TECHNISCHE UNIVERSITÄT MÜNCHEN
Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt

Lehrstuhl für Chemie Biogener Rohstoffe

Pichia pastoris as Enzyme Production Host: from Toolkit Design to Protein Expression

Ulrike Obst

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Wilfried Schwab
Prüfender der Dissertation:	1. Prof. Dr. Volker Sieber
	2. Dr. Tom Ellis, Imperial College, London
	3. Prof. Dr. Wolfgang Liebl

Die Dissertation wurde am 21.11.2018 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 15.04.2019 angenommen.

Abstract

Biological cell factories are capable of producing a wide range of products covering fuels to pharmaceuticals without the need for harsh chemistries. They support more environmentally friendly industrial processes and are a shift towards a bio-based economy. Currently, the development of cell factories is time consuming and laborious. A major step is ensuring that proteins encoding key enzymatic steps are sufficiently expressed. However, it is difficult to predict how regulatory elements and proteins will behave when used together in new ways. Furthermore, while yeasts like *Pichia pastoris* are a well-established host for industrial applications, they lack large libraries of genetic parts and efficient methodologies to identify optimal protein expression conditions in a systematic manner. This thesis addresses these limitations by developing an end-to-end platform for *P. pastoris* strain design with a focus on protein expression and secretion.

It begins by developing a toolkit of standardised genetic parts, which can be easily composed using a hierarchical assembly method to build large and diverse *P. pastoris* protein expression libraries. A toolkit of 42 parts enabling the control of gene expression, protein secretion and whether the genetic material is located on a self-replicating plasmid or integrated into the *P. pastoris* genome was made. To characterise the performance of the regulatory elements, 242 strains expressing two different fluorescent reporter proteins (RFP and yEGFP) in many different ways were constructed. Intracellular RFP and yEGFP expression were similar for the same combinations of promoters, terminators and whether expressed from a plasmid or genomically integrated. In contrast, secretion efficiency varied up to 100-fold between the fluorescent proteins for identical regulatory elements, suggesting that optimising the regulation and secretion tag for one type of protein will not translate to others.

Next, this thesis addresses the challenge of protein secretion by using the toolkit to find the optimal combination of regulatory elements and protein secretion tags to express and secrete several industrially relevant proteins. The necessary foundation to make effective randomised libraries and then search for optimal secretion conditions was established. The developed approach found new combinations of genetic parts that lead to improved secretion of several enzymes previously studied in *P. pastoris*. Attempts were also made to determine optimal expression conditions for several predicted unspecific peroxygenases and bacterial alcohol dehydrogenases, which have yet to be expressed in *P. pastoris*. Unfortunately, the performed screenings of these libraries did not identify the secretion of any functional enzymes.

Finally, to allow for the high-throughput screening of large and diverse libraries, it was attempted to develop a novel technique that encapsulated single *P. pastoris* cells in polyelectrolyte capsules. The approach works by integrating a single cell into a solid template, which is subsequently coated with polyelectrolyte layers before the template is dissolved. Despite thorough characterisation of the template formation process and the testing of a variety of conditions, it was not possible to encapsulate single cells using this approach.

The platform created in this work is a valuable resource for future strain development projects. The flexible design of the expression toolkit allows for customisation and new extensions to tackle future challenges in protein expression and secretion in yeast, and when combined with high-throughput screening technologies enables the efficient optimisation of cell factories.

Zusammenfassung

Biologische Zellfabriken können eine große Auswahl an Produkten wie zum Beispiel Treibstoffe oder Pharmazeutika nachhaltig herstellen. Sie unterstützen umweltfreundliche Industrieprozesse und damit einen Wandel in Richtung einer biobasierten Wirtschaft. Die Entwicklung von Zellfabriken ist derzeit zeitaufwändig und arbeitsintensiv. Ein wichtiger Entwicklungsschritt ist die ausreichende Expression von Proteinen für die Katalyse von Schlüsselreaktionen. Das Zusammenwirken von regulatorischen Elementen und Proteinen ist jedoch schwierig vorherzusehen, besonders wenn diese in neuen Kombinationen getestet werden. Trotz dessen, dass die Hefe *Pichia pastoris* für industrielle Anwendungen weit verbreitet ist, fehlt eine Auswahl an genetischen Regulatoren und effizienten Methoden, die es erlauben, die optimalen Proteinexpressionsbedingungen systematisch zu identifizieren. Diese Arbeit widmet sich dieser Problematik, indem eine Plattform für die *P. pastoris* Stammentwicklung mit dem Fokus der Proteinexpression und -sekretion entwickeln wurde.

Als Erstes wurde ein Werkzeugkasten aus standardisierten genetischen Elementen entwickelt. Die Elemente können unkompliziert mittels einer hierarchischen Assemblierung zu großen und diversen *P. pastoris* Proteinexpressionsbibliotheken kombiniert werden. Der Werkzeugkasten enthält 42 Elemente für die Regulation der Geneexpression und Proteinsekretion sowie den Erhalt der heterologen DNA auf einem selbstreplizierendem Plasmid, oder für die Integration in das *P. pastoris* Genom. Für die Charakterisierung der regulatorischen Elemente wurden 242 Stämme mit unterschiedlichem Aufbau für die Expression von zwei verschieden Fluoreszensproteinen (RFP oder yEGFP) hergestellt. Intrazellulare RFP oder yEGFP Expression waren ähnlich, wenn die gleiche Promoter- und Terminatorkombination, sowie die Expression entweder vom Plasmid oder genomisch integrierter DNA genutzt wurde. Im Gegensatz dazu variierte die Sekretionseffizienz der beiden Fluoreszenzproteine bis zu 100-fach, wenn identische regulatorische Elemente genutzt wurden. Daraus lässt sich schließen, dass die optimale Kombination von regulatorischen Elementen und Sekretionssignalen von einem Protein nicht übertragbar auf ein anderes Protein ist.

Im nächsten Schritt dieser Arbeit wurde die Schwierigkeit der Proteinsekretion adressiert. Der Werkzeugkasten wurde für eine Auswahl an industriell relevanten Proteinen angewendet. Ziel war die Identifikation der optimalen Kombination an regulatorischen Elementen und Sekretionssignalen. Die notwendige Grundlage für die effiziente Herstellung der kombinatorischen Bibliotheken und deren Screening nach den optimalen Sekretionsbedingungen wurde in dieser

Arbeit gelegt. Mittels des entwickelten Verfahrens konnten neue Kombinationen an genetischen Elementen für zuvor in *P. pastoris* exprimierte Enzymen identifiziert werden, die zu einer gesteigerten Proteinsekretion führen. Zusätzlich wurde versucht, die optimalen Expressionsbedingungen für eine Auswahl an prognostizierten Unspezifischen Peroxygenasen sowie bakterielle Alkohol Dehydrogenasen zu ermitteln, jedoch konnten diese Enzyme nicht in *P. pastoris* exprimiert werden. Problematisch war hierbei, dass kein funktionales Enzym sekretiert werden konnte.

Um Hochdurchsatz-Screenings von großen und diversen Bibliotheken zu ermöglichen, wurde eine neue Methodik zur Verkapslung von individuellen *P. pastoris* Zellen in Polyelektrolytkapseln getestet. Für diesen Ansatz sollten einzelne Zellen in ein festes Template integriert werden, welches mit Polyelektrolytschichten umhüllt wird. Anschließend sollte das Template aufgelöst werden. Trotz der sorgfältigen Charakterisierung der Templatebildung und der Analyse variierender Prozessbedingungen konnten keine Einzelzellen mittels dieser Methodik verkapselt werden.

Die hier bereitgestellte Plattform ist eine wertvolle Ressource für zukünftige Stammentwicklungsvorhaben. Das flexible Design des Expressionswerkzeugkastens ermöglicht zukünftige Anpassungen und Erweiterungen für die Herausforderungen der Proteinexpression und -sekretion in Hefen. Die Kombination dieses Werkzeugkastens mit Hochdurchsatz-Screenings ermöglicht die effiziente Optimierung von Zellfabriken.

Dedication and acknowledgements

I would like to thank Prof. Volker Sieber for welcoming me in his group and trusting me with this extremely interesting project. He gave me the freedom, support and encouragement to explore and pursue the world of *Pichia*. I also would like to thank Dr. Tom Ellis and Prof. Wolfgang Liebl for agreeing to review this dissertation, and Prof. Wilfried Schwab for acting as the chief examiner. I wish to thank Prof. Tim K Lu for giving me the opportunity to visit his group and introducing me to the fascinating world of synthetic biology. I also wish to thank Prof. Jörg Königstorfer for helping to organise the financial support of my placement at MIT.

I would like to thank all the past and present members of CBR. In particular, I would like to thank Prof. Jochen Schmid, Dr. Josef Sperl, and Dr. Bigna Wörsdörfer for all the precious scientific advice they have given me. My special thanks to all the technicians, especially Melanie Speck, for sharing all her lab tips and tricks, Magdalena Haslbeck for some fun time with "our" microscope, and Manuel Döring for all the support with the robot gear. A huge "thank you" to Elisabeth Aichner for guiding me through all the administrative challenges. I wish to thank all fellow PhDs who contributed in some way or another, Irina Funk, Dr.-Ing. Sumanth Ranganathan, André Pick, Samed Güner, Ioannis Zachos, Janine Simon and everyone I did not name here. I would like to thank the members of the Lu Lab for their kind welcome in Boston, as well as their valuable advice and support.

I wish to thank my friends who helped me on this journey. "Thank you" to Melli and Irina, for our cooking evenings and so much more. "Thanks" to my Berlin-friends for the lovely times whenever we met.

My thanks and gratitude to my family and especially my parents. Ihr wart immer für mich da, ganz besonders in den turbulenten Zeiten. Ihr seid eine riesige Unterstützung und habt immer an mich geglaubt. Vielen Dank.

Finally to Tom. Despite the miles that were often between us, you have always been on my side. I thank you for your patience, support, understanding, and the countless "just keep goings". This time has been exceptional and I can not thank you enough.

Table of Contents

Abbreviations

AA	Amino acid
ABTS	2,2´-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)
ad	to
ADH	Alcohol dehydrogenase
Amp	Ampicillin
AOX	Alcohol oxidase
APO	Aromatic peroxygenase
APS	Ammonium persulfate
ATP	Adenosine triphosphate
a.u.	Arbitrary unit
B	Biotin
BASIC	Biopart assembly standard for idempotent cloning
BiP	Binding protein
BLAST	Basic local alignment search tool
BMD	Buffered minimal dextrose
BMGY	Buffered complex glycerol
BMM	Buffered minimal methanol
BMMY	Buffered complex methanol
bp	Base pair
BSA	Bovine serum albumin
CAD	Cinnamyl alcohol dehydrogenase
CAI	Codon adaption index
CPEC	Circular polymerase extension cloning
CRISPR/Cas9	Clustered regularly interspaced short palindromic repeats/ CRISPR associated protein 9
C-terminus	Carboxyl-terminus
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Mixture of deoxyribonucleotides
DoE	Design of experiments
DTT	1,4-Dithiothreitol
dwp	Deep-well plates
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmatic reticulum
FACS	Fluorescence activated cell sorting

FDA	Food and drug administration
FITC	Fluorescein isothiocyanate
FSC	Forward scatter
GFP	Green fluorescent protein
GOI	Gene of interest
GRAS	Generally recognised as safe
GY	Glycerol
HR	Homologous recombination
Hsp70	Heat shock protein 70 kDA
IVC	*In vitro* compartmentalisation
kb	Kilo basepairs
kDa	Kilo dalton
LB	Lysogeny broth
LbL	Layer-by-layer
LCR	Ligase cycling reaction
MASTER	Methylation-assisted tailorable ends rational
MDR	Medium-chain dehydrogenases/reductases
MM	Minimal medium
MoClo	Modular cloning
MOPS	3-[N-morpholino]propane-sulfonic acid
mRNA	Messenger RNA
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NE-SLICE	Nicking endonuclease for sequence and ligation-independent cloning
NHEJ	Non-homologous end joining
N-terminus	Amino-terminus
OE-PCR	Overlap extension polymerase chain reaction
OFAT	One factor at time
PAGE	Polyacrylamide gel electrophoresis
PAH	Poly(allylamine hydrochloride)
PARS	*Pichia* autonomous replication sequence
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PERS	*Pichia* electroporation recovery solution
Phyre	Protein homology/analogy recognition engine
PSA	Pairwise selection assembly
PSS	Poly(sodium 4-styrenesulfonate)
RFP	Red fluorescent protein
RNA	Ribonucleic acid
rpm	Rounds per minute
RT	Room temperature
SBOL	Synthetic biology open language
SCORE	Sequence consistency of the overall residue evolution
SD	Standard deviation
SDS	Sodium dodecyl sulfate

SLIC	Sequence and ligation-independent cloning
SLiCE	Seamless ligation cloning extract
SOB	Super optimal broth
SOC	Super optimal broth with catabolite repression
SRP	Signal recognition particle
SSC	Side scatter
TEMED	N,N,N,N-Tetramethylethane-1,2-diamine
Tet	Tetracycline
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
tRNA	Transfer RNA
TU	Transcriptional unit
U	Units
UPO	Unspecific peroxygenase
UPR	Unfolded protein response
USER	Uracil-specific excision reagent cloning
v/v	Volume per volume
WT	Wild type
w/v	Weight per volume
YE	Yeast extract
yEGFP	Yeast enhanced green fluorecent protein
YOGE	Yeast oligo-mediated genome engineering
YPD	Yeast extract peptone dextrose medium
YNB	Yeast nitrogen base

Amino Acid Codes

3-Letter code	Name
Ala	Alanine
Cys	Cysteine
Asp	Aspartic acid
Glu	Glutamic acid
Phe	Phenylalanine
Gly	Glycine
His	Histidine
Ile	Isoleucine
Lys	Lysine
Leu	Leucine
Met	Methionine
Asn	Asparagine
Pro	Proline
Gln	Glutamine
Arg	Arginine
Ser	Serine
Thr	Threonine
Val	Valine
Trp	Tryptophan
Tyr	Tyrosine

List of Tables

List of Figures

1

Introduction

The global population is rapidly growing, demanding more food, fuel, pharmaceuticals, and other goods [27]. While the economy has managed to expand and meet these demands, declining fossil resources and pollution require that a transition is made to more sustainable production methods with reduced environmental impact. One such approach is to exploit living systems rather than harsh chemistries to support a transition to a bio-based economy [39]. Key to such a transition will be the use of efficient, robust, and versatile biological cell factories to produce the wide array of chemicals and products we rely on [57].

Living systems provide a unique solution to build chemical units from inexpensive and renewable carbon sources [316]. Until now, the engineering of microbes has already enabled the production of fuels (alcohols, isoprenoids) [115], pharmaceuticals (insulin, vaccines) [269], food/feed supplements and textile/detergent additives [158] in a sustainable and clean manner [99, 151]. The global enzyme market is expected to grow further, from $5.01 billion in 2016 to $6.32 billion in 2021 [54]. Although many products have been successfully produced using microbes, they are often not economically competitive in the initial stages and require significant optimisation to ensure robust and efficient production. This process is currently time consuming and costly [16, 309], and additional methods are needed to rapidly engineer cell factories [5].

The major challenge in the development of a cell factory is that naturally occurring microbiological hosts producing a desired chemical unit are not yet suitable for efficient industrial production processes [316]. In order to exploit the complexity and diversity of nature, modification of these production hosts in necessary. Advances in metabolic engineering and synthetic biology offer the opportunity to develop new production routes for these products in industrially applicable and well-studied hosts [26].

Recombinant deoxyribonucleic acid (DNA) technologies cross the natural boundaries of organisms and allow reconstruction and integration of production pathways into heterologous, usually microbial, production hosts. This circumvents the constraints of the native producers such as slow growth, high nutrition needs, or low product yields. Furthermore, it allows us to build cell factories that can utilise carbon feedstocks from waste streams to produce a value added product [26].

Suitable organisms for engineering are bacteria, yeast, filamentous fungi, plant and mammalian host systems [204]. The choice of the production organism depends on the specific application and must consider the desired product quality and quantity, existing industrial infrastructure and cost. For example, to produce biopharmaceutical proteins, the quality of the final product is of crucial importance, whereas for bio-fuels, high product yields are of greater interest [241]. No single host is currently available that addresses all these requirements at once, but understanding the cellular physiology and methods to alter genetic information helps us tailor the host of choice towards the desired properties [115].

The gram-negative bacterium *Escherichia coli* is probably the best-studied organism due to its role as a prokaryotic model organism and its potential as a microbial cell factory. *E. coli* offers robust growth characteristics on a variety of monomeric carbon substrates and presently the largest toolbox of synthetic biological parts to genetically engineer it [26, 292]. The gram-positive bacterium *Bacillus subtilis* is also commonly used in industry. Approximately 60 % of the commercially available technical enzymes are produced using this bacterium [320]. Its ability to efficiently secrete proteins outside the cell make it an interesting cell factory for simplifying the purification of products, but genetic tools for *B. subtilis* are more limited and the currently produced enzymes are mainly of homologous origin [241, 320]. Bacterial expression systems are simple and cheap, but they may not always be ideal, as they lack the option for complex post-translational modifications of proteins, and highly over-expressed proteins may aggregate, which adds processing costs for protein recovery and refolding [241]. Eukaryotic hosts are more challenging to cultivate but have cell organelles that permit natural compartmentalisation of biosynthesis and perform protein secretion that allows post-translational modifications [295]. The yeast *Saccharomyces cerevisiae* has probably the longest tradition as a cell factory due to its application in baking and brewing. Nowadays, it is a well-characterised eukaryotic model organism [26]. The yeasts *Klyveromyces lactis* and *Pichia pastoris* or the filamentous fungi *Trichoderma reesei* and *Aspergillus niger* are also established industrial hosts with tools for genetic manipulations being available [218]. When high-quality proteins with human post-translational modifications are necessary, mammalian expression systems may be required, even through they are comparatively complicated and expensive to maintain [28].

Once a host system, or chassis as it is often known, is selected, the next step is to engineer this organism towards the aimed production process [309]. An important consideration is to balance the burden of product formation and ensure host viability [50, 51].

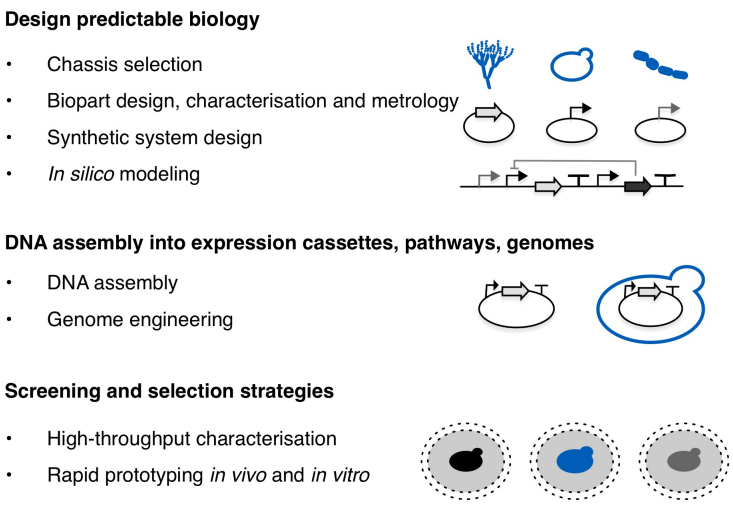

FIGURE 1.1. Overview of the steps to create a new microbial cell factory. A microbial chassis is selected and engineered using characterised biological parts. First, the desired production route is designed and modelled *in silico*, second, the various modules built from biological parts are integrated into the chassis and finally these hosts are screened to find the best modules for product formation.

The field of synthetic biology approaches this with an engineering perspective to replace the existing cellular architecture in a wholesale manner and construct genetic expression systems from scratch [247, 262]. Well characterised regulatory elements and genes act as a foundation that are combined to produce functional modules, which can then be integrated into a cell factory to tune gene expression in a desired way [22, 309]. Many genetic elements have been designed and tested, but because each production task is unique requiring different combinations of components, predicting the behaviour of these modules in the biological environment is challenging [115, 247]. For some organisms a catalogue of well studied genetic parts have been developed and can be "*mixed and matched*" [247] to identify modules that function the best. For other organisms, there is a lack of characterised elements and new methods are needed to allow genome manipulation [247]. To tackle this issue, screening techniques are necessary, to characterise more biological parts *en masse* and study the performance of parts under various conditions [15]. This will support the effective development of cell factories that can account for the unknowns in the native biological system and allow for the robust production of valuable goods [247]. Figure 1.1 summarises the overall work flow necessary to design a cell factory following a synthetic biology design approach.

In this thesis, the yeast *Pichia pastoris* was the chosen chassis for the development as a cell factory for the production of enzymes. This Chapter will first introduce *P. pastoris* and some of its history, then discuss the engineering of a cell factory in detail, and finally explain available screening techniques to help improve cell factory design.

1.1 The Yeast *Pichia pastoris*

P. pastoris (also known as *Komagataella phaffi*) is a unicellular eukaryote that offers advantages of bacteria as well as higher eukaryotes and is now the most frequently used yeast for the production of single proteins, surpassing even *S. cerevisiae* [30, 258]. For example, up to 2009 it has been successfully used for the production of over 1,000 industrial enzymes and 500 biopharmaceuticals [96, 329].

The application of yeasts for food processing and fermentation has a long history with roots in the time of early biotechnology dating back around 6,000 years [195]. The discovery of yeast as living matter in the nineteenth century started the scientific approach of studying yeast mono-cultures and target-oriented processes [41, 201]. In 1969, Ogata *et al.* first described *P. pastoris*, a strain that can use methanol as a sole carbon and energy source [225]. At that time, the application of this organism was interesting for the production of yeast biomass and single-cell proteins for human consumption or as animal feeds because methanol could be cheaply synthesised from natural gas [67]. The Phillips Petroleum Company commercialised the application of *P. pastoris* for the production of single cell protein as an animal feed additive, but in 1973, the oil crisis made the process uneconomical [4]. In 1985, Cregg *et al.* developed a heterologous protein expression system [65], which made *P. pastoris* an attractive industrial host again. Their expression system was used for the production of a plant hydroxynitrile lyase [125], which could be combined with the established Phillips Petroleum fermentation process [313]. This resulted in one of the first large-scale heterologous protein production process and to date still holds the record for a heterologously and intracellular protein, reaching 22 g/L [125]. For secreted protein, the record is held for expressing non-hydroxylated gelatins reaching 14.8 g/L [319].

The patented expression system from Cregg *et al.* [65] was released to the scientific community and since then *P. pastoris* has become a highly successful host for the heterologous production of many proteins [30, 67]. The success of this expression system mainly stems from the application of the strong and tightly regulated alcohol oxidase (AOX) promoter and the respiratory growth behaviour. To grow on methanol, *P. pastoris* gathers energy through a specific metabolic pathway, converting methanol into formaldehyde and later into dihydroxyacetone and glyceraldehyde-3-phosphate [96]. The first step of this pathway is catalysed by the AOX which reaches levels of >30 % of the total soluble protein within the cell when grown on methanol, but is not detectable when grown on other carbon sources. This characteristic makes this regulatory promoter very attractive for the expression of other genes [49]. The preference for a respiratory growth is also industrially interesting, as it allows high cell density fermentations. In contrast to *S. cerevisiae*, *P. pastoris* does not produce ethanol or other by-products that become toxic to the cell and thus limit the cell density that can be reached in fermentations [67].

P. pastoris is classified as a non-conventional yeast (e.g. non- *S. cerevisiae* yeast), which summarises all yeasts with a lower degree of fermentative overflow metabolism [201]. Beside the useful characteristics that made *P. pastoris* stand out, working with *P. pastoris* brings multiple other advantages such as [4, 67, 96]:

- Ease of genetic manipulation and similar techniques to those of *S. cerevisiae*;
- Rapid growth to high yield on inexpensive media (up to 150 g dry cell weight/L);
- Crabtree-negative metabolism allowing high biomass yields in fermentation processes;
- Efficient secretory capabilities;
- Ability to perform post-translational modifications typical for higher eukaryotic cells such as proteolytic processing, folding, disulphide bond formation and glycosylation;
- Engineered *P. pastoris* strains that allow production of fully humanised sialylated glycoproteins [73, 121];
- Food and Drug Administration (FDA) generally recognised as safe (GRAS) status;
- High quality genome sequences [74, 166].

One distinct feature that makes *P. pastoris* a successful cell factory is its ability to secrete proteins into the extracellular environment. This is advantageous because product purification does not require expensive cell rupture, denaturation or refolding processes of the product [269]. Product clean-up is simplified because *P. pastoris* secretes only a few proteases and endogenous proteins, which results to a fermentation supernatant that contains a nearly pure product [96, 155]. Processing of a protein through the secretion pathway also enables post-translational modifications such as glycosylation. Sugar moieties are added to specific amino acids and are often essential to produce a functional protein [73] or provide benefits such as increased protein solubility, stability, and osmotolerance [81].

Protein secretion is a multi-step mechanism to process a pre-protein into a mature active protein whilst guiding it out of the cell (Figure 1.2). To enter the secretion pathway, an N-terminal signal sequence in the nascent protein is necessary. In the cytosol, the tagged pre-protein is recognised by a signal recognition particle (SRP). The SRP binds first to the pre-protein and then to its receptor, the Sec61p complex, which is an integral membrane protein in the endoplasmatic reticulum (ER). A conducting channel is formed and the nascent protein is transferred through the channel, either co-translationally (ribosome coupled) or post-translationally (ribosome-uncoupled). During this process the signal peptide is removed by a membrane bound peptidase, Ste13p [127].

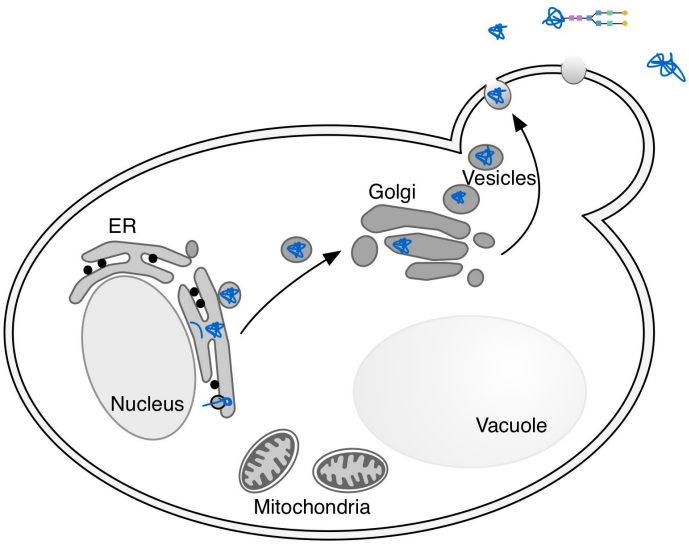

FIGURE 1.2. Protein secretion in yeast. Pathway of protein secretion in eukaryotic organisms guiding the nascent protein to the endoplasmic reticulum (ER) by default, then further to the Golgi apparatus, and finally the protein is released into the exterior space. Post-translational modifications are added in the ER and further processed in the Golgi.

In the ER, protein folding and post-translational modifications such as glycosylation, and peptide cleavage occur. A molecular chaperone, the heat shock protein 70 kDA (Hsp70) guides protein folding in addition to protein degradation, protein-protein interactions, and protein translocation. N-glycosylations occurs if the protein has the necessary recognition sites (amino acid sequence asparagine-X-serine/threonine, with X being any amino acid, except proline). An oligosaccharide precursor ($Glc_3Man_9GlcNAc_2$) is added to the asparagine residue and modified by removal and (re-)addition of saccharide units. For O-glycosylations, mannose from dolichyl phosphate is added to serine or threonine residues of the protein and processed like N-glycosylations [81]. Next, the protein is transported to the Golgi apparatus (Golgi) where it is further processed. Once it is correctly folded and passes the protein quality control, the mature protein is released into the medium [248, 276].

The only requirement for a protein to enter the secretion process is the signal peptide at the N-terminus of the pre-protein. Signal peptides appear to be simple and interchangeable domains, just like a “zip code”, which direct a protein to its designated location [91]. However, their structure and function is far more complex and how precisely these N-terminal sequences function as an export signal is unclear [52]. Secretion signal peptides are typically 20-30 amino acids long and have a distinct three-domain structure, the N-domain, H-domain and C-domain. The N-terminal (amino-terminal) domain is usually polar with a net positive charge, the H-domain is a hydrophobic core, which is not interrupted by charged residues.

The C-terminal (carboxyl-terminal) domain is also polar, containing proline and glycine residues as well as uncharged amino acids at position -3 and -1 to guide cleavage of the signal peptide from the protein [198]. In *P. pastoris*, the *S. cerevisiae* α-mating factor (αMF) signal sequence is the most widely used signal. A variety of other exogenous or endogenous signal peptides have also been used in *P. pastoris*, but finding an effective signal peptide for a protein of interest often requires much "trial and error" [74].

Despite the challenges in developing *P. pastoris* cell factories, a great variety of industrial enzymes and biopharmaceuticals produced in *P. pastoris* have already found their way to market [4]. The number of *P. pastoris* derived enzymes is also expected to increase further, as market trends predict increased productions especially of technical enzymes in food, feed and textile industries [54]. Available industrial enzymes are for example: a recombinant phytase product used as a feed additive (Phytex, USA), recombinant trypsin for proteomic research to obtain peptide patterns for MS analysis (Roche Applied Science, Germany), a nitrate reductase for water testing and treatment (The Nitrate Elimination Co., USA), and a phospholipase C for the degumming of vegetable oils (Verenium, USA and DSM, The Netherlands) [4]. *P. pastoris* is also an attractive host for the production of biopharmaceuticals, as it is able to perform post-translational modifications which are vital to ensure therapeutics acceptance of the human immune system [204]. In 2009, the first biopharmaceutical produced in *P. pastoris* was FDA approved, this was Kalbitor® a recombinant kallikrein inhibitor protein used to treat hereditary angioedema, a rare and potentially life-threatening genetic condition (Dyax, USA) [4, 271]. In 2012, Jetrea® was approved by the FDA and the European Commission, which is a drug to treat symptomatic vitreomacular adhesion (Alcon, USA). Other biopharmaceuticals expressed using *P. pastoris* include Insugen®, a recombinant human insulin (Biocon, India), Shanvac™, a recombinant hepatitis B vaccine (Shantha/Sanofi, India), and Nanobody® ALX-0061 a rheumatoid arthritis treatment (Ablynx, Belgium) [4, 214].

To further develop *P. pastoris* as an effective general-purpose cell factory, the obstacle of unpredictable protein secretion has to be tackled. Currently, signal peptides may not be functional in combination with a desired protein, or non-functional proteins become stuck in the secretion pathway limiting yield [80, 190]. Many of these problems stem from our limited understanding of the properties that drive protein secretion. New biological engineering approaches offer a means to explore these hurdles and provide support for future development of more efficient and robust processes using this cell factory [74].

1.2 Engineering Yeast Cell Factories

The field of synthetic biology attempts to apply engineering principles to biological systems with the aim of creating novel organisms with desired behaviour [247]. This approach is founded upon the idea that biological systems can be broken down into functional parts that can be reassembled in a modular way to generate new functionalities. The goals and methods of this concept are analogous to the field of computer engineering [11].

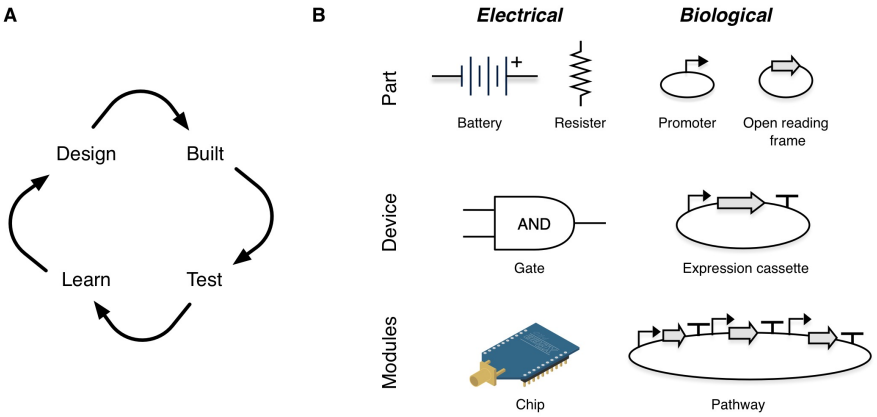

FIGURE 1.3. Summarising the principles of synthetic biology. **A** Development of new systems follows a design-built-test-learn cycle until the desired functionality is reached [11]. **B** Comparison of electrical and biological engineering construction definition. Parts are assembled into modules, which are further assembled into devices [17].

The design of a complex system, electrical or biological, requires the assembly of well characterised components or parts (e.g. transistors and capacitors or promoters and open reading frame). Using these parts, devices are build (e.g. logic gates or expression cassette) that are combined into modules (e.g. electrical circuits or metabolic pathways), which are finally integrated into a host system (computer/cell) to modify its behaviour and add new functionalities. To ensure a system responds as we expect, the individual parts should be well-characterised for the context (e.g. the particular host) they are used in [11, 17, 90, 259, 330]. To allow interchangeability of biological parts, common approaches from electrical and mechanical engineering such as standardisation, abstraction, modularity, predictability, reliability and uniformity are applied to the field of synthetic biology [11]. It is important to note that biological parts function in the environment of a living cell, making reliable predictions problematic due to the effect of a continually changing cellular environment, unavoidable expression noise, and unintended interactions of any new devices with existing host machineries (Figure 1.3). Therefore, synthetic biological construct need to be tested and characterised *in vivo*. Based on these measurements, the system may require re-design to tweak its function following the design-build-test-learn cycle, until a working system is established (Figure 1.3 A) [11, 17].

1.2.1 Biological Parts to Control Enzyme Production

Synthetic biologists and bioengineers rely on toolboxes of biological parts that are modular and well-characterised. Canton *et al.* defined a standard biological part as: "*a genetically encoded object that performs a biological function and that has been engineered to meet specified design or performance requirements*" [45]. Synthetic biological parts generally interact and control various cellular processes of a host and are generally specific for one type of microorganism. The most widely used hosts in the field of synthetic biology are the bacterium *E. coli* and the yeast *S. cerevisiae* [22]. Although the majority of synthetic biology tools have been developed for *E. coli* [33], there are many advantages associated with yeast and eukaryotic cells [99, 155]. In this section, an overview of the biological parts necessary to engineer protein production is given. The focus here will be the parameters specific for the design of expression constructs in yeast like *P. pastoris*. An overall summary of these parts is given in Figure 1.4, following the SBOL (Synthetic Biology Open Language) visualisation recommendations [64].

The protein of interest

The goal of any protein production process is to gain high amounts of product (i.e. protein). To achieve this goal, the design of the gene encoding the protein of interest is vitally important with many possible options.

First, codon usage compatibility between the natural host from which the gene was sourced and the new cell factory must be analysed. The genetic code is degenerate allowing the translation to the same protein from different messenger ribonucleic acid (mRNA) sequences (e.g. differing DNA sequences) because of synonymous codons for a given amino acid [17, 154]. Each organism has a characteristic codon usage which often corresponds to its aminoacyl-tRNA availability. For the expression of heterologous genes, the DNA sequence is ideally modified to better meet the codon usage of the host to improve translation efficiency and product yields. However, varying the DNA sequence only according to the most abundant synonymous codons does not consider the mRNA secondary structure and other processes like translational pausing, which are known to also influence the overall transitional efficiency [17, 336]. For *P. pastoris*, many studies have shown an increased expression of heterologous proteins when a codon optimised DNA sequence was applied [7, 25, 222], but this approach may not always work for the reasons described above.

Second, balancing the adenosine and thymine / guanine and cytosine (A+T/G+C) content of the DNA sequence influences expression efficiency. *P. pastoris* struggles with A+T rich regions as they can cause termination of transcription. Altering the A+T/G+C content can be used to increase protein expression and has been shown, for example, for a glucocerebrosidase [278] and an immunotoxin [322].

Third, mutations of the DNA that alter the amino acid sequence can be used to improve protein folding and activity and overcome natural limitations. Many enzyme evolution strategies are available to change substrate specificity and improve kinetics and thermostability [17, 227].

Generally, mutant libraries are made supported by advanced bioinformatic tools and screened using high-throughput methods. This results in tailor-made enzymes that are optimised for a particular industrial process [59, 70].

Fourth, protein expression using eukaryotic chassis requires the consideration of post-translational modifications, such as disulphide bond formation or O- and N-linked glycosylations. N-linked glycosylations can be added when the consensus sequence is present. The addition of N-glycosylation sites has shown improved protein secretion of a cutinase in *S. cerevisiae* and *P. pastoris* [266] and improved enzyme activity of an elastase from *P. pastoris* [122]. Glycosylation only occurs when the protein is secreted and processed through the secretion machinery. To guide the nascent protein into the secretion machinery, signal peptides are necessary.

Finally, many different tags, peptides, domains or even proteins can be fused to the protein N- or/and C-terminal of a product of interest. The purpose of the tags can be very diverse. Signal peptides can be used to guide the pre-proteins to specific subcellular compartments, organelles, or to direct secretion outside the cell [17]. To allow recognition of the N-glycosylation sites, the pre-protein must enter the secretion pathway. The success of protein secretion in *P. pastoris* is not always guaranteed. Therefore, the testing of a selection of tags is generally required to reach high protein titres in the supernatant [4]. Other tags can be used for protein purification or detection and quantification. Affinity tags like the poly(His)tag and Streptag are used to purify the protein from the crude cellular extracts. Epitope tags like c-myc and FLAG are mainly used for protein detection such as western blotting [332]. Live protein detection can also be performed in the cell using fluoresce tags like green fluorescent protein (GFP) fused to a protein of interest [42].

Regulatory elements for protein expression

To build a functional transcriptional unit (TU), the gene of interest must be flanked upstream by a promoter which initiates transcription and downstream by a terminator which stops transcription. Promoters recruit the transcriptional machinery to the gene of interest and their activity can be constitutive or induced. Constitutive promoters allow the binding of the RNA polymerase to the operator-binding site without a transcription factor. In contrast, inducible promoters are regulated by activators or repressors, which assist or block the binding of the RNA polymerase [17, 309]. Activators or repressors must be added during fermentation to initiate and tune gene expression. The use of inducible promoters allows the stages of cell growth and protein expression to be separated, which can be important for the expression of toxic products. The effect of the inducer, activator and repressor on normal cellular behaviours, and basic function of these parts under different conditions should also be validated. This includes the determination of the basal promoter activity and the activity with/without an inducer present [17]. Promoters also vary in their expression strength, resulting in differing amounts of protein being expressed. For the production of recombinant proteins, strong and controllable promoters are generally preferred to maximise product production [301]. However, high expression levels may not always be beneficial.

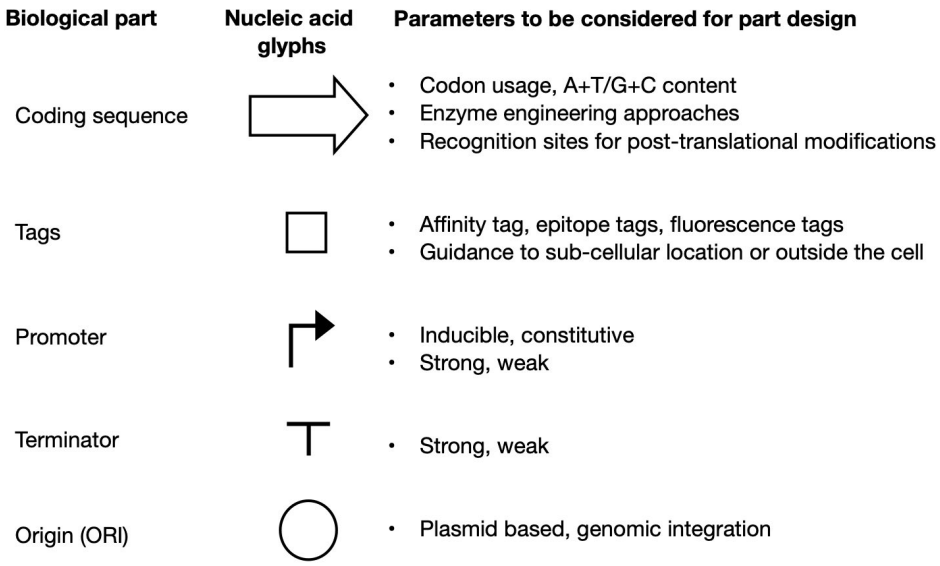

FIGURE 1.4. Selection of biological parts to tune cellular processes. For the design of a functional protein expression construct, the role and impact of many biological functions needs to be considered.

For example, overexpression of a secreted protein can overload the secretion machinery and lead to the aggregation of misfolded intracellular protein [155].

In *P. pastoris* the strong and tightly controlled methanol inducible alcohol oxidase promoter pAOX1 or the constitutive glyceraldehyde-3-phosphate dehydrogenase promoter pGAP are most commonly used. Both promoters have been studied extensively [242, 328] and several libraries varying transcription factor binding sites [18, 124] or using sequence mutagenesis [250] have been developed. Since methanol is inflammable and toxic, the application of the pAOX1 is not always suitable for large-scale applications. Alternatives include synthetic promoters which allow induction with glycerol or glucose [265], the nitrogen source dependent promoter pFLD1 [275], or strong promoters which can be repressed by methionine or inorganic phosphate [81].

Terminators are necessary to avoid transcriptional read-through and can also ensure a stable mRNA is produced [33]. Despite the importance of the terminator on overall protein expression, due to its impact on mRNA half-life [68], they are rarely considered in synthetic biology toolboxes [259]. For *S. cerevisiae* expression constructs, the strong effect of terminator choice was recently shown [212, 314]. However, in *P. pastoris*, the work of Vogl *et al*. [305] showed only minor effects in the expression of eGFP when varying the terminator used.

1.2.2 Assembling DNA Parts to Create Expression Constructs

In order to create a cell factory, genetic parts must be combined to form a functional expression module. This is achieved by physically piecing together DNA fragments encoding the relevant parts. DNA assembly generally requires cloning methods that are often performed *in vitro* using enzymes. The final DNA products are then inserted into and maintained by the microbial host [113].

In 1970, the first generation of cloning techniques was developed which enabled the "*cut and paste*" of DNA fragments by using restriction enzymes to cleave the DNA at specific sites and DNA ligases to connect DNA fragments together [61, 152, 343]. This biotechnological revolution opened up opportunities for genetic engineering, but the methods created often lacked flexibility and required sequential workflows that hindered the rapid assembly of multi-fragment designs [31]. Second generation cloning technologies have been developed over the past few decades to address this limitation and allow high-throughput and combinatorial construction of recombinant DNA in a flexible, fast, and precise manner [31, 176]. These assembly techniques extend both the number and the size of fragments that can be assembled and fall into four broad categories: (i) restriction enzyme-based methods, (ii) *in vivo* and (iii) *in vitro* sequence homology-based methods, and (iv) bridging oligo-based methods [53] (Figure 1.5).

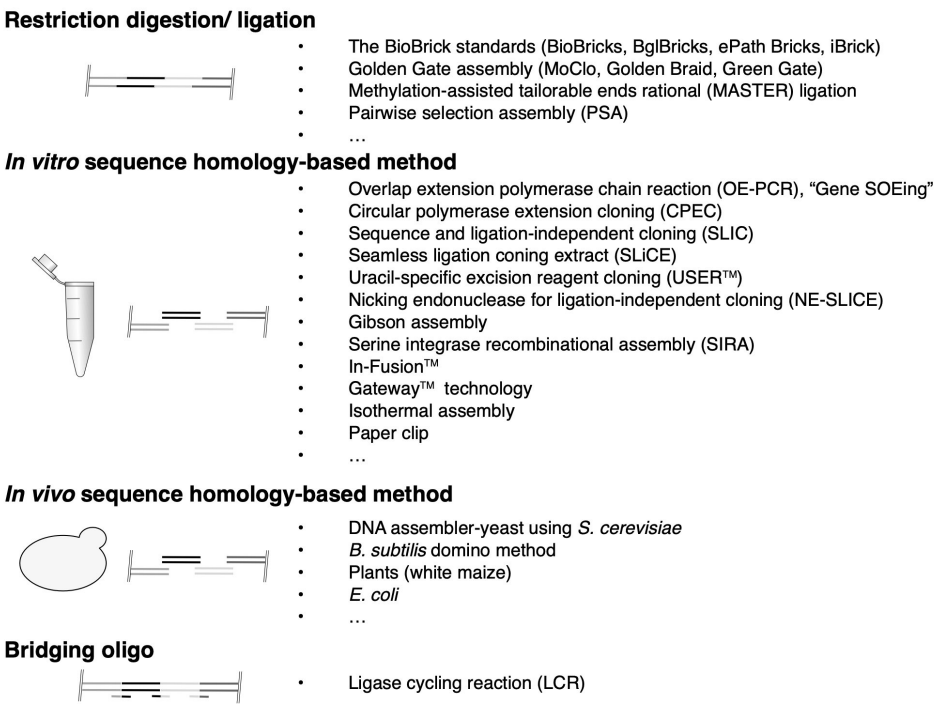

FIGURE 1.5. Overview of DNA assembly methods.

Physical standards defining how DNA parts should be encoded have also been developed to ensure DNA parts act as modular entities that are easily interchanged [48]. For example, standards exist for parts to be assembled using restriction digestion and ligation to ensure common DNA overhangs that can be used to mediate the correct sequence in the ligated product. Similar to other engineering disciplines, the standardisation of parts allows for large collections to be created that can be used across the field [15]. Based on this idea, "The Registry of Standard Biological Parts" (http://www.partsregistry.org) was founded to collect information on parts adhering to the BioBrick standard and share them globally. Even though these efforts have had some success, there is still much debate regarding the supporting information that should accompany these parts [255], with Canton *et al.* suggesting a comprehensive datasheet like in engineering disciplines [45]. Comprehensive reviews on the development of part and assembly standards can be found in the works of Casini *et al.* [48], Tsai *et al.* [295], Chao *et al.* [53], as well as Li and Borodina [175].

DNA cloning methods based on DNA restriction digestion and ligation

Basic restriction digestion cloning techniques have been improved by standardising the prefix and suffix of parts. The BioBrick™ assembly was one of the first attempts [161, 232] using common flanking sequences to enable sequential cycles of restriction digestion and *in vitro* ligation. In each step, two parts are assembled to create a new composite part that has the same prefix and suffix. This allows composite parts to then be assembled together using an identical set of steps. The prefix and suffix each define two restriction enzyme recognition sequences to facilitate this process. Due to the nature of restriction digestion, all BioBrick compatible parts must have the recognition sequences present in the prefix and suffix (i.e. EcoRI, XbaI, SpeI and PstI) removed within the rest of the part. Variations on the BioBrick standard are also available which differ in the enzymes used, e.g. BglBricks [9], ePath Brick [327] and iBrick [184]. A limitation of this approach is that scars remain between assembled parts and the number of parts that can be assembled in one reaction is limited to two.

This issue was overcome with the development of the Golden Gate assembly standard [93, 94], which uses a one-pot *in vitro* reaction to simultaneous perform restriction and ligation steps. Type IIs restriction enzymes are used, which cut outside their recognition site and allow for user-defined four bp overhangs to be produced. By designing complementary overhangs between parts that should be connected, the process allows large numbers of elements to be simultaneously assembled. Similar to the BioBrick standard, parts must have recognition sites for the specific TypeIIs restriction enzymes removed. Broader standards that exploit this approach have been developed that extend to the permissible DNA overhangs, the specific Type IIs restriction enzyme used, and vector backbones used. The modular cloning (MoClo) standard [312] uses two sets of backbone vectors that contain different Type IIs restriction sites. By alternating the backbone used to hold the product of an assembly reaction, hierarchical assembly becomes possible.

MoClo has seen wide acceptance with the development of toolkits including: mammalian MoClo [89], Yeast MoClo (YTK) [172] as well as plant [95] and *E. coli* variants [210]. An alternative standard to Golden-Gate based approaches is the Golden Braid standard which evolved over time (Golden Braid 2.0 (GB2.0), Golden Braid 3.0 (GB3.0)) as an online platform for part design, assembly and characterisation (https://gbcloning.upv.es). Other standards that rely on Golden Gate assembly include Green Gate [168], Biopart Assembly Standard for Idempotent Cloning (BASIC) [287] and the Mobius Assembly [10].

DNA cloning methods not based on DNA restriction digestion and ligation

A disadvantage of restriction digestion/ligation dependent methods is the requirement to remove recognition sites from genetic parts. Sequence homology methods both *in vitro* and *in vivo* use overlapping regions between parts to facilitate scarless assembly and require no modifications of any parts [48].

Overlap Extension Polymerase Chain Reaction (OE-PCR) [133] uses linear parts with overlapping ends (15-25 bp) that are normally generated by PCR. These fragments are annealed to each other and are extended by a DNA polymerase. One major drawback is the need for a set of unique primers for each junction between parts. To allow for multiple inserts in a one-step reaction, Circular Polymerase Extension Cloning (CPEC) [252] was developed, which requires only a DNA polymerase making it a relatively cheap method. Multi-fragment assembly is also possible by applying Sequence and Ligation-Independent Cloning (SLIC) [177] where 3´ ends of linearised vector and a desired insert are chewed back by a T4 DNA polymerase, without dNTPs. The single stranded fragments are combined by the RecA protein in the presence of adenosine triphosphate (ATP) and after *E. coli* transformation, gaps are fixed *in vivo*. Zhang *et al*. [339] modified the SLIC method developing the Seamless Ligation Cloning Extract (SLiCE). *E. coli* cell extract is used to guide homology-mediated DNA assembly and to reduce the cost of enzymes.

Another widely used homology-based DNA assembly method is Gibson assembly [108, 163]. 5´ ends of a linearised vector and a desired insert are chewed back by a T5 exonuclease to make single-stranded overhangs. These usually are 40 bp long and are joined by a DNA polymerase and Taq DNA ligase. It is a one-step *in vitro* reaction which is suitable for multi-fragments with an efficiency of 90 % when assembling 3-4 fragments. It is also suitable for the assembly of large fragments, but is generally seen as less useful for large numbers of fragments as the complementary ends for these are generally generated using PCR, which requires a large number of unique primers. Fragments less than 250 bp should also be avoided as T5 exonuclease may completely chew through the fragment. Alternative methods that rely on a similar concept are Uracil-specific excision reagent cloning (USER™) [32, 105] or the commercially supported methods of In-Fusion® [280] (clontech manual Cat. No. 121416) and Gateway® Technology (Life technologies™), but they are expensive compared to other methods and less suitable for the assembly of multiple fragments.

DNA assembly of overlapping regions can also be facilitated *in vivo*. The homologous recombination machinery of microorganisms have been proven for DNA assembly of linear DNA fragments with overlapping sequences. A variety of organisms such as *S. cerevisiae* with a method called "DNA assembler-yeast" [107, 274], *B. subtilis* [138], white maize [342] or *E. coli* [100] have been shown suitable for this methodology.

One drawback of *in vitro* and *in vivo* sequence homology based methods is the requirement that each part must be designed with the defined homologous sequence of its neighbouring part. This hampers interchangeability of parts as these are often unique to a particular design. The Ligase Cycling Reaction (LCR) [163] circumvents this issue by applying single stranded 60-90 bp long bridging oligomers which are 5´-phosphorylated and span the two ends of the neighbouring DNA. A thermostable ligase joins the DNA backbones. It is a one-pot scarless reaction for up to 12 DNA parts with an efficiency of 60-100 %. However, a disadvantage is that the fragments must be 5´-phosphorylated via PCR using 5´-phosphorylated primers or enzymatic treatment.

1.2.3 Host Considerations

The key features of a useful host for biotechnology applications according to Kim *et al.* are (i) its lifestyle, in other words growth behaviour and low nutrition demands or metabolic side-products, (ii) a robust cell-envelope towards physical stress, (iii) the accessibility for genetic manipulations, and (iv) a balanced and desired interaction between the host and the new biological modules [156]. *P. pastoris* meets many of these demands and is already an established industrial host [96, 329]. In the past, *P. pastoris* had been engineered further, to address some of these properties [103, 156] and is an attractive alternative to the model organism *S. cerevisiae* [103]. The most widely used and commercially available *P. pastoris* strains are GS115 [288] and the prototrophic strain X-33 [66], which are both derived from *P. pastoris* NRRL Y-11430, deposition CBS7435 at the CBS (Centraalbureau voor Schimmelcultures). These strains, but not their parental strain CBS7435, are restricted by patent protection or materials ownership policy for commercial applications [4]. Both strains are fully sequenced, supporting strain engineering efforts [74, 166].

The lifestyle of *P. pastoris* using an aerobic mode of respiration has always been interesting as cell densities of up to 150 g/L (dry cell weight) [313] are possible and no cell toxic compounds are produced [144]. The growth of *P. pastoris* using methanol as the sole C-source was undesirable for a long time as it is a fire hazard, unsuitable for products of the food industries and can hamper the cell viability and protein yields [96]. However, recently the use of methanol is receiving considerable interest in the transition towards a bio-based economy [238]. For future biorefinery projects, methanol could be made from CO_2 by reduction with renewable energy and used during the growth of methylotrophic microorganisms such a *P. pastoris* [103].

The genetic manipulations of *P. pastoris* rely on biosynthetic markers or markers based on antibiotic resistance genes [96]. For *P. pastoris*, a selection of host strains containing the required autotrophies are available for various biosynthetic pathways or genes (histidine pathway, gene deletions *HIS1*, *HIS2*, *HIS5*, *HIS6*; arginine pathway, gene deletions *ARG1*, *ARG2*, *ARG3* [217]; uracil pathway, gene deletions *URA3*, *URA5*; homoserine-O-transacetylase gene, *PpMTE2* [293] and formaldehyde dehydrogenase, *FLD1* [290]). Antibiotic selection is most widely performed using the *Sh ble* gene from *Streptoalloteichus hindustanus* as it provides resistance in *E. coli* as well as *P. pastoris* using Zeocin, a bleomycin-like compound [66]. Alternatively, the kanamycin *Tn903 Kan*^{r} gene from *E. coli* provides resistance to the G418 antibiotic [272] and the blasticidin S-deaminase gene from *Aspergillus terreus* provides resistance to blasticidin [157]. To expand the availability of selection markers a variety of selection marker recycling plasmids are also available [174, 283].

Classic *P. pastoris* transformation procedures rely on homologous recombination (HR) and non-homologous end joining (NHEJ) [215], with the requirement of plasmid linearisation and selection based on the antibiotic markers or strain auxotrophy. Besides the variety of auxotrophic strains [4, 182], strains with improved homologous recombination due to an impaired NHEJ mechanism and *ade1* or *his4* knockouts for selection [215] are available. For the construction of more sophisticated libraries, transformation based on HR is unfavoured as multiple random integration events can occur, making the comparison of different expression levels challenging [234]. Perez-Pinera *et al.* [234] have therefore developed a strain with three recombinase "*landing pads*" for the single-copy integration of plasmids at a defined locus. At the Trp2 locus in the *P. pastoris* genome (Chromosom 2) *attP* sites for each of the recombinases BxbI, R4 and TP-901 were classically integrated via HR and kanamycin selection. For genetic engineering purposes, the desired plasmid for transformation must contain the respective *attB* site and can be co-transformed with the plasmid for the transient expression of BxbI, R4 or TP901. The recombinase mediates integration at the corresponding locus (Figure 1.6). Throughout the work presented in this thesis, this strain containing the recombinase "*landing pads*" is used.

Recently, technologies such as Yeast Oligo-Mediated Genome Engineering (YOGE) [84] and the CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/CRISPR associated protein 9) technology have also become established for *P. pastoris* [318] allowing more rapid gene insertions and deletions with high efficiency and accuracy.

For the integration of synthetic modules and their long-term stability in the host *P. pastoris*, they are preferably integrated into the yeast genome [309] using either HR [215], BxbI recombinase-mediated integration [234] or the CRISPR/Cas9 technology [318]. However, for rapid and intermediate testing, episomal plasmids are also useful. Applying the *Pichia* Autonomous Replication Sequence (PARS) as a vector replication origin, higher transformation efficiencies than with the classic HR are achieved, which can be attractive for library screenings [65, 309].

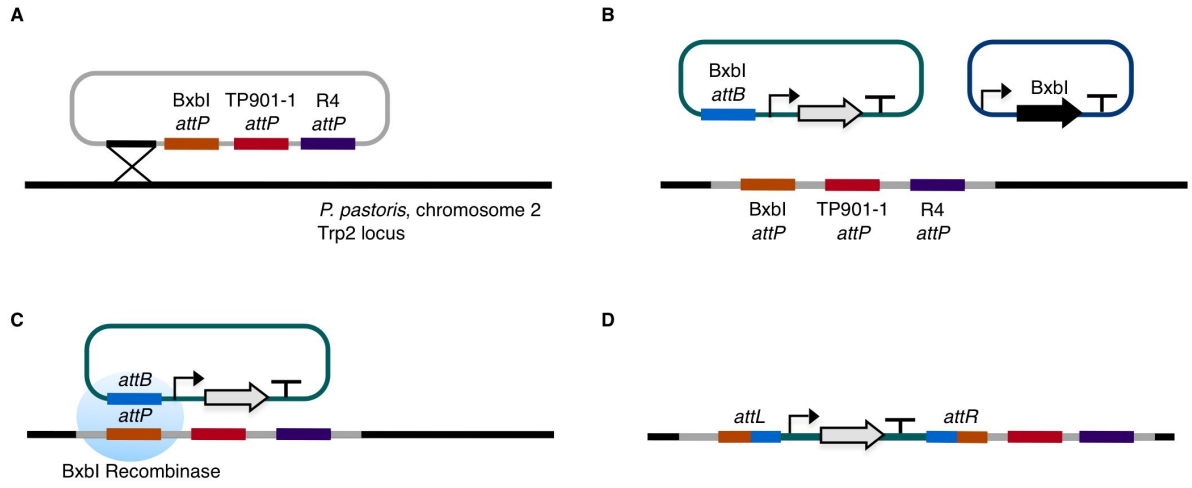

FIGURE 1.6. Recombinase based *P. pastoris* transformation. **A** Via HR three "*landing pads*" (*attP* sites) for the recombinases BxbI, TP901-1 and R4 are integrated at the Trp2 locus. **B** For the *P. pastoris* transformation, the strain is co-transformed with the desired vector containing a corresponding *attB* site and the expression vector for a recombinase, **C** which mediates single copy integration at the desired locus **D**.

The final key feature to address is the host capacity and capability to produce the desired product [156]. In addition to a synthetic construct being "readable" by the host (as discussed in Section 1.2.1), the host cell should also be able to provide sufficient amounts of precursors, co-factors, energy and enzyme processing capacity for product expression [175].

Endogenous proteases may cause difficulties when heterologous proteins are expressed in *P. pastoris*. Protein degradation by proteases occurs either intracellularly within the secretory pathway during transport in vesicles or in the extracellular space by secreted or cell-wall associated proteases [4]. These can degrade the protein of interest and lower the overall yield as well as complicate downstream purification. Protease-deficient strains have been engineered which lack known proteases such as the vacuolar aspartyl protease (*PEP4*), the carboxypeptidase Y (*PRC1*), or the proteinase B (*PRB1*) [4, 66]. The strains SMD1168 (*his4 pep4::URA3 ura3*), SMD1168H (Δ*pep4*) or the PichiaPink strains 2-4 (Δ*prb1* Δ*pep4*) are commercially available (Life technologies™) to circumvent this problem [137].

P. pastoris has been used for the production of several biopharmaceutical proteins, but is naturally incapable of many of the necessary modifications. For example, recombinant proteins for biopharmaceutical applications often require correctly humanised glycosylation for their function, their pharmacokinetic behaviour, and the acceptance in the human body [73]. *P. pastoris* adds high-mannose glycan structures to the protein, which are a severe problem as they can result in allergic reactions or a shorter *in vivo* half-life of the protein [4, 137].

Engineering *P. pastoris* to prevent N-glycosylation is challenging, as it is a complex multi-step process occurring during secretion. For the humanisation of N-glycosylation, first, the enzymes responsible for the hyperglycosylation must be deleted and additional glycosidases and glycosyltransferases as well as the biosynthetic pathways and transporter for missing sugars (e.g. sialic acid) added [302]. For all of these enzymes, enzyme localisation and concentration must be correctly tuned to ensure accurate processing and high overall efficiency. Hamilton *et al.* [120] performed this complex engineering feat by deleting four glycosylation genes and introducing 14 heterologous genes for human glycosylation patterns. This led to the successful recombinant expression of functional erythropoietin [120]. Additionally, Jacobs *et al.* [139] have engineered the N-glycosylation pathway of *P. pastoris* using their GlycoSwitch technology and shown successful expression of three murine cytokines. Their GlycoSwitch strain is commercially available (Pichia Technology from RCT) and can be used under licence.

1.3 Screenings to Optimise Design Choices

A major challenge when choosing a host and assembling a genetic construct is knowing which combination of parts and host will work best. This uncertainty has led to the widespread use of screening technologies to study large heterogeneous pools of cells where each contains a different combination of parts, overall design or varies other expression parameters. The information gained from these studies provides insight on the influential parameters and helps to guide future design-build-test iterations towards industrial applications [300]. Currently available screening technologies allow for the analysis of genetic part combinations to tune expression [75], but also can be used to engineer enzymes themselves to alter substrate specificity and increase stability or tolerance to exterior influences like temperature, pH and solvents [263].

The linkage between genotype and phenotype for each library member is essential during a screening. This is normally handled by spatially separating each library member. For enzymatic products, separated cells are incubated with a screening substance and analysed after a defined time. Results from substrate depletion or product formation are usually measured as a biochemical readout (e.g., fluorescence, absorption). Samples with beneficial characteristics then have their DNA construct sequenced to identify the genotype encoding the beneficial properties [186, 281].

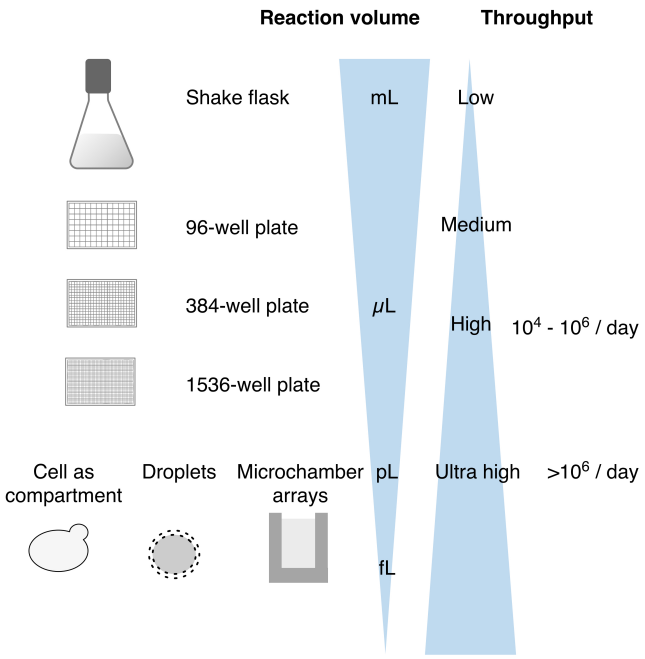

FIGURE 1.7. Screening methods with low- or high-throughput. As higher the throughput for any screening, as smaller the reaction volume is. Application of shake flasks only allows screening of few samples per day. However, reducing the reaction volume form mL to µL or pL, using deep-well plates or micrometric compartments allows screening of millions of samples.

With drops in the price of DNA synthesis, library sizes are continuously increasing. Classical screening methods such as expression in shake flasks, which allows only low-throughput experiments, are insufficient to handle these demands [300]. Currently, most screenings are performed using 96-well plates, but these also have limited throughput for complex libraries, require significant manpower or liquid handling equipment, and are costly due to the amount of reagents needed [62, 300] (Figure 1.7). 384-well plates have been developed to have an identical footprint as a 96-well plate, but scale screenings to four times the number of samples. The average working volume on 384-well plates is about 25-100 µL and many assays can be adapted to this format. Pharma and Biotech companies also prefer this format for compound storage [202]. Miniaturisation of screenings has continued and 1536-well and 3456-well plates have been developed with working volumes of 2.5-10 µL and 1-2 µL respectively. Some processes could be adapted for the 1536-well format [160, 202], but due to logistical challenges of these ultrahigh-density plates and practical challenges due to the high liquid volume to surface ratio, these plates are less commonly used [202]. To further increase throughput and reduce cost and time, other technologies have been developed to screen greater than 10 million variants per day using micrometric compartments. These reduce the reaction volume even further to a pico- or femtolitre scale [62]. These technologies can be grouped according to their separation technique: (i) cells as reaction compartments, (ii) synthetic droplets for *in vitro* compartmentalisation (IVC) and (iii) microchambers [186]. They differ significantly in their compartmentalisation strategy to maintain genotype and phenotype linkages, but separate individual variants; their screening method to detect a readout signal; and their sorting technique to separate remarkable hits for the identification of the genetic origin.

Cells themselves offer natural compartments for enzymatic reactions that can be used for measuring and sorting. Fluorescence activated cell sorting (FACS) devices are widely available to screen and sort cells based on an intracellular fluorescent signal. This requires that any assay is linked either to the expression, folding or trafficking of a fluorescent protein or fluorescently labelled compounds [186]. Such screenings can not only utilise fluorescence signals within the cell, but also allow for screenings based on the cell or phage surface display of reactions [3]. Surface display screenings allow unhindered access to the substrate and reaction conditions can be more easily tuned. However, maintaining the linkage between superior product formation and the enzyme variant producing it is challenging. Libraries of up to 10^7 variants are possible and depending on the transformation efficiency of the organism used, libraries of up to 10^{11} till 10^{12} are not out of reach [3].

Artificial compartments such as droplets and microchambers have been developed to overcome the limitation of cell-tethered signals. Droplet microfluidics are independent microcompartments of water-in-oil emulsions. These emulsions are made by dispersing an aqueous solution in an oil phase and are stabilised by surfactant molecules [62].

Encapsulation of microorganisms into the water droplets follows the Poisson distribution and allows tuning the number of cells per droplet by varying droplet size and cell density [300]. Droplets have an average volume of femto- to picolitre and an average of 10^7 to 10^9 droplets can be generated per experiment [62]. This reduces the cost by about 1-million-fold and increases the speed of the screenings by around 1000-fold [1]. Simple water in oil emulsions can be generated by mixing an aqueous phase and an hydrophobic oil phase using a stirring bar or emulsifier. These basic approaches generate polydisperse compartments, *i.e.* compartments with varying size that are suitable for screenings based on amplified DNA as the read-out information for screening. To enable analysis using flow cytometry, double emulsions have been developed, which utilise a second emulsification step to make water-in-oil-in-water droplets [62]. Again the droplets are coupled with a fluorescence read-out to allow FACS applications. A major drawback of this approach is the large size differences of the compartments, which results in varying signal strength and difficulties quantifying the droplet creation process [62]. More sophisticated technologies have been developed to generate monodisperse emulsion droplets with a defined and uniform size [186]. A lithographically-defined microfluidic device with precise channel dimensions and tunable fluid flow rates is generally used. Based on fluorescence or absorption signals, sorting can be performed using dielectrophoresis or acoustic waves at around 2 kHz [62]. For monodisperse compartments water-in-oil or also double emulsions are possible. To study intracellularly expressed proteins, lysing reagents can be added during the encapsulation process. To perform more complex studies, additional microfluidic technologies have been developed to add reagents later to the droplets via picoinjections or through the merging of multiple droplets [186]. Alternatively, gel-shell beads with uniform size have been developed by using a microfluidic device and generation of hydrogel beads coated in a polyelectrolyte shell [97]. The shell is size-selective allowing buffer exchange or supplementation of small molecules whilst capturing enzyme and coding DNA [97]. In summary, man-made droplets allow analysis and screening of a variety of enzymatic reactions, not only fluorescence coupled, and are compatible with a variety of protein expression systems. However, these screenings have to be tuned for each problem and if monodisperse droplets are required the approach is costly.

Microchambers are physically separate reaction vessels like a minimized microtiter plate. Two major types have been developed: microwell arrays and microcapillary arrays [186]. Microwell arrays are micron-scale wells etched into a glass slide where each well has an open top. The wells are loaded using discontinuous dewetting to place liquid droplets in the wells. Microcapillary arrays are bottom less high-aspect-ratio microcapillaries, loaded using capillary action. As for the microfluid droplets, Poisson statistics are applied to ensure controlled loading of microwell and microcapillary arrays with single cells [186]. Both microwell and microcapillary arrays allow signal detection over time and analysis of the library before deciding the precise sorting parameters. However, no later modification of the sample is possible and special equipment is needed to extract the desired samples [186].

1.4 Aim and Objectives

This thesis aim is to establish and explore a versatile platform for the generation of *Pichia pastoris* cell factories. Biological cell factories are becoming increasingly important for a transition to a bio-based economy, but current approaches using established model organisms are limited. The non-conventional yeast *P. pastoris* offers unique capabilities such as growth to high cell densities and efficient protein secretion. This makes it ideally suited for use in many biotechnology applications, but a lack of tools to control gene expression and identify effective expression conditions make cell engineering a challenge. This thesis has three major objectives that aim to address these issues (Figure 1.8).

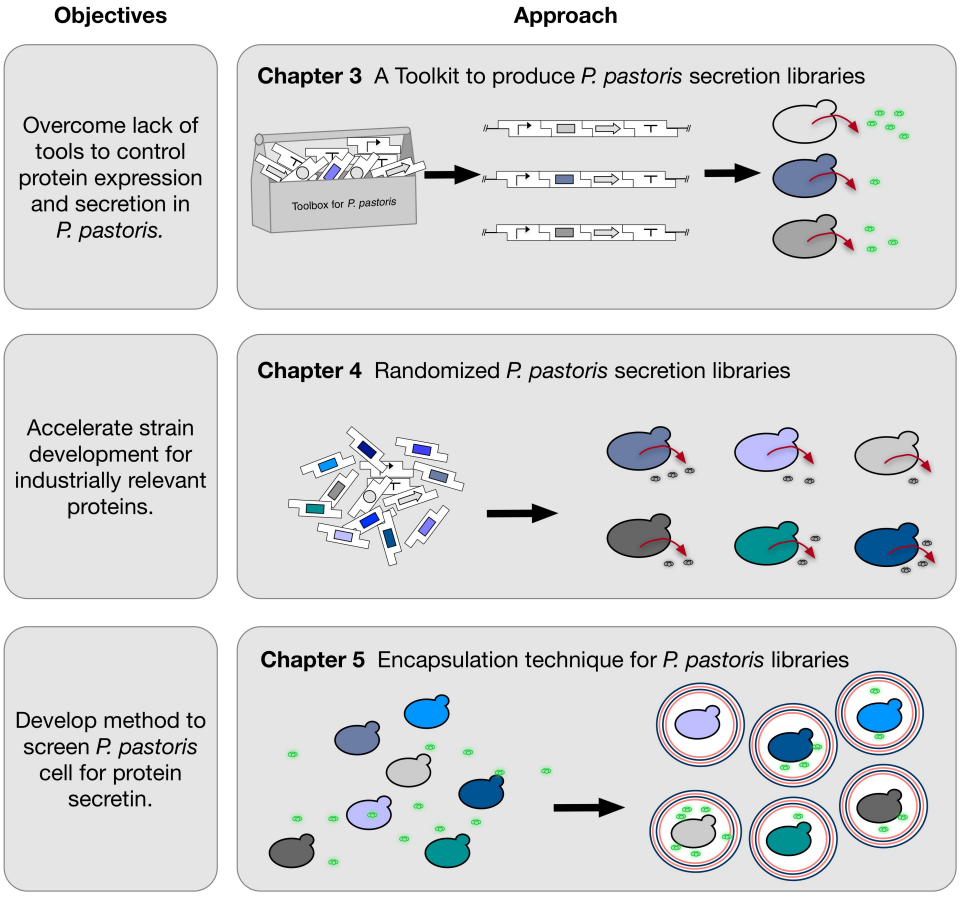

FIGURE 1.8. Objectives and approaches. The aim of a versatile *P. pastoris* cell factory is addressed by a toolkit enabling randomised library generation and the consideration of a novel screening technology.

The first objective is to overcome the lack of tools to control protein expression and secretion in *P. pastoris*. The aim is to develop a new modular and standardised toolkit of genetic parts that can be efficiently pieced together to create expression and secretion constructs. Crucially, the toolkit should enable rapid combinatorial assembly of many potential designs quickly and be easy to explore the potential design space, plus include a means of transforming *P. pastoris* at a high-efficiency. Where possible it is aimed to leverage existing synthetic biology assembly methodologies and part libraries (e.g., the Yeast Toolkit [172]) to expand the potential future scope of the toolkit.

The second objective is to demonstrate how such a toolkit opens up new ways of accelerating strain development for enzyme secretion. Specifically, it is planned to use the ability to generate diverse libraries of expression and secretion constructs to search for optimal genetic designs for producing industrially relevant enzymes. This will require the application of an approach where genetic parts like promoters, terminators and secretion tags are "shuffled" randomly to create a diverse library of designs. Furthermore, the ability to transform *P. pastoris* cells with this mixed library is required such that strains can be assayed and those performing well re-sequenced to find the precise combination of parts achieving a desired output.

The final objective is to develop a new screening method for *P. pastoris* that allows analysis of enzymes secreted from individual cells in high-throughput. This method must enable cultivation of physically isolated cells and the accumulation of secreted enzymes. Microencapsulation of single cells in polyelectrolyte capsules could be a mean to achieve this objective. These capsules offer tunable permeability allowing nutrients to be provided to the cells and can be assayed in bulk using high-throughput methods such as flow cytometry.

2

Materials and Methods

2.1 Materials

2.1.1 Devices, Consumables and Chemicals

Special devices used are given in Table 2.1. General consumables such as petri dishes, tubes (1.5 mL and 2.0 mL) and reaction tubes (15 mL and 50 mL) were purchased from Sarstedt AG & Co. KG (Nümbrecht, Germany). Pipettes and tips were purchased from BRAND GmbH + Co. KG (Wertheim, Germany). Other consumables are summarised in Table 2.2. All chemicals used were at least of analytical grade. Chemical suppliers are listed in Table 2.3. For every experiment, ultrapure water was utilised.

2.1.2 Software

Data analysis and plotting was performed using Phyton, Excel and Omnigraffle. The evaluation of microscopic images was done using ZEN 2.3, NIS Elements viewer, and Fiji. *In silico* cloning and DNA sequence alignments were performed using SnapGene, Benchling, and T-Coffee. Sequence codon optimisation was performed using the GeneOptimizer, and analysis thereof was done using the graphical codon usage analyser. Phylogenetic studies were done using phyloT (Phylogenetic Tree Generator). The FACS analysis was supported by the FACSDiva software and analysis of flow cytometry data was performed using FloJo. The experiments using the liquid handling platform were conducted using custom scripts written with the Freedom EVOware, Table 2.4.

TABLE 2.1. Devices

Device		**Company**
Autoklav	Varioklav 135S	Thermo Fisher Scientific, Waltham, MA, USA
Centrifuges	Rotanta 460R	Hettich, Tuttlingen, Germany
	Sorvall RC-6 Plus Rotors: SS-34, SH-3000, PN 11779	Thermo Fisher Scientific, Waltham, MA, USA
	Heraeus Fresco 21	Thermo Fisher Scientific, Waltham, MA, USA
	Heraeus Pico 17	Thermo Fisher Scientific, Waltham, MA, USA
	Galaxy Ministar	VWR International GmbH, Ismaning, Germany
Colony picker	CP-7200	Norgren Systems, Fairlea, WV, USA
Dispenser	MicroFlo Select	BioTek Instruments GmbH, Winooski, VT, USA
Electroporator	Micro Pulser	Bio-Rad Laboratories GmbH, München, Germany
Flow cytometer	BD LSR II HTS-2	BD Biosciences, San Jose, CA, USA
DNA Gel-documentation	Gel iX Imager	Intas Science Imaging Instruments GmbH, Göttingen, Germany
	Transiluminator DR-46B	MoBiTec GmbH, Göttingen, Germany
Incubators, shaking incubators	Heraeus BK 6160	Thermo Fisher Scientific, Waltham, MA, USA
	Heraeus B12	Thermo Fisher Scientific, Waltham, MA, USA
	Klimaschrank KBF 240 E5.1/C	BINDER GmbH, Tuttlingen, Germany
	KS4000ic control	IKA, Wilmington, NC, USA
	MaxQ Mini 4450	Thermo Fisher Scientific, Waltham, MA, USA
	Incubation room	Albert GmbH, Rain, Germany
	TiMix 5 control	Edmund Bühler GmbH, Hechingen, Germany
	HAT Minitron	Infors AG, Bottmingen, Switzerland
Liquid handling platform	FreedomEvo 200 Multi Channel Arm MCA96 Liquid Handling Arm LiHa	Tecan Group Ltd., Männedorf, Switzerland
	Incubator StoreX IC	LiCONiC Services Deutschland GmbH, Montabauer, Germany
Microscopes	Axio LabA.1	Zeiss, Oberkochen, Germany
	Axio ObserverZ.1	Zeiss, Oberkochen, Germany
	ECLIPS Ti-E	Nikon, Tokyo, Japan
Shakers	MaxQ 2000	Thermo Fisher Scientific, Waltham, MA, USA
	TiMix 5 control TiMix	Edmund Bühler GmbH, Hechingen, Germany
	Rocking Platform	VWR International GmbH, Ismaning, Germany
Plate reader	Infinite 200 pro	Tecan Group Ltd., Männedorf, Switzerland
Rotator	Rotator SB2	Stuart, Staffordshire, UK

TABLE 2.2. Consumables

Consumables	**Manufacturer**
Cryo pure tube 1.5 white	Sarstedt AG & Co., Nümbrecht, Germany
Millipore "V" Series Membranes (0.025 µm)	Merck KGaA, Darmstadt, Germany
Gene Pulser® Cuvette 0.1 cm (Catalog #165-2089)	Bio-Rad Laboratories GmbH, München, Germany
Gene Pulser® Cuvette 0.2 cm (Catalog #165-2086)	Bio-Rad Laboratories GmbH, München, Germany
PP-Masterblock, 2.0 mL® (#780271)	Greiner Bio-One GmbH, Frickenhausen, Germany
Nunc™ 96-Well MicroWell™ (#249944)	Thermo Fisher Scientific, Waltham, MA, USA
96-well plate (polystrol, F-ground)	Greiner Bio-One GmbH, Frickenhausen, Germany
96-well plate (UV-star)	Greiner Bio-One GmbH, Frickenhausen, Germany
96-well plate (ploystrol, U-ground, #655101)	Greiner Bio-One GmbH, Frickenhausen, Germany
96-well plate (black, F-ground, #732-2701)	Greiner Bio-One GmbH, Frickenhausen, Germany
Breathable adhesive film (#BF-400-S)	Axygen, Inc., Union City, CA, USA
Aluminium adhesive film PCR/AS	Axygen, Inc., Union City, CA, USA
His GraviTrap™	GE Healthcare, Little Chalfont, United Kingdom
PD-10 Desalting Column (#17-0851-01)	GE Healthcare, Little Chalfont, United Kingdom
96-Well SpinColumns™ (25-100 µL)	Harvard Apparatus, Holliston, MA USA
Centrifugal filters (516-0230)	VWR International GmbH, Darmstadt, Germany

TABLE 2.3. Chemical suppliers

Manufacturers
AppliChem GmbH, Darmstadt, Germany
Bio-Rad Laboratories GmbH, München, Germany
Biozym Scientific GmbH, Hess. Oldendorf, Germany
BODE Chemie GmbH, Hamburg, Germany
CARL ROTH GmbH & Co. KG, Karlsruhe, Germany
GE Healthcare Europe GmbH, Freiburg, Germany
GERBU Biotechnik GmbH, Gailberg, Germany
MERCK KGaA, Darmstadt, Germany
Rapidozym, Berlin, Germany
SERVA Electrophoresis GmbH, Heidelberg, Germany
Sigma-Aldrich, Deisenhofen, Germany
VWR International GmbH, Darmstadt, Germany

TABLE 2.4. Software

Software	**Origin**
Python version 2.7.11	Python Software Foundation, Wilmington, DE, United States
Excel for Mac version 16.12	Microsoft, Redmond, WA, USA
OmniGraffle	The Omni Group, Seattle, WA, USA
ZEN 2.3 (blue edition)	Zeiss, Jena, Germany
NIS Elements Viewer	Nikon, Tokyo, Japan
Fiji version 2.0.0-rc-54/1.51h	Image J [268]
SnapGene Viewer 4.2.4	GSL Biotech LLC, Chicago, IL, USA
Benchling	Benchling, San Francisco, CA, USA
T-Coffee	Notredame *et al.* [221]
GeneOptimizer	Thermo Fisher Scientific, Waltham, MA, USA
Graphical codon usage analyser	http://gcua.schoedl.de
phyloT	https://phylot.biobyte.de
FACSDiva	BD Biosciences, San Jose, CA, USA
FloJo	BD Biosciences, San Jose, CA, USA
Freedom EVOware 2 PLUS	Tecan Group Ltd., Männedorf, Switzerland

2.2 Cell Manipulation

2.2.1 *E. coli* and *P. pastoris* Strains

E. coli strains DH5α and NEB Turbo were used for plasmid construction purposes. *P. pastoris* containing the *attP* site for BxbI guided recombination [234] was used for expression studies, Table 2.5, and it will be referred to as the wild type throughout this work. *P. pastoris* strains designed in this study are listed in the Appendix Tables A.5, A.4, A.6, A.7, A.8 and A.9.

TABLE 2.5. *E. coli* and *P. pastoris* strains

Strain	Genotype	Manufacturer
E. coli DH5α	F^- Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17(r_k^-, m_k^+) phoA supE44 thi-1 gyrA96 relA1 λ^-	Thermo Fisher Scientific, Waltham, MA, USA
E. coli Turbo	F' $proA^+B^+$ $lacI^q$ Δ lacZ M15/ fhuA2 Δ(lac-proAB) glnV gal R(zgb- 210::Tn10)Tet^S endA1 thi-1 Δ(hsdS-mcrB)5	New England Biolabs, Ipswich, MA, USA
P. pastoris attP	NRRL Y-11430 + *attP*	Lu Lab, Massachusetts Institute of Technology, Cambridge, USA

2.2.2 Media and Growth Conditions for *E. coli* Cultivation

E. coli cultivation was performed in lysogeny broth (LB) for cloning purposes, Table 2.7. All solutions were autoclaved (121 °C, 20 min, 2 bar) and stored at room temperature (RT), if not stated otherwise. Depending on the strain or resistance, antibiotics were added, Table 2.6. Antibiotics dissolved in water were sterile filtered (0.2 µm filter) and all antibiotics were stored at - 20 °C. Liquid cultures were cultivated in shaking flasks (37 °C, 160 rpm, 12 h) or cultivation tubes (37 °C, 300 rpm, 12 h) on an orbital shaker. Cryo-cultures were prepared by mixing liquid cultures with a sterile 60 % (v/v) glycerol solution to a final glycerol concentration of 30 % (v/v) and stored at -80 °C.

TABLE 2.6. *E. coli* antibiotics for selection and cultivation

Solution	Stock concentration	Final concentration	Solvent
Ampicillin	100 mg/mL	100 µg/mL	ddH_2O
Carbenicillin	100 mg/mL	100 µg/mL	ddH_2O
Chloramphenicol	34 mg/mL	25 µg/mL	100 % (v/v) ethanol
Kanamycin sulfate	30 mg/mL	30 -50 µg/mL	ddH_2O
Zeocin	100 mg/mL	25 µg/mL	ddH_2O

TABLE 2.7. Media and stock solutions for *E. coli* cultivation.

Solution	Composition
LB	Lysogeny broth 10 g/L tryptone, 5 g/L yeast extract, 10 g/L sodium chloride (NaCl) Desired antibiotics were added prior use.
LB agar plates	10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, 15 g/L agar After autoclaving, the agar was cooled to 55 °C, the desired antibiotic was added and agar was poured into 10 cm petri plates. After solidification, plates were stored at 4 °C.
TSS buffer	10 % (w/v) polyethylene glycol (PEG) 6000, 30 mM magnesium chloride ($MgCl_2$), 5 % (v/v) dimethyl sulfoxide (DMSO) in LB medium Filter sterilised (0.2 µm filter) and stored at 4°C.
SOB	Super optimal broth 20 g/L tryptone, 5 g/L yeast extract, 10 mM NaCl, 2.5 mM potassium chloride (KCl), 10 mM $MgCl_2$
SOC	Super optimal broth with catabolite repression SOB-medium with 20 mM glucose

2.2.3 Media and Growth Conditions for *P. pastoris* Cultivation

P. pastoris strains were cultivated at 30 °C and shaken on an orbital shaker at 130 rpm - 160 rpm for shaking flasks with baffles, 300 rpm for cultivation tubes and 900 rpm for deep-well plates (dwp). Cultivation was done in YPD for growth and in BMGY/BMMY (buffered complex glycerol/methanol) or BMD/BMM (buffered minima dextrose/methanol) for protein expression, Table 2.8. To improve readability throughout this thesis, media names are shortened from 1 % (v/v) BMM to 1 % BMM, 0.2 % (w/v) or 1 % (w/v) BDM to 0.2 % or 1 % BDM and respectively 0.4 % (v/v) or 2 % (v/v) BMGY to 0.4 % or 2 % BMGY.

