LUND UNIVERSITY

LTH

FACULTY OF ENGINEERING

Evaluation of strategies for construction of protease libraries in yeast.

Sidharth Jaya Sankar

Supervisors: Magnus Carlquist, Tova Lindh
Examiner: Marie Gorwa Grauslund

Master Thesis
Applied Microbiology (TMB)
Lund University
Spring 2022

Acknowledgments

This master's degree project was conducted at the Applied Microbiology Department at Lund University during the spring semester of 2022.

I want to say a special thank you to Dr. Magnus Carlquist, who supervised my thesis. I especially want to pay gratitude for helpful suggestions, support, constructive feedback, and teaching me all the aspects required for successfully completing this project.

I would like to thank Dr. Marie Gorwa Grauslund, for accepting the role of examiner.

Many thanks to Tova Lindh, for co-supervising and supporting my work. I am very thankful for teaching me methods and underlying principles, providing constructive feedback, and helpful recommendations.

Last but not the least, I would like to thank all the people at TMB for the fun times and supporting me throughout the project.

Abstract

Proteases are a group of enzymes that are involved in breaking down larger proteins into smaller polypeptides or single amino acids, and they are used in various applications. In this project, the potential of budding yeast *Saccharomyces cerevisiae* as host for the construction of protease libraries is evaluated. Three model proteases of bacterial origin are used to evaluate transformation efficiency, protease gene expression level, and yeast cell fitness. A plasmid-based system was compared with a CRISPR/Cas9-based chromosomal integration system in terms of transformation efficiency. It was found that the plasmid system had 5-fold higher transformation efficiency. To study expression of the protease genes they were fused to green fluorescent protein (GFP), enabling easy detection of the proteins by measuring fluorescence. A new experimental protocol to screen the GFP expression level of the yeast colonies in microtiter plates based on fluorometry and spectrometry, to measure fluorescence and optical density, respectively, was developed. Gene expression and viability for yeast strains were also studied using the flow cytometry. Growth and metabolic profiles of yeast strains were studied in different conditions to understand possible effects of proteases on yeast fitness. It was found that two out of three proteases were expressed at high level and yeast still displayed high fitness, demonstrating its potential as host for proteases with a certain substrate specificity. On the other hand, the results indicate that recombinant proteases can also have a detrimental effect on yeast fitness thereby restricting the possibility to build a broad specificity protease gene expression library.

Table of contents

Abbreviations

OD- optical Density

PBS- phosphate buffered saline

PPI- protein-protein interaction

1.Introduction

Human genome comprises 2% of protease genes and this group of enzymes has important roles in numerous pathways and functions in our body. Any abnormal functioning of proteases in our body leads to serious health concerns like neurodegenerative diseases, cancer, microbial infections and hence it is a very vital enzyme group [1]. Cleaving peptides bond catalytically is a unique property of proteases and this ability increases its commercial value in various sectors including biopharmaceuticals, food or detergent industries [2]. Protease-targeted libraries are increasing in demand due to the ability for screening for candidates of interest in numerous pharmaceutical development projects. For constructing and implementing studies on collection of modified strains (libraries) in this project, yeasts was used as the host organism [3]. *Saccharomyces cerevisiae* is an important yeast in terms of industrial application and has immense potential in recombinant protein production. They are the best-known expression system within yeasts and possess significant demand in nutritional as well as biopharmaceutical industries. Several advantages are observed for *S. cerevisiae* as an expression host, which includes pathogen-free production, easy manipulation of genetic makeup, high biomass formation; these characteristics attract the industries to use this yeast as a host. [4].

1.1 Libraries in yeasts

Libraries are collections of strains where different genes are modified [6]. For studies relying on gene libraries, yeast *S. cerevisiae* is an important tool. There are numerous libraries of different types currently available in *S. cerevisiae* (Table 1). They are used for various purposes, including genomic and proteomic studies, understanding gene function, drug development, understanding protein-protein interactions and so on. Methods to construct libraries covering large variation in genetic space is needed.

Table 1: Examples of libraries constructed in *S. cerevisiae* and their purpose.

Library constructed	**Purpose**	**References**
Protein-fragment complemented assay library	in vivo screening for protein-protein interactions.	Tarassov et al.,2008
Protein Fusion library using bimolecular fluorescence complementation assay	Analyse protein-protein interactions in biological processes	Sung et al.,2013
Genome-wide library using SWAp-Tag approach	Further explore yeast cell biology	Weill et al.,2018

Yeast double transporter gene deletion library	Drug designing and development	Almeida et al.,2021
GFP tagged fusion protein library	Characterize protein subcellular localization in eukaryotic cells	Huh et al.,2002
Gene deletion library	Determine effect of gene deletion to understand gene function	Giaever et al.,2002

Protein fragment complementation assays are used for the detection of protein-protein interactions. Tarassov et al., (2008) focused on building library using protein fragment complementation assay in order to study protein-protein interactions in detail. Similarly, Sung et al., (2013) aimed to understand protein-protein interaction in *S. cerevisiae,* and used bimolecular fluorescent assay to detect PPI and the library used was protein fusion library for the study. Almeida et al.,2021 and Giaever et al.,2002, are both deletion libraries for the purpose of the library to understand the novel drug developments and study the effect of gene deletion, respectively. The GFP tagged fusion protein library and the fluorescence of each strain obtained was used to analyse the protein subcellular localization in *S. cerevisiae* by Huh et al.,2002.

1.2 Gene Expression systems

Inducible promoter systems are commonly used in yeasts as they help in attaining controllable gene expression. This also allows in maximizing growth before inducing the expression phase; hence, separation of growth phase from protein production phase is observed. There are generally two different strategies adopted by which the gene constructs for recombinant protein expression can be introduced in *S. cerevisiae*: plasmid-based systems and chromosomal integration systems.

The discovery of the 2-micron DNA prompted more advancements in studies relying on yeast episomal plasmid DNA in *S. cerevisiae* [12]. In plasmid constructs, the requirement of selectable marker genes is essential to enable the selection of transformed cells. The marker genes generally used are wild type alleles of yeast genes encoding enzymes involved in biosynthesis of essential metabolites or amino acids. These types of markers, for example *URA3*, *LEU2*, *HIS3,* are restricted to auxotrophic yeast strains. The auxotrophic strains propagate only in the media containing the growth factors required, until the strains are transformed with a functional allele of the marker gene [13]. Therefore, auxotrophic markers are important set of genes in plasmid-based systems. However, the plasmid-based systems are known to have issues related to stability [12].

Chromosomal integration systems focus instead on DNA insertion into the host genome performed using homologous recombination. To achieve significantly higher gene expression, several factors should be considered. One among such factors is to avoid non-targeted

integration, as it leads to unwanted side effects, such as altered expression of fitness-related genes. Consequently, it is important to integrate the gene construct at a position in the genome where the deletion does not affect the yeast growth. Another factor to consider when selecting the integration site is the level of protein expression, as it can significantly change depending on the location on the genome [14]. Nowadays, the homologous recombination in this system is routinely achieved through the CRISPR-Cas9 technique, detailed below.

1.3 CRISPR/Cas-9 gene editing technique in yeast

CRISPR stands for Clustered regularly interspaced short palindromic repeats. CRISPR and associated Cas proteins were discovered in 2007. This precise genome editing technology was invented by Emmanuelle Charpentier and Jennifer Doudna and they were awarded Nobel prize in 2020 in the field of chemistry for their discovery. For CRISPR/Cas9 system to function in *S. cerevisiae,* Cas9 gene is required to be expressed in yeast cells and this is performed with an episomal vector [15]. The Cas9 proteins use an RNA molecule termed single guide RNA (sgRNA) to cleave at the specified position in the genome. The sgRNA consists of tracrRNA and crRNA, which together binds to the cas9 proteins to create double stranded breakage of the DNA at the specified position. The PAM sequence NGG that is recognised by Cas9, is situated 3-4 nucleotides upstream of the position where Cas9 cleaves. The DNA repair mechanism occurs after the cleavage by two different ways in yeast, namely non-homologous end joining (NHEJ) and Homology directed repair (HDR). In our experiments, the donor DNA (DNA fragment contains genes to be integrated into genome) is provided to the cells and repair the double strand break by homologous recombination.

1.4 Green fluorescent protein as reporter gene for screening

Green fluorescent protein (GFP) was sourced from jellyfish *Aequorea Victoria* and is used as a non-invasive reporter tool for studying gene expression in vivo. By fusing GFP with the target gene it is possible to detect the fused GFP-protein product by measuring the fluorescence of the cells. GFP-based fusion technology in yeast enables simple screening of expression level of target genes. Screening of clones can be performed to determine the expression levels of protein formation using different approaches based on for example fluorometry and flow cytometry. While flow cytometry measures the fluorescence of individual cells, fluorometry instead measures fluorescence of cell cultures. Therefore, fluorometry requires the normalization of fluorescence levels to the cell density value, often measured by OD_{600}. This high-throughput method using GFP was used to characterize protein expression from a library of strains in previous studies [33]. For fluorometry, the specific cell density normalization approach is critical to reach reliable comparison between strains and/or conditions. Often the measurement is made for a single culture dilution (at a single OD-value); however, this gives relatively poor resolution and may hinder the distinction of possible subtle differences between strains. Hence, the normalisation procedure for fluorometry to analyse yeast expressing GFP-protease fusions was investigated in detail in the present project. An attempt to increase the resolution was made by measuring fluorescence of yeast at multiple cell densities.

1.5 Aim

In this master thesis project, the objective was to evaluate different parameters of relevance for successfully constructing and screening of protease-GFP fusion library in yeast. Three aspects were specifically considered (Figure 2): (a) transformation efficiency of a plasmid system vs. chromosomal integration system (strain comparison) (b) the screening set-up using fluorometry (GFP-based slope methodology development), and (c) characterisation of yeast strains expressing GFP-protease fusions.

Throughout the project, yeast strains expressing GFP-fusions with three different proteases, namely protease 1, protease 2 and protease 3, and strains with or without expression of only GFP were used. All the proteases used are sourced from bacteria.

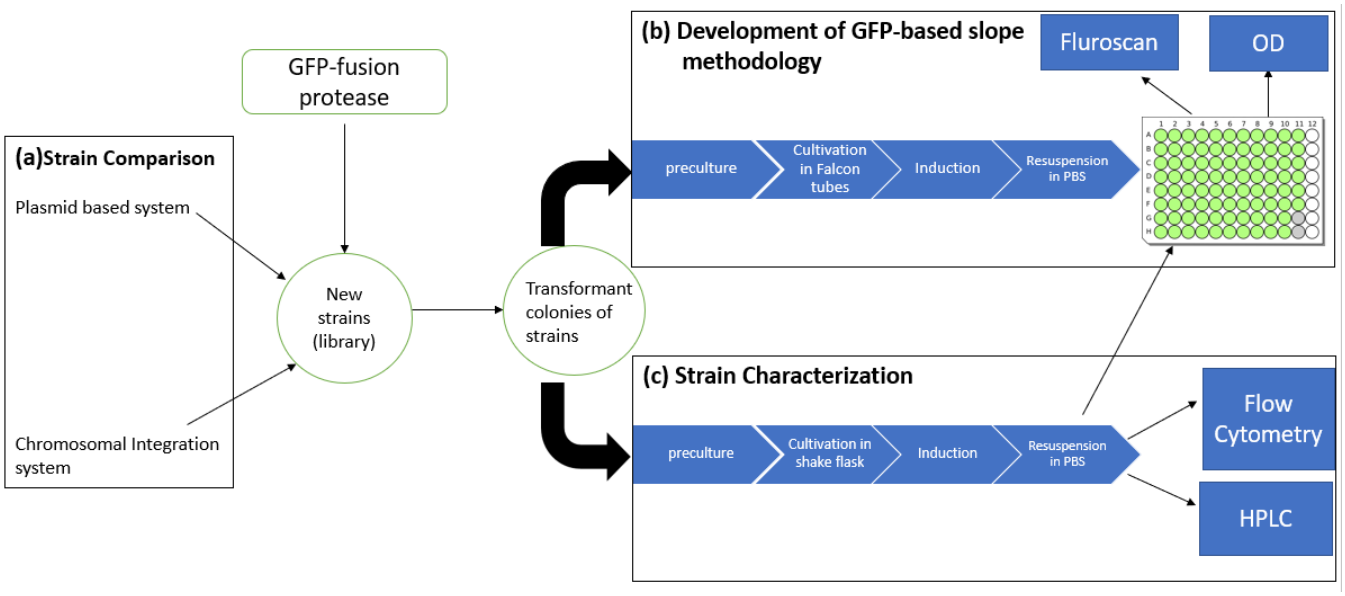

Figure 1: Overall project pipeline summary consisting of three different sections;(a) Comparison of plasmid-based systems and Chromosomal integration between using transformation efficiency, (b) Development of the GFP-based slope methodology validated using GFP positive and negative strains (positive and negative Control) and performed for different yeast strains used, (c) Characterization of different yeast strains using flow cytometry, HPLC and for validating slope method.

