separated and I1P, GlucA and G1P and G6P slightly co-eluting.

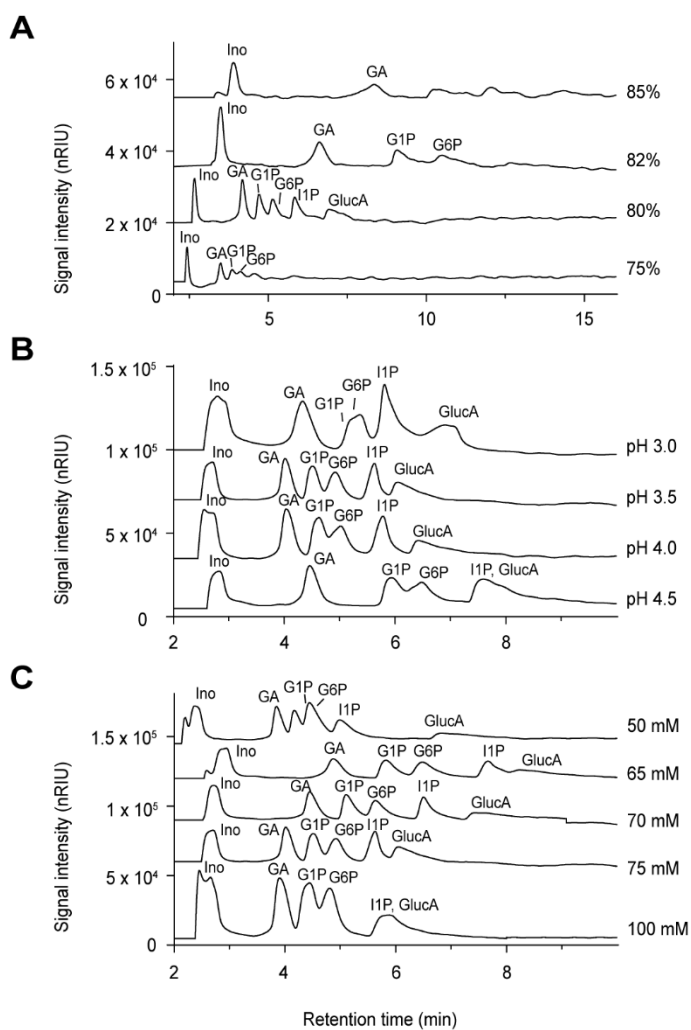


Figure 2. Effect of acetonitrile (ACN) content (**A**), eluent pH (**B**) and ammonium formate concentration in the mobile phase (**C**) on the chromatographic separation of the compounds listed in Table 1. Invariable conditions: **A**, ammonium formate concentration of 70 mM, pH 3.5; **B**, ACN:ammonium formate (75 mM, 80:20, v/v); **C**, ACN:ammonium formate (80:20, v/v), the pH of the aqueous phase was adjusted to 3.5. The column temperature was 40 °C and the flow rate was 0.3 ml min⁻¹. nRIU, nano Refractive Index Units.

Counterion concentration

The effect of the counterion concentration on the chromatographic separation was investigated by maintaining the aqueous buffer at pH 3.5 and the organic modifier content at 80% vol. ACN, while the ammonium formate concentration was altered gradually between 100 and 50 mM (Fig. 2C). As expected, neutral *myo*-inositol was barely affected by the changes in the counterion concentration whereas the sugar acids and the sugar phosphates showed the lowest retention at 100 mM and 50 mM and the highest retention at 65 mM ammonium formate. Counterions (here, the formate ion) compete with the analytes to bind the stationary phase and therefore, a higher counterion concentration causes the compounds to elute more rapidly [49]. The reduced retention observed with all compounds at 50 mM ammonium formate may be attributed to a thinning of the water layer on the stationary phase due to a decreased ion concentration in the aqueous phase [32, 44]. Generally, increased peak tailing occurred with the acidic compounds, most notably with GlucA at lower counterion concentrations, while peaks were more symmetrical at higher concentration (Fig. 2C). Peak tailing often results from secondary interactions of the analytes with the stationary phase which might be prevented by competing interactions of the counterion. However, as peaks co-elute at such high counterion concentrations, the best separation of all compounds was achieved using an ammonium formate concentration of 70 mM (Fig. 2C). As a result, this counterion concentration was applied in all subsequent analyses.

Temperature and flow rate

The early stages of the method development were performed at 40 °C but we found that reducing the column temperature to 25 °C resulted in more symmetrical peaks. The maximal flow rate for optimal resolution and peak shape was found to be 0.3 ml min⁻¹.

3.1.3 Optimisation of the sample solvent composition

We investigated the effects of the sample solvent on the compounds' RT, peak widths and signal intensities further to improve the separation and peak shapes observed with the optimised mobile phase (Fig. 3). Most compounds were retained similarly (indicated by their RT relative to the RT achieved for each compound at 0% vol. ACN in the sample solvent) and eluted with consistent peak widths independent of the ACN content in the sample solvent (Fig. 3A and B). In contrast, the RT of both GlucA and IIP decreased while the peak width of GlucA increased significantly upon raising the content of ACN which eventually caused their peaks to overlap partially and completely at 60 and 75% vol. ACN, respectively. These findings contradict previous studies where under HILIC, adding organic content to the sample solvent > 50% vol. improved the solute's partitioning in the aqueous layer, prevented peak broadening, column overloading and resulted in improved retention [49]. The effects observed here appeared to be compound-specific and were dependent upon the compounds' polarity. For example, GlucA - whose retention and peak width were affected the most from high ACN content, precipitated visibly at concentrations above 20 mM at > 30% vol. ACN in the sample. This effect could be minimised by adding TFA to a final concentration of 0.02% vol. which suppresses the dissociation of the carboxyl groups and thereby increases the analyte's hydrophobicity. Adding TFA to the sample solvent resulted in a decreased retention of all compounds and caused a slight increase of the signal intensities of mainly the acidic compounds. Therefore, it appeared that by increasing the compounds' hydrophobicity with TFA, the blending of the compounds into the organic-rich mobile phase is facilitated at the expense of their dissociation and electrostatic interactions with the column. Similarly, upon increasing the ACN content in the sample from 60 to 75% vol., the signal intensities of all compounds increased drastically which may be attributed to improved blending into the mobile phase due to the initial higher ACN content (Fig. 3C). Based on these results, we decided to set the sample ACN content at 30% vol., because these conditions are the best compromise between RT, peak width and signal

intensities for all compounds. However, to ensure that GlucA remains soluble at high concentrations and 30% vol. ACN, we added TFA to all samples to a final concentration of 0.02% vol. It is noteworthy that it was crucial to keep the samples at 4 °C once they contained 30% vol. ACN, as the organic solvent evaporates over time in the small sample volumes used (12 μ l for a 4 μ l injection) causing apparent increases in analyte concentrations.

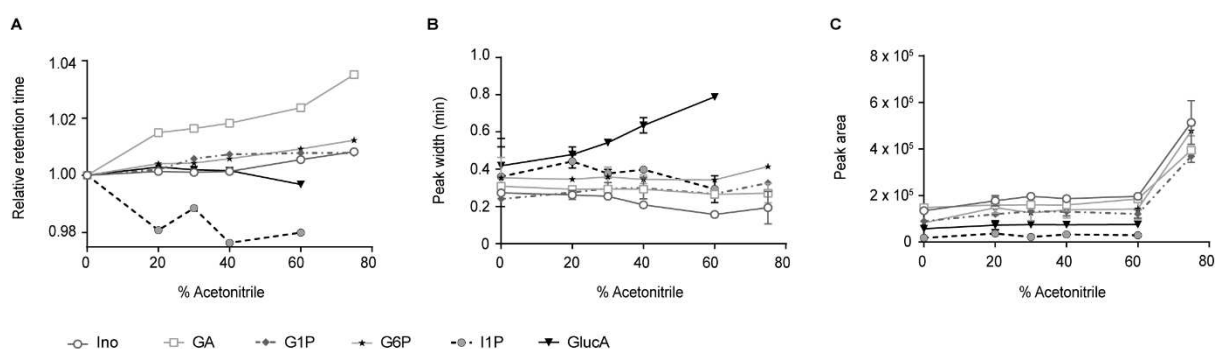


Figure 3. Effect of the acetonitrile (ACN) content in the sample solvent on relative retention time (A), the peak widths at 50% peak height (B) and the signal intensities (C) of the compounds described in Table 1. The mobile phase was ACN:ammonium formate (70 mM, pH 3.5, 80:20, v/v), the flowrate was 0.3 ml min⁻¹ and the column temperature was set to 40 °C. A mixture of standard compounds (10 mM) was injected in triplicate. Data shown as mean \pm standard deviation.

Next, we investigated the effect of different buffers that are commonly used for the preparation of bacterial cell lysates on the separation and analysis of the analytes. Standard mixtures of all compounds were prepared in sample solvents containing 30% vol. ACN and either HEPES, Tris-HCL, phosphate buffer (all at 50 mM, pH 8) or 50 mM MES (pH 6). We found that the buffer composition had marginal effect on the compounds' RT, peak widths, and signal responses when compared to standard solutions prepared in 50 mM ammonium formate.

However, due to the high concentrations of the buffers used, HEPES, Tris and phosphate eluted in very broad peaks. As a result, HEPES and Tris completely and partially co-eluted with *myo*-inositol, respectively (Fig. 4), while phosphate partially co-eluted with G1P. Only MES buffer did not interfere with the analysis of our compounds but MES may be unsuitable for enzymes that operate optimally at pH 7 or 8, as they buffer only around pH 6. Therefore, the choice of buffer depends on the enzymes that are being tested and may compromise the analysis of *myo*-inositol.

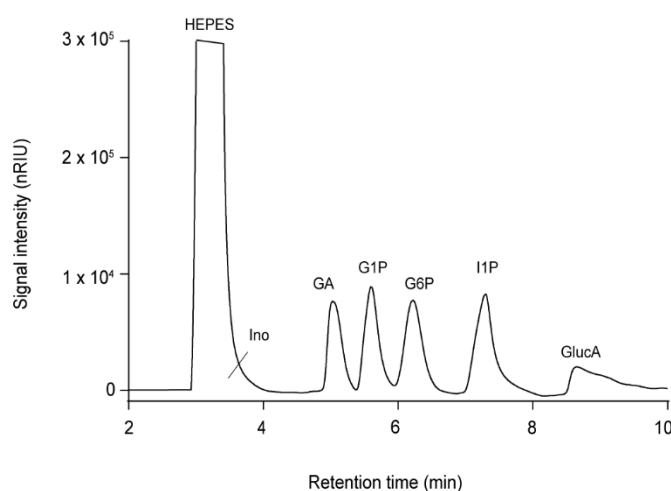


Figure 4. UHPLC-RID chromatogram of a 5 mM standard mixture of all compounds described in Table 1. Sample solvent was 50 mM HEPES, pH 8 and 30% vol. ACN. The mobile phase was ACN:ammonium formate (70 mM, pH 3.5, 80:20, v/v) with a flowrate of 0.3 ml min⁻¹ and column temperature of 25°C. nRIU, nano Refractive Index Units.

3.2 Method validation

The quantitation of all analytes was validated in 50 mM HEPES buffer pH 8, as most synthetic enzyme pathways for the production of bio-based GlucA would be operated at pH 7 or 8. The quantitation of *myo*-inositol was validated in 50 mM MES buffer, pH 6 to avoid interference from HEPES. Linearity, sensitivity and repeatability of the method are summarised in Table 2,

and the accuracy and precision of the method is described in Table 3. The calibration curves (five points from 5 to 40 mM) for each metabolite showed a linear correlation ($r^2 > 0.99$) over the concentration ranges tested. With an injection volume of 4 μl , the LODs and LOQs of all compounds were in the lower millimolar range (e.g. LODs ranged from 1.5 to 2 mM and LOQs ranged from 5 to 7 mM). Expressed in g L^{-1} , the LODs were between 0.3 and 0.4 g L^{-1} while the LOQs were between 0.9 and 1.5 g L^{-1} . The LOD and LOQs reported here are comparable to studies published previously using refractive index detection [50-52]. However, compared to studies using HILIC-UV-detection, the LOD and LOQ of GlucA were about 32 and 75 times higher by RID detection, respectively ([39], and about 7 times higher than those obtained using LC-MS [40]. The RSDs of the RTs obtained in six repeated measurements of standard solutions with analyte concentrations at 10 mM were less than 0.2%. In these measurements, the RSDs of the peak areas were below 8% for all compounds except for GlucA, which was 11.4%. This high variation for GlucA detection as compared to all other compounds is attributed to the suboptimal peak shape of GlucA under the conditions tested. In contrast, the RSDs of the peak areas obtained at higher concentrations (e.g. 30 mM) showed variations below 10% for all compounds. These results demonstrate a satisfactory repeatability of RTs and peak areas.

Table 2. Validation data for the developed UHPLC-RID method for the quantitation of all compounds described in Table 1. r^2 , coefficient of determination; y = peak area; x = analyte concentration (mM).

Compound	Linearity		Sensitivity					Repeatability			
	Calibration equation	r^2	Calibrated range (mM)	LOD		LOQ		Retention time		Peak area RSD (%)	
				mM	$\mu\text{g L}^{-1}$	mM	$\mu\text{g L}^{-1}$	Mean (min)	RSD (%)	10 mM ^a	30 mM ^a
Ino	$y = 31869x + 59933$	0.991	5 - 40	1.5	0.3	5.0	0.9	2.75	0.04	1.84	5.7
GA	$y = 27050x + 9351$	0.996	5 - 40	1.5	0.3	5.0	1.0	5.04	0.04	3.1	1.7
G1P	$y = 30183x + 22723$	0.997	5 - 40	1.5	0.4	5.0	1.3	5.67	0.06	3.2	3.0
G6P	$y = 32431x + 35131$	0.998	5 - 40	1.5	0.3	5.0	1.0	6.32	0.14	3.9	0.5
I1P	$y = 32785x + 21426$	0.996	5 - 40	1.5	0.4	5.0	1.3	7.55	0.02	7.9	1.9
GlucA	$y = 26674x + 11017$	0.993	5 - 40	2.0	0.4	7.0	1.5	8.26	0.19	11.4	9.5

^a Concentration of the analytes at which the repeatability was tested.

The accuracy and precision of the method were evaluated by the three-fold repeated preparation and measurement of spiked empty matrices. Recoveries of the spiking experiments at 10 mM ranged from 94.2 to 103.4% with RSDs being between 0.4 and 12.4%, and recoveries of the spiking experiments at 20 mM ranged from 98.2 to 106.8% with RSDs being between 0.5 and 6.1 (Table 3). These results indicated a very high degree of accuracy and a good degree of precision and confirm that the method developed was suitable for the analysis of all compounds in the bacterial cell lysate as presented in Table 1. It should be noted that the recovery of I1P was performed only at 10 mM due to low availability and high costs of this substrate.

Table 3. Accuracy (recovery (%)) and precision (RSD (%)) are shown as the average from three independent replicates of all compounds spiked in *E. coli* cell lysates.

Compound	Accuracy (Recovery (%))		Precision (RSD (%))	
	10 mM	20 mM	10 mM	20 mM
Ino	96.2	104.0	2.6	0.5
GA	94.2	106.4	0.4	4.0
G1P	96.2	101.5	12.4	5.7
G6P	103.4	106.8	5.8	6.1
I1P	102.3	ND ^a	1.0	ND ^a
GlucA	101.3	98.2	5.0	1.8

^a ND; not determined.

3.3 Application of the method for monitoring cell-free GlucA production

The method developed was applied to the final reaction (catalysed by Udh) of the GlucA pathway (Fig. 1) to validate its ability to monitor enzyme reactions within GlucA production pathways. Crude enzyme lysates from *E. coli* cells that produced Udh from *Fulvimarina pelagi* heterologously were incubated with 40 mM GA and NAD⁺ at 40 °C. At the time of sample preparation, Udh had converted 40% of GA to GlucA, and after 2h of incubation 96.7% of GA was converted to GlucA (Fig. 5B, Table 4). In contrast, the control samples (lysates from *E. coli* cells without Udh) did not contain GlucA at detectable concentrations (Fig. 5A, Table 4). The maximum concentration of GlucA produced was 44 mM which translates to a titre of 9.2 g L⁻¹ after 2h. The fast conversion rate of Udh is in agreement with a previous study [53]. The presence of GlucA after only a few minutes may be due to the high reaction rate of Udh, which was probably initiated during sample preparation and before the samples were boiled. Boiling the samples resulted in the degradation of heat-labile NAD⁺ but neither NAD⁺ nor its degradation product interfere with the analytes' detection as they elute after GlucA (Fig. 5).

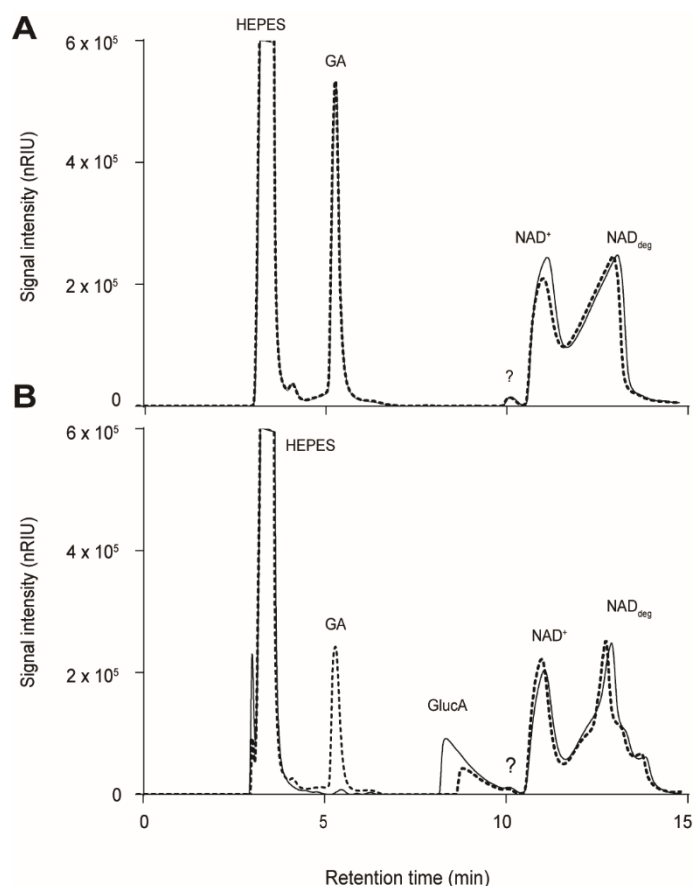


Figure 5. UHPLC-RID chromatograms of lysates from **(A)** *E. coli* cells and **(B)** *E. coli* cells expressing Udh from *Fulvimarina pelagi*. Samples were incubated with 40 mM GA at 40 °C for 0h (dashed line) and 2h (straight line) to mediate the conversion of GA to GlucA. nRIU, nano Refractive Index Units.

Surprisingly, when the samples were not boiled prior to UHPLC-RID analysis, GlucA was not detected while after boiling the samples the peak associated with GlucA clearly appeared in the chromatogram. This observation may indicate that Udh catalysed the conversion of GA to the glucaric acid 1,4-lactone instead of the desired open form of GlucA. It was anticipated that the GlucA lactone would exhibit a different chromatographic behaviour to GlucA and may have resulted in the co-elution with HEPES or other compounds present in the chromatogram, e.g. NAD^+ . Evidence for the assumption of the lactone formation is provided by a previous study where Udh from *Agrobacterium tumefaciens* converted galacturonic acid to the galactaric acid 1,4-lactone and not the open galactaric acid form [54]. The Udh in that study and the Udh employed here are both promiscuous towards galacturonic acid and GA, suggesting that upon

the reaction with GA, Udh may form the glucaric acid 1,4 lactone [53]. Accordingly, boiling the samples may result in the hydrolysis of the lactone to the open GlucA form, an explanation that has been suggested elsewhere [55].

Table 4. Titre (in mM) and conversion rate of GlucA produced from 40 mM GA in *E. coli* cell lysates that do not produce Udh (negative control) and *E. coli* cells that produce Udh heterologously, tested at various time points and determined by UHPLC-RID. ND, not detected.

Samples	Incubation (h)	GlucA (mM)	Conversion rate (%)
<i>E. coli</i>	0	ND	0
	1	ND	0
	2	ND	0
Udh in <i>E. coli</i>	0	17.8	39.2
	1	36.7	84.9
	2	44.0	96.7

4 Conclusions

To the best of our knowledge, this is the first method reported for the simultaneous analysis of GlucA and its five intermediate compounds (G1P, G6P, I1P, *myo*-inositol and GA) that are generated during the enzymatic synthesis of GlucA from glucose, or potentially could be generated during the enzymatic synthesis of GlucA from amylose/maltodextrin which has not been attempted here (Fig. 1). The UHPLC-RID method developed is simple, rapid, involves minimal sample preparation, allows quantitation with high accuracy and precision and enough sensitivity for the concentrations obtained in enzymatic reactions. The method was applied to analyse the enzymatic conversion of GA to GlucA in Udh-containing *E. coli* cell-free lysates expressing the GlucA pathway. The results showed that GlucA was produced with near 100% yield and a titre of 9.2 g L⁻¹ GlucA. There was no interference from endogenous *E. coli* compounds in the chromatogram and therefore, the method is expected to be suitable for the analysis of the GlucA pathway generated in similar cell lysates, e.g. from yeast. Consequently,

this technology provides a novel tool for the assessment and optimisation of biocatalytic GlucA production from renewable resources in metabolically-engineered microorganisms and cell-free systems.

One limitation of the method is the co-elution of *myo*-inositol with HEPES and Tris, which are buffers commonly used at high concentrations in the preparation of cell lysates to maintain a physiological pH. Alternative buffers can be used, such as MES, which does not interfere with the *myo*-inositol detection, but as MES buffers around pH 6 it may not be suitable for certain enzyme reactions. Gradient elution may be suitable for improvement of the separation of *myo*-inositol and HEPES. Another drawback is the suboptimal peak shape of GlucA under the conditions tested, which - although it does not greatly impact on the accuracy and precision - might be improvable through gradient elution. RID is incompatible with gradient elution as it is subject to changes in the mobile phase. However, since the method is composed of volatile eluent constituents and is operated at a low flow rate, it is well-suited for online coupling to MS or tandem MS which is applicable to gradient elution. Further, MS can be applied to increase the method's sensitivity if required.

This study provides detailed insights into the effects of the mobile phase and sample solvent on the retention mechanisms present in mixed mode HILIC/WAX and serves as a basis for further methodological developments. The predominant type of chromatographic mechanism under MM column chemistry is largely influenced by the organic modifier content in the mobile phase, while pH and ionic strength contribute to compound-selective interactions that can be tailored for high resolution. At 80% vol. ACN in the mobile phase, superimposition of HILIC and WAX mechanisms were most effective in retaining the neutral compound while separating structurally highly similar organic anions. Further, this study shows that the sample solvent had opposite effects on the compound's behaviour under MM HILIC/WAX than those previously

reported for HILIC mode alone, especially with regards to RT and peak shape. These effects largely depended on the compound's solubility and for highly polar compounds such as GlucA, it was beneficial to decrease the amount of organic content in the sample solvent or to add TFA as this compound facilitates the compound's solubility and appears to increase its blending into the organic-rich mobile phase.

The capability of such MM multimodality to separate several acidic compounds without compromising HILIC selectivity for neutral compounds makes this MM approach an attractive tool for the development of novel analysis protocols for multiple sugar derivatives in complex matrices which are suitable for detection by both RI or MS.

Acknowledgements

KP is 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) and the Australian Research Council (CE140100003). We thank Dr Morten Andersen for technical advice and supplying of UHPLC column.

References

- [1] T. Werpy, G. Petersen, Top value added chemicals from biomass volume I-results of screening for potential candidates from sugars and synthesis gas, Department of Energy, Washington, DC, 2004.
- [2] R.S. Koefod, Corrosion-inhibiting deicer composition, Google Patents, 2010.
- [3] T.R. Boussie, G.M. Diamond, E. Dias, V. Murphy, Synthesis of Adipic Acid Starting from Renewable Raw Materials, Chemicals and Fuels from Bio-Based Building Blocks, Wiley-VCH Verlag GmbH & Co. KGaA2016, pp. 151-172.
- [4] D.E. Kiely, L. Chen, T.H. Lin, Hydroxylated nylons based on unprotected esterified D-glucaric acid by simple condensation reactions, Journal of the American Chemical Society, 116 (1994) 571-578.
- [5] M. Hanausek, Z. Walaszek, T.J. Slaga, Detoxifying Cancer Causing Agents to Prevent Cancer, Integrative Cancer Therapies, 2 (2003) 139-144.
- [6] M. Fittkau, W. Voigt, H.-J. Holzhausen, H.-J. Schmoll, Saccharic acid 1.4-lactone protects against CPT-11-induced mucosa damage in rats, Journal of Cancer Research and Clinical Oncology, 130 (2004) 388-394.
- [7] T.S. Moon, S.H. Yoon, A.M. Lanza, J.D. Roy-Mayhew, K.L. Prather, Production of glucaric acid from a synthetic pathway in recombinant *Escherichia coli*, Applied and Environmental Microbiology, 75 (2009) 589-595.
- [8] T.S. Moon, J.E. Dueber, E. Shiue, K.L.J. Prather, Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*, Metabolic Engineering, 12 (2010) 298-305.
- [9] S. Raman, J.K. Rogers, N.D. Taylor, G.M. Church, Evolution-guided optimization of biosynthetic pathways, Proceedings of the National Academy of Sciences, 111 (2014) 17803-17808.
- [10] I.M.B. Reizman, A.R. Stenger, C.R. Reisch, A. Gupta, N.C. Connors, K.L.J. Prather, Improvement of glucaric acid production in *E. coli* via dynamic control of metabolic fluxes, Metabolic Engineering Communications, 2 (2015) 109-116.
- [11] Y.-N. Qu, H.-J. Yan, Q. Guo, J.-L. Li, Y.-C. Ruan, X.-Z. Yue, W.-X. Zheng, T.-W. Tan, L.-H. Fan, Biosynthesis of D-glucaric acid from sucrose with routed carbon distribution in metabolically engineered *Escherichia coli*, Metabolic Engineering, 47 (2018) 393-400.
- [12] E. Shiue, K.L. Prather, Improving D-glucaric acid production from *myo*-inositol in *E. coli* by increasing MIOX stability and *myo*-inositol transport, Metabolic Engineering, 22 (2014) 22-31.
- [13] A. Gupta, M.A. Hicks, S.P. Manchester, K.L. Prather, Porting the synthetic D-glucaric acid pathway from *Escherichia coli* to *Saccharomyces cerevisiae*, Biotechnology Journal, 11 (2016) 1201-1208.

- [14] N. Chen, J. Wang, Y. Zhao, Y. Deng, Metabolic engineering of *Saccharomyces cerevisiae* for efficient production of glucaric acid at high titer, *Microbial Cell Factories*, 17 (2018) 67.
- [15] Y. Liu, X. Gong, C. Wang, G. Du, J. Chen, Z. Kang, Production of glucaric acid from *myo*-inositol in engineered *Pichia pastoris*, *Enzyme and microbial technology*, 91 (2016) 8-16.
- [16] C. You, T. Shi, Y. Li, P. Han, X. Zhou, Y.P. Zhang, An *in vitro* synthetic biology platform for the industrial biomanufacturing of *myo*-inositol from starch, *Biotechnology and Bioengineering*, (2017).
- [17] T. Fujisawa, S. Fujinaga, H. Atomi, An *in vitro* enzyme system for the production of *myo*-inositol from starch, *Applied and Environmental Microbiology*, (2017).
- [18] C. Petzold, L.J. Chan, M. Nhan, P. Adams, Analytics for metabolic engineering, *Frontiers in Bioengineering and Biotechnology*, 3 (2015).
- [19] J.R. Swartz, Transforming biochemical engineering with cell-free biology, *AIChE Journal*, 58 (2012) 5-13.
- [20] Q.-T. Nguyen, M.E. Merlo, M.H. Medema, A. Jankevics, R. Breitling, E. Takano, Metabolomics methods for the synthetic biology of secondary metabolism, *FEBS Letters*, 586 (2012) 2177-2183.
- [21] C. Hold, S. Billerbeck, S. Panke, Forward design of a complex enzyme cascade reaction, *Nature Communications*, 7 (2016) 12971.
- [22] R.A. Sheldon, P.C. Pereira, Biocatalysis engineering: the big picture, *Chemical Society Reviews*, (2017).
- [23] I.H.T. Guideline, Validation of analytical procedures: text and methodology Q2 (R1), International Conference on Harmonization, Geneva, Switzerland, 2005, pp. 11-12.
- [24] R. Poon, D.C. Villeneuve, I. Chu, R. Kinach, HPLC determination of D-glucaric acid in human urine, *Journal of Analytical Toxicology*, 17 (1993) 146-150.
- [25] C. Corradini, D. Corradini, C.G. Huber, G.K. Bonn, High-performance anion-exchange chromatography of carbohydrates using a new resin and pulsed amperometric detection, *Chromatographia*, 41 (1995) 511-515.
- [26] E. Marcellin, L.K. Nielsen, P. Abeydeera, J.O. Krömer, Quantitative analysis of intracellular sugar phosphates and sugar nucleotides in encapsulated streptococci using HPAEC-PAD, *Biotechnology Journal*, 4 (2009) 58-63.
- [27] D.B. Chu, K. Klavins, G. Koellensperger, S. Hann, Speciation analysis of sugar phosphates via anion exchange chromatography combined with inductively coupled plasma dynamic reaction cell mass spectrometry - optimization for the analysis of yeast cell extracts, *Journal of Analytical Atomic Spectrometry*, 29 (2014) 915-925.

- [28] Y. Sekiguchi, N. Mitsuhashi, T. Kokaji, H. Miyakoda, T. Mimura, Development of a comprehensive analytical method for phosphate metabolites in plants by ion chromatography coupled with tandem mass spectrometry, *Journal of Chromatography A*, 1085 (2005) 131-136.
- [29] E. Groussac, M. Ortiz, J. François, Improved protocols for quantitative determination of metabolites from biological samples using high performance ionic-exchange chromatography with conductimetric and pulsed amperometric detection, *Enzyme and Microbial Technology*, 26 (2000) 715-723.
- [30] X. Liu, P.W. Villalta, S.J. Sturla, Simultaneous determination of inositol and inositol phosphates in complex biological matrices: quantitative ion-exchange chromatography/tandem mass spectrometry, *Rapid Communications in Mass Spectrometry*, 23 (2009) 705-712.
- [31] D.V. McCalley, Study of the selectivity, retention mechanisms and performance of alternative silica-based stationary phases for separation of ionised solutes in hydrophilic interaction chromatography, *Journal of Chromatography A*, 1217 (2010) 3408-3417.
- [32] B. Buszewski, S. Noga, Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique, *Analytical and Bioanalytical Chemistry*, 402 (2012) 231-247.
- [33] J. Pazourek, Fast separation and determination of free *myo*-inositol by hydrophilic liquid chromatography, *Carbohydrate Research*, 391 (2014) 55-60.
- [34] B. Preinerstorfer, S. Schiesel, M. Lämmerhofer, W. Lindner, Metabolic profiling of intracellular metabolites in fermentation broths from β -lactam antibiotics production by liquid chromatography–tandem mass spectrometry methods, *Journal of Chromatography A*, 1217 (2010) 312-328.
- [35] C. Mathon, G.A. Barding, C.K. Larive, Separation of ten phosphorylated mono- and disaccharides using HILIC and ion-pairing interactions, *Analytica Chimica Acta*, 972 (2017) 102-110.
- [36] H. Hinterwirth, M. Lämmerhofer, B. Preinerstorfer, A. Gargano, R. Reischl, W. Bicker, O. Trapp, L. Brecker, W. Lindner, Selectivity issues in targeted metabolomics: Separation of phosphorylated carbohydrate isomers by mixed-mode hydrophilic interaction/weak anion exchange chromatography, *Journal of Separation Science*, 33 (2010) 3273-3282.
- [37] W. Yuan, R.-H. Ding, H. Ge, P.-L. Zhu, S.-S. Ma, B. Zhang, X.-M. Song, Solid-phase extraction of d-glucaric acid from aqueous solution, *Separation and Purification Technology*, 175 (2017) 352-357.
- [38] T.R. Cataldi, C. Campa, G.E. De Benedetto, Carbohydrate analysis by high-performance anion-exchange chromatography with pulsed amperometric detection: the potential is still growing, *Fresenius' Journal of Analytical Chemistry*, 368 (2000) 739-758.
- [39] B. Xie, Y. Liu, H. Zou, Y. Son, H. Wang, H. Wang, J. Shao, Determination of d-glucaric acid and/or d-glucaro-1,4-lactone in different apple varieties through hydrophilic interaction chromatography, *Food Chemistry*, 203 (2016) 1-7.

- [40] J.L. Perez, G.K. Jayaprakasha, K.S. Yoo, B.S. Patil, Development of a method for the quantification of d-glucaric acid in different varieties of grapefruits by high-performance liquid chromatography and mass spectra, *Journal of Chromatography A*, 1190 (2008) 394-397.
- [41] M. Swartz, HPLC detectors: a brief review, *Journal of Liquid Chromatography & Related Technologies*, 33 (2010) 1130-1150.
- [42] E. Apfelthaler, W. Bicker, M. Lämmerhofer, M. Sulyok, R. Krska, W. Lindner, R. Schuhmacher, Retention pattern profiling of fungal metabolites on mixed-mode reversed-phase/weak anion exchange stationary phases in comparison to reversed-phase and weak anion exchange separation materials by liquid chromatography–electrospray ionisation–tandem mass spectrometry, *Journal of Chromatography A*, 1191 (2008) 171-181.
- [43] K. Zhang, X. Liu, Mixed-mode chromatography in pharmaceutical and biopharmaceutical applications, *Journal of Pharmaceutical and Biomedical Analysis*, 128 (2016) 73-88.
- [44] N.T.H. Bui, J.J. Verhage, K. Irgum, Tris(hydroxymethyl)aminomethane-functionalized silica particles and their application for hydrophilic interaction chromatography, *Journal of Separation Science*, 33 (2010) 2965-2976.
- [45] F. Gritti, A. dos Santos Pereira, P. Sandra, G. Guiochon, Efficiency of the same neat silica column in hydrophilic interaction chromatography and per aqueous liquid chromatography, *Journal of Chromatography A*, 1217 (2010) 683-688.
- [46] B.A. Olsen, Hydrophilic interaction chromatography using amino and silica columns for the determination of polar pharmaceuticals and impurities, *Journal of Chromatography A*, 913 (2001) 113-122.
- [47] W. Kopaciewicz, M.A. Rounds, F.E. Regnier, Stationary phase contributions to retention in high-performance anion-exchange protein chromatography: ligand density and mixed mode effects, *Journal of Chromatography A*, 318 (1985) 157-172.
- [48] G. Shi, J.-t. Wu, Y. Li, R. Geleziunas, K. Gallagher, T. Emm, T. Olah, S. Unger, Novel direct detection method for quantitative determination of intracellular nucleoside triphosphates using weak anion exchange liquid chromatography/tandem mass spectrometry, *Rapid Communications in Mass Spectrometry*, 16 (2002) 1092-1099.
- [49] P. Jandera, Stationary and mobile phases in hydrophilic interaction chromatography: a review, *Analytica Chimica Acta*, 692 (2011) 1-25.
- [50] U. Başaran, M. Akkik, H. Mut, E. Gülümser, M. Çopur Doğrusöz, S. Koçoğlu, High-Performance Liquid Chromatography with Refractive Index Detection for the Determination of Inulin in Chicory Roots, *Analytical Letters*, 51 (2018) 83-95.
- [51] A. Trani, G. Gambacorta, P. Loizzo, A. Cassone, C. Fasciano, A.V. Zambrini, M. Faccia, Comparison of HPLC-RI, LC/MS-MS and enzymatic assays for the analysis of residual lactose in lactose-free milk, *Food Chemistry*, 233 (2017) 385-390.
- [52] J.L. Chávez-Servín, A.I. Castellote, M.C. López-Sabater, Analysis of mono- and disaccharides in milk-based formulae by high-performance liquid chromatography with refractive index detection, *Journal of Chromatography A*, 1043 (2004) 211-215.

- [53] A. Pick, J. Schmid, V. Sieber, Characterization of uronate dehydrogenases catalysing the initial step in an oxidative pathway, *Microbial biotechnology*, 8 (2015) 633-643.
- [54] H. Boer, H. Maaheimo, A. Koivula, M. Penttilä, P. Richard, Identification in *Agrobacterium tumefaciens* of the D-galacturonic acid dehydrogenase gene, *Applied Microbiology and Biotechnology*, 86 (2010) 901-909.
- [55] C. Marsh, Metabolism of d-glucuronolactone in mammalian systems. Identification of d-glucaric acid as a normal constituent of urine, *Biochemical Journal*, 86 (1963) 77.

Chapter 4:

A novel framework for the cell-free enzymatic production of glucaric acid

4.1 Introduction

The production of GlucA has been performed typically by chemical and metabolic synthesis. However, these processes are either environmental unfriendly or too inefficient for commercial scale production of GlucA. Therefore, it was investigated whether the production of bio-based GlucA achieved with cell-free biocatalysis could be a viable alternative. Accordingly, rational pathway design and enzyme selection, cofactor regeneration, pathway immobilisation and metabolic flux analysis were employed and incorporated into the cell-free system and tested for efficient and sustainable GlucA production.

The results of this work have been submitted to the high-impact journal *Green Chemistry* for publication as a peer-reviewed paper.

4.2 Contribution to manuscript 3

The concept of this publication was developed in partnership with my supervisors Anwar Sunna and Andrew Care and with support from Peter Bergquist. They were also involved in designing experiments and troubleshooting. I completed all of the experimental work and data analysis. The manuscript was prepared by me and was reviewed by all co-authors. The contributions of each author are detailed in Table 4.1

Table 4.1 Author contribution summary for Publication 2.

	Kerstin Petroll	Andrew Care	Peter Bergquist	Anwar Sunna
Conception	•	•	•	•
Experiment Design	•	•	•	•
Data Collection	•			
Data Analysis	•			
Manuscript	•	•	•	•

Abstract

Glucaric acid (GlucA) is a valuable glucose-derived chemical with promising applications as biodegradable and biocompatible chemical in the manufacturing of plastics, detergents and drugs. In recent years, there has been a considerable focus on producing GlucA metabolically from renewable materials such as glucose, sucrose and *myo*-inositol. However, these *in vivo* GlucA production processes generally lack efficiency due to toxicity problems, metabolic repression and suboptimal enzyme ratios. Synthetic biology at the interface of cell-free biocatalysis has been proposed as a viable approach to overcome many of these limitations. However, cell-free biocatalysis faces its own limitations for industrial applications due to high enzyme costs and cofactor consumption. In this work, we constructed the first cell-free GlucA pathway and demonstrated a framework to overcome limitations of cell-free biocatalysis by immobilisation of the enzyme pathway and integration of a cofactor regeneration system. The cell-free production of GlucA was achieved from glucose-1-phosphate at a concentration of 5.7 ± 1.1 mM (1.2 ± 1.1 g/L) with a yield of 14.3% and a total reaction time of only 8h. Enzymes were recycled for several cycles without substantial loss of activity. Integrating an immobilised NADH regeneration system maintained the GlucA production at 70% of the initial production but required four times less NAD^+ than without the regeneration system.

1 Introduction

Glucaric acid (GlucA or saccharic acid) is a glucose-derived emergent platform chemical with the potential for applications such as a biodegradable and biocompatible chelating agent for the degradation of organic contaminants¹, removal of heavy metals in soil², calcium sequestration for phosphate-free detergent component³, as an adhesive⁴, anti-corrosive⁵, anti-plasticizer and plastic-strengtheners⁶, as a starting material for biodegradable polymers/poly(amide)s for example, hydroxylated nylon^{7,8} and to replace the petroleum-derived adipic acid in nylon 6.6⁹. In addition, GlucA could be manufactured as a cancer-preventive agent¹⁰ or as preventive agent for diarrhoea that often arises from the treatment with anticancer drugs such as CPT-11.¹¹

Increasing environmental concerns and resource shortages have put pressure on industry, policy makers and stock-markets to invest in ‘greener’ synthesis routes for biocommodities. Accordingly, the bio-based production of GlucA has become an attractive target. So far, three companies have piloted the production of GlucA from glucose by platinum-catalysed synthesis (Rennovia and Johnson Matthey Process Technologies^{12,9}, by nitric acid oxidation (Rivertop renewables¹³, <http://www.marketwired.com/press-release/rivertop-renewables-and-dti-announce-commercial-production-of-glucaric-acid-2061527.htm>, last checked on 09/08.2018) or by metabolically-engineered *Escherichia coli* (Kalion Inc. (<http://www.kalioninc.com/portfolio.html>)¹⁴. However, major commercial success has not been achieved and plummeting oil-prices, lack of financial incentives and the risks of changing technologies has resulted in a lack of investment and the demise of Rennovia despite its near-commercial developmental stage (<http://www.biofuelsdigest.com/bdigest/2018/03/09/rennovias-demise-the-rule-of-thirds-and-the-pursuit-of-sustainable-nylon-in-a-world-of-low-oil-prices/>, last checked on 09/08.2018).

Nevertheless, continuous research in its bio-based synthesis has focused on improving the efficiency and economics of the current technologies. Amongst those, microbial production of GlucA has been attempted repeatedly as fermentation-based manufacturing is considered to be cost-effective and environmentally friendly.^{15,16} A summary from the published literature of recent microbial production is given in Table 1. All processes in Table 1 are based on a synthetic pathway designed by Moon et al.¹⁴ which included the heterologous expression of three enzymes in *E. coli*: *myo*-inositol-1-phosphate synthase (Ino1) from *Saccharomyces cerevisiae*, *myo*-inositol oxygenase (MIOX) from *Mus musculus*, and uronate dehydrogenase (Udh) from *Pseudomonas syringae*. This initial pathway was modified subsequently and engineered into different *E. coli* strains to improve GlucA titres and yields (Table 1). However, the maximum GlucA titres obtained from feeding with glucose was 2.5 g/L and a molar yield of maximal 25% which is not viable economically. The reasons for this result are i) high GlucA concentrations lower the cellular pH which is toxic, ii) the microbes' endogenous enzymes metabolise the pathway intermediates for their own growth at limiting glucose concentrations (= metabolic repression) and iii) optimal enzyme ratios are difficult to control at a genetic and cellular level but essential for achieving high carbon flux through the pathway. A range of solutions have been proposed to overcome these limitations, such as using the more acid tolerant yeasts *S. cerevisiae* and *Pichia pastoris*, the continuous supply of glucose to avoid growth-limiting conditions, the use of synthetic scaffolds to control enzyme ratios and enzyme proximity, enzyme engineering to improve enzyme stabilities, genetic knockout of enzymes and dynamic knockdown of enzymes that otherwise metabolise the intermediates, controlled expression of endogenous enzymes to improve the flux, and supplementation of the GlucA pre-precursor *myo*-inositol to support the flux towards GlucA (Table 1). However, due to high genetic and metabolic complexity it is extremely challenging and tedious to improve microbial performances. The highest titre of 6.6 g/l has been achieved only with fed-batch fermentation of engineered *P. pastoris* and the supply of both glucose and *myo*-inositol (Table 1).

