IDENTIFICATION BY SATURATION MUTAGENESIS OF A SINGLE RESIDUE INVOLVED IN THE α-GALACTOSIDASE AGAB REGIOSELECTIVITY

By

Oyamo George Osanjo

B. Pharm. (UoN)

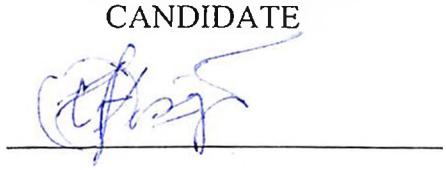
A thesis submitted in partial fulfilment of the requirements for the award of the degree of Master of Science of the University of Nairobi

2002

DECLARATION

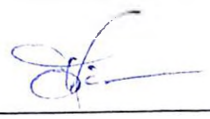
I, Oyamo George Osanjo, hereby declare that this thesis is my original work and has not been presented for a degree in any other University.

OYAMO GEORGE OSANJO
CANDIDATE

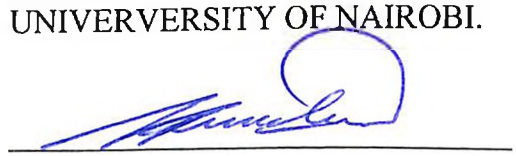


This thesis has been submitted for examination with our approval as supervisors :

DR. M. DION
SUPERVISOR,
UNIVERSITY OF NANTES.



DR. F. MULAA
SUPERVISOR,
UNIVERVERSITY OF NAIROBI.



PROF. D. MAKAWITI
CHAIRMAN,
DEPT. OF BIOCHEMISTRY.

11 June 2002

ACKNOWLEDGEMENTS

I am greatly indebted to the French Ministry of Foreign Affairs and to the University of Nairobi for the scholarship which made this work possible.

My deep appreciation to my principal supervisor, Dr Michel Dion, without whose guidance and constructive criticism this project would not have been accomplished. I am also grateful to Prof. Charles Tellier for welcoming me to his laboratory. I equally recognise the assistance accorded to me by members of Prof. Claude Rabiller's team; more so Corinne and Petra, on chemical techniques especially on NMR.

I acknowledge my supervisor Dr. Francis Mulaa for his helpful suggestions and the chairman of biochemistry department, Prof. D. Makawiti for his advice and encouragement during the project.

Special thanks to all the members of the "Unité de Recherche sur la Biocatalyse" for the warm ambience in and out of the laboratory that made my stay in Nantes so memorable.

Merci à tous!

CONTENTS

LIST OF FIGURES

FIGURE **PAGE**

SCHEMES PAGE

TABLES

ABBREVIATIONS

APS	Ammonium persulphate
BCA	Bicichoninic acid
BSA	Bovine serum albumin
DMSO	Dimethylsulfoxide
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
DTT	Dithiothreitol
EDTA	Ethylene diamine tetraacetic acid
Galα-F	1-Fluoro α-D-galactopyranoside
Gal-α-(1→2)-Gal-α-O-*p*NP	*p*-Nitrophenyl α-D-galactopyranosyl- (1→2)- α-D-galactopyranoside
Gal-α-(1→3)-Gal-α-O-*p*NP	*p*-Nitrophenyl α-D-galactopyranosyl- (1→3)- α-D-galactopyranoside
Gal-α-(1→6)-Gal-α-O-*p*NP	*p*-Nitrophenyl α-D-galactopyranosyl- (1→6)- α-D-Galactopyranoside
Gal-α-(1→6)-Gal-β-O-*p*NP	*p*-Nitrophenyl β-D-galactopyranosyl- (1→6)- α-D-galactopyranoside
IMAC	Immobilized metal affinity chromatography
Kd	Dissociation constant
LB	Luria Bertani (medium)
NMR	Nuclear magnetic resonance

NTA	Nitriloacetic acid
*o*NPαGal	*o*-Nitrophenyl α-D-galactopyranoside
*o*NPβGal	*o*-Nitrophenyl β-D-galactopyranoside
PCR	Polymerase chain reaction
*p*NPαGal	*p*-Nitrophenyl α-D-galactopyranoside
*p*NPαGlc	*p*-Nitrophenyl α-D-glucopyranoside
Rf	Relative front
RBS	Ribosome binding site
RNA	Ribonucleic acid
RNase	Ribonuclease
SDS	Sodium dodecyl sulphate
SDS PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	Simple and efficient method (of transformation)
TAE	Tris-acetate EDTA buffer
TE	Tris-EDTA buffer
TLC	Thin layer chromatography
TMS	3- (trimethylsilyl)-propanesulfonic acid
X-α-Gal	5-bromo-4-chloro-3-indolyl α-D-galactopyranoside
X-β-Gal	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside

UNIT ABBREVIATIONS

bp	base pair
cfu	colony forming unit(s)
hr	Hours
Kb	Kilo base
kD	kilo dalton
M	Molar
mg	Milligram
MHz	Megahertz
min	Minutes
mM	Millimolar
MW	Molecular weight
μg	Microgram
μl	Microlitre
nm	Nanometre
ppm	parts per million
rpm	Revolution per minute
sec	Seconds
v/v	Volume by volume

ABSTRACT

α-Galactosidase AgaB from *Bacillus stearothermophilus* displays a major α(1→6) and a minor α(1→3) regioselectivity in hydrolysis and transglycosylation. Its corresponding gene, *agaB*, was subjected to saturation mutagenesis at codon 442 in order to change its regioselectivity. The mutant genes were cloned and expressed in *Escherichia coli* using pBTac2 as vector. The regioselective activity of the mutants was determined using thin layer chromatography and nuclear magnetic resonance spectroscopy. A single point mutation, G442R, resulted in a mutant displaying an α(1→2) regioselectivity. Other amino acid substitutions at this site also gave mutants with altered regioselectivity and transglycosylation profiles. This is the first demonstration that single point mutations can lead to a strong modification of the regioselectivity of a glycosyl hydrolase. The kinetic parameters of the enzyme variants were determined and a preliminary investigation of possible substrates for condensation reactions conducted.

1.0 INTRODUCTION AND LITERATURE REVIEW

1.1 GENERAL INTRODUCTION

Oligosaccharides have been recognized in recent years as important information mediators in biomolecular recognition. The oligosaccharide components of glycoproteins are implicated in tumoral invasion mechanisms (Stroud *et al.*, 1994 ; Muller *et al.*, 1999); protein stability (Wang *et al.*, 1996) and tethering of leukocytes to endothelial cells during the initial stages of the inflammatory response (Phillips *et al.*, 1990 ; Lowe *et al.*, 1990; Dwek, 1996; Somers *et al.*, 2000).

The oligosaccharide Galα(1→3)Galβ(1→4)GlcNac has been the subject of many studies. It is known to be responsible for hyperacute rejection to porcine xenografts in man. About 1% of human IgG recognize the trisaccharide epitope and cause the rejection (Vaughan *et al.*, 1994). It has been shown that soluble Galα(1→3)Gal disaccharide could block the human anti-pig xenoantibodies in a pig kidney *in vitro* perfusion model (Magnusson *et al.*, 2000). Various methodologies for the synthesis of carbohydrates and especially galactose oligosaccharides having various linkages are consequently being developed, to tap this source of novel drugs.

The chemical synthesis of carbohydrates at an industrial scale is however expensive because the presence of multiple hydroxyl groups on carbohydrates necessitates numerous protection and deprotection steps, followed by a tedious purification procedure to remove the toxic reagents (Schmidt, 1986).

The enzymatic synthesis of saccharides using glycosidases is a powerful alternative (Nilsson, 1987; Ichikawa *et al.*, 1992; Nilsson, 1997; Scigelova *et al.*,1999). The glycosidases hydrolyse glycosidic bonds but they can also catalyse the formation of glycosidic linkages via their transferase activity. Moreover, these enzymes are highly stable, stereoselective and inexpensive (Yoon *et al.*, 1996; Vetere *et al.*, 1996). However, the enzymes often lack the desired regioselectivity or give a mixture of regioisomers during transglycosylation (Leloir, 1971; Toone *et al.*, 1989; Chiffoleau-Giraud *et al.*, 1997; Chiffoleau-Giraud *et al.*, 1999; Petzelbauer *et al.*, 2000).

α-Galactosidase AgaB is a glycosyl hydrolase isolated from the thermophile *Bacillus stearothermophilus*. Being a thermostable enzyme, it can be potentially used for industrial production of oligosaccharides with therapeutic or commercial value. However, in transglycosylation reactions it displays a major α-(1→6) and minor α-(1→3) regioselectivity, thus it cannot be used to synthesise compounds having the latter linkage. In previous studies (Nisole, 1999 ; Glottin-Fleury, 2000; Dion *et al.*, 2001), AgaB was subjected to a directed molecular evolution process using random mutagenesis and staggered extension process (StEP). Consequently, mutant enzymes, E901 and E500, which displayed an improved α-(1→3) and a major α-(1→2) regioselectivity respectively, were obtained.

In this study, the directed *in vitro* molecular evolution process of α-galactosidase AgaB was continued by using the strategy of saturation mutagenesis. This method enabled the determination of the amino acid substitution involved in the regioselectivity observed in E500 and in addition, gave another mutant, F1 with a better α-(1→3) regioisomer yield compared to the parent AgaB and the mutant E901.

The kinetic parameters of AgaB and its mutants were also determined in an attempt to understand the transglycosylation reaction catalysed by these enzymes.

1.2 LITERATURE REVIEW

1.2.1 The Glycosyl hydrolases

The glycosyl hydrolases, also called glycosidases, are the enzymes that catalyse the hydrolysis of glycosidic bonds of di- or polysaccharides. The glycosidases are classified as exoglycosidases or endoglycosidases depending on whether they hydrolyse terminal or interior glycosidic bonds of polysaccharides. They can also be classified as α-D-glycosidases or β-D-glycosidases. The former hydrolyse the α glycosidic bonds while the latter attack the β glycosidic linkages of polysaccharides.

1.2.2 Classification of glycosyl hydrolases according to sequence homology

The enzyme commission classifies the glycosyl hydrolases according to substrate specificity. The code given to the glycosyl hydrolases is EC 3.2.*X.Y* where X represents the anomeric nature of the linkage recognised and Y is the substrate hydrolysed (for example: EC 3.2.2.23 for α-galactosidases and EC 3.2.1.23 for β -galactosidases).

The enzyme commission classification is simple and widely used but it is too constraining for the glycosyl hydrolases because these enzymes usually act on multiple substrates. A more useful classification system of glycosyl hydrolases was developed by Henrissat and Davies (1997). According to this system the glycosyl hydrolases are grouped into 81 families based on amino acid sequence homology. Most of the bacterial α-galactosidases display a significant degree of sequence homology and have been placed in family 36 (Henrissat, 1991 ; Henrissat *et al.*, 1993). With the exceptions of fungal α-galactosidases AGLII from *Trichoderma reesei* (Margolles-Clark *et al.*, 1996) and AglB from *Aspergillus niger* (de Vries *et al.*, 1999), which resemble bacterial α-galactosidases, the

eukaryotic enzymes are grouped in family 27. The *E.coli* α-galactosidases MelA (or Melibiase) is a notable exception; it has no resemblance to the bacterial α-galactosidases in family 36 (Burstein *et al.*, 1971) and has been grouped in family 4. Moreover, MelA unlike the bacterial and eukaryotic α-galactosidases requires NAD^+ and manganese ions as cofactors.

The amino acid consensus sequence of the bacterial α-galactosidases belonging to family 36 is located within the central region of the enzyme. The presence of this consensus sequence indicates a similar reaction mechanism or substrate binding site.

α-Galactosidase AgaB from *Bacillus stearothermophilus* is grouped in family 36. Its consensus amino acid sequence, thought to correspond to its active site, is located between the amino acid 360 and 369. Its consensus sequence together with those of other α-galactosidases are presented in table 1.

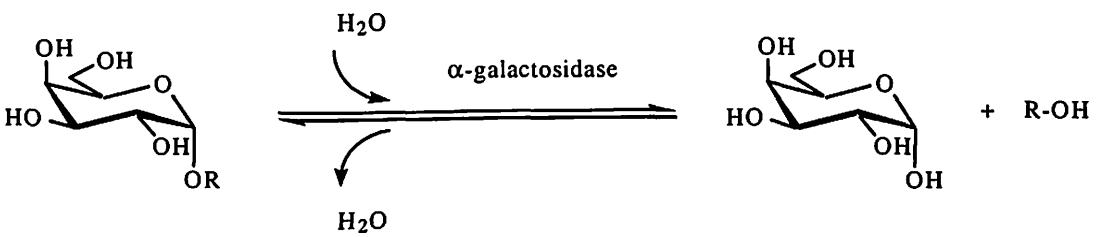
Enzyme	**Consensus sequence**
Aga1 (*Streptococcus mutans*)	I E L F V L D D G W
RafA (*Escherichia coli*)	V E R F I I D D G W
AgaR (*Pediococcus pentosaceus*)	I E M F V L D D G W
AgaS (*P. pentosaceus*)	L Q M L V L D D G W
AGLII (*Trichoderma reesei*)	I C L F V L D D G W
AgaN (*Bacillus stearothermophilus*)	I E L F V L D D G W
AgaB (*B. stearothermophilus*)	**I E L F V L D D G W**

Table 1 : The consensus sequence of some α-galactosidases belonging to family 36 (Adapted from Fridjonsson *et al.*, 1999).

1.2.3 Reactions catalysed by α-galactosidases

1.2.3.1 Hydrolysis reactions

α-galactosidases, *in vivo*, hydrolyse α-galactose residues from oligosaccharides and polymeric galactomannans (Luonteri *et al.*, 1998; Margolles-Clark *et al.*, 1996 ; Shibuya *et al.*, 1995).

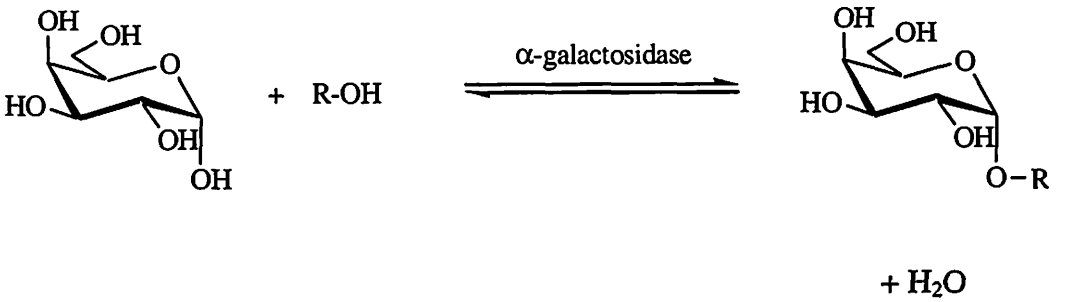


Scheme 1 : Hydrolysis catalysed by α-galactosidase

Given that all steps involved in the reactions catalysed by glycosidases are reversible, the enzymes can be used to catalyse the synthesis of glycosidic linkages (Scheme 1). There are two main ways of achieving preparation of glycoconjugates by the glycosidases : either by using a thermodynamically controlled synthesis reaction, or by employing a kinetically controlled transglycosylation reaction: when temperature of the reaction medium is raised, the tendency is to form a thermodynamic product rather than a kinetic product.

1.2.3.2 Reactions of synthesis

The synthetic reaction is the opposite of hydrolysis reaction, with the free monosaccharide as the substrate of the enzyme (Scheme 2).

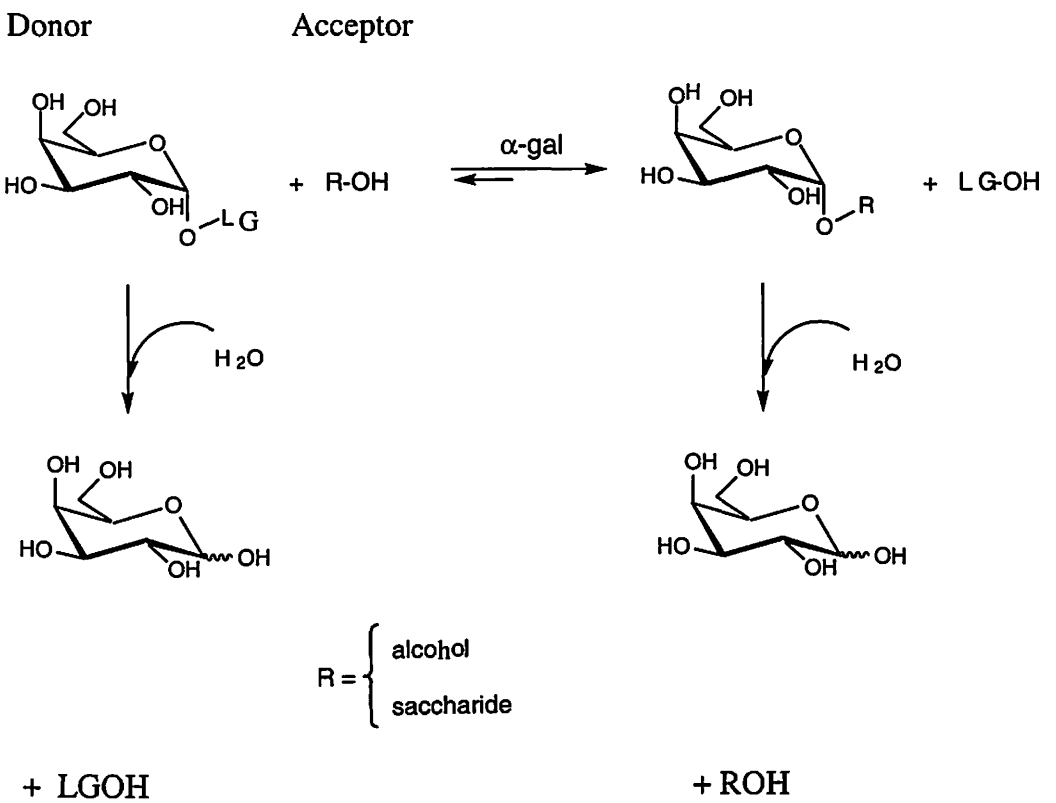


Scheme 2: The glycosylation reaction catalysed by α-galactosidase

The thermodynamic equilibrium of the reaction is in favour of the hydrolysis reaction, resulting in poor yield for the synthesis reaction (of about 10%). The reaction yield can be improved by removing the product of the reaction during the course of the experiment (Ajisaka *et al.*, 1987) or by diminishing the activity of water. However the yield normally remains low.