2. Materials and methods

2.2 Plasmids, strains, media and culture conditions

Plasmids and strains used in this study are listed in *Table 1*, including the relevant genotype. For plasmid preparation (Gene JET Plasmid Miniprep Kit, ThermoFisher Scientific.), *Escherichia coli* was cultivated in 50 mL falcon tubes containing 5 ml LB media (5 g/L yeast extract, 10 g/L peptone, 5 g/L NaCl) supplemented with ampicillin (100 mg/L) on a shaking table in 37°C incubator for 16 hours. The *S. cerevisiae* yeast strains cultivated were grown in the orbital shaker at 30°C and 180 r.p.m with Yeast nitrogen base (YNB) (Becton, Dickinson and company, USA) glucose medium (20 g/L glucose, 25 mM potassium phosphate buffer pH 6.5 ,6.7 g/L Yeast nitrogen base without amino acids). The volume of media added to the shake flasks was 10% of the total volume to ensure aerobic conditions. For induction of the GAL1 promotor, YNB galactose (20 g/L galactose, 25 mM potassium phosphate buffer pH 6.5, 6.7 g/L yeast nitrogen base without amino acids) was used. When performing CRISPR/Cas9 system the transformed cells were plated on solid yeast Peptone Dextrose (YPD) medium (20 g/L glucose, 20 g/L peptone and 10 g/L yeast extract) supplemented with two different antibiotics, namely geneticin (g418) and nourseothricin (cloNAT) having a concentration of 200 mg/L and 100mg/L respectively. These antibiotics provide selection for Cas-9 and gRNA plasmids, respectively.

Table-2: Plasmids and strains used in this study along with their relevant genotype and references.

Plasmid Name	**Relevant Genotype**	**Reference**
YEplac181	LEU2p-LEU2-LEU2t	(De La Cruz et al.,1997)
pCfB3045	gRNA XI-3; *nat*MX	(Jessop-Fabre et al., 2016)
pCfB3020	gRNA X-2; *nat*MX	(Jessop-Fabre et al., 2016)
pRS42N	TEF1p-natMX_ADH1t	(Taxis and Knop, 2006)
pCfB2904	XI-3 MarkerFree backbone	(Jessop-Fabre et al., 2016)
pRP005	pCfB2904; TEF1p-yEGFP3-ADH1t	Perruca Foncillas (2022)
pCfB2312	TEF1p-Cas9-CYC1t_Kan MX	(Jessop-Fabre et al., 2016)
pCfB2899	X-2 Markerfree backbone	(Jessop-Fabre et al., 2016)

pTL11	GAL1p-yEGFP-protease1-CYC1t_Leu2	Lindh (unpublished)
pTL21	Gal1p-yEGFP-protease2-CYC1t_Leu2	Lindh (unpublished)
pTL31	Gal1p-yEGFP-protease 3-CYC1t_Leu2	Lindh (unpublished)
pTL70	Gal1p-yEGFP-CYC1t_ Leu2	Lindh (unpublished)
pTL22	pCfB2899; Gal1p-yEGFP-protease2-CYC1t_Leu2	Lindh (unpublished)
pTL72	pCfB2899; Gal1p-yEGFP-CYC1t_ Leu2	Lindh (unpublished)
Strain Name	**Relevant Genotype**	**Reference**
TMBTL010 (Background strain)	CEN.PK113.16B; YEplac181	Lindh (unpublished)
TMB RP001	CEN.PK113-7D; XI-3::TEF1p-yEGFP3-ADH1t	Perruca Foncillas (2022)
CENPK113.16B	Matα;Leu2	Euroscarf.
TMBTL016	CEN.PK113.16B; pTL70	Lindh (unpublished)
TMBTL017	CEN.PK113.16B; pTL11	Lindh (unpublished)
TMBTL018	CEN.PK113.16B; pTL21	Lindh (unpublished)
TMBTL019	CEN.PK113.16B; pTL31	Lindh (unpublished)
TMBTL024	CEN.PK113.16B; pTL22	Lindh (unpublished)
TMBTL020	CEN.PK113.16B; pTL72	Lindh (unpublished)

2.3 Yeast Transformation

Plasmid transformation of background strain (TMBTL 010) with plasmids; pTL70, pTL11, pTL21 and pTL31 resulted in the strains TMBTL016, TMBTL017, TMBTL018, and TMBTL019 respectively. This was performed using high efficiency LiAc/ss carrier DNA/PEG method (20), with a modification in the protocol, where 40µL volume of Dimethyl Sulfoxide (DMSO) was added before heat shock treatment at 42°C (used for all transformation in this project).

Addition of DMSO showed enhanced permeability of the cells and this step has shown to increase the efficiency of the transformation process [22]. The transformed cells were plated on YPD media with different volumes (50 µL, 100 µL, 200 µL and remaining samples centrifuged for 10-12 seconds at 13500 rpm followed by 150 µL of the resuspended with pellet and plated) and sample without any plasmid was used as negative control. The transformation efficiency was calculated for all the yeast strains and the colonies obtained were used for further analysis. The number of colony forming units (CFU) was calculated and thereby the transformation efficiencies were found out using the equation [23]:

$$\text{Transformation Efficiency} = \frac{\text{CFU Formed X Volume added at recovery phase}}{\text{µg of plasmid DNA X Volume plated}}$$

CRISPR/Cas-9 genome editing tool protocol was designed according to Jessop-Fabre et al. (2016). *TMBTL016 and TMBTL018* were used for chromosomal integration studies. After transforming the cells, adequate amount colony from the plate was picked and inoculated for a pre-culture in 5 mL YPD supplemented with 5 µL of geneticin (g418) (200mg/L). The pre-culture was used to inoculate the 250 mL shake flask to an initial OD of 0.5 with the cultivation supplemented with 25µL of g418. The pCfB3045 gRNA plasmids (200mg/ml) and the 1µg donor DNA from plasmids (cut using restriction enzyme *Not*I) transformed into the yeast causes double stranded breaks at the XI-4 genomic locus. The transformed cells were plated on the YPD+g418+cloNAT in three different dilutions (100 µL, 200 µL and remaining samples centrifuged followed by 150 µL of the resuspended with pellet and plated) to provide selection for the Cas9 and gRNA plasmids. Three controls were also prepared with water, gRNA plasmid and plasmid pRS42N, where plasmid pRS42N was used as the positive control and the other two were negative controls. Plates were incubated for 3 days and the number of colonies forming units are found to calculate the transformation efficiency using the equation mentioned. The integrated strains were named TMBTL024 and TMBTL020 (refer *table-2*)

2.4. Yeast Growth and Induction

2.4.1 Cultivation in Falcon Tubes

50mL falcon tubes were used for the experiment and maximum volume of media used was 5mL (10% of the total volume). Five random clones/colonies were picked for each of the four strains, namely TMBTL017, TMBTL018, TMBTL019 and TMBTL016 from their respective transformed plates and were inoculated in 5mL YNB glucose for pre-culture. The inoculated pre-culture tubes were kept overnight in the shaking incubator at 30°C and 180 r.p.m. Pre-culture was used to inoculate the shake flask where the initial OD_{620nm} was set to 0.2. All the 20 tubes (5 colonies from each strain) were used for cultivation with each having 5mL total volume and was incubated in the incubator shaker at 30°C and 180 r.p.m. Optical density was measured at regular intervals of time.

Induction was performed using YNB Galactose media. After 24 hours, the cells containing YNB glucose was centrifuged at 4000 r.p.m for 5 minutes. The supernatant was discarded, and the pellet was resuspended in 25mL YNB galactose media. The tubes were kept at temperature

of 30°C and 120 rpm after resuspension and the cultivation progressed for 6 hours from time of induction. The cells were then harvested by centrifuging tubes at 4000 rpm for 5 minutes, followed by resuspending the pellet in 5 mL PBS (KCl, 10mM Na_2HPO_4,2mM KH_2PO_4). Serial dilution was performed on PBS containing cells obtained from the cultivation experiments and dilutions until the range of 1/32 was performed in Eppendorf tubes, which were transferred to 96-well plate. Fluorescence was measured (Thermolab systems Fluroscan FL microtiterplate reader) between an emission wavelength of 485nm and excited wavelength 538 nm and Optical density at 620nm (measured using Thermo lab systems 354 Multiskan Ascent microtiter reader) for the different dilution values.

2.4.2 Cultivation in Shake flask

Yeast cultivation was performed in 250 mL aerated baffled Erlenmeyer flasks and the shake flasks contained YNB glucose medium having a volume 10% of the total flask volume. Precultures for TMBTL017, TMBTL018, TMBTL019 and TMBTL016 was prepared which consisted of one colony (C1) picked from a plate and inoculated in 5mL of YNB glucose media. Replicate of the procedure was performed using another colony (C2) for all the four strains. The precultures were grown overnight in a shaking incubator at 30°C and 180 r.p.m and the cells from the preculture were then used to inoculate the shake flask to an initial OD_{620nm} of 0.2. The cultivation with 4 strains, each colony were grown in separate shake flasks was performed in incubator shaker at 30°C and 180 r.p.m. Samples were taken at regular interval to measure optical density at 620 nm wavelength using Ultrospec 2100 Pro spectrophotometer (Amersham Biosciences Corp., USA).

After 24 hours of cultivation, the glucose media was removed, and galactose was added to induce the GAL1 promoter. YNB glucose media containing the yeast cells were transferred into 50mL falcon tubes for four different strains used. Tubes were centrifuged at 4000 r.p.m for 5 minutes, Supernatant was discarded, and the pellet obtained was resuspended in fresh 25mL YNB galactose media. Upon resuspension, the tubes were kept in the incubator shaker and measurements of OD_{620nm} were taken until the 30^{th} hour of cultivation at 1-hour intervals. This was followed by harvesting the cells produced. For harvesting, the cells were pelleted by centrifugation of the falcon tubes at 4000rpm for 5 minutes, supernatant was discarded, and the cells were again resuspended in 5 mL PBS. Serial dilution was performed on PBS containing cells obtained until the range of 1/32 in Eppendorf tubes, which were transferred to 96-well plate. Fluorescence was measured (Thermolab systems Fluroscan FL microtiter plate reader) between an emission wavelength of 485nm and excited wavelength 538 nm and Optical density at 620nm (measured using Thermo lab systems 354 Multiskan Ascent microtiter reader) for the different dilution values.

The samples from the pre-culture of each strain and clone were also taken with equal volume of 50%(v/v) glycerol stock and stored in -80°C freezer.

2.5. HPLC

To analyze the formation of extracellular metabolites produced during cultivation, HPLC measurements were taken. 1mL of the samples from all the strains at different time intervals were taken in a sterile Eppendorf tube and centrifuged at 3200 rpm for 2 minutes. The supernatant obtained was transferred to a fresh Eppendorf tube and stored at -20°C until

analysis. Acetate, ethanol, galactose, glucose and glycerol levels were measured and analyzed using HPLC (Alliance 2695, Waters, Milford, MA, USA). A flow rate of 0.6mL/min was used, and each sample was run for a period of 30 minutes. Aminex HPX-87H column (300X7.8mm,9µm, BIO-RAD, Hercules, CA, USA) was used at a temperature of 60°C along with a mobile phase of 5mM H_2SO_4 and the metabolites were detected using a Refractive Index detector. The concentrations were calculated using an external eight-point calibration curve.

