

**Directed Evolution of *Escherichia coli* LacZ gene to  
create diversity in glycosidic bonds hydrolysis.**

Master Thesis in Medical Biology

By  
**Pallavi Bohra**



Submitted to  
**Faculty of Health science  
Institute of Medical Biology,  
University of Tromsø,  
Norway  
16 May 2011**

## **ACKNOWLEDGEMENTS**

I would first and foremost like to thank my supervisor Prof. Raafat El-Gewely for his patience, encouraging words and guidance throughout during my entire masters.

I would like to thank my husband Amit and my good friend Kari Jenssen Lægred for their support during my studies. I could not have made it without both of them.

I would also like to thank my parents, sister and brother for their motivational words during my thesis writing.

Finally, I would like to thank my daughter Arunima for her enormous naughtiness during my studies.

Pallavi Bohra

16 May 2011 Tromsø

## TABLE OF CONTENTS

Abbreviations.....	7
Abstract.....	9
1. Introduction.....	10
1.1 In vitro evolution of proteins & enzymes.....	10
1.2 Glycosyl hydrolase.....	13
1.3 $\beta$ -galactosidase.....	13
1.4 $\beta$ -xylosidase.....	16
1.5 $\beta$ -glucosidase.....	18
1.6 Random mutagenesis.....	19
1.7 Error-prone PCR.....	19
1.8 Screening.....	23
1.8.1 Semi-quantitative visual screening.....	23
1.8.2 Screening using 96-well plate format.....	23
1.9 Selection.....	24
2. Aims of the study.....	26
3. Material.....	27
3.1 Strains and Plasmid.....	27
3.2 Freeze and MgCl <sub>2</sub> Stock Solutions.....	27
3.3 Oligonucleotides Primers.....	28
3.4 Construction of Random Mutagenesis Library using Error-prone PCR.....	29
3.5 Screening Methods.....	30
3.6 Sequencing.....	31
3.7 Enzymatic Assay.....	31
3.8 Agarose Gel Electrophoresis.....	32
4. Methods.....	33
4.1 Plasmid DNA Isolation.....	33
4.2 PCR Purification.....	33
4.3 DNA Quantification.....	33
4.4 Agarose Gel Electrophoresis.....	34
4.5 Primer Design.....	34
4.6 DNA Sequencing.....	34
4.7 Whole Plasmid Mutagenesis.....	35
4.7.1 Sequencing of Wildtype Plasmid.....	36
4.7.2 Optimization for Whole Plasmid Random Mutagenesis PCR Conditions.....	37
4.7.3 Optimization of Primers Concentrations.....	38
4.7.4 Developing of PCR protocol with various Mn and dITP concentrations.....	39
4.8 Screening for $\beta$ -Glucosidase on Plates.....	40
4.8.1 Creating Random Library of LacZ Gene.....	40
4.8.2 Transformation.....	41
4.8.3 Screening on M9-cellobiose plates.....	41
4.8.4 Screening on LB- cellobiose plates.....	41
4.8.5 Screening with Filter Lift Assay.....	41
4.9 Screening for $\beta$ -Xylosidase on plates.....	42

4.9.1 Creating Random Library .....	42
4.9.2 Second round of mutagenesis .....	43
4.9.3 Second round for colony 9 (F9) DNA .....	43
4.9.4 Screening and Selection.....	43
4.9.5 Sequencing Analysis.....	43
4.9.6 Colorimetric $\beta$ -xylosidase Enzyme Assay .....	44
4.10 Screening for improved $\beta$ -Galactosidase Activity.....	45
4.10.1 Creating Random library and screening .....	46
4.10.2 Colorimetric $\beta$ -galactosidase Enzyme assay.....	46
4.11 Mutagenesis in the coding sequence of LacZ gene .....	46
4.11.1 Phenol chloroform DNA Extraction and Precipitation.....	47
4.11.2 Restriction Digestion and Ligation.....	48
4.11.3 Screening with Filter Lift Assay .....	49
4.12 Preparation of competent cells.....	49
4.12.1 CaCl <sub>2</sub> method.....	49
4.12.2 RbCl method.....	50
5. Results.....	51
5.1 Sequencing of pTZ1 Plasmid Encoding Wild Type LacZ Gene .....	51
5.2 Optimization for Mutagenesis PCR Conditions .....	61
5.3 Optimization of Primer concentration in random mutagenesis PCR.....	62
5.4 Development of PCR protocol with various Mn and dITP concentration.....	63
5.5 Screening and Selection for $\beta$ -glucosidase Activity .....	64
5.6 Screening and Selection for $\beta$ -xylosidase Activity.....	65
5.7 Screening of Second Round of Mutagenesis .....	66
5.8 Screening and Selection for Improved $\beta$ -galactosidase Activity.....	67
5.9 Colorimetric Enzyme Assays.....	68
5.10 Sequencing Analysis.....	77
6. Discussion.....	82
6.1 Sequencing of Plasmid pTZ1 Encoding Wildtype LacZ Gene.....	82
6.2 Random Mutagenesis Libraries of LacZ Coding Sequence .....	83
6.3 Screening and Selection.....	83
6.4 Colorimetric Enzyme Assay .....	87
6.5 Sequencing of Selected Mutants.....	88
7. Concluding Remarks.....	90
8. Future Perspectives .....	91
9. References.....	92
9.1 Articles.....	92
9.2 Websites.....	97
10. Appendix.....	98
10.1 Domain Structure of $\beta$ -galactosidase.....	98

## TABLES

Table 1: Some examples of enzymes that were optimized by using direct evolution .....	12
Table 2 : Average rate of mispairing among DNA polymerases during PCR .....	19
Table 3 : The bacterial strain and plasmid used during this study .....	27
Table 4 : Primer used in this study .....	28
Table 5: Machines and chemicals used to create Random Mutagenesis library .....	29
Table 6 : Chemicals and plates used for screening .....	30
Table 7: Chemicals used for preparing M9 minimal medium and M9 cellobiose plates .	30
Table 8: Machines and kits used for sequencing .....	31
Table 9: Machine and chemicals used for enzymatic assays .....	31
Table 10: Machines and chemicals used for agarose gel electrophoresis .....	32
Table 11: Twelve different PCR conditions for optimization .....	37
Table 12: Effects of dITP and Mn <sup>2+</sup> on PCR Product .....	64
Table 13: Xylosidase activity measured for variants of first and second generation .....	69
Table 14: Measurement of Beta-xylosidase with PNPX at 405nm .....	71
Table 15: Measurement with blue substrate X-xyl at 750nm .....	71
Table 16: Measurement with ONPG for $\beta$ -galactosidase .....	74
Table 17: Mutation in the LacZ coding gene of the selected variants .....	77

## FIGURES

Figure 1: Overview of directed enzyme evolution .....	11
Figure 2 : Lactose ( $\beta$ form) shown as Haworth perspectives .....	14
Figure 3 : General presentation of $\beta$ -galactosidase on its natural substrate .....	14
Figure 4: Showing tertiary structure of <i>E.coli</i> $\beta$ -galactosidase tetramer .....	15
Figure 5: An overview of xylan structure (comprise of D-xylose sugar) and microbial xylanolytic enzymes .....	17
Figure 6 : Structure of cellulose showing glycosidic bond .....	18
Figure 7: Proposed mechanisms for infidelity during DNA replication by metal ions ....	20
Figure 8: Schematic Presentation of the DNA shuffling .....	21
Figure 9 : Different method used in direct evolution to create libraries .....	22
Figure 10: Direct evolution of enzyme activity on a multiple substrate .....	24
Figure 11 : Hypothetical evolution presentation via uphill walk .....	25
Figure 12: Chemical structure of 5-Bromo-4-chloro-3-indolyl $\beta$ -D-glucopyranoside, <i>p</i> -Nitrophenyl $\beta$ -D-glucopyranoside and cellobiose .....	40
Figure 13: Chemical structure of 5-Bromo-4-chloro-3-indolyl $\beta$ -D-xylopyranoside and 4-Nitrophenyl $\beta$ -D-xylopyranoside .....	42
Figure 14: Sequencing primers and their locations on lacZ gene .....	44
Figure 15: Showing the visible range of the spectrum .....	44
Figure 16: Chemical structure of 5-Bromo-4-chloro-3-indolyl $\beta$ -D-galactoside (X-Gal) and 2-Nitrophenyl $\beta$ -D-galactopyranoside .....	46
Figure 17: Overview of random mutagenesis using coding sequence of LacZ .....	47
Figure 18: Entire 5,502 base pairs of plasmid pTZ1 .....	53
Figure 19: Restriction map of plasmid pTZ1 .....	58

Figure 20: Upstream region of LacZ gene .....	59
Figure 21: Two silent mutations in LacZ coding sequence of plasmid pTZ1 .....	60
Figure 22: Analysis PCR product with various conditions.....	61
Figure 23: Optimization of primers concentration for random mutagenesis PCR .....	62
Figure 24: Optimization of mutagenesis PCR with various concentrations of Mn and dITP.....	63
Figure 25: Blue colony appearing on chromogenic substrate for $\beta$ -glucosidase .....	64
Figure 26: Blue colonies after retransformation of plasmid DNA on substrate X-xyl.....	66
Figure 27: Retransformation of plasmids DNA isolated from the first and second generation.....	67
Figure 28: Screening and selection method of higher $\beta$ -galactosidase activity.....	68
Figure 29: Showing a mutation at primer binding site in colony 9 (F9) plasmid.....	70
Figure 30: $\beta$ -xylosidase activity with 4-Nitrophenyl $\beta$ -D-xylopyranoside at 405nm.....	72
Figure 31: $\beta$ -xylosidase activity using substrate 5-Bromo-4-chloro-3-indolyl $\beta$ -D-xylopyranoside at 705nm.....	73
Figure 32: A picture of microtiter plate with the performed colorimetric assay for $\beta$ -xylosidase activity.....	74
Figure 33: Showing the $\beta$ -galactosidase activity in compare with Wildtype .....	76
Figure 34: Chromatograms showing mutation in selected colonies .....	81

## Abbreviations

°C-Degree celsius

μl-Microliter

AMP- Ampicillin

β- Beta

bp-Base pair

BLAST- Basic local alignment search tool

*bla*- Beta- lactamase gene

CaCl<sub>2</sub>-Calcium Chloride

DMSO- Dimethylsulphoxide

DNA-Deoxyribonucleic acid

DTT-1,4-Dithiothreitol

dITP-2'-Deoxyinosine triphosphate

dNTP- Deoxy nucleoside triphosphate

EtBr-Ethidium bromide

FG – First Generation

HF- High fidelity

hr.-Hours

ITPG-Isopropyl-β-thiogalactopyranoside

Kb-Kilobase

LB-Lysogeny broth/Luria broth

LacZ-β-galactosidase

MgCl<sub>2</sub>-Magnesium chloride

MnCl<sub>2</sub>-Manganese chloride

mg- milligram

ml- milliliter

mM-Millimoles per liter

NaCl-Sodium Chloride

nm-Nanometer

NCBI- National centre for biotechnology information

OD<sub>600</sub>-Optical density at 600nm

ONPG-*o*-nitrophenol-Beta-D-galactopyranoside  
PCR-Polymerase chain reaction  
PD-1,2-Propandiol  
PNPX- *p*-nitrophenol-beta-D-xylopyranoside  
RPM- Revolution per minute  
RBS- Ribosomal binding site  
SG- Second Generation  
TAE-Tris-Acetic EDTA  
WT-Wildtype  
X-Gal-5Bromo-4chloro-3indolyl- $\beta$ -D-galactopyranoside  
X-gluc-5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronic acid  
X-Xyl -5-bromo-4-chloro-3-indolyl-beta-D-Xylopyranoside  
U- Unit (enzyme units)



## Abstract

Starting with LacZ of *Escherichia coli*, coding for  $\beta$ -galactosidase, the aim of the thesis project is to apply in vitro directed evolution techniques to help create other glycosidic bond hydrolysis activities. This was done using the main  $\beta$ -galactosidase backbone with limited amino acid sequence change. Any altered glycosyl hydrolase activity would lead to changed substrate specificity. Moreover, genetic changes leading to improved beta-galactosidase activity was also investigated.

Error-prone PCR was applied to the LacZ gene ( $\beta$ -galactosidase) to achieve the desired aims. The technique used to introduce random mutagenesis was based on modifications of method developed by Xu et al., 1999. Optimization was performed with DNA polymerase selection, PCR conditions and various Mn and dITP concentrations to obtain best amplified PCR product for random mutagenesis library construction.

Plasmid pTZ1 containing the entire coding sequence of LacZ was used a whole plasmid random Mutagenesis library construction strategy. The complete pTZ1 plasmid sequence had to be done in order to help establish a framework for primer design and establish a complete restriction map of the plasmid including the lacZ gene. The sequence analysis of the plasmid revealed that it has 5,502bp.

Screening of random mutagenesis libraries was based on the colour development resulting from the glycosidic hydrolysis of chromogenic substrate to identify any glycosidic activity towards particular glycosyl hydrolase on LB plates or M9 plates. We have screened random mutagenesis libraries for any possible activity for  $\beta$ -glucosidase,  $\beta$ -xylosidase or for an improved  $\beta$ -galactosidase activity.

Colonies that showed colour development on substrate even after retransformation of plasmid DNA for  $\beta$ -xylosidase activity were selected and its mutated plasmid DNA was sequenced. Two of the variants in which one has mutation at K552E position and another at N959Y were isolated, from two different clear blue colonies on  $\beta$ -xylosidase substrate.

However, to the issue of change in substrate specificity (colour development on plates) was not clear. The direct evolution method applied here is seems simpler and promising in creating random mutagenesis libraries in order to select variants with useful novel properties.

# 1. Introduction

## 1.1 In vitro evolution of proteins & enzymes

Direct evolution is a powerful approach for improving enzyme activity, alter substrate specificity and enhance stability of enzymes with potential applications in number of fields including agriculture, therapeutics and chemistry (Cole and Gaucher, 2011), that is in addition to the academic understanding of structure-function relationship of proteins in general. It mimics the natural Darwinian evolutionary process which takes millions of years to adopt new functions, but in vitro evolution provide opportunity to develop novel functional properties of protein and enzyme within much shorter time (Fenton et al., 2002b).

Number of specific alterations can be achieved by direct evolution for improving substrate range of enzymes, enantioselectivity, thermostability, protein solubility and expression, enzymatic properties of oxidases, lyases, recombinases, polymerases (Kaur and Sharma, 2006). Naturally occurring enzymes has its own catalytic activity, but for another specific utility, further tailoring is needed (Rubin-Pitel and Zhao, 2006). New enzyme function and property obtained by either searching among unknown natural species for different needed changes in that enzyme, or by conducting experimental in vitro approach to alter the known protein and enzyme, coupled with a designed selection scheme. (Zhang et al., 1997).

It is not necessary that small number of amino acid changes in active site can only affect the protein function but changes far away from active site residues can also give novel functional properties. In such case irrational design approaches and direct evolution are the efficient way to engineer a new enzyme property (Zhao and Arnold, 1997).

Directed evolution technique is based on three fundamental steps, production of mutation library from parental protein molecule by introducing sequence diversity, identifying the desire variants by efficient screening and selection and finally further mutagenesis and recombination of selected variants for further improvement of protein molecule (Cole and Gaucher, 2011).

Sequence diversity and screening/selection both are important to obtain a desired property of an enzyme catalyst using in vitro evolution techniques. Designing large libraries followed by high throughput assays to select a desired variant is preferable among the researchers. A variant with proxy desired function can be captured by using high throughput assay. To avoid these unwanted variants, much accurate low throughput assays should be considered if the library size can be reduced without losing functional diversity, otherwise it is worth to remember ‘you get what you screen for’(Cole and Gaucher, 2011, You and Arnold, 1996).

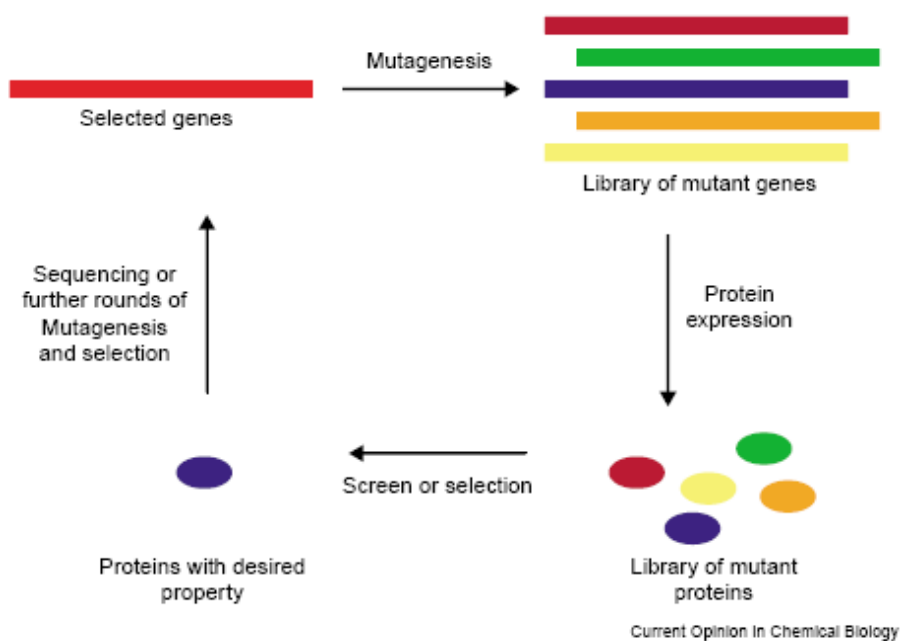


Figure 1: Overview of directed enzyme evolution  
(Tao and Cornish, 2002)

Table 1: Some examples of enzymes that were optimized by using direct evolution

Target enzyme	Evolved property	change	reference
$\beta$ -glucoronidase CelB ( <i>Pyrococcus furiosus</i> )	Increased catalytic activity at 20°C	3-fold	(Lebbink et al., 2000)
Galactosidase	Conversion to fucosidase	1000-fold increase in specificity towards <i>p</i> -nitrophenyl furanoside, 66-fold increase in specific activity	(Zhang et al., 1997)
$\beta$ - glucosidase ( <i>Paenibacillus polymyxa</i> )	Thermostable glucosidase	$\beta$ - 11-fold increase in the half-life of thermo inactivation at 50°C	(Liu et al., 2009)
$\beta$ -lactamase	Increased activity towards cefotaxime	32000-fold	(Stemmer, 1994)
Amidase ( <i>B.stearothermophilus</i> )	Increased expression in <i>E. coli</i>	23-fold	(Cheong and Oriel, 2000)

## 1.2 Glycosyl hydrolase

Enzymes are protein molecule that catalyzes chemical reactions in biological systems and this catalysis takes place at the active site of the enzyme. Enzymes are classified into six main groups. These groups are sub-classified and further subdivided according to the nomenclature of NC-IUBMB (Nomenclature Committee of the International Union of Biochemistry and Molecular Biology). Among all classes class third (EC 3) of the enzymes are the hydrolases that includes several subclasses and one of the subclass is glycoside hydrolase or glycosylases (EC 3.2). (Berg et al., 2001).

Glycosyl hydrolase are the enzyme that hydrolyze glycosidic bond between two or more carbohydrate or between carbohydrate and non-carbohydrate moiety (Henrissat et al., 1995). They are further divided into 3 sub-classes 1.) Enzymes hydrolyzing *O*- and *S*-glycosyl compounds (EC 3.2.1) 2.) Hydrolysing *N*-glycosyl compounds (EC 3.2.2) and 3.) Hydrolyzing *S*-Glycosyl compounds (EC 3.2.3)

(<http://www.chem.qmul.ac.uk/iubmb/enzyme/EC3/>).

Further classification is based on both chemical reaction and substrate specificity by IUBMB and amino acid similarities that are available and updated regularly at CAZy web server. In 2000 total 82 glycoside hydrolyses was identified (Henrissat and Davies, 2000). In 2009, 113 (Cantarel et al., 2009) and much more glycosyl hydrolases are made available at CAZy web server (<http://www.cazy.org/Glycoside-Hydrolases.html>).

## 1.3 $\beta$ -galactosidase

$\beta$ -galactosidase (LacZ) (EC 3.2.1.23) is the protein specified by the first structural gene of lactose (*lac*) operon proposed by Francois Jacob and Jacques Monod (Fowler and Zabin, 1970).  $\beta$ -galactosidase hydrolyzes  $\beta$ -1 $\rightarrow$ 4 linkage of lactose (a disaccharides joined covalently by O-glycosidic bond, abbreviated as Gal( $\beta$ 1 $\rightarrow$ 4)Glc) to D-galactose and D-glucose.(Matthews, 2005, Berg et al., 2001).

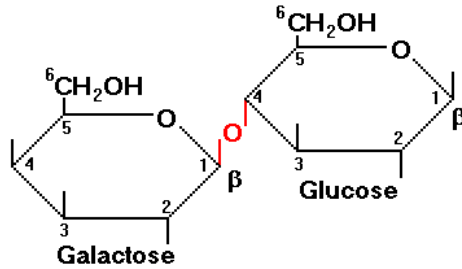


Figure 2 : Lactose ( $\beta$  form) shown as Haworth perspectives

(<http://www.rpi.edu/dept/chem-eng/Biotech-Environ/IMMOB/lactose.html>)

Lac system consists of two genetic elements that control the synthesis of proteins  $\beta$ -galactosidase, structural and regulatory elements. The structural element encoded two more proteins along with LacZ, LacY and LacA. LacY encoded permease allows bacteria to survive on lactose and LacA codes thiogalactoside transacetylase which involve in detoxification. Other regulatory elements (lacI gene) encoded repressor protein a disaccharide allolactose which binds to the operator site and prevent the expression. Transcription start when repressor protein removed from operator by inducer (Andrews and Lin, 1976, Berg et al., 2001).

Artificial inducer IPTG a most acceptable inducer (Isopropyl-thio- $\beta$ -D-galactoside) can also be used for induction of lac type promoter systems due to its structural similarity with allolactose. It is not a substrate and its concentration remains constant due to inability of *E.coli* to metabolized it (Hansen et al., 1998).

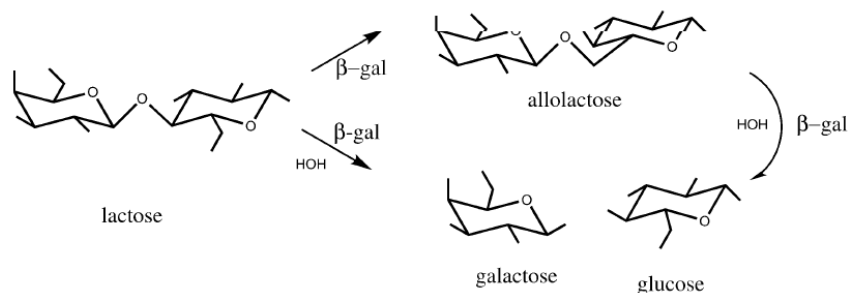


Figure 3 : General presentation of  $\beta$ -galactosidase on its natural substrate

(Matthews, 2005)

Four identical polypeptide chains (A to D) form the  $\beta$ -galactosidase tetramer each of the chain has 1023 amino acids. Each polypeptide (monomer) is made up of five domains and 50 residues at the N-terminal that contribute in the activating interface.

Different monomer segments contribute to the active site of  $\beta$ -galactosidase. Domain 3 that has alpha/beta barrel structure and loops comes from first and fifth domains of the same monomer to form active site. To complete four functional active sites, domain 2 of different monomers extends into the neighboring active sites called activating interface. As shown in Figure 4. monomer A donates, its Domain 2 loop to complete the active site of monomer D and in reciprocal monomer D donates its Domain 2 to complete the active site of monomer A. Similar kind of contribution of domains happen between monomer B and C (Matthews, 2005, Jacobson et al., 1994).

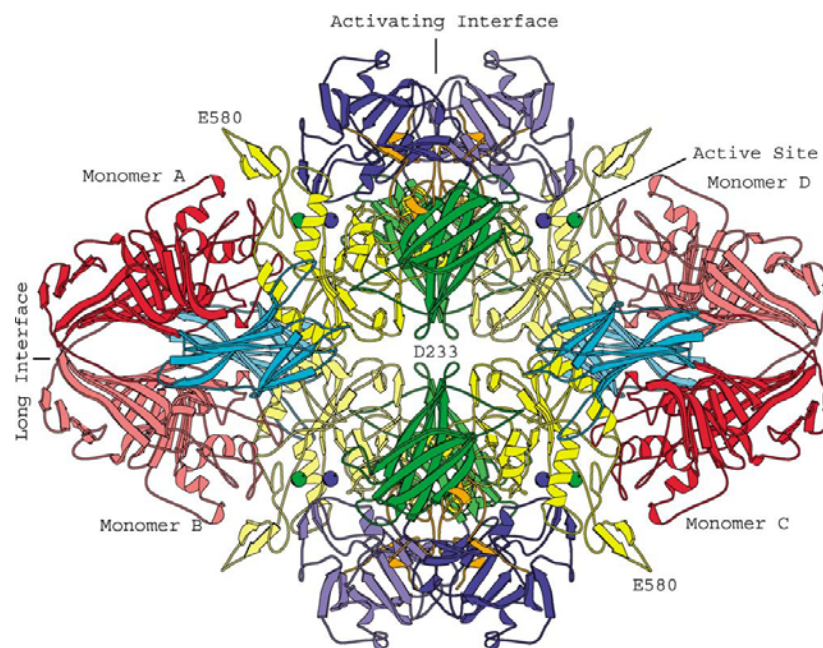


Figure 4: Showing tertiary structure of *E.coli*  $\beta$ -galactosidase tetramer

Colour indicates different domain, orange domain 1, blue domain 2, yellow domain 3, cyan domain 4 and red domain 5. Light and dark colour shades is to separate the same domain in different subunits (Matthews, 2005).

$\beta$ -galactosidase has huge utility in molecular biology. This enzyme also known as reporter gene, a basis of  $\alpha$ -complementation, a phenomenon widely used for blue/white screening (Langley et al., 1975). Various chromogenic substrate 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside (X-gal) and *o*-nitrophenyl- $\beta$ -D-galactopyranoside (ONPG) (Figure: 16) are available to indicate the presence of  $\beta$ -galactosidase during cell expression. ONPG hydrolyzed by  $\beta$ -galactosidase produce glycoside and *o*-nitrophenol.

Glycoside has no optical density at visible wave length while free *o*-nitrophenol gives yellow colour at 420nm (Lederberg, 1950, Matthews, 2005).

#### 1.4 $\beta$ -xylosidase

1.4  $\beta$ -D-xylosidases (E.C.3.2.1.37) also called xylobiase catalyzes the hydrolysis of D-xylose by attacking its  $\beta$ -1,4 glycosidic bond. 1,4  $\beta$ -D-xylosidases is one of the enzyme from microbial xylanolytic system require for complete degradation of xylan (Jordan et al., 2007). Three major component constitute plant cell wall that are cellulose (35-50%), hemicellulose (20-30%) and lignin (20-30%). Xylan is the major component of hemicellulose and key of plant biomass and carbon flow. Xylan is a heteropolysaccharides and its backbone consists of five carbon sugar D-xylose ( $\beta$ -1,4-linked xylopyranosyl residues) as well as it contains groups of 4-O-methyl-D-glucuronosyl, D-glucuronic acid both linked to O-2 position of D-xylose, acetyl and  $\alpha$ -arabinofuranosyl linked to O-3 position of D-xylose. Plant cell wall xylan linked to cellulose and lignin with covalent and non-covalent interaction (Subramaniyan and Prema, 2002, Biely, 1985).