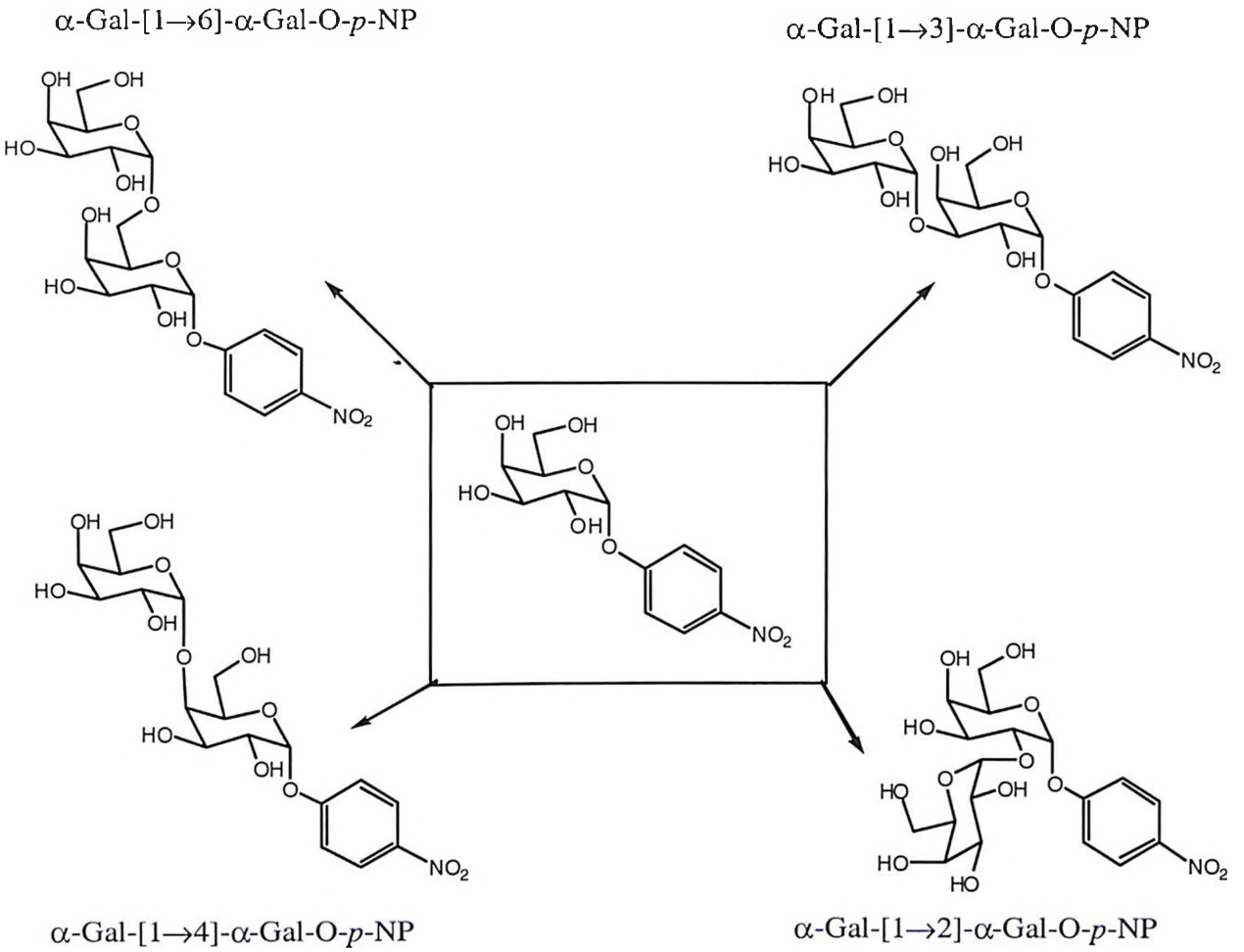
1.2.3.3 Transglycosylation reactions

Transglycosylation (Scheme 3) constitutes a kinetic approach of achieving enzymatic synthesis of glycosidic bonds. It requires a galactose donor molecule activated by a good leaving group (LG) and an acceptor molecule (ROH) having a free hydroxyl group such as another monosaccharide or an aliphatic alcohol. In this reaction the acceptor acts as a nucleophile and attacks the hydroxyl group in the anomeric position of the donor, in an overall nucleophilic substitution reaction.



Scheme 3: Transglycosylation reaction in competition with hydrolysis.

In aqueous media, the transglycosylation reaction is in competition with hydrolysis of activated galactose as well as hydrolysis of the transglycosylated product. The yield is higher compared to the synthesis reaction, but it can be improved further by increasing the rate of the reaction or reducing the hydrolytic activity. Whichever the approach used, the yield of transglycosylation reaction and the regioselectivity of the product highly depend on the nature of the enzyme. This explains the considerable effort undertaken to purify new enzymes or to use protein engineering techniques to improve existing ones. A commonly used donor in these *in vitro* studies is *p*-nitrophenyl α-D-galactopyranoside (*p*NPαGal). Scheme 4 summarises the possible disaccharides that can result from self condensation of *p*NPαGal catalysed by an α-galactosidase.



Scheme 4 : The possible disaccharides that can result from self condensation of *p*NPαGal catalysed by an α-galactosidase.

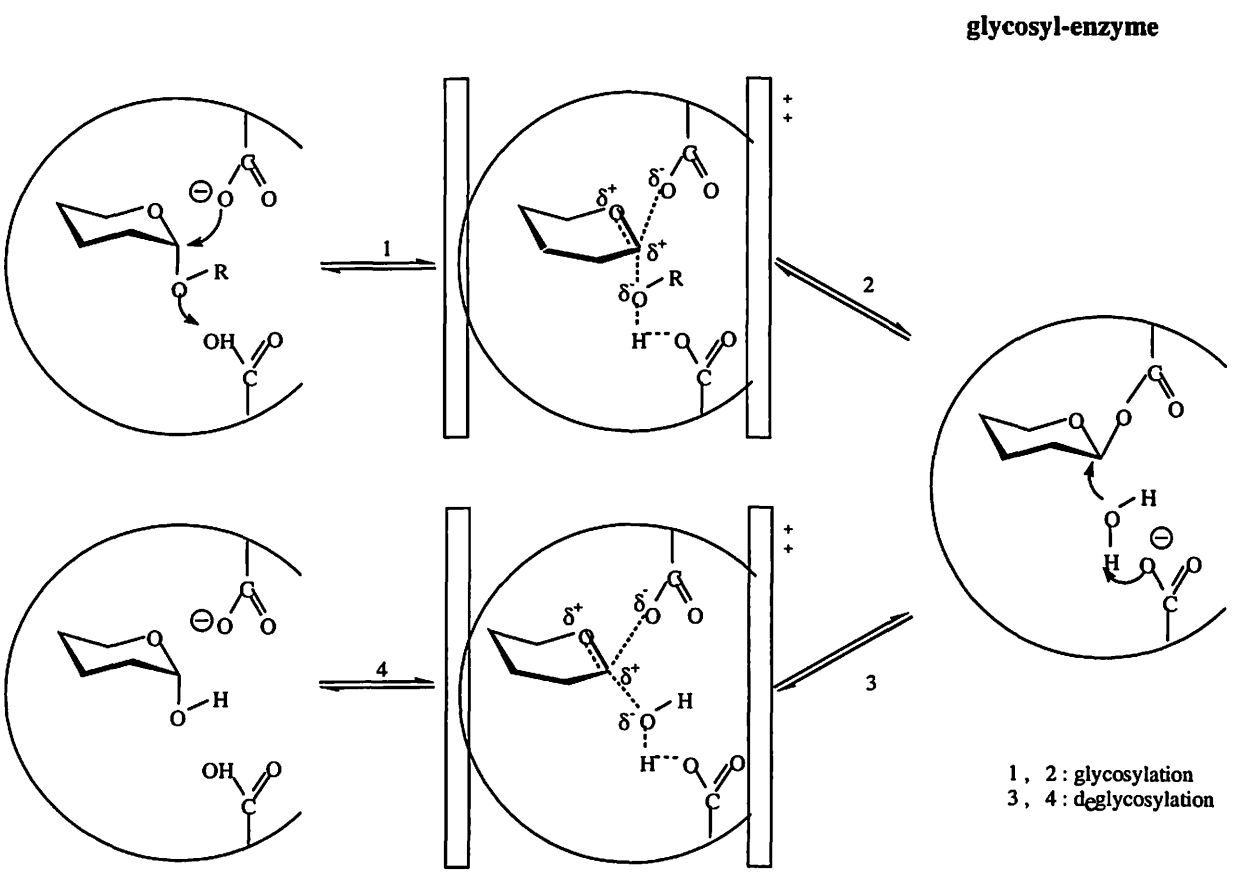
Most α-galactosidases present a major α(1→6) regioselectivity in synthesis due to the higher reactivity of the primary C_6 hydroxyl group compared to the other secondary hydroxyl groups.

1.2.4 Mechanism of catalysis

Glycosidases are thought to have two carboxylic acid groups contributed by aspartate or glutamate residues in their active sites (Stutz, 1996; Withers *et al.*, 1990; Ly *et al.*, 1999). The catalysis by the glycosidases can proceed either with retention or inversion of the configuration of the substrate (Wang *et al.*, 1995; Withers, 1995)

The reactions involving retention of configurations, such as seen in α- galactosidases, are two step reactions where one of the carboxylic acid groups play the role of a nucleophile enabling the attack of the anomeric carbon of the saccharide substrate. The other carboxylic acid behaves like an acid in the first step then subsequently as a base (Scheme 5).

During the first step it protonates the leaving group OR yielding an alcohol ROH. In the next step, it deprotonates the nucleophile (for example H_2O) that is attacking the glycosyl. The structures of the two transition states are believed to be high energy oxocarboniums. The mechanism of action therefore proceeds through two successive steps: the glycosylation step (1 and 2) and the deglycosylation step (3 and 4).



Scheme 5: Catalytic mechanism of glycosidases proceeding through retention of configuration (Wang *et al.*, 1995).

1.2.5 Thermostable proteins

The thermostability of α-galactosidase AgaB was an important factor in selecting it in favour of the other glycosidases of the same family given above, for the study. The source of the enzyme is the thermophile *B. stearothermophilus.* Thermophiles are examples of extremophiles. Extremophiles are microorganisms living in harsh ecosystems and can be divided into:

- **Thermophiles :** microorganisms living in high temperature environments (50-100°C).
- **Psychrophiles :** microorganisms living in low temperature environments (sometimes in conditions below 0°C).
- **Acidophiles and alkalophiles :** microorganisms living in low pH or high pH environments respectively.
- **Barophiles :** microorganisms living in high pressure environments.
- **Halophiles :** microorganisms living in highly salty environments.

The extremophiles have become important sources of enzymes used in research and industry. A celebrated example is the Taq DNA polymerase isolated from *Thermus aquaticus,* which revolutionised molecular biology by replacing the thermolabile DNA polymerase, in polymerase chain reactions (PCR).

A few glycosidases have also been obtained from thermophiles. Thus α-galactosidase AgaB was isolated from *B. stearothermophilus* strain KVE39 from Icelandic hot springs by Ganter *et al.* (1988). The native AgaB is thermostable, having optimum activity at 50°C. In this study, the screening procedure allowed the retention of only thermostable mutants (cf. Materials and Methods).

The principles underlying thermostability are now beginning to be understood due to

concerted research in the last decade. Stability at high temperatures require high stabilisation of the three dimensional structure of a protein even at high temperatures. 3-D stability is well correlated to a stable secondary structure made mainly of α-helices. This was demonstrated by Mathews (1996) during his study of T4 lysozyme folding. The number of hydrophobic amino acids, particularly alanine, have been found to be higher in the α-helices of these proteins (Menendez-Arias *et al.*, 1989). The presence of positive charges on the protein surface has also been found to be an important factor in thermostability (Borders *et al.*, 1994). The positive charges are due to a high presence of arginine residues and increase stability by forming ionic bridges or hydrogen bonds between the positive centres on arginine and the carbonyl group of the peptide bonds. Another factor considered to be important in stabilisation of thermostable proteins is their "compactness". The compactness of a protein increases : if there is a high proportion of proline residues; increase in hydrophobic interactions at the protein core and a diminution of the exposed protein surface (Watanabe *et al.*, 1991). Kirino *et al* (1994) also demonstrated, by directed mutagenesis, that the thermostability of isopropylmalate dehydrogenase was due to hydrophobic interactions between its two sub units.

1.2.6 Directed molecular evolution of AgaB

Directed molecular evolution of proteins is a technique that has been developed in recent years by protein engineers for efficient modification of protein structure and function. *In vitro* molecular evolution techniques require an iterative Darwinian optimisation algorithm: the target is to create molecular diversity of a protein and then to identify variants showing improved property of interest. The method chosen is typically mutagenesis together with recombination of the mutant genes (Zhao *et al.*, 1998; Jaeger *et al.*, 2000; Kuchner *et al.*, 1997). The genes encoding improved enzymes may be subjected to a next round of evolution.

It is important to establish a screening procedure that accurately selects the desired property (Ionna *et al.,* 2000).

Directed molecular evolution has been used to improve or alter enzymatic properties in the absence of structural or mechanistic information. Recent advances include the increase of the specific activity of β-lactamase (3200 times) or the fluorescence of Green Fluorescent Protein (45 times) (Stemmer, 1994 ; Crameri *et al.,* 1996).

To eliminate the requirement of a cofactor and thereby design a single-enzyme hydroxylation catalyst, Joo *et al.* (1999) directed the evolution of cytochrome $P450_{cam}$ to function via a peroxide shunt (where electrons required at different steps of the catalytic cycle are derived from peroxide instead of NAD[P]H). One round of random mutagenesis and recombination was sufficient to achieve this spectacular result.

Maxwell *et al.*(1999) engineered the increase in solubility of HIV integrase. By plating HIV integrase expression hosts on high levels of chloramphenicol, these workers selected a soluble HIV integrase mutant that had become fused with chloramphenicol acetyl transferase (CAT). When CAT is fused with insoluble proteins, lower resistance to chloramphenicol is obtained compared to fusions with soluble proteins, hence this is a very effective strategy in screening for soluble protein variants.

Thymidine kinase (TK) is an important target in the design of antiviral drugs because of the ability of the viral enzyme to phosphorylate thymidine analogues not recognised by the human enzyme. The specific activity of TK for the phosphorylation of zidovudine (AZT) was increased while at the same time reducing its specificity for thymidine. Christians and co-workers (1999) used family shuffling (*in vitro* recombination of naturally occurring homologous genes) to identify chimeras with a 32 times enhanced ability to phosphorylate AZT compared with the parent enzymes.

Directed evolution therefore offers a way to optimize enzyme function and the

technique has gained notable applications in enzyme engineering, being exploited to create improved industrial biocatalysts (Moore,1997).

With the desire to obtain an efficient biocatalyst for transglycosylation reactions to achieve synthesis of oligosaccharides, α-galactosidase AgaB has been subjected to directed molecular evolution strategies.

AgaB displays a major 1,6 and a minor 1,3 regioselectivity. In a previous study, it was subjected to molecular evolution using random mutagenesis and the staggered extension process (Nisole, 1999; Glottin-Fleury, 2000). The objective was to obtain AgaB mutants showing an improved 1,3 regioselectivity which is industrially important and if possible other regioselectivities for example 1,2 and 1,4 which are rarely found in naturally occurring α-galactosidases. To achieve the initial random mutagenesis and recombination, different techniques were used :

- **Error Prone PCR (EPP) technique**. This approach is based on the low fidelity of *Taq* polymerase during its amplification reaction (Cadwell, 1994). The enzyme incurs errors during the polymerisation phase generating random mutations. The parameters affecting *Taq* polymerase activity include: the concentration of $MgCl_2$, of $MnCl_2$ and each of the dNTPs.
- **The Staggered Extension Process (StEP)** is based on the *in vitro* recombination of entire genes during repeated cycles of denaturation and very brief renaturation-extension phases (Zhao *et al,* 1998). Denaturation is followed by random annealing of the newly synthesised DNA fragments with the matrix. These fragments are then prolonged by partial extension. The process is repeated several times over until full length DNA strands are obtained. This technique thus permits the generation of pool of genes having new combinations of mutations.

After cloning the AgaB mutants E901 (Nisole, 1999) and E500 (Glottin-Fleury, 2000) were obtained. The mutant E901 (6 mutations) possessed a higher α(1→3) regioselectivity relative to AgaB while E500 displayed a major α(1→2) regioselectivity. The enzyme E500 was found to have 10 amino acid changes: A2S, T4A, I34V, E46A, R48A, S50P, N11S, A198E, K399E and G442R (the sequence of AgaB is shown in figure 7) . The K399E and G422R substitutions can be argued to be the possible determinants of this regioselectivity due to their proximity to the AgaB consensus sequence. The other mutations are located farther away towards the N terminal of AgaB and, in addition, the high mutation of this region suggest that it is not important for substrate recognition.

The K399E mutation was, however, present in several other mutants not displaying α(1→2) regioselectivity (Dion *et al.*,2000). Hence, the G442R mutation is the most logical source of this regioselectivity.

In this project, this evolutionary process was continued, this time by introducing mutations at the site 442 which could be involved in specifying regioselectivity or transferase properties of the enzyme.

The technique of saturation mutagenesis was chosen because only a limited set of amino acid substitutions can be reached by classical PCR mutagenesis at low error rates. The technique has already achieved notable results. For instance Miyazaki *et al.* (1999) have used it to identify highly stabilising mutations in a key loop position of a psychrophilic protease.

Recently, saturation mutagenesis was applied in a directed evolution of an enantioselective catalyst (May *et al.,* 2000). These researchers succeeded in inverting the enantioselectivity of hydantoinase, a key enzyme in a multi-enzyme pathway for the production of L-methionine by introducing an I95F substitution in a saturation mutagenesis strategy. This technique was therefore chosen for the evolution strategy.

2.0 OBJECTIVES

The aim of this study was to :

i) determine whether G442R substitution is involved in the α-(1→2) regioselectivity; saturation mutagenesis being used as the method to introduce this mutation.

ii) determine whether other substitutions at the same position, could result in enzyme mutants with regioselectivity other than α-(1→6) or α-(1→2).

iii) analyse the transglycosylation product of mutant enzymes with substrates other than *p*NPαGal.

3.0 MATERIALS AND METHODS

3.1 MATERIALS

3.1.1 Bacterial strain

E. coli XL1 blue MRF' : Δ(*mcrA*)183 Δ(*mcrCB-hsdSMR-mrr*) 173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F' *proAB lacI*^{q} *ZΔM15* Tn10 (Tet^{r})]

This bacterial strain was used as host for cloning and expressing the wild type and mutant *agaB* genes.

3.1.2 Plasmid vectors

- **pBTac2 (Boehringer) :** An expression vector of 4.6 kb carrying the Tac promoter followed by a multiple cloning site comprising *Eco*RI and *Pst*I restriction sites as well as ampicillin resistance gene.
- **pAMG22:** This plasmid (Figure 1) was constructed by ligating *agaB* gene between the *Eco*RI and *Pst*I restriction sites (Ganter et al., 1988). It was prepared by Ralf Mattes' research team at the university of Stuttgart (Germany).

The vector was used as the template to prepare *agaB* mutant genes.