The antibiotic zeocin was added to a final concentration of 75 µg/mL to all media except for the growth of the wild type *P. pastoris attP*. All solutions were autoclaved (121 °C, 20 min, 2 bar) and stored at room temperature, if not stated otherwise. Cryo-cultures were prepared in cryo-tubes by mixing liquid cultures grown in YPD with 60 % (v/v) glycerol to a final concentration of 30 % (v/v) and are stored at -80 °C. Cryo-culture back-ups of randomised screenings were prepared by mixing 180 µL 16.66 % (v/v) glycerol with 20 µL sample in flat bottom assay plates. These screening backup plates were first used to determine OD_{600} and subsequently frozen at -80 °C for backup purposes.

Modifications to the medium were done depending on the expression strategy and enzyme being expressed. For all screenings, cultivation was carried out in 96-deep-well plates at 30 °C, 900 rpm on a microtiter plate shaker (Edmund Bühler GmbH, Germany). A pre-culture was inoculated from a glycerol stock master plate or a 96-well backup agar plate using the colony picker (Norgren Systems, Fairlea, WV, USA) and grown until saturation for 48 hours. The expression-culture, either for constitutive expression or induced expression, was inoculated from the pre-culture. Inoculation was done using the liquid handling platform and applying its multi-channel arm MCA96 (Tecan Group Ltd., Männedorf, Switzerland), which was controlled using custom scripts written using the Freedom EVOware software (Tecan Group Ltd., Männedorf,

Switzerland). The expression plates were incubated and when required, methanol induction was performed every 24 h to maintain 1 % (v/v) methanol in the expression medium.

Screening for expression of RFP and yEGFP was performed as described by Qin *et al.* [249], with modifications, Table A.5 and A.4. A 900 µL 0.2 % BMD glucose pre-culture was inoculated manually from a glycerol stock master plate. The expression-culture in 1 % BMD glucose or 1 % BMM methanol was inoculated with 30 µL of the pre-culture and incubated, Table 2.9. Measurements were performed after 24, 48 and 72 hours, with a set of plates for each time point.

The cultivation conditions for the phytase screening were as described by Hesampour *et al.* [129], with modifications. 1.2 mL pre-culture was inoculated from a 96-well backup agar plate using the colony picker. 100 µL of the pre-culture was used to inoculate 900 µL expression-culture and incubated for 48 h. For the phytase screening with inducible promoters, the pre-culture was 2 % BMGY-PP, the expression-culture 1 % BMMY-PP and to keep methanol induction, 10 % BMMY-PP was added after 24 h of expression to maintain 1 % (v/v) methanol. For the screening using constitutive promoters, the pre-culture was 0.4 % BMGY-PP and the expression-culture 2 % BMGY-PP, Table 2.10. The control strains for screening establishment can be found in Table A.6; strains identified via the screening and integrated to the strain collection are listed in Table A.7.

The unspecific peroxygenase (UPO) screening cultivation conditions were similar to Molina-Espeja *et al.* [209] and Krainer *et al.* [164], with modifications. For screenings utilizing inducible promoters, 1.2 mL 1 % BMD were inoculated from 96-well backup agar plates using a colony picker. In a new dwp, 500 µL pre-culture was used to inoculate 500 µL 2 % BMM+2H to reach a final concentration of 1 % (v/v) methanol, 10 µM hemine, and 3 mM $MgSO_4$. After 24 h of incubation, 100 µL 10 % BMM was added to maintain induction, Table 2.11. For the constitutive expression of UPOs, 1.2 mL 0.2 % BMD was inoculated as before. 100 µL of the pre-culture was used to inoculate 900 µL 1 % BMD+H, containing 1 % (w/v) glucose, 10 µM hemine and 3 mM $MgSO_4$, and grown for 48 h, Table 2.12. The control strains for the UPO screening establishment can be found in Table A.6, strains identified via the screening and integrated to the strain collection are listed in Table A.8.

UPO expression studies performed in baffled shaking flasks were performed following the same scheme as for the screening. A pre-culture was made either in cultivation tubes or baffled shaking flasks and used to inoculate the expression-culture. A typical inoculation volume was 10 % (v/v) of the expression-culture volume. The media used were identically to the media of the screening, except for the induced expression of UPO´s, where 1 % BMM+H was used.

TABLE 2.8. Media and stock solutions for *P. pastoris* cultivation.

Solution	**Composition**
Zeocin	75 µg/mL dissolved in water
20 % (w/v) Dextrose	200 g/L D-glucose dissolved in water
2X YP	20 g/L yeast extract, 40 g/L peptone dissolved in water
YPD medium	Yeast extract peptone dextrose medium 10 g/L yeast extract, 20 g/L peptone, 20 g/L dextrose To make 100 mL YPD: 50 mL 2X YP, 40 mL sterile water, 10 mL 20 % (w/v) dextrose
YPD Agar	10 g/L yeast extract, 20 g/L pepton, 20 g/L agar, 20 g/L dextrose 10 g yeast extract, 20 g peptone and 20 g agar were dissolved in 900 mL deionised water and autoclaved (121 °C, 20 min, 2 bar). Once cooled down to 55 °C, 20 % (w/v) dextrose and if desired zeocin were added and agar was poured into 10 cm petri plates. After solidification plates were stored at 4 °C in the dark.
10X YNB	13.4 % (w/v) yeast nitrogen base with ammonium sulfate without amino acids 100 g/L ammonium sulfate , 34 g/L yeast nitrogen base The solution was heated to dissolve completely in water, filter sterilised (0.2 µm filter) and stored at 4 °C.
500X B	0.02 % (w/v) biotin The solution was filter sterilised (0.2 µm filter) and stored at 4 °C.
1 M phosphate buffer	132 mL of 1 M dipotassium phosphate (K_2HPO_4) were combined with 868 mL of 1 M monopotassium phosphate (KH_2PO_4) pH = 6.0 ± 0.1, adjusted if necessary using phosphoric acid or potassium hydroxide (KOH).
10X GY	10 % (v/v) glycerol
30 mM $MgSO_4$	Magnesium sulfate dissolved in water.
1.25 mM Hemin	Hemine dissolved in 10 mM KOH. Filter sterilised (0.2 µm filter) and stored at 4 °C.
1 M Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid pH 6.8, autoclaved (121 °C, 20 min, 2 bar) and stored at RT.
1 M DTT	1,4-Dithiothreitol Freshly prepared and filter sterilised (0.2 µm filter)
PERS	*Pichia* electroporation recovery solution YPD : 1 M sorbitol (1:1 v/v)
PBS	Phosphate buffer saline 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4 * 7 H_20, 1.4 mM KH_2PO_4

TABLE 2.9. *P. pastoris* media for RFP and yEGFP expression

Stock	**Final concentration**	**0.2 % BMD**	**1 % BMD**	**1 % BMM**
1 M Phosphate buffer	100 mM potassium phosphate	100 mL	100 mL	100 mL
10X YNB	1.34 % (w/v) YNB	100 mL	100 mL	100 mL
500X B	4×10^{-5} % (w/v) biotin	2 mL	2 mL	2 mL
20 % (w/v) Dextrose	0.2 % - 1 % (w/v) glucose	10 mL	50 mL	-
Methanol	1 % (v/v) methanol	-	-	10 mL
ddH_2O		*ad* 1 L	*ad* 1 L	*ad* 1 L
Zeocin	75 µg/mL	750 µL	750 µL	750 µL

TABLE 2.10. *P. pastoris* media for phytase expression

Stock	**Final concentration**	**0.4 % BMGY-PP**	**2 % BMGY-PP**	**1 % BMMY-PP**	**10 % BMMY-PP**
2X YP	1 % (w/v) yeast extract, 2 % peptone	500 mL	500 mL	500 mL	500 mL
10X YNB	1.34 % (w/v) YNB	100 mL	100 mL	100 mL	100 mL
500X B	4×10^{-5} % biotin	2 mL	2 mL	2 mL	2 mL
10X GY	0.4 % - 2 % (v/v) glycerol	40 mL	200 mL	-	-
Methanol	1 % - 10 % (v/v) methanol	-	-	10 mL	100 mL
ddH_2O		*ad* 1 L	*ad* 1 L	*ad* 1L	*ad* 1L
Zeocin	75 µg/mL	750 µL	750 µL	750 µL	750 µL

TABLE 2.11. *P. pastoris* media for induced UPO expression

Stock	**Final concentration**	**1 % BMM+H**	**2 % BMM+2H**	**10 % BMM**
1 M Phosphate buffer	100 mM potassium phosphate	100 mL	100 mL	100 mL
10X YNB	1.34 % (w/v) YNB	100 mL	100 mL	100 mL
500X B	4×10^{-5} % (w/v) biotin	2 mL	2 mL	2 mL
Methanol	1 % - 10 % (v/v) methanol	10 mL	20 mL	100 mL
30 mM $MgSO_4$	6 mM $MgSO_4$	100 mL	200 mL	-
1.25 mM Hemine	10 - 20 µM hemine	8 mL	16 mL	-
ddH_2O		*ad* 1 L	*ad* 1 L	*ad* 1L
Zeocin	75 µg/mL	750 µL	750 µL	-

TABLE 2.12. *P. pastoris* media for constitutive UPO expression

Stock	**Final concentration**	**0.2 % BMD**	**1 % BMD+H**
1 M Phosphate buffer	100 mM potassium phosphate	100 mL	100 mL
10X YNB	1.34 % (w/v) YNB	100 mL	100 mL
500X B	4×10^{-5} % (w/v) biotin	2 mL	2 mL
20 % (w/v) Dextrose	0.2 % - 1 % (w/v) glucose	10	50
30 mM $MgSO_4$	3 mM $MgSO_4$	-	100 mL
1.25 mM Hemine	10 µM hemine	-	8 mL
ddH_2O		*ad* 1 L	*ad* 1L
Zeocin	75 µg/mL	750 µL	750 µL

2.2.4 Preparation of *E. coli* Competent Cells and Transformation

Preparation of *E. coli* competent cells was done as described by Chung and Miller [60]. A 5 mL overnight pre-culture was grown in LB medium to inoculate 50 mL LB medium to an OD_{600} 0.1. The 50 mL culture was grown until an OD_{600} 0.2-0.5 was reached (37 °C, 150 rpm). The cells were transferred to a sterile 50 mL reaction tube, incubated on ice for 10 minutes, and subsequently centrifuged (4 °C, 1800 x *g*, 10 min). The supernatant was discarded and remaining medium was pipetted out. The pellet was gently dissolved in 5 mL chilled TSS buffer, Table 2.7. 100 µL aliquots were prepared in chilled 1.5 mL reaction cups and stored at -80°C until further use.

For *E. coli* transformation, cells were thawed on ice, DNA (about 200 ng) was added and the cup was kept on ice for 30 minutes. Next, a heat shock at 42 °C for 90 seconds was performed and immediately afterwards the cells were placed back on ice. 1 mL of SOC or LB medium (Table 2.7) was added, and recovery of cells was performed (37 °C, 200 rpm, 1 h). Respective dilutions were plated on LB agar plates with the appropriate antibiotic.

2.2.5 Preparation of *P. pastoris* Competent Cells and Transformation

Transformation of *P. pastoris* was performed via electroporation either as described by Perez-Pinera *et al.* [234] or Madden *et al.* [192]. *P. pastoris* strains pUO_pL001 - pUO_pL124 and pUO_pL600 - pUO_pL662, Table A.5 and Table A.4, were prepared according to Perez-Pinera *et al.* [234]. For the preparation of competent cells, 5 mL YPD was inoculated in a 50 mL flask from a single, freshly grown colony (30 °C, 140 rpm). 50 µL of this pre-culture was used to inoculate 100 mL YPD in a 1 L flask and grown to an OD_{600} 1.3 - 1.5 (30 °C, 140 rpm). The cells were spun down (4 °C, 1500 x *g*, 5 min) and the pellet was gently washed as follows:

- 100 mL sterile, ice-cold water (centrifugation: 4 °C, 1500 x *g*, 5 min);
- 50 mL sterile, ice-cold water (centrifugation: 4 °C, 1500 x *g*, 5 min);
- 20 mL sterile, ice-cold 1 M sorbitol (centrifugation: 4 °C, 1500 x *g*, 5 min).

After the final centrifugation, the supernatant was decanted and remaining liquid on the pellet was carefully pipetted off. The pellet was dissolved in 1 mL 1 M sorbitol. 80 µL cell solution was transferred to chilled 1.5 mL reaction cups and directly used for transformation. The competent cells were mixed with 5 µg circular recombinase expression vector (BxbI plasmid) and 5 µg circular cassette vector. The solution was transferred into an electroporation cuvette (2 mm), incubated on ice (5 min) and then pulsed according to the parameters for *S. cerevisiae* as suggested by the manufacturer (1500 V, 25 µF, 200 Ohm). Immediately afterwards, 1 mL 1 M sorbitol was added to the cells and transferred into a 2 mL cup containing 0.8 mL 2X YP. After recovery (30 °C, 100 rpm, 8 h), 50 µL of cell culture was plated on YPD supplemented with 0.75 µg/mL zeocin. Transformation of *P. pastoris* based on the PARS (*Pichia* autonomously replication sequence) was performed as described above, but no BxbI plasmid was added and recovery was 2 h.

For all other *P. pastoris* strains, preparation of competent cells and transformation was performed according to Madden *et al.* [192]. A single fresh colony was used to inoculate 10 mL YPD in a 50 mL flask and grown over night (30 °C, 140 rpm). This pre-culture was used to inoculate 25 mL YPD in a baffled 100 mL flask to an OD_{600} 0.2 and grown to an OD_{600} 1.5 (30 °C, 140 rpm). Once the cells were grown sufficiently, the culture was placed on ice for 15 minutes. The cells were harvested in a sterile 50 mL tubes by centrifugation (4 °C, 1000 x *g*, 5 min) and resuspended in 2 mL YPD-Hepes (1.6 mL YPD + 0.4 mL 1 M Hepes, pH 6.8, Table 2.8). Carefully 75 µL of freshly prepared 1 M DTT (Table 2.8) was added and the sample was incubated (30 °C, 100 rpm, 25 min). After incubation, the sample was diluted by adding sterile, ice-cold water to make a final volume of 50 mL. The cells were centrifuged (4 °C, 1000 x *g*, 5 min) and the pellet was gently washed as following.

- 50 mL sterile, ice-cold water (centrifugation: 4 °C, 1000 x *g*, 5 min);
- 20 mL sterile, ice-cold 1 M sorbitol (centrifugation: 4 °C, 1000 x *g*, 5 min);
- 20 mL sterile, ice-cold 1 M sorbitol (centrifugation: 4 °C, 1000 x *g*, 5 min).

After the final centrifugation, the supernatant was poured off, and the remaining solution on the pellet was carefully pipetted off. The pellet was dissolved in 300 µL 1 M sorbitol. 30 µL E-competent cells aliquots were prepared in 1.5 mL reaction tubes, which were either directly used for transformation or stored at -80 °C until use. For *P. pastoris* transformation, the competent cells were mixed with 200-500 ng BxbI plasmid and 200-500 ng circular expression vector and transferred into an electroporation cuvette (1 mm). The mix was pulsed according to the parameters for *E. coli* as suggested by the manufacturer (1800 V, 25 µF, 200 Ohm). Immediately afterwards, 1 mL PERS, Table 2.8, was added to the cuvette and the mix was transferred into a 2 mL cup. After recovery (30 °C, 100 rpm, 3.5 h), 100 µL of the cell solution and the cell pellet (8000 rpm, 30 s) were plated on YPD supplemented with 0.75 µg/mL zeocin. The strains pUO_pL730 - pUO_pL771 were prepared by directly using the Golden Gate reaction for transformation, Table A.4.

2.3 DNA Manipulation

2.3.1 Plasmids and Primers

A summary of the part plasmids used can be found in the supplemental information A.1, along with the information of the DNA sequence. The DNA for part plasmid design was either obtained via PCR or DNA synthesis, and fragments were subsequently integrated into the pYTK001 entry vector, Table A.2, via Golden Gate reaction. For the PCR amplification, template plasmids are listed in Table A.12 and the respective primers are given in Table A.3. DNA fragments or complete ready to use part plasmids were ordered at Eurofins Genomics GmbH (Ebersberg, Germany), Twist Bioscience (Santa Clara, CA, USA) or Integrated DNA Technologies (Coralville, IA, USA). The MoClo-YTK was obtained from Addgene Kit: 1000000061 (Cambridge, MA, USA) Lee *et al.* [172]. Plasmids used from the YTK are listed in Table A.2. Expression plasmids made via Golden Gate reaction from part plasmids are listed in Tables A.5, A.4, A.6 and A.9. Oligonucleotides used for plasmid construction or sequence verification are listed in Table A.3 and were ordered from Integrated DNA Technologies (Coralville, IA, USA) or biomers.net GmbH (Ulm, Germany).

2.3.2 Isolation of Plasmid DNA from *E. coli*

For plasmid preparation, 7 mL cell culture was harvested by centrifugation (RT, 3500 x *g*, 5 min) and prepared as described in the corresponding kit (GeneJET™ Plasmid Miniprep Kit #K0503 (Thermo Fisher Scientific, Waltham, MA, USA) or Pure Yield Plasmid Miniprep System #A1222 (Promega, Madison, WI, USA)). DNA concentration and purity was measured by photometric absorption detection using the nanophotometer. The plasmid DNA was stored at 4 °C for frequent usage, otherwise at -20 °C.

2.3.3 Isolation of Genomic DNA from *P. pastoris*

Extraction of genomic DNA was performed as described by Looke *et al.* (2011) [187], with modifications. A single colony was resuspended in 100 µL genomic DNA extraction buffer (0.2 M lithium acetate, 1 % (w/v) SDS) and incubated (70 °C, 5 min). 300 µL 96 % (v/v) ethanol was added, the sample vigorously vortexed, and the cell debris were spun down (RT, 15000 x *g*, 3 min). The pellet was washed with 300 µL 70 % (v/v) ethanol (RT, 15000 x *g*, 3 min) and the dried pellet (70 °C, 5 min) was dissolved in 100 µL water. 2 µL solution was used for PCR applications.

2.3.4 Gel Electrophoresis

To separate DNA fragments based on size, electrophoresis was carried out as described by Ausubel *et al.* [21] using the Mini-Sub® Cell GT Cell system (Bio-Rad Laboratories GmbH, München, Germany). Samples were mixed with 5X loading dye and loaded to the solidified 1 % (w/v) agarose gel with a nucleic acids stain, Table 2.13.

To determine the DNA fragment size, a marker was added (Quick-Load® Purple 1 kb DNA Ladder #N0552G or Quick-Load® Purple 2-Log DNA Ladder #N0550S (NEB, Ipswich, MA, USA)) and the sample was run at 120 V for 20 minutes. Preparative visualization was performed using blue light, gel documentation occurred under UV-radiation. To purify distinct DNA fragments, the respective fragment was extracted from the agarose and prepared according to the corresponding kit (NucleoSpin® Gel and PCR Clean-up Kit #740609.250 (Macherey-Nagel GmbH & Co. KG, Düren, Germany) or Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA)).

TABLE 2.13. Solutions for DNA gel electrophoresis

Solution	**Composition**
TAE buffer	Tris-acetate-EDTA 40 mM Tris (pH 7.6), 20 mM acetic acid, 1 mM Ethylenediaminetetraacetic acid (EDTA) A 50X TAE stock solution was prepared and prior the electrophoresis the 1X TAE buffer was made.
1 % (w/v) Agarose	10 g/L agarose in 1X TAE buffer The solution was brought to a boil and stored at 60 °C. A nucleic acid stain was added prior use.
5X Loading dye	0.075 mM Tris pH 7.6, 0.05 mM EDTA, 50 % (v/v) glycerol, 0.025 % (w/v) bromphenol blue, 0.025 % (w/v) xylene cyanol
Nucleic acids stain	20000X SERVA DNA Stain Clear G (Serva, Heidelberg, Germany) or 10000X SYBR Safe DNA Gel Stain (Thermo Fisher Scientific, Waltham, MA, USA)

2.3.5 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was performed as described by Ausubel *et al.* [21]. For gene amplification and fusion of fragments, the Phusion® High-Fidelity DNA Polymerase #M0530 (New England Biolabs (NEB), Ipswich, MA, USA) or the KAPA HiFi DNA Polymerase (Kapa Biosystems, Wilmington, MA, USA) were used as described by their manufacturer, Table 2.14 and Table 2.15. To verify *E. coli* transformation, the Taq DNA Polymerase with ThermoPol® Buffer #M0267 (NEB, Ipswich, MA, USA) was applied and a single colony was directly dissolved in the PCR reaction, Table 2.16. To check *P. pastoris* transformation, genomic DNA was extracted according to Section 2.3.3 and the GoTaq® G2 DNA Polymerase #M7845 (Promega, Madison, WI, USA) was applied according to the manufacturer, Table 2.17. The KAPA2G Robust DNA Polymerase (Kapa Biosystems, Wilmington, MA, USA) was also used to verify *E. coli* and *P. pastoris* transformation. For *E. coli*, a single colony was directly resuspended in the PCR reaction. For *P. pastoris* a colony was resuspended in 20 mM NaOH, heated (95 °C, 10 min) and 1 µL of supernatant after centrifugation (RT, 6200 x *g*, 30 s) was used, Table 2.18.

Table 2.14: Conditions for the Phusion® High-Fidelity DNA Polymerase

Composition		**Program**			
Component	Amount	Step	[s]	[°C]	Cycles
5X HF-buffer	10.0 µL	Initial denaturation	30	98	1
10 mM dNTP mix	1.0 µL	Denaturation	10	98	
10 µM Primer fw	2.5 µL	Primer annealing	30	60	30
10 µM Primer rev	2.5 µL	Elongation	20/kb	72	
DNA template	100 ng	Final elongation	600	72	1
Phusion polymerase	0.5 µL	Hold		12	∞
ddH_2O	*ad* 50 µL				

Table 2.15: Conditions for the KAPA HiFi Polymerase

Composition		**Program**			
Component	Amount	Step	[s]	[°C]	Cycles
5X Kapa HiFi buffer	10.0 µL	Initial denaturation	180	95	1
10 mM dNTP mix	1.0 µL	Denaturation	20	98	
10 µM Primer fw	1.5 µL	Primer annealing	15	60	25
10 µM Primer rev	1.5 µL	Elongation	12/kb	72	
DNA template	100 ng	Final elongation	60	72	1
Kapa polymerase	1 µL	Hold		12	∞
ddH_2O	*ad* 50 µL				

Table 2.16: Conditions for the Taq DNA Polymerase

Composition		**Program**			
Component	Amount	Step	[s]	[°C]	Cycles
10X ThermoPol	2.5 µL	Initial denaturation	30	95	1
10 mM dNTP mix	0.5 µL	Denaturation	30	95	
10 µM Primer fw	0.5 µL	Primer annealing	30	60	30
10 µM Primer rev	0.5 µL	Elongation	60/kb	68	
DNA template	variable	Final elongation	300	72	1
Taq polymerase	0.125 µL	Hold		12	∞
ddH_2O	*ad* 25 µL				

Table 2.17: Conditions for the GoTaq® G2 DNA Polymerase

Composition		**Program**			
Component	Amount	Step	[s]	[°C]	Cycles
5X Green Buffer	5.0 µL	Initial denaturation	120	95	1
10 mM dNTP mix	0.5 µL	Denaturation	30	95	
10 µM Primer fw	1.25 µL	Primer annealing	30	60	30
10 µM Primer rev	1.25 µL	Elongation	60/kb	72	
DNA template	variable	Final elongation	300	72	1
GoTaq G2 polymerase	0.125 µL	Hold		12	∞
ddH_2O	*ad* 25 µL				

Table 2.18: Conditions for the KAPA2G Robust DNA Polymerase

Composition		**Program**			
Component	Amount	Step	[s]	[°C]	Cycles
5X Kapa 2G buffer	5.0 µL	Initial denaturation	180	95	1
10 mM dNTP mix	0.5 µL	Denaturation	15	95	
10 µM Primer fw	1.25 µL	Primer annealing	16	60	35
10 µM Primer rev	1.25 µL	Elongation	20/kb	72	
DNA template	1.0 µL	Final elongation	90	72	1
Kapa Robust polymerase	0.1 µL	Hold		12	∞
ddH_2O	*ad* 25 µL				

2.3.6 Digestion, Ligation and Golden Gate Assembly

For each restriction, 1 µg DNA was digested by 10 units of restriction enzyme at the respective digestion temperature, using the suitable buffer for each enzyme. The digestion duration was 2-12 hours. The enzymes used were BsaI, BsmBI, NdeI and XhoI (NEB, Ipswich, MA, USA). After agarose gel purification of the DNA fragment, DNA ligation occurred as described by the manufacturer. T4 and T7 ligase (NEB, Ipswich, MA, USA) were utilised. DNA, the respective ligase buffer, and water were mixed and ligase was added last. After incubation (RT, 30 min) the ligation mixture was directly used for chemical transformation of *E. coli*, Section 2.2.4. When required for the cloning strategy, linear DNA ends were treated with the Alkaline Phosphatase, Calf Intestinal (CIP) #M0290S or the Blunt Enzyme Mix #E1201S (NEB, Ipswich, MA, USA). Plasmids following the MoClo standard were designed using Golden Gate assembly [312]. The design of the *Pichia* Toolkit (PTK) was performed using the specific design rules and characteristic overhangs as suggested for the Yeast Toolkit(YTK) by Lee *et al.* [172]. Part plasmids were assembled using the part plasmid entry vector pYTK001 and the insert originated from a PCR fragment, gBlock or plasmid (Section 2.3.1), using the conditions in Table 2.19.

Expression vectors or expression vector libraries were assembled from part plasmids, ensuring a plasmid for each functional element. Various assembly strategies were utilised. Plasmids pUO_pL001-pUO_pL124 were prepared by individual restriction enzyme digestion and ligation as described above, from each individual part plasmid. All other expression vectors as well as the randomised libraries utilised the backbone vector pUO_pL501. Expression vector assembly was performed according to Table 2.19 and the Golden Gate reaction was used for chemical transformation of *E. coli* or for plasmids pUO_pL730 – pUO_pL771 directly for *P. pastoris* transformation (Section 2.2.5). For randomised libraries, the Golden Gate reaction was performed as described in Table 2.20 [93]. The total amount of insert for each functional element was twice as much as the backbone vector pUO_pL501. For each functional element which was aimed to be shuffled, the total amount of insert was divided by the number of plasmids for the shuffling.

E.g. for a Golden Gate shuffling with 300 fmol insert and 3 plasmids to be shuffled, 100 fmol of each plasmid were used for the assembly reaction. After the Golden Gate shuffling, chemical transformation of *E. coli* was performed.

Table 2.19: Golden Gate reaction for part plasmid design

Composition		**Program**			
Component	Amount	Step	[min]	[°C]	Cycles
Insert or plasmid	0.5 µL	Digestion	2	42	25
T4 ligase buffer	1.0 µL	Ligation	5	16	
T7 ligase	0.5 µL	Final digestion	10	60	1
Restriction enzyme	0.5 µL	Heat inactivation	10	80	1
ddH_2O	*ad* 10 µL	Hold		12	∞

Table 2.20: Golden Gate reaction for shuffling approach

Composition		**Program**			
Component	Amount	Step	[min]	[°C]	Cycles
Total insert	300 fmol	Initial digestion	10	37	1
Plasmid backbone	150 fmol	Digestion	5	37	50
T4 ligase buffer	1.0 µL	Ligation	5	16	
T4 ligase	0.5 µL	Final digestion	10	37	1
Restriction enzyme	0.5 µL	Heat inactivation	10	65	1
ddH_2O	*ad* 20 µL	Hold		12	∞

2.3.7 DNA Sequencing

Sequencing was done externally (GATC Biotech AG, Konstanz, Germany or Eurofins Genomics GmbH, Ebersberg, Germany). The primer was directly added into the sample, primers are listed in Table A.3. Evaluation was done using SnapGene (GSL Biotech LLC, Chicago, IL, USA) or Benchling (San Francisco, CA, USA). Information on the genomic DNA of *P. pastoris* was gathered from National Center for Biotechnology Information (NCBI) (Bethesda, MD, USA).

2.4 Protein Purification and Characterisation

2.4.1 SDS-PAGE

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) was performed for protein separation as described by Ausubel *et al.* [21] with modifications. To prepare polyacrylamide gels, the Mini/PROTEAN® Tetra Cell Casting Stand and Clamps (Bio-Rad Laboratories GmbH, München, Germany) were used. For 2 gels, a separating gel was prepared according to Table 2.21, the solution was swirled gently, immediately used to load the chambers and covered with water. After polymerization (RT, 60 min), the water layer was removed, the stacking gel layer added, and the desired comb inserted. After polymerization (RT, 60 min), the gels were stored at 4 °C until usage, wrapped in wet paper towels. To analyse a protein sample, it was mixed with 5X loading buffer (50 % (v/v) glycerol, 12.5 % (v/v) β-mercaptoethanol, 7.5 % (w/v) SDS, 0.25 g/l bromphenol blue) and heated (95 °C, 5 min). The sample was loaded to the SDS-PAGE, 1X SDS electrophoresis buffer (0.1 % (w/v) SDS, 25 mM Tris, 192 mM glycine) was added to the Mini-PROTEAN® Tetra Vertical Electrophoresis Cell and the gel was run for about 45 min with 40 mA for each gel. To determine the protein size, the PageRuler Unstained Protein Ladder #26614 or the PageRuler Prestained Protein Ladder #26616 (Thermo Fisher Scientific, Waltham, MA, USA) were applied. After the electrophoretic run, the gel was washed with water, stained using a coomassie-staining solution (0.2 % (w/v) coomassie brilliant blue G250 and R250, 50 % (v/v) ethanol, 10 % (v/v) acetic acid, filtered and stored protected from light) and discoloured using water. To study the glycosylation content of a protein, the sample was treated with PNGase F #P0705S (NEB, Ipswich, MA, USA) as described by the manufacturer and loaded to a SDS-PAGE.

TABLE 2.21. Solutions for SDS-PAGE

Solution	**Separating gel**	**Stacking gel**	**Composition**
Acrylamide concentration	12 %	5 %	
Acrylamide	4 mL	0.83 mL	30 % (v/v) acrylamide / 0.8 % (v/v) bisacrylamide
ddH_2O	3.29 mL	2.77 mL	
Separating gel buffer 4X	2.5 mL	-	0.8 % (w/v) SDS, 1.5 M Tris/HCl, pH 8.8 using hydrochloric acid (HCl)
Staking gel buffer 4X	-	1.25 mL	0.8 % (w/v) SDS, 0.5 M Tris/HCl, pH 6.8 using HCl
APS	100 µL	50 µL	10 % (w/v) ammoniumpersulfate
TEMED	10 µL	5 µL	N,N,N',N'-tetramethylethane-1,2-diamine

2.4.2 Protein Quantification Assay According to Bradford

Quantification of the protein amount was performed according to the method of Bradford [37], using Roti-Quant (CARL ROTH GmbH & Co. KG, Karlsruhe, Germany) according to the manu-

facturers description. A calibration curve was made using bovine serum albumin (BSA) in a range from 20 to 100 µg/mL. In a Greiner 96 Flat Transparent plate, 50 µL sample or BSA reference was added first and subsequently the Roti-Quant assay solution (2 volume Roti-Quant, 5.5 volume ddH_2O) was added . After 5 minutes of incubation at RT, the absorption of the Coomassie Brilliant Blue Dye-G250 was measured at 595 nm. The BSA reference absorption values were used to determine the protein concentration in the sample.

2.4.3 Fluorescence Assay for RFP and yEGFP Determination

For fluorescence measurements, 100 µL culture or 100 µL culture supernatant (centrifugation: 4 °C, 500 x *g*, 10 min) were measured, Table 2.22 for plate reader settings. Fluorescent intensity of each sample was normalized to its OD_{600}, if not stated otherwise. For OD_{600} determination, samples were diluted using Phosphate-buffered saline (PBS) buffer, pH 7.4. Background fluorescence was determined from the parent *P. pastoris* strain not expressing any fluorescent protein. This strain was assayed along with the constructs on each plate. The mean background was calculated over all parent *P. pastoris* strains and subtracted from the expression strains.

TABLE 2.22. Plate reader settings for fluorescence measurements

Parameter	**Setting**
Name	UO_Fluoreszenz_RFP_GFP_BMD.mth
Plate	Nunclon 96 Flat Black
Fluorescence	Label: GFP
Excitation wavelength	488 nm
Excitation bandwidth	9 nm
Emission wavelength	515 nm
Emission bandwidth	20 nm
Measuring mode	From top
Delay time	0 µs
Integration time	20 µs
Flashes	25
Gain	50
Fluorescence	Label: RFP
Excitation wavelength	561 nm
Excitation bandwidth	9 nm
Emission wavelength	600 nm
Emission bandwidth	20 nm
Measuring mode	From top
Delay time	0 µs
Integration time	20 µs
Flashes	25
Gain	100
Name	UO_OD600.mth
Plate	Greiner 96 Flat Transparent
Absorption	One measurement
Wavelength	600 nm
Bandwidth	9 nm
Flashes	25

2.4.4 Colorimetric Assay Determining Phytase Activity

The phytase activity was determined in the *P. pastoris* culture supernatant by assessing the free phosphate released from phytate using ammonium molybdate as colouring agent to perform colorimetric quantification. Cells were separated from the culturing broth via centrifugation (4 °C, 500 x *g*, 10 min). The supernatant was then subjected to gel-filtration chromatography to purify the secreted protein from smaller molecules such as free phosphate. 96-Well SpinColumns (25-100 µL) (Harvard Apparatus, Holliston, MA USA) were used as described by the manufacturer. New columns or columns stored in 20 % (v/v) ethanol were centrifuged (RT, 2000 x *g*, 2 min), hydrated for 20 min using 200 µL water per well and centrifuged again (RT, 2000 x *g*, 2 min). The columns were washed 3 x using 150 µL water and were centrifuged again (RT, 2000 x *g* for 2 min), except for the last centrifugation step where 1000 x *g* were used. 35 µL *P. pastoris* supernatant was loaded to each of the SpinColumn wells and the plate was centrifuged (RT, 1000 x *g*, 2 min) using a Nunc 96-Well MicroWell plate to collect the filtrate. The SpinColumn plate was washed as before using 3 x 150 µL water (centrifugation: RT, 2000 x *g*, 2 min) and either used again for gel filtration or stored with 150 µL 20 % (v/v) ethanol in each well at 4 °C.

Phytase activity was measured as described by Bae *et al.* [24], and adapted for 96-well-plate applications. 13.5 µL of the gel filtration filtrate, or a 1:10 dilution thereof using water to dilute, was transferred into a 96-well assay plate. 53.5 µL freshly prepared phytate substrate solution was added to each well and the plate was incubated (37 °C, 30 min). To terminate the reaction, 66.6 µL stop solution as well as 66.6 µL colouring solution were added and absorbency was measured at 700 nm, Tables 2.23. To determine phytase units (U), a calibration curve using potassium phosphate in a range from 0.8 - 20 mM was used.

TABLE 2.23. Colorimetric phytase assay, solutions and plate reader conditions

Solution	**Composition**
Substrate solution	1.5 mM sodium phytate in 0.1 M sodium acetate buffer pH 5.0 using acetic acid Always prepared fresh.
Stop solution	5 % (w/v) trichloroacetic acid (TCA) Stored at RT.
Colouring reagent A	1.5 % (w/v) ammonium molybdate in 5.5 % (v/v) sulfuric acid solution Stored at 4 °C in the dark.
Colouring reagent B	2.7 % (w/v) ferrous sulfate solution Stored at 4 °C in the dark.
Colouring solution	4 volumes colouring reagent A and 1 volume colouring reagent B Always prepared fresh.
Parameter	**Setting**
Name	UO_Phytase_700_Endpunkt.mth
Plate	Greiner 96 Flat Transparent
Absorption	One measurement
Wavelength	700 nm
Bandwidth	9 nm
Flashes	5

2.4.5 Colorimetric Assay Determining Unspecific Peroxygenase Activity

TABLE 2.24. Colorimetric UPO assay, stock solutions and plate reader conditions

Stock Solution	**Composition**
100 mM sodium phosphate / citrate puffer	27.8 mL 0.1 M citric acid, 22.2 mL 0.2 M dibasic sodium phosphate, 50 mL water, pH 4.4 stored at RT
3 mM ABTS	2,2'-azino-bis(3-ethylbenzthiazoline-6-sulphonic acid) in 100 mM sodium phosphate / citrate puffer Stored at - 20 °C.
20 mM H_2O_2	Hydrogen peroxide in 100 mM sodium phosphate / citrate puffer, prepared freshly.
Parameter	**Setting**
Name	UO_UPO_418_Steigung_15Min.mth
Plate	Greiner 96 Flat Transparent
Kinetics	30 cycles, wait: 30 sec
Absorption	One measurement
Wavelength	418 nm
Bandwidth	9 nm
Flashes	5
Name	UO_UPO_418_480_Endpunkt.mth
Plate	Greiner 96 Flat Transparent
Absorption	Measurement with reference
Wavelength	418 nm
Bandwidth	9 nm
Reference wavelength	480 nm
Reference bandwidth	9 nm

Unspecific peroxygenases (UPOs) were assayed for their peroxidase activity. The oxidation of ABTS in the presence of the sole cofactor H_2O_2 was monitored photometrically at 418 nm. The expression supernatant was prepared by centrifugation (4 °C, 500 x *g*, 10 min) and used without further purification. The enzymatic assay was performed as described by Molina-Espeja *et al.* [208] with modifications.

The UPO assay solution was prepared freshly from the stock solutions (Table 2.24). For the screening studying induced expression of UPOs, 20 µL of supernatant was added to a 96-well assay plate and the reaction was started by addition of 180 µL assay solution. The final assay conditions were 0.3 mM ABTS and 0.2 mM H_2O_2 in 100 mM sodium phosphate / citrate puffer. For the screening studying constitutive UPO expression, 50 µL *P. pastoris* supernatant was used, and respectively 150 µL UPO assay solution was added. The final assay conditions were 0.75 mM ABTS and 0.2 mM H_2O_2 in 100 mM sodium phosphate / citrate puffer. For both screenings, the absorption was measured at 418 nm for 15 min, every 30 s. After total reaction times of 1 h or 12 h endpoint measurements were performed, detecting absorption at 418 nm and 480 nm Table 2.24.

To concentrate the expression supernatant, centrifugal filters (PES membrane, pore size 10 K) were used. 400 µL supernatant was added to the column and centrifuged (RT, 10000 x *g*, 5 min). The flow through was discarded and more sample was loaded to the column and centrifuged as before. This process was repeated until the supernatant was concentrated as desired. The concentrated sample was removed from the column filter and used for UPO peroxidase activity determination or SDS-PAGE.

2.4.6 Alcohol Dehydrogenase Purification and Analysis

Alcohol dehydrogenase (ADH) activity was studied following the cofactor turnover for the reduction of an aldehyde to an alcohol. The substrate butyraldehyde was applied and the NADPH or NADH decrease was photometrically monitored at 340 nm. The expression supernatant was prepared by centrifugation (4 °C, 500 x *g*, 10 min) and either used directly or a affinity chromatography and subsequent sample clean-up using a desalting column of the supernatant was performed.

For the ADH purification, a His GraviTrap pre-packed, gravity-flow column containing precharged Ni Sepharose 6 Fast Flow was used according to the manufacturers description. 10 mL His column binding buffer, Table 2.25, were used for column equilibration. Subsequently, 300 mL expression supernatant, followed by 10 mL His column binding buffer were added to the column and the flow through was discarded. Protein elution was performed in two consecutive steps each using 3 mL His column elution buffer and the elution fraction were collected in a fresh tube. To wash the column before storage, one after the other solution was added to the column: 20 mL His column elution buffer, 20 mL ddH_2O, 20 mL 20 % (v/v) ethanol.

To desalt the His GraviTrap eluate, a PD-10 Desalting Column containing Sephadex G-25 was used. The column was washed using 10 mL ddH_2O and subsequently 25 mL 50 mM Bis-Tris buffer, pH 6. The first His GraviTrap elution fraction (approximately 3 mL) was loaded to the column. For sample elution, in two consecutive steps, 3 mL 50 mM Bis-Tris buffer, pH 6 were added to the column and each faction was collected in a fresh tube. The column was washed using 30 mL ddH_2O and 30 mL 20 % (v/v) ethanol before storage. The elution fractions were used for ADH activity assay or loaded on a SDS-PAGE for protein size analysis.

For the ADH assay, the cofactor NADPH was used for AdhZ2_Ec or AdhZ2_7476, and the cofactor NADH was used for AdhZ2_DIN, AdhZ3_LND or AdhZ3_242. For the measurements, 50 µL supernatant or purified supernatant were added to a 96-well assay plate and the reaction was started by the addition of freshly made and to 37 °C pre-warmed assay solution, containing 5 mM butyraldehyde, 0.5 mM NADH or 0.3 mM NADPH in 50 mM Bis-Tris buffer. The cofactor turnover was followed at 340 nm for 15 min, every 30 s, Table 2.25.

TABLE 2.25. Alcohol dehydrogenase purification and assay

Purification buffer	**Composition**
His column binding buffer	50 mM potassium phosphate, 500 mM NaCl, 10 % (v/v) glycerol, 20 mM imidazole, pH 8
His column elution buffer	50 mM potassium phosphate, 500 mM NaCl, 10 % (v/v) glycerol, 500 mM imidazole, pH 8
Desalting buffer	50 mM Bis-Tris buffer, pH 6
ADH Assay stock solution	**Composition**
50 mM Bis-Tris buffer	pH 6.0, stored at RT
50 mM Butyraldehyde	in 50 mM Bis-Tris buffer
5 mM NADH	in 50 mM Bis-Tris buffer
3 mM NADPH	in 50 mM Bis-Tris buffer
Parameter	**Setting**
Name	UO_ADH_340_Steigung_15Min.mth
Plate	Greiner 96 Flat Transparent
Temperature	37 °C
Shaking	10 s, 44.3 U/min
Kinetics	Duration: 15 min, wait: 30 s
Absorption	One measurement
Wavelength	340 nm
Bandwidth	9 nm
Flashes	10

2.4.7 SYPRO Orange Thermal Shift Assay

The detection of enzymes (aiming alcohol dehydrogenases) in the *P. pastoris* expression-culture supernatant was performed using a SYPRO® Orange thermal shift assay. SYPRO Orange is 10-100 times more sensitive than the commonly used colorimetric assays [136] and can be used to monitor protein unfolding during increasing temperatures [185]. A SYPRO Orange working solution was made using 3 µL SYPRO Orange Protein Gel Stain (Sigma-Aldrich, Deisenhofen, Germany, S5692-50UL) and 237 µL ddH_2O. 23 µL expression supernatant were combined with 2 µL SYPRO Orange working solution in Hard-Shell® PCR Plates (thin wall, #HSP9601) and sealed with Microseal® 'B' (MSB1001), (Bio-Rad Laboratories GmbH, München, Germany). A C1000™ Thermal cycler (Bio-Rad Laboratories GmbH, München, Germany) was used to monitor change of the fluorescence using the FRET Scan mode (extinktion 485/20nm, emission: 530/30nm). The temperature was keep at 5 °C for 5 minutes and then increased by 1 °C per min to a final temperature of 95 °C.

2.5 Cellular Imaging Techniques

2.5.1 Microscopy

Microscopy was performed either using a Axio Lab.A1, Axio ObserverZ.1 (Filter sets 38 HE and 43 HE) or ECLIPS Ti-E Microscope (Table 2.1). Subsequent image analysis was performed using the corresponding software ZEN 2.3 (blue edition), NIS Elements Viewer, or Fiji version 2.0.0-rc-54/1.51h. For comparative cell imaging of RFP and yEGFP expression, the conditions were kept identical such as for TL Bright: 15000 msec, 3.5 V, processing 20-15000, for RFP (mCherry): 5000 msec, 3.5 V, processing 300-8000, and for yEGFP (EGFP): 500 msec, 3.5 V, processing 300-15000.

2.5.2 Flow Cytometry

Analysis of cells and capsules was performed using a BD LSR II HTS-2 flow cytometer (Table 2.1). For data acquisition, the corresponding FACS Diva software was used and the parameters for the individual channel were: FSC: 220 V, SSC: 150 V, FITC: 500 V, PETexRed: 600 V. A threshold was set for FSC at 200 and for each run, 10.000 events were measured. Data analysis was performed using the software FloJo.

2.6 Cell Encapsulation Technique

Calcium carbonate template formation and the layer-by-layer (LbL) technology was studied based on protocols previously described in literature [72, 85, 98, 236]. The overall methodology was performed as described below, however modifications to this protocol were made and are discussed in detail in Chapter 5. The solutions used are listed in Table 2.26.

TABLE 2.26. Solution for polyelectrolyte microcapsules

Name	**Final concentration**	**Solvent**	**Comment**
Calcium chloride, $CaCl_2$	0.66 M	ddH_2O	Filtered, 0.2 µm
Ethylenediaminetetraacetic acid, EDTA	0.25 M	ddH_2O	pH 8 using NaOH
Glycerol	60 % (v/v)	ddH_2O	Filtered, 0.2 µm
Poly(allylamine hydrochloride), PAH	5 mg/mL	0.05 M NaCl	Filtered, 0.2 µm
Poly(sodium 4-styrenesulfonate), PSS	5 mg/mL	0.05 M NaCl	Filtered, 0.2 µm
Poly(sodium 4-styrenesulfonate), PSS	5 mg/mL	0.66 M $CaCl_2$	Filtered, 0.2 µm
Sodium carbonate, Na_2CO_3	0.33 M	ddH_2O	Filtered, 0.2 µm
Sodium chloride, NaCl	0.05 M	ddH_2O	pH 6.5 using NaOH, autoclaved
Sodium hydroxide, NaOH	10 M	ddH_2O	Filtered, 0.2 µm

For cell co-precipitation studies, 1 mL of *P. pastoris* culture with an OD_{600} 5 was washed three times with 1 mL 0.05 M NaCl (each centrifugation: RT, 9000 x *g*, 1 min). The final cell pellet was resuspended in 0.5 mL 60 % (v/v) glycerol and transferred into a small beaker containing a stirring bar. 0.5 mL 0.66 M $CaCl_2$ containing 5 mg/mL PSS were added while mixing at 500 rpm. For the calcium carbonate precipitation, 1 mL 0.33 M Na_2CO_3 was added while mixing (RT, 500 rpm, 1 min). This solution was transferred to a 2 mL reaction tube and taped to a rotator for incubation (RT, 20 rpm, 10 min). These templates were washed two times with 1 mL 0.05 M NaCl (each centrifugation: RT, 500 x *g*, 30 s).

For the LbL coating, the polyelectrolytes were alternatively adsorbed with thorough washing between each adsorption step. Adsorption was performed as following, until a total of 8 layers $(PAH/PSS)_4$ resulted.

- 1 mL 5 mg/mL PAH, incubation on rotator: RT, 20 rpm, 10 min;
- 3 x washing using each 1 mL 0.05 M NaCl, each centrifugation: RT, 500 x *g*, 30 s;
- 1 mL 5 mg/mL PSS, incubation on rotator: RT, 20 rpm, 10 min;
- 3 x washing using 1 mL 0.05 M NaCl for each centrifugation: RT, 500 x *g*, 30 s.

Optionally, an additional layer of FITC-PAH was added. To achieve this, the coated templates were incubated in 2 mg/mL FITC-PAH (RT, 20 rpm, 10 min), centrifuged (RT, 500 x *g*, 30 s) and washed 3 x using 1 mL 0.05 M NaCl (each centrifugation: RT, 500 x *g*, 30 s).

Finally, the calcium carbonate template was dissolved to make hollow capsules. The coated templates were incubated in 1 mL 0.25 M EDTA, pH 8.0 (RT, 20 rpm, 3 min) and the coated templates/capsules were centrifuged (RT, 1500 x *g*, 3 min). To ensure complete dissolving of the template, the coated templates/capsules were again incubated in EDTA and centrifuged as before. The capsules were washed twice using 1 mL 0.05 M NaCl (each centrifugation RT, 1500 x *g*, 3 min) and studied using microscopy of flow cytometry.

3

A Modular Toolkit for *P. pastoris* Expression Libraries

The yeast *Pichia pastoris* is a well-established expression host, but optimisation of protein secretion in this host remains a challenge due to the multiple steps involved during secretion and a lack of genetic tools to tune this process. A toolkit of standardised regulatory elements specific for *P. pastoris* was developed to overcome this issue, allowing the tuning of gene expression and choice of protein secretion tag. As protein secretion is a complex process, these parts are compatible with a hierarchical assembly method to enable the generation of large and diverse libraries in order to explore a wide range of secretion constructs, achieve successful secretion, and better understand the regulatory factors of importance to specific proteins of interest. In this Chapter, the performance of the parts is validated by constructing 242 *P. pastoris* strains to study different regulatory elements for the expression and secretion efficiency of two established fluorescent reporter proteins (RFP, yEGFP). These parts are used more extensively in Chapter 4, to build randomised secretion libraries for industrially relevant proteins. Excerpts of this work were previously published and are reproduced herein with permission from Obst U., Lu T.K. & Sieber V. (2017) "A Modular Toolkit for Generating *Pichia pastoris* Secretion Libraries." ACS Synth. Biol. 6, 6, 1016-1025. Copyright 2017 American Chemical Society [224].

3.1 Design of the Toolkit

Genetic parts often have varying performance when used in different contexts [46, 267], which makes it difficult to predict the parts necessary to achieve optimal expression for a particular gene of interest. High-throughput screening techniques are widely employed in bacterial systems to tackle this problem and explore many different combinations of genetic parts to tune gene expression. In contrast, yeast expression libraries are relatively small, with limited diversity in their sequence and functionality [259].

To allow effective *P. pastoris* screenings, the Yeast Toolkit (YTK) [172] was extended with additional elements that allow the control of protein expression and secretion in *P. pastoris*. Furthermore, functional elements of the YTK were characterised in *P. pastoris*.

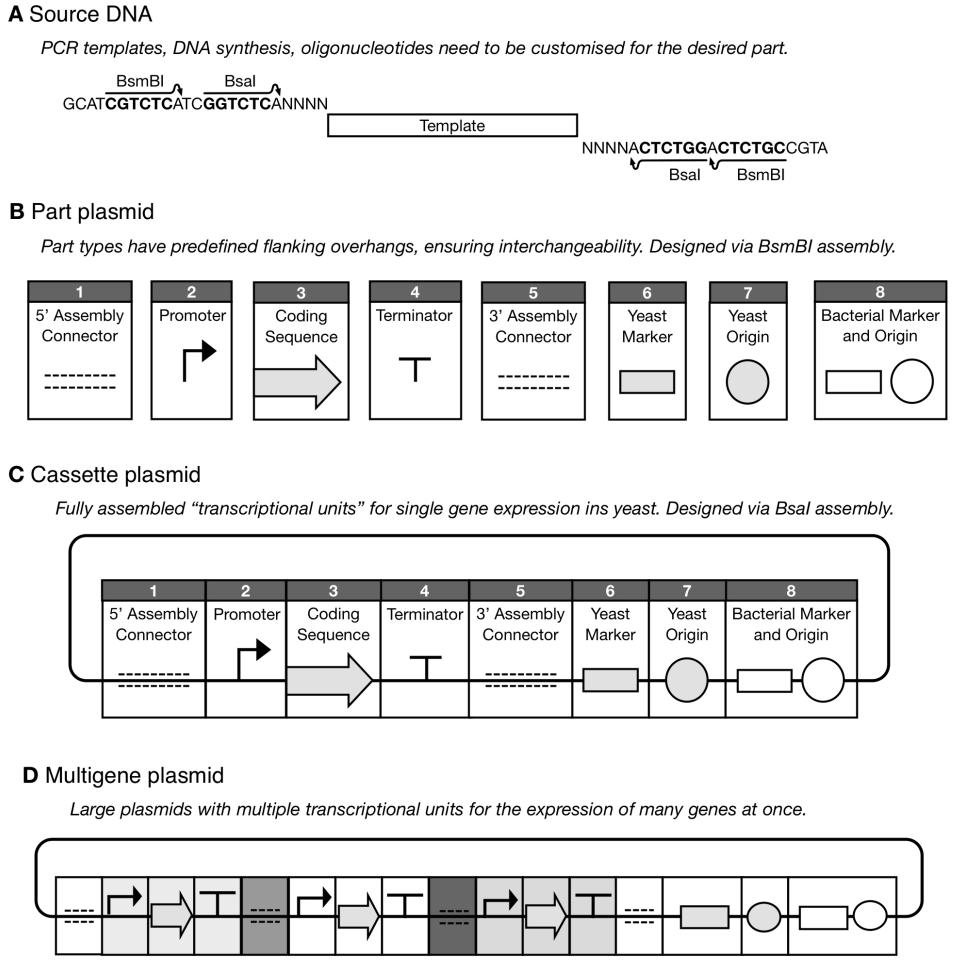

FIGURE 3.1. The *S. cerevisiae* yeast toolkit (YTK) hierarchical assembly strategy. **A** Template DNA needs to be assigned to a certain part type to define the characterising 4 base overhang and can be made via PCR or DNA synthesis. **B** Thereof the part plasmid is assembled using a BsmBI Golden Gate reaction together with the entry vector. All 'part plasmids' of one specific type have the identical overhang to allow interchangeability. **C** A functional expression vector is then assembled via a BsaI Golden Gate reaction and due to the defined overhangs correct assembly is ensured. Multigene plasmids containing up to 6 transcriptional units can be designed applying 'assembly connectors' via a BsmBI Golden Gate reaction. Figure adapted from Lee *et al.* 2015 [172].

The YTK was originally developed for *S. cerevisiae* to offer a simple and robust method for hierarchical bottom-up design and construction of expression vectors (Figure 3.1). It consists of 96 'part plasmids' that encode the basic elements of the expression vector. These are combined to produce 'cassette plasmids' that contain elements necessary to express a single gene, and optionally, these can be assembled into 'multigene plasmids' that encode an entire metabolic pathway or genetic circuit. The assembly method relies on MoClo [94, 312] using Type IIs restriction enzymes. A core feature of the YTK is the carefully defined structure of 8 primary encoding parts, which combine in a pre-defined order to a functional plasmid. Three of the parts are transcriptional units (TU) comprising of a promoter, gene of interest, and terminator, which allow the fine-tuning of gene expression. Five other types of parts control aspects such as the position of integration in the yeast genome or *E. coli* marker and origin (Figure 3.2 C) [172]. Several of these types of parts can be split into sub-parts to allow for cases where tags might need to be appended (e.g., His-tagged coding regions). When designing cassette plasmids for *P. pastoris*, some parts (assembly connectors and *E. coli* marker and origin) can be directly applied from the YTK. However, parts of the transcriptional unit, as well as the marker and origin, have to be designed specifically for the application in *P. pastoris*.

42 new *P. pastoris* specific control elements (10 promoters, 29 secretion tags, 1 terminator, and 2 origins of replication) as well as red fluorescent protein (RFP) and yeast enhanced green fluorescent protein (yEGFP) genes that can be expressed either intracellularly or secreted were designed. DNA for these parts was obtained via PCR or synthesised and where necessary (see Table A.1), BsmBI or BsaI restriction sites were removed to ensure compatibility with the assembly method. These new *P. pastoris* parts, combined with the selection of tested and functional YTK parts in *P. pastoris*, offer the ability to express a gene of interest in 580 different ways. By combining the *P. pastoris* parts with all YTK's existing *S. cerevisiae* parts (that may also function in *P. pastoris*), over 13,000 different possibilities could be obtained. Regulatory elements were successfully applied from the *S. cerevisiae* YTK in *P. pastoris*, highlighting the flexible platform design that offers the potential for further extensions to other yeast hosts. The combinatorial possibilities increase even further when considering multi-gene expression cassettes [172]. These multi-gene cassettes can open up opportunities for the co-expression of secretion-related proteins. The overexpression of chaperones, foldases and trafficking proteins may enhance secretion capacities by improving protein-trafficking for ER-to-Golgi and Golgi-to-plasma membrane processes or by supporting the folding and modification processes in the ER [137, 155]. For the part characterisation, 242 cassette plasmids were designed and characterised for RFP and yEGFP expression (Tables A.5, A.4).

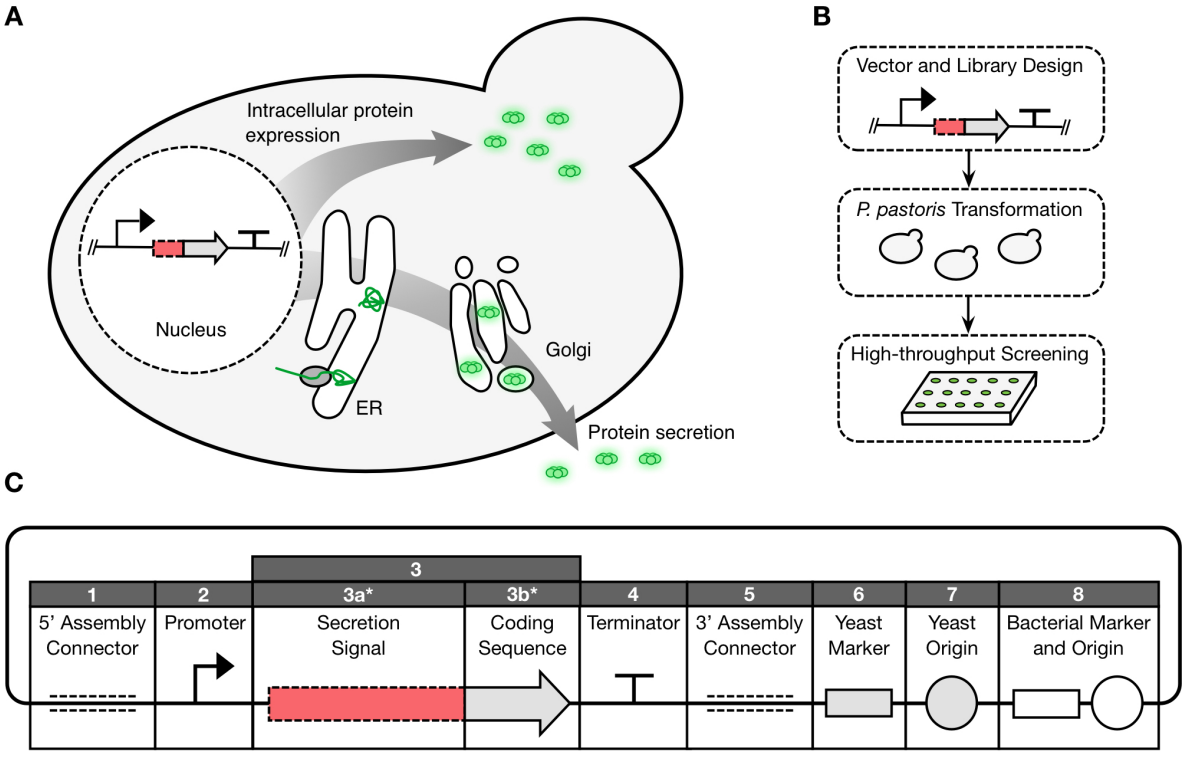

FIGURE 3.2. *P. pastoris* protein secretion platform. **A** Proteins produced in *P. pastoris* may either be expressed intracellularly or secreted extracellularly. The secretion pathway offers the possibility of protein modifications such as proteolytic maturation, glycosylation or disulfide bonds [4]. **B** After *P. pastoris* transformation, high-throughput screening of expression construct libraries can be performed in 96-well plates to find the optimal expression constructs. **C** Standardised and combinatorial assembly strategy for designing expression vectors. Parts are defined by flanking overhangs and are flexible and interchangeable. (*) For intracellular protein expression the coding sequence must be designed as a part with type 3 to ensure the correct overhang to the previous part type 2. For protein secretion, the secretion signal (part type 3a) and the coding sequence (part type 3b) are needed.

3.2 Characterisation of Promoter Strengths

Strong production of a protein requires precise tuning of gene expression. Promoter elements that control the rate of transcription initiation are important to this process. When choosing a suitable promoter, many intrinsic properties are to be considered. These include: whether the promoter is constitutive or inducible, inducer concentrations that are necessary for activation, the promoter's leakiness when inactive and overall promoter strength (e.g., induced expression fold-change) [17]. Furthermore, the gene itself may have toxic effects on the host, making selection of appropriate expression strength critical for successful expression and cell viability. Ten new promoter parts which are compatible with the YTK and are specific for *P. pastoris* were developed [246, 301, 304]. These included three strong and tightly regulated methanol inducible promoters (pAOX1, pDAS1, pPMP20) and seven constitutive promoters (pGAP, pENO1, pTPI1, pPET9, pG1, pG6, pADH2). These were selected to cover a broad range of expression levels (Supplementary Table A.1). In addition, nine existing YTK promoters from *S. cerevisiae* were included to test their function in *P. pastoris* (pTDH3, pCCW12, pPGK1, pHHF2, pTEF1, pTEF2, pHHF1, pHTB2, pRPL18B), (Supplementary Table A.2). Two of these *S. cerevisiae* promoters (pTEF1, pPGK1) have previously been shown to be functional in *P. pastoris*, too [286]. To ensure efficient termination of transcription, which can also affect gene expression levels, a new terminator part taken from the alcohol oxidase gene (tAOX1) of *P. pastoris* was constructed [296].

The strength of promoters in *S. cerevisiae*, in contrast to bacteria, are mostly independent of the downstream coding region [171]. To verify this for *P. pastoris*, each promoter expressing two different fluorescent reporter genes (RFP and yEGFP, intracellular variants) were tested (Figure 3.3). All promoters except pTEF1 and pPGK1 were combined with the terminator tAOX1 and the constructs were genomically integrated. The promoters pTEF1 and pPGK1 were combined with the terminator tPGK1 and being genomically integrated, too. Promoter strength was determined from bulk fluorescence measurements from a plate reader (Materials and Methods 2.2.3). Promoter strengths were found to span over three orders of magnitude and is relatively independent of downstream sequence context, with both RFP and yEGFP genes showing a similar ranking of expression levels, as it was previously shown for *S. cerevisiae* [171]. The inducible promoter pAOX1 results to highest expression levels for RFP and yEGFP. The other inducible promoters (pDAS1, pPMP20) and two constitutive promoters (pPET9, pGAP) show equally strong expression. The *P. pastoris* promoters pADH2, pG6, pG1 and the *S. cerevisiae* promoters pTEF1, pHHF1 and pTEF2 show intermediate expression levels with all promoters resulting in higher RFP expression levels than yEGFP expression levels. The *S. cerevisiae* promoter pPGK1 also results in intermediate expression, but having higher yEGFP than RFP expression levels. Only weak expression results from the other promoters (pTDH3, pRPL18B, pENO1, pTPI, pHTB2, pCCW12).

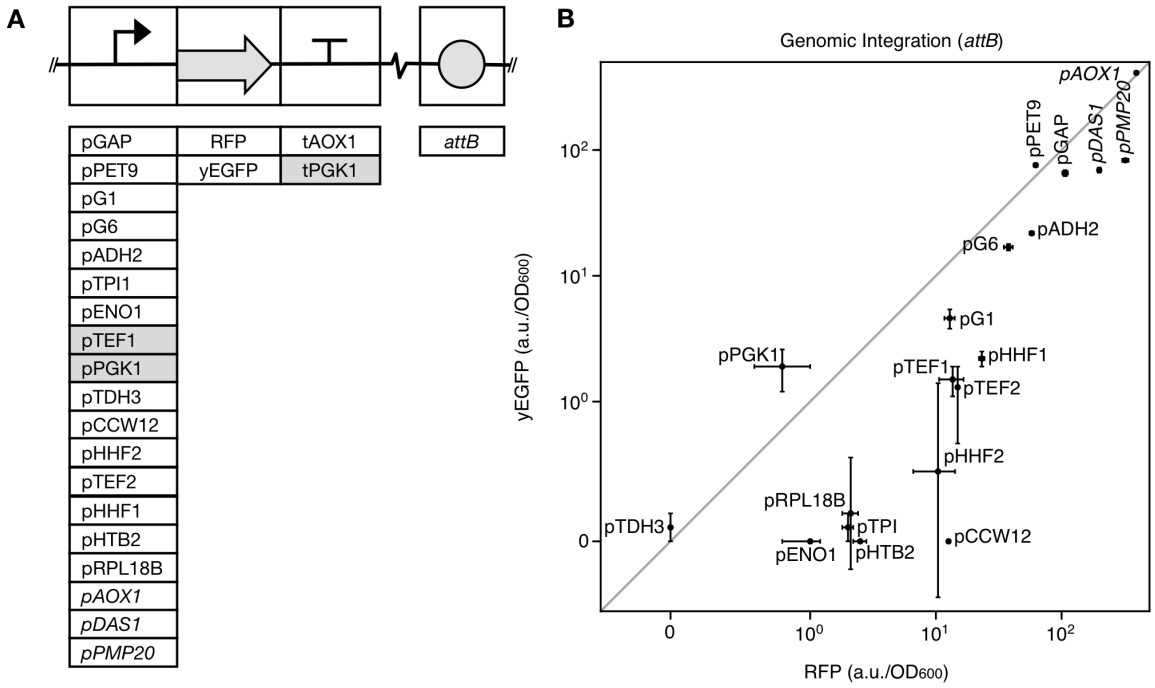

FIGURE 3.3. Characterisation of constitutive promoters. **A** Design of characterisation constructs. All promoters were characterised for RFP and yEGFP expression with the tAOX1 terminator, except the grey shaded parts (pTEF1 and pPGK1) were combined with tPGK1. **B** The relative strength of 19 promoters was evaluated for intracellular RFP and yEGFP expression. Three biological replicates were analysed and the mean fluorescence normalised to the OD_{600} is presented with error bars that denote ± 1 standard deviation. Characterisation was performed in 1 % BMM for pAOX1, pDAS1, pPMP20 and 1 % BMD for all other constitutive promoters.

To vary the overall expression mode, gene expression either from a genomically integrated construct (*attB*) or plasmid-based expression system (PARS) (see Section 3.7) was studied for a selection of promoters. For plasmid-based constructs, promoters also displayed similar rankings of strength for RFP and yEGFP (Figure 3.4 B). However, weaker promoters had a tendency for stronger production of RFP than yEGFP when genomically integrated, and the weak pPGK1, pENO1 and pTPI1 promoters showed higher overall RFP and yEGFP expression when expressed from a plasmid.

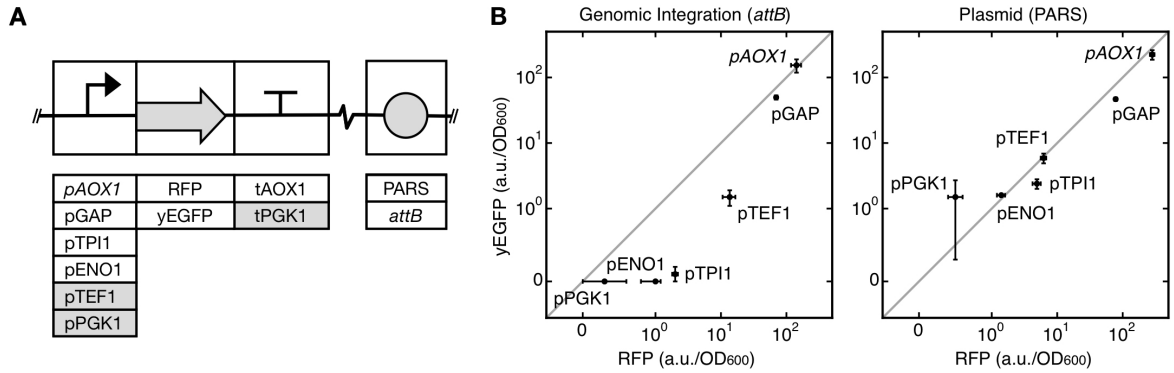

FIGURE 3.4. Characterisation of constitutive promoters. **A** Design of characterisation construct. For 6 promoters the expression strength of the construct being genomically integrated (*attB*) was compared to the expression level from the plasmid-based expression (PARS). The promoters pTEF1 and pPGK1 were coupled with tPGK1, all others with tAOX1. **B** The relative strength was evaluated for intracellular RFP and yEGFP expression. Three biological replicates were analysed and the mean fluorescence normalised to the OD_{600} is presented with error bars that denote ± 1 standard deviation. Characterisation was performed in 1 % BMM for pAOX1 and 1 % BMD for all other constitutive promoters.

The bulk fluorescence measurements using the plate reader are a simple and straight forward approach to study expression levels. However, to compare gene expression at a cellular level, a closer look with techniques such as flow cytometry or fluorescence microscopy is necessary and a selection of strains were studied with a fluorescence microscope. Samples were expressed in a deep-well plate format in 1 % BMD for all pGAP strains and the pAOX1 strains studying leaky expression. For induced pAOX1 expression, strains were also cultivated in 1 % BMM. To ensure comparability of the images, excitation conditions were kept constant for brightfield (15000 ms, 3.5 V), RFP (Filter set 43 HE, 5000 ms, 3.5 V) and yEGFP (Filter set 38 HE, 500 ms, 3.5 V) imaging. Figure 3.6 and 3.5 compare fluorescence levels across the cell population from genomically integrated constructs to plasmid-based expression. For pGAP as well as pAOX1 driven intracellular RFP and yEGFP expression, cells with the genomically integrated construct have similar expression strength, whereas cells with plasmid-based expression show a more diverse fluorescence output across the cell population. In the study of Camattari *et al.* [44], the copy number of expression plasmids carrying the PARS sequence was determined and varied from 6-18 copies per cell and consequently resulted in varying expression levels. The aim of the PTK is to obtain a shuffled screening of expression constructs to find the optimal expression condition. Screenings at the cellular level are aimed, to increase the throughput.

Therefore, the application of plasmid-based expression is not favoured, as expression variation from cell to cell is already substantial. Expression from the genomically integrated construct offers more homogenous expression levels and only variations which stem from differing transcriptional units are of interest.

The leakiness of the pAOX1 promoter was studied in Figure 3.5. The pAOX1 promoter is highly inducible in the presence of methanol and tightly repressed in the presence of other carbon sources such as glucose, glycerol, or ethanol. The microscopical evaluation does show some basal expression for RFP and yEGFP, but only in the smaller cells.

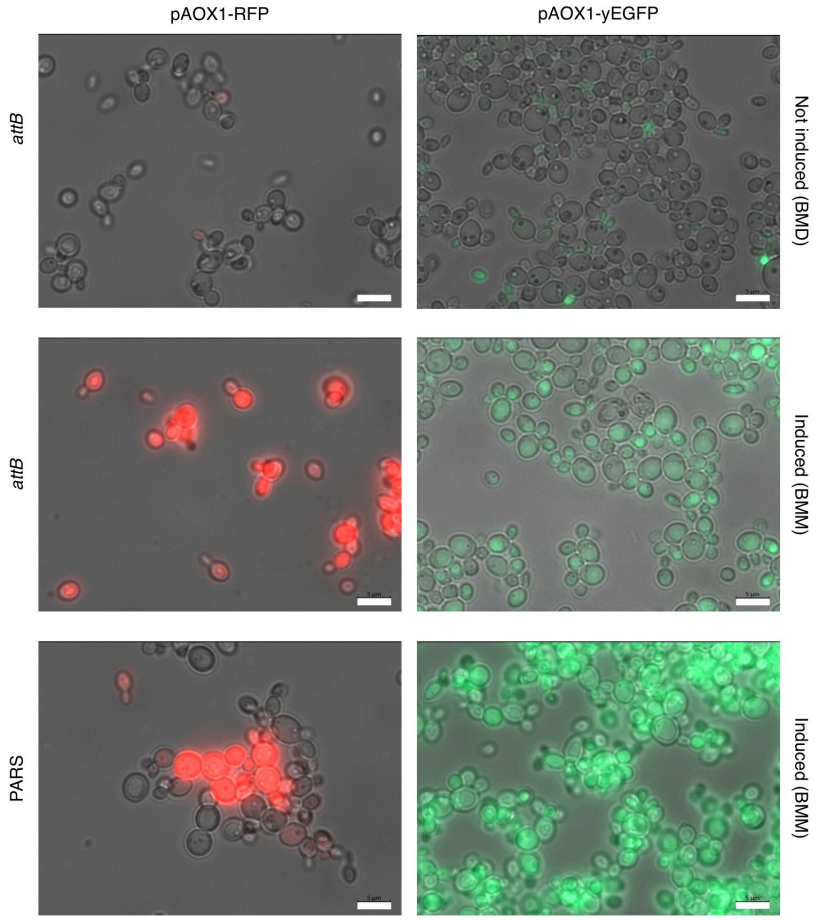

FIGURE 3.5. Microscopical comparison of pAOX1 driven RFP and yEGFP expression. Leaky expression in 1 % BMD and methanol induced expression in 1 % BMM for the expression construct being genomically integrated was studied. Induced RFP and yEGFP expression is also studied for plasmid-based expression (PARS). The scale bar denotes 5 µm.

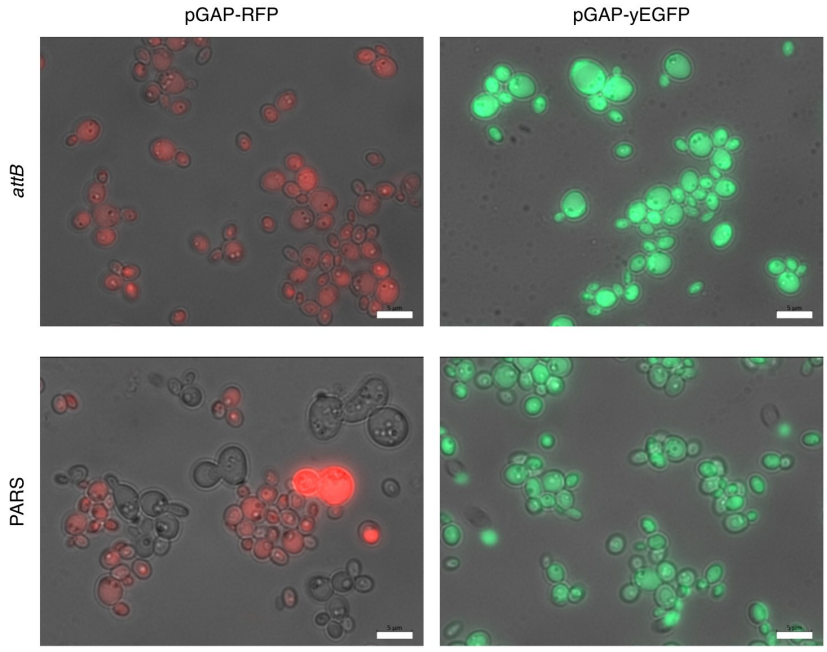

FIGURE 3.6. Microscopical comparison of pGAP driven RFP and yEGFP expression. Constitutive expression of RFP or yEGFP either from the plasmid (PARS) or being genomically integrated (*attB*). The scale bar denotes 5 µm.

For the methylotrophic yeast *P. pastoris*, methanol-inducible promoters are typically used to control protein expression and pAOX1 is one of the most commonly used inducible promoter systems [301]. It is strongly activated in the presence of methanol and repressed in the presence of glucose, glycerol or ethanol [96]. Additionally, the strong and tightly regulated promoters pDAS1 and pPMP20 were characterised for varying methanol concentrations and expression times to optimise protein expression (Figure 3.7). pAOX1 was characterised for the construct being genomically integrated as well as plasmid-based expression, pDAS1 and pPMP20 were solely characterised for being genomically integrated. For the analysis, methanol was the sole carbon source and its concentration was found to strongly influence cell growth. High methanol concentrations (> 4 % (v/v)) were found to be toxic, leading to decreased cell growth (Figure 3.7). The highest OD_{600} was reached for all strains after 72 hours and at a 2 % (v/v) methanol concentration for all pAOX1 constructs and at a 4 % (v/v) methanol concentration for pDAS1 and pPMP20. Highest fluorescence levels were reached at 2 % (v/v) methanol and 72 hours expression time for the genomically integrated RFP and yEGFP, and the plasmid-based yEGFP. RFP expression from plasmids was highest at 1 % (v/v) methanol at 72 hours. For pAOX1 the genomic integration was found to generally lead to higher expression levels than for plasmid-based expression. For pDAS1 and pPMP2 highest RFP and yEGFP expression was found at 4 % (v/v) methanol after 72 hours expression time.

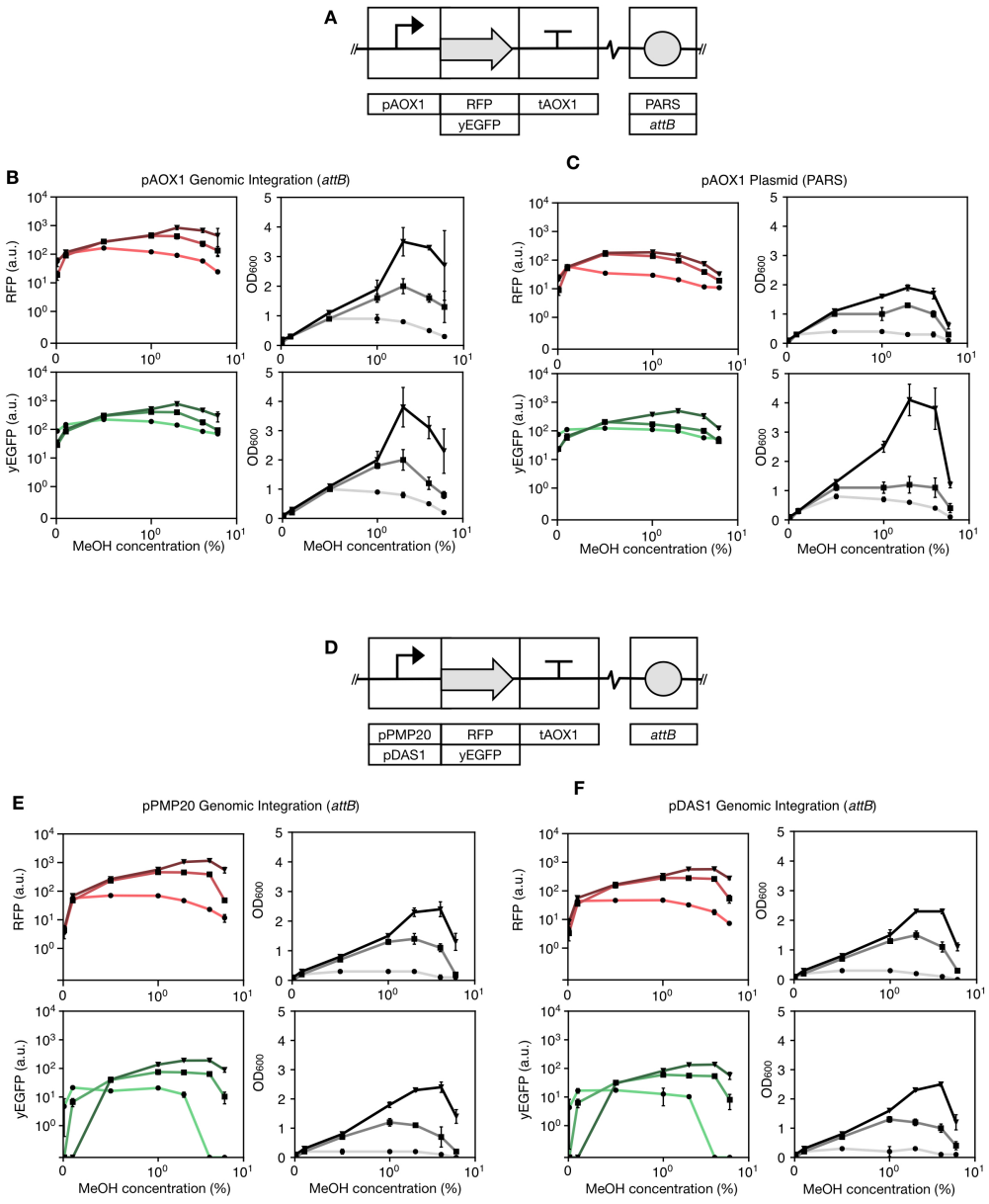

FIGURE 3.7. Characterisation of methanol-inducible promoters. **A, D** Design of pAOX1, pPMP20, pDAS1 promoter characterisation constructs. **B, C, E, F** The relative strength of the promoters under varying methanol concentrations was evaluated. RFP and yEGFP were expressed from the genomically integrated constructs and for pAOX1 also from the plasmid-based construct. Sampling was performed at three different time points: 24 hours (circle, light red/green/grey), 48 hours (square, medium red/green/grey) and 72 hours (triangle, dark red/green/grey). Total fluorescence in arbitrary units as well as the corresponding OD_{600} is shown. Three biological replicates were analysed with the mean shown and error bars denoting ± 1 standard deviation.

3.3 Characterisation of Terminator Strengths

To control the protein expression level of heterologous expressed proteins, the promoter driving gene expression is generally altered. However, the terminator sequence has been shown to play an important role in protein expression by influencing mRNA cleavage, mRNA stability, and mature RNA levels in *S. cerevisiae* [68, 212]. Terminator characterisations in *S. cerevisiae* have shown the importance of terminator choice [68, 314], with up to 35-fold varying expression strength caused by the terminator. The effect of the terminator for *P. pastoris* expression was previously studied for the eGFP expression under the control of pAOX1 for 20 terminators [304]. Vogl *et al.* [304] describe only minor differences in expression between the studied terminators. The study describes 57-89 % expression level variation for the *P. pastoris* terminators and 62-78 % expression level variation of the *S. cerevisiae* terminators in comparison to the tAOX1 terminator.

Here, intracellular expression levels of RFP were measured for eight *P. pastoris* promoters coupled with six different *S. cerevisiae* terminators of the YTK and the *P. pastoris* terminator tAOX1. Fluorescence measurements were performed as described for the promoters (Section 2.4.3). The terminators studied do alter the expression level considerably more than previously described [304]. For each promoter the expression strength is tuned additionally by the terminator with 5-fold differences in expression for all promoter-terminator combinations except for pAOX1 where differences are up to 10-fold high (Table 3.1). The choice of the terminator can also be ordered according to its influence on expression modulation and this ranking is valid for almost all promoters, with few exceptions (pGAP-tPGK1, pPET9-tENO2, pADH2-tTDH1) (Figure 3.8). Here, tAOX1 leads to higher expression levels while tSSA1 leads to lower expression levels.

TABLE 3.1. Characterisation of promoter-terminator combinations. Values given in a.u. (mean fluorescence normalised to the OD_{600})

	pPMP20	**pAOX1**	**pDAS1**	**pGAP**	**pPET9**	**pADH2**	**pG6**	**pG1**
tAOX1	320.3	391.5	197.7	106.7	62.2	57.8	37.8	12.9
tTDH1	302.1	101.2	98.4	46.6	30.8	90.0	16.4	9.7
tADH1	133.9	77.3	76.2	41.2	30.1	24.0	16.9	5.6
tPGK1	147.6	72.5	91.8	112.7	30.8	18.9	12.6	4.8
tENO2	124.8	66.6	78.7	47.0	88.4	16.9	12.6	3.7
tENO1	108.8	58.7	72.7	38.7	25.1	13.6	11.8	4.4
tSSA1	62.1	33.4	37.6	22.2	15.7	9.1	6.0	2.5

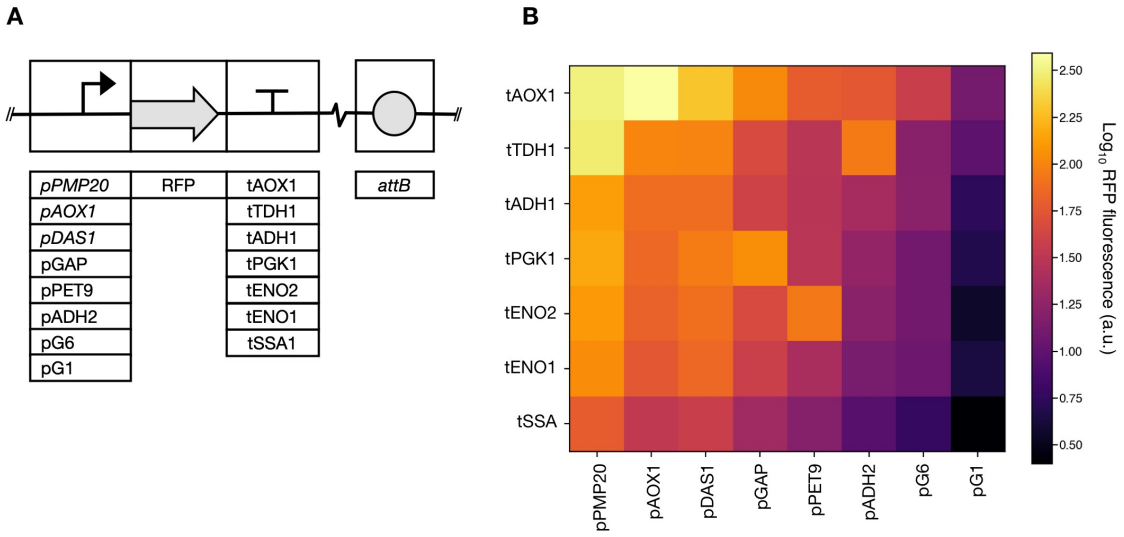

FIGURE 3.8. Characterisation of promoter-terminator combinations. **A** Design of the terminator characterisation constructs. All transcriptional units were expressed from genomically integrated constructs. **B** Heat map of RFP fluorescence values given in a.u. (fluorescence normalised to the OD_{600}) in a logarithmic scale.