2.6 Flow Cytometry

Flow cytometry (BD Accuri C6 flow cytometer, Becton Dickinson) was used to determine the GFP expression in various yeast strains used in the study. Samples were taken from the shake flask experiment of each yeast strain and the controls. Samples were prepared with a total volume of 500µL in Eppendorf tubes having an OD value of 0.5 with a mixture of PBS and 1.32µg/mL Propidium iodide (PI). The tubes were placed in dark for an incubation time of 10 minutes, followed by the analysis in the flow cytometer using 533/30 band pass filter (commonly referred to as FL-1H channel) for GFP fluorescence measurements. To estimate the percentage of live and dead cells, 670LP bandpass filter (FL-3H channel) was used (25).

3. Results and discussion

3.1 Comparison Between Plasmid-Based systems and Chromosomal-Based systems

Plasmids are used frequently for construction of libraries in yeast but may be less stable due to plasmid copy number variation during growth or even loss of plasmid. Chromosomal integration may therefore be more suitable choice for the construction of a gene expression library. On the other hand, chromosomal integration is frequently mentioned to have less transformation efficiency than plasmid-based system. To shed light into this matter, a 2-micron plasmid carrying the LEU2 marker was transformed into a LEU2 auxotrophic strain and compared with a previously developed CRISPR/Cas9 system for chromosomal integration. To measure the effectiveness of cloning, transformation efficiencies were used as a comparison factor.

Strains TMBTL016, TMBTL017, TMBTL018, TMBTL019 were obtained by transforming the leucine auxotrophic strain CEN.PK113.16B with plasmids pTL70 (GFP), pTL11 (GFP-protease 1), pTL21 (GFP-protease 2) and pTL31 (GFP-protease 3) respectively. TMBTL017 (GFP-protease 1) and TMBTL019 (GFP-protease 3) showed transformation efficiency of $5.07x10^3$ and $5.2x10^3$ CFU per µg plasmid DNA, respectively. These values were lower than for TMBTL018 (GFP-protease 2) as this strain had a slightly increased value of $5.63x10^3$ CFU per µg plasmid DNA. Highest efficiency was shown by TMBTL016 (GFP) strain having a value of $5.8\ x10^3$ CFU per µg plasmid DNA.

For comparison with chromosomal integration strategy, CRISPR-Cas9 approach was adopted for the gene integration using Cas-9, gRNA, and two strains TMBTL 024 and TMBTL 020 were Chromosomally integrated at the XI-4 genomic locus. Transformation efficiency of $1\ x10^3$ CFU per µg plasmid were obtained for TMBTL 024 and TMBTL 024 had $0.7\ X\ 10^3$ CFU per µg plasmid obtained.

As expected, transformation efficiency using plasmid-based system showed the highest numbers when compared to the Chromosomal integration from the results obtained in this study (*figure 2*).

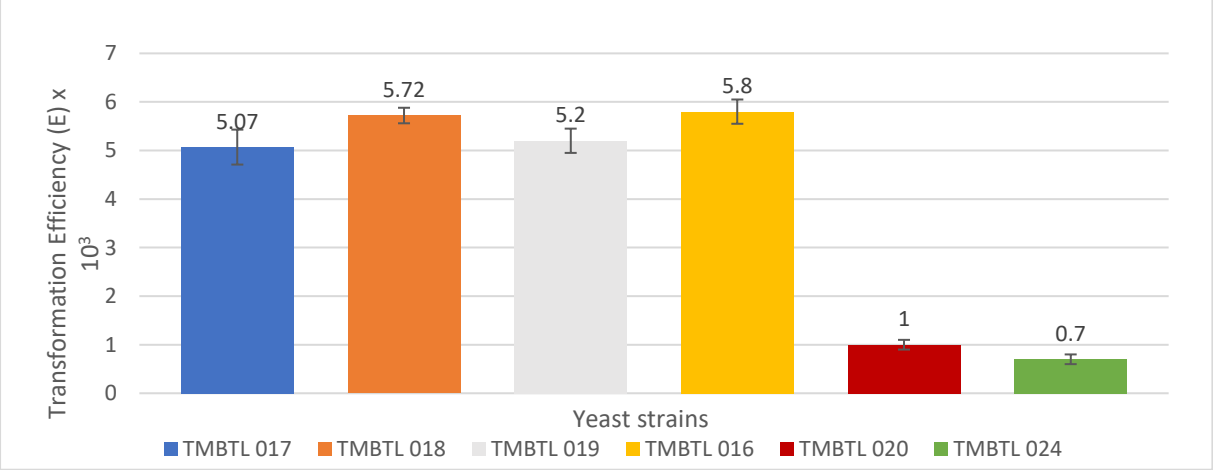

Figure 2: shows the bar plot between the transformation efficiencies of different strains. TMBTL 020 and TMBTL 024 are chromosomally integrated yeast strains and rest of the strains developed by plasmid-based systems.

3.2 Development of the slope methodology for screening yeast colonies

In this project, GFP-fusions are used as reporter tool to measure expression level of proteases. This technique enabled screening of multiple variants to identify suitable clones. However, a new protocol was needed to measure fluorescence of yeast cell transformants in a potential protease expression library in a high-throughput set-up. The protocol includes procedures for cell growth, medium exchange, induction of gene expression, dilution of yeast cells, and finally measurement of fluorescence and optical density (*figure 3*).

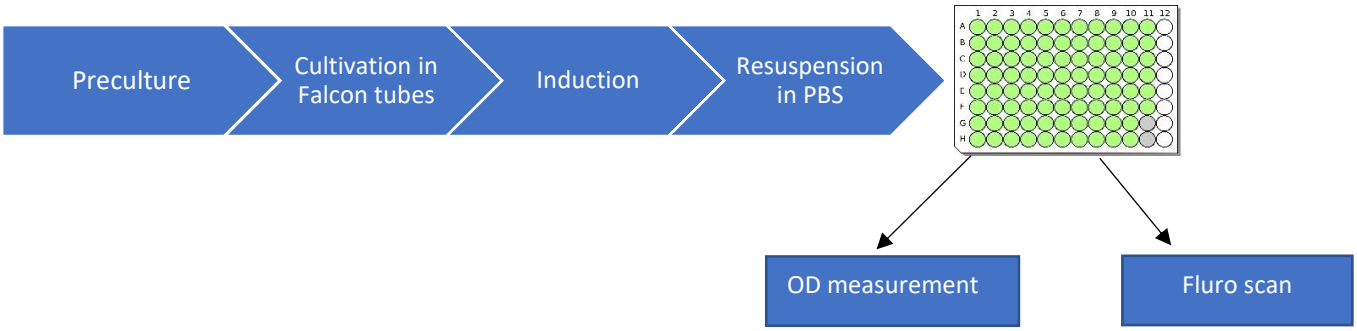

Figure 3: Pipeline for the slope methodology developed for screening of colonies in this experiment conducted. Steps in the pipeline involves pre-culture, cultivation in falcon tubes, Induction, resuspension in PBS, sampling in different dilutions on 96-well plate, analysing the plate on Fluro scan plate reader and spectrophotometer plate reader for obtaining OD measurement.

GFP-protease expression is induced by a GAL1 promoter, thus it relies on high galactose and low glucose in the medium since the promoter is repressed by glucose and induced by galactose. In presence of galactose the inactive genes are highly transcribed, resulting in a series of regulatory GAL gene expression. The transcription and translation lead to the production of recombinant protein.
Measuring GFP-fusion fluorescence in cell cultures depends on the density of the cell culture, which may vary depending on how well the strains have grown. To circumvent this, the fluorescence is typically normalised to the OD. Relying on a single measurement point makes the assay sensitive to differences in cell growth, and the value obtained (typically fluorescence divided by OD) may differ not be sufficient to distinguish strains with different specific GFP expression levels. To increase the resolution and enable more subtle differences to be distinguished, a serial dilution step and measurement of fluorescence at multiple different cell densities per strain/colony was included.

As a first step in the experiment, colonies from transformed plates were picked and inoculated in the YNB glucose medium. Inoculum from pre-culture was used to start the cultivation at an initial OD value of 0.2. When measuring the kinetics of GFP, highest level of protein production was achieved in stationary phase (after 24 hours of cultivation) [34]. Upon 24 hours, the YNB glucose media containing cells were removed and cells were then induced in YNB galactose media. Since yeast strains expressing different proteases may have different growth kinetics, which thus leads to different amount of glucose in the medium at the intended point of analysis, there was a need to exchange growth medium (YNB-glucose) with induction medium (YNB-galactose) to achieve induction. The exchange of medium was achieved by centrifugation and resuspending cells in new medium. The induction progressed for 6 hours as it has been previously shown that expression level peaks around 6th hour after induction; however, the time may be significantly longer depending on the protease [Lindh unpublished]. The supernatant from the YNB galactose removed after cultivation was taken up for fluorescence measurement for its possibility to be used directly for quantification. However, YNB galactose media showed high autofluorescence, affecting accuracy in the measurement of fluorescence and hence improper slope values. In addition, the cause of autofluorescence was due to the presence of flavin/riboflavin compound in YNB. PBS media was more suitable for measurements of the fluorescence as it emits low autofluorescence activity (*figure 4*). As a consequence, YNB galactose media was removed by centrifugation, followed by resuspension of cells in PBS.

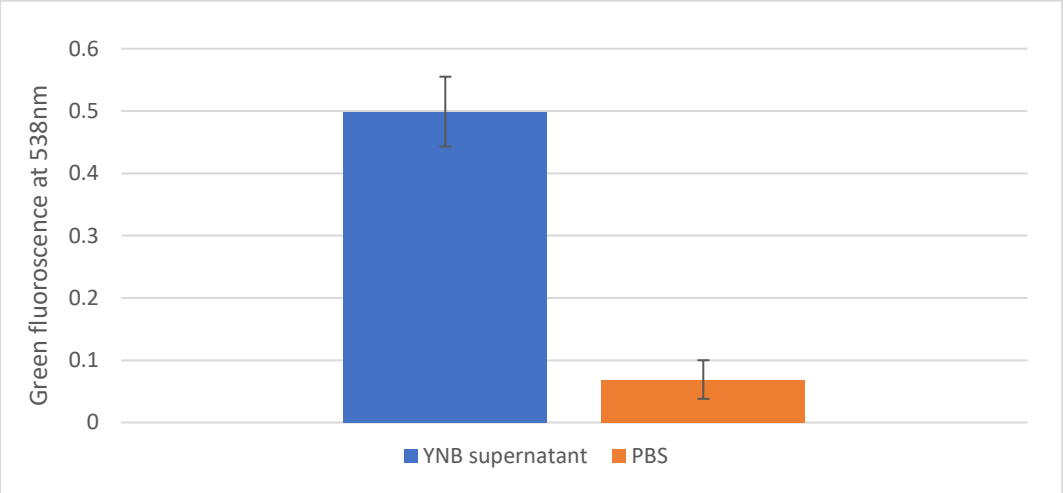

Figure 4: Fluorescence (ex./em. 485/538) of spent YNB galactose supernatant (blue) and PBS (orange) measured with the Fluoroscan.

The sample collected at a single time point was diluted in different concentration to measure fluorescence and OD values using Fluroscan plate reader and spectrophotometer plate reader respectively. The slope values obtained from the correlation curves when plotting fluorescence to OD was used as a measurement of GFP expression. The slope value for each curve was then calculated for each strain/condition. Hence the normalization approach was optimized and as a result this methodology was developed. Further, this method could be used as a tool to study and compare the protein expression of different yeast strains.

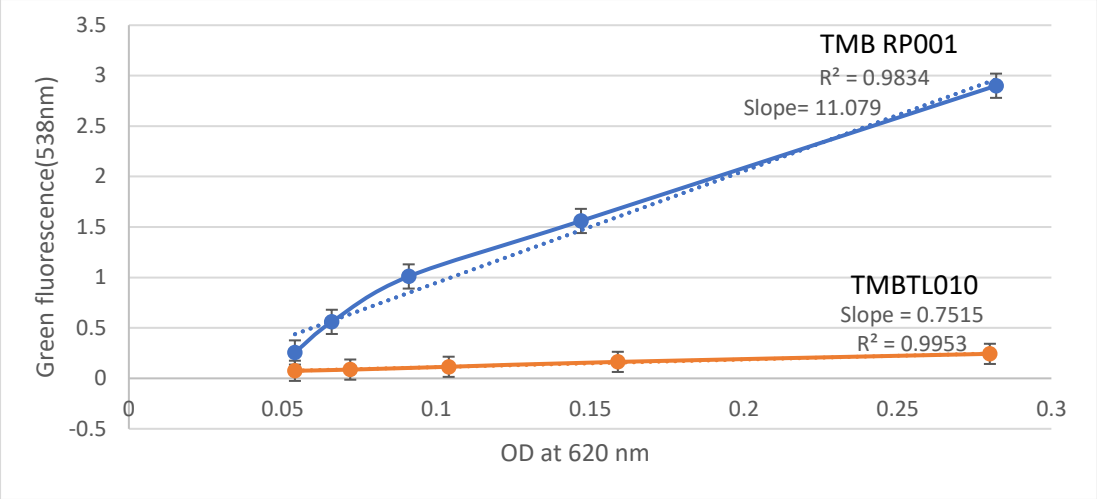

Figure-5: Fluorescence (ex./em. 485/538) vs. OD at 620nm for gfp positive strain (TMB RP001) (blue line) and gfp negative strain (TMBTL010) (orange line) measured with the Fluoroscan.