Conversion of xylan into monosaccharides can be achieved by several hydrolytic enzymes most commonly hemicellulases, a widespread group of glycoside hydrolases. Endo-1, 4- $\beta$ -xylanases (E.C.3.2.1.8) which attack on polysaccharide backbone by random hydrolysis and 1,4  $\beta$ -D-xylosidases (E.C.3.2.1.37) hydrolyses by endwise attack of

Xylooligosaccharides to D-xylose. The rest of the side chain is hydrolysed by  $\alpha$ -L-arabinofuranosidase,  $\alpha$ -D-glucuronidase, galactosidase and acetyl xylan esterase. All these enzymes together are called as xylanolytic enzymes. (Brux et al., 2006, Biely, 1985, Subramaniyan and Prema, 2002).



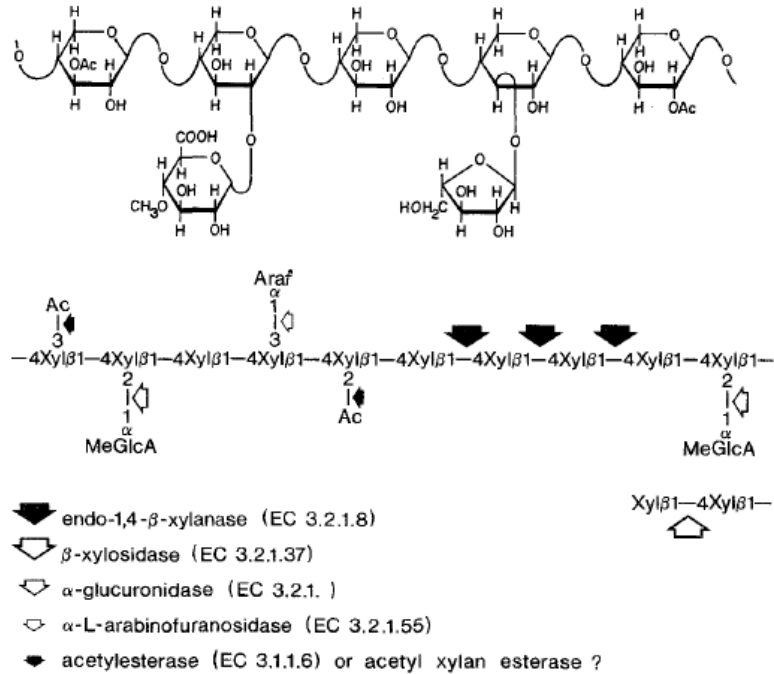


Figure 5: An overview of xylan structure (comprise of D-xylose sugar) and microbial xylanolytic enzymes

(Biely, 1985)

Few applications are available for xylanolytic enzymes in industry and improving these efficient hydrolytic enzymes can offer new prospects for breaking down hemicellulose in waste-material, as well as in other industries.

- These xylanolytic enzymes have the importance in paper and pulp industries to reduce the use of chlorine as the bleaching agent used for converting residual lignin into pulp while hydrolyzing the xylan (Subramaniyan and Prema, 2002).
- Xylanolytic enzymes along with cellulase and pectinase can be used for clarifying juices, liquifying fruits and vegetables (Biely, 1985).
- Preparation of the sugars xylose, xylobiose and xylooligomers by hydrolysis of xylan (Subramaniyan and Prema, 2002, Wong et al., 1988).
- Maximal hydrolysis of low value feedstock from agriculture forestry and municipal solid waste (Subramaniyan and Prema, 2002, Lynd et al., 1991).
- Production of ethanol from agro-waste by xylanolytic enzyme treatment (Subramaniyan and Prema, 2002).

## 1.5 $\beta$ -glucosidase

Complete degradation of cellulose requires three main classes of enzymes exo-cellulase ( $\beta$ -1,4-o-glucan cellobiohydrolase), endo-cellulase ( $\beta$ -1,4-D-glucan glucanohydrolase) and  $\beta$ -1, 4-D-glucosidase. Initial two enzymes hydrolyze cellulose into cellobiose and  $\beta$ -1, 4-D-glucosidase (EC 3.2.1.21) also called cellobiase convert cellobiose into glucose by hydrolyzing  $\beta$ -1-4 linkage (Love et al., 1988). Cellulose is a polymer in pyranose form of glucose units linked with  $\beta$ -1 $\rightarrow$ 4 bond (Beguin and Aubert, 1994). Cellulose is the main component of plant cell wall (15-40%) and the most abundant renewable resource in the nature. Due to its hydrogen bonded crystalline structure and insolubility it is difficult to degrade in compare to hemicellulose, lignin and pectin (Doi and Kosugi, 2004).

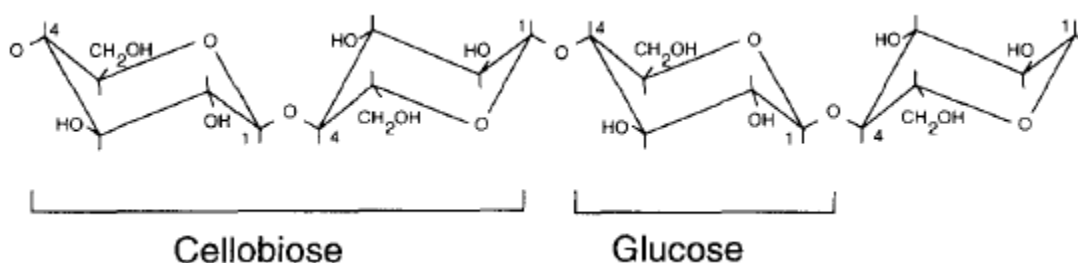


Figure 6 : Structure of cellulose showing glycosidic bond  
(Beguin and Aubert, 1994)

Due to its part in carbon cycle and enormous industrial utility cellulolytic enzymes become part of interest among the researchers. Improvement in cellulose degrading enzymes can open potential application in number of biotechnological field including food, brewery and wine, animal feed, textile and laundry, pulp and paper, agriculture. Textile waste contains 40% of cellulosic part and degradation of this material by natural enzymatic treatment can produce cost effective valuable products such as ethanol. Biogas and methane can also be produced from efficient enzymatic treatment of cellulose.  $\beta$ -glucosidase involve in improvement of wine aroma by modifying glycosylated precursors (Bhat, 2000).

## 1.6 Random mutagenesis

Direct evolution methods introduce mutations and recombination in order to develop biocatalysts with improved properties without knowing the complete catalytic and structural properties in details.

Identification and isolation of coding sequence of the wildtype gene followed by random mutagenesis, such as error-prone PCR, DNA shuffling and incremental truncation together with selecting the needed phenotype are the main steps of in vitro evolution. The random library expressed in host cell and subsequently screened and select. (Kaur and Sharma, 2006). Some of random mutagenesis methods are mention below.

## 1.7 Error-prone PCR

It is the first method described to achieve random mutagenesis. The technique based on the fact of Taq DNA polymerases that lack the proof reading that incorporate mispairing at the frequency of  $0.1 \times 10^{-4}$  to  $2 \times 10^{-4}$  per nucleotide during the extension of strand in PCR reaction (Kaur and Sharma, 2006). Several DNA polymerase fidelity has been identified and amongst all, Taq polymerase has the lowest fidelity, which makes Taq the best candidate for an in vitro mutagenesis (Cadwell and Joyce, 1992).

Table 2 : Average rate of mispairing among DNA polymerases during PCR

(Cline et al., 1996)

DNA polymerase	No. of PCR	Target (ng)	Template doubling	Error rate ( $10^{-6}$ +or – SD)
Pfu	10	24	9.7	1.3+or- 0.2
	2	2	12.7	0.7+ or-0.1
	2	0.2	16.0	0.8+ or -0.02
	2	0.02	19.4	1.0+or-0.04
Deep Vent	4	24	9.7-10	2.7+or-0.2
Vent	6	24	8.7-10	2.8+or-0.9
Taq	11	24	8.7-11	8.0+or-3.9
UITma	2	24	9.7	55+or-2

In addition to the Taq DNA polymerase, increasing the concentration of  $MgCl_2$  nucleotide analogs, and  $MnCl_2$  can incorporate mispairing during PCR (Kaur and Sharma, 2006). DNA polymerase has one binding site for template, one for dNTP and one for dNMP (Mildvan and Loeb, 1979). Binding of  $Mn^{+2}$  effect base-pairing properties by altering template and substrate molecule. It also interact with DNA polymerase reducing the selection priority of nucleotides before they insert (Beckman et al., 1985).

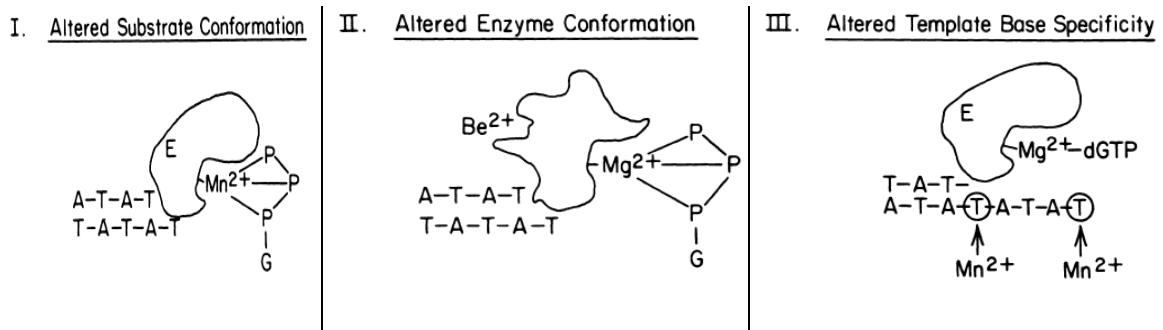


Figure 7: Proposed mechanisms for infidelity during DNA replication by metal ions (Zakour et al., 1981)

dITP is a natural occurring base analog which occasionally found at the first position of tRNA anticodon. It can pair with or without hydrogen bonds to any of the four nucleotides. In tRNA anticodon it pair with A, C, G and U and with poly(Aristarkhova et al.) they make a stable complex (Ohtsuka et al., 1985).

The important point to be considered in error-prone PCR technique is that the beneficial mutations are rare in comparison to the deleterious. It is possible that the combination of beneficial and deleterious mutation form an inactive enzyme. It is necessary in this technique that frequency should be maintained at low to obtain high number of desired variants (Harayama, 1998). The protocols available for error-prone are mostly not random enough. They mostly favor transitional point mutations over transversional mutations. Transitional point mutations exchange one pyrimidine with another pyrimidine, or one purine with another purine ( $AT \leftrightarrow GC$  and  $TA \leftrightarrow CG$ ) while in transversional exchange occur between purine to pyrimidine and pyrimidine to purine ( $AT \leftrightarrow CG, AT \leftrightarrow TA, GC \leftrightarrow CG, GC \leftrightarrow TA$ ) (Fenton et al., 2002b). Better results were obtained when a new method was developed by (Xu et al., 1999).

A number of other methods are published for recombining mutations from successful selected candidates into one or few more successful candidates (see figure: 8 and 9). DNA shuffling method involves DNAaseI treatment on large gene to create a random fragments pool of related DNA sequences. These fragments than reassembled by self-priming PCR and extension with DNA polymerase. As the product positive mutant accumulate from the sequence pool and negative eliminate (Zhang et al., 1997).

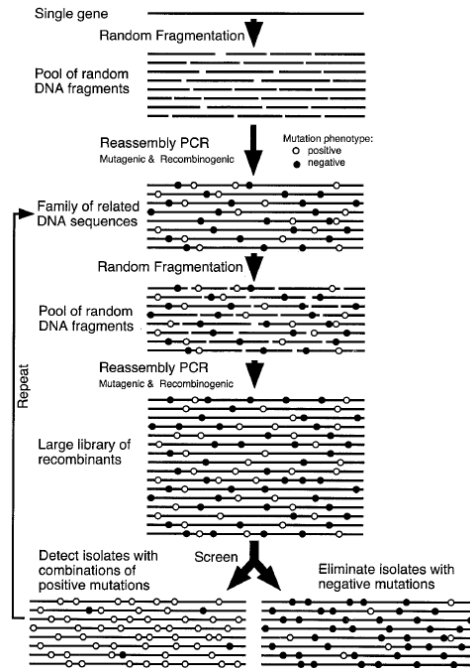


Figure 8: Schematic Presentation of the DNA shuffling  
(Zhang et al., 1997)

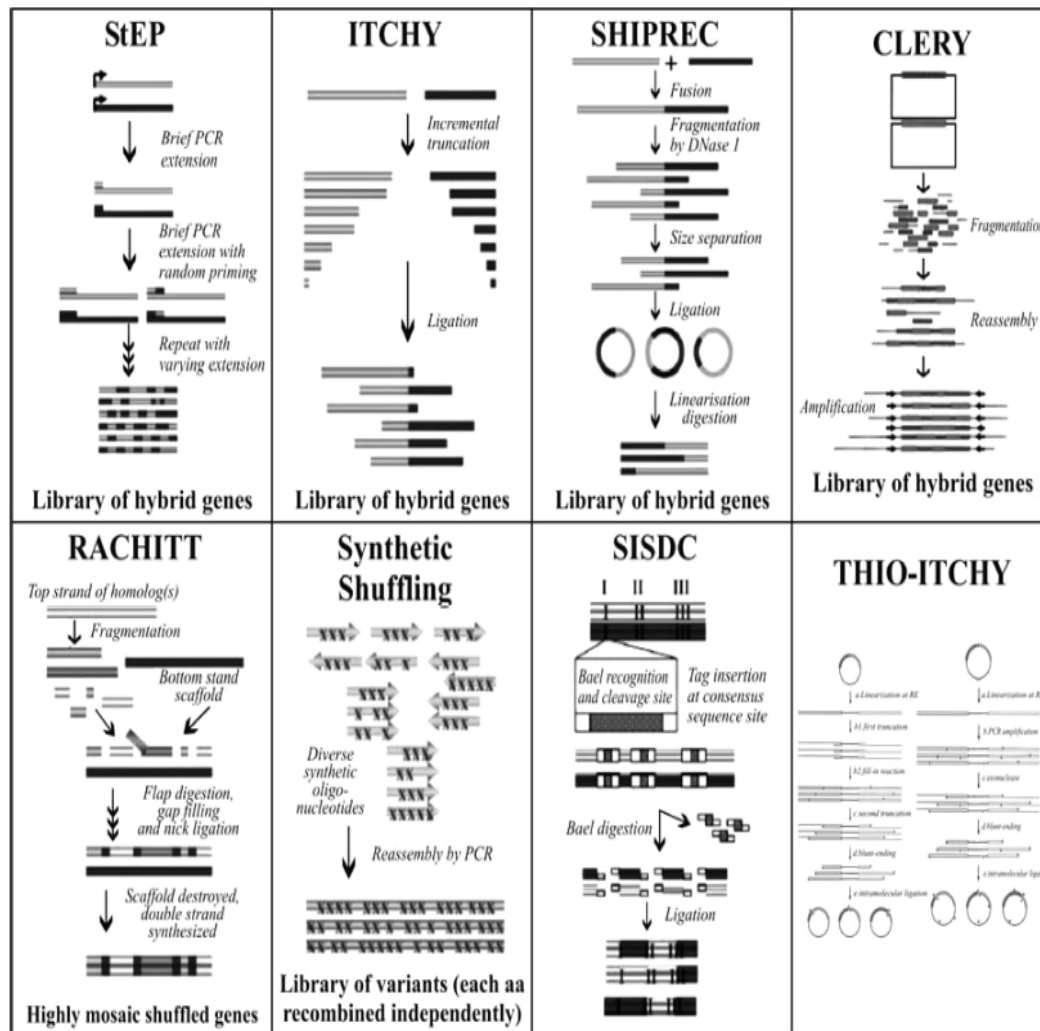


Figure 9 : Different method used in direct evolution to create libraries

(Kaur and Sharma, 2006).

StEP- Staggered Extension Protocol, ITCHY- Iterative Truncation for the Creation of Hybrid enzymes, SHIPREC- Sequence Homology Independent Protein RE Combination, RACHITT- Random CHImeragenesis on Transient Template, Synthetic Shuffling, SISDC- Sequence-Independent Site-Directed Chimeragenesis, CLERY- Combination Libraries Enhanced by Recombination in Yeast, THIO-ITCHY.

## **1.8 Screening**

Identification of desirable variants from a mutant library is the most challenging step in the directed evolution technique. Using commercial or model substrate that develops colour or fluorescence while screening is still in question. These substrates are easy to work with, but enzyme that is selected on a model substrate may not be active on another more commonly used substrate (Leemhuis et al., 2009).

### **1.8.1 Semi-quantitative visual screening**

In this screening is based on the visual signal generated by the gene of interest (Zhao and Arnold, 1997). Total of 36 variants were selected on chromogenic substrate X-gal for higher  $\beta$ -galactosidase activity from the wildtype GUS and LacZ-deficient *E.coli* (Matsumura and Ellington, 2001). Total of 10,000 transformants colonies were visually screened on X-fuc (5-bromo-4-chloro-3-indolyl- $\beta$ -D-fucopyranoside) and the bluest 20-40 colonies were selected in each round of screening (Zhang et al., 1997).

### **1.8.2 Screening using 96-well plate format**

It is a standard format of high throughput screening, with quantitative analysis and with increased sensitivity. It is another approach of screening in order to overcome the limitation of visual-based screening. Visual screens are non-quantitative and insensitive to the small changes (Zhao and Arnold, 1997). Number of screening are based on 96-well plates, one example is in the direct evolution of esterase, using the chromogenic substrate *p*-nitrobenzyl ester (pNB) was carried out in this format after the sequential generation of random mutagenesis (JC and FH, 1996, Zhao and Arnold, 1997).

## 1.9 Selection

A direct link of cell growth and improved enzyme function has been the basis of the selection process. The mutation library is transformed in host cells followed by plating on selective medium (Leemhuis et al., 2009).

To obtain a desired variant with high specificity for a new substrate can be achieved by increasing the substrate range. The possibility to remove variants with wildtype activity and undesired mutants by negative selection can also enhance the selection of positive variants. Positive mutants can be followed with further mutagenesis to improve the desired activity. Selection for particular enzyme specificity can also be achieved by continuous pressure for higher activity on a new substrate with recognition of older selection(Tracewell and Arnold, 2009).

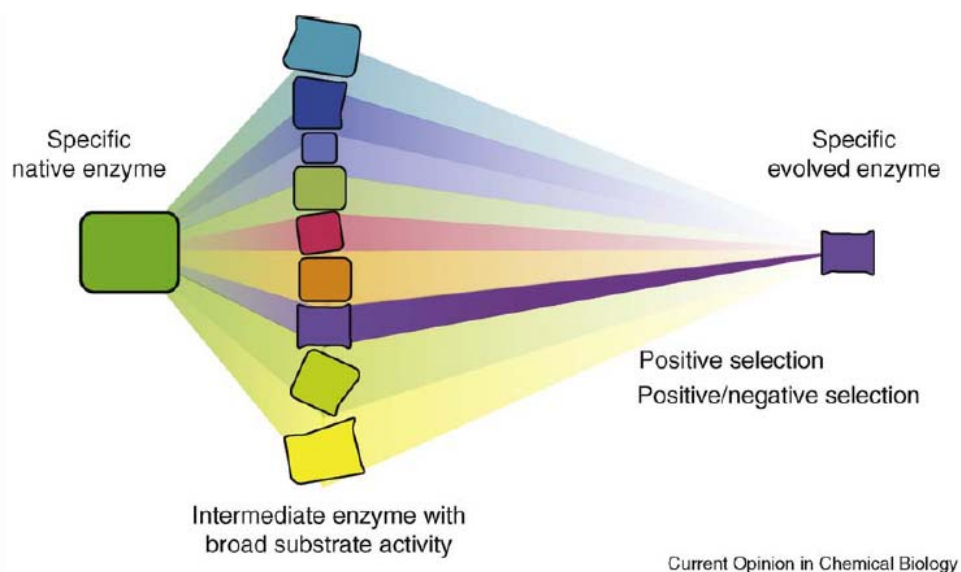


Figure 10: Direct evolution of enzyme activity on a multiple substrate

Selection of positive variant by eliminating unwanted mutants  
(Tracewell and Arnold, 2009).



A single beneficial mutation accumulates slowly during direct evolution process, along with efficient selection pressure. High mutation rate during random mutagenesis can convert the beneficial mutation into deleterious. Substitution of 1 or 2 mutation are preferred if entire gene is mutated. Two important situations are important for uphill walk through single beneficial mutations. First an intermediate should exist in the path from starting point to desired improved property. Second, the path should not reach to the dead-end or acquire any destabilizing mutations (Tracewell and Arnold, 2009).

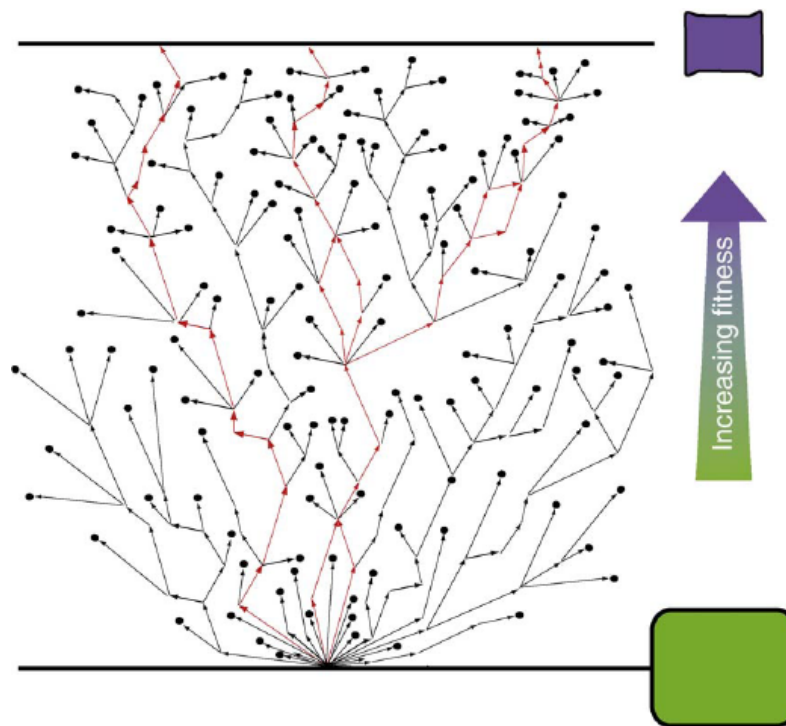


Figure 11 : Hypothetical evolution presentation via uphill walk

Arrow indicates the single amino acid substitution for increase in fitness. Number of sequences can have desired fitness and multiple paths can lead to those sequences. Restricting the mutations at particular position and destabilizing of protein can end up into the dead end (Tracewell and Arnold, 2009).

## **2. Aims of the study**

1. Sequencing of plasmid pTZ1 with entire coding sequence of LacZ gene.
2. Apply directed evolution technique to introduce random mutations in LacZ gene coding for  $\beta$ -galactosidase in order to change the natural substrate specificity towards the hydrolysis of other glycosides' as substrates.
3. Obtain a variant of  $\beta$ -galactosidase with improved activity towards its natural substrate.

### 3. Material

#### 3.1 Strains and Plasmid

Table 3 : The bacterial strain and plasmid used during this study

Strains and Plasmids	Description	Reference or source
Super competent <i>E. coli</i> XL1 blue cells/XL1Blue competent cells	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> (F' <i>proAB lacIqZΔM15 Tn10</i> (Sambrook et al., 1989)	Stratagene/Lab prepared
pTZ1 plasmid	Amp <sup>r</sup> , <i>tacI</i> , complete LacZ fragment	Provided by my supervisor Prof. Raafat El-Gewely (Su et al., 1990)

#### 3.2 Freeze and MgCl<sub>2</sub> Stock Solutions

*E. coli* host cell strain (XL1Blue competent cells), plasmids, were streaked on LB+ plates containing appropriate antibiotics and incubated at 37°C for overnight. Freeze stock was prepared by collecting single colony from LB+ plates and suspending the colony in 3-5 ml LB+ media with require antibiotics. To prepare glycerol stock, a volume of 1000μl of the overnight culture was mixed with 800μl of 70% glycerol (Final concentration of 31%) and stored at -70°C. Two separate primer concentrations were prepared, 100pmoles/μl and 10 pmoles/μl and stored at -70 °C freezers.

Along with the freeze stock, 1 ml of overnight culture was also stored in 10 mM of MgCl<sub>2</sub> at 4°C (Silhavy et al., 1984).

