VILNIUS UNIVERSITY
LIFE SCIENCES CENTER

AUŠRA KONDRATAITĖ
Molecular Biotechnology study programme

Master Thesis

DEVELOPMENT OF HETEROLOGOUS GENE EXPRESSION SYSTEM IN YEASTS FOR THE PRODUCTION OF BIOCATALYSTS

Supervisor Inga Matijošytė, PhD

_____
(signature)
Consultant Rimantas Šiekštelė
Student _____
(signature)

Vilnius, 2022

TABLE OF CONTENTS

ABBREVIATIONS

EDTA – ethylenediaminetetraacetic acid

EPG – endopolygalacturonase

GRAVY – grand average of hydropathy

HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

IB – inclusion bodies

Lacc – laccase

LB medium – Luria-Bertani medium

LIC – ligation-independent cloning

MBP – maltose-binding protein

MCO – multicopper oxidase

OE-PCR – overlap extension polymerase chain reaction

Ori – origin of replication

Sp – signal sequence

TAE – tris-acetate-EDTA

Tm – primer melting temperature

TPI – triosephosphate isomerase

Tris – tris(hydroxymethyl)aminomethane, 2-amino-2-(hydroxymethyl)-1,3-propanediol

YNB – yeast nitrogen base

β-CA – β-carbonic anhydrase

INTRODUCTION

Biocatalysis refers to the use of enzymes or whole cells to perform various types of chemical reactions. In comparison with chemical reactions, biocatalytic processes are considered to be a more environment-friendly alternative. In addition, biocatalysis ensures stereo-, chemo-, regioselectivity and high effectiveness as well as mild operational conditions; therefore, there is a growing demand to replace a conventional chemical synthesis with biocatalysis. Since the application in industrial processes require a large amounts of biocatalysts, an effective gene expression system is considered a key aspect in protein production. It consists of biological environment, a vector, and an expression cassette containing gene that codes a desired biocatalyst. The development of expression system is often a time-consuming process and requires a lot of resources, including enzymes, chemical reagents, laboratory equipment, *etc*. Therefore, it is important to search for ways to facilitate the process of developing and optimizing expression system for various biocatalysts.

The purpose of this study was to develop a convenient and effective system for expressing proteins of interest using *Kluyveromyces marxianus* as a host. The research strategy was based on the construction of expression cassettes, that could be easily adapted for the production of any enzyme. The constructs also comprised fusion protein genes fused to target genes in order to avoid protein insolubility – a very common limitation during protein overexpression. Two biocatalysts – laccase (Lacc) (*Bacillus pumilus*) and β-carbonic anhydrase (β-CA) (*Bacillus mojavensis*) – were chosen to serve as target proteins, that have never been expressed in the yeast of *K. marxianus* strain before.

Aim of the study: To develop a heterologous gene expression system in yeasts for the production of Lacc and β-CA.

Objectives:

- To assemble expression cassettes for Lacc and β-CA production;
- To clone the obtained expression cassettes into the cloning vector pCAST;
- To assemble expression vector KLEF containing Lacc and β-CA expression cassettes;
- To perform a transformation of *K. marxianus* with KLEF vectors containing Lacc expression cassettes.

1. LITERATURE REVIEW

1.1 Biocatalysis

Biocatalysis is considered as the use of isolated enzymes or whole cells to perform chemical transformations of organic compounds. Scientific and technological advances have made an impact on evolution of biocatalysis, a practical and environmentally friendly alternative to traditional metallo- and organo-catalysis in chemical synthesis, both in the laboratory and on an industrial scale (Bornscheuer et al., 2012). The process of biocatalysis is currently an extremely important part of industrial manufacturing due to several valuable discoveries and insights in the past (Heckmann & Paradisi, 2020).

More than a century ago, during the first breakthrough wave of biocatalysis, scientists recognized that apart from the fermentation, components of living cells can be exploited in useful chemical conversions. For example, hydroxylation of steroids in microbial cells, the use of proteases in laundry detergents or penicillin G acylase for making semisynthetic antibiotic. However, the problem of limited stability of biocatalysts was still relevant, which was overcome by immobilization of the enzymes (Bornscheuer et al., 2012).

Later, around the 1980s, gene technology was explored allowing cloning and performing expression of enzymes in a suitable microbial host. Most importantly, researchers were able to subject enzymes to site-directed mutagenesis in order to enhance operational properties. Based on that, several companies could market useful enzymes immobilized on carriers on a large scale enabling a broad range of industrialized processes. Examples include the synthesis of intermediates for herbicides catalyzed by hydroxy-nitrile lyase, the synthesis of enantiopure alcohols for cholesterol-lowering statin drugs catalyzed by carbonyl reductase, *etc*. Besides the stabilization, further tasks involved optimization of biocatalysts with their catalytic activity for the non-natural substrates (Bornscheuer, 2018).

The third wave of biocatalysis started with discoveries by Pim Stemmer and Frances Arnold in the mid and late 1990s. They developed molecular biology methods that rapidly and extensively alter biocatalysts via an *in vitro* version of Darwinian evolution. The methods such as DNA shuffling and error-prone polymerase chain reaction (EP-PCR) together with high-throughput screening techniques, now are known as directed evolution. Since then, biocatalysis has profited remarkably from developments in molecular biology, rapid and cheap genome sequencing, and especially bioinformatics as well as computer modelling. Nowadays, engineering a native enzyme may take only a few months instead of years (Bornscheuer, 2018).

For the past few years, biocatalysis has been considered as a key process for synthesizing chemical compounds. This was determined by several significant advantages of biocatalyts (Bell et al., 2021). One of the most important features of biocatalysts is its selectivity (stereo-, regio- and

chemo-) enabling synthesis of highly specific compounds without the use of protecting groups or undesirable side reactions (Dev et al., 2018). In contrast to organometallic catalysts, developed using limited metals, enzymes are renewable, meaning that resources are not exhausted during the process. Moreover, enzymatic reactions proceed under mild conditions with high catalytic activity. The latter feature together with the selectivity can be improved through protein engineering (Truppo, 2017; Pyser et al., 2021).

Nowadays, enzymes are used for various applications in agriculture, food, polymer synthesis, organic synthesis of fine chemicals, and particularly in the production of pharmaceutical intermediates (**Table 1.1**). Besides the use of single isolated enzymes, the combination of several enzymes in cascade reactions or in microbial strains subjected to metabolic engineering leads to the synthesis of more complex compounds. Thus, enzymes were applied in simple chemical synthesis routes as well as used for improving chemical process economics (Chapman et al., 2018).

Table 1.1 Examples of industrial applications of enzymes (Chapman et al., 2018)

Sector	Enzymes	Applications
Pharmaceuticals	Nitrile hydratase, monoamine oxidase, transaminase, penicillin acylase, lipase	Production of intermediates for synthesis of active pharmaceuticals
Food processing	Glucose isomerase, trypsin, amylase, papain, pectinase	Production of high fructose corn syrup; conversion of starch to glucose; synthesis of prebiotics; debittering of fruit juice
Detergent	Lipase, amylase, protease, cellulase	Removal of fats/oils and stains; colour retention
Biofuel	Lipase, xylanase, cellulase	Transformation of lignocellulose for the production of bioethanol; production of fatty acid methyl esters
Paper/pulp	Lipase, xylanase, cellulase	Lignin removal for enhanced bleaching; improvement in fiber properties

Table 1.1 overviews the variety of fields that enzymes have been employed in to produce industrially important compounds (bioethanol, glucose, or pharmaceutical intermediates) as well as to solve relevant problems. Detergent industry, in particular, utilizes enzymes for breakdown of fat-, starch- and protein-based stains which is considered a key application area of biocatalysts. Among biological detergents, lipases have been used the most frequently and are involved in most industries (Singh et al., 2016).

1.2 Protein expression systems and their development

In nature proteins are expressed in low quantities sufficient for the organism. However, due to the needs from various industries, the production of proteins must be performed in sufficient amounts. The main goal of R&D in biocatalysis is to obtain pure enzymes in a short time and in high quantities. Another area of improvement is the low cost of production (Lestari & Novientri, 2021). All of these requirements for the final product can be achieved by developing an effective recombinant protein

expression system. Recombinant proteins are encoded by recombinant DNA and expressed in a system enabling the transcription and translation of that DNA. Therefore, by using molecular methods to genetically engineer DNA from various organisms, it is possible to obtain the required amounts of target protein under cost-effective conditions (Dev et al., 2018).

Typically, a protein expression system consists of three main parts operating as one unit. Host or a biological environment is necessary for providing energy and the machinery to synthesize a protein. A cell is not obligatory in this process as there is an alternative cell-free expression system comprising ions (buffer) and all the required macromolecules for translation. The second part is a vector, which enables the introduction of genetic material into the host cell. Typically, vectors contain regulatory parts for the replication and selection markers for maintenance (**Fig. 1.1 a**). The third part of the system is an expression cassette inserted into the vector containing the open reading frame that encodes the target protein. Expression cassette also provides other necessary components for transcription and translation, such as promoter, terminator, ribosome-binding site, start/stop codons (**Fig. 1.1 b**) (Fisher et al., 2016).

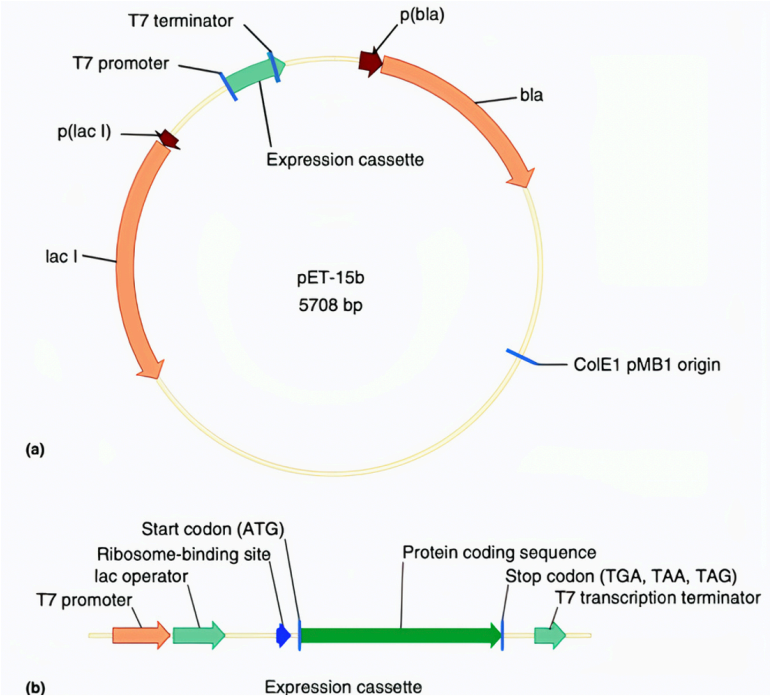


Figure 1.1 The example of expression vector and expression cassette. (a) pET-15b expression vector comprising *lacI* gene (encodes repressor protein controlling the expression from the T7 promoter), *bla* gene (encodes β-lactamase for ampicillin resistance), ColE1 pMB1 origin of replication (*ori,* a site of the start of replication) and expression cassette. (b) Expression cassette containing T7 promoter (a site of transcription initiation), lac operator (prevents transcription from T7 promoter, when lac repressor is bound to it), ribosome-binding site, protein coding sequence and T7 terminator (a site of the end of transcription) (Fisher et al., 2016)

A variety of organisms serve as hosts for the production of recombinant proteins. It includes bacteria, yeasts, filamentous fungi, insect cells, protozoa, and mammalian cells. Although, all of these organisms are applied depending on the characteristics of the target protein, bacteria and yeasts are the most common. Bacterial expression systems are preferred for their simple genetic manipulation,

rapid growth and the ability to grow in low cost and chemically defined media (Fisher et al., 2016). *Escherichia coli* is the most widely used host microorganism in biotechnology. Due to the complete characterization of its genome and a broad knowledge about its biochemistry, *E. coli* has been used as a model organism for over 60 years (Fisher et al., 2016; Lozano Terol et al., 2021). Besides the mentioned advantages, this bacterium also is associated with several problems, such as protein aggregation, metabolic burden or inefficient translocation/transport system of expressed proteins (Lozano Terol et al., 2021). In addition, *E. coli* does not proceed a lot of post-translational modifications that, on the contrary, are the property of eukaryotes. Among them, yeasts combine the simplicity of a unicellular organism, its relatively low nutritional requirements, the mentioned post-translational modifications. *Saccharomyces cerevisiae*, *Klyyveromyces lactis*, *Komagataella phaffii* and *Yarrowia lipolytica* are several widely used yeast species with the *S. cerevisiae* being the main choice in for recombinant protein production. Various biopharmaceutical proteins, produced by *S. cerevisiae,* have been marketed (Huang et al., 2014; Jozala et al., 2016).

One of the most common barriers in the development of protein expression system is protein insolubility or inclusion body (IB) formation. Proteins fail at folding properly resulting in aggregation due to high expression rates (Gutiérrez-González et al., 2019). However, to overcome this obstacle, several strategies were developed, such as lowering cultivation temperature or selecting inducer concentration, regulating relative codon abundance or fusing target protein to a more soluble protein (Zhu et al., 2013). The latter method is considered as a highly promising approach, yet it is still in a trial-and-error state. The use of fusion tag enhances the expression level and/or solubility of proteins fused to it. They are usually highly expressed in their native organisms and help to achieve natural folding resulting in a proper functionality and high expression of a protein (Ki & Pack, 2020). However, fusion tag, expressed with a target protein, can potentially affect the structure and biological function of that protein. In this case, removal of a fusion tag can be performed by inserting a protease recognition site between the target protein and the fused solubility tag. After downstream processing the final product would contain only the protein of interest (Yadav et al., 2016). Hence, by using fusion tags, the possibility of protein aggregation can be reduced significantly.

There are various solubility-enhancing tags available: maltose-binding protein (MBP), N-Utilization substance A (NusA), glutathione-S-transferase (GST), thioredoxin (Trx) and some other. The grand average of hydropathy (GRAVY) value is the parameter used to measure the hydrophobicity (positive values) or hydrophilicity (negative values) of the protein (Kyte & Doolittle, 1982). **Table 1.2** summarizes GRAVY values and molecular weights of some solubility enhancers.

Table 1.2 Examples of protein fusion tags to enhance heterologous protein expression (Ki & Pack, 2020)

Tag	Full name	Size, kDa	GRAVY
MBP	Maltose-binding protein	40	– 0.340
GST	Glutathione-S-transferase	26	– 0.369
NusA	N-Utilization substance A	55	– 0.278
SUMO	Small ubiquitin-related modifier	11	– 0.978
Trx	Thioredoxin	12	– 0.004
HE-MBP	Truncated maltotriose-binding protein with modified histidine tag	45	– 0.336
sfGFP	Superfolder green fluorescent protein	27	– 0.534
CBD	Cellulose binding domain	11	– 0.013
PDI	Disulfide isomerase I	55	– 0.496

Most of them have negative GRAVY values and one of the most frequently used is MBP. It functions as a molecular chaperone leading to a correct folding of the fused protein. The mechanism behind this could be based on interaction with hydrophobic amino acid residues located in unfolded proteins to avoid aggregation or proteolysis (Sachdev & Chirgwin, 2000; Needle & Waugh, 2014).

1.3 Molecular cloning methods

Molecular cloning refers to the isolation of DNA from any species and the subsequent insertion of that DNA into a vector to proceed the multiplication. As a result, this methodology generates population of organisms carrying the same unaltered molecule of recombinant DNA. The process starts with an *in vitro* construction of vector containing isolated DNA of interest, which is then transferred to a host, capable of replicating the inserted target gene. A non-pathogenic and easy-to-grow laboratory bacterial strain of *E. coli* is the most common choice for molecular cloning (Bertero et al., 2017).

The history of molecular cloning started with the discovery of bacterial enzymes, commonly known as “restriction endonucleases” (Bertero et al., 2017). Researchers have found that methylation of phage DNA by host methyltransferases protected DNA from digestion by host restriction nucleases. On the contrary, foreign DNA molecules, without host methylation pattern or the ones that were unmethylated, were recognized as foreign and were degraded by host restriction enzymes (Luria & Human, 1952; Bertani & Weigle, 1953). Later Meselson and Yuan characterized the first rectriction nucleases (Meselson & Yuan, 1968). This led to the subsequent discovery of restriction endonuclease ability to recognize and cleave DNA at specific sites, which is currently a widely applied property of restriction enzymes in genetic engineering (Kelly & Smith, 1970). Ligases were another significant enzymes isolated and characterized for the first time in 1967. The discovery of the latter enzymes enabled the creation of recombinant DNA by merging two separate DNA molecules (Cozzarelli et al., 1967). Moreover, molecular cloning has been improved remarkably after the invention of several methods, such as transformation, PCR, Sanger sequencing, as well as gene synthesis.

An accumulated knowledge about the molecular cloning itself led to the discovery of a number of strategies for developing a recombinant vector. For years, restriction and ligation based method,

known as conventional cloning, seemed to be the most suitable choice. However, due to inconvenient use of restriction enzymes, several new strategies were created, including PCR cloning, seamless cloning, ligation-independent cloning (LIC) or recombinational (Gateway) cloning (Bertero et al., 2017; El Qaidi & Hardwidge, 2019).

Traditional cloning. This method involves the use of restriction enzymes and ligases as the fragment of interest is inserted into the vector by restriction and ligation of DNA (**Fig. 1.2**). First, the target sequence is amplified during the PCR adding restriction sites to the ends of the sequences.

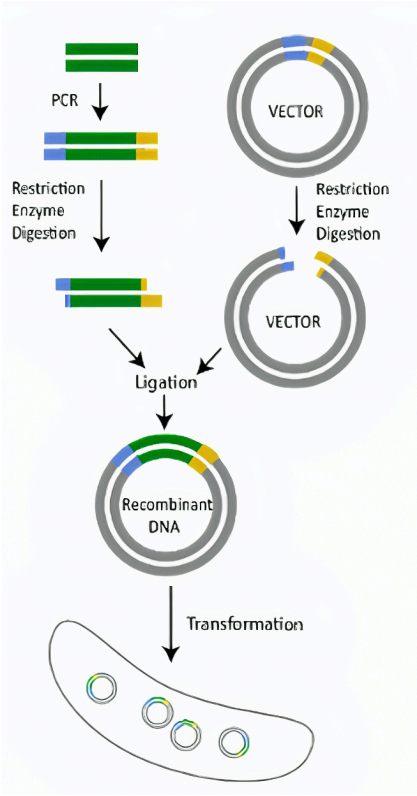


Fig. 1.2 The scheme of conventional cloning method[1]

Restriction enzymes can then generate "sticky ends" that contain single-stranded overhangs (either on the 3' or 5' ends) or "blunt ends" having no overhangs. Both of these types of ends can be consolidated by ligation enzymes. The blunt type of ends is considered to be more versatile, since it does not require the complementary Watson-Crick base pairing, which is a mandatory factor in terms of a sticky-end fragment. However, blunt-end ligation is less efficient as the binding stability is much weaker in comparison with the complementary overhangs. Also, "sticky ends" can be enzymatically affected to form "blunt ends" and *vice versa*. Such modifications of overhangs can be achieved by filling the gaps of single-stranded DNA or removing the overhangs, while new overhangs can be created by 3'-5' or 5'-3' exonucleases (Bertero et al., 2017).

PCR cloning. Direct ligation of a PCR-generated fragment without the use of restriction endonucleases is the basis of PCR cloning (**Fig. 1.3**) (Bertero et al., 2017). In this method vector is

[1] https://www.goldbio.com/articles/article/cloning-overview

prepared by developing a T-overhang at the 3'-end under the activity of deoxynucleotidyl-transferase. Such vectors are commonly called "T-vectors". Inserts are also modified by adding 3' adenosine overhangs using Taq DNA polymerase (Motohashi, 2019).

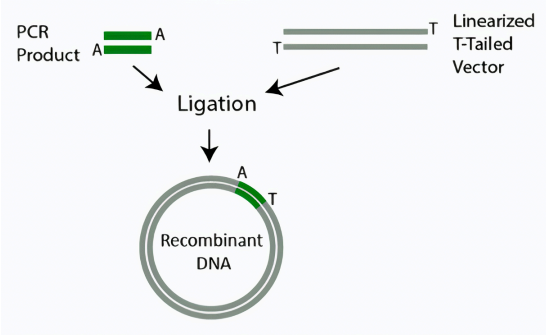


Fig. 1.3 The scheme of the TA cloning method[2]

This strategy is especially useful when the required restriction sites are not available. Despite the simplicity of the method, it has some disadvantages. One of them is the lack of polymerases that would provide both, proof-reading activity and an ability to create an A overhang. Also, the final recombinant vector could contain insert in sense or antisence orientation which complicates the analysis of the experimental results after cloning (Clark & Pazdernik, 2013; Carson et al., 2019).