In a first attempt to explain the comparison of expressing cells, TMB RP001 strain (GFP positive) and TMBTL010 (GFP negative) were analysed using this methodology. However, both strains were not induced as mentioned in the pipeline (*figure 3*). This was because, constitutive promoter TEF1 was used in gene construct of TMBTL RP001, instead of GAL1 promoter and hence induction with galactose had no effect on the strain. The control strain TMBTL 010 had no GFP encoding genes in its construct nor protease expression. So, neither of the strains needed galactose induction.

With this procedure, different levels of GFP were clearly observed for strains with and without GFP expression (slope= 11.09 vs 0.75) (Figure 5). The correlation coefficient 0.9834 and 0.9953 for GFP positive and GFP negative strains respectively, are result of linear regression of the emitted fluorescence as a function of optical density at 620nm. This indicates that emitted fluorescence was directly proportional to the cells produced. Looking at TMB RP001, the gfp positive strain with highest slope value, there was a huge difference when compared to TMBTL 010. These results confirm high slope values for gfp positive cells and low slope values for gfp negative cells.

3.3 Evaluation of GFP-protease fusion expression levels

The above-mentioned protocol for evaluating fluorescence of yeast cells using fluorometry was applied to analyse potential differences between strains. The strains TMBTL 016(plasmid pTL70), TMBTL 017(protease1, plasmid pTL11), TMBTL 018 (protease2 plasmid pTL11) and TMBTL 019 (protease3 plasmid pTL31) express the genes under the control of the GAL1 promoter from a 2µ plasmid carrying an LEU2 marker gene for selection. TMBTL010 is the negative control without a protease or GFP gene in its construct. For the positive control (TMB RP001), a constitutive promoter TEF1 is used to control the expression of GFP which is chromosomally integrated. Induction step was performed here as strains had GAL gene encoded in their gene construct.

Five random colonies were used from each yeast strain in falcon tubes. The protocol was followed according to the pipeline (Figure 2). All the colonies showed similar slope values when compared with each other for their respective yeast strains. (Table-A1, appendix). The slope values of various yeast strains used are shown in *figures 6-7*. The figures describe the slope values for all the different strains used in the study. Two different strains, namely TMB RP001 strain (used as the positive control) and TMBTL 016 had high slope values, indicating high expression of gfp for both strains. TMBTL 019 showed the lowest slope value and therefore least gfp expression. Interestingly, the background strain (TMBTL 010) had higher slope value than TMBTL 019. TMBTL 018 also had lower slope value than the best strains and hence low protease expression when compared to TMBTL 016. The results obtained suggest low protease expression in strains expressing proteases 1 and 2. Overall, this indicates the inhibiting effect of some of the proteases on the expression.

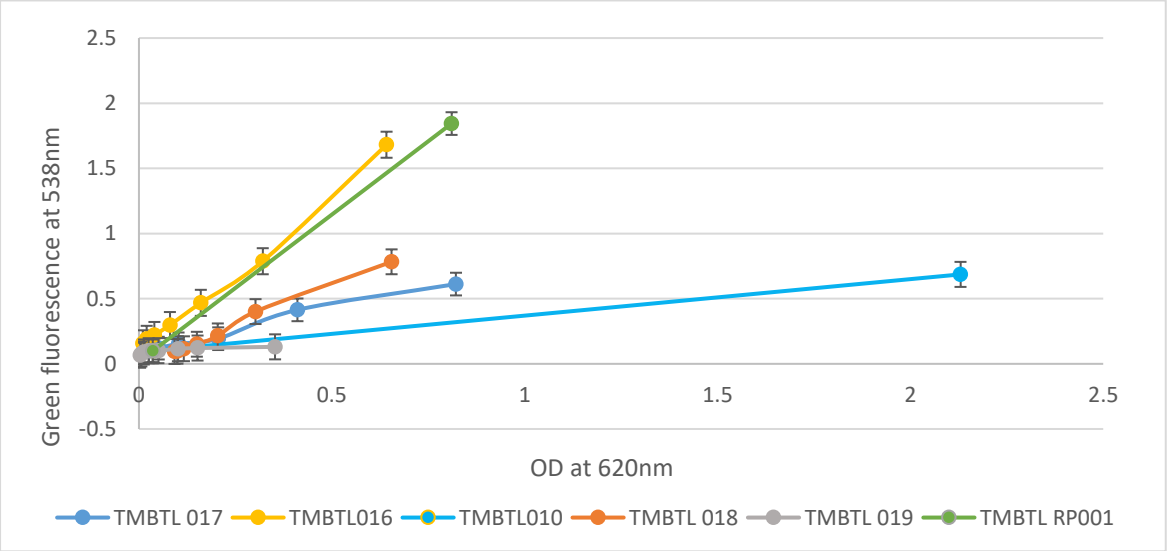

Figure-6: The plots between fluorescence measured between excitation wavelength of 485nm and emission wavelength of 538nm against optical density measured at 620nm for the different yeast strains and the controls TMBTL 016 (no protease, plasmid pTL70), TMBTL 017 (protease1, plasmid pTL11), TMBTL 018 (protease2, plasmid pTL11), TMBTL 019 (protease3, plasmid pTL31) and TMB RP001 (positive control). These readings mentioned are average values for five random colonies C1,

C2, C3, C4 and C5. These are the average values from two experimental replicates performed in falcon tubes.

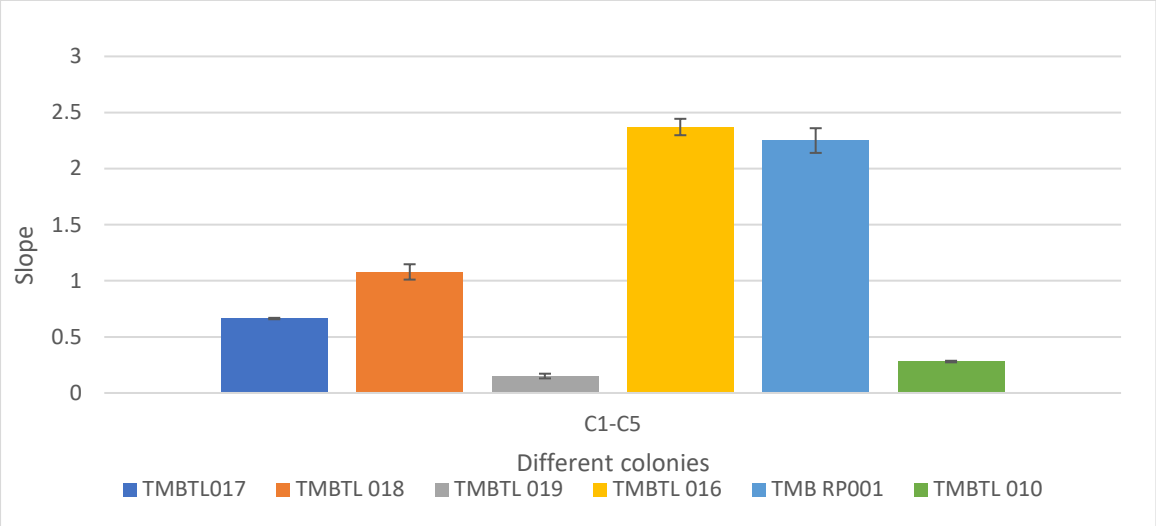

Figure-7: Bar plot mentioning the slope values obtained from Flu/OD plots for different strains and controls used plotted on y-axis and on x-axis are different clones/colonies C1, C2, C3, C4 and C5. These are the average values from several experimental replicates.

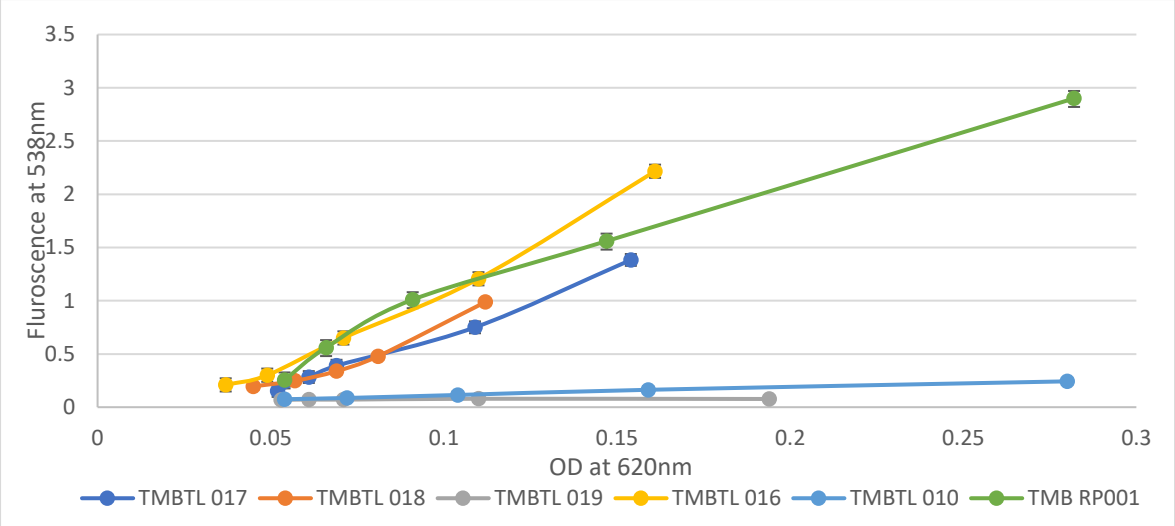

Figure-8: The plots between fluorescence measured at 485nm and 538nm against optical density measured at 620nm for the different yeast strains and the controls. These readings mentioned are average values from colonies C1 and C2. These are the average values from several experimental replicates performed in shake flasks.

Similar experiment was carried out in shake flasks and the slope values were compared with those obtained in the experiments in falcon tubes above. In shake flask experiments only two colonies were analysed for each strain. All the colonies showed similar slope values when compared with each other for their respective yeast strains. (Table-A4 and table-A5,

appendix). Figure-8 depicts the average value plotted for all the colonies used for each yeast strain.

The slope values, displayed in Figure 9, differed significantly from those in falcon tubes (Fig. 7), with increased value in shake-flasks (Fig. 7). For instance, the slope value denoted for TMBTL 017 in figure-9 was 13.6 whereas it was 0.664 in figure-7. This difference is explained by the varying cell densities in both the experiments. In addition, TMBTL 017 and TMBTL 018 had also shown high difference between both shake flask and falcon tube experiments. This arise because shake flasks experiment has more cells formed due to increase in cultivation volume and proper aeration. This shows that high difference of GFP expression can be obtained using slope methodology, depending on the cultivation conditions.