### 3.3 Oligonucleotides Primers

Table 4 : Primer used in this study

Primer name	Sequence from 5' to 3'	Bases	TM (Bauch et al.)	Reference	Supplier
U1Z1	ggtcatagctgttctctgtg	20	60	This study	Invitrogen
U2Z1	gctgcaaggcgattaagttg	20	60	This study	Invitrogen
D1Z1	agcaactgatggaaccagc	20	60	This study	Invitrogen
D2Z1	cagtatcggcggaattccag	20	62	This study	Invitrogen
Gap1	gatcatgtaactgccttg	19	56	This study	Invitrogen
Gap2	ggtagacacgacttatcg	19	56	This study	Invitrogen
Mut1	caggaacaggatccgacaaccga	24	74	This study	Invitrogen
Mut2	ctgcaggtcgacccccctgcc	21	74	This study	Invitrogen
Mut 3	tcggtgtcggatcctgttctctg	24	74	This study	Invitrogen
LacZ-seq1	gaattctcatgtttgacagct	21	58	This study	Invitrogen
LacZ-seq2	gcggattaacgattactcgt	20	58	This study	Invitrogen
LacZwpa	agctgtcaaacatgagaattc	21	58	This study	Invitrogen
LacZII2669(F)	gtgcccgaattctcatgtttg	21	62	This study	Invitrogen
LacZII1572(F)	gctgcataaacggactacaca	21	62	This study	Invitrogen
LacZII1772(R)	tcatacgataattcaccgccg	21	62	This study	Invitrogen
LacZII2577(F)	gctaaataactggcaggcggt	20	60	This study	Invitrogen
LacZII2740(R)	tcatacagaactggcgatcg	20	60	This study	Invitrogen
LacZII3600(F)	ggtagcagagcgggtaaac	20	62	This study	Invitrogen
LacZII3770(R)	ccataattcaattcgcgcgtc	21	60	This study	Invitrogen
LacZII4100(R)	agacctgcggttcagcagttc	20	62	This study	Invitrogen

### 3.4 Construction of Random Mutagenesis Library using Error-prone PCR

Table 5: Machines and chemicals used to create Random Mutagenesis library

Machines	Name	Suppliers
Programmable thermal cycler	GeneAmp PCR system 9700	Applied Biosystems
Transformation/XL1Blue Supercompetent Cells	Chemical Method	www.genomics.agilent.com
Plasmid Isolation	Qiagen mini/midi/maxi DNA purification System	www.Qiagen.com
Filter lift assay	Hybond™ filter (82mm)	www.gelifesciences.com
DNA Quantification	Nanodrop® 1000	www.nanodrop.com

Chemicals	Supplier
MnCl <sub>2</sub>	Sigma-Aldrich
MgCl <sub>2</sub>	Merck
10XPCR buffer	Qiagen
Hot start Taq plus polymerase	Qiagen
dNTP	Finnzymes
dITP	Fermentus
DpNI	New England Biolabs
NE buffer 3	New England Biolabs
BamHI	New England Biolabs
Sall	Takara Bio Inc
NEbuffer 4	New England Biolabs
Shrimp alkaline phosphatase	Promega
T4 DNA ligase	Takara Bio Inc
N,N-Dimethylformamide	Sigma-Aldrich
Methanol	Sigma-Aldrich (Sigma chemical. co)
Beta-merceptoethanol	Sigma-Aldrich

### 3.5 Screening Methods

Table 6 : Chemicals and plates used for screening

Chemicals	Suppliers/composition
LB-Media (per litre)	SUMP, UNN, Tromsø/10g tryptone,5g yeast extract and10g NaCl
LB-Plates(per litre)	SUMP, UNN, Tromsø/10g tryptone,5g yeast extract,10g NaCl and 20g agar
5-Bromo-4-Chloro-3-indolyl-β-D-Glucopyranoside (X-glu)	Sigma-Aldrich
5-Bromo-4-Chloro-3-indolyl-β-D-Galactopyranoside (X-gal)	Sigma-Aldrich
5-Bromo-4-Chloro-3-indolyl-β-D-Xylopyranoside (X-xyl)	Sigma-Aldrich
IPTG (Isopropyl-β-D-thiogalactopyranoside)	Sigma-Aldrich
Ampicillin	SUMP, UNN, Tromsø

Table 7: Chemicals used for preparing M9 minimal medium and M9 cellobiose plates

Chemical	Suppliers
Na <sub>2</sub> HPO <sub>4</sub>	Merck
KH <sub>2</sub> PO <sub>4</sub>	Merck
NaCl	Sigma-Aldrich
NH <sub>4</sub> Cl	Sigma-Aldrich
MgSO <sub>4</sub>	Sigma-Aldrich
CaCl <sub>2</sub>	Merck
Agar	Sigma-Aldrich
Cellobiose	Sigma-Aldrich

### 3.6 Sequencing

Table 8: Machines and kits used for sequencing

Machines and kits	Name	Suppliers
Thermal Cycler	Gene amp PCR systems 9700	Applied Biosystems
Sequencing machine	ABI 3130XL genetic analyzer	Applied Biosystems
PCR reaction	ABI PRISM BigDye™ kit 5X sequencing buffer	Applied Biosystems

### 3.7 Enzymatic Assay

Table 9: Machine and chemicals used for enzymatic assays

Machine	Name	Supplier
Spectrophotometer	Softmax® version 2.3	Molecular Devices GmbH

Chemical	Suppliers
Ortho-Nitrophenyl-β-galactoside (ONPG)	Sigma-Aldrich
4-Nitrophenol-Xylopyranoside	Sigma-Aldrich
Bugbuster (10X protein extraction reagent)	Novagen
CTAB	Sigma-Aldrich
Na <sub>2</sub> HPO <sub>4</sub>	Merck
NaH <sub>2</sub> PO <sub>4</sub>	Merck
KCl	Merck
DTT (1,4-Dithiothreitol)	Amershan Biosciences

### 3.8 Agarose Gel Electrophoresis

**Table 10:** Machines and chemicals used for agarose gel electrophoresis

Machines and kits	Name	Suppliers
UV-camera	GelDoc 2000	Bio-Rad

Chemicals	Supplier/ composition
Ultra pure <sup>TM</sup> Agarose	Invitrogen
Ethidium Bromide (EtBr)	Sigma-Aldrich (Sigma chemical. co)
TAE	Tris-Hcl (40mM, pH 8.0)-acetic acid (20mM) EDTA(1mM) buffer
1kb ladder	New England Biolabs
6X Loading dye	New England Biolabs



## **4. Methods**

### **4.1 Plasmid DNA Isolation**

The isolation of plasmid DNA (mini Plasmid purification and Maxi Plasmid purification) was carried by Qiagen plasmid purification kit (Qiagen, Hilden, Germany). According the user manual provided by manufacturer the method is based on two main steps, alkaline lysis procedure, binding of plasmid DNA to an anion-exchange resin under low salt and pH conditions. During the procedure other impurities like RNA, proteins, dyes and low molecular weight impurities are removed by a medium salt wash. Elution of plasmid DNA and precipitation is done with high salt buffer and isopropanol respectively ([www.qiagen.com](http://www.qiagen.com)).

All the Plasmid DNA isolation during this project was carried out by using Qiagen (Hilden, Germany) plasmid mini/midi/maxi DNA isolation Kit.

### **4.2 PCR Purification**

Technique is based on absorption of nucleic acid to the silica membrane in presence of high salt concentration. Undesired impurities like primers, salts, enzymes, nucleotides, dyes and oils do not bind to the silica membrane and release in flow through. Elution prefer in low salt concentration of E.B buffer contain 10mM of Tris-HCl pH 8.5 ([www.Qiagen.com](http://www.Qiagen.com)).

PCR purification in this project was carried out by using Qiagen (Hilden, Germany) PCR Purification Kit.

### **4.3 DNA Quantification**

DNA quantification was done by Nanodrop® ND-1000 (NanoDrop Technologies INC, Wilmington, DE, USA) at A260nm using 1µL of sample. The ration at 230/260nm and 260/280nm are calculated by the NanoDrop software and displayed on the attached computer ([www.nanodrop.com](http://www.nanodrop.com)).

#### **4.4 Agarose Gel Electrophoresis**

Gel electrophoresis is based on a simple electrochemical principle. The technique relies on negative charge of the phosphate back bone of nucleic acid and ability to distribute voltage gradient in a sieving matrix (Brody and Kern, 2004). Meaning when electric field applied to the agarose gel, DNA molecule migrate towards the positive charge according to there size. Smaller fragment migrate faster than the larger on the gel. Gel is prepared by a polysaccharide called agarose which is consist of D-galactose and 3,6-anhydro-L-galactosidase (Johansson, 1972, Primrose et al., 2006) It is prepared by boiling with TAE buffer and allowed to cool down up to 50°C. After cooling down ethidium bromide (EtBr) was added to the concentration of 0.5µg/ml to agarose gel (Sambrook et al., 1989).EtBr is a commonly used dye and fluoresce under UV light due to intercalating into a DNA strand (Primrose et al., 2006). Gels are solidified by casting on trays. Samples were loaded after mixing with loading dye containing 2.5 % Ficoll 400, 11 mM EDTA, 3.3 mM Tris-HCl, 0.017 % SDS, 0.015 % Bromophenol Blue, pH 8.0 at 25°C (for 1X) ([www.neb.com](http://www.neb.com)).

All PCR products in this project were analyzed on 0.8% agarose gel with TAE buffer for 45-90 minutes at 90V of current. PCR products were compared with 1Kb of ladder.

#### **4.5 Primer Design**

The amplification of DNA strand requires primer that attaches to the complementary template strand and serves as elongation starting points for the DNA polymerase. Two main points were considered during designing the primer specificity and efficiency. The primers were used for this study was designed manually and there  $T_m$  was calculated by standard formula  $T_m=2A/T+4G/C$  (Dieffenbach et al., 1993).

#### **4.6 DNA Sequencing**

DNA sequencing was made in order to confirm plasmid constructs sequence as well as to identify the sequence of any generated variant of interest by using Sanger's (1977) dideoxy chain terminating method.

The sequencing reaction was performed using ABI big dye 3.1v cyclic sequencing terminator reaction (Applied Biosystems) in total 20µl of volume. The reaction contains 250µg of template, 3µl of 5X sequencing buffer, 10pmole concentration of primers.

The ABI3130XL genetic analyzer (Applied Biosystems) at sequencing laboratory of Institute of Medical biology (Medical Genetics) at the University of Tromsø was used to read the nucleotide base pair and the obtained sequencing results were analyzed and manually edited by the sequence scanner and finch TV software. All generated sequences were aligned and compared to the wild type sequences of β-galctosidase (product of lacZ gene) by using the tools available at National Centre of Biotechnology Information website (<http://www.ncbi.nlm.nih.gov>).

#### **4.7 Whole Plasmid Mutagenesis**

Our group previously developed a useful protocol for the generation of error prone PCR libraries of genes of interest (Xu et al., 1999) and(Fenton et al., 2002b) .In this method primers flanking the gene of interest were designed and optimized error prone PCR was conducted, subsequently the generated repertoire of the coding sequence with the generated mutations were cloned downstream of an appropriate promoter in a suitable plasmid.

In this project we wanted to further simplify the method by designing complementary primers in opposite direction in order to amplify the entire recombinant plasmid containing the gene of interest. Error Prone PCR was conducted as before using combinations of Manganese(Rahme et al.), and dITP using plasmid pTZ1 containing the entire coding sequence for LacZ gene as a template for PCR mutagenesis See detailed steps in 4.7.4

### 4.7.1 Sequencing of Wildtype Plasmid

In this project, Wild type plasmid pTZ1 containing  $\beta$ -galactosidase gene (Provided by my supervisor Prof. Raafat El-Gewely (Su et al., 1990) was used for creating whole plasmid random libraries. In order to design primers for whole plasmid amplification and to identify exact restriction site for creating mutation in coding gene, we had to verify the sequence of plasmid pTZ1. Since the derivative plasmid pZ1981 and its sequence was known and available in data base (<http://www.ncbi.nlm.nih.gov/nuccore/434652>), we prepared sequencing primer from pZ1981 to sequence plasmid pTZ1. Four primers were designed, two for upstream (U1Z1, U2Z1) and two for downstream (D1Z1, D2Z1), such that a two overlapping part of the complete plasmid were sequenced with four Polymerase chain reaction. Two another primers Gap1 and Gap 2 were designed for sequence the remaining part of the plasmid. Sequencing for LacZ portion was done with 8 primers separately. The primers were synthesized by Invitrogen, Norway (see section 3.3).

#### Procedure:

1. Total 20 $\mu$ l of reaction was prepared for each of the primer (UIZI, U2Z1, DIZI, D2Z1, Gap1, Gap2 and 8 primers for LacZ coding gene of the plasmid).  
The reaction mixture for sequencing was performed as mention in section 4.6.
2. All four samples were loaded on thermal cycler in which template was denatured for 5 minutes at 96°C and 30 cycles for Step1.) 96°C for 10 seconds, Step 2.) 50°C for 5 seconds, Step 3.) 60°C for 4 minutes and finally the reaction was terminated at 4°C.
3. Precipitation and analysis of sequencing reactions products were performed in the sequencing laboratory of institute of Medical biology at the University of Tromsø as mention in section 4.6.

#### 4.7.2 Optimization for Whole Plasmid Random Mutagenesis PCR Conditions

To get best amplified PCR product, optimization was performed for different primers sets (Mut1,Mut2 and Mut1, Mut3), two different polymerase Hot start plus Taq polymerase , PFU high fidelity polymerase and PCR condition with water, Q solution (Qiagen, Hilden, Germany) and PD solution.

Table 11: Twelve different PCR conditions for optimization

Template DNA	Primer	DNA polymerase	Conditions		
pTZ1: WT LacZ	Mut1&Mut2	Hot star plus Taq	Water	Q sol <sup>n</sup>	PD
pTZ1: WT LacZ	Mut1&Mut2	PFU HF	Water	Q sol <sup>n</sup>	PD
pTZ1: WT LacZ	Mut1&Mut3	Hot star plus Taq	Water	Q sol <sup>n</sup>	PD
	Mut1&Mut3	PFU HF	Water	Q sol <sup>n</sup>	PD

#### Procedure:

Four different sets of PCR conditions were prepared to analyze each of the combination in total 25µl volume of reaction. 4 X 80µl of master mix were prepared for each of the polymerase and out of this 20µl was distributed for various PCR conditions.

Reagents	Volume
10x PCR buffer or PFU 10 x rx <sup>n</sup> buffer	10µl
DNA(75ng)	10µl
dNTP(200µM)	8µl
Mut1(15pmole/µl)	1µl
Mut2(15pmole/µl)	1µl
Hot star plus Taq polymerase or PFU HF polymerase	1µl
H2O	49µl
Total	80µl

1. 20µl of master mix was distributed in to 12 PCR reaction tubes contains different combination of polymerase and primers (see section 4.7.2).
2. In each of the 3 PCR sets, 5µl of water, 5µl of Q solution and 5µl PD was added separately.
3. All the samples were loaded on thermo cycler with activation of 95°C for 5 minutes and 30 cycles of step 1.) Denaturation 94°C for 20 minutes step 2.) Annealing at 72°C for 4 minutes, 1 cycle for additional extension at 72°C for 10 minutes and reaction was terminated at 4°C.
4. PCR products were analyzed on 0.8% agarose as mention previously in section 4.4

### **4.7.3 Optimization of Primers Concentrations**

Four different concentration of primers Mut1 and Mut3 and addition of Mg<sup>2+</sup> were optimized and analyzed for better PCR product.

#### **Procedure:**

1. 100µl of reaction mix was prepared with 12.5µl of 10xPCR buffer, 12.2µl 75ng DNA, 10µl dNTP (200µMmM), 1.25µl of Hot Start plus Taq polymerase (2.5U) and 63.75µl of water.
2. 20µl of master mix was distributed in to four 0.2ml tubes contained different primer concentration of 2.5pmole, 10pmole, 15pmole and 20pmole.
3. PCR cycle was same as in section.4.7.2.
4. 5µl of each sample was analyzed on 0.8% agarose gel in TAE buffer and compare with 1kb ladder as mention in section 4.4.

#### **4.7.4 Developing of PCR protocol with various Mn and dITP concentrations**

Different mutagenesis PCR conditions with various combinations of Manganese and dITP were used to obtain a suitable concentration for construction of random library.

Various Mn and dITP concentration were as follows:

- a) 40 $\mu$ M Mn
- b) 80 $\mu$ M Mn
- c) 80 $\mu$ M Mn+40 $\mu$ M dITP
- d) 80 $\mu$ M Mn+80 $\mu$ M dITP
- e) 120 $\mu$ M + 80 $\mu$ MdITP
- f) 120 $\mu$ MMn+ 120 $\mu$ MdITP

#### **Procedure:**

1. The amplification was done in a programmable thermal cycler by the following protocol: 200 $\mu$ M dNTP, 0.5mM MgCl<sub>2</sub>, 1xPCR buffer, 15pmole of each primer Mut1 and Mut3, 2.5 U Hot star Taq polymerase and 150ng template DNA (plasmid pTZ1).
2. 5 $\mu$ l of each sample was analyzed on 0.8% agarose gel in TAE buffer and compare with 1kb ladder as mention in section 4.4

## 4.8 Screening for $\beta$ -Glucosidase on Plates

$\beta$ -glucosidase is an enzyme hydrolyzes  $\beta 1 \rightarrow 4$  bonds (Beta-glucosidic) linking two glucose molecule such as cellobiose (cello-oligosaccharides) (Shewale, 1982).

$\beta$ -Glucosidase activity can be measured in two ways.

Monitor the release of glucose from a natural substrate cellobiose

Monitor the release of a product (color) from analogs of cellobiose for example

5-Bromo-4-chloro-3-indolyl  $\beta$ -D-glucopyranoside and *p*-nitrophenyl-beta-D-1,4-glucopyranoside (*p*NPG) (Liu et al., 2009).

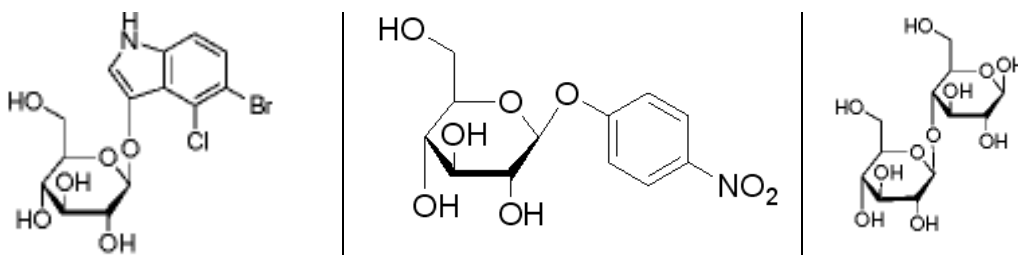


Figure 12: Chemical structure of 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-glucopyranoside, *p*-Nitrophenyl  $\beta$ -D-glucopyranoside and cellobiose

([www.sigmaaldrich.com](http://www.sigmaaldrich.com))

### 4.8.1 Creating Random Library of LacZ Gene

#### Procedure:

1. Plasmid containing LacZ gene was amplified with various Mn and dITP concentration as mention in section 4.7.4
2. Two sets of reaction for the above (see section 4.7.4) concentration were prepared. The amplification was done in a programmable thermal cycler for the following reaction 200 $\mu$ M dNTP, 0.5mM MgCl<sub>2</sub>, 1xPCR buffer, 15pmole of primers Mut1 and Mut3, 2.5 U Hot star Taq polymerase and 150ng template DNA.
3. PCR program was same as mention in section 4.7.2
4. 5 $\mu$ l of each sample was analyzed on 0.8% agarose gel in TAE buffer and compare with 1kb ladder as mention in section 4.4
5. Each of the set was treated with 0.5 $\mu$ l of DpnI enzyme to cleave methylated parental DNA.



6. One of the amplified plasmid set was purified using Qiagen PCR purification kit and another set DNA was precipitated after the addition of 100 $\mu$ l of T.E buffer.

#### **4.8.2 Transformation**

*E.coli* Xl1Blue competent cells was transformed with purified PCR product by chemical method and subsequently plated onto different screening plates containing different substrates corresponding to enzymes of interest.

#### **4.8.3 Screening on M9-cellobiose plates**

This screening method was based on the assumption that the survival of the cells is dependent on the degradation of cellobiose into glucose molecule. Transformed cells were grown on a minimal media M9 with 0.4% cellobiose (Liu et al., 2009) and IPTG were added to the final concentration of 0.05 mM.

#### **4.8.4 Screening on LB- cellobiose plates**

Screening on Luria Bertani media (LB-media) containing cellobiose depend on the fast growth of *E.coli* and a small activity of beta-glucosidase can be pushed with the help of other nutrient present in the media. The library was plated on LB- plates with 0.4% cellobiose and chromogenic substrate X-glu in concentration of 40 $\mu$ g/ml.

#### **4.8.5 Screening with Filter Lift Assay**

PCR product treated with 40 $\mu$ M Mn, 80 $\mu$ l Mn and 80 $\mu$ M Mn+ 40 $\mu$ M dITP was screened for  $\beta$ -Glucosidase activity. Amplification was performed as same as mention in section 4.7.2. The library was transformed on LB+ Plates with 200 $\mu$ g/ml ampicillin and IPTG.

Plates were incubated at 37 $^{\circ}$ C for overnight and next day colonies were screened for beta- glucosidase activity by filter lift assay using 82nm nitrocellulose filter (Amershan Hybond-N, GE Healthcare). Colonies were allowed to absorb by filter and transfer to another plate containing LB+, 200 $\mu$ g Ampicillin, IPTG (0.05nM) and 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-glucopyranoside (40 $\mu$ g/ml).

## 4.9 Screening for $\beta$ -Xylosidase on plates

$\beta$ -Xylosidase (1,4- $\beta$ -D-Xylosidase, E.C.3.2.1.37) are the hemicellulase that hydrolyze short xylooligosaccharides into xylose units or hydrolyze  $\beta$ -1,4 glycosidic bonds linking D-xylose residues (Shallom et al., 2005).

Activity of  $\beta$ -Xylosidase can be measured by production of color from chromogenic substrates 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside and 4-Nitrophenyl  $\beta$ -D-xylopyranoside. These substrate analogs consisting of xylose O-linked to a chromophore. The improved variants from the library could be selected on basis of hydrolysis of these substrate analogs (Wagschal et al., 2009).

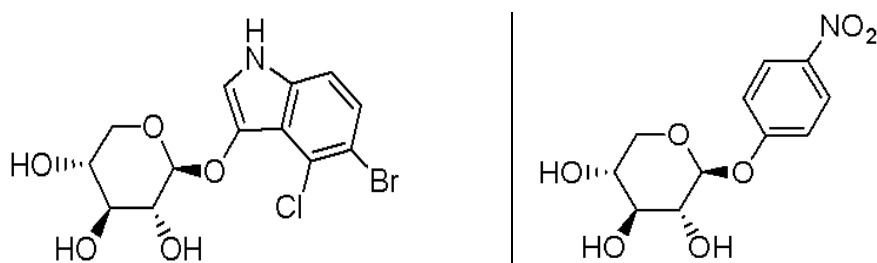


Figure 13: Chemical structure of 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside and 4-Nitrophenyl  $\beta$ -D-xylopyranoside

([www.sigmaldrich.com](http://www.sigmaldrich.com)).

### 4.9.1 Creating Random Library

#### Procedure:

1. Plasmid containing the coding sequence for beta- galactosidase (LacZ gene) was amplified with 2 different Mn and dITP concentrations ( $40\mu\text{M Mn} + 40\mu\text{M dITP}$  and  $80\mu\text{M Mn} + 40\mu\text{M dITP}$ ).
2. The amplification was done in a programmable thermal cycler by the following protocol:  $200\mu\text{M dNTP}$ ,  $0.5\text{mM MgCl}_2$ , 1xPCR buffer, 15pmole of primer Mut1 and Mut3, 2.5 U Hot Start plus Taq polymerase and 150ng template DNA.
3. PCR program was same as mention in section 4.7.2.
4.  $5\mu\text{l}$  of each sample was analyzed on 0.8% low melting agarose gel in TAE buffer and compare with 1kb ladder.
5. Each of the set was treated with  $0.5\mu\text{l}$  of DpnI enzyme to cleave methylated parental DNA.

#### **4.9.2 Second round of mutagenesis**

Plasmid DNA isolated from colony 1, 2, 3 was used as a template for the second round of mutagenesis. Error Prone PCR was performed on the DNA templates isolated from colony 2, 3 using concentration of 80 $\mu$ M Mn<sup>+</sup> 40 $\mu$ ldITP. Colony 1 as a template was treated with 40 $\mu$ MMn +40 $\mu$ ldITP. PCR cycling parameters were identical to the first PCR (section 4.7.2).

#### **4.9.3 Second round for colony 9 (F9) DNA**

DNA isolated from plasmid F9 was used as a template for the second round of mutagenesis. PCR reaction was performed with 80 $\mu$ M Mn<sup>+</sup> 40 $\mu$ M dITP with 10pmole of primers LacZseqI and LacZWPA. PCR cycle was activated at 95°C for 5minutes, 30 cycle of denaturation 94°C for 20 second, annealing 55°C for 45 second, extension 72°C for 4 minutes, 1 cycle of additional extension at 72°C for 12 minutes and reaction was terminated by holding at 4°C.

#### **4.9.4 Screening and Selection**

100 $\mu$ l of TE buffer was added to each of 20 $\mu$ l of PCR product. *E. coli* transformation was performed by adding 100 $\mu$ l of XL1Blue super competent cells to each of the total 120 $\mu$ l of PCR product. 1 $\mu$ l of DMSO and 1.7 $\mu$ l (1.42M) beta-mercepto ethanol were added, tubes were incubated for 30 minutes on ice. Cells were heat shocked by incubating at 42°C for 45 second and transferred back on ice for 2 minute. 3ml of SOC medium was transferred to each of the tube and incubated at 37°C for 1 hour and subsequently plated on the LB+ plates with 200 $\mu$ g of ampicillin, 0.05mM IPTG and X-Xyl (5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside) with the final concentration of 40 $\mu$ g/ml.

#### **4.9.5 Sequencing Analysis**

Sequencing of the selected mutants was performed according to the Big-dye protocol described earlier in section 4.6. The sequencing reaction performed with 8 primers (LacZII 669F, LacZII 1572F, LacZII 2577F, LacZII 3600F, LacZII 4100R,

LacZII 3770R, LacZII 2740R, LacZII 1768R) and the reaction was performed same as mention in section 4.6.

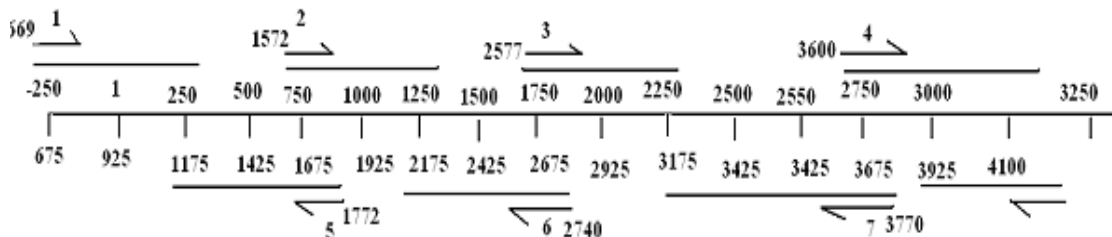
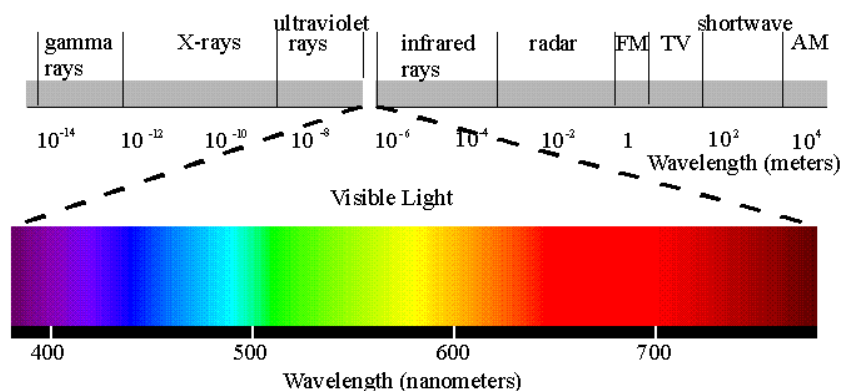


Figure 14: Sequencing primers and their locations on lacZ gene

#### 4.9.6 Colorimetric $\beta$ -xylosidase Enzyme Assay

Variation of the color which makes the change in concentration of some components called the colorimetric analysis. Natural or artificial white light (Visible spectrum 400-760nm) is used in visual colorimeter and the visual appearance of color obtain from the absorption of certain wavelengths of incident light by the colored substance (Vogel and Mendham, 2000). Activity of Beta-Xylosidase was measured by the release of 4-nitrophenol from the chromogenic substrate 4-nitrophenyl-Beta-D-xylopyranoside (NPh-xyl). One unit of the Beta-Xylosidase activity was defined as the amount of the enzyme releasing 1 mole of *p*- nitrophenol from NPh-xyl per minute (Eneyskaya et al., 2007).



[Table of Contents](#)

[Visual Stimulus](#)

Figure 15: Showing the visible range of the spectrum

(<http://www.yorku.ca/eye/spectru.htm>)

**Procedure:**

1. Selected mutants were inoculated from MgCl<sub>2</sub> stock in 4ml of LB+ media with 200µg of ampicillin.
2. Tubes were allowed to shake for 5-6 hours at 37°C
3. 1mM IPTG was added after 5-6 hours of incubation at 37°C.
4. Optical density at 590nm was measured for all the overnight culture.
5. 1ml of each variant transferred into 3 separate tubes (triplicate).
6. Cells were spin down at 13000rpm for 20 minutes at 4°C.
7. Supernatant was removed and pellet was resuspended in 50µl of bugbuster with benzonase (1µl/1ml).
8. Pellets were resuspended properly by vortexing followed by incubating on mild shaker for 20 minutes.
9. Cell lysate was centrifuged at 13000rpm for 20 minutes at 4°C.
10. 50µl of clear cell lysate was transferred to the 96 well micro-titer plates in triplicates.
11. 150µl of substrate solution (60mM Na<sub>2</sub>HPO<sub>4</sub>, 40mM NaH<sub>2</sub>PO<sub>4</sub>, 10mM KCl, 20µg/ml CTAB, 6mM substrate (4-nitrophenyl-Beta-D-xylopyranoside) and 10mM DTT) was transferred to the corresponding well.
12. Absorbance was measured in soft max microplate reader at 405nm for 30 minutes at 37°C.