Seamless cloning. As in PCR cloning, seamless cloning does not require the use of restriction enzymes and, additionally, allows a sequence-independent and scar-free insertion of several DNA fragments into a vector (**Fig. 1.4**).

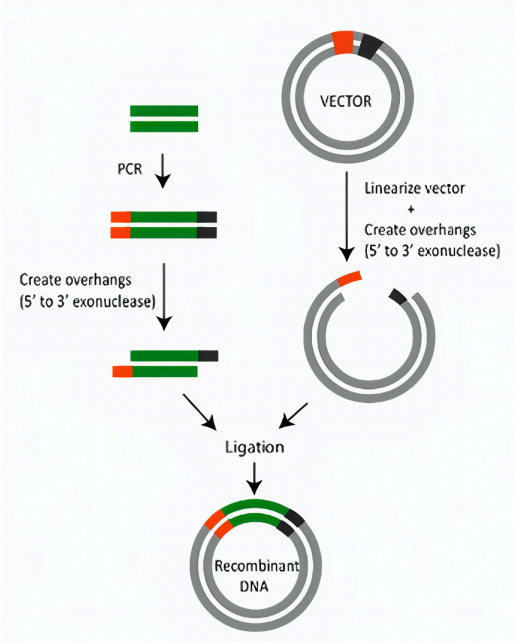


Fig. 1.4 The method of seamless cloning[3]

[2] https://www.goldbio.com/articles/article/cloning-overview
[3] https://www.goldbio.com/articles/article/cloning-overview

Such methodology can be advantageous when an insert of interest contains a number of restriction sites, which makes it difficult to find restriction enzymes that will not cut anywhere in a target gene during the cloning. The most common example of seamless cloning is the Gibson Assembly Method which is based on the addition of homologous regions at each end of the fragments that are to be cloned. The process includes combination of activities of an exonuclease, DNA polymerase and DNA ligase that allows integrating several inserts into the same vector (Bertero et al., 2017).

Ligation independent cloning (LIC). As the name of the method implies, it allows the cloning of the insert in the absence of DNA ligase. It is proceeded by adding short sequences of DNA to the insert, which will result in sequences at the ends of the DNA fragment that are complement to the ends of a destination vector (**Fig. 1.5**).

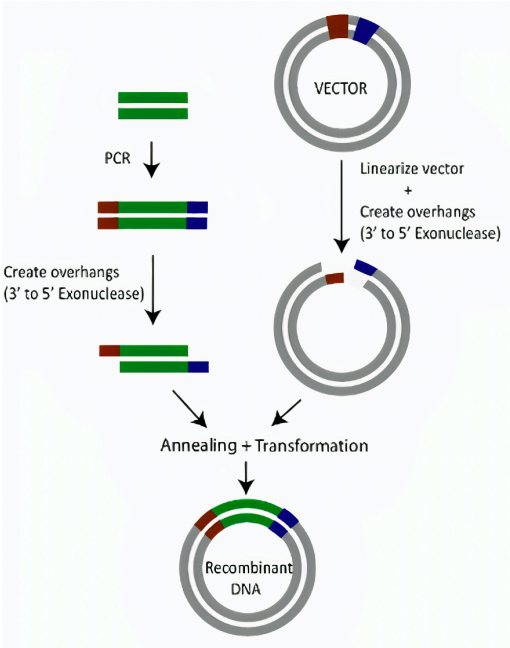


Fig. 1.5 The method of ligation independent cloning (LIC)[4]

Further step includes generation of complementary cohesive ends for both – insert and a vector, which is accomplished by using enzymes with 3'-5' exonuclease activity. The obtained DNA products are then combined to proceed the annealing and the subsequent repairing of the nicks on the recombinant DNA construct by the host organism. Importantly, the final product does not contain any unwanted sequences or restriction sites (Li & Elledge, 2007; Bertero et al., 2017).

Recombinational cloning. This method uses site-specific DNA recombinase enzymes that exchange and recombine sequences of DNA between two molecules containing particular recombination sites. First of all, the fragment of interest is prepared by adding appropriate recombination sites on both sides of the insert during PCR. Next, the obtained DNA product is recombined with a *donor vector* creating an *entry clone* which is finally recombined with a *destination vector* leading to a final

[4] https://www.goldbio.com/articles/article/cloning-overview

construct. The most used system based on recombinational cloming is Gateway cloning, which is widely commercialized to facilitate the application of this method (Walhout et al., 2000; Bertero et al., 2017).

A number of molecular cloning methods are available to this date, however, each of them has its own advantages and limitations (**Table 1.3**).

Table 1.3 Comparison of the main molecular cloning methods (Bertero et al., 2017)

Cloning method	Cost	Sequence dependency	Throughput	Assembly of multiple fragments	Directional Cloning	Examples of commercially available products
Traditional cloning	Low	Yes (restriction enzyme sites)	Low to mid	Difficult for more that two fragments	Possible	-
PCR cloning	Medium (vector)	No	High	Challenging (requires special modifications)	Difficult	TOPO TA
Ligation independent cloning	Medium (reagents)	Limited (vector)	Low	Yes	Yes	In-Fusion
Seamless cloning	High (reagents)	No	Low	Yes	Yes	Gibson Assembly GeneArt
Recombinatio nal cloning	High (reagents and vectors)	No	High	Challenging (requires special modifications)	Yes	Gateway Echo Cloning Creator

The choice of a method depends on the time, resources as well as the the target DNA, in particular. Considering the importance of a quick and efficient procedure that could be widely applied without high costs, homologous recombination seems to be the best alternative of all. Such homology-based method employ PCR products which are flanked on both sides by a relatively short sequences (from 15 bp to 60 bp in length) that match the ends of a linearized vector (Jacobus & Gross, 2015). Nevertheless, regardless of the technique used to generate recombinant DNA of interest, molecular cloning is unambigously a milestone of most research laboratories (Bertero et al., 2017).

1.4 Laccase and its expression systems in yeasts and bacteria

Since the discovery of laccases (Laccs) (EC 1.10.3.2) in 1883, a lot of research was conducted on the structure, function, mechanism of action as well as a variety of biotechnological applications of these enzymes. Laccs are found to be widely distributed across all kingdoms of life, yet they are mostly produced by fungi and some bacteria, insects, plants. Almost 150 Laccs have already been fully characterized with the fungal Laccs being the most studied. These biocatalysts belong to multicopper oxidases (MCOs) superfamily and catalyze oxidation reaction of various phenolic and non-phenolic compounds by reducing oxygen to water. Due to the presence of cupredoxin-like domains, all MCOs proceed the reduction of oxygen without producing any harmful byproducts (Janusz et al., 2020). Laccs contain three structural cupredoxin-like domains and four copper ions (T1-T3) located in two catalytic sites (**Fig. 1.6**). The mononuclear T1, the blue copper, serves as the primary electron acceptor site for oxidation of substrate, while the mononuclear T2 and binuclear T3

copper ions function as electron acceptors from T1, before the reduction of oxygen (Hakulinen & Rouvinen, 2015; Jones & Solomon, 2015; Lucas et al., 2017). Notwithstanding, Laccs are considered as an eco-friendly enzymes having a number of industrial applications including decolorization of dyes in textile industry, delignification of pulp paper, bioremediation, food processing, organic synthesis of various medications, biosensor technology, *etc* (Mayolo-Deloisa et al., 2020; Brugnari et al., 2021).

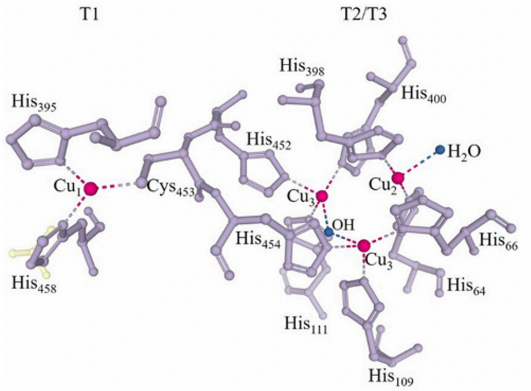


Fig. 1.6 Structure of the copper catalytic site of Lacc (*Trametes versicolor*). His – histidines, Cys – cysteine (Brugnari et al., 2021)

Considering an industrial potential of Laccs, more research on the production of this biocatalyst is requested. Expression of Laccs using native sources often can not meet the demand from the market. The abilities (growth rate, optimal growth temperature, *etc*) of many microorganisms do not satisfy the requirements of an industrial fermentation process. Therefore, efforts have been made to lower the cost of Lacc production by the use of recombinant organisms or screening for strains that naturally hypersecrete Lacc (Piscitelli et al., 2010; Yang et al., 2017). Laccs from various sources, such as plants, bacteria, fungi or actinobacteria, were expressed in a number of hosts, including *E. coli* BL21, *E. coli* DH5α, *Kluyveromyces lactis, Yarrowia lipolytica, Saccharomyces cerevisiae, Komagatarlla phaffii, etc* (Brugnari et al., 2021).

Recombinant Lacc production in bacteria

Researchers have been investigating recombinant Lacc production in non-fungi hosts for decades. To date, there are several different Laccs expressed in bacterial strains, mostly *E. coli*, which is often used due to their easy and cheap manipulation (Antošová & Sychrová, 2016).

Enzymatic activity of recombinant Laccs produced in bacteria ranges in 2–5600 U/L and Lacc protein amount of 10–350 mg/L. *E. coli* was the main choice for the heterologous expression of Laccs that originated from bacteria such as *Bacillus subtilis, Bacillus clausii, Streptomyces griseorubens, Klebsiella* sp., fungi of *Cyathus bulleri*, *Rigidoporus lignosus* (Antošová & Sychrová, 2016). Expression levels in *E. coli* were often relatively low and recombinant enzymes formed aggregates leading to a difficult purification process. To overcome this problem, the method for refolding Laccs

was developed. Overexpression of Lacc from *Bacillus* sp. HR03 followed by the solubilization of IB in an optimized refolding buffer, resulted in a significantly increase of Lacc activity (Mollania et al., 2013). A few years ago, Fang and its collegues implemented the modification of Lacc structure in order to enhance the enzyme activity. Deletion of His-tag and several enzyme residues, based on its crystal structure obtained in previous experiments, enabled Lacc expression in a completely soluble form even at higher temperatures (Fang et al., 2014). Another approach to improve Lacc production in bacteria was related to the use of oxygen-limited conditions. This way copper incorporation into the active center of enzyme was increased. An example of such conditions application would be the increased production of Laccs and Lacc-like multi-copper oxidases (LMCOs) from *Bacillus* sp. in *E. coli* after changing conditions from aerobic to microaerobic (Durão et al., 2008; Ihssen et al., 2015).

Recombinant Lacc production in the yeast

Yeasts are known for their rapid growth, easy handling of cells, facile genetic manipulation, an ability to perform eukaryote-specific post-translational modifications, namely proteolytic processing, glycosylation and disulfide bridge formation (Piscitelli et al., 2010; Antošová & Sychrová, 2016). Considering these advantages, yeasts were widely investigated in terms of heterologous protein expression. In literature there are a number of examples of recombinant Lacc expression in yeast hosts. However, only several of them revealed to be successful and beneficial.

Lacc production in yeasts varies in the range of 0.034 to 380000 U/L of enzymatic activity and Lacc protein amount of 2–130 mg/L (**Table 1.4**). A significantly high Lacc activity of 140000 U/L was reached after optimization of *K. phaffii* cultivation conditions in a fermentator. *lcc1* cDNA was isolated from *T. versicolor*. Important outcomes of this investigation, leading to an effective production of an enzyme, was the decrease of cultivation temperature and lowered methanol concentration (Hong et al., 2002). Even higher enzyme activity has been obtained in *Cryptococcus* sp. S-2, which ensured the expression of Lacc gene containing high guanine-cytosine (GC) content. *K. phaffii* was used in the same experiment in order to compare the expression level. In this study, the more beneficial expression system of *Cryptococcus* sp. S-2 exposed Lacc activity of 380000 U/L (Nishibori et al., 2013). Another yeast – *S. cerevisiae* – has also been extensively applied as an expression host for Lacc production. For example, expression of Lacc gene from *Trametes* sp. C30 resulted in 1200 U/L after cultivation in a fermentor (Klonowska et al., 2005). To increase the expression level of another Lacc originating from *Myceliophthora thermophila*, directed evolution was used to generate a more efficient variant. After error-prone PCR and the subsequent *in vivo* shuffling, the obtained libraries were transformed in to *S. cerevisiae* BJ5465 and screened for activity. Incubation in shake flasks resulted in Lacc protein concentration of 18 mg/mL (Bulter et al., 2003). Even though *K. phaffii* and *S. cerevisiae* have often been the first choice for recombinant Lacc expression, other hosts have shown potential as well (**Table 1.4**).

Table 1.4 Examples of Lacc expression systems in yeast hosts

Expression host	Source	Lacc	Promoter	Culture mode	Inducer	Max. Lacc activity (U/L)	Reference
Cryptococcus sp. S-2	*Gaeumannomyces graminis*	LAC2	Xylanase promoter	2 L fermentor	6 % xylose	380000	(Nishibori et al., 2013)
Kluyveromyces lactis	*Trametes trogii*	LCC1	*KlPDC1*	NR** fermentor	Induced in response to a decrease in oxygen concentration	192	(Ranieri et al., 2009)
Komagataella phaffii	*Thermus thermophilus*SG0.5JP17-16	LACTt	*AOX1*	10 L fermentor	1 % methanol	6130	(Liu et al., 2015)
	Yarrowia lipolytica	YlLAC	*AOX1*	Shake flasks (0.25 L)	0.5 % methanol	1290	(Kalyani et al., 2015)
	Trametes versicolor	LCC1	*AOX1*	2.5 L fermentor	0.5 % methanol	140000	(Hong et al., 2002)
Saccharomyces cerevisiae	*Cryphonectria parasitica*	LAC3	*GAL1*	Shake flasks (0.3 L)	2 % galactose	41.5	(Kim et al., 2010)
	Trametes sp. C30	LAC1,2,3	*PGK1*	3 L fermentor	No	1200	(Klonowska et al., 2005)
Yarrowia lipolytica	*Trametes versicolor* DSM11269	LCC1*	*TEF1*	Shake flask (NR**)	No	1000	(Theerachat et al., 2012)

*genetically engineered
** not reported

Although recombinant Lacc expression can provide relatively high levels of active enzymes, natural hosts may serve as a beneficial producents. Remarkably, there are cases where Lacc levels produced by native hosts exceed levels in recombinant yeasts or filamentous fungi. The reason for such achievements is related to the optimized cultivation parameters: cultivation temperature, optimal pH, carbon and nitrogen concentrations, also the amount of Lacc co-factor copper, addition of inducers, *etc*. As an example, fungus *T. versicolor* was able to produce 692000 U/L of Lacc in shake flasks with an optimized medium, containing additional copper as well as several aromatic inducers (2,5-xylidine, ferulic acid, *etc*) (Revankar & Lele, 2006).

1.5 β-Carbonic anhydrase and its expression systems in yeasts and bacteria

Carbonic anhydrases (CAs) (EC 4.2.1.1) belongs to the superfamily of metalloenzymes and are physiologically involved in multiple processes, such as respiration, photosynthesis, CO_2 and bicarbonate transport, pH and CO_2 homeostasis, *etc*. CAs are categorized into eight families, namely

α, β, γ, δ, ε, ζ, η, θ, ι, and are widely distributed among all kingdoms. In particular, β-CAs, have been identified in bateria, fungi, archaea, algae, and chloroplasts of both mono- and dicotyledons. Regardless of such variety of CA families, they are all responsible for the same reversible hydration of carbon dioxide (CO_2) resulting in bicarbonate (HCO_3^-) and a proton (H^+) (Nocentini & Supuran, 2019; Nocentini et al., 2021). Reaction mechanism of CAs consists of two main steps (**Fig. 1.7**). Firstly, the nucleophilic attack occurs between the Zn^{2+}-bound hydroxide ion to a CO_2 forming the enzyme-HCO_3 adduct. Such compound is then removed from the active site by a water molecule. During the second step, a kinetically rate limiting one, the catalytically active Zn^{2+}-bound hydroxide ion is regenerated through a proton transfer reaction from the Zn^{2+}-bound water to an exogenous proton (Angeli et al., 2020).

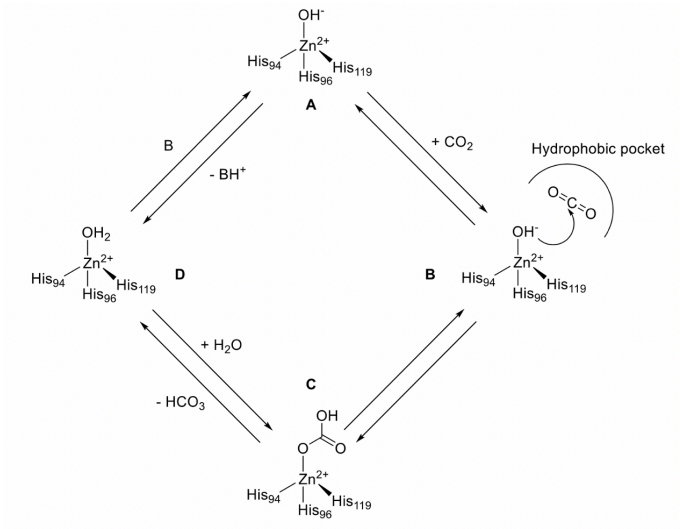


Fig. 1.7 Catalytic mechanism of CAs' reversible hydration of carbon dioxide (Angeli et al., 2020)

Recently, CAs ability to capture CO_2 has been highly investigated mainly due to the critically increased accumulation of the mentioned gases in the atmosphere. Physiologically essential molecule of CO_2 is one of the principal products in combustion reactions and is being produced in large quantities. Therefore, CAs have shown a potential to accelerate CO_2 capture from large combustion emitters (Chu, 2009). In addition to CO_2 capture and sequestration, CAs have been reported to participate in several other hydrolytic reactions, such as the hydrolysis of esters, thioesters, other molecules, such as cyanamide. Also the hydrolysis reaction of these enzymes have been employed to produce hydrogen sulfide (H_2S). Until now, only α-CAs and a few β-CAs were known to have other catalytic functions besides the CO_2 hydration (Angeli et al., 2020).

Recombinant β-CA production in bacteria

To date, a variety of CAs from different families have been cloned and expressed in a prokaryotic microorganism *E. coli*. A convenient system of *E. coli* BL21(DE3) strain is probably the most widely used host for high-level expression of recombinant proteins. It contains a prophage DE3

from a bacteriophage λ that carries T7 polymerase gene controlled by the *lac*UV5 promoter (Jeong et al., 2015).

In literature, *E. coli* BL21 strain was recognized as the main choice for β-CAs expression as well. A few years ago a successful expression of β-CA from *Enterobacter* sp. B13 was described. A newly isolated strain showed CO_2 hydratase activity indicating the presence of CA gene, which was later amplified, cloned, and expressed in *E. coli* BL21 (DE3)*pLysS*. The obtained concentration of the purified enzyme was 2,4 mg/mL (Eminoğlu et al., 2016). β-CA and γ-CA production using the same *E. coli* BL21 (DE3) strain is another successful example. After screening of *Bacillus* sp. SS105 for CA activity, it was selected for further investigation. The detected CA genes, homologous to β-CA and γ-CA, were used in pET30b (+) expression vector construction. Both recombinant proteins, containing His-tag at the N-terminus, were purified by Ni-NTA affinity chromatography. The evaluated specific activity of β-CA and γ-CA enzymes were 1449.1 U/mg and 1067.9 U/mg, respectively. The obtained biocatalysts were later used for biomineralization of CO_2 into highly relevant mineral of calcite (Maheshwari et al., 2019). A chemolithotrophic strain, identified as *Serratia marcescens* Wy064, also revealed the presence of four CA genes, named *CA1–CA4*. Two of them – CA1 and CA3 – were recognized as β-CAs after structure modelling combined with sequencing and both were expressed without formation of IB in *E. coli* BL21 (DE3) (Chen et al., 2019).

Recombinant β-CA production in yeasts

Even though in many cases *E. coli* has been and still is the most reliable and effective expression host, yeasts are widely implemented in recombinant protein production as well. Several yeast species, such as *S. cerevisiae*, *K. phaffii*, *K. lactis*, *Y. lipolytica*, were used for the production of a number of recombinant proteins, originating from both eukaryotes and prokaryotes (Vieira Gomes et al., 2018). Since *E. coli* is not recognized as GRAS (generally recognized as safe) organism, it is crucial to continue the development of yeasts expression systems for other prokaryotic enzymes, including β-CA.