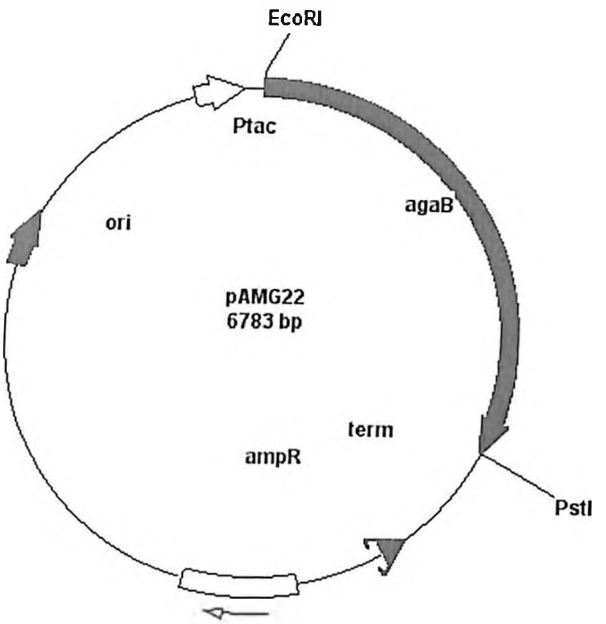


Figure 1 : The plasmid pAMG22

- **pUC18 :** A cloning vector carrying an ampicillin resistance gene, Plac promoter and *lac Z'* gene. It was used to determine the competence of *E. coli* cells.

3.1.3 Oligonucleotide primers

The following primers were used to amplify the genes cloned or the genes which were to be cloned in the pBTac2 vector:

Primer D: 5'-CAATTAATCATCGGCTCG-3'

Primer F : 5'-GGCGGATGAGAGAAGATT-3'

The primer D anneal upstream to the multiple cloning site of the pBTac2, while primer

F anneal downstream to the *Pst*I restriction site. This enables generation of PCR products that can be easily digested by *Eco*RI and *Pst*I at both extremities of the gene. *Eco*RI removes a 75 bp DNA fragment while *Pst*I cleaves out DNA of 60 bp.

Primer G442N$_1$:

5'-GCCGCGTTCGGAA (AGCT) (AGCT) (AGCT) CGAAACCAGCTTG-3'

Primer G442N$_2$:

5'-CAAGCTGGTTTCG (AGCT) (AGCT) (AGCT) TTCCGAACGCGGGC-3'

These primers were used for saturation mutagenesis at the 442 codon position of *AgaB*.

Primer R443N :

5'-CCGCGTTCGGAAGGA (AGCT) (AGCT) (AGCT) AACCAGCTTGTGC-3'

Primer for saturation mutagenesis at the 443 codon position of *agaB*.

Primer AHISP :

5'-ATATAT**CTGCAG**TTAGTGGTGGTGGTGGTGGTGTTGTTGAACAGCTTTCAA TCGCC-3'

This primer was used to prepare native and mutant agaB enzymes with a 6X histidine tail. The *Pst*I restriction site is underlined.

All the primers were synthesised by MWG Biotech.

3.1.4 Enzymes for DNA manipulation

The restriction enzymes and enzymes of DNA manipulation : *Eco*RI, *Pst*I, DyNazyme, *Pfu* polymerase and DNA ligase were supplied by Boerhinger Manheim, Stratagene, Promega, Finnzymes or TaKaRa.

3.2 METHODS

3.2.1 Isolation of plasmid DNA

For routine plasmid preparation, the alkaline lysis method (Sambrook *et al.*, 1989) was used. 5 ml of Luria Bertani medium (LB liquid medium : Tryptone 10 g/l, Yeast extract 5 g/l, NaCl 10 g/l, pH 7) supplemented with ampicillin (100 µg/ml) and tetracycline (20 µg/ml) was inoculated with a single bacterial colony, and incubated at 37°C overnight with vigorous shaking. 1.5 ml of the culture was centrifuged for 1 min. The bacterial pellet was resuspended by vortexing in 100 µl of solution I (50 mM Glucose, 10 mM EDTA, 25 mM Tris-HCl, pH 8), incubated for 5 min at room temperature then lysed by addition of 200 µl of solution II (NaOH 0.2 M, SDS 1%) and incubated for 5 min on ice. 150 µl of an ice-cold solution of solution III (sodium acetate 3 M, pH 5.2) was added, then vortexed and incubated on ice for 5 min. After centrifugation at 13000 rpm, 4°C, for 10 min, plasmid DNA was extracted from the supernatant by adding an equal volume of phenol/chloroform /isoamyl alcohol (50/48/2) and mixed by vortexing and followed by centrifugation for 2 min in a microfuge. Two volumes of cold absolute ethanol was added to the supernatant and centrifuged for 10 min at 4°C. 1 ml of 70% ethanol was added to the pellet and recentrifuged. The pellet was then dried in a vacuum dessicator and resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). Residual RNA was eliminated by adding RNAseA (200 µg/ml) and incubating for 10 min at 37°C.

Plasmids for sequencing and for cloning were isolated using the Concert™ Rapid Plasmid Purification System (Life Technologies). According to this protocol the cell lysate was applied to a spin cartridge containing silica-based membranes on which the plasmid DNA

is selectively adsorbed. The plasmid DNA was eluted at high concentration in TE buffer preheated at 65°C.

3.2.2 Mutagenesis of *agaB* gene

3.2.2.1 Introduction of a point mutation at codon 442 of *agaB* by overlapping extension.

Synthetic oligonucleotide primers designed to incorporate a point mutation at one end were used to generate two fragments of the *agaB* gene which mutually anneal around the position 442. A subsequent extension PCR was then used to amplify a full length *agaB*, mutated at codon 442. The primers were designed so as to specify all the 20 amino acids at this position (Figure 1, under Results)

For the first round of PCR amplification, 100 ng of each of the primers, D and G442N2 were used to generate the fragment A of 1300 bp (A1300); while F and G442N1 were used to generate the fragment B of 860 bp (B860). The primers were mixed with 50 ng of pAMG22 in a 50 μl PCR. The reaction conditions were : 1X *Pfu* buffer, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 2.5 U of *Pfu* DNA polymerase. The reaction was thermocycled for 5 min at 94°C , then for 30 cycles at 94°C, 1 min; 50°C, 1 min; 72°C, 3 min and a final extension step of 10 min at 72°C.

For overlapping extension 100 ng of each primer (D and F) were mixed with about 100 ng of A1300 and B860 in a 50 μl PCR. The reaction conditions were : 1X *DyNazyme* buffer, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 2.5 U of *DyNazyme* DNA polymerase. The reaction was thermocycled for 5 min at 94°C , then for 30 cycles at 94°C, 1 min; 50°C, 1 min; 72°C, 2.5 min and a final extension step of 10 min at 72°C.

3.2.2.2 Oligonucleotide-Directed mutagenesis at the 443 codon position of *agaB* using sequential PCR

A synthetic oligonucleotide primer containing an altered sequence at codon 443 of agaB was used. In the first PCR amplification the primers, R443N carrying the mutations corresponding to the codon to be changed, and primer F were employed to generate the mutated 3' coding region of *agaB* gene, 860 bp long (C860). The primer R443N was designed so as to specify all the 20 amino acids at this codon. In the second PCR amplification, the C860 fragment was used as primer together with primer D to create full length *agaB* gene from pAMG22 which was employed as template in both amplifications.

For the first PCR amplification 100 ng of each primer, R443N and F were mixed with 50 ng of pAMG22 in a 50 μl PCR. The reaction conditions were : 1X *Pfu* buffer, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 2.5 U of *Pfu* DNA polymerase. The reaction was thermocycled for 5 min at 94°C , then for 25 cycles at 94°C, 1 min; 50°C, 30 sec; 72°C, 2.5 min and a final extension step of 5 min at 72°C.

For the second PCR reaction 100 ng of primer D and about the same quantity of the C860 DNA fragment were mixed with 50 ng of pAMG22 in a 50 μL PCR. The reaction conditions were : 1X *DyNazyme* buffer, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 2.5 U of *DyNazyme* DNA polymerase. The reaction was thermocycled for 5 min at 94°C , then for 30 cycles at 94°C, 1 min; 50°C, 1 min; 72°C, 2.5 min and a final extension step of 10 min at 72°C.

3.2.2.3 Construction of his-tagged AgaB enzymes

To introduce a carboxy terminal hexahistidine tag (his-tag), a primer designed to carry 6 contiguous histidine codons, AHISP, was designed.

6 PCR reactions were set up with the different plasmids: pAMG22, pAMGL3,

pAMGE500, pAMGF1, pAMGF24 and pAMGF47 serving as templates to generate his-tagged enzymes : AgaB, L3, E500, F1, F24 and F47 respectively (Figure 10, Results).
Each PCR reaction mixture was composed of : about 50 ng of template DNA (i.e. the corresponding plasmid); 100 ng of each primer, D and AHISP; 1X *DyNazyme* buffer; 0.2 mM of each dNTP; 1.5 mM $MgCl_2$; 2.5 U of *DyNazyme* DNA polymerase. The reaction was thermocycled for 5 min at 94°C , then for 30 cycles at 94°C, 1 min; 50°C, 1 min; 72°C, 2.5 min and a final extension step of 10 min at 72°C.

3.2.3 Cloning techniques

3.2.3.1 Digestion of DNA with Restriction Endonucleases

The cleavage of mutant *agaB* genes and pBTac2 vector was accomplished by incubating the DNA with *Eco*RI and *Pst*I at 37°C for 90 min. 2 U of each restriction enzyme was added per 5 µg DNA in a reaction volume of 20 µl. The restriction buffer comprised 100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9.

The restriction enzymes were removed by phenol/chloroform extraction : an equal volume of TE-saturated phenol/chloroform/isoamyl alcohol (25/24/1) was added to the DNA sample and vigorously vortexed for 10 seconds. The sample was centrifuged for 5 min (10 000 *g*, 4°C) to separate the aqueous and organic phases. The aqueous phase, containing the DNA, was carefully removed into a new tube. An equal volume of chloroform was added, the sample vortexed and centrifuged (10 000 *g*, 4°C). The supernatant was subjected to ethanol precipitation : 1/10 volume of 3 M sodium acetate, pH 5.2 was added and inverted to mix. Two volumes of 100% ethanol was added and mixed. DNA was precipitated by placing the sample at -70°C for 30 minutes. The sample was centrifuged at 10 000 *g* for 10 min, at 4°C. The supernatant was discarded, the pellet washed with 1 ml ethanol (70%) and dried in a vacuum desiccator (Speed vac). Finally, the DNA was resuspended in TE buffer, pH 8.0.

3.2.3.2 Agarose gel electrophoresis

The separation, identification and quantification of DNA fragments were routinely done by performing agarose (0.8%) gel electrophoresis in TAE buffer (40 mM Tris-acetate, 2 mM EDTA, pH 8.0). DNA suspended in TE buffer was mixed with loading buffer (0.25% v/v Bromophenol blue, 0.25% v/v Xylene cyanol, 30% v/v Glycerol) before being deposited onto the gel. To allow visualisation under ultraviolet rays, ethidium bromide (0.5 μg/ml) was added to the gel. Molecular weight markers (Smartladder, Eurogentec) were used to estimate DNA size. The migration was carried out in Mupid electrophoresis apparatus (Eurogentec) at 100 V for 30 min.

3.2.3.3 Ligation Reactions

pBTac2 vector digested with *EcoR*I and *Pst*I was mixed with the *agaB* genes obtained by PCR and also digested with the same restriction endonucleases, in a total volume of 10 μl. One volume of DNA ligase (Solution I, TaKaRa) was added, thoroughly mixed then incubated at 16°C for 30 min. This ligation reaction mixture was used directly for bacterial transformation.

3.2.3.4 Preparation and transformation of competent *E.coli* cells

Competent cells were prepared and transformed based on the Simple and Efficient Method (SEM) (Hiroaki *et al.*, 1990). 12 large colonies of XL1 blue *E. coli* cells were isolated and inoculated to 250 ml of SOB medium (2% bactotryptone, 0.5% yeast extract, 2.5 mM KCl and 20 mM $MgCl_2$, added after the rest of the medium has been sterilised by autoclaving; pH 6.8) in a 1 litre flask, and grown to an A_{600} of 0.6 at 18°C, with vigorous shaking. Due to the low temperature this takes about 48 hr. The flask was then placed on ice for 10 min. The

culture was spun at 2500 X g for 10 min at 4°C. The pellet was resuspended in cold transformation buffer TB (10 mM Pipes, 15 mM $CaCl_2$, 250 mM KCl, 55 mM Mn^{2+}, pH 6.7) and incubated on ice for 10 min and recentrifuged. The pellet was resuspended in 20 ml TB, and DMSO was added to a final concentration of 7%. The cell suspension was further incubated on ice then divided in 500 µl aliquots for storage at -80°C after chilling in liquid nitrogen.

3.2.3.5 Transformation of competent cells

For transformation 10-20 µl of the ligation mixture was added to 200µl of the freshly thawed competent cells. After incubation on ice for 30 min, the cell suspension was heat shocked for 45 seconds at 42°C and transferred to an ice bath for 2 min. 1 ml of SOC (SOB containing 20 mM glucose) was added and the tubes were shaken vigorously at 37°C for 1 hour to enable expression of antibiotic resistance genes. The cells were spread on LB solid agar (LB liquid medium solidified by adding 1.5% bacteriological agar) supplemented with ampicillin (100 µg/ml) and tetracycline (20 µg/ml). Colonies were counted after an overnight incubation at 37°C.

To determine the competence level, 200 µl of the cells was also transformed with 0.1 ng of pUC18. This test gave typically 1000 colonies, therefore a competence of 10^7 transformants/µg DNA.

3.3 Isolation of active mutants

After an overnight culture of the transformed XL1 blue cells on LB medium, the colonies were picked up and spotted with sterile toothpicks onto two LB agar plates containing 20 µg/ml of tetracycline and 100 µg/ml of ampicillin. One plate was stored and the

other incubated for 2 hr at 55°C to denature the thermolabile MelA α-galactosidase of *E. coli.* The clones retaining a thermostable α-galactosidase activity were identified by pouring an X-α-Gal solution on the plate and the blue colonies were further characterised. X-α-Gal (5-bromo-4-chloro-3-indolyl α-D-galactopyranoside) solution is made by first preparing a 20 mg/ml solution in DMF then diluting to 0.75 mg/ml in sodium phosphate buffer (100 mM pH 7).

3.4.1 Determination of α-galactosidase activity of the mutant enzymes

For NMR analysis, the enzyme α-galactosidase activity was first determined to ensure equal enzyme units were used for each of the reactions.

The activity was determined on enzyme extracts prepared as follows: 50 ml of overnight culture was centrifuged, resuspended in 500 µl of sodium phosphate buffer 100 mM, pH 7.0 and sonicated briefly to lyse the cells. The cell lysate was centrifuged for 5 min at 13000 rpm at 4°C and enzyme activity was assayed in the supernatant. The endogenous α- galactosidase, Mel A, of *E.coli* is no longer active in the cellular extracts since it requires NAD+ and manganese ions as cofactors (Burstein and Kepes, 1971). Enzyme assays were performed by addition of 2 µl of the enzyme extract in 500 µl of a 10 mM *p*NPαGal solution (pH 7.0) at 25°C. The total activity was determined by following change in absorbance at 405 nm (ε_{pNP} = 14860 $l.mol^{-1}.cm^{-1}$).

3.4.2 Screening for transglycosylation activity by means of TLC

For the analysis of transglycosylation activity of the α-galactosidase by means of

TLC, enzyme extracts were prepared as follows: After an overnight growth, bacterial colonies were picked directly from an LB agar plate, and suspended in 100µl of 100 mM sodium phosphate buffer, pH 7. Brief sonication was done to lyse the cells and cell lysate was centrifuged for 1 min at 10 000 *g* at 4°C. The enzyme transglycosylation activity was then determined in the supernatant. 3 µl of the enzyme extract was added to 25µl of the substrate, 90mM *p*NPαGal or 90mM *o*NPαGal dissolved in sodium phosphate buffer (100 mM, pH 7.0). The reaction was allowed to proceed at room temperature for 16 hr. 0.5µl of the reaction mixture was analysed by thin layer chromatography (TLC) on precoated silica gel 60 sheets Merck F254. The mobile phase used was the Seymour eluent (ethanol/Chloroform/Acetic acid/Water, 60/30/3/5). After migration of the compounds, the TLC plates were dried then dipped in the revelation solution made up of 3.48% (w/v) ammonium molybdate $[(NH4)_6Mo_7O_{24}]$, 1.66% (w/v) Cerium sulphate and 19% (v/v) concentrated sulphuric acid. Finally the plates were dried and heated to reveal the spots corresponding to the different glycosides.

3.5 Protein analysis by SDS-PAGE

Polyacrylamide gel electrophoresis under denaturing conditions (SDS PAGE) was performed, according to the method of Laemmli (1970). The gels consisted of 3% stacking and 10% separating gels. The protein samples were denatured by heating for 5 minutes at 95°C in sample buffer (Sample buffer 4X: 62.5 mM Tris-HCl, 20% glycerol, 2% SDS, 5% β-mercaptoethanol, 0.025% w/v bromophenol blue, pH 6.8) before loading and running at constant current of 15 mA (Mini-PROTEAN II apparatus, Bio Rad). The running buffer was Tris-glycine (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3). Molecular weight markers phosphorylase b (97.4 kD), serum albumin (66.2 kD), ovalbumin (45 kD), carbonic

anhydrase (31 kD), trypsin inhibitor (21.5 kD) and lysozyme (14.4 kD) from Bio-Rad were run along with the samples for estimation of molecular weights. Protein bands were stained with coomassie brilliant blue solution (0.1% Coomassie brilliant blue, 40% methanol, 10% acetic acid) and destained by successive washing in destaining solution (40% methanol, 10% acetic acid).

3.6 Purification of the his-tagged AgaB enzymes

The *E. coli* strain containing the appropriate plasmid was grown overnight in 50 ml LB medium supplemented with ampicillin and tetracycline. The cells were pelleted by centrifugation (5000 rpm, 10 min, 4°C), and resuspended in lysis buffer (50 mM NaH_2PO4 pH 8, 300 mM NaCl, 20 mM imidazole). The cells were lysed by adding lysozyme (1 mg/ml). DNase (5μg/ml) was added to reduce lysate viscosity, incubated at 37°C for 30 min and sonicated for a further 30 min. The lysate was then cleared by centrifugation (14 000 rpm, 15 min, 4°C).