**4.10 Screening for improved β-Galactosidase Activity**

Beta galactosidase is a hydrolyzing enzyme and has two catalytic activities. It hydrolyzes lactose into galactose and glucose. It also catalyzes lactose to a natural inducer and a disaccharide allolactose (Matthews, 2005). Detection of beta galactosidase activity can be achieved by chromogenic or histochemical substrate 5-Bromo-4-chloro-3-indolyl -D-galactopyranoside which is cleaved by Beta-galactosidase to produce insoluble blue precipitate and colorimetric substrate *ortho*-Nitrophenyl-β-galactoside (ONPG) hydrolyzed by beta galactosidase produce galactose and *ortho* nitrophenol (WC et al., 1990, Matsumura and Rowe, 2005)

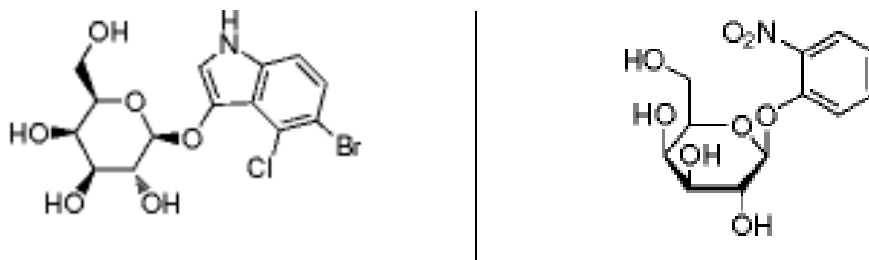


Figure 16: Chemical structure of 5-Bromo-4-chloro-3-indolyl β-D-galactoside (X-Gal) and 2-Nitrophenyl β-D-galactopyranoside ([www.sigmaaldrich.com](http://www.sigmaaldrich.com)).

#### 4.10.1 Creating Random library and screening

The error prone library of LacZ gene was created with 40μMMn+40μMdITP and 80μMMn+40μMdITP. All parameters for PCR, gel electrophoresis and transformation are same as mention earlier. Transformed cells were plated on LB+ plates with 200μg/ml of ampicillin 0.05mM IPTG and X-gal in concentration of 40μg/ml for screening of better beta galactosidase activity in compare to wild type.

#### 4.10.2 Colorimetric β-galactosidase Enzyme assay

Beta gal assay was performed same as xylosidase assay (see section 4.9.6). The difference was in substrate for beta gal detection. The substrate reaction buffer was include with 1mg/ml of ONPG, 60mM Na<sub>2</sub>HPO<sub>4</sub>, 40Mm NaH<sub>2</sub>PO<sub>4</sub>, 10Mm KCl, 20μg/ml CTAB, 6mM substrate and 10mM DTT.

#### 4.11 Mutagenesis in the coding sequence of LacZ gene

Our method to generate molecular diversity of coding sequences by error prone PCR (Xu et al., 1999, Fenton et al., 2002a) was also used in this project. The Error prone PCR of the Beta galactosidase gene done using various concentrations of Mn and dITP. Generated PCR sequences were cloned into the vector (pMREvec2). After transformation in *E. coli* X11Blue competent cells, screening for colonies exhibiting different glycoside hydrolyses activities were performed.

Two sets of each of the PCR reaction was performed with 40μMMn+40μMdITP and 80μMMn+ 40μMdITP, 0.5mMMgCl<sub>2</sub>, 1xPCR buffer, 2.5U of Hot Start Taq polymerase,

200 $\mu$ M dNTP, 300ng of template DNA, 10pmole of primer Mut1 and Mut2 in total 100 $\mu$ l of reaction.

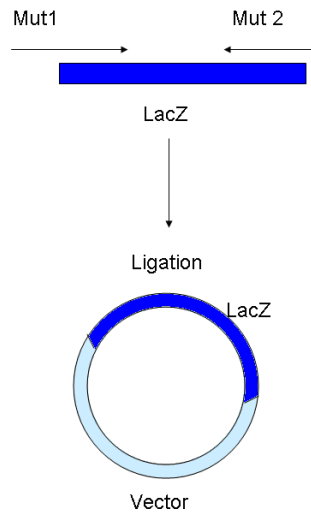


Figure 17: Overview of random mutagenesis using coding sequence of LacZ  
(Adopted from (Xu et al., 1999))

#### 4.11.1 Phenol chloroform DNA Extraction and Precipitation

Phenol chloroform mixture (1:1) was used occasionally to further purify DNA. Phenol works to denature the protein completely but not RNase. This problem can be solved by combination of phenol: chlorophorm. Phenol traces can be removes by further extraction with chloroform from the nucleic acid (Sambrook et al., 1989).

##### Procedure:

1. 100 $\mu$ l of each set of PCR product was combined in a single 1.5 ml tube.
2. 200 $\mu$ l of phenol: chloroform was added in 1:1 ratio.
3. Tubes were mixed and incubated for 10 minutes on bench.
4. Tubes were spin down for 10 minutes at 13000rpm
5. Two layer were visible, organic phase (bottom layer) was removed carefully.
6. Another 200 $\mu$ l of chloroform was added, incubated for 10 minutes and centrifuged for another 10 minutes at 13000 rpm.
7. Top aqueous layer containing DNA, was removed in a fresh 1.5ml tube

### **Precipitation:**

This technique is used to concentrate and removal of salt from the nucleic acid in aqueous solution. Using salt as sodium acetate (CH<sub>3</sub>COONa) neutralizes the charge on the sugar phosphate backbone. Na ion as positive charge neutralizes the negative charge on the phosphate group of DNA that makes molecule less water soluble. Ethanol helps in interacting sodium ion with phosphate group and help in making DNA more hydrophobic as well it also help in removing residual salt from DNA pellet (Sambrook et al., 1989) and [\[9.2F\]](#)

Procedure:

1. 20µl of sodium acetate and 600µl of 100% ethanol was added.
2. Tubes were mixed and incubated on dry-ice+ ethanol for 15 minutes.
3. Tubes were centrifuged for 20 minutes at 13000 rpm.
4. Supernatant was decanted and pellet was washed with 400µl of 70% ethanol.
5. Supernatant was decanted and pellet was allowed to dry.
6. Air-dry pellet was resuspended in 100µl of TE buffer.
7. DNA concentration was measured on Nanodrop<sup>®</sup>1000.

### **4.11.2 Restriction Digestion and Ligation**

1. 4µg of vector (pMREVec2) was digested with 2µl of BamHI and 2µl of SalI with 4µl of NE buffer 3 in total 40µl of reaction.
2. 9.6µg of insert DNA for both of the PCR product was digested with BamHI and SalI in total reaction of 30µl.
3. Reaction tubes were incubated at 37°C for 3 hours.
4. After 2 hour SAP (Shrimp alkaline phosphatase) was added to the vector and incubated further for another 1 hour at 37°C.
5. Enzymes were heat and inactivated by incubating at 85°C for 15 minutes.
6. Ligation reaction was performed with 1:6 vector insert ratio.
7. Precipitation reaction followed by air drying the tubes to obtain concentrated DNA.
8. Precipitated DNA (Vector: insert) was resuspended in 12µl of T.E buffer.
9. Reaction was incubated at 65°C warm water and allow to cool down to 20-25°C.



10. 1.5µl of ligase buffer and 1.5µl of ligase enzyme added and incubated at 16°C for overnight.
11. Next day tubes were placed on ice and transformation into *E. coli* XL1Blue competent cells. The procedure performed same as mentioned earlier.

#### **4.11.3 Screening with Filter Lift Assay**

Transformation mix was plated on LB+, 200µg of ampicillin, 0.05 mM IPTG and 40µg/ml of 5-Bromo-4-chloro-3-indolyl β-D-glucopyranoside (X-Glu) to screen any mutant with Beta-glucosidase activity.

Next day colonies were absorbed to nitrocellulose filters (82nm) and transferred (colony side up) to another LB+ ampicillin plate containing 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-Gal).

#### **4.11.4 Colorimetric Enzyme assay**

Assay was performed same as mentioned in section 4.9.6 with 1mg/ml concentration of ONPG as a substrate

### **4.12 Preparation of competent cells**

Transformation is a technique to introduce DNA into the cells and the state of bacteria in which they can take up naked DNA from the surrounding called competence (Snyder and Champness, 2007). Mandel and Higa (1970) discovered that CaCl<sub>2</sub> treatment allows the *E. coli* to uptake bacteriophage DNA (Primrose et al., 2006). Methods were provided by personal communication with Prof. Raafat El-Gewely.

#### **4.12.1 CaCl<sub>2</sub> method**

##### **Procedure:**

1. Single colony of XL1Blue was inoculated in 20ml of LB+ media with 15µg/ml of tetracycline and incubated for overnight at 37°C.
2. Subculture 50ml LB+ containing 10mM MgCl<sub>2</sub> with 2ml of overnight culture and shake at 37°C until OD reached 0.25-0.35 at absorbance 600nm.
3. Flask was chilled on ice.
4. Cell was centrifuged at 7000rpm for five minutes at 4°C using pre-chilled tubes.

5. Cells were resuspended in 25ml of cold 0.1 M  $MgCl_2$  and again centrifuged as same as above.
6. 12.5ml of ice cold T-salt solution containing 75mM  $CaCl_2$ , 6mM  $MgCl_2$  was added and incubated on ice for 30 minute.
7. Cells were centrifuges and resuspended in 1.5ml T salts containing 20% of glycerol.
8. 100ml and 200 $\mu$ l of cells suspension was transferred to freezing tubes and stored at  $-70^\circ C$ .

#### **4.12. 2 RbCl method**

##### **Procedure:**

Day 1: XL1Blue colony from freezer stock was streak on LB+ plates with 30 $\mu$ g of tetracycline and incubated for overnight at  $37^\circ C$ .

Day 2: Single colony was inoculated in 10ml of LB+ media with 15 $\mu$ g/ml of tetracycline and incubated at  $37^\circ C$  on shaker for 6-8 hours. From this culture another 10ml of LB+ media was inoculated and incubated for overnight at the same temperature.

Day3: 2.5ml of overnight culture was inoculated into 250ml of prewarmed Psi media and incubate at  $37^\circ C$  till the OD reach upto 0.4 at 600nm absorbance. Tubes were incubated on ice for 15 minutes followed by 5 minutes spinning at 4500 rpm. Cells were resuspended in 100ml of cold TFBI buffer (30mM KAc, 100mM RbCl, 10mM  $CaCl_2$ , 50mM  $MnCl_2$  and 15% glycerol). Cells were incubated on ice for 5 minutes and centrifuged at 4500rpm for another 5minutes. Cells were resuspended in 7.5ml of TFB2 buffer (10mM MOPS, 75mM  $CaCl_2$ , 10mM RbCl and 15% glycerol). 100 $\mu$ l of resuspended cells were transformed in 1.5ml of microcentrifuge tubes and stored at  $-70^\circ C$ .

## 5. Results

### 5.1 Sequencing of pTZ1 Plasmid Encoding Wild Type LacZ Gene

In order to investigate the complete sequence, location of restriction sites and to design primers for whole plasmid mutagenesis as well as for creating random library of LacZ itself, sequencing of plasmid pTZ1 (Su et al., 1990) was our first step in this project. The sequencing primers were initially designed from plasmid pZ1918, a derivative of pTZ1 (<http://www.ncbi.nlm.nih.gov/nuccore/434652>). We have designed 6 primers for the vector and 8 primers for LacZ coding sequence to sequence the entire complete pTZ1 plasmid. Su et al., 1990 have indicated that the size of the plasmid is 5.2kb, (Su et al., 1990) but our sequence results showed that the plasmid in fact is 5,502bp.

```
CGAATTCTCATGTTTGACAGCTTATCATCGGGAGCTGCATGTGTCAGAGGTTT
TCACCGTCATCACCGAAACGCGCGAGGCAGCTGTTGACAATTAATCATCGGC
TCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACACAGGAAACAGG
ATCCGACAACCGATGAAAGCGGCGACGCGCAGTTAATCCCACAGCCGCCAGT
TCCGCTGGCGGCATTTTAACTTTCTTTATCACACAGGAAACAGCTATGACCAT
GATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTG
GCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGT
AATAGCGAAGAGGCCCGCACCGATCGCCCTTCCAACAGTTGCGCAGCCTGA
ATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCCAGAAGCGGTGCCGAAAG
CTGGCTGGAGTGCATCTTCTGAGGCCGATACTGTCGTCGTCCCCTCAAAC
GGCAGATGCACGGTTACGATGCGCCCATCTACACCAACGTGACCTATCCCAT
TACGGTCAATCCGCCGTTTGTTCACGGAGAATCCGACGGGTTGTTACTCGC
TCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTAT
TTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTCG
GTTACGGCCAGGACAGTCGTTTGCCGCTGAATTTGACCTGAGCGCATTTTTA
CGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTGCGCTGGAGTGACGGCA
GTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGT
CTCGTTGCTGCATAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTC
GCTTTAATGATGATTTACGCCGCGCTGTACTGGAGGCTGAAGTTCAGATGTGC
GGCGAGTTGCGTGACTACCTACGGGTAACAGTTTCTTTATGGCAGGGTGAAA
CGCAGGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGAGCG
TGGTGGTTATGCCGATCGCGTCACACTACGTCTGAACGTCGAAAACCCGAAA
CTGTGGAGCGCCGAAATCCCGAATCTCTATCGTGCGGTGGTTGAACTGCACA
CCGCCGACGGCACGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGA
GGTGCGGATTGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATT
CGAGGCGTTAACCGTCACGAGCATCATCCTCTGCATGGTCAGGTCATGGATG
AGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAACAACCTTTAACGC
CGTGCGCTGTTTCGCATTATCCGAACCATCCGCTGTGGTACACGCTGTGCGACC
```

GCTACGGCCTGTATGTGGTGGATGAAGCCAATATTGAAACCCACGGCATGGT  
GCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGATGAGCGAA  
CGCGTAACGCGAATGGTGCAGCGCGATCGTAATCACCCGAGTGTGATCATCT  
GGTCGCTGGGGAATGAATCAGGCCACGGCGCTAATCACGACGCGCTGTATCG  
CTGGATCAAATCTGTGATCCTTCCC GCCCGGTGCAGTATGAAGGCGGGCGGA  
GCCGACACCACGGCCACCGATATTATTTGCCGATGTACGCGCGCGTGGATG  
AAGACCAGCCCTTCCC GGCTGTGCCGAAATGGTCCATCAAAAAATGGCTTTC  
GCTACCTGGAGAGACGCGCCCCTGATCCTTTGCGAATACGCCACGCGATG  
GGTAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGCGTTTCGTCAGTATCC  
CCGTTTACAGGGCGGCTTCGTCTGGGACTGGGTGGATCAGTCGCTGATTAAT  
ATGATGAAAACGGCAACCCGTGGTCGGCTTACGGCGGTGATTTTGGCGATAC  
GCCGAACGATCGCCAGTTCTGTATGAACGGTCTGGTCTTTGCCGACCGCACG  
CCGCATCCAGCGCTGACGGAAGCAAAACACCAGCAGCAGTTTTTCCAGTTCC  
GTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAG  
CGATAACGAGCTCCTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCA  
AGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAGGTAAACAGTTGATTGAAC  
TGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGGCTCACAGTACG  
CGTAGTGCAACCGAACGCGACCGCATGGTCAGAAGCCGGGCACATCAGCGC  
CTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCGCCGCG  
TCCCACGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCGAGCT  
GGGTAATAAGCGTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTCACAGATGT  
GGATTGGCGATAAAAAACAACCTGCTGACGCCGCTGCGCGATCAGTTCACCCG  
TGCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGCATTGACCCT  
AACGCCTGGGTGCAACGCTGGAAGGCGGGCGGGCCATTACCAGGCCGAAGCA  
GCGTTGTTGAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATTACGA  
CCGCTCACGCGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGGAAAAC  
CTACCGGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTG  
GCGAGCGATACACCGCATCCGGCGCGGATTGGCCTGAACTGCCAGCTGGCGC  
AGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCGCAAGAAAACCTATC  
CCGACCGCCTTACTGCCGCCTGTTTTGACCGCTGGGATCTGCCATTGTCAGAC  
ATGTATACCCCGTACGTCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGC  
GCGAATTGAATTATGGCCCACACCAGTGGCGCGGGCGACTTCCAGTTCAACAT  
CAGCCGCTACAGTCAACAGCAACTGATGGAAACCAGCCATCGCCATCTGCTG  
CACGCGGAAGAAGGCACATGGCTGAATATCGACGGTTTCCATATGGGGATTG  
GTGGCGACGACTCCTGGAGCCCGTCAGTATCGGCGGAATTCCAGCTGAGCGC  
CGGTCGCTACCATTACCAGTTGGTCTGGTGTCAAAAATAATAATAACCGGGC  
AGGGGGGATCCGTCGACCTGCAGGTCCGGCTGCACCGCAGGTGACTGCAGAA  
GACGCATCTGCCAGCCTGGCAGAACTGCTGAACGCAGGTCTGGGCGGTTCTG  
ATAACGAGTAATCGTTAATCCGCAAATAACGTAAAAACCCGCTTCGGCGGGT  
TTTTTTATGGGGGGAGTTTAGGGAAAGAGCATTTGTCAGAATATTTAAGGAA  
TTTCTGAATACTCATAATCAATGTAGAGATGACTAATATCCTGAAACTGACTG  
AACTAATTGAGTCAAACCTCGGCAAGGATTGATACTATTCTGTGTAACCTTTC  
TTAAGGAACGAGAATGAAACAGGAAGTGGAAAAGTGGCGACCTTTTGGACA  
TCCGACGGATCTCGACCTGCAGCCAAGCTTGGCACTGGCGTTGCTGGCGTTTT  
TCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCA  
GAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGA

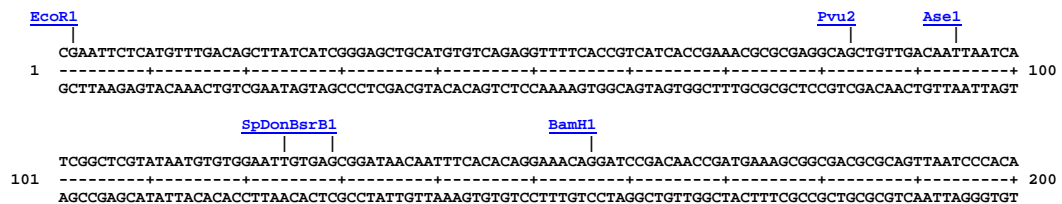
AGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC  
 CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGT  
 ATCTCAGTTCGGTGTAGGTTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCC  
 CCCGTTACGCCCAGCCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAA  
 CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATT  
 AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCTA  
 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCC  
 AGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACC  
 GCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAA  
 AAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG  
 AACGAAAACCTCACGTTAAGGGATTTTGGTTCATGAGATTATCAAAAAGGATCT  
 TCACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATA  
 TATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTAT  
 CTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA  
 GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATA  
 CCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAG  
 CCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCA  
 GTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTT  
 TCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTACGCTCGTCGTTT  
 GGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATC  
 CCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCA  
 GAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA  
 TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC  
 AACCAAGTCATTCTGAGAATAGTGTATGCGGGCAGCCGAGTTGCTCTTGCCCG  
 GCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCA  
 TCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTG  
 AGATCCAGTTCGATGTAACCCACTCGTGCACCCAAGTATCTTCAGCATCTTT  
 TACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA  
 AAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTT  
 TTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATA  
 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCC  
 GAAAAGTGCC

Figure 18: Entire 5,502 base pairs of plasmid pTZ1

Colour indicate-yellow vector sequences, blue-*bla* gene and remaining LacZ

Map created at [http://pga.mgh.harvard.edu/web\\_apps/web\\_map/start](http://pga.mgh.harvard.edu/web_apps/web_map/start) Editor at [http://pga.mgh.harvard.edu/web\\_apps/dna\\_utilities.html](http://pga.mgh.harvard.edu/web_apps/dna_utilities.html)

\*\*\* Sites protected by E. coli methylation are not shown in this map \*\*\*



```

                                SpDon
201  GCGCCAGTTCGCGTGGCGGCATTTAACTTTCTTTATCACACAGGAAACAGCTATGACCATGATTACGGATTCAGTGGCCGTCGTTTTACAACGTCGTG 300
orf 1 > CGGGCGTCAAGGCGACCGCGTAAATTAAGAAAGAAATAGTGTGTCCTTTGTCGATACTGGTACTAATGCCTAAGTGACCGGCAGCAAATGTTGCAGCAC
                                M T M I T D S L A V V L Q R R D
                                SpAcc  Bmr1  Pvu2  Ear1  Pvu1
201  ACTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGACGACATCCCCCTTTCGCGCAGCTGGCGTAATAGCGAAGAGGCCGACCGATCGCCCTTC 400
orf 1 > TGACCCTTTGGGACCGCAATGGGTTGAATTAGCGGAAACGTCGTTAGGGGAAAGCGGTCGACCGCATTATCGCTTCTCCGGGCGTGGCTAGCGGGAAG
                                FspI  BglI  Bsu36  BaeIb  BpmI
401  CCAACAGTTGCGCAGCTGAATGGCGAATGGCGCTTTGCTGGTTCCGGCACCAGAAGCGGTGCCGAAAGCTGGCTGGAGTGCATCTTCTGAGGCC 500
orf 1 > GGTGTCAACGCGTCGACTTACCGCTTACCGCGAAACGGACCAAAGCCGTTGCTTTCGCGCAGGCTTTCGACCGACCTCAGCTAGAAGGACTCCGG
                                Q Q L R S L N G E W R F A W F P A P E A V P E S W L E C D L P E A
                                BaeIa  AleI  EciI
501  GATACTGTCGTGCTCCCTCAAACCTGGCAGATGCACGGTTACAGTGCGCCATCTACACCAACGTGACCTATCCATTACGGTCAATCCGCGCTTTGTC 600
orf 1 > CTATGACAGCAGCAGGGGAGTTGACCGCTACGTGCAATGCTACGCGGGTAGATGTTGGTTCAGTGGATAGGGTAATGCCAGTTAGCGGCAAAACAG
                                D T V V V P S N W Q M H G Y D A P I Y T N V T Y P I T V N P P F V P
                                SpDon  SpAcc  HpaI
601  CCACGGAAATCCGACGGTGTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCAATTAATTTTGTAGGCGTTAACTC 700
orf 1 > GGTGCTCTTAGGCTGCCCAACAATGAGCGAGTGAATAACTACTTTCGACCGATGTCCTCCGGTCTGCGCTTAATAAAAACTACCGAATTGAG
                                T E N P T G C Y S L T F N V D E S W L Q E G Q T R I I F D G V N S
                                BseYI  Bpu10
701  GCGCTTTCATCTGTTGGTGAACGGCGCTGGGTGCGTTACGGCCAGGACAGTCGTTTGGCGTCTGAATTTGACCTGAGCGCATTTTACGCGCCGGAGAA 800
orf 1 > CCGCAAAGTAGACACCACTTGGCCGCGACCCAGCCAATGCCGCTCTGTGACGAAACGGCAGACTTAACTGGACTCGCTAAATGCGCGGCTCTT
                                A F H L W C N G R W V G Y G Q D S R L P S E F D L S A F L R A G E
                                BpmI  BsrB1  EciI  Aat2  BsmB1
801  AACCCCTCGCGGTGATGGTGTGCGTGGAGTGACGGCAGTTATCTGGAAGATCAGGATAATGCGCGGATGAGCGGCATTTCCGTGACGCTCTCGTTC 900
orf 1 > TTGGCGAGCGCCACTACCAGCAGCGACCTCACTGCCGTCAATAGACCTTCTAGTCTATACCCGCTACTCGCGTAAAGGCACTGCAGAGCAACG
                                N R L A V M V L R W S D G S Y L E D Q D M W R M S G I F R D V S L L
                                BpmI
1000  TGCATAAACCGACTACACAAATCAGCGATTCCATGTTGCCACTCGCTTTAATGATGATTTCAGCCGCGTACTGAGGCTGAAGTTTCAGATGTGCGG
orf 1 > ACGTATTTGGCTGATGTTTGTAGTCGCTAAAGGTACAACGGTGAAGCAAAATTAATAAGTTCGCGCGACATGACCTCCGACTTCAAGTCTACACGCC
                                H K P T T Q I S D F H V A T R F N D D F S R A V L E A E V Q M C G
                                Eco57  BspM1  ClaI
1100  CGAGTTGCGTACTACCTACGGTAAACAGTTCTTTATGCGAGGGTGAACGCGAGTGCACGCGCCAGCGCCTTTCGGCGGTGAATTAATCGATGAG 1200
orf 1 > GCTCAACGCACTGATGATGCCATTGTCAAAGAAATACCGTCCACTTTGCGTCCAGCGGTGCGGTGGCGCGGAAAGCCGCACTTTAATAGCTACTC
                                E L R D Y L R V T V S L W Q G E T Q V A S G T A P F G G E I I D E
                                PvuI  BsgI
1200  CGTGGTGGTTATGCCGATCGCGTCACACTACGTCTGAACGTCGAAACCCGAACTGGAGCGCCGAAATCCGATCTCTATCGTGGGTGGTTGAAC 1300
orf 1 > GCACCACAAATACGGCTAGCGCAGTGTGATGCAGACTTGACGCTTTTGGGCTTTGACACCTCGCGCTTTAGGGCTTAGAGATAGCACGCCACCAACTTG
                                R G G Y A D R V T L R L N V E N P K L W S A E I P N L Y R A V V E L
                                BtgZ1
1300  TGCACCCGCCGACGGCAGCTGATTGAAGCAGAAGCCTGCGATGTCGTTTCCGCGAGGTGCGGATTGAAATGGTCTGCTGCTGTAACGGCAAGCC 1400
orf 1 > ACGTGTGGGGCTCCGCTGCGACTAACTTCGCTTCGGACGCTACAGCCAAAGGCGTCCACGCTAACTTTTACCAGCAGCAGGACTTCCGCTTCGG
                                H T A D G T L I E A E A C D V G F R E V R I E N G L L L N G K P
                                HpaI  BssS1  EcoRV  BsgI
1400  GTTGCTGATTCGAGGCGTTAACCGTCAGGACATCATCTCTGATGGTCAGGTCATGGATGAGCAGAGATGGTGCAGGATATCCTGCTGATGAAGCAG 1500
orf 1 > CAACGACTAAGCTCCGCAATTGGCAGTGTCTAGTAGGAGACGTACAGTCCAGTACCTACTCGTCTGCTACCACTCTATAGGACGACTACTTCGTC
                                L L I R G V N R H E H H P L H G Q V M D E Q T M V Q D I L L M K Q
                                Dra3  SspI
AACAACTTTAAACCGCTGCGCTGTTTCGCAATTATCCGAACCTCCGCTGTTGACAGCTGTGCGACCGCTACGGCTGTATGGTGGATGAAGCCAATA