A non-methylotrophic yeast *S. cerevisiae,* which is considered to be the first and a very well-known yeast expression host, was used for the expression of β-CA gene from *E. coli* a few decades ago. In this investigation, *S. cerevisiae* was genetically modified by insertional mutagenesis to knockout the native β-CA gene. The obtained mutant strain of *S. cerevisiae* displayed an oxygen-sensitive growth-defect phenotype as well as increased sensitivity to oxidative stress. Yeasts, transformed with a β-CA expression vector, were able to grow under aerobic conditions indicating a successful expression of β-CA (Cronk et al., 2001). Another application of *S. cerevisiae* for the expression of prokaryotic enzyme was shown with one of the most valuable enzymes, that is now widely used in molecular biology procedures – chloramphenicol acetyltransferase. Originating from

a prokaryote *E. coli* Tn9, its gene was introduced into yeast cells as a part of yeast/*E. coli* shuttle vector and produced in high quantities (Cohen et al., 1983). Yeasts were not considered the first choice for the production of β-CA, since there are not many examples of such application in literature. However, several other yeast species were developed and implemented in recombinant protein production, including prokaryotic ones. A reliable and versatile expression system of *Y. lipolytica* have been investigated and applied for academic and commercial uses. Due to its low overglycosylation, high secretory efficiency and product yield, performance reproducibility, *Y. lipolytica* can be an effective choice for the expression of a number of biocatalysts (Coelho et al., 2010). In addition, another yeast – *K. lactis* – is known for its ability to achieve high levels of protein secretion, which makes it an attractive gene expression host alternative (Spohner et al., 2016). Evidences of yeasts application to express recombinant prokaryotic enzymes and their properties reveal the potential for effective bacterial β-CA expression as well.

1.6 Insights into the research project

The relevance of biocatalysis in a context of ecology as well as high demand of enzymes in industry lead to the need of a cost-effective production of biocatalysts (DiCosimo et al., 2013). Development of a gene expression system is one of the main steps in obtaining any protein of interest. The preparation of such system for expression of every new protein is often a time-consuming process and requires reagents and/or materials resulting in relatively high final cost. The use of restriction enzymes, ligases or the addition of special DNA sequences for cloning can be avoided when using homologous recombination. The mentioned method is based on homologous sequences between the insert and the vector and does not require any DNA preparation steps (Jacobus & Gross, 2015). Therefore, the best choice would be a convenient and versatile system to assemble an expression vector, that could effectively functionize in an expression host. Another important aspect of the enzyme production is the insolubility of proteins. Aggregation of proteins is a very common barrier limiting the expression (Schramm et al., 2020). One of the ways to avoid protein insolubility is to fuse target proteins with highly soluble proteins in order to enhance their folding during the expression (Kwon et al., 2021).

Lacc and β-CA in particular, are both industrially important enzymes, that have limited application due to high production cost. Therefore, the development of expression system for these enzymes is an important field of science worth focusing on. Recombinant Lacc and β-CA enzymes are known to be expressed in various heterologous bacterial and yeast expression systems. However, none of them have ever been produced in *K. marxianus* yeasts, which are the fastest growing eukaryotes having GRAS status and a wide spectrum of possible carbon sources (Mo et al., 2019; Karim et al., 2020).

2. MATERIALS AND METHODS

2.1 Devices and tools

Table 2.1 The list of equipment used during the investigation

Analytical balance	Kern ABJ
Balance	Kern E6
Centrifuges	Eppendorf 54 (rotor r = 84 mm) Labnet Prism R (rotor r = 84 mm) Labnet Prism Mini
Horizontal electrophoresis system	BioRad
Spectrophotometer	Nanodrop 2000; Spekol 2000
Thermocycler	BioRad iCycler
Thermostat	Binder
Vortexes	Vibramax 100
Lab shakers	Stuart Orbital Incubator, Biosan
pH meter	Mettler Toledo
Transilluminator	Biostep
Biological safety cabinet	Flow Fast H
Laboratory water bath	Thermo Haake P5
Magnetic stirrer	Biosan MMS-3000
Electroporator	BioRad

2.2 Materials and solutions

2.2.1 Microorganisms and vectors

Microorganisms: *Escherichia coli* DH5α *recA, endA*, *Kluyveromyces marxianus* BKM Y-719 *ura3 pec1.*

Vectors: cloning vector pCAST, expression vector KLEF.

2.2.2 Primers

Table 2.2 Primers and their DNA sequences (Metabion)

P619	5’-ATGAACATGGTGATGACGATGACAAGAACCTAGAAAAATTTGTTGACG-3’
P513	5’-ACGCGTGTCGAGAATCAAAC-3’
P663	5’- ATGAACATGGTGATGACGATGACAAGAGTCTTCTCAATGATATTCTTGAC-3’
P579	5’-CCTTTGTTTTTCTGTTCCCCTTACGCGTGTCGAGAATCAAACAAAACTG-3’
P620	5’-TCATCACCATGTTCATGTTCATGTTCATGCTTGGTGATACGAGTCTGC-3’
P621	5’-ACCATTGGAAAAGAGAAAAATCGAAGAAGGTAAACTGG-3’
P622	5’-CCTTCTTCGATTTTTCTCTTTTCCAATGGTGATGC-3’
P520	5’-CTAGAGGATCCCCCGACGCGTAATCATCATGGAGCTGGATG-3’
P625	5’-GTTGCTGAATAACATGTTTGATAGATTATGTGTATGTGTGTATG-3’
P623	5’-CACATAATCTATCAAACATGTTATTCAGCAACACCTTATTG-3’
P624	5’-TCATCACCATGTTCATGTTCATGTTCATGCTTGCTACCTGGGGTACC-3’
P627	5’-CACCATTGGAAAAGAGAGCTAGAACATTCTTTATTGGTGG-3’
P626	5’-ATAAAGAATGTTCTAGCTCTCTTTTCCAATGGTGATGC-3’
P628	5’-TCATCACCATGTTCATGTTCATGTTCATGGACTCTGGAGTTGATGATGTC-3’
P539	5’- CACGAGGCCCTTTCGTCTCGAATGCATCTAGATAACTTGG-3’
P540	5’- AGCTGTGACCGTCTCCGGGAGGATCCGATTGTATACAATG-3’

2.2.3 Reagents

"Acros Organics": agar, acetic acid (97 %), ammonium sulfate (99 %), D-sorbitol (97 %), calcium chloride (96 %).

"BioRad": tris(hydroxymethyl)aminomethane (Tris) (≥ 99.8 %).

"Fisher Bioreagents": LB medium (10 g/L peptone from casein, 10 g/L NaCl and 5 g/L yeast extract) molecular genetics grade.

"Merck & Co., Inc.": yeast extract, hydrogen chloride, magnesium chloride hexahydrate (≥ 95 %), agarose, ampicillin (sodium salt) (≥ 90 %), D(+)gliucose monohydrate, sodium chloride (≥ 99.5 %),

"Roth": etidium bromide.

"Sigma-Aldrich": PEG-3350, lithium acetate dihydrate (≥ 95 %, reagent grade).

"Thermo Fisher Scientific": 6×DNA Gel Loading Dye, 1 Kb Plus DNA Ladder, GeneRuler 100 bp DNA Ladder, carrier DNA TopVision Agarose, DTT, ethylenediaminetetraacetic acid (EDTA) (≥ 99 %), DMSO (100 %), Yeast Nitrogen Base (YNB) (without amino acids and ammonium sulfate), peptone, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (≥ 99 %).

2.2.4 Kits and enzymes

Kits (Thermo Fisher Scientific): "GeneJET Plasmid Miniprep Kit", "GeneJET PCR Purification Kit", "GeneJET Gel Extraction Kit".

PCR Master Mixes (Thermo Fisher Scientific): "Phusion Hot Start II High-Fidelity PCR Master Mix (2×)", "DreamTaq Green PCR Master Mix (2×)".

Restriction endonucleases (10 U/μL) (Thermo Fisher Scientific): *Eco*88I, *Nde*I, *Bam*HI, *Alw*44I, *Aat*II, *Sma*I, *Sal*I, *Pdi*I, *Eco*105I.

Reaction buffers for restriction endonucleases (Thermo Fisher Scientific): 10×Buffer O, 10×Buffer BamHI, 10×Buffer Tango.

2.2.5 Cultivation media for *E. coli*

LB (Luria-Bertani) liquid medium: 2.5 g of LB is dissolved in 100 mL of distilled H_2O and autoclaved. If the medium is used for the selection of transformants, 100 μg/ml of ampicillin wass added.

LB agar medium: 2.5 g of LB and 2 g of agar are dissolved in 100 mL of distilled H_2O and autoclaved. If the medium is used for the selection of transformants, 100 μg/ml of ampicillin was added.

100 mg/mL ampicillin solution: 0.1 g of ampicillin is dissolved in 1 mL of distilled H_2O.

2.2.6 Cultivation media for *K. marxianus*

YEPD medium: 2 g of glucose, 1 g of yeast extract, and 2 g of peptone are dissolved in 100 mL of distilled H_2O and autoclaved.

YEPD agar medium: 2 g of glucose, 1 g of yeast extract, 2 g of peptone, and 2 g of agar are dissolved in 100 mL of distilled H_2O and autoclaved.

YNGKR medium: 0.17 g of YNB (without amino acids and ammonium sulfate), 0.5 g of ammonium sulfate, 2 g of glucose, and 2 g of agar are dissolved in 100 mL of distilled H_2O and autoclaved.

40 mM HEPES/YEPD medium: 96 mL of YEPD medium is supplemented with 4 mL of 1 M HEPES.

1 M Sorbitol/YEPD medium: 60 mL of YEPD medium is supplemented with 18.22 g of sorbitol and YEPD until the final volume of 100 mL. The medium is autoclaved before use.

2.2.7 Solutions for preparation of competent *E. coli* cells and transformation

NaCl solution: 0.06 g of Tris, 0.1 g of $MgCl_2 \times 6H_2O$, and 0.58 g of NaCl are dissolved in 80 mL of distilled H_2O. pH of the solution is set to pH 8.0 by adding HCl. Volume of solution is brought to 100 mL and autoclaved.

$CaCl_2$ solution: 0.06 g of Tris, 0.1 g of $MgCl_2 \times 6H_2O$, and 1.1 g of $CaCl_2$ are dissolved in 80 mL of distilled H_2O. pH of the solution is set to pH 8.0 by adding HCl. Volume of solution is brought to 100 mL and autoclaved.

2.2.8 Solutions for preparation of competent *K. marxianus* cells and chemical transformation

10×TE buffer (pH 7.5): 1.21 g of Tris and 0.29 g of EDTA are dissolved in 80 mL of distilled H_2O. pH of the solution is set to pH 7.5 by adding HCl. Volume of solution is brought to 100 mL and autoclaved.

10× (1 M) lithium acetate solution (pH 7.5): 10.2 g of lithium acetate is dissolved in 80 mL of TE. pH of the solution is set to pH 7.5 by adding HCl. Volume of solution is brought to 100 mL and autoclaved.

50 % (w/v) PEG solution: 50 g of PEG-3350 is dissolved in 80 mL of TE buffer. Volume of solution is brought to 100 mL and sterilization using 0.22 μm filter is performed.

40 % PEG solution: 80 mL of filter-sterilized 50 % PEG-3350 solution is mixed with 10 mL of 10×lithium acetate solution and 10 mL of 10×TE buffer.

2.2.9 Solutions for preparation of electrocompetent *K. marxianus* cells and electroporation

Electroporation buffer (pH 8.0): 0.03 g of Tris and 0.02 g of $MgCl_2$ are dissolved in 80 mL of distilled H_2O. pH of the solution is set to pH 8.0 by adding HCl. Volume of the solution is brought to 100 mL and autoclaved.

1 M HEPES solution (pH 8.0): 23.8 g of HEPES is dissolved in 60 mL of distilled H_2O. pH of the solution is set to pH 8.0 by adding NaOH. Volume of the solution is brought to 100 mL and sterilized using 0.22 μm filter.

1 M DTT solution: 15.4 g of DTT is dissolved in 80 mL distilled H_2O. Volume of solution is brought to 100 mL and sterilized using 0.22 μm filter.

2.2.10 Solutions for electrophoresis

0.5 M EDTA (pH 8.0): 14.6 g of EDTA is dissolved in 80 mL of distilled H_2O. pH of the solution is set to pH 8.0. Volume of solution is brought to 100 mL and autoclaved.

10×TAE buffer: 48.5 g of Tris, 11.4 mL of glacial acetic acid and 20 mL of 0.5 M EDTA (pH 8.0) are dissolved in 1 L of distilled H_2O. 1×TAE buffer is used for electrophoresis.

1 mg/mL etidium bromide solution: 0.1 g of etidium bromide is dissolved in 100 mL of distilled H_2O.

2.3 Methods

2.3.1 DNA electrophoresis

DNA electrophoresis is performed in 1 % agarose gel (the percentage of agarose in a gel depends on the size of DNA fragments). 1 g of agarose (2.2.3) is mixed with 100 mL of 1×TAE (2.2.10) buffer and heated in a microwave until the agarose is completely dissolved (the solution should not be overboiled to avoid the alteration of the final percentage of agarose in a gel). The solution is left to cool down for about 5 min and supplemented with etidium bromide (2.2.10) to a final concentration of 0.5 µg/mL. The obtained solution is poured to a gel casting tray. Once solidified, agarose gel is placed in a horizontal electrophoresis unit, which is then filled with 1×TAE buffer (2.2.10) until the gel is completely covered. Samples are mixed with 6× loading dye (2.2.3) in a ratio 5:1 and loaded to wells of the gel. Molecular weight marker of 5 µL volume is loaded to at least one well for the determination DNA sizes. Duration of electrophoresis is about 25 min and the voltage is 7 V/cm. The results are analyzed using UV transilluminator.

Preparative electrophoresis is used for extraction of the DNA fragments from gel. 1 % agarose gel containing 0.5 µg/mL of etidium bromide (2.2.10) is prepared in a smaller horizontal electrophoresis chamber, which is filled with fresh 1×TAE buffer (2.2.10) to avoid any contamination. The voltage in preparative electrophoresis is 7 V/cm.

2.3.2 Purification of plasmid DNA from cells and DNA after enzymatic reactions

For purification of plasmid DNA from bacteria, a commercial kit "GeneJET Plasmid Miniprep Kit" (2.2.4) was used[5]. All cells are collected by centrifugation at 12000×g. The pelleted cells from 1-10 mL of culture are suspended in 250 µL of "Resuspension Solution" by vortexing and pipetting up and down. 250 µL of "Lysis Solution" is added to the suspension by carefully inverting the tube several times (the incubation can not be longer that 5 min). The obtained mixture is supplemented with "Neutralization Solution" and immediately mixed by inverting the tube a few times. Tubes with lysed cells are centrifuged for 5 min at 12000×g. After centrifugation the supernatant, containing plasmid DNA, is transferred to the "GeneJET" spin column without disturbing the precipitate. The column is centrifuged for 1 min at 12000×g and the flow-through is discarded. 500 µL of "Wash Solution" is added to the same column, which is subsequently centrifuged for 1 min at 12000×g and the flow-through is also discarded. The latter wash procedure is repeated one more time. Residual "Wash Solution" is removed by centrifuging the column for an additional 1 min. After that the column

[5] https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2FMAN0012655_GeneJET_Plasmid_Miniprep_UG.pdf

is transferred into a new 1.5 mL tube. It is then incubated at 50 °C with a column left open to ensure a more effective elimination of ethanol from "Wash Solution". 50 µL of "Elution Buffer" is added to the center of the column membrane avoiding any contact to the membrane. After 2 min incubation at room temperature the tube with a column is centrifuged for 2 min at 12000×g. The eluted solution, containing plasmid DNA, is stored at -20 °C.

"GeneJET Plasmid Miniprep Kit" was also used for purification of DNA after enzymatic reactions, such as PCR or DNA restriction. The reaction mixture is supplemented with "Neutralization Solution" in a ratio of 10:7, respectively. The obtained solution (up to 800 µL) is transferred to a "GeneJET" spin column and centrifuged for 2 min at 12000×g. After the discarge of flow-through, the following steps (starting with the addition of "Wash Solution") are the same as during plasmid purification from cells.

2.3.3 Extraction of DNA from agarose gel

In order to purify DNA from agarose gel, a commercial kit "GeneJET Gel Extraction Kit" (2.2.4) was used[6]. DNA fragment of interest is excised from a gel after preparative electrophoresis. A slice of gel is placed in a pre-weighed 1.5 mL tube. The calculated weight of a slice is mixed with a "Binding Buffer" in a ratio of 1:1 (e.g., 100 mg of gel requires 100 µL of "Binding Buffer"). A gel mixture is incubated at 50-60 °C for 10 min depending on the size of a gel slice. After gel is completely dissolved, the color of the solution is determined – if the solution is yellow, pH is suitable for the binding reaction. In case of orange or violet solution, 3 M of sodium acetate (pH 5.2) must be added. Up to 800 µL of mixture is transferred to a "GeneJET" purification column and centrifuged for 1 min at 12000×g. The flow-through is discarded. The column is filled with 700 µL of "Wash Buffer" and centrifuged for 1 min at 12000×g. After centrifugation, the flow-through is discarded again and empty purification column is centrifuged for 1 min at 12000×g to remove any residual "Wash Buffer". The column is placed in a new 1.5 mL tube and incubated at 50 °C with column left open to completely eliminate the remainings of "Wash Buffer". 50 µL of "Elution Buffer" is added to the center of the column membrane without a contact with the membrane. The tube with a column is centrifuged for 1 min at 12000×g. The obtained solution, containing DNA, is stored at -20 °C.

2.3.4 Measurement of DNA concentration

DNA concentration is measured using spectrophotometer „Nanodrop" (Thermo Fisher Scientific) according to the manufacturer's instructions[7].

[6] https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2FMAN0012661_GeneJET_Gel_Extraction_UG.pdf

[7] https://assets.thermofisher.com/TFS-Assets/CAD/manuals/NanoDrop-2000-User-Manual-EN.pdf

2.3.5 Polymerase chain reaction (PCR)

PCR was used to amplify DNA fragments. Reaction conditions were selected according to the instructions provided by polymerase's manufacturer[8]. The reaction is performed in a thermocycler, which is programmed depending on the fragment to be amplified (**Table 2.4**). The composition of reaction is given below in **Table 2.3**. Total reaction volume – 20 µL.

Table 2.3 Composition of the PCR reaction mixture including Phusion polymerase

Reagent	Volume, µL
MiliQ H_2O	7
Phusion Hot Start II High-Fidelity PCR Master Mix (2×) (2.2.4)	10
Forward primer	1
Reverse primer	1
DNA*	1

*Optimal DNA concentration is 1 pg-10 ng per reaction volume 50 µL

Table 2.4 Thermocycler program for PCR using Phusion polymerase

Cycle step	Temperature	Time	Cycles
Initial denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	30
Annealing	**X**	30 s	
Extension	72 °C	**Y**	
Final extension	72 °C	5 min	1
Hold	20 °C	∞	∞

X – primer annealing temperature is calculated using manufacturer's calculator[9]
Y – sequence extension time depends on the speed of the polymerase (Phusion – 15-30 s/kb) and the length of the amplified sequence

2.3.6 Overlap extension PCR

The method of overlap extension PCR (OE-PCR) was used to assemble gene expression cassettes. Firstly, individual DNA fragments are amplified during conventional PCR (2.3.5). Primers used in this step are synthesized to amplify DNA fragments that would have overlapping sequences with each other. Generated fragments are purified separating them from salts, proteins and other impurities (2.3.2). Next step involves the fusion of DNA fragments in order to assemble full-length DNA cassette. The reaction mixture contains all fragments that need to be fused, as they prime each other during the PCR reaction were performed using primers, that bind at the ends of the DNA cassette.

2.3.7 Preparation of competent *E. coli* cells

5 mL of sterilized LB medium (2.2.5) is inoculated with *E. coli* DH5α and incubated in a shaking incubator at 37 °C and 200 rpm for about 16-18 h. 50 µL of overnight culture is added to 5 mL of fresh LB medium (2.2.5) and incubated for 1.5-2 h at 37 °C and 200 rpm until the optical density (OD) at 600 nm reaches 0.4-0.5. The culture is cooled down on ice keeping it for 10-15 min.