The his-tagged enzymes were purified by immobilized metal affinity chromatography (IMAC). This technique (also called metal-chelate affinity chromatography or MCAC) is based on the ability of histidine to bind reversibly to transitional metal ions that have been immobilized by a chelating group covalently bound to a solid support (Porath *et al.*, 1975; Hutchens *et al.*, 1990). Nitriloacetic acid (NTA, Qiagen) has a quadridentate moiety covalently coupled via a spacer arm to agarose forming a stable complex with Ni^{2+} . After binding to NTA, the Ni^{2+} ion has two sites within its octahedral coordination sphere available for binding to electron donor groups. The affinity of histidine residues for imobilized Ni^{2+} ions is significantly enhanced by designing proteins to contain short stretches of histidines. This his-tag enables the proteins to bind to the Ni^{2+} -NTA complex with a K_d of 10^{-13} M at pH 8 (Hoffman *et al.*, 1991; Hochuli *et al.*, 1988).

To purify the recombinant protein, 1 ml of 50% Ni- NTA slurry was added to 4 ml cleared lysate and gently mixed by shaking at 4°C for 60 minutes. The lysate- Ni- NTA mixture was loaded into a column with the bottom outlet capped. The bottom cap was removed and the column flow-through discarded. To eliminate proteins with non specific interactions the column was washed with 8 ml wash buffer (50 mM NaH_2PO_4,, 300 mM NaCl, 20 mM imidazole, pH 8.0). The protein was finally eluted with 2 ml elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0) in four 0.5 ml fractions.

3.7 Determination of protein concentration

Protein concentration was determined by the method of Smith *et al.*, (1985) using bicinchoninic acid (BCA). This assay is based on the interaction between bicinchoninic acid, copper ions, and proteins. Peptide bonds react with Cu^{2+} salts resulting in the formation of a Cu^{+} complex. Two molecules of bicinchoninic acid interact specifically with Cu^{+} to produce a purple coloration whose intensity is proportional to the protein concentration in the sample. The optical density was measured at 562 nm (Perkin Elmer lambda 20 U.V- Visible spectrophotometer). Bovine serum albumin (BSA) was used as protein standard to estimate protein concentration.

3.8 Kinetic studies of hydrolysis reactions

Kinetic experiments to determine K_m and k_{cat} with pNPαGal as substrate were performed in microtitration plates by following changes in absorbance using a microplate reader (iEMS, Labsystem), or in a Kontron Uvikon 860 spectrophotometer equipped with a cell holder connected to a circulating water bath at 25°C. The rate of release of *p*-nitrophenol was determined in a continuous assay. The buffer employed for all kinetic experiments with

wild type and mutant α-galactosidase AgaB was 100 mM sodium phosphate buffer pH 7.0. The reaction mixtures containing substrate and buffer (100 μl), were preincubated at 25°C prior to addition of 100 μl of diluted enzymes (volumes of 200μl was used in assays performed with Kontron spectrophotometer). The substrate was used in concentrations ranging from 0.15 K_m to 5 K_m for each enzymes.

The parameters K_m and k_{cat} were calculated. The K_m value gives the substrate concentration at which the initial rate of the reaction is one-half of the maximal rate and is thus a measure of the enzyme affinity. k_{cat} is the catalytic constant, calculated as:

k_{cat} = *Vmax* / Enzyme concentration (expressed in molar concentration of catalytic sites);

Vmax is the maximal velocity.

k_{cat} therefore measures the turnover of each active site per unit time. The parameter, k_{cat}/K_m, indicative of the catalytic specificity of the enzymes was also calculated.

3.9 Kinetic studies of transglycosylation reactions by means of *in situ* proton NMR spectroscopy

The transglycosylation reactions were carried out *in situ* (directly in NMR tubes). Heavy water ($^2H_2{}^{16}O$ or D_2O) was used in place of H_2O since water (H_2O) gives a strong peak in the same region as anomeric protons and so would obscure important signals. D_2O also prevents saturation of the receptor site by the protons of H_2O.

Kinetic studies were performed using a 500 MHz Bruker spectrometer according to the method of Spangenberg *et al.*, (1999). In an NMR tube, 90 mM of *p*NPαGal and the amount of salts necessary for 0.78 mL of 100 mM (pH 7.0) sodium phosphate buffer were lyophilised, dissolved in D_2O and lyophilised once more to eliminate the residual H_2O. The buffer reference mixture containing the silylated reference [0.78 mL of a 10 mmol/L of 3-

(trimethylsilyl)-propanesulfonic acid sodium salt solution] was prepared as above and adjusted to pD=7.0. The α-galactosidase preparation (0.48 units) was dried in a desiccator containing P_2O_5 and redissolved in 0.18 mL of the phosphate buffer reference mixture. The solutions were warmed to 25°C and the enzymatic preparation added to the reactant solution. The resulting mixture was immediately filtered in an NMR tube and the reaction allowed to proceed in the magnet of the spectrometer pre-adjusted to the temperature of the reaction.

4.0 RESULTS

4.1 Saturation mutagenesis

Saturation mutagenesis was the strategy chosen to examine the role of codon 442 in AgaB regioselectivity. This process consists of replacing the amino acid residue at the site of interest with all the different 20 amino acids. To effect this we designed degenerate primers coding for all the amino acids at this position, and subsequently synthesised the corresponding mutant *agaB* gene using overlapping PCR as shown in Figure 2.

Initial PCR amplification to generate 1300 bp (A1300) and 860 bp (B860) DNA fragments

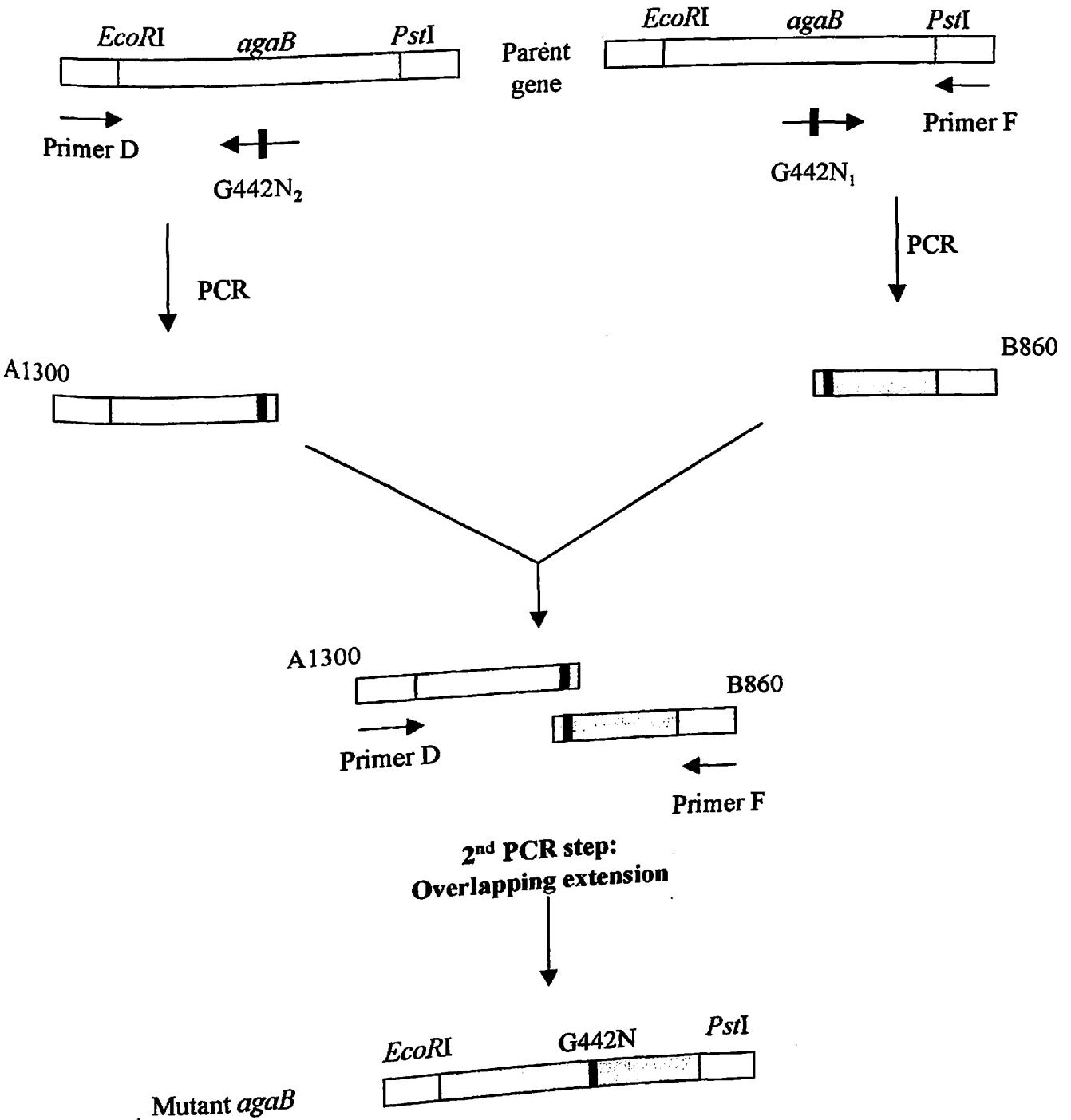


Figure 2 : Strategy of introducing G442N mutation by overlapping exension

The most critical parameter for this strategy to succeed is that the melting temperature (T_m) of the mutagenic oligonucleotides should be high enough to allow efficient annealing. 30-mers with about 56% GC content were therefore designed to ensure a high T_m and hence allowing base mismatch pairing at the point of mutation.

As the template for saturation mutagenesis, the plasmid pAMG22, which consists of the 2.2 kb wild type α-galactosidase *agaB* gene (Ganter *et al.*, 1988) from *Bacillus stearothermophilus* under the control of Ptac promoter in the pBtac2 vector, was used. The oligonucleotide primers which were used to introduce the G442N mutation are $G442N_1$ (30-mer) and $G442N_2$ (30-mer). Primer D (18-mer) and primer F (18-mer) enabled the amplification of the 1300 bp (5' coding region) and 860 bp (3' coding region) fragments respectively. To avoid introducing unintended mutations high fidelity DNA polymerases were used. *Pfu* DNA polymerase was used in the first amplification to generate the 1300 bp and the 860 bp DNA fragments because it does not add non-templated adenine residues at the 3' end of the DNA fragment. *DyNazyme* polymerase was employed in the subsequent overlapping extension reaction due to its robustness in amplification even of long DNA fragments.

The sizes of the various fragments obtained from the PCR reaction were checked on agarose gel (Figure 3). The DNA fragments obtained with *Pfu* DNA polymerase and *DyNazyme* polymerase were of the expected sizes.

The second amplification (overlapping PCR) also yielded a DNA fragment of the desired length (2.2 kb) which corresponds to the size of *agaB* gene (Figure 3).

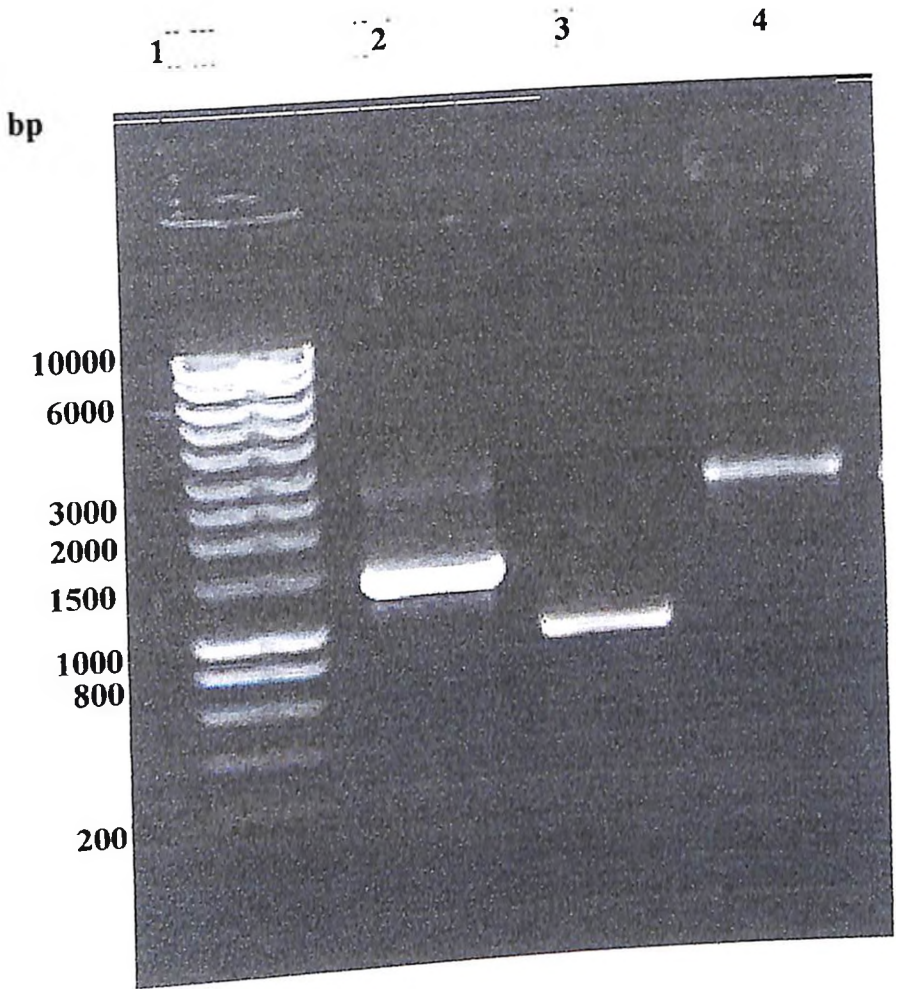


Figure 3 : Agarose gel electrophoresis of PCR amplification products.

Lane 1 : DNA size marker; Lane 2 : Amplification product of AgaB 5' coding region (1300 bp); Lane 3 : Amplification product of AgaB 3' coding region (860 bp);
Lane 4 : Overlapping extension product (2.2 kb).

The amplification conditions are described in Materials and Methods.

4.2 Cloning of the mutant genes

The plasmid pBTac2 was selected as the cloning vector. This plasmid possesses a Ptac promoter, a *lacZ* ribosome binding site (RBS) followed by the ATG initiation codon and strong transcription terminators. The RBS is followed by an *Eco*RI site which is upstream of the ATG initiation codon in the polylinker segment. A downstream *Pst*I site is also present in the multiple cloning site (Figure 1, Materials and Methods)

The *agaB* mutant genes and the pBTac2 vector were digested by *Eco*RI and *Pst*I restriction enzymes. A double digestion reaction was set up by incubating the DNA sample with the two enzymes in the *Pst*I buffer since *Pst*I is the more sensitive of the two enzymes to digestion conditions. The simultaneous digestion by the two enzymes avoided two digestion reactions which would require extra phenol/chloroform extraction and ethanol precipitation steps. This avoided loss of DNA during ethanol precipitation and limited exposure of the DNA to phenol which has deleterious effects.

The product of restriction digest was analysed on agarose gel. The size of DNA attained was of the correct size. *EcoR*I removes 75 bp oligonucleotide fragment while *Pst*I cleaves 60 bp from the 2.2 kb parent DNA thus generating two sticky ends.

Following digestion with the restriction endonucleases, mutated *agaB* genes were cloned back into the pBTac2 vector, also digested by the same enzymes (Figure 4). 25 ng of the vector and about 250 ng of the DNA to be inserted were used, giving a vector/insert molar ratio of about 0.2. This optimised recombinant plasmid production and reduced the chances of the vector closing back on itself.

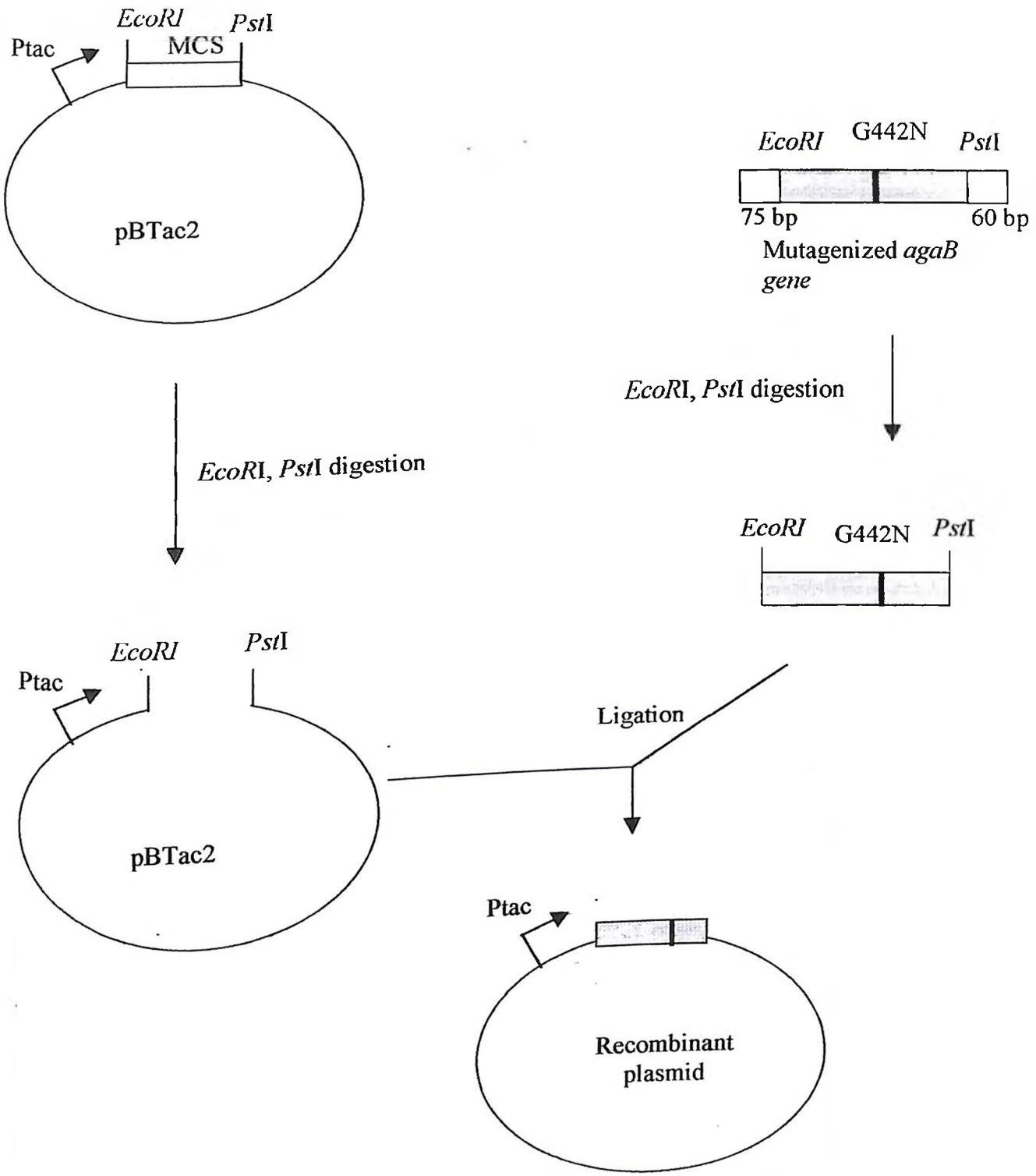


Figure 4 : Cloning of mutagenized *agaB* genes in pBTac2 vector.