An alternative to the cell-based approach is cell-free biocatalysis (or *in vitro* biocatalysis), which allows decoupling of the production pathway from the cellular growth and optimisation of pathway flux towards maximum yields by freely mixing enzymes until ideal ratios and process conditions have been identified.^{17,18} Although cell-free biocatalysis offers greater engineering flexibility and system control than the *in vivo* approach, limitations such as high enzyme and coenzyme costs have challenged its current integration into industrial processes.¹⁹

Here, we describe the use of synthetic biology to construct the first cell-free GlucA pathway and to demonstrate how cell-free systems can overcome some of the constraints encountered in whole cell microbial systems. Further, we demonstrate a viable framework to overcome the limitations ascribed to cell-free biocatalysis by employing a combination of thermostable and non-thermostable enzymes; immobilising the pathway for recycling, and incorporation of a cofactor regeneration system. Specifically, we divided the pathway into two separate reactions, which were operated at high (85 °C) or medium temperatures (40 °C). All six enzymes of the pathway were fused genetically to a synthetic silica-binding peptide²⁰ which allowed their facile immobilisation onto a silica-based matrix for easy collection and reuse without considerable loss of activity. A NADH oxidase also was immobilised for *in situ* cofactor regeneration to avoid the continuous supply of the expensive coenzyme, nicotinamide adenine dinucleotide (NAD⁺).

Table 1. Summary of synthetic GlucA production in whole cell or cell-free systems.

Production host	Genes	Carbon source	Engineering tools	Titer (g/L)	Yield (%)	Reference	Year
<i>E.coli</i> (BL21 derivative)	<i>Ino1</i> <i>MIOX</i> <i>Udh</i>	Glucose	Synthetic pathway with enzymes from various organism	1.13	11.3	Moon et al. (Prather Group)	2009
<i>E.coli</i> (BL21 derivative)	<i>Ino1</i> <i>MIOX</i> <i>Udh</i>	Glucose	Protein scaffolds to mediate enzyme proximity and control enzyme ratios	~2.50	25.0	Moon et al. (Prather Group)	2010
<i>E.coli</i> (K12 derivative)	<i>Ino1</i> <i>MIOX</i> <i>Udh</i>	Glucose	Multiplex automated genome engineering (MAGE)	0.012	2.4	Raman et al.	2014
<i>Gluconoacetobacter</i> sp. A4 (Kombucha)	-	Glucose	-	3.2	3.2	Wang et al.	2014
<i>E.coli</i> (K12 derivative)	<i>Ino1</i> <i>MIOX</i> <i>Udh</i>	Glucose	Dynamic knockdown of gene expression.	1.56	12.4	Reizman et al. (Prather Group)	2015
<i>S. cerevisiae</i>	<i>Ino1</i> <i>INM1</i> <i>MIOX</i> <i>Udh</i>	Glucose	Genetic knockouts, overexpression of endogenous enzymes	0.98	3.3	Gupta et al. (Prather group)	2016
<i>E.coli</i> (BL21 derivative)	<i>cscB</i> <i>cscA</i> <i>cscK</i> <i>ino1</i> <i>suhB</i> <i>MIOX</i> <i>Udh</i>	Sucrose	Synthetic pathway, genetic knockouts, genetic circuit	1.42	27	Qu et al.	2018
<i>E.coli</i> (BL21 derivative)	<i>Ino1</i> <i>SUMO-MIOX</i> <i>Udh</i>	myo-inositol	MIOX fusion tag to improve MIOX solubility	4.85	44.9	Shiue et al. (Prather group)	2014
<i>P. pastoris</i>	<i>MIOX</i> <i>Udh</i>	myo-inositol and glucose	Protein fusion of MIOX and Udh.	6.61	NA	Liu et al.	2016
<i>S. cerevisiae</i>	<i>Ino1</i> <i>MIOX4 (wheat)</i> <i>Udh</i>	myo-inositol	Gene knockout and genome integration of <i>MIOX4</i> and <i>Udh</i> .	6.00	55.5	Chen et al.	2018
<i>E.coli</i> cell-free lysate	<i>PGM1</i> <i>IPS</i> <i>IMP</i> <i>MIOX</i> <i>Udh</i> <i>Nox</i>	Glucose-1-phosphate	-	1.2	14.3	This work	

2 Materials and methods

2.1 Reagents

All chemicals were from Sigma-Aldrich (St. Louis, MO, USA) unless otherwise stated.

2.2 Recombinant enzyme preparation

The detailed construction of all expression plasmids and the primers used here can be found in Table 2 and the Supplementary Information of this article (S1 and Table S1). For recombinant enzyme production, freshly transformed *E. coli* BL21(DE3) harbouring the expression plasmid of interest were cultured in LB medium supplemented with 50 mg/ml carbenicillin. Cells were grown at 37 °C with shaking (250 rpm) until the A_{600} was approximately 0.6. Recombinant protein synthesis of PGM-L, IPS-L and IPS were induced by the addition of 0.4 mM isopropyl β -d-thiogalactoside (IPTG) and further incubation for another 4h. Protein synthesis of IMP-L, IMP, L-*Ta*MIOX, *Ta*MIOX, L-*Mm*MIOX, *Mm*MIOX, Udh-L, Udh and Nox-L and Nox was induced by cooling cells to 16 °C, before adding 0.4 mM IPTG and followed by cultivating the cells at 16 °C for 21h. Cells were harvested by centrifugation at 4000xg for 20 min at 4 °C and stored at -30 °C. They were resuspended in ice-cold 50 mM phosphate buffer or 50 mM HEPES-NaOH buffer (pH 8) except for Nox and Nox-L which were resuspended in 50 mM MES buffer or HEPES buffer pH 6.5. Cells were disrupted by three passages through a French pressure cell, supplemented with 25 U Benzonase (Novagen, Merck) and incubated on ice for 20 min. Cellular debris was removed by centrifugation at 11,000xg for 20 min at 4 °C to yield crude lysates. The crude lysates containing PGM-L, IMP-L, IMP, IPS-L and IPS were incubated further at 85 °C for 20 min to denature host proteins and to partially purify the thermostable recombinant protein. Heat-denatured proteins were removed by centrifugation at 11,000xg for 30 min at 4 °C. The supernatants obtained from the crude or partially purified

lysates were filtered through a 0.2 μm sterile filter (Millipore, Massachusetts, USA), divided into 200 μl portions and stored at $-80\text{ }^{\circ}\text{C}$. Cell lysates that contained L-*Ta*MIOX, *Ta*MIOX, L-*Mm*MIOX, *Mm*MIOX were obtained as describe above but were supplemented with 0.05 mM of the serine protease inhibitor Pefabloc and 1 mM glutathione and the lysates were prepared freshly prior to use. Protein concentrations of all preparations were determined by infrared spectroscopy (Direct Detect, Merck) and all enzyme samples were analysed by SDS-PAGE on a 4–15 or 4-20% Mini-PROTEAN TGX Precast Gels (BioRad, California, USA) or NuPage 4–12% Bis–Tris Gels (Thermofisher Scientific, California, USA) followed by staining with Coomassie Brilliant Blue. The purification values of the partially purified enzymes as compared to the crude lysates were derived on the basis of densitometric analysis of the digitized electrophoresis gels as described elsewhere.²¹

Table 2. Review of enzyme constructions used, enzyme origin, position of the linker fusion, binding to zeolite occurring and retention of enzyme activity upon immobilisation.

Plasmid	Enzyme	Function	Source	Linker position	Zeolite binding	Enzyme activity retained
pET22b-PGM-L	PGM-L	phosphoglucumutase	<i>Thermococcus kodakarensis</i>	C ¹	+ ²	+
pET22b-IPS-L	IPS-L	myo-inositol-3-phosphate synthase	<i>Archaeoglobus fulgidus</i>	C	+	+
pET22b-IPS	IPS	myo-inositol-3-phosphate synthase	<i>Archaeoglobus fulgidus</i>	NA ³	- ⁴	NA
pET22b-IMP-L	IMP-L	inositol-1-monophosphatase	<i>Archaeoglobus fulgidus</i>	C	+	+
pET22b-IMP	IMP	inositol-1-monophosphatase	<i>Archaeoglobus fulgidus</i>	NA	-	+
pET22b-L-TaMIOX	L-TaMIOX	myo-inositol oxygenase	<i>Triticum aestivum</i>	N ⁵	+	+
pET22b-TaMIOX	TaMIOX	myo-inositol oxygenase	<i>Triticum aestivum</i>	NA	-	NA
pET22b-L-MmMIOX	L-MmMIOX	myo-inositol oxygenase	<i>Mus musculus</i>	N	+	-
pET22b-MmMIOX	MmMIOX	myo-inositol oxygenase	<i>Mus musculus</i>	NA	-	NA
pET22b-L-Udh	L-Udh	uronate dehydrogenase	<i>Fulvimarina pelagi</i>	N	+	+
pET22b-Udh-L	Udh-L	uronate dehydrogenase	<i>Fulvimarina pelagi</i>	C	NA	NA
pET22b-Udh	Udh	uronate dehydrogenase	<i>Fulvimarina pelagi</i>	NA	-	NA
pET22b-Nox-L	Nox-L	NADH oxidase	<i>Lactobacillus rhamnosus</i>	C	+	+
pET22b-Nox	Nox	NADH oxidase	<i>Lactobacillus rhamnosus</i>	NA	-	NA

¹ Linker is at C-terminus of fusion protein.

² +, Yes.

³ NA, not applicable.

⁴ -, No.

⁵ Linker is at N-terminus of fusion protein.

2.3 Enzyme activity assays

All assays were performed in triplicate in 50 mM phosphate buffer or 50 mM HEPES-NaOH buffer (pH 8.0). Control reactions were prepared without substrate or without enzyme. One enzyme unit (U) was defined as the amount of enzyme required to form 1 μmol of product per minute in substrate- and coenzyme-saturating conditions and at a steady-state (initial velocity) catalysis.

Phosphoglucomutase activity was determined by measuring the product (G6P) formation in a discontinuous assay coupled with glucose 6-phosphate dehydrogenase (G6PDH) as described elsewhere.²² The initial reaction mixture (100 μl) contained 5 mM G1P, 0.05 mM glucose 1,6-bisphosphate (G16BP) and 10 mM MgCl_2 and was initiated with the addition of 40 μg of PGM-L. The reaction was performed at 85 °C and stopped by cooling on ice to avoid acid-catalysed G6P hydrolysis. The amount of G6P was measured by adding 20 μl of the reaction mixture to a G6PDH reagent (140 μl) consisting of 2U of G6PDH from *Leuconostoc mesenteroides* and 0.5 mM oxidated NAD^+ . G6P was determined spectrophotometrically by monitoring the amount of NADH at 340 nm being produced at 40 °C using a SPECTROstar Nano spectrophotometer (BMG Labtech, Ortenberg, Germany) and the NADH molar extinction coefficient according to Lambert-Beer's law.

Myo-inositol-3-phosphate synthase activity was determined by measuring the substrate (G6P) consumption in a discontinuous assay coupled with G6PDH. The initial reaction mixture (100 μl) consisted of 5 mM G6P, 0.5 mM MgCl_2 and 1 mM NAD^+ and was initiated with the addition of 80 μg of heat-purified enzyme. The reaction was performed at 85 °C and stopped by cooling on ice to avoid acid-catalysed G6P hydrolysis. The reaction mixture was diluted 1:25 in H_2O

and the amount of G6P was measured by adding 20 μ l of the dilution to the G6PDH reagent as described above.

Inositol-1-monophosphatase activity was determined colorimetrically using a malachite green reagent (MLG) as described elsewhere.^{23,24} The reaction mixture (20 μ l) consisted of 2 mM 1L-*myo*-inositol-1-phosphate (IIP), 5 mM MgCl₂ and 1 mM EDTA and was initiated by the addition of 45 μ g of heat-purified enzyme. The reaction was performed at 85 °C and stopped by adding H₂SO₄ to a final concentration of 0.225 M. The amount of enzymatically-released inorganic phosphate (P_i) was determined in 120 μ l of the acid-diluted reaction mixture after adding freshly prepared MLG (30 μ l) and measuring the OD₆₂₀ after 15 min incubation at room temperature (RT). A standard curve of potassium phosphate was used to convert OD₆₂₀ to μ mol of P_i produced.

Uronate dehydrogenase activity was determined spectrophotometrically by measuring the formation of NADH at 340 nm. The reaction mixture (120 μ l) contained 0.5 mM glucuronic acid (GA) and NAD⁺. The reaction was initiated with the addition of 0.2 - 2 μ g of enzyme and was monitored on-line at 40 °C. For the end-point detection, the enzyme reaction was incubated at 40 °C and stopped by adding ice-cold potassium hydroxide to a final concentration of 0.1 M prior to NADH measurement.

Myo-inositol oxygenase activity was determined by measuring the product formation (GA) in a continuous assay coupled with L-Udh. MIOX was activated by incubation with 1 mM ferrous sulphate (Fe) and 2 mM L-cysteine (Cys) for 15 min at room temperature prior to activity measurements. The reaction mixture (180 μ l) contained 48 mM *myo*-inositol, 0.48 mM NAD⁺ and 3 ng of L-Udh. The reaction was initiated with the addition of 150 –300 μ g of enzyme and was monitored on-line at 40 °C. For the end-point detection, the enzyme reaction (125 μ l) was

incubated with 60 mM *myo*-inositol at 40 °C and stopped by placing the reaction on ice. 70 µl of the reaction were transferred to a 96 well-plate, mixed with 100 µg of crude lysate containing the enzyme and 0.45 mM NAD⁺ and the NADH production was monitored on-line at 40 °C.

NADH oxidase activity was determined spectrophotometrically by measuring the depletion of NADH at 340 nm. Nox was activated by incubation with 1mM dithiothreitol (DTT) and 10 µM flavin adenine dinucleotide (FAD) for 10 min at RT as described previously.²⁵ Unbound FAD in the sample was removed by size exclusion chromatography using PD MidiTrap G-10 columns (GE Healthcare, Illinois, United States) prior to activity measurements.²⁶ The reaction mixture (120 µl) contained 0.5 mM NADH and was initiated with the addition of 17.6 - 176 µg of enzyme and was monitored on-line at 40 °C. For end-point detection, the enzyme reaction was incubated at 40 °C and stopped by adding ice-cold potassium hydroxide to a final concentration of 0.1 M prior to NADH measurement. A portion of the reaction was transferred to a 96 well plate and NADH concentration was measured in a spectrophotometer.

2.4 Enzyme immobilisation

Enzymes were immobilised onto synthetic zeolite CBV-100 (Zeolyst International, Conshohocken, PA, USA) as described previously.^{20,27} 5 mg of zeolite was washed three times with 500 µl washing buffer+T (1% Triton X-100 in 50 mM phosphate buffer or HEPES-NaOH buffer, pH 8.0). The starting material (SM) was partially purified enzyme (208 µg PGM-L, 480 µg IPS-L, 50 µg IMP-L) or enzyme in crude cell lysate (120 µg *Mm*MIOX, 60 µg *L-Mm*MIOX, 50 µg of *L-Udh*, 57.6 µg *Nox-L*) in 100 µl of phosphate or HEPES-NaOH buffer (pH 8) (except for immobilisation of *Nox-L* which was performed at pH 6.5). SM was incubated with the washed zeolite with rotation for 30 min at RT. After centrifugation at 10,000×g for 30 s, the supernatant containing the unbound (UB) protein fraction was removed. The zeolite pellet was washed three times with 100 µl of the washing buffer (without Triton X-100) by vortexing and

the resulting supernatants were collected as wash fractions (W) after centrifugation. The washed zeolite pellet was resuspended in 90 μ l washing buffer to yield 100 μ l of zeolite-bound enzyme (B). Enzyme activity was determined in the SM, UB, W and B fractions using the single enzyme assays described above. The immobilisation yield (%) (I_Y) was the amount of immobilised enzyme (derived from its U) relative to the amount of enzyme used (SM) and was calculated as follows: $I_Y = (a_{sm} - a_{uw}) / (a_{sm}) \times 100$, where a_{sm} was the activity in the starting material and a_{uw} was the total residual activity determined in the unbound and wash fractions (W).²⁸ The immobilisation efficiency (%) (I_E) describes the amount of enzyme (derived from its U) that remained *active* upon immobilisation relative to the SM and is calculated as follows: $I_E = (a_b) / (a_{sm} - a_{uw}) \times 100$, where a_b is the observed activity in the bound fraction.²⁸ The immobilisation recovery (%) (I_R) describes the total activity of the active immobilised enzyme (derived from its U) relative to the activity of the SM and is calculated as follows: $I_R = a_b / a_{sm} \times 100$.²⁸ The activities of bound enzymes per mg carrier were calculated as follows: $U/mg \text{ carrier} = a_b / w_c$ where w_c is the amount of enzyme-bound zeolite used for the activity measurements. Immobilised proteins were eluted from the zeolite using 100 μ l 1x SDS-PAGE loading buffer and incubation at 99 °C for 8 min. All fractions were analysed by SDS-PAGE as described above.

2.5 Inhibition test of IPS-L

The inhibitory effect of the substrate G6P on IPS-L activity was tested in presence of cofactors $MnCl_2$ and $MgCl_2$. 0.12 U/mL of IPS-L was mixed with 1 mM NAD^+ at concentrations of G6P ranging from 4 to 40 mM in presence of 0.4 mM $MnCl_2$ or 0.4 mM $MgCl_2$, respectively, in 100 mM HEPES-NaOH buffer, pH 8. The reaction was performed for 0, 10, 30 and 60 min at 85 °C and stopped by cooling on ice and diluted appropriately so that the final concentration of G6P was below 0.5 mM. The amount of G6P was measured in a discontinuous assay coupled with glucose 6-phosphate dehydrogenase (G6PDH). The data was fitted to the rate equation

that describes substrate inhibition (Eq. (1)) where v is the observed initial (steady-state) catalysis rate, V_{\max} is the maximum enzyme velocity under no substrate inhibition conditions, $[S]$ is the substrate concentration, K_m is the Michaelis-Menten constant and K_i is the dissociation constant for the inhibitory substrate-enzyme-substrate complex (e.g. when two substrates bind the enzymes).²⁹

$$v = \frac{V_{\max} [S]}{K_m + [S] \left(1 + \frac{[S]}{K_i}\right)} \quad (1)$$

2.6 Production of glucaric acid from G1P

The one pot reactions with PGM-L, IPS-L and IMP-L was performed by varying the amount of each enzyme between 0.05 and 1 U. Each reaction consisted of 100 mM HEPES-NaOH buffer, pH 8, 5 mM MgCl_2 , 10 μM G16BP and 4 mM NAD^+ in a final reaction volume of 400 or 800 μl unless otherwise stated. Reactions were incubated at 85 °C with varying amounts of G1P and NAD^+ added to the reaction every two hours to a final concentration of 4 mM.

The one pot reactions with *Mm*MIOX, L-Udh and Nox-L were performed by varying the amount of each enzyme between 0.01 and 0.05 U. *Mm*MIOX was incubated with 1 mM Fe and 2 mM Cys at room temperature for 15 min prior to the reaction. Each reaction consisted of 100 mM HEPES-NaOH buffer, pH 8 and varying amounts of NAD^+ in a final reaction volume of 500 or 800 μl , unless otherwise stated. Reactions were incubated at 40 °C with varying amounts of *myo*-inositol or samples from the one pot reaction of PGM-L, IPS-L and IMP-L which were concentrated using a vacuum concentrator (SpeedVac, ThermoFisher Scientific).

2.7 Pathway immobilisation and recycling

PGM-L, IPS-L and IMP-L, L-Udh and Nox-L were immobilised onto zeolite as described in section 2.4. For the single immobilisation of the upper pathway, 0.3 U of PGM-L, 0.24 U of IPS-L and 0.7 or 1.2 U of IMP-L were immobilised separately onto 10, 10 and 30 mg zeolite, respectively, and mixed together subsequently, washed as described in section 2.4 and centrifuged to remove supernatant. For co-immobilisation, the same units of enzyme were immobilised simultaneously onto 50 mg zeolite and treated as described above. 400 μ l of the reaction mixture which included 100 mM HEPES-NaOH buffer, pH 8, 20 mM G1P, 10 μ M G16BP and 4 mM NAD^+ was added to the zeolite-enzyme mixture and incubated for one hour at 85 °C. After the reaction, zeolite-bound enzyme was collected by centrifugation, washed three times and mixed with fresh reaction mixture. This procedure was repeated for a total of 5 times. For the immobilisation of the lower pathway, 0.3 U of L-Udh was immobilised onto 5 mg zeolite or 0.3 U of L-Udh and 0.15 U of Nox-L were co-immobilised onto 20 mg zeolite. 500 μ l of the reaction mixture which comprised 100 mM HEPES-NaOH buffer, pH 8, 20 mM *myo*-inositol and 20 mM NAD^+ or 5 mM NAD^+ was added to the zeolite-enzyme mix. During the first cycle, the immobilised enzymes were incubated for one hour at 40 °C and for the following cycles enzymes were incubated for two hours at 40 °C. After the reaction, zeolite-bound enzyme was removed by centrifugation, washed three times and mixed with fresh reaction mixture. This procedure was repeated for a total of 3 times

2.8 Metabolite analysis

2.8.1 Metabolite analysis using UHPLC-RID

The quantification of G1P, G6P, I1P, GA and GlucA was performed on a GlycanPac AXH-1 column (Dionex, Sunnyvale, CA, USA) (2.1 x 150 mm, 1.9 μ m particle diameter) by ultra-high-performance liquid chromatography (UHPLC) coupled to a refractive index detector (RID) as

described previously (Petroll, et al., Anal Chim Acta, submitted). Briefly, samples (20 μ l) from the reactions were collected at multiple timepoints and ultrafiltrated through a 10 kDa cut-off filterplate (ACROPREP ADV.350UL 10K, Pall Corporation, New York, USA) using a multi-plate vacuum manifold (Pall Corporation) to remove proteins and stop the reaction. Acetonitrile (ACN) was added to the filtrate at a final content of 30 vol. %. and 4 μ l of resulting sample were applied to UHPLC-RID analysis. Where reactions contained Udh or L-Udh, the samples were first boiled for 15 min and trifluoroacetic acid (TFA) was added to a final concentration of 30% vol. prior to ultrafiltration and UHPLC-RID analysis.

2.8.2 Analysis of the Udh reaction with UHPLC-MS

The L-Udh reaction was analysed by mass spectrometry (MS) using the GlycanPac AXH-1 column under the same chromatographic conditions as described above for the UHPLC-RID analysis. Samples were prepared by incubating 1 U of L-Udh crude cell lysate (100 μ l in 50 mM HEPES buffer, pH 8) with 15 mM GA and 15 mM NAD⁺ at 40 °C for 60 min. Samples were ultrafiltered and mixed with ACN as described in section 2.8.1. UHPLC was performed on a VanquishTM (ThermoFisher Scientific) UHPLC system composed of a binary gradient pump, autosampler, a degasser and a temperature-controlled column compartment (TCC). Mass spectrometric (MS) analysis was carried out using a Q ExactiveTM Plus OrbitrapTM mass spectrometer (ThermoFisher Scientific) equipped with a heated electrospray ionisation source (HESI-II). MS was operated in full scan under positive and negative polarity conditions using electron spray ionisation. The parameter settings for the conditions were: sheath and auxiliary gas flow rate at 55 and 10, respectively, spray voltage 3.5 kV, capillary temperature 350 °C, S-lens RF level 50, auxiliary gas heater temperature 400 °C.

2.8.3 Udh product analysis with NMR

Samples were prepared by incubating 1 U of Udh crude cell lysate (100 μ l in 50 mM phosphate buffer, pH 8) with 5 mM GA, 5 mM NAD⁺ in 30 mM phosphate buffer (pH 8) and a total volume of 500 μ l at 40 °C for 150 min. After incubation, samples were ultrafiltered as described above or were boiled for 15 min, then mixed with 0.2% TFA and ultrafiltered. A negative control was obtained by mixing heat-inactivated Udh with 5 mM GA and NAD⁺ without incubation at 40 °C. NMR samples were prepared by diluting 500 μ L of the ultrafiltered reactions with 50 μ L of deuterium oxide. 15 mM of standards of GA, GlucA and glucaro-1,4-lactone were subjected to NMR analysis for reference purposes. All NMR spectra were recorded in 10% deuterium oxide (D₂O) on a Bruker AVIIIHD 500 MHz NMR Spectrometer equipped with a 5mm BBFO SmartProbe™ at 25°C. ¹H NMR experiments were performed using the standard Bruker pre-saturation pulse sequence for water suppression. The chemical shifts were referenced to the internal dimethyl sulfoxide (DMSO) singlet. The acquisition time was set to 2 seconds, and for each sample, 32 scans were collected into a time domain of 32k points using a spectral width of 16.0204 ppm and a receiver gain of 90.5.

3 Results and Discussion

3.1 Pathway design and enzyme selection

The pathway design was adapted from the original synthetic *in vivo* GlucA pathway³⁰ but was modified further to be more sustainable and economically viable by i) using G1P as starting material instead of glucose, which can be derived by enzymatic phosphorolysis from renewable and abundant biomass (e.g. cellulose or starch) without consumption of the expensive coenzyme adenosine 5'-triphosphate (ATP)^{31,32} and ii) by incorporating an NADH regeneration system to minimise the supply of expensive coenzyme (Fig. 1A).

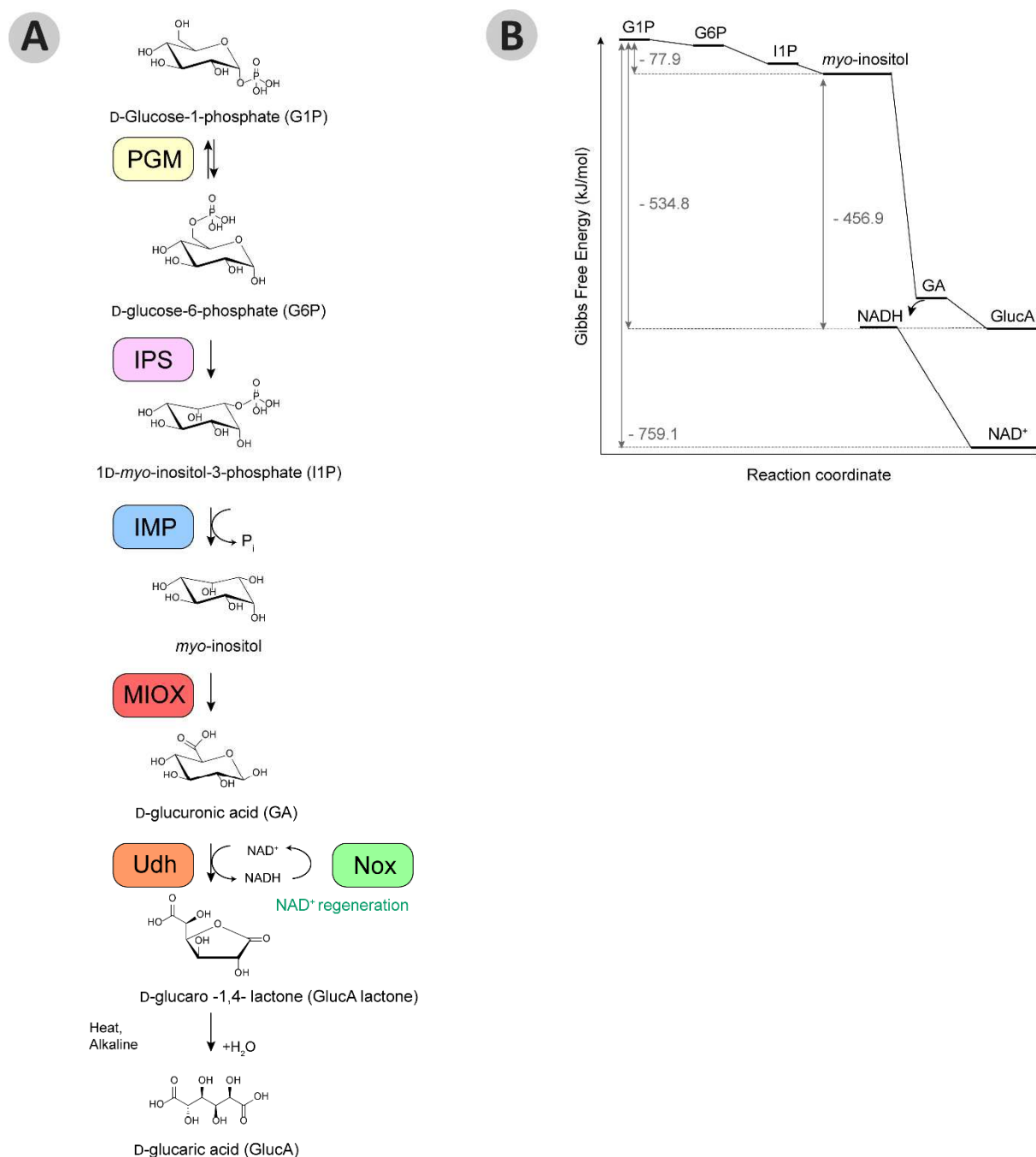


Figure 1. Synthetic pathway design (A) and thermodynamic driving force (B) of the cell-free glucaric acid (GlucA) production from glucose-1-phosphate (G1P). (A) The interconversion of G1P to GlucA and NADH regeneration via a six enzyme pathway. (B) The thermodynamic driving force of the pathway is described by the Gibbs Free Energy ($\Delta_r G'^{\circ}$) that is dissipated during the pathway at 25 °C, pH 8.0 at the standard substrate concentration of 1 M and an ionic strength of 0.1 M. The values were derived using the eQuilibrator.³³ $\Delta_r G'^{\circ}$ for the Udh catalysed reaction describes the conversion of GA to GlucA as $\Delta_r G'$ of the GlucA lactone is not included in the database.

The cell-free pathway consists of six enzymes: i) a phosphoglucumutase (PGM) which catalyses the intermolecular phosphate transfer of G1P to G6P, ii) a *myo*-inositol-3-phosphate synthase (IPS) which catalyses the multi-step oxidation, cyclisation and reduction of G6P to I1P, iii) a inositol-1-monophosphatase (IMP) which catalyses the dephosphorylation of I1P to *myo*-inositol, iv) a *myo*-inositol oxygenase (MIOX) which mediates the oxygen-catalysed oxidation of *myo*-inositol to GA, v) uronate dehydrogenase (Udh) which catalyses the oxidation of GA to the glucaro-1,4 – lactone (GlucA lactone) under reduction of NAD^+ to NADH and vi) a NADH oxidase (Nox) for NAD^+ regeneration as it catalyses the oxidation of NADH to NAD^+ . The GlucA lactone is stable in solution and only opens to the GlucA with heat and alkaline conditions (Fig. 1A). The six enzymes were selected for high catalytic activity (k_{cat}/K_m), substrate selectivity, compatibility, i.e. similar cofactor, pH and temperature preferences, and high stability (i.e. thermostability) (Table 3). Thermostable enzymes were considered good candidates for the assembly of the cell-free synthetic pathway because they often exhibit higher stability than their mesophilic analogues, survive higher process temperatures which improves reactant solubility and mass transfer in favour of higher reaction rates and can be partially (heat) –purified without additional costs.³²⁻³⁶

PGM from *Thermococcus kodakarensis* was selected as it exhibited the highest thermostability and second highest catalytic efficiency as compared to other thermostable PGM homologues²². IPS from *A. fulgidus* was chosen as it exhibits very high thermostability and a higher catalytic efficiency than the mesophilic IPS from *S. cerevisiae* which was the candidate mostly used in previous GlucA pathways.^{31,32,37} IMP from *A. fulgidus* was selected due to its similar temperature optimum to IPS and lower promiscuity to the upstream metabolite G6P than for example the IMP from *Thermotoga maritima*, which was reported previously.^{31,32,38} Promiscuous enzymes are highly disruptive for the metabolic flux of a pathway as they remove

metabolites as dead-end products and result in lower conversion yields unless additional enzymes are included for the byproducts re-entry into the pathway.^{34,39} However, by choosing the more selective IMP from *A. fulgidus* we had to compromise on efficiency, since IMP exhibits a lower catalytic activity towards its substrate (IIP) than the IMP from *T. maritima*.^{24,40} In other studies, MIOX was reported to be the main bottleneck in the metabolic production of GlucA due to its labile nature and catalytic inefficiency.^{14,41} This enzyme was replaced by two well-characterised and catalytically-efficient variants of from *Mus musculus* (MmMIOX)⁴² and *Triticum aestivum* (TaMIOX)⁴³. In general, MIOXs seem to be highly conserved and unique among eukaryotic organisms and no thermostable homologue has been reported. We also investigated and cloned an open reading frame (XP_013327840.1) encoding a putative MIOX from the thermostable fungus, *Rasamsonia emersonii*. The recombinant MIOX was overexpressed in *E. coli* but the enzyme displayed lower activity than MmMIOX and TaMIOX (data not shown). For conversion of GA to GlucA, we selected Udh from *Fulvimarina pelagi* as this was the variant with the highest thermostability reported and had a similar catalytic efficiency to the Udh from *P. syringae*⁴⁴, which was used previously in GlucA pathways.^{14,45} ⁴⁶ Nox from the mesophilic host *Lactobacillus rhamnosus* was used for NADH regeneration as it has been described as a thermostable variant with high catalytic efficiency and capable of generating H₂O as by-product instead of hydrogen peroxide.²⁵ Hydrogen peroxide is often formed by thermostable Nox enzymes from archaeal origin and can be detrimental to enzyme systems.^{47,48}

Table 3. Characteristics of selected enzymes with respect to molecular mass, tertiary structure, temperature optimum, thermostability, pH optimum and catalytic efficiency (K_{cat}/K_m).

Enzyme	Source	M ¹	Subunits ²	T _{opt} (°C)	Thermostability (Half-life)	pH _{opt}	K _{cat} /K _m (mM ⁻¹ s ⁻¹)	Cofactors	Ref
PGM	<i>Thermococcus kodakarensis</i>	50	4	90	> 2.5 h (90°C)	7	192	Mg ²⁺ , G16BP ³	Rashid, et al. 2004
IPS	<i>Archaeoglobus fulgidus</i>	44	4	90	40 min (98°C)	8	80	NAD ⁺ , Zn ²⁺ , Mn ²⁺ or Mg ²⁺	Stieglitz, et al. 2015
IMP	<i>Archaeoglobus fulgidus</i>	28	2	85	NA ⁴	8	2.12	Mg ²⁺	Morgan, et al. 2004; Wang, et al. 2005
MIOX	<i>Mus musculus</i>	33	1	NA	NA	NA	3.5E-4	Fe ²⁺ , L-cysteine	Arner, et al. 2004
	<i>Triticum aestivum</i>	35	1	35	NA	8	ND ⁵	Fe ²⁺ , L-cysteine	Alok, et al. 2015
Udh	<i>Fulvimarina pelagi</i>	29	6	60	3 h (37°C)	8	274	NAD ⁺	Pick, et al. 2015
Nox	<i>Lactobacillus rhamnosus</i>	49	2	65	6.8 h (60°C)	6	3.3E4	FAD ⁶	Zhang, et al. 2012

¹ M; molecular mass of the proteins' subunits

² Number of homologue protein subunits in quaternary protein structure

³ Glucose-1,6-bisphosphate

⁴ NA; not available

⁵ ND; not determined as MIOX protein concentrations used in the assay are unknown. K_m = 5.6 mM and activity is 3.47 μ mol s⁻¹ at pH 8, 35 °C.

⁶ FAD; flavin adenine dinucleotide

The pathway is composed of single thermodynamically-feasible reactions as indicated by a negative change of the Gibbs Free Energy ($\Delta rG'^{\circ}$) for each reaction (Fig 1B). The sum of $\Delta rG'^{\circ}$ for each reaction of the whole pathway including the NAD⁺ regeneration is -759 ± 38.1 kJ/mol. Thus, if optimal metabolite concentrations occur throughout the pathway, the pathway as a whole is anticipated to be thermodynamically feasible.⁴⁹

3.2 Solubility, stability and activity of L-enzymes

The 10 kDa silica-binding peptide sequence (linker, L) was incorporated into the seven enzymes of the pathway (Table 2 and 4). The resulting fusion proteins (L-enzymes) and enzymes (without linker) were expressed heterologously in *E. coli* and compared to investigate the effect of the linker fusion on the solubility, operational stability and activity of the L-enzymes (Fig. S1-S3 and Fig. S5 and Fig. S7).

Most L-enzymes derivatives (N- or C-terminus linker) and enzymes without the linker were produced in *E. coli* with a similar level of solubility (Fig. S1). Udh-L was not expressed at all, regardless of the expression temperature or the IPTG concentration (data not shown). However, when the linker was fused to the N-terminus of Udh (L-Udh), the expression level was improved drastically but the solubility remained very low when compared to Udh (Fig. S1). L-Udh solubility remained low irrespective of the expression temperature, the expression host (*E. coli* B121 or T7 shuffle express cells, Fig. S2), the ionic strength or the pH of the lysis buffer. Udh is a homohexamer of six 29 kDa large subunits (Table 3) whose correct folding and interaction may be impaired by the six 10 kDa-long linker sequences (one per subunit). However, the linker did not affect the correct folding of IPS (Fig. S1) which is a homotetramer of four 40 kDa large subunits (Table 3). The higher stability of IPS and lower impairment from the linker may be attributed to the higher thermostability of IPS and accordingly, a higher rigidity of its protein structure.

In contrast to previous studies^{14,42,43}, supplementation of the expression medium of *Mm*MIOX, *L-Mm*MIOX, *Ta*MIOX and *L-Ta*MIOX *myo*-inositol was not required since the solubility levels of all recombinant MIOX enzymes were acceptably high.

The linker had no effect on the thermostability of PGM-L, IPS-L and IMP-L since no degradation was observed upon their heat-treatment at 85 °C (data not shown) as previously reported for other L-enzymes.²⁷ In contrast, Nox-L was completely degraded after incubation at 60 °C (30 min) and at 80 °C (10 min) (data not shown). Nox was described previously as a thermostable enzyme with a half-life of 6.8h at 60 °C and 2.1h at 80 °C under conditions similar to those used here.²⁵ However, Nox enzymes from different species of *Lactobacillus* have been shown to exhibit temperature optima of only 37 (*L. pentosus*²⁶, or 40 °C (*L. brevis*⁵⁰) and a half-life of 3h at 37 °C.²⁶ Therefore, it was assumed that Nox from *L. rhamnosus* exhibits similar mesophilic characteristics as those from *L. brevis* and *L. pentosus* which explains why the

previous result of a thermostable Nox were not reproducible here and that this finding is independent of linker fusion.

The relative activities (U/mg) of enzymes and L-enzymes were compared under the same assay conditions (except for PGM-L which was only studied as the linker-fusion protein) (Fig. S3). While the linker did not affect the apparent activity of IPS, the linker decreased the activity of L-IMP and L-*Mm*MIOX on average by 31 – 36 and lowered the activity of L-*Ta*MIOX, L-Udh and L-Nox by 90, 92 and 87%, as compared to the enzymes without linker (Fig. S3). The poor activity of L-Udh in comparison to Udh was very likely to result from its drastically decreased solubility, and similarly, it would appear to be the case for Nox-L and IMP-L (Fig. S1). In contrast, L-*Ta*MIOX exhibits similar protein levels as *Ta*MIOX (Fig. S1), and thus the loss of activity of L-*Ta*MIOX is unlikely to result from the amount of protein but rather indicates an altered enzyme activity. The activity of L-*Mm*MIOX was similar to the activity of *Mm*MIOX but different from L-*Ta*MIOX which may suggest a higher robustness of the mouse variant. When comparing the temperature dependences and activities of L-*Ta*MIOX and L-*Mm*MIOX, L-*Mm*MIOX displays a higher temperature optimum and catalytic activity than L-*Ta*MIOX (Fig. S5). From this data, it was decided that L-*Mm*MIOX was the better candidate for cell-free biocatalysis and the study was continued with only L-*Mm*MIOX.

Despite the functional impairments through the linker fusion observed for IMP, Nox and Udh, the resulting activities of all recombinant L-enzymes were sufficient to observe product formation after only a few minutes and accordingly, the following studies were continued with the L-enzymes only.

3.3 Immobilisation of L-Enzymes onto zeolite

L-enzymes from partially purified or crude lysates were immobilised onto zeolite via affinity binding of the linker to silica-based matrices²⁰ (Fig. 2 and Fig. S5). All enzymes that were fused to the linker displayed strong affinity for the zeolite carrier ('bound' fraction, B) while enzymes without the linker did not bind to zeolite ('unbound' fraction, UB) and were removed from the carrier after a stringent wash procedure ('wash' fraction, W1).

To evaluate the effect of the immobilisation onto the activities of the L-enzymes, standard assays were performed with the starting material (SM), unbound, wash and bound fractions (Fig. 2 B shows the results from standard zeolite binding assays of IMP-L and L-Udh) and used to assess the performance of the binding (Table 4). The amount of L-enzyme that bound to zeolite (immobilisation yield, I_Y) was between 83 and 100% regardless of whether the starting material was partially purified (PGM-L, IPS-L or IMP-L) or a crude cell lysate (L-*Mm*MIOX, L-Udh and Nox-L) (Table 4). The high I_Y obtained for all enzymes indicated a high selectivity with which the linker directs the binding of its fusion protein onto zeolite. The binding appears to be independent of the pH as Nox-L was immobilised at a pH of 6.5 while all other L-enzymes were immobilised at a pH of 8. The fraction of enzyme that remained active upon immobilisation (immobilisation efficiency, I_E) was between 0 and 114 % (Table 4). The greatest loss of activity occurred for L-*Mm*MIOX (I_E = 0%) and PGM-L (I_E = 30%), while Nox-L (I_E = 68%) and IMP-L (I_E = 88%) remained largely active. The activity of IPS-L (I_E = 101%) was completely preserved and the activity of L-Udh (I_E = 113%) was enhanced upon immobilisation. Accordingly, these results reflect a total activity of active and bound enzyme (immobilisation recovery, I_R) as 32% (PGM-L), 84% (IPS-L), 85% (IMP-L), 0 % L-*Mm*MIOX, 105% (L-Udh) and 68% (Nox-L).