[8] https://www.thermofisher.com/lt/en/home/brands/thermo-scientific/molecular-biology/thermo-scientific-restriction-modifying-enzymes/restriction-enzymes-thermo-scientific/conventional-restriction-enzymes-thermo-scientific/reaction-conditions-for-restriction-enzymes.html
[9] https://www.thermofisher.com/lt/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html

E. coli cells are separated into several sterile 1.5 mL tubes and centrifuged at 4 °C and 6000×g for 30 s until all cells are pelleted. Supernatant is discarded and the pellet is resuspended with sterile ice-cold NaCl solution (2.2.7) and centrifuged at 4 °C and 6000×g for 30 s. After supernatant is discarded, cells are resuspended with ice-cold $CaCl_2$ solution (2.2.7) and centrifuged at 4 °C and 6000×g for 30 s. The obtained supernatant is discarded and the pelleted cells are resuspended in a $CaCl_2$ solution (2.2.7). The suspension is left to incubate on ice for about 1 hour. After incubation another centrifugation is performed at 4 °C and 6000×g for 30 s. The supernatant is discarded and cells are mixed with 50 µL of $CaCl_2$ solution (2.2.7).

2.3.8 Transformation of *E. coli* cells

The prepared competent *E. coli* cells are mixed with 5-10 µL of DNA and the mixture is incubated on ice for 1 h. After the incubation, tubes containing cells are placed in a water bath at 42 °C for 45 s. After the heat shock, cells are returned on ice and incubated for 2 min. Regeneration of cells is proceeded by suspending cells with a 1 mL of fresh LB medium (2.2.5). The suspension is incubated in a shaking incubator at 37 °C and 200 rpm for 1 h. Afterwards, cells are centrifuged (6000×g), the supernatant is discarded and cells are spreaded on the LB agar plates (2.2.5) containing 100 µg/ml of ampicillin (2.2.5). The plates are incubated at 37 °C for 16-18 h.

2.3.9 Preparation of competent *K. marxianus* cells and chemical transformation

5 mL of sterilized YEPD medium is inoculated with *K. marxianus* cells from a freshly prepared *K. marxianus* culture plate. Inoculate is incubated in a shaking incubator at 30 °C and 200 rpm for 18 h. 1 mL of overnight culture is added to a fresh 10 mL of YEPD medium (2.2.6) and left to grow for 3 h at 30 °C and 200 rpm. After incubation yeast cells are pelleted by centrifugation at 17530×g for 20 s. Supernatant is discarded leaving several microliters for suspension. 50-100 µg of single-stranded carrier DNA (2.2.3) and about 100 ng of plasmid DNA. The mixture of cells and DNA is gently vortexed and is supplemented with 0.5 mL of 40 % PEG solution (2.2.8) and DMSO (2.2.3) to the final concentration of 10 %. Cells are incubated in a shaker at room temperature for 15 min. Next, tubes containing cells are placed in a water bath for a heat shock at 42 °C for 15 min. After that, cells are centrifuged for 20 s at 17530×g. The cell pellet is washed once with 1×TE buffer (2.2.8), supernatant is discarded. Cells are spreaded on the YNGKR plates (2.2.6), which are incubated at 30 °C for at least 48 h.

2.3.10 Preparation of competent *K. marxianus* cells and electroporation

5 mL of sterilized YEPD medium is inoculated with *K. marxianus* cells and incubated at 30 °C and 200 rpm for 18 h. Cells are transferred to 1.5 mL tubes and centrifuged at 4 °C and 3000×g for 3 min. Supernatant is discarded and cells are suspended in 1 mL of 40 mM HEPES/YPD medium (2.2.6) and DTT (2.2.9) to a final concentration of 10 mM. The mixture is gently vortexed and incubated at 30 °C for 5 min without shaking. After that, cells are washed once with cold

electroporation buffer (2.2.9). The obtained supernatant is discarded and cell pellet is suspended in 50 µL of electroporation buffer. Cells are incubated on ice together with the electroporation cuvettes for 5-10 min. Cells are supplemented with plasmid DNA and transferred to a cold cuvette without introducing any bubbles and assuring all cells reached the bottom of the cuvette. Cuvette is placed in an electroporator for a high-voltage electric shock (2500 V/cm, 20 ms, square wave). Immediately after the shock cells are suspended in a cold 1 M sorbitol/YEPD medium (2.2.6) and left for 1 h regeneration at 30 °C. After that cells are centrifuged and plated on YNGKR plates (2.2.6) that are incubated at 30 °C for 48 h.

2.3.11 Colony PCR

Colony PCR is applied for screening colonies after genetic transformation of cells. It verifies the presence of the desired DNA construct. Biomass from each colony is put into the reaction mixture, which is prepared using DreamTaq Green PCR Master Mix (2×) according to the table below (**Table 2.5**). Reaction conditions are selected according to the DNA template of interest (**Table 2.6**).

Table 2.5 Composition of the PCR reaction mixture including *Taq* polymerase

Reagent	Volume, µL
MiliQ H_2O	7
DreamTaq Green PCR Master Mix (2×) (2.2.4)	10
Forward primer	1
Reverse primer	1
DNA*	1

*DNA of the colonies from the plate

Table 2.6 Thermocycler program for PCR using *Taq* polymerase

Cycle step	Temperature	Time	Cycles
Initial denaturation	95 °C	1 min	1
Denaturation	95 °C	30 s	30
Annealing	**X**	30 s	
Extension	72 °C	**Y**	
Final extension	72 °C	5 min	1
Hold	20 °C	∞	∞

X – primer annealing temperature is calculated using manufacturer's calculator[10]
Y – sequence extension time depends on the speed of the polymerase (*Taq* – 1 min/kb) and the length of the amplified sequence

2.3.12 Restriction analysis

Restriction analysis was performed in order to verify the obtained DNA construct. Restriction endonucleases (2.2.4) were used to digest the DNA and the reaction mixture (10 µL) consisted of the following components: 1 µL of 10×reaction buffer (depending on the restriction endonuclease) (2.2.4), 1 µL of restriction endonuclease (10 U/µL) (2.2.4), 0.5-1.0 µL of DNA, MiliQ H_2O to the final reaction volume (10 µL). Reaction conditions were selected according to the instructions provided by the manufacturers [11].

[10] https://www.thermofisher.com/lt/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html
[11] https://www.thermofisher.com/lt/en/home/brands/thermo-scientific/molecular-biology/thermo-scientific-restriction-modifying-enzymes/restriction-enzymes-thermo-scientific/conventional-restriction-enzymes-thermo-scientific/reaction-conditions-for-restriction-enzymes.html

3. RESULTS AND DISCUSSION

The goal of this study was to develop an improved heterologous expression system, which could be applied for production of various biocatalysts in *Kluyveromyces marxianus* yeasts. Laccase (Lacc) and β-carbonic anhydrase (β-CA) genes, that served as target genes during the investigation, were isolated from two bacteria – *B. pumilus* and *B. mojavensis,* respectively, by laboratory colleagues. Lacc and β-CA have never been expressed in *K. marxianus*, which was selected as an expression host for the study. This yeast strain is considered as a promising microorganism to be applied in industry due to its rapid growth, thermotolerance, GRAS status, as well as capability to use various carbon sources besides glucose, such as xylose, inulin, lactose (Mo et al., 2019).

To enhance heterologous expression of biocatalysts, three genes such as a coding maltose-binding protein (MBP), endopolygalacturonase (EPG) and triosephosphate isomerase (TPI) were used as fusion partners-chaperones. During protein expression, such molecular chaperones lead to an accurate folding of target protein preventing it from aggregation (Ki & Pack, 2020). The use of fusion tags is one of effective strategies to increase the solubility of proteins. MBP is already a very well-known fusion protein with a number of successful applications. One of the most remarkable examples has shown, that fusing MBP to a human oncostatin M (OSM) increased soluble protein production by 79 % at 37 °C in *E. coli* BL21(DE3). Lowering the induction temperature in this case increased the solubility up to 91 % (Nguyen et al., 2019). Another heterologous expression of MBP-tagged human vascular endothelial growth factor (VEGF) resulted in a remarkably increased solubility. Induction at 18 °C in *E. coli* cells led to 92.8 % of soluble targeted protein (Nguyen et al., 2016). The other two proteins – EPG and TPI – were selected for this investigation considering their high expression and secretion in *K. marxianus* yeasts. Strain of *K. marxianus* BKM Y-719 is known to produce effectively the pectin-degrading enzyme EPG, that has been successfully expressed in other yeast strains as *Kluyveromyces lactis*, *Saccharomyces cerevisiae,* and *K. marxianus* (Bartkevičiūtė et al., 2000; Šiekštelė et al., 1999). The glycolysis enzyme TPI is expressed in *K. marxianus* under a strong promoter, which has been used in expression of heterologous bacterial lipase from *Serratia* sp. in yeasts of *K. lactis* (Šiekštelė et al., 2015). Due to an efficient expression and secretion of *K. marxianus* enzymes EPG and TPI, these proteins were selected for the construction of expression cassettes for Lacc and β-CA production.

First of all, all three fusion proteins were evaluated for their hydropathy values by applying Kyte and Doolittle algorithm, taken from an Expasy ProtScale online tool[12] (Kyte & Doolittle, 1982) (**Fig. 3.1**).

[12] https://web.expasy.org/protscale/

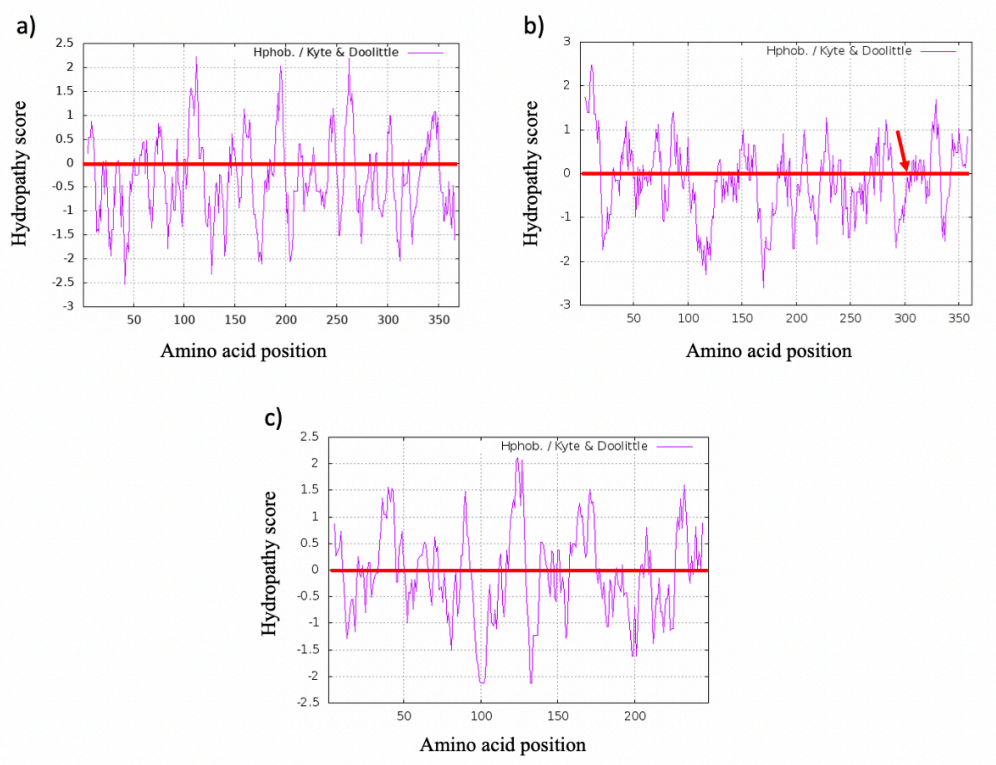


Fig. 3.1 Graphs showing the results from hydropathy score evaluation of the three chaperone proteins: **a)** MBP (GRAVY value – -0.34); **b)** EPG (GRAVY value – -0.18); **c)** TPI (GRAVY value – -0.08). Red arrow separates a part of the amino acid sequence, that was eliminated (final length of the sequence was 302 amino acids, GRAVY value – -0.252). The GRAVY value is calculated by the sum of hydropathy values of all amino acids divided by the length of protein[13]

Amino acid sequence of MBP showed very high hydrophilicity (highly negative hydropathy scores), which is a key characteristic of chaperone proteins (Ki & Pack, 2020). However, EPG and TPI had higher GRAVY values indicating about lower hydrophilicity of proteins. Therefore, EPG amino acid sequence was shortened by removing amino acids, that are highly hydrophobic. The modified variant of EPG, having a more negative GRAVY value, was used in further study. TPI sequence was not changed, since in this case, theoretical removal of hydrophobic amino acids from the sequence resulted only inconsiderable change of GRAVY value. The goal of altering the sequence of highly expressed and secreted proteins according to their hydropathy scores, was to obtain more hydrophilic variants. The use of such modified variant as fusion protein could possibly result in an improved solubility of target protein as well.

3.1 Construction of gene expression cassettes for Lacc and β-CA

The study was started with construction of the gene expression cassettes for both Lacc and β-CA enzymes. Each of the genes were fused to the different phusion partners – MBP, EPG, and TPI. In total, six DNA cassettes were constructed. The DNA constructs for Lacc were composed of a strong TPI promoter and terminator, EPG signal sequence (for extracellular protein secretion),

[13] http://www.gravy-calculator.de

chaperone genes (either MBP, EPG or TPI), 7x(HE) purification tag, and target *lacc* gene. By analogy, the expression cassette for β-CA was constructed, only replacing the target gene (**Fig 3.2**).

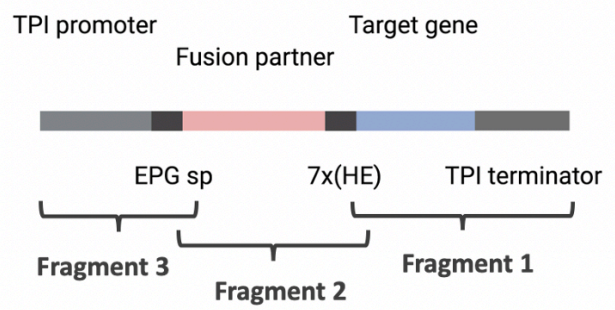


Fig. 3.2 General scheme of the expression cassette for this investigation. Target gene indicates either Lacc or β-CA gene, whereas fusion partner – maltose-binsing protein (MBP), endopolygalacturonase (EPG) or triosephosphate isomerase (TPI) genes. Sp – signal sequence, 7x(HE) – modified polyhistidine tag. Cassette is separated to *Fragment 1*, *Fragment 2*, and *Fragment 3* – the initial sequences, that were fused to each other to result in final expression cassette

In this study, the method of overlap extension PCR (OE-PCR) (2.3.6) was used for the assembly of the DNA constructs. This strategy is based on overlapping sequences between the fragments that need to be fused. Complementary sequences are added to the ends of fragments by conventional PCR using hybrid primers. Thus, the resulting fragments contain short sequences, that can further act as priming sites for DNA polymerase during overlap extension PCR. The main advantage of this method is the ability to avoid the insertion of additional nucleotides into the sequence to create targets for restriction endonucleases (Nelson & Fitch, 2012). Considering the convenience and versatility of OE-PCR, it has been applied not only in combining individual fragments, but also directly to mutate genes, that are being fused. In such experiments of site-directed mutagenesis, initial PCR is used to embed overlapping segments containing nucleotide insertions, substitutions or deletions to the ends of the fragments. During the subsequent PCR the obtained intermediate products hybridize at the complementary region. The final full-length product is generated by amplification with flanking primers (Heckman & Pease, 2007).

For the construction of DNA cassettes, three fragments (*Fragment 1, Fragment 2, and Fragment 3*) needed to be amplified (**Fig. 3.2**) (2.3.5). The amplification was done using hybrid primers, which insert short sequences to the ends of the amplified fragments, that are complementary to sequences to be fused with. All fragments were amplified without any non-specific products (**Fig. 3.3**). The obtained PCR products included:

- *Fragment 1*: Lacc (primers – P619, P579), CA (primers – P513, P663).
- *Fragment 2*: MBP (primers – P620, P621), EPG (primers – P623, P624), TPI (primers – P627, P628).
- *Fragment 3*: TPIprM (primers – P520, P622), TPIprE (primers – P520, P625), TPIprT (primers – P520, P626).

The obtained PCR products were purified to eliminate enzymes, primers, salts, and other unnecessary components left from the amplification reaction (2.3.2).

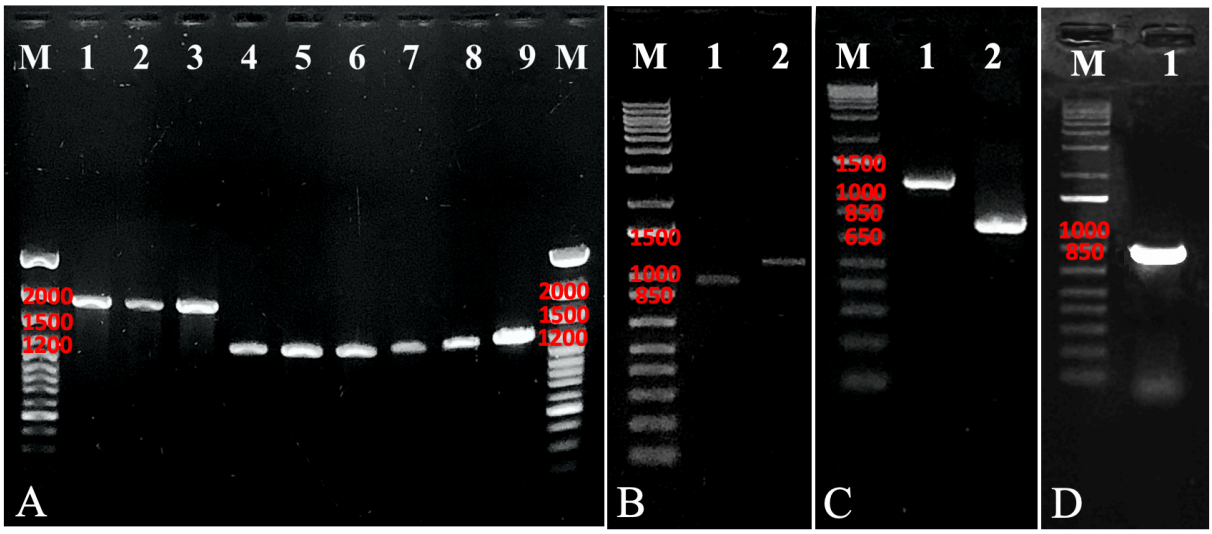


Fig. 3.3 Amplified fragments for the construction of **Lacc** and **β-CA** expression cassettes:
(A) Lane M – GeneRuler 100 bp DNA Ladder, **lanes 1-3** – Lacc (1850 bp), **lanes 4-6** – MBP (1155 bp), **lanes 7-9** – TPIprM (1198 bp),
(B) Lane M – 1 Kb Plus DNA Ladder, **lane 1** – EPG (952 bp), **lane 2** – TPIprE (1124 bp),
(C) Lane M – 1 Kb Plus DNA Ladder, **lane 1** – TPIprT (1201 bp) **lane 2** – TPI (787 bp),
(D) Lane M – 1 Kb Plus DNA Ladder, **lane 1** – CA (859 bp)

To obtain full gene expression cassettes, a single PCR reaction containing all three fragments as templates was performed (2.3.6). Such variant resulted in very low quantities or no DNA constructs of expected length and showed a lot of non-specific products. In fact, the assembly of large DNA constructs from three or more individual fragments in one reaction is possible and often is very effective in terms of the need for optimization and time. However, the efficiency of such one-step fusion reaction to assemble several DNA fragments might depend on the length of homologous sequences between fragments, that needed to be fused. There are examples, where at least 50 bp of sequence homology is required in order to perform a successful fusion of multiple fragments (Kadkhodaei et al., 2016). Whereas the overlapping sequences between *Fragment 1* and *Fragment 2*, as well as *Fragment 2* and *Fragment 3* in construction of Lacc and β-CA cassettes were shorter – 16 bp and 30 bp, respectively. Therefore, another attempts to assemble the cassettes consisted of two separate fusion steps: i) fusion to obtain intermediate constructions (that resulted in fusion of two fragments); ii) fusion to obtain full-length DNA cassette. Regarding construction of Lacc expression cassettes, six intermediate DNA constructs were obtained:

- TPIprM-**MBP** (primers – P520, P620) and **MBP**-Lacc (primers – P579, P621),
- TPIprE-**EPG** (primers – P520, P624) and **EPG**-Lacc (primers – P579, P623),
- TPIprT-**TPI** (primers – P520, P628) and **TPI**-Lacc (primers – P579, P627).