```

```

1500 1401 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
orf 1 > TTGTTGAAATTCGGGCACGCGACAAGCGTAATAGGCTTGGTAGGCGACACCATGTGCGACACGCTGGCGATGCCGGACATACACACCTACTCTGGTTAT
> N N F N A V R C S H Y P N H P L W Y T L C D R Y G L Y V V D E A N I

                                     PflM1                                     Bcg1b
                                     MluI                                     BtgZ1
                                     PvuI
1501 TTGAAACCCACGGCATGGTGCCTAATGAATCGTCTGACCGATGATCCGCGTGGCTACCGCGATGAGCGAACCGGTAACGCGAATGGTGCAGCGCGATCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1600 AACTTTGGGTGCCGTACCACGGTTACTTAGCAGACTGGCTACTAGGCGGACCGATGGCCCGTACTCGCTTGCGCATTGCGCTTACCACGTCCGCGTAGC
orf 1 > E T H G M V P M N R L T D D P R W L P A M S E R V T R M V Q R D R

                                     Bcg1a
                                     AleI
                                     Bsg1                                     BseY1
1601 TAATCACCCGAGTGTGATCATCTGGTCGCTGGGGAATGAATCAGGCCACGGCGCTAATCAGACGCGCTGTATCGCTGGATCAAATCTGTCGATCCTTCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1700 ATTAGTGGGCTCACACTAGTAGACCGACCCCTTACTTAGTCCGCTGCCGCGATTAGTGTGCGCGACATAGCGACCTAGTTTAGACAGCTAGGAAGG
orf 1 > N H P S V I I W S L G N E S G H G A N H D A L Y R W I K S V D P S

                                     Bsg1                                     Eci1
                                     BssH2                                     Bbs1
1701 CGCCCGGTGCAGTATGAAGCGCGGAGCCGACACCAACGGCCACCGATATTATTTGCCCGATGTACCGCGCGTGGATGAAGACACCGCCCTTCCCGGCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1800 GCGGGCCACGTCTACTTCCGCCCTCGGCTGTGGTGGCTATAATAAACGGGTACATGCGCGCGCACCTACTTCTGGTGGGAAGGGCCGAC
orf 1 > R P V Q Y E G G G A D T T A T D I I C P M Y A R V D E D Q P F P A V

                                     SpAcc
                                     BsmB1
                                     PflM1                                     BtgZ1
1801 TGCCGAAATGGTCCATCAAAAAATGGCTTTCGCTACCTGGAGAGACCGCCCGCTGATCCTTTGCGAATACGCCACGCGATGGGTAACAGTCTTGGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1900 ACGGCTTTACCAGGTAGTTTTTACCAGAAAGCGATGGACCTCTCTGCGGGCGGACTAGGAAACGCTTATGCGGGTGCCTACCCATTGTCAGAACCAGCC
orf 1 > P K W S I K K W L S L P G E T R P L I L C E Y A H A M G N S L G G

                                     Bcg1b
                                     BciV1
                                     Bcg1a
                                     Bmr1 AlwN1
1901 TTTGCTAAATACTGGCAGGCCTTTTCGTAGTATCCCGCTTACAGGGCGGCTTCGTCTGGGACTGGGTGGATCAGTCGCTGATTAAATATGATGAAAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2000 AAAGCGATTTATGACCGTCCGCAAAGCAGTCAATAGGGCAAATGCCCGCAAGCAGACCCTGACCCACCTAGTCAGCGACTAATTTACTACTTTTG
orf 1 > F A K Y W Q A F R Q Y P R L Q G G F V W D W V D Q S L I K Y D E N

                                     PvuI
2001 GGCAACCCGTGGTCCGCTTACGGCGGTGATTTTGGCGATACCGCAACGATCGCCAGTCTGTATGAACGGTCTGGTCTTTGCCGACCGCAGCCGCGATC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2100 CCGTGGGCACCAGCCGAATGCCGCCACTAAAACCGCTATGCGGCTGTAGCGGTCAAGACATACTTCCAGACCAGAAACCGCTGGCGTGGCGGTAG
orf 1 > G N P W S A Y G G D F G D T P N D R Q F C M N G L V F A D R T P H P

                                     AfeI
                                     SpAcc
                                     Bsg1
2101 CAGCGCTGACGGAAGCAAAACACCAGCAGCAGTTTTCCAGTTCGGTTTACCGGGCAAACCTCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2200 GTCGCGACTGCCTTCGTTTTGTGGTCTGTCGTAAGGTCAGGCAAATAGGCCCGTTTGGTAGCTTCACTGGTCTGCTTATGACAAAGGCAGTATCGCT
orf 1 > A L T E A K H Q Q Q F F Q F R L S G Q T I E V T S E Y L F R H S D

                                     SacI
                                     Ecl2
                                     SpDon
                                     SpDon
2201 TAACGAGCTCCTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCCTGGATGTCGCTCCACAAGGTAACAGTTGATTGAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2300 ATTGCTGAGGACGTGACCTACCACCGCAGCTACCATTCGGCGACCGTTCCGCACTTACGGAGACCTACAGCGAGGTGTTCCATTGTCACAACTACT
orf 1 > N E L L H W M V A L D G K P L A S G E V P L D V A P Q G K Q L I E

                                     MluI
                                     Drd1
2301 CTGCCTGAATACCGCAGCGAGAGCGCCGGCACTCTGGCTCACAGTACGCGTAGTGCAACCGAACCGACCGCATGGTTCAGAAGCCGGGCACATCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2400 GACGGACTTGATGGCTCGGCCTCTCCGGCCCGTTGAGACCGAGTGTATCGGCATCAGTTGGCTTGGCTTGGCTGGCGTACCAGTCTTCGGCCCGGTAGT
orf 1 > L P E L P Q P E S A G Q L W L T V R V V Q P N A T A W S E A G H I S

                                     Bts1
                                     Eci1
                                     Chi
                                     XcmI
                                     BstX1
2401 GCGCCTGGCAGCTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCGCCGCGTCCCACGCCATCCCGCATCTGACCACCGGAAATGGATTTTTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2500 CGCGGACCGTCTCACCAGACCGCCTTTTGGAGTCACTGCGAGGGGGCGGCGAGGGTCCGGTAGGGCGTAGACTGGTGGTTCGCTTTACTAAAAAC
orf 1 > A W Q Q W R L A E N L S V T L P A A S H A I P H L T T S E M D F C

                                     BseY1
                                     BglI
                                     SpAcc
                                     AleI

```

```

CATCGAGCTGGGTAATAAGCGTTGGCAAATTAACCGCAGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACCTGCTGACGCCGCTGCGC
2501 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2600
GTAGCTCGACCCATTATTGCAACCGTTAAATTTGGCGGTGAGTCCGAAAGAAAGTGTCTACACCTAACCGCTATTTTTTTGTTGACGACTGCGGGCAGCGG
orf 1
> I E L G N K R W Q F N R Q S G F L S Q M W I G D K K Q L L T P L R

          ApaI
GATCAGTTACCCGTGCACCGCTGATAACGACATTGGCGTAAGTGAAGCGACCCGATTGACCTAACCGCTGGGTGCAACGCTGGAAGGCGGCGGCC
2601 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2700
CTAGTCAAGTGGGACGTTGGCGACTATTGCTGTAACCCGATTCACTTCGCTGGGCGTAAGTGGGATTGCGGACCCAGCTTGCAGCTTCCGCGCCGCGG
orf 1
> D Q F T R A P L D N D I G V S E A T R I D P N A W V E R W K A A G H

          BtsI          MluI
          ApaI          BsrBI
ATTACCAGGCCGAAGCAGCTTGTTCAGTGCACGGCAGATACACTTGTCTGATGCGGTGCTGATTACGACCCGCTCACGCTGGCAGCATCAGGGGAAAAAC
2701 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2800
TAATGGTCCGCTTCGTCGCAACAGTCAAGTCCGCTCTATGTGAACGACTACGCCACGACTAATGCTGGCGAGTGCACCCGTCGTAGTCCCTTTTG
orf 1
> Y Q A E A A L L Q C T A D T L A D A V L I T T A H A W Q H Q G K T

          Pvu2          BsrBI
          BspMI
ATTGGCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAACTGGCTCGGATTAGGGCCCAAGAAAATATCCCGACCCGCTTACTGCCGCTGTT
2901 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3000
TAACCGGACTTGACGGTCGACCGCGTCCATCGTCTGCCCATTTGACCGAGCCTAATCCCGCGTCTTTTGGATAGGGTGGCGGAATGACGGCGGACAA
orf 1
> I G L N C Q L A Q V A E R V N W L G L G P Q E N Y P D R L T A A C F

          BseI          BspLU BstZ1 BsiW1
          BspMI          BbsI
TTGACCGCTGGGATCTGCCATTGTTCAGACATGTATACCCCGTACGCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGCGCAATGAATTATGGCCC
3001 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3100
AATGGCGACCCCTAGACGGTAACAGTCTGTACATATGGGCATGAGAAGGCTCGCTTTTCCAGACGCGACCCCTGCGCGCTTAACCTTAATACCGGG
orf 1
> D R W D L P L S D M Y T P Y V F P S E N G L R C G T R E L N Y G P

          BstX1          BtgZ1
          AleI          AlwNI          BspI
ACACCAGTGGCGCGGCACTTCCAGTTCACATCAGCCGTACAGTCAACAGCAACTGATGGAAAACCGCATCGCCATCTGCTGCAACGCGGAAGAAGGC
3101 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3200
TGTGGTCAACCGCGCTGAAGTCAAGTTGTAGTCGGGATGTCAGTTGCTGTTGACTACCTTTGGTGGTAGCGGTAGACGACTGCGCTTCTTCCG
orf 1
> H Q W R G D F Q F N I S R Y S Q Q Q L M E T S H R H L L H A E E G

          NdeI XcmI          EcoRI          BspI          BspI
          Pvu2          EciI
ACATGGCTGAATATCGACGGTTTCCATATGGGGATTGGTGGCAGGACTCCTGGAGCCGTCAGTATCGGCGGAATTCAGCTGAGCGCGGTGCTTACC
3201 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3300
TGTACCAGCTTATAGCTGCCAAAGGTATACCCCTAACCCAGCTGCTGAGGACTCGGGCAGTATAGCCGCTTAAAGTGCAGCTCGCGCCAGCGATGG
orf 1
> T W L N I D G F H M G I G G D D S W S P S V S A E F Q L S A G R Y H

          BspMI          AarI
          BspMI          BspMI
          BspI          SbfI          PstI          DrdI
          BamHI BspMI          SpAac PetI          BbsI
ATTACCAGTTGCTGTTGTCAAAAATAATAAACCGGCGAGGGGATCCGTCGACCTGCAGGTGCGGCTGCACCGCAGGTGACTGAGAAGACGATC
3301 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3400
TAATGGTCAACCGACCCAGTTTATTATTATTGGCCGCTCCCCCTAGGCAGCTGGACGTCAGCGGACGTCGGCTCCACTGACGCTTCTTCTGCGTAG
orf 1
> Y Q L V W C Q K *

          BglI          BspMI
TGCCAGCTGGCAGAACTGCTGAACGAGGTCTGGCGGTCTGATAACGAGTAATCGTTAATCCGCAAATAACGTAACAAACCGCTTCGGCGGTTTTT
3401 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3500
ACGGTCGGACCGCTTTCGACTTTCGCTCCAGACCCGCAAGACTATTGCTCATTAGCAATTAGCGCTTTATTGCATTTTGGGCGAAGCCGCAAAAA
          SspI
TTATGGGGGAGTTTAGGGAAAGACATTGTCAGAATATTAAGGAATTTCTGAATACTCATAATCAATGTAGAGATGACTAATATCTGAACTGACT
3501 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
3600
AATACCCCTCAAATCCCTTCTCGTAACAGTCTTATAAATCCCTTAAGACTTATGAGTATTAGTTACATCTCTACTGATTATAGGACTTTGACTGA

```



3601 GAACTAATTGAGTCAAACCTCGCAAGGATTTCGATACTATTCTGTGTAACCTTCTTAAGGAACGAGAATGAAACAGGAAGTGGAAAAGTGGCGACCTTTT  
 3700 CTTGATTAACCTCAGTTTGAGCCGTTCTTAAGCTATGATAAGGACACATTGAAAGAATTCTCTTGCTCTTACTTTGCTCCTCACCTTTTCACCCTGGAAAA

3701 GGACATCCGACGGATCTCGACCTGCAGCCAAAGCTTGGCACTGGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGAC  
 3800 CCTGTAGGCTGCCTAGAGCTGGACGTCGGTTCGAACCGTGACCGCAACGACCGCAAAAAGGTATCCGAGGCGGGGGACTGCTCGTAGTGTCTTTAGCTG

3801 GCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACAGCGGTTTCCCTCGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTCGCCGT  
 3900 CGAGTTCAGTCTCCACCGCTTTGGGCTGCTGATATTTCTATGGTCCGCAAGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGGACGGCGCA

3901 TACCGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGTCTGCTCCAAG  
 4000 ATGGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCAGGAAAGATATCGAGTGCACATCCATAGAGTCAAGCCACATCCAGCAAGCGAGGTTT

4001 CTGGGCTGTGTGCAGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCAC  
 4100 GACCCGACACAGTGTCTGGGGGCAAGTCGGGCTGGCGACGCGGAATAGGCCATTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTG

4101 TGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGATAGGCGGTCTACAGGTTCTTGAAGTGGTGGCCTAACTACGGTACACTAGAAGGAC  
 4200 ACCGTCGTGGTGACCATTTGCTTAATCGTCTCGTCCATACATCCGCCAGTGTCTCAAGAACTTACCACCGGATTGATGCCGATGTGATCTTCTCTG

4201 AGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGCTAGCGGTGTTTTT  
 4300 TCATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGGCCGTTTTGTTGGTGGGACCATCGCCCAAAAA

4301 TTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACTCAC  
 4400 AAACAACGTTCTGCTCTAATGCGCGTCTTTTTTCTAGAGTCTTCTAGGAAACTAGAAAAGATGCCCCAGACTCGGAGTCACTTGTCTTTGAGTG

4401 GTTAAGGATTTTGGTCTAGGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATATAGTA  
 4500 CAATTCCTAAAACAGTACTCTAATAGTTTTTCTAGAGTGGATCTAGGAAAATTTAATTTTACTTCAAAATTTAGTTAGATTTTATATATACTCAT

4501 AACTTGTCTGACAGTTACCAATGCTTAATCAGTGAAGCCACTATCTCAGCGATCTGCTATTTCTGTTCCATAGTTGCTGACTCCCGCTCGTGTAG  
 4600 TTGAACCAGACTGCAATGGTTACGAATTAGTCACTCCGTTGAGTAGAGTGCCTAGACAGATAAAGCAAGTAGGTATCAACGGACTGAGGGGACGACATC  
 orf 2 < \* W H K I L S A G I E A I Q R N R E D M T A Q S G T T Y

4601 ATAACCTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCAGGCTCAGATTTATCAGCAATAAACCCAGC  
 4700 TATTGATGCTATGCCCCCTCCGAAATGGTAGACCGGGTTCACGAGTTACTATGGCGCTCTGGGTGCGAGTGGCCGAGGCTAAATAGTCTGTTATTTGGTGC  
 orf 2 < I V V I R S P K G D P G L A A I I G R S G R E G A G S K D A I F W

4701 CAGCCGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGCGGGAAGTAGAGTAAGTAGTTCGCCAGT  
 4800 GTCGGCTTCCCGCTCGGCTCTTACCAGGACGTTGAAATAGGCGGAGGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTATCAAGCGGTCA  
 orf 2 < G A P L A S R L L P G A V K D A E M W D I L Q Q R S A L T L L E G T

TAATAGTTTGGCACAAGTGTGTCATTGCTGCAGGCATCGTGTGTCAGCTCGTCTTGGTATGGCTTCATTGAGTCCGGTCCCAACGATCAAGG

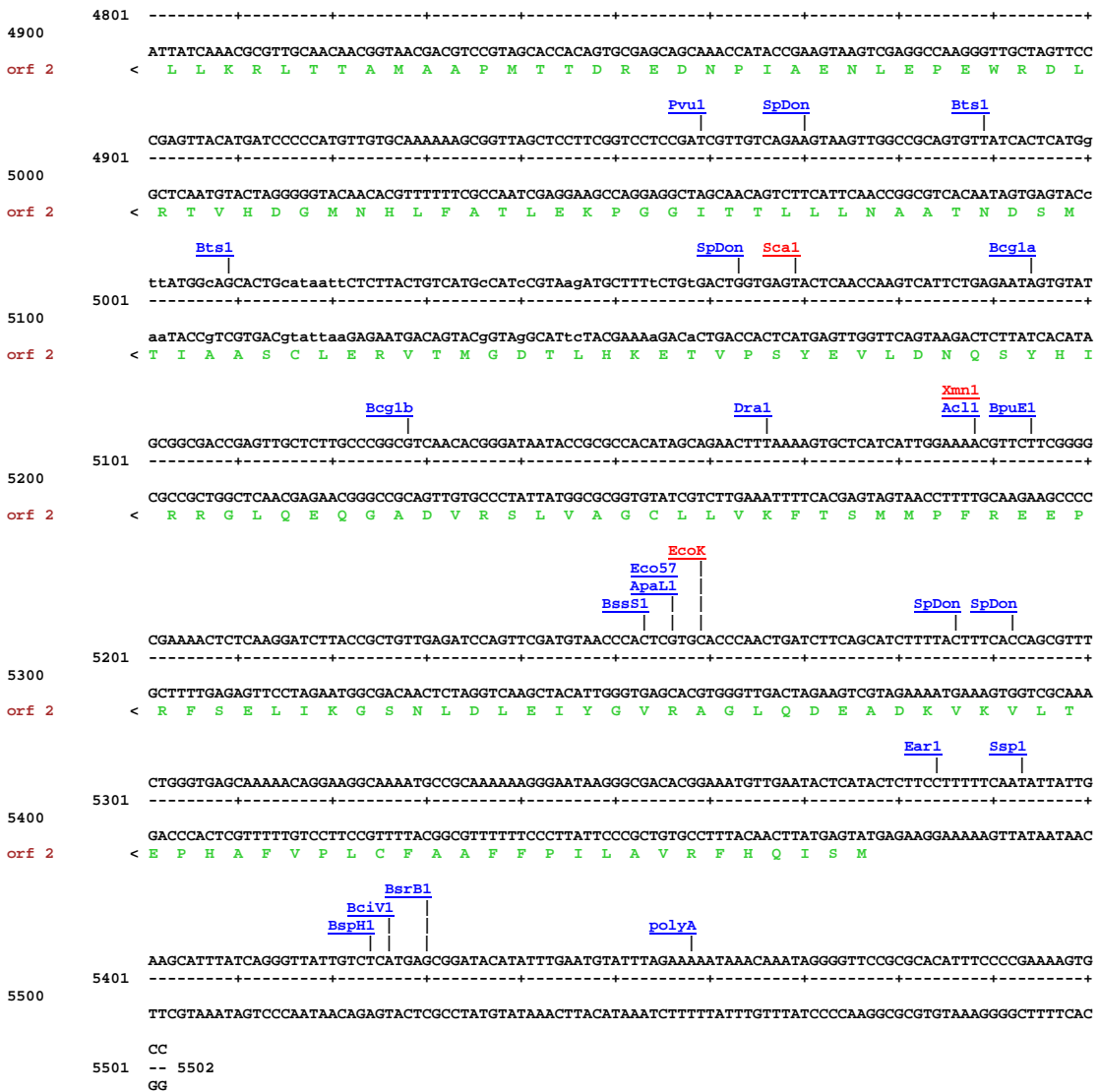


Figure 19: Restriction map of plasmid pTZ1  
 Map constructed with 'web map preference' tool at  
[http://pga.mgh.harvard.edu/web\\_apps/web\\_map/start](http://pga.mgh.harvard.edu/web_apps/web_map/start)



Two silent mutations were found in the LacZ coding sequence of plasmid pTZ1 while aligning to plasmid pZ1918. Silent mutations are GTA to GTG (Valine to Valine) at nucleotide position 312 amino acid position 104 and CGT to CGC (Arginine to Arginine) at 573 nucleotides and 191 amino acid.

```

Query 301 TACACCAACGTACCTATCCCATTACGGTCAATCCGCCGTTTGTGCCACGGAGAATCCG 360
          |||
Sbjct 301 TACACCAACGTACCTATCCCATTACGGTCAATCCGCCGTTTGTGCCACGGAGAATCCG 360

Query 361 ACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACG 420
          |||
Sbjct 361 ACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACG 420

Query 421 CGAATTATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTC 480
          |||
Sbjct 421 CGAATTATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTC 480

Query 481 GGTACGGCCAGGACAGTCGTTTGCCGCTCTGAATTTGACCTGAGCGCATTTTTACGCGCC 540
          |||
Sbjct 481 GGTACGGCCAGGACAGTCGTTTGCCGCTCTGAATTTGACCTGAGCGCATTTTTACGCGCC 540

Query 541 GGAGAAAACCGCCTCGCGGTGATGGTGCTGCGTTGGAGTGACGGCAGTTATCTGGAAGAT 600
          |||
Sbjct 541 GGAGAAAACCGCCTCGCGGTGATGGTGCTGCGTTGGAGTGACGGCAGTTATCTGGAAGAT 600

Query 601 CAGGATATGTGGCGGATGAGCGGCATTTCCGTGACGTCTCGTTGCTGCATAAACCGACT 660
          |||
Sbjct 601 CAGGATATGTGGCGGATGAGCGGCATTTCCGTGACGTCTCGTTGCTGCATAAACCGACT 660

```

Figure 21: Two silent mutations in LacZ coding sequence of plasmid pTZ1

Sequencing analysis revealed two silent mutations in LacZ gene of pTZ1 while comparison to the LacZ in pZ1918 plasmid, using the alignment tool at NCBI BLAST.

## 5.2 Optimization for Mutagenesis PCR Conditions

The complete sequence of plasmid pTZ1 with entire coding sequence of LacZ gene coding for  $\beta$ -galactosidase has helped us to design our mutagenesis experiments. We had to optimize the PCR conditions to achieve best results for amplified products for creating the random library of LacZ mutations. Two different types of DNA polymerase (Hot start plus Taq polymerase and PFU HF polymerase) were tested with two different set of primers (Mut1, Mut3 and Mut1, Mut2) and 3 different PCR solutions (sterile H<sub>2</sub>O, Q solution, PD solution) see table 11.

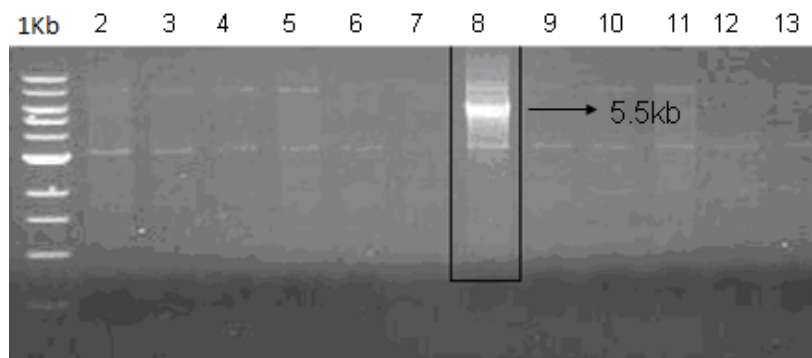


Figure 22: Analysis PCR product with various conditions

**Lane1:**1Kb Ladder,**Lane2:**mut1, mut2, water, Hot Star Taq plus Polymerase, **Lane3:** mut1, mut2, Q solution, Hot Star Taq plus Polymerase, **Lane4:** mut1, mut2, PD, Hot Star Taq plus Polymerase, **Lane5:** mut1, mut2, water, PFU HF polymerase, **Lane6:** mut1, mut2, Q solution, PFU HF polymerase, **Lane7:** mut1, mut2, PD, PFU HF polymerase, **Lane8:** mut1, mut3, water, Hot Star Taq plus Polymerase, **Lane9:** mut1, mut3, Q solution, Hot Star Taq plus Polymerase, **Lane10:** mut1, mut3, PD, Hot Star Taq plus Polymerase, **Lane11:** mut1, mut3, water, PFU HF polymerase, **Lane12:** mut1, mut3, Q solution, PFU HF polymerase, **Lane13:** mut1, mut3, PD, PFU HF polymerase.

PCR products amplified with all of the conditions were analyzed using 0.8% agarose gel electrophoresis with ethidium bromide (0.5 $\mu$ g/ml) and 1Kb DNA ladder was used as a base pair marker during the procedure. Best PCR results were obtained by using Hot Start Taq DNA polymerase, water and primer Mut1& Mut3. Primer Mut1 and Mut3 were designed to amplify the entire plasmid and as seen in gel picture (Figure: 22) the band intensity was clear and of the correct size of 5.5Kb.

### 5.3 Optimization of Primer concentration in random mutagenesis PCR

A suitable PCR conditions for the whole- plasmid mutagenesis with Hot start plus Taq polymerase, using Mut1and Mut3 primer. Further optimization was necessary to obtain a suitable primer concentration for best amplified PCR product. Four different primer-concentrations were tested in the amplifications of plasmid pTZ1, 25pmole, 10pmole, 15pmole and 20pmole.

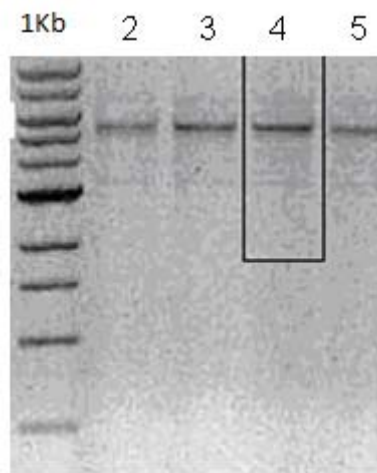


Figure 23: Optimization of primers concentration for random mutagenesis PCR

**Lane1:**1Kb ladder, **Lane2:**2.5pmole conc.**Lane3:** 10pmole conc., **Lane4:**15pmole  
**Lane5:**20pmole

Samples were analyzed on 0.8% agarose gel electrophoresis.1Kb ladder was used as base pair standard. Concentration of 15pmole appeared to be best for the whole plasmid PCR amplification (Figure: 23)

## 5.4 Development of PCR protocol with various Mn and dITP concentration

In order to develop an efficient and simple method for creating random mutagenesis to create libraries with error-prone PCR, we used the modified our previously developed protocol by Xu et.,al 1999. The initial method was based on two subsequent PCR steps using Mn in the first PCR followed by the second PCR containing dITP. This was made in order the created error-prone to overlies the bias in ratio of transition:transversion(Fenton et al., 2002b). Attempt to use the whole plasmid amplification with error-prone PCR, made the method simple and fast for constructing the random libraries in one single step. The experiment was done to obtain the suitable concentrations of Mn and dITP simultaneously, without jeopardizing the PCR product for generating library

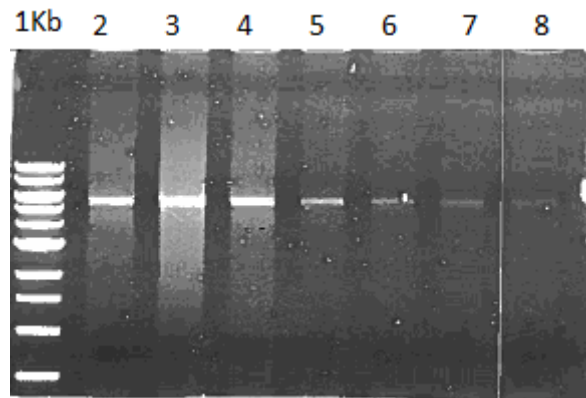


Figure 24: Optimization of mutagenesis PCR with various concentrations of Mn and dITP

Gell: **Lane1:**1Kb ladder, **lane2:** control, **lane3:**40 $\mu$ M Mn,**Lane4:**80 $\mu$ MMn, **Lane5:** 80 $\mu$ MMn+40 $\mu$ MdITP **Lane6:**80 $\mu$ MMn+80 $\mu$ MdITP, **Lane7:**120 $\mu$ MMn+80 $\mu$ MdITP,**lane6:**120 $\mu$ MMn+120 $\mu$ MdITP

5 $\mu$ l of PCR product was analyzed on 0.8% agarose gel electrophoresis by comparing the intensity of PCR product to the intensity of known amount of 1Kb DNA ladder. The comparisons can be seen in Table: 12. The addition was 40 $\mu$ M Mn, 80 $\mu$ M Mn, 80 $\mu$ M Mn+40 $\mu$ M dITP, 80 $\mu$ M Mn+80 $\mu$ M dITP generate a viable PCR product but 120 $\mu$ M Mn+80 $\mu$ M dITP and 120 $\mu$ M Mn+120 $\mu$ M dITP did not generated a suitable PCR product for random library construction.