3.4 Design and Characterisation of Synthetic Secretion Signal Peptides

The overproduction of recombinant proteins can be challenging if protein folding in the ER and secretion capacity are overloaded. This can lead to the accumulation of misfolded or unfolded protein and consequently to cellular stress and low protein production yields [81]. To overcome secretory bottlenecks and enhance secretion strength, optimisation of the secretion signal or overexpression of translocon components and stabilising cytosolic chaperones can be performed [80]. The only requirement for a protein to enter the secretion pathway is a secretion signal peptide to be present at the N-terminus of the nascent polypeptide [81, 178]. The secretion efficiency of a recombinant protein fused to a given signal peptide can differ strongly even for similar genes with only small differences in protein sequence, structure or gene expression regulation [81]. The PTK allows for the rapid testing of multiple combinations of promoters, secretion tags or any other functionality to find optimal secretion conditions.

The recognition of the secretion tag has a high degree of flexibility. A study where random sequence modifications were performed on the invertase signal sequence showed that 20 % of the signal peptides were still recognised [146].

Even though signal peptides appear to have a simple structure with interchangeable domains and a low information content, even minor variations can affect protein targeting, translocation, signal sequence cleavage, and further postcleavage processes [127]. The most commonly used signal peptide is the *S. cerevisiae* α-mating factor (αMF) (Figure 3.9). To enhance efficiency of the secretion tag, various studies have been performed. Random mutagenesis to substitute a single amino acid did not make a significant difference [183]. Site-directed mutagenesis and the generation of deletion mutants gave higher activities of the horseradish peroxidase (HRP) reporter protein [183]. Also, alternative signal sequences, including invertase, inulinase or glucoamylase signal peptides, are available (Figure 3.9; Table A.1) [4].

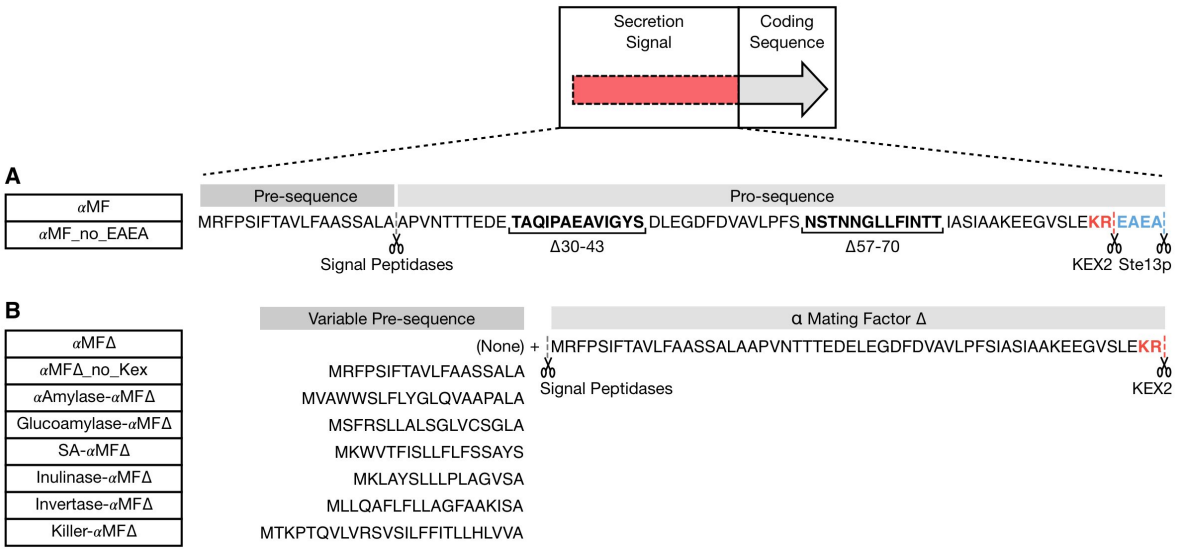

FIGURE 3.9. Design of synthetic secretion signal peptides. **A** Overall structure of the α-mating factor secretion signal. The secretion signal is composed of a pre- and pro-region. The pre-region interacts with the signal recognition particle to guide the protein into the ER. It consists of a basic amino terminus, a hydrophobic core as well as a polar region at the carboxy terminus. The pre-region is cleaved off by signal peptidases. The pro-region slows down transport to ensure proper protein folding and is thought to be important for the transport from the ER to the Golgi. It consists of hydrophobic amino acids, which are interrupted by charged or polar amino acids. The pro-region is cleaved at the dibasic KR site by the endoprotease Kex2p and the two EA repeats are removed by the Ste13p dipeptidy amino peptidase [4, 106, 183]. The Ste13p cleavage is often problematic and results in a non-native amino acid sequence at the N-terminus of the heterologous protein [74]. The αMF_no_EAEA sequence is the same as the αMF sequence except that the EAEA amino acids at the C-terminus are removed. **B** Design of new secretion tags. The newly designed tags were composed of a varying amino terminus followed by a deletion mutant of the α-mating factor [183].

Here, synthetic secretion tags were designed by combining established leader sequences (N-terminal) with an optimised α-mating factor secretion tag (C-terminal) (Figure 3.9). Modulation in the N-terminal region was expected to influence the signal recognition. As the Ste13p cleavage is often problematic, the C-terminus ends at the Kex2p recognition site [74]. Additionally, to prevent the tag from being too long, a double deletion variant of the α-mating factor that has been previously shown to have higher cleavage activity in comparison to the full length α-mating factor secretion peptide was used [183]. Additionally, various modified signal peptides of the α-mating factor were studied.

To test the suitability of this library and the feasibility of the combinatorial approach in identifying optimal combinations of elements for protein secretion, a library of 100 secretion constructs was designed. By applying the hierarchical assembly method with the recombinase-based system for integration into a defined *P. pastoris* locus, this library could be made in a matter of weeks. If more common *P. pastoris* cloning methods were used [67, 215], the same library would have taken months to complete. RFP and yEGFP were applied as reporters to allow for accurate, but indirect measurement of protein levels being expressed intracellularly or secreted (Figure 3.10 A). For each reporter, 10 different signal peptides, each under the control of 5 different promoters were analysed. Characterisation was performed in 1 % BMM for all pAOX1 constructs and in 1 % BMD for all constitutive constructs. Results of overall expression and secretion efficiency are shown in Figure 3.10 C. First, bulk measurements of the total fluorescence from individual wells were taken. Next, the cells were spun down and pelleted and the supernatant was re-measured to provide the secreted level. The secretion efficiency S_e was then calculated as $S_e = S_s/S_t$, where S_t is the total fluorescence measurement of the cells in culture in arbitrary units (a.u.), and S_s is the fluorescence measurement from the supernatant in a.u. The value of S_e ranges from 0, when no protein is secreted, to 1, when all protein is secreted (Table 3.2).

TABLE 3.2. Secretion efficiency for combinations of tags, promoters and genes.

	RFP					**yEGFP**				
Secretion tag	pAOX1	pGAP	TEF1	pENO1	pPGK1	pAOX1	pGAP	pTEF1	pENO1	pPGK1
None	0.01	0.02	0.00	0.00	0.08	0.06	0.00	0.00	0.00	0.00
αMFΔ	0.04	0.00	0.38	0.52	0.33	0.19	0.02	0.00	0.00	0.00
Invertase-αMFΔ	0.02	0.02	0.00	0.00	0.04	0.12	0.00	0.00	0.00	0.00
Killer-αMFΔ	0.02	0.01	0.04	0.00	0.00	0.13	0.00	0.00	0.00	0.00
Glucoamylase-αMFΔ	0.02	0.01	0.00	0.00	0.01	0.12	0.00	0.00	0.00	0.00
αMF	0.87	0.00	0.88	0.72	0.60	0.75	0.78	0.00	0.00	0.00
αMFΔ_no_Kex	0.01	0.00	0.00	0.00	0.00	0.11	0.02	0.00	0.00	0.00
αMF_no_EAEA	0.92	0.78	0.78	0.00	0.81	0.77	0.67	0.00	0.00	0.00
SA-αMFΔ	0.38	0.37	0.13	0.00	0.81	0.59	0.00	0.00	0.00	0.00
Inulinase-αMFΔ	0.25	0.00	0.13	0.00	0.00	0.53	0.00	0.00	0.00	0.00
αAmylase-αMFΔ	0.00	0.02	0.00	0.00	0.00	0.71	0.44	0.44	0.79	0.47

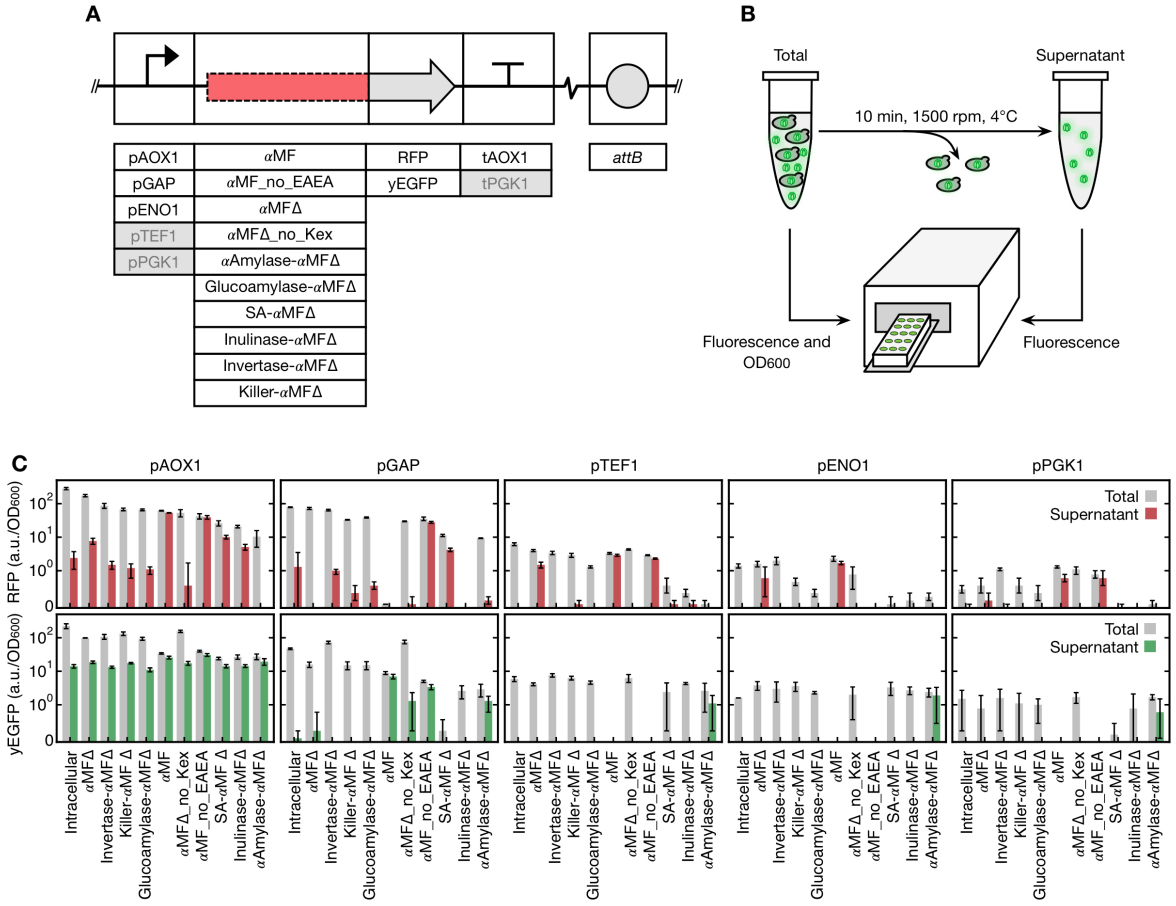

FIGURE 3.10. Characterisation of synthetic secretion tags. **A** Vector design for secretion analysis. Ten different secretion tags, as well as the intracellular variant, were analysed for fluorescent reporter expression under the control of 5 different promoters. White boxes correspond to new elements developed in this study; grey shaded boxes are taken from existing YTK. **B** For all samples, the total fluorescence as well as the fluorescence of the supernatant after centrifugation was determined. The characterisation was performed in 1 % BMM for all pAOX1 constructs and in 1 % BMD for all constitutive constructs. **C** Bars represent the mean fluorescence of three biological replicates normalised to OD_{600} (a.u./OD_{600}). Error bars denote ± 1 standard deviation.

Even though RFP and yEGFP are both beta-barrel proteins [337], their secretion was found to differ when expressed from identical regulatory elements. Furthermore, high overall protein expression did not necessarily result in high protein secretion. The most commonly used secretion tag is the αMF and its derivative, αMF_no_EAEA, in which the EA repeat is removed. Using these secretion tags generally led to the largest amount of secretion, especially for RFP, while for yEGFP secretion exhibited a rather reduced dependence on the tag used. Remarkably, using these secretion tags was either highly efficient, with almost all of the expressed protein being secreted, or there was no detectable protein expression. The αMF and αMF_no_EAEA secretion tags guided RFP secretion very efficiently for all promoters (S_e = 0.6-0.92) except for the combinations, pGAP-αMF and pENO-αMF_no_EAEA, where no RFP was detectable. For yEGFP, these tags strongly drove secretion under the regulation of pAOX1 and pGAP (S_e = 0.67-0.78) but for all weak promoters, no yEGFP was detectable.

The newly designed synthetic hybrid tags were less efficient, but achieved significant protein secretion, thus demonstrating that there is flexibility in the secretion tag design. In contrast to the most commonly used αMF and αMF_no_EAEA tags, a majority of constructs showed protein expression without successful secretion. However, the synthetic SA-αMFΔ tag secreted RFP well when this cassette was under the control of a strong promoter. For RFP secretion under the control of a weak promoter, there were only three tags leading to secretion: αMF, αMF_no_EAEA, and αMFΔ. For yEGFP expression, the synthetic αAmylase-αMFΔ tag was the only tag resulting in secretion in the presence of weak promoters. When using a strong promoter with the αAmylase-αMFΔ tag, secretion was not achieved for RFP but was seen for yEGFP. Thus, in contrast to promoter strength, the efficiency of these secretion tags was found to be unpredictable and varied based on the promoter and the downstream coding region used. It is unclear if these contextual effects arise due to interactions at DNA, RNA or, protein levels with components in the secretion process or between the gene and secretion tag used. However, the findings do highlight the importance of testing multiple tags when designing new secretion constructs.

3.5 Characterisation of Predicted Secretion Signal Peptides

The design and characterisation of the newly designed secretion signal peptides showed strong differences in efficiency for RFP and yEGFP. To find optimal secretion conditions for any protein, a wider selection of signal peptides should be available. To extend the toolkit, 9 endogenous *P. pastoris* (Table 3.3), 8 exogenous and 2 synthetic signal peptides (Table 3.4) were added and characterised for the secretion of RFP and yEGFP under the control of the strong constitutive promoter pGAP. The selection of these tags was based on predictions of potential *P. pastoris* signal peptides, previously established signal peptides or natural signal peptides of the proteins applied for the randomised library approach (Chapter 4).

TABLE 3.3. Predicted endogenous signal peptides for protein secretion.

Name	**Description of the coding region**	**Locus in *P. pastoris* genome**	**Amino acid sequence** (N-region **H-region** C-region)	**Reference**
SP_Cyclophilin	Peptidyl-prolyl cis-trans isomerase (cyclophilin) of the endoplasmic reticulum	PAS_chr1-1_0267	MKLLN**FLLSFVTLFGLL**SGSVFA	[200]
SP_Disulfide isomerase	Protein disulfide isomerase, multifunctional protein resident in the endoplasmic reticulum lumen	PAS_chr1-1_0160	MK**ILSALLLLF**TLAFA	[200]
SP_Peptidylprolyl isomerase	Hypothetical protein (Peptidylprolyl isomerase)	PAS_c131_0001	MKVSTTK**FLAVFLLV**RLVCA	[200]
SP_Cell wall protein	Cell wall protein that functions in the transfer of chitin to beta(1-6)glucan	PAS_chr1-1_0293	MRP**VLSLLLLLA**SSVLA	[200]
SP_C4R6P1	Hypothetical protein (C4R6P1)	PAS_chr4_0040	MW**SLFISGLLIFY**PLVLG	[200]
SP_C4R8H7	Hypothetical protein (C4R8H7)	PAS_chr4_0643	MS**TLTLLAVLLSLQ**NSALA	[200]
SP_Mucin	Mucin family member	PAS_chr1-3_0276	MINLNS**FLILTVTLL**SPALA LPKN-VLEEQQAKDDLAKR	[200]
SP_PHO1	Acid phosphatase PHO1	PPU28658	MFSPI**LSLEIILALATL**QSVFA	[128]
SP_Scw11p	Cell wall protein with similarity to glucanases	PAS_chr2-1_0052	MLS**TILNIFILLLFI**QASLQ	[178, 200]

TABLE 3.4. Predicted exogenous signal peptides for protein secretion.

Name	**Description of the coding region**	**Locus**	**Amino acid sequence** (N-region **H-region** C-region)	**Reference**
*α*MF_no_EAEA	Alpha mating factor no EAEA	*S. cerevisiae*, Synthetic	MR**FPSIFTAVLFAASS**ALAAPVNTTTEDETAQIPAEAV IGYSDLEGDFDVAVLPFSNSTNNGLLFINTTIASIAA-KEEGVSLEKR	[183]
*α*MF	Alpha mating factor	*S. cerevisiae* MFAL1_YEAST	MR**FPSIFTAVLFAASS**ALAAPVNTTTEDETAQIPAEAV IGY SDLEGDFDVAVLPFSNSTNNGLLFINTTIASIAA-KEEGVSLEKREAEA	[183]
SP_An_phyA	3-phytase A	*A. niger* PHYA_ASPNG	MG**VSAVLLPLYLL**SGVTSG(LAVP)	NYC
SP_Pl_phyA	Phytase	*P.lycii* Q96VH9_9HOMO	MVSSAFAPS**ILLSLMSSLALSTQFSFVAA**	NYC
SP_Th_phyA	Histidine acid phosphatase phytase	*T.heterothallica* V5M269_THIHE	MTG**LGVMVVMVGFLAI**ASLQS	NYC
SP_Aae_UPO	Aromatic peroxygenase	*A. aegerita* APO1_AGRAE	MKY**FPLFPTLVFAA**RVVAFPAYASLAGLSQQELDAII PTLEAR	[208]
SP_Aae_UPOeng	Aromatic peroxygenase tag engineered	*A. aegerita*, synthetic	MKY**FPLFPTLVYAV**GVVAFPDYASLAGLSQQELDAII PTLEAR	[208]
SP_SUC2	Invertase 2	*S. cerevisiae* INV2_YEAST	ML**LQAFLFLLAGFA**AKISA	[29, 167, 206, 228]
SP_ALB	Serum albumin	*H. sapiens* ALBU_HUMAN	MK**WVTFISLLFLF**SSAYS	[162, 326]
SP_CSN2	Beta-casein	*B. taurus* CASB_BOVIN	MK**VLILACLV**ALALA	[126]
SP_PHA-E	Phytohaemag-glutinin	*P. vulgaris* I0J8I4_PHAVU	MASSN**LLSLALFLVLLTHANS**	[254]
MF41	Synthetic signal peptide MF41	Synthetic	MAIPR**FPSIFIAVLFAASSALA**APVNTTTEDETAQIPAE AVIGYSDLEGDFDVAVLPFSNSTNNGLLEEAEAEAEP-KFINTTIASIAAKEEGVSLE	[325]

The prediction of signal peptides is complex due to the variation in amino acid sequence and the sequence length itself. Nevertheless, computational programs are available, looking for conserved regions within the amino acid sequence. Massahi *et al.* performed an *in-silico* analysis of 41 potential signal peptides [200] based on a proteomic analysis of the *P. pastoris* secretome [135]. In their study, five computer programs (SignalP4.1, Phobius, WolfPsort0.2, ProP1.0, NetNGlyc1.0) were used to predict the suitability of proteins from the secretome to be potential signal peptides [200]. One widely used program is SignalP4.1, which looks at the possible cleavage site and the discrimination of different segments to predict signal peptides. The likelihood for secretion is given as D-score (discrimination score), and values of D-score >0.7 implies that a peptide is very likely a secretory signal peptide [200]. To confirm their prediction, Massahi *et al.* showed the functionality of five of the predicted signal peptides for the secretion of recombinant human growth hormone (rhGH), the following year [199]. Here, these five predicted and confirmed signal peptides as well as three additionally predicted tags were designed and studied.

Another source of signal peptides can be the variety of signal peptides of proteins being secreted in their native organism. Heterologous protein expression has shown the suitability of exogenous tags with either the corresponding protein [162, 208] or another heterologous protein [326]. A selection of seven previously described exogenous signal peptides was studied. Additionally, three signal peptides from differing fungal hosts guiding the secretion of a phytase and a native as well as an engineered signal peptide of *Agrocybe aegerita* guiding the secretion of an Unspecific Peroxygenase (UPO) were included (for more detailed information regarding the phytase and UPO selection, see Chapter 4).

All signal peptides were characterised for the secretion of RFP and yEGFP, genomically integrated, and regulated by the strong constitutive pGAP promoter. The overall fluorescence, as well as fluorescence in the supernatant were measured using the plate reader. The fluorescence normalised to the OD_{600} is displayed in Figure 3.11, the secretion efficiency S_e ($S_e = S_s/S_t$) is given in Table 3.5. The D-score was determined suggesting the efficiency of all tags to be probable signal peptides (D-score > 0.7) except for SP_Aae_UPO and SP_Aae_UPOeng, which have a D-score of 0.56.

The efficiency of the tags varies for the secretion of RFP and yEGFP, as shown for the synthetic tags. Four RFP constructs with the signal peptides SP_An_phyA, SP_SUC2, SP_Aae_UPO, and SP_Pl_phyA express more RFP overall than the expression construct without a signal peptide. However, only two (SP_SUC2, SP_Aae_UPO) result to protein secretion with secretion levels higher than the commonly used αMF_no_EAEA tag. For yEGFP, none of the tested constructs expresses as much as the intracellular expression construct without a tag. The comparison of the secreted amount of yEGFP shows more protein for the tags SP_Scw11p, SP_ALB, SP_Cyclophilin, SP_Disulfide isomerase, SP_Peptidylprolyl isomerase, SP_Cell wall protein and SP_Aae_UPOeng in comparison to the widely applied αMF_no_EAEA or αMF.

Comparing the overall efficiency of the tags, there is no preference for either endogenous *P. pastoris* signal peptides or exogenous peptides. For the six newly characterised signal peptides that have not yet been studied in *P. pastoris* (grey boxed in Figure 3.11 A) two (SP_Cyclophilin and SP_Th_phyA) guide good secretion, two (SP_An_phyA, SP_Pl_phyA) result in very high overall protein amounts but no secretion, and two (SP_C4R8H7, SP_Mucin) guide only low amounts of RFP and yEGPF outside the cell. The tags SP_Aae_UPO and SP_Aae_UPOeng with a low D-score actually show high protein secretion, in contrast SP_PHO1 and SP_PHA-E had promising D-sores (> 0.79), but do not efficiently guide secretion.

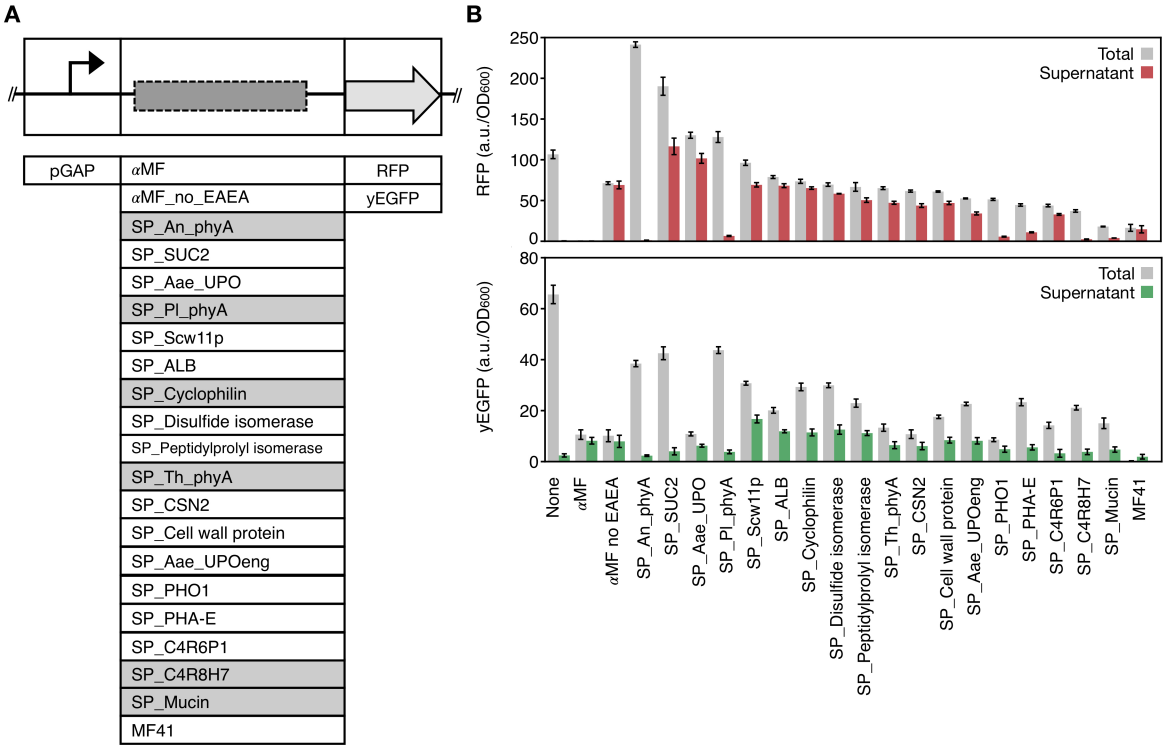

FIGURE 3.11. **A** Vector design for secretion analysis. 21 secretion tags under the control of pGAP were analysed for RFP and yEGFP expression and secretion. Grey boxed signal peptides highlight peptides which to our knowledge have not been tested before. **B** For all samples, the total fluorescence as well as the fluorescence of the supernatant after centrifugation was determined. The characterisation was performed in 1% BMD. Bars represent the mean fluorescence of three biological replicates normalised to OD_{600} (a.u./OD_{600}). Error bars denote ± 1 standard deviation.

TABLE 3.5. Predicted signal peptides for protein secretion.

Name	**D-score RFP prediction P4.1**	**D-score yEGFP prediction P4.1**	**S_e RFP and pGAP**	**S_e RFP and pGAP**
None	-	-	0.00	0.04
SP_Cyclophilin	0.87	0.893	0.89	0.39
SP_Disulfide isomerase	0.862	0.908	0.84	0.42
SP_Peptidylprolyl isomerase	0.858	0.893	0.76	0.49
SP_Cell wall protein	0.867	0.904	0.77	0.48
SP_C4R6P1	0.904	0.912	0.75	0.23
SP_C4R8H7	0.877	0.895	0.07	0.18
SP_Mucin	0.866	0.868	0.22	0.31
SP_PHO1	0.795	0.829	0.11	0.57
SP_Scw11p	0.797	0.854	0.72	0.54
αMF_no_EAEA	0.884	0.884	0.97	0.78
αMF	0.884	0.884	1.84	0.77
SP_An_phyA	0.815	0.807	0.01	0.06
SP_Pl_phyA	0.739	0.743	0.05	0.09
SP_Th_phyA	0.746	0.653	0.72	0.48
SP_Aae_UPO	0.56	0.568	0.78	0.58
SP_Aae_UPOeng	0.56	0.568	0.65	0.36
SP_SUC2	0.79	0.837	0.61	0.10
SP_ALB	0.817	0.868	0.87	0.59
SP_CSN2	0.838	0.899	0.71	0.57
SP_PHA-E	0.837	0.894	0.25	0.24
MF41	0.9	0.9	0.89	16.33

The strains pGAP–SP_SUC2–RFP and pGAP–SP_Scw11p–yEGFP which showed highest secretion fluorescence signals, and the strain pGAP–SP_An_phyA–RFP which showed high intracellular RFP expression were studied using fluorescence microscopy in comparison to the wild type (Figure 3.12). *P. pastoris* pGAP–SP_SUC2–RFP and pGAP–SP_Scw11p–yEGFP show as expected only moderate fluorescence within the cell. In contrast, the pGAP–SP_An_phyA–RFP strain shows very high RFP expression, as expected from the plate reader screening. The RFP distribution within the cell is not as homogenous across the cell as for intracellular expression, but denser at distinct locations, suggesting that the RFP enters the secretion pathway (e.g., endoplasmatic reticulum and Golgi apparatus), but gets stuck in it and is not released out of the cell.

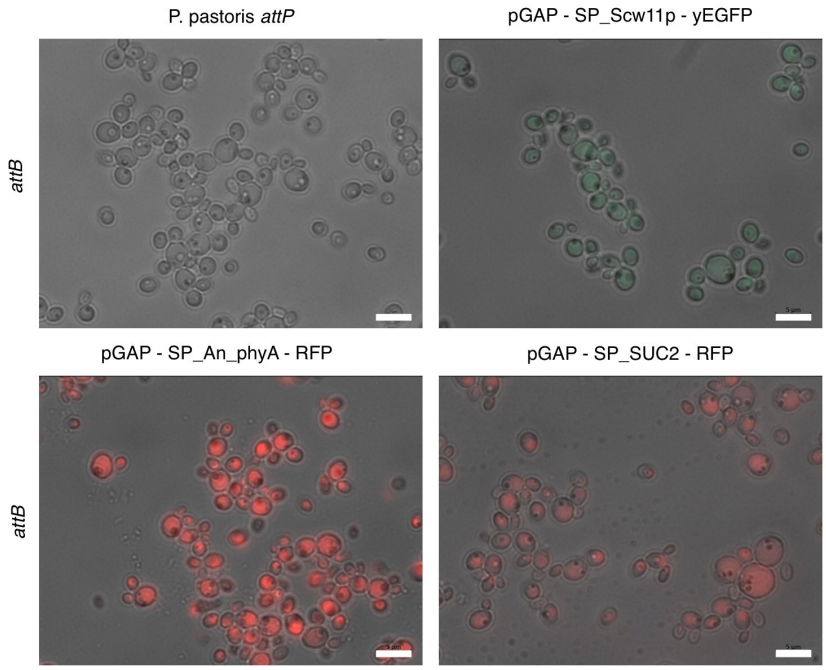

FIGURE 3.12. Microscopical study of the wild type and secretion strains. The expression constructs are all being genomically integrated and BMD was used for expression. The scale bar denotes 5 µm.

3.6 Effect of Methanol on Protein Secretion

An interesting result from the characterisation of the secretion tags was that while the two strongest promoters, pAOX1 and pGAP, were found to produce similar total amounts of protein, pAOX1 displayed significantly higher levels of secretion. This was puzzling because the protein itself should have been constant, and thus, a similar secretion was expected. A potential cause for this difference could be the use of BMM medium for the inducible pAOX1 promoter, which could affect normal cell physiology, versus the use of BMD medium for pGAP. In comparison to glucose utilization, growth on methanol has high oxygen demands that can be critical for high cell density cultivation [143, 153]. To understand whether higher fluorescent titres arose from the secretion construct itself or due to differences in medium composition, we further analysed all pGAP expression constructs in 1 % BMM medium (Figure 3.13). Overall cell growth was significantly higher in BMD (OD_{600} of 4.6-6.1) than in BMM (OD_{600} of 1.7-2.3), which led to overall higher fluorescence. However, fluorescence normalised to the OD_{600} was higher in BMM. Importantly, the fraction of secreted protein, when compared to the overall expressed protein, was much higher for samples grown in BMM.

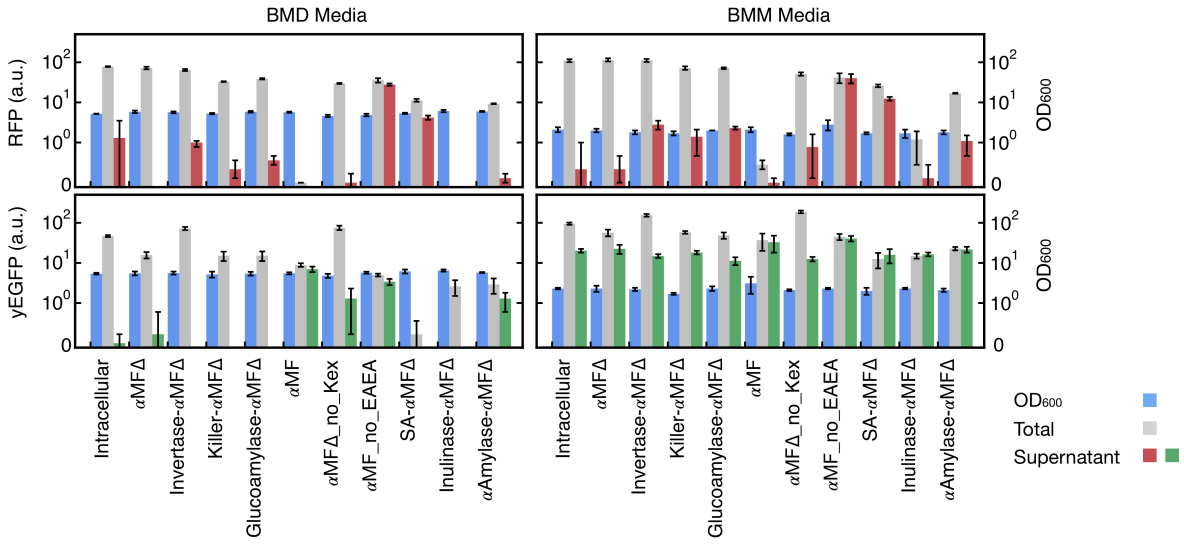

FIGURE 3.13. Comparison of pGAP-guided secretion in 1 % BMM and 1 % BMD medium. Total OD_{600} and fluorescence as well as fluorescence of the supernatant after centrifugation was determined from three biological replicates. The mean of these experiments is shown with error bars that denote ± 1 standard deviation.

The choice of carbon source has a strong impact on the metabolism of recombinant strains as it can result in different biomass yields and product concentrations. Prielhofer *et al.* [245] analysed the gene-specific response of *P. pastoris* to varying carbon sources and showed an upregulation of genes involved in protein production for cells grown on methanol compared to glycerol. Their study revealed that the power of methanol induction is not only due to the strong pAOX1 promoter itself, but it is also linked to differential growth effects of methanol. To achieve strong secretion, the central energy metabolism, the additional burden of heterologous protein production, and the capacity of the secretory machinery must be balanced [188, 190]. Therefore, the characterisation of genetic constructs must account for cultivation conditions in addition to regulatory elements. Growth on glucose allowed for higher specific growth rates and higher overall product titres while growth on methanol enabled fine tuning and higher secretory capacity [96].

P. pastoris is a widely-used host for the secretion of heterologous proteins because of few contaminants of the host cell protein. Furthermore, the majority of *P. pastoris* expression processes utilise methanol as a substrate and inducer as it offers tight gene regulation and high product titres. However, methanol is not always favoured for large-scale production. Jahic *et al.* [140] compared host cell protein release from cultures grown on methanol and glucose. The degree of cell lysis in cells grown on methanol and therefore contamination of the secreted protein was much higher.

This combined with the fact that it is a highly flammable substrate and results in strong heat production and oxygen consumption limits the use of methanol in the pAOX1 system and may not always be the ideal method of choice. However, pAOX1 does offer users of the PTK a way to tune the expression rate using a single construct before deciding on the strength of constitutive promoters to use in the final design.

3.7 Transformation of *Pichia pastoris*

High-throughput screening methods rely on efficient transformation strategies to ensure that a large portion of a potential range of designs can be tested. In general, yeast transformations have a significantly lower transformation efficiency (*P. pastoris* 10^3-10^5 transformants/µg DNA [65]) in comparison to *E. coli* (10^9-10^{10} transformants/µg DNA [86]). This can become a limiting factor for the generation of libraries.

For *P. pastoris*, expression can either be plasmid-based or from a genomically integrated gene of interest. Even though transformation efficiency of plasmid systems is one to two orders of magnitude higher than integrations, the latter is usually preferred as it is more stable [65, 215]. Here, a *P. pastoris* strain containing a *attP* site [234], which enables a precise and single-copy insertion of plasmid DNA into a known genomic region via BxbI recombinase-mediated integration was used. Therefore, the strain was co-transformed with the cassette plasmid encoding the genetic construct of interest and the *attB* site as well as a plasmid for the transient expression of BxbI. For a selection of promoters, plasmid-based expression was studied using the *Pichia* Autonomous Replication Sequence (PARS) [65]. For both expression strategies, no vector linearisation was required for transformation. Throughout this study, two different protocols for the generation of competent cells were applied, which significantly influenced the transformation efficiency.

Following the method of Perez-Pinera *et al.*, transformation efficiency was approximately 10^2 transformants/µg DNA for recombinase-based genomic integration and 2 x 10^2 transformants/µg DNA for plasmid-based constructs. By applying the method of Madden *et al.* [192] with modifications (see Methods Section 2.2.5), the transformation efficiency could be increased to 10^4 transformants/µg DNA. This method is suitable for the generation of comparatively large libraries and should be applied when the transformation efficiency is of concern.

To further accelerate the strain construction, the constructs for the terminator characterisation (42 strains, pUO_pL730 – pUO_pL771, Table A.4) were made by directly using the Golden Gate mixture for *P. pastoris* transformation. After the Golden Gate reaction, the solution should contain 40 fmol cassette plasmid (approximately 250 ng DNA) in 10 µL, which was used for *P. pastoris* electroporation (cells prepared according to Madden *et al.* [192]). This method can be applied when strain generation time is of concern as the protocol duration was shortened by one working day, but transformation efficiency was reduced (2 x 10^2 transformants/µg DNA) making this method less applicable for the generation of libraries.

All transformants were PCR verified and for *P. pastoris* strains made directly from the Golden Gate reaction, a selection of fragments was sequence verified. For genomically integrated constructs, primers spanning the defined BxbI locus were applied. For PARS-based constructs, primers within the expression construct were used to verify correct strains.

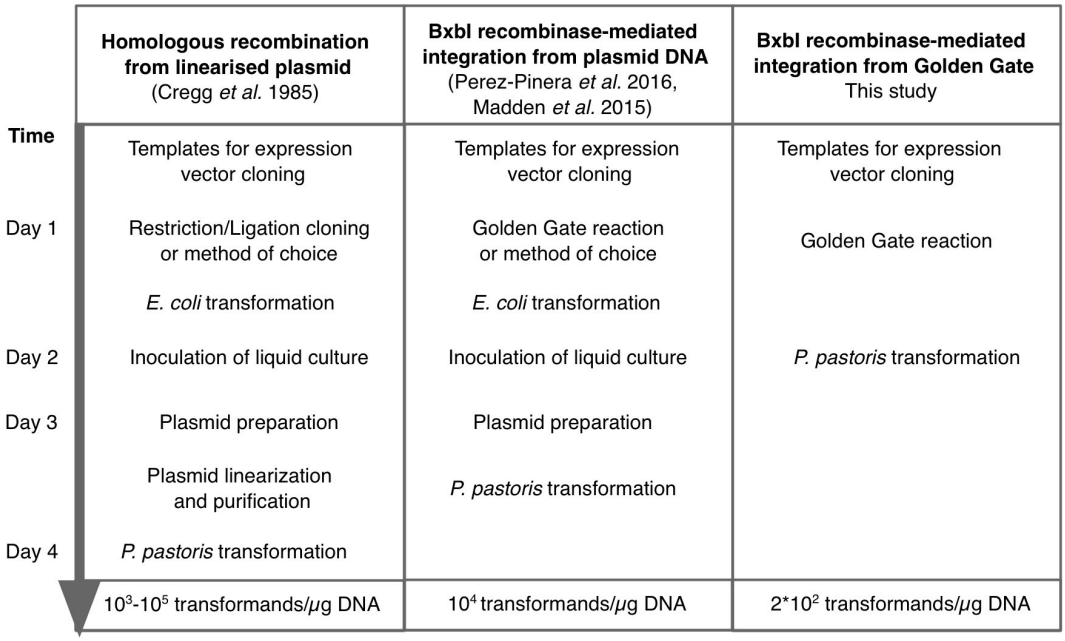

FIGURE 3.14. Overview of *P. pastoris* transformation procedures. Comparison of the laboratory effort for the different *P. pastoris* strain construction methods.

3.8 Conclusion

To economically engineer yeasts as a protein production platform, many parts have to be tuned and thus the fast generation of expression variations is necessary. Standards are essential to efficiently design expression vectors and engineer heterologous protein production using flexible libraries based on well characterised parts. To avoid the creation of different standards and increase the rational diversity of an existing toolbox, the YTK [172] was extended. Specifically, new functional modules that allow heterologous protein expression and secretion in *P. pastoris* were added. 242 expression constructs were built and characterised for protein expression and secretion, a library size that was made possible due to the efficient assembly protocol and recombinase-based integration procedure.

4

Application of the *P. pastoris* Toolkit for Secretion Libraries

The ideal conditions for protein secretion in *P. pastoris* are currently challenging to predict, especially the choice of the signal peptide to guide secretion. To achieve high amounts of secreted protein, various signal peptides (tags) should be tested for a gene of interest. As described in Chapter 3, the PTK allows a simple and straight forward cloning approach for the generation of defined *P. pastoris* strains. However, the generation of large libraries is still laborious and time consuming. In this chapter, the application of the PTK for the generation of randomised libraries is studied. Three industrially relevant enzymes are used as a proof-of-concept: phytases, unspecific peroxygenases (UPO) and alcohol dehydrogenases (ADH). Golden Gate shuffling is applied for phytases and UPOs, to generate libraries of varying promoters and secretion tags. A variety of phytases previously successfully expressed in *P. pastoris* and unspecific peroxygenases either previously heterologously expressed or solely predicted are used. First, the underlying studies allowing Golden Gate shuffling and library generation of *P. pastoris* were performed. Second, the screenings for phytase and UPO secretion were developed, and finally, screenings for constitutive and induced expression were performed for each enzyme. For the phytases and one unspecific peroxygenases the preferred secretion tags could be identified. Bacterial ADHs were studied for heterologous expression in *P. pastoris*, but no secreted ADH could be detected under the tested conditions to establish a screening set-up.

4.1 Prerequisites for *Pichia pastoris* Library Design

Golden Gate shuffling is a technique that allows for the generation of diverse genetic libraries in a single restriction-ligation reaction [93]. As for the Golden Gate reaction, this method relies on Type IIs restriction enzymes and carefully designed parts/modules, see Chapter 1. For the shuffling approach, multiple variations of a part are used in the restriction-ligation reaction. Because of the carefully designed overhangs, a part can only ligate to the designed consecutive part, ensuring the correct order of assembly, but allowing varying final plasmids as for a single part different options are possible. With each restriction-ligation cycle, the complete constructs will accumulate as they do not contain restriction recognition sites, however, the individual parts do and will be immediately digested. As the overhang is identical for each homologous part, ligation should occur with equal probability across all variations of one part to the matching part of a contiguous set. However, as recently reported by Potapov *et al.*, special care has to be taken while designing the overhang as ligation efficiency varies [243]. Previously, Golden Gate shuffling was applied for enzyme engineering purposes as described in the case of trypsinogen [93] or lipase A [251] assembly. Here, the enzyme of interest is assigned to one individual part as defined in the PTK and the parts of the transcriptional unit are varied by Golden Gate shuffling, Figure 4.1. In a one-pot reaction, the previously built promoters and secretion signal peptides (Chapter 3) are shuffled for the desired enzymes of interest. The limitation in transformation efficiency of *P. pastoris* still hampers the design of big libraries. Here, the aim is to overcome this bottleneck and take a step towards high-throughput screens directly in *P. pastoris*.

To use *P. pastoris* as the host for enzyme expression screenings, the generation of sufficient amounts of variants is the main bottleneck. The overall procedure for the generation of a *P. pastoris* library starts with the Golden Gate reaction which is used for the transformation of *E. coli*. From the *E. coli* colonies, a plasmid library is prepared, which is subsequently used for the transformation of *P. pastoris* (Figure 4.2). To address the bottleneck for the generation of large libraries, two procedures need to be considered. First, the Golden Gate shuffling needs to produce sufficient amounts of correctly assembled expression vector constructs, without leaving any backbone vector or falsely assembled vectors. Second, the *P. pastoris* transformation procedure must be efficient enough to produce a sufficient number of transformants to screen.

A variety of protocols are available for the Golden Gate shuffling [93, 270]. Different Golden Gate approaches were tested to study the influencing factors by varying the type of ligase as well as the final digestion and heat inactivation step (Table 4.1). T7 DNA ligase and T4 DNA ligase are commonly used for Golden Gate reactions, T7 ligase only ligates sticky-ends but no heat inactivation is possible. T4 ligase can be heat inactivated, but it will also ligate blunt-ends.

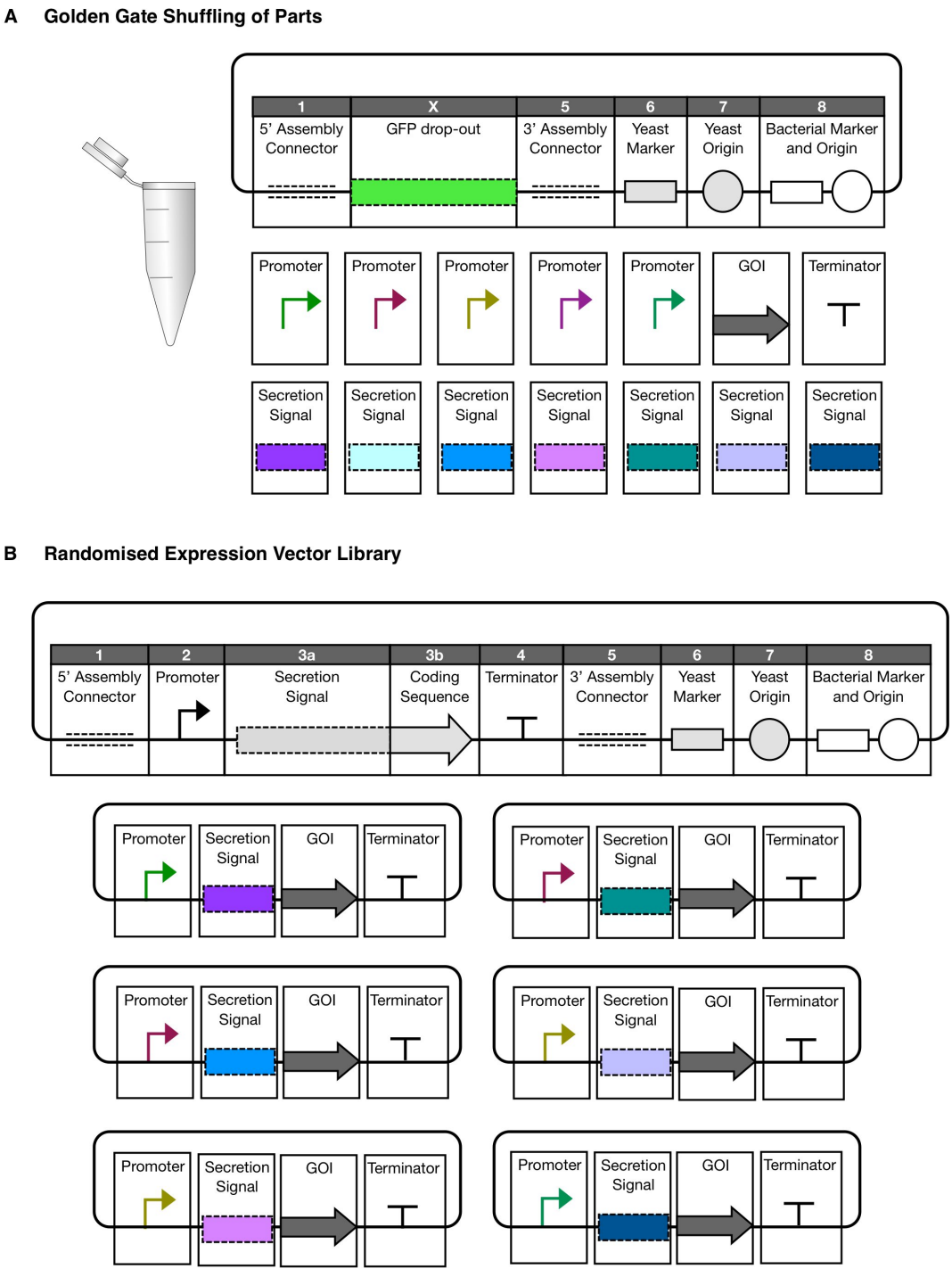

FIGURE 4.1. Golden Gate shuffling methodology. **A** In a one-pot reaction a backbone vector, the GOI, a terminator, as well as a selection of promoters and secretion signal peptides are merged to a randomised library **B**.

The consideration is, that the final step should be a restriction digestion, leaving only desired expression vectors. Therefore, at the end, DNA ligase activity should be reduced as much as possible to not re-ligate residual backbone fragments. For this analysis, the primary steps of the pre-digestion, and the cycling of digestion and ligation were performed as described by Engler *et al.* [93]. Then, a variety of digestion temperatures and heat inactivation procedures were studied. Furthermore, an extended digestion step was included after the cycling, by adding 1 µL fresh restriction enzyme (BsaI) to the Golden Gate reaction with incubation for 1 h at 37 °C. After *E. coli* transformation and incubation, the number of white colonies (expression vectors) and green colonies (backbone vector with GFP) were evaluated. Application of the T4 ligase, a final digestion at 37 °C, and heat inactivation at 65 °C lead to the highest amount of transformants and no backbone colony, and was therefore used for the library design. To decrease the probability of any backbone vectors in the screening even further, the backbone was separately digested with BsaI and the correct fragment was purified via an agarose gel. The Golden Gate shuffling was then performed using this linear backbone fragment and the plasmid parts of the transcriptional unit in equimolar ratios.

To ensure whether all plasmids generated via Golden Gate shuffling were screened, the *P. pastoris* transformation efficiency should allow generation of thousands of strains. First, the original protocol for generation of competent cells and *P. pastoris* transformation from Perez-Pinera *et al.* [234] was used, but transformation efficiency for recombinase-based genomic integration was approximately 2 x 10^2 transformants/µg DNA. Then, the method of Madden *et al.* [192] was tested and a transformation efficiency of approximately 10^4 transformants/µg DNA was achieved. To expand the understanding of influencing parameters for the transformation procedure, transformations of *P. pastoris* were performed by varying the DNA quantity, the DNA containing liquid volume, the cuvette, and electroporation settings, as suggested by Madden *et al.* [192].

TABLE 4.1. Golden Gate shuffling parameter variation.

Pre-digestion	37 °C, 10 min				
Digestion Ligation	37 °C, 5 min 16 °C, 5 min			50 x	
Experiment	**1** [93]	**2**	**3**	**4**	**5**
Ligase	T4	T4	T7	T4	T7
Digestion	50 °C, 10 min	37 °C, 10 min	37 °C, 10 min	+1 µL BsaI, 37 °C, 1 h	+1 µL BsaI, 37 °C, 1 h
Heat inactivation	80 °C, 20 min	65 °C, 20 min	65 °C, 20 min	-	-
White colonies	350	1000	800	80	800
Green colonies	0	0	20	0	5

The plasmids for constitutive and intracellular expression of RFP and yEGFP were used because of the change in colour of colonies the positive transformants can be directly counted on the transformation plate. Varying the DNA amount of the pGAP – RFP plasmid (pUO_pL_002) showed a higher transformation efficiency for 250 ng compared to 1 µg of DNA. Concentration variation e.g., variation of the amount of DNA and water to the competent cells did not affect the efficiency for pGAP – RFP, as opposed to pGAP – yEGFP (pUO_pL_008). However, there was no clear trend in the behaviour. These transformations were performed as suggested by Madden *et al.* [192] using a 1 mm cuvette and electroporation parameters for *E. coli* (1800 V, 25 µF, 200 Ohm). Transformation parameters as suggested by Perez-Pinera *et al.* using a 2 mm cuvette and electroporation settings for *S. cerevisiae* (1500 V, 25 µF, 200 Ohm) but with competent cells made as described by Madden *et al.* [192] reduced transformation efficiency. These results provide initial data on the influential parameters for the transformation of *P. pastoris*, as transformation was done only once for each reaction. To draw conclusions with statistical significance, the experiments would need to be performed in triplicate. However, sufficient understanding of the influential parameters for the *P. pastoris* transformation was gained to define the operating conditions.

TABLE 4.2. Transformation of *P. pastoris*

Sample	**DNA amount [ng]**	**DNA concentration [ng/µL]**	**Cuvette [mm]**	**Transformation efficiency [cfu/µg DNA]**
pGAP - RFP	1000 500 250	720	1	$1.5 * 10^4$ $3.4 * 10^4$ $6.1 * 10^4$
pGAP - RFP	500	500 250 125	1	$2.2 * 10^4$ $2.2 * 10^4$ $2.5 * 10^4$
pGAP - yEGFP	500	500 250 125	1	$8.0 * 10^4$ $1.3 * 10^4$ $4.0 * 10^4$
pGAP - RFP pGAP - yEGFP	500 500	500 500	2	$8.0 * 10^2$ $3.3 * 10^3$
Wild type	0	-	1	0

The procedure for the generation of a *P. pastoris* library with varying parts of the transcriptional unit using Golden Gate shuffling is shown in Figure 4.2. First, Golden Gate shuffling is performed and used for *E. coli* transformation. The *E. coli* colonies are washed off the transformation agar plate and used for plasmid preparation. Finally, 250 ng DNA is to be used for transformation of *P. pastoris*.

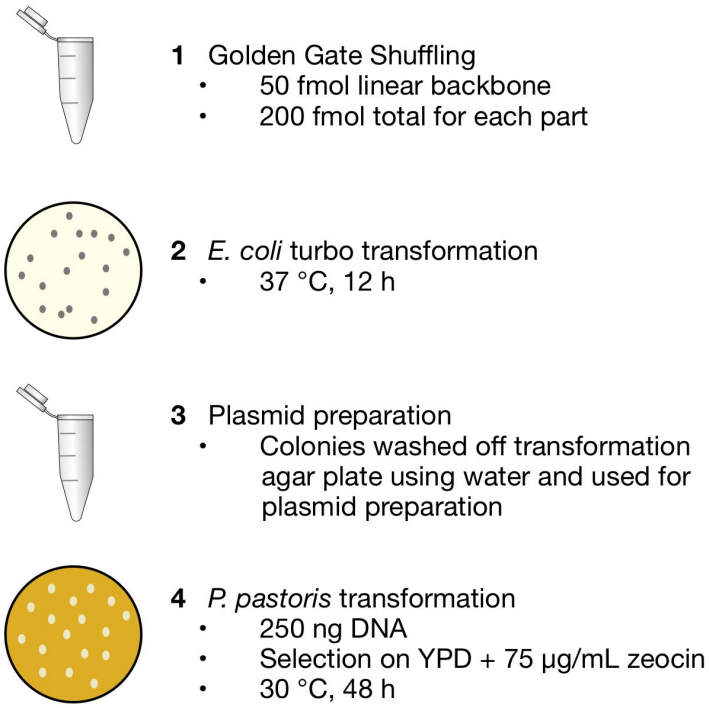

FIGURE 4.2. Generation of *P. pastoris* randomised library.

4.2 Phytases

Phytate (myo-inositol(1,2,3,4,5,6)hexakisphosphate) is the main storage form of phosphorus (P) and inositol in plants, especially in grains and legumes, the main sources for animal feeds. However, monogastric and agastric animals lack sufficient amounts of digestive enzymes to use phytate and utilise the phosphorus. Therefore, phytases are one of the most important enzyme additives of cereal and grain feed for monogastric animals such as poultry, swine, and fish [55, 257]. Phytases (myo-inositol hexakisphosphate phosphohydrolase) hydrolyse phytate ester bonds to less phosphorylated myo-inositol and inorganic phosphates, providing the main phosphorus source for animals [257]. Phytate also binds important minerals such as Ca^{2+}, Mg^{2+}, Zn^{2+}, Cu^{2+}, Fe^{2+}, Mn^{2+}, and proteins, reducing their availability [55]. In addition to phytate being an antinutritional compound for animal feed itself, undigested phytate results in unwanted ecological consequences. Phytate, when excreted by animals will be degraded by soil bacteria releasing high concentrations of $PO_4{}^{2-}$, leading to the eutrophication of lakes and rivers [110]. The industrial application of phytases in animal feed improves phosphate and mineral uptake, diminishing the need to supplement inorganic phosphorus necessary for the animal nutrition and significantly lowering environmental pollution [257].

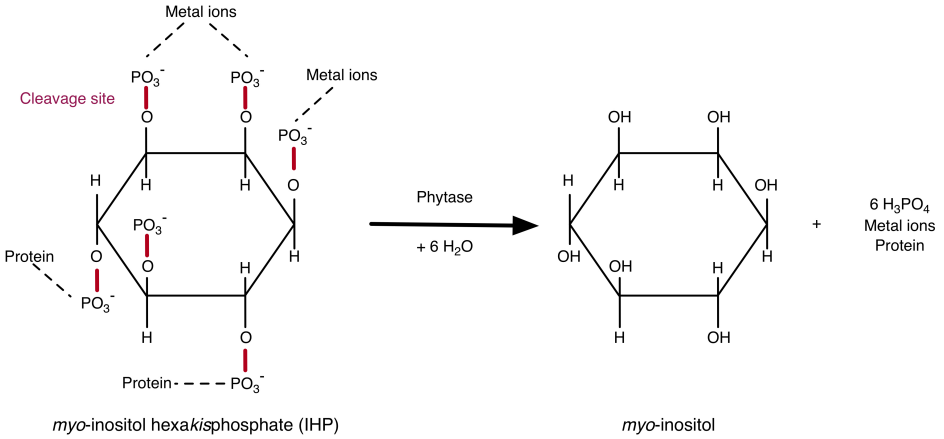

FIGURE 4.3. Reaction mechanism of phytate hydrolysis. Phytase cleaves the phosphoester bonds releasing phosphate groups, myo-inositol, metal ions, and bound proteins. Figure adapted from Chen *et al.* [55].

Phytases are wide spread across almost all forms of life such as bacteria, fungi, plants, and animals [55]. Just like their diverse origin, the enzymatic structure and the reaction mechanism cleaving phosphate groups vary considerably [213]. Phytases can be classified based on their optimal pH, the catalytic mechanism, or the initiation site of dephosphorylation [205]. Categorization according to pH yields acid, neutral, or alkaline phosphatases, with acid phosphatases being the most studied ones, as they will most likely be active in the digestive tract of monogastric animals [213]. Considering the reaction mechanism, three groups can be defined: (1) histidine acid phosphatases or acid phosphatases (EC 3.1.3.2), (2) β-propeller phytases (EC 3.1.3.8), and (3) purple acid phosphatases or cysteine phytases (EC 3.1.3.2) [257]. Finally, classification according to the initiation site of dephosphorylation defines 3-phytases, 6-phytases, and 5-phytases [257]. Ever since the commercialisation of phytases in 1991 by Gist-Brocades (now DSM) and sold by BASF [257], the phytase market has exceeded $300 million, growing about 10 % per year [55]. Hence, there is still substantial interest in improving catalytic efficiency, substrate specificity, thermostability, modification of the pH profile, and reduction in production cost [55].

To test the randomised library approach for finding suitable expression and secretion conditions in the yeast *P. pastoris*, a screening of 4 industrially interesting phytases was performed. These four phytases have previously been successfully secreted in *P. pastoris*, providing the ideal starting point to test the Golden Gate shuffling for randomised libraries. The selection of the phytases was done so as to ensure comparability of each phytase screening. The selected phytases are all histidine acid phosphatases or acid phosphatases with a conserved active site sequence (RHGXRXPT) [226, 257].

TABLE 4.3. Summary of phytases.

Taxa		**Fungi**		**Bacteria**
	Ascomycetes	Ascomycetes	Basidiomycetes	Gram-negative
Microorganism	*A. niger*	*T. heterothallica*	*P. lycii*	*E. coli*
NCBI accession number	CAA78904	KF535924	CAC48195	M58708
Molecular function	3-phytase activity	3-phytase activity	4-phytase activity	4-phytase activity
Gene	phyA	phyA	phyA	appA
bp	2665, 1 intron in tag	1521,1 intron in tag	1568	1296
Amino acids (signal peptide + chain)	467 (23+444)	487 (21+466)	439 (29+410)	432 (22+410)
N-linked glycosylation	10	4	10	3

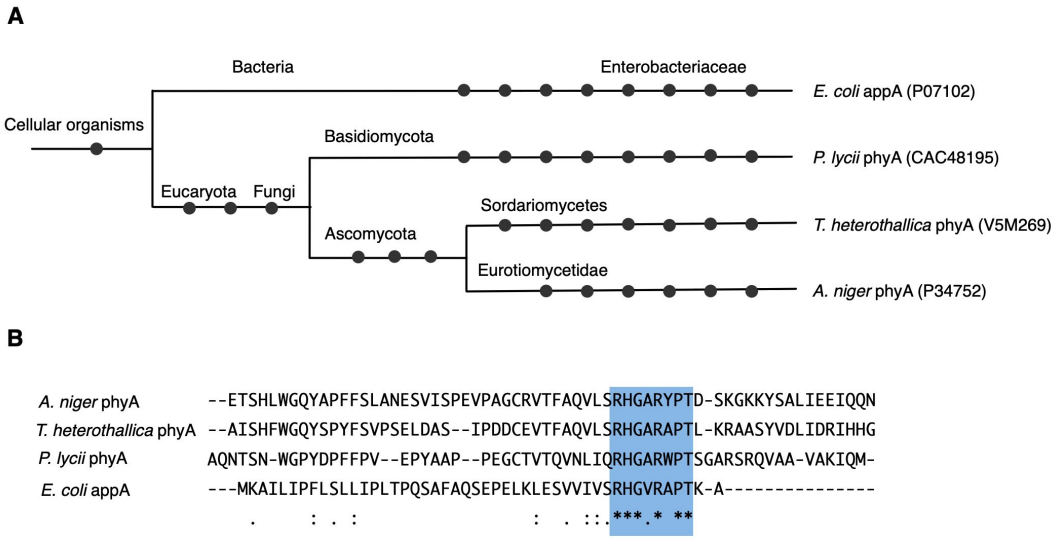

FIGURE 4.4. Phylogenetic tree and sequence alignment of phytase selection. **A** Phylogenetic tree generated using phyloT (https://phylot.biobyte.de). **B** Sequence alignment performed using T-Coffee Expresso multiple sequence alignment (http://tcoffee.crg.cat/apps/tcoffee/do:expresso) [221].

Three fungal phytases (phyA) from *Aspergillus niger* (*A. niger*), *Thielavia heterothallica* (*T. heterothallica*), *Peniophora lycii* (*P. lycii*) as well as a bacterial phytase (appA) from *Escherichia coli* (*E. coli*) were studied, Table 4.3, Figure 4.4 A showing the phylogenetic relationship of the phytases and B an excerpt of a sequence alignment along the conserved active site. *Aspergillus* phytases are mainly utilised for industrial applications as feed additives and a variety of phytase products are authorised by the European Union for this usage [118]. Han *et al.* [123] first expressed an *Aspergillus* phytase heterologous in *P. pastoris* and ever since, multiple studies have been performed to improve the thermostability [129, 180], the specific activity [181], or to reach higher expression levels in *P. pastoris* by varying expression conditions such as varying the secretion signal peptide or using a codon optimised sequence [324].

The phytase from the thermophilic mould *T. heterothallica* is industrially attractive as it offers high acid stability and thermostability, a wide substrate spectrum as well as resistance to proteolytic cleavage [279]. Expression using *P. pastoris* was previously done by Ranajan *et al.* [256] using a codon optimised sequence. The basidiomycete fungi *P. lycii* secretes a phytase showing 4-phytase activity [169] and a codon optimised version was successfully secreted in *P. pastoris* using a synthetic signal peptide (MF41) by Xiong *et al.* [325]. Lately, bacterial phytases have been of interest as their catalytic activity is competitively better than fungal phytases [55]. The phytase appA from *E. coli* has been expressed and secreted in *P. pastoris* using various expression strategies [6] or additional glycosylation sides [261].

To study these phytases, all sequences have been designed to be compatible with the PTK, see Figure 4.5. Each enzyme coding sequence was made to be a suitable part 3b for the Golden Gate assembly (removal of BsaI and BsmBI sides) and codon optimisation for *P. pastoris* using GeneOptimized (ThermoFischer). The annotated secretion signal sequences of the fungal phytases were designed as 3a parts. The signal peptides of *A. niger* and *T. heterothallica* each contained one intron, which was removed. For all tags, BsaI or BsmBI restriction sides were removed but the sequences were not codon optimised.

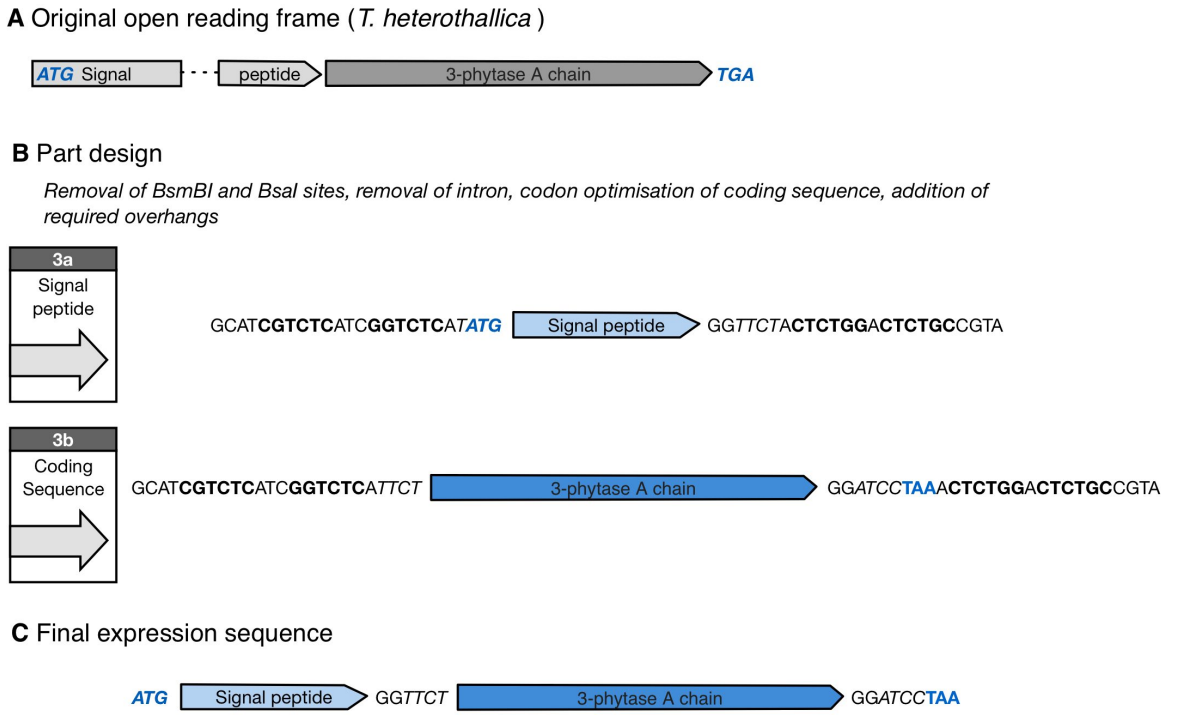

FIGURE 4.5. Part design for phytases. Based on the phytase open reading frame, a 3a part for the secretion signal sequence and a 3b part for the enzyme coding sequence were designed.

4.2.1 Development and Validation of a Screening Quantifying Phytase Activity

The fundamental requirement for any screening is the need to differentiate positive hits (expression of active enzyme) from negative variants (no expression or no active enzyme) and quantify the active enzyme. Therefore, the process from strain cultivation until the enzymatic assay must ensure enzyme expression and a sensitive measurement.

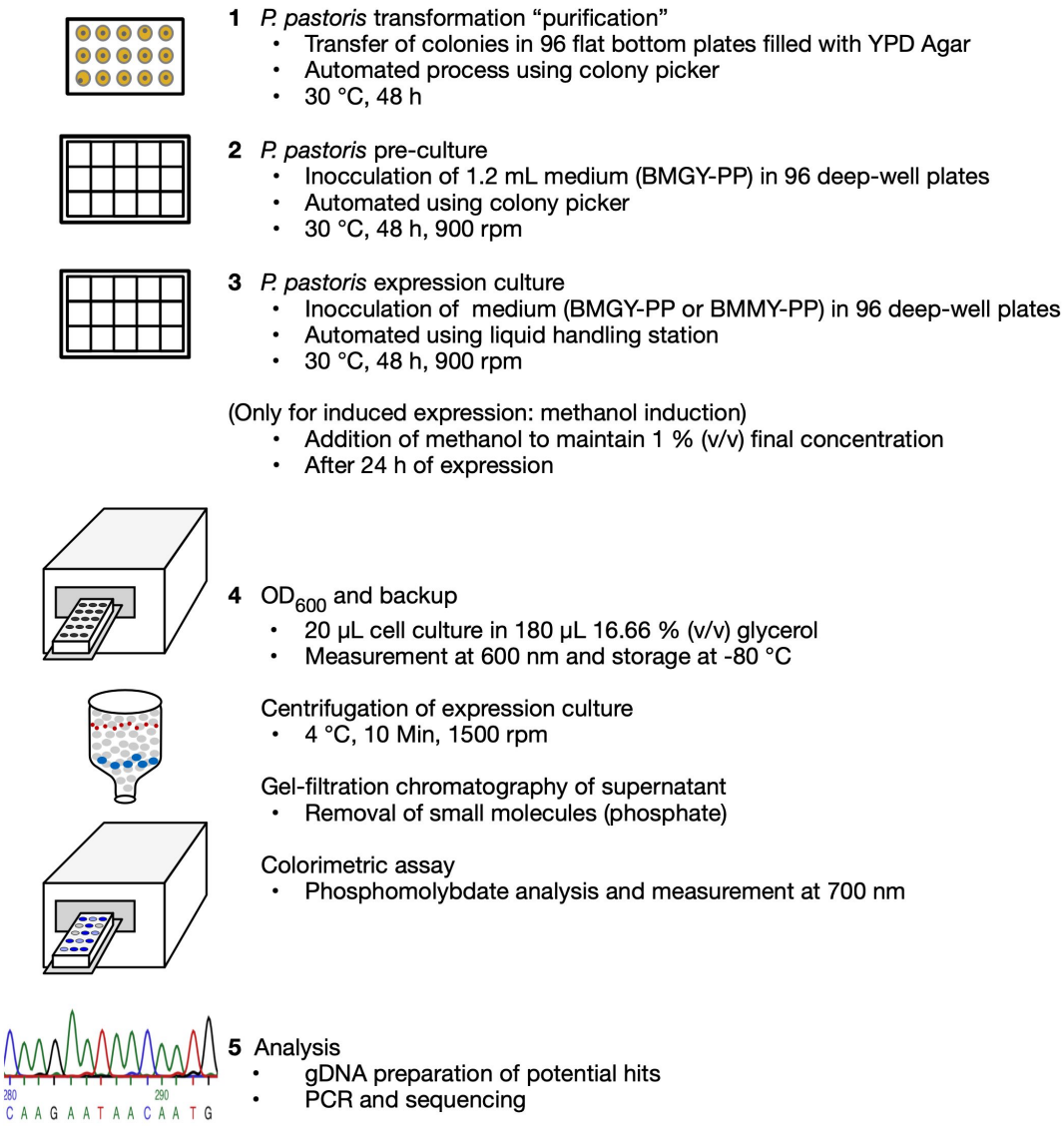

FIGURE 4.6. Workflow of *P. pastoris* screening for phytases.

Figure 4.6 summarises the final workflow of the phytase screening. To allow measurements of phytase activity in higher throughputs, an established colorimetric phosphate analysis was used measuring the released phosphate [253]. A major challenge for this approach was the occurrence of high phosphate in the commonly used media for *P. pastoris* cultivation. To asses this, first the colorimetric phosphomolybdate analysis method from Bae *et al.* [24] was adapted to suit the 96-well format and subsequently commonly used *P. pastoris* media, as well as their main components were measured (Figure 4.7 A). Except for water and biotin, all media and their components already resulted in high absorption signals. Therefore, a more detailed look at the screening procedure was needed and positive controls were made for each phytase as previously described in literature, but using the parts of the PTK/YTK [6, 123, 230, 325] Figure 4.7 B. These controls were expressed in a selection of media and gel-filtration chromatography was performed for the expression supernatant. This purification step helps to remove most of the phosphate from the medium and enables measurement of phosphate released from the substrate sodium phytate by the secreted phytase. Figure 4.7 C shows that even after the gel-filtration step, the absorption measured for the media BMM and BMMY is almost as much as for the expression constructs expressing active phytase and therefore these are unsuitable for the screening. However, 1 x YP + 1 % methanol as well as BMMY-PP (buffered complex methanol medium without phosphate buffer) are suitable for the screening as the phosphate release due to phytase activity can be differentiated from the phosphate content of the medium. BMMY-PP was further assessed (Figure 4.7 D) and the supernatant before and after gel filtrations and dilutions were assayed. A dilution step is necessary in order to differentiate the expression levels of different constructs as without the dilution step, the upper range of measurement is already reached.

The Z-factor was determined, to validate the suitability of the screening and determine the quality of the assay. The Z-factor was introduced by Zhang *et al.* as a simple statistical parameter for the evaluation of assays [338] and is ever since widely used [202, 335]. The screening procedure as described in Figure 4.8, was performed for one plate of wild type (i.e. control) and one plate of the *P. pastoris* strain pUO_pL666 expressing the *E. coli* phytase (i.e. sample), Figure 4.8. The Z-factor can be calculated according to Formula 4.1 (SD standard deviation). The determined Z-factor of 0.8268 stands for an excellent screening with a high degree of confidence and overall, a high-quality screening that can be applied for the targeted secretion libraries [338].

$$Z = 1 - \frac{3 \cdot \text{SD of sample} + 3 \cdot \text{SD of control}}{|\text{mean of sample} - \text{mean of control}|} \tag{4.1}$$

$$Z = 1 - \frac{3 \cdot 0.0347 + 3 \cdot 0.0017}{|0.7489 - 0.1186|} \tag{4.2}$$

$$Z = 0.8268 \tag{4.3}$$

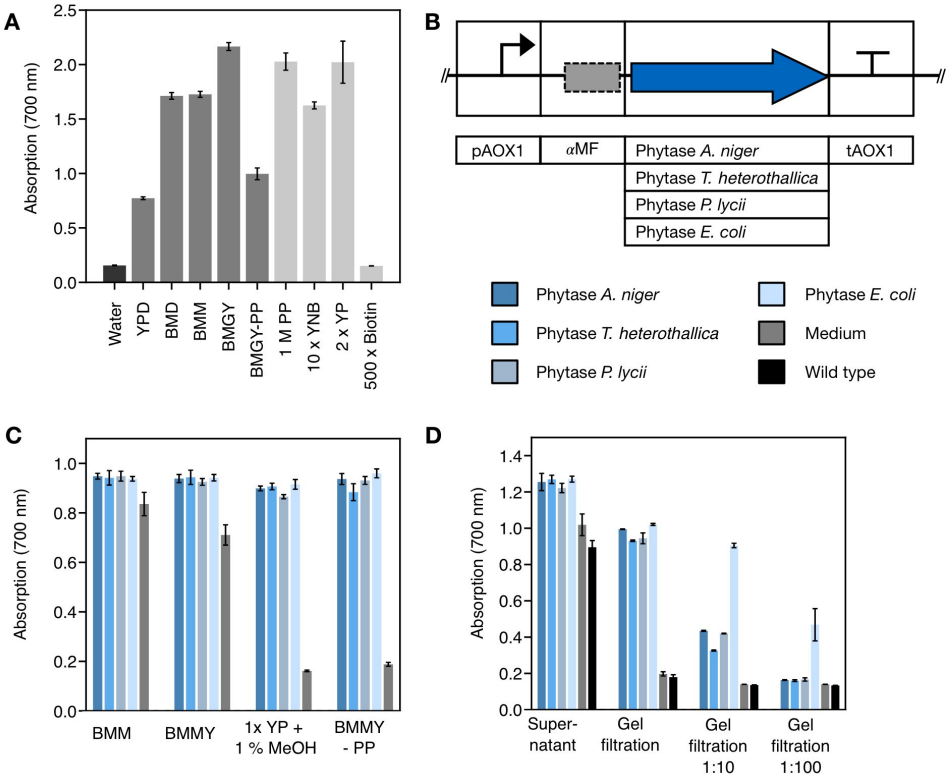

FIGURE 4.7. Colorimetric phosphate analysis of media and phytase expressions. **A** Widely used media and their components were evaluated for their initial phosphate content. **B** *P. pastoris* controls expressing each phytase were made and **C** the enzyme activity was measured after gel-filtration of the supernatant. **D** Phytase expression in BMMG-PP with gel-filtration of the supernatant and determination of the phytase activity was performed and different dilutions of the gel-filtration filtrate were assayed. Bars represent the mean absorption from the phytase assay of triplicates measured at 700 nm. Error bars denote ± 1 standard deviation.

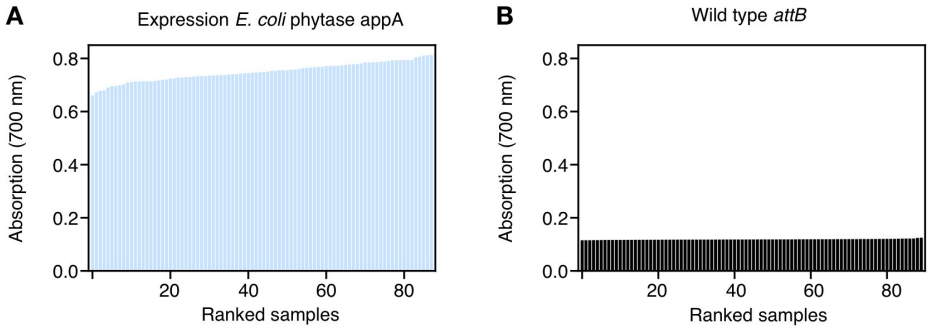

FIGURE 4.8. Wild type landscape to determine Z-factor. **A** Expression of positive control and **B** the wild type to determine quality of assay. Bars represent absorption at 700 nm after the phytase assay.

4.2.2 *P. pastoris* Screening of Phytase Secretion

The aim for the library generation using Golden Gate shuffling and subsequent screening in *P. pastoris* is to find good secretion conditions at the genetic level e.g., a suitable combination of promoter, tag and the desired gene of interest to reach high amounts of secreted protein. To do so, a secretion library for induced expression (3 different promoters) and constitutive expression (5 different promoters) were made for each phytase variant (*A. niger*, *T. heterothallica*, *P. lycii* and *E. coli*), Figure 4.9. The Golden Gate shuffling reactions were made by preparing a master mix containing the linearised backbone, 20 secretion signal peptides as well as the terminator part, Figure 4.9 A. This primary master mix was divided into two secondary master mixes, Figure 4.9 B, and the respective promoter parts for each induced and constitutive expression were added. These secondary master mixes were each split into 4 reactions for each phytase respectively. The final master mixes contained 50 fmol of backbone, 100 fmol for all parts of one functionality (part type), restriction enzyme BsaI as well as T4 ligase and assembly was performed as described in Section 2.3.6. The detailed list of parts in each shuffling set-up can be found in the Appendix Table A.10. After transformation of *E. coli* and plasmid preparation, transformation of *P. pastoris* was performed to obtain the *P. pastoris* library.

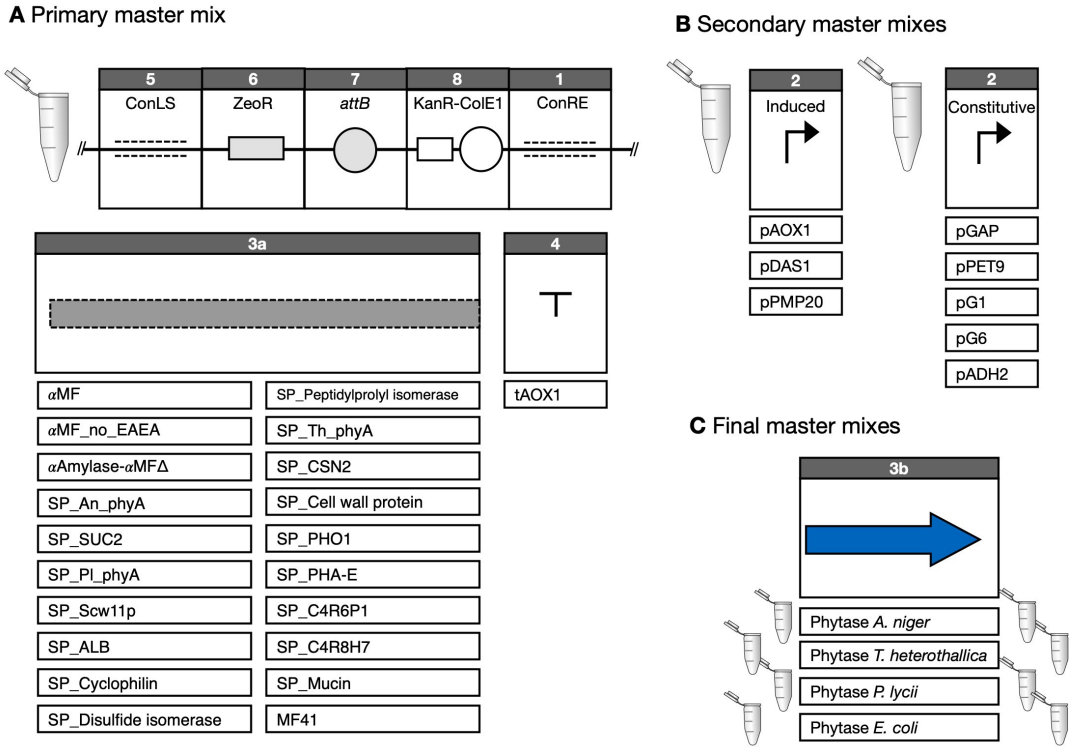

FIGURE 4.9. Parts for Golden Gate shuffling of phytase library.

To confirm sufficient statistical evidence, the expected coverage, i.e. the expected percentage of all possible variants that are represented in the library, was calculated according to Formula 4.4. The calculation relies on the assumption that the resulting randomised expression plasmids generated via Golden Gate shuffling are all likely with equal probability. For each constitutive phytase library, the number of equiprobable variants n = 100 and for the inducted phytase library n = 60. After *E. coli* transformation, the number of *E. coli* colonies used to prepare the plasmid library were counted and the expected coverage Y_i within *E. coli* was calculated. After transformation of *P. pastoris* with the respective plasmid library, the library size x was defined as 364 colonies were picked for each constitutive screening and 182 for each induced screening. The expected *P. pastoris* library coverage (Formula 4.4), as well as the final expected coverage (Formula 4.5) could be calculated [231]. For the constitutive library, the expected coverage of all 100 possible combinations represented in the library is > 97 % and for the induced library with 60 possible combinations the expected coverage is > 95 %.

$$Y_i = 1 - (1 - (\frac{1}{n}))^x \tag{4.4}$$

$$Y_i \text{ (final)} = Y_i \text{ (\textit{E. coli})} \cdot Y_i \text{ (\textit{P. pastoris})} \tag{4.5}$$

TABLE 4.4. Calculation of library coverage for constitutive library.

	Estimation of *E. coli* colonies	Y_i *E. coli*	*P. pastoris* library size	Y_i *P. pastoris*	Y_i (final)
Phytase *A. niger*	3200	100.00 %	364	97.42 %	97.42 %
Phytase *T. heterothallica*	1100	100.00 %		97.42 %	97.42 %
Phytase *P. lycii*	950	99.99 %		97.42 %	97.42 %
Phytase *E. coli*	950	99.99 %		97.42 %	97.42 %

TABLE 4.5. Calculation of library coverage for induced library.