All of the intermediate constructs had one *Fragment 2* (MBP, EPG, or TPI), that was common for both intermediates. Fusion reactions resulted in additional non-specific amplicons. Therefore,

intermediate DNA constructs for Lacc expression cassettes were purified from agarose gel to avoid the presence of any unwanted DNA sequences (2.3.3) (**Fig. 3.4**).

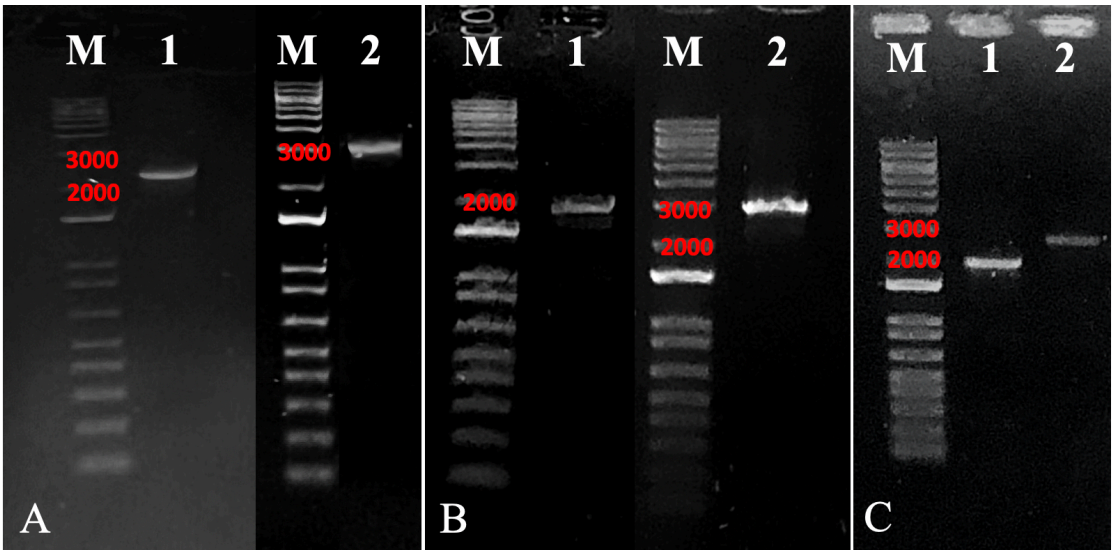


Fig. 3.4 Intermediate DNA constructs of **Lacc** expression cassettes after purification from agarose gel **(A) Lane M** – 1 Kb Plus DNA Ladder, **lane 1** – TPIprM-**MBP** (2323 bp), **lane 2** – **MBP**-Lacc (2989 bp), **(B) Lane M** – 1 Kb Plus DNA Ladder, **lane 1** – TPIprE-**EPG** (2044 bp), **lane 2** – **EPG**-Lacc (2786 bp), **(C) Lane M** – 1 Kb Plus DNA Ladder, **lane 1** – TPIprT-**TPI** (1939 bp), **lane 2** – **TPI**-Lacc (2621 bp).

Usually, the first cycles of OE-PCR are performed without primers – the overlapping parts of fragments serve as priming sites instead. Primers are subsequently used in PCR for amplifying the whole full-length construct (Nelson & Fitch, 2012). However, during the construction of intermediate constructs for Lacc cassette, primers were included in the reaction mixture. Such addition can be the reason for the production of non-specific amplicons during the OE-PCR. The result of the reaction in this case included target product – intermediate construct – and several non-specific fragments, that were possibly amplified by the added primers. Overall, this experiment resulted in higher amounts of the desired DNA construct despite the nonspecificity, which was successfully eliminated by purification from agarose gel (2.2.3).

For the assembly of β-CA expression cassettes, only CA (*Fragment 1*) needed to be fused to MBP, EPG and TPI (*Fragment 2*) (2.3.6). Such simplified construction of intermediate constructs for β-CA cassettes proved the versatility of the method. This fusion step also resulted in a mixture of target construct together with a few nonspecific products, that were later separated *via* agarose gel purification (2.3.3). The obtained intermediate DNA constructs for β-CA expression cassettes are presented in **Fig. 3.5**. Since expression cassettes for β-CA and Lacc differ only by the target gene (in *Fragment 1*), intermediate constructs TPIprM-**MBP**, TPIprE-**EPG**, and TPIprT-**TPI** were suitable to apply further for the construction of full-length β-CA cassettes.

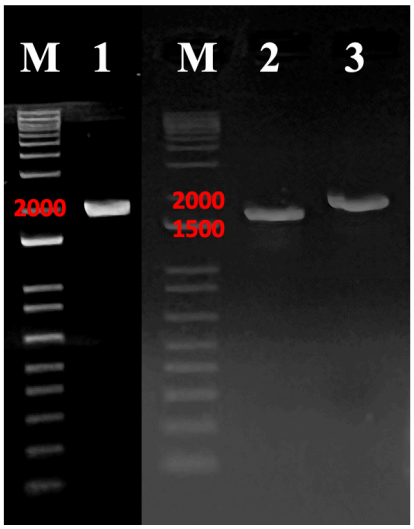


Fig. 3.5 Intermediate DNA constructs (*Fragment 1* fused to *Fragment 2*) for **β-CA** expression cassettes after purification from agarose gel: **lane M** – 1 Kb Plus DNA Ladder, **lane 1** – **MBP**-CA (1998 bp), **lane 2** – **TPI**-CA (1630 bp), **lane 3** – **EPG**-CA (1795 bp)

The final step for development of the Lacc and β-CA expression cassettes required purified intermediate constructs. This fusion reaction resulted in an effective formation of full-length cassettes containing three fragments (2.3.6) (**Fig. 3.2**). The amplification of the fused sequences was performed using P579 and P520 primers that bound at the ends the full-length expression cassettes. Six expression cassettes – TPIprM-**MBP**-Lacc, TPIprE-**EPG**-Lacc, TPIprT-**TPI**-Lacc, TPIprM-**MBP**-CA, TPIprE-**EPG**-CA, and TPIprT-**TPI**-CA were successfully constructed. This fusion step also resulted in some nonspecific products, that were eliminated by purification from gel (2.3.3). Purified six expression cassettes were assessed by gel electrophoresis (**Fig. 3.6**).

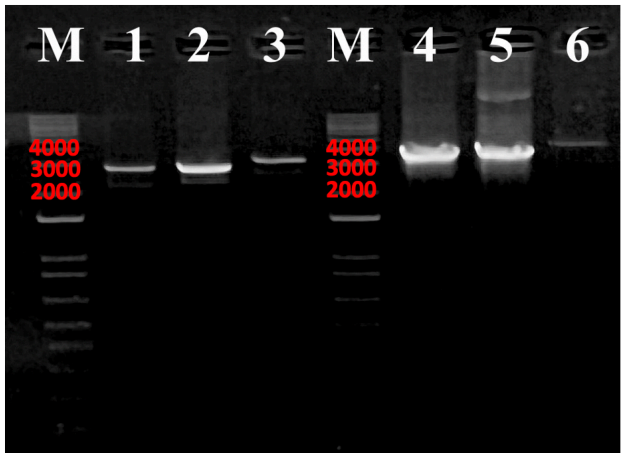


Fig. 3.6 Full-length **Lacc** and **β-CA** expression cassettes after purification from gel: **lane M** – 1 Kb Plus DNA Ladder, **lane 1** – TPIprT-**TPI**-CA (2797 bp), **lane 2** – TPIprE-**EPG**-CA (2887 bp), **lane 3** – TPIprM-**MBP**-CA (3166 bp), **lane 4** – TPIprT-**TPI**-Lacc (3788 bp), **lane 5** – TPIprE-**EPG**-Lacc (3878 bp), **lane 6** – TPIprM-**MBP**-Lacc (4157 bp)

The overlapping sequence between the intermediate constructs were long enough to make a strong junction that could successfully serve as prime site for DNA polymerase. Each cassette has rised from fusion of two intermediates, that had complementarity along the length of *Fragment 2*

(either MBP, EPG or TPI). The fusion of two fragments was possible in a presence of much shorter overlapping sequences. However, by using complementary sequence, that is above 1000 bp in length, the probability of fusion act could be increased.

It is important to emphasize the impact of primer melting temperature (Tm) during the both DNA construction steps (intermediate and final). In this study, fragments fusion reactions were performed under temperatures that were a few degrees lower than the calculated Tm of the primers. This way the resulting PCR products included relatively high amount of target construct, yet, a lot of non-specific products were present. The reason for this is related to the lower specificity of primer binding at lower annealing temperatures. However, using calculated Tms for particular pairs of primers during OE-PCR led to a lower concentration of full-length cassettes. A touchdown PCR protocol could be applied in this case – it decreases off-target priming and, therefore, increases the specificity. The idea of this method relies on the gradual decrease of the annealing temperature during PCR cycles starting with the temperature that is 5-10 °C higher than the calculated Tm. At the beginning of PCR, the annealing temperature is set to be high to compile the conditions for formation of highly specific primer-template hybrids. In the following cycles the temperature is decreased 0.5 °C per cycle. By applying this method the amount of non-specific products can be significantly lowered (Green & Sambrook, 2018).

In order to generate sufficient quantities of the verified DNA material, cloning and sequencing of the obtained cassettes had to be performed. To propagate and clone the constructed Lacc and β-CA expression cassettes, a vector was prepared for the homologous recombination. A pCAST cloning vector was selected for cloning of DNA constructions (**Fig. 3.7 A**). Vector consisted of *ori* from *E. coli*, AmpR promoter, AmpR gene coding a β-lactamase (provides resistance to ampicillin and is used as resistance marker), promoter and terminator of *K. marxianus TPI* gene. The presence of identical sequences in the vector and in the *Fragments 3* and *Fragment 1* (**Fig. 3.2**) allowed the insertion of the whole cassettes *via* homologous recombination in *E. coli*. First of all, *E. coli* cells were transformed with pCAST vector to propagate it (2.3.7, 2.3.8). It was extracted from positive transformants (2.3.2). Since pCAST contained TPI promoter and TPI terminator that are common with the constructed expression cassettes, homologous ends of vector were formed by cutting in the area of these sequences. The digestion of a vector was performed using *Eco*88I restriction enzyme (2.3.12) and resulted in a linearized pCAST vector (~3500 bp) (**Fig. 3.7 B**), that had homologous sequences (271 bp and 110 bp in length) indentical with the flanking ends of the expression cassettes. In literature, there are examples of effective homologous recombination even in the presence of only 20 bp homology between vector and insert (Jacobus & Gross, 2015). Therefore, having homology of above 100 bp allowed enhanced the possibility of recombination during the transformation of *E. coli*.

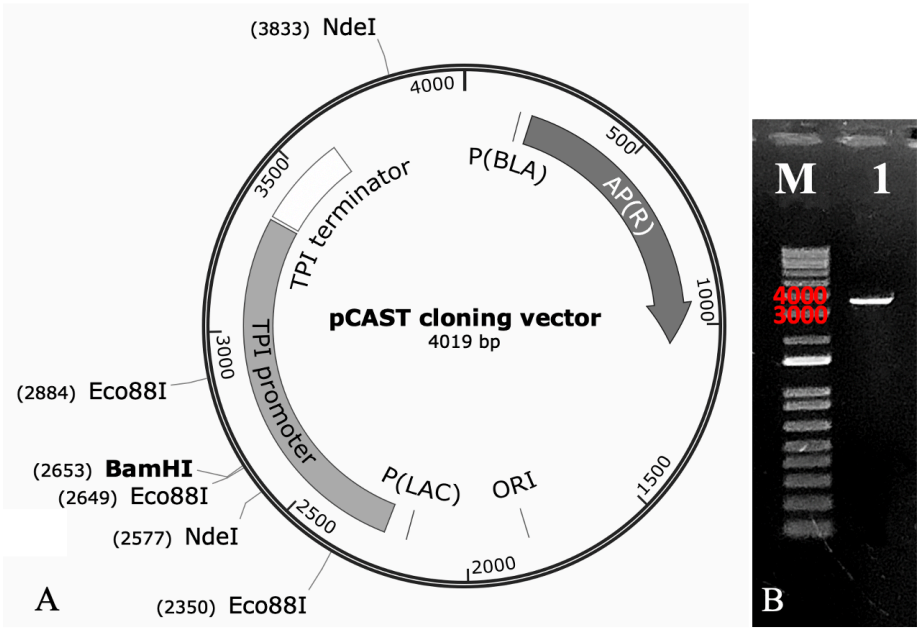


Fig. 3.7 Scheme of pCAST cloning vector used for cloning of **Lacc** and **β-CA** gene expression cassettes **(A)** and pCAST vector after digestion with *Eco*88I **(B)**

After the preparation of a cloning vector and all six expression cassettes, the construction of pCAST vector containing cassettes as inserts had to be performed. Prepared competent *E. coli* cells (2.3.7) were used for cloning of constructions. The ratio of vector/insert 1:3 was used to transform competent *E. coli* cells (2.3.8). In order to test the transformability of competent cells, a positive control was included in the experiment. Competent *E. coli* cells, without a contact with DNA, were used as negative control. Transformants were grown under the selective pressure of ampicillin (2.2.5). Relatively high amount of clones in positive control and in plates of the pCAST constructions were obtained, which indicated about an effective transformation. However, the plasmids from these colonies had to be tested for the presence of inserts – Lacc and β-CA cassettes (2.3.11).

PCR mixture included primer pairs for the amplification of *Fragment 2*. DNA samples of the cassettes each containing the *Fragments 2* (MBP, EPG, and TPI), served as positive controls to test the PCR system. Several positive clones were found based on colony PCR results. Expression cassettes (TPIprM-**MBP**-Lacc, TPIprE-**EPG**-Lacc, TPIprT-**TPI**-Lacc, TPIprM-**MB**P-CA, TPIprE-**EPG**-CA, and TPIprT-**TPI**-CA) were successfully inserted into pCAST vector by homologous recombination in *E. coli* cells (**Fig. 3.8**). The difference between intensity of bands in gel electrophoresis can be the consequence of non-equal amount of biomass put in the PCR reaction mixture. Also, nonspecific products in positive control samples appeared, probably, due to lower purity of control template.

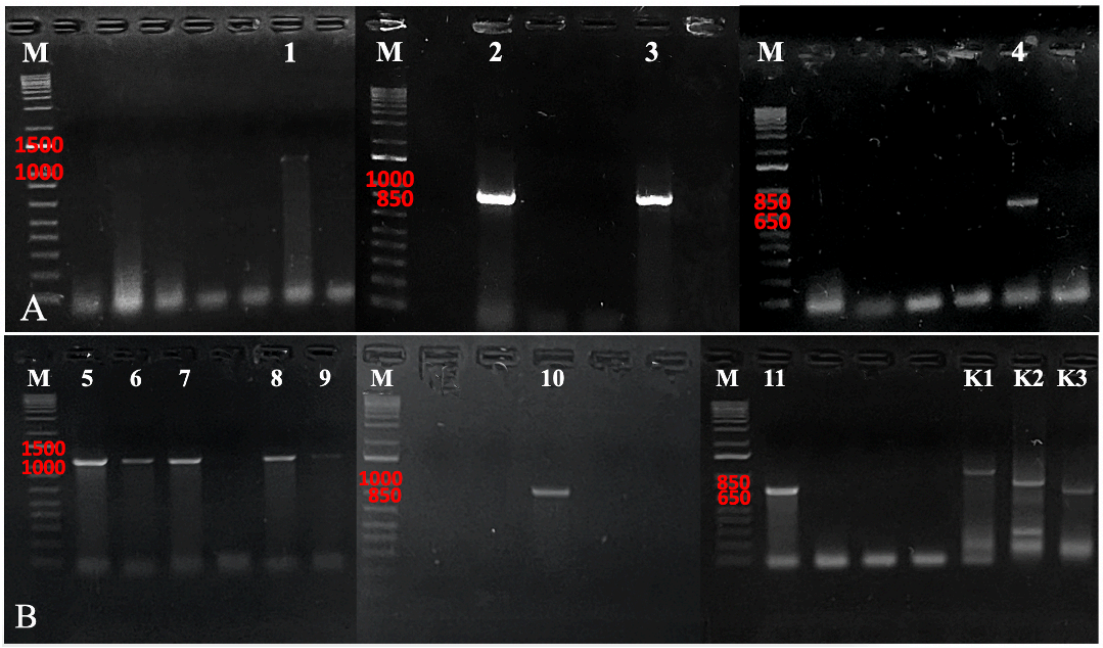


Fig. 3.8 Colony PCR analysis on *E. coli* transformants for the presence of **Lacc (A)** and **β-CA (B)** expression cassettes in pCAST vector. **Lane M** – 1 Kb Plus DNA Ladder, **lane 1** – clone containing TPIprM-**MBP**-Lacc, **lanes 2-3** – clones containing TPIprE-**EPG**-Lacc, **lane 4** – clone containing TPIprT-**TPI**-Lacc, **lanes 5-9** – clones containing TPIprM-**MBP**-CA, **lane 10** – clone containing TPIprE-**EPG**-CA, **lane 11** – clone containing TPIprT-**TPI**-CA. Positive controls: **lane K1** – cassette containing MBP, **lane K2** – cassette containing EPG, and **lane K3** – cassette containing TPI

Regarding the literature, a number of experiments on an ideal insert/vector ratio were done, however no strong findings could be formulated (Watson & García-Nafría, 2019). Some studies about homologous recombination in *E. coli* revealed little or no influence of insert/vector ratio to the efficiency of transformation. Contradictory, there are cases where 5:1 ratio showed the increase of colony formation significantly (Kostylev et al., 2015). Nevertheless, it shows that every homologous recombination requires the need of parameter optimization.

Additional analysis had to be done in order to ensure, that the insertion of cassettes was correct. Clones 1, 2, 3, 4, 5, 10, and 11 were examined for the presence of Lacc and β-CA expression cassettes in pCAST vector by digestion using *Nde*I and *Bam*HI restriction endonucleases (2.3.12). The obtained bands were of expected sizes, what reaffirmed the insertion of the targeted cassettes (**Fig. 3.9 A**). Digestion with *Bam*HI resulted in: 6763 bp (clone 1), 5794 bp (clone 5), 5425 bp (clone 11), 4925 bp and 1559 bp (clones 2, 3), 3956 bp and 1559 bp (clone 10) fragments. Digestion with *Nde*I resulted in: 4000 bp and 2763 bp (clone 1), 3721 bp and 2763 bp (clones 2, 3), 3631 bp and 2763 bp (clone 4), 3031 bp and 2763 bp (clone 5), 2752 bp and 2763 bp (clone 10), 2662 bp and 2763 bp (clone 11) fragments. A final construct of pCAST and TPIprM-**MBP**-Lacc is presented in **Fig. 3.9 B**. The rest of the constructions (pCAST_TPIprE-**EPG**-Lacc, pCAST_TPIprT-**TPI**-Lacc, pCAST_TPIprM-**MBP**-CA, pCAST_TPIprE-**EPG**-CA, pCAST_TPIprT-**TPI**-CA) are provided in Supplements section (Supplement 1, Supplement 2).