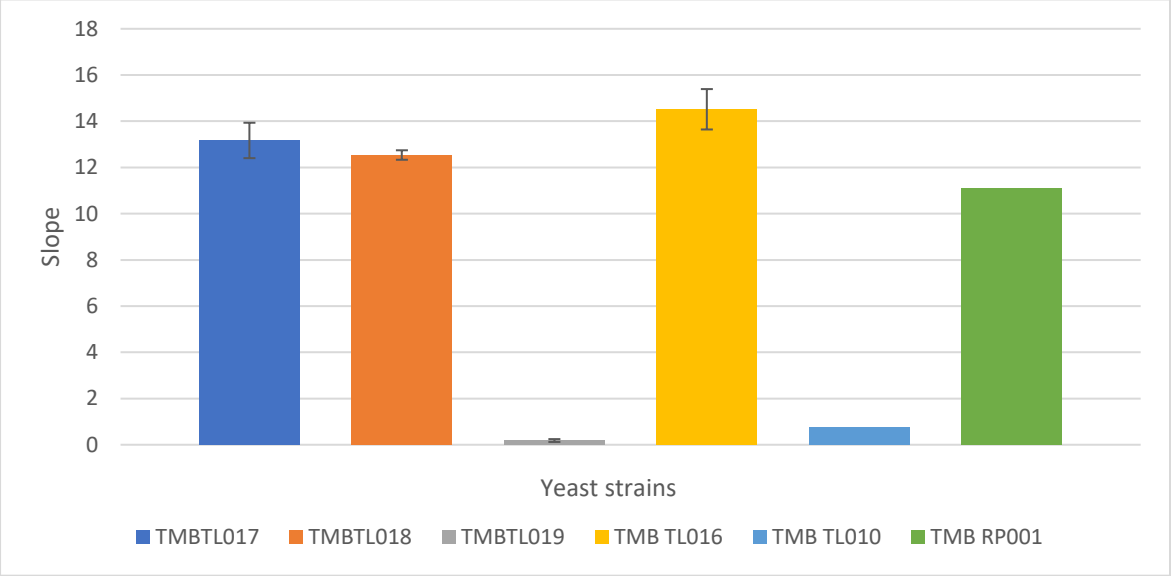

Figure-9: Bar plot mentioning the slope values obtained from Flu/OD plots for both colonies C1 as well as C2 plotted on y-axis and on x-axis are different yeasts strains used. These are the average values from two experimental replicates.

3.4 Flow Cytometry

3.4.1 GFP expression and comparison with slope method for different strains

Flow cytometry measures the fluorescence of individual cells using GFP encoded in its genetic construct. The analysis using flow cytometry was used to compare the GFP-protease fusion expression in various yeast strains (*Figure-10).* The highest mean fluorescence was measured in TMBTL 016 (Mean GFP = 23.5), even higher than the positive control, TMB RP001. Both strains have no protease gene in their genetic construct indicating that gfp expression can be affect by protease expression. The lowest fluorescence was shown by the negative control, i.e., TMBTL 010, indicating the level of autofluorescence, since gfp and protease gene are absent in its construct. TMBTL 018 and TMBTL 017 showed comparable mean fluorescent intensities having a value of 7.6 and 8.6, respectively (*table 3*).

The flow cytometry analysis revealed a significant population heterogeneity, with a large subpopulation displaying levels comparable to autofluorescence. The frequency of GFP positive cells, defined as having higher fluorescence than the negative control, varied between strains. TMBTL 017(protease1) had 55% gfp positive cells whereas TMBTL018 (protease2) had 52% of gfp positive cells. However, the fluorescence in the positive control TMB RP001 displayed a Gaussian distribution with more than 95% GFP positive cells, which can be explained by that GFP is stably integrated in the chromosome. The heterogeneity observed for the other strains is most likely explained by loss of plasmid in a subpopulation of the cells when compared with the other strains. Highest percentage of gfp positive cells were shown by TMB RP001 followed by TMBTL 016. TMBTL 019 had low values for both methods, indicating the protease in TMBTL 019 (Protease 3, plasmid pTL31) has higher inhibiting effect than other proteases. However, all the gfp-protease fusion yeast strains showed inhibition of protease expression.

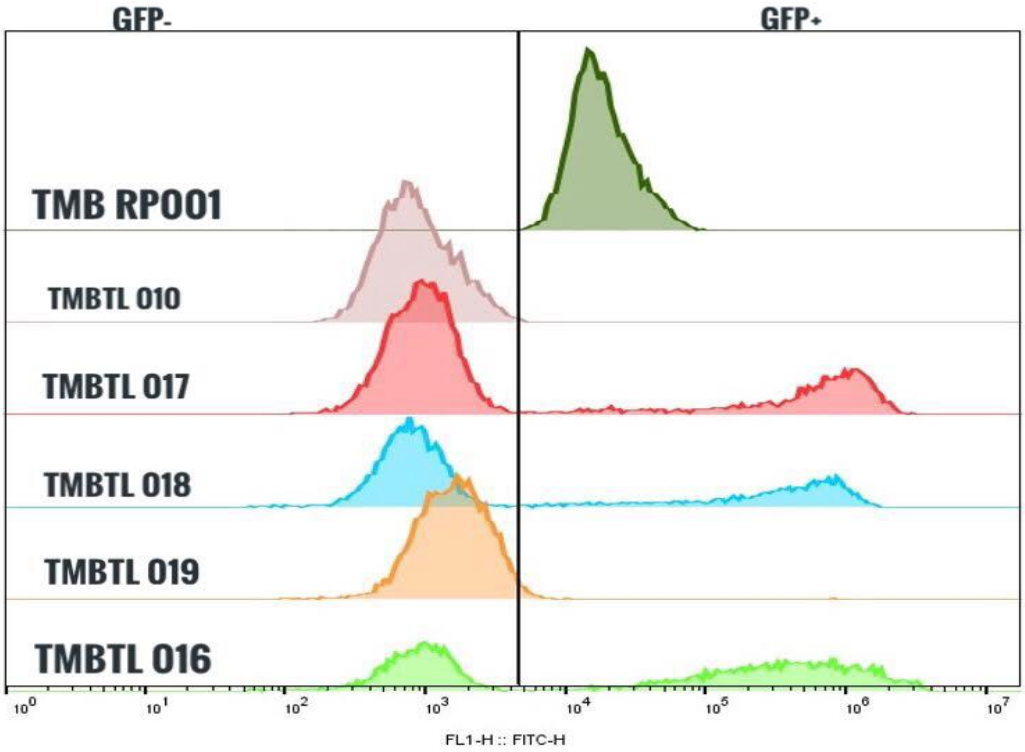

Figure 10: The histograms of flow cytometry from different yeast strains, namely (from top) TMB RP001, TMBTL 010, TMBTL 017, TMBTL 018, TMBTL 019, TMBTL 016. Black line represents autofluorescence of cells without GFP. Histogram show a single replicate performed that is representative of both technical and biological replicates performed.

Table 3: Displays the Mean fluorescence Intensity values of the total population, gfp positive values in percentage from the histogram obtained and slope method values shake flask experiment for different yeast strains.

Yeast strain	TMBTL 017	TMBTL 018	TMBTL 019	TMBTL 016	TMB TL010	TMB RP001
Mean fluorescence Intensity	8.6± 0.56	7.5± 0.7	0.38± 0.65	23.6± 1.2	0.09± 0.013	22.5± 0.9
GFP Positive %	55%	52%	19%	62%	5%	95%
Slope (gfp/od)	11.7	12.6	0.2	16.2	0.6	11.1

The flow cytometric mean fluorescence intensity measurements and the slopes obtained from the fluorometry method were compared (*table 3*). The gfp expression data obtained from both these experiments varied. However, it is visible from for both sets of values, strains with proteases in the construct displayed lower gfp expression.

3.4.2 Cell Viability of yeast strains

Using propidium iodide staining, flow cytometry was also used to monitor the cell viability. (Figure 11) after 30 hours of cultivation. This was performed to observe the percentage of live and dead cells for different strains. TMBTL 017 (protease 1) showed the lowest percentage of viable cells followed by TMBTL 018 (protease 2) having 71.6% and 86.2% respectively. Even though TMBTL 019 (protease 3) showed low performance in terms of growth, its viability was higher than TMBTL 016. Both control strains had high viability as well. These results give a glimpse of live cells at the end of 30 hours cultivation and all the strains had good percentage of viable cells.

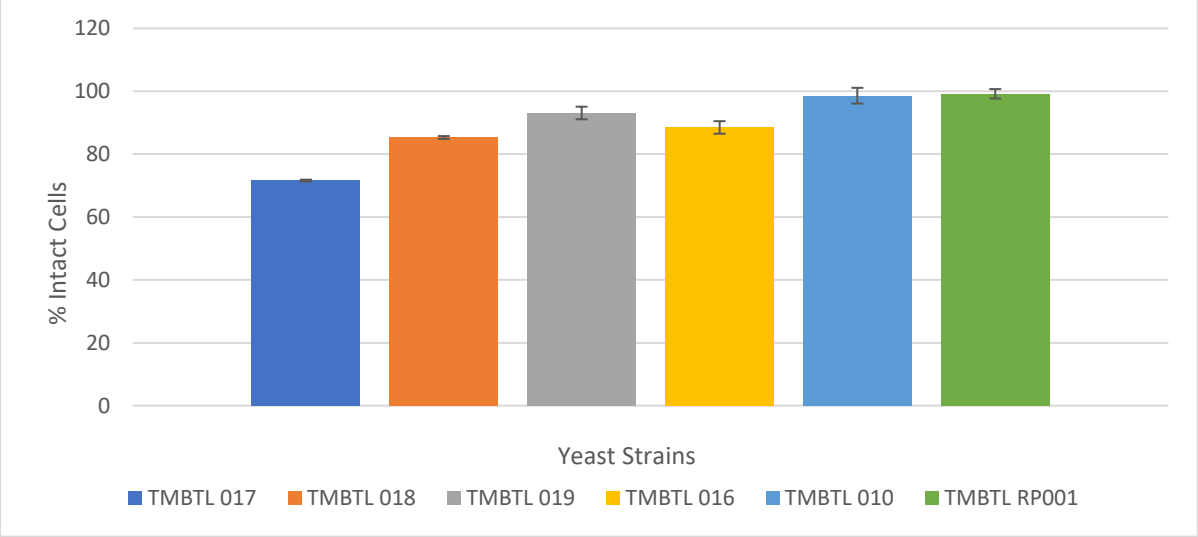

Figure-11: Percentage of viable cells for different yeast strains measured at 30th hour of cultivation.

3.5 Comparison of growth characteristics of yeast strains.

Since the recombinant proteases may hydrolyse yeast native proteins, they may pose a significant fitness burden to the cells. To study their potential effect on fitness, cell growth in shake flasks was monitored for all the genetically modified yeast strains (Figure 12-15). For TMBTL017 (protease 1), TMBTL 018 (protease 2), and TMBTL 019 (protease 3), the glucose concentration was shown to constantly decrease with an increasing optical density value measured at 620nm. TMBTL 018 did however not proliferate, and consumed neglectable amount of glucose. At 24 hours of cultivation, cells were induced with galactose by exchanging the growth medium (YNB glucose) to induction medium (YNB galactose). The next measurement taken at 40th hour of cultivation (16 hours after induction) in order to study the cell fitness showed an increase in the cell growth for all strains except TMBTL 019 (Protease 3).

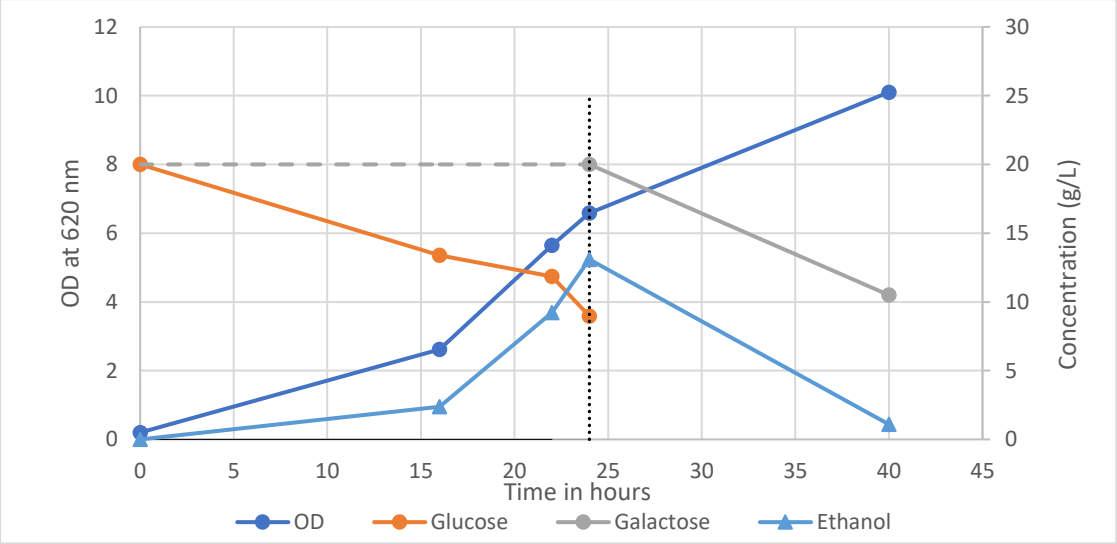

Figure 12: The growth rate of TMBTL 017 strain (protease1, plasmid pTL11) at an optical density measured at 620nm is depicted. Orange line represents the glucose concentration(g/L) for TMBTL 017, yellow line shows the galactose after induction. Black dotted line indicates time of induction.