Table 12: Effects of dITP and Mn+2 on PCR Product

dITP( $\mu\text{M}$ )	Mn+2( $\mu\text{M}$ )	PCR product (ng)	PCR Quality
		~40	
-	40 $\mu\text{M}$	~100	Good
-	80 $\mu\text{M}$	~50	Good
40 $\mu\text{M}$	80 $\mu\text{M}$	~25	Good
80 $\mu\text{M}$	80 $\mu\text{M}$	~12	Good
120 $\mu\text{M}$	80 $\mu\text{M}$	~6	Poor
120 $\mu\text{M}$	120 $\mu\text{M}$	~0	Poor

### 5.5 Screening and Selection for $\beta$ -glucosidase Activity

*E. coli* X11Blue host cells harboring the library, treated with various concentration of Mn and dITP (see section 4.8) were screened on different plates (M9/cellobiose, LB/cellobiose and LB+X-Glu) for beta-glucosidase activity. The screening and selection method was based on the concept that transfer of heterologous beta-glucosidase into *E.coli* permit the hydrolysis of cellobiose to glucose and allow the non-cellobiose utilizing *E.coli* to grow on these plates (Liu et al., 2009).

The attempts were also performed to screen the library on LB/cellobiose plates for the same enzyme activity. The screening power was evaluated by the proportion of colony appearance, where no colony appears on M9/cellobiose plates while all transformants grown on LB/cellobiose selection plates(Liu et al., 2009)

Although all transformants grown on LB/cellobiose and LB+ plates but no color development on chromogenic substrate 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-glucopyranoside(X-Glu) was notice in 24 hour of incubation.

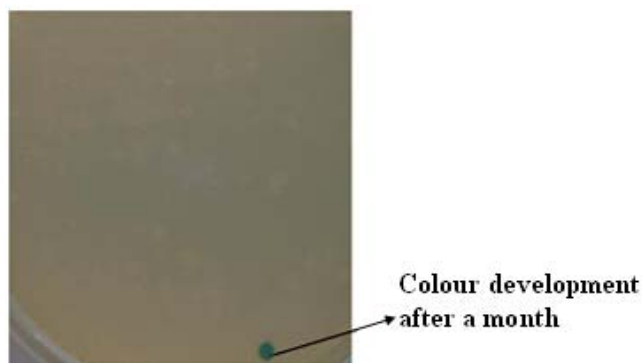


Figure 25: Blue colony appearing on chromogenic substrate for  $\beta$ -glucosidase



Several rounds of screening were done to select a single mutant with beta-glucosidase activity. No activity was seen in any round of screening but a blue colony appears after a month on LB/Cellobiose and X-Glu plate (See Figure: 25).

Retransformation of Plasmid DNA from this colony gave all white colonies, this possibly means that appearance of blue color cannot be due to any encoded enzyme activity in plasmid DNA, but can be due to mutation in chromosomal DNA.

## **5.6 Screening and Selection for $\beta$ -xylosidase Activity**

The objective was to screen for any possible  $\beta$ -xylosidase activity in the generated random libraries of LacZ gene. As the information available in the literature, most of the bacterial species including *E. coli*-K12 and its derivatives do not contain endogenous xylosidase activity (Whitehead, 1997, Geddie, 2004). This makes the screening easy by avoiding false positive due to endogenous xylosidase activity in host cells.

Around 1600 colonies were screened on LB/X-Xyl plates to select a mutant colony exhibiting xylosidase activity. Screening was based on colour development on the chromogenic substrate 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside (X-Xyl).

27 blue colonies were picked after ~24 hrs of incubation and compared with colonies expressing wildtype beta-glucosidase under the same growth condition and same type of plates.

First 10 blue colonies (Named Colony 1 to 10) and from this 5 (No.1, 2, 3, 8 and 9) variants were selected for further investigation. Plasmid from all the 5 variants as well as wildtype were retransformed into *E. coli* XL1Blue competent cells on same growth condition with x-xyl substrate to validate the enzyme activity is due to plasmid or chromosomal. Blue colour appeared in all 5 colonies but wild type shown no visible xylosidase activity. (See figure 26).

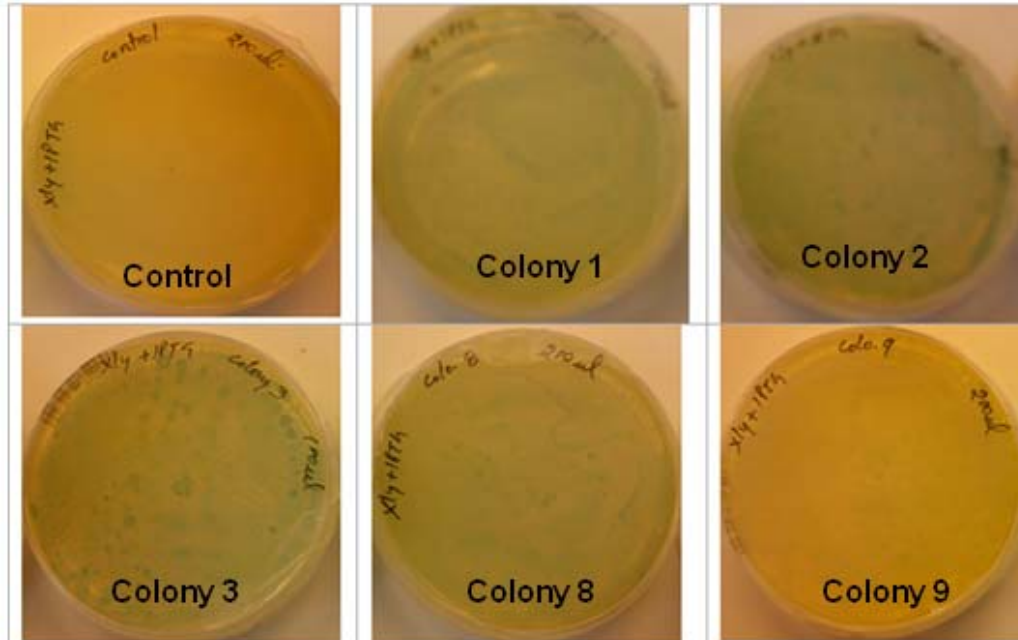


Figure 26: Blue colonies after retransformation of plasmid DNA on substrate X-xyl

Retransformation performed with plasmid DNA from 5 selected colonies (colony 1,2,3,8 and 9) and pTZ1 as a control on LB agar plates with 200 $\mu$ g/ml ampicillin, IPTG and 40 $\mu$ g/ml 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside as a indicator. Transformants from colony 1,2,3 and 8 show in clear blue colour in compare to wildtype. Colour intensity of colony 9 is lower than other mutants but showed clear difference than wild type.

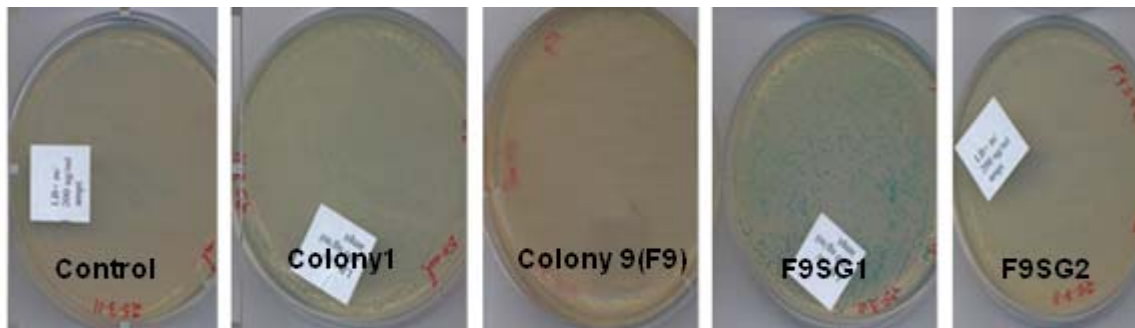
## 5.7 Screening of Second Round of Mutagenesis

Plasmid DNA from first three selected colonies 1, 2, 3 and colony 9 (Based on colorimetric enzyme assay (see section: 5.9.1) were used as a template in a second round of mutagenesis.

14 colonies (10 blue and 4 white) named SG1 to SG14 from parental colonies 2 and 3, 2 colonies were named F9SG1 and F9SG2 from parental colony 9 (F9) and a single colony named C1SG1 from colony 1 was selected for further investigation.

To select a fittest variant for  $\beta$ -xylosidase activity from the second generation of colony 2 and 3 (SG1 to SG14), an enzymatic assay performed with the chromogenic substrate 4-Nitrophenyl  $\beta$ -D-xylopyranoside (See section 5.9.1). None of the mutant were shown higher activity than colony 9 of first generation. Further validation for the selected colonies containing the selected phenotype was done by retransformation of the “mutated” plasmid DNA isolated from these selected colonies on X-Xyl plates. Colony 1, colony 9(F9), F9SG1 F9SG2 and wild type plasmid pTZ1 were retransformed (see

figure: 27). Transformation of WT plasmid which has unmutated  $\beta$ -galactosidase used as a control. Plasmid DNA isolated from colony C1SG1 was supposed to be retransformed, but was not done due to the limited time. Retransformation result shows that activity of  $\beta$ -xylosidase in colony 1(K552E), F9SG1 (N959Y), was higher on screening plates with x-xyl than wild type, but the result for colony 9(F9) was not same as previous transformation. The second retransformation of plasmid DNA from colony 9 showed all white colonies (see figure: 26 and 27). The sequencing result of plasmid DNA isolated from colony 9 (F9) showed that there was no mutation in the LacZ coding sequence, but a (G to A) substitution was present in the vector sequence of the plasmid (near RBS) (see Figure: 29).



**Figure 27:** Retransformation of plasmids DNA isolated from the first and second generation

Transformation performed on LB agar plates with 200 $\mu$ g of ampicillin, IPTG and 40 $\mu$ g/ml of X-Xyl substrate. Variant F9SG1 (N959Y) showing very clear blue colonies on X-Xyl plates.

## 5.8 Screening and Selection for Improved $\beta$ -galactosidase Activity

One of the goals in this study was to select variants from the random library with enhanced Beta-galactosidase enzyme activity in compare to the wild type.

A total of 2959 colonies were screened on LB/X-Gal plates. The screening was based on the extent of color development the chromogenic substrate 5-Bromo-4-chloro-3-indolyl - D-galactopyranoside (X-Gal).

165 (with deeper blue color) variants were picked from the larger pool of 2959 colonies for the further investigation. 18 variants were again selected out 165 by comparing phenotypically with the wildtype on LB/X-Gal plates.

8 mutants were also selected by the random library of LacZ coding gene on the same types of plates and by filter lift assay (See figure: 28).

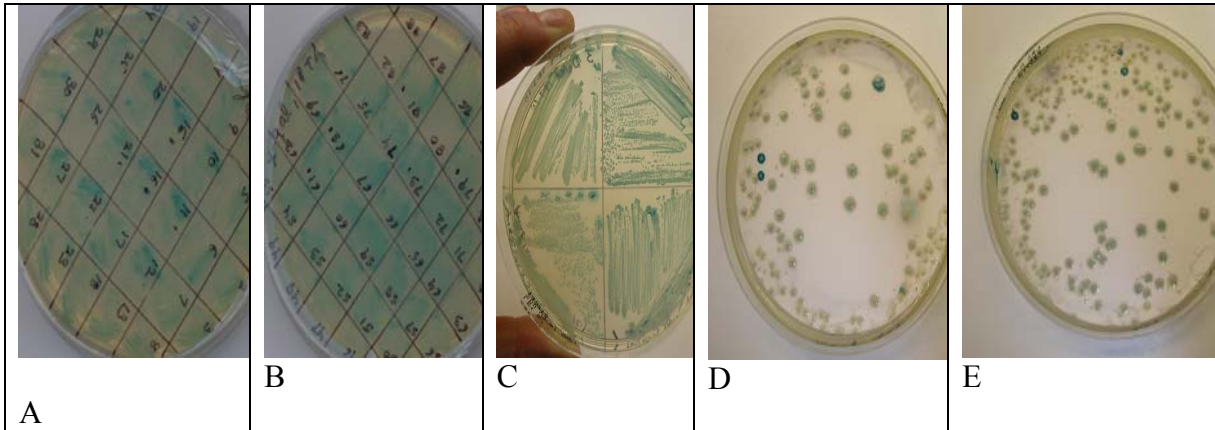


Figure 28: Screening and selection method of higher  $\beta$ -galactosidase activity

A-C By squares on x-gal plates and D, E Filter lift assay method and placing filters on X-gal plates.

## 5.9 Colorimetric Enzyme Assays

### 5.9.1 $\beta$ -Xylosidase Assay

Xylosidase enzyme activity was performed in order to determine if any of the selected colonies is expressing a higher beta-Xylosidase enzyme activity in comparison to the activity of colonies expressing the Wildtype beta-galactosidase. The assays were repeated few times to achieve reproducible results of the enzyme activity measurements. Enzyme activity was also measured in 30 minutes at real time, but the color development was very weak and slow. This necessitate to leave the plates overnight and all the reading were measured again and calculated according the hours of incubation at room temperature.

Initial 10 variants from the first round of random mutagenesis and 14 variants from second round of random mutagenesis of first generation colony 2 and 3 were assayed in triplicates for beta-Xylosidase activity in comparison to that of wildtype beta-galactosidase activity. Surprisingly the wild LacZ, although on plates containing x-xyl did not have any detectable activity, but in the colorimetric enzyme assay using 4-nitrophenyl-Beta-D-xylopyranoside substrate, it exhibited some noticeable activities Activity measured in units and calculated according to the standard Miller Unit formula.

A405nm

1000X 

---

A590nm x 1mL x time (hr)

Table 13: Xylosidase activity measured for variants of first and second generation

Mutants	Beta-xylosidase activity (U) After 15 hours	%Beta-xylosidase Mutant/Wildtype
Wild type (pTZ1)	17.63	100
colony1	18.70	106.03
colony2	24.74	140.28
colony3	18.75	106.30
colony4	27.57	156.34
colony5	22.34	126.65
colony6	32.77	185.81
colony7	18.22	103.30
colony8	21.84	123.84
colony9	42.25	239.51
colony10	23.83	135.10
SG1	20.99	119.01
SG2	25.24	143.10
SG3	25.21	142.94
SG4	29.48	167.14
SG5	19.31	109.46
SG6	26.19	148.49
SG7	24.89	141.13
SG8	16.07	91.12
SG9	23.56	133.59
SG10	28.62	162.26
SG11	20.57	116.61
SG12	21.66	122.82
SG13	17.26	97.86
SG14	22.79	129.21

Most colonies with mutated plasmids demonstrated a higher enzyme activity than the wildtype beta galactosidase. But the colony 9 (FG9) was shown highest activity than the rest of the colonies. Initially on screening plates containing X-xyl substrate and the above assay performed using substrate using 4-nitrophenyl-Beta-D-xylopyranoside (See figure: 13). This observation made us to go for second round of mutagenesis on the isolated plasmid DNA from colony 9 (FG9) (see section 5.7).

To investigate the corresponding coding sequence for the higher activity of colony 9 we sequence the LacZ portion as well as the whole plasmid. Sequencing result shown that a nucleotide substitution (G to A) was present in the vector portion (near RBS) of the plasmid. No other mutations can be located in coding sequence of LacZ (see Figure: 29). The presence of mutation near to the ribosomal binding site could explain the observed higher activity for Beta- Xylosidase as well as Beta –galactosidase activities, both on screen plates and in colorimetric enzymes assay. Further study is needed to validate the results.

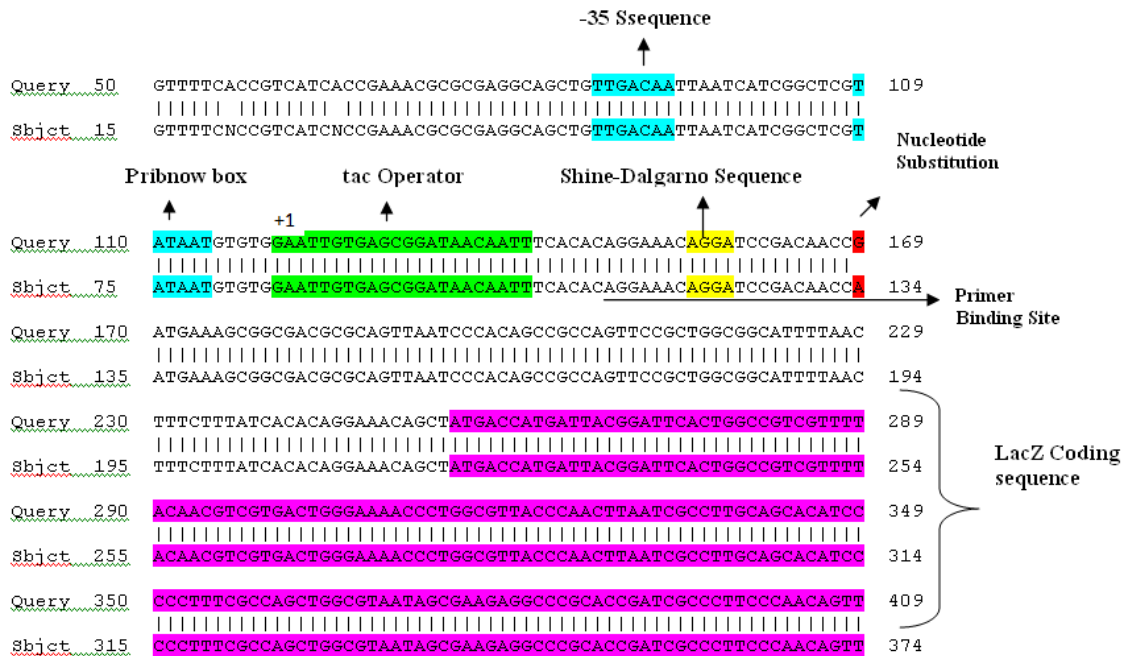


Figure 29: Showing a mutation at primer binding site in colony 9 (F9) plasmid

Mutation at +49 position close to +36 position of RBS. Same mutation prevented primer Mut1 (caggaaacagatccgacaaccga) to anneal resulting in lack of PCR product while combining with primer Mut3 for a second generation mutagenesis

The final selected and sequenced variants of first generation (Colony 1 and Colony 9) and second generation (C1SG1, F9SG1 and F9SG2) were again assayed in triplicates for Xylosidase activity. Second generation of colony 9 was obtained by using another set of primers Laczseq1 and laczWPA.

This time the activity was measured with two different chromogenic substrate 4-Nitrophenyl  $\beta$ -D-xylopyranoside (PNPX) and 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside(X-xyl) (see figure: 30 and 31).

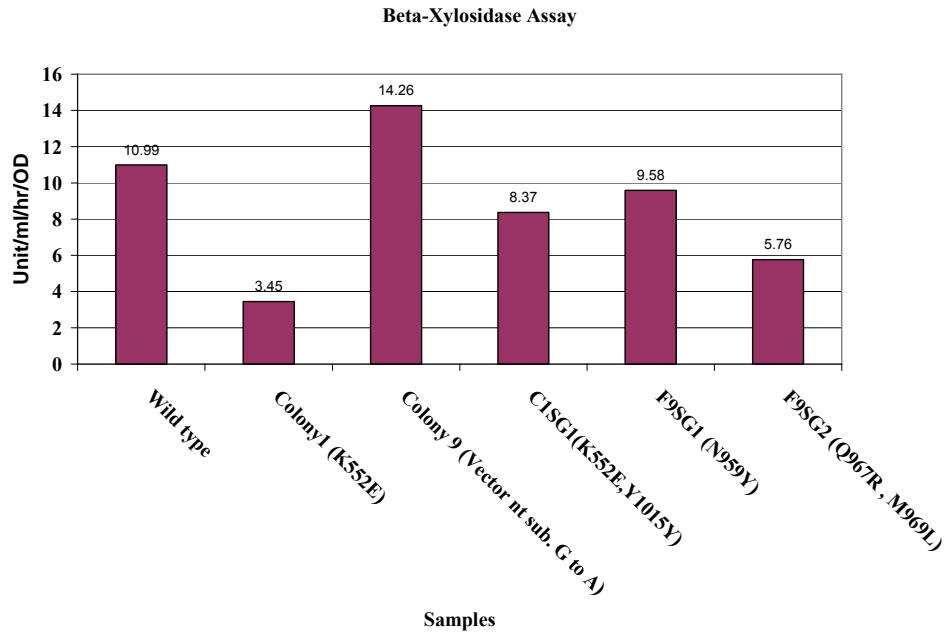
Table 14: Measurement of Beta-xylosidase with PNPX at 405nm

Initial assay		
Mutants	Beta-xylosidase activity (U) After 18 hours	% of Beta-xylosidase Mutant/Wildtype
<b>Wild type</b>	10.99	100
<b>Colony1(K552E)</b>	3.45	31.39
<b>Colony 9 (F9) (vector nt sub. G to A)</b>	14.26	129.75
<b>C1SG1 (K552E, Y1015Y)</b>	8.37	76.16
<b>F9SG1 (N959Y)</b>	9.58	87.17
<b>F9SG2 (Q967R, M969L)</b>	5.76	52.41
Repeating assay		
Mutants	Beta-xylosidase activity (U) After 16 hours	% Beta-xylosidase Mutant/Wildtype
<b>Wild type</b>	14.56	100
<b>Colony1(K552E)</b>	6.51	44.75
<b>Colony 9 (F9) (vector nt sub. G to A)</b>	10.11	69.47
<b>C1SG1 (K552E, Y1015Y)</b>	8.88	61.02
<b>F9SG1 (N959Y)</b>	7.88	54.16
<b>F9SG2 (Q967R, M969L)</b>	5.55	38.12

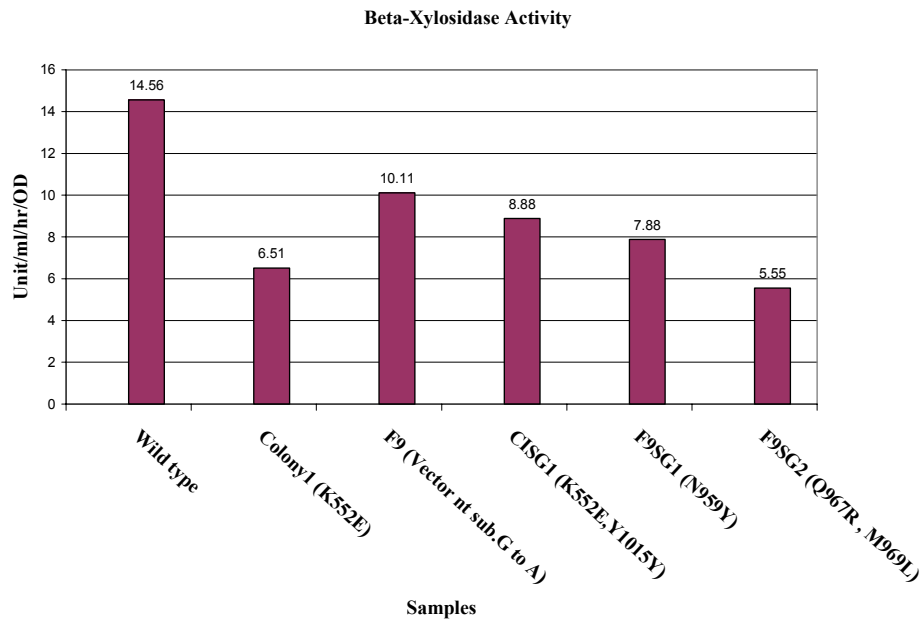
Table 15: Measurement with blue substrate X-xyl at 750nm

Mutants	Beta-xylosidase activity (U) After 22 hours	% Beta-xylosidase Mutant/Wildtype
<b>Wild type</b>	11.22	100
<b>Colony1 (K551E)</b>	3.74	33.30
<b>Colony 9 (F9)(nt sub.G to A)</b>	2.73	24.34
<b>F9SG1(N959Y)</b>	3.48	31.00
<b>F9SG2 (Q967R, M969L)</b>	3.47	30.97

Note:C1SG1-not measured



A



B

Figure 30:  $\beta$ -xylosidase activity with 4-Nitrophenyl  $\beta$ -D-xylopyranoside at 405nm

A: Initial assay B: Repeating assay for  $\beta$ -xylosidase activity in compare to wildtype



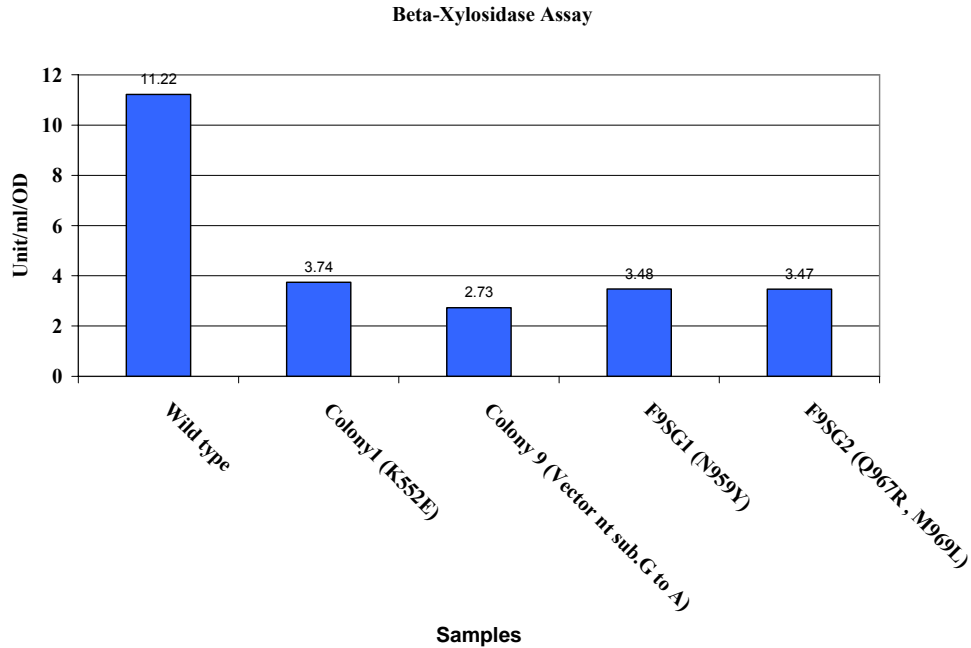


Figure 31:  $\beta$ -xylosidase activity using substrate 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside at 705nm

The enzyme activity was analyzed with two different substrates. Similar lysates of host cell XL1Blue were also assayed as a blank but due to high absorbance of the cloudy lysates, causing a large calculation error so we had to replace the blank XL1Blue with a blank buffer. According the colorimetric enzyme assays, the wildtype LacZ exhibited higher enzymatic activities during two measurements, reaping assay with PNPX and assay with x-xyI (see figure:30 and 31) but the initial measurement with PNPX (see figure:30) shown a bit higher activity in colony 9 in compare to wildtype. Activity for colony 9 was also higher when we assayed for all first and second generation (see table: 13).