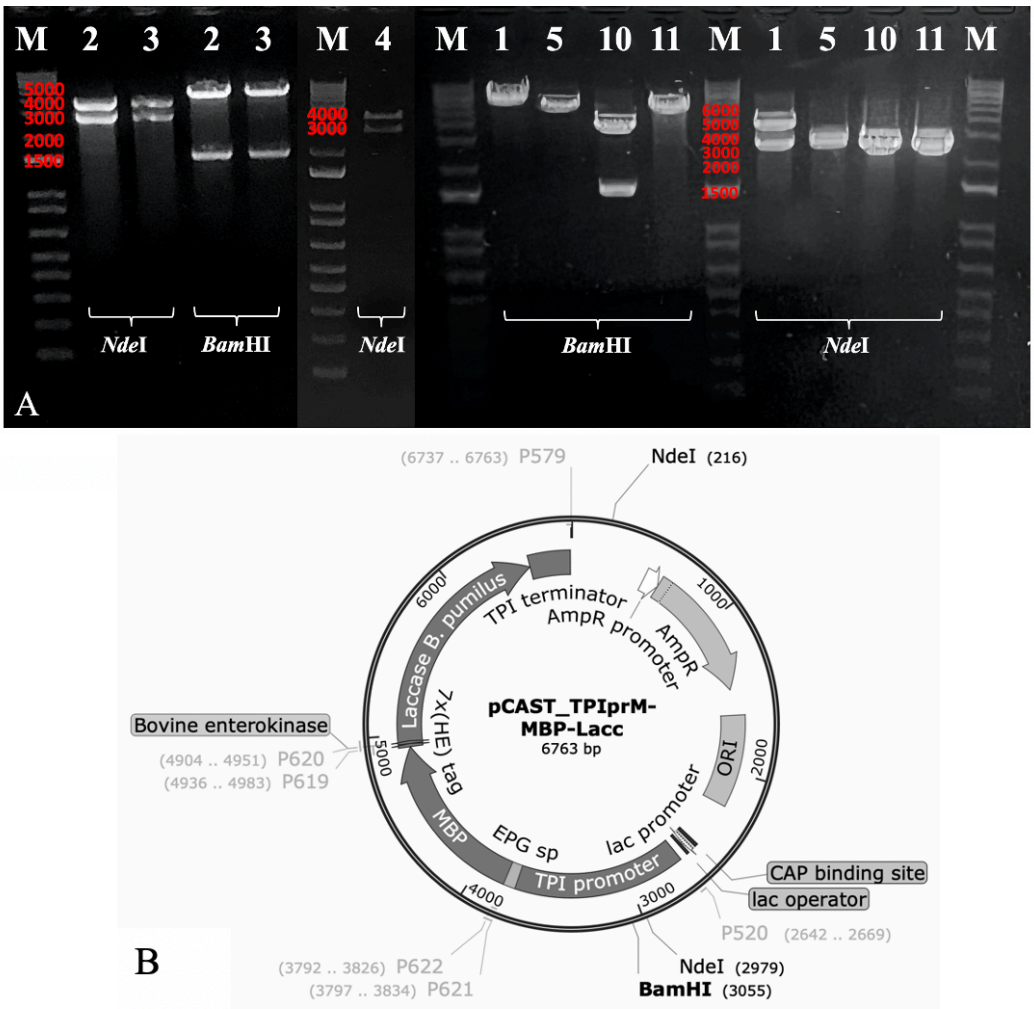


Fig. 3.9 Restriction analysis of clones 1, 2, 3, 4, 5, 10, and 11 using *Nde*I and *Bam*HI **(A)** and scheme of a complete construction of pCAST-TPIprM-**MBP**-Lacc **(B)**. **Lane M** – 1 Kb Plus DNA Ladder

All pCAST constructions were sequenced at DNA Sequencing Center, Institute of Biotechnology. The results proved, that all constructions for Lacc expression (TPIprM-**MBP**-Lacc, TPIprE-**EPG**-Lacc, TPIprT-TPI-Lacc) are correct in sequence and can be further used for the construction of expression vector. However, β-CA constructs – TPIprM-**MBP**-CA, TPIprE-**EPG**-CA, TPIprT-**TPI**-CA showed an alteration in DNA sequence – a one nucleotide deletion in a region of bovine enterokinase recognition site (**Fig. 3.9 B**). Since the mutation occurred in the same part of the sequence in all cassettes, it indicated that such error might have occurred during the PCR step (amplification of separate fragments) rather than homologous recombination. This way all of the cassettes, that were constructed using alterated *Fragment 3* (that contains bovine enterokinase recognition site), resulted in a mutated expression cassette. Mutation in protease recognition site is considered as critical point. Firstly, because it directly affects the performance of the enzyme, which recognizes that specific DNA sequence. Secondly, such alteration in DNA sequence usually results in shifting of the protein coding frame. Therefore, constructions containing this mutation can not be used, unless the alteration is eliminated. Due to a lack of time, further experiments were continued using correct Lacc expression constructions.

3.2 Development of Lacc expression system

The goal of this part of investigation was to insert Lacc expression cassettes (TPIprM-**MBP**-Lacc, TPIprE-**EPG**-Lacc, TPIprT-**TPI**-Lacc) into an expression vector KLEF (**Fig. 3.10 A**). It contained three selection markers – *bla* gene, *ura3* gene (from *K. marxianus*), and *fldA* gene (from *K. lactis*), that provided resistance to formaldehyde. For the assembly of the construction, a same strategy based on the cloning *via* homologous recombination in *E. coli* was used.

The first step included preparation of KLEF vector for homologous recombination. Competent *E. coli* cells were transformed with the mentioned vector in order to obtain a sufficient amount of plasmid for further experiments (2.3.7, 2.3.8). KLEF was extracted from the obtained positive transformants (2.3.2), and then was digested with *Sma*I, *Pdi*I, and *Eco*105I restriction endonucleases. The restriction resulted in a linearized KLEF (**Fig. 3.10 B**). The purpose of digestion with *Sma*I was to linearize vector, so it contained homology with the flanking ends of expression cassettes. The other two endonucleases were used to lower the possibility of false positive clones to appear due to transformation of *E. coli* with undigested KLEF. Recognition sites of *Sma*I, *Pdi*I, and *Eco*105I are located in region of TPI promoter and Lacc in vector (**Fig. 3.10 A**). It means that KLEF vectors, which would result in restriction by *Pdi*I, and *Eco*105I enzymes only, would have even longer homology needed for the subsequent homologous recombination. Digested KLEF was purified before the subsequent recombination step (2.3.3).

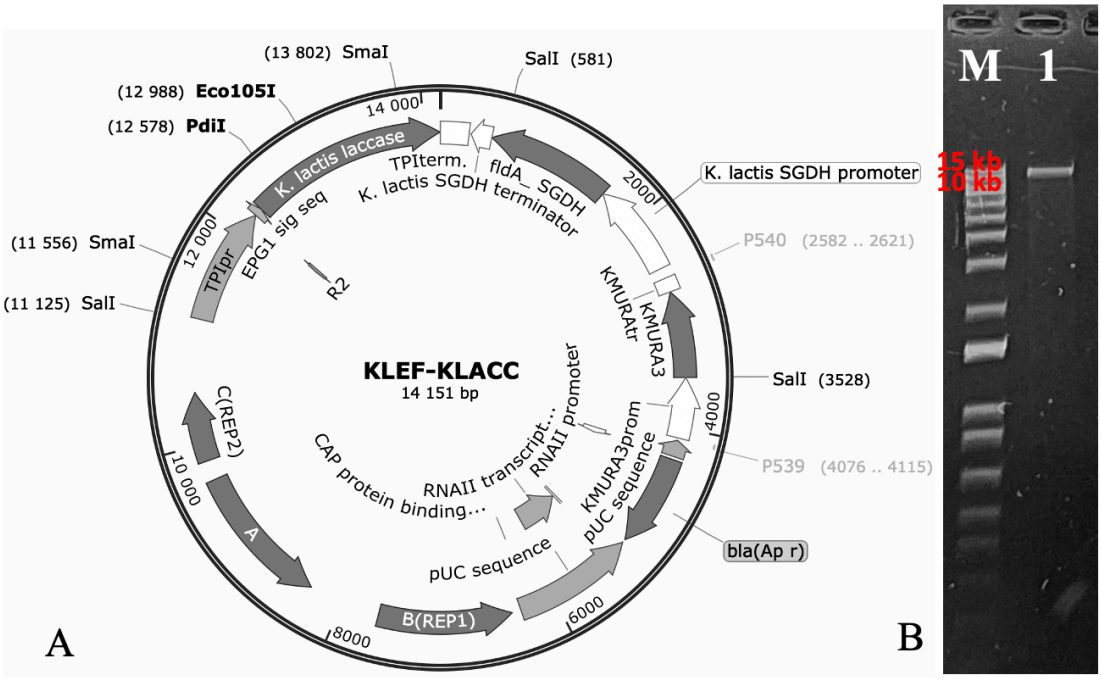


Fig. 3.10 Scheme of expression vector KLEF **(A)** and KLEF after digestion restriction endonucleases **(B)**. **Lane M** – 1 Kb Plus DNA Ladder, **lane 1** – KLEF digested with *Sma*I, *Pdi*I and *Eco*105I (~ 11.9 kb)

Expression cassettes were prepared for cloning in KLEF by amplification (2.3.5). pCAST constructions containing sequenced expression cassettes were used as templates for PCR. Template

concentration in PCR reaction mixture was reduced to the minimum in order to eliminate the clones with the original pCAST vector as a false positive transformation result. Additionally, to degrade the remainings of transformable pCAST constructions, the digestion of expression cassettes with *Alw*44I or *Aat*II was performed (**Fig. 3.11**). The latter digestion lowered the possibility of *E. coli* transformation with pCAST. Recognition sites of *Alw*44I and *Aat*II restriction endonucleases were located only in a region of pCAST vector, and no sites were present in expression cassettes. Purified cassettes (2.3.2) were further applied in homologous recombination experiments.

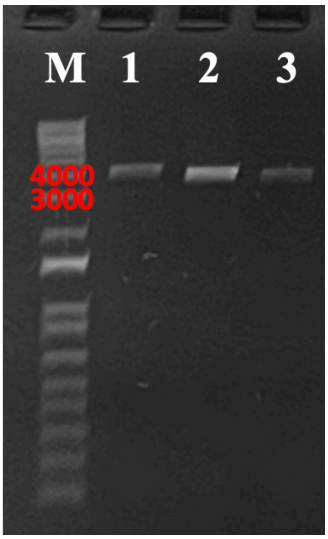


Fig. 3.11 Digestion of **Lacc** expression cassettes with *Alw*44I and *Aat*II: **lane M** – 1 Kb Plus DNA Ladder, **lane 1** – TPIprM-**MBP**-Lacc cassette digested with *Alw*44I, **lane 2** – TPIprE-**EPG**-Lacc cassette digested with *Aat*II, **lane 3** – TPIprT-**TPI**-Lacc cassette digested with *Alw*44I

Insertion of Lacc cassettes into KLEF vector was accomplished by homologous recombination in *E. coli*. Competent cells (2.3.7) were suspended with expression cassetes and linearized KLEF in insert/vector with a ratio of 3:1 (2.3.8). After the transformation, cells were grown under the selective pressure of ampicillin (2.2.5). Randomly selected portion of transformants were tested by colony multiple PCR (2.3.11). PCR mixture contained pairs of primers for *Fragment 2* (MBP, EPG ir TPI) amplification and another pair of primers (P539, P540) to amplify a part of vector KLEF (1534 bp sequence). By including two primer pairs instead, the obtained colony PCR results were more informative. It indicated that the colonies had DNA construct, consisting of the desired vector (KLEF) with the right insert (either TPIprM-**MBP**-Lacc, TPIprE-**EPG**-Lacc or TPIprT-**TPI**-Lacc cassettes). For a positive control, PCR reaction mixtures with two templates – empty KLEF vector and cassettes containing MBP, EPG, or TPI – were prepared (**Fig. 3.12 B**). Competent *E. coli* cells without a contact with DNA were served as negative control. The experiment resulted in several positive clones (appear as two bands in gel electrophoresis) (**Fig. 3.12**). Constructions, consisting of KLEF expression vector and each of Lacc expression cassettes, were successfully obtained.

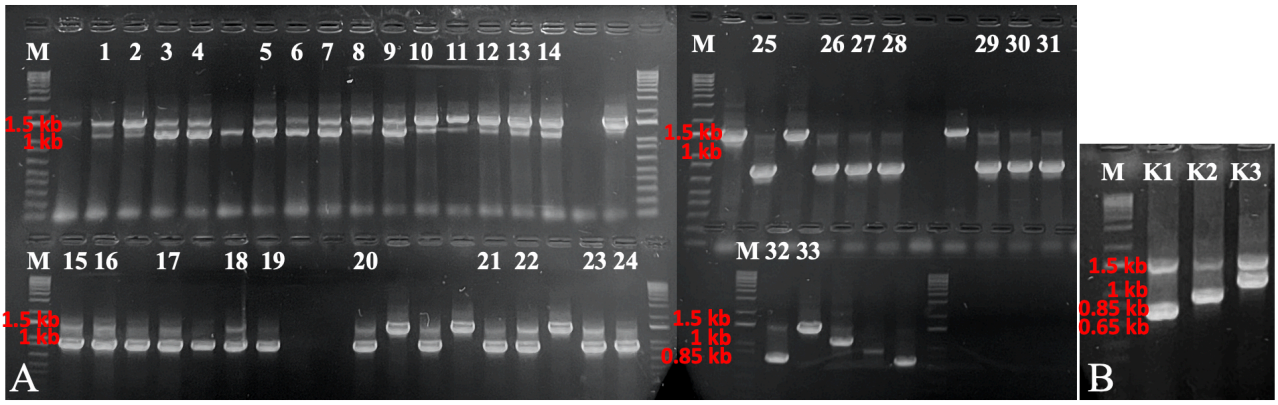


Fig. 3.12 Colony PCR analysis on *E. coli* transformants for the presence of **Lacc** expression cassettes in KLEF vector: **(A) lane M** – 1 Kb Plus DNA Ladder, **lanes 1-14** – clones containing KLEF-TPIprM-**MBP**-Lacc, **lanes 15-24** – clones containing KLEF-TPIprE-**EPG**-Lacc, **lanes 25-33** – clones containing KLEF-TPIprT-**TPI**-Lacc. Positive controls **(B): lane K1** – KLEF vector and cassette containing TPI, **lane K2** – KLEF vector and cassette containing EPG, **lane K3** – KLEF vector and cassette containing TPI

The colony PCR showed that the transformation resulted in a high amount of positive transformants. All three Lacc expression cassettes – TPIprM-**MBP**-Lacc, TPIprE-**EPG**-Lacc, and TPIprT-**TPI**-Lacc – were successfully inserted into KLEF vector.

However, a common result in transformations was the obtainment of colonies with empty vectors or any type of DNA that are not desired. In this study arba experiment, there were some clones containing the pCAST constructions were obtained. Also, a few clones had primary KLEF vector indicating, that some of it was left undigested during the preparation of homologous ends in vector. By this way, a part of *E. coli* cells were transformed with KLEF that did not have the inserted cassette. Clones containing primary KLEF or pCAST construction appeared as one band in electrophoresis (**Fig. 3.12 A**). Thus, in the transformation experiments it is important to use vector and cassettes, that are highly pure.

To verify, whether the construct was assembled correctly, a restriction analysis had to be performed. Clone 5 (with KLEF-TPIprM-**MBP**-Lacc cassette), clone 18 (with KLEF-TPIprE-**EPG**-Lacc cassette), and clone 32 (with KLEF-TPIprT-**TPI**-Lacc cassette) were further analyzed by restriction with *Sal*I restriction endonuclease (**Fig. 3.14 A**). Digestion occurred at three places in analyzed DNA constructs and resulted in three fragments, that were of correct length: clone 5 – 7597 bp, 4450 bp, 2947 bp, clone 18 – 7597 bp, 4171 bp, 2947 bp, and clone 32 – 7597 bp, 4081 bp, 2947 bp.

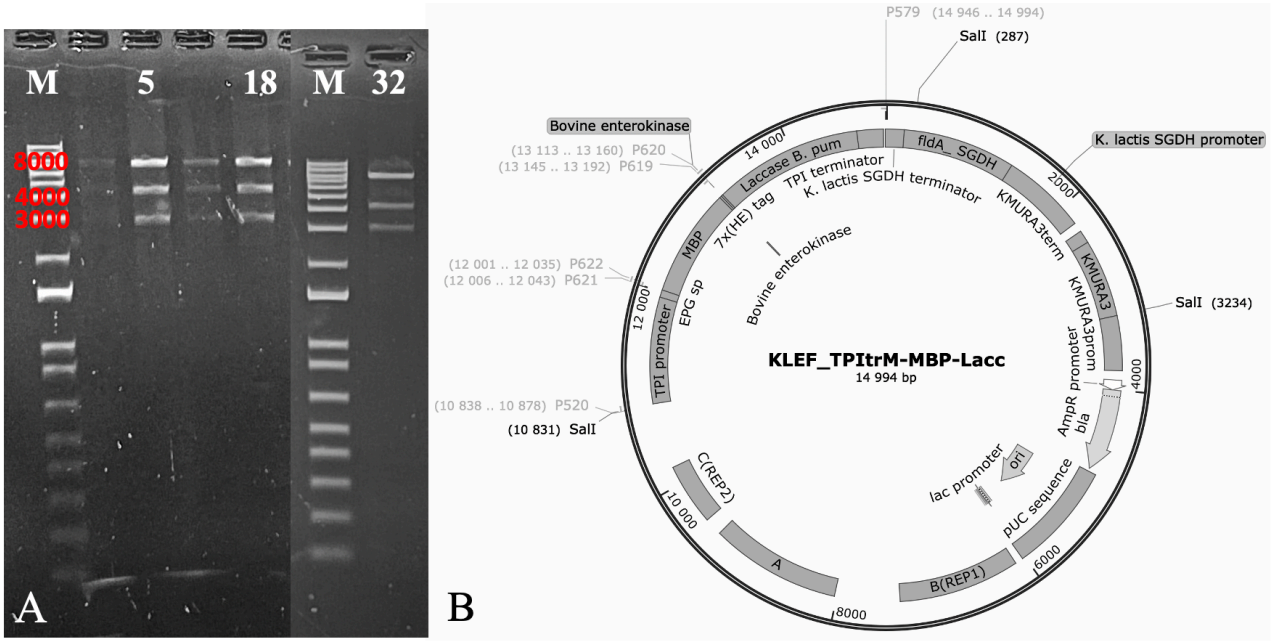


Fig. 3.14 Restriction analysis of clones 5, 18, 32 using *Sal*I **(A)** and scheme of a complete construction of KLEF-TPIprM-**MBP**-Lacc **(B). Lane M** – 1 Kb Plus DNA Ladder

All results verified the correct insertion of Lacc expression cassettes into KLEF vector using homology between the vector and cassettes. Recombination occurred in the regions of TPI promoter and TPI terminator in vector (**Fig. 3.14 B, Fig. 3.10 A**). Expression vector constructions with the other two Lacc cassettes are provided in a Supplementary section (Supplement 3). Lacc expression vectors were sequenced at DNA Sequencing Center, Institute of Biotechnology. The results confirmed, that the sequences of Lacc expression cassettes did not have any mutations and, therefore, can be applied in furher experiments.

The obtained and sequenced constructions of KLEF and Lacc expression cassettes – KLEF-TPIprM-**MBP**-Lacc, KLEF-TPIprE-**EPG**-Lacc, KLEF-TPIprT-**TPI**-Lacc – were used to transform *K. marxianus,* which was selected to serve as an expression host. Two transformation methods were applied – chemical lithium acetate transformation (2.3.9) and electroporation (2.3.10). One transformation using electroporation method was performed, yet it yielded in no transformants. However, by applying chemical transformation, eight clones were obtained, that were further analyzed by colony PCR (2.3.11). Seven of them were proved to have KLEF-TPIprT-**TPI**-Lacc DNA construct (**Fig. 3.15**).

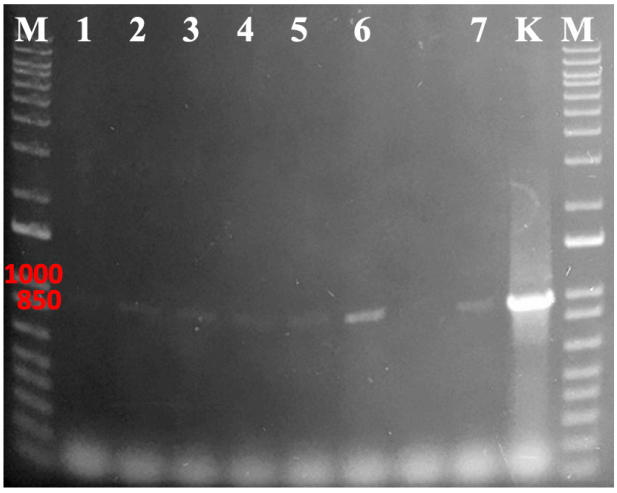


Fig. 3.15 Colony PCR analysis on *K. marxianus* transformants for the presence of KLEF-TPIprT-**TPI**-Lacc construct: **lane M** – 1 Kb Plus DNA Ladder, **lanes 1-7** – clones containing the target construct KLEF-TPIprT-**TPI**-Lacc, **lane K** – cassette containing TPI (positive control)

Colony PCR results indicated, that the transformation was very effective, since almost all transformants obtained were verified as positive. However, no transformants were obtained in cases of KLEF-TPIprM-**MBP**-Lacc and KLEF-TPIprE-**EPG**-Lacc constructs. This could have happened due to too low concentration of KLEF constructs in the mixture of competent yeast cells during the transformation. The differences in band intensity between the analyzed clones in the electrophoresis can be the reason of unequal amount of biomass of colonies that was put in the reaction mixture.

Regarding all of the results from this study, several further experiments have to be done. Since transformation of *K. marxianus* with DNA constructs of KLEF-TPIprM-**MBP**-Lacc and KLEF-TPIprE-**EPG**-Lacc have not yielded in any transformants, additional attemps are required. The increase in DNA concentration during transformations could be applied. Experiments with Lacc expression would help to make more conclusions, including the impact of fusion partners on enzyme solubility. Also, the correction of mutation in β-CA expression cassettes can be achieved by performing a site-directed mutagenesis, which is one of the applications of OE-PCR.