The efficiency of ligation was also determined on agarose gel. The resulting ligation mixture was used to transform competent XL1 blue *E. coli* cells which were prepared by the Simple and Efficient Method (SEM).

E. coli XL1 blue strain was chosen for transformation because of its high transformation efficiency. Since the XL1 blue strain is tetracycline resistant and pBTac2 vector carries ampicillin resistance gene, the transformants were selected on LB agar plates supplemented with the two antibiotics.

4.3 Screening of mutant library by *p*NPαGal self-condensation

About 5000 colonies were obtained after transformation of XL1 blue strain. Of these, 800 clones were spotted on two LB plates. One of the plates was tested with X-α-gal after preheating at 55°C for 2 hours to inactivate the thermolabile MelA enzyme of *E. coli*. Thermostable α-galactosidase of active mutants split X-α-gal giving rise to 354 blue colonies. Several factors explain the occurrence of white colonies : i) The pBTac2 vector may close back on itself. Since pBTac2 codes for ampicillin resistance, the cells transformed by it can grow on selective plates; ii) Some amino acid changes may result in inactivated mutant AgaB enzymes; and iii) Some mutations may result in highly thermolabile AgaB enzymes which are destroyed by the 2 hr exposure at 55°C.

Of the blue colonies, 120 were screened by analysing the transglycosylation using *p*NPαGal as substrate. This number of clones was selected for screening to ensure that all the codon combinations (64 combinations are possible) in the sample were screened. In such a reaction, *p*NPαGal acts as both donor and acceptor giving rise to self-condensation products.

The distinction on TLC of the products (cf. Literature review) of this reaction is based on the fact that the different regioisomers differ in their polarities and as such migrate at different rates depending on the partition coefficient between the mobile and stationary

phases. The relative front (Rf) is a convenient means of expressing these differences. The compounds with low polarity have high relative fronts. Table 2 summarises the Rf of the reactants and products of the reaction:

Reactant/Product	Relative Front (Rf)
Galα-(1→3) Gal*p*NP	0.35
Galα-(1→6)Gal*p*NP	0.26
Galα-(1→2)Gal*p*NP	0.30
*p*NPαGal	0.59
Galactose	0.14

Table 2: Relative fronts (Rf) of products and reactants from *p*NPαGal self-condensation

Migration conditions: Thin layer plates precoated with silica gel F254, Seymour mobile phase.

The results obtained on TLC for some of the more significant clones are shown on figure 5.

An analysis of the migration patterns on TLC plates revealed that 33.1% of the enzymes exhibited a transgalactosylation profile similar to the E500 mutant, that is synthesizing the α-(1→2) disaccharide as the major product. 43.1% of the enzymes showed AgaB (wild type) transgalactosylation profile : α-(1→6) regioisomer is the major product with a negligible amount if any of the α-(1→3) product. 23.8% of the enzymes screened showed an F1 mutant profile : they synthesise α-(1→6) and α-(1→3) disaccharides in almost equal amounts. Some of the enzymes displaying the AgaB profile (23.8% of total screened enzymes) synthesise the product in quantities less than those observed for AgaB.

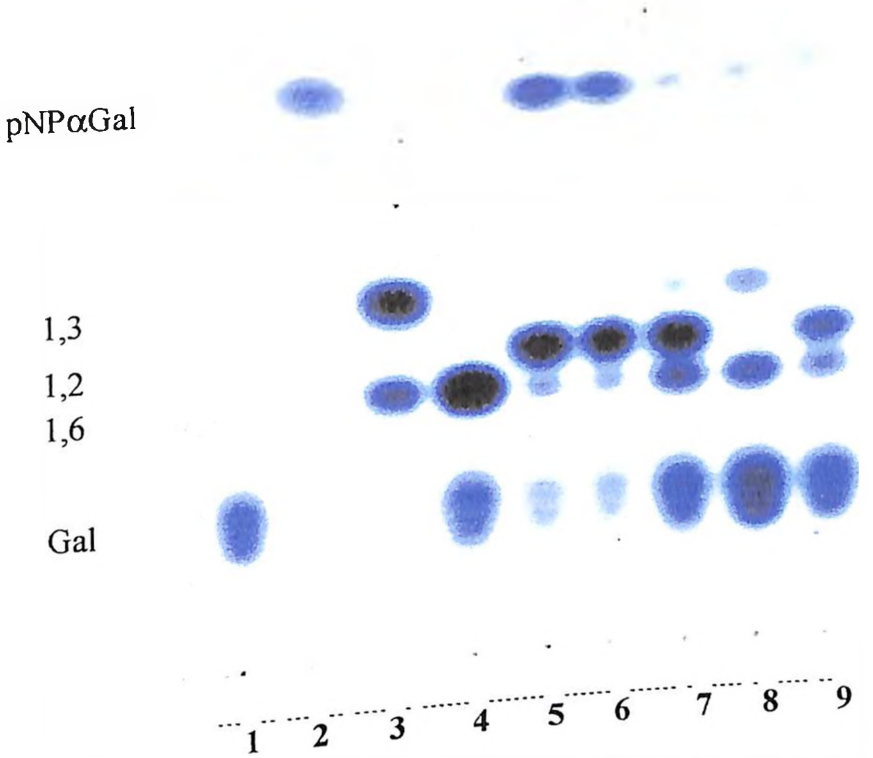


Figure 5 : TLC analysis of the transglycosylation products of G442N mutants.

1 µl of the reaction mixture is deposited on the TLC plate, dried and migrated in a preequilibrated TLC tank containing a Seymour mobile phase.

Key

*p*NPαGal : *p*-nitrophenyl α-D- galactopyranoside; 1,3 : Galα(1->3)Gal*p*NP; 1,2 : Galα(1->2)Gal*p*NP; 1,6 : Galα(1->6)Gal*p*NP; Gal : Galactose
Lane 1 : Galactose; Lane 2 : *p*NPαGal; Lane 3 : 1,3 and 1,6 mixture; Lane 4: AgaB reaction products; Lane 5 : E500 reaction products; Lane 6 : L3 reaction products; Lane 7 : F24 reaction products; Lane 8 : F1 reaction products; Lane 9 : F47 reaction products.

4.4 Sequencing of mutants

Representative clones which displayed striking transglycosylation profiles on TLC were selected for further analysis. It can be recalled that one of our objectives was to determine whether G442R mutation at this site was responsible for α(1→2) regioselectivity shown by E500: a hypothesis which was already supported by observing an E500 like activity in some of the clones in our library.

L3 and M38 were picked for sequence analysis on the basis of their ability to synthesise α(1→2) disaccharide almost at the exclusion of any other isomer. This activity was quite similar to what is observed with E500.

F24 and F47 were selected due to their relatively high activity while at the same time catalysing the synthesis of the α(1→2) disaccharide as the major product together with other regiomers in minor quantities. Therefore, there was a possibility that they could contain the G442R substitution. The relative activity of a mutant can be detected quite early during the screening process because colonies expressing highly active mutants turn blue first.

F1 was sequenced because of its remarkable ability to catalyse the synthesis of both the α(1→3) and α(1→6) regioisomers in similar quantities and at the exclusion of the α(1→2) regioisomer.

The *agaB* mutant genes of these clones were sequenced at the codon 442 level by MWG-Biotech. The plasmids purified for sequencing were analysed on agarose gel (Figure 6). The full sequence of AgaB gene is given in Figure 7 while the mutations found for the selected enzymes are summarized in Table 3.

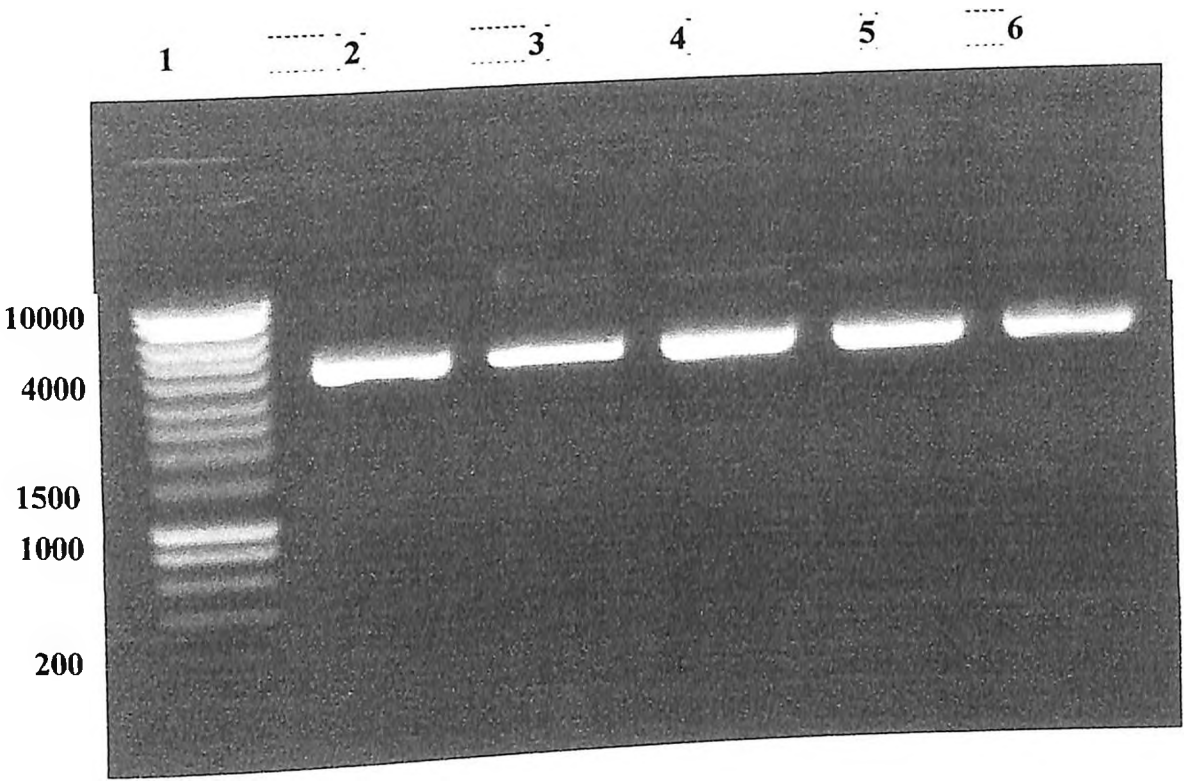


Figure 6 : Analysis on agarose gel of plasmids purified for sequencing.

2 µl of plasmid prepared by the Concert method (Life Technologies) was loaded on 0.8% agarose gel.

Lane 1 : DNA size marker; Lane 2 : pAMGF1; Lane 3: pAMGF24; Lane 4 : pAMGF47; Lane 5 : pAMGM38; Lane 6 : pAMGL3.

Figure 7 : Nucleotide and translated sequences of the *agaB* gene.

```
  M  A  V  T  Y  N  P  Q  T  K  Q  F  H  L  R  A  G  K  A  S  Y  V  M      23
ATG GCG GTT ACA TAC AAT CCG CAA ACG AAG CAG TTT CAT TTG CGG GCG GGC AAG GCA AGC TAC GTG ATG     69

  Q  L  F  R  S  G  Y  L  A  H  I  Y  W  G  K  A  V  R  D  V  R  G  E      46
CAG CTT TTC CGC TCC GGC TAT TTG GCT CAT ATC TAT TGG GGA AAA GCG GTG CGC GAT GTT CGG GGG GAG    138

  R  R  F  S  R  L  D  R  A  F  S  P  N  P  D  P  S  D  R  T  F  S  L      69
CGG AGG TTT TCG CGG CTG GAT CGC GCG TTT TCC CCC AAT CCC GAC CCG TCT GAC CGC ACG TTT TCG CTC    207

  D  T  L  P  Q  E  Y  P  A  Y  G  N  T  D  F  R  S  P  A  Y  Q  V  Q      92
GAT ACG CTG CCG CAA GAA TAT CCA GCC TAT GGG AAT ACC GAT TTT CGT TCT CCG GCG TAT CAA GTG CAG    276

  L  E  N  G  S  T  V  T  D  L  R  Y  K  T  H  R  I  Y  K  G  K  P  R     115
CTT GAG AAC GGC TCG ACC GTG ACC GAT TTG CGC TAC AAA ACG CAC CGC ATT TAC AAA GGG AAG CCT AGA    345

  L  N  G  L  P  A  T  Y  V  E  H  E  Q  E  A  E  T  L  E  I  V  L  G     138
CTT AAC GGA CTG CCG GCG ACA TAC GTG GAG CAT GAA CAG GAA GCG GAG ACG CTC GAA ATT GTC CTT GGA    414

  D  A  L  I  G  L  E  V  T  L  Q  Y  T  A  Y  E  K  W  N  V  I  T  R     161
GAT GCG CTG ATC GGT CTA GAG GTC ACG TTG CAG TAT ACG GCG TAT GAA AAA TGG AAC GTC ATC ACG CGC    483

  A  A  R  F  E  N  K  G  G  E  R  L  K  L  L  R  A  L  S  M  S  V  D     184
GCG GCC CGT TTT GAA AAC AAA GGC GGC GAG CGG TTG AAA CTG CTG CGT GCG CTC AGT ATG AGC GTT GAT    552

  F  P  T  A  D  Y  D  W  I  H  L  P  G  A  W  G  R  E  R  W  I  E  R     207
TTC CCA ACT GCT GAC TAT GAT TGG ATC CAT CTC CCC GGA GCG TGG GGG CGC GAG CGC TGG ATC GAG CGC    621

  R  P  L  V  T  G  V  Q  A  A  E  S  R  R  G  A  S  S  H  Q  Q  N  P     230
CGC CCG CTT GTG ACC GGC GTG CAA GCG GCG GAA AGC CGC CGC GGC GCG AGC AGC CAC CAA CAA AAC CCA    690

  F  I  A  L  V  A  K  N  A  D  E  H  Q  G  E  V  Y  G  F  S  F  V  Y     253
TTT ATC GCG CTC GTC GCC AAA AAC GCC GAT GAA CAC CAA GGC GAA GTA TAC GGG TTC AGC TTT GTG TAC    759

  S  G  N  F  L  A  Q  V  E  V  D  Q  F  H  T  A  R  V  S  M  G  I  N     276
AGC GGC AAT TTC CTC GCC CAA GTC GAG GTC GAC CAG TTT CAC ACC GCC CGT GTC TCG ATG GGA ATC AAC    828

  P  F  D  F  T  W  L  L  Q  P  G  E  S  F  Q  T  P  E  V  V  M  V  Y     299
CCG TTT GAT TTC ACA TGG CTG CTT CAG CCG GGT GAG TCG TTC CAA ACG CCG GAA GTG GTC ATG GTC TAC    897

  S  D  Q  G  L  N  G  M  S  Q  T  Y  H  E  L  Y  R  T  R  L  A  R  G     322
TCC GAC CAA GGG TTA AAT GGG ATG TCG CAA ACG TAC CAT GAA CTG TAC CGC ACC CGC CTG GCG CGC GGC    966

  A  F  R  D  R  E  R  P  I  L  I  N  N  W  E  A  T  Y  F  D  F  N  E     345
GCA TTC CGC GAC CGC GAA CGC CCG ATT TTG ATC AAC AAC TGG GAA GCA ACG TAC TTC GAT TTT AAC GAA   1035

  E  K  L  V  N  I  A  K  T  E  A  E  L  G  I  E  L  F  V  L  D  D  G     368
GAA AAA CTC GTC AAC ATT GCG AAA ACG GAA GCG GAA CTA GGC ATC GAA CTG TTT GTG CTT GAC GAT GGC   1104

  W  F  G  K  R  D  D  D  R  R  S  L  G  D  W  I  V  N  R  R  K  L  P     391
TGG TTT GGC AAG CGC GAT GAC GAC CGT CGT TCG CTC GGC GAT TGG ATC GTC AAC CGG CGC AAG CTT CCG   1173

  N  G  L  D  G  L  A  K  Q  V  N  E  L  G  M  Q  F  G  L  W  V  E  P     414
AAC GGC TTA GAC GGG TTG GCG AAA CAA GTG AAC GAA CTC GGG ATG CAG TTC GGC TTA TGG GTC GAA CCG   1242

  E  M  V  S  P  N  S  E  L  Y  R  K  H  P  D  W  C  L  H  V  P  N  R     437
GAA ATG GTG TCG CCA AAC AGC GAA CTG TAC CGG AAA CAC CCC GAC TGG TGT CTG CAT GTG CCC AAC CGC   1311

  P  R  S  E  G  R  N  Q  L  V  L  D  Y  S  R  E  D  V  C  D  Y  I  I     460
CCG CGT TCG GAA GGA CGA AAC CAG CTT GTG CTC GAT TAT TCC CGC GAA GAC GTT TGC GAC TAT ATC ATC   1380

  E  T  I  S  N  V  L  A  S  A  P  I  T  Y  V  K  W  D  M  N  R  H  M     483
GAG ACG ATC TCG AAC GTC CTC GCA AGC GCG CCG ATC ACG TAC GTG AAA TGG GAC ATG AAC CGC CAT ATG   1449

  T  E  I  G  S  S  A  L  P  P  E  R  Q  R  E  T  A  H  R  Y  M  L  G     506
ACG GAA ATC GGC TCC TCC GCC CTG CCG CCC GAG CGC CAG CGC GAA ACG GCG CAC CGC TAT ATG CTT GGG   1518

  L  Y  R  V  M  D  E  M  T  S  R  F  P  H  I  L  F  E  S  C  S  G  G     529
CTG TAC CGC GTC ATG GAC GAG ATG ACC TCG CGC TTT CCG CAC ATT TTG TTT GAA AGC TGT TCG GGA GGC   1587

  G  G  R  F  D  P  G  M  L  Y  Y  M  P  Q  T  W  T  S  D  N  T  D  A     552
GGG GGG CGG TTT GAC CCG GGG ATG CTG TAT TAT ATG CCG CAA ACG TGG ACG AGC GAC AAT ACC GAT GCC   1656

  V  S  R  L  K  I  Q  Y  G  T  S  L  V  Y  P  I  S  A  M  G  A  H  V     575
GTC TCG CGC CTG AAA ATT CAA TAC GGC ACG AGC CTC GTC TAT CCG ATT AGT GCG ATG GGC GCC CAC GTC   1725

  S  A  V  P  N  H  Q  V  G  R  V  A  S  L  K  A  R  G  H  V  A  M  S     598
TCG GCG GTG CCG AAC CAC CAA GTC GGG CGG GTG GCG TCG CTC AAG GCG CGC GGC CAT GTC GCG ATG TCA   1794

  G  N  F  G  Y  E  L  D  I  T  K  L  T  E  T  E  K  Q  M  I  K  Q  Q     621
GGC AAC TTC GGC TAT GAG CTC GAT ATC ACG AAA TTG ACG GAA ACA GAA AAA CAA ATG ATT AAG CAA CAA   1863

  V  A  F  Y  K  D  V  R  R  L  V  Q  F  G  T  F  Y  R  L  L  S  P  F     644
GTC GCG TTT TAC AAG GAC GTA CGC CGC CTC GTC CAG TTC GGC ACG TTT TAT CGA CTG CTA AGC CCG TTT   1932

  E  G  N  E  A  A  W  M  F  V  S  A  D  R  S  E  A  L  V  A  Y  F  R     667
GAA GGC AAC GAG GCG GCG TGG ATG TTC GTC TCT GCC GAC CGC TCG GAA GCG CTC GTC GCC TAC TTC CGC   2001

  V  L  A  E  A  N  A  P  L  S  Y  L  R  L  K  G  L  D  P  N  Q  D  Y     690
GTT CTG GCC GAA GCG AAC GCG CCG CTG TCA TAC CTG CGC TTA AAA GGG CTT GAC CCG AAT CAA GAC TAC   2070

  E  I  E  G  L  G  V  Y  G  G  D  E  L  M  Y  A  G  V  A  L  P  Y  R     713
GAA ATC GAG GGA CTT GGC GTT TAC GGC GGC GAC GAG CTG ATG TAT GCC GGA GTG GCC TTG CCG TAC CGC   2139

  S  G  D  F  I  S  M  M  W  R  L  K  A  V  Q  Q  *                     730
TCT GGC GAT TTT ATC AGC ATG ATG TGG CGA TTG AAA GCT GTT CAA CAA TAA       2190
```

Enzyme	Regioselectivity	Mutation
M38	1,2>1,6>1,3	G442R
L3	1,2>1,6>1,3	G442R
F24	1,2>1,6>1,3	G442I
F47	1,2>1,6	G442L
F1	1,3 = 1,6	G442Y

Table 3 : Amino acid substitutions at codon 442 in the selected AgaB mutants, and corresponding regioselectivities in self-condensation with *p*NPαGal.