	Estimation of *E. coli* colonies	Y_i *E. coli*	*P. pastoris* library size	Y_i *P. pastoris*	Y_i (final)
Phytase *A. niger*	1050	100.00 %	182	95.31 %	95.31 %
Phytase *T. heterothallica*	400	99.88 %		95.31 %	95.19 %
Phytase *P. lycii*	500	99.98 %		95.31 %	95.29 %
Phytase *E. coli*	800	100.00 %		95.31 %	95.31 %

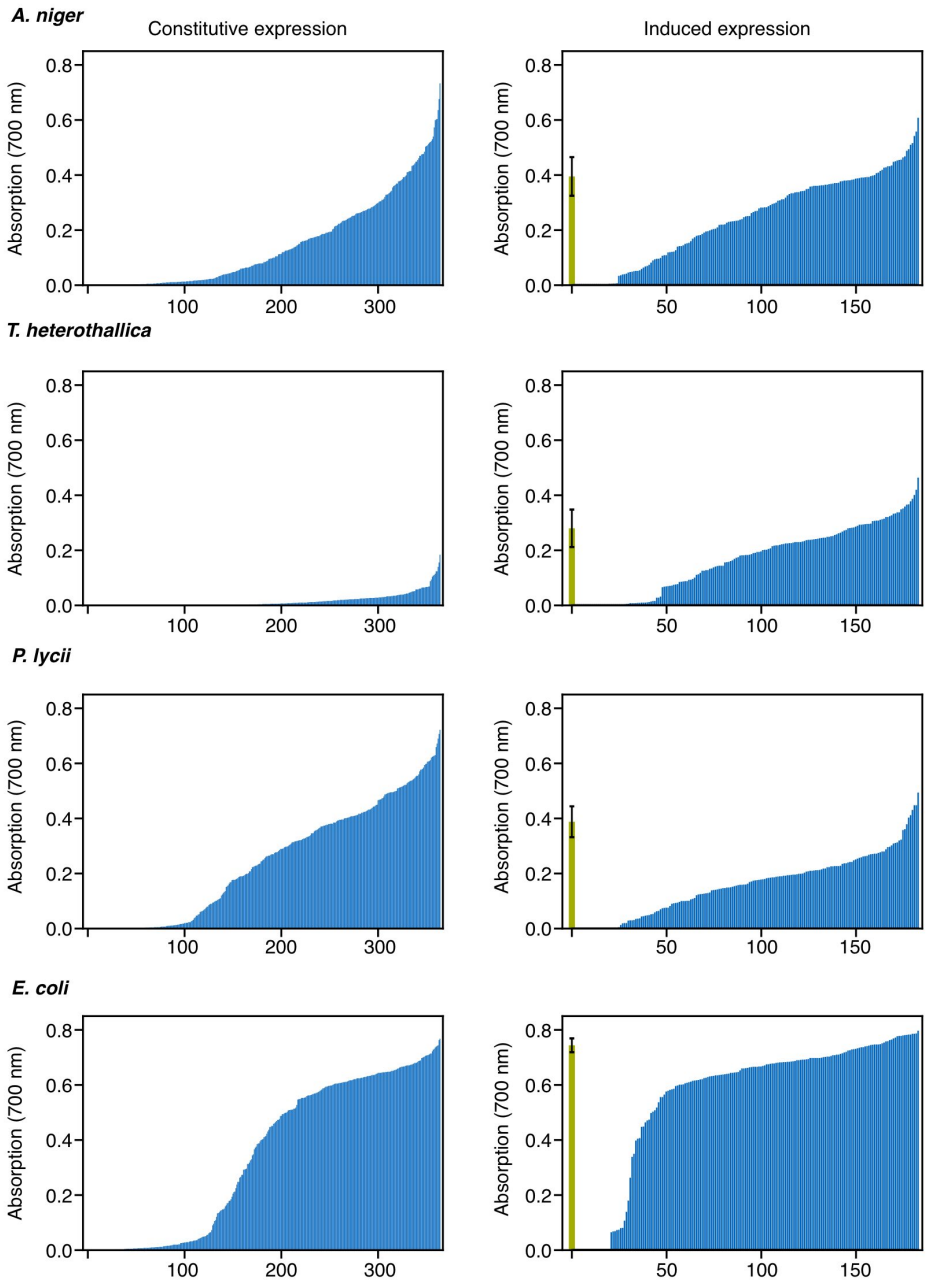

FIGURE 4.10. Screening of phytases. For each phytase a screening for constitutive and induced expression was performed. Constitutive expression was performed in 2 % BMGY-PP and induced expression in 1 % BMMY-PP. For induced expression, the controls for each phytase (Figure 4.7 B) were assayed along with each respective phytase plate (green bar, error bar denoting ± 1 standard deviation of all phyatse specific controls measured). To remove background noise, absorption values presented are measured absorption for each well subtracted by the average absorption of 3 medium controls which were assayed on each plate. The x-axis defines the number of samples screened in a ranked order.

The screenings were performed as described in Section 2.2.3 and absorption was measured at 700 nm, Figure 4.10. For all screenings, strains showing active enzyme could be found, however, the overall distribution of high or low protein secretion and measured activity varies significantly between the phytases. The overall trend for each landscape shows similarities between the constitutive and induced screening. For the screening of the *A. niger* and *P. lycii* phytases, the gradually decreasing slope indicates a clear preference of the promoter and secretion tag. For the *E. coli* phytase, the landscape slope is much smaller with many strains showing high activity, indicating a much higher degree of flexibility in the choice of secretion tag and promoter. The overall measured activity of the *T. heterothallica* phytase is less then all the other phytases, especially for its constitutive expression only few secreting strains could be detected. For induced *T. heterothallica* phytase secretion, the slope is similar to the one of the *A. niger* phytase secretion with the performance decreasing more quickly than for *E. coli* phytase secretion.

To confirm the results of this screening, a re-screening was done for each library. The highest performers of the induced (10 colonies) and constitutive screening (22 colonies), as well as colonies not expressing phytase (3 and 5 colonies, respectively) were picked onto fresh agar plates. For these samples, the screening using three biological replicates was repeated and the genetic origin was determined for a selection of these strains via sequencing, Figure 4.11 and 4.12. All the hits with the highest phytase activity prefere a small selection of signal peptides.

For the *A. niger* phytase, the signal peptides SP_Cyclophillin, SP_C4R8H7 and SP_PHO1 in combination with the pGAP promoter lead to the highest constitutive expression, while SP_CSH2, SP_Peptidyl isomerase and also SP_C4R8H7 in combination with pAOX1 lead to the highest induced expression. The tag SP_CSH2 results in 22 % more activity than the strain using the αMF secretion tag as previously described in the literature [340].

For the *T. heterothallica* phytase, the pGAP promoter drives highest expression levels in combination with SP_Cyclophillin, SP_PHO1 and SP_Disulfide isomerase. For the induced *T. heterothallica* phytase expression the three best hits all use SP_C4R8H7 to drive protein secretion in combination with the pPMP20 or the pAOX1 promoter. In comparison to the control pAOX1-αMF, the pPMP20 - SP_C4R8H7 construct leads to 28 % higher activity.

P. lycii phytase expression and protein secretion do not favour the most widely used and very strong pGAP or pAOX1 promoter, but instead prefer the weaker pG6 and pAHD2 promoter for constitute expression and pPMP20 and pDAS1 for induced expression. For constitutive expression, the tags SP_Disulfide iosmerase, SP_Scw11p, and SP_Mucin guide highest amounts of phytase outside the cell and for induced expression SP_Mucin and αMF_no_EAEA were found to be the best during the screening. Nevertheless, for the induced expression, the control combination pAOX1-αMF results in 10 % more measured phytase activity, a combination that was not found by the screening.

Also for the induced *E. coli* phytase expression, no construct being better than the control could be found, however the control combination pAOX1-αMF was identified within the screening.

Almost identical amounts were determined by applying pAOX1-SP_PHO1 and pDAS1-MF41. The signal peptide MF41 in combination with pGAP also results in the highest measured phytase activity for the constitutive expression too.

Summarising the appearance and impact of the tags, each phytase secretion process prefers certain tags over others and some tags are not suitable for any of the sequenced phytases. 60 strains have been prepared for sequencing and six of the twenty tags did not appear in any strain. These are the three phytase signal peptides SP_An_phyA, SP_Pl_phyA, SP_Th_phyA as well as αAmylase-αMFΔ, SP_SUC2 and SP_PHA-E. For the secretion of RFP and yEGFP, the tags SP_An_phyA and SP_Pl_phyA have been identified to be unsuitable, Chapter 3.5, indicating that *P. pastoris* may not be capable of recognising these heterologous tags for its own secretion machinery. As the other tags have guided proteins out of the cell previously (Chapter 3), their functionality might have been hampered by the phytatse. The tag SP_PHO1 stands out, as for all constitutive phytase expression strains, it guides secretion very well. Otherwise, the more frequently appearing tags have preferences for certain phytases, such as the SP_C4R8H7 for *A. niger* and *T. heterothallica*, SP_Mucin for *P. lycii*, and MF41 or the related αMF and αMF_no_EAEA tags for *E. coli*. For constitutive expression, the pGAP promoter often results in highest expression levels, except for the expression of the *P. lycii* phytase, where pG6 and pAHD2 result in highest expression levels. Five different constitutive promoters have been applied in the constitutive screening and all except the pPET9 promoter appeared in the sequenced samples. For the induced expression, all three promoters resulted in high phytase expression levels, but pAOX1 and pPMP20 generally guided highest expression levels, as seen before for RFP and yEGFP (Chapter 3.2).

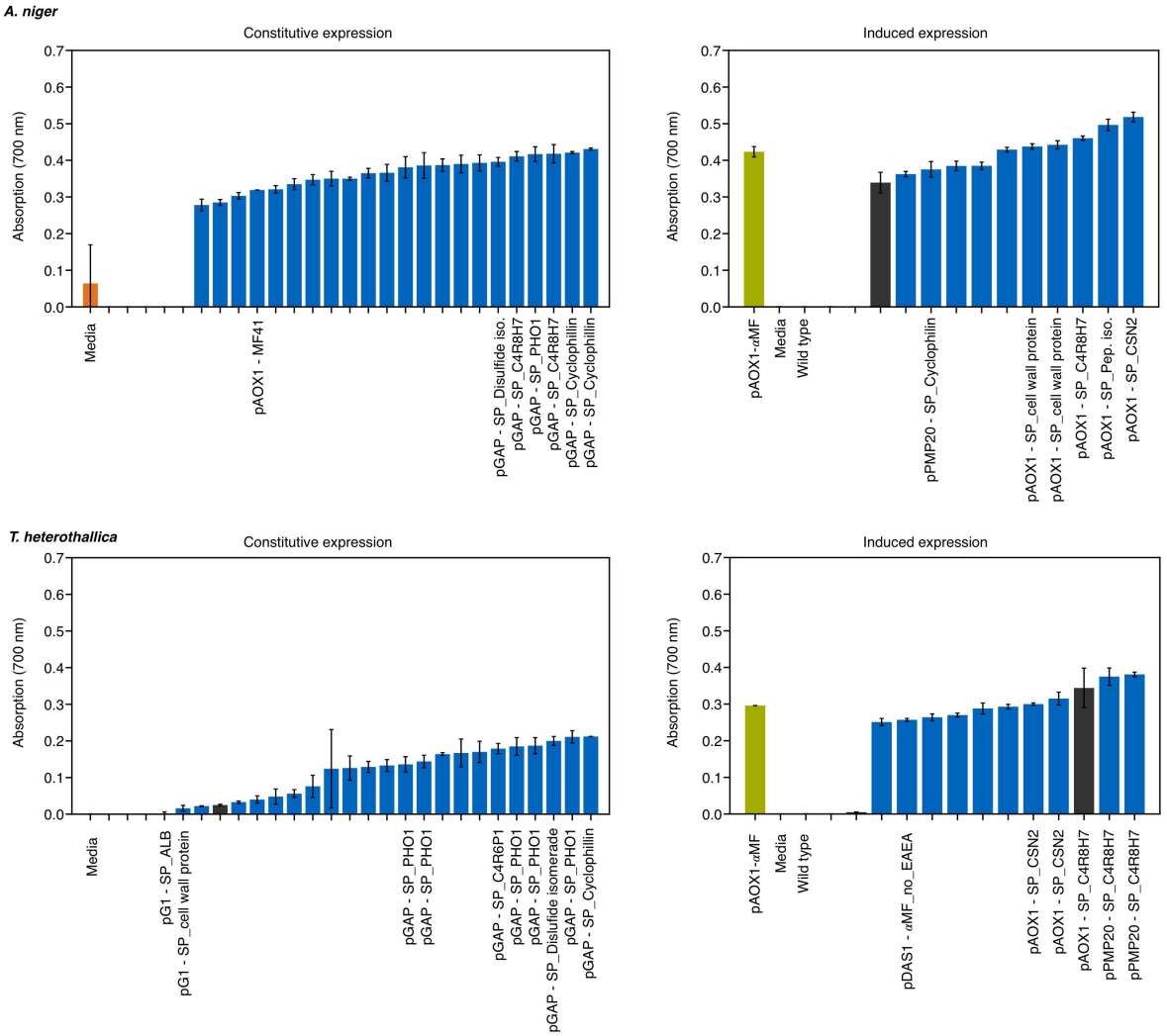

FIGURE 4.11. Re-screening of *A. niger* and *T. heterothallica* phytases. For each phytase, a re-screening of the constitutive and induced expression constructs was performed. Bars represent the mean absorption from the phytase assay of three biological replicates, subtracted by the mean absorption of three medium samples. Error bars denote ± 1 standard deviation. For the samples that were characterised via sequencing, the genetic composition of promoter and secretion tag are indicated. The green bars describe the respective *A. niger* and *T. heterothallica* positive controls (Figure 4.7 B), the medium sample is highlighted in orange and strains not showing activity in the first screening are highlighted with dark grey bars, the selected hits from the first screening are presented with blue bars.

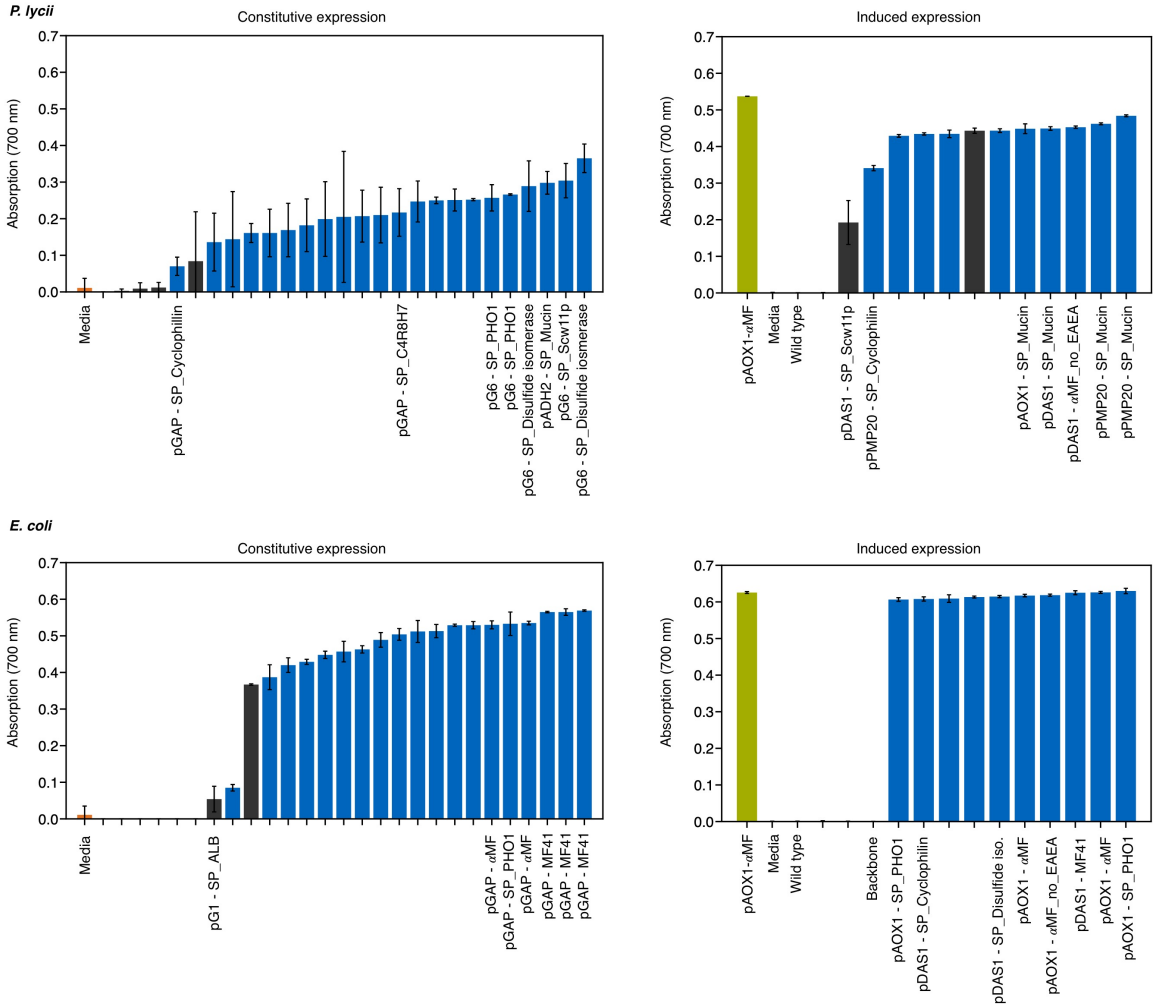

FIGURE 4.12. Re-screening of *P. lycii* and *E. coli* phytases. A re-screening of the constitutive and induced expression constructs was performed. Bars represent the mean absorption from the phytase assay of three biological replicates, subtracted by the mean absorption of three medium samples. Error bars denote ± 1 standard deviation. For the samples that were characterised via sequencing, the genetic composition of promoter and secretion tag are indicated. Colour coordination of bars is done as in Figure 4.11. The respective positive controls for induced *P. lycii* and *E. coli* phytase expression are presented in green, the genetic composition is given in Figure 4.7 B.

Finally, a selection of 16 different phytase secreting strains were selected and the reference strains for each phytase were expressed in a shake flask format. The expression was done according to the same scheme as for the screening, using a pre-culture to inoculate the expression-culture. The expression was performed twice and each measurement was performed in triplicate. Figure 4.13 shows the average of these measurements. The phytase activity was determined in phytase units (U), i.e. the amount of phytase that releases 1 µmol of inorganic phosphate per minute at 37 °C, pH 5.0 from sodium phytate [82, 118, 123]. The inorganic phosphate concentration was determined from a calibration curve using potassium phosphate. Additionally, the OD_{600} of the culture and the protein concentration of the supernatant was determined. The protein concentration was measured according to the method of Bradford [37] and for the calibration curve bovine serum albumin was used. A fraction of supernatant was concentrated 15X using 'centrifugal filters' (Section 2.4.5) and loaded on an SDS-PAGE for protein separation.

The determined enzyme activity in the screening could be reproduced for the shake flask experiment. However, for the samples pAOX1–SP_C4R8H7–Phytase_T_heterothallica the values of the re-screening could not be reproduced. This was surprising as the tag SP_C4R8H7 was also found for the best hit of this enzyme and should be suitable for the secretion of this enzyme. Surprisingly, the OD_{600} is overall low, with strains grown in BMMY-PP having an average OD_{600} of 3 and in BMGY-PP of 5. The average OD_{600} for expression in the deep-well plate format was 3-4 and 5-9 for BMMY-PP and BMGY-PP, respectively. This may explain the comparatively low activity determined for secreted phytases in the supernatant but certainly highlights the importance of optimising the expression conditions to make use of the high cell density fermentation capacity of *P. pastoris*.

Akbarzadeh *et al.* had previously optimised the induced heterologous *E. coli* phytase expression (codon optimised) in *P. pastoris* and reported 237.2 U/mL activity, when cultivation was performed at pH 6, 1 % methanol, 20 °C for 3 days [6]. Han *et al.* reported 25–65 U/ml for the secretion of the *A. niger* phytase (phyA) in *P. pastoris* [123] and Rodriguez *et al.* 128 U/mL for a *A. fumigatus* phytase [260]. Here, cultivation was done for 2 days at 30 °C to prove the first scale up from a deep-well plate format to the shake flak format. Except for one strain, the measured phytase activity could be reproduced in this varied setting and for each phytase, a combination of promoter and tag could be found that is better than the most widely used pAOX1-αMF combination.

The novel tags are of particular interest, as the highest secreted phytase activity measured in the supernatant was found from constitutively expressed *A. niger* and *T. heterothallica* phytases with SP_Cyclophilin, and induced expression of *T. heterothallica* phytase with SP_C4R8H7. The induced expression of *P. lycii* phytase resulted in the highest measured activity with SP_Mucin. For the *E. coli* phytase, the established αMF or related tags (αMF_no_EAEA or MF41) were the most suitable for this phytase.

The protein amount determined from the supernatant shows greater levels for strains cultivated in BMMY-PP medium with methanol than in the BMGY-PP medium, without methanol, indicating either secretion of other proteins or as seen in Chapter 3, a higher degree of cell lysis. The SDS-PAGE analysis clearly shows the presence of many proteins other than the expected phytases. The *A. niger* phytase has an expected protein size of 85 kDa and is highly glycosylated [226], as the sequence contains ten N-linked glycosylation recognition sites (Table 4.3). The SDS-PAGE also indicates the high degree of glycosylation as the phytases are visible as a smear from 75 kDa to 120 kDA. Despite the fact, that an identical protein is being expressed, the protein size varies between the five different *A. niger* phytases, indicating a potential influence of the secretion tag to the degree of glycosylation. The *P. lycii* phytase also has ten N-linked glycosylation recognition sites and an expected protein size varying from 65 kDA to 120 kDA [324]. The varying size for this phytases seems even more pronounced as the SDS-PAGE does not show a strong protein band that can be assigned to this phytase. The *T. heterothallica* and *E. coli* phytase have each four N-glycosylation recognition sites. The lower degree of glycosylation for this phytases is indicated by the distinct protein bands for each phytase. The *T. heterothallica* phytase has an expected protein size of about 70 kDA [230] which could be confirmed in the SDS-PAGE. The *E. coli* phytase has an expected protein size of 42 kDA [226], but the SDS-PAGE shows it to be a bit bigger as a strong protein band is at 50 kDa. Surprisingly, the strain pGAP-MF41-*E. coli*_phytase does not have a protein band at this size, despite resulting in similar measured protein activity.

For future applications, the purity here is not a major concern as phytases are mainly used as a feed additive. These strains could now be used for fermentation studies like those by Parashar *et al.* to perform high cell density cultivation of *P. pastoris* and reach phytase expressions of 780 U/mL [230].

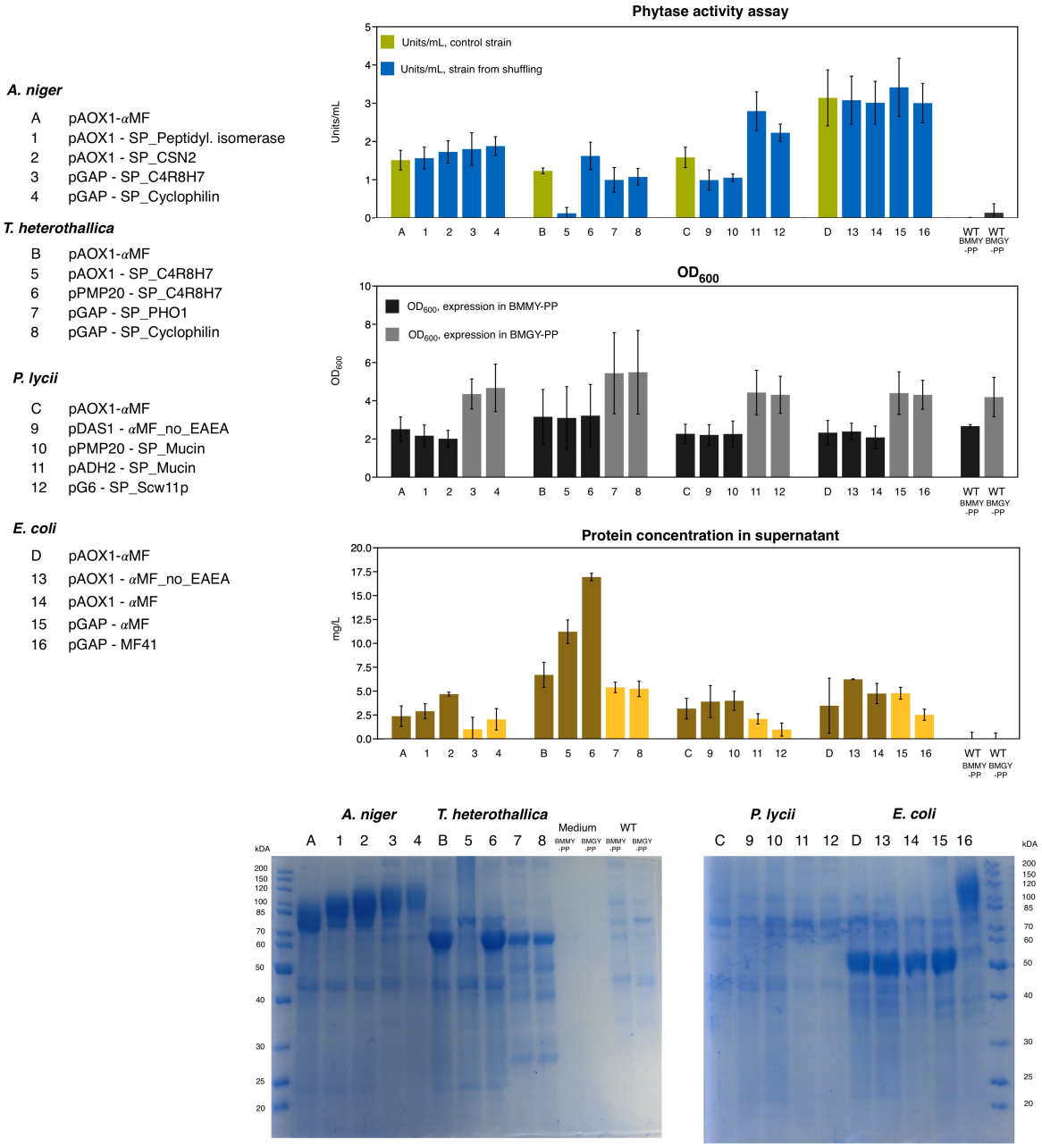

FIGURE 4.13. Expression of phytases in shake flask format. The reference phytase expression strains each having pAOX1-αMF as well as the best two strains of each screening (induced and constitutive) were used for phytase expression in 75 mL medium in shake flasks. The shake flask expression was done twice and values presented are the average absorption of the both expressions measured in triplicate. The error bars denote ± 1 standard deviation. For each expression the phytase activity, OD_{600}, and the protein concentration of the supernatant was measured. The supernatant was concentrated 15 x for the SDS-PAGE analysis.

4.3 Unspecific Peroxygenases

Unspecific peroxygenases (UPOs) are heme-containing oxidoreductases secreted by fungi that catalyse the oxyfunctionalisation of a variety of organic compounds, requiring only hydrogen peroxide (H_2O_2) as a cofactor [207, 235]. Naturally, UPOs belong to the ligninolytic enzyme consortium involved in the biodegradation of lignocelluloses and lignins or the conversion of humic substances, organohalogens, or other complex compounds [131]. Due to their exceptional and diverse catalytic properties, many applications of this biocatalyst are of interest to substitute harsh organic and pharmaceutical syntheses for more environmentally friendly and sustainable processes [197]. UPOs are also known as aromatic peroxygenases (APOs) and form a separate peroxidase subclass (EC 1.11.2.1). They perform mono-peroxygenase reactions to transfer an oxygen atom from hydrogen peroxide to a variety of organic substrates such as aromatics, linear and cyclic alkanes/alkenes, fatty acids, and others [130]. This substrate spectrum and the product pattern of UPOs are very similar to those of cytochrome P450 monooxygenases (P450s) and classic heme peroxygenases. P450s are intracellular proteins with low stability, moderate turnover numbers, and requirements for expensive NAD(P)H as electron donor and auxiliary flavin-reductases for electron transfer to O_2 [23]. UPOs are a "*self-sufficient*" alternative [196] that only require H_2O_2 [130].

In 2004, the first natural peroxygenase (*Aae*UPO) was isolated from the basidiomycete *Agrocybe aegerita* [298] and ever since *Aae*UPO has been used for the conversion of over 100 substrates with high regio- and enantio-selectivity [149, 159, 235]. Two other UPOs have been identified and characterised one from *Coprinellus radians* (*Cra*UPO) [12] and another from *Marasmius rotula* (*Mro*UPO) [114]. Furthermore, one *Coprinopsis cinereal* peroxygenase was heterologous expressed by Novozymes A/S (Bagsvaerd, Denmark) [23, 191]. The three characterised UPOs (*Aae*UPO, *Cra*UPO, *Mro*UPO) all contain a heme moiety as the prosthetic group, are highly glycosylated, and have a molecular mass of 29 – 46 kDA. Within the kingdom of fungi, about one thousand putative UPOs have been predicted from BLAST searches in GenBank and public sequence databases [130, 131]. These sequences are mainly from the basidiomycota and ascomycota phylum, fewer representatives are also form mucoromycotina, chytridiomycota, glomeromycota and the oomycota (fungus-like eukaryotic microorganisms), but no homologous sequence was found in ascomycoteos yeast [131]. UPOs can be classified as short or long UPOs depending on the enzyme structure. Short UPOs have an average mass of 29 kDA and are distributed amongst all fungal phyla, like the *Mro*UPO. Their active site contains a conserved histidine as charge stabilizer and an overall conserved PCP-EHD-E motif. Long UPOs, like the *Aae*UPO and *Cra*UPO, have an average mass of 44 kDA and are solely found in ascomycetes and basidiomycetes. These have an internal disulphide bridge, the charge stabilizing amino acid is an arginine, and the overall conserved motif is PCP-EGD-R-E. However, variations are possible when not part of the *Aae*UPO subgroup [131, 150].

The main bottleneck for industrial applications of UPOs are the low amounts of this enzyme being secreted from their natural host and laborious cultivation conditions of these hosts. To overcome this, Molina-Espeja *et al.* performed directed evolution experiments of the *Aae*UPO for heterologous expression in *S. cerevisiae* introducing 9 mutations to increase activity by 3250-fold [208]. Later they used this engineered UPO variant for expression in *P. pastoris*, with a fed-batch fermentation reaching 217 mg/L product yield [209].

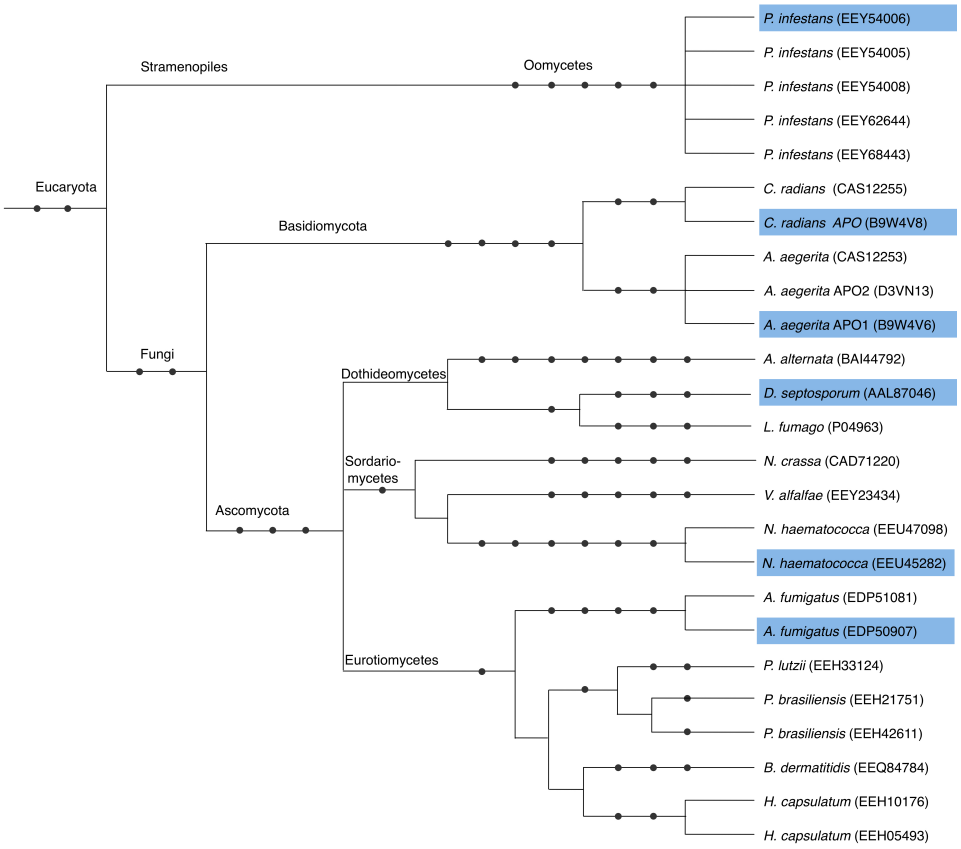

FIGURE 4.14. Phylogenetic tree of UPOs. A selection of UPO sequences provided by Ullrich *et al.* [131] was studied using phyloT (https://phylot.biobyte.de). UPOs selected for the *P. pastoris* screening are highlighted in colour.

The aim of the work presented here is to apply the improved *Aae*UPO (*Aae*UPOeng) for UPO expression as well as to make use of the great variety of predicted UPOs, and utilise them for the Golden Gate shuffling approach of the PTK. Molina-Espeja *et al.* [209] kindly provided the *Aae*UPOeng expressing *P. pastoris* strain, which was used for the development and validation of the screening. Furthermore, a *P. pastoris* codon optimised version of the *Aae*UPOeng (*Aae*UPOeng_co) was designed as part 3b type to be compatible with the PTK and the original secretion signal peptide (SP_Aae_UPO), as well as the engineered secretion peptide (SP_Aae_UPOeng) were designed as type 3a parts. To expand the screening, five additional UPO sequences were designed as type 3b parts: the previously described, but not heterologously expressed *C. radians Cra*UPO, as well as four other putative UPOs from *P. infestans*, *D. septosporum*, *N. haematococca* and *A. fumigatus* (Table 4.6).

FIGURE 4.15. Sequence alignment of selected UPOs. A section of the UPO amino acid sequences is given, presenting the conserved amino acid residues, highlighted in blue. Engineered amino acid position form *Aae*UPO and *Aae*UPOeng are framed.

All sequences have been codon optimised for *P. pastoris*, the BsaI and BsmBI sites were removed when necessary, annotated introns were removed as Molina-Espeja *et al.* [208] have performed for *Aae*UPOeng, and the annotated tags of *P. infestans* and *D. septosporum* have been removed as well. For the selection of the four predicted UPOs, the detailed BLAST research of heme-thiolate peroxidases e.g. chloroperoxidases (CPOs) and aromatic peroxidases (APOs) kindly provided by Ullrich *et al.* [131] was studied. From the 98 provided sequences, 25 sequences were unspecific peroxygenases with sufficiently annotated sequence information at the National Center for Biotechnology Information (NCBI) GenBank and the Universal Protein Resource (UniProt). A phylogenetic tree was generated (Figure 4.14) and besides the UPOs form *A. aegerita* and *C. radians*, the additional UPOs were selected such as to cover a broad species variety. Alignment of the selected sequences (Figure 4.15) highlights the consensus PCP motif across all these UPOs and the EGD amino acid motif of long UPOs for the original as well as the engineered *A. aegerita* UPO, the *C. radians* UPO and the *N. haematococca* UPO. The *D. septosporum* long UPO has a varying ESD amino acid motif. For the short UPOs of *P. infestans* and *A. fumigatus*, the conserved amino acids at the active site are DHD and EHD respectively.

TABLE 4.6. Summary of unspecific peroxygenases.

Organism	Protein (Gene)	Taxa	Enzyme structure	NCBI GenBank	Genomic sequence	Protein UniProt	Amino acids
Phytophthora infestans T30-4	Hypothetical protein (PITG_07698)	Oomycetes	short	EEY54006.1 (Genom: DSO28128)	792 bp, no introns	D0N8X1	236 (+ 27 signal peptide)
Agrocybe aegerita	Aromatic peroxygenase (*Aae*UPO)	Basidiomycota	long	FM872457	1374 bp, 5 Introns	B9W4V6	328 (+ 43 signal peptide)
Coprinellus radians	Aromatic peroxygenase (*Cra*UPO)	Basidiomycota	long	FM872459	687 bp, no introns	B9W4V8	261 (no tag found)
Dothistroma septosporum	Putative oxidase (dotB)	Ascomycota	long	AAL87046.1, AF448056	1245 bp, no introns	Q8TFD4	396 (+ 18 signal peptide)
Nectria haematococca mpVI 77-13-4	Hypothetical protein (NECHADRAFT _80813)	Ascomycota	long	EEU45282.1	1734 bp, 4 Introns	C7YSQ2	471 (no tag found)
Aspergillus fumigatus A1163	Peroxidase, putative (AFUB_050830)	Ascomycota	short	EDP51081.1, (Genom: DS499597)	828 bp, no introns	B0Y2B1	275 (no tag found)

4.3.1 Development and Validation of an UPO Screening

A screening of induced *P. pastoris* unspecific peroxygenases expression was previously performed by Molina-Espeja *et al.* [209] and these screening conditions offered an ideal starting point for the intended screening. Following the established overall scheme of agar backup, pre-culture, expression-culture, enzyme activity determination and identification of hits (Figure 4.16), enzyme specific parameters such as the medium and enzyme assay had to be assessed.

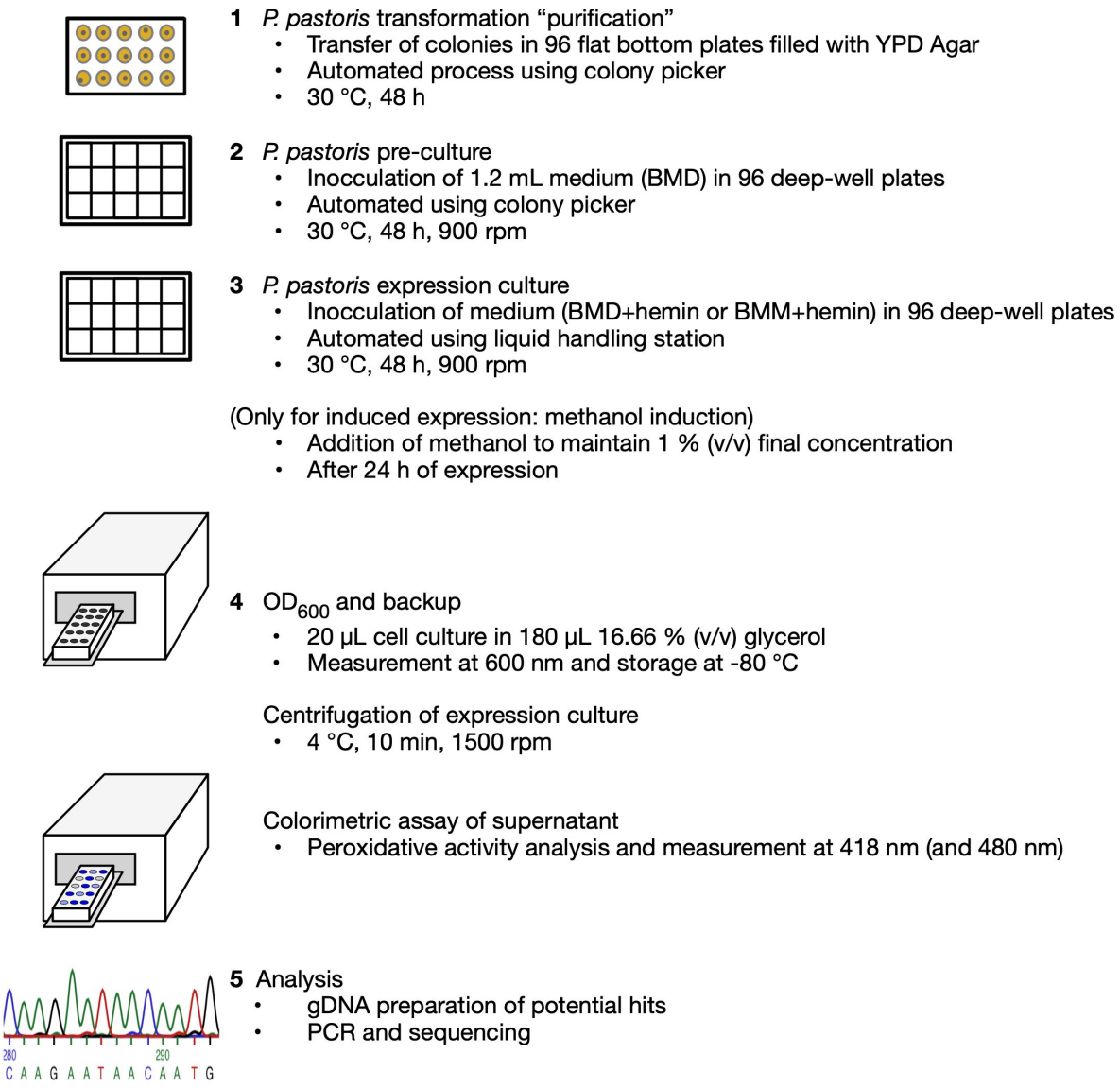

FIGURE 4.16. Workflow of *P. pastoris* screening for UPO activity.

Expression of active heme proteins such as the UPOs require successful incorporation of the heme group. Krainer *et al.* [164] have studied strategies to increase the yield of actively expressed heme peroxidase in *P. pastoris* and described the supplementation of the medium with hemine to be most effective. In contrast to the studies with supplementation of 5-aminolevulinic acid or the co-overexpression for genes of the endogenous heme biosynthesis pathway, recombinant enzyme production was not improved. Therefore, the UPO screening was performed using a final concentration of 10 µM hemin [164] and not as described by Molina-Espeja *et al.* [209] 20 mg/L haemoglobin.

To screen the UPO library, a suitable high-throughput assay is necessary. Previous UPO screenings did not screen for their mono-oxygenase activity but their side activity, which is the peroxidase activity [209]. For the peroxidase reaction of UPOs, hydrogen peroxide (H_2O_2) or other hydroperoxides are used as co-substrate to catalyse one-electron oxidations of substrates [131]. In contrast to the peroxygenase reaction, the peroxidase reaction can be followed with a colorimetric assay, where the substrate 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) and the co-substrate H_2O_2 is used. To develop the suitable assay conditions, the *P. pastoris* strain expressing *Aae*UPOeng provided by Molina-Espeja *et al.* [209] was cultivated as described in Chapter 2.2.3 and the supernatant was used to define the assay conditions. Too high H_2O_2 concentrations may inhibit the enzyme activity [235], and therefore a variety of H_2O_2 concentrations were tested (Figure 4.17 A). A concentration of 0.2 mM H_2O_2 resulted in optimal conditions. Subsequently, various *Aae*UPOeng enzyme concentrations were studied for 0.2 mM H_2O_2 and a suitable correlation of absorption after 12.5 minutes can be assumed for further screenings.

The Z-factor [338] (Section 4.2.1) was determined for the absorption after 14.5 minutes (Figure 4.17 E) and 45 minutes with 0.371 and 0.708 respectively (Figure 4.17 C and D). A Z-factor of 0.371 allows separation of sample and controls, but is not ideal. Meanwhile, the Z-factor of 0.708 allows better separation of sample and control, however, here high absorption values may not be in the required linear range of the photometer. Therefore, the screening was performed measuring the absorption increase for 15 minutes to allow correlation of high expressing samples and the final absorption was measured after 45 minutes, to identify samples expressing less enzyme in order to be more sensitive.

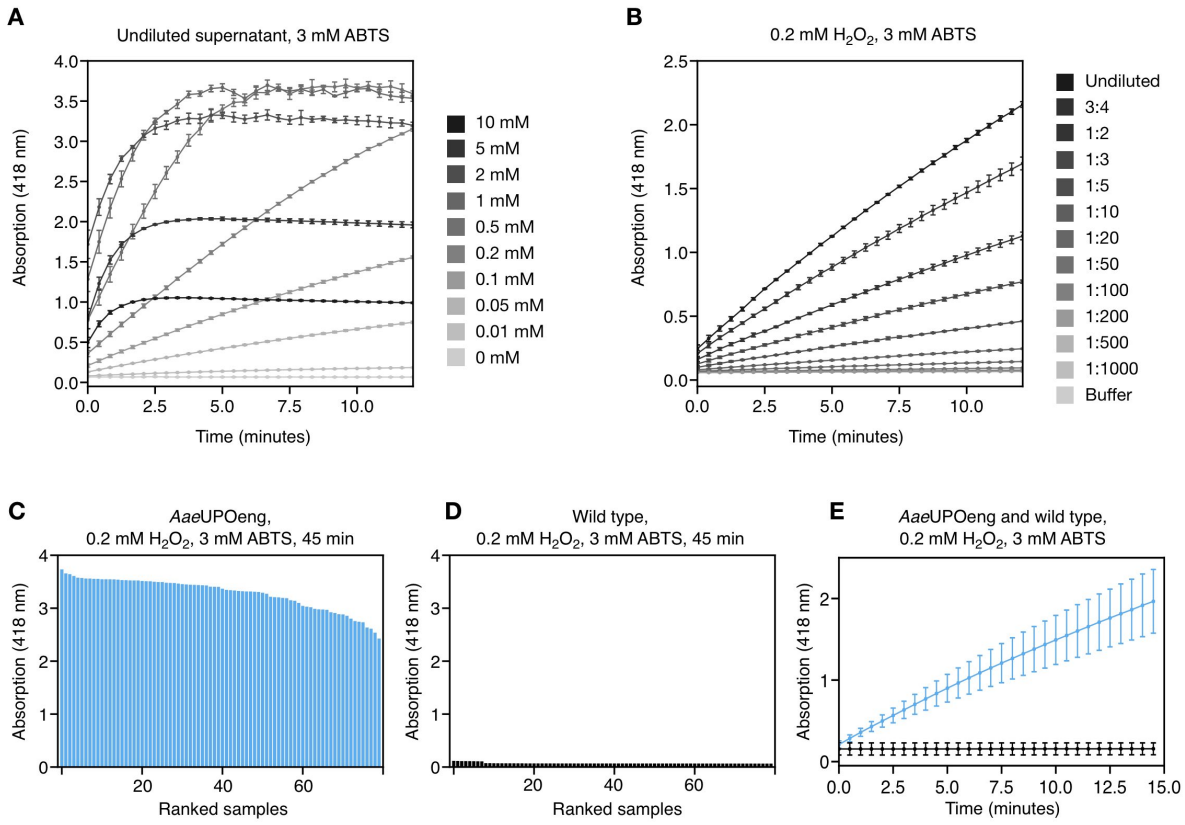

FIGURE 4.17. UPO assay development. *Aae*UPOeng enzyme activity. **A** Activity was studied for differing H_2O_2 concentrations and **B** for a H_2O_2 concentration of 0.2 mM for various enzyme concentrations [209]. With the assay conditions of 0.2 mM H_2O_2 and 0.3 mM ABTS a landscape of the *Aae*UPOeng [209] and the wild type was performed, (**C, D**) measuring the final absorption after 45 min and **E** the activity slope for 14.5 minutes. For **A, B** the mean of three replicates is shown and the error bar denotes ± 1 standard deviation

4.3.2 *P. pastoris* Screening of UPO Secretion

The aim of the UPO screening is to gain a better understanding of the applicability of the PTK shuffling for the development of novel enzyme production processes. The identification of novel UPOs is industrially very interesting and the developed library generation approach could provide a suitable platform for the identification of novel heterologous secretion processes. Additionally, the greatly engineered *A. aegerita* UPO will be studied to discuss the highly fruitful directed evolution approach performed by Molina-Espeja *et al.* [208] in context with the Golden Gate shuffling. First, a single library was made to shuffle all UPOs, utilizing 19 secretion tags and three promoters for induced expression, Figure 4.18. The library generation was performed as in Chapter 4.2, a detailed list of the shuffling setup can be found in Table A.11, Formula 4.5 was used to determine the expected coverage of this library. For the plasmid preparation, 2420 *E. coli* colonies were used, and it was subsequently used to make the *P. pastoris* library where 1092 colonies were screened resulting in an expected coverage of 95.83 %.

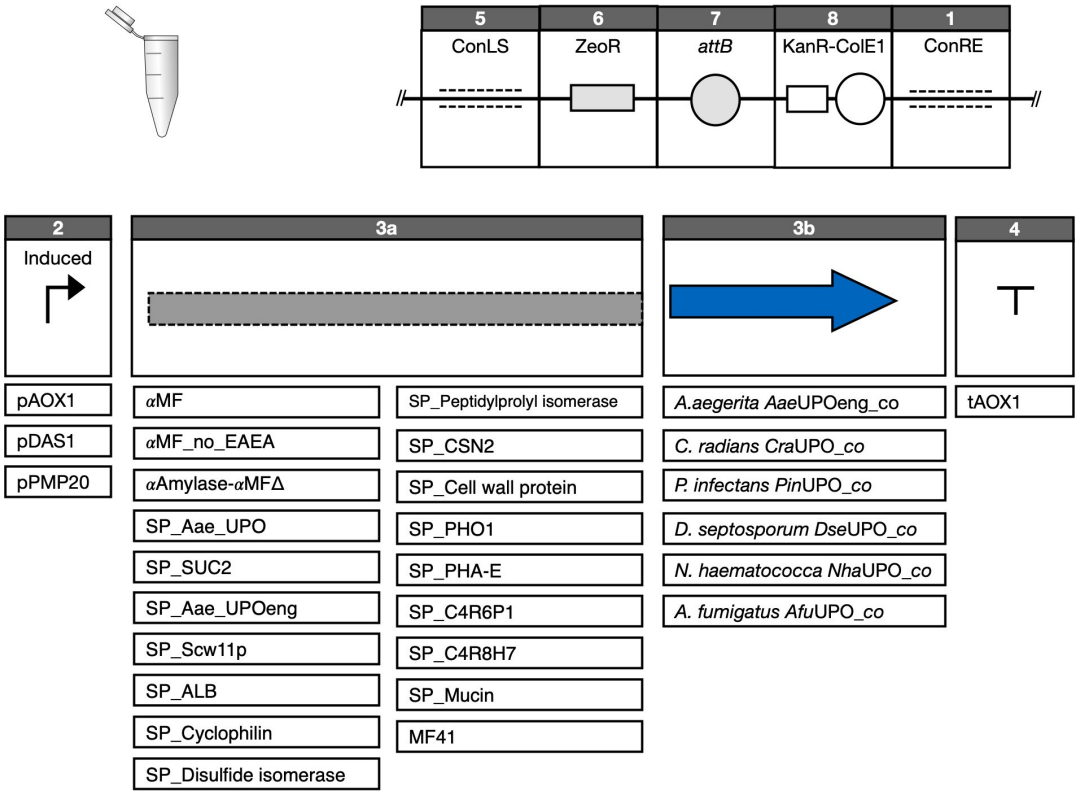

FIGURE 4.18. Parts for Golden Gate shuffling of induced UPO expression library. One library was generated for all UPOs combined.

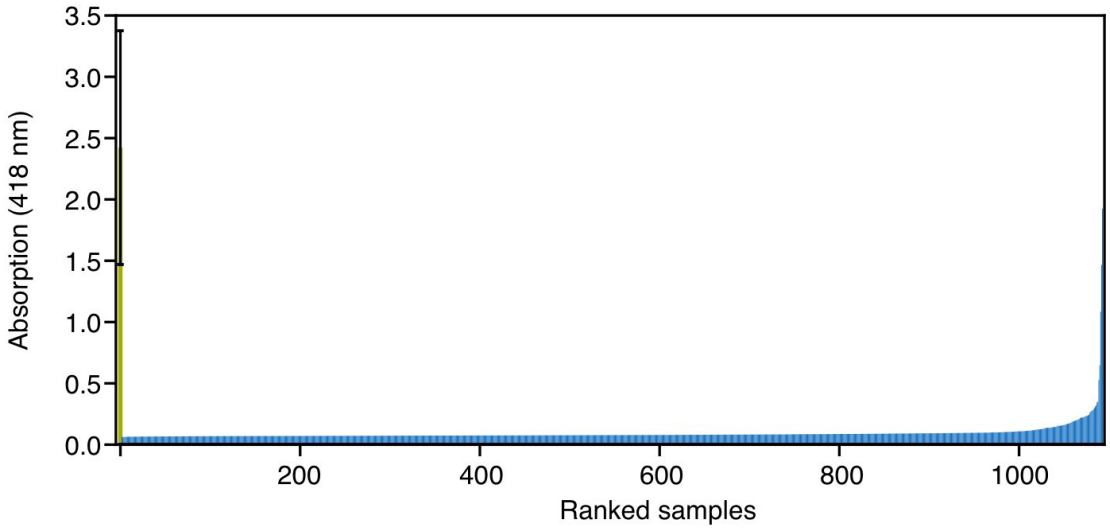

FIGURE 4.19. Screening of induced UPO expression library. 1274 *P. pastoris* variants were screened for induced UPO secretion, blue bars. Along on each plate, three replicates of the engineered *A. aegerita* UPO [208] were screened. The overall average and ± 1 standard deviation thereof are depicted as the green bar (left). For the assay 0.2 mM H_2O_2 and 0.3 mM ABTS were applied.

Figure 4.19, presents the outcome of the first screening showing the measured adsorption after 45 minutes at 418 nm for the screened samples and the average of the engineered *A. aegerita* UPO [208], screened along in triplicate on each plate. As planned for the screening development, the kinetics were measured as well, but adsorption values are so low that the evaluation can be based on the 45 minutes measurement. 1087 of the 1092 samples show only very low absorption, the remaining 5 samples result in considerable activity, but all below the engineered *A. aegerita* UPO [208].

To identify the samples showing activity, a re-screening was performed looking at 25 samples of the screening in triplicates. To be able to assess the high-performance and low-performance variants, again a kinetic for 15 minutes was measured and a final measurement after 45 minutes was performed, Figure 4.20. In Figure 4.20 C the data of the final measurement is presented and samples have been identified using sequencing. The high-performance samples are the codon optimised variants of the *A. aegerita* UPO in combination with the *A. aegerita* UPO signal peptide either in its original form or engineered. Evaluating the kinetics in Figure 4.20 B, a superior activity of the control *A. aegerita* UPO can be seen. Nevertheless, the signal peptides SP_PHO1, SP_Peptidylprolyl isomerase and SP_C4R6P1, and the novel tags SP_Cyclophillin and SP_C4R8H7 also mediate secretion, but with much lower efficiency. Two samples other than the *Aae*UPOeng_co appear in this screening, the *Afu*UPO and the *Dse*UPO, both combined with the novel tag SP_C4R8H7 and the pDAS1 and pAOX1 promoter, respectively.

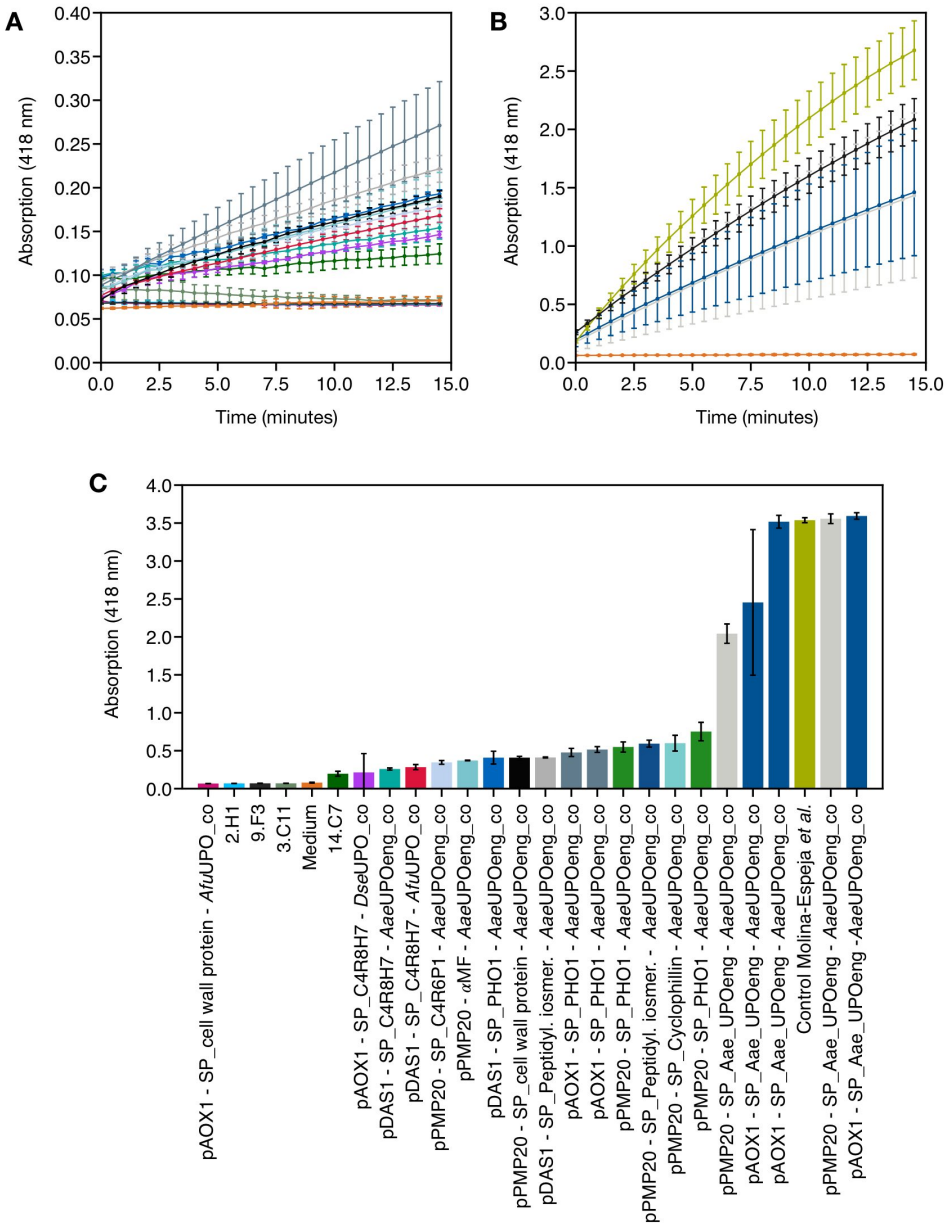

FIGURE 4.20. Re-screening of induced UPO expression library. 25 samples of the screening as well as the engineered control *A. aegerita* UPO were measured in triplicate. To meet the requirements for low- and high-performance samples, the kinetic was determined for 15 minutes, **A** and **B** , and a final measurement was done after 45 minutes, **C**. For the assay, 0.2 mM H_2O_2 and 0.3 mM ABTS were used. Samples were identified using sequencing and identical samples based on the sequencing are presented in the identical colour.

This screening shows once again that for a specific enzyme there are preferences of which tag works best. This was already seen in Chapter 3 and has been reported in the literature as well [81]. A key finding from this screening was that one screening for all UPOs is inadequate. The data is challenging to present as for low- and high-performance samples, the adsorption values are across such a large range (0.0 - 4.0), that the laboratory methods must be expanded to address these requirements. More importantly, a comparison of the highly towards the substrate ABTS engineered *Aae*UPOeng with none engineered putative UPOs is not realistic. If a novel UPO is heterologous expressed, we expect the expression level to be much lower than for *Aae*UPOeng, which would not allow their identification in a screening containing the *Aae*UPOeng. Therefore, a screening for constitutive expression was performed individually for each UPO.

A screening was performed for each UPO as depicted in Figure 4.21. A master mix containing the linearised backbone, 19 tags, 5 constitutive promoters and the tAOX1 terminator was made, separated for six individual master mixes to each added the corresponding UPO plasmid. After Golden Gate shuffling, they were used for *E. coli* transformation and at least 1500 colonies with equal size were used for plasmid preparation. These plasmid libraries were used for *P. pastoris* transformation and for each library 297 colonies were screened, resulting in an expected coverage of 95.68 %.

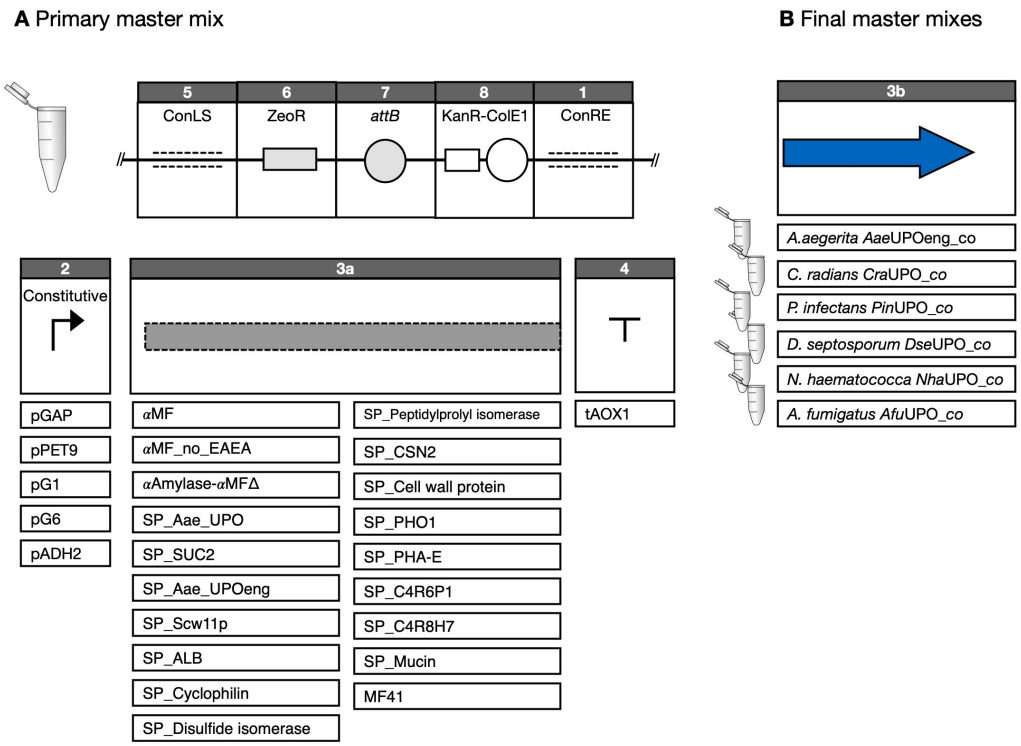

FIGURE 4.21. Parts for Golden Gate shuffling of constitutive UPO secretion library. For each UPO an individual library was made, based from one master mix containing the linearised backbone vector, 5 promoters, 19 secretion tags and one terminator.

To improve the screening sensitivity, a higher ABTS concentration was used and the assay contained final concentrations of 0.2 mM H_2O_2 and 0.75 mM ABTS. Again, enzyme kinetics were measured at 418 nm for 15 minutes, and the absorption was measured at 418 nm and 480 nm after 1 h. Additionally, the absorption was measured after 16 h of incubation at RT, to test whether the sensitivity of the screening approach can be increased further. Figure 4.22 shows the results of the screening for the secretion of constitutive *Aae*UPOeng_co expression after 1 h and 16 h. For 16 samples, the measured absorption is already out of the suitable linear range of the photometer, highlighted as overflow measurement and show in comparison to the *Aae*UPOeng higher absorption. This indicates comparatively more secreted protein. For the measurement after 16 h, values are presented in an identical order showing a similar trend, but much higher measured absorption.

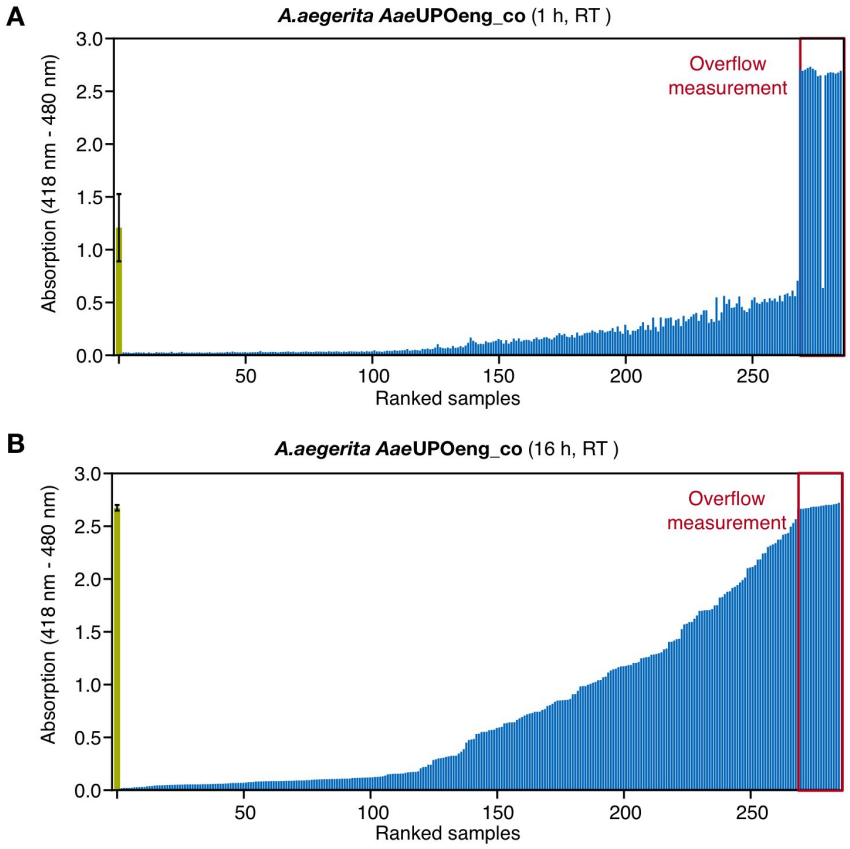

FIGURE 4.22. Screening for *Aae*UPOeng_co UPO secretion from constitutive expression. Absorption measured at 418 nm subtracted by the reference value at 480 nm are shown with incubation times of 1 h **A** and 16 h **B** , blue bars. The control strain of induced expression of the engineered *A. aegerita* [208] was screened on each plate as triplicate and the overall absorption with ± 1 standard deviation is given, green bar at left. The overall ordering of the screened samples of **A** and **B** is identical. The assay contained final concentrations of 0.2 mM H_2O_2 and 0.75 mM ABTS.

Here, a differentiation of the low-performing samples to samples with no secreted protein and no measured activity is much clearer. As the kinetics for the other UPOs measured after 1 h did not show activity, the value after 16 h was considered to be suitable to find small amounts of potential UPO activity. Figure 4.23 shows the results of the activity assay for UPOs of *C. radians*, *P. infestans*, *D. septosporum*, *A. fumigatus* and *N. haematococca*. The measurements show no measurable activity, except potentially very low activity for one *C. radians* sample and two *D. septosporum* samples. Also, *N. haematococca* does not show any activity except one, where an absorption three times higher than the average is measured.

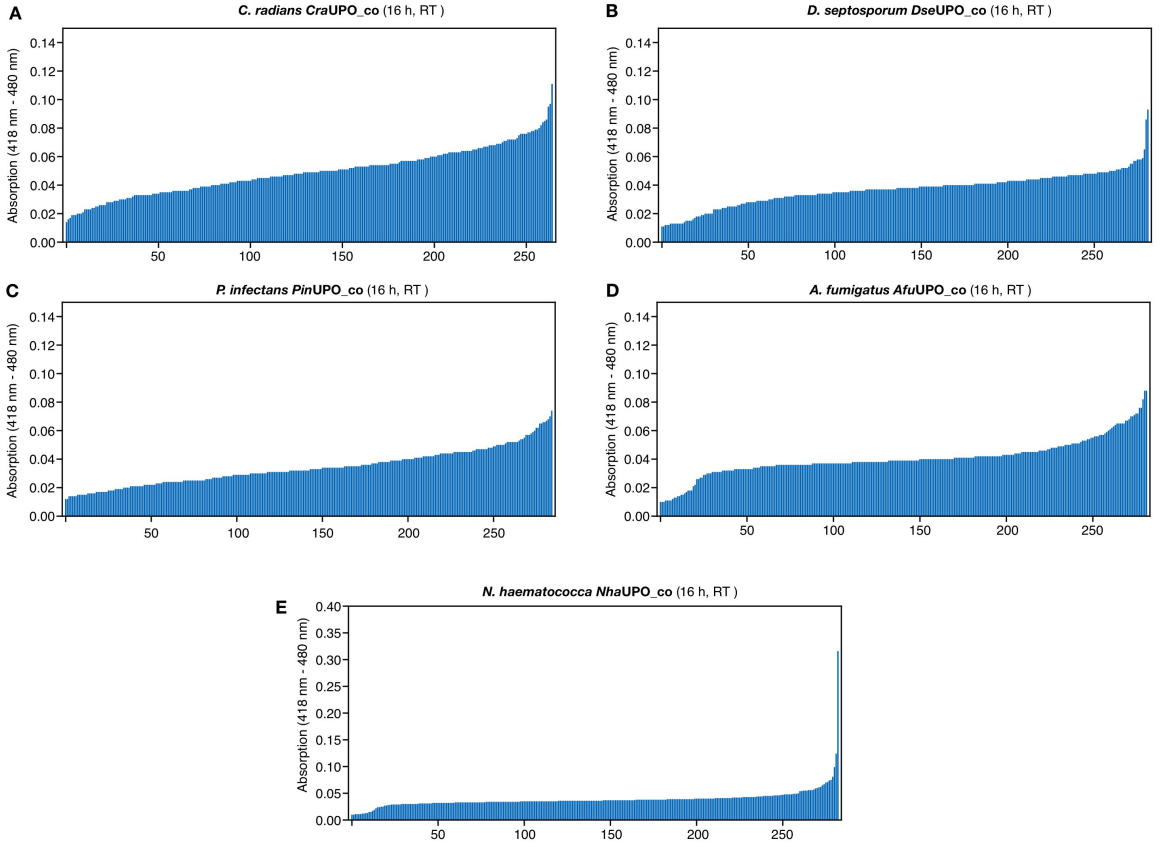

FIGURE 4.23. Screening of constitutive UPO secretion. ABTS assay evaluation for **A** *C. radians*, and the putative UPOs **B** *D. septosporum*, **C** *P. infestans*, **D** *A. fumigatus* and, **E** *N. haematococca*. Absorption at 418 nm was subtracted by absorption at 480 nm. The assay contained final concentrations of 0.2 mM H_2O_2 and 0.75 mM ABTS. The x-axis defines the number of samples screened in a ranked order.

A re-screening of 28 *Aae*UPOeng_co samples as well as a selection of 21 putative UPOs was performed, Figures 4.24 and 4.25. For the secretion of *Aae*UPOeng_co guided from constitutive promoters, three strong expression conditions could be identified that resulted in higher measured activity than the reference strain from Molina-Espeja *et al*. The reference strain was assayed along with the screened samples in triplicate, using the necessary conditions for induced expression. The beneficial conditions for *Aae*UPOeng_co are the promoter pG6 in combination with SP_AaeUPO, and the promoters pG6 and pADH2 for the engineered tag SP_AaeUPOeng. Lower levels of secretion could be detected with other tags such as, SP_Peptidylprolyl isomerase, SP_CSN2, αMF_no_EAEA, SP_C4R6P1, and the novel tag SP_Mucin.

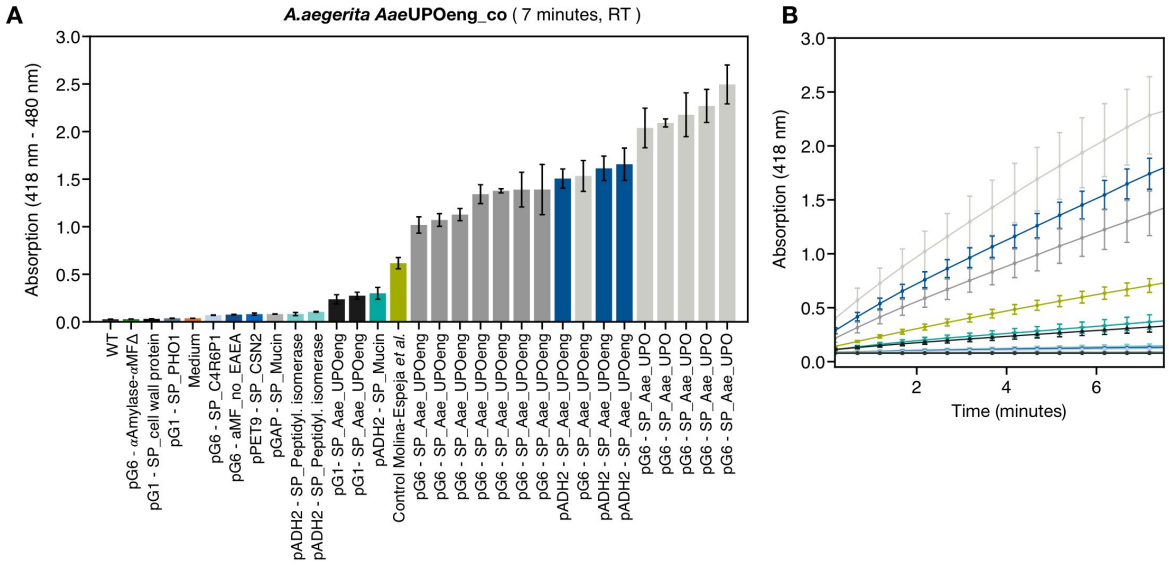

FIGURE 4.24. Re-screening for constitutive expression of *Aae*UPOeng_co. Comparison of absorption detected from the ABTS assay with 0.2 mM H_2O_2 and 0.75 mM ABTS, **A** showing the absorption after 7 minutes for 418 nm subtracted by the reference absorption at 480 nm and **B** the kinetic at 418 nm. **A** Bars of samples identified via the screening as identical are coloured using the identical colour, each bar showing the mean for a triplicate determination and the error bar denoting ± 1 standard deviation. **B** identical samples are combined, and the error bard denotes ± 1 standard deviation across all samples from the identical origin. The control from Molina-Espeja *et al.* was expressed on the screening plate too, using the respective conditions for induced expression.

The re-screening of the putative UPOs did not show any activity for any of the selected samples, Figure 4.25 A. The difficulty in screening putative UPOs stems from four main challenges:

- The selected UPOs are solely predictions, meaning the selected sequences could potentially not be real enzymes, even in the natural host.
- The conditions to express the enzyme heterologously could be not suitable.
- The detection method could not be sensitive enough, when the secreted protein amounts are too low.
- The detection method could not be specificity enough. The applied assay screened for their side activity, which is the peroxidase activity. The predicted UPOs may lack this side activity and only have peroxygenase activity, which was not considered in this screening.

To address the concern of low sensitivity, the supernatant of the three replicates of the putative UPOs was merged and concentrated 10X using 'centrifugal filters' (Section 2.4.5) and the final assay in a total reaction volume of 100 µL with final concentrations of 0.2 mM H_2O_2 and 0.75 mM ABTS was performed and evaluated after 24 h, Figure 4.25 B. Besides one sample from *P. infestans*, all samples show lower absorption values than the wild type not expressing any UPO.

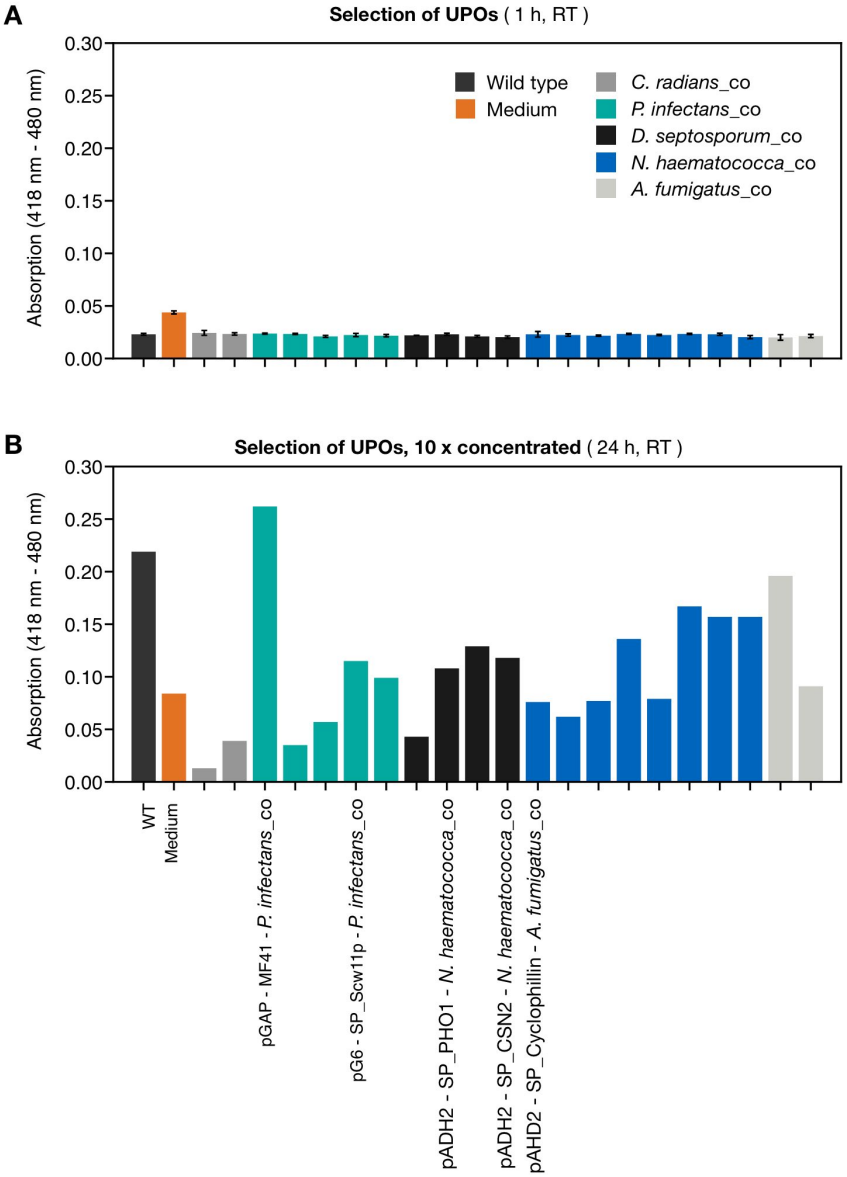

FIGURE 4.25. Re-screening for constitutive UPO expression. Comparison of absorption from putative UPOs using the ABTS assay with final concentrations of 0.2 mM H_2O_2 and 0.75 mM ABTS. **A** Triplicates for each sample were measured after 1 h of reaction time and the error bare denotes ± 1 standard. **B** The triplicate samples were combined, 10 times concentrated, and absorption was measured after 24 h of incubation at RT.

4.4 Alcohol Dehydrogenases

Alcohol dehydrogenases (ADHs, EC 1.1.1.1 and EC 1.1.1.2) are oxidoreductases that catalyse the reversible reduction of ketones and aldehydes to their corresponding alcohols. They are nicotinamide adenine dinucleotide (NAD^+)- or nicotinamide adenine dinucleotide phosphate ($NADP^+$)- dependent enzymes and are widely distributed across all domains of life, such as archaea, bacteria and eukaryotes [117, 216]. Dehydrogenases are highly chemo-, regio-, and stereo-selective, which is very interesting for the production of chiral alcohols, an important pre-cursor for the production of pharmaceuticals and fine chemicals [219]. Furthermore, ADHs convert various primary and secondary alcohols to aldehydes and ketones generating $NADP^+$ from NADPH, which can be used as a comparatively flexible cofactor recycling system [87]. ADHs have also attracted a large research interest for the production of gasoline substituents to meet our increasing demand of energy and reduce environmental pollution [20]. Production of biofuels from renewable resources is a likely step and alcohols (e.g., isobutanol, 1-butanol, 2-methyl-1-butanol, 3-methyl-1-butanol and 2-phenylethanol) could be favourable over the traditional biofuel, ethanol, due to the higher energy density, lower hygroscopicity and higher-octane numbers [19, 20]. Despite the production of these alcohols being challenging [20], metabolic engineering approaches for *E. coli* have been shown, for example at DuPont, the production of 1,3-propanediol using the native alcohol dehydrogenase YqhD [92, 141], the production of isobutanol from glucose using an engineered *E. coli* strain [19], or a cell-free production system [116]. The final step of these approaches, reduction of an aldehyde to an alcohol, is generally catalysed by an alcohol dehydrogenase.

4.4.1 Studies for an ADH screening

Here, two alcohol dehydrogenases AdhZ2 (previously named YahK, UniProt P75691) and AdhZ3 (previously named YjgB, UniProt: A1YN40), which were identified after purification from the crude extract of *E. coli* K12 by Pick *et al.* [240] were selected for heterologous expression and secretion in *P. pastoris*. These ADHs have a similar substrate specificity as the industrially applied YqhD and are interesting candidates for the production of alcohols from lignocellulosic hydrolysates [240]. For AdhZ3, an engineered enzyme was applied (AndZ3_LND) having an improved cofactor preference towards NADH [239].

When expressing a non-glycosylated, natural bacterial protein heterologous in yeast, the possibility of protein glycosylation has to be considered. Yeasts perform N-glycosylation at asparagine residues, when the consensus sequence N-X-S/T is present or O-glycosylation on the hydroxy groups of threonine or serine residues, with no recognition motif necessary [38]. This can be of great interest as glycosylations may increase the protein secretion [266], enhance the proteins thermostability [122] or are simply necessary for the functionality of therapeutic proteins [73].

TABLE 4.7. Summary of alcohol dehydrogenases.

Name	Cofactor preference	Glycosylation motifs [a]	Reference
AdhZ2_Ec	NADPH	One natural [335-337 (NRT)]	[240]
AdhZ2_DIN	NADH (shift [204-206 (DIN)]	One natural [335-337 (NRT)]	[239, 240]
AdhZ2_7476	NADPH	One natural [335-337 (NRT)], one added [74-76 (NCT)]	[223]
AdhZ3_LND	NADH (shift [199-201(LND)]	One natural [240-242 (NVS)]	[239, 240]
AdhZ3_242	NADH (shift [199-201(LND)]	None, removed [240-242(NVA)]	[223]

[a] Specific position of the mutations are given in brackets: [position (amino acid sequence)]

However, if glycosylation of a bacterial enzyme is performed without being desired, it may hamper the process of protein secretion or the enzyme functionality.

The selected ADHs each contain one naturally occurring N-glycosylation motif. For the AdhZ2, this is almost at the C-terminal end of the enzyme, and for the AdhZ3 it is in middle of the sequence, with the amino acids present on the enzymatic surface. In a previous study, an enzyme engineering approach was taken to include additional N-glycosylation motifs or remove the AdhZ3 glycosylation motif, while still maintaining enzyme activity when expressed in *E. coli* [223]. The enzymes AdhZ2_7476, an engineered version of AdhZ2_Ec, with an additional N-glycosylation motif, and AdhZ3_242 an engineered version of AdhZ3_LND with the glycosylation motif removed were used for expression studies in *P. pastoris*. Additionally, AdhZ2_DIN was used along with an engineered version of AdhZ2_Ec with a changed cofactor preference towards NADH [239].

To study these enzymes, part plasmids compatible with the PTK have been designed for each ADH, in the format of a 3b part. Each ADH was already cloned with an N-terminal hexahistidine tag (His-tag) suitable for protein purifications when expressed in *E. coli*, which was kept to the designed part (His-Adh). As stated previously for the phytases (Section 4.2), *P. pastoris* reference strains were built to establish a sensitive screening. For each ADH, three reference strains for constitutive expression (pGAP) in combination with each of the three signal peptides αMF, αMF_noEAEA or SA-αMFΔ (Figure 4.26 A) were built. Test expressions for each strain in either the complex medium BMGY or the minimal medium BMD were performed. The enzyme activity of ADH was measured using the substrate butyraldehyde and the cofactor conversion was observed at 340 nm (Chapter 2.4.6). Unfortunately, no activity could be determined. As for the UPOs, the main question to address was whether any protein was secreted or not, and whether the assay was sensitive enough to detect the low amount of protein being secreted.

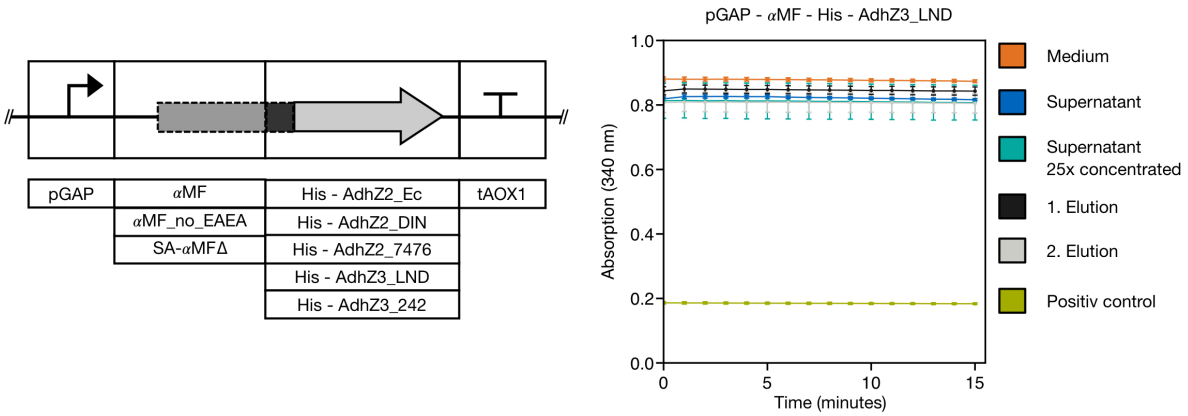

FIGURE 4.26. *P. pastoris* strains for AdhZ2 and AdhZ3 expression and assay after affinity chromatography. **A** For each ADH, three reference strains were built for constitutive expression using three different secretion tags. **B** Enzyme activity was measured following the cofactor conversion (NADH) and 340 nm. As positive control AdhZ3_LND expressed in *E. coli* with a concentration of 629 µg/mL was applied. The expression supernatant, the supernatant 25 x concentrated and the two first elution fraction from the desalting column after affinity chromatography were measured. Expression was performed once, triplicates were measured, and the mean is shown, the error bar denotes ± 1 standard deviation.