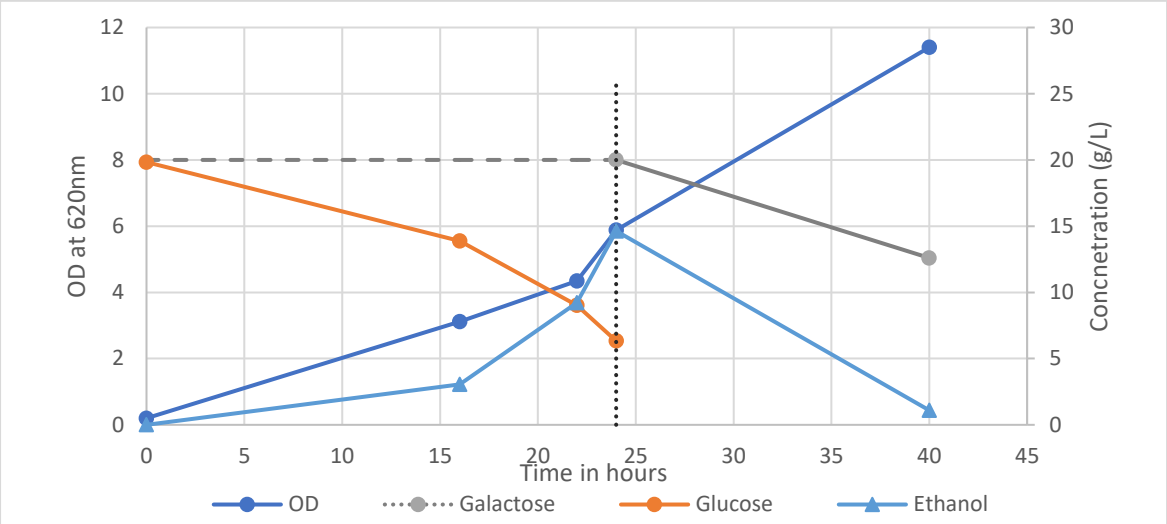

Figure 13: The growth rate of TMBTL 018 strain (protease2, plasmid pTL21) at an optical density measured at 620nm is depicted. Orange line represents the glucose concentration(g/L) for TMBTL 018. Black dotted line indicates time of induction.

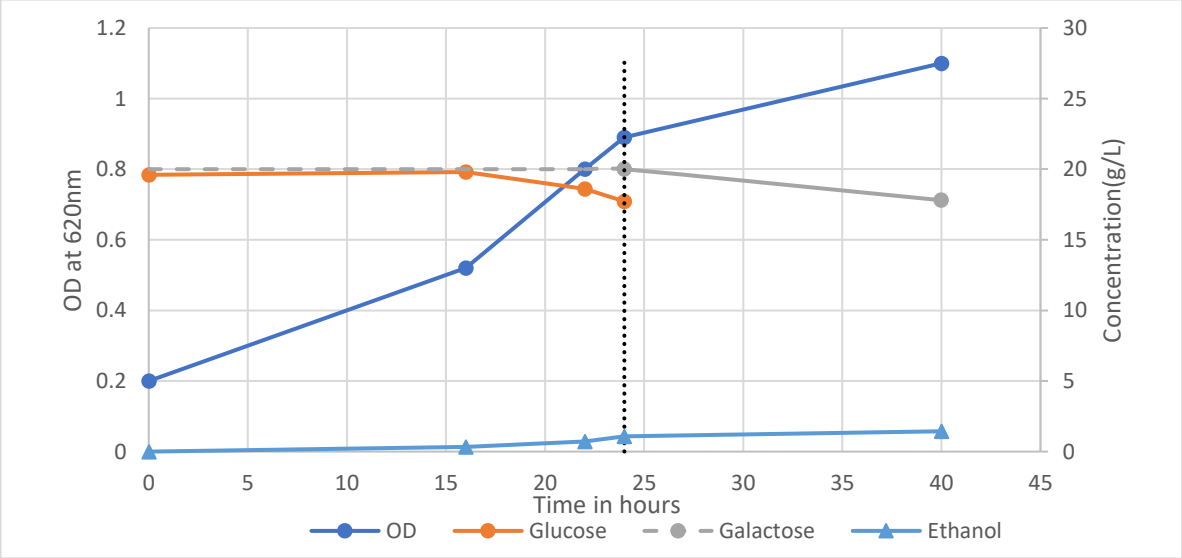

Figure 14: The growth rate of TMBTL 019 strain (protease1, plasmid pTL31) at an optical density measured at 620nm is depicted. Orange line represents the glucose concentration(g/L) for TMBTL 019.Black dotted line indicates time of induction.

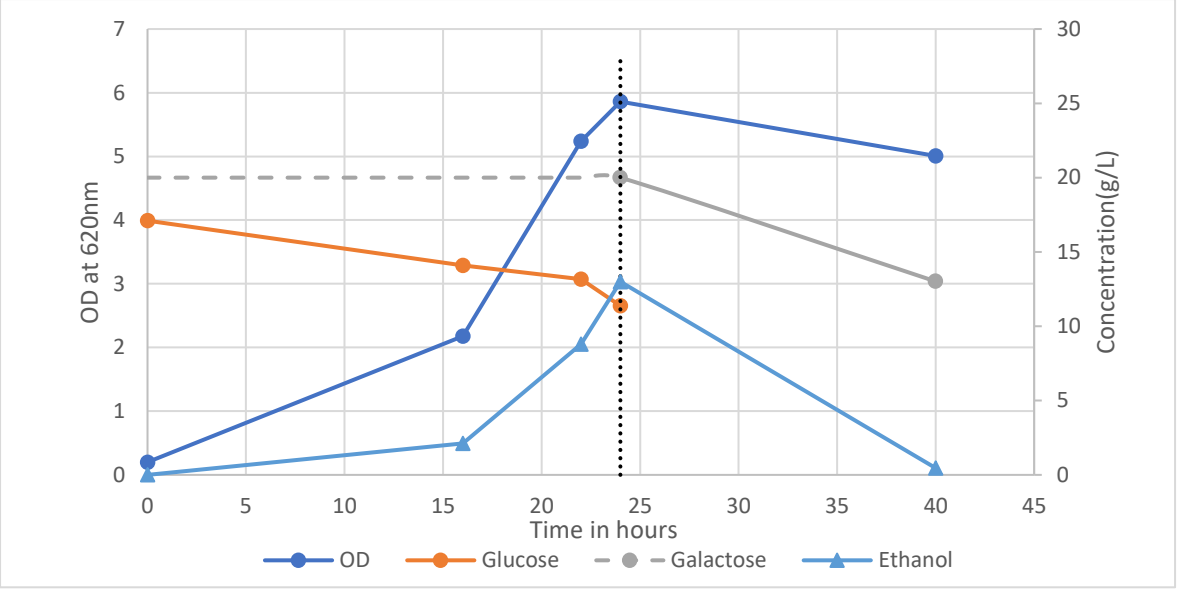

Figure 15: The growth rate of TMBTL 016 strain (plasmid pTL70) at an optical density measured at 620nm is depicted. Orange line represents the glucose concentration(g/L) for TMBTL 016. Black dotted line indicates time of induction.

The profile of TMBTL 018 showed a decline in the concentration of glucose as the time progressed with a corresponding observed growth. TMBTL019 had a unique growth pattern when compared to growth profile of other yeast strains. This strain had almost no cell growth throughout the cultivation, and this was assumed to be due to the produced protease that would have a toxic effect on the strain, thereby inhibiting its growth. Low consumption of glucose through-out the cultivation was shown as the final concentration after 24 hours of cultivation was 17.3g/L as this arises due to low growth rate of TMBTL019 strain as observed. The growth profile of TMBTL016 showed constant increase in the cell growth with glucose depleted to a final concentration of 12.6g/L at 24 hours. After induction, the growth was shown to be constant and a slight decrease was shown. This is possibly because that the strain had reached the stationary phase of its growth.

The galactose concentration also declined when measured at 40th hour of cultivation for all the yeast strains. Highest galactose consumption was indicated by TMBTL017 when measured at the 40th hour of cultivation (16th hour of induction). On the contrary, TMBTL 019 had least galactose consumption, where a concentration of 17.31g/L was remaining at the end of 40th hour of cultivation. In some of the strains, the growth rate using galactose as a sole carbon source was low, for example in TMBTL 016. This is assumed to arise due to high affinity of yeast strains for glucose or due to growth saturation. However, strains encoding proteases 1 and 2 had high optical density growth after induction.

Ethanol production for all strains except TMBTL 019 showed an increase right from the start of the cultivation as the optical density increased. From analysing results obtained, the glucose for all the strains was not entirely depleted after 24 hours of cultivation, that is when the media was changed for induction. Hence effect of ethanol on the cultivation was not further analysed. However, ethanol concentration at the end of 40th hour of cultivation drastically declined for all the strains. Other metabolites such as glycerol and acetate produced

were less relevant to this study performed. (figure A1 and A2, appendix). Overall, the metabolite profile of TMBTL 017, TMBTL 018 and TMBTL 016 were almost similar to each other, and this is probably because of these strains originating from the same background strain TMBTL 010.

3.7 TMBTL 019 strain extended growth.

To investigate the impaired growth of TMBTL 019, another set of growth experiments were performed on TMBTL 019 using YNB glucose and YNB galactose. The growth characteristics of the TMBTL 019 was studied until the 150th hour of the cultivation. Out of both shake flasks used in the experiment, one was induced at 48th hour of cultivation, wherein YNB glucose media was exchanged with YNB galactose, and the other shake flask was not induced, thereby the cultivation progressed with YNB glucose. Growth profile for both shake flasks was very similar (Figure 20). There showed an extended lag phase for the strain cultivated until 48 hours of cultivation, thereafter a constant increase in growth until 150th hour for both the shake flasks. The OD value reached as high as 25 by the end of the cultivation for induced as well as non-induced cells.

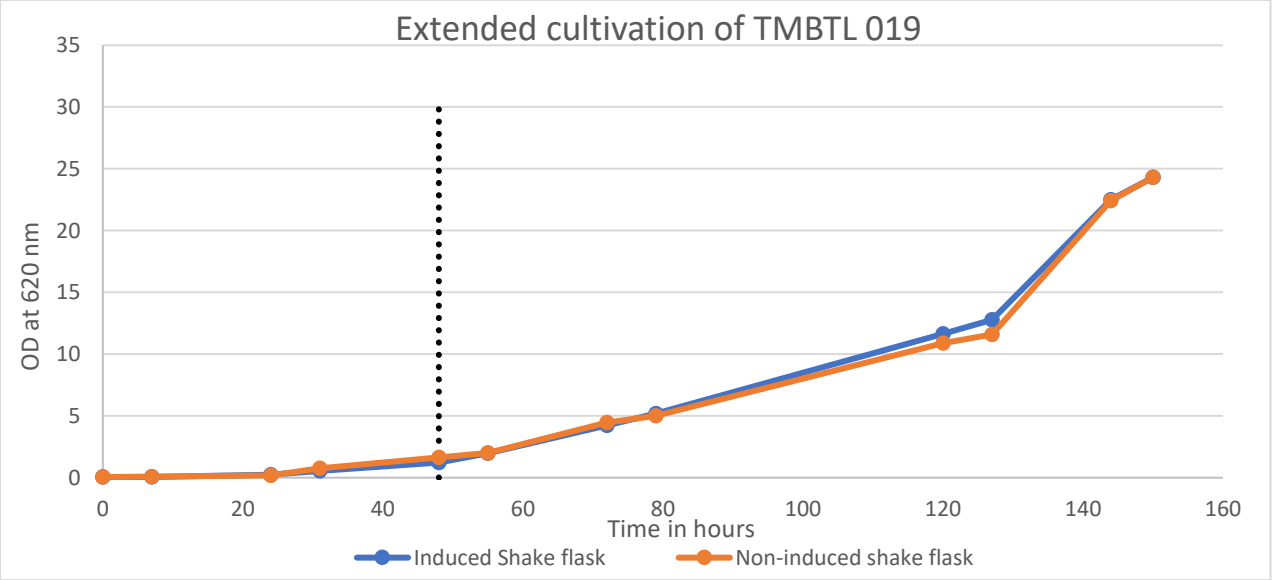

Figure 16: Shows growth curve for TMBTL 019 for cultivation time of 150 hours with OD at 620nm at y-axis and time(hours) in x-axis. Black dotted line represents the time of induction.