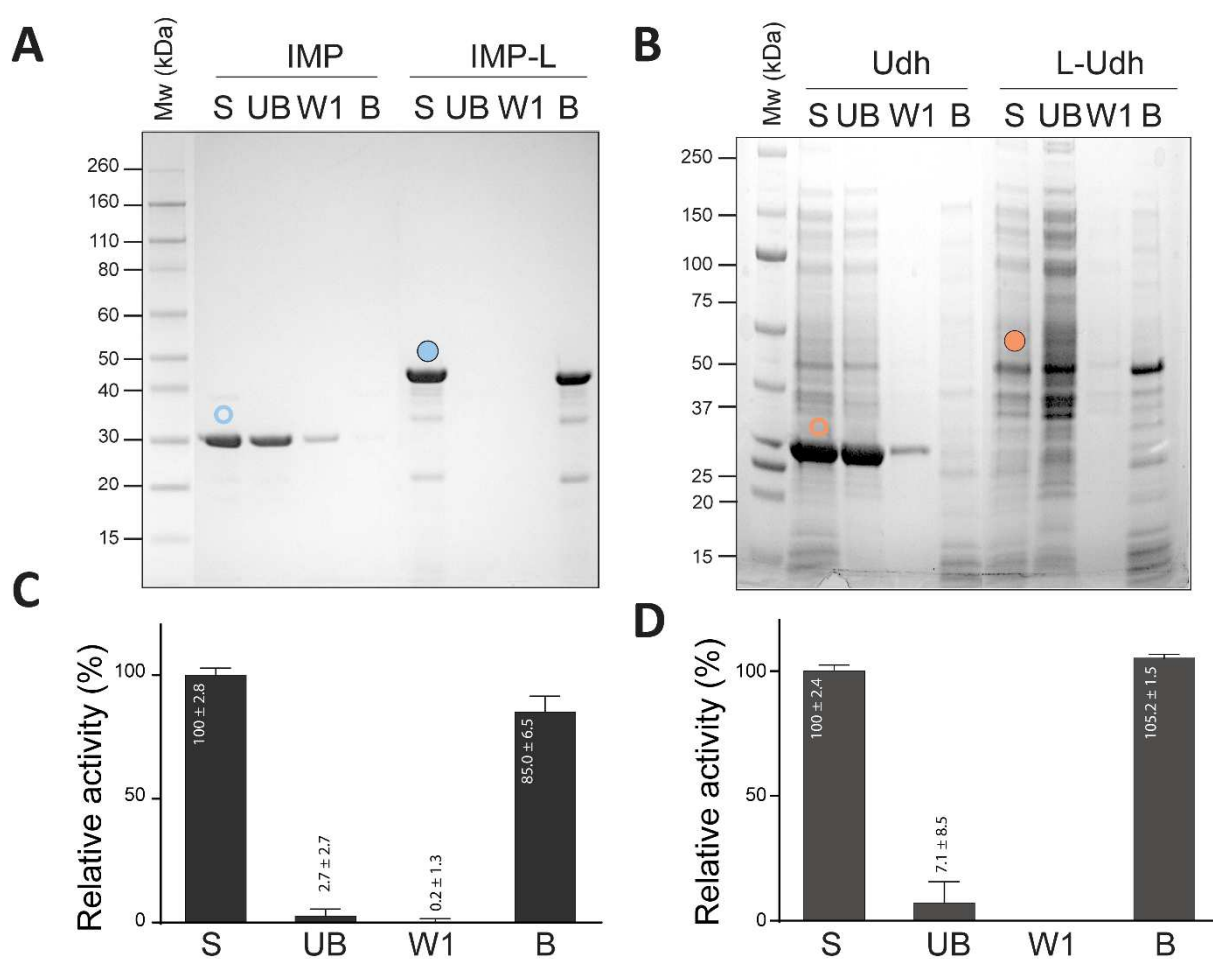





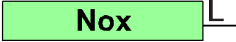


Figure 2. SDS-PAGE analysis showing the zeolite-binding capacity of IMP (without linker) (*empty circle*) and IMP-L (*filled circle*) (**A**), Udh (without linker) (*empty circle*) and L-Udh (*filled circle*) (**B**). Standard zeolite binding assays were performed with partially (heat)-purified protein (IMP and IMP-L) or crude cell lysate (Udh and L-Udh) as described in the “Methods” section. S starting protein sample, UB unbound protein fraction, W1 wash fractions (1 of 3 total), (B) zeolite-bound protein fraction. Enzyme activity profile for the immobilisation of IMP-L (**C**) and L-Udh (**D**) on zeolite, respectively. The averages and standard deviations from three independent replicates are shown.

Table 4. Characteristics of L-enzyme immobilisation. Activity of the free L-enzymes (starting material, SM), (U per mg total protein), immobilisation yield I_Y (%), efficiency I_E (%) and recovery I_R (%), ‘specific activity’ of the immobilised L-enzymes (U per mg carrier). Specific activities are shown as the averages and standard deviations from three independent replicates. I_Y , I_E and I_R (%) represent the averages from three independent replicates.

Enzyme	Construction	Activity ^a (U/mg)	I_Y (%)	I_E (%)	I_R (%)	Activity ^b (U/mg carrier)
PGM-L		1.05 ± 0.12^c	98.9	32.5	32.1	$1.4E^{-3}$
IPS-L		$5.8E^{-1} \pm 0.07^c$	83.3	101.2	84.3	$3.9E^{-3}$
IMP-L		$6.7E^{-1} \pm 0.26^c$	97.1	87.6	85.0	$5.6E^{-3}$
L- <i>Mm</i> MIOX		$4.5E^{-2} \pm 0.07^c$	100	0	0	0
L-Udh		$9.6E^{-1} \pm 0.04^d$	92.9	113.2	105.2	$1.1E^{-2}$
Nox-L		$3.6E^{-2} \pm 0.002^d$	100	67.7	67.6	$3.3E^{-4}$

^a Units of the free L-enzymes per total protein amount in partially (heat)- purified or crude lysates

^b Units of the bound L-enzymes per mg zeolite

^c Activity was tested at 85 °C in 50 mM HEPES-NaOH, pH 8

^d Activity was tested at 40 °C in 50 mM HEPES-NaOH, pH 8

The loss of activity is usually assumed to occur due to interactions of proteins with the matrix altering the flexibility/protein assembly required for functional substrate/product interaction and is largely empirical. Thus, it can only be speculated why *Mm*MIOX lost all of its activity, PGM-L lost most of its activity and L-Udh gained activity. *Mm*MIOX was demonstrated previously to adopt three different structural forms in solution due to its high flexibility and which corresponded to largely differing kinetic properties.⁴² Accordingly, the immobilisation of L-*Mm*MIOX onto zeolite may cause a loss in flexibility that is crucial to maintain high enzyme activity, may lock the enzyme irreversibly in a less active form or hinder the enzyme active site being accessed by the substrate. The activity of *Mm*MIOX (which does not bind to zeolite, Fig. S5 C) was not affected in presence of zeolite confirming that zeolite matrix is not causing the loss of activity but the binding to zeolite was affected (Fig. S6).

PGM-L was reported as a highly ‘sticky’ enzyme (Dr. André Pick, Technical University of Munich, Germany, personal communication) which may have resulted in its partial unfolding upon contact with the matrix or burying of the active site and thus giving lower functionality. The activity enhancement of L-Udh is intriguing as Udh is a homohexameric protein with small subunits whose solubility is impaired drastically by linker-fusion when compared to Udh without the linker (Fig. S1 and S2). The activities observed translate to a ‘specific’ immobilised activity (U/mg carrier) of 0.003 – 0.014 U/mg which is about 100 times lower than the activity per total protein amount (U/mg) (Table 4). The amount of enzyme that was immobilised was non-saturating to ensure high immobilisation yields onto the zeolite.

In summary, linker fusion facilitated nearly 100% binding of all L-enzymes onto zeolite with activity enhancement (105%) of L-Udh, high recoveries (78 – 85%) of IMP-L, IPS-L and Nox-L activities and satisfactory recovery of PGM-L activity (32%) upon immobilisation (Table 4). The only exception was L-*Mm*MIOX showed a complete loss of activity upon immobilisation

3.4 Two-pot reaction for the production of glucaric acid from G1P

The reaction conditions have to be optimised for all enzymes to ensure high pathway productivity including the pH, temperature and cofactors. PGM-L, IPS-L and IMP-L operate preferably at high temperatures (e.g. 85 °C), *Mm*MIOX is active optimally at 40 °C (Fig. S4) while L-Udh and Nox-L exhibit their highest activity at temperatures around 60 °C (Table 3). To compromise between the different temperature dependencies of the thermostable and the mesophilic enzymes, we decided to divide the pathway into two reactions (‘two-pot’) with PGM-L, IPS-L and IMP-L being operated at 85 °C, pH 8 and *Mm*MIOX, L-Udh and Nox-L being operated at 40 °C, at the same pH. Upon completion of the first reaction, the temperature was decreased to 40 °C and enzymes and cofactors for the second reaction were added. Since

all enzymes operate preferably at pH 8, except for Nox-L which shows higher activities at a pH of 6, the buffers of the two-pot reactions are reasonably compatible.

We optimised the two-pot reactions separately to achieve high conversion rates and used the UHPLC-RID method to test various reaction conditions and their impact on the metabolic flux of the pathway rapidly. Prior to this, we determined whether the L-enzymes and *Mm*MIOX can be stored at -80 °C without loss of activity to allow repeated enzyme preparation to be omitted and to ensure uniform enzyme activities for pathway optimisation (A detailed experimental description and results can be found in the Supplementary Information of this article under S5 and Fig. S7). Except for *Mm*MIOX, all enzymes showed high stability when stored at -80 °C and were prepared in large batches, stored at -80 °C and thawed as needed. L-*Mm*MIOX was expressed in large batches but was stored as whole cells at -80 °C and was lysed immediately prior to use.

3.4.1 Optimising the production of the intermediate *myo*-inositol

The enzymatic rates of PGM-L, IPS-L and IMP-L exhibited a high degree of control on the metabolic flux (i.e. the turnover rate) of the pathway. For example, a suitable amount of PGM-L and IPS-L (e.g. 0.1 U each), or a lower amount of PGM-L (e.g. 0.025 U) and four times the amount of IPS-L (e.g. 0.1 U) were required to achieve a conversion of G1P to G6P of close to 80% (Fig S8). This result was expected as the reaction catalysed by PGM is reversible and constitutes a low thermodynamic driving force with a Gibbs Free Energy change of only 7 kJ/mol (Fig. 1B). Accordingly, a considerable flux is carried in the reverse direction and ‘wastes’ enzyme units.⁵¹ By increasing the PGM-L concentration, more enzyme units are available to facilitate the forward reaction and to re-adjust the G1P-G6P equilibrium more rapidly upon the continuous removal of G6P by IPS-L. Likewise, by increasing IPS-L concentration and enhancing the conversion rate of G6P to I1P, the reactant equilibrium of G1P and G6P is reached in a faster fashion which ultimately results in an increased flux in the

forward direction. Despite IPS-L positively controlling the flux, we observed that the conversion of G6P to IIP was arrested at G6P concentrations near 10 mM (Fig S8). First, it was assumed that the reaction of IPS ceases due to degradation of its coenzyme NAD^+ . Although IPS regenerates NAD^+ within its catalytic cycle, NAD^+ is degraded rapidly at high temperatures (e.g. $> 85^\circ\text{C}$ ⁵²) and an hourly supplementation of mM NAD^+ into a cell-free biocatalytic process was able to maintain longer reaction rates of IPS at 90°C .³¹ Therefore, NAD^+ was added every two hours at a concentration of 4 mM. Nevertheless, IIP production reached a maximum of 10 mM (data not shown), so we investigated the impact of substrate concentration on IPS activity to rule out substrate inhibition. IPS from *A. fulgidus* is highly dependent on divalent cations and thus Mn^{2+} was added to the reaction as it was described as a better activator than Mg^{2+} .³⁷ However, we observed different enzyme kinetics in presence of Mn^{2+} compared to Mg^{2+} (Fig. 3). At low G6P concentrations (5 mM) the reaction rate was higher with MnCl_2 than with MgCl_2 , similar to a previous study.³⁷ In presence of MnCl_2 , the reaction rate decreased dramatically to 20% of the initial rate when the G6P concentration was greater than 8 mM. In contrast, in presence of MgCl_2 the reaction rate increased more slowly with increasing G6P concentration and only decreased to 60% of the initial rate at a G6P concentration higher than 12 mM. The kinetic behaviour observed in presence of MgCl_2 resembled the effect of substrate inhibition (Fig. 3, grey and black dotted lines). Substrate inhibition primarily is caused by two substrates instead of one binding to the active site of an enzyme resulting in an inhibitory substrate-enzyme-substrate complex (K_i).²⁹ The kinetic behaviour observed in presence of MgCl_2 occurs when the dissociation constant K_i of the inhibitor complex is higher (thus less stable) than the Michaelis-Menten constant K_m (e.g. $K_i/K_m = 10$). In contrast, the reaction rate in presence of MnCl_2 behaves differently than is predicted with classical substrate inhibition models (Fig. 3, black, dotted line). Instead, it resembled a model where K_i approximates K_m and thus is more stable (e.g. $K_i/K_m = 1$).²⁹ Manganese is a multivalent metal with chelating properties and it may alter the active site architecture of IPS differently from magnesium, which

results in an increased stability of the inhibitory substrate complex reflected through a lower K_i . Consequently, only Mg^{2+} was supplied in the following reactions and resulted in the flux proceeding even at high G6P concentrations (Fig. S9).

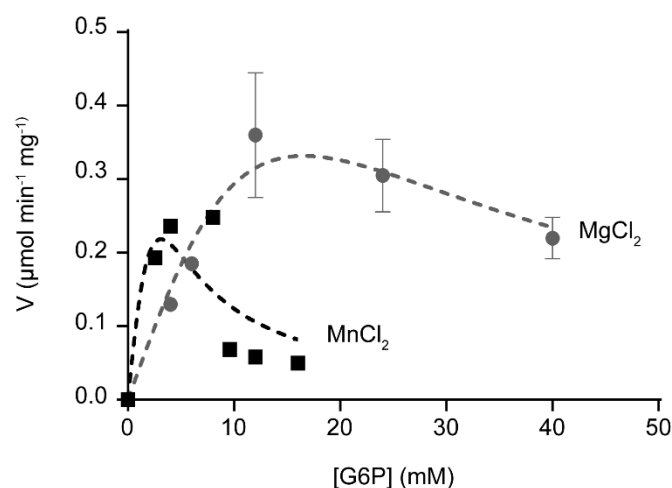


Figure 3. Effects of 0.4 mM $MgCl_2$ (grey circles) and $MnCl_2$ (black squares) on the substrate inhibition of IPS-L at saturating NAD^+ concentrations (1 mM) in 100 mM HEPES-NaOH buffer, pH 8 at 85 °C. Single values or averages and standard deviation from three independent replicates are shown. Data were fitted to Eq. (1) of substrate inhibition ($MgCl_2$: grey, dotted line; $MnCl_2$ black, dotted line).

We observed that IMP-L controlled the flux of the IPS-L and PGM-L reactions firmly. The more IMP-L that was present, the more G6P was converted to I1P which in turn improved the conversion of G1P to G6P to 90% (Fig. 4A and Fig. S9). The multi-step reaction of G6P to I1P catalysed by IPS-L is virtually irreversible with a large thermodynamic driving force (Fig. 1B). However, the first step of the reaction is reversible⁵³ and may be susceptible to flux control by changes in the product concentration. Although IMP-L increased the flux noticeably through IPS-L and PGM-L, I1P accumulated faster than it was transformed to *myo*-inositol (Fig. 4A and Fig. S10). In contrast, when the IMP-L concentration was decreased and the IPS-L

concentration was increased relative to PGM-L, the flux through PGM-L reached nearly 100%, while less I1P accumulated and more *myo*-inositol was produced (Fig. 4B). These results are unexpected as one would expect less *myo*-inositol production when less IMP-L is present. It appeared that by lowering the force on the IPS-L reaction through limiting IMP-L concentrations, less G6P was converted to I1P, while the high IPS-L concentration ensured a high flux through PGM-L. Likewise, less I1P accumulated when the concentration of IMP-L was increased substantially relative to PGM-L and IPS-L, indicating that although a high driving force on the IPS-L flux was exerted, IMP-L was active enough to counterbalance the rapid production of I1P by conversion to *myo*-inositol (Fig. 4C). However, to achieve high amounts of IMP-L required the addition of large sample volumes, resulting in higher sample volumes than in previous experiments. Consequently, the concentration of PGM-L and IPS-L was reduced and was inadequate to drive the flux effectively towards G6P.

We conclude that a high flux through PGM-L and a low flux through IPS-L appear to play a major role for high reaction rates of IMP-L and production of *myo*-inositol.

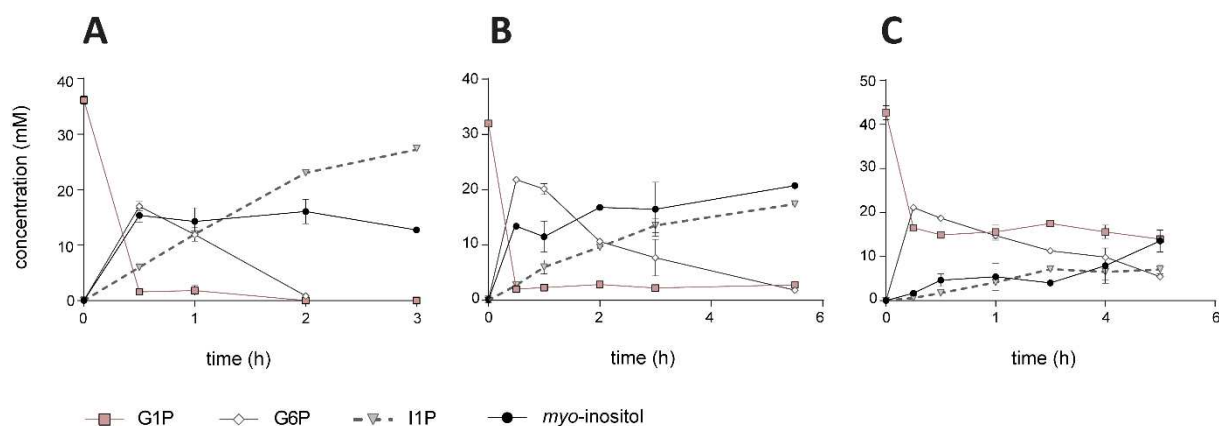


Figure 4. Effect of the enzyme ratios of PGM-L, IPS-L and IMP-L on the cell-free synthesis of *myo*-inositol from glucose-1-phosphate (G1P). (A) 0.1 U PGM-L, 0.2 U IPS-L and 0.6 U IMP-L, (B) 0.1 U PGM-L, 0.3 U IPS-L and 0.1 U IMP-L and (C) 0.1 U PGM-L, 0.2 U IPS-L and 1 U IMP-L were incubated with 40 mM *myo*-inositol, 5 mM MgCl₂ and 10 μM G16BP at 85 °C in 100 mM HEPES buffer, pH 8 (A and B: total reaction volume was 400 μl; C: total reaction volume was 800 μl). 4 mM NAD⁺ were added to the reaction every two hours. G1P (squares), glucose-6-phosphate (G6P, diamonds) and inositol-1-phosphate (I1P, triangles) were measured by UHPLC-RID and *myo*-inositol (circles) was calculated as the difference between the added amount of G1P and the sum of all measured metabolites. The averages and standard deviation from two technical replicates are shown. *myo*-inositol production in (A) was 12.7 ± 0.1 after 3h, in (B) was 20.7 ± 0.4 after 5.5h and in (C) was 13.5 ± 2.5 mM after 5.5h.

The yields of *myo*-inositol were lower than in previous studies, where > 200 mM *myo*-inositol was produced from amylose³¹ and > 500 mM from maltodextrin with near 100% yields.³² Similar enzymes to those in the present study were used with the difference that an amylose/maltodextrin phosphorylating enzyme was added upstream of PGM for the production of G1P and the IMP was derived from *T. maritima* and not *A. fulgidus*. The phosphorylating enzyme was feeding G1P into the reaction at a slow but continuous rate which may have been a key factor for a high flux through PGM-L, IPS-L and IMP-L. Therefore, we tested whether substrate feed in facilitates higher *myo*-inositol production. When G1P was fed slowly into the reaction composed of 0.1 U PGM-L, 0.2 U IPS-L and 0.6 U IMP-L at concentrations of 6 mM every 10 min, 6 mM every 30 min, or 20 mM every 20 min to a final concentration of 40 mM, I1P did not accumulate, supporting the view that the slow production of G6P and I1P improved

the flux through IMP-L (data not shown). The rate of feed-in imposed a high control on the flux as a too low f rate resulted in a significant accumulation of G1P (data not shown). We conclude that an optimal feed-in rate of G1P that mediates a high flux through PGM-L and a low flux through IPS-L appears to be crucial for high reaction rates of IMP-L. The use of the catalytically more efficient IMP from *T. maritima* might also mitigate the low *myo*-inositol yield by increasing reaction rates.

3.4.2 Optimising the GlucA production from *myo*-inositol

The MIOX reaction is thermodynamically highly favourable (Fig. 1B) while the rate with which *Mm*MIOX catalyses the reaction rate is very slow (Table 3, Table 4 shown as U/mg of L-*Mm*MIOX and compared to *Mm*MIOX in Fig. S3). These facts, in addition to the labile nature of MIOX, have been a major bottleneck in previous GlucA production pathways that were conducted *in vivo*.^{14,41,46} In contrast, the L-Udh reaction proceeds very rapidly at high substrate concentrations and in presence of equimolar amounts of NAD⁺ (Table 3, 4 and Petroll et al., Anal Chim Acta, submitted).

Operation in a cell-free manner has allowed the free increase of the *Mm*MIOX units present in the reaction vessel to achieve high yield conversion of *myo*-inositol to GA in short times without being limited by degradation of the enzyme. For example, in presence of *Mm*MIOX, L-Udh, and 40 mM NAD⁺, 40 mM of *myo*-inositol was converted to 29.7 mM GlucA after three hours at 40 °C with a molar yield of 74.25% (Fig. 5A and Table 5A). Low amounts of GA were measured in the reaction which are attributable to the rapid reaction rates of L-Udh. This conversion was largely dependent on the activation of *Mm*MIOX with ferrous (Fe) and L-cysteine (Cys) prior to incubation with *myo*-inositol. Without the pre-activation only 4.6 ± 0.41 mM GlucA were produced after three hours at 40 °C (Fig. 5B). It was reported elsewhere that where *Mm*MIOX was purified from cell-free extracts, it required activation by ferrous and

cysteine.⁴² However, the findings here suggest that the cell-free environment is not supplying sufficient Fe and Cys for MIOX, which is crucial for a high yield of GlucA production.

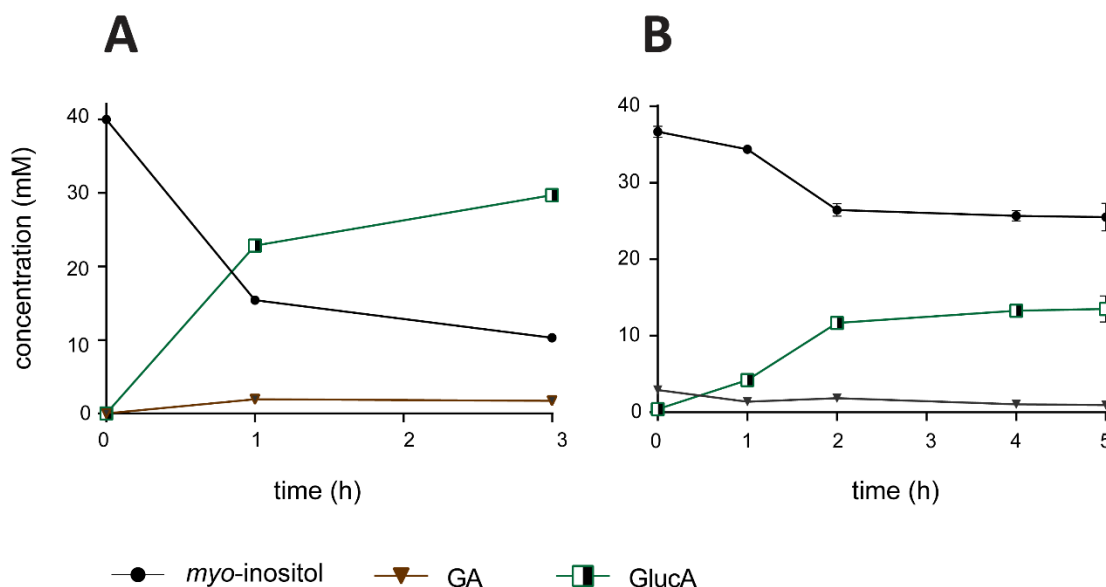


Figure 5. Effect of the cofactors ferrous (Fe) and L-cysteine (Cys) on the cell-free synthesis of glucaric acid (GlucA) from *myo*-inositol. 0.02 U of *Mm*MIOX which was pre-incubated with 1 mM Fe and 2 mM Cys (A) and 0.02 U of *Mm*MIOX without pre-incubation (B) were incubated with 40 mM of *myo*-inositol, 0.3 U of L-Udh and 40 mM NAD⁺ in 100 mM HEPES buffer, pH 8 at 40 °C (total volume 500 μ l). Glucuronic acid (GA, triangles) and GlucA (half-filled squares) were measured by UHPLC-RID and *myo*-inositol (circles) was calculated as the difference between the added amount of *myo*-inositol and the sum of all measured metabolites. Single values (A) or averages and standard deviation from two technical replicates are shown. (B). *myo*-inositol production in (A) was 29 mM after 3h and in (B) was 13.5 ± 1.7 mM.

In contrast to the upper pathway (section 3.4.1), changing the number of enzyme units of *Mm*MIOX and L-Udh only altered the reaction rates (kinetics) of the pathway, while little control was exerted on the metabolic flux. This finding is expected as the pathway from *myo*-inositol to GlucA is largely unidirectional and thus less susceptible to changes in metabolite concentration.

3.4.3 Production of GlucA from G1P in a two-pot reaction

When the samples from the PGM-L, IPS-L and IMP-L one-pot reactions were combined with *Mm*MIOX and L-Udh, a general impediment was the high dilution of the reactions due to the high volume of sample needed to accommodate 0.04 U of *Mm*MIOX (as U/mg were very low, Table 4). Therefore, the one-pot samples were concentrated first and then mixed with *Mm*MIOX and L-Udh. The highest amount of GlucA produced from 40 mM G1P was 5.69 ± 1.06 mM after 3h hours of incubation at 40 °C, which translates to a molar yield of 14.3% (Table 5C). Thus, the total reaction time for the cell-free GlucA production was eight hours and was achieved by mixing 0.1 U of PGM-L, 0.2 U of IPS-L and 1 U of IMP-L in one pot at 85 °C, and 0.04 U of *Mm*MIOX and 0.3 U of L-Udh in a second pot at 40 °C.

3.4.4 NADH regeneration

The Udh reaction requires equimolar amounts of NAD^+ and GA to produce GlucA. For example, the one-pot reaction of *Mm*MIOX and L-Udh with 40 mM *myo*-inositol and 5 mM NAD^+ resulted in the formation of only 13.6 mM GlucA after 3 hours, whereas twice as much GlucA (29.7 mM) was produced when 40 mM NAD^+ was added (Fig. 6A). The production of more than 5 mM GlucA (despite only 5 mM NAD^+ added to the reaction) was attributed to the fact that the enzyme samples were crude cell lysates which contained endogenous NAD^+ . In addition, we have observed previously that crude lysates from *E. coli* BL21 cells oxidise NADH to some extent upon its addition and incubation at 40 °C and may exhibit NADH regeneration to some extent (data not shown). Achieving a high GlucA titre requires the addition of high amounts of NADH which is expensive and reduces the cost-effectiveness of the cell-free approach. Therefore, we tested whether incorporating an additional NADH regeneration system would minimise the cofactor requirements. Nox-L catalyses the oxidation of NADH to NAD^+

and regenerates NAD^+ that is otherwise consumed by Udh. When 0.05 U Nox-L were added to the *Mm*MIOX and L-Udh one-pot reaction, 19.8 mM GlucA were produced from 40 mM *myo*-inositol in presence of only 5 mM NAD^+ (Fig. 6A and Table 5B). Thus, the Nox-L regeneration system maintained the GlucA production at 70% of the initial system (which is the control reaction in Fig 6A and where equivalent amounts of *myo*-inositol and NAD^+ were provided), while it required an eight times lower amount of NAD^+ than the initial system. The finding that only 70% of the initial GlucA titre was produced indicates that NAD^+ was not regenerated at adequate rates. Increasing the amount of Nox-L may increase the NAD^+ regeneration rate and result in higher GlucA formation.

3.5 Pathway immobilisation and recycling

All enzymes from the two-pot reactions were immobilised onto zeolite (including the NADH regeneration system) to minimise the use of enzymes and costs of the cell-free pathway. As described in section 3.3, L-*Mm*MIOX lost all activity upon immobilisation onto zeolite and was supplied freshly into the reaction for each cycle. First, it was tested whether the productivity is the same when PGM-L, IPS-L and IMP-L are immobilised separately ('singly') or co-immobilised (Fig. 6B). Adequate enzyme units were used to account for the percentage of enzyme that loses activity upon immobilisation to allow comparison with the results from the free enzyme system (Table 4). For example, to obtain 0.1 U of immobilised and active PGM-L, 0.3 units were used for immobilisation as described in section 2.7. There was no significant difference between the *myo*-inositol production from the singly immobilised enzyme as compared to the co-immobilised enzymes as 4.5 ± 3.2 and 6.9 ± 1.5 mM *myo*-inositol were produced from 40 mM G1P after 2h incubation at 85 °C, respectively (Fig. 6B and Fig. S10). Protein analysis of the immobilised enzymes confirmed that they were all bound to the zeolite, indicating that the desired enzyme ratios were achieved, and no competing effects occurred (data not shown).

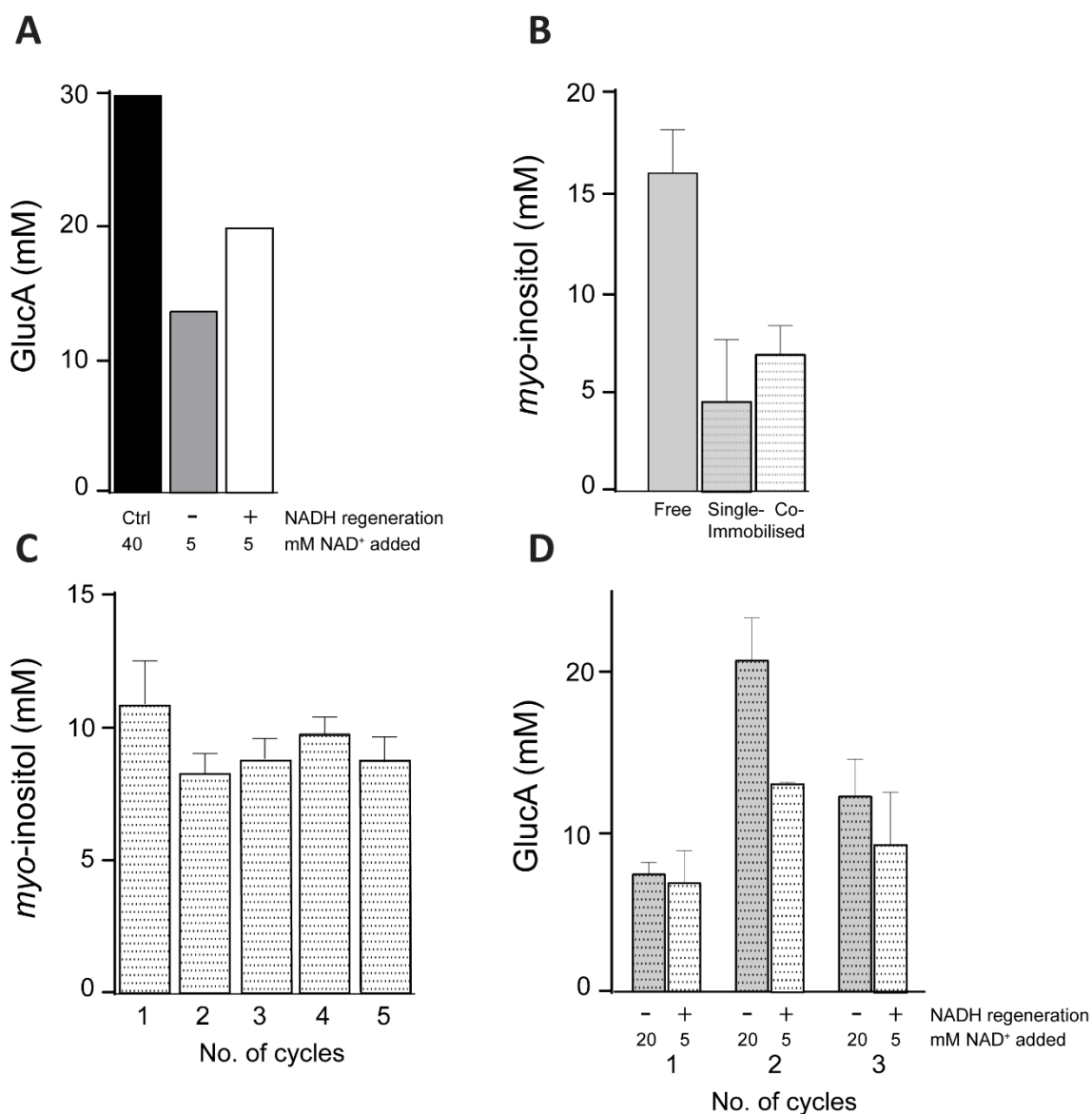


Figure 6. Effect of integrating Nox-L for NADH regeneration on the production of glucaric acid (GlucA) (A), effect of single or co-immobilisation of PGM-L, IPS-L and IMP-L onto zeolite on *myo*-inositol production (B), reusability of co-immobilised PGM-L, IPS-L and IMP-L for the cell-free synthesis of *myo*-inositol (C) and reusability of immobilised L-Udh (dark shaded bars) or co-immobilised L-Udh and Nox-L (white shaded bars) for synthesis of GlucA (D). (A) *Mm*MIOX and L-Udh were incubated at 40 °C for 3h with 40 mM *myo*-inositol and either 40 mM NAD⁺ (ctrl), 5 mM NAD⁺ or 5 mM NAD⁺ together with Nox-L for NADH regeneration. (B) 0.1 U PGM-L, 0.2 U IPS-L and 0.6 U IMP-L were employed either as free enzymes, singly immobilised onto zeolite and subsequently mixed or co-immobilised onto zeolite for *myo*-inositol synthesis from 40 mM G1P at 85 °C in 2h. (C) 0.1 U PGM-L, 0.2 U IPS-L and 1 U IMP-L were co-immobilised onto zeolite and incubated with 20 mM *myo*-inositol at 85 °C for five times at one hour each. (D) *Mm*MIOX was incubated with 20 mM *myo*-inositol at 85 °C for five times at one hour each.

20 mM *myo*-inositol and 20 or 5 mM NAD⁺ and zeolite-immobilised L-Udh or co-immobilised L-Udh and Nox-L, respectively, at 40°C for in total three cycles that lasted 1, 2 and 2 hours. The single values (**A**) or averages and standard deviation from two technical replicates (**B**, **C** and **D**) are shown.

In comparison to the free enzyme system in which 16.6 ± 2.2 mM *myo*-inositol was generated, the productivity of the immobilised system was approximately 60% lower (Fig. 6B). As shown in Fig. S11, G6P accumulates within the reaction of the single and co-immobilised pathway which hindered the flux proceeding to *myo*-inositol. It appears that the kinetics of the immobilised pathway may differ from the free enzyme pathway, indicating that these two systems are not necessarily transferable and require separate optimisation. Improved reaction rates may be achieved through enhancing the amount of IPS-L that is immobilised as IPS-L was the bottleneck in the immobilised systems.

Next, we tested whether the immobilised enzymes can be re-used. As there was no difference in productivity between the separate and co-immobilised enzymes and since co-immobilising enzymes is more economically-attractive, they were incubated with G1P at 85 °C for a total five cycles that lasted one hour each. This time, 1 U IMP were co-immobilised with 0.1 U PGM-L and 0.2 U IPS-L without increasing the sample volume as the immobilisation of enzymes onto zeolites allows a significant reduction of the sample volume. The immobilised enzymes were collected by centrifugation after each cycle and incubated with fresh substrate. The highest production was achieved in the first cycle with 9.1 ± 2.2 mM *myo*-inositol from 20 mM G1P, whereas in the following four cycles the production of *myo*-inositol reached 6 to 7 mM (Fig. 6C). As the result of the immobilisation of higher amounts of IMP-L, after one hour of incubation with G1P, the productivity of *myo*-inositol was already greater than reported for the free, single and co-immobilised enzyme reactions after two hours of incubation (Fig. 6B)

which were composed of only 0.6 U IMP-L. In total, 36.0 ± 7.9 mM *myo*-inositol was produced from 100 mM G1P in five hours.

Myo-inositol production after the fifth cycle was 71% of the first cycle, indicating a highly robust immobilisation and high thermostability of all three enzymes at 85 °C. The loss of activity after the first cycle may be attributed to an initial stripping of loosely bound enzyme at the high reaction temperature and the subsequent centrifugation step. These results suggest that the enzymes can be re-used for more than five cycles and are likely to be sustainable for even longer reaction times for higher *myo*-inositol formation. The samples from the five cycles were combined, concentrated and mixed with *Mm*MIOX and L-Udh and resulted in the production of 6.7 ± 0.5 mM GlucA after three hours of incubation at 40 °C with a yield of 6.7% (Table 5D).

We tested whether L-Udh and Nox-L can be recycled for economical GlucA production. Fresh *Mm*MIOX was incubated with 20 mM *myo*-inositol, equimolar amounts of NAD^+ and 0.3 U of immobilised L-Udh for in total three cycles which lasted one, two and two hours, respectively. In the first cycle, 7.4 ± 0.8 mM GlucA was produced while in the second and third cycles 20.7 ± 2.7 mM and 12.2 ± 2.33 mM GlucA were produced (Fig. 6D, grey bars). We investigated whether the co-immobilisation of L-Udh with Nox-L for NAD^+ regeneration was capable of sustaining a high yield of GlucA production with only 5 mM instead of 20 mM NAD^+ in the reaction. In the first cycle, 6.8 ± 2.0 mM GlucA was produced while in the second and third cycles 13.09 ± 0.1 mM and 9.1 ± 3.3 mM GlucA were produced. Although more Nox-L was immobilised than in the previous experiments (0.01 U), GlucA formation was again about 1.5 times lower than when equimolar amounts of NAD^+ and *myo*-inositol were supplied, indicating that the reaction rate of Nox-L was still not sufficient to maintain a constant NAD^+ concentration (Fig. 6D, white bars). Protein analysis of the co-immobilised enzymes showed

that only L-Udh bound to zeolite while Nox-L was almost not visible on the SDS-PAGE. L-Udh is a homohexamer and thus is constituted of six linkers per enzyme whereas Nox-L is a homodimer with only two linkers per enzyme. The affinity of L-Udh to zeolite may be higher than that of Nox-L and therefore, L-Udh may compete with the binding of Nox-L. In this respect, it may be beneficial to singly immobilise both enzymes and subsequently mix the samples. Thus, by incorporating cofactor regeneration, a GlucA yield of near 50% was achieved while the dependency on NADH was reduced four times as compared to having no regeneration system.

In addition, it was found that enzymes can be stored after their immobilisation on zeolite either in solution or lyophilised without alteration of zeolite binding for multiple weeks at 4 °C enzymes (described in Supplementary Results S5 and Fig. S11). Enzyme inactivation due to long-term storage was similar in the immobilised and non-immobilised samples. Accordingly, immobilised enzymes can be recovered from the reaction, stored and re-used as means to conserve costly enzymes. A detailed experimental description and results can be found in the Supplementary Information in S7 and Fig. S11.

Table 5. Summary of the cell-free production of GlucA from *myo*-inositol (A, B) or glucose-1-phosphate (G1P) (C, D) using free or immobilised enzymes (Zeo).

#	Enzyme units	Reaction			Final GlucA titre (mM)	GlucA molar yield (%)
		volume (ml)	time (h)	composition		
A	0.02 U <i>Mm</i> MIOX 0.3 U L-Udh	0.8	3	40 mM <i>myo</i> -inositol, 40 mM NAD ⁺ 100 mM HEPES	29.7	74.3
B	0.02 U <i>Mm</i> MIOX 0.3 U L-Udh 0.05 U Nox-L	0.8	3	40 mM <i>myo</i> -inositol, 5 mM NAD ⁺ 100 mM HEPES	19.8	49.5
C	0.1 U PGM-L 0.2 U IPS-L 1.0 U IMP-L 0.04 U <i>Mm</i> MIOX 0.3 U L-Udh	0.8	8 ^a	40 mM G1P, 5 mM MgCl ₂ 0.01 mM G16BP 56 mM NAD ^{+b} 100 mM HEPES	5.7 ± 1.1	14.3
D	0.1 U PGM-L _{Zeo} 0.2 U IPS-L _{Zeo} 1.0 U IMP-L _{Zeo} 0.04 U <i>Mm</i> MIOX 0.3 U L-Udh	2.0	8 ^c	100 mM G1P, 5 mM MgCl ₂ 0.01 mM G16BP 60 mM NAD ^{+d} 100 mM HEPES	6.7 ± 0.5	6.7

^a 5 hours with PGM-L, IPS-L and IMP-L plus three hours with *Mm*MIOX and L-Udh.

^b 4 mM NAD⁺ every two hours for the first five hours plus 40 mM NAD⁺ for the reaction with *Mm*MIOX and L-Udh.

^c 5 cycles of one hour each with co-immobilised PGM-L, IPS-L and IMP-L plus 3h with *Mm*MIOX and L-Udh.