To overcome the potential issue of too low protein amounts to detect, the use of the N-terminal His-tag was tried. For AdhZ3_LND, the strains pGAP - αMF and pGAP – αMF_no_EAEA were used for expression in BMGY and BMD, and the supernatant was purified using affinity chromatography and a Ni-Sepharose™ column, as well as subsequent sample-clean-up using a PD-10 desalting column. Additionally, the supernatant was concentrated 25X using 'centrifugal filters'. The ADH activity was measured once again, but no activity was detected. Figure 4.26 B shows examples for pGAP – αMF – His-AdhZ3_LND from BMGY expression.

Since no enzymatic activity could be determined even after affinity chromatography, it indicates that no ADH is being secreted. In contrast to the UPOs, in this case it is absolutely certain, that the enzyme itself is an active enzyme and the assay in general is suitable to determine ADH activity [240]. Therefore, the expression conditions chosen are not suitable. The first concern was the presence of the N-terminal His-tag to potentially interfere with the process of protein secretion. Therefore, a variaty of new 3b parts for AdhZ3 were designed without the His-tag. Subsequently, expression strains were build without any His-tag and with a C-terminal His-tag using the 4a part 6xHis_3xFlag from the YTK and a 4b part tAOX1 terminator designed in this thesis.

The enzyme activity for both strains either newly made or previously used was determined as shown in Figure 4.28. However, none of the tested strains showed any activity.

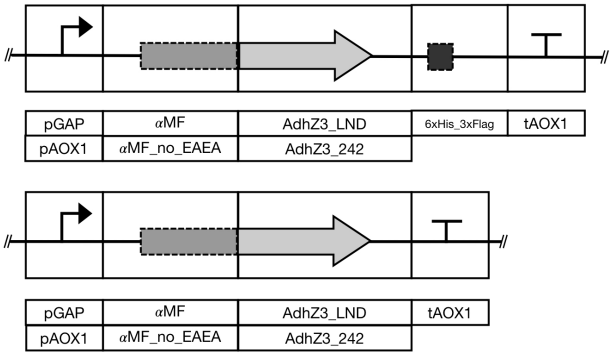

FIGURE 4.27. *P. pastoris* strains for AdhZ3 expression. For AdhZ3_LND and AdhZ3_242 in total 16 additional strains were built to remove the N-terminal His-tag and allow for induced or constitutive expression, either without a His-tag or with a C-terminal His-tag. The 6xHis_3xFlag is a 4a part from the YTK was used with the suitable 4b tAOX1. For expression without a His-tag, the 4 part terminator could be used.

Finally, the supernatant was used for a SYPRO Orange thermal shift assay, Figure 4.29 (Section 2.4.7). As *P. pastoris* is supposedly not secreting many endogenous proteins [96, 155], the supernatant should contain mainly the protein of interest. Therefore, a difference of the protein amount in the supernatant from the *P. pastoris* wild type to the strain secreting protein should be detectable, if any protein is being secreted. SYPRO Orange is a fluorescent dye used for protein detection, which is 10-100 times more sensitive than more commonly used colorimetric assays [136]. For the thermal shift assay, the unfolding of proteins is monitored during a continuous increase of the incubation temperature. The unfolding exposes hydrophobic regions of the protein and results in an increase in the measured SYPRO Orange fluorescence [185]. The determined melt curve has characteristic features for each protein and can help in the study of impure supernatant to identify ADHs. Figure 4.29 A and B present the melting curves of all AdhZ2 and AdhZ3 reference strains under the control of the constitutive promoter pGAP and being expressed in BMGY. None of the supposed ADH expression samples, the wild type, or the medium sample show any peak, whereas the ADH reference has a distinctive peak at 50-55 °C. The AdhZ3 expression samples under the control of the methanol induced pAOX1 promoter were expressed in BMMY, a complex medium containing 1 % (v/v) methanol, and all strains show a peak at around 60-65 °C. As this peak is largest for the wild type and not at the desired location of the ADH sample, this peak does also not indicate ADH secretion.

Therefore, it has to be concluded that none of the designed strains secreted alcohol dehydrogenases and the foundation necessary to establish a screening for the Golden Gate shuffling was not possible. Unfortunately, the process of protein secretion in *P. pastoris* is still poorly understood and the list of dials to tune is long. For example, a codon optimised version of the ADHs could be tried or strain engineering approaches could be used to aid secretion of the desired enzyme. As this was outside the scope of this work, the ADHs were not further investigated.

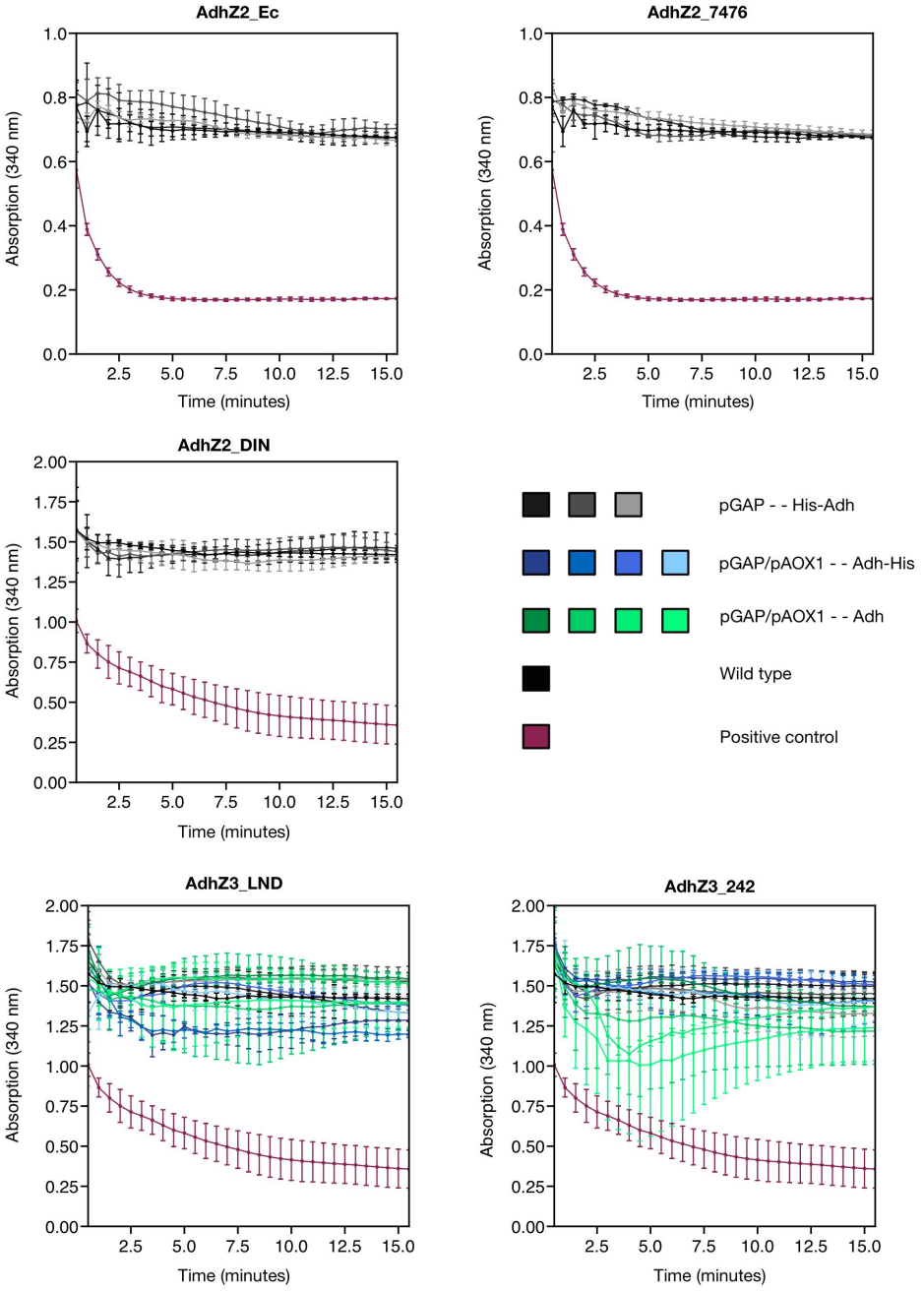

FIGURE 4.28. ADH enzyme activity assay for ADH reference strains. For all ADH reference strains, cultivation was performed in a dwp format in BMGY and the supernatant was used for ADH enzyme activity determination using the substrate butyraldehyde and the cofactor NADPH for AdhZ2_Ec and AdhZ2_7476 as well as NADH for AdHZ2_DIN, AdhZ3_LND and AdhZ3_242. Mean absorption values of three biological replicates are presented, with error bars denoting ± 1 standard deviation. The positive control AdhZ3_LND expressed using *E. coli* was also measured (315 µg/ml).

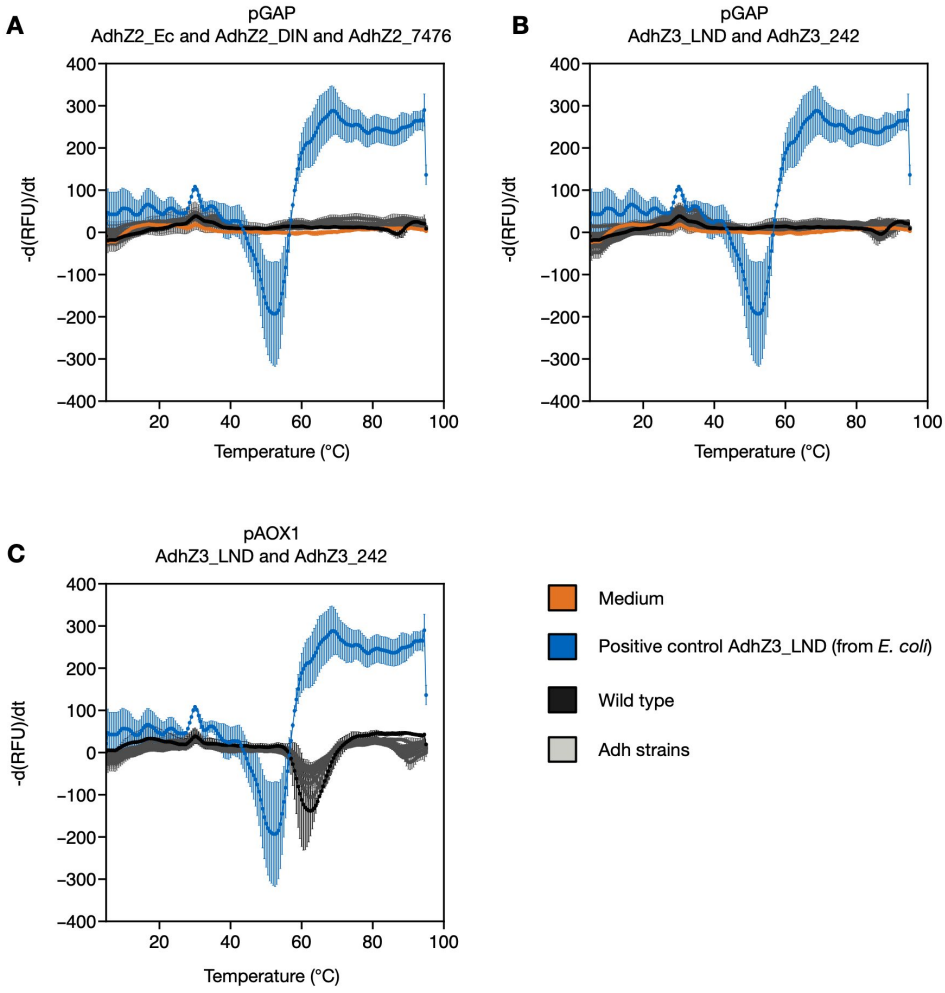

FIGURE 4.29. SYPRO Orange thermal shift assay of ADH expression supernatant. **A** and **B** Expression of all reference samples for AdhZ2 and AdhZ3 under the control of the pGAP promoter as well as the wild type in BMGY medium. **C** Expression of all AdhZ3 reference strains with the pAOX1 promoter and the wild type in BMMY. As positive control AdhZ3_LND expressed in *E. coli* (concentration in assay 144.9 µg/ml) was used. The mean fluorescent signal of three biological replicates is presented and the error bar denotes ± 1 standard deviation.

4.5 Conclusion

Application of the PTK to randomised library generation allows for the identification of suitable expression conditions at the genetic level. The prerequisite experiments performed in Chapter 3 paved way for increasing library size and simplify cloning for a desired library complexity. Industrially relevant enzymes were used to study and validate the approach. Libraries for four different phytases and six different (partly predicted) UPOs were built.

For the phytases a total number of 2184 colonies were screened and for the unspecific peroxygenases 2874 colonies were screened. The approach enabled the identification of beneficial combinations of promoters, secretion signal peptides, and the gene of interest for all studied phytases and the previously described *Aae*UPO. The screenings also highlighted the challenge of cell factory development. For predicted enzymes such as the selected UPOs or heterologous enzymes that are difficult to express in *P. pastoris* such as the ADHs, the cell factory design still remains a challenge.

5

Polyelectrolyte Microcapsules for Cell Screening

A great variety of screening technologies have been developed, however, most of these require special equipment and lack the ability to easily change the environment in which the cells are grown. Advanced polyelectrolyte microcapsules could act as smart micro-reactors for short-term cultivation and screening of individually isolated cells, thereby overcoming some of the limitations of existing approaches. Their manufacturing requires only standard laboratory equipment and their tunable permeability is ideal for enabling desired media transfer and thus cell cultivation. Here, polyelectrolyte microcapsules are considered for yeast encapsulation. First, calcium carbonate templates are studied with the goal of trapping single *P. pastoris* cells such that subsequent calcium carbonate crystallisation occurs around them. A variety of influencing parameters are studied and a design of experiments approach is used to systematically identify influencing factors. Despite the variety of parameters tested, cells were not found to be the major origin of precipitation and thus were rarely coated in calcium carbonate. Second, a layer-by-layer approach was used in an attempt to cover encapsulated cells in a permeable shell such that the calcium carbonate template could be later dissolved. We found that cells alone could be encapsulated in a polymer shell and studied their properties using flow cytometry and microscopy.

5.1 Microencapsulation using Layer-by-layer Techniques

Directed evolution approaches for generating cell factories require suitable screening methods to select fitter individuals from a mixed population. The key for any cell screening approach is the linkage of a desired phenotype to the underlying genotype that supports it. In terms of cell factories this generally involves isolating individual cells (e.g., by compartmentalisation) and measuring the critical enzymatic reactions using high-throughput optical technologies such that those cells with a good performance can be separated and their genotype identified (e.g., through re-sequencing) [186].

To meet the increasing demands of screening in terms of the numbers of cells that need to be tested, ultrahigh-throughput screening methodologies such as microchamber arrays, droplet based *in vitro* compartmentalisation, or screenings using the cells directly have been developed [300]. A more detailed description of these screening methods can be found in Chapter 1.3. High and ultrahigh-throughput screening approaches have been successfully applied for strain engineering purposes or screenings of environmental samples [300]. Despite the great scope of applications, limitations remain with every method and alternative approaches may help to address some of these bottlenecks for specific use cases.

Here, advanced polyelectrolyte microcapsules are considered as a method for yeast encapsulation and subsequent screening. In 1998, Donath *et al.* [85] first described hollow polymer shells which were made by applying the layer-by-layer (LbL) technique. In an alternating fashion, polyanions and polycations are adsorbed on a colloid template, which is dissolved once the polyelectrolyte shell is completely assembled. The key benefit of these micro- and nano-sized capsules is the potential to tune the permeability of the shell and to vary the number of layers and incorporation of functional groups to further modify their characteristics [237]. The overall polyelectrolyte adsorption and final core removal for the generation of polyelectrolyte microcapsules is summarised in Figure 5.1.

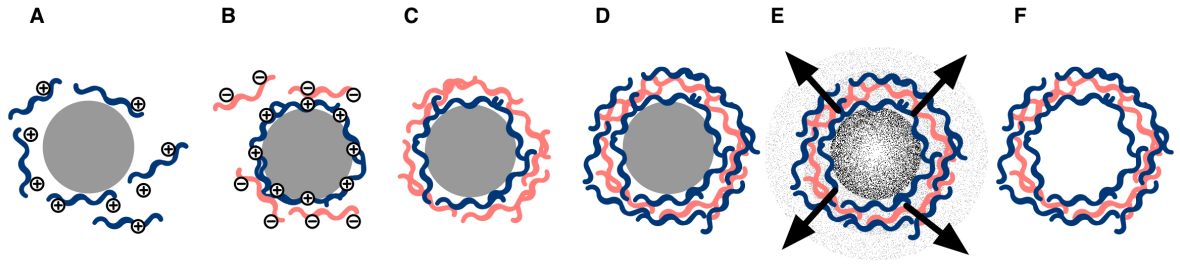

FIGURE 5.1. Layer-by-layer technique for polyelectrolyte microcapsules. **A** Cationic and **B** anionic polyelectrolytes are adsorbed in an alternating fashion on a colloid core. Once the desired number of layers has been added **D**, the core is decomposed **E** and a hollow capsule is obtained **F**. Figure adapted from Donath *et al*. [85].

Multi-layered capsules have mainly been studied as a carrier for drug delivery purposes where the size, shape, charge, and surface characteristics are of high importance [331]. In this context the aim is to create “smart” micro- and nano-containers that protect the drug from external influences and release it at the desired location upon triggering by the external surroundings [79, 142]. As polymeric capsules formed by the LbL approach are responsive to a range of stimuli such as temperature, pH and ionic strength, their use is promising as a material for other life science applications [237]. For example, polyelectrolyte capsules were built to be pH and salt sensors that could be evaluated using flow cytometry [119] or to rupture under specific conditions for pulsed drug delivery [71].

For the design of new cell factory screening technologies, polyelectrolyte microcapsules may offer interesting characteristics as these monodisperse vesicles have a tunable permeability and various functionalities can be added to their shell. For example, the polyelectrolyte can be tuned to allow diffusion of fresh medium into the capsules while incubating the entrapped cells. Furthermore, reacting agents or proteins can be co-precipitated with the template core or adsorbed onto the core to be trapped in the capsules. A reaction of the additional reactant with the cell or potentially secreted proteins will only occur once the encapsulation process is complete. The PTK for protein expression and secretion libraries described in Chapters 3 and 4 requires a suitable high-throughput screening technique. A potential capsule-based approach to assay secretion efficiency would be to tag the secreted product with part of a split-GFP [42, 43]. By encapsulating the cells with the complementary part of the split-GFP and cultivating cells in the capsules for sufficient time to allow for expression and secretion, proteins secreted would allow for a functional GFP protein to be produced causing the capsule to become fluorescent. Furthermore, the level of fluorescence is proportional to the amount of secreted protein. High-throughout fluorescence activated cell sorting (FACS) and then sequencing of capsules would then allow the genotype of effective secretion constructs to be recovered.

Polyelectrolyte capsules could offer exciting screening opportunities to study cell factories. The aim of this chapter is to validate the applicability of polyelectrolyte microcapsules for yeast cell screenings. Therefore, cells need to be integrated into the calcium carbonate template which is subsequently coated by polymer layers. Under conditions favourable for cell growth, the template has to be dissolvable to finally obtain a single cell in a polymer capsule. The desired process is summarised in Figure 5.2.

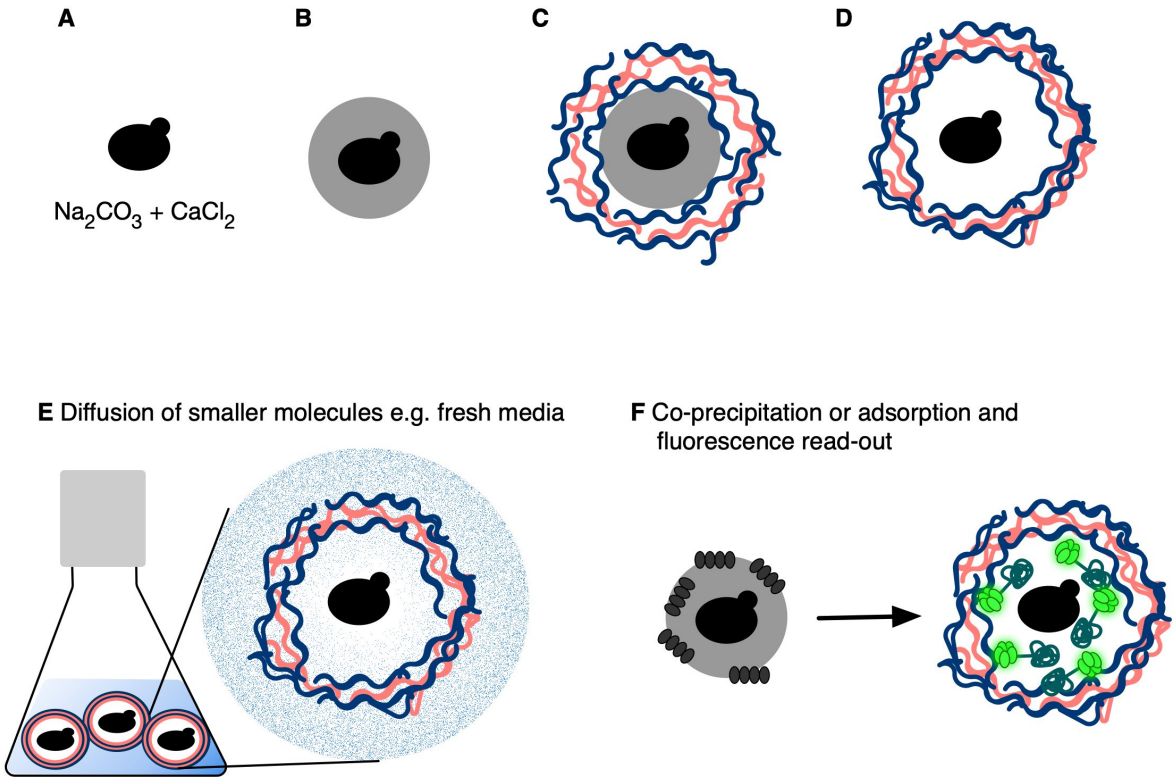

FIGURE 5.2. The process of cell polyelectrolyte encapsulation. The goal of encapsulation is to obtain a calcium carbonate template using calcium chloride and sodium carbonate within which a single yeast cell is integrated **A, B**. This template is then applied for a layer-by-layer encapsulation, **C**, after which, the template is removed **D**. Cells within the capsules are accessible for screening approaches. Polyelectrolyte capsules are interesting screening vessels as **E** the capsules can be used as small incubation vessels. If the polyelectrolyte layers are tuned correctly, medium will diffuse into the capsules, but the cells or secreted proteins will be entrapped. Furthermore, **F**, additional reactants or screening additives such as the GFP detector fragment can be added during template formation and assembly of the GFP fragment with secreted proteins linked to the GFP tag will result in a functional GFP protein and a readable fluorescent signal.

5.2 Calcium Carbonate Templates for Cell Enclosure

The generation of capsules is a multistep process where the choice of polyelectrolytes predominantly defines the properties of the final product. The choice of the core template is of equal importance for successful capsule formation [13]. Suitable templates should ideally be spherical and must be stable under the LbL process. To generate hollow capsules comprised solely of the LbL coating, the core template must be dissolvable under conditions that are not harmful to the structure [237]. To use such capsules for cell screenings, the template must also incorporate individual yeast cells and dissolving the template should also not hurt cell viability.

Commonly used templates can be grouped according to their properties and dissolution products [13]:

- Organic cores, dissolved into water-insoluble oligomers
- Ionic or molecular crystals, dissolved into small molecules or ions
- Biological cells

Organic cores such as melamine formaldehyde [102], polystyrene [76] or polylactic acid/ polylactic-co-glycolic acid are widely used templates, but unsuitable for cell encapsulation and template formation, as their removal from the capsules are not compatible with living cells due to the solutions and pH conditions used [13]. Ionic or molecular crystals are more promising, as removal occurs under either acidic or basic conditions, or even in organic solvents. This is not ideal, but the core can be removed completely without causing significant osmotic stress to the cells. Various carbon particles ($CaCO_3$, $CdCO_3$, $MnCO_3$) or silicon particles (SiO_2) have also been used as templates [76]. Lastly, biological cells have been used as templates themselves for layer-by-layer coating to either modify the outer surface of the cells, (e.g., for the generation of multi-layered tissues) [220], or to act purely as a template where the cell is finally dissolved using deproteinizing sodium hypochlorite [104].

Here the aim of the screening approach is to identify living cells, therefore, a closer look at crystal particles was taken. $CdCO_3$, $MnCO_3$ as well as silicon particles were dismissed as their respective crystal precursors are very toxic or crystallisation occurs under harsh conditions. Nevertheless, calcium carbonate ($CaCO_3$) a mineral abundantly found in nature, can be simply crystallized from supersaturated solutions and due to its industrial importance, crystallisation conditions are highly studied in geo-, bio- and material science [306].

Calcium carbonate possesses three different anhydrous polymorphs: calcite, aragonite and vaterite, which have a rhombohedral, needle-like or spheroidal morphology, respectively. The morphological characteristics strongly influence the properties of calcium carbonate, and the parameters for the precipitation are well-studied [8, 34, 211, 284]. The overall process of calcium carbonate formation at a lab-scale is shown in Figure 5.3. The parameters are discussed in more detail below, nevertheless the necessity to work with living cells excludes some approaches for homogenous crystal formation.

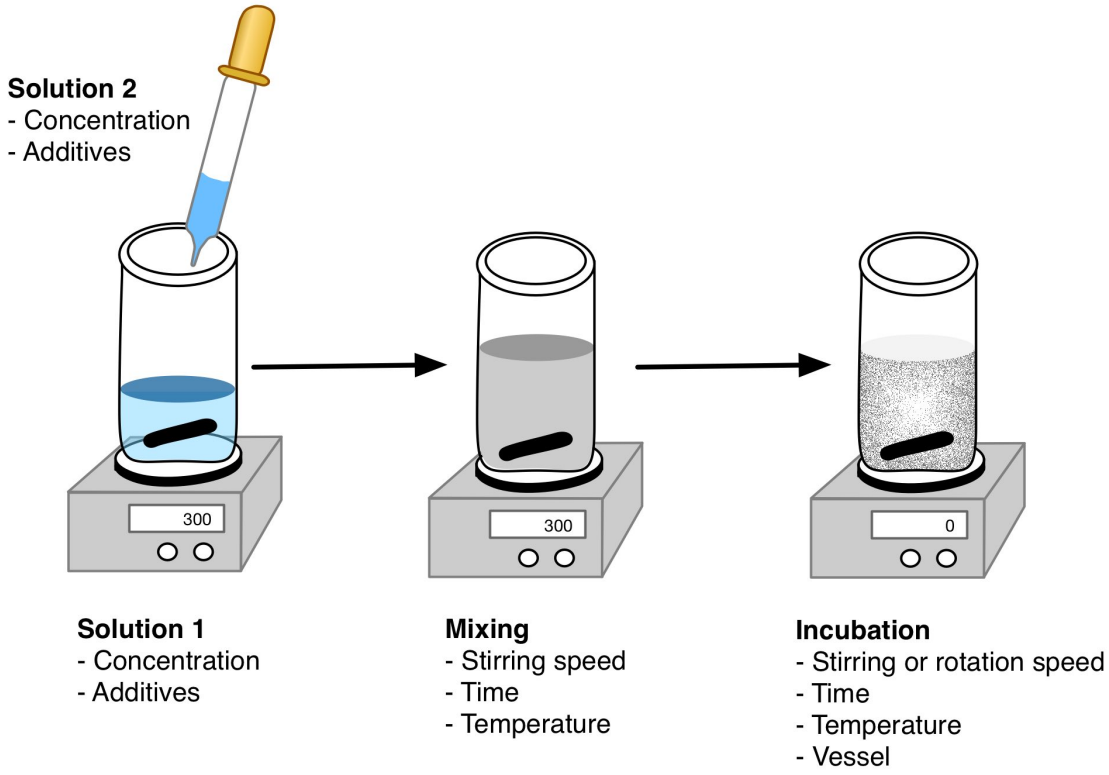

FIGURE 5.3. Lab scale calcium carbonate precipitation. For the precipitation of calcium carbonate, supersaturated solutions of calcium chloride (solution 2) and sodium carbonate (solution 1) are combined. The salt concentration and additives to the solutions will influence the precipitation behaviour. The combined solutions are vigorously mixed and subsequently incubated for crystal formation, were the operating conditions (time, temperature, reaction set-up) will influence the final crystal structure.

5.2.1 Survey of Operational Parameters for Calcium Carbonate Precipitation

Calcium carbonate crystals are mainly generated using saturated 0.33 M solutions of calcium chloride and sodium carbonate [77, 78, 98, 236, 306, 331], but higher concentrations such as 0.5 M [333] and 1 M [165] have been used. Calcium carbonate crystals may also be made using ammonium bicarbonate and calcium chloride [56], but less information on influencing parameters are available. A wide variety of additives has been tested to alter the availability of Ca^{2+} and CO_3^{2-} ions and their interaction behaviour to build morphologically different structures [8].

Previous works with additives such as polymers (polyacrylic acid, polyvinyl alcohol) or organic halides (cetrimonium bromide, polydiallyldimethylammonium chloride, cetyltrimethylammonium bromide) [8, 333] did not show significant effects at 30 °C or 25 °C, but led to a reduced aragonite formation and more monodisperse particles at 80 °C. The addition of glycol, glycerol ethers or aldehydes was studied at 80 °C and resulted in aragonite shape or aragonite and calcite shape particles, as expected without any additives [173]. The influence of polyethylene glycol was studied and did not modify the particle structure at 25 °C or 80 °C [333]. However, the addition of ethylenediaminetetraacetic acid (EDTA), which contains four carboxylic and two amine groups and is also a known complexing agent of Ca^{2+}, resulted in an apple core-type morphology at ambient conditions [8].

For the mixing of calcium chloride and sodium carbonate, varying mixing speeds (250 – 1000 rpm) [77, 78, 331] and times (20 s - 60 s) [77, 78, 98, 165, 236, 331] are published but the effect is to our knowledge not studied in detail. In contrast, the conditions for the incubation and crystal ageing are extensively studied. The incubation temperature is very important for the crystal morphology, with temperatures from 20 °C to 50 °C reported to result in calcite and vaterite crystals and temperatures higher than 50 °C resulting in aragonite particles [8, 56]. Incubation times from 0 h to 24 h of incubation are reported. Using ambient and cell-compatible conditions, calcite or vaterite templates with rhombohedral or spheroidal shape of around 2 µm - 8 µm could be generated [165]. Differing incubation times resulted in differing particle size, for example to 4 µm - 6 µm (no incubation) [165], 3.5 µm - 4 µm particles (3-4 min incubation) [77] or to 2.5 µm - 3.5 µm particles (4 min incubation) [78]. For the generation of calcium carbonate templates containing single cells, the main challenge is how to integrate the cells into spheroidal templates, while still maintaining the cell viability during the encapsulation procedure and cell screening [132]. This reduces the number of tuning parameters (e.g., no crystallization at 80 °C), but still leaves plenty of influencing parameters that can be examined. Before the integration of cells into the template, the calcium carbonate precipitation was solely studied by varying the main influencing parameters such as the incubation temperature and time, Figure 5.4.

For at least 500 templates under each condition, the diameter was determined using Fiji (plugin: automatic particle counting). As cell viability is pivotal, incubation was performed at room temperature or on ice for 10 – 90 min. Particle diameter was as previously described in the literature [165] at around 2 µm - 4 µm with outliers. Shorter incubation time resulted in slightly larger particle sizes. This preliminary test of calcium carbonate formation without cells shows that the generation of spherical templates at room temperature is possible. However, the challenge of integrating *P. pastoris* cells which have an oval shape with an average size of 2~3 x 4.5~6.8 µm [225] remains.

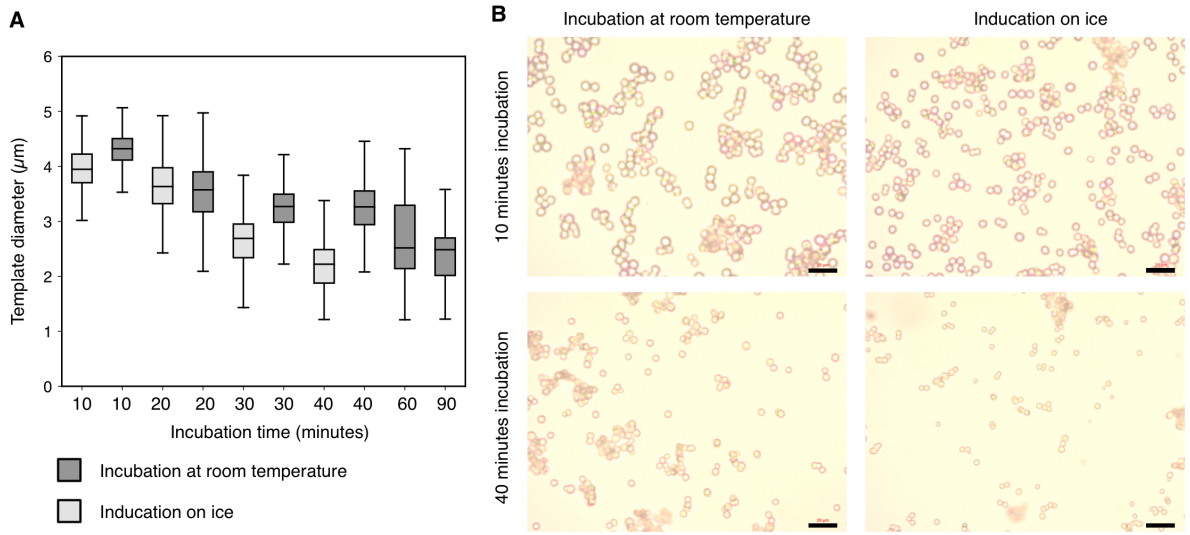

FIGURE 5.4. Calcium carbonate template formation at varying incubation times and temperatures. **A** Box plot for at least 500 templates for each depicted condition and **B** templates incubated at room temperature and on ice for 10 minutes and 40 minutes respectively. The scale bar denotes 20 µm. The precise conditions for these test were: 1. solution: 0.3 M $CaCl_2$, 2. solution: 0.3 M Na_2CO_3, stirring at: RT, 300 rpm, 30 sec, incubation: varying temperature, 0 rpm, varying time.

5.2.2 Identification of Suitable Tuning Parameters for Templates

To apply polyelectrolyte microcapsules for cell screening purposes, individual cells must be integrated into the core template. Therefore, *P. pastoris* strains either constitutively expressing yEGFP (pGAP – yEGFP, pUO_pL_008) or methanol induced expressing RFP (pAOX1 - RFP, pUO_pL_001) are used to ensure good microscopic visualisation of the cells. For cell integration, the precipitation has to start on the yeast cell surface and form a uniform layer around the whole cell. The yeast cell wall is composed of mannose, glucose polymers, and N-acetylglucosamine polymers, resulting in an overall uncharged surface and a low density of electronic charge [2].

This is not beneficial as for the nucleation the spontaneous formation of the solid phase from the supersaturated solutions a minimal electronic charge is needed to form chemical bonds [194]. Wang *et al.* [310] studied artificial biomimetic mineralisation with the aim of encapsulating *S. cerevisiae* in a calcium phosphate shell. Scanning electron microscopy revealed that co-precipitation of bare cells was not efficient to cover the cells as the calcium phosphate crystals precipitated either separately or non-uniformly at parts of the cell surface. The researchers performed a layer-by-layer coating with 8 polyelectrolyte layers on the cells and used these charged cells successfully for cell mineralisation. However, the enclosed cells became biologically inactive throughout the procedure [310].

To make use of the advantageous polyelectrolyte microcapsules for cell screenings, the cells should still be viable and not directly coated in the polymer. Therefore, extensive layers on the cells are unsuitable. Also, the template should act as scaffold around the cell to provide sufficient reaction volume. However, the low density of electronic charge on the cell surface has to be considered, when performing the calcium carbonate precipitation. Therefore, first precipitation experiments were performed with cells having the polymer poly(sodium 4-styrenesulfonate) (PSS) added to the $CaCl_2$ solution along with the cells. PSS is an anionic polyelectrolyte widely applied for the layer-by-layer technology [308]. *S. cerevisiae* cells have previously been directly coated in poly(allylamine hydrochloride) (PAH) and PSS, suggesting the polymers will adsorb onto the cell surface [83]. Figure 5.5, shows the first precipitation experiments together with cells, by varying the parameters important for crystal growth such as the incubation time, temperature, and the stirring speed while mixing the two solutions. All crystals formed did not incorporate cells. Faster stirring and shorter incubation resulted in rhombohedral crystals with up to 10 µm in size, whereas incubation for 30 min with faster or slower mixing and incubation at RT or on ice resulted in spheroidal crystals with up to 8 µm in size.

Different numbers of cells were tested and evaluated via microscopy, Figure 5.6. Differentiation of templates and cells is important, but the fluorescence signal of free cells can potentially overpower the weaker signal of cells within a crystal. Additionally, the formation of calcium carbonate crystals and cell cubes make the evaluation challenging. However, areas with lower cell numbers allowed for suitable evaluation without too many free cells disturbing the imaging.

A study from Mann *et al.* [194] of biological environments performing biomineralisation shows the effect of increasing the supersaturation to perform nucleation. A reduction in the activation energy takes place and nucleation is favoured [194]. Hence, different $CaCl_2$ and Na_2CO_3 concentrations were tested, see Figure 5.7. None of the tested concentration resulted in the incorporation of cells, however, the widely applied 0.3 M for each reaction results in spherical templates and a higher concentration of 0.5 M and 1 M yielded smaller and less distinctive crystals.

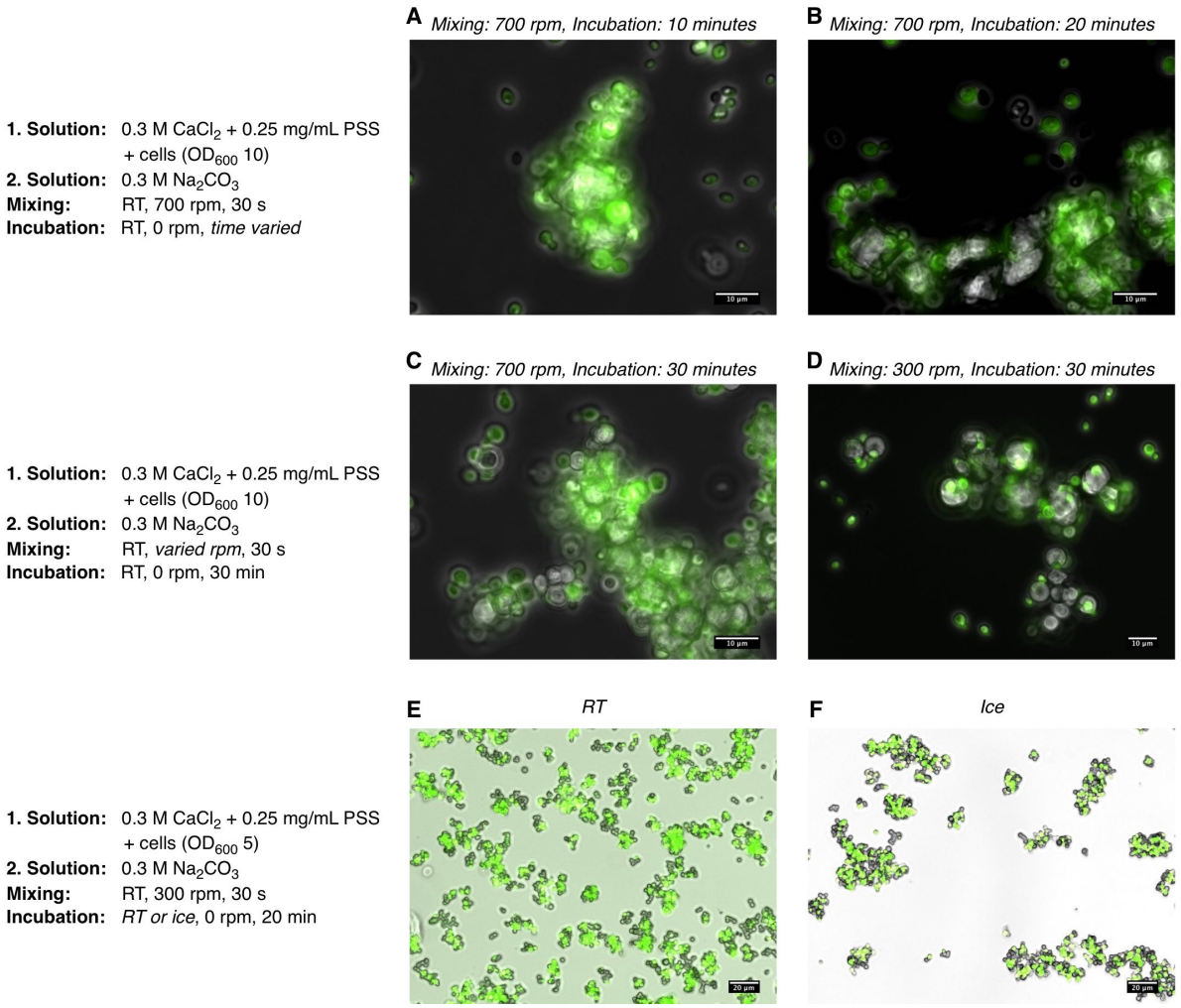

FIGURE 5.5. Template formation varying incubation time (**A, B, C**), stirring (**C, D**), and incubation temperature (**E, F**). Two saturated solutions with PSS and *P. pastoris* cells expressing yEGFP were combined in a beaker with vigorous stirring and was incubated in this beaker. Subsequently, templates and cells were analysed microscopically. The scale bar denotes 10 µm for **A, B, C, D** and 20 µm for **E, F**.

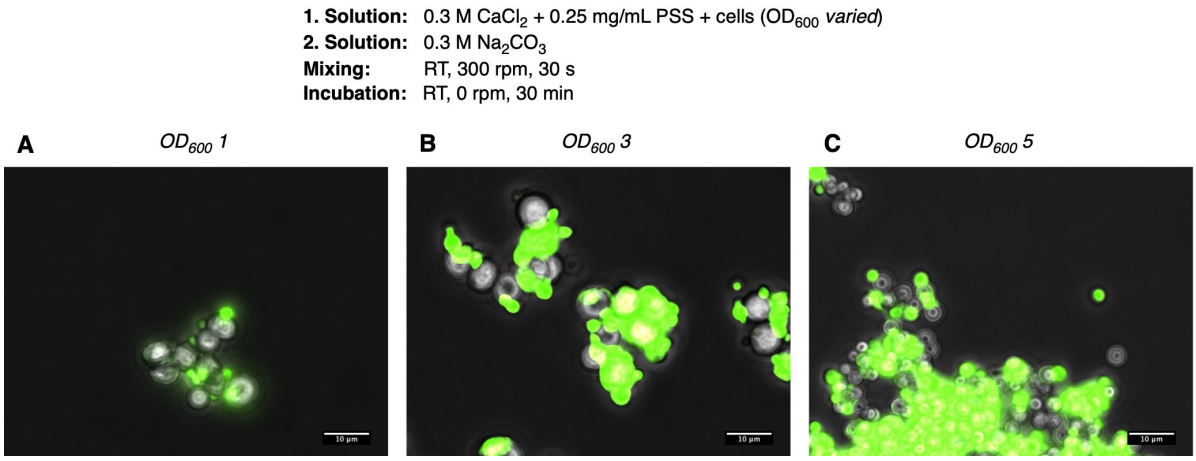

FIGURE 5.6. Calcium carbonate template formation with varying cell amount. To study different OD_{600}, *P. pastoris* cell constitutively expressing yEGFP were resuspended in 0.3 M $CaCl_2$. 0.3 M Na_2CO_3 was added while vigorously mixing, the templates were incubated for crystal formation and subsequently analysed microscopically. The scale bar denotes 10 µm.

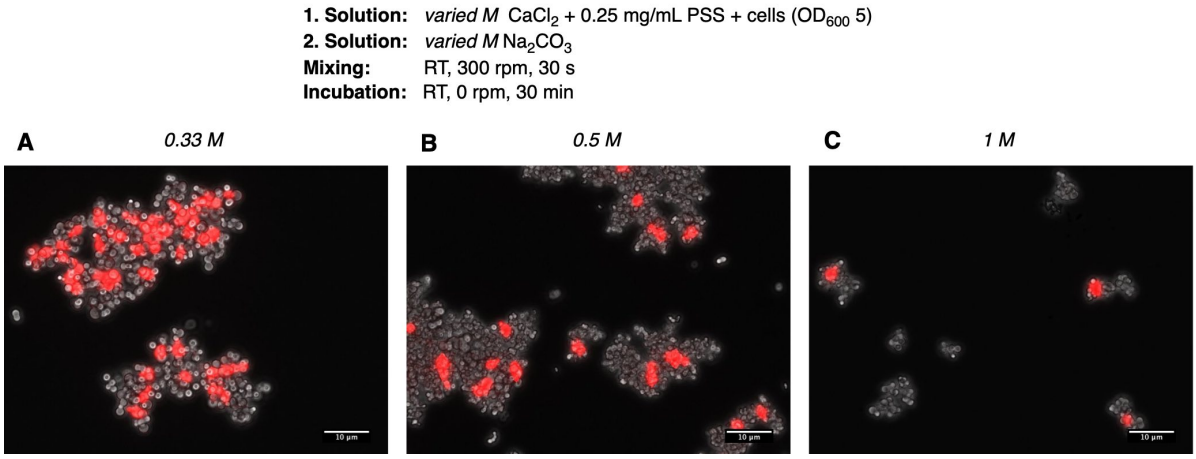

FIGURE 5.7. Calcium carbonate template formation using varying salt concentrations. *P. pastoris* cells expressing RFP after MeOH induction were resuspended in $CaCl_2$ with varying concentration. Na_2CO_3 with the respective concentration was added, the solution vigorously mixed and then incubated. Analysis was performed microscopically with the scale bar denoting 10 µm.

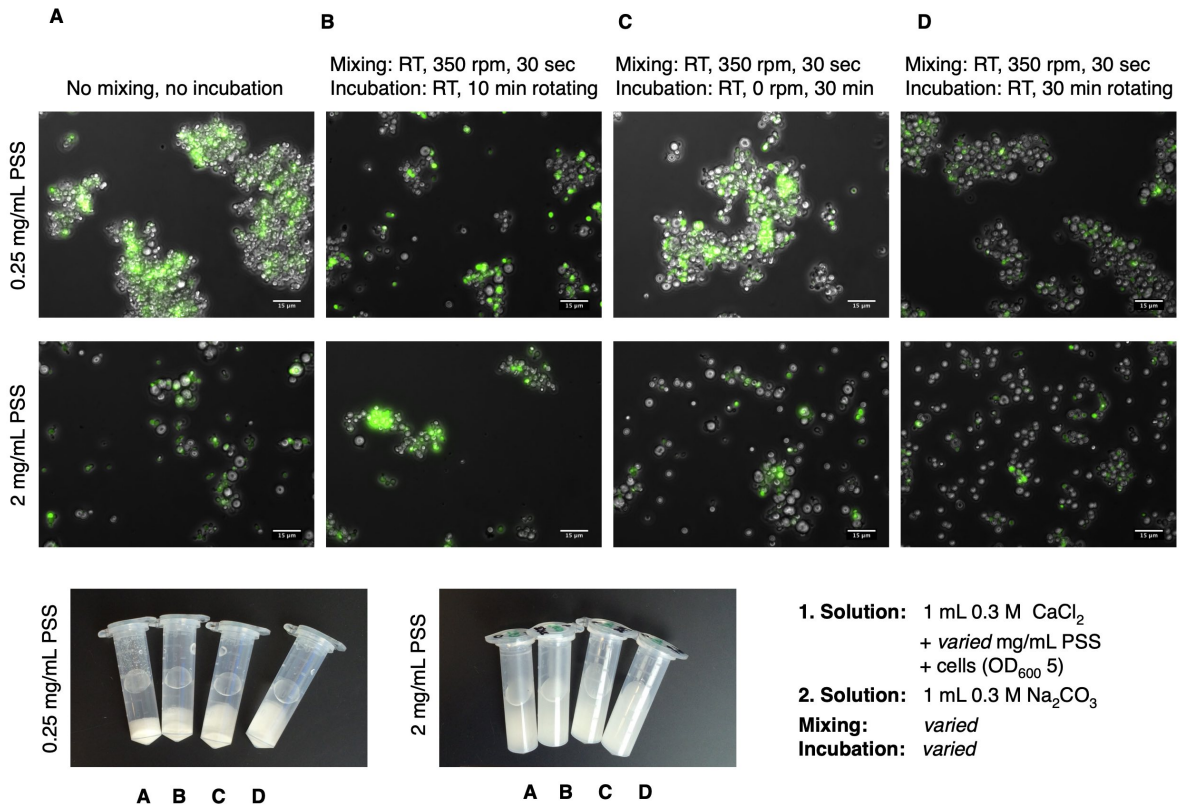

FIGURE 5.8. Template formation varying mixing and incubation conditions for different PSS concentrations. *P. pastoris* yEGFP expressing cells were co-precipitated and studied microscopically. For each condition, the appearance of the washed templates was documented. The scale bar denotes 15 µm.

For the first trial, PSS was added at a concentration of 0.25 mg/mL to mediate the nucleation of calcium carbonate on the cell, but no cells could be incorporated into the templates. Then, more PSS (2 mg/mL) was tested under different mixing and incubation conditions, see Figure 5.8. 2 mg/ml PSS results in more homogenous sized templates, however, none of these conditions resulted in suitable cell templates.

Other additives were considered to influence the nucleation process of calcium carbonate. EDTA is a complexing agent of Ca^{2+} and the presence should reduce the amount of free Ca^{2+} ions. Thus, the driving force for nucleation should be lowered [321]. Altay *et al.* [8] observed a change of morphology for calcium carbonate precipitates at various temperatures (30 °C – 90 °C) in the presence of EDTA, Figure 5.9. The addition of glycerol was also studied, as organic solvents will influence the solubility [173]. Again, neither the addition of EDTA nor the addition of glycerol supported cell integration, but templates with glycerol are homogeneous and spherical.

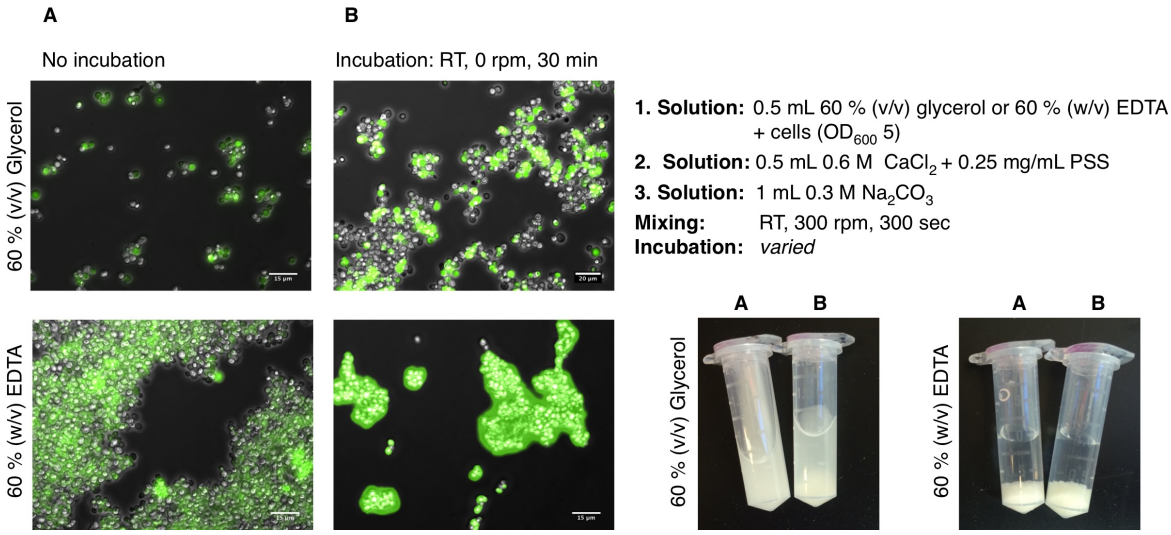

FIGURE 5.9. Template formation with cells and addition of glycerol or EDTA. *P. pastoris* yEGFP expressing cells were resuspended in glycerol or EDTA. Precipitation was performed using $CaCl_2$ with 0.25 mg/mL PSS and Na_2CO_3. Templates were washed before the microscopic evaluation.

Of all parameters studied so far, some had an influence on the calcium carbonate structure, but none mediated the nucleation on the cell surface encapsulating the cell. This highlights the requirement of modifying the cell charge. Therefore, the cells were pre-incubated in PAH, a cationic polyelectrolyte also applied for layer-by-layer technology [308] and applied for template formation with various salt, glycerol, and PSS concentrations. Despite small variations in template size, no templates contained cells, Figure 5.10.

In summary, the cell amount, cell pre-incubation condition, the salt concentration, additives to the cells containing solution (EDTA, glycerol, PSS), as well as conditions of the mixing and incubation procedure were studied, but none resulted in the incorporation of cells. Table 5.1 summarises the tested parameters and the subsequent observation. Due to the number of possible influencing parameters, a variation of one parameter at a time is not feasible and a more systematic approach must be taken. Therefore, a systematic screening to narrow down the influential parameters was necessary.

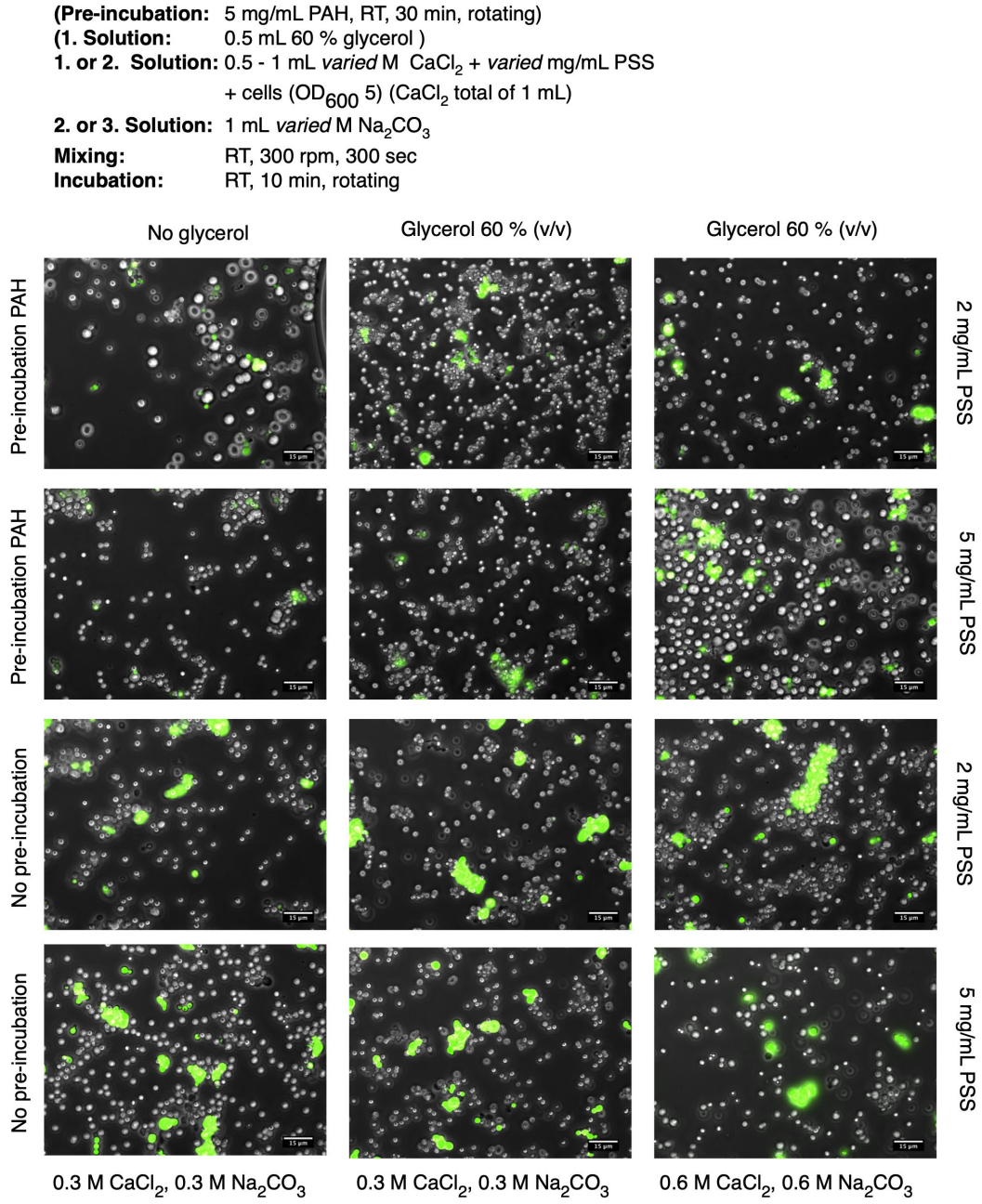

FIGURE 5.10. Template formation with cells studying cell pre-incubation while varying the concentration of glycerol, PSS and salts. Pre-incubation of *P. pastoris* expressing yEGFP was studied using PAH. If cells were incubated in PAH, cells were incubated in 1 mL 5 mg/mL PAH for 30 min, while rotating (20 rpm). The sample was centrifuged and the pellet was resuspended in glycerol and then the $CaCl_2$ part was added or resuspended directly in the $CaCl_2$ solution to make 1 mL. During addition of 1 mL Na_2CO_3 the sample was mixed and then incubated for crystal formation. The scale bare denotes 15 µm.

TABLE 5.1. Influential factors for template formation.

Parameter	Tested conditions[a]	Observation single variation approach
Salt		
Concentration	0.3 M, 0.5 M, 1 M	Strong effect, 0.3 M results in spherical templates, concentrations higher form less round and less defined templates
Cells		
Pre-incubation	30 min in PAH	No visible effect
OD_{600}	1, 3, 5	Evaluation using a fluorescent microscope possible with either concentration
Additives		
EDTA	60 % (w/v)	No visible effect
Glycerol	30 % (v/v), 60 % (v/v), 90 % (v/v)	Effect on size and homogeneity
PSS	0.25 mg/mL, 2 mg/mL, 5 mg/mL	Strong effect on template shape, size and homogeneity
Mixing		
Time	0 min, 0.5 min	Effect on template size and shape
Speed	300 rpm, 350 rpm, 700 rpm	No effect visible
Incubation		
Time	None, 10 min, 20 min, 30 min	Strong effect on template shape, homogeneity and shape
Temperature	0 °C, RT	No obvious effect
Condition	In beaker, tube or tube rotating (20 rpm)	Effect on template homogeneity

[a] Concentrations given of the operating solution.

5.2.3 Design of Experiments Approach for Template Formation

The task of cell integration into the calcium carbonate templates was not as straight forward as hoped. The main challenges were the many independent variables and time-consuming evaluation using microscopy. Additionally, the traditional approach of altering a single parameter while keeping all others constant – one-factor-at-the-time (OFAT) – was not feasible. Therefore, a more efficient strategy was needed to understand the system and narrow down the influential parameters [40]. The Design of Experiments (DoE) approach is a statistical methodology for process optimisation to quickly determine the important factors affecting yield or quality [111]. It can be of use during process development to identify those parameters that are critical for an established processes and to optimise conditions (e.g., towards a more robust or a more environmental friendly process) [315]. For biological systems that are often multifactorial, the DoE approach is a valuable method as these systems have responses that are not linear and often noisy [40].

When using a DoE approach, a limited number of experiments is performed with the aim of gaining the maximum amount of information about the parameters that can be varied in a system. Each parameter can relate to an environmental factor (e.g. temperature or media), or in the case of an expression construct some aspect of the genetic design (e.g. the promoter or secretion tag used). Each parameter is allowed to take a limited number of settings or levels as they are also known. For example, for a promoter parameter it might be possible to select a strong, medium and weak variant. There are numerous ways of selecting which combinations of settings/levels should be chosen for each parameter, but one of the most widely used approaches is to create fractional factorial designs using the Taguchi method. This relies on precomputed matrices called orthogonal arrays that define for a number of parameters and associated levels, the specific combinations of parameters that should be tested (i.e., the experiments performed). The matrix is composed of rows corresponding to specific experiments, where each column denotes a parameter. The number in each cell of the matrix provides the level that the associated parameter should have in that experiment. By performing this limited number of experiments it is possible to predict the influence of each parameter on the desired output and thus target future experiments towards those factors to optimise the process further [58, 145, 277].

As an example, for a system that has four parameters where each has three different settings/levels, an OFAT approach would require 81 trials to analyse the full combinatorial effects and 243 experiments when performed as triplicates. In contrast, when using the Taguchi approach the number of trials is reduced to nine (defined by what is termed the L9 orthogonal array) and thus if performed in triplicate a total of 27 experiments would be required.

For the identification of influencing factors of the template formation, three to four parameters were studied each using an L9 orthogonal array (four three-level factors), with each trial performed in triplicate. The L9 orthogonal array is shown in Table 5.2. The influence of the additives glycerol and PSS, the mixing time as well as incubation time, temperature and experimental set-up were studied, and the 3 levels for each parameter are shown in Table 5.3, Table 5.4 and Table 5.5. The results of the respective microscopy evaluations are shown in Figure 5.11, Figure 5.12 and Figure 5.13. To identify to important parameters of the calcium carbonate crystallisation for the cell encapsulation, the following output criteria are considered: (i) the resulting template shape being homogenous spheres and (ii) cells entrapped within the template.

TABLE 5.2. Taguchi L9 orthogonal array.

Experiment No.	Level for each Parameter A	B	C	D
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

TABLE 5.3. Taguchi DoE parameters I.

Parameter	Level 1	Level 2	Level 3
A Glycerol concentration (v/v)	30 %	60 %	90 %
B PSS concentration	1 mg/mL	5 mg/mL	7.5 mg/mL
C Incubation time	-	10 min	30 min
D Incubation condition	0 rpm, RT	0 rpm, 4 °C	Rotating, 20 rpm, RT

TABLE 5.4. Taguchi DoE parameters II.

Parameter	Level 1	Level 2	Level 3
A PSS concentration	0 mg/mL	2.5 mg/mL	7.5 mg/mL
B Stirring time/condition	Manual mixing	10 s, 750 rpm	1 min, 750 rpm
C Incubation time	-	10 min	60 min

TABLE 5.5. Taguchi DoE parameters III.

Parameter	Level 1	Level 2	Level 3
A Stirring time	1 min	10 min	30 min
B Incubation time	10 min	30 min	60 min
C Incubation condition	0 rpm, in beaker	0 rpm, in tube	Rotating, 20 rpm, in tube

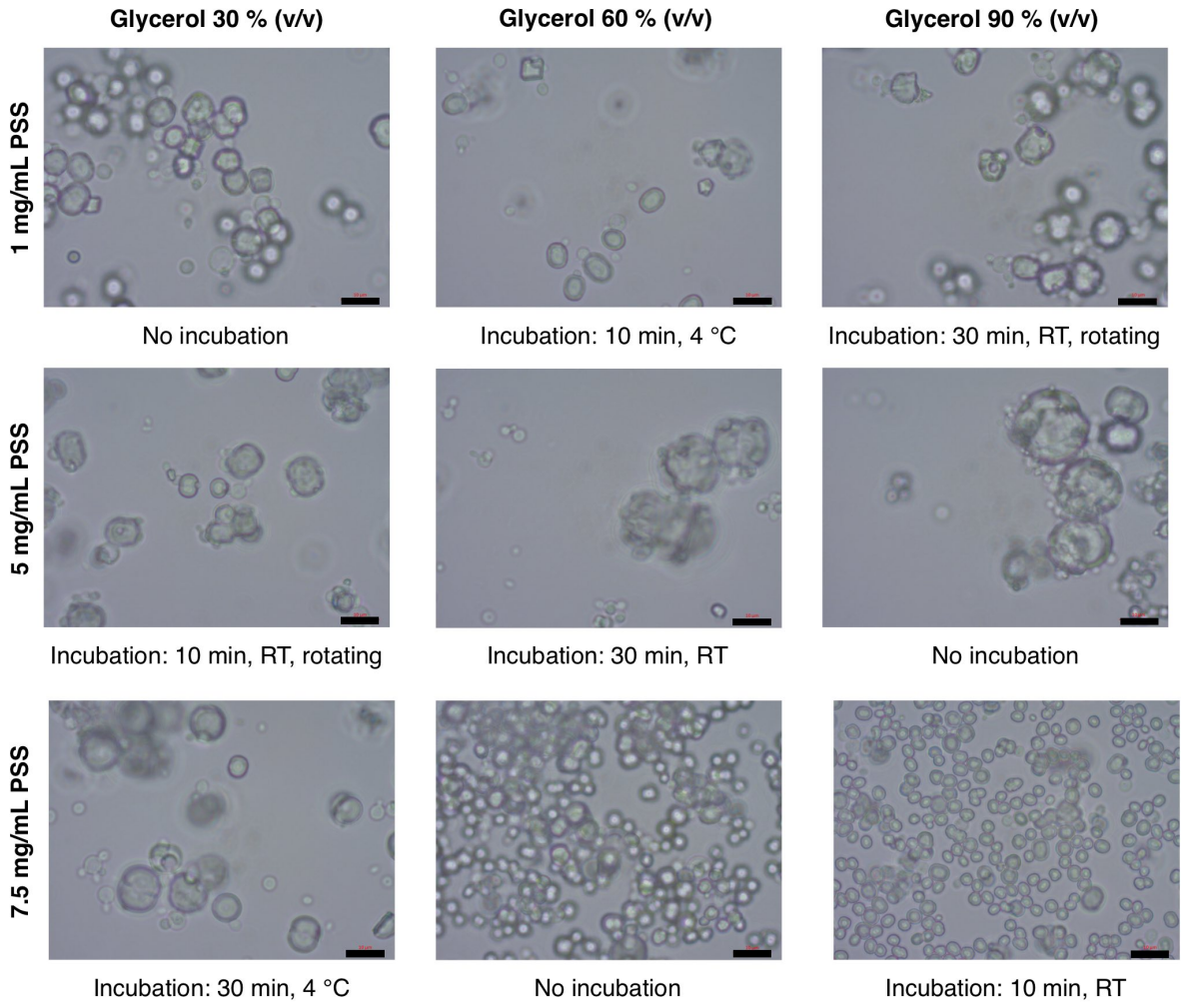

FIGURE 5.11. Taguchi DoE approach studying template formation with cells I. Cells were resuspended in 0.5 mL glycerol (30-90 %), then 0.5 mL 0.66 M $CaCl_2$ with PSS (1-7.5 mg/mL) was added while vigorously stirring. 1 mL 0.33 M Na_2CO_3 was added (750 rpm, 20 s) and the beaker was place for incubation at the defined condition. The scale bar denotes 10 µm.

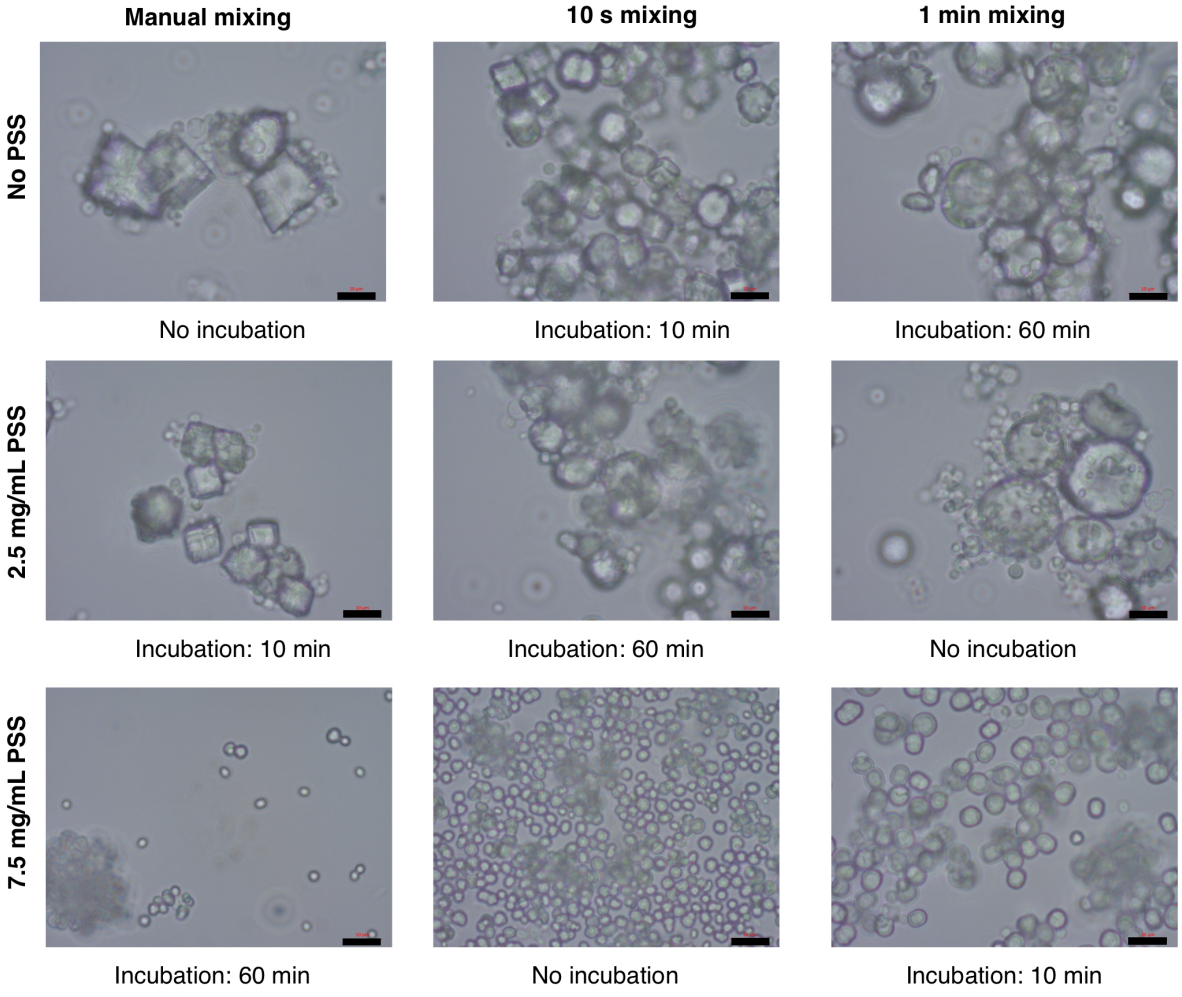

FIGURE 5.12. Taguchi DoE approach studying template formation with cells II. Cells were resuspended in 0.5 mL 60 % glycerol, 0.5 mL 0.66 M $CaCl_2$ with or without PSS (0-7.5 mg/mL) added. Subsequently, 1 mL 0.33 M Na_2CO_3 was added while mixing (manually by swivelling the beaker or at 750 rpm, 10-60 s) and the sample in the beaker was either analysed directly or the beaker was placed for incubation (0 rpm, 10 min or 1 h). The scale bar denotes 10 µm.

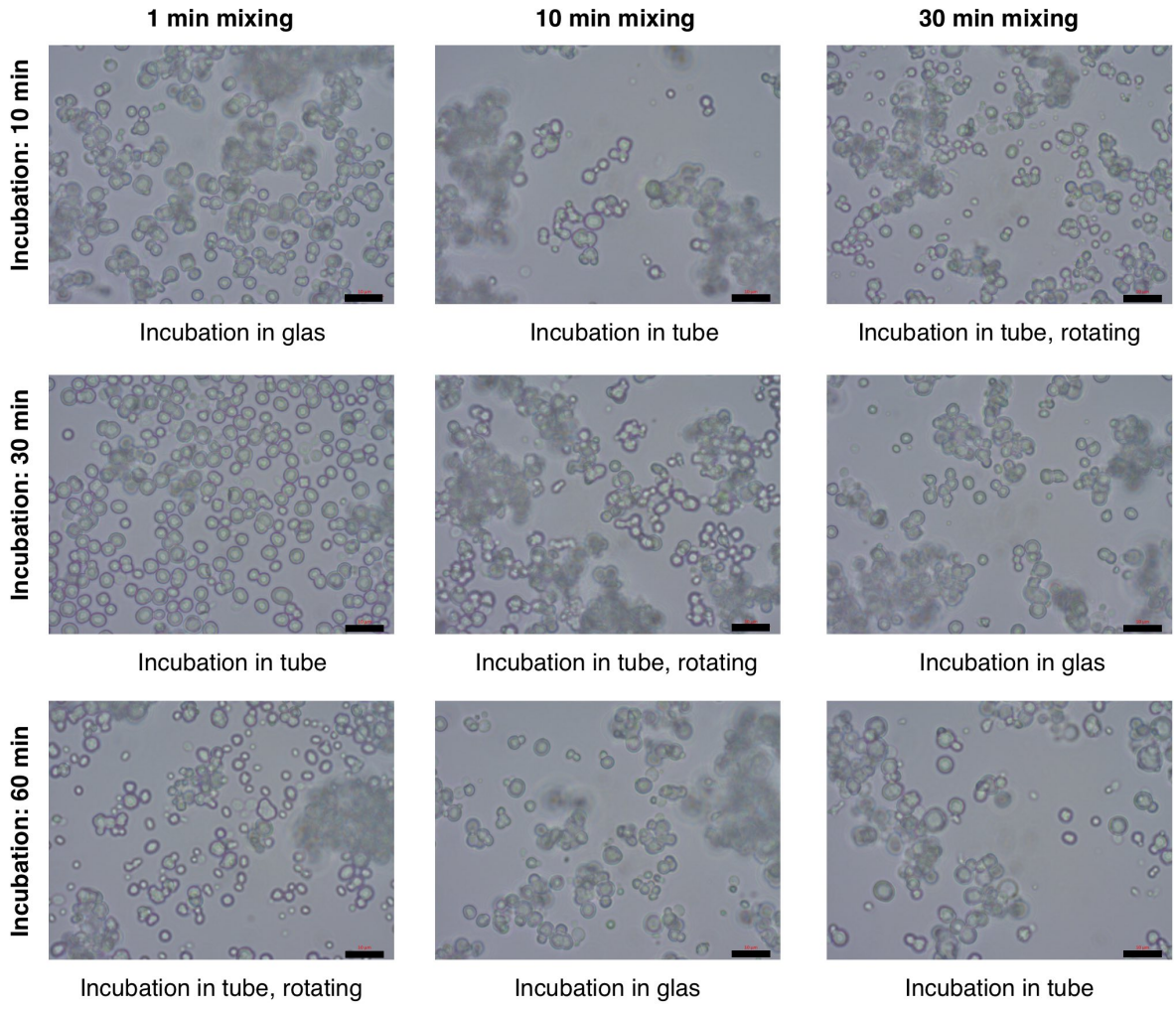

FIGURE 5.13. Calcium carbonate template formation at varying incubation times and temperatures. First, cells were resuspended in 0.5 mL 60 % glycerol, second, 0.5 mL 0.66 M $CaCl_2$ with 2.5 mg/mL PSS was added and third, 1 mL 0.33 M Na_2CO_3 was added while stirring (750 rpm, 1 min to 30 min). Finally, the beaker was placed at RT for sample incubation (10 min to 60 min). The scale bar denotes 10 µm.

To determine the effect of the parameters tested at the three different levels, the templates were studied using a light microscope. Unfortunately, free cells are clearly differentiated from the templates under all conditions and most settings show significant numbers of free cells. Therefore, the output variable (ii), addressing cells entrapped within the template does not allow the identification of conditions suitable for forming a crystallisation mantle around the cells. However, the approach did enable the identification of parameters that result in templates which are homogenous spheres, addressing output variable (i). The addition of glycerol and PSS has a strong influence on crystal shape and homogeneity. 30 % (v/v) glycerol resulted mostly in consistent sized templates. The addition of PSS significantly supports the formation of spherical shaped templates and 7.5 mg/mL PSS always resulted in spherical templates irrespective of all other conditions. The mixing condition strongly influenced template formation. Shorter mixing times with defined mixing conditions using a stirring bar were better for spherical, homogenous templates. The longer the mixing, the less defined the templates were, and it appeared that they had become grounded. With increasing incubation times, less uniform templates were obtained, but the temperature does not influence the procedure significantly. Furthermore, regardless of whether the incubation was done in a beaker without stirring or in a tube being slowly rotated (20 rpm), the outcome was similar.

Despite no suitable conditions for cell integration into a calcium carbonate template being found, parameters important to reliably produce spherical and homogenous sized templates could be identified. These templates are suitable to study the layer-by-layer coating.

5.3 Layer-by-layer Approach for Capsule Generation

The layer-by-layer (LbL) assembly is a promising technology for thin-film research and the coating of microparticles [72]. Highly tailored layers can be generated by the alternation of charged cationic and anionic polyelectrolytes on a charged surface, driven by electrostatic interactions at each step [237]. The LbL technique is comparatively simple and does not require expensive equipment, but offers great flexibility in terms of the final chemical and physical properties of the three-dimensional structures generated [119]. The choice of polyelectrolytes and number of assembled layers precisely tunes the shell thickness and consequently, the permeability. Generally, capsules are semipermeable allowing the bi-directional diffusion of small molecules in and out, but do not allow the diffusion of high molecular mass molecules. Exterior conditions such as the pH, ionic strength, or solvents can be used to tune the permeability. The addition of functional polymers like weak polyelectrolytes or thermosensitive polymers can furthermore link the permeability of the capsule to external stimuli [13]. The most widely used polyelectrolye couple are the strong polyanion poly (sodium styrene sulfonate) (PSS) and the relatively weak polycation poly (allylamine hydrochloride) (PAH). They have been extensively studied for layer formation on flat and spherical surfaces and are sensitive to pH changes [14]. The laboratory method for capsules formation is depicted in Figure 5.14.

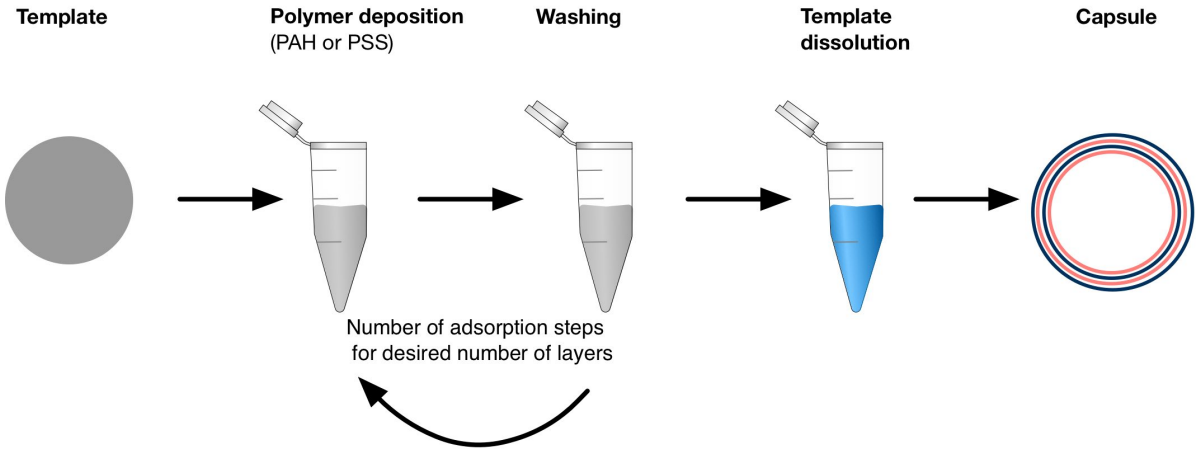

FIGURE 5.14. Lab scale layer-by-layer technology on spherical particles. In a reaction tube scale, charged polyelectrolytes are adsorbed in an alternating fashion. For the first polymer deposition, templates are incubated in PAH. These coated templates are then centrifuged and washed multiple times to remove all free polymer before the next adsorption step of the oppositely charged polymer PSS. Afterwards, the coated templates are washed thoroughly. Once the desired number of layers is reached, the core template is removed via EDTA dissolution of $CaCO_3$ and an empty capsule remains.

Despite the challenges in template formation with cells integrated into the calcium carbonate core, the LbL technology was tested using the spherical templates produced. Plenty of protocols for the LbL formation are published [14, 47, 78, 236, 307], all applying the "*centrifugation technique*" [289], to spin down the coated templates for thorough washing and addition of the next polymer layer. First, templates were made by resuspending *P. pastoris* cells in 1 mL 0.3 M $CaCl_2$ + 0.25 mg/mL PSS, and adding 1 mL 0.3 M Na_2CO_3 while mixing at RT, 300 rpm, 30 sec in a beaker with the string bar and subsequently incubation at RT, 10 min without stirring. The templates were washed 3 times (1 mL 0.05 M NaCl, centrifugation RT, 500 x *g*, 30 sec) before the coating. For the coating, the templates were first incubated in 1 mL PAH (5 mg/mL) (RT, varied mixing, 10 min). Next, the coated templates were spun down (RT, 500 x *g*, 30 sec) and thoroughly washed using 1 mL 0.05 M NaCl 3 times. Subsequently, a layer of PSS was added (1 mL PSS (5 mg/mL)), incubated (RT, varied mixing, 10 min) and finally the excess polyelectrolyte was removed by washing three times using NaCl as before. Eight layers of polymer were added $(PAH/PSS)_4$, before the calcium carbonate template was dissolved using 1 mL 0.25 M EDTA each for two consecutive incubations (RT, mixing varied, 3 min) and lastly washing three times using 1 mL 0.05 M NaCl (all dissolution centrifugations: RT, 1500 x *g*, 3 min) [98, 236, 307]. A more detailed description can be found in Chapter 2.6.

Template incubation in the polyelectrolyte solution, was performed in two different experimental set-ups. First, in a beaker while mixing with a stirring bar at 300 rpm [98], which resulted in only few intact spherical capsules, see Figure 5.15. Second, in 1.5 mL reaction tubes while rolling on a rotator (20 rpm), see Figure 5.16. Despite the fact that cells were not integrated into the templates (Figure 5.16), spherical capsules were made via this approach.

Spherically intact, but empty capsules can be clearly evaluated using the microscope, however, evaluation of polymer coated or potentially encapsulated cells is challenging. Therefore, a layer of fluorescent labelled PAH was added using fluorescein isothiocyanate – PAH (FITC-PAH) [308]. Unfortunately, the polymer capsules stick together and microscopy images of individual capsules were rare, see Figure 5.17 A and B. Nevertheless, a few individual capsules could be visualised, see Figure 5.17 C, and as indicated with the arrow a cell was coated in polymer. As the RFP signal of the intracellularly expressed RFP fills the whole capsule, it is likely that the cell was directly coated with the polymer, as opposed to the assumption of first being entrapped in calcium carbonate and then encapsulated. Nevertheless, the encapsulation of some cells could be performed as desired (Figure 5.18), with comparatively small cells enclosed in capsules with sufficient reaction volume around the cells.

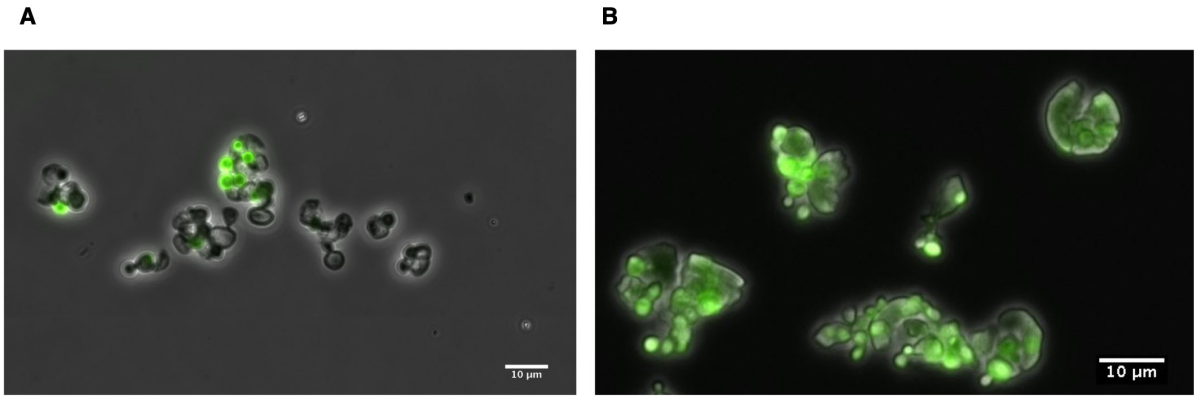

FIGURE 5.15. Layer-by-layer approach with mixing in a beaker. The incubation of the templates in the polyelectrolyte solutions as well as the dissolution of calcium carbonate via incubation of the capsules in EDTA was performed in a beaker and mixing was performed through a stirring bar. The templates coated in a $(PAH/PSS)_4$ polyelectrolyte layer **A** as well as $(PAH/PSS)_4$ polyelectrolyte capsules **B**, each with GFP expressing.