To see if the cells had lost the plasmid, and dead cell debris were sufficient to sustain leucine needed for the auxotrophic yeast, 5µl of sample obtained at 150th hour of cultivation from each of the shake flasks were inoculated in both YPD and YNB plates. (Figure-17). Significantly more growth was indeed observed in the YPD plate compared to the YNB plate (without amino acids). There were however colonies formed on the YNB plates for both the shake flask samples. The growth on YNB plate indicates the presence of leucine during cultivation obtained due to cell lysis through which the autotrophic yeast survives. On the contrary, massive growth obtained on YPD plates was due to the presence of amino acids in the plate,

hence absence of selection pressure. Moreover, there were less clones noted in YNB plate with induced shake flask sample and comparatively large number in non-induced YNB plate.

Altogether, the low growth of TMBTL 019 strain is accounted to be due to the protease (protease3) that gives an inhibiting effect on the yeast strain. Assumption of cell lysis occurring resulting in the Leucine in the medium enable the cells to grow in the YNB medium lacking amino acids. Also, the results from the YPD and YNB plate points out the plasmid loss from the cells, which is why the cells begin to grow. Hence, from the results obtained, plasmid-based system strategy for TMBTL 019 is highly unstable due to plasmid loss as the cells grow. Alternatively, chromosomal-integration strategy should be adopted as it provides better stability for TMBTL 019. However, further experiments are required to confirm this.

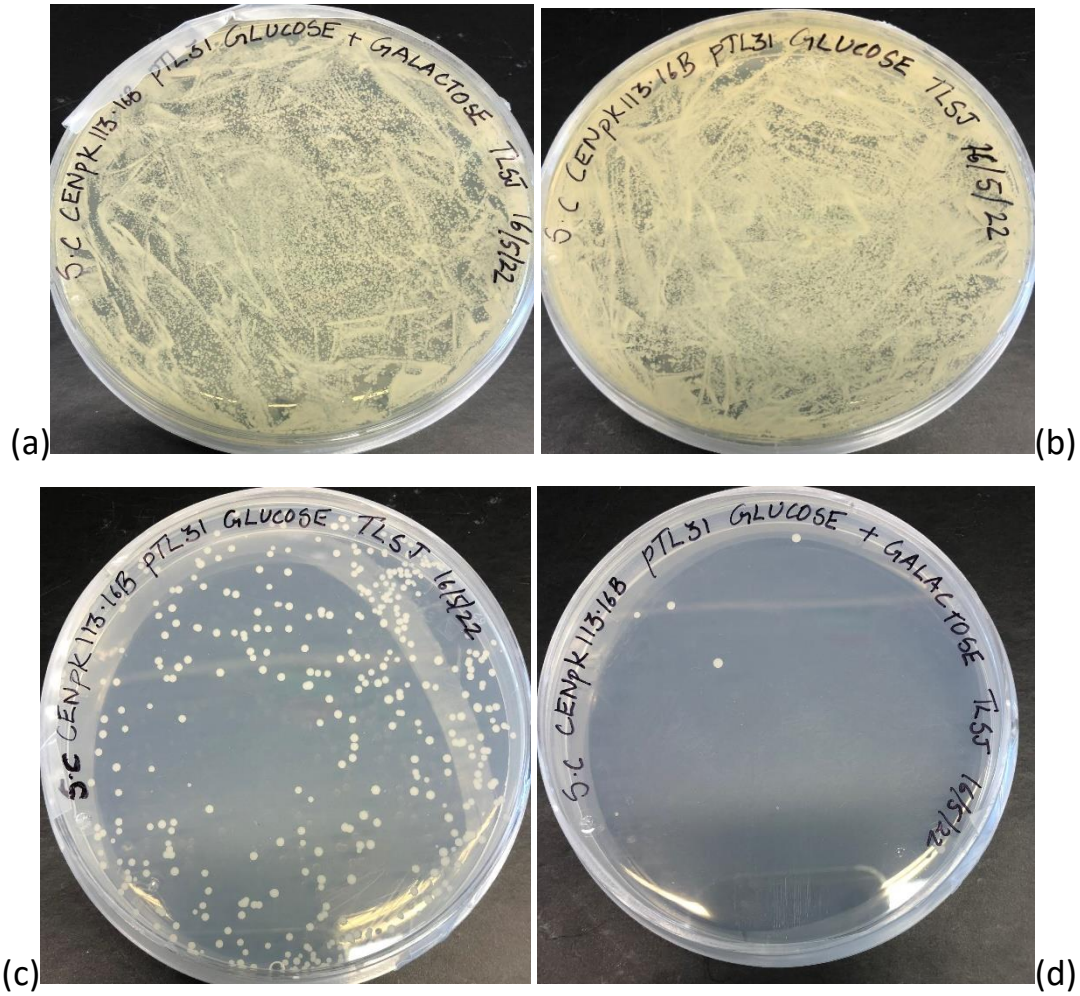

Figure-17: depicts the colonies formed on different plates from samples of pTL31 strain collected at 150th hour of cultivation. (a) shows colonies formed from sample induced with galactose streaked on a YPD plate, (b) shows colonies formed from sample with only glucose on YPD plate, (c) sample with only glucose plated on a YNB plate (without amino acids), (d) colonies formed from sample induced with galactose streaked on a YNB plate (without amino acids).

4.Conclusion

Comparison of plasmid-based systems with chromosomal integration systems indicated differences in transformation efficiency between both strategies. The chromosomal integration technique was shown to produce lower transformation efficiency compared to plasmid-based systems for the yeast strains used in the study.

New experimental protocol referred to as slope methodology was developed to screen the GFP expression level of the yeast colonies in microtiter plates based on fluorometry and spectrometry, to measure fluorescence and optical density, respectively. An attempt to optimize the methodology for screening of different clones was successfully performed. Screening of GFP fusion libraries in future research work could be performed using the developed protocol based on fluorometry and spectrophotometry. The slope values of correlation curves provides the gfp expression of the yeast strains and this value could be used as a basis for comparison. However, the difference between experiments where cells were grown in falcon tubes and shake flasks indicates that the gfp expression may be influenced by the different growing conditions. This is assumed to be due to the difference in cells produced in both the growing conditions (falcon tube and shake flask). The protease expression obtained on flow cytometry was compared to the slope method and is another method that may be applied for screening. It should though be noted that the flow cytometry analysis indicated a significant heterogeneity in the plasmid-based system, which may result in issues with both production and possible loss of genetic variation of a potential gene library. TMBTL 019 expressing protease 3 showed impaired growth due to high inhibiting effect (toxic effect) on the yeast strain. Experiments were performed to understand the impaired growth of TMBTL 019 and conclusive result regarding plasmid loss from strain was observed after long time of cultivations. TMBTL 018 and TMBTL 017 was shown to have protease inhibition, but lower than TMBTL 019. TMBTL 016, which lacked protease gene in its construct was found to have high GFP expression. These results indicate that recombinant proteases can have a detrimental effect on yeast fitness thereby restricting the possibility to build a broad specificity protease gene expression library.

5.References

[1] Harish, B. S., & Uppuluri, K. B. (2018). Microbial serine protease inhibitors and their therapeutic applications. *International journal of biological macromolecules*, *107*, 1373-1387.

[2] Guerrero, J. L., Daugherty, P. S., & O'Malley, M. A. (2017). Emerging technologies for protease engineering: new tools to clear out disease. *Biotechnology and Bioengineering*, *114*(1), 33-38.

[3] Yofe, I., Weill, U., Meurer, M. *et al.* One library to make them all: streamlining the creation of yeast libraries via a SWAp-Tag strategy. *Nat Methods* **13,** 371–378 (2016). https://doi.org/10.1038/nmeth.3795

[4] Baghban, R., Farajnia, S., Rajabibazl, M., Ghasemi, Y., Mafi, A., Hoseinpoor, R., ... & Aria, M. (2019). Yeast expression systems: overview and recent advances. *Molecular biotechnology*, *61*(5), 365-384.

[5] Jessop-Fabre, M. M., Jakočiūnas, T., Stovicek, V., Dai, Z., Jensen, M. K., Keasling, J. D., & amp; Borodina, I. (2016). EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae via CRISPR-Cas9.Biotechnology journal, 11(8), 1110-1117.

[6] Weill, U., Yofe, I., Sass, E., Stynen, B., Davidi, D., Natarajan, J., ... & Schuldiner, M. (2018). Genome-wide SWAp-Tag yeast libraries for proteome exploration. *Nature methods*, *15*(8), 617-622.

[7.] Sung, M. K., Lim, G., Yi, D. G., Chang, Y. J., Yang, E. B., Lee, K., & Huh, W. K. (2013). Genome-wide bimolecular fluorescence complementation analysis of SUMO interactome in yeast. *Genome research*, *23*(4), 736-746.

[8] Tarassov, K., Messier, V., Landry, C. R., Radinovic, S., Molina, M. M. S., Shames, I., ... & Michnick, S. W. (2008). An in vivo map of the yeast protein interactome. *Science*, *320*(5882), 1465-1470.

[9.] Almeida, L. D., Silva, A. S. F., Mota, D. C., Vasconcelos, A. A., Camargo, A. P., Pires, G. S., ... & Bilsland, E. (2021). Yeast Double Transporter Gene Deletion Library for Identification of Xenobiotic Carriers in Low or High Throughput. *Mbio*, *12*(6), e03221-21.

[10] Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., & O'Shea, E. K. (2003). Global analysis of protein localization in budding yeast. *Nature*, *425*(6959), 686-691.

[11] Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., ... & Johnston, M. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *nature*, *418*(6896), 387-391.

[12] Rizvi, S. M. A., Prajapati, H. K., & Ghosh, S. K. (2018). The 2 micron plasmid: a selfish genetic element with an optimized survival strategy within Saccharomyces cerevisiae. *Current genetics*, *64*(1), 25-42.

[13] Pronk, J. T. (2002). Auxotrophic yeast strains in fundamental and applied research. *Applied and environmental microbiology*, *68*(5), 2095-2100.

[14] Thak, E. J., Yoo, S. J., Moon, H. Y., & Kang, H. A. (2020). Yeast synthetic biology for designed cell factories producing secretory recombinant proteins. *FEMS Yeast Research*, *20*(2), foaa009.

[15] Davidsson, J. (2020). Assessment of the Fluorescent Activities of Different Fluorescent Proteins for Potential Use as Biosensors in Saccharomyces cerevisiae.

[16] Bai Flagfeldt, D., Siewers, V., Huang, L., & Nielsen, J. (2009). Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast*, *26*(10), 545-551.

[17] Stovicek, V., Borja, G. M., Forster, J., & Borodina, I. (2015). EasyClone 2.0: expanded toolkit of integrative vectors for stable gene expression in industrial Saccharomyces cerevisiae strains. *Journal of Industrial Microbiology and Biotechnology*, *42*(11), 1519-1531.

[18] Jessop-Fabre, M. M., Jakočiūnas, T., Stovicek, V., Dai, Z., Jensen, M. K., Keasling, J. D., & Borodina, I. (2016). EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae via CRISPR-Cas9. *Biotechnology journal*, *11*(8), 1110-1117.

[19] Jensen, N. B., Strucko, T., Kildegaard, K. R., David, F., Maury, J., Mortensen, U. H., ... & Borodina, I. (2014). EasyClone: method for iterative chromosomal integration of multiple genes Saccharomyces cerevisiae. *FEMS yeast research*, *14*(2), 238-248.

[20] Gietz, R. D., & Woods, R. A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. In *Methods in enzymology* (Vol. 350, pp. 87-96). Academic Press.