From the sequencing results M38 and L3 were found to carry the same mutation , G442R. This result confirm the hypothesis that this mutation conferred the α(1→2) regioselectivity to the E500 mutant. F24 and F47 had the G442I and G442L substitutions respectively. Both isoleucine and leucine are hydrophobic amino acids and structural isomers. F1 had glycine being substituted for tyrosine, an aromatic amino acid. All these results show that position 442 is a critical site in determining AgaB regioselectivity. The structural diversity of the residues that can replace glycine at this site while still giving active enzymes is also remarkable.

TLC is a reliable method for identification of compounds when their migration is compared to those of known reference compounds chromatographed alongside the sample on the TLC plate. However, it was necessary to confirm the identity of the regioisomers and quantify the self condensation products. To attain this goal Nuclear Magnetic Resonance (NMR) spectroscopy was used (section 5.1).

5. Kinetic characterisation of mutants

5.1 Kinetics of transglycosylation by means of NMR spectroscopy

An accurate analysis method was needed to achieve two goals :

i) confirmation of the identity of the regioisomers

ii) following quantitatively and kinetically the transglycosylation and hydrolysis reactions.

High performance liquid chromatography (HPLC) and capillary electrophoresis could have been suitable alternative methods to attain these objectives given that all the products of the self condensation reaction, as depicted in scheme 4 (cf. Literature Review) can be reliably separated by these techniques. However the limitation of these techniques is that the detection systems that go with them (e.g. UV-visible detector) do not permit the characterisation of the glycosidic linkages of the oligosaccharides.

Proton NMR spectroscopy is an alternative method to follow the progress of the transglycosylation reaction. It is based upon the change of the spin state of a nuclear magnetic moment when the nucleus absorbs electromagnetic radiation in a strong magnetic field. The proton NMR instrument therefore measures the intensity of absorption of radiowaves when a magnetic field is applied. Different hydrogen atoms will give different signals (peaks) in the spectrum. The peak integral is dependent upon the number of hydrogen atoms present (hence concentration).

The advantage of this spectroscopic method is that it gives signals that are directly proportional to the concentration of the compound being analysed. Further it is non-invasive making it possible to study the transglycosylation reactions *in situ*. NMR studies require, however, use of heavy water (D_2O) in place of light water (H_2O) because the protons of H_2O resonate between 4.6 and 4.8 ppm, a spectral region very rich on information. Figure 8 shows

a proton NMR spectrum obtained during *p*NPαGal self condensation reaction catalysed by native AgaB. The chemical shifts of the protons are given in Table 4.

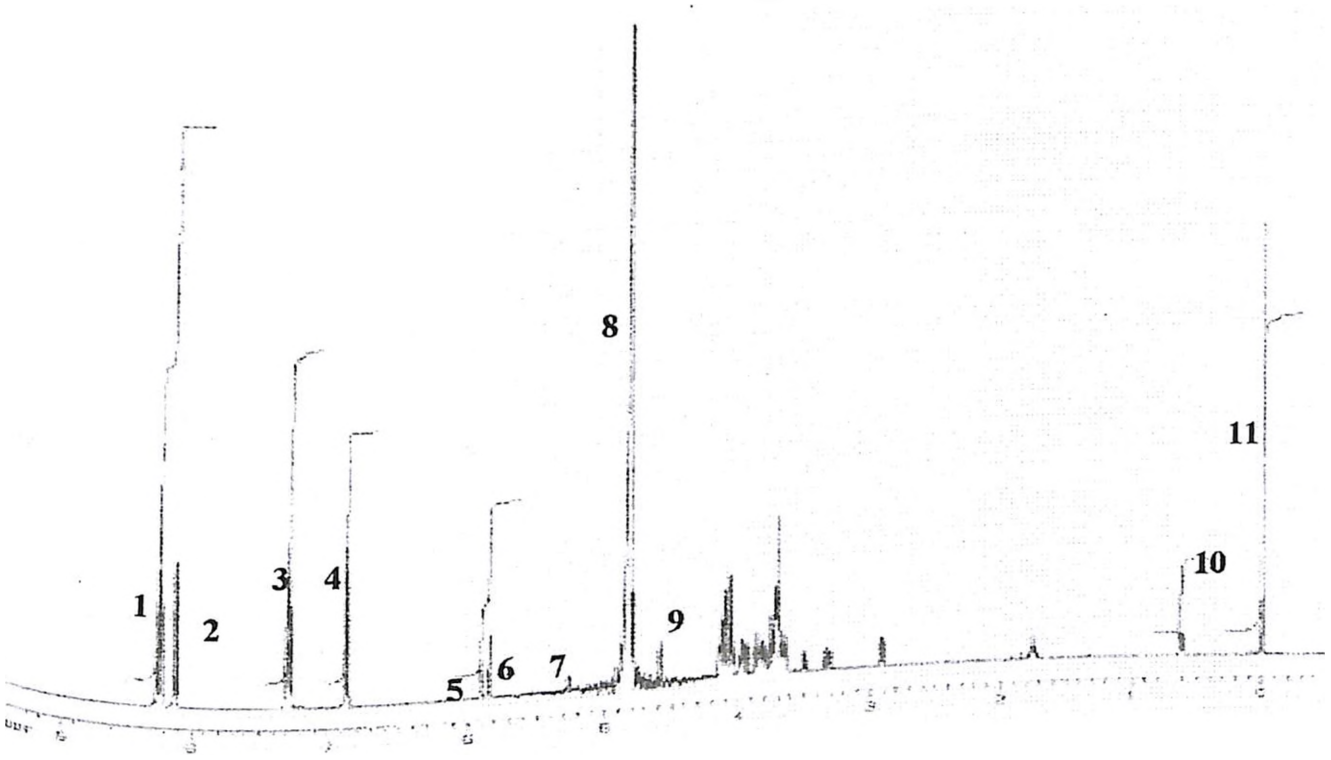


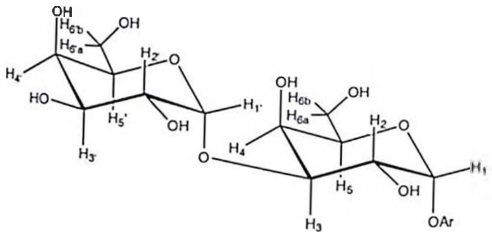
Figure 8 : ^{1}H NMR (500 MHz) spectrum of the reaction mixture during self-condensation of *p*NPαGal catalysed by AgaB.

The spectrum was registered *in situ* after 8 hr of incubation of 0.48 units of AgaB with 90 mM substrate solution. The reaction temperature was 25°C.

Key :

Peaks 1-4 are peaks of aromatic protons; peak 5 : 1,3 anomeric protons; peak 6: anomeric proton of *p*NPαGal; peak 7 : α anomeric proton of galactose; peak 8 : protons of water;

peak 9 : β anomeric proton of galactose; peak 10 and 11 are TMS peaks



Proton		Galα(1→2) GalO*p*NP	Galα(1→3) GalO*p*NP	Galα(1→6) GalO*p*NP
Glycoside acceptor	**1**	6.09	5.88	5.90
	2	4.14	4.20	4.04
	3	3.83	4.25	3.84
	4	3.96	4.32	4.13
	5	4.07	4.27	3.80
	6	3.73/3.73	3.73/3.73	4.11/4.09
Glycosyl donor	**1'**	5.13	5.24	4.80
	2'	3.77	3.91	3.29
	3'	4.19	4.00	3.64
	4'	3.75	4.06	3.68
	5'	3.98	3.97	3.61
	6'	3.70	3.74	3.79
Aromatic protons		6.8-8.5	6.8-8.5	6.8-8.5

Table 4 : Chemical shifts of ^{1}H of the disaccharide products(Adapted from Spangenberg, 2000). The various carbon positions are indicated by the numbers 1-6 and 1'-6'.

5.1.1 Analysis of NMR spectra

The concentration of the various components of the reaction mixture is given by integration of the corresponding peaks. The region of resonance of the anomeric protons that is between 4.3 and 5.9 ppm relative to the TMS peak (0.0), was used to follow the reaction. The molar percentage of each substituent represents the amount present at each time. This percentage is determined by calculating the ratio of a substituent's integration to the sum of all the substituents of the reaction mixture: residual donor, hydrolysis product and product of transglycosylation (i.e. the self condensation product).

The formula (Dion *et al.*, 2001b) used was therefore :

$$I_{\%} = \frac{I_x}{I_{Donor} + I_{Hydrolysis\ product} + \sum I_{Acceptor\ TG\ product} + 2 \times \sum I_{Self\ condensation\ product}} \times 100$$

Where:

$\%_x$: Molecular percentage of substituent x

I_x : Integration of substituent x protons

I_{Donor} : Integration of donor protons

$I_{Hydrolysis\ product}$: Integration of hydrolysis product protons

$I_{Acceptor\ TG\ product}$: Integration of transglycosylation product protons with an acceptor (for example if there is synthesis of a trisaccharide).

$I_{Self\ condensation\ product}$: Integration of self condensation product protons (This integration must be multiplied by 2 in the case of a self condensation reaction because each molecule of the product is made from 2 molecules of the donor).

The yield Y, of self condensation can also be expressed as the percentage of the donor which has reacted. When a disaccharide is the product of such a reaction , then Y = 2 I % or

$$Y_{\%} = \frac{2 \times I_x}{I_{Donor} + I_{Hydrolysis\ product} + \sum I_{Acceptor\ TG\ product} + 2 \times \sum I_{Self\ condensation\ product}} \times 100$$

Where:

$\%_x$: Molecular percentage of substituent x (that is self condensation product only)

I_x : Integration subsituent x

I_{Donor} : Integration of donor protons

$I_{Hydrolysis\ product}$: Integration of hydrolysis product protons

$I_{Acceptor\ TG\ product}$: Integration of transglycosylation product protons with an acceptor.

$I_{Self\ condensation\ product}$: Integration of self condensation product protons

Figure 9 shows the results obtained for each of the enzymes. For the spectral analysis; 0.48 units of each enzyme and 90 mM *p*NPαGal substrate in 100 mM sodium phosphate buffer, pH 7.0, were treated as described under Materials and Methods.

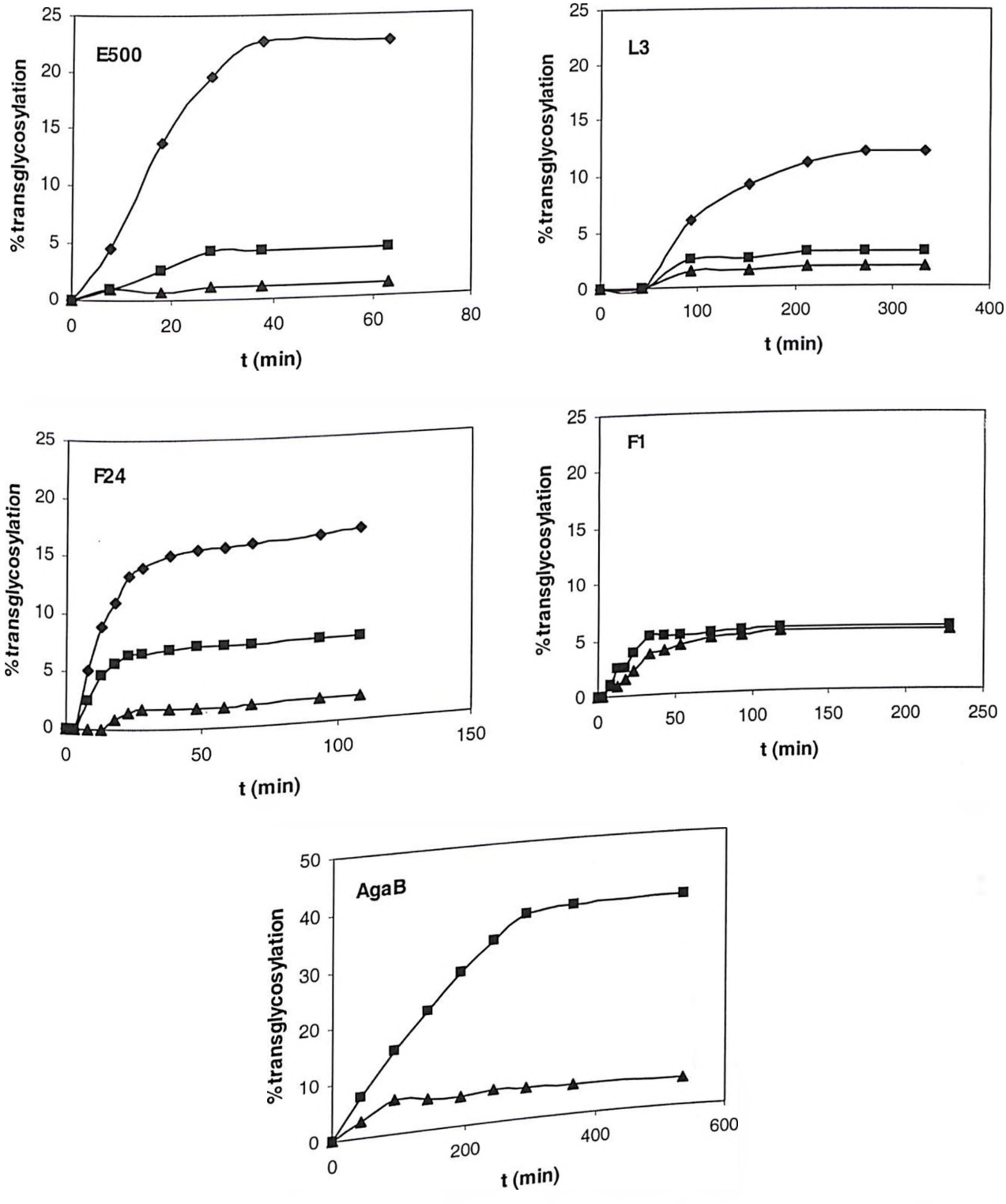


Figure 9 : Kinetics of synthesis of the self condensation products by mutant enzymes.

Kinetics of the synthesis of the self-condensation disaccharides: α-Gal-[1→6]-α-Gal-O-*p*-NP (■), α-Gal-[1→3]-α-Gal-O-*p*-NP (▲) and α-Gal-[1→2]-α-Gal-O-*p*-NP (◆); catalysed by AgaB and the α-galactosidase mutants E500, L3, F24 and F1 at 25°C. The molar percentages represent the proportions of *p*NPαGal which have formed a given disaccharide and thus are also the yields for the formation of this compound. For all kinetics, 0.48 units of enzyme was used in a solution of *p*NPαGal (90 mM).

Results similar to those already observed on TLC were obtained. AgaB shows a high capacity to synthesise the α(1→6) disaccharide (38 %) while the α(1→3) regioisomer is synthesised in negligible amounts (1 %). E500 synthesises α(1→2) product in quantities (25 %) much higher than α(1→6) and α(1→3) disaccharides (5 % and 1 % respectively). Similarly, L 3 synthesises α (1→2) regioisomer as the major product but with a lower transferase activity (12 %) and at a reaction rate lower than that observed for E500. F24 produces high levels of α(1→2) regioisomer (16 %) but it has a lower regioselectivity than E500 or L3. The lowest transglycosylation activity was shown by F1, which produced the α(1→6) and α(1→3) regioisomers at nearly equal amounts at 5 % each.

5.2 Construction of AgaB mutants tagged with an histidine tail

The determination of kinetic parameters of the mutant enzymes, especially k_{cat}, required their purification. To do this, a carboxy terminal hexahistidine tag was added to the enzymes : AgaB, E500, L3, F1, F24 and F47. Further analysis of the M38 was stopped because of its identity to L3. Figure 10 summarizes the method used to introduce the his-tag. Six histidine codons and a *Pst*I restriction site were integrated in the AHISP primer. Through two PCR reactions the additional sequence was added to the each of the genes encoding the various enzymes. The PCR products were cleaved at the *Eco*RI and *Pst*I sites as previously done for enzymes without his-tag (Figure 10).