^d 4 mM NAD⁺ per cycle plus 40 mM NAD⁺ for the reaction with *Mm*MIOX and immobilised L-Udh.

3.6 Analysis of the Udh reaction with LC-MS and NMR

In initial experiments with *Mm*MIOX and L-Udh, GlucA was not detectable by UHPLC-RID unless the samples were boiled for 15 min and subsequently mixed with trifluoroacetic acid (TFA) as described previously (Petroll, et al., Anal Chim Acta, submitted). LC-ESI-MS analysis revealed that the L-Udh reaction forms primarily the GlucA lactone ($m/z = 191$) (Fig 7, lower panel) and not the open form GlucA $m/z = 209$ (Fig 7, upper panel) as the intensity of GlucA was two orders of magnitude lower than the intensity of the lactone. Signal intensities cannot be compared directly in ESI-MS as distinct compounds ionise differently. Nevertheless,

the low intensity of GlucA in the MS may explain why no signal has been observed in the UHPLC-RID analysis as the RID exhibits a much lower sensitivity than MS. In addition, the GlucA lactone was observed to co-elute with HEPES and cannot be detected by the UHPLC-RID method as the sample buffer is 100 mM HEPES (data not shown).

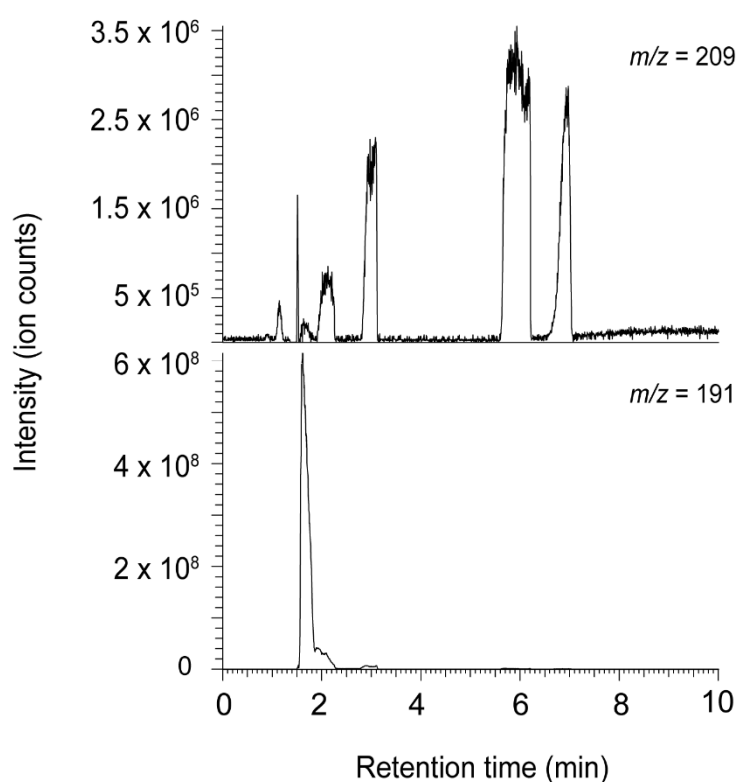


Figure 7. Extracted ion chromatograms at $m/z = 209$ (GlucA, upper panel) and 191 (GlucA lactone, lower panel) obtained on GlycanPac AXH-1 LC-ESI-MS analysis of the L-Udh reaction after incubation with GA and NAD^+ at 40 °C for 60 min.

NMR spectroscopy was performed on the Udh reaction (Udh without the linker) to validate that Udh produced predominantly the glucaro-1,4-lactone and not the open GlucA form and to investigate whether this effect is due to the presence of the linker or due to Udh itself (Fig. 8D). Heat-inactivated Udh mixed with GA and NAD^+ but without incubation at 40 °C was used as

a negative control (Fig. 8B). To investigate the processes that occur when boiling the Udh samples and mixing with TFA, Udh was first incubated with GA and NAD^+ at 40 °C, boiled for 15 min and mixed with TFA at a final concentration of 0.2 % vol. prior to NMR analysis (Fig. 8F). NMR signals were assigned using standards from GA (Fig. 8A), glucaro-1,4-lactone (Fig. 8C) and GlucA (Fig. 8E) and previous NMR studies on GA and GlucA.⁵⁴⁻⁵⁶

NMR analysis confirmed that GA was completely metabolised after 150 min and that the main product was the glucaro-1,4-lactone and not the open form GlucA. After boiling the samples, the lactone disappeared while signals corresponding to GlucA appeared in the spectrum. It has been suggested previously that as a result of boiling of the samples, glucaro-1,4-lactone is converted to the open form (Petroll, et al., Anal Chim Acta, submitted). It has been established that GlucA in solution is in an equilibrium with two different lactones and one dilactone and that acid catalyses the conversion of the lactones to the open form.⁵⁶ However, the finding that high amounts of the GlucA lactone and only little GlucA were formed in the Udh sample indicates a catalytic mode rather than equilibration. In agreement, the Udh from *A. tumefaciens* (which is promiscuous to G-galacturonic acid and D-glucuronic acid) was identified to produce the D-galactaric acid lactone and not the open form.⁵⁷ The lactone produced did not spontaneously open unless slightly basic conditions were provided in which it equilibrated slowly to the galactaric acid form over several days. Thus, a similar mechanism may occur with the Udh from *F. pelagi* as it also exhibits promiscuity towards D-galacuronate.⁴⁴

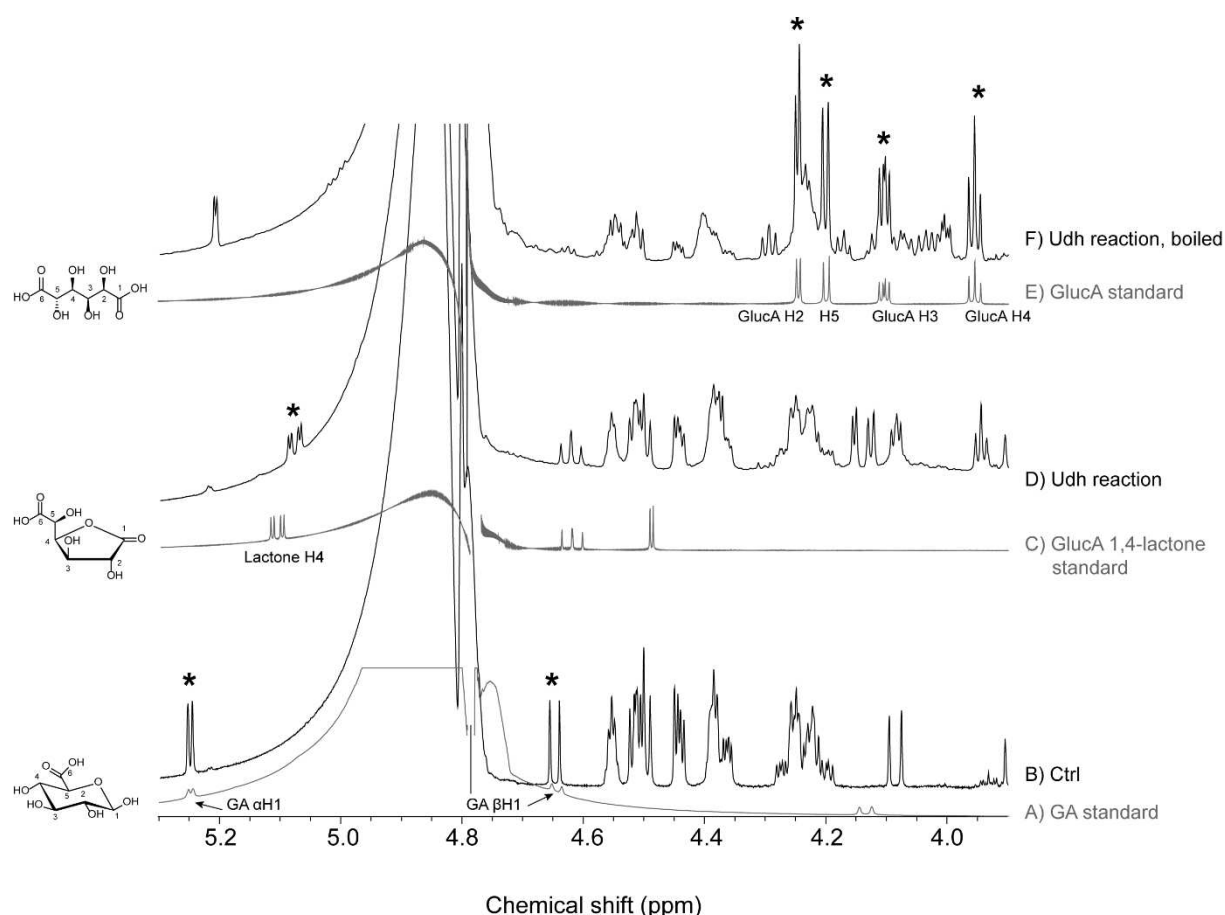


Figure 8. Portion of ^1H NMR spectra obtained from standards of GA (A), of glucaro-1,4-lactone (C) and of GlucA (E), and ^1H NMR spectra obtained from the reaction of Udh with GA and NAD^+ at the start of the reaction (time 0, negative control) (B), after incubation of the reaction at 40°C for 150 min (D) and after incubation of the reaction at 40°C for 150 min, boiling the samples and adding TFA prior to NMR analysis (F). The signals that were used for product identification are marked by asterisks.

Evidence for the GlucA-1,4-lactone being produced *in vivo* was originally established by the Marsh group.^{58,59} They found that the primary product from the catabolism of glucuronic acid lactone and glucuronic acid in the liver is the D-glucaro-1,4-3,6-dilactone which decomposes spontaneously in solution into the GlucA-1,4-lactone.^{56,59} What has been described earlier as a D-glucuronolactone dehydrogenase in fact may be the catalytic action of Udh (the former EC number of the D-glucuronolactone dehydrogenase was 1.1.1.70 which has been deleted and

incorporated in the class of aldehyde dehydrogenase (EC 1.2.1.3)). Of relevance is that these authors identified that boiling the liver extracts which contained the GlucA-1,4-lactone in slightly basic conditions resulted in removal of the lactone to a different compound.⁵⁸ This finding agrees with our observations and supports our view that boiling the samples under slightly basic conditions catalyses ring opening of the GlucA-1,4-lactone to GlucA (Fig. 8 D and F).

The finding that Udh expressed in crude *E. coli* cell lysates catalyses the formation of the GlucA lactone (and does not result in the open form) that does not open spontaneously is significant since many biocatalytic applications utilise Udh for the production of GlucA. For example, the enzyme glucarate dehydratase was employed for the cell-free production of α -ketoglutarate from GlucA and was assayed with the open GlucA form instead of the lactone.^{60,61} However, an earlier study indicated that the partially-purified glucarate dehydratase from *A. tumefaciens* was only catabolising the GlucA lactone and not the open form while cell-free extracts of *A. tumefaciens* were able to utilise both forms.⁶² Those results may suggest that the downstream enzyme of GlucA, the glucarate dehydratase, requires the open form GlucA and the Udh-derived GlucA lactone is converted to GlucA by endogenous lactonases. However, a GlucA lactonase has never been identified. Therefore, further investigation of whether endogenous enzymes present in *E. coli* or *S. cerevisiae* catalyse the ring opening of GlucA lactone is of considerable interest for all applications that rely on GlucA production as this observation affects product analysis, product purification and downstream enzyme activities.

4 Conclusion

We have shown the first cell-free production of GlucA using a synthetic six enzyme pathway. GlucA was synthesised to a final concentration of 5.7 ± 1.1 mM GlucA from G1P as a starting material after only eight hours reaction time. These results are similar to recent studies on the *in vivo* production of GlucA with respect to titre and yields obtained from glucose as starting material (Table 1). However, due to the short time that was required to achieve these titres in the cell-free system as compared to the *in vivo* system (fermentation was for ≥ 3 days^{14,46,63}) the productivity observed here is higher. Therefore, cell-free biocatalysis for the production of GlucA represents an effective alternative to microbially-based production systems.

We have demonstrated a framework that could improve the cost-effectiveness and efficiency of cell-free biocatalysis as shown by the immobilisation of the pathway for re-use of the enzymes and with the incorporation of a cofactor recycling system. The feasibility of long-term storage of all pathway enzymes which facilitated the rapid construction, testing and optimisation of the cell-free pathway has been shown. Flux analysis using a previously-developed UHPLC method (Petroll, et al., Anal Chim Acta, submitted) has allowed close monitoring of the effect of enzyme ratios and metabolite concentrations on the metabolic flux of the pathway. As a result, we have identified thermodynamic and kinetic bottlenecks for all enzymes of the upper part of the pathway that required delicate fine-tuning of enzyme amounts and reaction conditions. We have established that the final reaction of the pathway catalysed by Udh generates the GlucA lactone and not, as previously suggested, the open GlucA form. Whether endogenous enzymes are responsible for the lactonisation of the GlucA lactone in previous *in vivo* studies remains to be identified. This finding may have an impact on product analysis and downstream applications of the GlucA pathway both in cell-free and *in vivo* applications.

Compared to recent studies on cell-free biocatalysis, the final titre achieved is satisfactory, but yields are much lower than theoretically possible, which may be explained by the high degree of complexity of this pathway and the high flux control exerted by multiple enzymes. It is possible that the use of different enzyme variants may improve the flux. Another major bottleneck in the pathway was the MIOX enzyme. Three different variants were tested, MIOX from mouse, MIOX from wheat and a putative MIOX from *Rasamsonia emersonii* but all were found to be catalytically inefficient and unstable. Although Mouse MIOX was the best candidate, it exhibited the lowest activity compared to the other pathway enzymes and completely lost activity upon immobilisation to zeolite. Thus, MIOX was supplied into the reaction freshly and not recycled, which imposes a substantial burden on the process costs. Therefore, successful protein engineering of MIOX towards higher activity and stability will be a key driver for efficient and economic GlucA production in cell-free and whole cell processes. Expanded kinetic characterisation of all linker enzymes would facilitate kinetic modelling of the pathway to reduce enzyme usage and cofactor requirements.

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

KP is 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) and the Australian Research Council (CE140100003). We thank Atul Bhatnagara and Matthew J. McKay from the Australian Proteome Analysis Facility

(Macquarie University, Sydney, Australia) for their support in performing LC-MS experiments and Nicole Cordina from the Macquarie University NMR facility from the Department of Molecular Sciences for her help in performing NMR studies.

Supplementary material

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

References

1. G. Subramanian and G. Madras, *Water Research*, 2016, **104**, 168-177.
2. K. Fischer and H.-P. Bipp, *Water, Air, and Soil Pollution*, 2002, **138**, 271-288.
3. A. Abbadi, K. F. Gotlieb, J. B. M. Meiberg, J. A. Peters and H. Van Bekkum, *Green Chemistry*, 1999, **1**, 231-235.
4. M. D. W. and K. D. E., *Journal of Applied Polymer Science*, 2000, **77**, 3085-3092.
5. R. S. Koefod, U.S. Patent No. 7,658,861. 9.
6. C. Lu and E. Ford, *Macromolecular Materials and Engineering*, 2018, **303**, n/a-n/a.
7. D. E. Kiely, L. Chen and T. H. Lin, *Journal of the American Chemical Society*, 1994, **116**, 571-578.
8. E. Chiellini, S. Dantone and R. Solaro, *Macromolecular Symposia.*, 1997, **123**, 25-44.
9. T. R. Boussie, G. M. Diamond, E. Dias and V. Murphy, in *Chemicals and Fuels from Bio-Based Building Blocks*, Wiley-VCH Verlag GmbH & Co. KGaA, 2016, DOI: 10.1002/9783527698202.ch6, 151-172.
10. M. Hanausek, Z. Walaszek and T. J. Slaga, *Integrative Cancer Therapies*, 2003, **2**, 139-144.
11. M. Fittkau, W. Voigt, H.-J. Holzhausen and H.-J. Schmoll, *Journal of Cancer Research and Clinical Oncology*, 2004, **130**, 388-394.
12. G. Diamond, V. Murphy and T. R. Boussie. Bentham Science Publishers: Sharjah (UAE), 2014.
13. T. N. Smith, K. Hash, C. L. Davey, H. Mills, H. Williams and D. E. Kiely, *Carbohydr Res*, 2012, **350**, 6-13.
14. T. S. Moon, S. H. Yoon, A. M. Lanza, J. D. Roy-Mayhew and K. L. Prather, *Applied and Environmental Microbiology*, 2009, **75**, 589-595.
15. A. Pellis, S. Cantone, C. Ebert and L. Gardossi, *New Biotechnology*, 2018, **40**, 154-169.
16. J. D. Keasling, A. Mendoza and P. S. Baran, *Nature*, 2012, **492**, 188.
17. J. M. Sperl and V. Sieber, *ACS Catalysis*, 2018, **8**, 2385-2396.
18. R. A. Sheldon and D. Brady, *Chemical Communications*, 2018, **54**, 6088-6104.
19. Q. M. Dudley, A. S. Karim and M. C. Jewett, *Biotechnology Journal*, 2015, **10**, 69-82.
20. A. Sunna, F. Chi and P. L. Bergquist, *New Biotechnology*, 2013, **30**, 485-492.
21. P. Rehbein and H. Schwalbe, *Protein Expression and Purification*, 2015, **110**, 1-6.
22. N. Rashid, T. Kanai, H. Atomi and T. Imanaka, *Journal of Bacteriology*, 2004, **186**, 6070-6076.
23. A. A. Baykov, O. A. Evtushenko and S. M. Avaeva, *Analytical Biochemistry*, 1988, **171**, 266-270.
24. Y. K. Wang, A. Morgan, K. Stieglitz, B. Stec, B. Thompson, S. J. Miller and M. F. Roberts, *Biochemistry*, 2006, **45**, 3307-3314.
25. Y.-W. Zhang, M. K. Tiwari, H. Gao, S. S. Dhiman, M. Jeya and J.-K. Lee, *Enzyme and Microbial Technology*, 2012, **50**, 255-262.
26. C. Nowak, B. Beer, A. Pick, T. Roth, P. Lommès and V. Sieber, *Frontiers in Microbiology*, 2015, **6**.
27. A. Care, K. Petroll, E. S. Y. Gibson, P. L. Bergquist and A. Sunna, *Biotechnology for Biofuels*, 2017, **10**, 29.
28. R. A. Sheldon and S. van Pelt, *Chemical Society Reviews*, 2013, **42**, 6223-6235.
29. M. F. Chaplin and C. Bucke, *Enzyme Technology*, CUP Archive, 1990.
30. T. S. Moon, S.-H. Yoon, M.-J. Tsang Mui Ching, A. M. Lanza and K. L. J. Prather, *Analytical Biochemistry*, 2009, **392**, 183-185.
31. T. Fujisawa, S. Fujinaga and H. Atomi, *Applied and Environmental Microbiology*, 2017, DOI: 10.1128/aem.00550-17.

32. C. You, T. Shi, Y. Li, P. Han, X. Zhou and Y. P. Zhang, *Biotechnology and Bioengineering*, 2017, DOI: 10.1002/bit.26314.
33. A. Flamholz, E. Noor, A. Bar-Even and R. Milo, *Nucleic Acids Research*, 2012, **40**, D770-775.
34. J. K. Guterl, D. Garbe, J. Carsten, F. Steffler, B. Sommer, S. Reisse, A. Philipp, M. Haack, B. Ruhmann, A. Koltermann, U. Kettling, T. Bruck and V. Sieber, *ChemSusChem*, 2012, **5**, 2165-2172.
35. R. M. Daniel, M. J. Danson, D. W. Hough, C. K. Lee, M. E. Peterson and D. A. Cowan, *Protein Adaptation in Extremophiles*, 2008, 1-34.
36. K. Cheng, F. Zhang, F. Sun, H. Chen and Y. H. Percival Zhang, *Scientific Reports*, 2015, **5**, 13184.
37. L. Chen, C. Zhou, H. Yang and M. F. Roberts, *Biochemistry*, 2000, **39**, 12415-12423.
38. L. Chen and M. F. Roberts, *Applied and Environmental Microbiology*, 1999, **65**, 4559-4567.
39. T. Schwander, L. Schada von Borzyskowski, S. Burgener, N. S. Cortina and T. J. Erb, *Science (New York, N.Y.)*, 2016, **354**, 900-904.
40. A. J. Morgan, Y. K. Wang, M. F. Roberts and S. J. Miller, *Journal of the American Chemical Society*, 2004, **126**, 15370-15371.
41. T. S. Moon, J. E. Dueber, E. Shiue and K. L. J. Prather, *Metabolic Engineering*, 2010, **12**, 298-305.
42. R. J. Arner, K. S. Prabhu and C. C. Reddy, *Biochemical and Biophysical Research Communications*, 2004, **324**, 1386-1392.
43. A. Alok, H. Kaur, K. K. Bhati, J. Kumar, P. Pandey, S. K. Upadhyay, A. Pandey, N. C. Sharma, A. K. Pandey and S. Tiwari, *Plant Gene*, 2015, **4**, 10-19.
44. A. Pick, J. Schmid and V. Sieber, *Microbial Biotechnology*, 2015, **8**, 633-643.
45. N. Chen, J. Wang, Y. Zhao and Y. Deng, *Microbial Cell Factories*, 2018, **17**, 67.
46. A. Gupta, M. A. Hicks, S. P. Manchester and K. L. Prather, *Biotechnology Journal*, 2016, **11**, 1201-1208.
47. B. Jia, S.-C. Park, S. Lee, B. P. Pham, R. Yu, T. L. Le, S. W. Han, J.-K. Yang, M.-S. Choi, W. Baumeister and G.-W. Cheong, *FEBS Journal*, 2008, **275**, 5355-5366.
48. M. A. Nisar, N. Rashid, Q. Bashir, Q. T. Gardner, M. H. Shafiq and M. Akhtar, *Journal of Bioscience and Bioengineering*, 2013, **116**, 39-44.
49. M. L. Mavrovouniotis, *Chemical Engineering Science*, 1996, **51**, 1495-1507.
50. B. Geueke, B. Riebel and W. Hummel, *Enzyme and Microbial Technology*, 2003, **32**, 205-211.
51. E. Noor, A. Bar-Even, A. Flamholz, E. Reznik, W. Liebermeister and R. Milo, *PLOS Computational Biology*, 2014, **10**, e1003483.
52. R. M. Daniel and D. A. Cowan, *Cellular and Molecular Life Sciences : CMLS*, 2000, **57**, 250-264.
53. J. Geiger and X. Jin, in *Biology of Inositols and Phosphoinositides*, eds. A. L. Majumder and B. B. Biswas, Springer US, 2006, **39**, ch. 7, pp. 157-180.
54. T. Bezabeh, O. B. Ijare, N. Albiin, U. Arnelo, B. Lindberg and I. C. Smith, *Magma (New York, N.Y.)*, 2009, **22**, 267-275.
55. T. T. Denton, K. I. Hardcastle, M. K. Dowd and D. E. Kiely, *Carbohydrate Research*, 2011, **346**, 2551-2557.
56. J. M. Brown, M. Manley-Harris, R. J. Field and D. E. Kiely, *Journal of Carbohydrate Chemistry*, 2007, **26**, 455-467.
57. H. Boer, H. Maaheimo, A. Koivula, M. Penttila and P. Richard, *Applied Microbiology and Biotechnology*, 2010, **86**, 901-909.
58. C. Marsh, *Biochemical Journal*, 1963, **86**, 77-86.
59. C. A. Marsh, *Biochemical Journal*, 1966, **99**, 22-27.

- 60. B. Beer, A. Pick and V. Sieber, *Metabolic Engineering*, 2017, **40**, 5-13.
- 61. A. Pick, B. Beer, R. Hemmi, R. Momma, J. Schmid, K. Miyamoto and V. Sieber, *BMC Biotechnology*, 2016, **16**, 80.
- 62. Y. F. Chang and D. S. Feingold, *Journal of Bacteriology*, 1970, **102**, 85-96.
- 63. I. M. B. Reizman, A. R. Stenger, C. R. Reisch, A. Gupta, N. C. Connors and K. L. J. Prather, *Metabolic Engineering Communications*, 2015, **2**, 109-116.

4.4 Supplementary

A novel framework for the cell-free enzymatic production of glucaric acid

Supplementary information

Kerstin Petroll^a, Andrew Care^{a,b}, Peter L Bergquist^{a,b,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

*Corresponding author Anwar Sunna, anwar.sunna@mq.edu.au

S1 Plasmid construction

Plasmid construction and propagation was performed in *Escherichia coli* strain DH5 α which was cultivated at 37°C in Luria-Bertani (LB) medium supplemented with 50 mg/ml carbenicillin. DNA sequences for the genes for inositol-1-monophosphatase (IMP) from *Archaeoglobus fulgidus* DSM 4304 (protein sequence GenBank WP_010879859.1) and of myo-inositol-3-phosphate synthase (IPS, AF_RS09020) from *A. fulgidus* DSM 4304 (protein sequence GenBank WP_010879290.1) were synthesised with flanking restriction sites (*NdeI/SalI*) and optimised codon usage for expression in *E. coli* (GeneArt, Thermo Fisher Scientific, Massachusetts, USA). The coding sequence (CDS) of myo-inositol oxygenase from *Triticum aestivum* (*TaMIOX*) (GenBank accession number AK334637.1) was derived from the full-length open reading frame as described by (Alok, Kaur et al. 2015). The DNA sequences for the corresponding genes of *TaMIOX*, *MIOX* from *Mus musculus* (*MmMIOX*) (protein sequence GenBank NP_064361.2) and uronate dehydrogenase (*Udh*) from *Fulvimarina pelagi* HTCC2506 (protein sequence GenBank™ Q0FXI7) were synthesised as gBlocks® Gene Fragments (Integrated DNA Technologies, Illinois, USA) with flanking restriction sites (*BamHI/SalI*, *BamHI/SalI* and *NdeI/SalI*) and optimised codon usage for expression in *E. coli*. The CDS for NADH oxidase (*Nox*) from *Lactobacillus rhamnosus* ATCC 53103 (GenBank accession number AP011548.1) was derived from the genome sequence¹. The DNA sequences for the corresponding genes of *Nox* and phosphoglucomutase (*PGM*) from *Thermococcus kodakarensis* KOD1 (protein sequence GenBank BAD42440.1) were synthesised with flanking restriction sites (*NdeI/SalI*) and optimised codon usage for expression in *E. coli* (Integrated DNA Technologies).

For enzyme immobilisation, the genes encoding *PGM*, *IMP*, *IPS* and *Udh* were ligated into pET22b (Novagen, Merck, New Jersey, USA) plasmids with a C-terminal linker sequence (L)² to give the expression plasmids pET22b-*PGM*-L, pET22b-*IMP*-L, pET22b-*IPS*-L and

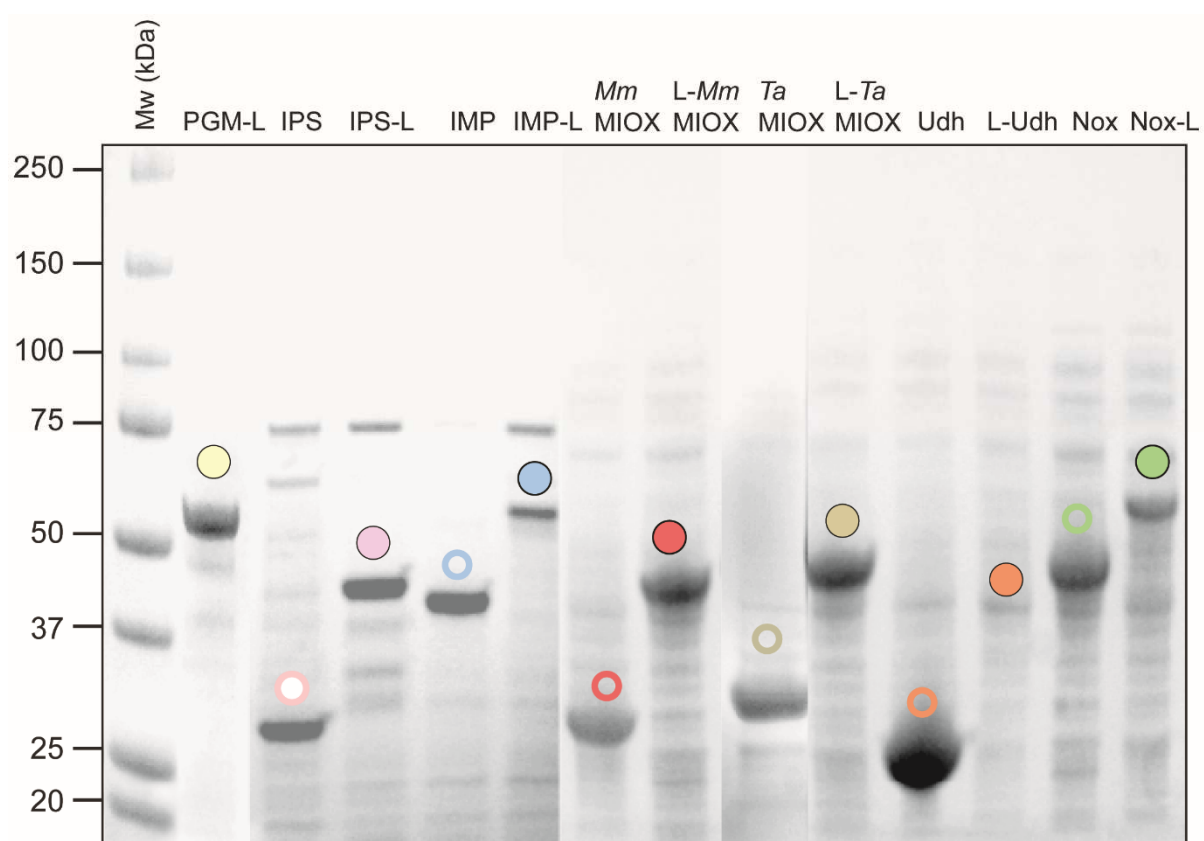
pET22b-Udh-L (Table 2). The genes encoding *TaMIOX*, *MmMIOX* and *Nox* were ligated into pET22b plasmids with an N-terminal L² to give the expression plasmids pET22b-L-*TaMIOX*, pET22b-L-*MmMIOX* and pET22b-L-*Nox* (Table 2). To obtain enzymes without the L, the genes that were genetically fused to the L at their C- or N-terminus were amplified by polymerase chain reaction (PCR) and flanked with restriction sites by using the primers listed in Table S1. The amplicons were ligated into pET22b plasmids without the L to give pET22b-IMP, pET22b-IPS, pET22b-Udh, pET22b-*MmMIOX*, pET22b-*TaMIOX* and pET22b-*Nox* (Table 2). The plasmid pET22b-L-Udh was obtained through amplification of pET22b-Udh and ligation into the pET22b plasmid with the N-terminal L using the primer listed in Table S1. All recombinant constructions were sequenced to validate their correctness (Macrogen, Seoul, Korea).

Supplementary Table 1. Primers used in this study. The restriction sites are underlined and stated behind the sequence. Start and stop codons are marked in bold.

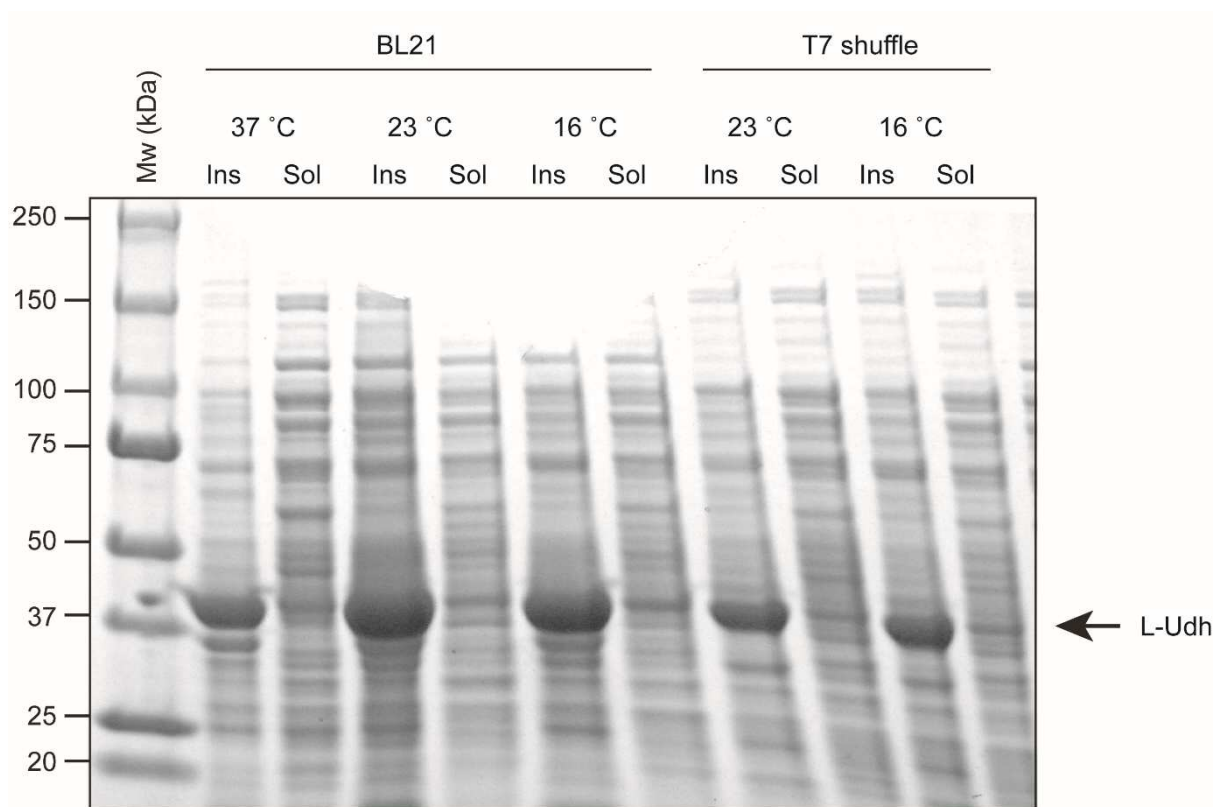
Primers	Nucleotide sequence (5' – 3') and (restriction site)
IPS-F	GATATAC <u>CATATG</u> AAGTCTGGTTAGTTGG (<i>NdeI</i>)
IPS-R	GTTTGA <u>AAGCTT</u> CTATTATTCAGGTTTCGAG (<i>HindIII</i>)
IMP-F	GATATAC <u>CATATG</u> GATGAAAGGGATGCG (<i>NdeI</i>)
IMP-R	GGTTTGGAATTCCTATTATTTAATCAGTTCCAGC (<i>EcoRI</i>)
MmMIOX-F	GGAATTCC <u>CATATG</u> AAGGTAGATGTCGGACCCGACCCATC (<i>NdeI</i>)
MmMIOX-R	GGAATTC <u>CGTCGACT</u> CACCAGCTAAGGGTTCCG (<i>Sall</i>)
TaMIOX-F	GGAATTCC <u>CATATG</u> ACTATTATCATCGAGCAGCCCCAATTG (<i>NdeI</i>)
TaMIOX-R	GGAATTC <u>CGTCGACT</u> CACCAACGTAATTTACCG (<i>Sall</i>)
Udh-F	GATATAC <u>CATATG</u> TTAGAGAACTTCTTATAACAGG (<i>NdeI</i>)
Udh-R	ATATGGATCCCTATTATGAGTGGATGGGCTC (<i>BamHI</i>)
L-Udh-F	CGAGGATCCGTTAGAGAACTTCTTATAACAGGGGCA (<i>BamHI</i>)
L-Udh-R	GGAATTC <u>CGTCGAC</u> CTATTATGAGTGGATGGGCTCCG (<i>Sall</i>)
Nox-F	GGAATTCC <u>CATATG</u> TTGATGAAAGTCGTCGTGGTAG (<i>NdeI</i>)
Nox-R	ATTCA <u>AAGCTT</u> CTATTACACTGCTGAATTATCGGCCTG (<i>HindIII</i>)

S2 Effect of the linker fusion on the solubility and activity of the L-enzymes

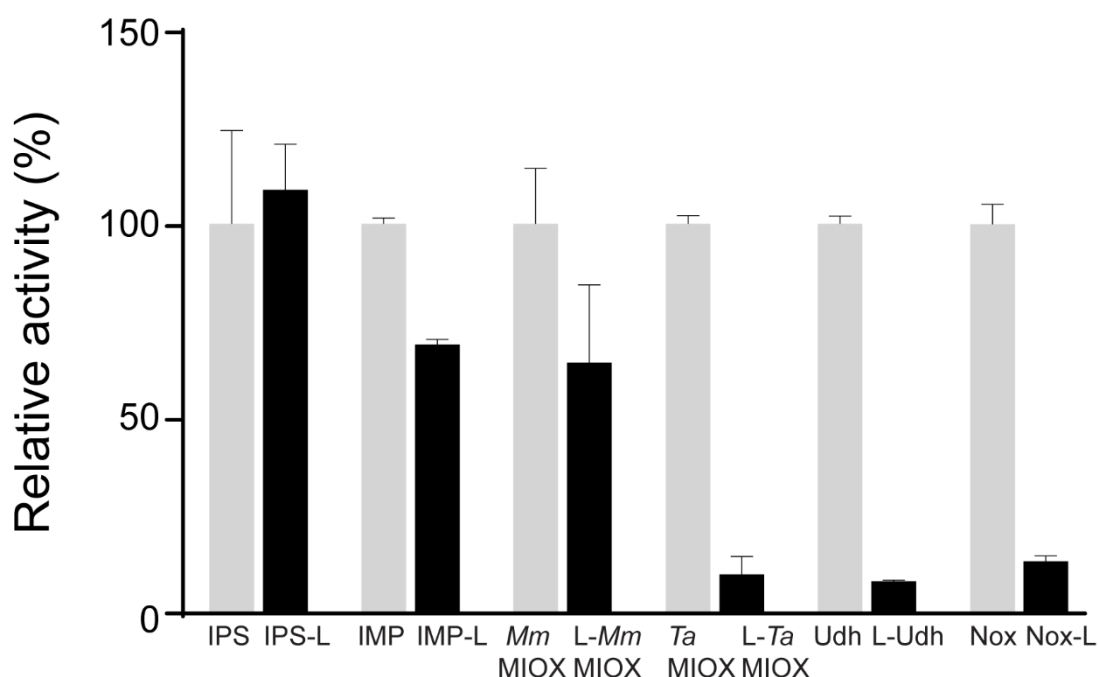
Experimental details can be found in the ‘Methods’ (Section 2.2 and 2.3) of the main publication.



Supplementary Figure 1. SDS-PAGE analysis of the effect of the linker enzyme fusion on the enzyme's solubility. Samples are soluble, partial (heat)- purified or crude cell lysates with enzymes (*empty circles*) or L-enzymes (*filled circles*).



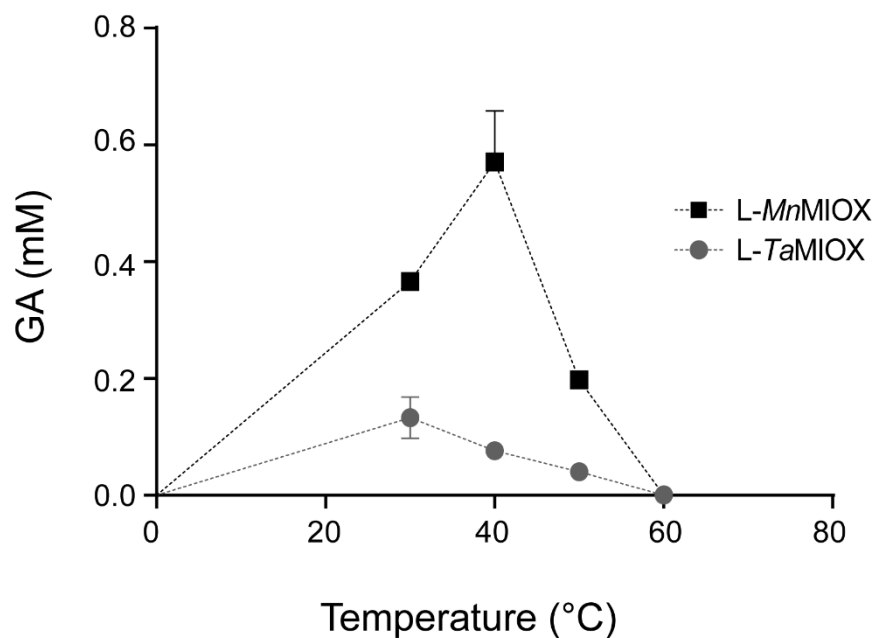
Supplementary Figure 2. SDS-PAGE analysis of the effect of expression conditions and expression host on L-Udh solubility. Protein expression was induced with 0.4 mM IPTG at 37°C for 4h, 23°C for 18h or 16°C for 19h. Expression hosts were either *E. coli* BL21 or *E. coli* T7 shuffle express cells which provide additional assistance to the folding of proteins due to increased chaperone activity. After protein expression, cells were resuspended in 50 mM phosphate buffer pH 8 supplemented with 1 mg/ml lysozyme and 25 U Benzonase and incubated for 1h on ice. After separation of insoluble (Ins) and soluble (Sol) cell extracts by centrifugation at 11,000 x g for 20 min at 4°C, the samples were subjected to SDS-PAGE. The molecular mass of L-Udh is 39kDa.



Supplementary Figure 3. Relative activities of enzymes (without linker) (white bars) and L-enzymes (black bars) under the same reaction conditions as described in the Methods section. Enzyme samples were partially purified (IPS, IPS-L, IMP, IMP-L) or from crude cell lysates (MmMIOX, L-MmMIOX, Udh, L-Udh, Nox, Nox-L) and the activities were relative to total protein amount. The graph shows the averages and standard deviations from three independent replicates. Relative activity (%) of **IPS**, 100.0 ± 24.1 ; **IPS-L**, 108.8 ± 11.7 ; **IMP**, 100.0 ± 1.5 ; **IMP-L**, 108.8 ± 11.7 ; **MmMIOX**, 100.0 ± 14.3 ; **L-MmMIOX**, 64.3 ± 20.0 ; **TaMIOX**, 100.0 ± 2.2 ; **L-TaMIOX**, 9.9 ± 4.6 ; **Udh**, 100.0 ± 2.1 ; **L-Udh**, 8.1 ± 0.3 ; **Nox**, 100.0 ± 5.0 ; **Nox-L**, 13.2 ± 1.5 .