[21] Gietz, R. D., & Schiestl, R. H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature protocols*, *2*(1), 31-34.

[22] Hill, J., K.A. Donald, and D.E. Griffiths, DMSO-enhanced whole cell yeast transformation. Nucleic acids research, 1991. 19(20): p. 5791-5791.

[23] Duan, X., Ma, X., Li, S., & Zhou, Y. J. (2019). Free fatty acids promote transformation efficiency of yeast. *FEMS Yeast Research*, *19*(7), foz069.

[24] Mazumder, M., & McMillen, D. R. (2014). Design and characterization of a dual-mode promoter with activation and repression capability for tuning gene expression in yeast. *Nucleic acids research*, *42*(14), 9514-9522.

[25]. Persson, V. (2020). Evaluating Sugar Signaling during Xylose Oxidation in Saccharomyces cerevisiae.

[25]. Jessop-Fabre, M. M., Jakočiūnas, T., Stovicek, V., Dai, Z., Jensen, M. K., Keasling, J. D., & Borodina, I. (2016). EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae via CRISPR-Cas9. *Biotechnology journal*, *11*(8), 1110-1117.

[26] De La Cruz, J., Iost, I., Kressler, D., & Linder, P. (1997). The p20 and Ded1 proteins have antagonistic roles in eIF4E-dependent translation in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences*, *94*(10), 5201-5206.

[27] Foncillas, R. P., Davidsson, J., Carlquist, M., & Gorwa-Grauslund, M. F. (2022). Assessment of fluorescent protein candidates for multi-color flow cytometry analysis of Saccharomyces cerevisiae. *Biotechnology Reports*, e00735.

[28] Entian, K. D., & Kötter, P. (2007). 25 yeast genetic strain and plasmid collections. *Methods in microbiology*, *36*, 629-666.

[29] Yu, S. C., Dawson, A., Henderson, A. C., Lockyer, E. J., Read, E., Sritharan, G., ... & Spanu, P. D. (2016). Nutrient supplements boost yeast transformation efficiency. *Scientific reports*, *6*(1), 1-7.

[30] Siddiqui, M. S., Choksi, A., & Smolke, C. D. (2014). A system for multilocus chromosomal integration and transformation-free selection marker rescue. *FEMS yeast research*, *14*(8), 1171–1185.

[31] Zhang, J. H., Chung, T. D., & Oldenburg, K. R. (1999). A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *Journal of biomolecular screening*, *4*(2), 67–73.

[32] Hoepfner, D., Helliwell, S. B., Sadlish, H., Schuierer, S., Filipuzzi, I., Brachat, S., ... & Movva, N. R. (2014). High-resolution chemical dissection of a model eukaryote reveals targets, pathways and gene functions. *Microbiological research*, *169*(2-3), 107-120.

[33] Sheng, J., Flick, H., & Feng, X. (2017). Systematic optimization of protein secretory pathways in Saccharomyces cerevisiae to increase expression of hepatitis B small antigen. *Frontiers in microbiology*, *8*, 875.

[34] Štagoj, M. N., Comino, A., & Komel, R. (2005). Fluorescence based assay of GAL system in yeast Saccharomyces cerevisiae. *FEMS microbiology letters*, *244*(1), 105-110.

I. Appendix

Table-A1: Contains the slope values obtained from flu/OD (figure-8) for each yeast strain and control for colony-1(C1). These are the average values from several experimental replicates.

Strains	pTL11	pTL21	pTL31	pTL70	CENPK113.16B	RP001
Slope	11.637	12.125	0.058	16.265	0.7515	11.09

Table-A2: Contains the slope values obtained from flu/OD (figure-8) for each yeast strain and control for colony-1(C4). These are the average values from several experimental replicates.

Strains	pTL11	pTL21	pTL31	pTL70	CENPK113.16B	RP001
Slope	13.168	12.537	0.1813	14.517	0.7515	11.079

Table-A3: Contains the slope values obtained from flu/OD (figure-6) for each yeast strain and control for colony C1, C2, C3, C4 and C5. These are the average values from several experimental replicates.

Strains	Colony1	Colony2	Colony3	Colony4	Colony5
pTL11	0.6634	0.6639	0.6478	0.6568	0.06634
pTL21	1.0791	1.275	1.134	1.143	1.1002
pTL31	0.151	0.2113	0.176	0.184	0.2015
pTL70	2.3706	2.40	2.25	2.2067	2.3367
TMB RF001	2.25	2.34	2.56	2.24	2.29
CENPK113.16B	0.276	0.287	0.298	0.287	0.290

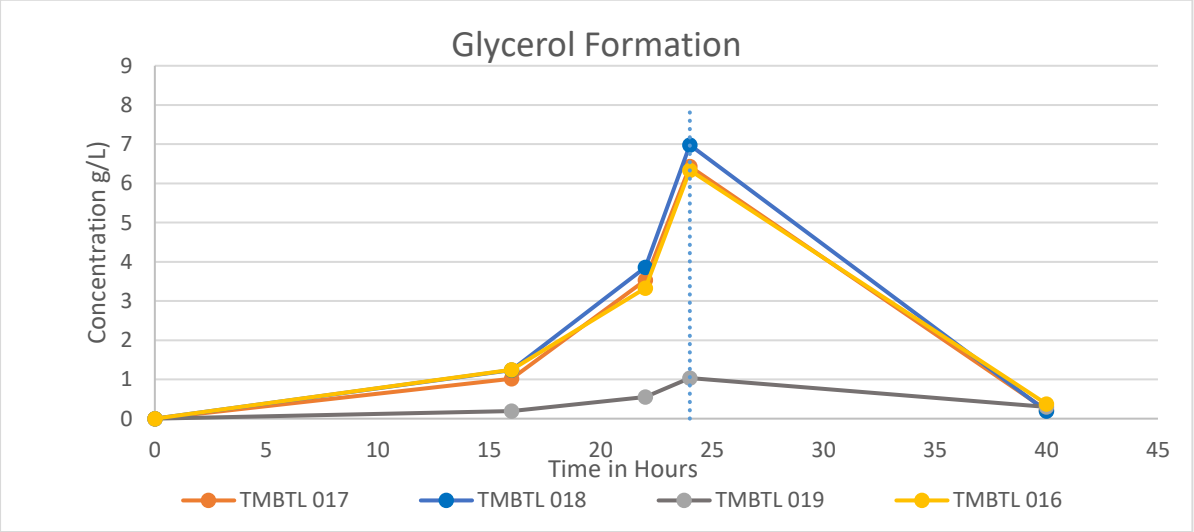

Figure A1: Depicts the glycerol formation of all yeast strain during cultivation measured at different time intervals. Blue dotted line represents the time of induction.

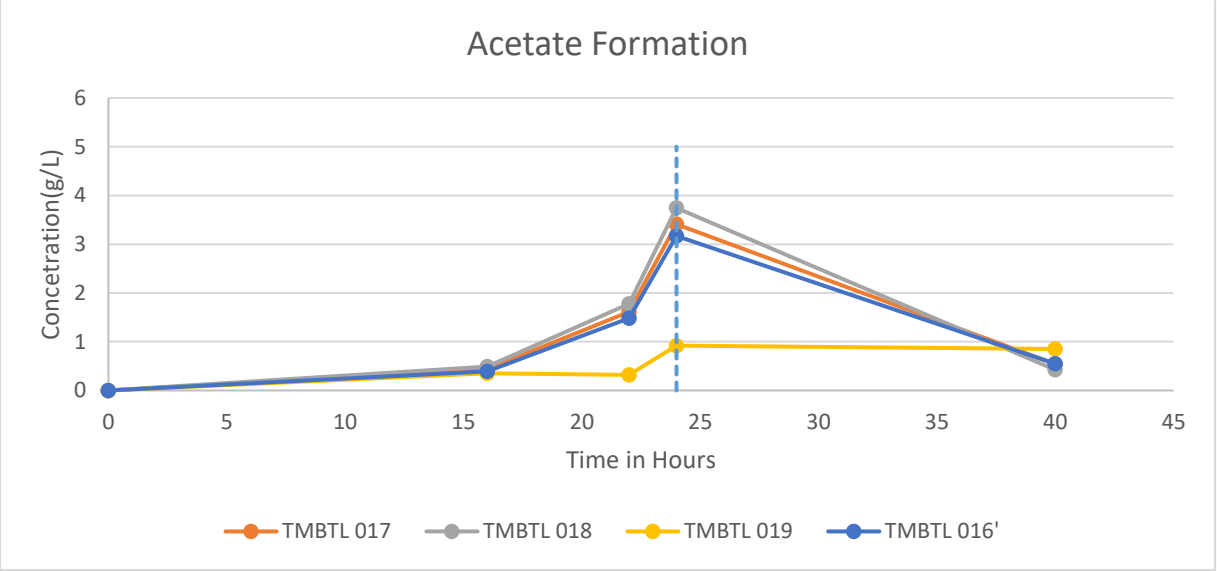

Figure A2: Depicts the acetate formation of all yeast strain during cultivation measured at different time intervals. Blue dotted line represents the time of induction.

Protocol for slope methodology

Step1: colony picked from a plate and inoculated in 5mL of YNB glucose media.

Step2: The precultures were grown overnight in a shaking incubator at 30°C and 180 r.p.m and the cells from the preculture were then used to inoculate the shake flask to an initial OD_{620nm} of 0.2.

Step3: Cultivation performed in a shaking incubator at 30°C and 180 r.p.m for 24 hours.

Step 4: YNB glucose media containing the yeast cells were transferred into 50mL falcon tubes. Tubes were centrifuged at 4000 r.p.m for 5 minutes, Supernatant was discarded, and the pellet obtained was resuspended in fresh 25mL YNB galactose media. The falcon tubes were kept in the incubator shaker and measurements of OD_{620nm} were taken until the 30th hour of cultivation.

Step5: Tubes were centrifuged at 4000 r.p.m for 5 minutes, Supernatant was discarded, and the pellet obtained was resuspended in fresh 25mL YNB galactose media

Step 6: Resuspension of cells in 5 mL PBS and the dilution series were made in 96-well plate

Step7: Measuring the plate in fluroscan plate reader and spectrophotometer plate reader

Step8: Fluroscence values were plotted against OD value to obtain curves and the slope of the curve is obtained and compared.

II. Popular Science summary

Yeasts is used in a lot of day-to-day applications including in various baking and beverages products. But how many of us know its utilization in pharmaceutical industry for protein production like protease? Unfortunately, optimal production of protease in yeasts still can be improved and strategies need to be developed to produce protease with higher yield. My master's project involves evaluation of such strategies expression of protease in yeast strain. Also, my project is part of a bigger project that aims to figure out the possibility of finding novel enzymes which have specific function and applications in various industries.

Yeasts *Saccharomyces cerevisiae* known as brewer's yeast has been widely used in food industry and is an important ingredient in baking as well as brewing products. However, these organisms also have applications pharmaceutical and molecular biology sector. In addition, due to ease in genetic manipulation of yeast, they are used as an excellent host for expressing specific proteins. These proteins formed are referred to as recombinant proteins since they are formed by genetic manipulation and have wide range of application including in synthesis of pharmaceutical products, enzymes for treatment of diseases and utilization in drug delivery systems.

Proteases have wide range of application in food, detergent, and pharmaceutical industries. In this project the expression of proteases in yeast was performed with the help of GFP-based fusion technology. Let's understand what this term is! so green fluorescent protein (GFP) is isolated from jellyfish and are actively used in research studies. This has a unique ability to absorb blue light and re-emit green light, producing bioluminescence. The gene coding for GFP in combination with gene that produces protease is fused together (hence called GFP-fusions) and is inserted into yeast cell with the help of different strategies. Since GFP produces fluorescence, by measuring it we could infer that the protease gene is being expressed by the yeast cells. In my master's thesis project, different approaches/strategies to express protease for a collection of different yeast strains was evaluated. GFP-fusions having different protease gene encoded, were used in this study to evaluate the strategies for high protease expression. Our results showed that, when different strategies were adopted, different protease expression was observed in each GFP-fusion protease, meaning that, each protease showed varying expression on the yeast strains. However, this work could be an aid when constructing collection of genetically modified yeast strains (called genetic library) which express protease for future research work. Furthermore, the results obtained represent an important step towards building a platform for producing novel enzymes with high specificity.