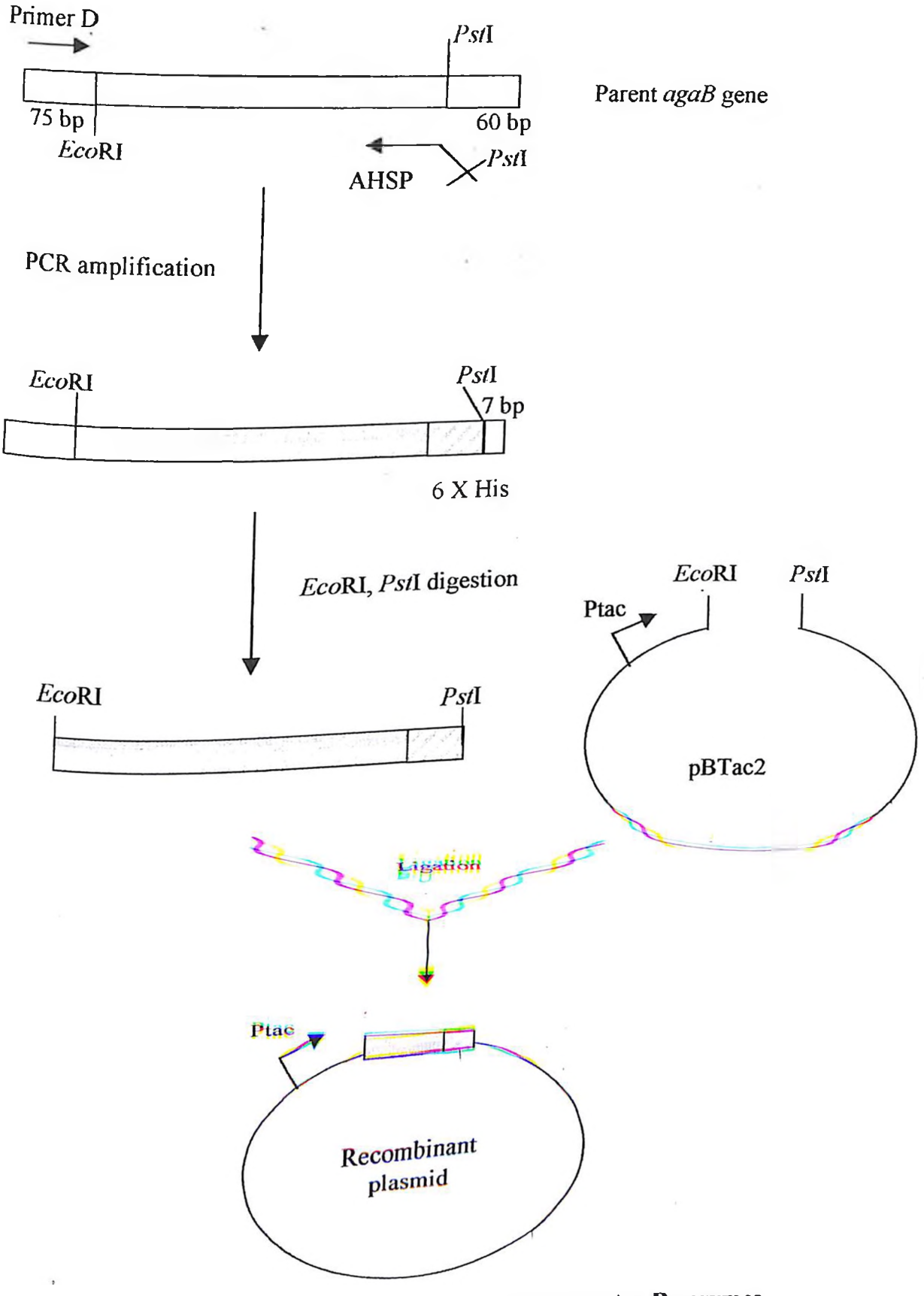


Figure 10 : Strategy of introducing his-tag in AgaB enzymes

After the amplification of the corresponding genes using the primers D and AHISP, a double digestion reaction with *Eco*RI and *Pst*I restriction enzymes, was set up. Initially an efficient cleavage could not be achieved. It was finally established that the *Pst*I site was not cut efficiently, probably due to its location near the end of DNA fragment. This restriction site was only 7 bp from the DNA terminal. In an attempt to improve the digestion, the sample was incubated for longer periods (3 hr 30 min).

The digested DNA was cloned in the pBtac2 vector and the constructs used to transform *E.coli* XL1 blue competent cells. About 5000 transformants were obtained for all the mutants. However the number of blue colonies after heating the colonies for 2 hr at 55°C and testing with X-α-gal were few. To determine if the blue colonies expressed his-tagged enzyme, small scale protein purification, was carried out to determine if the protein was being retained by the affinity resin. The clones having the correct plasmids were isolated and stored. Some of the blue colonies did not express his-tagged proteins. This can be explained by the original template used for the PCR reaction and which was still present in the ligation mixture being able to transform *E .coli* cells.

5.3 Purification of enzymes

In order to prepare a substantial amount of his-tagged enzymes, an overnight culture of 50 ml of each mutant was carried out in LB liquid medium. The cells were lysed as described in Materials and Methods and the his-tagged enzymes were purified on a Ni^{+2}- NTA column. A sample of the crude lysate and the corresponding purified enzyme was used for analysis on SDS- PAGE (Figure 8).

Using this method protein of quite high purity (about 95 %) was normally obtained in a single step purification process. Few contaminants appeared when higher amount of total protein was loaded. Typically protein yield of 6 mg was obtained in a volume of 2 ml. The expression level of AgaB enzymes can therefore be estimated to be at least 120 mg/l of culture.

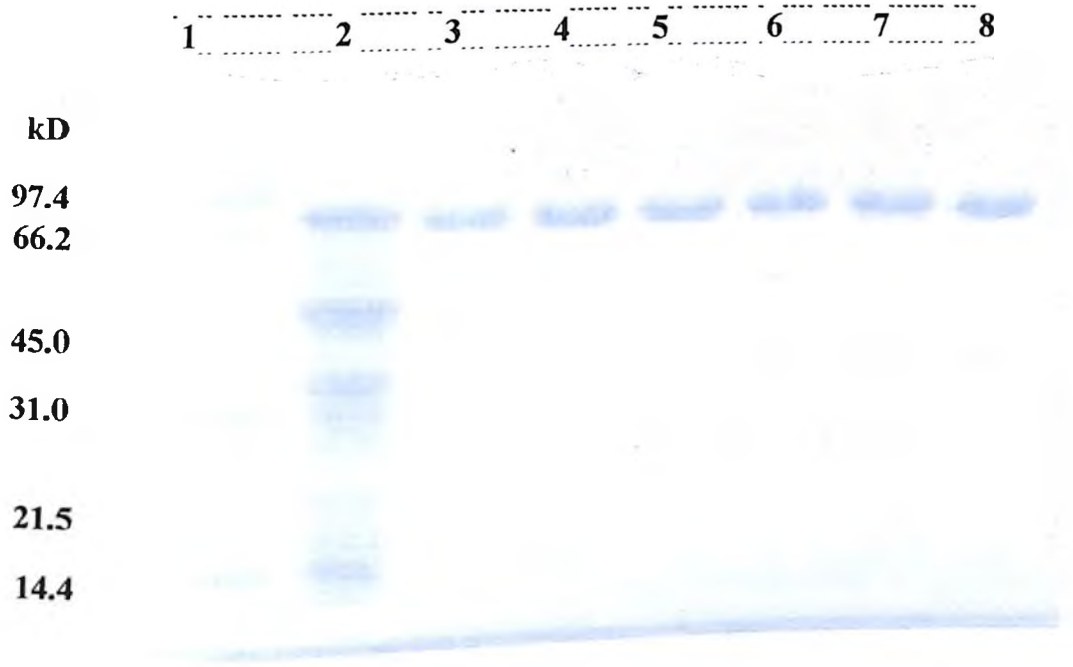


Figure 11 : SDS-PAGE analysis of purified proteins.

The purified protein samples were subjected to SDS-PAGE (10%) until the tracker dye reached the bottom of the gel. Bio Rad low range protein molecular weight standard was used.
Lanes : 1 : Molecular weight standard; 2 : Bacterial cell lysate; 3-8 : Purified enzymes : 3 : AgaB; 4 : E500; 5 : L3; 6 : F1; 7 : F24; 8 : F47.

5.4 Comparison of transglycosylation regioselectivity of his-tagged and non his-tagged enzymes

*p*NPαGal self-condensation products in the reactions catalysed by the enzymes with his-tag and those without were analysed. 3 μl enzyme preparation was added to 25 μl of 90 mM *p*NPαGal substrate solution, and 0.5 μl of the reaction mixture charged on TLC plates before migrating in Seymour reagent.

The tagged enzymes were found to display similar regioselectivity pattern to the non tagged enzymes. However, for E500, there was an increase in the amount of the α(1→6) regioisomer synthesised (figure 12).

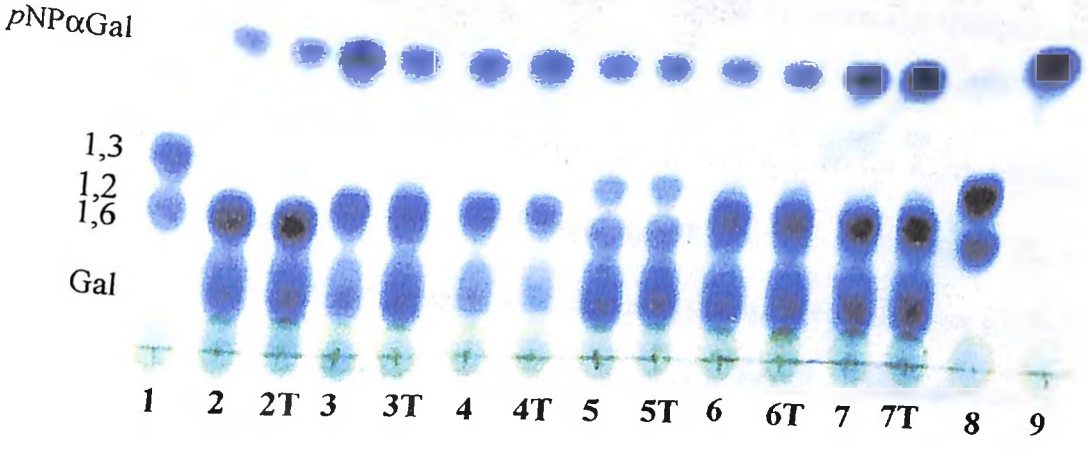


Figure 12 : Comparison on TLC of the *p*NPαGal self condensation products of his-tagged and non his-tagged enzymes.

1 µl of the reaction mixture is deposited on the TLC plate, dried and migrated in a preequilibrated TLC tank containing a Seymour mobile phase. Spots containing saccharides are stained blue.

Key

Lane 1 : Disaccharide standards (1,3 and 1,6); **Lane 2** : AgaB; **Lane 3** : E500; **Lane 4**: L3; **Lane 5** : F1; **Lane 6** : F24; **Lane 7** : F47; **Lane 8** : Disaccharide standards (1,3 and 1,6); **Lane 9** : *p*NPαGal.

T : denotes his-tagged enzymes.

5.5 Determination of kinetic parameters of the mutant enzymes during hydrolysis

In an attempt to understand how the mutations affected the catalytic behaviour of the enzymes, further investigations were carried out to determine their kinetic parameters in hydrolysis. The kinetic studies were performed at 25°C by following changes in absorbance at 405 nm.

Substrates such as *p*NP-disaccharides could not be used for K_m determination because of their low aqueous solubilities and high K_m values (for instance the K_m of AgaB for melibiose is 120 mM). This prevents kinetic determination because even solubility equalling K_m values may not be attained in some cases. K_m and k_{cat} (Table 5) were therefore determined using the substrate *p*NPαGal.

The determination of kinetic parameters was based on the release of *p*-nitrophenol during self- condensation reaction and hydrolysis of *p*NPαGal (cf. Materials and Methods).

All the enzymes exhibited Michaelis - Menten kinetics (Figure 13).

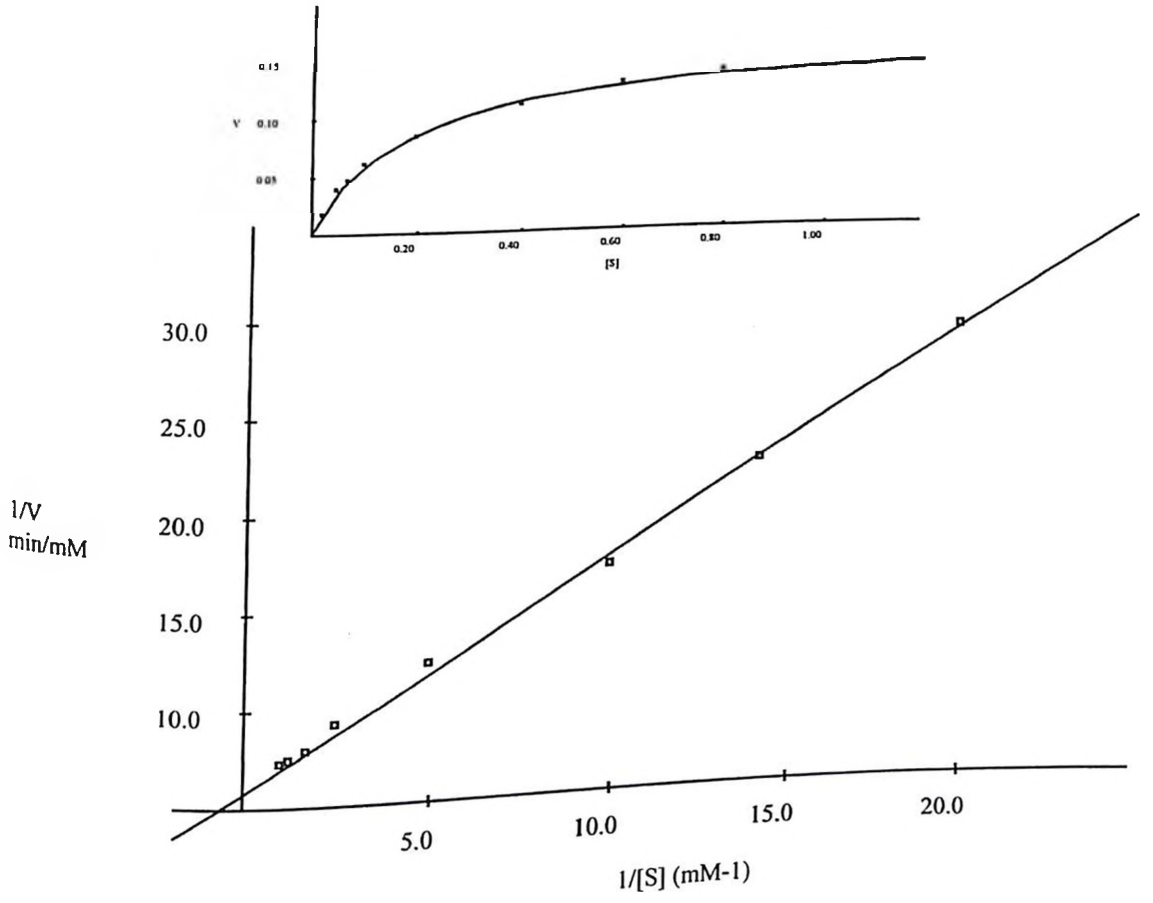


Figure 13 : Lineweaver-Burk plot for L3 mutant; inset is the corresponding plot of initial velocity against substrate concentration (Lineweaver, 1934).

This is an example of the plots obtained and analysed by EnzymeKinetics (Trinity software).

Enzyme	Mutations	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
AgaB		0.92 ± 0.01	26.5 ± 0.25	28.7
E500	A2S,T4A, I34V, E46A, R48A, S50P, N117S, A198E, K399E **G442R**	0.06 ± 0.003	0.52 ± 0.025	8.55
L3	G442R	0.19 ± 0.01	0.69 ± 0.005	3.7
F24	G442I	13.3 ± 0.15	17.5 ± 0.05	1.31
F47	G442L	14.3 ± 0.1	3.8 ± 0.05	0.27
F1	G442Y	6.3 ± 0.1	4.05 ± 0.05	0.64

Table 5 : Kinetic parameters of *p*NPαGal hydrolysis by AgaB and by mutant α-galactosidases (His-tagged)

The kinetic measurements were done at least three times and the entries represent the mean values. The data were statistically analysed by EnzymeKinetics™ a Trinity software.

The kinetic parameters of the untagged enzymes were also determined to check if the histidine tag had any effects (Table 6).

Enzyme	K_m of his-tagged enzymes (Purified enzymes) (mM)	K_m of untagged enzymes (Crude extract) (mM)
AgaB	0.92 ± 0.01	0.98 ± 0.01
E500	0.06 ± 0.003	0.25 ± 0.01
L3	0.19 ± 0.01	0.16 ± 0.01
F24	13.3 ± 0.15	12.9 ± 0.1
F47	14.3 ± 0.1	15.1 ± 0.15
F1	6.3 ± 0.1	7.3 ± 0.1

Table 6 : Comparison of the K_m of enzymes with and without his-tag.

For most enzymes, the addition of the carboxy terminal hexahistidine tag did not introduce significant changes on the K_m of the enzymes. E500 however shows a marked reduction on its K_m from 0.25 mM for the untagged enzyme to 0.06 mM for the his-tagged enzyme.

The results show that the wild type enzyme AgaB has the highest catalytic efficiency (k_{cat} of 26.5 s^{-1}). F24 at 17.5 s^{-1} displays the second highest k_{cat} after native AgaB enzyme. F47 and F1 also show k_{cat} significantly higher than those of L3 and E500. The catalytic parameters of L3 more closely match those ones of E500. L3 and E500 both have low k_{cat} values (0.69 and 0.52 s^{-1}, respectively) but high affinities for *p*NPαGal (K_m values of 0.19 and 0.6 mM, respectively). The similarity between E500 and L3 was expected given their similar mutation of G442R. AgaB also has a relatively high affinity at a K_m value of 0.92 mM. F1,

F24 and F47 have high K_m (low affinities) at 6.3, 13.3 and 14.3 mM, respectively.

5.6 Results of preliminary investigation of alternative donors

*p*NPαGal was used as the donor (cf. Literature Review) because of its convenience for measurements at the ultraviolet-visible region, and its classic use by many synthetic chemists. However, *p*-nitrophenol released by the substrate during hydrolysis, was shown in a previous study to inhibit some enzymes such as E901 (Dion *et al.*, 2000). It was also demonstrated in a previous study that E500 was inhibited by high concentrations (40 mM) of *p*-nitrophenol (data not shown). This inhibition is exacerbated by the good solubility of *p*-nitrophenol in aqueous media. This phenomenon is not without advantage because it prevents the hydrolysis of the transglycosylation product. However it could also display drawbacks in that it could hamper obtaining a high yield of transglycosylation products. Other substrates were therefore tested in order to further study the self-condensation and transglycosylation reactions catalysed by the mutants. Given the utility of TLC in reliable visualisation of saccharide products, it was employed to investigate the new substrates.

5.6.1 *o*-Nitrophenyl α-D-galactopyranoside (*o*NPα-Gal)

*o*NPα-Gal was first to be tested. The only difference between this compound and *p*NPα-Gal is the *ortho* location of the nitro group on the benzene ring. This results in different polarities at the phenolic function. By a resonance stabilisation process, the oxygen atom on the hydroxyl group of *p*-nitrophenol is more negatively charged hence the compound has higher polarity and solubity in polar solvents. *o*-nitrophenol, in contrast, precipitates easily in aqueous media. As a consequence, minimal inhibition of the native and mutant enzymes is observed.