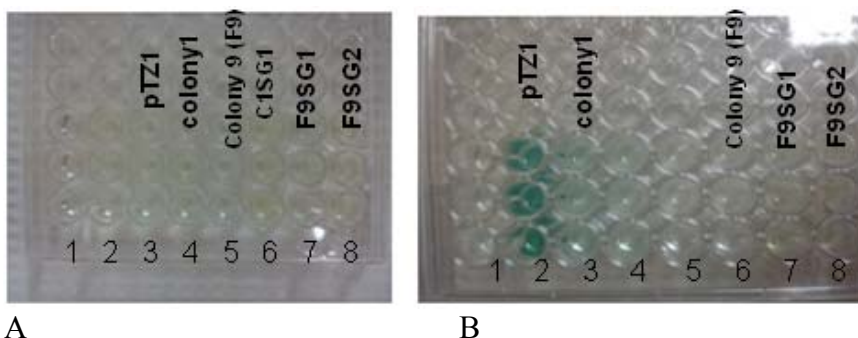


Figure 32: A picture of microtiter plate with the performed colorimetric assay for  $\beta$ -xylosidase activity

**A:** Activity measurement with the yellow substrate 4-Nitrophenyl  $\beta$ -D-xylopyranoside (repeating assay). **B:** Activity measurement with 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside: visible yellow colour appeared in many of the mutants after time duration when using substrate PNPX but only wildtype showed clear visible colour on plates with x-xyl as a substrate.

### 5.9.2 $\beta$ -galactosidase Assay

A total of 26 selected colonies (darker blue than wildtype by comparing on plates using x-gal as a substrate) (see section: 5.8).18 selected from whole plasmid random library and 8 variants from random library of only LacZ gene were assayed in duplicates for higher  $\beta$ -galactosidase activity. None of the variant showing higher activity in compare to wild type beta galactosidase.

The variants that were screened and selected on x-xyl plates and exhibited a noticeable  $\beta$ -xylosidase were also assayed for  $\beta$ -galactosidase activities and were compared to that of wild type. Assays for beta-galactosidase were also repeated twice and measured in Unit/ml/hr/OD

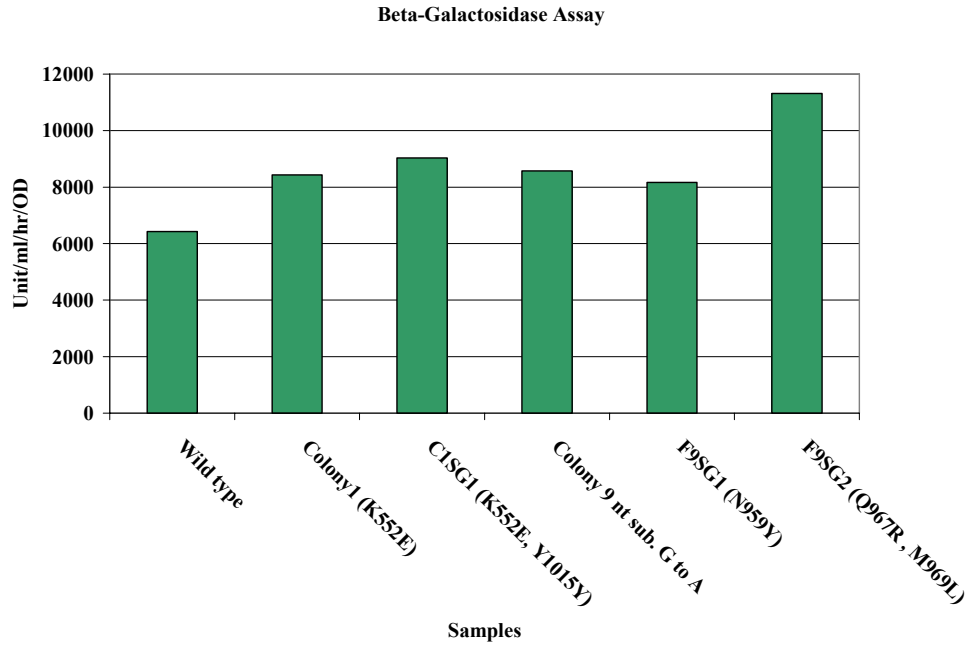
Table 16: Measurement with ONPG for  $\beta$ -galactosidase

Initial assay		
Mutants	Beta-Galactosidase activity (U)	% Beta-galactosidase Mutant/Wildtype
<b>Wild type</b>	6427.08	100
<b>Colony1(K552E)</b>	8429.33	131.15
<b>C1SG1 (K552E, Y1015Y)</b>	9034.28	140.56
<b>Colony 9(F9) (nt sub. G to A)</b>	8576	133.43
<b>F9SG1 (N959Y)</b>	8168.83	127.10
<b>F9SG2 (Q967R , M969L)</b>	11312.28	176.00

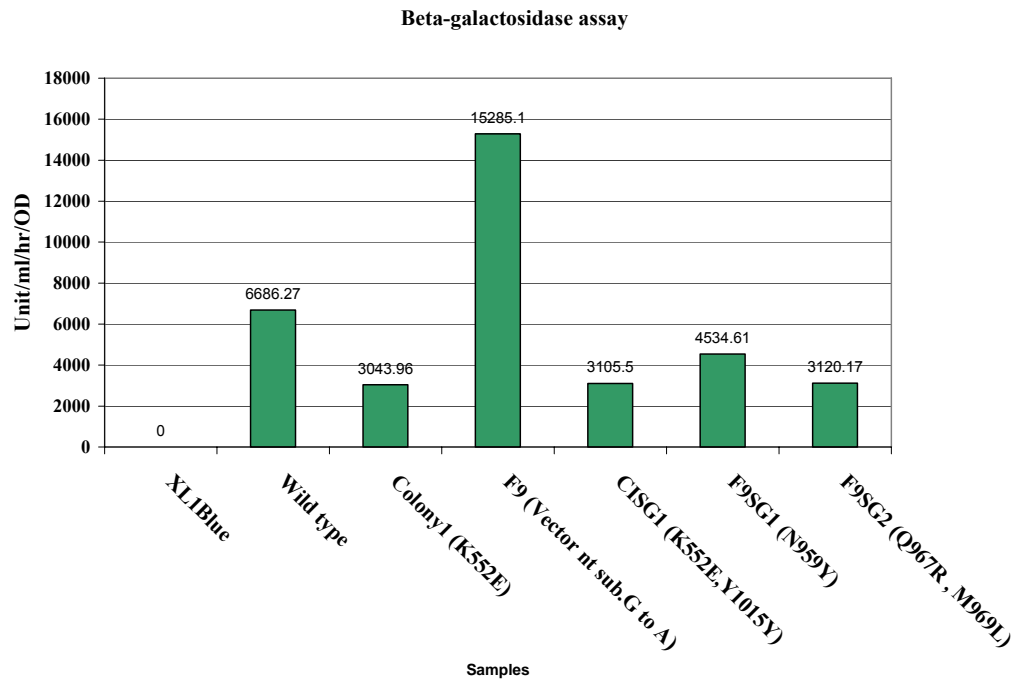
Repeating assay

Mutants	Beta-Galactosidase activity (U)	% of Beta-galactosidase Mutant/Wildtype
<b>Wild type</b>	6686.27	100
<b>Colony1(K552E)</b>	3043.96	45.52
<b>C1SG1 (K552E, Y1015Y)</b>	3105.5	46.44
<b>Colony 9(F9) (nt sub. G to A)</b>	15285.1	228.6
<b>F9SG1 (N959Y)</b>	4534.61	67.81
<b>F9SG2 (Q967R , M969L)</b>	3120.17	46.66

Assay data was also not reproducible for of beta galactosidase activity as for Beta-xylosidase. The results of first assay show higher enzyme activity in all the variants than Wildtype while the second assay with the same variants, showed a low activity in all mutant variants except colony 9. This Beta-galactosidase activity data is also demonstrated in histograms in compare to Wildtype (See figure: 33).



A



B

Figure 33: Showing the  $\beta$ -galactosidase activity in compare with Wildtype

A: Initial assay B repeating assay for beta-galactosidase in compare to wildtype.

## 5.10 Sequencing Analysis

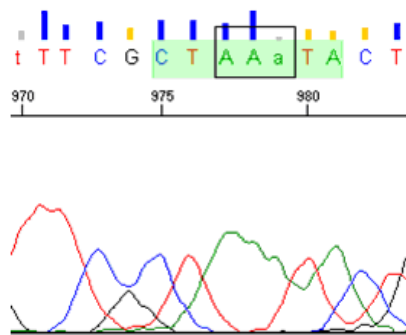
Colonies with mutant plasmids were selected for sequencing during screening of  $\beta$ -Xylosidase activities on x-xyl plates. Plasmid DNA of the selected colonies were sequenced in order to identify the mutations in the LacZ coding sequence that could be associated with the selected phenotype (colour development on chromogenic substrate) on the 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside as a substrate.

Table 17: Mutation in the LacZ coding gene of the selected variants

Colony with Mutated plasmid	Amino acid substitution	Nucleotide Number substitution	Silent Mutation
Colony1(First generation)	K552E		
C1SG1 (Second generation)	K552E		Y1015Y
Colony 9(F9)(First eneration)		G to A sub in vector (see Figure: 29)	
F9SG1 (Second generation)	N959Y	G to A sub in vector	
F9SG2 (Second generation)	Q967R And M969L	G to A sub in vector	

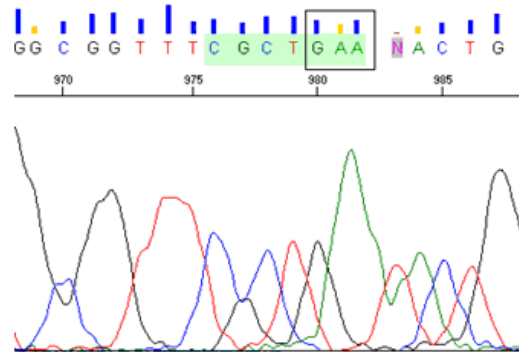
Colony 1 (First Generation)

**K552E**



Wild Type

AAA (Lysine)

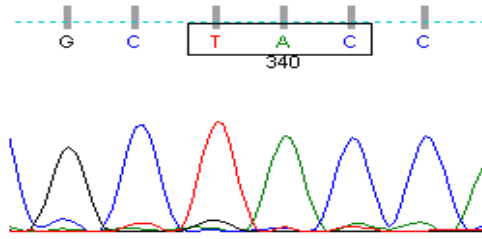


Mutant

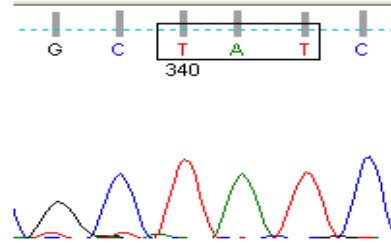
GAA (Glutamic acid)

C1SG1 (Second generation)

### Y1022Y

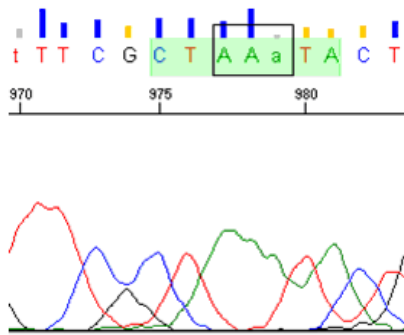


**Wild type**  
**TAC (Tyrosine)**



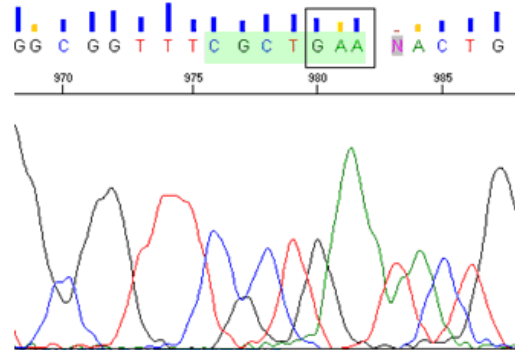
**Mutant**  
**TAT (Tyrosine)**

### K552E



**Wild Type**

**AAA (Lysine)**

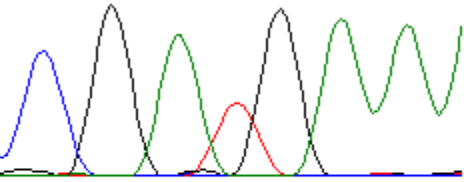
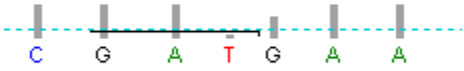


**Mutant**

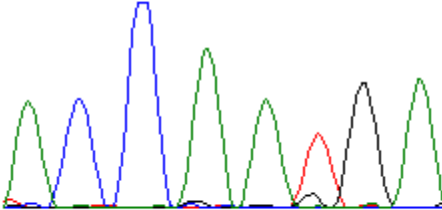
**GAA (Glutamic acid)**

Colony 9 (F9) First Generation at nucleotide position 169 (+49 position)

**Nucleotide substitution G to A at primer binding site in vector portion**



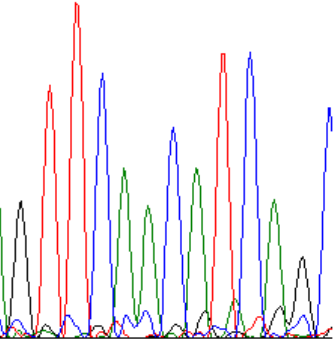
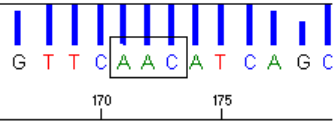
**Wild Type**



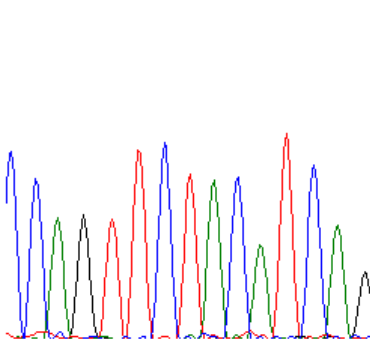
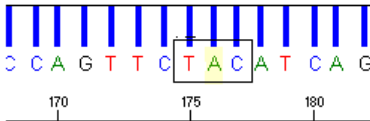
**Mutant**

F9SG1 (Second Generation)

**N959Y**



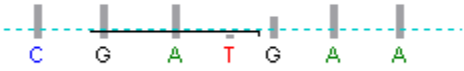
**Wild Type**  
**AAC (Asparagine)**



**Mutant**  
**TAC (Tyrosine)**

Nucleotide position 169 (+49 position)

**Nucleotide substitution G to A at primer binding site in vector portion**



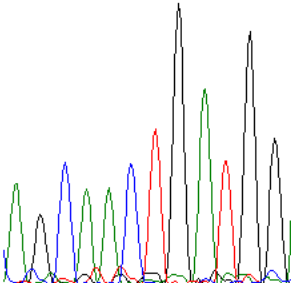
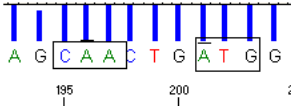
**Wild Type**



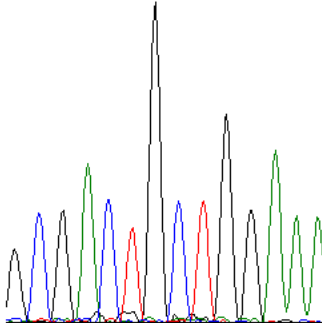
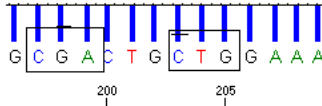
**Mutant**

F9SG 2 (Second generation)

**Q967R and M969L**



**Wild Type**  
**CAA Glutamine**  
**ATG Methionine**



**Mutant**  
**CGA Arginine**  
**CTG leucine**

And



Nucleotide position 169 (+49 position)

**Nucleotide substitution G to A at primer binding site in vector portion**

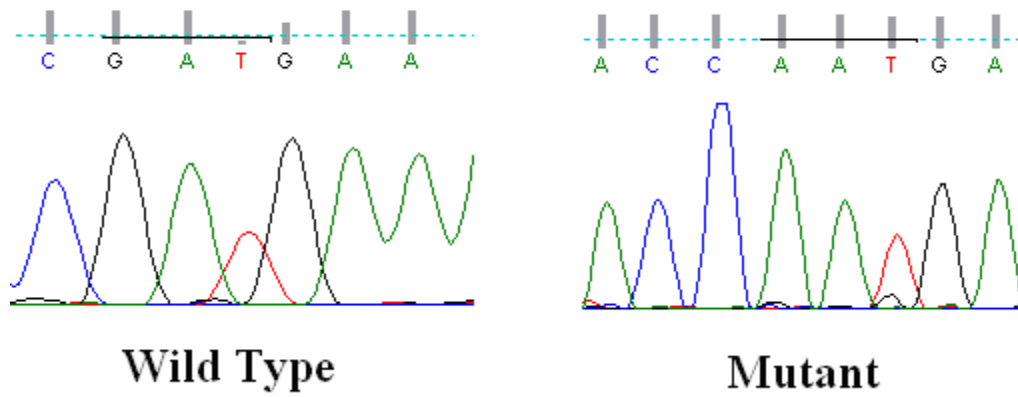


Figure 34: Chromatograms showing mutation in selected colonies

## 6. Discussion

Directed evolution of proteins in general and enzymes in particular has been a very powerful tool in introducing landscape diversity in the coding sequence of the proteins coupled with a rigorous selection approaches for the desired phenotypes, characteristics, and kinetics and enzyme substrate specificities with the obvious biotechnological applications in addition to the benefit of structure-function relationship of proteins (Fasan et al., 2011, Romero and Arnold, 2009, Tracewell and Arnold, 2009) Improving enzyme properties through directed evolution is a rapid and versatile approach in protein engineering due to its independence of prior knowledge of enzyme structure, function or enzyme-substrate interaction (Liu et al., 2009). In this project our goal was to use to apply directed evolution tools to introduce random mutations in LacZ gene coding for  $\beta$ -galactosidase in order to change the natural substrate specificity towards the hydrolysis of other glycosides' as substrates. Also, we wanted to improve the enzyme activity of  $\beta$ -Galactosidase.

### 6.1 Sequencing of Plasmid pTZ1 Encoding Wildtype LacZ Gene

The complete sequence of the plasmid pTZ1, containing the coding sequence of LacZ was investigated in order to help us design the needed primers for mutagenesis as well as to design sequencing primers needed for the validation of the obtained mutants. The sequencing analysis showed that plasmid pTZ1 is 5,502bp, while according to its estimated size using restriction enzyme mapping pTZ1 was estimated to be 5.2 kb (Su et al., 1990). The sequence analysis of pTZ1, also demonstrated that LacZ gene has two silent mutations, GTA to GTG (Valine to Valine) at nucleotide position 312 and amino acid position 104 as well as CGT to CGC (Arginine to Arginine) at the 573 nucleotide and amino acid 191 in Beta-galactosidase (see figure: 21). These silent amino acids substitutions were in comparison to a published derivative of pTZ1, which is plasmid pZ1918 (<http://www.ncbi.nlm.nih.gov/nuccore/434652>). Based on our completed sequence data we have constructed, a restriction enzyme map (see figure: 19), the worked as the frame work for all of our subsequent primers design.

## 6.2 Random Mutagenesis Libraries of LacZ Coding Sequence

In this project we have done further optimization to the previously developed two step method by Xu *et.,al* 1999. We have succeeded here in using only one single step PCR method for creating random library using whole plasmid amplification, containing the desired target gene (*lacZ*) using  $Mn^{+2}$  and dITP. Different combinations of DNA polymerases, primers and PCR buffers and cycles (see figure: 22). Best results were obtained with Hot start Taq plus DNA polymerase, primers Mut1, Mut3, and with water as a PCR solution. PCR product with Q solution and PD were showed low band intensity in compare to band with water as a PCR solution. Primers concentrations were also optimized using various concentrations (2.5pmole, 10pmole, 15pmole and 20pmole). Concentration of 15pmole/ $\mu$ l has given the best PCR result (See figure: 23).

Various concentrations of  $Mn^{+2}$  and dITP were also optimized to figure out viable PCR product for library construction. PCR quality and viability was good enough with the concentration of 40 $\mu$ M  $Mn^{+2}$ , 80 $\mu$ M  $Mn^{+2}$ , 40 $\mu$ M  $Mn^{+2}$  +80 $\mu$ MdITP, 80 $\mu$ M  $Mn^{+2}$  + 80 $\mu$ MdITP but PCR product quality was not suitable for library construction with 120 $\mu$ M  $Mn^{+2}$  +80 $\mu$ M dITP, 120 $\mu$ M  $Mn^{+2}$  +120 $\mu$ M dITP (see table: 12 and figure: 24).

Hot start Taq plus DNA polymerase lack 3'→5' exonuclease activity which is a proofreading activity and remove mispaired based during replication. DNA polymerase with non-proofreading activity exhibit mispair at the rate of 10<sup>-2</sup> to 10<sup>-6</sup> (Cline et al., 1996). Addition of  $Mn^{+2}$  to the reaction mixture can further reduce fidelity of PCR. Substitution of  $Mn^{+2}$  to  $Mg^{+2}$  at the substrate binding site and ancillary binding sites (distant sites) of DNA polymerase change the substrate conformation and affect fidelity of polymerase and increases misincorporation (Zakour et al., 1981). In addition rate of misincorporation also induced by nucleotide analogue dITP in comparison to normal dNTP. It has been also reported that termination induced by dITP is also reduced when reaction started at 70°C (Innis et al., 1988, Spee et al., 1993).

## 6.3 Screening and Selection

1600 colonies were used for Beta-Xylosidase activity screening in the first round and a similar in second round. A total of 2959 colonies were used for  $\beta$ -galactosidase

screening. Additional several rounds of mutant colonies were used for for  $\beta$ -glucosidase enzyme activity screening.

The theoretical Estimation of the total number of variants that should be screened for the new properties and altered specificity of an enzyme is a big question (Lin and Cornish, 2002). In library size of 3000 a variant was selected with changed substrate specificity from phospholipase to 1000-fold increase lipase activity by using a new technique MURA (Mutagenic and unidirectional reassembly) a combination of error-prone PCR and DNA shuffling.(Song et al., 2002).An improved endo- $\beta$ -1, 4-glucanase III from *Trichoderma reesei* was obtained by using two round of error-prone PCR and each round had 9000 and 2000 colonies respectively (Nakazawa et al., 2009). In the application of the whole plasmid (pTZ1) mutagenesis for library creation with error-prone PCR, the number of obtained colonies was very limited. It is possible that mutation could have taken place in ampicillin ( $\beta$ -lactamase) and origin of replication which not allow us to screen entire diversity. Moreover the efficiency of transformation of generated DNA library into competent *E. coli* was not high enough. Higher number of mutations per gene coding for larger protein molecule requires higher number of colonies to be screened (Hansen MS and LT, 2002).

We used *tacI* promoter for the expression of enzymes. *tacI* promoter is a hybrid promoter and derived from *trp* and *lac* UV5 promoter (de Boer et al., 1983). It is an inducible promoter by IPTG .IPTG was on plates and medium in order to induce the expression of mutated and WT LacZ, coding for beta-galactosidase.

To screen for Beta-glucosidase activity, different types of plates were used, namely, M9/cellobiose plate, LB plate and LB+. The advantage of minimal M9 plates is to minimize false positive which normally occur when screen on LB plates (due to other nutritional source that promote the cell growth) since the carbon source here is mainly cellobiose. It also connects the relation between survival of cell and enzyme activity (Hansen MS and LT, 2002). The idea here is in order for the cell to survive on the M9/cellobiose plates *E.coli* cells that encoded enzyme has to have been mutated to beta-glucosidase that would hydrolyze cellobiose to glucose.

However, no colonies on M9/Cellobiose plates were observed, indicating the lack of  $\beta$ -glucosidase enzyme activity which made them unable to hydrolyze cellobiose to glucose for cell survival (Liu et al., 2009).

When all transformants of the library were tested again on LB/X-Glu plates, they all grew, but no blue colour was developed in any of the variant. A blue colour develops in one of the colony after one month in the cold room, on LB/cellobiose and X-Glu plates. *E.coli* Wildtype strain normally not able to grow on Beta-Glucoside likes cellobiose (Schnetz et al., 1987). The cellobiose utilizing genes are cryptic in wildtype *E.coli* and they normally not expressed, but it was reported that spontaneous mutation can activate and allow them to utilize cellobiose. Cryptic genes of *bgl* operon, *citA*, *citB* involve in citrate transport and catabolism of *p*-glucoside sugar arbutin and salicin. A another identified regulatory system *Cel* gene cluster involved in cellobiose metabolism (Hall et al., 1986). *E.coli* map is located at 37.8min (Hall et al., 1986). Study reveals that the mutation in *Cel* gene can also hydrolyze *p*-nitrophenyl- $\beta$ -D-glucoside which is an artificial substrate giving a blue color (Hall et al., 1986). The possible reason of developing blue color in colony after a month on synthetic chromogenic substrate 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-glucopyranoside could be the spontaneous mutation of the cryptic gene *Cel* cluster in the host *E.coli* XL1Blue cell.

Beta-Xylosidase activity was screened on LB plates with chromogenic substrate 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside (X-Xyl) as an indicator.

Most of the bacterial species including *E. coli*-K12 and its derivatives do not contain endogenous xylosidase activity (Whitehead, 1997, Geddie, 2004). This makes the screening easy by avoiding false positive due to endogenous xylosidase activity in host cells. Screening showed 10 early blue colonies were picked after 24 hours, from LB/X-Xyl plates. From this 5 variants were again selected based on color intensity on plates in order to compare with the wildtype  $\beta$ -Galactosidase. Plasmid DNA from these selected 5 variants together with DNA of the wildtype plasmid pTZ1 was retransformed in order to validate that the presence of blue color were due to plasmid DNA. The plasmid DNA transformation of all mutants has resulted in clear blue colonies but not the wildtype (very faint blue).

A second round of mutagenesis plasmids DNA from colony 2 and 3 were initially chosen on basis of their higher blue colour intensity on plates in comparison to wild type. Later colony 1 was also chosen due to its first appearance on the X-Xyl plate after 24hour of incubation. A total 17 mutants were selected from the second round of random mutagenesis (14 variant from parent mutant colony 2 and 3, 1 variant C1SG1 from colony 1 and 2 mutants F9SG1, F9SG2 from colony 9) for  $\beta$ -Xylosidase activity on LB/X-Xyl plates. 14 variants from parental colony 2, 3 and 10 clones of first round were further screened on colorimetric enzyme assay along with wildtype pTZ1 for selection of a best variant, but none of them shown higher activity than colony 9 of first generation on x-xyl plates.