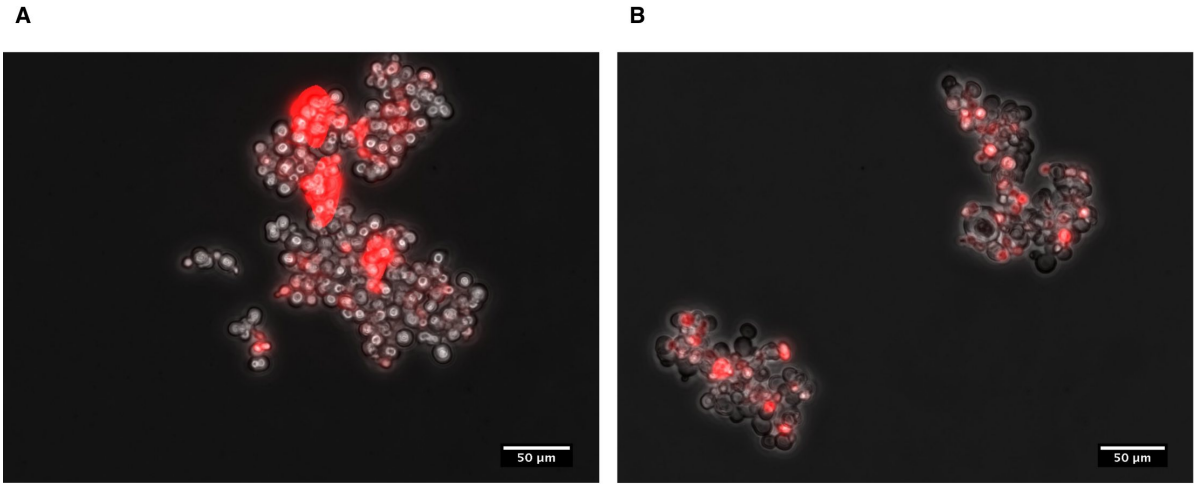

FIGURE 5.16. Layer-by-layer approach with rolling mixing using a rotator. **A** Round templates are made via calcium carbonate precipitation, in the presence of RFP expressing cell. **B** Templates are coated and the templates were dissolved using EDTA. All incubations were performed in the reaction tube being gently mixed via rolling on a rotator (20 rpm).

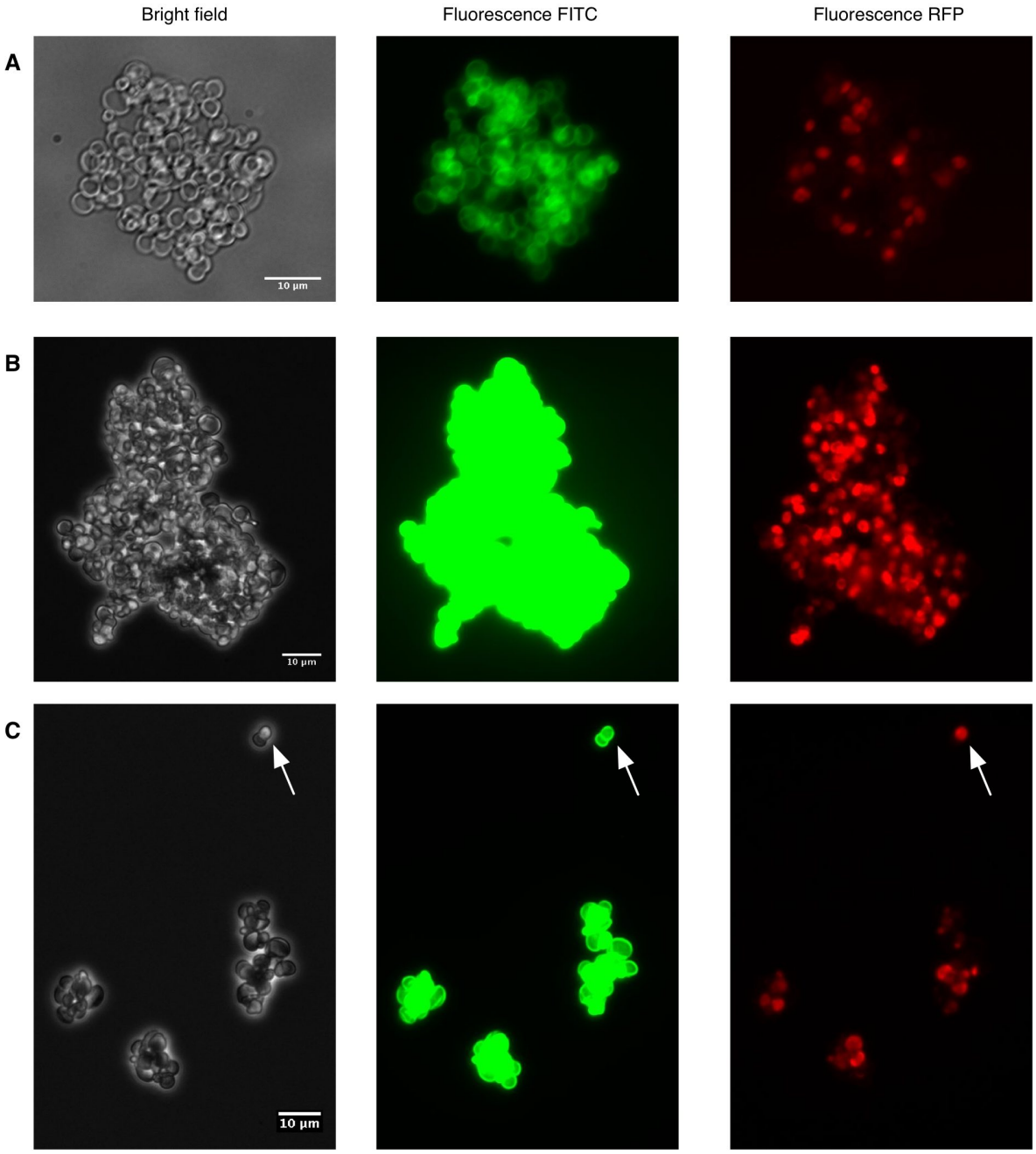

FIGURE 5.17. Layer-by-layer approach for cell encapsulation and FITC labelled polymer shell I. **A** and **B** Imaging of individual capsules is challenging as the capsules stick together and the overlay of fluorescent signals does not support evaluation. **C** A individual cell was coated with polymer, as the whole interior capsule space is taken up by the cell. The scale bar denotes 10 µm.

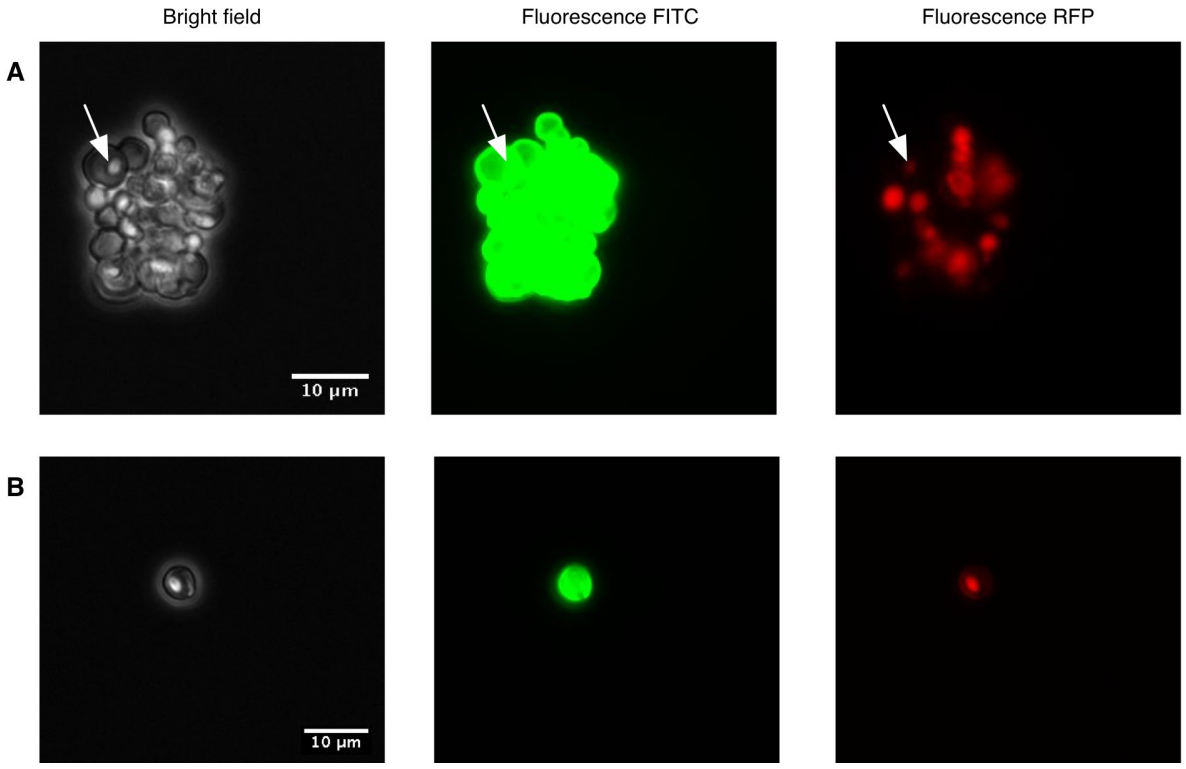

FIGURE 5.18. Layer-by-layer approach for cell encapsulation and FITC labelled polymer shell II. Individual cells could be encapsulated, allowing reaction volume around the cell as desired for this approach. The scale bar denotes 10 µm.

The final goal of the polyelectrolyte microcapsules for cell screening is to perform a fluorescent activated cell screening (FACS) for identification. The capsules have previously been shown to be suitable for flow-cytometry analysis and therefore were subjected to this procedure [273]. In Figure 5.19, capsules of $(PAH/PSS)_4$ and $(PAH/PSS)_4$+FITC-PAH without cells were studied. For both samples, the distribution of capsules represented by the forward scatter (FSC) and side scatter (SSC) shows one population. For the capsules with FITC-PAH labelling, the respective FITC signal is strongly detectable.

FACS was performed on the RFP expressing cells, free cells (non-encapsulated), cells exposed to LbL technology for $(PAH/PSS)_4$ capsules and $(PAH/PSS)_4$+FITC-PAH capsules, Figure 5.20. Free cells form one population according to the evaluation of SSC and FSC for the size of the cells. Two populations were detected for the RFP signal, the main group with RFP expression but also some cells without RFP expression. Exposure to the LbL method for $(PAH/PSS)_4$ capsules modifies the size distribution, but a single population remains. The measured RFP signal still shows the two populations despite seeing an increased RFP signal for the population expressing the protein.

The cells exposed to the approach for $(PAH/PSS)_4$+FITC-PAH capsules also show one population considering the size distribution, which is wider than for the $(PAH/PSS)_4$ capsules. The overlay of fluorescence signals shows the largest population having both the RFP and FITC signal, meaning FITC coated or encapsulated cells. Only a small population of empty capsules corresponding to the FITC signal were measured, and an even smaller population of free cells is seen. Thus, the flow cytometry analysis of encapsulated single cells was possible, despite what was seen under the microscope. Cells exposed to the LbL approach could be successfully coated in polyelectrolytes, however, successful encapsulation as desired, remains questionable.

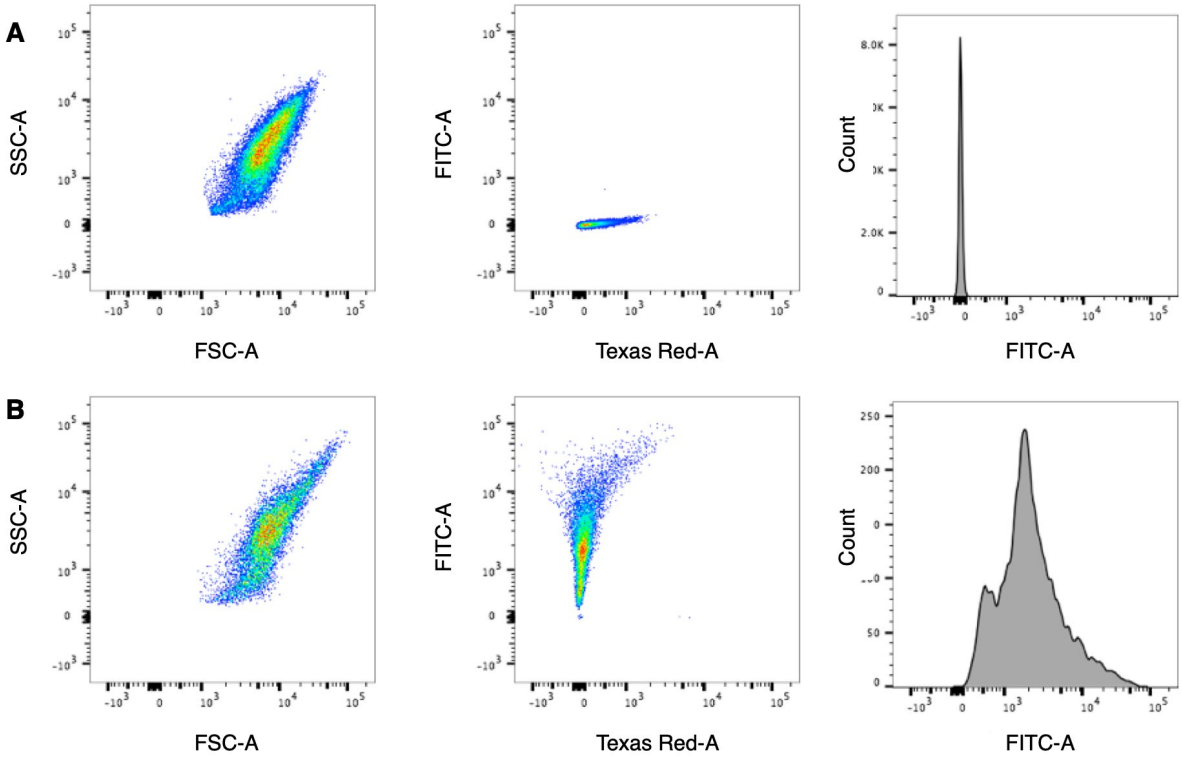

FIGURE 5.19. Flow cytometry analysis of capsules. **A** $(PAH/PSS)_4$ and **B** $(PAH/PSS)_4$ + FITC-PAH capsules were analysed, for their size distribution (SSC and FCS), and fluorescence signals of FITC and RFP. Analyses performed using FlowJo.

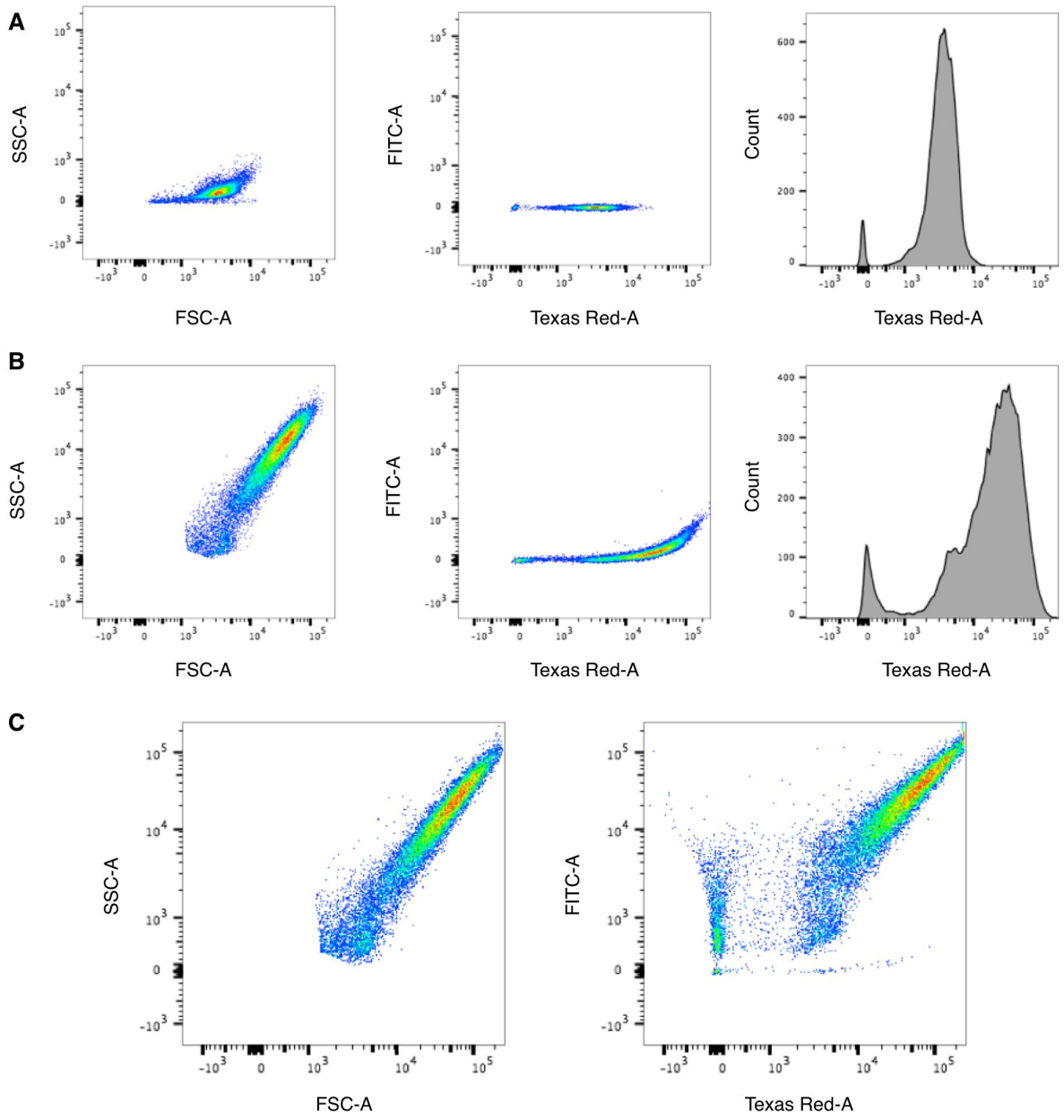

FIGURE 5.20. Flow cytometry analysis of *P. pastoris* cells and capsules. **A** RFP expressing *P. pastoris* cells and **B** RFP expressing cells in $(PAH/PSS)_4$ capsules. **C** RFP expressing *P. pastoris* in $(PAH/PSS)_4$ + FITC-PAH capsules. Analyses performed using FlowJo.

5.4 Conclusion

Polyelectrolyte microcapsules offer a great scope of application covering drug delivery purposes or sensitive pH sensors. The objective of this chapter was to test their applicability for cell encapsulation and screening. The approach was to study the integration of cells into round calcium carbonate crystals and perform subsequent coatings via a layer-by-layer methodology with polyelectrolytes.

The calcium carbonate crystallisation process has previously been intensively characterised due to its industrial use. In this study, we investigated the crystallisation at ambient conditions to form a coating of calcium carbonate around single yeast cells. Despite the variety of conditions tested, we did not succeed in forming a crystalline shell around the cell. Hence, it was not possible to develop a novel screening technology as desired for *P. pastoris* cultivation and screening in polyelectrolyte microcapsules. However, *P. pastoris* cells could directly be coated in polyelectrolytes and were viable for flow cytometry analysis. The polyelectrolyte layer should be permeable for nutrients to feed the cells, but not permeable for proteins being secreted from the cell. A modified approach for a cell screening technology could consider these directly coated cells, but future investigations into their viability would be required.

6

Discussion and Outlook

6.1 Discussion

This thesis set out to develop an end-to-end platform for creating and optimising *P. pastoris* cell factories for protein expression and secretion. The motivation was to overcome the difficulties identifying genetic constructs for efficient protein secretion and to simplify the creation and testing of diverse libraries of genetic designs. The work strived to highlight the potential of *P. pastoris* as a protein expression platform and provide the tools and methods needed to expand its use in biotechnology.

6.1.1 A Toolkit to Build Expression and Secretion Libraries

Acting as a foundation for this work the PTK was developed to allow for *P. pastoris* expression vectors to be built in a modular way (Chapter 3). This enables the design of diverse expression constructs to meet varying demands [144]. The toolkit consists of 42 regulatory elements covering promoters, terminators, secretion signal peptides, and a yeast origin of replication for *P. pastoris*. It builds upon the widely used YTK from Lee *et al.* [172] providing a wealth of parts from *S. cerevisiae* as well as expanding the core functionality to include elements controlling secretion and parts that are compatible with *P. pastoris*.

The modular structure of the toolkit allowed for the flexible assembly of different combinations of genetic parts. The PTK was used to build 242 strains expressing either RFP or yEGFP as a reporter protein to systematically characterise the new regulatory components as well as a selection of existing YTK parts in *P. pastoris*. To our knowledge, previous work solely addressed the characterisation of promoters [124, 250, 286, 303] and terminators [304] separately.

Here with the PTK more complex interactions between promoters, secretion tags, and terminators, and their impact on recombinant protein expression in *P. pastoris* were studied. The studied constitutive *P. pastoris* and *S. cerevisiae* promoters were found to allow tuning of expression across three orders of magnitude with their strength being mostly independent of the downstream coding region for intracellularly expressed proteins. The three methanol inducible promoters resulted in high protein expression independent of the downstream coding region and facilitated fine-tuning of protein expression with varying methanol concentrations. Testing of a broad range of expression levels is often necessary for the identification of optimal expression conditions [250] and the PTK promoters provide the necessary tools to carry out such studies.

The combinatorial capabilities of the PTK were useful for the characterisation of terminators to study the influence of promoter-terminator combinations. Terminators from *S. cerevisiae* and *P. pastoris* were studied in combination with a variety of promoters to assess intracellular expression of RFP. It was found that these terminators provide an additional dial to tune expression strength. For a single promoter the expression level could be tuned by five- to ten-fold depending on the terminator used. This is in contrast to a previous study that compared yEGFP expression from a pAOX1 promoter with 20 different terminators, where similar fluorescence was seen in all cases [304].

Overall, the consistent behaviour of these regulatory elements enables expression levels to be directly related to DNA sequences, enabling the design of expression vectors with a more predictable outcome. Furthermore, the ability for *S. cerevisiae* regulatory parts to function in *P. pastoris* highlights the benefit of compatibility with the YTK, opening up access to a range of existing parts that can be directly used to accelerate strain development.

Stable expression is essential for any protein production process [65] and therefore the influence of expression constructs being genomically integrated or placed on a self-replicating plasmid was studied. Overall, expression strength did not vary significantly between the two modes. However, microscopy showed strong variation of intracellular RFP and yEGFP expression across the population when expressed from a plasmid. This finding is important to consider when high and ultrahigh-throughput screenings are performed [186] as changes in expression due to plasmid variability may hamper an accurate assessment of the average expression level for a particular design.

The major reason for developing the PTK was to study protein secretion. A bottleneck when developing strains that secrete a desired protein is identifying the optimal expression conditions and secretion tag [69]. To address this, 29 secretion tags from endogenous, exogenous and synthetic sources were designed, and their efficiency to secrete RFP and yEGFP in *P. pastoris* was systematically tested. It was found that commonly used combinations of promoter and secretion tag did not always result in optimal expression, highlighting the importance of testing many options when optimising secretion [200]. Another interesting observation was that the secretion tags not only influenced the amount of protein secreted, but also the total amount of protein

expressed (intra and extracellular).

Six of the secretion tags tested in this work have not been studied before. Some of these would make good candidates for designing future secretion constructs. Previously, the endogenous *P. pastoris* tags SP_Cyclophilin, SP_C4R8H7, and SP_Mucin have been predicted to work effectively [200] and in this thesis were shown to successfully drive secretion of a variety of proteins. The exogenous signal peptide SP_An_phyA did not guide RFP secretion, but resulted in an almost 2.5-fold increase in measured intracellular RFP expression compared to strains constructed for intracellular expression (Section 3.5). The application of this tag to solely increase the overall expression yield could be an interesting future direction, but further testing for a wider range of proteins is required. The phytase secretion signals SP_Pl_phyA and SP_Th_phyA did not enable efficient protein secretion, but RFP and yEGFP were still expressed intracellularly.

In addition to natively occurring secretion tags, synthetic hybrid tags were constructed. These were designed to have an initial recognised leader sequence which was then coupled to a shortened αMF secretion tag. The idea was that the leader sequence guides secretion and the αMF part supports the correct processing and cleavage of the tag from the protein to ensure successful protein folding once secreted [276]. Despite this design being built with a limited understanding of the secretion process, some did guide secretion of RFP and yEGFP, demonstrating the flexibility and modularity of secretion tag design.

Improving protein secretion still remains to be one of the most important goals for engineering in *P. pastoris* [4, 144, 329]. The demonstrated potential of novel endogenous and synthetic secretion tags will further support the development of protein secretion cell factories.

Previously, the integration of expression constructs into the genome of *P. pastoris* during strain development was challenging. Procedures dependent on homologous recombination often result in varying numbers of integration events within the genome or ectopic integration, i.e., integrations at a locus other then the one desired. This makes it difficult to directly compare strains when studying diverse libraries of expression constructs [65]. Here, this issue is addressed by using a recombinase guided transformation procedure developed by Perez-Pinera *et al.* [234]. This method allowed for a single construct to be integrated at a specific locus in the genome and enabled direct comparisons between strains. In this work, the transformation procedure is further improved to enable quicker strain generation and higher transformation efficiency. The ability to now efficiently and accurately generate *P. pastoris* expression strains containing a defined combination of genetic parts opens up possibilities to perform high-throughput screenings to find optimal designs and potentially learn key design constraints via machine learning.

The published toolkit containing a selection of the parts described in this thesis (4 promoters, 10 secretion tags, 1 terminator and 2 origin of replication) has already gained recognition within the *Pichia* community [63, 103, 233, 264, 271, 334] as Schwarzhans *et al.* write: "*The publication* [Obst *et al.*] *marks the first P. pastoris library of characterized, sequenced and standardized DNA parts for modular vector construction.*" [271].

The availability of these parts through the Addgene repository (#1000000108) further promotes the application and development of the toolkit. At present (27.08.2018), the provided toolkit has been requested seven times and seventeen individual part plasmids have been acquired by different groups world-wide. The compatibility of the PTK with the YTK makes it attractive to researchers who may already be familiar with the YTK. The PTK is also a step towards more standardised methods for engineering yeast, of any sort, to broaden the availability of tools for the expression of desired proteins [103].

Since the publication of the PTK, two alternative Golden Gate based toolkits for *P. pastoris* have been developed [244, 270]. The toolkit by Schreiber *et al.* was developed for the production of antimicrobial peptides and allows the flexible design of the transcriptional unit into a defined backbone vector [270]. The GoldenPiCS toolkit from Prielhofer *et al.* supports, like the PTK, the design of complete expression vectors and the design of multi-gene constructs [244].

6.1.2 Toolkit Application for Enzyme Secretion

To explore the potential of the PTK for developing cell factories, randomised protein secretion libraries were constructed, where genetic regulatory elements driving both expression and secretion of a protein were varied (Chapter 4). The aim was to achieve high-levels of secretion for three industrially relevant proteins. Establishing a suitable approach for library generation in *P. pastoris* and sensitive screenings for each enzyme was crucial. It has already been shown that Golden Gate shuffling enables the creation of randomised expression construct libraries [93], and this technique was successfully applied in this work to assemble the PTK parts. Important for the successful generation of *P. pastoris* libraries was ensuring that correctly assembled expression plasmids were produced and a sufficiently high transformation efficiency could be reached [323] to guarantee full library coverage. This was tested in this thesis for three different enzymes: phytases, unspecific peroxygenases (UPOs), and alcohol dehydrogenases (ADHs).

Four different phytases were used for each screening with varying secretion tags and promoters for induced or constitutive expression. For each phytase, suitable combinations of promoters and tags were identified that enabled protein secretion. The finding was not only the identification of suitable secretion tags, but also the overall applicability of signal peptides for certain enzymes. In particular, it was found that enzymes differ strongly in the number of tags that enable efficient secretion. The fungal phytases from *A. niger*, *T. heterothallica* and *P. lycii* showed a clear preference for specific tags indicated by the significant slope of the screening landscape (Section 4.2). In contrast, the *E. coli* phytase landscape slope is much shallower for the 150 best expression strains, indicating that tag choice has less of an effect. As already seen in Chapter 3 for RFP ad yEGFP, the efficiency of tags vary between the four phytases, despite these enzymes have a similar amino acid sequence and structure. Generally, the strong pGAP and pAOX1 promoters resulted the in highest protein secretion levels except for the *P. lycii* phytase, where the pPMP20, pDAS1, pG6 and pAHD2 promoters yielded the best results.

These findings should be considered in future studies as the construction of libraries with fewer promoters may miss beneficial combinations. Although the promoter strengths were found to be independent of the downstream coding region for intracellular expression, secretion is strongly effected by the combination of promoter, tag, and gene of interest. The most widely used tag αMF [183] showed high levels of secretion for the proteins tested here, but novel tags such as SP_C4R8H7, SP_Cyclophillin or the less commonly used tags such as SP_CSN2, or SP_PHO1 [112, 126, 148] also resulted in high protein secretion and thus are viable candidates during strain development. The application of the SP_CSN2 instead of the tag αMF increased the amount of secreted *A. niger* phytase by 20 %. Considering that the global phosphatase enzyme market has exceeded 300 million US$, an increase in production of 20 % by merely choosing an alternative secretion signal is an attractive means for strain improvement.

Scale-up experiments were performed to test expression between deep-well plates and shake flasks. *P. pastoris* is known for its ability to achieve high cell density cultivations [103]. However, the procedure used here resulted in low cell densities and thus low protein yields, leaving space for optimisation. To use the identified strains for enzyme expression applications, further cultivation studies or high-cell density fermentations must be considered to improve enzyme yield. The pharmaceutical company Huvepharma (Sofia, Bulgaria) already uses *P. pastoris* for the production of a 6-phytase from *E. coli*. Their product OptiPhos® was launched in 2008 and is currently the fastest phytase in the market [299].

Six different UPOs were tested: the *Aae*UPO previously engineered and successfully secreted by *P. pastoris* [209], the *Cra*UPO not yet heterologously expressed [207], and four predicted UPOs never expressed before [131]. A single library with all UPOs for induced expression, as well as 6 libraries for each UPO using constitutive expression were constructed and screened. The library containing all UPOs for induced expression found that none of the novel UPOs could be identified as expressible, but did demonstrate the approach as suitable for generating complex libraries when searching for high expression strains. Previously described combinations of secretion tags for the *Aae*UPO [209] as well as suitable combinations with other tags for *Aae*UPO did result in measurable secretion. *Aae*UPO expression levels were highest for *Aae*UPO with SP_Aae_UPO or SP_Aae_UPOeng. However, SP_PHO1, SP_Cyclophillin, or SP_Mucin were found to also guide *Aae*UPO secretion.

At this point, a limitation of the shuffling approach became clear. The *Aae*UPO was engineered by Molina *et al.* [208] using multiple rounds of directed evolution via mutagenic PCR and *in vivo* DNA shuffling to increase activity by 3250-fold. The other UPOs have not been heterologously expressed before or even confirmed as active enzymes. Our work demonstrates that our shuffling methodology supports the identification of suitable enzyme secretion condition. It allows high-throughput screenings for a great variety of expression conditions and can contribute to our understanding of protein secretion. However, the expression conditions alone do not enable enzyme secretion. Enzyme engineering approaches are still required to improve enzyme activity,

increase expression levels, and support protein secretion, necessary for the development of a cell factory [36].

An enzyme engineering approach may also support the secretion of the alcohol dehydrogenases (ADHs) tested. Two different ADHs, each engineered with varying number of glycosylation recognition sites, were studied with individually designed secretion strains. Despite varying numerous genetic parameters (signal peptides, promoters, or the location of a His-tag), and the medium for enzyme expression, no condition seemed to yield effective secretion. To enable ADH secretion, a codon optimised sequence might have worked better. Alternatively, *P. pastoris* itself expresses the ADHs ADH3 (XM_002491337) and ADH (FN392323) [147] that could be studied instead.

An objective of this work was to identify protein expressing strains from a diverse library, and then sequence the expression construct from strains with high enzymatic activity to identify beneficial combinations of genetic parts. Studying the randomised libraries directly would also be interesting, to gain further insight into the DNA assembly process and to calculate precise statistics of part ligations. Deep-sequencing of the libraries would provide this information [291].

The application of the PTK provides a useful solution for rapid strain engineering by enabling the construction of diverse expression libraries. The development of this approach and the supporting screening methods provides a means of identifying useful promoter and secretion tag combinations for a particular enzyme of interest and will contribute to our understanding of protein secretion. For industrial applications of *P. pastoris* cell factories, the presented approach supports ongoing enzyme engineering efforts and the optimisation of expression conditions.

6.1.3 Polyelectrolyte Capsules for Cell Screening

As our ability to assemble diverse libraries of strains improves, there is growing demand for flexible high-throughput screening technologies. Polyelectrolyte microcapsules could provide a novel cell screening platform due to the flexible design of the polyelectrolyte capsule and the modification opportunities that are available. Therefore, the ability to use polyelectrolyte microcapsules for yeast cell encapsulation was studied (Chapter 5).

Initially it was intended to encapsulate *P. pastoris* cells in calcium carbonate templates and coat these via a layer-by-layer technique using polyelectrolytes. However, the development of a screening platform for *P. pastoris* cells using polyelectrolyte capsules could not be implemented due to the inability to incorporate cells into calcium carbonate templates. The uncharged yeast surface did not mediate nucleation around the cell, which was required to create a micro-reaction vessel. Nevertheless, suitable conditions for the formation of reliably round templates and stable polyelectrolyte capsules were determined (see Chapter 2.6 for the final protocol). A few cells could be encapsulated as desired, but presumably these cells were encapsulated together with a neighbouring template. Interestingly, evaluation with flow cytometry highlighted the successful coating of most cells with a polymer shell. To make use of this alternative encapsulation

approach, more research is necessary. First, the coated cells should be tested for protein secretion, as the polyelectrolyte capsule should still allow the mass transfer of sufficient nutrients and secreted proteins could still be secreted very close to the cell surface. Tests would also need to be performed that address cell viability and recovery time, as well as the diffusion properties for the desired medium. Also, a suitable read-out approach must be developed, to be able to differentiate intracellular and secreted proteins.

Alternatively, other templates and other microorganisms could be tested. Magnesium phosphate or calcium phosphate could be an alternative crystalline template that may be more accessible to nucleation around the uncharged cell surface. To circumvent the precipitation bottleneck, microgels such as degradable dextran-based gels, alginate gels, or agarose beads could be used as reported elsewhere [71, 88, 132, 203, 229, 282]. Other commonly used expression hosts such as *E. coli* or *B. subtilis* could be also considered as their differing cell wall structure may be more accessible for template formation.

To conclude, the polyelectrolyte capsules have outstanding characteristics useful for cell screening, but many challenges remain that hamper their application. Despite the fact that only standard laboratory equipment is needed, the method is fairly laborious, requiring more than 40 centrifugation steps and a resulting hands-on bench time of >4 h. This is very stressful for the cells and could hamper the applicability of this method for screenings.

Other ultrahigh-throughput screening technologies such as microfluidic droplets may therefore offer a better approach [101]. Microfluidic droplets are less harmful to the cells, as cells are suspended in water droplets present in a water-in-oil emulsion. These droplets can be sorted according to fluorescence or adsorption signals. Microfluidic screenings also offer short-term cultivation options in incubation-chips or off-chip reservoirs and addition of nutrients and other supplements via pico-injection. However, the drawback of this approach in contrast to the polyelectrolyte microcapsules is the price of each device needed for the screening, incubation, and pico-injection of microfluidic droplets [300].

6.2 Future Perspectives

The work presented in this thesis opens up many future directions to develop cell factories important for industrial applications. New approaches could use the PTK and the developed library generation approaches to establish screenings with higher throughput, perform deeper analyses to improve *P. pastoris* as cell factory, or expand the reach of this toolkit to other hosts.

6.2.1 Future Direction for the *P. pastoris* Toolkit

An immediate future application of the PTK is its use for protein expression and secretion screenings of other industrially relevant enzymes. Our work provides the necessary methods and tools that can directly be applied to built new *P. pastoris* expression strains. The secretion of enzymes with varying characteristics such as codon usage or glycosylation patterns will further deepen our understanding of parameters influencing secretion. Due to the improved transformation procedure, higher transformation efficiencies enable enzyme engineering experiments with protein expression and screening performed directly in *P. pastoris* cells. Previously, enzymes engineering approaches required tandem expression systems, where *S. cerevisiae* was used to perform directed evolution studies [208] and once a suitable enzyme was found it was used in *P. pastoris* for fermentation studies [209]. This was mostly due to the low transformation efficiency of *P. pastoris*. In the future, this indirect engineering approach is not necessary and plasmid libraries from mutagenic PCR studied or DNA shuffling experiments can directly be used for screenings in *P. pastoris*.

The value of the modular toolkit is that more regulatory elements can be easily added to expand the functionality. Since the secretion tags have such an important influence on the resulting expression level, more tags could be investigated. Many more secretion tags are predicted [200] that provide an ideal starting point for future investigations. Additional promoters and terminators could also be incorporated to improve strain engineering. To broaden expression levels, additional promoters from other well-characterised libraries can be easily integrated. For example, promoter libraries of pGAP [250], pAOX1 [124], and synthetic promoters [303] are available for *P. pastoris*. Currently, the PTK does not include regulatable promoters that do not rely on methanol. Unfortunately, methanol is unsuitable for large-scale processes due to is toxicity and flammability [301]. Therefore, the inclusion of other regulatable promoters would also be a valuable future direction.

6.2.2 Improving *P. pastoris* as a Cell Factory

Our understanding of the *P. pastoris* metabolism is currently limited, which hampers the predictable design of efficient protein expression strains. To exploit the full potential of *P. pastoris*, a better understanding of the whole-cell metabolism under differing conditions is necessary. The tools to manage the cellular machinery in a holistic manner and tune all precursors precisely as

desired will take time to develop, but ultimately support a more holistic cell factory development [201].

Currently, the overexpression of chaperones has proven valuable in improving protein expression levels [201, 334]. New approaches could implement the overexpression of chaperones such as the heat shock factor 1, which may improve protein secretion by preventing the risk of protein accumulation [155] or trafficking proteins (e.g., Sly1p and Sec1p) which support guiding proteins from the ER to the Golgi and further outside the cell to improve protein secretion [134]. These strain engineering approaches could be realised using the CRISPRS/Cas9 system [318] or by making multi-gene constructs with the PTK.

6.2.3 Extending the *P. pastoris* Toolkit to Other Hosts

The PTK shows the ability to effectively expand the capabilities of the widely used YTK to a new host and new functionality. To take this approach even further, the extension of these toolkits to a wider spectrum of organisms is proposed. For example, other non-conventional yeasts such as *Kluyveromyces lactis*, *Hansenula polymorpha*, and *Yarrowia lipolytica* are interesting cell factories. They are designated as GRAS [189], grow fast on cheap carbon sources, and have high secretion capacities [155]. These yeasts still lack the genetic toolkits needed to control gene expression efficiently [309], but strain design faces the same challenges as for *P. pastoris*. Well-characterised promoters and a selection of secretion signal peptides is necessary to build functional expression vectors [309]. For *K. lactis* and *K. marxianus* the widely applied αMF secretion tag from *S. cerevisiae* is known to be functional, which indicates the broad functionality of secretion tags across many types of yeast [285, 330, 341]. The most commonly used procedures for transformation of *K. marxianus* or *Y. lipolytica* are still based on vector linearisation and construct integration via NHEJ, which has comparatively low transformation efficiencies [109, 193].

In this thesis, the advantages of using a recombinase-based transformation procedure for *P. pastoris* was demonstrated and its consideration for other non-conventional yeast is suggested. Alternatively, the CRISPR/Cas9 system has been implemented and improved extensively for conventional and non-conventional yeasts and allows rapid and highly efficient genome editing [258, 317]. How the CRISPR/Cas9 system is set up in each yeast varies and further optimisation may be required that consider the nuclear localization, codon optimisation and expression of the Cas9 DNA endonuclease, as well as the expression and processing of the guide RNA [258].

The expansion of the toolkit to other hosts may also open up future research in comparing types of yeast to find the optimal host organism for a protein of interest. A shared pool of genetic parts that can be easily used in a wide variety of hosts expands the available platform to tackle challenges beyond the capacity of a single type of cell.

6.3 Conclusion

This thesis describes a versatile platform for the generation of *P. pastoris* cell factories. The power of taking a synthetic biology approach to develop a toolkit that allows for the control of protein expression and secretion in *P. pastoris* was shown. This toolkit and the newly developed approach to generate "shuffled" genetic libraries of expression constructs helps to accelerate strain engineering and lays the foundation for future *P. pastoris* protein expression and secretion studies. Combining the work in this thesis with ongoing developments in related fields of enzyme engineering and high-throughput screening technologies will pave the way for systematic approaches to cell factory design.

Appendix A

A.1 Detailed Information about Part Plasmids

Table A.1: List of part plasmids

Plasmid	Part Name	Type	Origin	Plasmid	Description	pPTK	References
pUO-pp-302	pGAP	2	gBlock	Entry vector	Glyceraldehyde-3-phosphate dehydrogenase, BsaI site removed	pPTK002	[311]
pUO-pp-303	pAOX1	2	PCR product from plasmid PP117 (primer 1, 5)	Entry vector	Alcohol oxidase 1	pPTK001	[296]
pUO-pp-321	pENO1	2	gBlock	Entry vector	Enolase 1, BsaI and BsmBI site removed	pPTK003	[297]
pUO-pp-322	pTPI1	2	gBlock	Entry vector	Triose phosphate isomerase 1	pPTK004	[206]
pUO-pp-353	pPET9	2	PCR product from genomic DNA (primer 92, 93, 94, 95)	Entry vector	ADP, ATP carrier	-	[286]
pUO-pp-354	pG1	2	PCR product from genomic DNA (primer 96, 97)	Entry vector	Hypothetical protein (PAS_chr1-3_0011)	-	[246]
pUO-pp-355	pG6	2	PCR product from genomic DNA (primer 114, 109) and gBlock	Entry vector	Hypothetical protein (PAS_chr2-1_0853)	-	[246]
pUO-pp-356	pADH2	2	PCR product from genomic DNA (primer 100, 101, 102, 103)	Entry vector	Methylformate synthase, alcohol dehydrogenase	-	[304]
pUO-pp-357	pDAS1	2	PCR product from genomic DNA (primer 104, 105, 106, 107)	Entry vector	Dihydroxyacetone synthase, isoenzymes	-	[304]
pUO-pp-358	pPMP20	2	PCR product from genomic DNA (primer 98, 99)	Entry vector	Peroxisomal glutathione peroxidase	-	[304]
pUO-pp-328	αAmylase-αMFΔ	3a	PCR product from plasmid PP280 (primer 52, 61)	Entry vector	α-Amylase followed by aMFΔ	pPTK009	[228]
pUO-pp-329	αMFΔ	3a	PCR product from plasmid PP168 (primer 53, 61)	Entry vector	α-mating factor, pre-sequence shortened	pPTK007	[106, 183]
pUO-pp-330	αMFΔ _no_Kex	3a	PCR product from plasmid PP264 (primer 54, 61)	Entry vector	α-mating factor no Kex recognition site	pPTK008	[106, 183]
pUO-pp-331	Glucoamylase-αMFΔ	3a	PCR product from plasmid PP265 (primer 47, 58)	Entry vector	Glucoamylase followed by αMFΔ	pPTK010	[170]
pUO-pp-332	SA-αMFΔ	3a	PCR product from plasmid PP270 (primer 48, 61)	Entry vector	Serum albumin followed by αMFΔ	pPTK011	[162]
pUO-pp-333	Inulinase-αMFΔ	3a	PCR product from plasmid PP266 (primer 49, 61)	Entry vector	Inulinase followed by αMFΔ	pPTK012	[35]
pUO-pp-334	Invertase-αMFΔ	3a	PCR product from plasmid PP267 (primer 50, 61)	Entry vector	Invertase followed by αMFΔ	pPTK013	[29, 206]
pUO-pp-335	Killer-αMFΔ	3a	PCR product from plasmid PP281 (primer 51, 61)	Entry vector	Killer followed by αMFΔ	pPTK014	[294]

Continued on next page

Table A.1 – continued from previous page

Plasmid	Part Name	Type	Origin	Plasmid	Description	pPTK	References
pUO-pp-336	αMF	3a	gBlock, PCR modification (primer 59, 60)	Entry vector	α-mating factor	pPTK005	[106, 183]
pUO-pp-337	αMF_no_EAEA	3a	gBlock, PCR modification (primer 59, 61)	Entry vector	α-mating factor no EAEA	pPTK006	[106, 183]
pUO-pp-360	SP_Disulfide isomerase	3a	gBlock	Entry vector	Protein disulfide isomerase	-	[200]
pUO-pp-361	SP_C4R6P1	3a	gBlock	Entry vector	Hypothetical protein (C4R6P1)	-	[200]
pUO-pp-362	SP_Cell wall protein	3a	Complete DNA synthesis	Twist vector	Cell wall protein	-	[200]
pUO-pp-363	SP_Cyclophillin	3a	Complete DNA synthesis	Twist vector	Peptidyl-prolyl cis-trans isomerase	-	[200]
pUO-pp-364	SP_CSN2	3a	Complete DNA synthesis	Twist vector	Beta-casein	-	[126]
pUO-pp-365	SP_PHA-E	3a	gBlock	Entry vector	Phytohaemag-glutinin	-	[254]
pUO-pp-366	MF41	3a	Complete DNA synthesis	Twist vector	Synthetic signal peptide MF41	-	[325]
pUO-pp-367	SP_C4R8H7	3a	Complete DNA synthesis	Twist vector	Hypothetical protein (C4R8H7)	-	[200]
pUO-pp-368	SP_Peptidylprolyl isomerase	3a	gBlock	Entry vector	Hypothetical protein (Peptidylprolyl isomerase)	-	[200]
pUO-pp-369	SP_ALB	3a	PCR product, plasmid pUO-pp-332 (primer 48, 111)	Entry vector	Serum albumin	-	[162, 326]
pUO-pp-370	SP_Scw11p	3a	gBlock	Entry vector	Cell wall protein	-	[179, 200]
pUO-pp-371	SP_Mucin	3a	gBlock	Entry vector	Mucin family member	-	[200]
pUO-pp-372	SP_An_phyA	3a	gBlock	Entry vector	*A. niger* 3-phytase A	-	NYC
pUO-pp-373	SP_Pl_phyA	3a	Complete DNA synthesis	Twist vector	*P.lycii* phytase	-	NYC
pUO-pp-374	SP_Th_phyA	3a	Complete DNA synthesis	Twist vector	*T.heterothallica* phytase	-	NYC
pUO-pp-375	SP_Aae_UPO	3a	Complete DNA synthesis	Twist vector	Aromatic peroxygenase	-	[208]
pUO-pp-376	SP_Aae_UPOeng	3a	PCR product from plasmid pPICB_UPOeng (primer 112, 113)	Entry vector	Aromatic peroxygenase tag engineered	-	[208]
pUO-pp-377	SP_PHO1	3a	gBlock	Entry vector	Acid phosphatase PHO1	-	[128]
pUO-pp-378	SP_Suc2	3a	PCR product from plasmid pUO-pp-334 (primer 50, 110)	Entry vector	Invertase 2	-	[29, 167, 206, 228]
pUO-pp-324	yEGFP	3	PCR product from plasmid pLC152 (primer 43, 57)	Entry vector	Green fluorescent protein, *P. pastoris* codon optimised	pPTK015	
pUO-pp-325	yEGFP	3b	PCR product from plasmid pLC152 (primer 44, 57)	Entry vector	Green fluorescent protein, *P. pastoris* codon optimised	pPTK017	
pUO-pp-326	RFP	3	gBlock, PCR modification (primer 45, 55)	Entry vector	Red fluorescent protein, BsaI removed, *P. pastoris* codon optimised	pPTK016	

Continued on next page

Table A.1 – continued from previous page

Plasmid	Part Name	Type	Origin	Plasmid	Description	pPTK	References
pUO-pp-327	RFP	3b	PCR product from plasmid pUO-pp-326 (primer 45, 56)	Entry vector	Red fluorescent protein, BsaI removed, *P. pastoris* codon optimised	pPTK018	
pUO-pp-338	yEGFP_1-10	3b	PCR product from plasmid pUO-pp-325 (primer 44, 81)	Entry vector	Green fluorescent protein β-barrel 1-10	-	[43]
pUO-pp-339	yEGFP_11	4a	PCR product from plasmid pUO-pp-325 (primer 82, 83)	Entry vector	Green fluorescent protein β-barrel 11	-	[43]
pUO-pp-380	*An*_phA	3b	Complete DNA synthesis	Twist vector	*A. niger* 3-phytase A, *P. pastoris* codon optimised	-	[123]
pUO-pp-381	*Th*_phA	3b	Complete DNA synthesis	Twist vector	*T.heterothallica* phytase, *P. pastoris* codon optimised	-	[279]
pUO-pp-382	*Pl*_phA	3b	Complete DNA synthesis	Twist vector	*P.lycii* phytase, *P. pastoris* codon optimised	-	[169]
pUO-pp-383	*Ec*_appA	3b	Complete DNA synthesis	Twist vector	*E. coli* appA, *P. pastoris* codon optimised	-	[55]
pUO-pp-384	*Aae*UPOeng_co	3b	Complete DNA synthesis	Twist vector	*A. aegerita* aromatic peroxygenase, *P. pastoris* codon optimised	-	[208]
pUO-pp-385	*Cra*UPO_co	3b	Complete DNA synthesis	Twist vector	*C. radians* potential aromatic peroxygenase, *P. pastoris* codon optimised , *P. pastoris* codon optimised , *P. pastoris* codon optimised	-	[131]
pUO-pp-386	*Pin*UPO_co	3b	Complete DNA synthesis	Twist vector	*P. infectans* potential aromatic peroxygenase	-	[131]
pUO-pp-387	*Dse*UPO_co	3b	Complete DNA synthesis	Twist vector	*D. septosporum* potential aromatic peroxygenase	-	[131]
pUO-pp-388	*Nha*UPO_co	3b	Complete DNA synthesis	Twist vector	*N. haematococcus* potential aromatic, *P. pastoris* codon optimised peroxygenase	-	[131]
pUO-pp-389	*Afu*UPO_co	3b	Complete DNA synthesis	Twist vector	*A. fumigatus* potential aromatic peroxygenase, *P. pastoris* codon optimised	-	[131]
pUO-pp-340	His+AdhZ2_Ec	3b	PCR product from plasmid CBR P_511 (primer 64, 65)	Entry vector	*E. coli* alcohol dehydrogenase	-	[239, 240]
pUO-pp-341	His+AdhZ2_DIN	3b	PCR product from plasmid CBR P_512 (primer 64, 65)	Entry vector	*E. coli* alcohol dehydrogenase, cofactor shift towards NADH	-	[239, 240]
pUO-pp-342	His+AdhZ2_7476	3b	PCR product from plasmid CBR P_513 (primer 64, 65)	Entry vector	*E. coli* alcohol dehydrogenase	-	[223]

Continued on next page

Table A.1 – continued from previous page

Plasmid	Part Name	Type	Origin	Plasmid	Description	pPTK	References
pUO-pp-343	His+AdhZ3_LND	3b	PCR product from plasmid CBR P_514 (primer 64, 66, 67, 68)	Entry vector	*E. coli* alcohol dehydrogenase	-	[239, 240]
pUO-pp-344	His+AdhZ3_242	3b	PCR product from plasmid CBR P_515 (primer 64, 66, 67, 68)	Entry vector	*E. coli* alcohol dehydrogenase	-	[223]
pUO-pp-347	AdhZ3_LND_no_HIS	3b	PCR product from plasmid pUO-pp-343 (primer 74, 77)	Entry vector	*E. coli* alcohol dehydrogenase	-	[239, 240]
pUO-pp-348	His+AdhZ3_LND	3	PCR product from plasmid pUO-pp-343 (primer 76, 77)	Entry vector	*E. coli* alcohol dehydrogenase	-	[239, 240]
pUO-pp-349	AdhZ3_LND_no_HIS	3	PCR product from plasmid pUO-pp-343 (primer 75, 77)	Entry vector	*E. coli* alcohol dehydrogenase	-	[239, 240]
pUO-pp-350	AdhZ3_242_no_HIS	3b	PCR product from plasmid pUO-pp-344 (primer 74, 77)	Entry vector	*E. coli* alcohol dehydrogenase	-	[223]
pUO-pp-351	His+AdhZ3_242	3	PCR product from plasmid pUO-pp-344 (primer 76, 77)	Entry vector	*E. coli* alcohol dehydrogenase	-	[223]
pUO-pp-352	AdhZ3_242_no_HIS	3	PCR product from plasmid pUO-pp-344 (primer 75, 77)	Entry vector	*E. coli* alcohol dehydrogenase	-	[223]
pUO-pp-308	tAOX1	4	PCR product from plasmid PP117 (primer 46, 89)	Entry vector	Alcohol oxidase 1	pPTK019	[296]
pUO-pp-307	tAOX1	4b	PCR product from plasmid PP117 (primer 88, 89)	Entry vector	Alcohol oxidase 1	-	[296]
pUO-pp-301	attB	7	gBlock	Entry vector	BxbI recognition site, BsaI site removed	pPTK020	[234]
pUO-pp-323	PARS	7	gBlock	Entry vector	*Pichia* autonomously replicating sequence	pPTK021	[65]
pUO-pp-501	ConLS-sfGFPdrop-ConE-ZeoR-attB-KanaR	Backbone	Golden gate	-	Backbone plasmid	-	-

DNA sequence of parts

attB, pUO-pp-301

TGGCCGTGGCCGTGCTCGTCCTCGTCGGCCGGCTTGTCGACGACGGCGGTCACCGTCGTCAGGATCATCCGGGCCACAAGCT
TGCTGACAGAAGCCTCAAGAAAAAAAAAAATTCTTCTTCGACTATGCTGGAGGCAGAGATGATCGAGCCGGTAGTTAACTATAT
ATAGCTAAATTGGTTCCATCAC

pGAP, pUO-pp-302

TTTTTGTAGAAATGTCTTGGTGTCCTCGTCCAATCAGGTAGCCATCTCTGAAATATCTGGCTCCGTTGCAACTCCGAACGACCT
GCTGGCAACGTAAAATTCTCCGGGGTAAAACTTAAATGTGGAGTAATGGAACCAGAAACGTGTCTTCCCTTCTCTCTCCTTCC
ACCGCCCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCCCCTTGCAGCAATGCTCTTCCCAGCAT
TACGTTGCGGGTAAAACGGAGGTCGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGGCAATAATA
GCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAGAATCGAATATAAAAGGCGAACACCTTTCCCAATTTTGGTTTCT
CCTGACCCAAAGACTTTAAATTTAATTTATTTGTCCCTATTTCAATCAATTGAACAACTAT

pAOX1, pUO-pp-303

GATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTCCATTCTCACACATAAGTGCCA
AACGCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAAACGCAGGACCTCCACTCCTCTTCTCCTCAACACCCACTTTTG
CCATCGAAAAACCAGCCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCAATTCCTTCTATTAGGCTACTAACACCATGAC
TTTATTAGCCTGTCTATCCTGGCCCCCCTGGCGAGGTTCATGTTTGTTTATTTCCGAATGCAACAAGCTCCGCATTACACCCGA
ACATCACTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCCAAATGGCCCAAAACTGACAGTTTAAAC
GCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAACTAAGTTTGGTTCGTTGAAATGCTAACGGCCAGT
TGGTCAAAAAGAAACTTCCAAAAGTCGGCATACCGTTTGTCTTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTA
ATGCTTAGCGCAGTCTCTCTATCGCTTCTGAACCCCGGTGCACCTGTGCCGAAACGCAAATGGGGAAACACCCGCTTTTTGGA
TGATTATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTTCATGATCAAAATTT
AACTGTTCTAACCCCTACTTGACAGCAATATATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTATCATCATTATTA
GCTTACTTTCATAATTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGACAACTTGAGAAGATCAAAAA
ACAACTAATTATTCGAAACG

tAOX1, pUO-pp-307 and pUO-pp-308

TCAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAG
GATTTTTTTTGTCATTTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTGGTAGG
GGTTTGGGAAAATCATTCGAGTTTGATGTTTTTCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAGA

pENO1, pUO-pp-321

AGAAAGCATACTATACTATTCGACATTCCTTTCAATCCTGGAATTAACAGTCACTTTTAAAAAAGACATCTACCGTGAAGGTGC
CGTAGAGTATCGCGTTACCATATCGCCAAAAACTGATATACGCCGCGGAAACCAGGCAAACAATTGAAAAGAAAAATTTTGAG
GAACTCTCTGCATCGAAGCCGTCTAGAGTTACCACTAGTCAGATGCCGCGGGCACTTGAGCACCTCATGCACAGCAATAACAC
AACACAATGGTTAGTAGCAACCTGAATTCGGTCATTGATGCATGCATGTGCCGTGAAGCGGGACAACCAGAAAAGTCGTCTAT
AAATGCCGGCACGTGCGATCATCGTGGCGGGGTTTTAAGAGTGCATATCACAAATTGTCGCATTACCGCGGAACCGCCAGATA
TTCATTACTTGACGCAAAAGCGTTTGAAATAATGACGAAAAAGAAGGAAGAAAAAAAAAGAAAAATACCGCTTCTAGGCGGGT
TATCTACTGATCCGAGCTTCCACTAGGATAGCACCCAAACACCTGCATATTTGGACGACCTTTACTTACACCACCAAAAACCAC
TTTCGCCTCTCCCGCCCCTGATAACGTCCACTAATTGAGCGATTACCTGAGCGGTCCTCTTTTGTTTGCAGCATGAGACTTGCA
TACTGCAAATCGTAAGTAGCAACcTCTCAAGGTCAAAACTGTATGGAAACCTTGTCACCTCACTTAATTCTAGCTAGCCTACCC
TGCAAGTCAAGAGcTCTCCGTGATTCCTAGCCACCTCAAGGTATGCCTCTCCCCGGAAACTGTGGCCTTTTCTGGCACACATGA
TCTCCACGATTTCAACATATAAATAGCTTTTGATAATGGCAATATTAATCAAATTTATTTTACTTCTTTCTTGTAACATCTCTCT
TGTAATCCCTTATTCCTTCTAGCTATTTTTCATAAAAAACCAAGCAACTGCTTATCAACACACAAACACTAAATCAAA

pTPI1, pUO-pp-322

GTGTTTAAAGATTACGGATATTTAACTTACTTAGAATAATGCCATTTTTTTGAGTTATAATAATCCTACGTTAGTGTGAGCGGG
ATTTAAACTGTGAGGACCTTAATACATTCAGACACTTCTGCGGTATCACCCTACTTATTCCCTTCGAGATTATATCTAGGAACC
CATCAGGTTGGTGGAAGATTACCCGTTCTAAGACTTTTCAGCTTCCTCTATTGATGTTACACCTGGACACCCCTTTTCTGGCAT
CCAGTTTTTAATCTTCAGTGGCATGTGAGATTCTCCGAAATTAACTAAAGCAATCACACAATTCTCTCGGATACCACCTCGGTT
GAAACTGACAGGTGGTTTGTTACGCATGCTAATGCAAAGGAGCCTATATACCTTTGGCTCGGCTGCTGTAACAGGGAATATAA
AGGGCAGCATAATTTAGGAGTTTAGTGAACTTGCAACATTTACTATTTTCCCTTCTTACGTAAATATTTTTCTTTTTAATTCTAA
ATCAATCTTTTTCAATTTTTTGTTTGTATTCTTTTCTTGCTTAAATCTATAACTACAAAAAACACATACATAAACTAAAA

PARS, pUO-pp-323

CCTTCGTTTGTGCGGATCCAATTAATATTTACTTATTTTGGTCAACCCCAAATAGGTTGATTTCATACTTGGTTCATTCAAAAAT
AAGTAGTCTTTTGAGATCTTTCAATATTATAATAAATATACTATAACAGCCGACTTGTTTCATTTTCGCGAATGTTCCCCCAGCT
TATCG

yEGFP, pUO-pp-324 and pUO-pp-325

GTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCA
GCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGT
GCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGAC
TTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGC
CGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACT
TCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCC
CGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGAGCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC
CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG

RFP, pUO-pp-326 and pUO-pp-327

GCAACTTCCGGTATGGTGTCAAAGGGAGAGGAAAATAATATGGCTATTATTAAGGAGTTTATGCGTTTTAAGGTACATATGGA
AGGTTCTGTCAACGGTCACGAATTCGAAATTGAAGGTGAGGGGGAGGGGAGGCCATACGAGGGAACTCAGACTGCTAAGTTA
AAGGTCACTAAAGGTGGTCCTTTACCTTTCGCCTGGGATATCCTGTCTCCACAGTTTATGTACGGTTCAAAGGCTTATGTGAAA
CATCCTGCCGATATCCCAGATTATCTTAAACTTTCTTTCCCTGAGGGTTTTAAGTGGGAGAGGGTAATGAACTTTGAAGACGG
TGGTGTGGTCACTGTTACTCAGGACTCAAGTCTGCAGGACGGTGAGTTCATCTACAAGGTGAAGCTGAGAGGTACCAATTTTC
CATCAGATGGTCCCGTGATGCAAAAAAAGACAATGGGTTGGGAAGCTTCTAGTGAACGTATGTATCCCGAAGATGGAGCTTTG
AAAGGTGAAATTAAGCAAAGACTAAAACTTAAGGATGGTGGACATTACGATGCTGAAGTTAAGACGACCTACAAGGCCAAAAA
GCCAGTCCAGTTGCCTGGAGCATACAATGTTAACATCAAATTGGATATAACTTCCCATAATGAAGACTATACCATCGTCGAGC
AATACGAACGAGCCGAAGGGAGACACAGTACTGGTGGTATGGATGAACTTTATAAAGGATCCGGAACCGCA

***α*Amylase-*α*MFΔ, pUO-pp-328**

ATGGTGGCATGGTGGTCCTTATTCTTATATGGTCTTCAAGTTGCTGCTCCTGCCCTTGCTATGAGATTTCCTAGTATTTTCACT
GCTGTGCTATTTGCCGCTAGTTCCGCTCTAGCTGCTCCAGTTAATACTACTACTGAAGATGAATTGGAGGGTGACTTCGATGT
TGCTGTTCTGCCTTTTTCCGCTTCTATCGCAGCCAAGGAAGAAGGTGTATCTCTAGAGAAGCGT

***α*MFΔ, pUO-pp-329**

ATGAGATTTCCTAGTATTTTCACTGCTGTGCTATTTGCCGCTAGTTCCGCTCTAGCTGCTCCAGTTAATACTACTACTGAAGAT
GAATTGGAGGGTGACTTCGATGTTGCTGTTCTGCCTTTTTCCGCTTCTATCGCAGCCAAGGAAGAAGGTGTATCTCTAGAGAA
GCGT

***α*MFΔ_no_Kex, pUO-pp-330**

ATGAGATTCCCATCAATTTTTACTGCTGTTCTGTTCGCCGCTTCTAGTGCACTTGCCATGAGATTTCCTAGTATTTTCACTGCT
GTGCTATTTGCCGCTAGTTCCGCTCTAGCTGCTCCAGTTAATACTACTACTGAAGATGAATTGGAGGGTGACTTCGATGTTGC
TGTTCTGCCTTTTTCCGCTTCTATCACAGCCAAGGAAGAAGGTGTATCTCTAGAGAAGCGT

Glucoamylase- *α*MFΔ, pUO-pp-331

ATGTCTTTCAGATCCCTATTGGCATTGTCAGGGTTGGTCTGTTCTGGATTGGCTATGAGATTTCCTAGTATTTTCACTGCTGTG
CTATTTGCCGCTAGTTCCGCTCTAGCTGCTCCAGTTAATACTACTACTGAAGATGAATTGGAGGGTGACTTCGATGTTGCTGTT
CTGCCTTTTTCCGCTTCTATCGCAGCCAAGGAAGAAGGTGTATCTCTAGAGAAGCGT

SA-*α*MFΔ, pUO-pp-332

ATGAAGTGGGTAACTTTCATCTCATTGTTATTCTTGTTCTCCTCTGCTTACTCTATGAGATTTCCTAGTATTTTCACTGCTGTGC
TATTTGCCTCTAGTTCCGCTCTAGCTGCTCCAGTTAATACTACTACTGAAGATGAATTGGAGGGTGACTTCGATGTTGCTGTTC
TGCCTTTTTCCGCTTCTATCGCAGCCAAGGAAGAAGGTGTATCTCTAGAGAAGCGT

Inulinase-*α*MFΔ, pUO-pp-333

ATGAAACTGGCTTACTCCCTGTTGCTACCTCTGGCTGGAGTTTCCGCTATGAGATTTCCTAGTATTTTCACTGCTGTGCTATTT

GCCGCTAGTTCCGCTCTAGCTGCTCCAGTTAATACTACTACTGAAGATGAATTGGAGGGTGACTTCGATGTTGCTGTTCTGCC
TTTTTCCGCTTCTATCGCAGCCAAGGAAGAAGGTGTATCTCTAGAGAAGCGT

Invertase-αMFΔ, pUO-pp-334

ATGTTATTGCAAGCTTTTTTATTTCTGCTGGCAGGTTTTGCAGCAAAGATTTCTGCCATGAGATTTCCTAGTATTTTCACTGCT
GTGCTATTTGCCGCTAGTTCCGCTCTAGCTGCTCCAGTTAATACTACTACTGAAGATGAATTGGAGGGTGACTTCGATGTTGC
TGTTCTGCCTTTTTCCGCTTCTATCGCAGCCAAGGAAGAAGGTGTATCTCTAGAGAAGCGT

Killer-αMFΔ, pUO-pp-335

ATGACCAAACCAACGCAAGTCTTAGTTCGTTCAGTCTCTATTTTATTCTTCATCACACTGTTGCACTTGGTTGTTGCAATGAGA
TTTCCTAGTATTTTCACTGCTGTGCTATTTGCCGCTAGTTCCGCTCTAGCTGCTCCAGTTAATACTACTACTGAAGATGAATTG
GAGGGTGACTTCGATGTTGCTGTTCTGCCTTTTTCCGCTTCTATCGCAGCCAAGGAAGAAGGTGTATCTCTAGAGAAGCGT

αMF, pUO-pp-336

ATGAGATTTCCTTCAATTTTTACTGCTGTTTTATTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGAT
GAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTC
CAACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGA
AAAGAGAGGCTGAAGCT

aMF_no_EAEA, pUO-pp-337

ATGAGATTTCCTTCAATTTTTACTGCTGTTTTATTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGAT
GAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTC
CAACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGA
AAAGAG

yEGFP_1-10, pUO-pp-338

GTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCA
GCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGT
GCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGAC
TTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGC
CGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACT
TCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCC
CGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGAGCGCCCTGAGCAAAGACCCCAACGAGAAG

yEGFP_11, pUO-pp-339

GATGGAGGGTCTGGTGGCGGATCAACAAGTCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACT

AdhZ2_Ec+His, pUO-pp-340

CATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGAAGATCAAAGCTGTTGGTGCATA
TTCCGCTAAACAACCACTTGAACCGATGGATATCACCCGGCGTGAACCGGGACCGAATGATGTCAAAATCGAAATCGCTTACT
GTGGCGTTTGCCATTCCGATCTCCACCAGGTCCGTTCCGAGTGGGCGGGGACGGTTTACCCCTGCGTGCCGGGTCATGAAAT
TGTGGGGCGTGTGGTAGCCGTTGGTGATCAGGTAGAAAAATATGCGCCGGGCGATCTGGTCGGTGTCGGCTGCATTGTCGAC
AGTTGTAAACATTGCGAAGAGTGTGAAGACGGGTTGGAAAACTACTGTGATCACATGACCGGCACCTATAACTCGCCGACGCC
GGACGAACCGGGCCATACTCTGGGCGGCTACTCACAACAGATCGTCGTTCATGAGCGATATGTTCTGCGTATTCGTCACCCGC
AAGAGCAGCTGGCGGCGGTGGCTCCTTTGTTGTGTGCAGGGATCACCACGTATTCGCCGCTACGTCACTGGCAGGCCGGGCC
GGGTAAAAAAGTGGGCGTGGTCGGCATCGGCGGTCTGGGACATATGGGGATTAAGCTGGCCCACGCGATGGGGGCACATGT
GGTGGCATTTACCACTTCTGAGGCAAAACGCGAAGCGGCAAAAGCCCTGGGGGCCGATGAAGTTGTTAACTCACGCAATGCC
GATGAGATGGCGGCTCATCTGAAGAGTTTCGATTTCATTTTGAATACAGTAGCTGCGCCACATAATCTCGACGATTTTACCAC
CTTGCTGAAGCGTGATGGCACCATGACGCTGGTTGGTGCGCCTGCGACACCGCATAAATCGCCGGAAGTTTTCAACCTGATCA
TGAAACGCCGTGCGATAGCCGGTTCTATGATTGGCGGCATTCCAGAAACTCAGGAGATGCTCGATTTTTGCGCCGAACATGGC
ATCGTGGCTGATATAGAGATGATTCGGGCCGATCAAATTAATGAAGCCTATGAGCGAATGCTGCGCGGTGATGTGAAATATCG
TTTTGTTATCGATAATCGCACACTAACAGAC

AdhZ2_DIN+His, pUO-pp-341

CATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGAAGATCAAAGCTGTTGGTGCATA
TTCCGCTAAACAACCACTTGAACCGATGGATATCACCCGGCGTGAACCGGGACCGAATGATGTCAAAATCGAAATCGCTTACT
GTGGCGTTTGCCATTCCGATCTCCACCAGGTCCGTTCCGAGTGGGCGGGGACGGTTTACCCCTGCGTGCCGGGTCATGAAAT
TGTGGGGCGTGTGGTAGCCGTTGGTGATCAGGTAGAAAAATATGCGCCGGGCGATCTGGTCGGTGTCGGCTGCATTGTCGAC
AGTTGTAAACATTGCGAAGAGTGTGAAGACGGGTTGGAAAACTACTGTGATCACATGACCGGCACCTATAACTCGCCGACGCC
GGACGAACCGGGCCATACTCTGGGCGGCTACTCACAACAGATCGTCGTTCATGAGCGATATGTTCTGCGTATTCGTCACCCGC
AAGAGCAGCTGGCGGCGGTGGCTCCTTTGTTGTGTGCAGGGATCACCACGTATTCGCCGCTACGTCACTGGCAGGCCGGGCC
GGGTAAAAAAGTGGGCGTGGTCGGCATCGGCGGTCTGGGACATATGGGGATTAAGCTGGCCCACGCGATGGGGGCACATGT
GGTGGCATTTGATATTAATGAGGCAAAACGCGAAGCGGCAAAAGCCCTGGGGGCCGATGAAGTTGTTAACTCACGCAATGCC
GATGAGATGGCGGCTCATCTGAAGAGTTTCGATTTCATTTTGAATACAGTAGCTGCGCCACATAATCTCGACGATTTTACCAC
CTTGCTGAAGCGTGATGGCACCATGACGCTGGTTGGTGCGCCTGCGACACCGCATAAATCGCCGGAAGTTTTCAACCTGATCA
TGAAACGCCGTGCGATAGCCGGTTCTATGATTGGCGGCATTCCAGAAACTCAGGAGATGCTCGATTTTTGCGCCGAACATGGC
ATCGTGGCTGATATAGAGATGATTCGGGCCGATCAAATTAATGAAGCCTATGAGCGAATGCTGCGCGGTGATGTGAAATATCG
TTTTGTTATCGATAATCGCACACTAACAGAC

AdhZ2_7476+His, pUO-pp-342

CATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGAAGATCAAAGCTGTTGGTGCATA
TTCCGCTAAACAACCACTTGAACCGATGGATATCACCCGGCGTGAACCGGGACCGAATGATGTCAAAATCGAAATCGCTTACT
GTGGCGTTTGCCATTCCGATCTCCACCAGGTCCGTTCCGAGTGGGCGGGGACGGTTTACCCCTGCGTGCCGGGTCATGAAAT
TGTGGGGCGTGTGGTAGCCGTTGGTGATAACTGTACAAAATATGCGCCGGGCGATCTGGTCGGTGTCGGCTGCATTGTCGAC
AGTTGTAAACATTGCGAAGAGTGTGAAGACGGGTTGGAAAACTACTGTGATCACATGACCGGCACCTATAACTCGCCGACGCC
GGACGAACCGGGCCATACTCTGGGCGGCTACTCACAACAGATCGTCGTTCATGAGCGATATGTTCTGCGTATTCGTCACCCGC
AAGAGCAGCTGGCGGCGGTGGCTCCTTTGTTGTGTGCAGGGATCACCACGTATTCGCCGCTACGTCACTGGCAGGCCGGGCC
GGGTAAAAAAGTGGGCGTGGTCGGCATCGGCGGTCTGGGACATATGGGGATTAAGCTGGCCCACGCGATGGGGGCACATGT
GGTGGCATTTACCACTTCTGAGGCAAAACGCGAAGCGGCAAAAGCCCTGGGGGCCGATGAAGTTGTTAACTCACGCAATGCC
GATGAGATGGCGGCTCATCTGAAGAGTTTCGATTTCATTTTGAATACAGTAGCTGCGCCACATAATCTCGACGATTTTACCAC
CTTGCTGAAGCGTGATGGCACCATGACGCTGGTTGGTGCGCCTGCGACACCGCATAAATCGCCGGAAGTTTTCAACCTGATCA
TGAAACGCCGTGCGATAGCCGGTTCTATGATTGGCGGCATTCCAGAAACTCAGGAGATGCTCGATTTTTGCGCCGAACATGGC
ATCGTGGCTGATATAGAGATGATTCGGGCCGATCAAATTAATGAAGCCTATGAGCGAATGCTGCGCGGTGATGTGAAATATCG
TTTTGTTATCGATAATCGCACACTAACAGAC

AdhZ3_LND+His, pUO-pp-348 and pUO-pp-343

CATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGTCGATGATAAAAAGCTATGCCGCAAAAGAAG
CGGGCGGCGAACTGGAAGTTTATGAGTACGATCCCGGTGAGCTGAGGCCACAAGATGTTGAAGTGCAGGTGGATTACTGCGG
GATCTGCCATTCCGATCTGTCGATGATCGATAACGAATGGGGATTTTCACAATATCCGCTGGTTGCCGGGCATGAGGTGATTG
GGCGCGTGGTGGCACTCGGGAGCGCCGCGCAGGATAAAGGTTTGCAGGTCGGTCAGCGTGTCGGGATTGGCTGGACGGCGC
GTAGCTGTGGTCACTGCGACGCCTGTATTAGCGGTAATCAGATCAACTGCGAGCAAGGTGCGGTGCCGACGATTATGAATCG
CGGTGGCTTTGCCGAGAAGTTGCGTGCGGACTGGCAATGGGTGATTCCACTGCCAGAAAATATTGATATCGAGTCCGCCGGG
CCGCTGTTGTGCGGCGGTATCACGGTCTTTAAACCACTGTTGATGCACCATATCACTGCTACCAGCCGCGTTGGGGTAATTGG
TATTGGCGGGCTGGGGCATATCGCTATAAAACTTCTGCACGCAATGGGATGCGAGGTGACAGCCTTTCTTAATGATCCGGCGA
AAGAGCAGGAAGTGCTGGCGATGGGTGCCGATAAAGTGGTGAATAGCCGCGATCCGCAGGCACTGAAAGCACTGGCGGGGC
AGTTTGATCTCATTATCAACACCGTCAACGTCAGCCTCGACTGGCAGCCCTATTTTGAGGCGCTGACCTATGGCGGTAATTTC
CATACGGTCGGTGCGGTTCTCACGCCGCTGTCTGTTCCGGCCTTTACGTTAATTGCGGGCGATCGCAGCGTgtctGGTTCTGCT
ACCGGCACGCCTTATGAGCTGCGTAAGCTGATGCGTTTTGCCGCCCGCAGCAAGGTTGCGCCGACCACCGAACTGTTCCCGAT
GTCGAAAATTAACGACGCCATCCAGCATGTGCGCGACGGTAAGGCGCGTTACCGCGTGGTGTTGAAAGCCGATTTT

AdhZ3_242+His, pUO-pp-351 and pUO-pp-344

CATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGTCGATGATAAAAAGCTATGCCGCAAAAGAAG
CGGGCGGCGAACTGGAAGTTTATGAGTACGATCCCGGTGAGCTGAGGCCACAAGATGTTGAAGTGCAGGTGGATTACTGCGG
GATCTGCCATTCCGATCTGTCGATGATCGATAACGAATGGGGATTTTCACAATATCCGCTGGTTGCCGGGCATGAGGTGATTG
GGCGCGTGGTGGCACTCGGGAGCGCCGCGCAGGATAAAGGTTTGCAGGTCGGTCAGCGTGTCGGGATTGGCTGGACGGCGC
GTAGCTGTGGTCACTGCGACGCCTGTATTAGCGGTAATCAGATCAACTGCGAGCAAGGTGCGGTGCCGACGATTATGAATCG
CGGTGGCTTTGCCGAGAAGTTGCGTGCGGACTGGCAATGGGTGATTCCACTGCCAGAAAATATTGATATCGAGTCCGCCGGG
CCGCTGTTGTGCGGCGGTATCACGGTCTTTAAACCACTGTTGATGCACCATATCACTGCTACCAGCCGCGTTGGGGTAATTGG

TATTGGCGGGCTGGGGCATATCGCTATAAAACTTCTGCACGCAATGGGATGCGAGGTGACAGCCTTTCTTAATGATCCGGCGA
AAGAGCAGGAAGTGCTGGCGATGGGTGCCGATAAAGTGGTGAATAGCCGCGATCCGCAGGCACTGAAAGCACTGGCGGGGC
AGTTTGATCTCATTATCAACACCGTCAACGTCGCCCTCGACTGGCAGCCCTATTTTGAGGCGCTGACCTATGGCGGTAATTTC
CATACGGTCGGTGCGGTTCTCACGCCGCTGTCTGTTCCGGCCTTTACGTTAATTGCGGGCGATCGCAGCGTgtctGGTTCTGCT
ACCGGCACGCCTTATGAGCTGCGTAAGCTGATGCGTTTTGCCGCCCGCAGCAAGGTTGCGCCGACCACCGAACTGTTCCCGAT
GTCGAAAATTAACGACGCCATCCAGCATGTGCGCGACGGTAAGGCGCGTTACCGCGTGGTGTTGAAAGCCGATTTT

AdhZ3_LND_no_HIS, pUO-pp-349 and pUO-pp-347

ATGTCGATGATAAAAAGCTATGCCGCAAAAGAAGCGGGCGGCGAACTGGAAGTTTATGAGTACGATCCCGGTGAGCTGAGGC
CACAAGATGTTGAAGTGCAGGTGGATTACTGCGGGATCTGCCATTCCGATCTGTCGATGATCGATAACGAATGGGGATTTTCA
CAATATCCGCTGGTTGCCGGGCATGAGGTGATTGGGCGCGTGGTGGCACTCGGGAGCGCCGCGCAGGATAAAGGTTTGCAG
GTCGGTCAGCGTGTCGGGATTGGCTGGACGGCGCGTAGCTGTGGTCACTGCGACGCCTGTATTAGCGGTAATCAGATCAACT
GCGAGCAAGGTGCGGTGCCGACGATTATGAATCGCGGTGGCTTTGCCGAGAAGTTGCGTGCGGACTGGCAATGGGTGATTCC
ACTGCCAGAAAATATTGATATCGAGTCCGCCGGGCCGCTGTTGTGCGGCGGTATCACGGTCTTTAAACCACTGTTGATGCACC
ATATCACTGCTACCAGCCGCGTTGGGGTAATTGGTATTGGCGGGCTGGGGCATATCGCTATAAAACTTCTGCACGCAATGGGA
TGCGAGGTGACAGCCTTTCTTAATGATCCGGCGAAAGAGCAGGAAGTGCTGGCGATGGGTGCCGATAAAGTGGTGAATAGCC
GCGATCCGCAGGCACTGAAAGCACTGGCGGGGCAGTTTGATCTCATTATCAACACCGTCAACGTCAGCCTCGACTGGCAGCC
CTATTTTGAGGCGCTGACCTATGGCGGTAATTTCCATACGGTCGGTGCGGTTCTCACGCCGCTGTCTGTTCCGGCCTTTACGT
TAATTGCGGGCGATCGCAGCGTGTCTGGTTCTGCTACCGGCACGCCTTATGAGCTGCGTAAGCTGATGCGTTTTGCCGCCCGC
AGCAAGGTTGCGCCGACCACCGAACTGTTCCCGATGTCGAAAATTAACGACGCCATCCAGCATGTGCGCGACGGTAAGGCGC
GTTACCGCGTGGTGTTGAAAGCCGATTTT

AdhZ3_242_no_HIS, pUO-pp-352 and pUO-pp-350

ATGTCGATGATAAAAAGCTATGCCGCAAAAGAAGCGGGCGGCGAACTGGAAGTTTATGAGTACGATCCCGGTGAGCTGAGGC
CACAAGATGTTGAAGTGCAGGTGGATTACTGCGGGATCTGCCATTCCGATCTGTCGATGATCGATAACGAATGGGGATTTTCA
CAATATCCGCTGGTTGCCGGGCATGAGGTGATTGGGCGCGTGGTGGCACTCGGGAGCGCCGCGCAGGATAAAGGTTTGCAG
GTCGGTCAGCGTGTCGGGATTGGCTGGACGGCGCGTAGCTGTGGTCACTGCGACGCCTGTATTAGCGGTAATCAGATCAACT
GCGAGCAAGGTGCGGTGCCGACGATTATGAATCGCGGTGGCTTTGCCGAGAAGTTGCGTGCGGACTGGCAATGGGTGATTCC
ACTGCCAGAAAATATTGATATCGAGTCCGCCGGGCCGCTGTTGTGCGGCGGTATCACGGTCTTTAAACCACTGTTGATGCACC
ATATCACTGCTACCAGCCGCGTTGGGGTAATTGGTATTGGCGGGCTGGGGCATATCGCTATAAAACTTCTGCACGCAATGGGA
TGCGAGGTGACAGCCTTTCTTAATGATCCGGCGAAAGAGCAGGAAGTGCTGGCGATGGGTGCCGATAAAGTGGTGAATAGCC
GCGATCCGCAGGCACTGAAAGCACTGGCGGGGCAGTTTGATCTCATTATCAACACCGTCAACGTCGCCCTCGACTGGCAGCC
CTATTTTGAGGCGCTGACCTATGGCGGTAATTTCCATACGGTCGGTGCGGTTCTCACGCCGCTGTCTGTTCCGGCCTTTACGT
TAATTGCGGGCGATCGCAGCGTgtctGGTTCTGCTACCGGCACGCCTTATGAGCTGCGTAAGCTGATGCGTTTTGCCGCCCGCA
GCAAGGTTGCGCCGACCACCGAACTGTTCCCGATGTCGAAAATTAACGACGCCATCCAGCATGTGCGCGACGGTAAGGCGCG
TTACCGCGTGGTGTTGAAAGCCGATTTT

pPET9, pUO-pp-353

TAGAAAATTCACCACTGTCGGAAAGTTGTCTACTTCCGTCGGTTGGAAATACGAGTCTGTTGTTGAGAAGTTGGAGGAGAAGA
GAAAGGCTGAGGAAGCTGAGTACCAGGAGAAGAAGAGAGCTTACACCCAGAGATTAGACGCAGCTAGTGCCGAGTTTGCCCA
AACCGAGGAGGGAAAGCAGTTGGCTGCCTTTGGTTACTAAATAGTAAAGTAGGGTATCTTCAAGTAATAGTATACTAACCATC
TGAAATAACCACCGTCCTGTAGTTTTTTTTCGATATCGAAGAGCCTATGCTAGTACTGTGGATTTGCGCTCCATCCAACATCTG
TGCGCAAACTAAAACTTCCGAGACTGACATCTACCATCGCTAGACCCTAAGTAAAACCAATCTCGCGTCCGAACTTTTAAATTT
CAGTCCTTAAAACTTCAGAGCATTGGTTGTAGTTTCCGGATCTGAGGGGTCGTATTGGAGTCAAGAcACGGAGCTGCCTCCAC
AGCGCGAAACGTCAACCCCAACACCAACCTGAATTTGCAATCACCATGGGGACAAGTTTCAGCAGTCAATGGGCAATTCAGAC
GTTGATACGGTACCCATTTGCTAAGCTCAATGACGATCCATCCAACTTCAGAGAAAGGCCTTTCTCTGGTATGCTCTGGTATTC
ATTCGTCTTTTATCACTCTCGTTGCACAATGCCCGGGTACTCCCGGAACAAGGGAGTCTTCCAGCCAAGCTGTACAGAGTGAA
AAATAGAAATACACCTTTGCAATCAAGACGCGCGTTGGCCAATCACAAGACTTAATCGGTGCAAAGAAGGATTACCAAATTTT
TTTTTCCCAAAATCGCTATATAGAAATAATGGAGGAAAAAGGGTTAATATAAAGGAGAATTCCCCCGTTTTTCTCCCCTTTTCT
TTTCTTCTTCAGGCTTTCTTACAAATCTATAATATTCCAAAATGGCTGACAACAACAAGTCTAACTTCTTCGTCGACTTC

pG1, pUO-pp-354

CAAACATTTGCTCCCCCTAGTCTCCAGGGAAATGTAAAATATACTGCTAATAGAAAACAGTAAGACGCTCAGTTGTCAGGATA
ATTACGTTCGACTGTAGTAAAACAGGAATCTGTATTGTTAGAAAGAACGAGAGTTTTTTACGGCGCCGCCATATTGGGCCGTG
TGAAAACAGCTTGAAACCCCACTACTTTCAAAGGTTCTGTTGCTATACACGAACCATGTTTAACCAACCTCGCTTTTGACTTGA