S3 Temperature optima of L-*Mm*MIOX and L-*Ta*MIOX

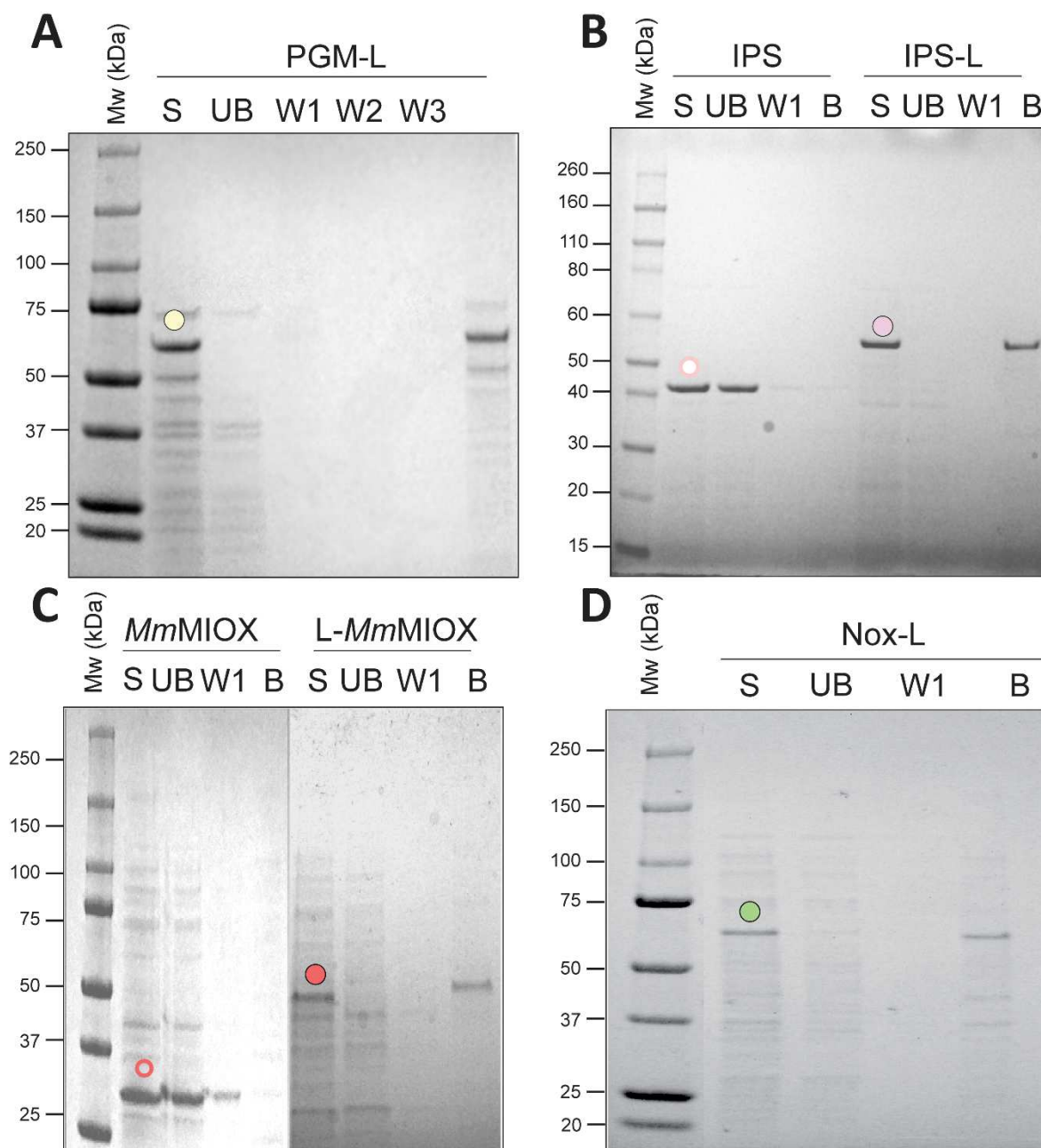
Experimental details can be found in the Methods (Section 2.3) of the main publication.



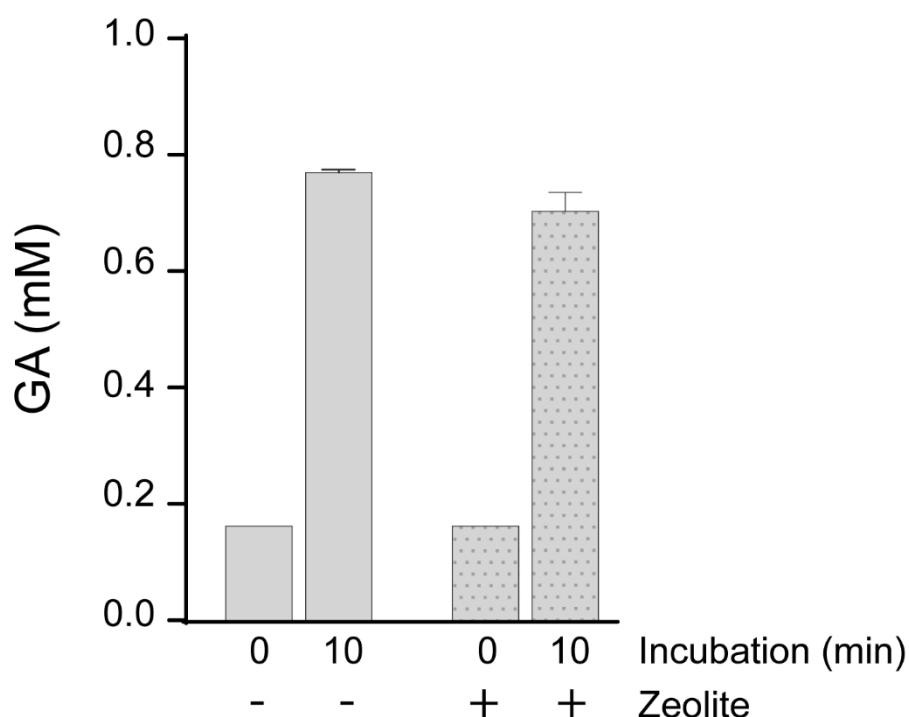
Supplementary Figure 4. Effect of temperature on the enzyme activity of L-*Mm*MIOX and L-*Ta*MIOX. L-*Mm*MIOX and L-*Ta*MIOX crude lysates were incubated with *myo*-inositol at temperatures ranging from 30 to 60 °C in 50 mM phosphate buffer, pH 8 under the same reaction conditions and protein concentration (0.6 mg/ml total protein). The mean and standard deviation of two individual replicates are shown.

S4 Enzyme immobilisation

Experimental details on the immobilisation of L-enzymes can be found in the Methods (Section 2.4) of the main publication.



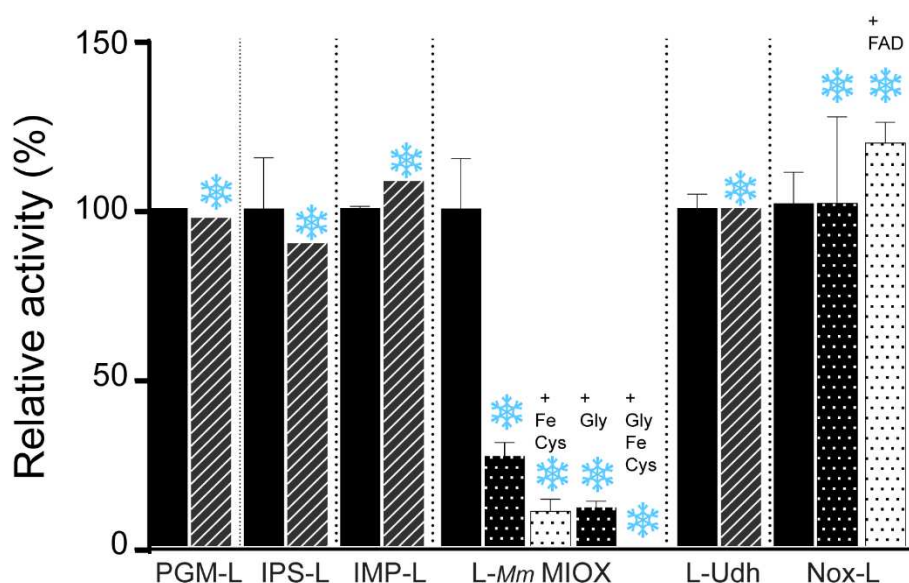
Supplementary Figure 5. SDS-PAGE analysis showing the zeolite-binding capacity of enzymes (without linker) (*empty circles*) and L-enzymes (*filled circles*) of PGM-L (A), IPS and IPS-L (B), MmMIOX and L-MmMIOX (C) and Nox-L (D). Standard zeolite binding assays were performed with partially-purified proteins of PGM-L, IPS and IPS-L and with crude lysates of MmMIOX, L-MmMIOX and Nox-L as described in the Methods section of the main publication. S, starting protein sample, UB, unbound protein fraction, W1, wash fractions (1 of 3 total), B, zeolite-bound protein fraction.



Supplementary Figure 6. Effect of zeolite on *Mm*MIOX activity. Crude lysates containing *Mm*MIOX (without the linker) were incubated with *myo*-inositol in presence (+) or absence (-) of zeolite at 40 °C for 0 (control) and 10 min. The reaction mixture (190 μ l) contained 76 mM *myo*-inositol, 150 μ g of *Mm*MIOX and 5 mg zeolite. The reaction was incubated at 40 °C and activity was determined using the end point method described in the the Methods section of the main publication. As a control, the enzyme was incubated with *myo*-inositol but without the addition of zeolite.

S5 Long-term storage of L-enzymes

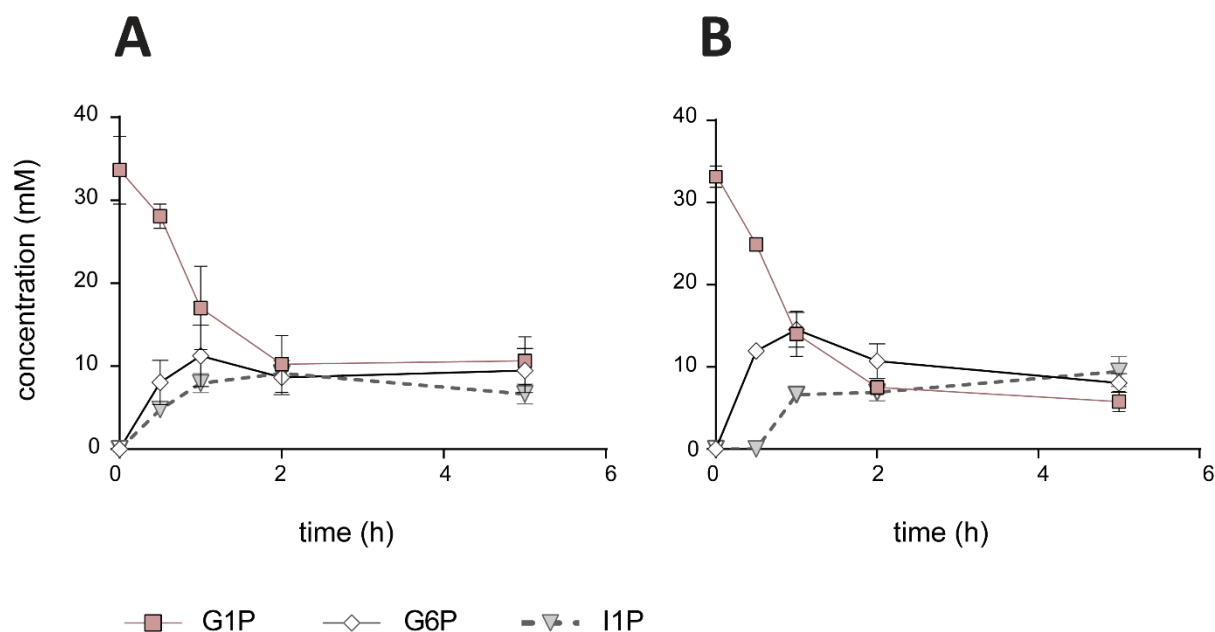
The effect of storing enzymes that were fused genetically to the linker (L-enzymes) at -80 °C was investigated by testing the enzyme activities directly after cell lysis and after the enzymes had undergone one freeze-thaw cycle at -80 °C. They were stored in 50 mM HEPES-NaOH buffer, pH 8 (pH 6.5 for Nox-L) in 1.5 ml Eppendorf tubes unless otherwise stated and placed in a -80 °C freezer (Thermofisher Scientific) in 0.2 mL portions. After 24 h, the enzymes were thawed by placing the tube in water at room temperature and enzyme activities were determined as described in the Methods (section 2.3) of the main publication.



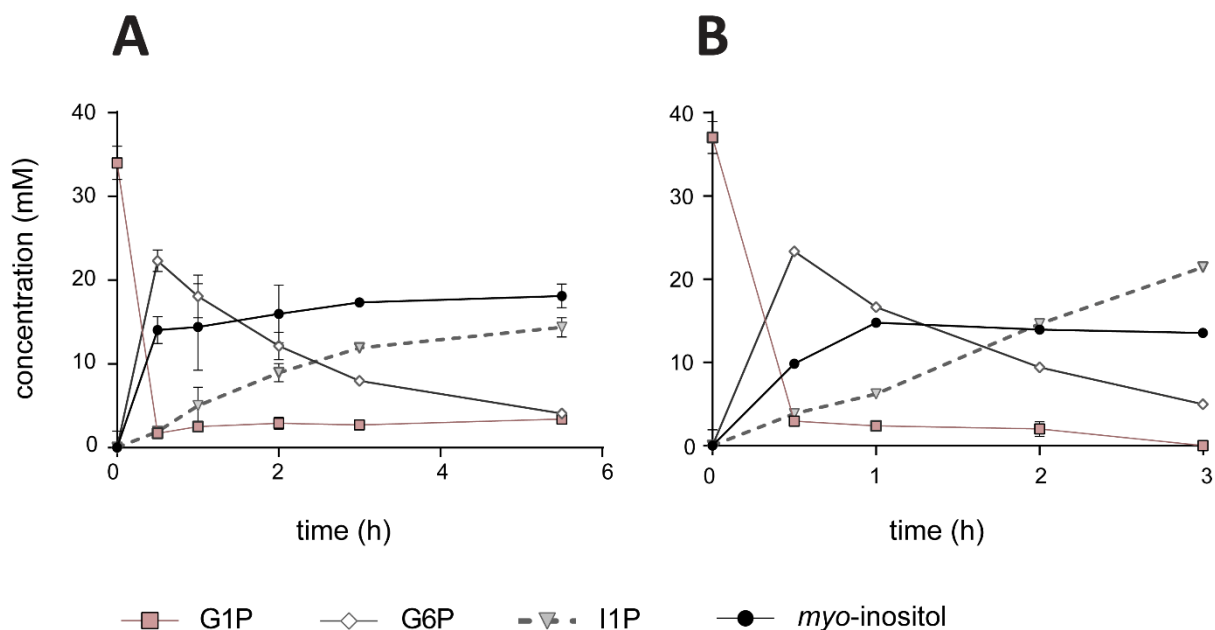
Supplementary Figure 7. Relative activities (%) of L-enzymes after cell lysis (black bars) and after one freeze/thaw cycle at -80°C (shaded bars with ice flake) under the same reaction conditions as described in the Methods section. L-*MmMIOX* was stored at -80°C as a crude cell lysate, as crude lysate plus presence of cofactors ferrous and L-cysteine (+ Fe, Cys), as a crude lysate diluted to 50% vol. glycerol (+ Gly), or as crude lysate with 50% vol. glycerol, Fe and Cys (+ Gly, Fe, Cys). Nox-L was stored as crude lysate or as crude lysate in presence of the coenzyme flavin adenine dinucleotide (+ FAD). Single values or the averages and standard deviations from three independent replicates are shown. Relative activities (%) of -80°C stored L-enzymes relative to fresh L-enzymes were: PGM-L, 97.14; IPS-L, 89.7; IMP-L, 108.085; L-*MmMIOX* (crude cell lysate, 26.7 ± 3.9 ; crude cell lysate + Fe, Cys, 10.3 ± 3.7 ; crude cell lysate + Gly, 11.5 ± 1.8 and crude cell lysate + Gly, Fe, Cys, 0.3 ± 0.1); L-Udh, 100; Nox-L (crude cell lysate, 101.4 ± 25.6 ; crude cell lysate + FAD, 119.4 ± 6.0).

S6 Optimising the production of *myo*-inositol from glucose-1-phosphate (G1P) from free enzymes and immobilised enzymes

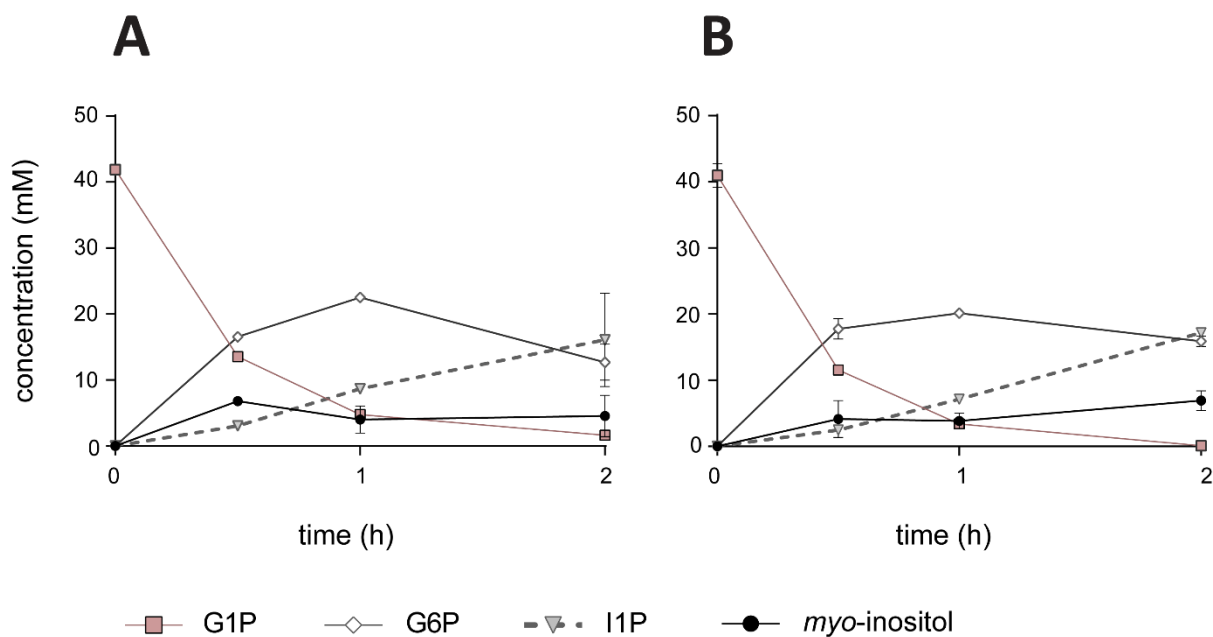
Experimental details can be found in the Methods (Section 2.6) of the main publication.



Supplementary Figure 8. Effect of the enzyme ratios of PGM-L and IPS-L on the cell-free synthesis of *myo*-inositol from glucose-1-phosphate (G1P). (A) 0.1 U PGM-L and 0.1 U IPS-L, or (B) 0.025 U PGM-L and 0.1 U IPS-L were incubated with 40 mM *myo*-inositol, 0.5 mM MnCl_2 and 10 μM glucose-1,6-bisphosphate (G1,6BP) at 85 °C in 100 mM HEPES buffer, pH 8 (total volume 400 μl). 4 mM NAD^+ was added into the reaction every two hours. G1P (squares), glucose-6-phosphate (G6P, diamonds) and inositol-1-phosphate (I1P, triangles) were measured by UHPLC-RID and *myo*-inositol (circles) was calculated as the difference between the amount of G1P added and the sum of all measured metabolites. The averages and standard deviation from two technical replicates are shown. I1P production after 5h in (A) was 6.6 ± 1.2 mM, and after 3h in (B) was 9.4 ± 1.8 mM.



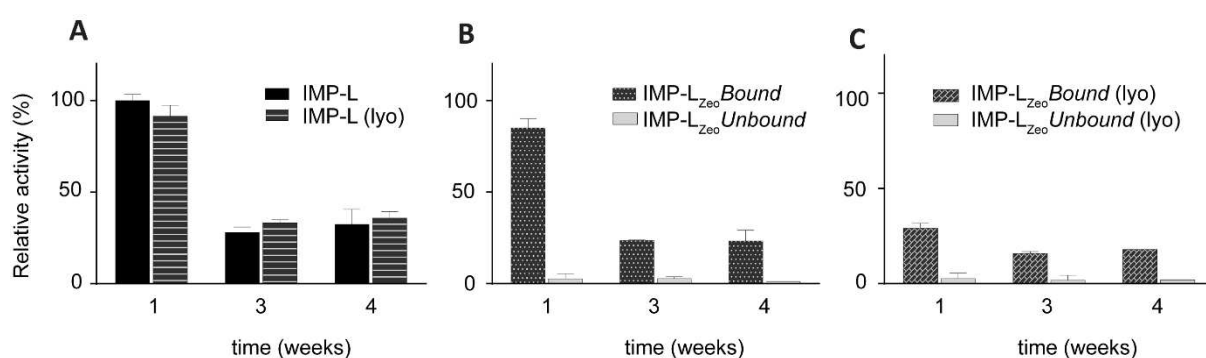
Supplementary Figure 9. Effect of increasing IMP-L concentration on the cell-free synthesis of *myo*-inositol from glucose-1-phosphate (G1P). (A) 0.1 U PGM-L, 0.2 U IPS-L and 0.1 U IMP-L and (B) 0.1 U PGM-L, 0.2 U IPS-L and 0.4 U IMP-L were incubated with 40 mM *myo*-inositol, 5 mM MgCl₂ and 10 μ M G1,6BP at 85 °C in 100 mM HEPES buffer, pH 8 (total volume 400 μ l). 4 mM NAD⁺ was added to the reaction every two hours. G1P (squares), glucose-6-phosphate (G6P, diamonds) and inositol-1-phosphate (I1P, triangles) were measured by UHPLC-RID and *myo*-inositol (circles) was calculated as the difference between the amount of G1P added and the sum of all measured metabolites. The averages and standard deviation from two technical replicates are shown. *myo*-inositol production after 5h in (A) was 18.1 ± 1.4 mM, and after 3h in (B) was 13.6 ± 0.4 mM.



Supplementary Figure 10. Effect of single (A) or co-immobilisation (B) of PGM-L, IPS-L and IMP-L onto the cell-free synthesis of *myo*-inositol. (A) 0.1 U PGM-L, 0.2 U IPS-L and 0.6 U IMP-L were singly immobilised onto 10, 10, and 30 mg zeolite, respectively and subsequently mixed for *myo*-inositol synthesis from 40 mM G1P at 85 °C. (B) 0.1 U PGM-L, 0.2 U IPS-L and 0.6 U IMP-L were co-immobilised onto 50 mg zeolite and incubated with 40 mM G1P at 85 °C. Both reactions were composed of 5 mM MgCl₂ and 10 μM G1,6BP and 100 mM HEPES buffer, pH 8 in a total volume of 400 μl without taking into account the volume contributed by the zeolite-L-enzymes. 4 mM NAD⁺ was added into the reaction every two hours. G1P (squares), glucose-6-phosphate (G6P, diamonds) and inositol-1-phosphate (I1P, triangles) were measured by UHPLC-RID and *myo*-inositol (circles) was calculated as the difference between the amount of G1P added and the sum of all measured metabolites. The averages and standard deviation from two technical replicates are shown. *myo*-inositol production after 2h in (A) was 4.5 ± 3.2 mM, and after 3h in (B) was 6.9 ± 1.5 mM.

S7 Long-term storage of immobilised enzymes

Whether the zeolite-bound enzymes can be stored in solution (Fig.S11 B, IMP-L_{Zeo}) or lyophilised (Fig.S11 C, IMP-L_{Zeo} (lyo)) at 4 °C without disintegration of the zeolite binding was tested with IMP-L only as an example. Non-immobilised IMP-L and non-immobilised, lyophilised IMP-L (Fig.S11 A) were used as controls. The test of activities during the four weeks of storage revealed that IMP-L in solution and lyophilised and immobilised IMP-L lost about 70% of the initial activity during the first week and then remained stable over the following three weeks, while the lyophilised immobilised IMP-L lost 70% of its initial activity during the lyophilisation process but sustained activity over the following four weeks. The minor activity in the unbound fractions observed throughout the experiment indicated that IMP-L binding to zeolite was highly stable regardless of being in solution or lyophilised.



Supplementary Figure 11. Stability of IMP-L and immobilised IMP-L when stored in solution (A, IMP-L and B, IMP-L_{Zeo}) or lyophilised (C, IMP-L (lyo) and D, IMP-L_{Zeo} (lyo)) at 4 °C over the course of four weeks. Inositol-1-monophosphate (IMP) activity was measured in the partial purified cell extracts of IMP-L in solution or lyophilised (A), in the bound fraction (patterned bars) and in the unbound fractions (plain bars) (B and C) and was normalised to the initial activity measured in the IMP-L sample at week 1. The averages and standard deviations from two independent replicates are shown.

References

1. Y.-W. Zhang, M. K. Tiwari, H. Gao, S. S. Dhiman, M. Jeya and J.-K. Lee, *Enzyme and Microbial Technology*, 2012, **50**, 255-262.
2. A. Sunna, F. Chi and P. L. Bergquist, *New Biotechnology*, 2013, **30**, 485-492.

Chapter 5:

Characterisation of a putative *myo*-inositol oxygenase from the moderately thermophilic fungus *Rasamsonia emersonii*

5.1 Introduction

Myo-inositol oxygenase (MIOX, EC 1.13.99.1) is a highly selective non-haem-iron oxidoreductase that catalyses the conversion of *myo*-inositol to D-glucuronic acid (GA) using molecular oxygen. This reaction combines *myo*-inositol and GA as the second to last step in the cell-free synthetic glucaric acid (GlucA) pathway described in Chapter 4. Unfortunately, the MIOX enzymes described so far have displayed low stability and low catalytic efficiency, impeding their biochemical characterisation and presenting a major bottleneck for *in vivo* and cell-free production of GlucA [1-3].

MIOX is highly conserved in eukaryotes (Fig. 1) and several representatives have been isolated from a range of sources including fungi (*Cryptococcus neoformans* [4] and *Cryptococcus lactativorus* [5]), plants (*Arabidopsis thaliana* [6] and *Triticum aestivum* L. [7] and mammals (mouse, rat, pig and human [3, 8]). Genomic sequencing of various bacterial species has revealed a number of putative MIOX sequences but the annotation of bacterial MIOX remains mostly hypothetical (Fig. 1). MIOX variants can be expressed heterologously in *Escherichia coli* and exhibit optimal catalytic temperatures at around 30 °C, but consistently display low solubility and poor stability in solution. To circumvent this instability, the functionality of MIOX is often characterised by performing enzyme activity assays with the whole cells in which it is expressed (e.g. *E. coli*) [4]. MIOX solubility can be enhanced also by the addition of solubility tags (i.e. protein fusion tags), for example mouse MIOX (*Mm*MIOX) was fused to a small ubiquitin-like modifier (SUMO) protein tag which dramatically increased its soluble expression in *E. coli*, resulting in a ~2-fold improvement in the MIOX-catalysed production of GA *in vivo* [9]. In contrast, the only report describing the successful protein engineering of *Mm*MIOX through directed evolution reported only a 1.3 improvement of cell-based MIOX enzyme activity [9].

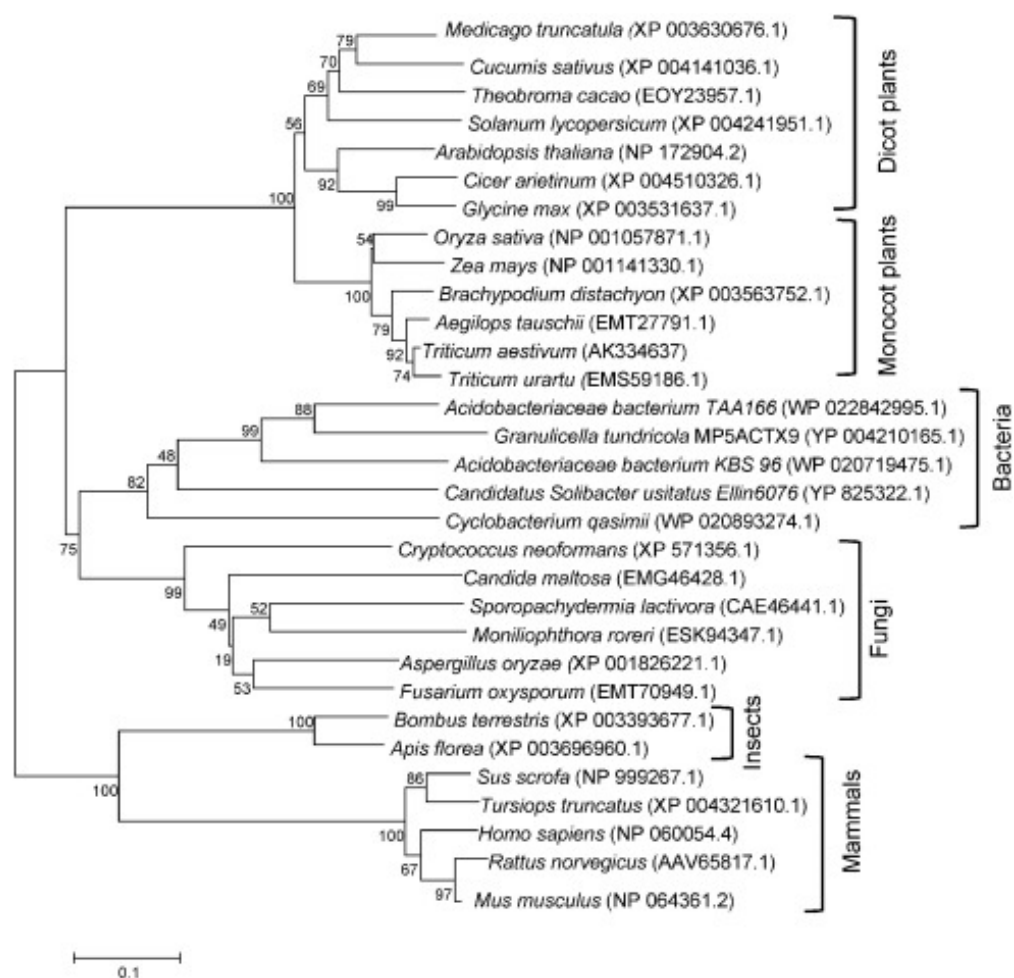


Figure 1. Phylogenetic tree of MIOX protein sequences from various species, including both characterised and hypothetical proteins [7].

The synthetic cell-free pathway presented in Chapter 4 was assembled using predominantly thermostable enzymes. The rationale of this approach was to exploit the intrinsic properties of thermostable enzymes to improve the stability and robustness of the system. Thermostable enzymes are considered good candidates compared to their mesophilic analogues for the assembly of the cell-free synthetic pathway because they often exhibit higher stability, remain active at higher process temperatures which improves reactant solubility and mass transfer in favour of higher reaction rates and can be partially purified by a simple heat denaturation step [10-13]

In the course of this study, we identified a putative MIOX gene from the moderately thermophilic fungus *Rasamsonia emersonii* (formerly *Talaromyces emersonii* [14] which is known to produce a variety of thermostable enzymes [14-20]. The open-reading frame (ORF) of the putative gene was discovered via homology modelling and its sequence displayed 44 to 65 % sequence similarity to previously characterised MIOX's. The putative gene, which was predicted to encode the *R. emersonii* MIOX (*ReMIOX*), was expressed successfully in *E. coli* with (*L-ReMIOX*) and without (*ReMIOX*) an N-terminal silica-binding peptide (linker). The linker was incorporated to facilitate the immobilisation of the recombinant fusion protein onto silica-based materials. The heterologous expression of both *ReMIOX* and *L-ReMIOX* was optimised in *E. coli* and their enzymatic activities were evaluated and compared to that of *MmMIOX* (see Chapter 3) and the fungus *C. neoformans* [4]. However, conflicting results from two different activity assays and the low solubility observed for both enzymes hampered their biochemical characterisation. Based on the results of this study, we suggest that improving *ReMIOX* solubility potentially would enable its characterisation and future applications.

5.2 Material and Methods

5.2.1 Reagents, vectors, strains and growth conditions

All reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and all *E. coli* strains were purchased from New England Biolabs (Ipswich, MA, USA), unless stated otherwise.

E. coli DH5 α was used for plasmid construction and propagation and was cultivated at 37°C in Luria-Bertani (LB) medium supplemented with 50 mg/ml carbenicillin. *E. coli* BL21(DE3) and SHuffle® T7 Express Competent cells were used as expression hosts for the production of *ReMIOX* and *L-ReMIOX*. Plasmids pET22b (Novagen, Merck, New Jersey, USA) and L-pET22b expressing a N-terminal silica-binding peptide sequence (L) [21] were used for sub-cloning and expression in *E. coli*.

5.2.2 Cloning of MIOX

The ORF of the putative MIOX gene from *R. emersonii* (ReMIOX) (GenBank accession number XP_013327840) was codon optimised for expression in *E. coli* and synthesised as a gBlocks® Gene Fragments (Integrated DNA Technologies, Illinois, USA) with flanking restriction sites (*NdeI/BamHI* or *NdeI/HindIII*) for insertion into pET22b or L-pET22b to give the expression plasmids pET22b-ReMIOX and pET22b-L-ReMIOX. The integrity of the plasmid constructions was confirmed by DNA sequencing (Macrogen, Seoul, Korea).

5.2.3 Identification and analysis of the putative MIOX sequence

A search of annotated MIOX gene sequences was conducted using the Pfamseq database (<http://pfam.xfam.org/family/PF05153/alignment/ncbi>). The protein sequence identified from the putative *R. emersonii* MIOX was derived from the National Center for Biotechnology Information (NCBI) protein database (www.ncbi.nlm.nih.gov). ReMIOX and L-ReMIOX protein sequences were analysed *in silico* using different bioinformatics tools. ExPASy proteomics server (<http://ca.expasy.org>) was used to compute the theoretical molecular masses (Mw) and isoelectric points (pI) of ReMIOX and L-ReMIOX, and Signal P 4.1 server (<http://www.cbs.dtu.dk/services/SignalP/>) was used for signal peptide predictions. Hypothetical or characterised MIOX protein sequences from animal, plant, bacteria and fungi species were retrieved from the NCBI protein database and compared to ReMIOX by Multiple Sequence Alignment using the ClustalW program (ClustalW - Multiple Sequence Alignment (EBI, United Kingdom), www.ebi.ac.uk/Tools/msa/clustalw2).

5.2.4 Recombinant expression of ReMIOX and L-ReMIOX

Freshly transformed *E. coli* BL21(DE3) or SHuffle® T7 Express Competent cells harbouring the pET22b-ReMIOX or pET22b-L-ReMIOX plasmids were cultured in LB medium supplemented with 50 mg/ml carbenicillin for the production of ReMIOX (with and without

the Linker). Cells were grown at 37 °C with shaking (250 rpm) until the A₆₀₀ was approximately 0.6 and recombinant protein synthesis was induced by the addition of 0.4 mM isopropyl β-d-thiogalactoside (IPTG). Cells were cultivated at either 16, 30 or 37 °C for 24h, 4h or 5h, respectively. Samples (200 µl) were collected each hour and analysed by SDS-PAGE on NuPage 4–12% Bis–Tris Gels (ThermoFisher Scientific, California, USA) and visualised with Coomassie Brilliant Blue staining. The expression medium was supplemented with *myo*-inositol at a final concentration of 0, 20, 60 or 80 mM. Cells were harvested by centrifugation at 4000 x g for 20 min at 4 °C and stored at -30 °C. For cell lysis, cells were resuspended in ice-cold 50 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris)-HCL, pH 7.0 supplemented with 1 mM glutathione and 0.05 mM of the serine protease inhibitor Pefabloc and then disrupted by three passages through a French pressure cell. The crude cell lysates (total protein fractions) obtained were incubated with 25 U Benzonase (Novagen, Merck) for 20 min at 4 °C and cellular debris was then removed by centrifugation at 11,000 × g for 20 min at 4 °C to yield soluble crude lysates. Total, soluble and insoluble protein fractions were analysed by SDS-PAGE and Coomassie staining. All protein concentrations were determined by infrared spectroscopy (Direct Detect, Merck).

5.2.5 Protein refolding

Insoluble protein fractions (5 mg) were resuspended in either non-denaturing reagent (Tween 100) or denaturing reagents; guanidinium chloride (GdnHCl) or urea. For solubilisation in Tween 100, insoluble proteins were washed with 0.75% potassium chloride and solubilised overnight at 4 °C with either 200 mM Tris-HCL buffer, pH 8 or 200 mM Tris-HCL buffer, pH 8 supplemented with 1% (v/v). Tween 100. For solubilisation in denaturing reagents, insoluble proteins were resuspended in 6 M GdnHCl or 8 M Urea, with 5 mM ethylenediaminetetraacetic acid (EDTA), 1 mM Pefabloc, 1 mM dithiothreitol (DTT) and 50 mM Tris-HCL, pH 7.0, and were then incubated at RT for 1h. The denatured, solubilised proteins were collected by

centrifugation (11, 000 x g, 30 min, 4 °C) and ‘refolded’ by dropwise addition of the sample into a 10 or 100 times higher volume of refolding buffer (0.25 M L-arginine, 10% sucrose, 1 mM EDTA and 0.1 mM Pefabloc) and incubation overnight at 4 °C. Alternatively, refolding was performed by dropwise dilution (1:10 v/v) in refolding buffer followed by overnight dialysis (Slide-A-Lyzer™ Dialysis Cassettes, 3.5K MWCO, ThermoFisher Scientific, California, USA) against 400 ml of refolding buffer with constant stirring at 4 °C. Any accumulated precipitate was removed by centrifugation and the remaining soluble protein was purified using gel filtration columns (PD MidiTrap G-25, GE Healthcare, Illinois, United States) and concentrated using an Amicon Ultra-10 centrifugal ultrafiltration unit (10 kDa cut-off, Merck, Darmstadt, Germany).

5.2.6 Enzyme activity measurements

Myo-inositol oxygenase activity was determined spectrophotometrically in whole cell extracts or protein fractions (total, insoluble and soluble) using a phenyl-phenol based colorimetric assay previously described by [22]. *ReMIOX* or *L-ReMIOX* were purified from protein fractions by size exclusion using Illustra NAP-25 Columns (GE Healthcare) and were concentrated on an Ultra-10 centrifugal ultrafiltration unit (10 kDa cut-off) prior to activity measurements. *ReMIOX* was activated by incubation with 1 mM ferrous sulphate (Fe) and 2 mM L-cysteine (Cys) for 15 min at room temperature before activity measurements. The final reaction mixture (300 µl) contained 60 mM *myo*-inositol, 1 mM glutathione and 50 mM Tris-HCL, pH 7.0 and the reactions were initiated by the addition of 100 µl whole cell extract or protein fractions. The reaction mixture was incubated at 20, 45 or 60 °C for between 1 and 4h. After incubation, 100 µl of sample or standard were mixed with 600 µl of 12.5 mM sodium tetraborate in 100% vol. sulphuric acid. The mixture was boiled at 100°C for 5 min to enable the complex formation of GA and tetraborate and was placed on ice and placed on ice. After chilling the sample, 10 µl of 0.5% NaOH (blank) or 0.15% m-phenylphenol in 0.5% NaOH

were added and 200 μ l of the resulting mixture were transferred into a 96-well plate and the absorbance of the pinkish-coloured solution was measured at 520 nm. Negative controls were performed with *E. coli* BL21 and *E. coli* T7 shuffle cells harbouring pET22b plasmids without the *ReMIOX* gene.

Alternatively, *ReMIOX* and L-*ReMIOX* activity were determined by measuring GA with the D-Glucuronic/D-Galacturonic Acid Assay Kit (Megazyme, Wicklow, Ireland). The assays were performed as described above and the reaction was stopped by adding 0.025% vol. perchloric acid to each sample. After neutralisation of the reaction mixture with 1 M potassium hydroxide, the samples were analysed according to the manufacturer's instructions.

5.3 Results and Discussion

5.3.1 Identification of a putative MIOX from *R. emersonii*

A database search of annotated MIOX genes resulted in the discovery of a putative MIOX protein sequence (A0A0F4YUD1) from *R. emersonii* (formerly *T. emersonii*) [14]. *R. emersonii* is a thermophilic filamentous fungus that inhabits natural soil and compost heaps and has been reported to produce several thermostable enzymes optimally active between 60 and 80 °C [14-20].

When compared to protein sequences of MIOX from various sources (Table 1), the putative MIOX from *R. emersonii* (*ReMIOX*) displayed sequence similarity to MIOX enzymes from the fungus *C. neoformans* (65%), the plant *Arabidopsis thaliana* (50%), *Homo sapiens* (human) (45%) and *Mus musculus* (mouse) (44%). As with other MIOX protein sequences, no putative signal peptide sequence was predicted in *ReMIOX*.

Table 1. The percentage identity matrix for the protein sequence alignment of *ReMIOX* (6) with MIOX protein sequences from *Homo sapiens* (1), *Mus musculus* (2), *Cyclobacterium qasimii* (3), *Arabidopsis thaliana* (4) and *Cryptococcus neoformans* (5).

	1	2	3	4	5	6
1 <i>Homo sapiens</i>	100	89.82	48.01	45.04	46.62	45.00
2 <i>Mus musculus</i>	89.82	100	47.65	44.33	46.26	34.57
3 <i>Cyclobacterium qasimii</i> (putative)	48.01	47.65	100	47.95	50.69	49.31
4 <i>Arabidopsis thaliana</i>	45.04	44.33	47.95	100	50.83	50.50
5 <i>Cryptococcus neoformans</i>	46.62	46.26	50.69	50.83	100	65.27
6 <i>Rasamsonia emersonii</i>	45.00	43.57	49.31	50.50	65.27	100

An alignment of the MIOX protein sequences listed in Table 1 revealed highly-conserved substrate and iron-binding sites within MIOX from all of the species analysed, including *R. emersonii* (Fig. 2). In accordance with previous reports, only the N-terminal region of the sequences showed high variability [7]. These results indicated that the putative *ReMIOX* had high homology to MIOX from multiple species and potentially identical substrate and iron-binding binding sites.