CONCLUSIONS

1. DNA expression cassettes for Lacc and β-CA – TPIprM-**MBP**-Lacc, TPIprE-**EPG**-Lacc, TPIprT-**TPI**-Lacc, TPIprM-**MBP**-CA, TPIprE-**EPG**-CA, TPIprT-**TPI**-CA – were fully assembled using homologous recombination in *E. coli*.
2. All of the obtained expression cassettes were successfully cloned into cloning vector pCAST.
3. KLEF expression vectors, containing TPIprM-**MBP**-Lacc, TPIprE-**EPG**-Lacc, TPIprT-**TPI**-Lacc expression cassettes, were constructed and sequenced.
4. *K. marxianus* yeasts were successfully transformed with KLEF-TPIprT-**TPI**-Lacc DNA construct.

DISSEMINATION OF RESULTS

Participation in conferences:

1. **Kondrataite A**., Šiekštelė R., Matijošytė I. *Investigation of Kluyveromyces marxianus as a host for the expression of the heterologous genes for the biocatalysts production.* March 27, 2021 – March 30, 2021. The COINS 2021, Vilnius, Lithuania. ***Poster presentation***.
2. **Kondrataite A**., Šiekštelė R., Matijošytė I. *Investigation of Kluyveromyces marxianus as a host for the expression of bacterial laccase.* July 19, 2021 – July 22, 2021. International Symposium on Biocatalysis and Biotransformations (Biotrans 2021), Graz, Austria. ***Poster presentation***.
3. **Kondrataite A**., Šiekštelė R., Matijošytė I. *Investigation of Kluyveromyces marxianus as a host for the expression of the heterologous genes for the biocatalysts production.* March 15, 2022 – March 18, 2022. Open Readings 2022, Vilnius, Lithuania. ***Poster presentation***.

Research funding:

1. Project **09.3.3-LMT-K-712-24-0100.** Funds from Lithuanian Research Council for the education of young scientists through laboratory practice (June 1, 2021 – August 31, 2021).
2. Project **09.3.3-LMT-K-712-25-0060.** Funds from Lithuanian Research Council for the education of young scientists through laboratory practice (September 1, 2021 – March 31, 2022).

VILNIUS UNIVERSITY
LIFE SCIENCES CENTER
Aušra Kondrataitė
Master Thesis

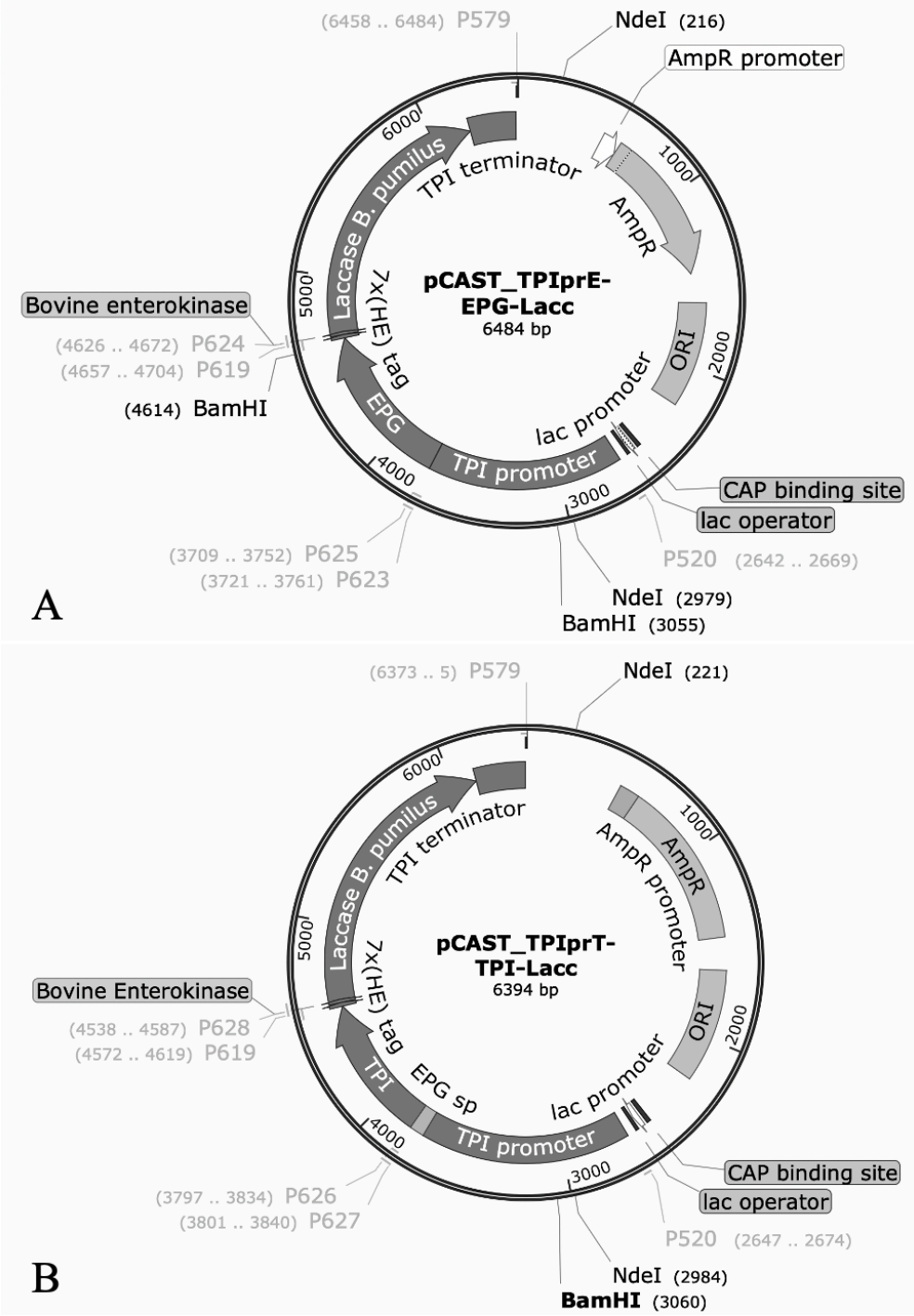
DEVELOPMENT OF HETEROLOGOUS GENE EXPRESSION SYSTEM IN YEASTS FOR THE PRODUCTION OF BIOCATALYSTS

SUMMARY

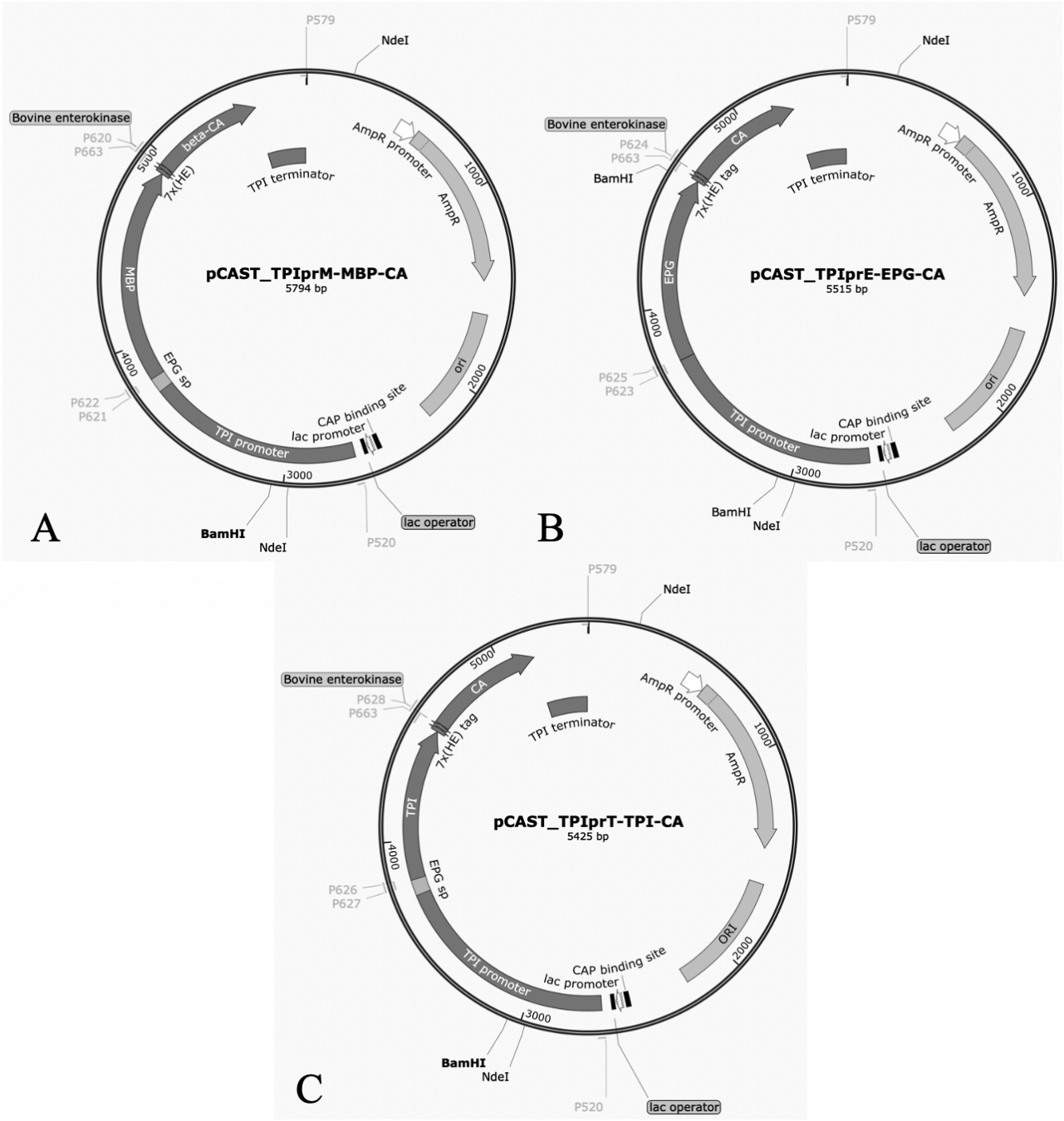
Biocatalysis is defined as the use of enzymes or whole cells to speed up chemical reactions. In order to meet the growing demand of biocatalysts in industry, it is crucial to have an effective gene expression system, that could be applied for a wide range of targeted enzymes. Since the problem of protein aggregation is still a very common issue in enzyme production, the focus of developing a prosperous expression systems is very relevant in academia and industry.

The aim of this study was to develop a convenient expression system in yeast *Kluyveromyces marxianus*. Laccase (Lacc) from *Bacillus pumilus* and β-carbonic anhydrase (β-CA) from *Bacilus mojavensis* were selected as target proteins, that have never been expressed in *K. marxianus*. The genes of three proteins – maltose-binding protein (MBP), endopolygalacturonase (EPG) and triosephosphate isomerase (TPI) – were used as chaperones for increasing the solubility. Lacc cassettes (TPIprM-**MBP**-Lacc, TPIprE-**EPG**-Lacc, TPIprT-**TPI**-Lacc) and β-CA cassettes (TPIprM-**MBP**-CA, TPIprE-**EPG**-CA, TPIprT-**TPI**-CA) were assembled using overlapping DNA sequences between the fusing fragments. All of them were successfully cloned into pCAST cloning vector by homologous recombination in *E. coli* and Lacc cassettes were verified by sequencing. The three sequenced Lacc expression cassettes were further inserted into KLEF expression vector resulting in final DNA constructs for the expression of Lacc. *K. marxianus* were successfully transformed with the construct of KLEF-TPIprT-**TPI**-Lacc.

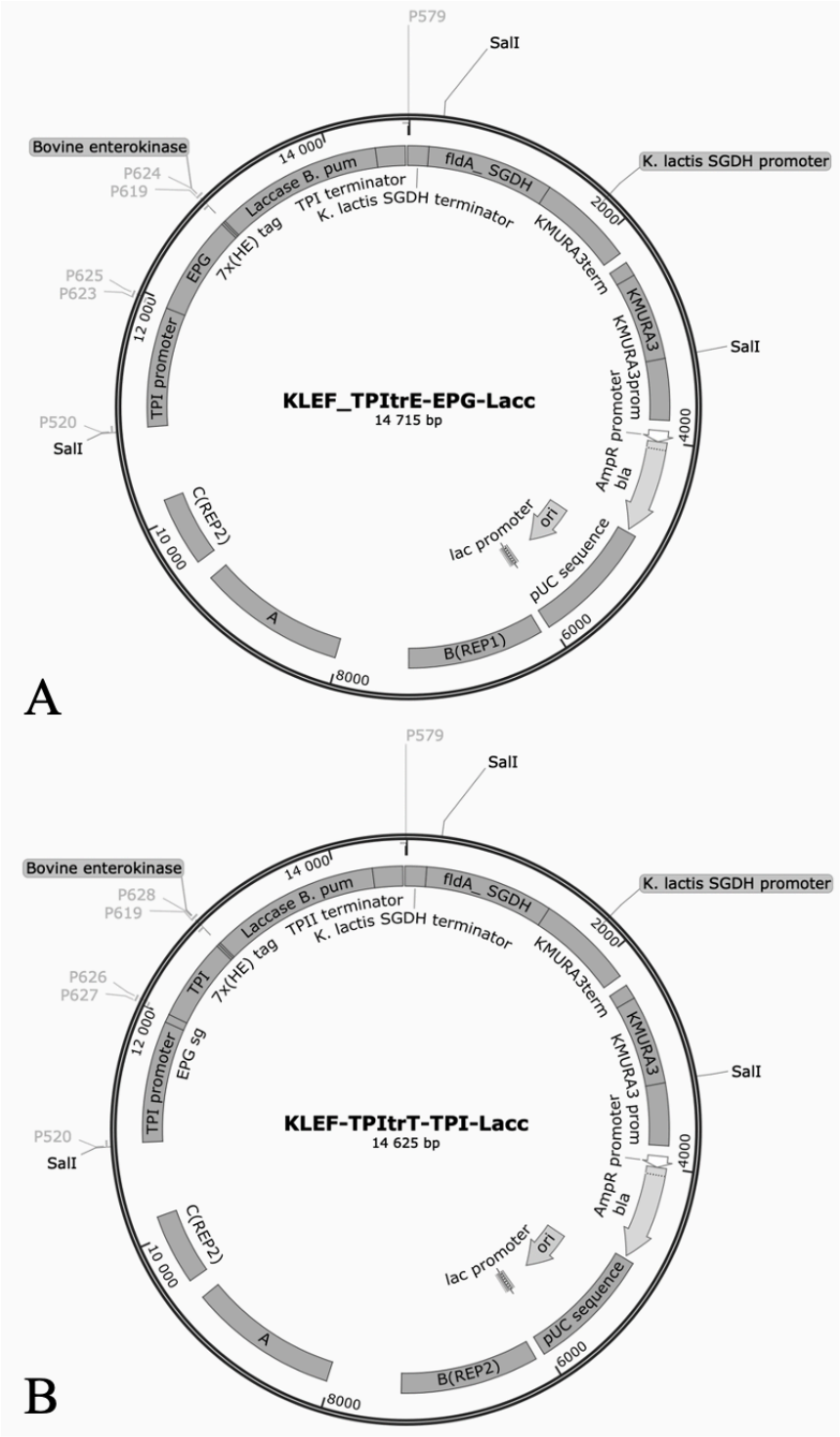
SUPPLEMENTAL MATERIAL



Supplement 1 Constructs of pCAST vector containing TPIprE-**EPG**-Lacc **(A)** and TPIprT-**TPI**-Lacc **(B)** expression cassettes



Supplement 2 Constructs of pCAST vector containing TPIprM-**MBP**-CA **(A)**, TPIprE-**EPG**-CA **(B)**, and TPIprT-**TPI**-CA **(C)** expression cassettes



Supplement 3 Constructions of KLEF vector containing TPIprE-**EPG**-Lacc **(A)** and TPIprT-**TPI**-Lacc **(B)** expression cassettes

ACKNOWLEDGEMENTS

I want to acknowledge Rimantas Šiekštelė for his accurate insights and great scientific collaboration. I also would like to thank to my supervisor Inga Matijošytė for the support and encouragement during the whole investigation. A special thanks goes to Jokūbas Krutkevičius, Dovilė Daunoraitė, and Justinas Babinskas, who were always open to help and shared valuable advices. I am truly grateful to the whole team of Sector of Applied Biocatalysis for a warm welcome and an opportunity to contribute to their work and achievements.

BIBLIOGRAPHY

Angeli, A., Carta, F., & Supuran, C. T. (2020). Carbonic anhydrases: versatile and useful biocatalysts in chemistry and biochemistry. *Catalysts*, *10*(9), 1008. https://doi.org/10.3390/catal10091008

Antošová, Z., & Sychrová, H. (2016). Yeast hosts for the production of recombinant laccases: a review. *Molecular Biotechnology*, *58*(2), 93–116. https://doi.org/10.1007/s12033-015-9910-1

Bartkevičiūtė, D., Šiekštelė, R., & Sasnauskas, K. (2000). Heterologous expression of the *Kluyveromyces marxianus* endopolygalacturonase gene (*epg1*) using versatile autonomously replicating vector for a wide range of host. *Enzyme and Microbial Technology*, *26*(9–10), 653–656. https://doi.org/10.1016/S0141-0229(00)00155-1

Bell, E. L., Finnigan, W., France, S. P., Green, A. P., Hayes, M. A., Hepworth, L. J., Lovelock, S. L., Niikura, H., Osuna, S., Romero, E., Ryan, K. S., Turner, N. J., & Flitsch, S. L. (2021). Biocatalysis. *Nature Reviews Methods Primers*, *1*(1), 46. https://doi.org/10.1038/s43586-021-00044-z

Bertani, G., & Weigle, J. J. (1953). Host controlled variation in bacterial viruses. *Journal of Bacteriology*, *65*(2), 113–121. https://doi.org/10.1128/jb.65.2.113-121.1953

Bertero, A., Brown, S., & Vallier, L. (2017). *Basic science methods for clinical researchers*. Elsevier. https://doi.org/10.1016/B978-0-12-803077-6.00002-3

Bornscheuer, U. T., Huisman, G. W., Kazlauskas, R. J., Lutz, S., Moore, J. C., & Robins, K. (2012). Engineering the third wave of biocatalysis. *Nature*, *485*(7397), 185–194. https://doi.org/10.1038/nature11117

Bornscheuer, U. T. (2018). The fourth wave of biocatalysis is approaching. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2110), 20170063. https://doi.org/10.1098/rsta.2017.0063

Brugnari, T., Braga, D. M., dos Santos, C. S. A., Torres, B. H. C., Modkovski, T. A., Haminiuk, C. W. I., & Maciel, G. M. (2021). Laccases as green and versatile biocatalysts: from lab to enzyme market – an overview. *Bioresources and Bioprocessing*, *8*(1), 131. https://doi.org/10.1186/s40643-021-00484-1

Bulter, T., Alcalde, M., Sieber, V., Meinhold, P., Schlachtbauer, C., & Arnold, F. H. (2003). Functional expression of a fungal laccase in *Saccharomyces cerevisiae* by directed evolution. *Applied and Environmental Microbiology*, *69*(2), 987–995. https://doi.org/10.1128/AEM.69.2.987-995.2003

Carson, S., Miller, H. B., Witherow, D. S., & Srougi, M. C. (2019). *Molecular Biology Techniques*. Elsevier. https://doi.org/10.1016/B978-0-12-815774-9.00003-4

Chapman, J., Ismail, A., & Dinu, C. (2018). Industrial applications of enzymes: recent advances, techniques, and outlooks. *Catalysts*, *8*(6), 238. https://doi.org/10.3390/catal8060238

Chen, F., Jin, W., Gao, H., Guo, Z., Lin, H., Li, J., Hu, K., Guan, X., Kalia, V. C., Lee, J.-K., Zhang, L., & Li, Y. (2019). Cloning, expression and characterization of two beta carbonic anhydrases from a newly isolated CO_2 fixer, *Serratia marcescens* Wy064. *Indian Journal of Microbiology*, *59*(1), 64–72. https://doi.org/10.1007/s12088-018-0773-6

Chu, S. (2009). Carbon capture and sequestration. *Science*, *325*(5948), 1599–1599. https://doi.org/10.1126/science.1181637

Clark, D. P., & Pazdernik, N. J. (2013). *Molecular Biology*. Elsevier. https://doi.org/10.1016/B978-0-12-378594-7.00007-X

Coelho, M. A. Z., Amaral, P. F. F., & Belo, I. (2010). *Yarrowia lipolytica*: an industrial workhorse. *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, *9*(2), 930–940.