CTGAAGTCATCGGTTAACAATCAAGTACCCTAGTCTGTCTGAATGCTCCTTTCCATATTCAGTAGGTGTTTCTTGCACTTTTGC
ATGCACTGCGGAAGAATTAGCCAATAGCGCGTTTCATATGCGCTTTTACCCCCTCTTTTGTCAAGCGCAAAATGCCTGTAAGA
TTTGGTGGGGGTGTGAGCCGTTAGCTGAAGTACAACAGGCTAATTCCCTGAAAAAACTGCAGATAGACTTCAAGATCTCAGGG
ATTCCCACTATTTGGTATTCTGATATGTTTTTCCTGATATGCATCAAAACTCTAATCTAAAACCTGAATCTCCGCTATTTTTTTT
TTTTTTTGATGACCCCGTTTTCGTGACAAATTAATTTCCAACGGGGTCTTGTCCGGATAAGAGAATTTTGTTTGATTATCCGTT
CGGATAAATGGACGCCTGCTCCATATTTTTCCGGTTATTACCCCACCTGGAAGTGCCCAGAATTTTCCGGGGATTACGGATAA
TACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACATTTTGTTGCAGTCTCGTGCGAGTATGTGCAATAATAAACAAGATGA
GCCAATTTATTGGATTAGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACTTGG
ATGCAGTGAGTTTTGGAGTATAAAAGATCCTTAAAATTCCACCCTT

pG6, pUO-pp-355

GACCAGCAGTTTAACTACGCAAATCCACAGGAATTTCTACATCACAATACCAATGGTAATACCACGACGTCAAGGAATGGAAA
CGACGACTTGGAGGAAGACTTCGTCAACCTCTTGCGGAGTACCCGAGGCTAAGACAATAAGAAGAAAAAAAAAAGAAAAGCGG
TGGGGGAGGGATTATTAAATAAGGATTATGTAACCCCAGGGTACCGTTCTATACATATTTAAGGATTATTTAGGACAATCGAT
GAAATCGGCATCAAACTGGATGGGAGTATAGTGTCCGGATAATCGGATAAATCATCTTGCGAGGAGCCGCTTGGTTGGTTGGT
GAGAGGAGTGAAATATGTGTCTCCTCACCCAAGAATCGCGATATCAGCACCCTGTGGGGGACACTATTGGCCTCCCTCCCAAA
CCTTCGATGTGGTAGTGCTTTATTATATTGATTACATTGATTACATAGCTAAACCCTGCCTGGTTGCAAGTTGAGCTCCGAATT
CCAATATTAGTAAAATGCCTGCAAGATAACCTCGGTATGGCGTCCGACCCCGCTTAATTATTTTAACTCCTTTCCAACGAGGAC
TTCGTAATTTTTGATTAGGGAGTTGAGAAACGGGGGGTCTTGATACCTCCTCGATTTCAGATCCCACCCCCTCTCAGTCCCAA
GTGGGACCCCCCTCGGCCGTGAAATGCGCGCACTTTAGTTTTTTTCGCATGTAAACGCCGGTGTCCGTCAATTAAAAGTCGCA
GACTAGGGTGAACTTTACCATTTTTGTCGCACTCCGTgTCCTCGGAATAGGGGTGTAGTAATTCTGCAGTAGTGCAATTTTTAC
CCCGCCAAGGGGGGGCGAAAAGAcACGACCTCATCACGCATTCTCCAGTCGCTCTCTACGCCTACAGCACCGACGTAGTTAAC
TTTCTCCCATATATAAAGCAATTGCCATTCCCCTGAAAACTTTAACCTCTGCTTTTTCTTGATTTTTCCTTGCCCAAAGAAAAG

pADH2, pUO-pp-356

CGCAGCGTTTTCTGACGGTACTAGAGGACTCTTAGGGGAAGGTAGAATCAATAAAGATCATATTAGGTAAGCAAATTTTGGAT
GGAATAGGAGACTAGGTGTGGATGCGCGATCTCGCCAAATTGCACGACCAGAGTGGATGCCGGATGGTGGTAAACCGTTTCT
TCCTTTTTACCACCCAAGTGCGAGTGAAACACCCCATGGCTGCTCTCCGATTGCCCCTCTACAGGCATAAGGGTGTGACTTTG
TGGGCTTGAATTTTACACCCCCTCCAACTTTTCTCGCATCAATTGATCCTGTTACCAATATTGCATGCCCGGAGGAGACTTGCC
CCCTAATTTCGCGGCGTCGTCCCGGATCGCAGGGTGAGACTGTAGAcACCCCACATAGTGACAATGATTATGTAAGAAGAGGG
GGGTGATTCGGCCGGCTATCGAACTCTAACAACTAGGGGGGTGAACAATGCCCAGCAGTCCTCCCCACTCTTTGACAAATCAG
TATCACCGATTAACACCCCAAATCTTATTCTCAACGGTCCCTCATCCTTGCACCCCTCTTTGGACAAATGGCAGTTAGCATTGG
TGCACTGACTGACTGCCCAACCTTAAACCCAAATTTCTTAGAAGGGGCCCATCTAGTTAGCGAGGGGTGAAAAATTCCTCCAT
CGGAGATGTATTGACCGTAAGTTGCTGCTTAAAAAAAATCAGTTCAGATAGCGAGACTTTTTTGATTTCGCAACGGGAGTGCC
TGTTCCATTCGATTGCAATTCTCACCCCTTCTGCCCAGTCCTGCCAATTGCCCATGAATCTGCTAATTTCGTTGATTCCCACCC
CCCTTTCCAACTCCACAAATTGTCCAATCTCGTTTTCCATTTGGGAGAATCTGCATGTCGACTACATAAAGCGACCGGTGTCCG
AAAAGATCTGTGTAGTTTTCAACATTTTGTGCTCCCCCCGCTGTTTGAAAACGGGGGTGAGCGCTCTCCGGGGTGCGAATTCG
TGCCCAATTCCTTTCACCCTGCCTATTGTAGACGTCAACCCGCATCTGGTGCGAATATAGCGCACCCCCAATGATCACACCAA
CAATTGGTCCACCCCTCCCCAATCTCTAATATTCACAATTCACCTCACTATAAATACCCCTGTCCTGCTCCCAAATTCTTTTTTC
CTTCTTCCATCAGCTACTAGCTTTTATCTTATTTACTTTACGAAA

pDAS1, pUO-pp-357

AATAAAAAAAACGTTATAGAAAGAAATTGGACTACGATATGCTCCAATCCAAATTGTCAAAATTGACCACCGAAAAAGAACAATT
GGAATTTGACAAGAGGAACAACTCACTAGATTCTCAAACGGAGCGTCACCTAGAGTCAGTTTCCAAGTCAATTACAGAAAGTT
TGGAAACAGAAGAGGAGTATCTACAATTGAATTCCAAACTTAAAGTCGAGCTGTCCGAATTCATGTCGCTAAGGCTTTCTTAC
TTGGACCCCATTTTTGAAAGTTTCATTAAAGTTCAGTCAAAAATTTTCATGGACATTTATGACACATTAAAGAGCGGACTACCT
TATGTTGATTCTCTATCCAAAGAGGATTATCAGTCCAAGATCTTGGACTCTAGAATAGATAACATTCTGTCGAAAATGGAAGCG
CTGAACCTTCAAGCTTACATTGATGATTAGAGCAATGATATAAACAACAATTGAGTGACAGGTCTACTTTGTTCTCAAAAGGCC
ATAACCATCTGTTTGCATCTCTTATCACCACACCATCCTCCTCATCTGGCCTTCAATTGTGGGGAACAACTAGCATCCCAACAC
CAGACTAACTCCACCCAGATGAAACCAGTTGTCGCTTACCAGTCAATGAATGTTGAGCTAACGTTCCTTGAAACTCGAATGAT
CCCAGCCTTGCTGCGTATCATCCCTCCGCTATTCCGCCGCTTGCTCCAACCATGTTTCCGCCTTTTTCGAACAAGTTCAAATAC
CTATCTTTGGCAGGACTTTTCCTCCTGCCTTTTTTAGCCTCAGGTgTCGGTTAGCCTCTAGGCAAATTCTGGTCTTCATACCTAT
ATCAACTTTTCATCAGATAGCCTTTGGGTTCAAAAAAGAACTAAAGCAGGATGCCTGATATATAAATCCCAGATGATCTGCTTT
TGAAACTATTTTCAGTATCTTGATTCGTTTACTTACAAACAACTATTGTTGATTTTATCTGGAGAATAATCGAACAAA

pPMP20, pUO-pp-358

TTCTGGAGTGTCAAAACAGTAGTGATAAAAGGCTATGAAGGAGGTTGTCTAGGGGCTCGCGGAGGAAAGTGATTCAAACAGA
CCTGCCAAAAAGAGAAAAAAGAGGGAATCCCTGTTCTTTCCAATGGAAATGACGTAACTTTAACTTGAAAAATACCCCAACCA
GAAGGGTTCAAACTCAACAAGGATTGCGTAATTCCTACAAGTAGCTTAGAGCTGGGGGAGAGACAACTGAAGGCAGCTTAAC
GATAACGCGGGGGGATTGGTGCACGACTCGAAAGGAGGTATCTTAGTCTTGTAACCTCTTTTTTCCAGAGGCTATTCAAGATT
CATAGGCGATATCGATGTGGAGAAGGGTGAACAATATAAAAGGCTGGAGAGATGTCAATGAAGCAGCTGGATAGATTTCAAA
TTTTCTAGATTTCAGAGTAATCGCACAAAACGAAGGAATCCCACCAAGCAAAAAAAAAAAATCTAAG

SP_Disulfide isomerade, pUO-pp-360

ATGAAAATATTAAGTGCATTGCTTCTTCTTTTTACGTTGGCCTTTGCT

SP_C4R6P1, pUO-pp-361

ATGTGGTCGCTGTTCATATCTGGACTATTAATCTTCTATCCTTTGGTCCTTGGA

SP_Cell wall protein, pUO-pp-362

ATGAGGCCAGTGCTTTCGTTATTACTCTTGCTGGCTTCTTCGGTACTCGCT

SP_Cyclophilin, pUO-pp-363

ATGAAATTGTTGAACTTTCTGCTTAGCTTCGTAACTCTGTTCGGACTATTATCAGGTTCTGTGTTTGCA

SP_CSN2, pUO-pp-364

ATGAAGGTCCTCATCCTTGCCTGCCTGGTGGCTCTGGCCCTTGCA

SP_PHA-E, pUO-pp-365

ATGGCTTCCTCCAACTTACTCTCCCTAGCCCTCTTCCTTGTGCTTCTCACCCACGCAAACTCA

MF41, pUO-pp-366

ATGGCTATTCCAAGATTCCCATCTATCTTCATCGCTGTCTTGTTCGCTGCTTCTTCTGCCTTGGCTGCTCCTGTCAACACTACT
ACCGAGGATGAAACTGCTCAAATCCCTGCTGAGGCTGTCATCGGTTACTCTGACCTGGAGGGTGACTTCGACGTCGCTGTCTT
GCCATTCTCTAACTCCACCAACAACGGTTTGTTGGAGGAGGCTGAAGCTGAAGCTGAACCTAAATTCATCAACACTACTATCG
CTTCTATCGCTGCTAAGGAGGAGGGTGTTTCCCTCGAG

SP_C4R8H7, pUO-pp-367

ATGAGCACCCTGACATTGCTGGCTGTGCTGTTGTCGCTTCAAAATTCAGCTCTTGCT

SP_Peptidylprolyl isomerade, pUO-pp-368

ATGAAAGTTTCTACGACCAAATTTCTGGCTGTGTTCTTATTAGTTAGACTCGTTTGCGCT

SP_ALB, pUO-pp-369

ATGAAGTGGGTAACTTTCATCTCATTGTTATTCTTGTTCTCCTCTGCTTACTCT

SP_Scw11p, pUO-pp-370

ATGCTATCAACTATCTTAAATATCTTTATCCTGTTGCTCTTCATACAGGCATCCCTACAG

SP_Mucin, pUO-pp-371

ATGATTAATTTAAACTCCTTTCTTATACTTACAGTAACACTGTTATCTCCAGCTTTGGCACTTCCAAAGAATGTTCTTGAAGAAC
AACAAGCAAAAGACGACCTTGCTAAAAGA

SP_An_phyA, pUO-pp-372

ATGGGCGTTTCTGCTGTTCTACTTCCTTTGTATCTCCTGTCTGGAGTCACCTCCGGA

SO_Pl_phyA, pUO-pp-373

ATGGTTTCTTCGGCATTCGCACCTTCCATCCTACTTAGCTTGATGTCGAGTCTTGCTTTGAGCACGCAGTTCAGCTTTGTTGCGGCG

SP_Th_phyA, pUO-pp-374

ATGACTGGCCTCGGAGTGATGGTGGTGATGGTCGGCTTCCTGGCGATCGCCTCTCTACAATCC

SP_Aae_UPO, pUO-pp-375

ATGAAATATTTTCCCCTGTTCCCAACCTTGGTCTTCGCAGCGAGGGTCGTTGCTTTTCCTGCCTACGCCTCATTGGCCGGCCTC
AGCCAGCAGGAATTGGACGCTATAATCCCAACACTCGAGGCCCGA

SO_Aae_UPOeng, pUO-pp-376

ATGAAATATTTTCCCCTGTTCCCAACCTTGGTCTACGCAGTGGGGGTCGTTGCTTTTCCTGACTACGCCTCATTGGCCGGCCTC
AGCCAGCAGGAATTGGACGCTATAATCCCAACACTCGAGGCCCGA

SP_PHO1, pUO-pp-377

ATGTTTTCTCCTATTCTAAGTCTGGAAATTATTCTCGCTTTGGCTACTCTCCAATCAGTCTTTGCG

SP_Suc2, pUO-pp-378

ATGTTATTGCAAGCTTTTTTATTTCTGCTGGCAGGTTTTGCAGCAAAGATTTCTGCC

Phytase *A.niger*, pUO-pp-380

TTGGCTGTTCCAGCCTCCAGAAATCAATCCTCCTGTGACACTGTTGACCAGGGTTACCAATGTTTCTCTGAGACTTCTCACCTG
TGGGGTCAATACGCTCCTTTCTTCTCTTTGGCCAACGAATCCGTTATCTCCCCAGAAGTTCCAGCCGGTTGTAGAGTTACTTTC
GCTCAGGTTTTGTCCAGACACGGTGCTAGATACCCAACTGACTCTAAGGGTAAGAAGTACTCCGCCTTGATCGAAGAGATTCA
GCAGAACGCTACTACCTTCGACGGAAAGTACGCTTTCCTTAAGACCTACAACTACTCCTTGGGTGCTGACGACTTGACTCCAT
TCGGTGAACAAGAGTTGGTCAACTCCGGTATCAAGTTCTACCAGAGATACGAGTCCCTGACCAGAAACATCGTCCCATTCATT
AGATCTTCCGGTTCCTCCAGAGTTATCGCCTCTGGAAAGAAGTTCATCGAGGGTTTCCAGTCCACCAAGTTGAAGGATCCAAG
AGCACAACCAGGTCAGTCCTCTCCAAAGATCGACGTTGTTATTTCTGAGGCCTCCTCCTCCAACAACACTTTGGATCCAGGTA
CTTGCACCGTTTTCGAGGACTCTGAATTGGCTGATACTGTCGAGGCTAACTTCACTGCTACTTTCGTCCCATCCATCAGACAGA
GATTGGAGAACGACTTGTCCGGTGTTACTTTGACTGACACCGAGGTCACTTACCTGATGGACATGTGTTCCTTCGACACTATC
TCCACTTCCACCGTCGACACTAAGTTGTCTCCATTCTGCGACTTGTTCACTCACGACGAGTGGATCAACTACGACTACTTGCAG
TCCCTGAAAAAGTACTACGGTCACGGTGCTGGTAACCCATTGGGTCCAACTCAAGGTGTTGGTTACGCTAACGAGTTGATCGC
TAGATTGACTCACTCTCCAGTTCACGACGACACCTCTTCTAACCATACTTTGGACTCTTCCCCAGCTACTTTCCCATTGAACTC
TACCTTGTACGCTGACTTCTCTCACGACAACGGTATCATCTCCATCTTGTTCGCTCTGGGTCTGTACAACGGTACTAAGCCATT
GTCCACTACTACCGTCGAGAACATCACTCAGACTGACGGTTTTTCTTCCGCTTGGACTGTTCCATTCGCTTCCAGACTTTACGT
CGAGATGATGCAATGTCAGGCTGAGCAAGAGCCATTGGTTAGAGTTTTGGTCAACGACAGAGTTGTCCCATTGCACGGTTGTC
CAGTTGATGCTTTGGGTAGATGTACCAGAGACTCCTTCGTTAGAGGTTTGTCCTTCGCTAGATCTGGTGGTGATTGGGCTGAA
TGTTTCGCT

Phytase *T. heterothallica*, pUO-pp-381

GAATCCAGACCATGTGACACTCCAGACTTGGGTTTCCAATGTGGTACTGCTATTTCCCACTTCTGGGGTCAATACTCCCCATA
CTTCTCTGTTCCATCTGAGTTGGACGCTTCCATTCCAGACGACTGTGAAGTTACTTTCGCCCAGGTTTTGTCCAGACATGGTGC
TAGAGCACCAACTTTGAAGAGAGCTGCTTCCTACGTTGACCTGATCGACAGAATTCACCACGGTGCTATTTCTTACGGTCCAG
GTTACGAGTTCCTGAGAACTTACGACTACACTTTGGGTGCTGACGAGTTGACTAGAACTGGTCAACAGCAGATGGTCAACTCC
GGTATCAAGTTCTACAGAAGATACAGAGCCCTGGCCAGAAAGTCCATTCCATTCGTTAGAACTGCTGGTCAGGACAGAGTTGT
TCACTCCGCTGAAAACTTCACTCAGGGTTTCCACTCTGCTTTGTTGGCTGACAGAGGTTCCACTGTTAGACCAACCTTGCCATA
CGACATGGTTGTCATCCCAGAAACAGCTGGTGCTAACAACACCTTGCACAACGACTTGTGTACTGCCTTCGAAGAAGGTCCAT
ACTCCACTATTGGTGACGACGCTCAAGACACTTACTTGTCCACTTTCGCTGGTCCAATCACCGCTAGAGTTAACGCTAATTTGC
CAGGTGCTAACTTGACTGACGCTGACACTGTTGCTTTGATGGACTTGTGTCCTTTCGAGACTGTTGCCTCCTCTTCTTCTGATC
CTGCTACTGCTGATGCTGGTGGTGGTAATGGTAGACCATTGTCTCCATTCTGCAGACTGTTCTCTGAATCTGAGTGGCGTGCC
TACGACTACTTGCAATCTGTTGGTAAATGGTACGGTTACGGTCCTGGTAACCCATTGGGTCCTACTCAAGGTGTCGGTTTTGT
CAACGAGTTGTTGGCTAGATTGGCCGGTGTTCCAGTTAGAGATGGTACTTCCACTAACAGAACCTTGGACGGTGACCCAAGAA
CCTTTCCATTGGGTAGACCACTGTACGCTGATTTCTCTCACGACAACGACATGATGGGTGTTTTGGGTGCATTGGGAGCTTAC
GATGGTGTTCCACCATTGGACAAGACTGCCAGAAGAGATCCTGAAGAACTTGGTGGTTACGCTGCTTCTTGGGCTGTTCCATT
TGCTGCCAGAATCTACGTCGAGAAGATGAGATGTTCTGGTGGCGGTGGTGGTGGCGGAGGCGGTGAAGGTAGACAAGAAAAG
GATGAAGAGATGGTCAGAGTCCTGGTCAACGACAGAGTCATGACCTTGAAAGGTTGTGGTGCTGATGAGAGGGGTATGTGTA
CTTTGGAGAGATTCATCGAGTCCATGGCTTTCGCAAGAGGTAACGGAAAGTGGGATTTGTGTTTCGCT

Phytase *P. lycii*, pUO-pp-382

CAGTTGCCAATTCCAGCTCAGAACACTTCTAACTGGGGTCCATACGATCCATTCTTCCCAGTTGAACCATACGCTGCTCCACCA
GAAGGTTGTACTGTTACCCAGGTTAACTTGATCCAGAGACACGGTGCTAGATGGCCAACTTCTGGTGCAAGATCTAGACAGGT
TGCTGCTGTTGCTAAGATCCAGATGGCTAGACCATTCACTGACCCAAAGTACGAGTTCCTGAACGACTTCGTCTACAAGTTCG
GTGTTGCTGACTTGTTGCCATTCGGTGCTAACCAGTCTCACCAAACTGGTACTGACATGTACACCAGATACTCCACCTTGTTC
GAAGGTGGTGACGTTCCATTTGTTAGAGCTGCTGGTGACCAGAGAGTTGTTGACTCTTCCACTAACTGGACTGCTGGTTTCGG
TGATGCTTCCGGTGAAACTGTTTTGCCAACCTTGCAGGTTGTCCTGCAAGAAGAGGGTAACTGTACCCTGTGTAACAACATGT
GTCCAAACGAGGTTGACGGTGACGAATCCACTACTTGGTTGGGTGTTTTCGCTCCAAACATCACCGCTAGATTGAACGCTGCT
GCTCCATCTGCTAACTTGTCTGATTCTGACGCCTTGACCTTGATGGACATGTGCCCTTTTGACACTTTGTCCTCTGGTAACGCT
TCCCCATTCTGTGACTTGTTCACTGCCGAAGAATACGTGTCCTACGAGTACTACTACGACCTGGACAAGTACTACGGTACTGG
TCCTGGTAATGCTTTGGGTCCAGTTCAAGGTGTTGGTTACGTCAACGAGTTGTTGGCCAGATTGACTGGTCAAGCCGTTAGAG
ATGAGACTCAGACCAACAGAACTCTGGATTCTGACCCAGCTACTTTCCCACTGAACAGAACTTTCTACGCTGACTTCTCCCAC
GACAACACCATGGTTCCAATTTTTGCTGCCCTGGGTTTGTTCAACGCTACTGCTTTGGATCCATTGAAGCCAGACGAGAACAG
ATTGTGGGTTGACTCCAAGTTGGTTCCATTCTCCGGTCACATGACCGTTGAAAAGTTGGCCTGTTCTGGTAAAGAGGCCGTCA
GAGTTTTGGTTAACGACGCTGTTCAACCATTGGAGTTCTGTGGTGGTGTTGATGGTGTTTGTGAGTTGTCCGCTTTCGTTGAG
TCCCAGACTTACGCTAGAGAAAACGGTCAAGGTGACTTCGCCAAGTGTGGTTTCGTTCCATCTGAA

Phytase *E. coli*, pUO-pp-383

CAATCTGAGCCAGAGTTGAAGTTGGAGTCCGTTGTCATCGTTTCCAGACATGGTGTTAGAGCCCCAACTAAGGCTACTCAATT
GATGCAGGATGTTACTCCAGACGCTTGGCCAACTTGGCCTGTTAAGTTAGGTTGGTTGACTCCAAGAGGTGGTGAGTTGATTG
CTTACTTGGGTCACTACCAGAGACAGAGATTGGTTGCTGATGGTTTGTTGGCTAAGAAGGGTTGTCCACAATCCGGTCAGGTT
GCTATTATTGCTGACGTTGACGAGAGAACCAGAAAGACTGGTGAAGCTTTCGCTGCTGGTTTGGCTCCAGATTGTGCTATCAC
TGTTCACACTCAAGCCGACACTTCATCCCCAGATCCTTTGTTCAACCCACTTAAGACCGGTGTTTGCCAGTTGGACAACGCTAA
CGTTACTGACGCTATCTTGTCTAGAGCCGGTGGTTCCATTGCTGATTTCACTGGTCATAGACAGACCGCCTTCAGAGAATTGG
AGAGAGTCTTGAACTTCCCACAGTCCAACCTGTGTTTGAAGAGAGAGAAGCAAGACGAGTCCTGCTCCTTGACTCAAGCTTTG
CCATCTGAGTTGAAGGTTTCCGCTGACAACGTTTCCTTGACTGGTGCTGTTTCTTTGGCCTCCATGCTGACCGAGATTTTCTTG
TTGCAGCAAGCTCAAGGTATGCCAGAACCAGGTTGGGGTAGAATTACTGACTCTCACCAGTGGAACACCTTGTTGTCCTTGCA
CAACGCTCAGTTCTACTTGCTGCAGAGAACTCCAGAAGTTGCTAGATCCAGAGCTACCCCTTTGTTGGACTTGATTAAGACCG
CTTTGACCCCACATCCACCACAAAAGCAAGCTTACGGTGTTACTTTGCCAACCTCCGTCTTGTTCATTGCTGGTCACGACACAA
ACCTGGCTAACCTTGGTGGTGCTTTGGAGTTGAACTGGACTTTGCCAGGTCAACCAGATAACACTCCACCAGGTGGTGAATTG
GTTTTCGAGAGATGGCGTAGATTGTCCGACAACTCCCAATGGATTCAGGTTTCCTTGGTGTTCCAAACCTTGCAGCAGATGAG
AGACAAGACCCCATTGTCCTTGAATACTCCTCCAGGTGAGGTTAAGTTGACCTTGGCTGGTTGTGAGGAAAGAAACGCTCAGG
GTATGTGTTCCTTGGCCGGTTTCACTCAAATTGTCAACGAGGCTAGAATCCCAGCCTGTTCTTTG

***Aae*UPOeng_co, pUO-pp-384**

GAACCAGGTTTACCACCAGGTCCATTGGAAAACTCTTCCGCCAAGTTGGTTAACGACGAAGCTCATCCATGGAAGCCACTTAG
ACCAGGTGATATCAGAGGTCCATGTCCAGGTTTGAACACTTTGGCTTCTCACGGTTACTTGCCAAGAAACGGTGTTGCTACTC
CAGCTCAGATCATCAACGCTGTTCAAGAGGGTTTCAACTTCGACAACCAGGCTGCTATCTTCGCTACTTACGCTGCTCATTTG
GTCGACGGTAACTTGATCACTGACTTGCTGTCCATCGGTAGAAAGACCAGATTGACTGGTCCAGATCCACCACCACCAGCTTC
AGTTGGTGGTTTGAACGAACACGGTACTTTCGAAGGTGACGCTTCCATGACTAGAGGTGATGCTTTCTTCGGTAACAACCACG
ACTTCAACGAGACTTTGTTCGAGCAGTTGGTCGACTACTCCAACAGATTCGGTGGTGGTAAGTACAACTTGACTGTTGCTGGT
GAGCTGAGGTTCAAGAGAATCCAAGACTCCATTGCCACCAATCCAAACTTCAGCTTCGTGGACTTCAGATTCTTCACTGCCTA
CGGTGAGACTACTTTCCCAGCCAACTTGTTCGTTGACGGTAGAAGAGATGACGGTCAGTTGGATATGGACGCTGCCAGATCAT
TCTTCCAGTTCTCTAGAATGCCAGACGACTTCTTCAGAGCCCCATCTCCAAGATCTGGTACTGGTGTTGAGGTTGTTGTTCAG
GCTCATCCAATGCAGCCAGGTAGAAACGTTGGTAAGATCAACTCCTATACTGTCGACCCAACTTCCTCCGACTTCTCCACTCC
ATGTTTGATGTACGAGAAGTTCGTCAACATCACCGTCAAGTCTCTGTACCCAAATCCAACCGTCCAGTTGAGAAAGGCCTTGA
ACACTAACCTGGACTTTCTGTTCCAAGGTGTTGCTGCTGGTTGTACCCAAGTTTTCCCATACGGTAGAGAT

***Cra*UPO_co, pUO-pp-385**

CCACCACCAGAATACGTTGGTCCAAAGTTGGTTAACGACGCTGATCATCCATGGGAGCCACTTAGACCAGGTGATATTAGAGG
TCCATGTCCAGGTTTGAACACCTTGGCTTCTCATGGTTACTTGCCAAGAAACGGTGTTGCTACTCCAGCTCAGATCATCAACG
CTATCGTTGAGGGTTTCAACTTCAACTACGAGGGTGCCGTTTTCGTTACTTACTTCGCTCACATCGTCGACGGTAACTTGGTTA
CTGACTTGTTGTCCATCGGTGGTAAGACCAACTTGACTGGTGAAGATACTGGTGCTCCAGCCATTATCGGTGGATTGAACACT
CACTCTGTTTTCGAAGGTGACGCCTCCATGACTAGAGATGATTTCCACTTCGGTGACAACCACTCCTTCAACCAGACTTTGTTC
GACCAGTTCGTCGAGTACTCCAACACTTACGGTGGTGGTTTCTACAATCAAGAGGTTGCCGGTCACCTGAGAAGAAGAAGGAT

TGAGCAATCCATTGCCACCAATCCAGAGTTCGACTTCACTTCCCCAAGATTCTTCACTGCTTTCGCCGAATCCTCCTTTCCATA
CTCATTCTTCGTGGACGGTAGAATCACCGAAAGACCTGGTGGTTTGTCTATGGAAAACGCCACCTTGTTCTTCAGGGACCACA
AGATGCCAGATGACTTCTGG

***Pin*UPO_co, pUO-pp-386**

CAACCTACTGCTCAACCATCCTTGACTACTGGTCAACACGAGTACTTCAGACCACACGACGATCAGGTTTACGGTTTGCCAGG
TAACAACTCCTCCACCTACTTCAGATCTCCATGTCCAGCTTTGAACACCTTGGCTAACCATGGTCACATCCCAAGAGATGGAAA
GTCCTTGACTCCAACCGTTCTTGGTGACGGTATCGTCAAGGTTTACAACTTCGACAAGAAGCTGCTGGACGTCATCTTCTTGG
CTTTGCCATCCAAGTTCACTTTGGCTGATTTGGGTGACCCCAACTTCATTGATCACGACGCTTCATTGGTCCACGATGACTCAT
TCTTCCAGGTCGAGCCATTCAAGGTCAACAAGACTTTGGTCGACGAGTTGTTGTCCTCCGCTGAAGATATTGGTGGTCACTCC
AACAGAGTCCTGACCAAGAATACTGTTGCCAGATTCAGACGTCACCGTGAGACTGAATGTGCCAGAACTAACCCAGAGTTCTC
TATGTCTGCTTTGGCTTCCTTCGTTGCTAACGGTGAAGCTTCTTTCGTCTTGCAAGGTTTGGGTGATTACTCCTCCGCTACTAT
TTCTGTTGACCACGCCAGATCTTTCTTGGTGGACGAGAGAATTCCAACCGATTTCAGACCATCCAAGACTCCCATCACCATCA
CTTCCGTTTTGTTGGCTATCGCCGAGTTGCAAATCAGAGCT

***Dse*UPO_co, pUO-pp-387**

TACCCAGCTTTGGAACAAGCTGCTTCTTCCGCTGAGTTCAAAGAGTACCAGAAGCAAGAGAAGAGACAGACCTTGGGTTTCGA
CGCTGCTTCTCAAATCGTTTCCACTACTGGTGATCACGCTTGGCAAGCTCCAGGTGCTAACGATATTAGAGGTCCATGTCCTG
GTTTGAACTCCATGGCTAACCACGGTTACATCCCAAGAAACGGTTACACTTCCGACGCTCAGATTATCGCTGCTATGCAGGCT
GTTTTCAACATCTCTCCAGACTTCGGTGGTTTCTTGACCGTTTTGGGTTCTGCTATGGGTGGTGATGGTTTGGGTTTTTCCATT
GGTGGTCCACCATCTGCTTCTCTGTTGACTGCTACTGGTTTGGTTGGTAAGCCACAGGGTATGTCCAACACTCACAACAGATT
CGAGTCCGACCAGTCCATCACTAGAGATGACTTGTACCAGACCGGTAACGACGTCACCTTGAACATGAACTTTTTCCAGGACC
TGCTGAACAGCTCCTTGCCAAAAGGTTGGTACGACATTGACGTCTTGGGTAACCACGCCGTTAAGAGATTCCAATACTCCGTC
GCTAACAACCCCTACTTCTTCAAGGGTTTGAACACCGCTTTTATCCCAGAGGCTACTTCCGCTTTGGTCACTTACTTGTTCGCT
AACCACTCTGCTGCTTGTCCAGCTGGTTGTTTGGACGCTACTAATCTGAAGTCCTTCTACTCCGTTACTGGTTCCGGTTCCACT
TTGAAGTACACTCCAGGTCACGAGAGAATCCCAGACAACTGGTACAAATACCCAGTCGGTTACGGTGTCGCTAACGTTTTCGC
TGACATGGTTACCGTCTACTCCAAGTACTCCAACCAGGCTGCTTTTGGTGGTAATACCGGTACTGTTAACTCCTTCACCGGTTT
GGACGTTGCTAACATCACTGGTGGTGTTTACAACGCCGAGACTTTGTTGCAGGGTAACAACTTGGGTTGCTTCCTGTTCAACG
GTATGGAATTCTTCATGCCCGACCTGATTTCCAACGGTGGTGTTATTGGTGACGTTTCCGGTGTTGTTTCCTCCTTGACTGGTA
CTATCACTTCCTTGTTGGCCCCATTCAACTGCCCAAAGTTGTCCGGTATTGACAAGAAGGCTTTCGCTATCTACCCAGGTTGGA
ACGATGGTAAGCCTAGAAAG

***Nha*UPO_co, pUO-pp-388**

GCTAGAAGGCCAACTCAAGGTGCTATGGAACCAGCTCCAAAGGCTGACTTGAACATTGGTGACCAATCCTTGGGTACTGCCCC
AGTTAGAAAGGGTGACTTGTGTCAATCCTTCTTGTCCGTCTTGCCATTCCTTGGTGTTGGTGTTAACGGTTTCCCATCCTACAT
GAACGGTCCAATCGGTGACATCGTTTCCTCTAGATTGTCCGACAAGGTCGAGGGTGTTTTTCACGAAGCTCACGAGAAGAGAT
TGCTGTTCGACCCATTGACTAAGCCAATCGACGTTTCTGGTGACCACAAGTTCATTGCTCCAGACTACTCTAGAGGTGCTCAG
AGAGGTCCTTGTCCAGGTTTGAACGCTTTGGCTAACCACGGTTACATCAACCGTAAGGGTGTTACTTCCTTGACCGAGGTTAC
TGGTGCCATCAACAAGATCTTCGGTATGGGTCTTGAGCTGTCCACTATCTTGTCTGTCATGGGTACTGTGTTCGTCGGTAACC
CATTGTCTTTGAACCCAGGTTTCTCTATTGGTGACACTGCTTCCGGTGCTCAAAACTTGTTGGGTAACTTGGCTGGTTTGCTGG
GTACTCCAAGAGGTTTGACTGGTTCCCACAACATTATTGAGGGTGACTCCTCCAACACTAGAGCCGACTTGTACGTTACAGGT
GACGCTTCTACTTTGGTCCTGCAACAGTTCAAGGACTTCTACGAAATGTCCTCCGGTGAAGGTGACTACAACTTCGACGTTTT
CGCTGAGAGAGCCTACATCAGATTCCACGAATCCGTTGCCACCAATCCAAACTTCTACTACGGTCCATTCACCGGTATGATCG
CTAGAAACGCTGGTTACTTGTTCGCCTGTAGAATGTTCGCTAACCACTCCTCTGAAAACCCATCCGGTTTCTTGAACAAAGAG
GTCCTGAAGTCATTCTTCGCCATCGAAGGTGAGGGTGATAAGCTGACTTACAAGAGAGGTTACGAGAGAATCCCAGAGAACT
GGTACAGAAGGCCAGTCGACTACTCTTTGGTCCACTTGAACTTGGACGTCTTGGCTTTGGCTGCTAAGTACCCAGAATTGGCT
TCCATCGGTGGTAACATGGGTGAAGTTAACTCCTACGCTGGTGTTGACTTGTCCGACTTGACTGGTGGTGTTTTGAACTTGAC
CAAGCTGTTGGAGGGTAACAACTTGCTGTGCTTCGTTTTCGAGATCGTCAAGACTGTTGCCCCAAACTCCTTGTCTACCCTGTT
CAAGATTATCGAGGTCCCATTGAAGTTGGTCACCGACACTTTGGGTGCTGCTATCTTGGATTTGACCTGTCCAGCCTTTAAGG
ACTTGACCGTTGGTGGTAAGTCTTTCGAAGAGGGTATCCAGGTTCAATTCCCAGGTGCTAAGTTGGGTGCTTCCGTTTTG

***Afu*UPO_co, pUO-pp-389**

CACAAGTACTGTGCTCTGTCCAACACCAGACCATGTTTGAGGCCATTGCAGACTACCATGTCTAAGATGGCTGAGTTGGAGAT
CATCGACTCCAAGGGTTTGCCATTGATCAAGGGTGATTACTGTCCAGCTGGTCCAGGTCATTTGAGAGCACCTTGTCCTGTTT
TGAACTCCTTGGCTAACCACGGTATTATCGCCAGATCCGGTAGAAACATTACTGCCGCTGAATTGAAGGCCGCCTTGAGATAT

CTTGGTATGGGTATCGACGTCATCACCATCTTGGTTAACGGTGCCTTCAAGGTTCACTCTGACGATCCAAAGAAGGGTGCCTT
GTTGGGTTTGAGAGACAAGGACCAAACTAACGAGGACGGTGTTCCAGTCTTGAACTTGGATCAAGTTGGTAGACCACACGCT
GTTGAACACGACGTTTCCGTTACTAGACAGGACAGAGCTTTGGGTGACTGCATGAGAGTTAACGCTGACCTGTTGGAGAGATT
CTTGGCTGCTCCAAAGACTGAGAGAGGTTTCTCTGCTTCTGCCTTCGGTAAGTACAGAAAGACCAGATACAACGAGCAGAAGA
GGGACAACCCAGCTTTGGAATTCGACAGATTCAACCACTTCTCCGGTTGTGCTGAATTGGGTGCTGTTCAGTGTATCTTCGGT
AGAGGTTTCCCATACAGAGTCCCAGAAGAGTACATCAGAGTCCTGTTCGGTGAAGAGAGACTGCCAATTGAAGAAGGTTGGA
AGCCAAGAAGATTGCCATTGCTGTTGCCAGAGTTGGCCCCAGTTATTTTGAGAATCTCCCACTTCGCTTCCCCATTT

Table A.2: List of part plasmids used from the YTK

Plasmid	**Type**	**Name**	***E. coli* Antibiotic Marker**
pYTK001	entry vector	Part Plasmid Entry Vector	Chloramphenicol
pYTK002	1	ConLS	Chloramphenicol
pYTK009	2	pTDH3	Chloramphenicol
pYTK010	2	pCCW12	Chloramphenicol
pYTK011	2	pPGK1	Chloramphenicol
pYTK012	2	pHHF2	Chloramphenicol
pYTK013	2	pTEF1	Chloramphenicol
pYTK014	2	pTEF2	Chloramphenicol
pYTK015	2	pHHF1	Chloramphenicol
pYTK016	2	pHTB2	Chloramphenicol
pYTK017	2	pRPL18B	Chloramphenicol
pYTK044	234r	GFP dropout	Chloramphenicol
pYTK051	4	tENO1	Chloramphenicol
pYTK052	4	tSSA1	Chloramphenicol
pYTK053	4	tADH1	Chloramphenicol
pYTK054	4	tPGK1	Chloramphenicol
pYTK055	4	tENO2	Chloramphenicol
pYTK056	4	tTDH1	Chloramphenicol
pYTK060	4a	6XHis_3XFlag	Chloramphenicol
pYTK072	5	ConRE	Chloramphenicol
pYTK080	6	ZeocinR	Chloramphenicol
pYTK084	8	KanR-ColE1	Kanamycin

A.2 List of Primers

Table A.3: Primer

Number	Name	Sequence 5´-3´
Cloning of promoters		
1	FW-2-pAOX1	gcatcgtctcatcggtctcaaacgGATCTAACATCCAAAGACGAAAGGTTGAATGA
5	REV-2-pAOX1	atgccgtctcaggtctcacataCGTTTCGAATAATTAGTTGTTTTTTGATCTTCTCAAGT
92	FW-2-pPET9	gcatcgtctcatcggtctcaaacgTAGAAAATTCACCACTGTCGGAAAGTTG
93	REV-2-pPET9	atgccgtctcaggtctcacatagatctGAAGTCGACGAAGAAGTTAGACTTGTTG
94	FW-Fix-pPET9	gcatcgtctcacACGGAGCTGCCTCCAC
95	REV-Fix-pPET9	atgccgtctcaCGTgTCTTGACTCCAATACGACCCCTC
96	FW-2-pG1	gcatcgtctcatcggtctcaaacgCAAACATTTGCTCCCCCTAGTCTC
97	REV-2-pG1	atgccgtctcaggtctcacatagatctAAGGGTGGAATTTTAAGGATCTTTTATACTC
98	FW-2-pPMP20	gcatcgtctcatcggtctcaaacgTTCTGGAGTGTCAAAACAGTAGTGATAAAAGG
99	REV-2-pPMP20	atgccgtctcaggtctcacatagatctCTTAGATTTTTTTTTTTGCTTGGTGGGATTCC
100	FW-2-pADH2	gcatcgtctcatcggtctcaaacgCGCAGCGTTTTCTGACGG
101	REV-2-pADH2	atgccgtctcaggtctcacatagatctTTTCGTAAAGTAAATAAGATAAAAGCTAGTAGCTG
102	FW-Fix-pADH2	gcatcgtctcacACCCCACATAGTGACAATGATTATGTAAGAAG
103	REV-Fix-pADH2	atgccgtctcaGGTgTCTACAGTCTCACCCTGCGATC
104	FW-2-pDAS1	gcatcgtctcatcggtctcaaacgAATAAAAAAACGTTATAGAAAGAAATTGGACTACG
105	REV-2-pDAS1	atgccgtctcaggtctcacatagatctTTTGTTCGATTATTCTCCAGATAAAATCAACAATA GTTG
106	FW-Fix-pDAS1	gcatcgtctcagTCGGTTAGCCTCTAGGCAAATTCTG
107	REV-Fix-pDAS1	atgccgtctcaCGAcACCTGAGGCTAAAAAAGGCAG
109	REV-Fix-pG6	atgccgtctcTTCACCCTAGTCTGCGACTTTTAATTG
114	FW-2-pG6	gcatcgtctcatcggtctcaaacgTGACCAGCAGTTTAACTACGCAAATC
Cloning of signal peptides		
47	FW-3a-pp-Gluco-new	gcatcgtctcatcggtctcatatgtctttcagatccctattggcattg
48	FW-3a-pp-hBA-new	gcatcgtctcatcggtctcatatgaagtgggtaactttcatctcattgttattc
49	FW-3a-pp-Inuli-new	gcatcgtctcatcggtctcatatgaaactggcttactccctgttgc
50	FW-3a-pp-Inver-new	gcatcgtctcatcggtctcatatgttattgcaagctttttttatttctgctgg
51	FW-3a-pp-Killer-new	gcatcgtctcatcggtctcatatgaccaaaccaacgcaagtcttag
52	FW-3a-pp-AlphaAmy-new	gcatcgtctcatcggtctcatatggtggcatggtggtcc
53	FW-3a-pp-AlphaKex-new	gcatcgtctcatcggtctcatatgagatttcctagtattttcactgctgtg
54	FW-3a-pp-AlphaT-new	gcatcgtctcatcggtctcatatgagattcccatcaatttttactgctgttc
58	REV-3a-pp-Gluco-Linker	atgccgtctcaggtctcaagaaccACGCTTCTCTAGAGATACACCTTCTTCC
59	FW-3a-pp-Alpha-Factor	gcatcgtctcatcggtctcatatgagatttccttcaatttttactgctgttttattc
60	REV-3a-pp-Alpha-Factor	atgccgtctcaggtctcaagaaccagcttcagcctctcttttctcgagagataccc
61	REV-3a-pp-Alpha-Factor-noEAEA	atgccgtctcaggtctcaagaacctcttttctcgagagataccccttc
110	REV-3a-Invertase	atgccgtctcaggtctcaagaaccGGCAGAAATCTTTGCTGCAAAACC
111	REV-3a-hSA	atgccgtctcaggtctcaagaaccAGAGTAAGCAGAGGAGAACAAGAATAACAATG
Cloning of RFP and yEGFP parts		
43	FW-3-pp-GFP-new	gcatcgtctcatcggtctcatatggtgagcaagggcgag
44	FW-3b-pp-GFP-new	gcatcgtctcatcggtctcattctgtgagcaagggcgaggag
45	FW-3b-pp-RFP-new	gcatcgtctcatcggtctcattctgcaacttccggtatggtgtcaaag
46	FW-4-tAOX1-new	gcatcgtctcatcggtctcaatcctaaTCAAGAGGATGTCAGAATGCC
55	FW-3-pp-RFP-Linker	gcatcgtctcatcggtctcatatgG
56	REV-3/3b-pp-RFP-Linker	atgccgtctcaggtctcaggatccTGCGGTTCCGG
57	REV-3/3b-pp-GFP-Linker	atgccgtctcaggtctcaggatcccttgtacagctcgtccatgc

Continued on next page

Table A.3 – continued from previous page

Number	Name	Sequence 5´-3´
Cloning of ADH parts		
64	FW-3b-Adh	gcatcgtctcatcggtctcattctCATCATCATCATCATCACAGCAGCG
65	REV-3/3b-AdhZ2	atgccgtctcaggtctcaggatccGTCTGTTAGTGTGCGATTATCGATAACAAAACG
66	REV-AdhZ3_FixBsmBI	atgccgtctcaagacACGCTGCGATCGCCCG
67	FW-AhdZ3_FixBsmBI	gcatcgtctcagtctGGTTCTGCTACCGGCACGCC
68	REV-3/3b-AdhZ3	atgccgtctcaggtctcaggatccAAAATCGGCTTTCAACACCACGC
74	FW-3b-AdhZ3_no HIS	gcatcgtctcatcggtctcattctATGTCGATGATAAAAAGCTATGCCGC
75	FW-3-AdhZ3_no HIS	gcatcgtctcatcggtctcatatgTCGATGATAAAAAGCTATGCCGC
76	FW-3-Adh	gcatcgtctcatcggtctcatatgCATCATCATCATCATCACAGCAGCG
77	REV-3/3b-AdhZ3 from part	atgccgtctcaggtctcaggatccAAAATCGGC
Cloning of other parts		
9	FW-JC071-BxbILocus	tggtttctcctgacccaaagactttaaattt
10	REV-JC073-BxbILocus	gaaccaatttagctatatatagttaactaccggctcg
88	FW-4b-tAOX1	gcatcgtctcatcggtctcatggcTCAAGAGGATGTCAGAATGCCATTTGCCTGAG
89	REV-4-AOX1terminator	atgccgtctcaggtctcacagcTCTCACTTAATCTTCTGTACTCTGAAGAGGAGTGGG
112	FW-3a-pp-AaeUPOeng	gcatcgtctcatcggtctcatatgAAATATTTTCCCCTGTTCCCAACC
113	REV-3a-pp-AaeUPOeng	atgccgtctcaggtctcaagaaccTCGGGCCTCGAGTGTTG
Primers for sequencing confirmation		
11	FW-Sequencing-EntryVector	ccttttgctggccttttgctc
12	REV-Sequencing-EntryVector	ccagtaatgacctcagaactcc
13	FW-Seq-ConS	cgacaacgtggcaattcgtcg
14	FW-Seq-Ruby	gttgtcatggagggctccg
15	FW-Seq-Turqu	ggtcacaaattttctgtctccgg
16	FW-Seq-ConE	gaaccagcgccggcgaac
17	FW-Seq-ZeoTerm	gcgaagttaagtgcgcagaaag
18	FW-Seq-ColE1	gcggagcctatggaaaaacgc
19	REV-Seq-ConS	cagatggtcctggagatcgttg
20	REV-Seq-Ruby	cggagccctccatgacaac
21	REV-Seq-Turqu	ccggagacagaaaatttgtgacc
22	REV-Seq-ConE	catcggtatgatctgtacatgattcg
23	REV-Seq-Kana	ctcaccggattcagtcgtcactc
62	REV-Sequ-Mid_RFP	caccttcaatttcgaattcgtgaccg
63	REV-Sequ-Mid_GFP	cagcttgccgtaggtggcatc
72	FW-Seq-pGAP	gcgaacacctttcccaattttggtttc
73	REV-Seq-tAOX1	gatcaggagcaagctcgtacgagaa
115	FW-Seq-pAOX1	gcgactggttccaattgacaag
116	FW-Seq-ConS (N-term)	ggtagagccacaaacagccg
117	Rev-Seq-tAOX1	gcctgcatctctcaggcaaatg
118	REV-Seq-Zeocin	caaggagggtattctgggcctc
119	REV-Seq-UPO-a	gttcaaacctggacatggacctc
120	REV-Seq-UPO-b	ccgtggttagccaaggagttcaaa
121	REV-Seq-UPO-c	ggttagccaaagcgttcaaagctggac
122	REV-Seq-Phytase_An	ctggtcagggactcgtatctctgg
123	REV-Seq-Phytase_Sth	ccaaagtgtagtcgtaagttctcaggaactc
124	REV-Seq-Phytase_Ply	cgaatggcaacaagtcagcaacac
125	REV-Seq-Phytase_Ec	gcaataatagcaacctgaccggattgtg

A.3 Detailed Information about Expression Plasmids

Table A.4: *P. pastoris* strains for intracellular expression of RFP or yEGFP

Name	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7	Type 8
pUO_pL001	ConLS	pAOX1	RFP	tAOX1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL002		pGAP		tAOX1				
pUO_pL003		pPGK1		tPGK1				
pUO_pL004		pTEF1		tPGK1				
pUO_pL005		pENO1		tAOX1				
pUO_pL006		pTPI1		tAOX1				
pUO_pL007	ConLS	pAOX1	yEGFP	tAOX1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL008		pGAP		tAOX1				
pUO_pL009		pPGK1		tPGK1				
pUO_pL010		pTEF1		tPGK1				
pUO_pL011		pENO1		tAOX1				
pUO_pL012		pTPI1		tAOX1				
pUO_pL093	ConLS	pAOX1	RFP	tAOX1	ConRE	ZeoR	PARS	KanR-ColE1
pUO_pL094		pGAP		tAOX1				
pUO_pL095		pPGK1		tPGK1				
pUO_pL096		pTEF1		tPGK1				
pUO_pL097		pENO1		tAOX1				
pUO_pL098		pTPI1		tAOX1				
pUO_pL099	ConLS	pAOX1	yEGFP	tAOX1	ConRE	ZeoR	PARS	KanR-ColE1
pUO_pL100		pGAP		tAOX1				
pUO_pL101		pPGK1		tPGK1				
pUO_pL102		pTEF1		tPGK1				
pUO_pL103		pENO1		tAOX1				
pUO_pL104		pTPI1		tAOX1				
pUO_pL621	ConLS	pTDH3	RFP	tAOX1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL622		pCCW12						
pUO_pL623		pHHF2						
pUO_pL624		pTEF2						
pUO_pL625		pHHF1						
pUO_pL626		pHTB2						
pUO_pL627		pRPL18B						
pUO_pL628	ConLS	pTDH3	yEGFP	tAOX1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL629		pCCW12						
pUO_pL630		pHHF2						
pUO_pL631		pTEF2						
pUO_pL632		pHHF1						
pUO_pL633		pHTB2						
pUO_pL634		pRPL18B						
pUO_pL635	ConLS	pGAP	RFP	tENO1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL636				tSSA1				
pUO_pL637				tADH1				
pUO_pL638				tPGK1				
pUO_pL639				tENO2				
pUO_pL640				tTDH1				

Continued on next page

Table A.4 – continued from previous page

Name	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7	Type 8
pUO_pL641	ConLS	pGAP	yEGFP	tENO1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL642				tSSA1				
pUO_pL643				tADH1				
pUO_pL644				tPGK1				
pUO_pL645				tENO2				
pUO_pL646				tTDH1				
pUO_pL680	ConLS	pPET9	RFP	tAOX1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL681		pG1						
pUO_pL682		pG6						
pUO_pL683		pADH2						
pUO_pL684		pDAS1						
pUO_pL685		pPMP20						
pUO_pL686	ConLS	pPET9	yEGFP	tAOX1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL687		pG1						
pUO_pL688		pG6						
pUO_pL689		pADH2						
pUO_pL690		pDAS1						
pUO_pL691		pPMP20						
pUO_pL730	ConLS	pAOX1	RFP	tENO1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL731				tSSA1				
pUO_pL732				tADH1				
pUO_pL733				tPGK1				
pUO_pL734				tENO2				
pUO_pL735				tTDH1				
pUO_pL736	ConLS	pDAS1	RFP	tENO1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL737				tSSA1				
pUO_pL738				tADH1				
pUO_pL739				tPGK1				
pUO_pL740				tENO2				
pUO_pL741				tTDH1				
pUO_pL742	ConLS	pPMP20	RFP	tENO1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL743				tSSA1				
pUO_pL744				tADH1				
pUO_pL745				tPGK1				
pUO_pL746				tENO2				
pUO_pL747				tTDH1				
pUO_pL748	ConLS	pPET9	RFP	tENO1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL749				tSSA1				
pUO_pL750				tADH1				
pUO_pL751				tPGK1				
pUO_pL752				tENO2				
pUO_pL753				tTDH1				
pUO_pL754	ConLS	pG1	RFP	tENO1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL755				tSSA1				
pUO_pL756				tADH1				
pUO_pL757				tPGK1				
pUO_pL758				tENO2				
pUO_pL759				tTDH1				

Continued on next page

Table A.4 – continued from previous page

Name	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7	Type 8
pUO_pL760	ConLS	pG6	RFP	tENO1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL761				tSSA1				
pUO_pL762				tADH1				
pUO_pL763				tPGK1				
pUO_pL764				tENO2				
pUO_pL765				tTDH1				
pUO_pL766	ConLS	pAHD2	RFP	tENO1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL767				tSSA1				
pUO_pL768				tADH1				
pUO_pL769				tPGK1				
pUO_pL770				tENO2				
pUO_pL771				tTDH1				

Table A.5: *P. pastoris* strains for secreted expression of RFP or yEGFP

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8
pUO_pL013	ConLS	pGAP	αAmylase-αMFΔ	RFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL014			αMFΔ						
pUO_pL015			αMFΔ_no_Kex						
pUO_pL016			Glucoamylase-αMFΔ						
pUO_pL017			SA-αMFΔ						
pUO_pL018			Inulinase-αMFΔ						
pUO_pL019			Invertase-αMFΔ						
pUO_pL020			Killer-αMFΔ						
pUO_pL021	ConLS	pGAP	αAmylase-αMFΔ	yEGFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL022			αMFΔ						
pUO_pL023			αMFΔ_no_Kex						
pUO_pL024			Glucoamylase-αMFΔ						
pUO_pL025			SA-αMFΔ						
pUO_pL026			Inulinase-αMFΔ						
pUO_pL027			Invertase-αMFΔ						
pUO_pL028			Killer-αMFΔ						
pUO_pL029	ConLS	pAOX1	αAmylase-αMFΔ	RFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL030			αMFΔ						
pUO_pL031			αMFΔ_no_Kex						
pUO_pL032			Glucoamylase-αMFΔ						
pUO_pL033			SA-αMFΔ						
pUO_pL034			Inulinase-αMFΔ						
pUO_pL035			Invertase-αMFΔ						
pUO_pL036			Killer-αMFΔ						
pUO_pL037	ConLS	pAOX1	αAmylase-αMFΔ	yEGFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL038			αMFΔ						
pUO_pL039			αMFΔ_no_Kex						
pUO_pL040			Glucoamylase-αMFΔ						
pUO_pL041			SA-αMFΔ						
pUO_pL042			Inulinase-αMFΔ						
pUO_pL043			Invertase-αMFΔ						
pUO_pL044			Killer-αMFΔ						

Continued on next page

Table A.5 – continued from previous page

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8
pUO_pL045	ConLS	pPGK1	αAmylase-αMFΔ	RFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL046			αMFΔ						
pUO_pL047			αMFΔ_no_Kex						
pUO_pL048			Glucoamylase-αMFΔ						
pUO_pL049			SA-αMFΔ						
pUO_pL050			Inulinase-αMFΔ						
pUO_pL051			Invertase-αMFΔ						
pUO_pL052			Killer-αMFΔ						
pUO_pL053	ConLS	pPGK1	αAmylase-αMFΔ	yEGFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL054			αMFΔ						
pUO_pL055			αMFΔ_no_Kex						
pUO_pL056			Glucoamylase-αMFΔ						
pUO_pL057			SA-αMFΔ						
pUO_pL058			Inulinase-αMFΔ						
pUO_pL059			Invertase-αMFΔ						
pUO_pL060			Killer-αMFΔ						
pUO_pL061	ConLS	pTEF1	αAmylase-αMFΔ	RFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL062			αMFΔ						
pUO_pL063			αMFΔ_no_Kex						
pUO_pL064			Glucoamylase-αMFΔ						
pUO_pL065			SA-αMFΔ						
pUO_pL066			Inulinase-αMFΔ						
pUO_pL067			Invertase-αMFΔ						
pUO_pL068			Killer-αMFΔ						
pUO_pL069	ConLS	pTEF1	αAmylase-αMFΔ	yEGFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL070			αMFΔ						
pUO_pL071			αMFΔ_no_Kex						
pUO_pL072			Glucoamylase-αMFΔ						
pUO_pL073			SA-αMFΔ						
pUO_pL074			Inulinase-αMFΔ						
pUO_pL075			Invertase-αMFΔ						
pUO_pL076			Killer-αMFΔ						

Continued on next page

Table A.5 – continued from previous page

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8
pUO_pL077	ConLS	pENO1	αAmylase-αMFΔ	RFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL078			αMFΔ						
pUO_pL079			αMFΔ_no_Kex						
pUO_pL080			Glucoamylase-αMFΔ						
pUO_pL081			SA-αMFΔ						
pUO_pL082			Inulinase-αMFΔ						
pUO_pL083			Invertase-αMFΔ						
pUO_pL084			Killer-αMFΔ						
pUO_pL085	ConLS	pENO1	αAmylase-αMFΔ	yEGFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL086			αMFΔ						
pUO_pL087			αMFΔ_no_Kex						
pUO_pL088			Glucoamylase-αMFΔ						
pUO_pL089			SA-αMFΔ						
pUO_pL090			Inulinase-αMFΔ						
pUO_pL091			Invertase-αMFΔ						
pUO_pL092			Killer-αMFΔ						
pUO_pL105	ConLS	pAOX1	αMF	RFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL106		pAOX1		yEGFP	tAOX1				
pUO_pL107		pGAP		RFP	tAOX1				
pUO_pL108		pGAP		yEGFP	tAOX1				
pUO_pL109		pPGK1		RFP	tPGK				
pUO_pL110		pPGK1		yEGFP	tPGK				
pUO_pL111		pTEF1		RFP	tPGK				
pUO_pL112		pTEF1		yEGFP	tPGK				
pUO_pL113		pENO1		RFP	tAOX1				
pUO_pL114		pENO1		yEGFP	tAOX1				
pUO_pL115	ConLS	pAOX1	αMF_no_EAEA	RFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL116		pAOX1		yEGFP	tAOX1				
pUO_pL117		pGAP		RFP	tAOX1				
pUO_pL118		pGAP		yEGFP	tAOX1				
pUO_pL119		pPGK1		RFP	tPGK				
pUO_pL120		pPGK1		yEGFP	tPGK				
pUO_pL121		pTEF1		RFP	tPGK				
pUO_pL122		pTEF1		yEGFP	tPGK				

Continued on next page

Table A.5 – continued from previous page

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8
pUO_pL123		pENO1		RFP	tAOX1				
pUO_pL124		pENO1		yEGFP	tAOX1				
pUO_pL692			SP_Disulfide isomerase						
pUO_pL693			SP_C4R6P1						
pUO_pL694			SP_Cell wall protein						
pUO_pL695			SP_Cyclophilin						
pUO_pL696			SP_CSN2						
pUO_pL697			SP_PHA-E						
pUO_pL698			MF41						
pUO_pL699			SP_C4R8H7						
pUO_pL700			SP_Peptidylprolyl isomerase						
pUO_pL701	ConLS	pGAP	SP_ALB	RFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL702			SP_Scw11p						
pUO_pL703			SP_Mucin						
pUO_pL704			SP_An_phyA						
pUO_pL705			SP_Pl_phyA						
pUO_pL706			SP_Th_phyA						
pUO_pL707			SP_Aae_UPO						
pUO_pL708			SP_Aae_UPOeng						
pUO_pL709			SP_PHO1						
pUO_pL710			SP_Suc2						
pUO_pL711			SP_Disulfide isomerase						
pUO_pL712			SP_C4R6P1						
pUO_pL713			SP_Cell wall protein						
pUO_pL714			SP_Cyclophilin						
pUO_pL715			SP_CSN2						
pUO_pL716			SP_PHA-E						
pUO_pL717			MF41						
pUO_pL718			SP_C4R8H7						
pUO_pL719			SP_Peptidylprolyl isomerase						
pUO_pL720	ConLS	pGAP	SP_ALB	yEGFP	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL721			SP_Scw11p						
pUO_pL722			SP_Mucin						
pUO_pL723			SP_An_phyA						
pUO_pL724			SP_Pl_phyA						

Continued on next page

Table A.5 – continued from previous page

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8
pUO_pL725			SP_Th_phyA						
pUO_pL726			SP_Aae_UPO						
pUO_pL727			SP_Aae_UPOeng						
pUO_pL728			SP_PHO1						
pUO_pL729			SP_Suc2						

Table A.6: *P. pastoris* strains for secreted expression of phytase or UPO

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8
pUO_pL663	ConLS	pAOX1	αMF	*An*_phyA	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL664				*Th*_phyA					
pUO_pL665				*Pl*_phyA					
pUO_pL666				*Ec*_appA					
pUO_pL667	ConLS	pAOX1	SP_UPOeng	*Aae*UPOeng_co	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1
pUO_pL668				*Cra*UPO_co					
pUO_pL669				*Pin*UPO_co					
pUO_pL670				*Dse*UPO_co					
pUO_pL671				*Nha*UPO_co					
pUO_pL672				*Afu*UPO_co					

Table A.7: *P. pastoris* strains from Golden Gate shuffling for phytase expression

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8	Shuffling origin
pUO_pL917	ConLS	pPMP20	SP_Cyclophilin	*An*_phyA	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 13
pUO_pL918		pAOX1	SP_Cell wall protein							
pUO_pL919		pAOX1	SP_Cell wall protein							
pUO_pL920		pAOX1	SP_C4R8H7							
pUO_pL921		pAOX1	SP_Peptidylisomerase							
pUO_pL922		pAOX1	SP_CSN2							
pUO_pL923	ConLS	pDAS1	αMF_no_EAEA	*Th*_phyA	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 14
pUO_pL924		pAOX1	SP_CSN2							
pUO_pL925		pAOX1	SP_CSN2							
pUO_pL926		pAOX1	SP_C4R8H7							
pUO_pL927		pPMP20	SP_C4R8H7							
pUO_pL928		pPMP20	SP_C4R8H7							
pUO_pL929	ConLS	pPMP20	SP_Cyclophilin	*Pl*_phyA	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 15
pUO_pL930		pAOX1	SP_Mucin							
pUO_pL931		pDAS1	SP_Mucin							
pUO_pL932		pDAS1	aMF_no_EAEA							
pUO_pL933		pPMP20	SP_Mucin							
pUO_pL934		pPMP20	SP_Mucin							
pUO_pL935	ConLS	pDAS1	SP_Disulfide isomerase	*Ec*_appA	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 16
pUO_pL936		pAOX1	αMF							
pUO_pL937		pAOX1	αMF_no_EAEA							
pUO_pL938		pDAS1	MF41							
pUO_pL939		pAOX1	αMF							
pUO_pL940		pAOX1	SP_PHO1							
pUO_pL941	ConLS	pGAP	SP_Disulfide isomerase	*An*_phyA	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 9
pUO_pL942		pGAP	SP_C4R8H7							
pUO_pL943		pGAP	SP_PHO1							
pUO_pL944		pGAP	SP_C4R8H7							
pUO_pL945		pGAP	SP_Cyclophilin							
pUO_pL946		pGAP	SP_Cyclophilin							

Continued on next page

Table A.7 – continued from previous page

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8	Shuffling origin
pUO_pL947										
pUO_pL948		pG1	SP_ALB							
pUO_pL949	ConLS	pG1	SP_Cell wall protein	*Th*_phyA	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 10
pUO_pL950		pGAP	SP_PHO1							
pUO_pL951		pGAP	SP_PHO1							
pUO_pL952		pGAP	SP_Cyclophilin							
pUO_pL953		pG6	SP_PHO1							
pUO_pL954		pG6	SP_PHO1							
pUO_pL955	ConLS	pG6	SP_Disulfide isomerase	*Pl*_phyA	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 11
pUO_pL956		pADH2	SP_Mucin							
pUO_pL957		pG6	SP_Scw11p							
pUO_pL958		pG6	SP_Disulfide isomerase							
pUO_pL959		pGAP	αMF							
pUO_pL960		pGAP	SP_PHO1							
pUO_pL961	ConLS	pGAP	αMF	*Ec*_appA	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 12
pUO_pL962		pGAP	MF41							
pUO_pL963		pGAP	MF41							
pUO_pL964		pGAP	MF41							

Table A.8: *P. pastoris* strains from Golden Gate shuffling for UPO expression

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8	Shuffling origin
pUO_pL900		pDAS1	SP_C4R8H7	*Afu*UPO_co						
pUO_pL901		pPMP20	SP_C4R6P1	*Aae*UPOeng_co						
pUO_pL902		pPMP20	aMF (336)	*Aae*UPOeng_co						
pUO_pL903		pDAS1	SP_PHO1	*Aae*UPOeng_co						
pUO_pL904		pPMP20	SP_cell wall protein	*Aae*UPOeng_co						
pUO_pL905		pDAS1	SP_Peptidyl. Iso.	*Aae*UPOeng_co						
pUO_pL906		pAOX1	SP_PHO1	*Aae*UPOeng_co						
pUO_pL907		pAOX1	SP_PHO1	*Aae*UPOeng_co						
pUO_pL908		pPMP20	SP_PHO1	*Aae*UPOeng_co						
pUO_pL909	ConLS	pPMP20	SP_Peptidyl. Iso.	*Aae*UPOeng_co	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 1
pUO_pL910		pPMP20	SP_Cyclophillin	*Aae*UPOeng_co						
pUO_pL911		pPMP20	SP_PHO1	*Aae*UPOeng_co						
pUO_pL912		pPMP20	SP_Aae_UPOeng	*Aae*UPOeng_co						
pUO_pL913		pAOX1	SP_Aae_UPOeng	*Aae*UPOeng_co						
pUO_pL914		pAOX1	SP_Aae_UPOeng	*Aae*UPOeng_co						
pUO_pL915		pPMP20	SP_Aae_UPOeng	*Aae*UPOeng_co						
pUO_pL916		pAOX1	SP_Aae_UPOeng	*Aae*UPOeng_co						
pUO_pL965		pPMP20	SP_AaeUPO	*Afu*UPO_co						
pUO_pL966		pDAS1	SP_Peptidylisomerase	*Aae*UPOeng_co						
pUO_pL967		pG6	SP_Aae_UPO	*Aae*UPOeng_co						
pUO_pL968		pG6	SP_Aae_UPO	*Aae*UPOeng_co						
pUO_pL969		pADH2	SP_Aae_UPOeng	*Aae*UPOeng_co						
pUO_pL970		pADH2	SP_Aae_UPOeng	*Aae*UPOeng_co						
pUO_pL971		pG6	SP_Aae_UPOeng	*Aae*UPOeng_co						
pUO_pL972		pADH2	SP_Aae_UPOeng	*Aae*UPOeng_co						
pUO_pL973	ConLS	pG1	SP_Aae_UPOeng	*Aae*UPOeng_co	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 3
pUO_pL974		pADH2	SP_Peptidylisomerase	*Aae*UPOeng_co						
pUO_pL975		pGAP	SP_Mucin	*Aae*UPOeng_co						
pUO_pL976		pPET9	SP_CSN2	*Aae*UPOeng_co						
pUO_pL977		pG6	αMF_no_EAEA	*Aae*UPOeng_co						
pUO_pL978		pG6	SP_C4R6P1	*Aae*UPOeng_co						
pUO_pL979		pG1	SP_PHO1	*Aae*UPOeng_co						
pUO_pL980		pG1	SP_cell wall protein	*Aae*UPOeng_co						

Continued on next page

Table A.8 – continued from previous page

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4	Type 5	Type 6	Type 7	Type 8	Shuffling origin
pUO_pL981	ConLS	pGAP	MF41	*Pin*UPO_co	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 5
pUO_pL982		pG6	SP_Scw11p	*Pin*UPO_co						
pUO_pL983	ConLS	pADH2	SP_PHO1	*Nha*UPO_co	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 7
pUO_pL984		pADH2	SP_CSN2	*Nha*UPO_co						
pUO_pL985	Con LS	pADH2	SP_Cyclophillin	*Afu*UPO_co	tAOX1	ConRE	ZeoR	BxbI	KanaR-ColE1	Shuffling Library 8

Table A.9: *P. pastoris* strains for secreted expression of ADHs

Name	Type 1	Type 2	Type 3a	Type 3b	Type 4a	Type 4b	Type 5	Type 6	Type 7	Type 8
pUO_pL600 pUO_pL601 pUO_pL602	ConLS	pGAP	αMF αMF_no_EAEA SA-αMFΔ	AdhZ2_Ec	tAOX1		ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL603 pUO_pL604 pUO_pL605	ConLS	pGAP	αMF αMF_no_EAEA SA-αMFΔ	AdhZ2_DIN	tAOX1		ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL606 pUO_pL607 pUO_pL608	ConLS	pGAP	αMF αMF_no_EAEA SA-αMFΔ	AdhZ2_7476	tAOX1		ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL609 pUO_pL610 pUO_pL611	ConLS	pGAP	αMF αMF_no_EAEA SA-αMFΔ	AdhZ3_LND	tAOX1		ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL612 pUO_pL613 pUO_pL614	ConLS	pGAP	αMF αMF_no_EAEA SA-αMFΔ	AdhZ3_242	tAOX1		ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL647 pUO_pL648 pUO_pL649 pUO_pL650	ConLS	pGAP	αMF αMF_no_EAEA αMF αMF_no_EAEA	AdhZ3_LND_ no HIS AdhZ3_242_no HIS	6XHis_3XFlag	tAOX1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL651 pUO_pL652 pUO_pL653 pUO_pL654	ConLS	pAOX1	αMF αMF_no_EAEA αMF αMF_no_EAEA	AdhZ3_LND_ no HIS AdhZ3_242_no HIS	6XHis_3XFlag	tAOX1	ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL655 pUO_pL656 pUO_pL657 pUO_pL658	ConLS	pGAP	αMF αMF_no_EAEA αMF αMF_no_EAEA	AdhZ3_LND_ no HIS AdhZ3_242_no HIS	tAOX1		ConRE	ZeoR	BxbI	KanR-ColE1
pUO_pL659 pUO_pL660 pUO_pL661 pUO_pL662	ConLS	pAOX1	αMF αMF_no_EAEA αMF αMF_no_EAEA	AdhZ3_LND_ no HIS AdhZ3_242_no HIS	tAOX1		ConRE	ZeoR	BxbI	KanR-ColE1

Table A.10: Shuffling set-up for phytases

Type	Plasmids
Shuffling Library 9, Constitutive, *A. niger* phytase	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF_no_EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_An_phyA, SP_Pl_phyA, SP_Th_phyA, SP_PHO1, SP_Suc2
3b	*An*_phyA
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 10, Constitutive, *T. heterothallica* phytase	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF_no_EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_An_phyA, SP_Pl_phyA, SP_Th_phyA, SP_PHO1, SP_Suc2
3b	*Th*_phyA
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 11, Constitutive, *P. lycii* phytase	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF_no_EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_An_phyA, SP_Pl_phyA, SP_Th_phyA, SP_PHO1, SP_Suc2
3b	*Pl*_phyA
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 12, Constitutive, *E. coli* phytase	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF_no_EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_An_phyA, SP_Pl_phyA, SP_Th_phyA, SP_PHO1, SP_Suc2
3b	*Ec*_appA
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 13, Induced, *A. niger* phytase	
2	pAOX1, pDAS1, pPMP20
3a	αMF, αMF_no_EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_An_phyA, SP_Pl_phyA, SP_Th_phyA, SP_PHO1, SP_Suc2
3b	*An*_phyA
4	tAOX1
5-6-7-8-1	Backbone

Continued on next page

Table A.10 – continued from previous page

Type	Plasmids
Shuffling Library 14, Induced, *T. heterothallica* phytase	
2	pAOX1, pDAS1, pPMP20
3a	αMF, αMF_no_EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_An_phyA, SP_Pl_phyA, SP_Th_phyA, SP_PHO1, SP_Suc2
3b	*Th*_phyA
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 15, Induced, *P. lycii* phytase	
2	pAOX1, pDAS1, pPMP20
3a	αMF, αMF_no_EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_An_phyA, SP_Pl_phyA, SP_Th_phyA, SP_PHO1, SP_Suc2
3b	*Pl*_phyA
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 16, Induced, *E. coli* phytase	
2	pAOX1, pDAS1, pPMP20
3a	αMF, αMF_no_EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_An_phyA, SP_Pl_phyA, SP_Th_phyA, SP_PHO1, SP_Suc2
3b	*Ec*_appA
4	tAOX1
5-6-7-8-1	Backbone

Table A.11: Shuffling set-up for UPOs

Type	Plasmids
Shuffling Library 1, Induced UPO´s	
2	pAOX1, pDAS1, pPMP20
3a	αMF, αMF no EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_Aae_UPO, SP_Aae_UPOeng, SP_PHO1, SP_Suc2
3b	*Aae*UPOeng_co, *Cra*UPO_co, *Pin*UPO_co, *Dse*UPO_co, *Nha*UPO_co, *Afu*UPO_co
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 3, *A. aegerita* UPO, engineered	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF no EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_Aae_UPO, SP_Aae_UPOeng, SP_PHO1, SP_Suc2
3b	*Aae*UPOeng_co
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 4, *C. radians* UPO	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF no EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_Aae_UPO, SP_Aae_UPOeng, SP_PHO1, SP_Suc2
3b	*Cra*UPO_co
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 5, *P. infectans* UPO	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF no EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_Aae_UPO, SP_Aae_UPOeng, SP_PHO1, SP_Suc2
3b	*Pin*UPO_co
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 6, *D. sepospорum* UPO	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF no EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_Aae_UPO, SP_Aae_UPOeng, SP_PHO1, SP_Suc2
3b	*Dse*UPO_co
4	tAOX1
5-6-7-8-1	Backbone

Continued on next page

Table A.11 – continued from previous page

Type	Plasmids
Shuffling Library 7, *N. haematococca* UPO	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF no EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_Aae_UPO, SP_Aae_UPOeng, SP_PHO1, SP_Suc2
3b	*Nha*UPO_co
4	tAOX1
5-6-7-8-1	Backbone
Shuffling Library 8, *A. fumigatus* UPO	
2	pGAP, pPET9, pG1, pG6, pAHD2
3a	αMF, αMF no EAEA, αAmylase-αMFΔ, SP_Disulfide isomerase, SP_C4R6P1, SP_Cell wall protein, SP_Cyclophillin, SP_CSN2, SP_PHA-E, MF41, SP_C4R8H7, SP_Peptidylprolyl isomerase, SP_ALB, SP_Scw11p, SP_Mucin, SP_Aae_UPO, SP_Aae_UPOeng, SP_PHO1, SP_Suc2
3b	*Afu*UPO_co
4	tAOX1
5-6-7-8-1	Backbone

TABLE A.12. Plasmids used or made not in PTK formate

Plasmid	**Description**	**Reference**
pJC043	BxbI plasmid (recombinase for transformation)	Lu Lab, MIT, Cambridge, USA, [234]
pJC117	pAOX1-yEGFP	Lu Lab, MIT, Cambridge, USA
pJC011	pGAP-αMF	Lu Lab, MIT, Cambridge, USA
PP264	pAOX1, αMFΔ signal peptide	Lu Lab, MIT, Cambridge, USA
PP265	pAOX1, glucoamylase signal peptide	Lu Lab, MIT, Cambridge, USA
PP266	pAOX1, inulinase signal peptide	Lu Lab, MIT, Cambridge, USA
PP267	pAOX1, invertase signal peptide	Lu Lab, MIT, Cambridge, USA
PP270	pAOX1, SA signal peptide	Lu Lab, MIT, Cambridge, USA
PP280	pAOX1, aAmylase signal peptide	Lu Lab, MIT, Cambridge, USA
PP281	pAOX1, killer signal peptide	Lu Lab, MIT, Cambridge, USA
PP168	pAOX1, αMFΔ signal peptide	Lu Lab, MIT, Cambridge, USA
pPICB_UPOeng	pPICC vector, pAOX1, αMF, *A. aegerita* UPO engineered	Alcalde Lab, Madrid, Spain, [208]
CBR P_511	pPICZαA+YahK E.c., alkohol dehydrogenase	Made in this study, [239, 240]
CBR P_512	pPICZαA+YahK DIN, alkohol dehydrogenase with cofaktor mutation	Made in this study, [239, 240]
CBR P_513	pPICZαA+YahK 74/76, alkohol dehydrogenase with glycosylation mutation	Made in this study, [223]
CBR P_514	pPICZαA+YjgB LND, alkohol dehydrogenase with cofaktor mutation	Made in this study, [239, 240]
CBR P_515	pPICZαA+YjgB 242, alkohol dehydrogenase with glycosylation mutation	Made in this study, [223]
CBR P_516	pPICZαB+galOX galactose oxidase	Made in this study
CBR P_704	Chromogenic protein, vector for *E. coli* expression, pT5, VixenPurple	DNA2.0, CPB-36-441
CBR P_705	Chromogenic protein, vector for *P. pastoris* expression, pAOX1, CupidPink	DNA2.0, CPB-41-902

BIBLIOGRAPHY

[1] J. J. AGRESTI, E. ANTIPOV, A. R. ABATE, K. AHN, A. C. ROWAT, J.-C. BARET, M. MARQUEZ, A. M. KLIBANOV, A. D. GRIFFITHS, AND D. A. WEITZ, *Ultrahigh-throughput screening in drop-based microfluidics for directed evolution*, Proceedings of the National Academy of Sciences, 107 (2010), pp. 4004–4009.

[2] B. AGUILAR-USCANGA AND J. M. FRANÇOIS, *A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation*, Letters in Applied Microbiology, 37 (2003), pp. 268–274.

[3] A. AHARONI, G. AMITAI, K. BERNATH, S. MAGDASSI, AND D. S. TAWFIK, *High-Throughput Screening of Enzyme Libraries: Thiolactonases Evolved by Fluorescence-Activated Sorting of Single Cells in Emulsion Compartments*, Chemistry and Biology, 12 (2005), pp. 1281–1289.

[4] M. AHMAD, M. HIRZ, H. PICHLER, AND H. SCHWAB, *Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production*, Applied Microbiology and Biotechnology, 98 (2014), pp. 5301–5317.

[5] M. K. AHMADI AND B. A. PFEIFER, *Recent progress in therapeutic natural product biosynthesis using Escherichia coli*, Current Opinion in Biotechnology, 42 (2016), pp. 7–12.

[6] A. AKBARZADEH, E. DEHNAVI, M. AGHAEEPOOR, AND J. AMANI, *Optimization of recombinant expression of synthetic bacterial phytase in Pichia pastoris using response surface methodology*, Jundishapur Journal of Microbiology, 8 (2015), p. e27553.

[7] G. B. AKCAPINAR, O. GUL, AND U. SEZERMAN, *Effect of Codon Optimization on the Expression of Trichoderma reesei Endoglucanase 1 in Pichia pastoris*, Biotechnology Progress, 27 (2011), pp. 1257–1263.

[8] E. ALTAY, T. SHAHWAN, AND M. TANOĞLU, *Morphosynthesis of CaCO3 at different reaction temperatures and the effects of PDDA, CTAB, and EDTA on the particle morphology and polymorph stability*, Powder Technology, 178 (2007), pp. 194–202.

[9] J. C. ANDERSON, J. E. DUEBER, M. LEGUIA, G. C. WU, J. A. GOLER, A. P. ARKIN, AND J. D. KEASLING, *BglBricks: A flexible standard for biological part assembly*, Journal of Biological Engineering, 4 (2010), pp. 1–12.

[10] A. I. ANDREOU AND N. NAKAYAMA, *Mobius Assembly: A versatile Golden-Gate framework towards universal DNA assembly*, PLoS ONE, 13 (2018), pp. 1–18.

[11] E. ANDRIANANTOANDRO, S. BASU, D. K. KARIG, AND R. WEISS, *Synthetic biology: new engineering rules for an emerging discipline*, Molecular Systems Biology, 2 (2006), pp. 1–14.

[12] D. H. ANH, R. ULLRICH, D. BENNDORF, A. SVATOŚ, A. MUCK, AND M. HOFRICHTER, *The Coprophilous Mushroom Coprinus radians Secretes a Haloperoxidase That Catalyzes Aromatic Peroxygenation*, Applied and Environmental Microbiology, 73 (2007), pp. 5477–5485.

[13] A. A. ANTIPOV AND G. B. SUKHORUKOV, *Polyelectrolyte multilayer capsules as vehicles with tunable permeability*, Advances in Colloid and Interface Science, 111 (2004), pp. 49–61.

[14] A. A. ANTIPOV, G. B. SUKHORUKOV, S. LEPORATTI, I. L. RADTCHENKO, E. DONATH, AND H. MÖHWALD, *Polyelectrolyte multilayer capsule permeability control*, Colloids and Surfaces A: Physicochemical and Engineering Aspects, 198-200 (2002), pp. 535–541.

[15] A. ARKIN, *Setting the standard in synthetic biology*, Nature Biotechnology, 26 (2008), pp. 771–774.

[16] A. P. ARKIN AND D. A. FLETCHER, *Fast, cheap and somewhat in control*, Genome Biology, 7 (2006), pp. 114.1–114.6.

[17] J. A. J. ARPINO, E. J. HANCOCK, J. ANDERSON, M. BARAHONA, G. B. V. STAN, A. PAPACHRISTODOULOU, AND K. POLIZZI, *Tuning the dials of Synthetic Biology*, Microbiology (United Kingdom), 159 (2013), pp. 1236–1253.

[18] Ö. ATA, R. PRIELHOFER, B. GASSER, D. MATTANOVICH, AND P. ÇALIK, *Transcriptional Engineering of the Glyceraldehyde-3-Phosphate Dehydrogenase Promoter for Improved Heterologous Protein Production in Pichia pastoris*, Biotechnology and Bioengineering, 114 (2017), pp. 2319–2327.

[19] S. ATSUMI, T. HANAI, AND J. C. LIAO, *Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels*, Nature, 451 (2008), pp. 86–90.

[20] S. ATSUMI, T.-Y. WU, E.-M. ECKL, S. D. HAWKINS, T. BUELTER, AND J. C. LIAO, *Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison*

of three aldehyde reductase/alcohol dehydrogenase genes, Applied Microbiology and Biotechnology, 85 (2010), pp. 651–657.

[21] F. M. AUSUBEL, R. BRENT, R. E. KINGSTON, D. D. MOORE, J. G. SEIDMAN, J. A. SMITH, K. STRUHL, C. J. WILEY, R. D. ALLISON, M. BITTNER, AND S. BLACKSHAW, *Current Protocols in Molecular Biology*, vol. 1, 2003.

[22] A. R. AWAN, W. M. SHAW, AND T. ELLIS, *Biosynthesis of therapeutic natural products using synthetic biology*, Advanced Drug Delivery Reviews, 105 (2016), pp. 96–106.

[23] E. D. BABOT, J. C. DEL RÍO, L. KALUM, A. T. MARTÍNEZ, AND A. GUTIÉRREZ, *Oxyfunctionalization of Aliphatic Compounds by a Recombinant Peroxygenase From Coprinopsis cinerea*, Biotechnology and Bioengineering, 110 (2013), pp. 2323–2332.

[24] H. D. BAE, L. J. YANKE, K. J. CHENG, AND L. B. SELINGER, *A novel staining method for detecting phytase activity*, Journal of Microbiological Methods, 39 (1999), pp. 17–22.