<i>H. sapiens</i>	-----MKVTVGPDP-----SLVYRPDVPDPEVAKDK
<i>M. musculus</i>	-----MKVDVGPDP-----SLVYRPDVPDPEMAKSK
<i>C. qasimii</i>	MRQELTDKE-----NAPLDNIEDWEAD-----LKVRYPKPKSEEAASK
<i>A. thaliana</i>	-MTILIDRHS-----DQNDAGDEIVEKNQGNK-----EETELVLDAGFEAPHTN
<i>R. emersonii</i>	MAPVAIEPTFQKRDGHALEDISEKIDTVNVLKAN---MQKEQKERGLYDETFDKNKDK
<i>C. neoformans</i>	MHAPEINNHI-KGEAVKLDKVSDEIDEVNVLLKQKDAVEKTQAEIDYDLASKFDQEKDK
	: : :
<i>H. sapiens</i>	A---SF ^R NYTSGP ^L LLDR---VF ^T TYKLMH ^T HQTVD ^F VR ^S KHAQ ^F GG ^F SYK ^K MTV ^M EAV ^D L
<i>M. musculus</i>	D---SF ^R NYTSGP ^L LLDR---VF ^T TYKLMH ^T HQTVD ^F VR ^S KRIQ ^F YS ^F SYK ^K MTI ^M EAV ^G M
<i>C. qasimii</i>	T---DY ^R NYHDS---ERVATVKEFYRLNHQYQTVDFVSQKQDEF ^L K ^F DK ^K EMPFWD ^A ID ^I Y
<i>A. thaliana</i>	SFGRTF ^R QYEEAC---ERRRGVEEFYRVNHIGQTVDFVRKMR ^E EY ^E KLN ^R TE ^M SIWE ^C CEL
<i>R. emersonii</i>	S---QF ^R QYEEAC---DR---VKNFYKEQHTKQTVAFNLKARNEFRSKTRAEMTVWEAIEK
<i>C. neoformans</i>	A---AF ^R QYEEAC---DR---VKNFYAEQHLKQTYEYNVKIRQEF ^R NTVRARMSIWEAMEL
	:*: *: * * * ** : . : : *.. :
<i>H. sapiens</i>	LDGLV ^D ESDPD ^V DFPNS ^F HAFQTAEGIRKAHPDKDWFHLVGLL ^H D ^L GK ^V L--ALFGE-PQ
<i>M. musculus</i>	LDDL ^V D ^S DPD ^V DFPNS ^F HAFQTAEGIRKAHPDKDWFHLVGLL ^H D ^L GK ^I M--ALWGE-PQ
<i>C. qasimii</i>	LNTLV ^D SDPD ^T SLDQL ^H LLQTSEAIRAD-GHPDWFVLTGFI ^H D ^L GK ^V L--CLFDE-PQ
<i>A. thaliana</i>	LNEFI ^D SDPD ^L DEPQIE ^H LLQTAEAIRKDYPDEDWLHTGLI ^H D ^L GK ^V LLHSSFGELPQ
<i>R. emersonii</i>	LNTLI ^D SDPD ^T ELSQID ^H ALQTAEAIRRD-GKPRWMQLTGLI ^H D ^L GK ^L LL--YFFDAQGG
<i>C. neoformans</i>	LDNLV ^D SDPD ^T SVGQIE ^H LLQTAEAIRRD-GKPEWMQLTGLI ^H D ^L GK ^L LL--CFFGADGQ
	: :: :**** . : * :*: :*. ** * : *. *: :****: : . *
<i>H. sapiens</i>	WAVV ^G DTFPVGCRPQASVVFCDSTFQDNPD ^L QDPRY ^S TELGMYPHCGLDRV ^L MSW ^G HDE
<i>M. musculus</i>	WAVV ^G DTFPVGCRPQASVVFCDSTFQDNPD ^L QDPRY ^S TELGMYPHCGLENV ^L MSW ^G HDE
<i>C. qasimii</i>	WAVV ^G DTFPVGCKHSDKIVY-PEYFDANPDSKDERYNTKYGIYEKNTGLDNVQMSW ^G HDE
<i>A. thaliana</i>	WAVV ^G DTFPVGCAFDESIVH-HKYFKENPDYDNPSYNSKYGIYTEGCLDNV ^L MSW ^G HDD
<i>R. emersonii</i>	WDVV ^G DTFPVGCAFDDRIIYGRESFKDNPDYGHPIYDTKFGIYTPGCLDNV ^L MSW ^G HDE
<i>C. neoformans</i>	WDVV ^G DTFVVGCFSDKIIY-PDTFKANPDYNNPKLNTKYGVYEPNCGLDNV ^L LSW ^G HDE
	* **** * * . :. . * . ** . : : *. * : : * : * : * : * : * : * : *
<i>H. sapiens</i>	YMYQVMKFNKFSLPPEAFYMI ^R F ^H SFYPWHTGRDYQQLCSQQLAML ^P WVREFNKF ^D LYT
<i>M. musculus</i>	YLYQMMKFNKFSLPPEAFYMI ^R F ^H SFYPWHTGGDYRQLCSQQLDML ^P WVQEFNKF ^D LYT
<i>C. qasimii</i>	YLYQIVK---DFLPEPALYMI ^R Y ^H SFYSHRENA ^D YH ^L MSD ^H DR ^E MF ^K W ^D KFN ^P YD ^L YS
<i>A. thaliana</i>	YMYLVAKENQTTLPSAGLFI ^R Y ^H SFYALHKSEAYK ^H LMN ^N EDREN ^M K ^W LK ^V FN ^K YD ^L YS
<i>R. emersonii</i>	YLYHVVK-DQSTLPPEALAMIRY ^H SFYPHRAGAYHELMNEHDK ^E MLKAVQAFN ^P YD ^L YS
<i>C. neoformans</i>	YMYEICK-NQSTLPQEALAMIRY ^H SFYPHREGAYQHLMNEKD ^H SQLKAVKAFN ^P YD ^L YS
	*. * : * ** . : : * : * : * * * * * * * : : * : * : *
<i>H. sapiens</i>	KCPDLPDVKLR ^P YYQGLIDKYCPGI-LSW
<i>M. musculus</i>	KCPDLPDVESLR ^P YYQGLIDKYCPGT-LSW
<i>C. qasimii</i>	KVPVPDAKALKPYYEDLVAKYLPAT-LKF
<i>A. thaliana</i>	KSKVRVNVEEVKPY ^L SLTNKYFPSK-LKW
<i>R. emersonii</i>	KSDDPPNVEELKPY ^Q ELIDEYFPKKVIKW
<i>C. neoformans</i>	KSDDPPKKEELKPY ^Q SIISKFFPAE-VQW
	* . . :.*** : : : * :. :

Figure 2. Alignment of the protein sequences from ReMIOX with MIOX from *Homo sapiens*, *Mus musculus*, *Cyclobacterium qasimii*, *Arabidopsis thaliana*, and *Cryptococcus neoformans* computed with CLUSTAL W. Blue and red coloured amino acids indicate conserved substrate- and iron-binding sites, respectively. Magenta coloured amino acids indicate those which join both substrate- and iron-binding sites according to a previous study [7]. (*) fully conserved amino acids; (:) highly conserved residues; (.) weakly conserved residues.

5.3.2 Recombinant expression of *ReMIOX* and *L-ReMIOX* in *E. coli*

A 10 kDa silica-binding peptide (referred to as “linker”) was fused to the N-terminus of *ReMIOX* to enable its immobilisation onto silica-based matrices (e.g., zeolite) for applications in biocatalysis. The linker has been shown previously to mediate effectively the co-immobilisation of several enzymes onto synthetic zeolite without any significant change in enzyme functionality [21, 23]. Similarly, as reported in Chapter 4, the fusion of the linker to mouse MIOX (*MmMIOX*) did not affect the activity of the resulting Linker-*MmMIOX* but it completely lost its activity upon immobilisation to zeolite. *L-ReMIOX* and *ReMIOX* (without linker) were expressed heterologously in *E. coli* and compared for solubility and activity to determine the effect of the linker fusion. *ReMIOX* and *L-ReMIOX* both displayed rapid and high expression at temperatures of 30 and 37 °C after 3h, and no observable expression below 30 °C (Fig. 3). This result differs from previous tests performed with *MmMIOX*, *L-MmMIOX*, *TaMIOX* and *L-TaMIOX* variants, where the highest expression levels were observed at 16 °C (Chapter 3), and which is different from MIOX from *C. neoformans* which showed higher expression only at temperatures below 30 °C [4].

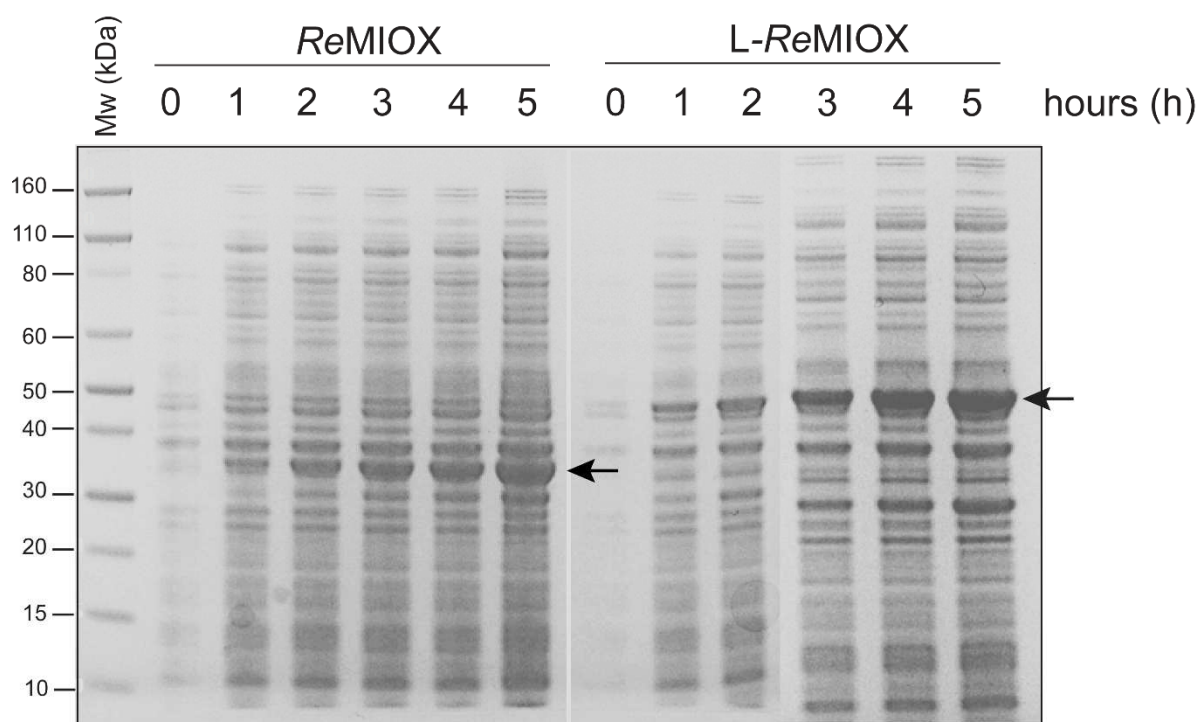


Figure 3. SDS-PAGE analysis of the total cell protein at various timepoints during protein expression of *ReMIOX* (left panel) and *L-ReMIOX* (right panel) in *E. coli* BL21 cells at 37 °C. The theoretical molecular weight (Mw) of *ReMIOX* is 36.9 kDa and of *L-ReMIOX* is 47.2 kDa (respective protein bands are indicated with arrows).

The majority of *ReMIOX* and *L-ReMIOX* proteins were insoluble after production in *E. coli* (data not shown). Supplementation of the medium with 60 mM *myo*-inositol during protein synthesis has been reported to improve the production and solubility of MIOX from various species [2, 3, 7]. However, on the contrary, no significant improvement was observed in the production and solubility of *ReMIOX* after the addition of 20, 60 or 80 mM *myo*-inositol to the growth medium (Fig. 4).

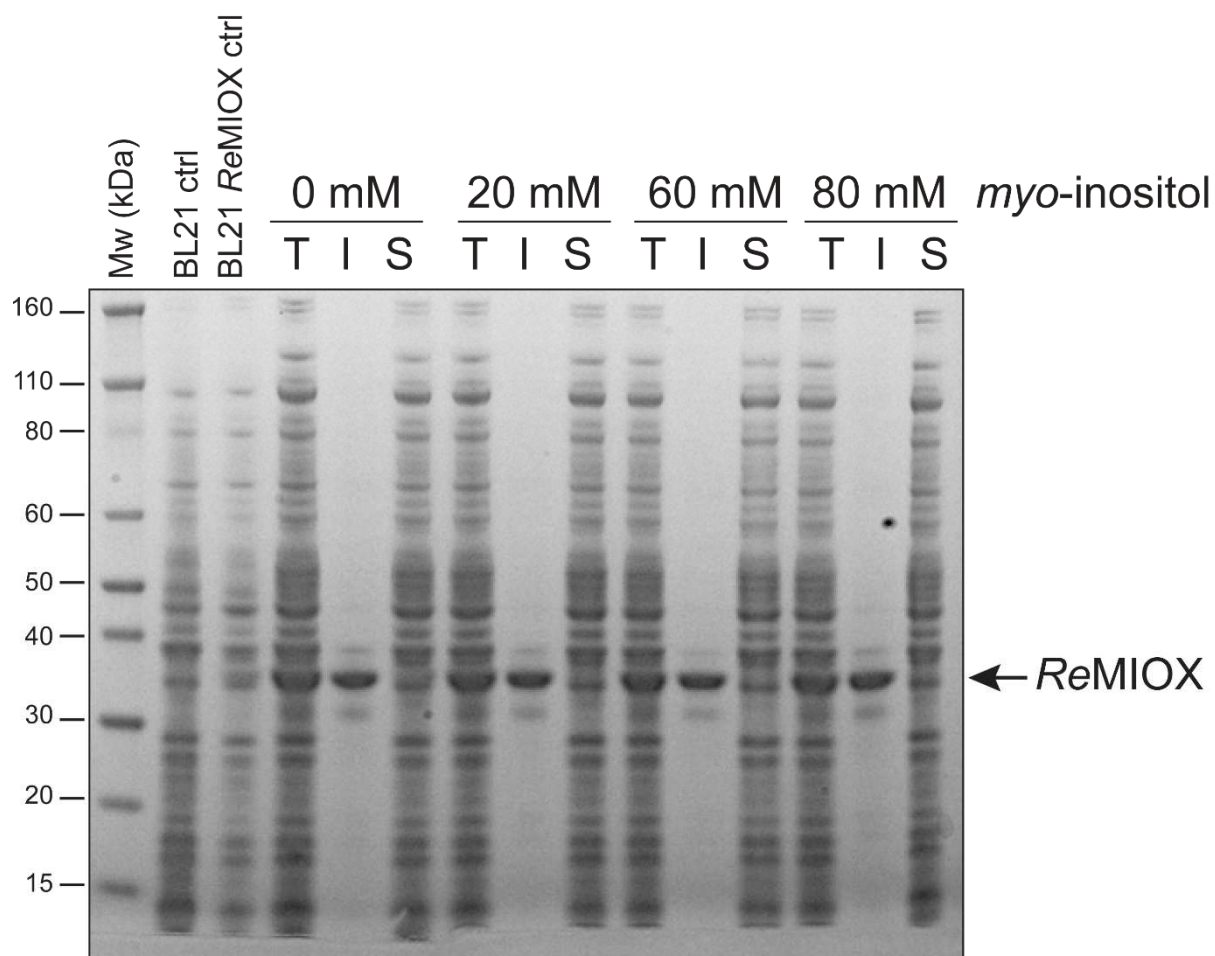


Figure 4. SDS-PAGE analysis of *ReMIOX* solubility (white circle) produced in *E. coli* BL21 cells with different concentrations of *myo*-inositol in the growth medium at 37°C. T, total cell lysate; I, insoluble protein fraction, S, soluble protein fraction. Controls: BL21 cells (BL21 ctrl) and total cell lysates of BL21 cells that contain *ReMIOX* (BL21 *ReMIOX* ctrl) but were collected prior to induction of recombinant protein expression.

To improve protein solubility, the expression of *ReMIOX* and *L-ReMIOX* was tested in *E. coli* T7 shuffle express cells (referred to as T7 cells) which are derived from the BL21 strain and display enhanced chaperone activity for improved protein folding. *ReMIOX* exhibited high expression levels at 30 °C when expressed in T7 cells while *L-ReMIOX* displayed the highest expression levels at 37 °C (data not shown). Nevertheless, most of *ReMIOX* and *L-ReMIOX* remained insoluble as inclusion bodies (Fig. 7). Changes to cell lysis conditions, such as

different buffers (phosphate buffer, 2-(N-morpholino) ethanesulfonic acid (MES) buffer or Tris-HCl) and varying the pH, did not improve *ReMIOX* or *L-ReMIOX* solubility (data not shown).

The solubilisation and refolding of *ReMIOX* from inclusion bodies was attempted in order to obtain soluble active enzyme. As shown in Fig. 5, solubilisation with GdnHCl or Urea yielded large amounts of soluble *ReMIOX* and high purity was obtained following refolding by stepwise dilution (10-fold or 100-fold) or dialysis with a refolding buffer as described in Materials and Methods. The SDS-PAGE analysis of the solubilised and subsequently refolded *ReMIOX* showed a predominant protein band with a smaller (denatured) molecular mass than *ReMIOX*. A protein of similar size was observed in the expression samples and cell lysates of *ReMIOX* and *L-ReMIOX* (Fig. 3 and 4, data for the cell lysates of *L-ReMIOX* are not shown). Similar observations have been reported for other MIOX variants (i.e. mouse, pig and rat) and was found to be the result of an N-terminal cleavage of MIOX after amino acid 32 [3]. Therefore, the protein band observed may be a truncated form of *ReMIOX* and *L-ReMIOX*. As a consequence, such a truncation would result in the loss of the linker fusion due to the N-terminal position of the Linker in *L-ReMIOX*.

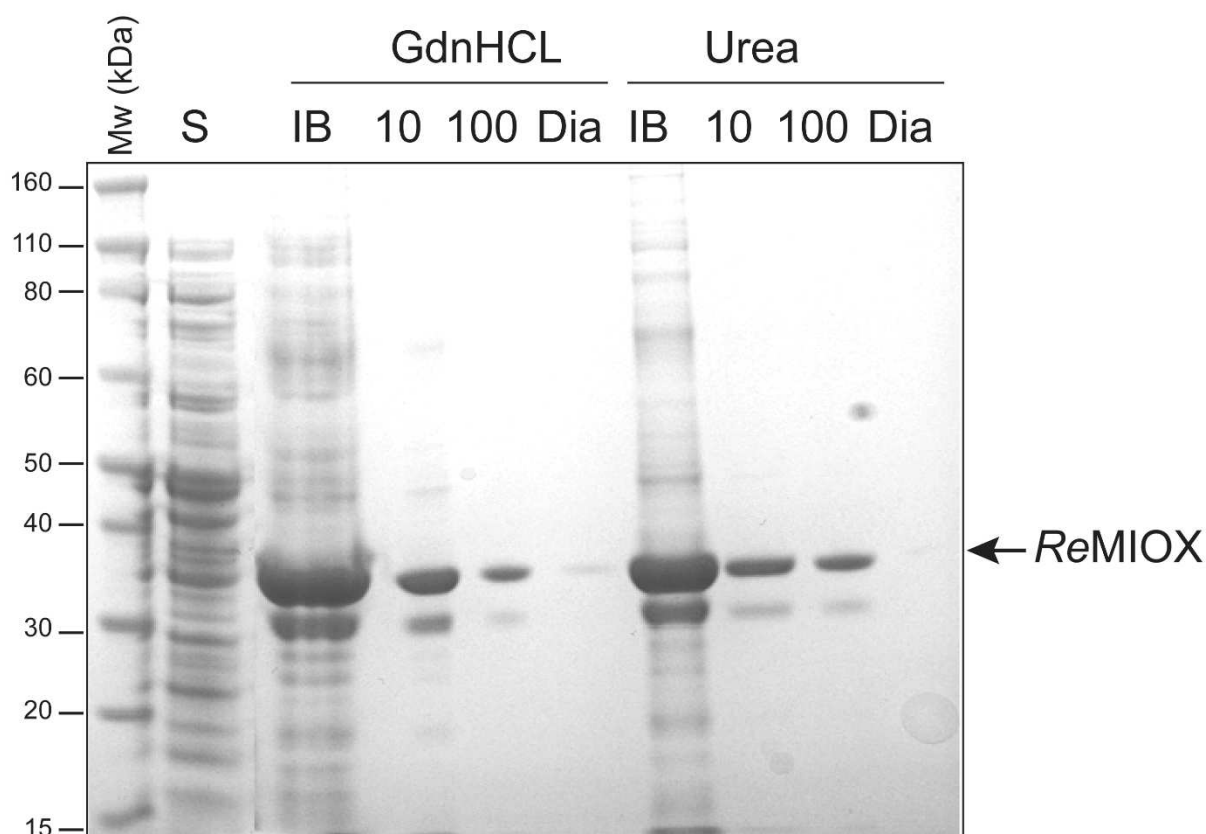


Figure 5. SDS-PAGE analysis of soluble *ReMIOX* cell extract (S) and inclusion bodies (IB). IB were solubilised with either guanidinium chloride (GdnHCl) or urea and then refolded by 10-fold, or 100-fold dilution or by dialysis (Dia).

5.3.3 Characterisation of MIOX

We first assessed the activity of *ReMIOX* and L- *ReMIOX* in whole cell activity assays to evaluate any potential enzyme instability. All activity assays were performed with whole BL21 and T7 cells at 45 °C. This temperature was selected because *R. emersonii*, the source of *ReMIOX* grows optimally at 45 °C [14].

ReMIOX showed higher activity in BL21 than T7 cells, whereas L-MIOX was more active in T7 cells, as shown in Fig. 6. The activity of the *ReMIOX* expressed in BL21 or T7 cells resulted in the production of 150 and 68 µM GA after 1 h of incubation, respectively. Under the same conditions, L-*ReMIOX* expressed in BL21 or T7 cells produced 146 and 180 µM GA, respectively. This translates to a production of 29 – 13 and 28 – 35 mg/L GA with *ReMIOX*

and L-*ReMIOX*. When compared to the whole cell activity of *MIOX* from *C. neoformans* [4], the production of GA observed here was about 60 times lower. However, these activities cannot be compared accurately when considering the differing cell densities and expression conditions.

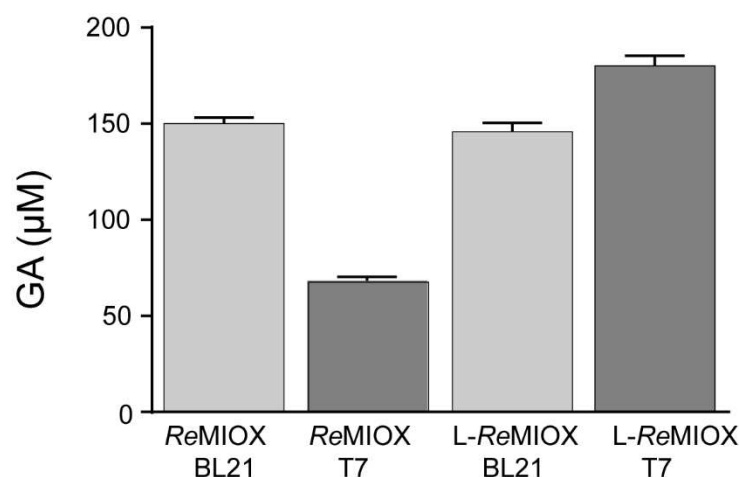


Figure 6. Production of glucuronic acid (GA) with *ReMIOX* and L-*ReMIOX* in whole *E. coli* BL21 cells or T7 shuffle express cells (T7) after expression at 37 °C for 5h. Cell extracts were incubated with 60 mM *myo*-inositol and 50 mM Tris-HCl buffer, pH 7.5 at 45 °C for 1h. The mean and standard deviation of two individual replicates are shown.

The activity of the protein fractions (total, insoluble or soluble) from lysed cells was evaluated (only with *ReMIOX*) following the whole cell activity tests. Surprisingly, activity tests performed with all protein fractions from lysed T7 cells contained no detectable *ReMIOX* activity. In contrast, *ReMIOX* activity was detected in all protein fractions from lysed BL21 cells that were expressed with and without *myo*-inositol supplementation (Fig. 7). The *myo*-inositol supplementation was conducted potentially to improve the enzyme solubility and activity. However, in agreement with our results in Section 3.2, soluble protein fractions from BL21 cells supplemented with *myo*-inositol showed no enhanced *ReMIOX* activity when

compared to cells without supplementation (Fig. 7). Nevertheless, higher *ReMIOX* activity was observed in the total cell extracts when *myo*-inositol was present as compared to when *myo*-inositol was absent during expression, with an average of 96.7 μM GA being produced after a one hour reaction (1.6 μM GA/min). In general, the soluble protein fractions showed no measurable activity or less activity than total protein fractions. Only in samples where no *myo*-inositol was supplemented did soluble protein yield more activity than the total or insoluble protein fractions, with 37.0 ± 7.9 μM GA being produced after one hour of reaction (0.5 μM GA/min). The least activity was observed in the insoluble protein fractions regardless of *myo*-inositol being present or absent during expression. This result was expected as insoluble proteins and inclusion bodies are often inactive [24]. No activities were measured in control samples of *E. coli* BL21 cells harbouring pET22b plasmids without the *ReMIOX* gene indicating that the activities observed were due to *ReMIOX*. The higher activity of *ReMIOX* in the total protein fractions that contained *myo*-inositol during expression may be attributed to the possibility that *ReMIOX* already converted *myo*-inositol to GA during protein expression which was performed for 4h at 37 °C. Upon cell lysis, GA would be freed into solution and detected together with the freshly produced GA. However, this suggestion does not explain the loss of *ReMIOX* activity observed in the soluble protein fractions, as GA is soluble in the medium and thus the GA produced during cell expression would be present in the soluble protein fractions. Instead, the loss of MIOX in the soluble protein fractions may result from the separation of soluble and insoluble cellular components rendering MIOX less stable and inactive. MIOX productivity was reported previously to be the highest in whole cells and was lost between 50 – 80 % upon different cell lysis procedures [4]. However, the *ReMIOX* activity observed in this study is about 1000 times lower than previous results obtained with soluble protein fractions of MIOX from mouse (Chapter 3).

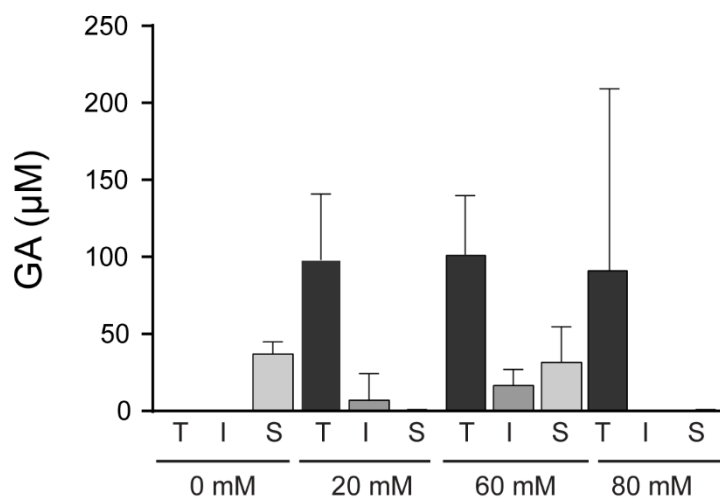


Figure 7. Production of glucuronic acid (GA) in *ReMIOX* BL21 cell lysates after supplementation of different concentrations of *myo*-inositol into the expression medium. Total (T), insoluble (I) or soluble (S) cell extracts were incubated with 60 mM *myo*-inositol and 50 mM Tris-HCl buffer, pH 7.5 at 45 °C for 1h. The means and standard deviations of two individual replicates are shown.

It should be noted, that none of the observations can explain the loss of MIOX activity upon lysis of T7 shuffle cells. In addition, the phenyl-phenol assay was found to be highly variable, which may be due to the non-specific detection of glucose when present in high amounts [22]. The assay may have detected GA within the cell wall of *E. coli* which could have been released when the sample was boiled with the sulphuric acid. Accordingly, the assay may have resulted in MIOX activity being measured incorrectly when conducted with whole cells and total cell extracts. The activity was measured with a different assay which was based on the highly selective conversion of GA into glucaric acid and NADH by urinate dehydrogenase (Udh) and can be monitored spectrophotometrically to rule out any false positives. This Udh assay detected no enzymatic activity in the whole cells or any of the protein fractions (total, insoluble and soluble) tested, and lowering or increasing the reaction temperature from 45 °C to 20 or 60 °C did not yield any detectable activity. Whether these conflicting observations indicate that the results obtained from the phenyl-phenol assay represented the GA concentration present in

E. coli cell wall rather than actual MIOX activity remains to be determined. Alternatively, the apparent loss of activity in the Udh assay may result from a lower sensitivity (detection limit = 17 mg/L or 88 μ M) than the phenyl-phenol assay (detection limit = 5 mg/L or 25 μ M). GA concentrations determined previously in total, insoluble or soluble protein fractions were near or below this detection limit (Fig. 7).

Inclusion bodies were resolubilised with either non-denaturing reagents (e.g. Tween100) and denaturing reagents (GdnHCl or urea) and refolded by using two different approaches, stepwise dilution or dialysis to obtain higher soluble enzyme levels and thus higher enzyme activity. The refolding buffer was supplemented with the chemical additives L-arginine and sucrose which were reported to stabilise proteins from aggregation during the refolding process [24]. No activity was measured in neither of the resulting protein fractions although highly purified *ReMIOX* was obtained (Fig. 5). MIOX might be truncated partially during cell lysis as described in Section 3.2 and this truncation may be enhanced throughout the refolding process. As suggested earlier, this truncation might be the main reason for the low activity observed in soluble cell extracts and a complete loss of activity observed in the refolded protein samples.

In conclusion, although the activity of *ReMIOX* and *L-ReMIOX* has been observed to some extent, they were much lower than those obtained with MIOX from mouse or the structurally highly similar MIOX from *C. neoformans*. This low activity and the low specificity of the phenyl-phenol activity assay together with conflicting results obtained when using the more selective but less sensitive Udh-based activity assay has hindered the in-depth biochemical characterisation of the *ReMIOX* enzyme and a confident statement about the true nature of the enzymatic activity of this protein.

5.4 Conclusion

A putative MIOX from the moderately thermostable fungus *R. emersonii* (*ReMIOX*) has been identified with high sequence similarity to previously characterised MIOXs. *R. emersonii* is well known for its production of thermostable enzymes and would be a promising host for producing a more thermostable MIOX enzyme than those that are derived from mesophilic mammals, plants or fungi.

Satisfactory expression levels were obtained upon heterologous expression of both *ReMIOX* and L-*ReMIOX* in *E. coli* and whole cell activity assays demonstrated slight MIOX activity for both enzymes. However, the very low solubility of *ReMIOX* and L-*ReMIOX* and accordingly, the resulting low activity observed in lysed cells has hampered their further biochemical characterisation. In addition, conflicting results from different activity assays has hindered a conclusive statement about the true activity of this putative enzyme. Attempts to improve the solubility of *ReMIOX* through different expression conditions and refolding of aggregated protein were not successful. The fusion of *ReMIOX* to protein tags such as GST or SUMO might help to increase *ReMIOX* solubility as previously reported for MIOX from mouse, or protein engineering might be effective in increasing MIOX activity [9, 25]. However, due to time constraints, further work pursued at improving solubility and activity of *ReMIOX* was not possible and outside the scope of this PhD project.

References

- [1] T.S. Moon, J.E. Dueber, E. Shiue, K.L.J. Prather, Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*, *Metabolic Engineering*, 12 (2010) 298-305.
- [2] T.S. Moon, S.H. Yoon, A.M. Lanza, J.D. Roy-Mayhew, K.L. Prather, Production of glucaric acid from a synthetic pathway in recombinant *Escherichia coli*, *Applied and Environmental Microbiology*, 75 (2009) 589-595.
- [3] R.J. Arner, K.S. Prabhu, C.C. Reddy, Molecular cloning, expression, and characterization of *myo*-inositol oxygenase from mouse, rat, and human kidney, *Biochemical and Biophysical Research Communications*, 324 (2004) 1386-1392.
- [4] S. Zheng, P. Wei, L. Huang, J. Cai, Z. Xu, Efficient expression of *myo*-inositol oxygenase in *Escherichia coli* and application for conversion of *myo*-inositol to glucuronic acid, *Food Science and Biotechnology*, 23 (2014) 445-450.
- [5] U. Kanter, M. Becker, E. Friauf, R. Tenhaken, Purification, characterization and functional cloning of inositol oxygenase from *Cryptococcus*, *Yeast (Chichester, England)*, 20 (2003) 1317-1329.
- [6] A. Lorence, B.I. Chevone, P. Mendes, C.L. Nessler, *myo*-inositol oxygenase offers a possible entry point into plant ascorbate biosynthesis, *Plant Physiology*, 134 (2004) 1200-1205.
- [7] A. Alok, H. Kaur, K.K. Bhati, J. Kumar, P. Pandey, S.K. Upadhyay, A. Pandey, N.C. Sharma, A.K. Pandey, S. Tiwari, Biochemical characterization and spatio-temporal expression of *myo*-inositol oxygenase (MIOX) from wheat (*Triticum aestivum* L.), *Plant Gene*, 4 (2015) 10-19.
- [8] R.J. Arner, K.S. Prabhu, J.T. Thompson, G.R. Hildenbrandt, A.D. Liken, C.C. Reddy, *myo*-Inositol oxygenase: molecular cloning and expression of a unique enzyme that oxidizes *myo*-inositol and D-*chiro*-inositol, *The Biochemical Journal*, 360 (2001) 313-320.
- [9] E. Shiue, K.L. Prather, Improving D-glucaric acid production from *myo*-inositol in *E. coli* by increasing MIOX stability and *myo*-inositol transport, *Metabolic Engineering*, 22 (2014) 22-31.
- [10] J.K. Guterl, D. Garbe, J. Carsten, F. Steffler, B. Sommer, S. Reisse, A. Philipp, M. Haack, B. Ruhmann, A. Koltermann, U. Kettling, T. Bruck, V. Sieber, Cell-free metabolic engineering: production of chemicals by minimized reaction cascades, *ChemSusChem*, 5 (2012) 2165-2172.
- [11] R.M. Daniel, M.J. Danson, D.W. Hough, C.K. Lee, M.E. Peterson, D.A. Cowan, Enzyme stability and activity at high temperatures, *Protein Adaptation in Extremophiles*, (2008) 1-34.
- [12] K. Cheng, F. Zhang, F. Sun, H. Chen, Y.H. Percival Zhang, Doubling Power Output of Starch Biobattery Treated by the Most Thermostable Isoamylase from an Archaeon *Sulfolobus tokodaii*, *Scientific Reports*, 5 (2015) 13184.

- [13] C. You, T. Shi, Y. Li, P. Han, X. Zhou, Y.P. Zhang, An *in vitro* synthetic biology platform for the industrial biomanufacturing of *myo*-inositol from starch, *Biotechnology and Bioengineering*, (2017).
- [14] J. Houbraken, H. Spierenburg, J.C. Frisvad, *Rasamsonia*, a new genus comprising thermotolerant and thermophilic *Talaromyces* and *Geosmithia* species, *Antonie Van Leeuwenhoek*, 101 (2012) 403-421.
- [15] B.R. Nielsen, J. Lehmbeck, T.P. Frandsen, Cloning, heterologous expression, and enzymatic characterization of a thermostable glucoamylase from *Talaromyces emersonii*, *Protein Expression and Purification*, 26 (2002) 1-8.
- [16] P.G. Murray, A. Grassick, C.D. Laffey, M.M. Cuffe, T. Higgins, A.V. Savage, A. Planas, M.G. Tuohy, Isolation and characterization of a thermostable endo-beta-glucanase active on 1,3-1,4-beta-D-glucans from the aerobic fungus *talaromyces emersonii* CBS 814.70, *Enzyme and Microbial Technology*, 29 (2001) 90-98.
- [17] C.M. Collins, P.G. Murray, S. Denman, J.P. Morrissey, L. Byrnes, T.T. Teeri, M.G. Tuohy, Molecular cloning and expression analysis of two distinct beta-glucosidase genes, *bg1* and *aven1*, with very different biological roles from the thermophilic, saprophytic fungus *Talaromyces emersonii*, *Mycological Research*, 111 (2007) 840-849.
- [18] D.M. Waters, P.G. Murray, L.A. Ryan, E.K. Arendt, M.G. Tuohy, *Talaromyces emersonii* Thermostable Enzyme Systems and Their Applications in Wheat Baking Systems, *Journal of Agricultural and Food Chemistry*, 58 (2010) 7415-7422.
- [19] H. Iwasa, A. Hiratsuka, K. Yokoyama, H. Uzawa, K. Orihara, H. Muguruma, Thermophilic *Talaromyces emersonii* Flavin Adenine Dinucleotide-Dependent Glucose Dehydrogenase Bioanode for Biosensor and Biofuel Cell Applications, *ACS Omega*, 2 (2017) 1660-1665.
- [20] R. Maheshwari, G. Bharadwaj, M.K. Bhat, Thermophilic fungi: their physiology and enzymes, *Microbiology and Molecular Biology Reviews : MMBR*, 64 (2000) 461-488.
- [21] A. Sunna, F. Chi, P.L. Bergquist, A linker peptide with high affinity towards silica-containing materials, *New Biotechnology*, 30 (2013) 485-492.
- [22] N. Blumenkrantz, G. Asboe-Hansen, New method for quantitative determination of uronic acids, *Analytical Biochemistry*, 54 (1973) 484-489.
- [23] A. Care, K. Petroll, E.S.Y. Gibson, P.L. Bergquist, A. Sunna, Solid-binding peptides for immobilisation of thermostable enzymes to hydrolyse biomass polysaccharides, *Biotechnology for Biofuels*, 10 (2017) 29.
- [24] H. Yamaguchi, M. Miyazaki, Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies, *Biomolecules*, 4 (2014) 235-251.
- [25] M. Lehmann, M. Wyss, Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution, *Current Opinion in Biotechnology*, 12 (2001).

Chapter 6:

Summary and Future Perspectives

A novel biocatalytic system was developed in this thesis, for the first cell-free synthesis of GlucA to be reported. The system is composed of a total six enzymes which were combined into a synthetic enzymatic pathway for the interconversion of glucose-1-phosphate (G1P) to GlucA. To accomplish the synthesis of GlucA in an economically-viable manner, four strategies were identified as being crucial to enable cell-free biocatalysis: i) a simple, rapid and reliable analytical method for analysing pathway metabolites, ii) the use of thermostable enzymes, iii) immobilisation and recycling of enzymes and iv) cofactor regeneration.

Chapter 1 sets up the scene by providing an up-to-date review of the current tools and strategies available for the construction of cell-free synthetic pathways. This review draws attention to the requirements for substrates, enzymes and cofactor regeneration systems to improve the effectiveness of cell-free biocatalysis. Especial attention has been given to the recent development of integrative tools such as *in silico* pathway modelling and high throughput flux analysis, emphasising their indispensable role in the future of cell-free biocatalytic pathways for industrial applications. Further discussion focused on the importance and challenges faced by several analytical methods for the measurement of products and intermediates generated by enzyme pathways, a subject usually neglected when in the discussion of synthetic biology pathway strategies.

Several synthetic enzyme pathways have been engineered to produce GlucA from glucose and *myo*-inositol in yeast and bacteria. However, major limitations for the optimisation of GlucA production via these pathways are the laborious analytical procedures required to detect the final product and pathway intermediates. Previously, GlucA and pathway intermediates were detected laboriously using a combination of enzymatic, colorimetric and chromatographic assays which hindered a rapid and comprehensive analysis of *in vivo* GlucA production [1-3].

Chapter 3 describes the development of a novel analytical method for the simultaneous

analysis of GlucA and all intermediates of the enzymatic pathway achieved in less than 10 minutes. The method was developed to facilitate the rapid metabolic flux analysis and identification of crucial pathway parameters such as enzyme ratios and cofactors. The GlucA pathway intermediates include glucose-1-phosphate (G1P), glucose-6-phosphate (G6P), inositol-1-phosphate (I1P), *myo*-inositol and glucuronic acid (GA) and their analysis was performed by ultra-high performance liquid chromatography (UHPLC) online-coupled to a refractive index detector (RID). The separation of these structurally highly similar compounds required the combination (i.e. mixed mode) of two chromatographic separation mechanisms, namely hydrophilic interaction liquid chromatography (HILIC) and weak anion exchange chromatography (WAX). The method is highly repeatable, precise and accurate and facilitated the simultaneous quantitation of all compounds down to a minimum concentration of 5 to 7 mM.

This novel method provides a valuable tool for the assessment and optimisation of GlucA production in cell-free systems which potentially can be applied to monitoring GlucA production in metabolically-engineered microorganisms. Furthermore, the high multimodality of mixed mode HILIX/WAX was shown to be effective in the simultaneous separation of various anionic sugar phosphates, sugar acids and neutral sugar alcohols, widely recognised as a challenging task amongst chromatographic-based separations [4-6]. The results presented in this research may serve as a basis for further development of new analytical tools for sugar-related applications.

The main limitation of this method was the impracticality of using gradient elution as RID is highly susceptible to changes in the mobile phase. Thus, isocratic elution was employed which was not sufficient to separate *myo*-inositol from typical enzyme buffers, such as HEPES or Tris-HCL, or to separate G1P from buffers (e. g., phosphate buffer) commonly used to

maintain the pH at which most enzymes reactions are performed. Similarly, reactions performed in HEPES buffer to maintain a pH of 8 in the analysis of *myo*-inositol were also compromised in this study. The potential development and integration of gradient elution for resolution of *myo*-inositol and HEPES and the use of a detector that is not subject to mobile phase composition (such as mass spectrometry) should address this limitation.

Chapter 4 describes the design, assembly and critical requirements of a cell-free biocatalytic system for production of GlucA which was adapted from a previously-published pathway for the *in vivo* (microbial) production of GlucA [7]. Specifically, the pathway was designed to start from G1P, instead of glucose, which can be derived by enzymatic phosphorolysis from renewable and more abundant biomass (e.g. cellulose or starch) without consumption of the expensive coenzyme, ATP [8, 9]. Further, the pathway was expanded with an NADH regeneration system to minimise the supply of expensive coenzymes, and where applicable, enzymes were derived mostly from (hyper)thermostable organisms to increase the robustness of the system and facilitate their simple and inexpensive heat purification.