First retransformation of plasmid DNA from colony 9 shown low blue colour intensity on x-xyl plates than rest of the mutants, but higher activity than wildtype. Activity measurements of colony 9 on another two assays with PNPX and x-xyl were also not consistent (see figure: 30 and 31).

To verify previous transformation, again retransformation was performed with plasmid DNA from colony 9 on LB/X-Xyl plates. The second retransformation showing no colour development on chromogenic substrate X-xyl, all colonies were white. Further verification was done by sequencing and the result shown a mutation outside the LacZ coding sequence at +49 position close to ribosomal binding site (+36). No mutation was identified on LacZ coding sequence. The reason of observed initial colour development colony 9 on plates and higher activity on assays (table: 13) was still not been determined. Further study also needed to investigate the validation of second generation mutants' transformation on LB/X-Xyl plates showing clear blue colonies on F9SG1 at ~24 hour of incubation, but no blue colour colonies resulting from F9SG2. Colony C1SG1 should have been retransformed but it was not performed. To obtain enhanced  $\beta$ -galactosidase enzyme activity the mutation library was screened on LB plates with chromogenic substrate 5-Bromo-4-chloro-3-indolyl -D-galactopyranoside (X-Gal) as indicator. A total of 26 variants were selected from such plates and by retesting on plates as well as by filter lift assay.

The disadvantage of screening of higher Beta-Galactosidase phenotypically on LB plates with chromogenic substrate is that all the mutants and wildtype turn blue. It was difficult to judge the highest blue colonies on LB plates.

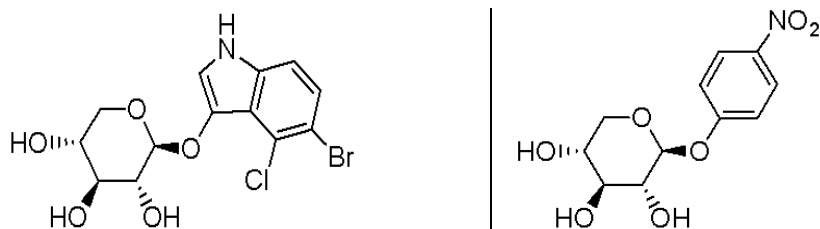
#### **6.4 Colorimetric Enzyme Assay**

Enzyme activity of 24 mutants variants (10 first generation and 14 second generation) have shown that the plasmid in colony 9 exhibited the highest activity with 42.25unit/ml (see table: 13) using 4-nitrophenyl-Beta-D-xylopyranoside as substrate.

The final selected mutants (Colony1, Colony9, C1SG1, and F9SG1and F9SG2) were again assayed on PNPX and x-xyl substrate. Assays were repeated several times, data of two recent assays are shown in the table 14. The activity of colony 9 in first assay showing again higher activity (14unit/ml/hr.) than wild type but repeating this assay with same mutants showing low enzyme activity with 10unit/ml/hr.). All remaining mutants (Colony1, C1SG1, and F9SG1and F9SG2) had lower enzyme activities than wildtype in both of the assays (see table: 14). However, colony 1 and F9SG1 demonstrated clear visual signal of  $\beta$ -xylosidase activity on plates. The reason of these two different results is still to be verified. But it could indicate that the two different substrates (5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside and 4-Nitrophenyl  $\beta$ -D-xylopyranoside) are recognized very differently by the mutants. Sequencing analysis of plasmid DNA from colony 9 were showed a mutation at +49 positions outside the LacZ coding sequence. This position is close to ribosomal binding site (+36 position) which could be the reason for the higher activities but retransformation of plasmid DNA from colony9 showed white colonies, making it necessary to study this further.

Use of microtiter plates for enzyme assays allowed their quantitative and kinetic measurements by absorbance of the hydrolyzed *p*-nitrophenyl dye linked to polysaccharides (Robertson and Steer, 2004). Enzyme activity was measured in Units (one unit of enzyme defined as the amount of enzyme activity that produce one  $\mu$ mole of *p*-nitrophenol (reducing sugar) in one minute (Khan et al., 1986).

For Beta-Xylosidase activities measurements, the substrate “5-bromo-4-chloro-3-indolyl- $\beta$ -D-Xylopyranside”was used on plates. We have used the substrate “4-nitrophenyl-Beta-D-xylopyranoside” for colorimetric assays.



Chemical structure of 5-Bromo-4-chloro-3-indolyl  $\beta$ -D-xylopyranoside and 4Nitrophenyl  $\beta$ -D-xylopyranoside

The use of two different substrates could be the reason of the different results between what was observed on the colorimetric assay data. This could possibly reflect different enzyme specificities for the different substrates. Using of this model substrate could be misleading in direct evolution studies. The selected mutants with improved activity on plates from the mutation library on substrate analog could be inactive to hydrolyze glycosidic bond of xylooligosaccharides or the possible converse situation (Wagschal et al., 2005). Therefore glycosidic bond hydrolysis in an analog substrate may not indicate activities of the enzyme of interest. Screening based on these method (colorimetric assay) should be used to identify the group of enzyme candidates and followed by more specific screening to obtain the ultimate gene sequence of interest (Robertson and Steer, 2004).

For improved  $\beta$ -galactosidase activity, 26 selected dark blue on LB plates with x-gal further screened on a 96-well microtiter plate using the substrate analog *o*-nitrophenyl  $\beta$ -D-galactopyranoside. The screening was based on quantitating the activities of beta-galactosidase as described by J. H. Miller (1972). The final measured beta-galactosidase activities values obtained were expressed in Miller units (Griffith and Wolf, 2002). Screening of higher  $\beta$ -galactosidase was also performed with the mutants screened for  $\beta$ -xylosidase activity on X-Xyl plates but assay data were not reproducible. The activity of the mutants was not reproducible.

## 6.5 Sequencing of Selected Mutants

Sequencing analysis was performed for all the five final selected mutants (Colony1, Colony9, C1SG1, and F9SG1and F9SG2). Colony 1 and C1SG1 as mutation



K552E (Lysine number 552 was replaced by glutamic) located at domain 3 (residues 334-627) of the wildtype *E.coli*  $\beta$ -galactosidase. In the native beta-galactosidase protein, the majority of amino acids in domain 3 form substrate binding pocket. Blue colour appearance on colony1 and C1SG1 on X-Xyl plates could be possibly due to this mutation that may have affected the enzyme's substrate specificity and develop colour on X-Xyl plates.

Mutant C1SG1 was not sequenced completely and not even retransformed for verification. These both procedures should have been done in order to reach the conclusion regarding substrate specificity of the mutant. It was reported that Glu537 is the nucleophile residue in the enzyme. The conserved active site for *E. coli* LacZ is the residues starting by at residue 534 I L C E Y A H A M G N(Gebler et al., 1992) .

Amino acid substitutions were concluded from the sequence analysis for the other selected variants during in this study as N959Y (Asparagine 959 to Tyrosine) in F9SG1 and Q967R (Glutamine 967 to Arginine), M969L (Methionine 969 to Leucine) in mutant F9SG2. These mutation are present in the domain 5 of LacZ gene and not a part of active site residues. F9SG1 gave clear blue colonies after retransformation on X-Xyl plates but mutant F9SG2 shown white colonies after retransformation.

Additional work is needed to analyze how this mutation supports the enzyme to develop colour on this substrate. Some of the validation experiments might involve site-directed mutagenesis. But it is clear that further improvement for the method to create a large library with diversity in the range of 1 to 2 substitutions per the coding sequence in order not to drastically deviate from an enzyme that can hydrolyze one or more glycosidic substrates. Moreover, experiments within a limited sequence coding for the substrate binding domain might be of interest as well.

## 7. Concluding Remarks

In the present study, in vitro evolution has been performed on the *E. coli* LacZ gene coding for  $\beta$ -galactosidase in order to create new variant enzymes with altered substrate specificities as well as enhanced  $\beta$ -galactosidase activity. The objective was to select a variant of  $\beta$ -galactosidase from random libraries created by error-prone PCR. We had to sequence the entire pTZ1 plasmid in order to set a clear complete sequence and map to make our plans more clear and to avoid ambiguities. This study has shown that the plasmid is 5,502 bp.

We have partially achieved our target by identifying a limited number of variants with altered substrate specificity of  $\beta$ -galactosidase to  $\beta$ -xylosidase based on colour development using the substrate 5-bromo-4-chloro-3-indolyl- $\beta$ -D-xylopyranside (X-Xyl). Retransformation of mutant plasmids DNA has also showed the clear blue colonies on LB/X-Xyl plates which is another validation. One of the selected mutant (on x-xyl plate) F9 had a mutation near RBS, thus it could explain its higher activities for  $\beta$ -xylosidase for assays.

Moreover, positive mutants selected from x-xyl plates, retransformed and sequenced for further validation. They showed mutations at Domain 3 (K552E) a substrate binding site and domain 5 (N959Y) of LacZ gene (1023 amino acids). We believe that the whole plasmid error-prone PCR method used during this project is the fast and efficient approach but it has to be further optimized.

Creating mutants with new enzyme substrate specificity for glycosidic bonds and to enhance activity of  $\beta$ -galactosidase from the random pool was not completed, but with some results that can be further improved and validated if we have more time.

## 8. Future Perspectives

- Enzyme kinetics study of  $\beta$ -xylosidase & beta -galactosidase and its new substrates with mutant colony 1(K552E, Lysine 552 to Glutamic acid) and F9SG1 (N959Y, Asparagine 959 to Tyrosine).
- Perform further round of random mutagenesis in order to obtain higher  $\beta$ -galactosidase activity and alter substrate specificity for various glycosidases.
- Perform other method of random mutagenesis such as sequence saturation mutagenesis (SeSaM), Staggered Extension Process (StEP), DNA shuffling and Mutagenic and unidirectional reassembly (MURA) to achieve the remaining targets.
- Random mutagenesis within the coding sequence for the substrate-binding domain.
- Validate any mutant of interest by site-directed mutagenesis

## 9. References

### 9.1 Articles

- ANDREWS, K. J. & LIN, E. C. 1976. Thiogalactoside transacetylase of the lactose operon as an enzyme for detoxification. *J. Bacteriol.*, 128, 510-513.
- ARISTARKHOVA, S. A., BURLAKOVA, E. B. & SHELUDCHENKO, N. I. 1979. [Effect of lecithin on liver microsomal lipid peroxidation]. *Biokhimiia*, 44, 125-9.
- BAUCH, T., BOCKER, W., MALLEK, U., MULLER, W. U. & STREFFER, C. 1999. Optimization and standardization of the "comet assay" for analyzing the repair of DNA damage in cells. *Strahlenther Onkol*, 175, 333-40.
- BECKMAN, R. A., MILDVAN, A. S. & LOEB, L. A. 1985. On the fidelity of DNA replication: manganese mutagenesis in vitro. *Biochemistry*, 24, 5810-7.
- BEGUIN, P. & AUBERT, J. P. 1994. The biological degradation of cellulose. *FEMS Microbiol Rev*, 13, 25-58.
- BERG, J. M., TYMOCZKO, J. L. & STRYER, L. 2001. *Biochemistry*, New York, Freeman.
- BHAT, M. K. 2000. Cellulases and related enzymes in biotechnology. *Biotechnol Adv*, 18, 355-83.
- BIELY, P. 1985. Microbial xylanolytic systems. *Trends in biotechnology*, 3, 286-290.
- BRODY, J. R. & KERN, S. E. 2004. History and principles of conductive media for standard DNA electrophoresis. *Anal Biochem*, 333, 1-13.
- BRUX, C., BEN-DAVID, A., SHALLOM-SHEZIFI, D., LEON, M., NIEFIND, K., SHOHAM, G., SHOHAM, Y. & SCHOMBURG, D. 2006. The structure of an inverting GH43 beta-xylosidase from *Geobacillus stearothermophilus* with its substrate reveals the role of the three catalytic residues. *J Mol Biol*, 359, 97-109.
- CADWELL, R. C. & JOYCE, G. F. 1992. Randomization of genes by PCR mutagenesis. *PCR Methods Appl*, 2, 28-33.
- CANTAREL, B. L., COUTINHO, P. M., RANCUREL, C., BERNARD, T., LOMBARD, V. & HENRISSAT, B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research*, 37, D233-D238.
- CHEONG, T. K. & ORIEL, P. J. 2000. Cloning of a wide-spectrum amidase from *Bacillus stearothermophilus* BR388 in *Escherichia coli* and marked enhancement of amidase expression using directed evolution\*. *Enzyme Microb Technol*, 26, 152-158.
- CLINE, J., BRAMAN, J. C. & HOGREFE, H. H. 1996. PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res*, 24, 3546-51.
- COLE, M. F. & GAUCHER, E. A. 2011. Exploiting models of molecular evolution to efficiently direct protein engineering. *J Mol Evol*, 72, 193-203.
- DE BOER, H. A., COMSTOCK, L. J. & VASSER, M. 1983. The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proc Natl Acad Sci U S A*, 80, 21-5.
- DIEFFENBACH, C. W., LOWE, T. M. & DVEKSLER, G. S. 1993. General concepts for PCR primer design. *PCR Methods Appl*, 3, S30-7.

- DOI, R. H. & KOSUGI, A. 2004. Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat Rev Microbiol*, 2, 541-51.
- ENEYSKAYA, E. V., IVANEN, D. R., BOBROV, K. S., ISAEVA-IVANOVA, L. S., SHABALIN, K. A., SAVEL'EV, A. N., GOLUBEV, A. M. & KULMINSKAYA, A. A. 2007. Biochemical and kinetic analysis of the GH3 family beta-xylosidase from *Aspergillus awamori* X-100. *Arch Biochem Biophys*, 457, 225-34.
- FASAN, R., CROOK, N. C., PETERS, M. W., MEINHOLD, P., BUELTER, T., LANDWEHR, M., CIRINO, P. C. & ARNOLD, F. H. 2011. Improved product-per-glucose yields in P450-dependent propane biotransformations using engineered *Escherichia coli*. *Biotechnol Bioeng*, 108, 500-10.
- FENTON, C., XU, H., PETERSEN, E. I., PETERSEN, S. B. & EL-GEWELY, M. R. 2002a. In Vitro Mutagenesis Protocols. In: BRAMAN, J. (ed.) *Methods Mol Biol*. second 2002/01/05 ed.: Humana Press.
- FENTON, C., XU, H., PETERSEN, E. I., PETERSEN, S. B. & EL-GEWELY, M. R. 2002b. Random mutagenesis for protein breeding. *Methods Mol Biol*, 182, 231-41.
- FOWLER, A. V. & ZABIN, I. 1970. The amino acid sequence of beta galactosidase. I. Isolation and composition of tryptic peptides. *J Biol Chem*, 245, 5032-41.
- GEBLER, J. C., AEBERSOLD, R. & WITHERS, S. G. 1992. Glu-537, not Glu-461, is the nucleophile in the active site of (lac Z) beta-galactosidase from *Escherichia coli*. *J Biol Chem*, 267, 11126-30.
- GEDDIE, M. L. 2004. Rapid Evolution of -Glucuronidase Specificity by Saturation Mutagenesis of an Active Site Loop. *Journal of Biological Chemistry*, 279, 26462-26468.
- GRIFFITH, K. L. & WOLF, R. E., JR. 2002. Measuring beta-galactosidase activity in bacteria: cell growth, permeabilization, and enzyme assays in 96-well arrays. *Biochem Biophys Res Commun*, 290, 397-402.
- HALL, B. G., BETTS, P. W. & KRICKER, M. 1986. Maintenance of the cellobiose utilization genes of *Escherichia coli* in a cryptic state. *Mol Biol Evol*, 3, 389-402.
- HANSEN, L. H., KNUDSEN, S. & SORENSEN, S. J. 1998. The effect of the lacY gene on the induction of IPTG inducible promoters, studied in *Escherichia coli* and *Pseudomonas fluorescens*. *Curr Microbiol*, 36, 341-7.
- HANSEN MS & LT, J. 2002. *Breeding Specificity of Cutinase from Fusarium Solani PISI Towards Long Chain Triacylglycerols*. Master Master, Aalborg University.
- HARAYAMA, S. 1998. Artificial evolution by DNA shuffling. *Trends Biotechnol*, 16, 76-82.
- HENRISSAT, B., CALLEBAUT, I., FABREGA, S., LEHN, P., MORNON, J. P. & DAVIES, G. 1995. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc Natl Acad Sci U S A*, 92, 7090-4.
- HENRISSAT, B. & DAVIES, G. J. 2000. Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. *Plant Physiol*, 124, 1515-9.
- INNIS, M. A., MYAMBO, K. B., GELFAND, D. H. & BROW, M. A. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc Natl Acad Sci U S A*, 85, 9436-40.

- JACOBSON, R. H., ZHANG, X. J., DUBOSE, R. F. & MATTHEWS, B. W. 1994. Three-dimensional structure of beta-galactosidase from *E. coli*. *Nature*, 369, 761-6.
- JC, M. & FH, A. 1996. - Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nat Biotechnol*, 14, 458-67.
- JOHANSSON, B. G. 1972. Agarose gel electrophoresis. *Scand J Clin Lab Invest Suppl*, 124, 7-19.
- JORDAN, D. B., LI, X. L., DUNLAP, C. A., WHITEHEAD, T. R. & COTTA, M. A. 2007. Structure-function relationships of a catalytically efficient beta-D-xylosidase. *Appl Biochem Biotechnol*, 141, 51-76.
- KAUR, J. & SHARMA, R. 2006. Directed evolution: an approach to engineer enzymes. *Crit Rev Biotechnol*, 26, 165-99.
- KHAN, A. W., TREMBLAY, D. & LEDUY, A. 1986. Assay of xylanase and xylosidase activities in bacterial and fungal cultures. *Enzyme and Microbial Technology*, 8, 373-377.
- LANGLEY, K. E., VILLAREJO, M. R., FOWLER, A. V., ZAMENHOF, P. J. & ZABIN, I. 1975. Molecular basis of beta-galactosidase alpha-complementation. *Proc Natl Acad Sci U S A*, 72, 1254-7.
- LEBBINK, J. H., KAPER, T., BRON, P., VAN DER OOST, J. & DE VOS, W. M. 2000. Improving low-temperature catalysis in the hyperthermostable *Pyrococcus furiosus* beta-glucosidase CelB by directed evolution. *Biochemistry*, 39, 3656-65.
- LEDERBERG, J. 1950. The beta-d-galactosidase of *Escherichia coli*, strain K-12. *J Bacteriol*, 60, 381-92.
- LEEMHUIS, H., KELLY, R. M. & DIJKHUIZEN, L. 2009. Directed evolution of enzymes: Library screening strategies. *IUBMB Life*, 61, 222-8.
- LIN, H. & CORNISH, V. W. 2002. Screening and selection methods for large-scale analysis of protein function. *Angew Chem Int Ed Engl*, 41, 4402-25.
- LIU, W., HONG, J., BEVAN, D. R. & ZHANG, Y. H. 2009. Fast identification of thermostable beta-glucosidase mutants on cellobiose by a novel combinatorial selection/screening approach. *Biotechnol Bioeng*, 103, 1087-94.
- LOVE, D. R., FISHER, R. & BERGQUIST, P. L. 1988. Sequence structure and expression of a cloned beta-glucosidase gene from an extreme thermophile. *Mol Gen Genet*, 213, 84-92.
- LYND, L. R., CUSHMAN, J. H., NICHOLS, R. J. & WYMAN, C. E. 1991. Fuel ethanol from cellulosic biomass. *Science*, 251, 1318-23.
- MATSUMURA, I. & ELLINGTON, A. D. 2001. In vitro Evolution of Beta-glucuronidase into a Beta-galactosidase Proceeds Through Non-specific Intermediates. *Journal of Molecular Biology*, 305, 331-339.
- MATSUMURA, I. & ROWE, L. A. 2005. Whole plasmid mutagenic PCR for directed protein evolution. *Biomol Eng*, 22, 73-9.
- MATTHEWS, B. W. 2005. The structure of *E. coli* beta-galactosidase. *C R Biol*, 328, 549-56.
- MILDVAN, A. S. & LOEB, L. A. 1979. The role of metal ions in the mechanisms of DNA and RNA polymerases. *CRC Crit Rev Biochem*, 6, 219-44.

- NAKAZAWA, H., OKADA, K., ONODERA, T., OGASAWARA, W., OKADA, H. & MORIKAWA, Y. 2009. Directed evolution of endoglucanase III (Cel12A) from *Trichoderma reesei*. *Appl Microbiol Biotechnol*, 83, 649-57.
- OHTSUKA, E., MATSUKI, S., IKEHARA, M., TAKAHASHI, Y. & MATSUBARA, K. 1985. An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. *Journal of Biological Chemistry*, 260, 2605-2608.
- PRIMROSE, S. B., TWYMAN, R. M. & OLD, R. W. 2006. *Principles of gene manipulation and genomics*, Malden, Mass., Blackwell.
- RAHME, L. G., AUSUBEL, F. M., CAO, H., DRENKARD, E., GOUMNEROV, B. C., LAU, G. W., MAHAJAN-MIKLOS, S., PLOTNIKOVA, J., TAN, M. W., TSONGALIS, J., WALENDZIEWICZ, C. L. & TOMPKINS, R. G. 2000. Plants and animals share functionally common bacterial virulence factors. *Proc Natl Acad Sci U S A*, 97, 8815-21.
- ROBERTSON, D. E. & STEER, B. A. 2004. Recent progress in biocatalyst discovery and optimization. *Curr Opin Chem Biol*, 8, 141-9.
- ROMERO, P. A. & ARNOLD, F. H. 2009. Exploring protein fitness landscapes by directed evolution. *Nat Rev Mol Cell Biol*, 10, 866-76.
- RUBIN-PITEL, S. B. & ZHAO, H. 2006. Recent advances in biocatalysis by directed enzyme evolution. *Comb Chem High Throughput Screen*, 9, 247-57.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. 1989. *Molecular cloning: a laboratory manual*, Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory.
- SCHNETZ, K., TOLOCZYKI, C. & RAK, B. 1987. Beta-glucoside (bgl) operon of *Escherichia coli* K-12: nucleotide sequence, genetic organization, and possible evolutionary relationship to regulatory components of two *Bacillus subtilis* genes. *J Bacteriol*, 169, 2579-90.
- SHALLOM, D., LEON, M., BRAVMAN, T., BEN-DAVID, A., ZAIDE, G., BELAKHOV, V., SHOHAM, G., SCHOMBURG, D., BAASOV, T. & SHOHAM, Y. 2005. Biochemical characterization and identification of the catalytic residues of a family 43 beta-D-xylosidase from *Geobacillus stearothermophilus* T-6. *Biochemistry*, 44, 387-97.
- SHEWALE, J. G. 1982. Beta-Glucosidase: its role in cellulase synthesis and hydrolysis of cellulose. *Int J Biochem*, 14, 435-43.
- SILHAVY, T. J., BERMAN, M. L. & ENQUIST, L. W. 1984. *Experiments with gene fusions*, Cold Spring Harbor, New York, Cold Spring Harbor Laboratory.
- SNYDER, L. & CHAMPNESS, W. 2007. *Molecular genetics of bacteria*, Washington D.C., ASM Press.
- SONG, J. K., CHUNG, B., OH, Y. H. & RHEE, J. S. 2002. Construction of DNA-shuffled and incrementally truncated libraries by a mutagenic and unidirectional reassembly method: changing from a substrate specificity of phospholipase to that of lipase. *Appl Environ Microbiol*, 68, 6146-51.
- SPEE, J. H., DE VOS, W. M. & KUIPERS, O. P. 1993. Efficient random mutagenesis method with adjustable mutation frequency by use of PCR and dITP. *Nucleic Acids Res*, 21, 777-8.
- STEMMER, W. P. 1994. Rapid evolution of a protein in vitro by DNA shuffling. *Nature*, 370, 389-91.

- SU, T. Z., SCHWEIZER, H. & OXENDER, D. L. 1990. A novel phosphate-regulated expression vector in *Escherichia coli*. *Gene*, 90, 129-33.
- SUBRAMANIYAN, S. & PREMA, P. 2002. Biotechnology of microbial xylanases: enzymology, molecular biology, and application. *Crit Rev Biotechnol*, 22, 33-64.
- TAO, H. & CORNISH, V. W. 2002. Milestones in directed enzyme evolution. *Curr Opin Chem Biol*, 6, 858-64.
- TRACEWELL, C. A. & ARNOLD, F. H. 2009. Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Curr Opin Chem Biol*, 13, 3-9.
- VOGEL, A. I. & MENDHAM, J. 2000. *Vogel's textbook of quantitative chemical analysis*, Harlow, Prentice Hall.
- WAGSCHAL, K., FRANQUI-ESPIET, D., LEE, C. C., ROBERTSON, G. H. & WONG, D. W. 2005. Enzyme-coupled assay for beta-xylosidase hydrolysis of natural substrates. *Appl Environ Microbiol*, 71, 5318-23.
- WAGSCHAL, K., HENG, C., LEE, C. C. & WONG, D. W. 2009. Biochemical characterization of a novel dual-function arabinofuranosidase/xylosidase isolated from a compost starter mixture. *Appl Microbiol Biotechnol*, 81, 855-63.
- WC, L., TP, P., PRETLOW TG, N. & LA, C. 1990. - Bacterial lacZ gene as a highly sensitive marker to detect micrometastasis. *Cancer Res*, 50, 2808-17.
- WHITEHEAD, T. R. 1997. Development of a bifunctional xylosidase/arabinosidase gene as a reporter gene for the gram-negative anaerobes *Bacteroides* and *Porphyromonas*, and *Escherichia coli*. *Curr Microbiol*, 35, 282-6.
- WONG, K. K., TAN, L. U. & SADDLER, J. N. 1988. Multiplicity of beta-1,4-xylanase in microorganisms: functions and applications. *Microbiol Rev*, 52, 305-17.
- XU, H., PETERSEN, E. I., PETERSEN, S. B. & EL-GEWELY, M. R. 1999. Random mutagenesis libraries: optimization and simplification by PCR. *Biotechniques*, 27, 1102-4, 1106, 1108.
- YOU, L. & ARNOLD, F. H. 1996. Directed evolution of subtilisin E in *Bacillus subtilis* to enhance total activity in aqueous dimethylformamide. *Protein Eng*, 9, 77-83.
- ZAKOUR, R. A., KUNKEL, T. A. & LOEB, L. A. 1981. Metal-induced infidelity of DNA synthesis. *Environ Health Perspect*, 40, 197-205.
- ZHANG, J. H., DAWES, G. & STEMMER, W. P. 1997. Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc Natl Acad Sci U S A*, 94, 4504-9.
- ZHAO, H. & ARNOLD, F. H. 1997. Combinatorial protein design: strategies for screening protein libraries. *Curr Opin Struct Biol*, 7, 480-5.



## 9.2 Websites

- A. [www.neb.com](http://www.neb.com)
- B. [www.biocompare.com](http://www.biocompare.com)
- C. <http://www.ncbi.nlm.nih.gov>
- D. <http://www.yorku.ca/eye/spectru.htm>
- E. <http://www.sigmaaldrich.com/norway.html>
- F. <http://bitesizebio.com/articles/the-basics-how-ethanol-precipitation-of-dna-and-rna-works/>
- G. <http://www.chem.qmul.ac.uk/iubmb/enzyme/EC3/>
- H. <http://www.cazy.org/Glycoside-Hydrolases.html>
- I. <http://www.ncbi.nlm.nih.gov/nuccore/434652>
- J.