[25] J. BAI, D. J. SWARTZ, I. I. PROTASEVICH, C. G. BROUILLETTE, P. M. HARRELL, E. HILDEBRANDT, B. GASSER, D. MATTANOVICH, A. WARD, G. CHANG, AND I. L. URBATSCH, *A Gene Optimization Strategy that Enhances Production of Fully Functional P-Glycoprotein in Pichia pastoris*, PLoS ONE, 6 (2011), p. e22577.

[26] J. BECKER AND C. WITTMANN, *Advanced Biotechnology: Metabolically Engineered Cells for the Bio-Based Production of Chemicals and Fuels, Materials, and Health-Care Products*, Angewandte Chemie - International Edition, 54 (2015), pp. 3328–3350.

[27] J. BELL, L. PAULA, T. DODD, S. NÉMETH, C. NANOU, V. MEGA, AND P. CAMPOS, *EU ambition to build the world's leading bioeconomy—Uncertain times demand innovative and sustainable solutions*, New Biotechnology, 40 (2018), pp. 25–30.

[28] A. BERLEC AND B. ŠTRUKELJ, *Current state and recent advances in biopharmaceutical production in Escherichia coli, yeasts and mammalian cells*, Journal of Industrial Microbiology and Biotechnology, 40 (2013), pp. 257–274.

[29] J.-G. BERRIN, G. WILLIAMSON, A. PUIGSERVER, J.-C. CHAIX, W. R. MCLAUCHLAN, AND N. JUGE, *High-Level Production of Recombinant Fungal Endo-β-1,4-xylanase in the Methylotrophic Yeast Pichia pastoris*, Protein Expression and Purification, 19 (2000), pp. 179–187.

[30] R. M. BILL, *Playing catch-up with Escherichia coli: using yeast to increase success rates in recombinant protein production experiments*, Frontiers in Microbiology, 5 (2014), pp. 1–5.

[31] A. BINDER, J. LAMBERT, R. MORBITZER, C. POPP, T. OTT, T. LAHAYE, AND M. PARNISKE, *A Modular Plasmid Assembly Kit for Multigene Expression, Gene Silencing and Silencing Rescue in Plants*, PLoS ONE, 9 (2014), p. e88218.

[32] J. BITINAITE, M. RUBINO, K. H. VARMA, I. SCHILDKRAUT, R. VAISVILA, AND R. VAISKUNAITE, *USER TM friendly DNA engineering and cloning method by uracil excision*, Nucleic Acids Research, 35 (2007), pp. 1992–2002.

[33] B. A. BLOUNT, T. WEENINK, AND T. ELLIS, *Construction of synthetic regulatory networks in yeast*, FEBS Letters, 586 (2012), pp. 2112–2121.

[34] C. R. BLUE, A. GIUFFRE, S. MERGELSBERG, N. HAN, J. J. DE YOREO, AND P. M. DOVE, *Chemical and physical controls on the transformation of amorphous calcium carbonate into crystalline CaCO3 polymorphs*, Geochimica et Cosmochimica Acta, 196 (2017), pp. 179–196.

[35] C. BONG HYUN, N. SOO WAN, K. BYUNG MOON, AND P. YOUNG HOON, *Highly Efficient Secretion of Heterologous Proteins From Saccharomyces cerevisiae Using Inulinase Signal Peptides*, Biotechnology and Bioengineering, 49 (1996), pp. 473–479.

[36] U. T. BORNSCHEUER, *The fourth wave of biocatalysis is approaching*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376 (2018), pp. 1–7.

[37] M. M. BRADFORD, *A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding*, Analytical Biochemistry, 72 (1976), pp. 248–254.

[38] R. K. BRETTHAUER AND F. J. CASTELLINO, *Glycosylation of Pichia pastoris-derived proteins*, Biotechnology and Applied Biochemistry, 30 (1999), pp. 193–200.

[39] M. L. BROEREN, M. C. ZIJP, S. L. WAAIJERS-VAN DER LOOP, E. H. HEUGENS, L. POSTHUMA, E. WORRELL, AND L. SHEN, *Environmental assessment of bio-based chemicals in early-stage development: a review of methods and indicators*, Biofuels, Bioproducts and Biorefining, 11 (2017), pp. 701–718.

[40] S. R. BROWN, M. STAFF, R. LEE, J. LOVE, D. PARKER, S. AVES, AND T. HOWARD, *Design of Experiments Methodology to Build a Multifactorial Statistical Model Describing the Metabolic Interactions of Alcohol Dehydrogenase Isozymes in the Ethanol Biosynthetic Pathway of the Yeast Saccharomyces cerevisiae*, ACS Synthetic Biology, 7 (2018), pp. 1676–1684.

[41] K. BUCHHOLZ AND J. COLLINS, *The roots–a short history of industrial microbiology and biotechnology*, Applied Microbiology and Biotechnology, 97 (2013), pp. 3747–3762.

[42] S. CABANTOUS, T. C. TERWILLIGER, AND G. S. WALDO, *Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein*, Nature Biotechnology, 23 (2005), pp. 102–107.

[43] S. CABANTOUS AND G. S. WALDO, *In vivo and in vitro protein solubility assays using split GFP*, Nature Methods, 3 (2006), pp. 845–854.

[44] A. CAMATTARI, A. GOH, L. Y. YIP, A. H. M. TAN, S. W. NG, A. TRAN, G. LIU, I. LIACHKO, M. J. DUNHAM, AND G. RANCATI, *Characterization of a panARS-based episomal vector in the methylotrophic yeast Pichia pastoris for recombinant protein production and synthetic biology applications*, Microbial Cell Factories, 15 (2016), pp. 1–11.

[45] B. CANTON, A. LABNO, AND D. ENDY, *Refinement and standardization of synthetic biological parts and devices*, Nature Biotechnology, 26 (2008), pp. 787–793.

[46] S. CARDINALE AND A. P. ARKIN, *Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems*, Biotechnology Journal, 7 (2012), pp. 856–866.

[47] F. CARUSO AND H. MÖHWALD, *Protein Multilayer Formation on Colloids through a Stepwise Self-Assembly Technique*, Science, (1999), pp. 6039–6046.

[48] A. CASINI, M. STORCH, G. S. BALDWIN, AND T. ELLIS, *Bricks and blueprints: methods and standards for DNA assembly*, Nature Reviews Molecular Cell Biology, 16 (2015), pp. 568–576.

[49] J. L. CEREGHINO AND J. M. CREGG, *Heterologous protein expression in the methylotrophic yeast Pichia pastoris*, FEMS Microbiology Reviews, 24 (2000), pp. 45–66.

[50] F. CERONI, A. BOO, S. FURINI, T. E. GOROCHOWSKI, O. BORKOWSKI, Y. N. LADAK, A. R. AWAN, C. GILBERT, G. B. STAN, AND T. ELLIS, *Burden-driven feedback control of gene expression*, Nature Methods, 15 (2018), pp. 387–393.

[51] F. CERONI AND T. ELLIS, *The challenges facing synthetic biology in eukaryotes*, Nature Reviews Molecular Cell Biology, 19 (2018), pp. 481–482.

[52] S. CHAHAL, P. WEI, P. MOUA, S. P. J. PARK, J. KWON, A. PATEL, A. T. VU, J. A. CATOLICO, Y. F. T. TSAI, N. SHAHEEN, T. T. CHU, V. TAM, Z.-E.-H. KHAN, H. H. JOO, L. XUE, J. LIN-CEREGHINO, J. W. TSAI, AND G. P. LIN-CEREGHINO, *Structural characterization of the α-mating factor prepro-peptide for secretion of recombinant proteins in Pichia pastoris*, Gene, 598 (2017), pp. 50–62.

[53] R. CHAO, Y. YUAN, AND H. ZHAO, *Recent advances in DNA assembly technologies*, FEMS Yeast Research, 15 (2015), pp. 1–9.

[54] J. CHAPMAN, A. ISMAIL, AND C. DINU, *Industrial Applications of Enzymes: Recent Advances, Techniques, and Outlooks*, Catalysts, 8 (2018), pp. 1–26.

[55] C.-C. CHEN, K.-J. CHENG, T.-P. KO, AND R.-T. GUO, *Current Progresses in Phytase Research: Three-Dimensional Structure and Protein Engineering*, ChemBioEng Reviews, 2 (2015), pp. 76–86.

[56] J. CHEN AND L. XIANG, *Controllable synthesis of calcium carbonate polymorphs at different temperatures*, Powder Technology, 189 (2009), pp. 64–69.

[57] Y. CHEN AND J. NIELSEN, *Advances in metabolic pathway and strain engineering paving the way for sustainable production of chemical building blocks*, Current Opinion in Biotechnology, 24 (2013), pp. 965–972.

[58] A. CHENTHAMARAKSHAN, N. PARAMBAYIL, N. MIZIRIYA, P. S. SOUMYA, M. S. LAKSHMI, A. RAMGOPAL, A. DILEEP, AND P. NAMBISAN, *Optimization of laccase production from Marasmiellus palmivorus LA1 by Taguchi method of Design of experiments*, BMC Biotechnology, 17 (2017), pp. 1–10.

[59] J. R. CHERRY AND A. L. FIDANTSEF, *Directed evolution of industrial enzymes: an update*, Current Opinion in Biotechnology, 14 (2003), pp. 438–443.

[60] C. T. CHUNG AND R. H. MILLER, *Preparation and Storage of Competent Escherichia coli Cells*, Methods in Enzymology, 218 (1993), pp. 621–627.

[61] S. N. COHEN, A. C. Y. CHANG, H. W. BOYER, AND R. B. HELLING, *Construction of Biologically Functional Bacterial Plasmids In Vitro*, Proceedings of the National Academy of Sciences, 70 (1973), pp. 3240–3244.

[62] P.-Y. COLIN, A. ZINCHENKO, AND F. HOLLFELDER, *Enzyme engineering in biomimetic compartments*, Current Opinion in Structural Biology, 33 (2015), pp. 42–51.

[63] J. H. COLLINS AND E. M. YOUNG, *Genetic engineering of host organisms for pharmaceutical synthesis*, Current Opinion in Biotechnology, 53 (2018), pp. 191–200.

[64] R. S. COX, C. MADSEN, J. MCLAUGHLIN, T. NGUYEN, J. BEAL, N. ROEHNER, B. BARTLEY, J. BEAL, M. BISSELL, K. CHOI, K. CLANCY, R. GRÜNBERG, C. MACKLIN, G. MISIRLI, E. OBERORTNER, M. POCOCK, M. SAMINENI, M. ZHANG, Z. ZHANG, Z. ZUNDEL, J. H. GENNARI, C. MYERS, H. SAURO, AND A. WIPAT, *Synthetic Biology Open Language (SBOL) Version 2.2.0*, Tech. Rep. 1, 2018.

[65] J. CREGG, K. BARRINGER, A. Y. HESSLER, AND K. R. MADDEN, *Pichia pastoris as a Host System for Transformations*, Molecular and Cellular Biology, 5 (1985), pp. 3376–3385.

[66] J. M. CREGG, *Pichia Protocols*, Totowa, 2 ed., 2007.

[67] J. M. CREGG, J. L. CEREGHINO, J. SHI, AND D. R. HIGGINS, *Recombinant Protein Expression in Pichia pastoris*, Molecular Biotechnology, 16 (2000), pp. 23–52.

[68] K. A. CURRAN, A. S. KARIM, A. GUPTA, AND H. S. ALPER, *Use of High Capacity Terminators in Saccharomyces cerevisiae to Increase mRNA half-life and Improve Gene Expression Control for Metabolic Engineering Applications*, Metabolic Engineering, 19 (2013), pp. 88–97.

[69] L. M. DAMASCENO, C.-J. HUANG, AND C. A. BATT, *Protein secretion in Pichia pastoris and advances in protein production.*, Applied Microbiology and Biotechnology, 93 (2012), pp. 31–39.

[70] T. DAVIDS, M. SCHMIDT, D. BÖTTCHER, AND U. T. BORNSCHEUER, *Strategies for the discovery and engineering of enzymes for biocatalysis*, Current Opinion in Chemical Biology, 17 (2013), pp. 215–220.

[71] B. G. DE GEEST, C. DÉJUGNAT, E. VERHOEVEN, G. B. SUKHORUKOV, A. M. JONAS, J. PLAIN, J. DEMEESTER, AND S. C. DE SMEDT, *Layer-by-layer coating of degradable microgels for pulsed drug delivery*, Journal of Controlled Release, 116 (2006), pp. 159–169.

[72] B. G. DE GEEST, R. E. VANDENBROUCKE, A. M. GUENTHER, G. B. SUKHORUKOV, W. E. HENNINK, N. N. SANDERS, J. DEMEESTER, AND S. C. DE SMEDT, *Intracellularly degradable polyelectrolyte microcapsules*, Advanced Materials, 18 (2006), pp. 1005–1009.

[73] K. DE POURCQ, K. DE SCHUTTER, AND N. CALLEWAERT, *Engineering of glycosylation in yeast and other fungi: current state and perspectives*, Applied Microbiology and Biotechnology, 87 (2010), pp. 1617–1631.

[74] K. DE SCHUTTER, Y.-C. LIN, P. TIELS, A. VAN HECKE, S. GLINKA, J. WEBER-LEHMANN, P. ROUZÉ, Y. VAN DE PEER, AND N. CALLEWAERT, *Genome sequence of the recombinant protein production host Pichia pastoris*, Nature Biotechnology, 27 (2009), pp. 561–566.

[75] T. DECOENE, B. DE PAEPE, J. MAERTENS, P. COUSSEMENT, G. PETERS, S. L. DE MAESENEIRE, AND M. DE MEY, *Standardization in synthetic biology: an engineering discipline coming of age*, Critical Reviews in Biotechnology, 0 (2017), pp. 1–10.

[76] C. DÉJUGNAT AND G. B. SUKHORUKOV, *pH-Responsive Properties of Hollow Polyelectrolyte Microcapsules Templated on Various Cores*, Langmuir, 20 (2004), pp. 7265–7269.

[77] L. L. DEL MERCATO, A. Z. ABBASI, M. OCHS, AND W. J. PARAK, *Multiplexed sensing of ions with barcoded polyelectrolyte capsules*, ACS Nano, 5 (2011), pp. 9668–9674.

[78] L. L. DEL MERCATO, A. Z. ABBASI, AND W. J. PARAK, *Synthesis and Characterization of Ratiometric Ion-Sensitive Polyelectrolyte Capsules*, Small, 7 (2011), pp. 351–363.

[79] M. DELCEA, H. MÖHWALD, AND A. G. SKIRTACH, *Stimuli-responsive LbL capsules and nanoshells for drug delivery*, Advanced Drug Delivery Reviews, 63 (2011), pp. 730–747.

[80] M. DELIC, R. GÖNGRICH, D. MATTANOVICH, AND B. GASSER, *Engineering of Protein Folding and Secretion-Strategies to Overcome Bottlenecks for Efficient Production of Recombinant Proteins.*, Antioxidants & Redox Signaling, 21 (2014), pp. 414–437.

[81] M. DELIC, M. VALLI, A. B. GRAF, M. PFEFFER, D. MATTANOVICH, AND B. GASSER, *The secretory pathway: exploring yeast diversity*, FEMS Microbiology Reviews, 37 (2013), pp. 872–914.

[82] Y. DERSJANT-LI, A. AWATI, H. SCHULZE, AND G. PARTRIDGE, *Phytase in non-ruminant animal nutrition: A critical review on phytase activities in the gastrointestinal tract and influencing factors*, Journal of the Science of Food and Agriculture, 95 (2015), pp. 878–896.

[83] A. DIASPRO, D. SILVANO, S. KROL, O. CAVALLERI, AND A. GLIOZZI, *Single Living Cell Encapsulation in Nano-organized Polyelectrolyte Shells*, Langmuir, 18 (2002), pp. 5047–5050.

[84] J. E. DICARLO, A. J. CONLEY, M. PENTTILÄ, J. JÄNTTI, H. H. WANG, AND G. M. CHURCH, *Yeast Oligo-Mediated Genome Engineering (YOGE)*, ACS Synthetic Biology, 2 (2013), pp. 741–749.

[85] E. DONATH, G. B. SUKHORUKOV, F. CARUSO, S. A. DAVIS, AND H. MÖHWALD, *Novel Hollow Polymer Shells by Colloid-Templated Assembly of Polyelectrolytes*, Angewandte Chemie - International Edition, 37 (1998), pp. 2201–2205.

[86] W. J. DOWER, J. F. MILLER, AND C. W. RAGSDALE, *High efficiency transformation of E. coli by high voltage electroporation*, Nucleic Acids Research, 16 (1988), pp. 6127–6145.

[87] M. DREIFKE, F. J. BRIELER, AND M. FRÖBA, *Immobilization of Alcohol Dehydrogenase from E. coli onto Mesoporous Silica for Application as a Cofactor Recycling System*, ChemCatChem, 9 (2017), pp. 1197–1210.

[88] J. M. DUARTE, I. BARBIER, AND Y. SCHAERLI, *Bacterial Microcolonies in Gel Beads for High-Throughput Screening of Libraries in Synthetic Biology*, ACS Synthetic Biology, 6 (2017), pp. 1988–1995.

[89] X. DUPORTET, L. WROBLEWSKA, P. GUYE, Y. LI, J. EYQUEM, J. RIEDERS, T. RIMCHALA, G. BATT, AND R. WEISS, *A platform for rapid prototyping of synthetic gene networks in mammalian cells.*, Nucleic Acids Research, 42 (2014), pp. 13440–13451.

[90] T. ELLIS, T. ADIE, AND G. S. BALDWIN, *DNA assembly for synthetic biology: from parts to pathways and beyond*, Integrative Biology, 3 (2011), pp. 109–118.

[91] O. EMANUELSSON, S. BRUNAK, G. VON HEIJNE, AND H. NIELSEN, *Locating proteins in the cell using TargetP, SignalP and related tools*, Nature Protocols, 2 (2007), pp. 953–971.

[92] M. EMPTAGE, S. L. HAYNIE, L. A. LAFFEND, J. P. PUCCI, AND G. M. WHITED, *Process for the biological production of 1,3-propanediol with high titer*, 2003.

[93] C. ENGLER, R. GRUETZNER, R. KANDZIA, AND S. MARILLONNET, *Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes*, PLoS ONE, 4 (2009), p. e5553.

[94] C. ENGLER, R. KANDZIA, AND S. MARILLONNET, *A One Pot, One Step, Precision Cloning Method with High Throughput Capability*, PLoS ONE, 3 (2008), p. e3647.

[95] C. ENGLER, M. YOULES, R. GRUETZNER, T. M. EHNERT, S. WERNER, J. D. G. JONES, N. J. PATRON, AND S. MARILLONNET, *A Golden Gate Modular Cloning Toolbox for Plants*, ACS Synthetic Biology, 3 (2014), pp. 839–843.

[96] P. FICKERS, *Pichia pastoris: a workhorse for recombinant protein production*, Current Research in Microbiology and Biotechnology, 2 (2014), pp. 354–363.

[97] M. FISCHLECHNER, Y. SCHAERLI, M. F. MOHAMED, S. PATIL, C. ABELL, AND F. HOLLFELDER, *Evolution of enzyme catalysts caged in biomimetic gel-shell beads*, Nature Chemistry, 6 (2014), pp. 791–796.

[98] J. FLEMKE, M. MAYWALD, AND V. SIEBER, *Encapsulation of Living E. coli Cells in Hollow Polymer Microspheres of Highly Defined Size*, Biomacromolecules, 14 (2013), pp. 207–214.

[99] E. FLETCHER, A. KRIVORUCHKO, AND J. NIELSEN, *Industrial Systems Biology and Its Impact on Synthetic Biology of Yeast Cell Factories*, Biotechnology and Bioengineering, 103 (2016), pp. 1164–1170.

[100] J. FU, X. BIAN, S. HU, H. WANG, F. HUANG, P. M. SEIBERT, A. PLAZA, L. XIA, R. MÜLLER, A. F. STEWART, AND Y. ZHANG, *Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting*, Nature Biotechnology, 30 (2012), pp. 440–446.

[101] P. C. GACH, K. IWAI, P. W. KIM, N. J. HILLSON, AND A. K. SINGH, *Droplet microfluidics for synthetic biology*, Lab on a Chip, 17 (2017), pp. 3388–3400.

[102] C. GAO, S. MOYA, E. DONATH, AND H. MÖHWALD, *Melamine Formaldehyde Core Decomposition as the Key Step Controlling Capsule Integrity: Optimizing the Polyelectrolyte Capsule Fabrication*, Macromolecular Chemistry and Physics, 203 (2002), pp. 953–960.

[103] B. GASSER AND D. MATTANOVICH, *A yeast for all seasons – is Pichia pastoris a suitable chassis organism for future bioproduction?*, FEMS Microbiology Letters, 365 (2018), pp. 1–4.

[104] R. GEORGIEVA, S. MOYA, E. DONATH, AND H. BÄUMLER, *Permeability and Conductivity of Red Blood Cell Templated Polyelectrolyte Capsules Coated with Supplementary Layers*, Langmuir, 20 (2004), pp. 1895–1900.

[105] F. GEU-FLORES, H. H. NOUR-ELDIN, M. T. NIELSEN, AND B. A. HALKIER, *USER fusion: A rapid and efficient method for simultaneous fusion and cloning of multiple PCR products*, Nucleic Acids Research, 35 (2007), p. e55.

[106] A. GHOSALKAR, V. SAHAI, AND A. SRIVASTAVA, *Secretory expression of interferon-alpha 2b in recombinant Pichia pastoris using three different secretion signals.*, Protein Expression and Purification, 60 (2008), pp. 103–109.

[107] D. G. GIBSON, G. A. BENDERS, C. ANDREWS-PFANNKOCH, E. A. DENISOVA, H. BADEN-TILLSON, J. ZAVERI, T. B. STOCKWELL, A. BROWNLEY, D. W. THOMAS, M. A. ALGIRE, C. MERRYMAN, L. YOUNG, V. N. NOSKOV, J. I. GLASS, J. C. VENTER, C. A. HUTCHISON, AND H. O. SMITH, *Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome.*, Science, 319 (2008), pp. 1215–1220.

[108] D. G. GIBSON, L. YOUNG, R.-Y. CHUANG, J. C. VENTER, C. A. HUTCHISON, H. O. SMITH, C. A. H. III, AND N. AMERICA, *Enzymatic assembly of DNA molecules up to several hundred kilobases.*, Nature Methods, 6 (2009), pp. 343–345.

[109] A. K. GOMBERT, J. V. MADEIRA, M.-E. CERDÁN, AND M.-I. GONZÁLEZ-SISO, *Kluyveromyces marxianus as a host for heterologous protein synthesis*, Applied Microbiology and Biotechnology, 100 (2016), pp. 6193–6208.

[110] I. GONTIA-MISHRA AND S. TIWARI, *Molecular Characterization and Comparative Phylogenetic Analysis of Phytases from Fungi with Their Prospective Applications*, Food Technology and Biotechnology, 51 (2013), pp. 313–326.

[111] T. E. GOROCHOWSKI AND T. ELLIS, *Designing efficient translation*, Nature biotechnology, 36 (2018), pp. 934–935.

[112] N. GOVINDAPPA, M. HANUMANTHAPPA, K. VENKATARANGAIAH, S. PERIYASAMY, S. SREENIVAS, R. SONI, AND K. SASTRY, *A new signal sequence for recombinant protein secretion in Pichia pastoris*, Journal of Microbiology and Biotechnology, 24 (2014), pp. 337–345.

[113] J. GRAW, *Genetik*, no. 5, Springer-Verlag, Heidelberg, 2010.

[114] G. GRÖBE, R. ULLRICH, M. J. PECYNA, D. KAPTURSKA, S. FRIEDRICH, M. HOFRICHTER, AND K. SCHEIBNER, *High-yield production of aromatic peroxygenase by the agaric fungus Marasmius rotula*, AMB Express, 1 (2011), pp. 1–11.

[115] L. S. GRONENBERG, R. J. MARCHESCHI, AND J. C. LIAO, *Next generation biofuel engineering in prokaryotes*, Current Opinion in Chemical Biology, 17 (2013), pp. 462–471.

[116] J. K. GUTERL, D. GARBE, J. CARSTEN, F. STEFFLER, B. SOMMER, S. REISSE, A. PHILIPP, M. HAACK, B. RÜHMANN, A. KOLTERMANN, U. KETTLING, T. BRÜCK, AND V. SIEBER, *Cell-Free Metabolic Engineering: Production of Chemicals by Minimized Reaction Cascades*, ChemSusChem, 5 (2012), pp. 2165–2172.

[117] J. E. GUY, M. N. ISUPOV, AND J. A. LITTLECHILD, *The Structure of an Alcohol Dehydrogenase from the Hyperthermophilic Archaeon Aeropyrum pernix*, Journal of Molecular Biology, 331 (2003), pp. 1041–1051.

[118] S. HAEFNER, A. KNIETSCH, E. SCHOLTEN, J. BRAUN, M. LOHSCHEIDT, AND O. ZELDER, *Biotechnological production and applications of phytases*, Applied Microbiology and Biotechnology, 68 (2005), pp. 588–597.

[119] D. HALOŽAN, U. RIEBENTANZ, M. BRUMEN, AND E. DONATH, *Polyelectrolyte microcapsules and coated CaCO3 particles as fluorescence activated sensors in flowmetry*, Colloids and Surfaces A: Physicochemical and Engineering Aspects, 342 (2009), pp. 115–121.

[120] S. R. HAMILTON, R. C. DAVIDSON, N. SETHURAMAN, J. H. NETT, Y. JIANG, S. RIOS, P. BOBROWICZ, T. A. STADHEIM, H. LI, B.-K. CHOI, D. HOPKINS, H. WISCHNEWSKI, J. ROSER, T. MITCHELL, R. R. STRAWBRIDGE, J. HOOPES, S. WILDT, AND T. U. GERNGROSS, *Humanization of Yeast to Produce Complex Terminally Sialyated Glycoproteins*, Science, 313 (2006), pp. 1441–1443.

[121] S. R. HAMILTON AND T. U. GERNGROSS, *Glycosylation engineering in yeast: the advent of fully humanized yeast*, Current Opinion in Biotechnology, 18 (2007), pp. 387–392.

[122] M. HAN, X. WANG, G. YAN, W. WANG, Y. TAO, X. LIU, H. CAO, AND X. YU, *Modification of recombinant elastase expressed in Pichia pastoris by introduction of N-glycosylation sites*, Journal of Biotechnology, 171 (2014), pp. 3–7.

[123] Y. HAN AND X. G. LEI, *Role of glycosylation in the functional expression of an Aspergillus niger phytase (phyA) in Pichia pastoris*, Archives of biochemistry and biophysics, 364 (1999), pp. 83–90.

[124] F. S. HARTNER, C. RUTH, D. LANGENEGGER, S. N. JOHNSON, P. HYKA, G. P. LIN-CEREGHINO, J. LIN-CEREGHINO, K. KOVAR, J. M. CREGG, AND A. GLIEDER, *Promoter library designed for fine-tuned gene expression in Pichia pastoris.*, Nucleic Acids Research, 36 (2008), p. e76.

[125] M. HASSLACHER, M. SCHALL, M. HAYN, R. BONA, K. RUMBOLD, J. LÜCKL, H. GRIENGL, S. D. KOHLWEIN, AND H. SCHWAB, *High-level Intracellular Expression of Hydroxynitrile Lyase from the Tropical Rubber Tree Hevea brasiliensis in Microbial Hosts*, Protein Expression and Purification, 11 (1997), pp. 61–71.

[126] Z. HE, Y. HUANG, Y. QIN, Z. LIU, D. MO, P. CONG, AND Y. CHEN, *Comparison of Alpha-Factor Preprosequence and a Classical Mammalian Signal Peptide for Secretion of Recombinant Xylanase xynB from Yeast Pichia pastoris*, Journal of Microbiology and Biotechnology, 22 (2012), pp. 479–483.

[127] R. S. HEGDE AND H. D. BERNSTEIN, *The surprising complexity of signal sequences*, Trends in Biochemical Sciences, 31 (2006), pp. 563–571.

[128] H. HEIMO, K. PALMU, AND I. SUOMINEN, *Expression in Pichia pastoris and Purification of Aspergillus awamori Glucoamylase Catalytic Domain*, Protein Expression and Purification, 10 (1997), pp. 70–79.

[129] A. HESAMPOUR, S. E. R. SIADAT, M. A. MALBOOBI, N. MOHANDESI, S. S. ARAB, AND M. M. GHAHREMANPOUR, *Enhancement of Thermostability and Kinetic Efficiency of Aspergillus niger PhyA Phytase by Site-Directed Mutagenesis*, Applied Biochemistry and Biotechnology, 175 (2015), pp. 2528–2541.

[130] M. HOFRICHTER AND R. ULLRICH, *Oxidations catalyzed by fungal peroxygenases*, Current Opinion in Chemical Biology, 19 (2014), pp. 116–125.

[131] M. HOFRICHTER, R. ULLRICH, M. J. PECYNA, C. LIERS, AND T. LUNDELL, *New and classic families of secreted fungal heme peroxidases*, Applied Microbiology and Biotechnology, 87 (2010), pp. 871–897.

[132] I. HOLZMEISTER, M. SCHAMEL, J. GROLL, U. GBURECK, AND E. VORNDRAN, *Artificial inorganic biohybrids: The functional combination of microorganisms and cells with inorganic materials*, Acta Biomaterialia, 74 (2018), pp. 17–35.

[133] R. M. HORTON, H. D. HUNT, S. N. HO, J. K. PULLEN, AND L. R. PEASE, *Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension*, Gene, 77 (1989), pp. 61–68.

[134] J. HOU, K. E. J. TYO, Z. LIU, D. PETRANOVIC, AND J. NIELSEN, *Metabolic engineering of recombinant protein secretion by Saccharomyces cerevisiae*, FEMS Yeast Research, 12 (2012), pp. 491–510.

[135] C.-J. HUANG, L. M. DAMASCENO, K. A. ANDERSON, S. ZHANG, L. J. OLD, AND C. A. BATT, *A proteomic analysis of the Pichia pastoris secretome in methanol-induced cultures*, Applied Microbiology and Biotechnology, 90 (2011), pp. 235–247.

[136] K. HUYNH AND C. L. PARTCH, *Analysis of protein stability and ligand interactions by thermal shift assay*, Current Protocols in Protein Science, (2016), pp. 1–19.

[137] A. IDIRIS, H. TOHDA, H. KUMAGAI, AND K. TAKEGAWA, *Engineering of protein secretion in yeast: strategies and impact on protein production*, Applied Microbiology and Biotechnology, 86 (2010), pp. 403–417.

[138] M. ITAYA, K. FUJITA, A. KUROKI, AND K. TSUGE, *Bottom-up genome assembly using the Bacillus subtilis genome vector*, Nature Methods, 5 (2008), pp. 41–43.

[139] P. P. JACOBS, S. GEYSENS, W. VERVECKEN, R. CONTRERAS, AND N. CALLEWAERT, *Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology*, Nature Protocols, 4 (2009), pp. 58–70.

[140] M. JAHIC, F. WALLBERG, M. BOLLOK, P. GARCIA, AND S.-O. ENFORS, *Temperature limited fed-batch technique for control of proteolysis in Pichia pastoris bioreactor cultures*, Microbial Cell Factories, 2 (2003), pp. 1–11.

[141] L. R. JARBOE, *YqhD: a broad-substrate range aldehyde reductase with various applications in production of biorenewable fuels and chemicals*, Applied Microbiology and Biotechnology, 89 (2011), pp. 249–257.

[142] A. P. R. JOHNSTON, C. CORTEZ, A. S. ANGELATOS, AND F. CARUSO, *Layer-by-layer engineered capsules and their applications*, Current Opinion in Colloid and Interface Science, 11 (2006), pp. 203–209.

[143] J. JORDÀ, P. JOUHTEN, E. CÁMARA, H. MAAHEIMO, J. ALBIOL, AND P. FERRER, *Metabolic flux profiling of recombinant protein secreting Pichia pastoris growing on glucose:methanol mixtures*, Microbial Cell Factories, 11 (2012), pp. 1–14.

[144] V. JUTURU AND J. C. WU, *Heterologous Protein Expression in Pichia pastoris: Latest Research Progress and Applications*, ChemBioChem, 19 (2018), pp. 7–21.

[145] R. KACKER, E. LAGERGREN, AND J. FILLIBEN, *Taguchi's orthogonal arrays are classical designs of experiments*, Journal of Research of the National Institute of Standards and Technology, 96 (1991), p. 577.

[146] C. A. KAISER, D. PREUSS, P. GRISAFI, AND D. BOTSTEIN, *Many Random Sequences Functionally Replace the Secretion Signal Sequence of Yeast Invertase*, Science, 235 (1987), pp. 312–317.

[147] M. KARAOGLAN, F. E. KARAOGLAN, AND M. INAN, *Functional analysis of alcohol dehydrogenase (ADH) genes in Pichia pastoris*, Biotechnology Letters, 38 (2016), pp. 463–469.

[148] M. KARAOGLAN, H. YILDIZ, AND M. INAN, *Screening of signal sequences for extracellular production of Aspergillus niger xylanase in Pichia pastoris*, Biochemical Engineering Journal, 92 (2014), pp. 16–21.

[149] A. KARICH, K. SCHEIBNER, R. ULLRICH, AND M. HOFRICHTER, *Exploring the catalase activity of unspecific peroxygenases and the mechanism of peroxide-dependent heme destruction*, Journal of Molecular Catalysis B: Enzymatic, 134 (2016), pp. 238–246.

[150] A. KARICH, R. ULLRICH, K. SCHEIBNER, AND M. HOFRICHTER, *Fungal Unspecific Peroxygenases Oxidize the Majority of Organic EPA Priority Pollutants*, Frontiers in Microbiology, 8 (2017), pp. 1–26.

[151] M. KAVŠČEK, M. STRAŽAR, T. CURK, K. NATTER, AND U. PETROVIČ, *Yeast as a cell factory: current state and perspectives.*, Microbial Cell Factories, 14 (2015), pp. 1–10.

[152] R. KELWICK, J. T. MACDONALD, A. J. WEBB, AND P. FREEMONT, *Developments in the tools and methodologies of synthetic biology*, Frontiers in Bioengineering and Biotechnology, 2 (2014), pp. 1–23.

[153] N. K. KHATRI AND F. HOFFMANN, *Impact of Methanol Concentration on Secreted Protein Production in Oxygen-Limited Cultures of Recombinant Pichia pastoris*, Biotechnology and Bioengineering, 93 (2006), pp. 871–879.

[154] E. KIM, B. S. MOORE, AND Y. J. YOON, *Reinvigorating natural product combinatorial biosynthesis with synthetic biology*, Nature Chemical Biology, 11 (2015), pp. 649–659.

[155] H. KIM, S. J. YOO, AND H. A. KANG, *Yeast synthetic biology for the production of recombinant therapeutic proteins*, FEMS Yeast Research, 15 (2015), pp. 1–16.

[156] J. KIM, M. SALVADOR, E. SAUNDERS, J. GONZÁLEZ, C. AVIGNONE-ROSSA, AND J. JIMÉNEZ, *Properties of alternative microbial hosts used in synthetic biology: towards the design of a modular chassis*, Essays In Biochemistry, 60 (2016), pp. 303–313.

[157] M. KIMURA, A. TAKATSUKI, AND I. YAMAGUCHI, *Blasticidin S deaminase gene from Aspergillus terreus (BSD): a new drug resistance gene for transfection of mammalian cells*, Biochimica et Biophysica Acta, 1219 (1994), pp. 653–659.

[158] O. KIRK, T. V. BORCHERT, AND C. C. FUGLSANG, *Industrial enzyme applications*, Current Opinion in Biotechnology, 13 (2002), pp. 345–351.

[159] M. KLUGE, R. ULLRICH, C. DOLGE, K. SCHEIBNER, AND M. HOFRICHTER, *Hydroxylation of naphthalene by aromatic peroxygenase from Agrocybe aegerita proceeds via oxygen transfer from H2O2 and intermediary epoxidation*, Applied Microbiology and Biotechnology, 81 (2009), pp. 1071–1076.

[160] M. KLUMPP, A. BOETTCHER, D. BECKER, G. MEDER, J. BLANK, L. LEDER, M. FORSTNER, J. OTTL, AND L. M. MAYR, *Readout Technologies for Highly Miniaturized Kinase Assays Applicable to High-Throughput Screening in a 1536-Well Format*, Journal of Biomolecular Screening, 11 (2006), pp. 617–633.

[161] T. KNIGHT, *Idempotent Vector Design for Standard Assembly of Biobricks*, tech. rep., 2003.

[162] K. KOBAYASHI, S. KUWAE, T. OHYA, T. OHDA, M. OHYAMA, H. OHI, K. TOMOMITSU, AND T. OHMURA, *High-level Expression of Recombinant Human Serum Albumin from the Methylotrophic Yeast Pichia pastoris with Minimal Protease Production and Activation*, Journal of Bioscience and Bioengineering, 89 (2000), pp. 55–61.

[163] S. D. KOK, L. H. STANTON, T. SLABY, M. DUROT, V. F. HOLMES, K. G. PATEL, D. PLATT, E. B. SHAPLAND, Z. SERBER, J. DEAN, J. D. NEWMAN, AND S. S. CHANDRAN, *Rapid and Reliable DNA Assembly via Ligase Cycling Reaction*, ACS Synthetic Biology, 3 (2014), pp. 97–106.

[164] F. W. KRAINER, S. CAPONE, M. JÄGER, T. VOGL, M. GERSTMANN, A. GLIEDER, C. HERWIG, AND O. SPADIUT, *Optimizing cofactor availability for the production of recombinant heme peroxidase in Pichia pastoris*, Microbial Cell Factories, 14 (2015), p. 4.

[165] O. KREFT, M. PREVOT, H. MÖHWALD, AND G. B. SUKHORUKOV, *Shell-in-Shell Microcapsules: A Novel Tool for Integrated, Spatially Confined Enzymatic Reactions*, Angewandte Chemie - International Edition, 46 (2007), pp. 5605–5608.

[166] A. KÜBERL, J. SCHNEIDER, G. G. THALLINGER, I. ANDERL, D. WIBBERG, T. HAJEK, S. JAENICKE, K. BRINKROLF, A. GOESMANN, R. SZCZEPANOWSKI, A. PÜHLER, H. SCHWAB, A. GLIEDER, AND H. PICHLER, *High-quality genome sequence of Pichia pastoris CBS7435*, Journal of Biotechnology, 154 (2011), pp. 312–320.

[167] S. KUWAE, M. OHYAMA, T. OHYA, H. OHI, AND K. KOBAYASHI, *Production of Recombinant Kuman Antithrombin by Pichia pastoris*, Journal of Bioscience and Bioengineering, 99 (2005), pp. 264–271.

[168] A. LAMPROPOULOS, Z. SUTIKOVIC, C. WENZL, I. MAEGELE, J. U. LOHMANN, AND J. FORNER, *GreenGate - A Novel, Versatile, and Efficient Cloning System for Plant Transgenesis*, PLoS ONE, 8 (2013), p. e83043.

[169] S. F. LASSEN, J. BREINHOLT, P. R. ØSTERGAARD, R. BRUGGER, A. BISCHOFF, M. WYSS, AND C. C. FUGLSANG, *Expression, Gene Cloning, and Characterization of Five Novel Phytases from Four Basidiomycete Fungi: Peniophora lycii, Agrocybe pediades, a Ceriporia sp., and Trametes pubescens*, Applied and Environmental Microbiology, 67 (2001), pp. 4701–4707.

[170] J. W. LEE, D. O. KANG, B. Y. KIM, W. K. OH, T. I. MHEEN, Y. R. PYUN, AND J. S. AHN, *Mutagenesis of the glucoamylase signal peptide of Saccharomyces diastaticus and functional analysis in Saccharomyces cerevisiae*, FEMS Microbiology Letters, 193 (2000), pp. 7–11.

[171] M. E. LEE, A. ASWANI, A. S. HAN, C. J. TOMLIN, AND J. E. DUEBER, *Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay*, Nucleic Acids Research, 41 (2013), pp. 10668–10678.

[172] M. E. LEE, W. C. DELOACHE, B. CERVANTES, AND J. E. DUEBER, *A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly*, ACS Synthetic Biology, 4 (2015), pp. 975–986.

[173] M. LEI, P. G. LI, Z. B. SUN, AND W. H. TANG, *Effects of organic additives on the morphology of calcium carbonate particles in the presence of CTAB*, Materials Letters, 60 (2006), pp. 1261–1264.

[174] C. LI, Y. LIN, X. ZHENG, Q. YUAN, N. PANG, AND X. LIAO, *Recycling of a selectable marker with a self-excisable plasmid in Pichia pastoris*, Scientific Reports, 7 (2017), pp. 1–10.

[175] M. LI AND I. BORODINA, *Application of synthetic biology for production of chemicals in yeast Saccharomyces cerevisiae*, FEMS Yeast Research, 15 (2015), pp. 1–12.

[176] M. V. LI, D. SHUKLA, B. H. RHODES, A. LALL, J. SHU, B. S. MORIARITY, AND D. A. LARGAESPADA, *HomeRun Vector Assembly System: A Flexible and Standardized Cloning System for Assembly of Multi-Modular DNA Constructs*, PLoS ONE, 9 (2014), p. e100948.

[177] M. Z. LI AND S. J. ELLEDGE, *Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC*, Nature Methods, 4 (2007), pp. 251–256.

[178] S. LIANG, C. LI, Y. YE, AND Y. LIN, *Endogenous signal peptides efficiently mediate the secretion of recombinant proteins in Pichia pastoris*, Biotechnology Letters, 35 (2013), pp. 97–105.

[179] S. LIANG, C. ZOU, Y. LIN, X. ZHANG, AND Y. YE, *Identification and characterization of PGCW14: A novel, strong constitutive promoter of Pichia pastoris*, Biotechnology Letters, 35 (2013), pp. 1865–1871.

[180] Y. LIAO, C. M. LI, H. CHEN, Q. WU, Z. SHAN, AND X. Y. HAN, *Site-Directed Mutagenesis Improves the Thermostability and Catalytic Efficiency of Aspergillus niger N25 Phytase Mutated by I44E and T252R*, Applied Biochemistry and Biotechnology, 171 (2013), pp. 900–915.

[181] Y. LIAO, M. ZENG, Z. F. WU, H. CHEN, H. N. WANG, Q. WU, Z. SHAN, AND X. Y. HAN, *Improving Phytase Enzyme Activity in a Recombinant phyA Mutant Phytase from Aspergillus niger N25 by Error-Prone PCR*, Applied Biochemistry and Biotechnology, 166 (2012), pp. 549–562.

[182] G. P. LIN-CEREGHINO, A. JAY, J. SUNGA, M. GLEESON, G. P. LIN, J. LIN, A. JAY, M. A. JOHNSON, M. LIM, M. A. G. GLEESON, AND J. M. CREGG, *New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of Pichia pastoris*, Gene, 263 (2001), pp. 159–169.

[183] G. P. LIN-CEREGHINO, C. M. STARK, D. KIM, J. CHANG, N. SHAHEEN, H. POERWANTO, K. AGARI, P. MOUA, L. K. LOW, N. TRAN, A. D. HUANG, M. NATTESTAD, K. T. OSHIRO, J. W. CHANG, A. CHAVAN, J. W. TSAI, AND J. LIN-CEREGHINO, *The effect of α-mating factor secretion signal mutations on recombinant protein expression in Pichia pastoris*, Gene, 519 (2013), pp. 311–317.

[184] J. K. LIU, W. H. CHEN, S. X. REN, G. P. ZHAO, AND J. WANG, *iBrick: A New Standard for Iterative Assembly of Biological Parts with Homing Endonucleases*, PLoS ONE, 9 (2014), p. e110852.

[185] M. C. LO, A. AULABAUGH, G. JIN, R. COWLING, J. BARD, M. MALAMAS, AND G. ELLESTAD, *Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery*, Analytical Biochemistry, 332 (2004), pp. 153–159.

[186] C. K. LONGWELL, L. LABANIEH, AND J. R. COCHRAN, *High-throughput screening technologies for enzyme engineering*, Current Opinion in Biotechnology, 48 (2017), pp. 196–202.

[187] M. LÖOKE, K. KRISTJUAHAN, AND A. KRISTJUHAN, *Extraction of Genomic DNA From Yeasts for PCR- Based Applications*, Biotechniques, 50 (2011), pp. 325–328.

[188] V. LOOSER, B. BRÜHLMANN, F. BUMBAK, C. STENGER, M. COSTA, A. CAMATTARI, D. FOTIADIS, AND K. KOVAR, *Cultivation strategies to enhance productivity of Pichia pastoris: A review*, Biotechnology Advances, 33 (2014), pp. 1177–1193.

[189] K. R. LOVE, N. C. DALVIE, AND J. C. LOVE, *The yeast stands alone: the future of protein biologic production*, Current Opinion in Biotechnology, 53 (2018), pp. 50–58.

[190] K. R. LOVE, T. J. POLITANO, V. PANAGIOTOU, B. JIANG, T. A. STADHEIM, AND C. LOVE, *Systematic Single-Cell Analysis of Pichia pastoris Reveals Secretory Capacity Limits Productivity*, PLoS ONE, 7 (2012), p. e37915.

[191] H. LUND, L. KALUM, M. HOFRICHTER, AND S. PETER, *Epoxidation using Peroxygenases*, 2017.

[192] K. MADDEN, I. TOLSTORUKOV, AND J. CREGG, *Electroporation of Pichia pastoris*, in Genetic Transformation Systems in Fungi, vol. 1, 2015, pp. 87–91.

[193] C. MADZAK, *Yarrowia lipolytica: recent achievements in heterologous protein expression and pathway engineering*, Applied Microbiology and Biotechnology, 99 (2015), pp. 4559–4577.

[194] S. MANN, *Molecular recognition in biomineralization*, Nature, 332 (1988), pp. 119–124.

[195] V. MAPELLI, *Yeast Metabolic Engineering*, Humana Press, 2014.

[196] A. T. MARTÍNEZ, F. J. RUIZ-DUEÑAS, S. CAMARERO, A. SERRANO, D. LINDE, H. LUND, J. VIND, M. TOVBORG, O. M. HEROLD-MAJUMDAR, M. HOFRICHTER, C. LIERS, R. ULLRICH, K. SCHEIBNER, G. SANNIA, A. PISCITELLI, C. PEZZELLA, M. E. SENER, S. KILIÇ, W. J. VAN BERKEL, V. GUALLAR, M. F. LUCAS, R. ZUHSE, R. LUDWIG, F. HOLLMANN, E. FERNÁNDEZ-FUEYO, E. RECORD, C. B. FAULDS, M. TORTAJADA, I. WINCKELMANN, J. A. RASMUSSEN, M. GELO-PUJIC, A. GUTIÉRREZ, J. C. DEL RÍO, J. RENCORET, AND M. ALCALDE, *Oxidoreductases on their way to industrial biotransformations*, Biotechnology Advances, 35 (2017), pp. 815–831.

[197] A. T. MARTÍNEZ, F. J. RUIZ-DUEÑAS, A. GUTIÉRREZ, J. C. DEL RÍO, M. ALCALDE, C. LIERS, R. ULLRICH, M. HOFRICHTER, K. SCHEIBNER, L. KALUM, J. VIND, AND H. LUND, *Search, engineering, and applications of new oxidative biocatalysts*, Biofuels, Bioproducts and Biorefining, 8 (2014), pp. 819–835.

[198] B. MARTOGLIO AND B. DOBBERSTEIN, *Signal sequences: more than just greasy peptides*, Trends in Cell Biology, 8 (1998), pp. 410–415.

[199] A. MASSAHI AND P. ÇALIK, *Endogenous signal peptides in recombinant protein production by Pichia pastoris: From in-silico analysis to fermentation*, Journal of Theoretical Biology, 408 (2016), pp. 22–33.

[200] A. MASSAHI AND P. ÇALIK, *In-silico determination of Pichia pastoris signal peptides for extracellular recombinant protein production*, Journal of Theoretical Biology, 364 (2015), pp. 179–188.

[201] D. MATTANOVICH, M. SAUER, AND B. GASSER, *Yeast biotechnology: teaching the old dog new tricks.*, Microbial Cell Factories, 13 (2014), p. 34.

[202] L. M. MAYR AND D. BOJANIC, *Novel trends in high-throughput screening*, Current Opinion in Pharmacology, 9 (2009), pp. 580–588.

[203] M. A. J. MAZUMDER, N. A. D. BURKE, F. SHEN, M. A. POTTER, AND H. D. H. STÖVER, *Core-Cross-Linked Alginate Microcapsules for Cell Encapsulation*, Biomacromolecules, 10 (2009), pp. 1365–1373.

[204] M. A. MEEHL AND T. A. STADHEIM, *Biopharmaceutical discovery and production in yeast*, Current Opinion in Biotechnology, 30 (2014), pp. 120–127.

[205] B. D. MEHTA, S. P. JOG, S. C. JOHNSON, AND P. P. N. MURTHY, *Lily pollen alkaline phytase is a histidine phosphatase similar to mammalian multiple inositol polyphosphate phosphatase (MINPP)*, Phytochemistry, 67 (2006), pp. 1874–1886.

[206] D. T. MOIR AND D. R. DUMAIS, *Glycosylation and secretion of human alpha-1-antitrypsin by yeast.*, Gene, 56 (1987), pp. 209–217.

[207] P. MOLINA-ESPEJA, P. G. DE SANTOS, AND M. ALCALDE, *Directed Evolution of Unspecific Peroxygenase*, in Directed Enzyme Evolution: Advances and Applications, M. Alcalde, ed., Cham, Switzerland, 2017, ch. 5, pp. 127–143.

[208] P. MOLINA-ESPEJA, E. GARCIA-RUIZ, D. GONZALEZ-PEREZ, R. ULLRICH, M. HOFRICHTER, AND M. ALCALDE, *Directed Evolution of Unspecific Peroxygenase from Agrocybe aegerita*, Applied and Environmental Microbiology, 80 (2014), pp. 3496–3507.

[209] P. MOLINA-ESPEJA, S. MA, D. M. MATE, R. LUDWIG, AND M. ALCALDE, *Tandem-yeast expression system for engineering and producing unspecific peroxygenase*, Enzyme and Microbial Technology, 73-74 (2015), pp. 29–33.

[210] S. J. MOORE, H. E. LAI, R. J. KELWICK, S. M. CHEE, D. J. BELL, K. M. POLIZZI, AND P. S. FREEMONT, *EcoFlex: A Multifunctional MoClo Kit for E. coli Synthetic Biology*, ACS Synthetic Biology, 5 (2016), pp. 1059–1069.

[211] J. W. MORSE, R. S. ARVIDSON, AND A. LÜTTGE, *Calcium Carbonate Formation and Dissolution*, Chemical Reviews, 107 (2007), pp. 342–381.

[212] N. J. MORSE, M. R. GOPAL, J. M. WAGNER, AND H. S. ALPER, *Yeast Terminator Function Can Be Modulated and Designed on the Basis of Predictions of Nucleosome Occupancy*, ACS Synthetic Biology, 6 (2017), pp. 2086–2095.

[213] E. J. MULLANEY AND A. H. ULLAH, *The term phytase comprises several different classes of enzymes*, Biochemical and Biophysical Research Communications, 312 (2003), pp. 179–184.

[214] A. MULLARD, *2012 FDA drug approvals*, Nature Reviews Drug Discovery, 12 (2013), pp. 87–90.

[215] L. NÄÄTSAARI, B. MISTLBERGER, C. RUTH, T. HAJEK, F. S. HARTNER, AND A. GLIEDER, *Deletion of the Pichia pastoris ku70 Homologue Facilitates Platform Strain Generation for Gene Expression and Synthetic Biology*, PLoS ONE, 7 (2012), p. e39720.

[216] C. M. NEALON, M. M. MUSA, J. M. PATEL, AND R. S. PHILLIPS, *Controlling Substrate Specificity and Stereospecificity of Alcohol Dehydrogenases*, ACS Catalysis, 5 (2015), pp. 2100–2114.

[217] J. H. NETT, N. HODEL, S. RAUSCH, AND S. WILDT, *Cloning and disruption of the Pichia pastoris ARG1, ARG2, ARG3, HIS1, HIS2, HIS5, HIS6 genes and their use as auxotrophic markers*, Yeast, 22 (2005), pp. 295–304.

[218] H. NEVALAINEN AND R. PETERSON, *Making recombinant proteins in filamentous fungi- are we expecting too much?*, Frontiers in Microbiology, 5 (2014), pp. 1–10.

[219] Y. NIE, S. WANG, Y. XU, S. LUO, Y. L. ZHAO, R. XIAO, G. T. MONTELIONE, J. F. HUNT, AND T. SZYPERSKI, *Enzyme Engineering Based on X-ray Structures and Kinetic Profiling of Substrate Libraries: Alcohol Dehydrogenases for Stereospecific Synthesis of a Broad Range of Chiral Alcohols*, ACS Catalysis, 8 (2018), pp. 5145–5152.

[220] A. NISHIGUCHI, H. YOSHIDA, M. MATSUSAKI, AND M. AKASHI, *Rapid Construction of Three-Dimensional Multilayered Tissues with Endothelial Tube Networks by the Cell-Accumulation Technique*, Advanced Materials, 23 (2011), pp. 3506–3510.

[221] C. NOTREDAME, D. G. HIGGINS, AND J. HERINGA, *T-Coffee: A novel method for fast and accurate multiple sequence alignment.*, Journal of Molecular Biology, 302 (2000), pp. 205–217.

[222] F. OBERG, J. SJÖHAMN, M. T. CONNER, R. M. BILL, AND K. HEDFALK, *Improving recombinant eukaryotic membrane protein yields in Pichia pastoris: The importance of codon optimization and clone selection*, Molecular Membrane Biology, 28 (2011), pp. 398–411.

[223] U. OBST, *Expression and Secretion of bacterial alcohol dehydrogenases in Aspergillus niger*, Tech. Rep. Diploma thesis, TU Berlin, 2013.

[224] U. OBST, T. K. LU, AND V. SIEBER, *A modular toolkit for generating Pichia pastoris secretion libraries*, ACS Synthetic Biology, 6 (2017), pp. 1016–1025.

[225] K. OGATA, H. NISHIKAWA, AND M. OHSUGI, *A Yeast Capable of Utilizing Methanol*, Agricultural and Biological Chemistry, 33 (1969), pp. 1519–1520.

[226] B. C. OH, W. C. CHOI, S. PARK, Y. O. KIM, AND T. K. OH, *Biochemical properties and substrate specificities of alkaline and histidine acid phytases*, Applied Microbiology and Biotechnology, 63 (2004), pp. 362–372.

[227] M. S. PACKER AND D. R. LIU, *Methods for the directed evolution of proteins*, Nature Reviews Genetics, 16 (2015), pp. 379–394.

[228] E. PAIFER, E. MARGOLLES, J. CREMATA, R. MONTESINO, L. HERRERA, AND J. M. DELGADO, *Efficient Expression and Secretion of Recombinant Alpha Amylase in Pichia pastoris using two different Signal Sequences*, Yeast, 10 (1994), pp. 1415–1419.

[229] R. M. PAPI, S. A. CHAITIDOU, F. A. TRIKKA, AND D. A. KYRIAKIDIS, *Encapsulated Escherichia coli in alginate beads capable of secreting a heterologous pectin lyase*, Microbial Cell Factories, 4 (2005).

[230] D. PARASHAR AND T. SATYANARAYANA, *Enhancing the production of recombinant acidic α-amylase and phytase in Pichia pastoris under dual promoters [constitutive (GAP) and inducible (AOX)] in mixed fed batch high cell density cultivation*, Process Biochemistry, 51 (2016), pp. 1315–1322.

[231] W. M. PATRICK, A. E. FIRTH, AND J. M. BLACKBURN, *User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries*, Protein Engineering Design and Selection, 16 (2003), pp. 451–457.

[232] N. J. PATRON, *DNA assembly for plant biology: Techniques and tools*, Current Opinion in Plant Biology, 19 (2014), pp. 14–19.

[233] D. A. PEÑA, B. GASSER, J. ZANGHELLINI, M. G. STEIGER, AND D. MATTANOVICH, *Metabolic engineering of Pichia pastoris*, Metabolic Engineering, (2018).

[234] P. PEREZ-PINERA, N. HAN, S. CLETO, J. CAO, O. PURCELL, K. A. SHAH, K. LEE, R. RAM, AND T. K. LU, *Synthetic biology and microbioreactor platforms for programmable production of biologics at the point-of-care*, Nature Communications, 7 (2016), pp. 1–10.

[235] S. PETER, A. KARICH, R. ULLRICH, G. GRÖBE, K. SCHEIBNER, AND M. HOFRICHTER, *Enzymatic one-pot conversion of cyclohexane into cyclohexanone: Comparison of four fungal peroxygenases*, Journal of Molecular Catalysis B: Enzymatic, 103 (2014), pp. 47–51.

[236] A. I. PETROV, D. V. VOLODKIN, AND G. B. SUKHORUKOV, *Protein-Calcium Carbonate Coprecipitation: A Tool for Protein Encapsulation*, Biotechnology Progress, 21 (2005), pp. 918–925.

[237] C. S. PEYRATOUT AND L. DÄHNE, *Tailor-made polyelectrolyte microcapsules: From multilayers to smart containers*, Angewandte Chemie - International Edition, 43 (2004), pp. 3762–3783.

[238] J. PFEIFENSCHNEIDER, T. BRAUTASET, AND VOLKER F. WENDISCH, *Methanol as carbon substrate in the bio-economy: Metabolic engineering of aerobic methylotrophic bacteria for production of value-added chemicals*, Biofuels, Bioproducts and Biorefining, 11 (2017), pp. 719–731.

[239] A. PICK, W. OTT, T. HOWE, J. SCHMID, AND V. SIEBER, *Improving the NADH-cofactor specificity of the highly active AdhZ3 and AdhZ2 from Escherichia coli K-12.*, Journal of Biotechnology, 189 (2014), pp. 157–165.

[240] A. PICK, B. RÜHMANN, J. SCHMID, AND V. SIEBER, *Novel CAD-like enzymes from Escherichia coli K-12 as additional tools in chemical production*, Applied Microbiology and Biotechnology, 97 (2012), pp. 5815–5824.

[241] D. PORRO, B. GASSER, T. FOSSATI, M. MAURER, P. BRANDUARDI, M. SAUER, AND D. MATTANOVICH, *Production of recombinant proteins and metabolites in yeasts*, Applied Microbiology and Biotechnology, 89 (2011), pp. 939–948.

[242] R. M. PORTELA, T. VOGL, C. KNIELY, J. E. FISCHER, R. OLIVEIRA, AND A. GLIEDER, *Synthetic Core Promoters as Universal Parts for Fine-Tuning Expression in Different Yeast Species*, ACS Synthetic Biology, 6 (2017), pp. 471–484.

[243] V. POTAPOV, J. L. ONG, R. B. KUCERA, B. W. LANGHORST, K. BILOTTI, J. M. PRYOR, E. J. CANTOR, B. CANTON, T. F. KNIGHT, T. C. EVANS, AND J. GREGORY, *Optimization of Golden Gate assembly through application of ligation sequence-dependent fidelity and bias profiling*, (2018).

[244] R. PRIELHOFER, J. J. BARRERO, S. STEUER, T. GASSLER, R. ZAHRL, K. BAUMANN, M. SAUER, D. MATTANOVICH, B. GASSER, AND H. MARX, *GoldenPiCS: A Golden Gate-derived modular cloning system for applied synthetic biology in the yeast Pichia pastoris*, BMC Systems Biology, 11 (2017), pp. 1–14.

[245] R. PRIELHOFER, S. P. CARTWRIGHT, A. B. GRAF, M. VALLI, R. M. BILL, D. MATTANOVICH, AND B. GASSER, *Pichia pastoris regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level.*, BMC Genomics, 16 (2015), pp. 1–17.

[246] R. PRIELHOFER, M. MAURER, J. KLEIN, J. WENGER, C. KIZIAK, B. GASSER, AND D. MATTANOVICH, *Induction without methanol: novel regulated promoters enable high-level expression in Pichia pastoris.*, Microbial Cell Factories, 12 (2013), pp. 1–10.

[247] P. E. M. PURNICK AND R. WEISS, *The second wave of synthetic biology: from modules to systems*, Nature Reviews Molecular Cell Biology, 10 (2009), pp. 410–422.

[248] V. PUXBAUM, D. MATTANOVICH, AND B. GASSER, *Quo vadis? The challenges of recombinant protein folding and secretion in Pichia pastoris*, Applied Microbiology and Biotechnology, 99 (2015), pp. 2925–2938.

[249] X. QIN, J. QIAN, C. XIAO, Y. ZHUANG, S. ZHANG, AND J. CHU, *Reliable high-throughput approach for screening of engineered constitutive promoters in the yeast Pichia pastoris*, Letters in Applied Microbiology, 52 (2011), pp. 634–641.

[250] X. QIN, J. QIAN, G. YAO, Y. ZHUANG, S. ZHANG, AND J. CHU, *GAP Promoter Library for Fine-Tuning of Gene Expression in Pichia pastoris.*, Applied and Environmental Microbiology, 77 (2011), pp. 3600–3608.

[251] D. QUAGLIA, M. C. EBERT, P. F. MUGFORD, AND J. N. PELLETIER, *Enzyme engineering: A synthetic biology approach for more effective library generation and automated high-throughput screening*, PLoS ONE, 12 (2017), p. e0171741.

[252] J. QUAN AND J. TIAN, *Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways*, Methods in Molecular Biology, 4 (2014), p. e6441.

[253] L. QVIRIST, N.-G. CARLSSON, AND T. ANDLID, *Assessing phytase activity–methods, definitions and pitfalls*, Journal of Biological Methods, 2 (2015), p. 16.

[254] R. J. M. RAEMAEKERS, L. DE MURO, J. A. GATEHOUSE, AND A. P. FORDHAM-SKELTON, *Functional phytohemagglutinin (PHA) and Galanthus nivalis agglutinin (GNA) expressed in Pichia pastoris. Correct N-terminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide*, European Journal of Biochemistry, 265 (1999), pp. 394–403.

[255] A. RAI, *Law and Technology Unstandard Standardization: The Case of Biology*, Communications of the ACM, 53 (2010), pp. 37–39.

[256] B. RANJAN AND T. SATYANARAYANA, *Recombinant HAP Phytase of the Thermophilic Mold Sporotrichum thermophile: Expression of the Codon-Optimized Phytase Gene in Pichia pastoris and Applications*, Molecular Biotechnology, 58 (2016), pp. 137–147.

[257] D. E. C. S. RAO, K. V. RAO, T. P. REDDY, AND V. D. REDDY, *Molecular characterization, physicochemical properties, known and potential applications of phytases: An overview*, Critical Reviews in Biotechnology, 29 (2009), pp. 182–198.

[258] H. RASCHMANOVÁ, A. WENINGER, A. GLIEDER, K. KOVAR, AND T. VOGL, *Implementing CRISPR-Cas technologies in conventional and non-conventional yeasts: Current state and future prospects*, Biotechnology Advances, 36 (2018), pp. 641–665.

[259] H. REDDEN, N. MORSE, AND H. S. ALPER, *The synthetic biology toolbox for tuning gene expression in yeast*, FEMS Yeast Research, 15 (2015), pp. 1–10.

[260] E. RODRIGUEZ, E. J. MULLANEY, AND X. G. LEI, *Expression of the Aspergillus fumigatus Phytase Gene in Pichia pastoris and Characterization of the Recombinant Enzyme*, Biochemical and Biophysical Research Communications, 268 (2000), pp. 373–378.

[261] E. RODRIGUEZ, Z. A. WOOD, P. KARPLUS, AND X. G. LEI, *Site-Directed Mutagenesis Improves Catalytic Efficiency and Thermostability of Escherichia coli pH 2.5 Acid Phosphatase/Phytase Expressed in Pichia pastoris*, Archives of Biochemistry and Biophysics, 382 (2000), pp. 105–112.

[262] S. ROLLIÉ, M. MANGOLD, AND K. SUNDMACHER, *Designing biological systems: Systems Engineering meets Synthetic Biology*, Chemical Engineering Science, 69 (2012), pp. 1–29.

[263] P. A. ROMERO AND F. H. ARNOLD, *Exploring protein fitness landscapes by directed evolution*, Nature Reviews Molecular Cell Biology, 10 (2009), pp. 866–876.

[264] K. E. ROYLE AND K. POLIZZI, *A streamlined cloning workflow minimising the time-to-strain pipeline for Pichia pastoris*, Scientific Reports, 7 (2017), pp. 1–10.

[265] C. RUTH, T. ZUELLIG, A. MELLITZER, R. WEIS, V. LOOSER, K. KOVAR, AND A. GLIEDER, *Variable production windows for porcine trypsinogen employing synthetic inducible promoter variants in Pichia pastoris*, Systems and Synthetic Biology, 4 (2010), pp. 181–191.

[266] C. M. J. SAGT, B. KLEIZEN, R. VERWAAL, M. D. M. DE JONG, W. H. MULLER, A. SMITS, C. VISSER, J. BOONSTRA, A. J. VERKLEIJ, AND C. T. VERRIPS, *Introduction of an N-Glycosylation Site Increases Secretion of Heterologous Proteins in Yeasts*, Applied and Environmental Microbiology, 66 (2000), pp. 4940–4944.

[267] H. M. SALIS, E. A. MIRSKY, AND C. A. VOIGT, *Automated design of synthetic ribosome binding sites to control protein expression*, Nature Biotechnology, 27 (2009), pp. 946–950.

[268] J. SCHINDELIN, I. ARGANDA-CARRERAS, E. FRISE, V. KAYNIG, M. LONGAIR, T. PIETZSCH, S. PREIBISCH, C. RUEDEN, S. SAALFELD, B. SCHMID, J. Y. TINEVEZ, D. J. WHITE, V. HARTENSTEIN, K. ELICEIRI, P. TOMANCAK, AND A. CARDONA, *Fiji: An open-source platform for biological-image analysis*, Nature Methods, 9 (2012), pp. 676–682.

[269] F. R. SCHMIDT, *Recombinant expression systems in the pharmaceutical industry*, Applied Microbiology and Biotechnology, 65 (2004), pp. 363–372.

[270] C. SCHREIBER, H. MÜLLER, O. BIRRENBACH, M. KLEIN, D. HEERD, T. WEIDNER, D. SALZIG, AND P. CZERMAK, *A high-throughput expression screening platform to optimize the production of antimicrobial peptides*, Microbial Cell Factories, 16 (2017), pp. 1–13.

[271] J. P. SCHWARZHANS, T. LUTTERMANN, M. GEIER, J. KALINOWSKI, AND K. FRIEHS, *Towards systems metabolic engineering in Pichia pastoris*, Biotechnology Advances, 35 (2017), pp. 681–710.

[272] C. A. SCORER, J. J. CLARE, W. R. MCCOMBIE, M. A. ROMANOS, AND K. SREEKRISHNA, *Rapid Selection Using G418 of High Copy Number Transformants of Pichia pastoris for High-level Foreign Gene Expression*, Nature Biotechnology, 12 (1994), pp. 181–184.

[273] M. SEMMLING, O. KREFT, A. M. JAVIER, G. B. SUKHORUKOV, J. KÄS, AND W. J. PARAK, *A Novel Flow-Cytometry-Based Assay for Cellular Uptake Studies of Polyelectrolyte Microcapsules*, Small, 4 (2008), pp. 1763–1768.

[274] Z. SHAO, H. ZHAO, AND H. ZHAO, *DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways*, Nucleic Acids Research, 37 (2009), p. e16.

[275] S. SHEN, G. SULTER, T. W. JEFFRIES, AND J. M. CREGG, *A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast Pichia pastoris*, Gene, 216 (1998), pp. 93–102.

[276] J. R. SHUSTER, *Gene expression in yeast: protein secretion*, Current Opinion in Biotechnology, 2 (1991), pp. 685–690.

[277] K. SIEBERTZ, D. VAN BEBBER, AND T. HOCHKIRCHEN, *Statistische Versuchsplanung*, VDI, 2010.

[278] G. SINCLAIR AND F. Y. CHOY, *Synonymous codon usage bias and the expression of human glucocerebrosidase in the methylotrophic yeast, Pichia pastoris*, Protein Expression and Purification, 26 (2002), pp. 96–105.

[279] B. SINGH AND T. SATYANARAYANA, *Characterization of a HAP–phytase from a thermophilic mould Sporotrichum thermophile*, Bioresource technology, 100 (2010), pp. 2046–2051.

[280] S. C. SLEIGHT, B. A. BARTLEY, J. A. LIEVIANT, AND H. M. SAURO, *In-fusion biobrick assembly and re-engineering*, Nucleic Acids Research, 38 (2010), pp. 2624–2636.

[281] M. J. SMANSKI, H. ZHOU, J. CLAESEN, B. SHEN, M. FISCHBACH, AND C. A. VOIGT, *Synthetic biology to access and expand nature's chemical diversity*, Nature Reviews Microbiology, 14 (2016), pp. 135–149.

[282] O. SMIDSRØD AND G. SKJÅK-BRÆK, *Alginate as immobilization matrix for cells*, Minerva Biotecnologica, 12 (2000), pp. 223–233.

[283] J. SODERHOLM, B. J. BEVIS, AND B. S. GLICK, *Vector for Pop-In / Pop-Out Gene Replacement in Pichia pastoris*, Biotechniques, 31 (2001), pp. 306–312.

[284] O. SÖHNEL AND J. W. MULLIN, *Precipitation of calcium carbonate*, Journal of Crystal Growth, 60 (1982), pp. 239–250.

[285] S. C. SPOHNER, V. SCHAUM, H. QUITMANN, AND P. CZERMAK, *Kluyveromyces lactis: An emerging tool in biotechnology*, Journal of Biotechnology, 222 (2016), pp. 104–116.

[286] G. STADLMAYR, A. MECKLENBRÄUKER, M. ROTHMÜLLER, M. MAURER, M. SAUER, D. MATTANOVICH, AND B. GASSER, *Identification and characterisation of novel Pichia pastoris promoters for heterologous protein production*, Journal of Biotechnology, 150 (2010), pp. 519–529.

[287] M. STORCH, A. CASINI, B. MACKROW, T. FLEMING, H. TREWHITT, T. ELLIS, AND G. S. BALDWIN, *BASIC: A New Biopart Assembly Standard for Idempotent Cloning Provides Accurate, Single-Tier DNA Assembly for Synthetic Biology*, ACS Synthetic Biology, 4 (2015), pp. 781–787.

[288] D. W. STROMAN, J. M. CREGG, M. M. HARPOLD, AND G. T. SPERL, *Transformation of yeasts of the genus pichia*, 1989.

[289] G. B. SUKHORUKOV, E. DONATH, H. LICHTENFELD, E. KNIPPEL, M. KNIPPEL, A. BUDDE, AND H. MÖHWALD, *Layer-by-layer self assembly of polyelectrolytes on colloidal particles*, Colloids and Surfaces A: Physicochemical and Engineering Aspects, 137 (1998), pp. 253–266.

[290] A. J. SUNGA AND J. M. CREGG, *The Pichia pastoris formaldehyde dehydrogenase gene (FLD1) as a marker for selection of multicopy expression strains of P. pastoris*, Gene, 330 (2004), pp. 39–47.

[291] S. S. TEREKHOV, I. V. SMIRNOV, A. V. STEPANOVA, T. V. BOBIK, Y. A. MOKRUSHINA, N. A. PONOMARENKO, A. A. BELOGUROV, M. P. RUBTSOVA, O. V. KARTSEVA, M. O. GOMZIKOVA, A. A. MOSKOVTSEV, A. S. BUKATIN, M. V. DUBINA, E. S. KOSTRYUKOVA, V. V. BABENKO, M. T. VAKHITOVA, A. I. MANOLOV, M. V. MALAKHOVA, M. A. KORNIENKO, A. V. TYAKHT, A. A. VANYUSHKINA, E. N. ILINA, P. MASSON, A. G. GABIBOV, AND S. ALTMAN, *Microfluidic droplet platform for ultrahigh-throughput single-cell screening of biodiversity*, Proceedings of the National Academy of Sciences, 114 (2017), pp. 2550–2555.

[292] K. TERPE, *Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems.*, Applied Microbiology and Biotechnology, 72 (2006), pp. 211–222.

[293] D. THOR, S. XIONG, C. C. ORAZEM, A. C. KWAN, J. M. CREGG, J. LIN-CEREGHINO, AND G. P. LIN-CEREGHINO, *Cloning and characterization of the Pichia pastoris MET2 gene as a selectable marker*, FEMS Yeast Research, 5 (2005), pp. 935–942.

[294] M. TOKUNAGA, N. WADA, AND F. HISHINUMA, *A novel yeast secretion vector utilizing secretion signal of killer toxin encoded on the yeast linear DNA plasmid pGKL1*, Biochemical and Biophysical Research Communications, 144 (1987), pp. 613–619.

[295] C.-S. TSAI, S. KWAK, T. L. TURNER, AND Y.-S. JIN, *Yeast synthetic biology toolbox and applications for biofuel production*, FEMS Yeast Research, 15 (2015), pp. 1–15.

[296] J. TSCHOPP, G. SVERLOW, R. KOSSON, W. CRAIG, AND L. GRINNA, *High-level secretion of glycosylated invertase in the methylotrophic yeast, Pichia pastoris*, Nature Biotechnology, 5 (1987), pp. 1305–1308.

[297] H. UEMURA, Y. JIGAMI, H. TANAKA, N. TOSHIMITSU, M. PATERSON, AND S. NAKASATO, *Nucleotide sequence of the 5' flanking region responsible for the enhancement of the expression of yeast enolase 1 gene.*, Journal of biochemistry, 98 (1985), pp. 859–862.

[298] R. ULLRICH, J. NÜSKE, K. SCHEIBNER, J. SPANTZEL, AND M. HOFRICHTER, *Novel Haloperoxidase from the Agaric Basidiomycete Agrocybe aegerita Oxidizes Aryl Alcohols and Aldehydes*, Applied and Environmental Microbiology, 70 (2004), pp. 4575–4581.

[299] M. V. USHASREE, K. SHYAM, J. VIDYA, AND A. PANDEY, *Microbial phytase: Impact of advances in genetic engineering in revolutionizing its properties and applications*, Bioresource Technology, 245 (2017), pp. 1790–1799.

[300] Y. VERVOORT, A. G. LINARES, M. RONCORONI, C. LIU, J. STEENSELS, AND K. J. VERSTREPEN, *High-throughput system-wide engineering and screening for microbial biotechnology*, Current Opinion in Biotechnology, 46 (2017), pp. 120–125.

[301] T. VOGL AND A. GLIEDER, *Regulation of Pichia pastoris promoters and its consequences for protein production*, New Biotechnology, 30 (2013), pp. 385–404.

[302] T. VOGL, F. S. HARTNER, AND A. GLIEDER, *New opportunities by synthetic biology for biopharmaceutical production in Pichia pastoris*, Current Opinion in Biotechnology, 24 (2013), pp. 1094–1101.

[303] T. VOGL, C. RUTH, J. PITZER, T. KICKENWEIZ, AND A. GLIEDER, *Synthetic core promoters for Pichia pastoris*, ACS Synthetic Biology, 3 (2014), pp. 188–191.

[304] T. VOGL, L. STURMBERGER, T. KICKENWEIZ, R. WASMAYER, C. SCHMID, A. M. HATZL, M. A. GERSTMANN, J. PITZER, M. WAGNER, G. G. THALLINGER, M. GEIER, AND A. GLIEDER, *A Toolbox of Diverse Promoters Related to Methanol Utilization: Functionally Verified Parts for Heterologous Pathway Expression in Pichia pastoris*, ACS Synthetic Biology, 5 (2016), pp. 172–186.

[305] T. VOGL, L. STURMBERGER, T. KICKENWEIZ, R. WASMAYER, C. SCHMID, A.-M. HATZL, M. A. GERSTMANN, J. PITZER, M. WAGNER, G. G. THALLINGER, M. GEIER, AND A. GLIEDER, *A toolbox of diverse promoters related to methanol utilization – functionally verified parts for heterologous pathway expression in Pichia pastoris*, ACS Synthetic Biology, 5 (2016), pp. 172–186.

[306] D. V. VOLODKIN, N. I. LARIONOVA, AND G. B. SUKHORUKOV, *Protein Encapsulation via Porous CaCO3 Microparticles Templating*, Biomacromolecules, 5 (2004), pp. 1962–1972.

[307] D. V. VOLODKIN, A. I. PETROV, M. PREVOT, AND G. B. SUKHORUKOV, *Matrix Polyelectrolyte Microcapsules: New System for Macromolecule Encapsulation*, Langmuir, 20 (2004), pp. 3398–3406.

[308] R. VON KLITZING AND H. MÖHWALD, *A Realistic Diffusion Model for Ultrathin Polyelectrolyte Films*, Macromolecules, 29 (1996), pp. 6901–6906.

[309] J. M. WAGNER AND H. S. ALPER, *Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances*, Fungal Genetics and Biology, 89 (2016), pp. 126–136.

[310] B. WANG, P. LIU, W. JIANG, H. PAN, X. XU, AND R. TANG, *Yeast Cells with an Artificial Mineral Shell: Protection and Modification of Living Cells by Biomimetic Mineralization*, Angewandte Chemie - International Edition, 47 (2008), pp. 3560–3564.

[311] H. R. WATERHAM, M. E. DIGAN, P. J. KOUTZ, S. V. LAIR, AND J. M. CREGG, *Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter.*, Gene, 186 (1997), pp. 37–44.

[312] E. WEBER, C. ENGLER, R. GRUETZNER, S. WERNER, AND S. MARILLONNET, *A Modular Cloning System for Standardized Assembly of Multigene Constructs*, PLoS ONE, 6 (2011), p. e16765.

[313] G. WEGNER, *Emerging applications of the methylotrophic yeasts*, FEMS Microbiology Reviews, 87 (1990), pp. 279–284.

[314] L. WEI, Z. WANG, G. ZHANG, AND B. YE, *Characterization of Terminators in Saccharomyces cerevisiae and an Exploration of Factors Affecting Their Strength*, ChemBioChem, 18 (2017), pp. 2422–2427.

[315] S. A. WEISSMAN AND N. G. ANDERSON, *Design of Experiments (DoE) and Process Optimization. A Review of Recent Publications*, Organic Process Research and Development, 19 (2015), pp. 1605–1633.

[316] M. WEN, B. B. BOND-WATTS, AND M. C. CHANG, *Production of advanced biofuels in engineered E. coli*, Current Opinion in Chemical Biology, 17 (2013), pp. 472–479.

[317] A. WENINGER, J. E. FISCHER, H. RASCHMANOVÁ, C. KNIELY, T. VOGL, AND A. GLIEDER, *Expanding the CRISPR/Cas9 toolkit for Pichia pastoris with efficient donor integration and alternative resistance markers*, Journal of Cellular Biochemistry, 119 (2018), pp. 3183–3198.

[318] A. WENINGER, A.-M. HATZL, C. SCHMID, T. VOGL, AND A. GLIEDER, *Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast Pichia pastoris*, Journal of Biotechnology, 235 (2016), pp. 139–149.

[319] M. W. T. WERTEN, T. J. VAN DEN BOSCH, R. D. WIND, H. MOOIBROEK, AND F. A. DE WOLF, *High-yield secretion of recombinant gelatins by Pichia pastoris*, Yeast, 15 (1999), pp. 1087–1096.

[320] L. WESTERS, H. WESTERS, AND W. J. QUAX, *Bacillus subtilis as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism*, Biochimica et Biophysica Acta - Molecular Cell Research, 1694 (2004), pp. 299–310.

[321] K. J. WESTIN AND Å. C. RASMUSON, *Crystal growth of aragonite and calcite in presence of citric acid, DTPA, EDTA and pyromellitic acid*, Journal of Colloid and Interface Science, 282 (2005), pp. 359–369.

[322] J. H. WOO, Y. Y. LIU, A. MATHIAS, S. STAVROU, Z. WANG, J. THOMPSON, AND D. M. NEVILLE, *Gene optimization is necessary to express a bivalent anti-human anti-T cell immunotoxin in Pichia pastoris*, Protein Expression and Purification, 25 (2002), pp. 270–282.

[323] S. WU AND G. J. LETCHWORTH, *High efficiency transformation by electroporation of Pichia pastoris pretreated with lithium acetate and dithiothreitol.*, BioTechniques, 36 (2004), pp. 152–154.

[324] A. S. XIONG, Q. H. YAO, R. H. PENG, P. L. HAN, Z. M. CHENG, AND Y. LI, *High level expression of a recombinant acid phytase gene in Pichia pastoris*, Journal of Applied Microbiology, 98 (2005), pp. 418–428.

[325] A.-S. XIONG, Q.-H. YAO, R.-H. PENG, Z. ZHANG, F. XU, J.-G. LIU, P.-L. HAN, AND J.-M. CHEN, *High level expression of a synthetic gene encoding Peniophora lycii phytase in methylotrophic yeast Pichia pastoris*, Applied Microbiology and Biotechnology, 72 (2006), pp. 1039–1047.

[326] R. XIONG, J. CHEN, AND J. CHEN, *Secreted expression of human lysozyme in the yeast Pichia pastoris under the direction of the signal peptide from human serum albumin*, Biotechnology and Applied Biochemistry, 51 (2008), pp. 129–134.

[327] P. XU, A. VANSIRI, N. BHAN, AND M. A. KOFFAS, *ePathBrick: A synthetic biology platform for engineering metabolic pathways in E. coli*, ACS Synthetic Biology, 1 (2012), pp. 256–266.

[328] Y. XUAN, X. ZHOU, W. ZHANG, X. ZHANG, Z. SONG, AND Y. ZHANG, *An upstream activation sequence controls the expression of AOX1 gene in Pichia pastoris*, FEMS Yeast Research, 9 (2009), pp. 1271–1282.

[329] Z. YANG AND Z. ZHANG, *Engineering strategies for enhanced production of protein and bio-products in Pichia pastoris: A review*, Biotechnology Advances, 36 (2017), pp. 182–195.

[330] T. YARIMIZU, M. NAKAMURA, H. HOSHIDA, AND R. AKADA, *Synthetic signal sequences that enable efficient secretory protein production in the yeast Kluyveromyces marxianus.*, Microbial Cell Factories, 14 (2015), pp. 1–14.

[331] A. YASHCHENOK, B. PARAKHONSKIY, S. DONATAN, D. KOHLER, A. SKIRTACH, AND H. MÖHWALD, *Polyelectrolyte multilayer microcapsules templated on spherical, elliptical and square calcium carbonate particles*, Journal of Materials Chemistry B, 1 (2013), pp. 1223–1228.

[332] C. L. YOUNG, Z. T. BRITTON, AND A. S. ROBINSON, *Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications*, Biotechnology Journal, 7 (2012), pp. 620–634.

[333] J. YU, M. LEI, B. CHENG, AND X. ZHAO, *Facile preparation of calcium carbonate particles with unusual morphologies by precipitation reaction*, Journal of Crystal Growth, 261 (2004), pp. 566–570.

[334] R. J. ZAHRL, D. A. PEÑA, D. MATTANOVICH, AND B. GASSER, *Systems biotechnology for protein production in Pichia pastoris*, FEMS Yeast Research, 17 (2017), pp. 1–15.

[335] F. ZANELLA, J. B. LORENS, AND W. LINK, *High content screening: seeing is believing*, Trends in Biotechnology, 28 (2010), pp. 237–245.

[336] G. ZHANG, M. HUBALEWSKA, AND Z. IGNATOVA, *Transient ribosomal attenuation coordinates protein synthesis and co-translational folding*, Nature Structural and Molecular Biology, 16 (2009), pp. 274–280.

[337] J. ZHANG, R. E. CAMPBELL, A. Y. TING, AND R. Y. TSIEN, *Creating new fluorescent probes for cell biology.*, Nature Reviews Molecular Cell Biology, 3 (2002), pp. 906–918.

[338] J. ZHANG, T. CHUNG, AND K. OLDENBURG, *A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays*, Journal of Biomolecular Screening, 4 (1999), pp. 67–73.

[339] Y. ZHANG, U. WERLING, AND W. EDELMANN, *SLiCE: A novel bacterial cell extract-based DNA cloning method*, Nucleic Acids Research, 40 (2012), p. e55.

[340] D. M. ZHAO, M. WANG, X. J. MU, M. L. SUN, AND X. Y. WANG, *Screening, cloning and overexpression of Aspergillus niger phytase (phyA) in Pichia pastoris with favourable characteristics*, Letters in Applied Microbiology, 45 (2007), pp. 522–528.

[341] J. ZHOU, P. ZHU, X. HU, H. LU, AND Y. YU, *Improved secretory expression of lignocellulolytic enzymes in Kluyveromyces marxianus by promoter and signal sequence engineering*, Biotechnology for Biofuels, 11 (2018), pp. 1–14.

[342] C. ZHU, S. NAQVI, J. BREITENBACH, G. SANDMANN, P. CHRISTOU, AND T. CAPELL, *Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize*, Proceedings of the National Academy of Sciences, 105 (2008), pp. 18232–18237.

[343] S. B. ZIMMERMAN, J. W. LITTLE, C. K. OSHINSKY, AND M. GELLERT, *Enzymatic joining of DNA strands: a novel reaction of diphosphopyridine nucleotide*, Proceedings of the National Academy of Sciences of the United States of America, 57 (1967), pp. 1841–1848.

Curriculum Vitae