The resulting multi-enzyme pathway includes a total of six enzymes which mediated successfully the cell-free production of 5.7 ± 1.1 mM GlucA from G1P after 8 h reaction time. However, further characterisation of the final product of the pathway with NMR and MS revealed that the actual product from the last reaction is the GlucA lactone and not the open form GlucA, which had to be interconverted first with heat and alkaline. This finding stresses that a structural characterisation of the product from a biocatalytic reaction through unambiguous analytical technologies such as NMR or MS is essential to validate reliably the product identity, as this will ultimately affect product analysis and downstream applications.

Compared to recently reported *in vivo* productions of GlucA, the cell-free system performed similarly with respect to titre and yield but demonstrated higher productivity (product formation over time) [3, 10-12]. Critical requirements for achieving adequate titres and yields were the delicate fine-tuning of enzyme ratios and cofactor requirements to maintain thermodynamically favourable conditions and enhance the reaction rates (kinetics) of the whole pathway. The performance of the pathway was optimised by successively changing the reaction conditions and monitoring the metabolic flux using the analytical method developed in **Chapter 3**. Nevertheless, the pathway conversion yields and titres achieved were lower than theoretically possible which indicates that the metabolic flux is still suboptimal. It is possible that the use of enzymes from different organisms and with different kinetic properties might improve the flux and final product titres.

It should be relatively simple to test this hypothesis with cell-free biocatalysis as compared to *in vivo* biocatalysis as there is no need for complex genetic modifications of the host organism and less interference from intracellular processes. However, cell-free biocatalysis faces its own limitations for industrial applications due to high enzyme costs and cofactor consumption. Accordingly, the enzymes of the cell-free GlucA pathway were immobilised via a silica-binding peptide onto a low-cost and commercially-available silica-based material (zeolite), enabling their recycling for several cycles without substantial loss of activity. Furthermore, incorporation of an immobilised NADH regeneration system (Nox enzyme) maintained the GlucA production at 70% of the initial production but required four times less NAD^+ than without the regeneration system.

The immobilised enzymes of the GlucA pathway exhibited different kinetic behaviours to those of their free enzyme counterparts. Future studies will be required for characterisation of the kinetic properties (e.g. K_{cat} and K_{m}) of the immobilised enzymes to achieve similar titres to those displayed by the free enzymes. The availability of more accurate pathway kinetic data

will allow for *in silico* kinetic modelling to minimise enzyme, substrate and cofactor usage while maintaining high pathway productivity [13-15].

The main limitation of the pathway is the mouse-derived *myo*-inositol oxygenase (MIOX) which catalyses the second to last step in the GlucA pathway. MIOX is catalytically-inefficient and highly unstable in solution. Although MIOX activity was not altered by incorporation of the silica-binding peptide and still displayed binding to the zeolite matrix, the enzyme completely lost activity upon immobilisation. In addition, the low temperature optimum of MIOX (40 °C) dictated a lower reaction temperature for all other enzymes which was circumvented by dividing the GlucA synthesis into a two-pot reaction. However, the two-pot solution demands higher system requirements due to double vessel maintenance, two separate temperature conditions and complicates system automation and scale-up.

Another critical limitation of MIOX is its low catalytic activity and instability which demands high amounts and a continuous replenishment of the enzyme for high pathway productivity. MIOX has been recognised previously to be the major bottleneck in the microbial GlucA production [3, 11, 16]. Therefore, the discovery or engineering of a more catalytically-efficient and thermostable enzyme is of high relevance and could improve the system performance of both the cell-free and microbial GlucA production significantly.

In **Chapter 5**, the preliminary attempts to characterise a putative thermostable MIOX are presented. MIOX is highly conserved among eukaryotes and was identified previously only in mesophilic organisms such as fungi, plants, or mammals. A putative MIOX gene was identified from the moderately thermophilic fungus *Rasamsonia emersonii* (referred to as *ReMIOX*) which is known to produce thermostable enzymes optimally active between 60 and 80 °C [17-21]. The putative *ReMIOX* exhibited between 65 and 44 % sequence similarity to MIOX from

the fungus *Cryptococcus neoformans* and mouse, respectively. The sequence also showed the highly conserved substrate and iron-binding sites present among MIOX from different species. This putative *ReMIOX* gene was expressed heterologously in *E. coli* with satisfactory levels but very low solubility. Attempts to enhance the solubility or refold active *ReMIOX* from inclusion bodies were unsuccessful. The low solubility and/or activity of *ReMIOX* was responsible for the conflicting results obtained in the activity assays with whole cells and cell lysates and did not allow for confirmation of the enzymatic identity of the recombinant enzyme. Unfortunately, due to time constraints, further work aimed at confirming the MIOX activity of *ReMIOX* was not possible and outside the scope of this PhD project.

Fusion of the putative *ReMIOX* to protein tags such as GST or SUMO that were shown to increase enzyme solubility could be a possible strategy to achieve higher soluble protein levels and facilitate the biochemical characterisation of *ReMIOX*. Solubility improvement have been reported for MIOX from mouse upon fusion to the SUMO tag [16]. However, a general problem was found to be the truncation of the MIOX protein at its N-terminus which renders it inactive [22]. Similarly, in this study a truncated protein was observed consistently in cell lysates of *ReMIOX*. Therefore, improving *ReMIOX* solubility might not necessarily improve the enzyme's stability and activity.

Alternatively, engineering a more thermostable and catalytically-efficient MIOX enzyme may be a more promising route for stabilisation. So far, only one study has applied directed evolution to engineer MIOX but did not succeed in producing a more stable or active variant [16]. Thus, the intrinsic lability of the MIOX enzyme might be a poor template for protein engineering by random mutagenesis. Better results may be obtained by using rational design and the introduction of 'thermostabilising mutations' or by a combination of directed evolution and rational design for incorporating additive mutations [23].

The SpyTag/SpyCatcher conjugation technique was recently reported in which SpyTag can form a covalent isopeptide bond spontaneously with its protein partner SpyCatcher [24]. This system could be investigated as an alternative approach to increase the thermostability of the MIOX enzyme since the stability of several enzymes has been reported by way of SpyTag/SpyCatcher cyclization to improve their performance in industrial biocatalysts [25-28]. For example, a trehalose synthase from *Thermomonospora curvata* was cyclized after the fusion of a SpyTag to its C-terminus and SpyCatcher to the N-terminus. The cyclized enzyme retained more than 85% of its activity at temperatures ranging from 40 to 50°C while the wild type kept only 60 and 80% of its activity under the same conditions [25]. Firefly luciferase from *Photinus pyralis* was cyclized *in vivo* via the SpyTag/SpyCatcher. The resulting enzyme was more thermostable and alkali-tolerant than the wild type, without compromising the specific activity [28]. The SpyTag/SpyCatcher cyclized lichenase from *Bacillus subtilis* was shown to retain 80% of its activity after 100 °C exposure, whereas the linear form of the enzyme lost almost all of its activity [26]. The cyclized variant not only displayed a significantly higher thermal stability but also 1.7-fold lower K_m value than that of the linear enzyme, indicating a higher affinity for its substrate [26].

The synthesis of GlucA was conceptualised with G1P as starting material as it provided the possibility of linking the GlucA pathway to two naturally-abundant biopolymers, e.g. cellulose or starch. It has been shown previously that G1P can be produced enzymatically from starch (including amylose and maltodextrin or potatoes) and was linked to the cell-free production of *myo*-inositol, which is a precursor in the synthetic GlucA pathway [8, 9]. Further, multiple enzymes have been described that catalyse the phosphorolytic degradation of cellobiose and cellodextrins (which are cellulose oligomers) to G1P and potentially could be exploited for the cell-free production of GlucA from cellulose [15, 29-31]). However, as those enzymes do only

accept the cellulose oligomers with a polymerisation degree of up to 7 (celloheptaose), acid-or enzyme catalysed degradation of cellulose is a critical requisite.

As shown in the **Appendix**, preliminary work in this thesis has demonstrated that thermostable cellulases and hemicellulases were immobilised and recycled effectively using the linker technology. The combination of cellulases and hemicellulases has promising applications in the degradation of plant biomass and thus could be potentially exploited to produce GlucA directly from cellulosic food and paper waste.

In conclusion, this work provides the first cell-free production of GlucA which is environmental friendly, similarly efficient as the microbial GlucA production and proposes valuable solutions for its low-cost operation. In addition, the work identifies some of the remaining bottlenecks that needs to be addressed to improve the systems performance. Specifically, future work is needed to extend the synthesis of GlucA from biomass such as cellulose and to reach the maximum theoretical conversion yields and product titres. Kinetic modelling of the pathway might further reduce enzyme and substrate loading. Improvements in regard to the MIOX enzyme are necessary to further optimise system efficiency and sustainability as they pose a major bottleneck in the cell-free and microbial production of GlucA. In addition, to meet the requirements of a bio-based industry, the laboratory process would need to demonstrate commercial scale-ability.

References

- [1] Y.-N. Qu, H.-J. Yan, Q. Guo, J.-L. Li, Y.-C. Ruan, X.-Z. Yue, W.-X. Zheng, T.-W. Tan, L.-H. Fan, Biosynthesis of D-glucaric acid from sucrose with routed carbon distribution in metabolically engineered *Escherichia coli*, *Metabolic Engineering*, 47 (2018) 393-400.
- [2] Y. Liu, X. Gong, C. Wang, G. Du, J. Chen, Z. Kang, Production of glucaric acid from *myo*-inositol in engineered *Pichia pastoris*, *Enzyme and Microbial Technology*, 91 (2016) 8-16.
- [3] T.S. Moon, S.H. Yoon, A.M. Lanza, J.D. Roy-Mayhew, K.L. Prather, Production of glucaric acid from a synthetic pathway in recombinant *Escherichia coli*, *Applied and Environmental Microbiology*, 75 (2009) 589-595.
- [4] H. Hinterwirth, M. Lämmerhofer, B. Preinerstorfer, A. Gargano, R. Reischl, W. Bicker, O. Trapp, L. Brecker, W. Lindner, Selectivity issues in targeted metabolomics: Separation of phosphorylated carbohydrate isomers by mixed-mode hydrophilic interaction/weak anion exchange chromatography, *Journal of Separation Science*, 33 (2010) 3273-3282.
- [5] B. Preinerstorfer, S. Schiesel, M. Lämmerhofer, W. Lindner, Metabolic profiling of intracellular metabolites in fermentation broths from β -lactam antibiotics production by liquid chromatography–tandem mass spectrometry methods, *Journal of Chromatography A*, 1217 (2010) 312-328.
- [6] C. Mathon, G.A. Barding, C.K. Larive, Separation of ten phosphorylated mono- and disaccharides using HILIC and ion-pairing interactions, *Analytica Chimica Acta*, 972 (2017) 102-110.
- [7] T.S. Moon, S.-H. Yoon, M.-J. Tsang Mui Ching, A.M. Lanza, K.L.J. Prather, Enzymatic assay of d-glucuronate using uronate dehydrogenase, *Analytical Biochemistry*, 392 (2009) 183-185.
- [8] T. Fujisawa, S. Fujinaga, H. Atomi, An *in vitro* enzyme system for the production of *myo*-inositol from starch, *Applied and Environmental Microbiology*, (2017).
- [9] C. You, T. Shi, Y. Li, P. Han, X. Zhou, Y.P. Zhang, An *in vitro* synthetic biology platform for the industrial biomanufacturing of *myo*-inositol from starch, *Biotechnology and Bioengineering*, (2017).
- [10] I.M.B. Reizman, A.R. Stenger, C.R. Reisch, A. Gupta, N.C. Connors, K.L.J. Prather, Improvement of glucaric acid production in *E. coli* via dynamic control of metabolic fluxes, *Metabolic Engineering Communications*, 2 (2015) 109-116.
- [11] T.S. Moon, J.E. Dueber, E. Shiue, K.L.J. Prather, Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*, *Metabolic Engineering*, 12 (2010) 298-305.
- [12] A. Gupta, M.A. Hicks, S.P. Manchester, K.L. Prather, Porting the synthetic D-glucaric acid pathway from *Escherichia coli* to *Saccharomyces cerevisiae*, *Biotechnology Journal*, 11 (2016) 1201-1208.

- [13] C. Zhong, P. Wei, Y.-H.P. Zhang, A kinetic model of one-pot rapid biotransformation of cellobiose from sucrose catalysed by three thermophilic enzymes, *Chemical Engineering Science*, 161 (2017) 159-166.
- [14] J.A. Rollin, J. Martin del Campo, S. Myung, F. Sun, C. You, A. Bakovic, R. Castro, S.K. Chandrayan, C.-H. Wu, M.W.W. Adams, R.S. Senger, Y.-H.P. Zhang, High-yield hydrogen production from biomass by *in vitro* metabolic engineering: Mixed sugars coutilisation and kinetic modeling, *Proceedings of the National Academy of Sciences*, 112 (2015) 4964-4969.
- [15] X. Ye, Y. Wang, R.C. Hopkins, M.W. Adams, B.R. Evans, J.R. Mielenz, Y.H. Zhang, Spontaneous high-yield production of hydrogen from cellulosic materials and water catalysed by enzyme cocktails, *ChemSusChem*, 2 (2009) 149-152.
- [16] E. Shiue, K.L. Prather, Improving D-glucaric acid production from *myo*-inositol in *E. coli* by increasing MIOX stability and *myo*-inositol transport, *Metabolic Engineering*, 22 (2014) 22-31.
- [17] B.R. Nielsen, J. Lehmbeck, T.P. Frandsen, Cloning, heterologous expression, and enzymatic characterization of a thermostable glucoamylase from *Talaromyces emersonii*, *Protein Expression and Purification*, 26 (2002) 1-8.
- [18] D.M. Waters, P.G. Murray, L.A. Ryan, E.K. Arendt, M.G. Tuohy, *Talaromyces emersonii* Thermostable Enzyme Systems and Their Applications in Wheat Baking Systems, *Journal of Agricultural and Food Chemistry*, 58 (2010) 7415-7422.
- [19] H. Iwasa, A. Hiratsuka, K. Yokoyama, H. Uzawa, K. Orihara, H. Muguruma, Thermophilic *Talaromyces emersonii* Flavine Adenine Dinucleotide-Dependent Glucose Dehydrogenase Bioanode for Biosensor and Biofuel Cell Applications, *ACS Omega*, 2 (2017) 1660-1665.
- [20] R. Maheshwari, G. Bharadwaj, M.K. Bhat, Thermophilic fungi: their physiology and enzymes, *Microbiology and Molecular Biology Reviews : MMBR*, 64 (2000) 461-488.
- [21] J. Houbraken, H. Spierenburg, J.C. Frisvad, Rasamsonia, a new genus comprising thermotolerant and thermophilic *Talaromyces* and *Geosmithia* species, *Antonie Van Leeuwenhoek*, 101 (2012) 403-421.
- [22] R.J. Arner, K.S. Prabhu, C.C. Reddy, Molecular cloning, expression, and characterization of *myo*-inositol oxygenase from mouse, rat, and human kidney, *Biochemical and Biophysical Research Communications*, 324 (2004) 1386-1392.
- [23] M. Lehmann, M. Wyss, Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution, *Current Opinion in Biotechnology*, 12 (2001).
- [24] B. Zakeri, J.O. Fierer, E. Celik, E.C. Chittock, U. Schwarz-Linek, V.T. Moy, M. Howarth, Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin, *Proceedings of the National Academy of Sciences of the United States*, 109 (2012) E690-697.
- [25] C. Xu, Q. Xu, H. Huang, L. Jiang, Enhancing the stability of trehalose synthase via SpyTag/SpyCatcher cyclization to improve its performance in industrial biocatalysts, *Bioscience, Biotechnology, and Biochemistry*, (2018) 1-7.

- [26] J. Wang, Y. Wang, X. Wang, D. Zhang, S. Wu, G. Zhang, Enhanced thermal stability of lichenase from *Bacillus subtilis* 168 by SpyTag/SpyCatcher-mediated spontaneous cyclization, *Biotechnology for Biofuels*, 9 (2016) 79.
- [27] C. Schoene, J.O. Fierer, S.P. Bennett, M. Howarth, SpyTag/SpyCatcher cyclization confers resilience to boiling on a mesophilic enzyme, *Angewandte Chemie (International ed. in English)*, 53 (2014) 6101-6104.
- [28] M. Si, Q. Xu, L. Jiang, H. Huang, SpyTag/SpyCatcher Cyclization Enhances the Thermostability of Firefly Luciferase, *PLOS ONE*, 11 (2016) e0162318.
- [29] M. Arai, K. Tanaka, T. Kawaguchi, Purification and properties of cellodextrin phosphorylase from *Clostridium thermocellum*, *Journal of Fermentation and Bioengineering*, 77 (1994) 239-242.
- [30] M. Reichenbecher, F. Lottspeich, K. Bronnenmeier, Purification and Properties of a Cellobiose Phosphorylase (CepA) and a Cellodextrin Phosphorylase (CepB) from the Cellulolytic Thermophile *Clostridium Stercorarium*, *European Journal of Biochemistry*, 247 (1997) 262-267.
- [31] Y. Wu, G. Mao, H. Fan, A. Song, Y.-H.P. Zhang, H. Chen, Biochemical properties of GH94 cellodextrin phosphorylase THA_1941 from a thermophilic eubacterium *Thermosiphon africanus* TCF52B with cellobiose phosphorylase activity, *Scientific Reports*, 7 (2017) 4849.

Appendix A: Publication 4

Care, A., Petroll, K., Gibson, E. SY., Bergquist, P.L., Sunna, A. (2017). Solid-binding peptides for immobilisation of thermostable enzymes to hydrolyse biomass polysaccharides. *Biotechnology for Biofuels* 10 (1): 29. <https://doi.org/10.1186/s13068-017-715-2>.

Care et al. *Biotechnol Biofuels* (2017) 10:29
DOI 10.1186/s13068-017-0715-2

Biotechnology for Biofuels

RESEARCH

Open Access



Solid-binding peptides for immobilisation of thermostable enzymes to hydrolyse biomass polysaccharides

Andrew Care^{1,2}, Kerstin Petroll¹, Emily S. Y. Gibson¹, Peter L. Bergquist^{1,3,4} and Anwar Sunna^{1,4*}

Abstract

Background: Solid-binding peptides (SBPs) bind strongly to a diverse range of solid materials without the need for any chemical reactions. They have been used mainly for the functionalisation of nanomaterials but little is known about their use for the immobilisation of thermostable enzymes and their feasibility in industrial-scale biocatalysis.

Results: A silica-binding SBP sequence was fused genetically to three thermostable hemicellulases. The resulting enzymes were active after fusion and exhibited identical pH and temperature optima but differing thermostabilities when compared to their corresponding unmodified enzymes. The silica-binding peptide mediated the efficient immobilisation of each enzyme onto zeolite, demonstrating the construction of single enzyme biocatalytic modules. Cross-linked enzyme aggregates (CLEAs) of enzyme preparations either with or without zeolite immobilisation displayed greater activity retention during enzyme recycling than those of free enzymes (without silica-binding peptide) or zeolite-bound enzymes without any crosslinking. CLEA preparations comprising all three enzymes simultaneously immobilised onto zeolite enabled the formation of multiple enzyme biocatalytic modules which were shown to degrade several hemicellulosic substrates.

Conclusions: The current work introduced the construction of functional biocatalytic modules for the hydrolysis of simple and complex polysaccharides. This technology exploited a silica-binding SBP to mediate effectively the rapid and simple immobilisation of thermostable enzymes onto readily-available and inexpensive silica-based matrices. A conceptual application of biocatalytic modules consisting of single or multiple enzymes was validated by hydrolysing various hemicellulosic polysaccharides.

Keywords: Solid-binding peptide, Thermostable enzymes, CLEAs, Enzyme immobilisation, Biocatalytic modules

Background

The immobilisation of enzymes onto solid supports can be achieved using various chemical and physical techniques, such as adsorption, covalent attachment, entrapment/encapsulation, and crosslinking [1]. Compared to soluble enzymes, immobilised enzymes offer enhanced stability and easy removal from a reaction mixture, enabling repetitive use in batch and continuous bioprocesses, rapid termination of reactions, controlled product formation and the provision of flexibility

to industrial bioprocesses (e.g. biofuel production) [2, 3]. Unfortunately, conventional enzyme immobilisation methods often result in the non-uniform orientation of enzymes and unwanted conformational changes that distort active sites and thus reduce the catalytic activity of the enzyme [4–7].

Solid-binding peptides (SBPs) exhibit binding affinity and selectivity to the surfaces of solid materials, including support materials used in biocatalysis e.g., silica, zeolite, glass, metals and polymers [8]. They bind strongly to their cognate solids by adopting specific conformations that support numerous non-covalent interactions (e.g. electrostatic bonding) between key amino acids and the solid surface, often resulting in binding affinities (K_d) in the sub-micromolar and nanomolar range [9].

*Correspondence: anwar.sunna@mq.edu.au

¹ Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, Australia
Full list of author information is available at the end of the article



© The Author(s) 2017. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Furthermore, SBPs can be integrated easily into the permissive regions of proteins (e.g. N- or C-terminal) using genetic engineering techniques to produce bifunctional fusion proteins. Consequently, SBPs are employed frequently as molecular linkers and/or anchoring molecules for the immobilisation of functional proteins onto solid surfaces in a single-step and without the need for any complicated chemical reactions or physical treatments [8].

SBPs have been shown to direct the oriented immobilisation of enzymes onto solid supports with retention of native enzyme structure and catalytic activity (e.g. k_{cat}/K_m) [10, 11]. For example, the immobilisation of alkaline phosphatase (AP) onto gold surfaces via a gold-binding peptide (GBP-1) has been reported [12]. When compared to native AP, a GBP1-AP fusion displayed high binding affinity to gold surfaces and self-assembled into densely packed, uniform protein monolayers that exhibited improved enzyme orientation and greater enzymatic activity per unit of area than AP without GBP1. Additionally, biophysical studies have shown that GBP1 is capable of facilitating the immobilisation of enzymes without any observable effect on their native structural conformation [13].

SBPs have been used primarily as molecular tools for the functionalisation of nanomaterials and the practical application of SBP-enzyme fusions are limited currently to chemo/biosensors and bioassays that require exotic and expensive laboratory-based matrices that may not be realistic economically for large-scale biotechnological processes [8, 10, 13]. Silica, zeolite and other mesoporous silica-based materials are used commonly in industry as support matrices in a number of applications and the provision for bulk supply is well established. These inorganic matrices are usually considered the most suitable for the immobilisation of enzymes [14, 15]. They do not require any chemical functionalisation and can be synthesised or mined cheaply. In particular, silica can be tailored easily to provide a large range of porous surfaces, surface functionalities and processing conditions [14]. However, use of unmodified silica as an inorganic support is unusual, as supports are usually functionalised prior to immobilisation of a biomolecule, both in single and multiple enzyme immobilisation [16]. Mesoporous nanomaterials are used because of their high surface area, controlled porosity and simple adsorption and desorption [17]. Silica-based supports show high mechanical strength and microbial resistance compared to other polysaccharide-based adsorbents [15, 18]. These considerations make silica-based materials well-suited for industrial-scale processes. There are few reports regarding inexpensive matrices and little is known about the use of SBPs for the immobilisation of

thermophilic enzymes and their feasibility in industrial-scale biocatalysis.

We reported previously the application of a unique SBP of amino acid sequence (VKTQATSREEPPRLPSKHRPG)₄ VKTQTAS (referred to as the linker, L) that binds strongly to a range of inexpensive silica-based materials (e.g. natural or synthetic zeolites and silica) [16, 19–22]. Upon binding to silica surfaces, the linker is thought to undergo conformational changes that promote multiple electrostatic interactions between its many positively-charged residues (K and R) and the negatively-charged silica. Using this technology, a number of linker-fusion proteins were constructed for direct and effective functionalisation of nanomaterial surfaces [16, 19, 21, 22].

Further extensions of this technology to single enzymes and enzyme complexes that can be designed as biocatalytic modules using a silica-binding SBP and low cost bulk silica-based matrices can be envisaged. Single biocatalytic modules consist of one type of enzyme immobilised per matrix molecule, whereas multiple enzyme biocatalytic modules consist of more than one enzyme. Biocatalytic modules can be used to replace industrial chemical catalysts due to their high specificity and high efficiency under mild reaction conditions [17].

Here we focus particularly on the effect that incorporating the silica-binding peptide has on the activity and function of three industrially-relevant thermostable enzymes, providing basic groundwork for future studies on the concept of immobilisation of a biomass degradation pathway consisting of multiple enzymes. β -Glucosidase is the last enzyme in the classical cellulose degradation pathway and unlike the other cellulases, it cannot be recycled by exposure to fresh substrate so it is a priority candidate enzyme for immobilisation [23]. To further complement this work, a thermostable β -mannanase (ManA) and β -xylanase (XynB) from *Dictyoglomus thermophilum* were used to supply additional data on the functionality of the silica-binding peptide. Both enzymes are hemicellulases that catalyse the hydrolysis of the polysaccharides mannan (1,4-linked β -D-mannopyranosyl residues) and xylan (1,4-linked β -D-xylopyranosyl residues), respectively. Hemicellulases, in combination with cellulose-degrading enzymes, play an important role in the efficient degradation of plant biomass.

This study demonstrated the application of the silica-binding SBP-mediated immobilisation of industrially-relevant enzymes onto a bulk and low-cost solid zeolite matrix. Additional introduction of crosslinking into the immobilised enzymes to form single and multiple enzyme biocatalytic modules further highlighted the potential of this platform technology for its translation into industrial-scale processes.

Results

Construction of thermophilic genes as linker fusions and their effect on enzyme activity

The linker (L) sequence was incorporated at the N- or C-terminus of three thermostable hemicellulases isolated from thermophilic bacteria (see Table 1). The resulting fusion proteins (L-enzymes) and enzymes (without linker) were expressed in *Escherichia coli* and partially purified by heat treatment as described in Methods.

The optimum pH, temperature and thermostability of each enzyme (with and without linker) was compared to investigate the effect of linker fusion on the operational activity and stability of the enzymes. Table 2 shows that all L-enzymes remained active after fusion and displayed maximal activity at the same pH and temperature optimum as their counterpart enzymes without linker. Thermostability experiments in the absence of substrate showed that L-XynB and ManA-L retained 100% of their enzymatic activity after 6 h incubation at 70 °C. In earlier reports, XynB and ManA (without linker) showed no inactivation following 6 h incubation at 80 °C [24, 25]. However, under similar conditions, L-XynB and ManA-L retained only 73 and 19% of their initial activities,







respectively. BglA was the least thermostable enzyme tested as it retained only 69% of its activity after 1 h of incubation at 80 °C whilst L-BglA retained 35%. BglA was reported previously to have lost 50% of its initial activity at 80 °C after 15 and 24 min [26, 27]. Similarly, a recombinantly-expressed and purified β -glucosidase (CbBgl1A) from another *Caldicellulosiruptor* species, *Cs. bescii*, showed a half-life of 20 min at 80 °C [28]. Initially the half-life at 80 °C of the BglA produced in *E. coli* was reported to be 14 h [29].

Enzyme immobilisation

Zeolite-bound enzymes (zeo-L-enzymes)

The linker has been reported to mediate the immobilisation of proteins onto various silica-based materials, including natural and synthetic zeolites [16, 22]. Standard zeolite binding assays were performed with partially purified enzymes (with and without the linker, Fig. 1). All enzymes carrying the linker (L-BglA, L-XynB and ManA-L) displayed strong binding affinity for the zeolite matrix with all of their initial protein present in the final zeolite-bound fraction after a stringent washing procedure. In contrast, the enzymes without linker (BglA, XynB and

Table 1 Enzyme constructions used

Plasmid	Protein	Construction	Function	Source	Linker (L) position
pET22-L-bglA	L-BglA		β -Glucosidase	<i>Caldicellulosiruptor saccharolyticus</i>	N
pET22-L-xynB	L-XynB		Xylanase	<i>Dictyoglomus thermophilum</i>	N
pET22-manA-L	ManA-L		Mannanase	<i>Dictyoglomus thermophilum</i>	C
pET22-bglA	BglA		β -Glucosidase	<i>Caldicellulosiruptor saccharolyticus</i>	N/A
pJLA602-xynB	XynB		Xylanase	<i>Dictyoglomus thermophilum</i>	N/A
pET22-manA	ManA		Mannanase	<i>Dictyoglomus thermophilum</i>	N/A

N Linker is at N-terminus of fusion protein, C linker is at C-terminus of fusion protein

Table 2 Effect of linker fusion on the optimal pH, apparent temperature optima and thermostability of BglA, XynB, and ManA

	BglA	L-BglA	XynB	L-XynB	ManA	ManA-L
Optimum pH	6.0	6.0	6.5	6.5	5.0	5.0
Optimum temperature (°C)	85	85	80	80	85	85
Thermostability at (°C)	Relative activity (%)					
70	81 ^a	54 ^a	100 ^b	100	100 ^c	100
80	69 ^a	35 ^a	100 ^b	73	100 ^c	19

^a Samples incubated for 1 h at the specified temperature. All other samples were incubated for 6 h

^b Morris et al. [24]

^c Gibbs et al. [25]

storage and operational stability and are easy to recover and recycle.

CLEAs of free enzymes (with and without linker) were prepared by ethanol precipitation and subsequent crosslinking with glutaraldehyde at 4 °C (Fig. 2a, b). Several protein precipitants were tested, including ammonium sulfate and isopropanol, but ethanol provided the most rapid and complete precipitation. Allowing the crosslinking reaction to occur overnight without shaking gave the best results, resulting in low enzyme activity in the supernatant and high activity for the resulting CLEAs.

CLEAs of zeolite-bound enzymes (zeo-L-enzyme_{CLEAs})

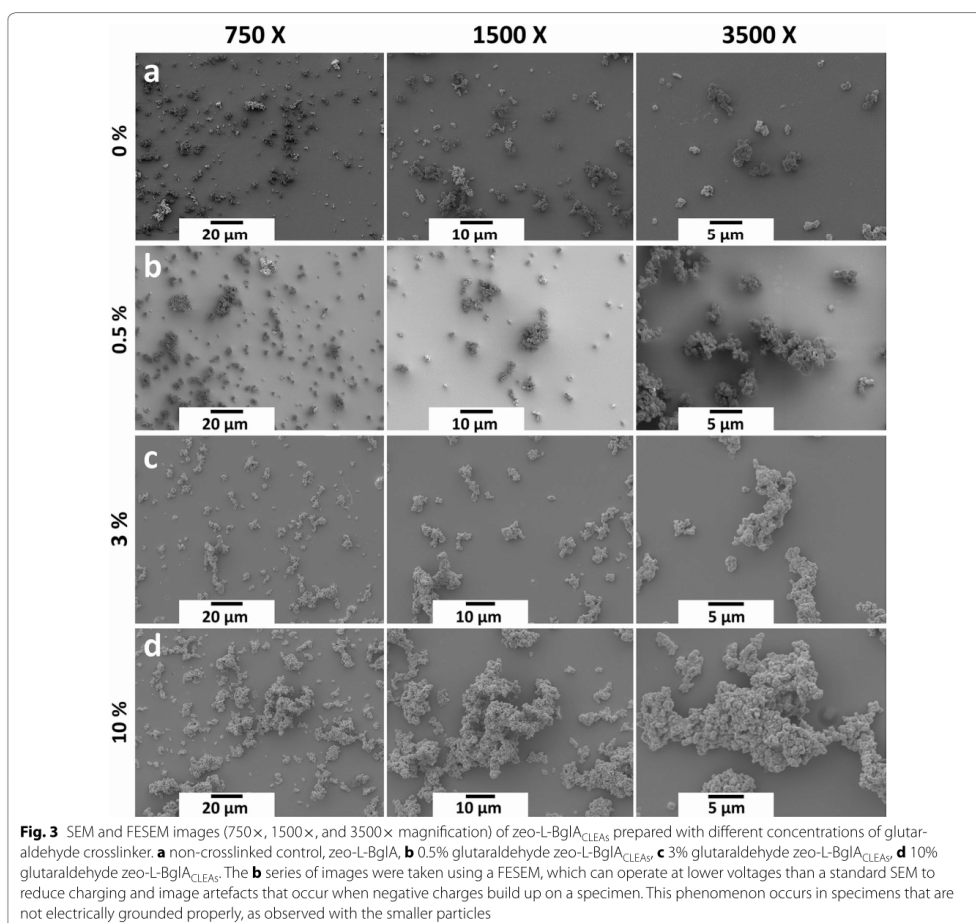
Immobilisation onto solid supports imparts mechanical and operational stability to CLEAs and allows their separation from the reaction mixture [32]. The possibility that zeo-L-enzymes could be stabilised further by forming CLEAs was examined as follows. First, a zeo-L-enzyme of BglA was prepared using the standard binding assay without the final SDS heat elution step. Then, covalent crosslinks were introduced by treatment with glutaraldehyde to form a zeo-L-enzyme_{CLEAs} (Fig. 2d). Changes in the crosslinker concentration is known to influence CLEA particle size. Therefore, it was expected that CLEA particles would increase in size with increasing concentration of glutaraldehyde as a result of enhanced crosslinking between protein-coated particles. This notion was tested by crosslinking zeo-L-BglA samples with different glutaraldehyde concentrations and comparing the resulting zeo-L-BglA_{CLEAs} particles by scanning electron microscopy (SEM) or field emission SEM (FESEM). Without crosslinking, the particles of the zeo-L-enzymes were well dispersed and less than ~2.5 µm in size (Fig. 3a). Small increases in particle size could be observed after crosslinking with 0.5% glutaraldehyde (Fig. 3b). At 3% glutaraldehyde (Fig. 3c), large particles ~20 µm in size were produced, and at 10% glutaraldehyde, bulky amorphous particles were obtained with dimensions that exceeded ~40 µm (Fig. 3d). Accordingly, higher glutaraldehyde concentrations initiated a greater degree of crosslinking between the protein coatings of zeo-L-BglA samples, resulting in the formation of larger zeo-L-BglA_{CLEAs}.

Enzymatic activity of zeo-L-enzyme_{CLEAs}

A successful crosslinking reaction should prevent elution of L-enzymes from the zeolite matrix in the zeo-L-enzyme_{CLEAs} through the strength of the covalent bonds between the enzyme molecules. Various solutions have been formulated for the elution of proteins immobilised onto zeolite via the linker technology [22].

Several of these solutions were tested for their ability to release L-enzymes from zeolite in their active form and 1 M betaine (pH 6) was determined to be the most suitable elution agent (data not shown). The crosslinking reaction was verified by determining the glutaraldehyde concentration required for the effective formation of an enzymatically active zeo-L-enzyme_{CLEAs} and the enzyme activity profile of the resulting zeo-L-enzyme_{CLEAs} was measured after treatment with 1 M betaine. Three different glutaraldehyde concentrations were tested using L-BglA. After the zeo-L-BglA_{CLEAs} were washed with phosphate buffer, 100 µl of 1 M betaine was added and incubated at room temperature for 5 min with rotation. This step was repeated three times to allow for enzyme elution. Standard enzyme activity assays on all fractions and the fractions that displayed activity are shown in Fig. 4. The relative activity of each fraction was calculated as a percentage of the starting material (100%). All the activity of the 10% glutaraldehyde CLEAs was retained in the CLEAs fraction and none was present in the 1 M betaine elution fractions. CLEAs prepared with the lower concentrations of glutaraldehyde, 0.5 and 3%, displayed residual enzyme activities of 55 and 20% less than the 10% glutaraldehyde CLEAs fraction. The remaining activity of the lower glutaraldehyde concentrations was found in the betaine elution fractions. This result indicated that at 0.5 and 3% glutaraldehyde linker-enzymes were not efficiently crosslinked and were consequently eluted from the zeo-L-enzyme_{CLEAs}. Although the effective formation of CLEAs by the addition of 0.5 and 3% glutaraldehyde has been reported previously [33], our results showed that these concentrations were insufficient for protein crosslinking and led to enzyme leaching from the CLEAs. In comparison, 10% glutaraldehyde was the most effective concentration, providing a sufficient degree of protein–protein crosslinking that ensured no leaching occurred. Accordingly, 10% (v/v) glutaraldehyde was used for producing CLEAs in all subsequent experiments (including CLEAs from ethanol-precipitated free enzymes).

The process was repeated with L-XynB and ManA-L once 10% was established as the most suitable glutaraldehyde concentration for crosslinking. Activity assays also were performed on all fractions with the resulting zeo-L-XynA_{CLEAs} and zeo-ManA-L_{CLEAs} as described above (results not shown) and protein bands from the different fractions were visualised using SDS-PAGE and Coomassie staining (Fig. 5). Corresponding zeo-L-enzymes (zeolite-bound but not crosslinked) were treated similarly with 1 M betaine elution and used as controls. As shown in Fig. 5, without crosslinking (zeo-L-enzymes), all of the L-enzymes were released from the surface of zeolite upon

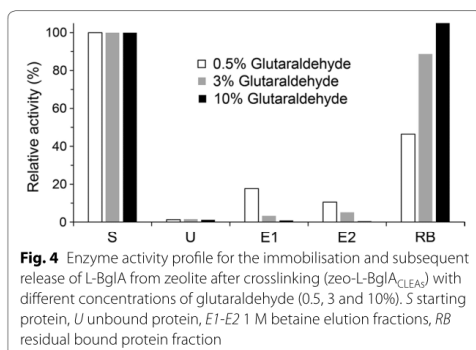


the addition of the elution agent with the main portion of the released protein located in the first elution fraction (E1) and lesser amounts found in E2. In contrast, the L-enzymes were not eluted from the zeo-L-enzyme_{CLEAs} after betaine treatment with no proteins observed in both the E1 and E2 fractions.

Characterisation of CLEAs

Both CLEAs and carriers are known to enhance the stability and maintain the activity of enzymes [34–36]. The activities of each thermophilic enzyme were compared as CLEAs prepared from free enzymes without (enzyme_{CLEAs}) and with (L-enzyme_{CLEAs}) linker and from L-enzymes bound to zeolite (zeo-L-enzyme_{CLEAs}).

As shown in Table 3, there was no variation in the optimal pH of the XynB and ManA enzyme_{CLEAs} and L-enzyme_{CLEAs} preparations. However, an increase in pH optimum for activity was observed in BglA prepared similarly. The zeo-L-enzyme_{CLEAs} preparations of L-BglA and L-XynB showed a decrease in their pH optimum when compared to their respective L-enzyme_{CLEAs} whereas the zeo-ManA-L_{CLEAs} showed an increase in pH optimum for activity when compared directly to its L-enzyme_{CLEAs} preparation. The CLEAs preparations of all three enzymes displayed an apparent optimum temperature for activity of 80 °C. In the case of BglA and ManA CLEAs preparations, this is 5 °C below the optimum temperature for activity measured for their respective free enzyme with and without linker (Table 2).



Recycling of immobilised enzymes

Enzyme immobilisation allows the easy recovery of both enzymes and products, thus significantly reducing overall production costs in biocatalytic processes [31]. The re-usability of the three enzymes was

evaluated by using repeated cycles of activity and recovery to determine the degree to which the various CLEAs preparations (enzyme_{CLEAs}, L-enzyme_{CLEAs}, and zeo-L-enzyme_{CLEAs}) and the zeolite-bound enzymes (zeo-L-enzymes) improved activity retention during enzyme recycling (Fig. 6). After 12 cycles of 10 min each, the CLEAs formed from the L-enzymes and zeo-L-enzymes were more resistant to denaturation and retained greater enzyme activity than the enzymes formulated as CLEAs or as zeolite-bound enzymes (zeo-L-enzymes). The zeo-L-BglA_{CLEAs} (Fig. 6a) and zeo-L-XynB_{CLEAs} (Fig. 6b) retained over 60% of their initial activity at the end of the 12 rounds of recycling while zeo-ManA-L_{CLEAs} (Fig. 6c) retained over 40% of its initial activity.