Cohen, J. D., Abrams, E., Eccleshall, T. R., Buchferer, B., & Marmur, J. (1983). Expression of a prokaryotic gene in yeast: isolation and characterization of mutants with increased expression.

Molecular and General Genetics MGG, *191*(3), 451–459. https://doi.org/10.1007/BF00425762

Cozzarelli, N. R., Melechen, N. E., Jovin, T. M., & Kornberg, A. (1967). Polynucleotide cellulose as a substrate for a polynucleotide ligase induced by phage T4. *Biochemical and Biophysical Research Communications*, *28*(4), 578–586. https://doi.org/10.1016/0006-291X(67)90353-1

Cronk, J. D., Endrizzi, J. A., Cronk, M. R., O'neill, J. W., & Zhang, K. Y. J. (2001). Crystal structure of *E. coli* β-carbonic anhydrase, an enzyme with an unusual pH-dependent activity. *Protein Science*, *10*(5), 911–922. https://doi.org/10.1110/ps.46301

Dev, A., Srivastava, A. K., & Karmakar, S. (2018). *Handbook of Nanomaterials for Industrial Applications*. Elsevier. https://doi.org/10.1016/B978-0-12-813351-4.00013-4

DiCosimo, R., McAuliffe, J., Poulose, A. J., & Bohlmann, G. (2013). Industrial use of immobilized enzymes. *Chemical Society Reviews*, *42*(15), 6437. https://doi.org/10.1039/c3cs35506c

Durão, P., Chen, Z., Fernandes, A. T., Hildebrandt, P., Murgida, D. H., Todorovic, S., Pereira, M. M., Melo, E. P., & Martins, L. O. (2008). Copper incorporation into recombinant CotA laccase from *Bacillus subtilis*: characterization of fully copper loaded enzymes. *JBIC Journal of Biological Inorganic Chemistry*, *13*(2), 183–193. https://doi.org/10.1007/s00775-007-0312-0

El Qaidi, S., & Hardwidge, P. R. (2019). ABC cloning: an efficient, simple, and rapid restriction/ligase-free method. *MethodsX*, *6*, 316–321. https://doi.org/10.1016/j.mex.2019.02.007

Eminoğlu, A., Vullo, D., Aşık, A., Çolak, D. N., Supuran, C. T., Çanakçı, S., & Osman Beldüz, A. (2016). Cloning, expression and biochemical characterization of a β-carbonic anhydrase from the soil bacterium *Enterobacter* sp. B13. *Journal of Enzyme Inhibition and Medicinal Chemistry*, *31*(6), 1111–1118. https://doi.org/10.3109/14756366.2015.1100176

Fang, Z., Zhou, P., Chang, F., Yin, Q., Fang, W., Yuan, J., Zhang, X., & Xiao, Y. (2014). Structure-based rational design to enhance the solubility and thermostability of a bacterial laccase Lac15. *PLOS ONE*, *9*(7), e102423. https://doi.org/10.1371/journal.pone.0102423

Fisher, D. I., Mayr, L. M., & Roth, R. G. (2016). *Encyclopedia of Cell Biology*. Elsevier. https://doi.org/10.1016/B978-0-12-394447-4.10009-4

Gomes, A. M. V., Carmo, T. S., Carvalho, L. S., Bahia, F. M., & Parachin, N. S. (2018). Comparison of yeasts as hosts for recombinant protein production. *Microorganisms*, *6*(2), 38. https://doi.org/10.3390/microorganisms6020038

Green, M. R., & Sambrook, J. (2018). Touchdown polymerase chain reaction (PCR). *Cold Spring Harbor Protocols*, *2018*(5), pdb.prot095133. https://doi.org/10.1101/pdb.prot095133

Gutiérrez-González, M., Farías, C., Tello, S., Pérez-Etcheverry, D., Romero, A., Zúñiga, R., Ribeiro, C. H., Lorenzo-Ferreiro, C., & Molina, M. C. (2019). Optimization of culture conditions for the expression of three different insoluble proteins in *Escherichia coli*. *Scientific Reports*, *9*(1), 16850. https://doi.org/10.1038/s41598-019-53200-7

Hakulinen, N., & Rouvinen, J. (2015). Three-dimensional structures of laccases. *Cellular and Molecular Life Sciences*, *72*(5), 857–868. https://doi.org/10.1007/s00018-014-1827-5

Heckman, K. L., & Pease, L. R. (2007). Gene splicing and mutagenesis by PCR-driven overlap extension. *Nature Protocols*, *2*(4), 924–932. https://doi.org/10.1038/nprot.2007.132

Heckmann, C. M., & Paradisi, F. (2020). Looking back: a short history of the discovery of enzymes and how they became powerful chemical tools. *ChemCatChem*, *12*(24), 6082–6102. https://doi.org/10.1002/cctc.202001107

Hong, F., Meinander, N. Q., & Jönsson, L. J. (2002). Fermentation strategies for improved heterologous expression of laccase in *Pichia pastoris*. *Biotechnology and Bioengineering*, *79*(4), 438–449. https://doi.org/10.1002/bit.10297

Huang, M., Bao, J., & Nielsen, J. (2014). Biopharmaceutical protein production by *Saccharomyces cerevisiae*: current state and future prospects. *Pharmaceutical Bioprocessing*, *2*(2), 167–182. https://doi.org/10.4155/pbp.14.8

Ihssen, J., Reiss, R., Luchsinger, R., Thöny-Meyer, L., & Richter, M. (2015). Biochemical properties and yields of diverse bacterial laccase-like multicopper oxidases expressed in *Escherichia coli*. *Scientific Reports*, *5*(1), 10465. https://doi.org/10.1038/srep10465

Jacobus, A. P., & Gross, J. (2015). Optimal cloning of PCR fragments by homologous recombination in *Escherichia coli. PLOS ONE*, *10*(3), e0119221. https://doi.org/10.1371/journal.pone.0119221

Janusz, G., Pawlik, A., Świderska-Burek, U., Polak, J., Sulej, J., Jarosz-Wilkołazka, A., & Paszczyński, A. (2020). Laccase properties, physiological functions, and evolution. *International Journal of Molecular Sciences*, *21*(3), 966. https://doi.org/10.3390/ijms21030966

Jeong, H., Kim, H. J., & Lee, S. J. (2015). Complete genome sequence of *Escherichia coli* strain BL21. *Genome Announcements*, *3*(2), e00134-15. https://doi.org/10.1128/genomeA.00134-15

Jones, S. M., & Solomon, E. I. (2015). Electron transfer and reaction mechanism of laccases. *Cellular and Molecular Life Sciences*, *72*(5), 869–883. https://doi.org/10.1007/s00018-014-1826-6

Jozala, A. F., Geraldes, D. C., Tundisi, L. L., de Araújo Feitosa, V., Breyer, C. A., Cardoso, S. L., Mazzola, P. G., de Oliveira-Nascimento, L., de Oviveira Rangel-Yagui, C., de Oliveira Magalhães, P., de Oliveira, M. A. & Pessoa, A. (2016). Biopharmaceuticals from microorganisms: from production to purification. *Brazilian Journal of Microbiology*, *47*, 51–63. https://doi.org/10.1016/j.bjm.2016.10.007

Kadkhodaei, S., Memari, H. R., Abbasiliasi, S., Rezaei, M. A., Movahedi, A., Shun, T. J., & Ariff, A. B. (2016). Multiple overlap extension PCR (MOE-PCR): an effective technical shortcut to high throughput synthetic biology. *RSC Advances*, *6*(71), 66682–66694. https://doi.org/10.1039/C6RA13172G

Kalyani, D., Tiwari, M. K., Li, J., Kim, S. C., Kalia, V. C., Kang, Y. C., & Lee, J.-K. (2015). A highly efficient recombinant laccase from the yeast *Yarrowia lipolytica* and its application in the hydrolysis of biomass. *PLOS ONE*, *10*(3), e0120156. https://doi.org/10.1371/journal.pone.0120156

Karim, A., Gerliani, N., & Aïder, M. (2020). *Kluyveromyces marxianus*: an emerging yeast cell factory for applications in food and biotechnology. *International Journal of Food Microbiology*, *333*, 108818. https://doi.org/10.1016/j.ijfoodmicro.2020.108818

Kelly, T. J., & Smith, H. O. (1970). A restriction enzyme from *Hemophilus influenzae*. *Journal of Molecular Biology*, *51*(2), 393–409. https://doi.org/10.1016/0022-2836(70)90150-6

Ki, M.-R., & Pack, S. P. (2020). Fusion tags to enhance heterologous protein expression. *Applied Microbiology and Biotechnology*, *104*(6), 2411–2425. https://doi.org/10.1007/s00253-020-10402-8

Kim, J.-M., Park, S.-M., & Kim, D.-H. (2010). Heterologous expression of a tannic acid-inducible laccase3 of *Cryphonectria parasitica* in *Saccharomyces cerevisiae*. *BMC Biotechnology*, *10*(1), 18. https://doi.org/10.1186/1472-6750-10-18

Klonowska, A., Gaudin, C., Asso, M., Fournel, A., Réglier, M., & Tron, T. (2005). LAC3, a new low redox potential laccase from *Trametes* sp. strain C30 obtained as a recombinant protein in yeast. *Enzyme and Microbial Technology*, *36*(1), 34–41. https://doi.org/10.1016/j.enzmictec.2004.03.022

Kostylev, M., Otwell, A. E., Richardson, R. E., & Suzuki, Y. (2015). Cloning should be simple: *Escherichia coli* DH5α-mediated assembly of multiple DNA fragments with short end homologies. *PLOS ONE*, *10*(9), e0137466. https://doi.org/10.1371/journal.pone.0137466

Kwon, J., Cho, H., Kim, S., Ryu, Y., & Lee, J. (2021). A combination strategy of solubility enhancers for effective production of soluble and bioactive human enterokinase. *Journal of Biotechnology*, *340*, 57–63. https://doi.org/10.1016/j.jbiotec.2021.09.002

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, *157*(1), 105–132. https://doi.org/10.1016/0022-2836(82)90515-0

Lestari, C. S. W., & Novientri, G. (2021). Advantages of yeast-based recombinant protein technology as vaccine products against infectious diseases. *IOP Conference Series: Earth and Environmental Science*, *913*, 012099. https://doi.org/10.1088/1755-1315/913/1/012099

Li, M. Z., & Elledge, S. J. (2007). Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nature Methods*, *4*(3), 251–256. https://doi.org/10.1038/nmeth1010

Liu, H., Cheng, Y., Du, B., Tong, C., Liang, S., Han, S., Zheng, S., & Lin, Y. (2015). Overexpression of a novel thermostable and chloride-tolerant laccase from *Thermus thermophilus* SG0.5JP17-16 in *Pichia pastoris* and its application in synthetic dye decolorization. *PLOS ONE*, *10*(3), e0119833. https://doi.org/10.1371/journal.pone.0119833

Terol, G. L., Gallego-Jara, J., Martínez, R. A. S., Vivancos, A. M., Díaz, M. C., & de Diego Puente, T. (2021). Impact of the expression system on recombinant protein production in *Escherichia coli* BL21. *Frontiers in Microbiology*, *12*, 682001. https://doi.org/10.3389/fmicb.2021.682001

Lucas, M. F., Monza, E., Jørgensen, L. J., Ernst, H. A., Piontek, K., Bjerrum, M. J., Martinez, Á. T., Camarero, S., & Guallar, V. (2017). Simulating substrate recognition and oxidation in laccases: from description to design. *Journal of Chemical Theory and Computation*, *13*(3), 1462–1467. https://doi.org/10.1021/acs.jctc.6b01158

Luria, S. E., & Human, M. L. (1952). A nonhereditary, host-induced variation of bacterial viruses. *Journal of Bacteriology*, *64*(4), 557–569. https://doi.org/10.1128/jb.64.4.557-569.1952

Maheshwari, N., Kumar, M., Thakur, I. S., & Srivastava, S. (2019). Cloning, expression and characterization of β- and γ-carbonic anhydrase from *Bacillus* sp. SS105 for biomimetic sequestration of CO_2. *International Journal of Biological Macromolecules*, *131*, 445–452. https://doi.org/10.1016/j.ijbiomac.2019.03.082

Mayolo-Deloisa, K., González-González, M., & Rito-Palomares, M. (2020). Laccases in food industry: bioprocessing, potential industrial and biotechnological applications. *Frontiers in Bioengineering and Biotechnology*, *8*, 222. https://doi.org/10.3389/fbioe.2020.00222

Meselson, M., & Yuan, R. (1968). DNA restriction enzyme from *E. coli*. *Nature*, *217*(5134), 1110–1114. https://doi.org/10.1038/2171110a0

Mo, W., Wang, M., Zhan, R., Yu, Y., He, Y., & Lu, H. (2019). *Kluyveromyces marxianus* developing ethanol tolerance during adaptive evolution with significant improvements of multiple pathways. *Biotechnology for Biofuels*, *12*(1), 63. https://doi.org/10.1186/s13068-019-1393-z

Mollania, N., Khajeh, K., Ranjbar, B., Rashno, F., Akbari, N., & Fathi-Roudsari, M. (2013). An efficient *in vitro* refolding of recombinant bacterial laccase in *Escherichia coli*. *Enzyme and Microbial Technology*, *52*(6–7), 325–330. https://doi.org/10.1016/j.enzmictec.2013.03.006

Motohashi, K. (2019). A novel series of high-efficiency vectors for TA cloning and blunt-end cloning of PCR products. *Scientific Reports*, *9*(1), 6417. https://doi.org/10.1038/s41598-019-42868-6

Needle, D., & Waugh, D. S. (2014). *Protein Affinity Tags*. Springer New York. https://doi.org/10.1007/978-1-4939-1034-2_7

Nelson, M. D., & Fitch, D. H. A. (2012). *Molecular Methods for Evolutionary Genetics*. Humana Press. https://doi.org/10.1007/978-1-61779-228-1_27

Nguyen, M. T., Krupa, M., Koo, B.-K., Song, J.-A., Vu, T. T. T., Do, B. H., Nguyen, A. N., Seo, T., Yoo, J., Jeong, B., Jin, J., Lee, K. J., Oh, H.-B., & Choe, H. (2016). Prokaryotic soluble

overexpression and purification of human VEGF165 by fusion to a maltose binding protein tag. *PLOS ONE*, *11*(5), e0156296. https://doi.org/10.1371/journal.pone.0156296

Nguyen, M. T., Prima, M. J., Song, J.-A., Kim, J., Do, B. H., Yoo, J., Park, S., Jang, J., Lee, S., Lee, E., Novais, M. de P., Seo, H.-B., Lee, S., Cho, M.-L., Kim, C. J., Jang, Y. J., & Choe, H. (2019). Prokaryotic soluble overexpression and purification of oncostatin M using a fusion approach and genetically engineered *E. coli* strains. *Scientific Reports*, *9*(1), 13706. https://doi.org/10.1038/s41598-019-50110-6

Nishibori, N., Masaki, K., Tsuchioka, H., Fujii, T., & Iefuji, H. (2013). Comparison of laccase production levels in *Pichia pastoris* and *Cryptococcus* sp. S-2. *Journal of Bioscience and Bioengineering*, *115*(4), 394–399. https://doi.org/10.1016/j.jbiosc.2012.10.025

Nocentini, A., & Supuran, C. T. (2019). *Carbonic Anhydrases*. Elsevier. https://doi.org/10.1016/B978-0-12-816476-1.00001-0

Nocentini, A., Supuran, C. T., & Capasso, C. (2021). An overview on the recently discovered iota-carbonic anhydrases. *Journal of Enzyme Inhibition and Medicinal Chemistry*, *36*(1), 1988–1995. https://doi.org/10.1080/14756366.2021.1972995

Piscitelli, A., Pezzella, C., Giardina, P., Faraco, V., & Sannia, G. (2010). Heterologous laccase production and its role in industrial applications. *Bioengineered Bugs*, *1*(4), 254–264. https://doi.org/10.4161/bbug.1.4.11438

Pyser, J. B., Chakrabarty, S., Romero, E. O., & Narayan, A. R. H. (2021). State-of-the-art biocatalysis. *ACS Central Science*, *7*(7), 1105–1116. https://doi.org/10.1021/acscentsci.1c00273

Ranieri, D., Colao, M. C., Ruzzi, M., Romagnoli, G., & Bianchi, M. M. (2009). Optimization of recombinant fungal laccase production with strains of the yeast *Kluyveromyces lactis* from the pyruvate decarboxylase promoter. *FEMS Yeast Research*, *9*(6), 892–902. https://doi.org/10.1111/j.1567-1364.2009.00532.x

Revankar, M. S., & Lele, S. S. (2006). Enhanced production of laccase using a new isolate of white rot fungus WR-1. *Process Biochemistry*, *41*(3), 581–588. https://doi.org/10.1016/j.procbio.2005.07.019

Sachdev, D., & Chirgwin, J. M. (2000). *Methods in Enzymology*. Elsevier. https://doi.org/10.1016/S0076-6879(00)26062-X

Schramm, F. D., Schroeder, K., & Jonas, K. (2020). Protein aggregation in bacteria. *FEMS Microbiology Reviews*, *44*(1), 54–72. https://doi.org/10.1093/femsre/fuz026

Šiekštelė, R., Bartkevičitūtė, D., & Sasnauskas, K. (1999). Cloning, targeted disruption and heterologous expression of the *Kluyveromyces marxianus* endopolygalacturonase gene (EPG1). *Yeast*, *15*(4), 311–322. https://doi.org/10.1002/(SICI)1097-0061(19990315)15:4<311::AID-YEA379>3.0.CO;2-9

Šiekštelė, R., Veteikytė, A., Tvaska, B., & Matijošytė, I. (2015). Yeast *Kluyveromyces lactis* as host for expression of the bacterial lipase: cloning and adaptation of the new lipase gene from *Serratia* sp. *Journal of Industrial Microbiology and Biotechnology*, *42*(10), 1309–1317. https://doi.org/10.1007/s10295-015-1655-0

Singh, R., Kumar, M., Mittal, A., & Mehta, P. K. (2016). Microbial enzymes: industrial progress in 21st century. *3 Biotech*, *6*(2), 174. https://doi.org/10.1007/s13205-016-0485-8

Spohner, S. C., Schaum, V., Quitmann, H., & Czermak, P. (2016). *Kluyveromyces lactis*: an emerging tool in biotechnology. *Journal of Biotechnology*, *222*, 104–116. https://doi.org/10.1016/j.jbiotec.2016.02.023

Theerachat, M., Emond, S., Cambon, E., Bordes, F., Marty, A., Nicaud, J.-M., Chulalaksananukul, W., Guieysse, D., Remaud-Siméon, M., & Morel, S. (2012). Engineering and production of laccase from *Trametes versicolor* in the yeast *Yarrowia lipolytica*. *Bioresource Technology*, *125*, 267–274. https://doi.org/10.1016/j.biortech.2012.07.117

Truppo, M. D. (2017). Biocatalysis in the pharmaceutical industry: the need for speed. *ACS Medicinal Chemistry Letters*, *8*(5), 476–480. https://doi.org/10.1021/acsmedchemlett.7b00114

Walhout, A. J. M., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., & Vidal, M. (2000). *Methods in Enzymology*. Elsevier. https://doi.org/10.1016/S0076-6879(00)28419-X

Watson, J. F., & García-Nafría, J. (2019). In vivo DNA assembly using common laboratory bacteria: a re-emerging tool to simplify molecular cloning. *Journal of Biological Chemistry*, *294*(42), 15271–15281. https://doi.org/10.1074/jbc.REV119.009109

Yadav, D. K., Yadav, N., Yadav, S., Haque, S., & Tuteja, N. (2016). An insight into fusion technology aiding efficient recombinant protein production for functional proteomics. *Archives of Biochemistry and Biophysics*, *612*, 57–77. https://doi.org/10.1016/j.abb.2016.10.012

Yang, J., Li, W., Ng, T. B., Deng, X., Lin, J., & Ye, X. (2017). Laccases: production, expression regulation, and applications in pharmaceutical biodegradation. *Frontiers in Microbiology*, *8*, 832. https://doi.org/10.3389/fmicb.2017.00832

Zhu, S., Gong, C., Ren, L., Li, X., Song, D., & Zheng, G. (2013). A simple and effective strategy for solving the problem of inclusion bodies in recombinant protein technology: his-tag deletions enhance soluble expression. *Applied Microbiology and Biotechnology*, *97*(2), 837–845. https://doi.org/10.1007/s00253-012-4630-y