In general, the L-enzyme_{CLEAs} were the next most resistant and stable preparations after the full recycling period with L-BglA_{CLEAs} and L-XynB_{CLEAs} retaining 50% of their initial activity while ManA-L_{CLEAs} retained 30%. In most cases, the CLEAs prepared from the free enzymes without linker (enzyme_{CLEAs}) were the least stable preparations with XynB_{CLEAs} and ManA_{CLEAs} losing over 80% of their initial activity after only one

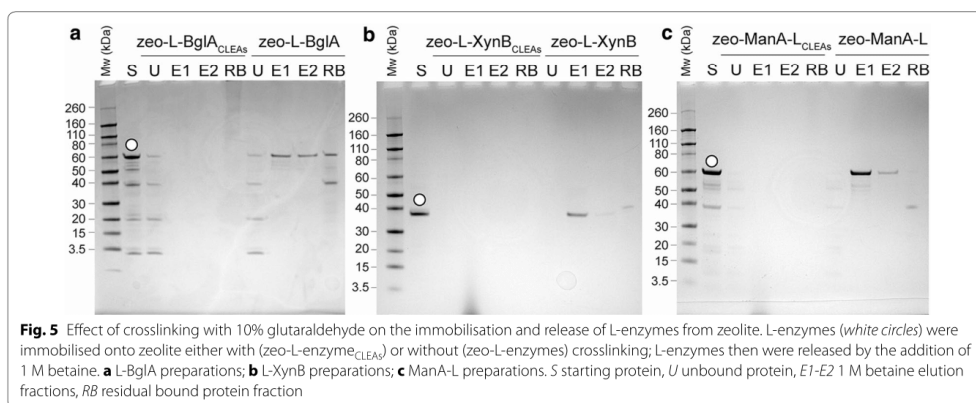
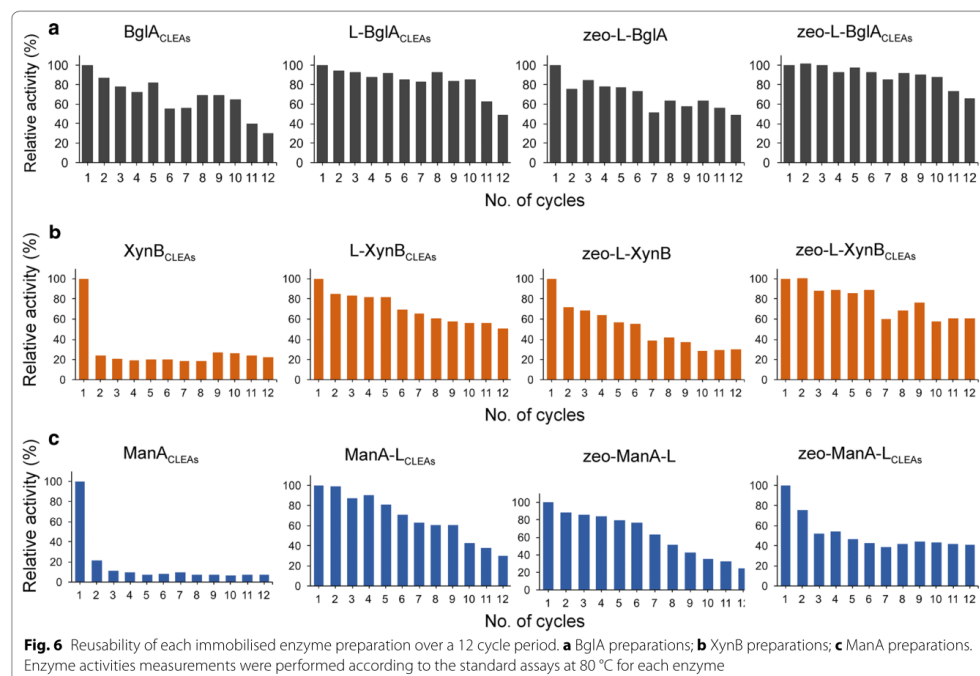


Table 3 Effect of forming CLEAs on the optimal pH and temperature of the thermostable enzymes

	BglA	L-BglA	XynB	L-XynB	ManA	ManA-L
Optimum pH						
enzyme _{CLEAs}	6.0		6.5		6.0	
L-enzyme _{CLEAs}		7.0		6.5		6.0
zeo-L-enzyme _{CLEAs}		5.0		6.0		6.5
Optimum temperature (°C)						
enzyme _{CLEAs}	80		80		80	
L-enzyme _{CLEAs}		80		80		80
zeo-L-enzyme _{CLEAs}		80		80		80



incubation cycle. It has been reported previously that the recombinant ManA enzyme exists in both a full and truncated form [20]. Thus, the rapid loss of activity for the ManA_{CLEAs} may be due to an increase in ManA truncation and subsequent release and/or inactivation of the enzyme during recycling. The same CLEAs preparation of BglA (BglA_{CLEAs}), on the contrary, was stable for several cycles and still retained 30% of its initial activity after 12 cycles. All of the zeolite-bound enzyme (zeo-L-enzymes) preparations were stable during the recycling period and displayed higher stability profiles than their respective enzyme_{CLEAs} counterparts. The zeo-ManA-L was the next most stable of the ManA preparations, outperforming the zeo-ManA-L_{CLEAs} for the first 8 cycles of the testing period. In general, a combination of linker-mediated zeolite binding and glutaraldehyde crosslinking to form the zeo-L-enzyme_{CLEAs} preparations was the most effective method with which to immobilise these thermostable enzymes as they provided greater enzyme activity retention during the complete recycling period.

Assembly of single and multiple enzyme biocatalytic modules

Recently, it was shown that CLEAs are able to catalyse a sequence of reactions. These forms of CLEAs have been called combi-CLEAs by some authors [37]. In combi-CLEAs, heterogeneous populations of enzymes can be confined simultaneously in the same aggregate and they perform a versatile cascade or non-cascade bioconversion. For example, thermostable CLEAs of L-arabinose isomerase from *Thermoanaerobacter mathranii* and a β -glycosidase from *Sulfolobus solfataricus* were immobilised by crosslinking with glutaraldehyde and polyethylenimine [38]. In combination, these CLEAs were shown to mediate the two-step conversion of D-galactose to D-tagatose at 65 °C in a single reaction with a 38% conversion.

Recently, the use of thermostable enzymes immobilised on solid supports has been presented as an alternative approach to produce biocatalytic modules for low-cost biomanufacturing [17, 39, 40]. Biocatalytic modules can be constructed and used to replace industrial chemical

catalysts due to their high specificity and high efficiency under mild reaction conditions [17]. In this context, zeo-L-enzyme_{CLEAs} can be considered to be biocatalytic modules with immense potential for industrial applications. The biocatalytic module design relies on the affinity of the linker for low-cost bulk silica-based matrices with additional introduction of crosslinking of the immobilised enzyme. Thus, single biocatalytic modules can be prepared easily consisting of one type of enzyme immobilised per matrix as well as multiple enzyme modules (Fig. 7a).

To demonstrate the potential application of the multiple enzymes module concept, all three enzymes without (BglA, XynB and ManA) and with (L-BglA, L-XynB and ManA-L) the linker were mixed in equal amounts, bound to zeolite using the standard zeolite binding assay and then analysed by SDS-PAGE (Fig. 7b). As expected, all of the L-enzymes exhibited high binding affinity towards zeolite as seen in single binding assays (Fig. 1) with most of the initial protein remaining in the bound fraction. In the zeolite binding assays with the mixed enzymes without the linker, the majority of each of the enzymes was observed in the unbound fraction.

After confirmation that all three mixed L-enzymes were successfully immobilised simultaneously onto zeolite, single and multiple enzyme biocatalytic modules were prepared as described above for zeo-L-enzyme_{CLEAs}. Single biocatalytic modules were tested against all three substrates and no module showed significant activity against any substrate except that expected for its constituent enzyme (Fig. 7c). The activity of the multiple enzymes module for each substrate was measured as a percentage of activity of the corresponding single module. The multiple enzymes module showed activity against all three substrates with comparable activity to the single module (L-BglA) against the pNPGlu substrate but only 44 and 59% of the corresponding single module activity against OSX and LBG substrates, respectively.

Discussion

A silica-binding linker peptide was fused genetically to three thermostable polysaccharide-degrading enzymes for applications in industrial-scale biocatalysis (Table 1). The linker showed high affinity for silica-containing support materials, allowing for directional immobilisation of the three enzymes onto a zeolite matrix (Fig. 1). The enzymes as partially-purified fusion proteins retained both biological activity and binding affinity for zeolite. These results were independent of the position of the linker (N- or C-terminal). The integration of the linker did not have negative effects on the pH optima but lowered the temperature stability of the enzymes (Table 2). SBPs have been shown to have varying effects on the

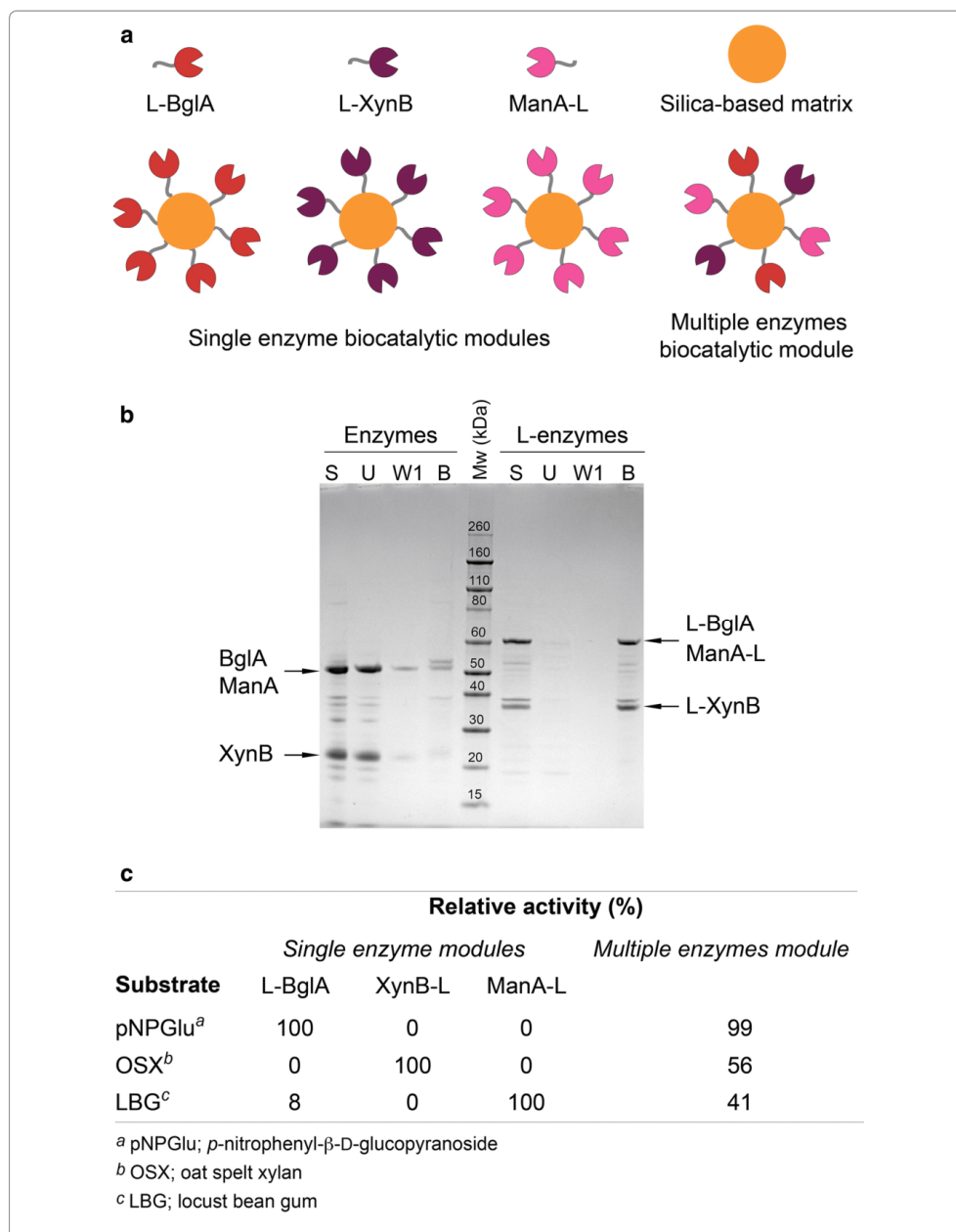
secondary structure of their fusion partners [11]. Hence, a decrease in thermostability may be explained by small changes in the structural conformation of the enzyme after linker fusion, resulting in a loss of thermal stability at high temperatures.

The L-enzymes in several conformations, including being bound to zeolite, were crosslinked individually with glutaraldehyde to form CLEAs and their general properties and activities were essentially unchanged in all cases (Fig. 2; Table 3). For example, protein was not seen in the 1 M betaine elution fractions of the zeo-L-enzyme_{CLEAs}, indicating that the crosslinking was effective and no enzyme leakage was observed. The absence of residual bound protein in the SDS-PAGE visualisation of the zeo-L-enzyme_{CLEAs} can be attributed to SDS at high temperatures failing to elute the enzymes due to the presence of the covalent bonds. Additionally, it was found that 0.5 and 3% glutaraldehyde concentrations were not sufficient for the crosslinking reaction to proceed to completion, as evidenced by the activity seen in the elution fractions in Fig. 5. However, our optimisation showed that 10% glutaraldehyde was the most effective concentration for immobilisation.

Previous research comparing CLEAs to native enzymes has shown that enzyme crosslinking broadens the pH values at which the enzyme is active, as well as increasing total activity in U/g/ml. CLEAs also were shown to increase the temperature range at which the enzyme was active [41]. The broadening of activity ranges seen in CLEAs (Table 3) was due to the stabilising nature of the crosslinking reaction. When the enzymes were crosslinked, they were in their native form and the covalent bonds of the crosslinking reaction protected the enzyme from any tertiary structural distortion caused by heat. This effect was shown further by the lack of optimum temperature broadening seen in the free enzymes between the linker and non-linker enzymes, indicating that it was the crosslinking and not the presence of the linker that was responsible for the increased activity range at sub-optimal pH and temperature conditions.

CLEAs made using enzymes immobilised prior to crosslinking display higher activity than both free enzymes and CLEAs made from precipitated free enzymes (Fig. 7) [42]. Research on xylanolytic enzymes immobilised onto silanised magnetic particles and crosslinked showed that the magnetic CLEAs displayed superior thermostability and activity in comparison to both free xylanase and precipitated free enzyme CLEAs.

An advantage of using the enzymes as CLEAs is that it avoids expensive solid supports. However, the linker enables immobilisation onto inexpensive bulk silica-based materials as well as allowing additional properties to be added to the support, for example by employing



(See figure on previous page.)

Fig. 7 Design, formation and enzymatic activity of biocatalytic modules. **a** Illustration showing the linker-mediated formation of biocatalytic modules consisting of single or multiple types of enzymes; **b** formation of multiple enzyme modules. SDS-PAGE analysis of zeolite binding assays performed with mixtures containing equal amounts of all three enzymes or L-enzymes. *S* starting protein sample, *U* unbound protein fraction, *W1* wash fractions (1 of 3 total), *B* zeolite-bound protein fraction; **c** relative enzymatic activity of single enzyme and multiple enzyme modules against different substrates

magnetic silica beads. The linker also adds additional Lys residues for amine-reactive crosslinking and thus provides a benefit to proteins possessing a low Lys content for the crosslinking reaction. It provides directionality for immobilisation so that the individual enzyme molecules are uniformly orientated rather than being bound randomly [43] and eliminates the need for protein precipitants as in the example of zeo-L-enzyme_{CLEAs}.

In this work, three thermostable glycoside-hydrolysing enzymes were selected as a model system because there is sufficient published data to allow us to compare the performance of our system against studies reported previously [24, 25]. Immobilisation of polysaccharide-degrading enzymes has significant economic and environmental benefits due to its potential for industrial-scale monosaccharide production from minimally-treated plant biomass. Downstream applications include fermentation for cellulosic biofuels, a renewable resource with a much smaller carbon footprint than both traditional petroleum-based fuels and starch-based biofuels [43]. Additionally, the use of cellulosic materials for biofuel production does not compete with food production, unlike starch-based biofuels [44]. The cost of enzymes is one of the most significant expense associated with industrial biocatalysis. Enzyme immobilisation allows the easy recovery of both enzymes and products, multiple reuse of enzymes and the continuous operation of enzymatic processes, thus significantly reducing overall production costs that are frequently a limiting factor in biomass conversion and biocatalytic processes. Enhanced operational performance and the reuse of immobilised enzymes leads to greater catalytic productivity, which reduces the enzyme cost per kg of end product [31]. Large-scale industrial applications require enzyme immobilisation to be simple, without the need for highly purified enzyme preparations or expensive solid supports that limit the commercial viability of the bioprocess. Industrial operations using enzyme immobilisation on inexpensive silica-based supports (e.g. zeolite) combined with enzyme reuse have the capacity to reduce operating and production costs as well as improving the yield of a process.

The zeo-L-enzyme_{CLEAs} of all enzymes display the most activity retention after twelve 10-min cycles, equivalent to 2 h spent at the optimum temperature (Fig. 6). The

linker contains 23% Lys or Arg residues, which participate in the crosslinking reaction. The stabilisation provided by these extra residues is a possible reason as to why the zeo-L-enzyme_{CLEAs} are the most stable over 12 recycling assays. The stabilising effect provided by the linker also could provide an explanation as to why the enzyme_{CLEAs} without linkers were the least recyclable of the enzyme preparations. Without the extra stabilisation provided by the linker, the enzyme_{CLEAs} could be more prone to denaturation and loss of activity than immobilised enzymes with linkers. Zeo-L-enzymes display better activity retention than the enzyme_{CLEAs}, but showed less recyclability than the L-enzyme_{CLEAs} with and without zeolite. This result implies that the crosslinking reaction plays an important role in stabilising the immobilised enzymes against heat-induced denaturation over subsequent uses. Each of the single zeo-L-enzyme_{CLEAs} showed activity against the substrate of the constituent enzyme and negligible activity towards the other two substrates (Fig. 7c). In contrast, the mixed zeo-L-enzyme_{CLEAs} showed activity towards all three substrates at much lower levels than the single zeo-L-enzyme_{CLEAs} (Fig. 7c).

These results have shown that it is possible for multiple thermostable polysaccharide-degrading enzymes to be immobilised simultaneously onto the same support and retain specific hydrolytic activity. Crosslinking previously immobilised zeo-L-enzymes was the immobilisation method that shows the most potential for industrial applications. The zeo-L-enzyme_{CLEAs} have comparable activity to the free enzymes and superior activity to the enzyme_{CLEAs} with the added ability of being the most recyclable (i.e. retaining the most activity after multiple reactions) and the easiest to remove from the reaction mixture. This fact may have implications in the enzymatic breakdown of lignocellulosic substrates for industrial purposes as well as wider implications in further enzyme immobilisation for production of industrially-useful biomolecules.

In previous studies, cellulolytic enzymes have been immobilised using gold-coated nanoparticles immobilised on thiolated magnetic silica nanoparticles. Cysteine-tagged cellulases were bound to the gold nanoparticles using thiol-gold chemistry [45]. The yields of glucose and cellobiose were found to increase by 179 and 158%, respectively, when using immobilised cellulases

compared to the free ones. Stability and activity of the immobilised cellulases remained after four reuses, while the free enzymes dropped to 40% of their previous activity after a single reuse. Plasma immersion ion implantation (PIII) is a technique that reduces the time and amount of steps compared to conventional covalent immobilisation techniques. PIII involves extracting ions from plasma by applying a high voltage direct current which directs the ions to a polymeric substrate covered by a semiconductor wafer. The ions break bonds in the polymer chains and create free radicals that react covalently with the enzyme [46]. However, PIII treatment must be done under vacuum conditions and plasma is created by superheating gas or applying the gas in a strong electromagnetic field. Flat PIII-treated polystyrene was used as a support to immobilise a thermostable β -glucosidase and a commercially-available β -glucosidase. It was found that the thermostable β -glucosidase displayed higher activity when immobilised onto the PIII-treated polystyrene than when adsorbed onto untreated polystyrene [47]. The immobilised thermophilic β -glucosidase also showed 20 times higher activity than the immobilised commercially-available β -glucosidase. PIII has also been used to immobilise a thermophilic β -glucosidase on a curved surface. Using a plastic polymer granules support treated with PIII, a thermophilic β -glucosidase from *Caldicellulosiruptor saccharolyticus* was immobilised and retained activity for 2 weeks, compared to 6 days with untreated granules [48].

The immobilisation technologies described above are elegant and ideally suited for laboratory experiments but do not translate well into industrial-scale processes. Our system, based on low-cost bulk matrices and a simple immobilisation technology, is a viable alternative that requires further investigation and optimisation before application at an industrial scale. Further downstream applications of immobilised polysaccharide-degrading enzymes using this platform technology are in the industrial production of glucose from cellulose in minimally-treated plant biomass. The results reported here indicate that it is possible to set up a degradative pathway for biomass to produce fermentable sugars using inexpensive reagents and matrices that allow the reuse of the enzymatic component at high temperatures and over a number of cycles. In future studies, biocatalytic modules will be designed and constructed for the cell-free assembly of natural and non-natural pathways for the efficient degradation and conversion of biomass into higher-value products.

Conclusions

The linker technology (and potentially, other solid-binding peptides) represents a novel and inexpensive immobilisation strategy for industrially-significant enzymes. The

linker sequence has the capacity to impart directionality and orientation to the immobilisation of enzymes and allows enzyme immobilisation onto cheap bulk matrices for example, zeolite and silica. It provides highly accessible lysine/arginine residues (free amino groups) making it amenable to crosslinking reactions (such as glutaraldehyde) to form CLEAs. Furthermore, the linker system minimises the time spent selecting precipitants and crosslinking reagents. The structural and compositional properties of the linker allow it to impart directionality and orientation to the enzymes after crosslinking (L-enzyme_{CLEAs}) resulting in enhanced enzyme reusability. It may be instrumental in allowing crosslinking in a protein–protein only context to form a protein layer around zeolite particles that can enhance thermostability and enzyme activity. The L-enzyme CLEAs lend themselves to facile pathway formation and immobilisation for the optimal hydrolysis of diverse lignin-based biomass sources.

Methods

Construction of linker-enzymes

The gene encoding *Caldicellulosiruptor saccharolyticus* β -glucosidase A (*bglA*) was amplified with flanking restriction sites (*Bam*HI and *Hind*III) from pProEXTHe-cKanBglA [26] by polymerase chain reaction (PCR) using the primers (forward, 5'-aaaGGATCCGAGTTCCCAA AAGGA-3'; reverse, 5'-TTTAATCTGTATCAGGCTG A-3') and ligated into pET22b (Novagen) plasmids with and without an N-terminal linker sequence [16] to give the expression plasmids pET22-bglA and pET22-L-bglA, respectively.

Transformation of *Escherichia coli* competent cells

The expression plasmids pET22-L-bglA and pET22-L-xynB contained the linker sequence (L) genetically fused to the N-terminus of *Caldicellulosiruptor saccharolyticus* β -glucosidase A (*bglA*) and *Dictyoglomus thermophilum* β -xylanase B (*xynB*), respectively. The pET22-ManA-L expression plasmid contained the linker sequence fused to the C-terminus of *D. thermophilum* β -mannanase A (*manA*) [20]. The expression plasmids pet22-bglA, pJLA602-xynB [24], and pET22-manA [20] contained each enzyme without the linker sequence. All recombinant constructions and expression hosts are represented diagrammatically in Table 1.

Recombinant protein expression

Escherichia coli Tuner (DE3) or BL21 harbouring the expression plasmid of interest were cultured in Luria-Bertani (LB) medium supplemented with 50 μ g/ml carbenicillin for the production of all recombinant enzymes with and without the linker. Cells were grown at 37 °C

with shaking (250 rpm) until the A_{600} was between 0.7 and 0.9. Recombinant protein synthesis (except for XynB) was induced by the addition of 0.4 mM isopropyl β -D-thiogalactoside (IPTG). The expression plasmid pJLA602-XynB is heat-inducible and requires an increase in growth temperature to 42 °C to initiate the expression of XynB [49]. After a further 3–4 h of cultivation at 37 °C (42 °C for XynB), cells were harvested by centrifugation for 15 min at 4000 \times g at 4 °C and stored at –20 °C.

Cells were resuspended in ice-cold lysis buffer (25 mM Tris–HCl, pH 8.0, 100 mM NaCl, 1.25 mM EDTA and 0.05% Tween 20) supplemented with 4 mM of the serine protease inhibitor Pefabloc (Sigma) and 1.5 mg lysozyme (Sigma). Cells were disrupted by three passages through a French pressure cell and then 25 U Benzonase (Novagen) and a further 4 mM of Pefabloc were added. Cellular debris was removed by centrifugation for 30 min at 4000 \times g at 4 °C following incubation on ice for 20 min. The supernatant obtained was incubated for 30 min at 70 °C to denature host proteins and to partially purify the thermostable recombinant protein. Heat-denatured proteins were removed by centrifugation at 4000 \times g for 30 min at 4 °C. The supernatant obtained was filtered through a 0.2 μ m sterile filter (Millipore), supplemented with 0.05% sodium azide and stored at 4 °C. All partially purified protein concentrations were determined by measuring the absorbance of each sample at 280 nm on a Nanodrop 2000c spectrophotometer (Thermo Scientific).

Standard zeolite binding assay

The synthetic zeolite CBV-100 (Zeolyst International, USA) was used as the silica-based matrix in this study. Standard zeolite binding assays were performed as described previously [16]. Briefly, 5 mg of zeolite was washed three times with 500 μ l washing buffer (1% Triton X-100 in 50 mM phosphate buffer, pH 6.0). Soluble, partially purified enzyme (20 μ g in 100 μ l of phosphate buffer) was mixed with the washed zeolite and incubated with rotation for 30 min at room temperature. After centrifugation at 10,000 \times g for 30 s, the supernatant containing the unbound fraction was removed. The zeolite pellet then was washed three times with 100 μ l of 50 mM phosphate buffer (pH 6.0) by vortexing and centrifugation. Zeolite-bound proteins were eluted from the matrix using 100 μ l SDS-PAGE loading buffer and incubation at 99 °C for 10 min. Fractions containing protein were separated on NuPage 4–12% Bis–Tris Gels (Life Technologies) by SDS-PAGE and identified with Coomassie Brilliant Blue staining.

Standard enzyme assays

All reaction mixtures consisted of an appropriate amount of enzyme (20 μ l) and their respective substrate (80 μ l).

Reactions were prepared in triplicate and incubated at 80 °C for 10 min. Each enzyme was diluted to a concentration that was not substrate limited during the assay reaction. In controls, each enzyme was added after the termination of the reaction. For comparison and convenience, all enzyme activities are expressed as a percentage of the maximal activity.

β -Glucosidase activity was assayed using 0.5 mg/ml *p*-nitrophenyl- β -D-glucopyranoside (pNPGLuc, MP Bio-medicals) substrate in 40 mM universal buffer [50]. The assay reactions were terminated by addition of 100 μ l 1 M Na₂CO₃ and the absorbance of the released *p*-nitrophenol was measured at 405 nm using a PHERAstar FS 96-well plate reader (BMG Labtech, Germany).

β -Xylanase and β -mannanase activities were assayed using 0.5% (w/v) oat spelt xylan (OSX, Sigma) and locust bean gum (LBG, Sigma) substrates in 120 mM universal buffer, respectively. The reducing sugars released were measured using the dinitrosalicylic acid (DNS) assay procedure [51] and the absorbance increase generated by the release of 3-amino-5-nitrosalicylic acid was measured at 540 nm using a PHERAstar FS 96-well plate reader.

Activity characterisation of free enzymes

The effects of pH and temperature on the free enzyme activities were examined using the partially purified recombinant proteins. The effect of temperature on the reaction rate was determined by incubating the free enzyme with the appropriate substrate at temperatures ranging from 40 to 90 °C under standard assay conditions. To determine the optimal pH, the appropriate substrate was prepared in universal buffer and assayed at different pH values (all pH values were adjusted at the temperature of the assay) at the optimal temperature for activity. To determine the thermostability of the free enzymes, they were incubated between 70 and 90 °C, in the absence of substrate. Samples were removed at different time intervals and the residual enzyme activity was measured under standard enzyme assay conditions with an assay time of 10 min.

Enzyme immobilisation

Zeolite-bound linker-enzymes (zeo-L-enzyme)

Zeolite-bound linker-enzymes were prepared using the standard zeolite binding assay without the final SDS heat elution step. The resulting final insoluble zeolite-bound fraction was resuspended in 50 mM phosphate buffer (pH 6.0).

CLEAs of free enzymes (enzyme_{CLEAs} and linker-enzyme_{CLEAs})

CLEAs of each free enzyme (with and without linker) were prepared by a combined one-step ethanol precipitation and crosslinking strategy. Glutaraldehyde (0.5, 3.0

and 10% v/v) in chilled 100% ethanol to a total volume of 9 ml was added to 1 ml of each of the partially purified free enzymes and incubated overnight at 4 °C without shaking. The enzyme_{CLEAs} formed were centrifuged for 4000×g for 30 min at 4 °C, and the supernatant tested for residual enzyme activity. The absence of detectable activity in the supernatant was taken as an indication the aggregation and crosslinking had gone to completion. The enzyme_{CLEAs} were washed three times in 50 mM phosphate buffer (pH 6.0) to remove any unreacted glutaraldehyde.

CLEAs of zeolite-bound linker-enzymes (zeo-L-enzyme_{CLEAs})

Single enzyme biocatalytic modules: CLEAs of each zeolite-bound linker-enzyme were prepared using standard zeolite binding assays without the final SDS heat elution step (see Fig. 6a), followed by crosslinking with 100 µl glutaraldehyde (0.5, 3.0 and 10% v/v). After incubation at 4 °C with rotation for 1.5 h, the resulting zeo-L-enzyme_{CLEAs} were washed three times with 500 µl of 50 mM phosphate buffer (pH 6.0) to remove any unreacted glutaraldehyde.

Multiple enzyme biocatalytic modules: A mixed zeo-L-enzyme_{CLEAs} was prepared using an initial combined mixture containing 100 µl of each of the three linker-enzymes in a single zeolite standard binding assay (see Fig. 6a) as described above for Single enzyme biocatalytic modules. The resulting mixed zeo-L-enzyme_{CLEAs} were resuspended in 100 µl 50 mM phosphate buffer (pH 6.0). A 10 µl sample was wet-mounted onto a glass slide for light microscopy to ascertain that the crosslinking reaction had been successful.

Recycling immobilised enzymes

Recycling assays were performed with the four forms of the immobilised enzymes (enzyme_{CLEAs}, L-enzyme_{CLEAs}, zeo-L-enzymes and zeo-L-enzyme_{CLEAs}). Each enzyme was incubated with its appropriate substrate under standard assay conditions. After an incubation period of 10 min, each reaction was centrifuged briefly to pellet the immobilised enzyme and the supernatant obtained was transferred to a fresh tube for the addition of the stopping reagents DNS or Na₂CO₃. Fresh substrate was then added to each immobilised enzyme and the standard assay was repeated. The immobilised enzymes were recycled in this manner for a total of 12 cycles.

Multiple substrate hydrolysis

Each of the constituent enzymes in the mixed and individual zeo-L-enzyme_{CLEAs} were tested for activity against each substrate (OSX, LBG and pNPGlu) under the standard assay conditions to determine whether the

mixed zeo-L-enzyme_{CLEAs} were capable of carrying out their hydrolysis reactions. The relative activity of the mixed zeo-L-enzyme_{CLEAs} against each substrate was calculated as a percentage of the activity of the individual zeo-L-enzyme_{CLEAs} against the same substrate.

Scanning electron microscopy

Samples of zeo-L-BglA_{CLEAs} (5 µl) were snap frozen and lyophilised overnight onto electron microscopy coverslips and mounted onto adhesive coverslip stands. The samples were gently agitated with air to remove loose particles, after which they were coated in gold using an Emitech K550 Gold Sputter Coater Unit. A JEOL-SM-6480 scanning electron microscope was used to image the samples.

Abbreviations

AP: alkaline phosphatase; BglA: β-glucosidase; CLEAs: cross-linked enzyme aggregates; DNS: dinitrosalicylic acid; EDTA: ethylenediaminetetraacetic acid; enzyme_{CLEAs}: CLEAs of enzymes without the linker; FESEM: field emission scanning electron microscopy; GBP: gold-binding peptide; IPTG: isopropyl β-D-thiogalactoside; L: linker; LB: Luria-Bertani medium; L-enzymes: enzymes with the linker; L-enzyme_{CLEAs}: CLEAs of L-enzymes; LBG: locust bean gum; ManA: β-mannanase; OSX: oat spelt xylan; PIII: plasma immersion ion implantation; pNPGlu: p-Nitrophenyl-β-D-glucopyranoside; SBP: solid-binding peptide; SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis; SEM: scanning electron microscopy; zeo-L-enzymes: zeolite-bound L-enzymes; zeo-L-enzyme_{CLEAs}: CLEAs of zeolite-bound L-enzymes; XynB: β-xylanase.

Authors' contributions

AC, KP, ESYG and AS designed the research. AC, KP and ESYG performed the experiments and acquired the data. AC, KP, ESYG, PLB and AS analysed and interpreted the data. AC, PLB and AS drafted the manuscript. All authors read and approved the final manuscript.

Author details

¹ Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, Australia. ² ARC Centre of Excellence for Nanoscale BioPhotonics (CNBP), Macquarie University, Sydney, Australia. ³ Department of Molecular Medicine & Pathology, Medical School, University of Auckland, Auckland, New Zealand. ⁴ Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, Australia.

Acknowledgements

We thank Nicole Vella from the microscopy unit at Macquarie University for her assistance with electron microscopy.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Data will be made available upon request.

Consent for publication

All authors have approved the manuscript for publication.

Funding

This research was supported by a Macquarie University Vice-Chancellor's Innovation Fellowship (MQVCIF) to AS.

Received: 13 October 2016 Accepted: 19 January 2017

Published online: 02 February 2017

References

- Mateo C, Palomo JM, Fernandez-Lorente G, Guisan JM, Fernandez-Lafuente R. Improvement of enzyme activity, stability and selectivity via immobilization techniques. *Enzyme Microb Technol*. 2007;40:1451–63.
- Raghuvanshi S, Gupta R. Advantages of the immobilization of lipase on porous supports over free enzyme. *Protein Pept Lett*. 2010;17:1412–6.
- Kumar G, Mudhoo A, Sivagurunathan P, Nagarajan D, Ghimire A, Lay C-H, Lin C-Y, Lee D-J, Chang J-S. Recent insights into the cell immobilization technology applied for dark fermentative hydrogen production. *Bioreour Technol*. 2016;219:725–37.
- Camarero JA. Recent developments in the site-specific immobilization of proteins onto solid supports. *Pept Sci*. 2008;90:450–8.
- Hanefeld U, Gardossi L, Magner E. Understanding enzyme immobilisation. *Chem Soc Rev*. 2009;38:453–68.
- Mohamad NR, Marzuki NHC, Buang NA, Huyop F, Wahab RA. An overview of technologies for immobilization of enzymes and surface analysis techniques for immobilized enzymes. *Biotechnol Biotechnol Equip*. 2015;29:205–20.
- Datta S, Christena LR, Rajaram YRS. Enzyme immobilization: an overview on techniques and support materials. *3 Biotech*. 2013;3:1–9.
- Care A, Bergquist PL, Sunna A. Solid-binding peptides: smart tools for nanobiotechnology. *Trends Biotechnol*. 2015;33:259–68.
- Tang Z, Palafox-Hernandez JP, Law W-C, Hughes ZE, Swihart MT, Prasad PN, et al. Biomolecular recognition principles for bionanocombinatorics: an integrated approach to elucidate enthalpic and entropic factors. *ACS Nano*. 2013;7:9632–46.
- Cetinel S, Caliskan HB, Yucseyot DT, Donatan AS, Yuca E, Urgen M, et al. Addressable self-immobilization of lactate dehydrogenase across multiple length scales. *Biotechnol J*. 2013;8:262–72.
- Chen X, Wang Y, Wang P. Peptide-induced affinity binding of carbonic anhydrase to carbon nanotubes. *Langmuir*. 2015;31:397–403.
- Kacar T, Zin MT, So C, Wilson B, Ma H, Gul-Karaguler N, et al. Directed self-immobilization of alkaline phosphatase on micro-patterned substrates via genetically fused metal-binding peptide. *Biotechnol Bioeng*. 2009;103:696–705.
- Yang M, Choi BG, Park TJ, Heo NS, Hong WH, Lee SY. Site-specific immobilization of gold binding polypeptide on gold nanoparticle-coated graphene sheet for biosensor application. *Nanoscale*. 2011;3:2950–6.
- Pierre A. The sol-gel encapsulation of enzymes. *Biocatal Biotransform*. 2004;22:145–70.
- Ho L-F, Li S-Y, Lin S-C, Hsu W-H. Integrated enzyme purification and immobilization processes with immobilized metal affinity adsorbents. *Process Biochem*. 2004;39:1573–81.
- Sunna A, Chi F, Bergquist PL. A linker peptide with high affinity towards silica-containing materials. *N Biotechnol*. 2013;30:485–92.
- Zhang Y-HP, Myung S, You C, Zhu Z, Rollin JA. Toward low-cost biomanufacturing through in vitro synthetic biology: bottom-up design. *J Mater Chem*. 2011;21:18877–86.
- Sivagurunathan P, Kumar G, Sen B, Lin C-Y. Development of a novel hybrid immobilization material (HY-IM) for fermentative biohydrogen production from beverage wastewater. *J Chin Chem Soc*. 2014;61:827–30.
- Care A, Chi F, Bergquist PL, Sunna A. Biofunctionalization of silica-coated magnetic particles mediated by a peptide. *J Nanopart Res*. 2014;16:1–9.
- Care A, Nevalainen H, Bergquist PL, Sunna A. Effect of *Trichoderma reesei* proteinases on the affinity of an inorganic-binding peptide. *Appl Biochem Biotechnol*. 2014;173:2225–40.
- Liang L, Care A, Zhang R, Lu Y, Packer NH, Sunna A, et al. Facile assembly of functional upconversion nanoparticles for targeted cancer imaging and photodynamic therapy. *ACS Appl Mater Interfaces*. 2016;8:11945.
- Sunna A, Chi F, Bergquist PL. Efficient capture of pathogens with a zeolite matrix. *Parasitol Res*. 2013;112:2441–52.
- Rosales-Calderon O, Trajano HL, Duff SJ. Stability of commercial glucanase and β -glucosidase preparations under hydrolysis conditions. *PeerJ*. 2014;2:e402.
- Morris DD, Gibbs MD, Chin CW, Koh M-H, Wong KK, Allison RW, et al. Cloning of the *xynB* Gene from *Dictyoglomus thermophilum* Rt46B. 1 and action of the gene product on kraft pulp. *Appl Environ Microbiol*. 1998;64:1759–65.
- Gibbs MD, Reeves RA, Sunna A, Bergquist PL. Sequencing and expression of a β -mannanase gene from the extreme thermophile *Dictyoglomus thermophilum* Rt46B. 1, and characteristics of the recombinant enzyme. *Curr Microbiol*. 1999;39:351–7.
- Hardiman E, Gibbs M, Reeves R, Bergquist P. Directed evolution of a thermophilic β -glucosidase for cellulosic bioethanol production. *Appl Biochem Biotechnol*. 2010;161:301–12.
- Hong M-R, Kim Y-S, Park C-S, Lee J-K, Kim Y-S, Oh D-K. Characterization of a recombinant β -glucosidase from the thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *J Biosci Bioeng*. 2009;108:36–40.
- Bai A, Zhao X, Jin Y, Yang G, Feng Y. A novel thermophilic β -glucosidase from *Caldicellulosiruptor bescii*: characterization and its synergistic catalysis with other cellulases. *J Mol Catal B Enzym*. 2013;85:248–56.
- Love DR, Streiff MB. Molecular cloning of a β -glucosidase gene from an extremely thermophilic anaerobe in *E. coli* and *B. subtilis*. *Nat Biotechnol*. 1987;5:384–7.
- Sheldon RA. Characteristic features and biotechnological applications of cross-linked enzyme aggregates (CLEAs). *Appl Microbiol Biotechnol*. 2011;92:467–77.
- Sheldon RA, van Pelt S. Enzyme immobilisation in biocatalysis: why, what and how. *Chem Soc Rev*. 2013;42:6223–35.
- Talekar S, Ghodake V, Ghotage T, Rathod P, Deshmukh P, Nadar S, et al. Novel magnetic cross-linked enzyme aggregates (magnetic CLEAs) of alpha amylase. *Bioreour Technol*. 2012;123:542–7.
- Yu H, Chen H, Wang X, Yang Y, Ching C. Cross-linked enzyme aggregates (CLEAs) with controlled particles: application to *Candida rugosa* lipase. *J Mol Catal B Enzym*. 2006;43:124–7.
- Cao L, van Rantwijk F, Sheldon RA. Cross-linked enzyme aggregates: a simple and effective method for the immobilization of penicillin acylase. *Org Lett*. 2000;2:1361–4.
- Sheldon R, Schoevaart R, Van Langen L. Cross-linked enzyme aggregates (CLEAs): a novel and versatile method for enzyme immobilization (a review). *Biocatal Biotransformation*. 2005;23:141–7.
- Cui JD, Jia SR. Optimization protocols and improved strategies of cross-linked enzyme aggregates technology: current development and future challenges. *Crit Rev Biotechnol*. 2015;35:15–28.
- Taboada-Puig R, Junghanns C, Demarche P, Moreira M, Feijoo G, Lema J, et al. Combined cross-linked enzyme aggregates from versatile peroxidase and glucose oxidase: production, partial characterization and application for the elimination of endocrine disruptors. *Bioreour Technol*. 2011;102:6593–9.
- Jørgensen F, Hansen O, Stougaard P. Enzymatic conversion of D-galactose to D-tagatose: heterologous expression and characterisation of a thermostable L-arabinose isomerase from *Thermoanaerobacter mathranii*. *Appl Microbiol Biotechnol*. 2004;64:816–22.
- You C, Zhang Y-HP. Cell-free biosystems for biomanufacturing. *Future trends in biotechnology*. Berlin: Springer; 2012. p. 89–119.
- Care A, Bergquist PL, Sunna A. Solid-binding peptides: Immobilisation strategies for extremophile biocatalysis in biotechnology. *Biotechnology of extremophiles*. Berlin: Springer; 2016. p. 637–74.
- Sangeetha K, Abraham TE. Preparation and characterization of cross-linked enzyme aggregates (CLEA) of subtilisin for controlled release applications. *Int J Biol Macromol*. 2008;43:314–9.
- Bhattacharya A, Pletschke BI. Magnetic cross-linked enzyme aggregates (CLEAs): a novel concept towards carrier free immobilization of lignocellulolytic enzymes. *Enzyme Microb Technol*. 2014;61:17–27.
- Larson ED. A review of life-cycle analysis studies on liquid biofuel systems for the transport sector. *Energy Sustain Dev*. 2006;10:109–26.
- Carroll A, Somerville C. Cellulosic biofuels. *Annu Rev Plant Biol*. 2009;60:165–82.
- Cho EJ, Jung S, Kim HJ, Lee YG, Nam KC, Lee H-J, et al. Co-immobilization of three cellulases on Au-doped magnetic silica nanoparticles for the degradation of cellulose. *Chem Commun*. 2012;48:886–8.
- Tran CT, Nosworthy NJ, Kondyurin A, McKenzie DR, Bilek MM. CelB and β -glucosidase immobilization for carboxymethyl cellulose hydrolysis. *RSC Adv*. 2013;3:23604–11.
- Hirsh S, Nosworthy N, Kondyurin A, dos Remedios C, McKenzie D, Bilek M. Linker-free covalent thermophilic β -glucosidase functionalized polymeric surfaces. *J Mater Chem*. 2011;21:17832–41.

48. Nosworthy NJ, Kondyurin A, Bilek MM, McKenzie DR. Ion implantation treatment of beads for covalent binding of molecules: application to bioethanol production using thermophilic beta-glucosidase. *Enzyme Microb Technol*. 2014;54:20–4.
49. Schauder B, Blöcker H, Frank R, McCarthy JE. Inducible expression vectors incorporating the *Escherichia coli* atpE translational initiation region. *Gene*. 1987;52:279–83.
50. Britton HTS, Robinson RA. CXCVIII.—Universal buffer solutions and the dissociation constant of veronal. *J Chem Soc*. 1931:1456–62.
51. Bailey MJ, Biely P, Poutanen K. Interlaboratory testing of methods for assay of xylanase activity. *J Biotechnol*. 1992;23:257–70.

Submit your next manuscript to BioMed Central
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



Appendix B of this thesis has been removed as it may contain sensitive/confidential content