To test *o*NPα-Gal, similar conditions as for *p*NPα-Gal were used : 3 μl of enzyme

preparation was added to 25 μl of 90 mM *o*NPα-Gal solution in 100 mM sodium phosphate buffer. 1μl aliquots were taken during the course of the reaction to test the transglycosylation products on TLC. This procedure allows determination of the time at which the substrate is consumed.

Figure 14a shows an example of the plate obtained on testing *o*NPα-Gal with AgaB. Corresponding *p*NP disaccharides were used as standards because their Rf was the same as those of *o*NP disaccharides on the TLC plate; even though *o*NPα-Gal show an Rf of 0.61 while for *p*NPα-Gal it is 0.59.

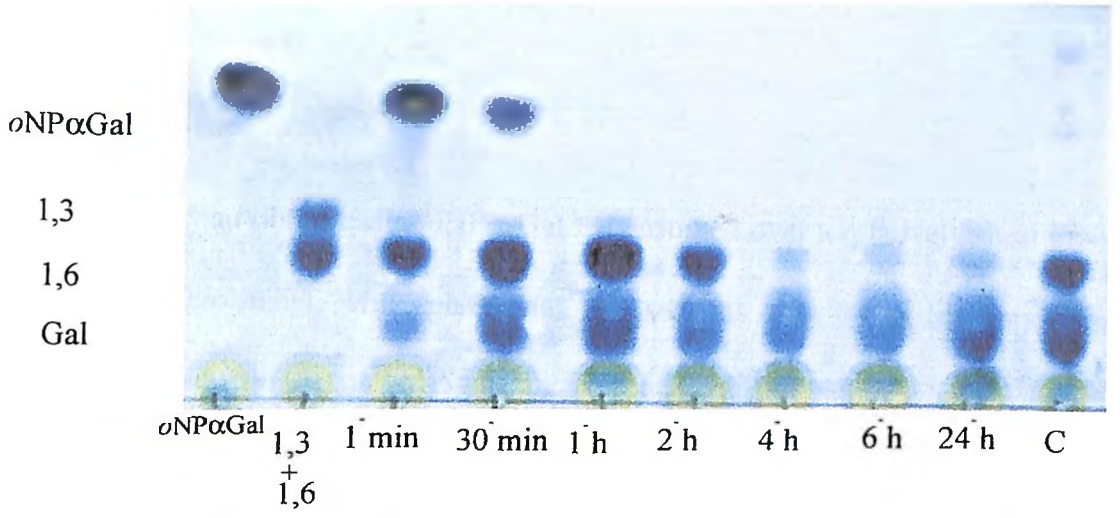


Figure 14a : Kinetics of self-condensation of *o*NPαGal catalysed by AgaB

The marked time indicates when the sample was aliquoted.
C is a sample from a *p*NPαGal self condensation reaction catalysed by AgaB aliquoted at 24 h.

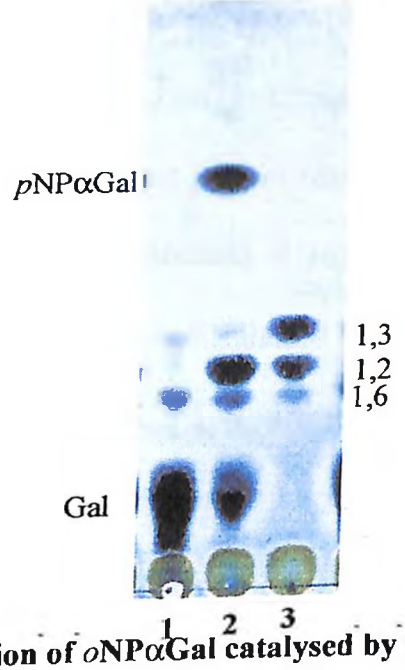


Figure 14b : Self-condensation of *o*NPαGal catalysed by E500

Lane 1 : *o*NPαGal self condensation; Lane 2 : *p*NPαGal self condensation;
Lane 3 : *p*NPdisaccharide standards.

One can observe that *o*NPα-Gal is consumed over a relatively short time (within 6 hr for all the enzymes), while *p*NPα-Gal still remains after 24 hr incubation with the same amount of enzyme (for native AgaB, there is total consumption of *p*NPα-Gal before the 24 hr time limit). Self-condensation with *o*NPα-Gal resulted in a higher hydrolysis of the disaccharide product, this is demonstrated by the marked rate of disappearance of the self-condensation product. In addition there is some change in regioselectivity profile of the products obtained. E500 synthesised α(1→6) and α(1→3) regioisomers and not α(1→2) product, obtained from *p*NPα-Gal self-condensation (figure 14b). The other mutants L3, F24 and F47 also synthesised α(1→6) and α(1→3) regioisomers on using *o*NPα-Gal as a substrate, while α(1→2) product was not present (data not shown). For F1, there was no change in the regioselectivity profile because it still synthesised α(1→6) and α(1→3) regioisomers in equal amounts. AgaB also maintained its major α(1→6) and minor α(1→3) regioselectivity. These results mean that the substrate also exert a strong influence on the regioselectivity. A major disadvantage when using *o*NPα-Gal is the precipitation of *o*-nitrophenol, which makes quantitative analyses by means of spectroscopic methods difficult.

5.6.2 Galα-F and *o*NPβGal

α-D-galactopyranosyl fluoride (Gal-α-F) was also investigated as a substrate with the mutants. Gal-α-F is a well known substrate for galactosidases. It has good water solubility and is also able to induce transglycosylation reactions (Spangenberg, 2000).

Galα-F self-condensation was initially investigated in the presence of the mutants.

This reaction was studied using 100 mM Galα-F in 10 mM, 30 mM 100 mM and 300 mM sodium phosphate buffer, as follows: 1 μl of the enzyme solution was added to 10 μl of 100 mM Galα-F dissolved in the sodium phosphate buffers of different concentrations and the reaction left to continue overnight at room temperature. 0.5 μl of the sample was analysed on TLC using Seymour's mobile phase.

Analysis of the plates revealed that AgaB gave no self-condensation products (TLC plate not shown), thus confirming the observation made by Spangenberg (2000) that this product appears at a low level and is quickly hydrolysed.

Condensation reactions of the new mutants were then studied with Galα-F as donor and *o*NPβGal as an acceptor. Total Galα-F consumption took place only in the 100 and 300 mM sodium phosphate buffer. In the buffers of weaker concentration (hence low buffering capacity), there is inhibition of the enzyme because Galα-F undergoes hydrolysis to give HF, a strong acid. The low pH that results inhibits the enzyme. However this also has the advantage of quenching the reaction and preventing product hydrolysis by the enzymes.

A buffer of low concentration was therefore used to stop the reaction before the products are hydrolysed. The reactions set up by adding 1 μl enzyme preparation in a solution containing 50 mM Galα-F and 50 mM *o*NPβGal dissolved in 10 mM sodium phosphate buffer. After overnight incubation, all the reactions stopped spontaneously. The TLC profile revealed the presence of condensation products. No self-condensation products could be detected. To confirm that the spots on the TLC plates were not Galα-F self-condensation products the plates were viewed under UV light and the fluorescent spots circled in pencil before dipping the plate in the colouring solution.

AgaB gave two condensation products corresponding to Rf of 0.33 (major product A) and 0.23 (minor product B) (Figure 15). The product occurring at the Rf of 0.33 has been identified to be Gal-α(1→6)- Galβ-*o*NP using NMR studies. The second product at Rf 0.24

could not be completely characterised due to persistent low yields.

E500 gave a major product which comigrates with the regioisomer B. F24 and F47 both gave equal amounts of the two regioisomers, A and B. L3 and F1 showed very weak activities.

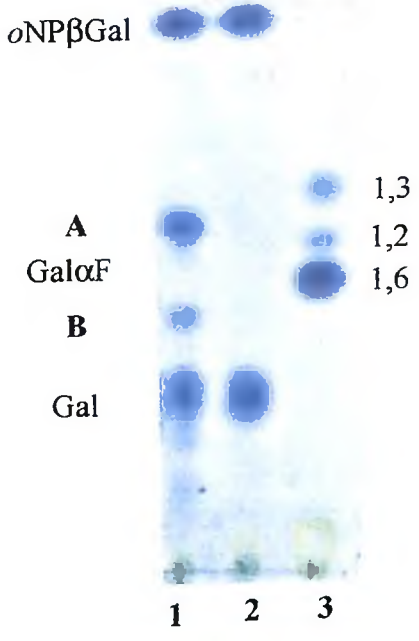


Figure 15 : Condensation reaction of GalαF and *o*NPβGal catalysed by AgaB.

Lane 1: Reaction in sodium phosphate buffer (10 mM); Lane 2: Reaction in sodium phosphate buffer (100 mM); Lane 3 : Disaccharides standards having α linkages were run alongside for comparison (abbreviations as previously indicated). 50 mM of each substrate was used in the reaction.

5.6.3 *p*NPαGlucose

An investigation of *p*NPαGlucose (*p*NPαGlc) in self-condensation reactions with AgaB and the mutant enzymes yielded no detectable self-condensation products. The same reaction conditions as for *p*NPαGal self condensation reactions were used (cf. Materials and Methods).

6.0 DISCUSSION

The objective of this study was to introduce single point mutations at the position 442 of AgaB and hence determine whether its regioselectivity would be changed to α(1→2) or other regioselectivity such as α(1→3) and α(1→4). A previous study had yielded the enzyme E500 which possessed G442R mutation in addition to other mutations. E500 displayed a major α(1→2) regioselectivity in transglycosylation with *p*NPαGal as substrate.

The mutant enzymes emerging from the saturation mutagenesis process showed properties also remarkably different from the parent enzyme. The mutants exhibited regioselective and kinetic differences from the parent AgaB. α(1→2), α(1→3), and α(1→6) products were obtained in significantly different proportions in reactions catalysed by the different mutants. The α(1→2) and α(1→3) regioselectivities are interesting given their rare occurrence in natural glycosyl hydrolases. The enzymes engineered to have these properties would find applications in the industrial sector notably in synthesis of pharmaceutical products (Magnusson *et al.*, 2000).

G442R substitution gave the L3 mutant displaying a major α(1→2) regioselectivity. This shows that arginine 442 is highly involved in the α(1→2) regioselectivity of E500. Arginine residue has a free guanidino group, normally positively charged. This group could bind to the phenyl group of the substrate through cation-π interaction hence increasing affinity for substrate, and, further hold it in place to induce α(1→2) linkage synthesis (Dion *et al.*, 2001b). This can also explain the low K_m of E500 and L3. Some of the remaining 9 amino acid mutations found in E500 could still be playing some role given that the transglycosylation level of E500 is higher than that of L3.

In contrast, F1 had a tyrosinyl residue at position 442. Tyrosine is an aromatic amino acid with a phenolic group. Its phenolic group could interact with the substrate inducing the formation of both the α(1→3) and α(1→6) disaccharides. The bulkiness of its aromatic side chain may be responsible for the low catalytic efficiency by causing steric hindrance at the enzyme active site. However, the fact that this type of regioselectivity was also found in the E901 mutant (carrying 6 mutations), demonstrates that it is possible to obtain particular regioselectivity through a single point mutation or via several mutations at different sites.

F24 and F47 carried G442I and G442L substitutions respectively. Both isomeric isoleucine and leucine are hydrophobic amino acids. The similarity of these amino acids explains the closeness in their properties when compared to the other mutants analysed.

Surprisingly, the α(1→4) regioisomer was not obtained. Two explanations for this are possible : This isomer could have been synthesised but migrated with a similar Rf to any of the other disaccharide regiosomers. The other possibility is that this product was not formed. An explanation to this may be found on comparing the reactivity of AgaB to *p*NPαGlc and *p*NPαGal. When AgaB and the mutant enzymes were used to catalyse *p*NPαGlucose self-condensation, no self-condensation products were obtained. Such a result was not surprising given the high specificity of the α-galactosidases. But it is notable that the only difference between glucose and galactose is the position of the hydroxyl group on C_4. Fourage (2000) found that the stereochemistry of this position was highly important in determining specificity of β-glycosidase of *Thermus thermophilus* because it forms a hydrogen bond with the nitrogen of a tryptophan residue in the active site. In the case of the α-galactosidase AgaB, this group may have a similar essential role in substrate emplacement within the enzyme active site. It is therefore not glycosylated because of its "unavailability" during catalysis.

The enzymes analysed had varying kinetic parameters. The native enzyme, AgaB was found to have the highest catalytic efficiency, k_{cat} of 26.5 s^{-1}. This is about 50 times higher

than that of E500 and L3; and 7 times higher than F1 and F47. F24 had a k_{cat} of 17.45 s^{-1} and was the one nearest to AgaB in activity. However, a low catalytic efficiency does not necessarily mean a low specificity of transglycosylation. This is because k_{cat} demonstrates the turnover at the active site, while the specificity constant, k_{cat} /*Km*, is a second order rate constant (in a Michaelis-Menten system) and reflects the rate of enzyme and substrate encounter. Thus an enzyme with a low k_{cat} may still have a high specificity for the substrate. It appears that the presence of arginine at position 442 enhances the enzyme-substrate (*p*NPαGal) encounter as also shown by the elevated specificity (k_{cat} /*Km*) of L3 and E500.

It can be noted that the catalytic constants, do not present any direct relationship with the observed regioselectivities. Enzymes showing different kinetic parameters had similar behaviour in regioselective and transferase activities. F24 and L3 had k_{cat} of 17.5 and 0.69 s^{-1} respectively yet they both synthesised the α(1→2) regioisomer in a similar yield.

Another observation of this study is the high dependence of regioselective synthesis on the enzyme and substrates used. When *o*NPαGal was used as the substrate, the α(1→2) regioisomer disappeared as product in the reactions catalysed by E500, L3, F24 and F47. The isomers obtained instead were α(1→3) and α(1→6) disaccharides. For E500, the addition of his-tag also lowered the α(1→2) regioisomer yield with an accompanying increase in the α(1→6) disaccharide. The high rate of hydrolysis of the transglycosylation product when this substrate is used is also noteworthy. Consequently, for industrial production processes, the acceptors and enzymes should be carefully chosen to maximise the synthesis of the desired regioisomer.

The analysis of the condensation products obtained with Galα-F and *o*NPβGal as substrates also provided some insight concerning the development of a screening method that enables isolation of mutants with high transferase activity. It would be desirable to evolve further the enzyme mutants to obtain biocatalysts with even higher transferase level. A

screening method that enables easy detection of condensation products would be critical for this. In a proposed new screening system, AgaB enzyme mutants would catalyse condensation of Galα-F and *o*NPβGal ; Galα-F acting as the donor. When a weak buffer is used for the reaction, HF released as the reaction by-product lowers medium pH thus stopping the reaction. To determine the transglycosylation activity of the enzyme, the remaining amount of *o*NPβGal is assayed by adding a β-galactosidase which hydrolyses the β-glycosidic bond. The *o*-nitrophenol produced could be assayed by reading the absorbance at 405 nm. The higher the transferase activity of an enzyme is, the lower the amount of the remaining *o*NPβGal and therefore the lower the absorbance.

The important role played by the residue 442 in AgaB regioselectivity determination prompted the examination of a possible involvement of proximal amino acids in significant function. The amino acid residue at the neighbouring position 443 in AgaB is arginine. Interestingly, this arginine is conserved in most α-galactosidases of the family 36. For these reasons, saturation mutagenesis was carried out at the codon 443 of *agaB* gene. It was found that all the 120 mutants screened showed same regioselectivity as AgaB in *p*NPαGal condensation, albeit at lower transglycosylation level for some of the clones. It can therefore be inferred that this residue is not involved in regioselectivity determination. The reduction in reactivity due to mutation at this site nonetheless suggests that it is involved in binding of the substrate or of *p*-nitrophenol.

In the final analysis, a higher understanding of the molecular processes involved in regioselectivity determination can only be understood when the three dimensional structure of AgaB is elucidated. This will give insights on the role of the residues in the active site, including the contribution by the residue 442. This would enable a more rational approach in development of AgaB mutants with improved regioselectivities.

In conclusion, this study has showed that G442R is responsible for the α(1→2) regioselectivity seen in the E500, an α-galactosidase mutant, when *p*-NPαGal is used as substrate. It was further demonstrated that a single point mutation is sufficient to change the regioselectivity of native AgaB. This has important implications for enzyme engineering because it offers a simpler alternative to random mutagenesis in directed molecular evolution of enzyme properties. However, It was also shown that regioselective modification of the product can also be achieved through a careful choice of substrate and enzyme. More work in structural modelling and directed molecular evolution is also required to enable construction of α-galactosidase AgaB variants with higher transferase activity than the mutants obtained in this study.

7.0 REFERENCES

Ajisaka, K., Nishida, H. and Fujimoto, H. (1987). Use of activated carbon column for the synthesis of disaccharides by use of reversed hydrolysis activity of β- galactosidases. *Biotechnol. Lett.* **9** : 387-392.

Borders, C.L. Jr., Broadwater, J.A., Bekeny, P.A., Salmon, J.E., Lee, A.S., Elridge, A.M. and Pett, V.B. (1994). A structural role for arginine in proteins : multiple H-bonds to backbone carbonyl oxygens. *Protein Sci* **3** : 541-548.

Burstein, C. and Kepes, A. (1971). The α-galactosidase from *Escherichia coli* K12. *Biochim. Biophys. Acta* **230:** 52-63.

Cadwell, R. and Joyce, G. (1994). Mutagenic PCR. *PCR Methods and Applications* **2** : 28-23.

Chiffoleau-Giraud V., Spangenberg, P. and Rabiller, C. (1997). β-Galactosidase transferase activity in ice and use of vinyl-β-D-galactoside as donor. *Tetrahedron : Asymmetry* **8** : 2017-2023.

Chiffoleau-Giraud V., Spangenberg P., Dion M. and Rabiller, C. (1999). Transferase activity of a β-glycosidase from *Thermus thermophilus* : specificities and limits-application to the synthesis of β-[1→3]-disaccharides. *Eur. J. Org. Chem.* **8** : 757-763.

Christians, F.C., Scapozza, L., Crameri, A., Folkers, G. and Stemmer, W.P.C. (1999). Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling. *Nature Biotechnology* **17** : 259-264.

Crameri, A., Whitehorn, E.A., Tate, A. and Stemmer, W.P.C. (1996). Improved Green Fluorescent Protein by Molecular Evolution using DNA shuffling. *Nature Biotechnology* **14** : 315-319.

de Vries, R.P., van den Broeck, H.C., Dekkers, E., Manzanares, P., De Graaff, L.H. and Visser, J. (1999). Differential expression of three α-galactosidase genes and a single β-galactosidase gene from *Aspergillus niger. Appl Environ Microbiol* **65** : 2453-2460.

Dion, M., Nisole, A., Spangenberg, P., André, C., Glottin-Fleury, A., Mattes, R., Tellier, C. and Rabiller, C. (2001). Modification of the regioselectivity of a Bacillus α-galactosidase by directed evolution. Accepted in *Glycoconjugate journal.*

Dion, M., Osanjo, G.O., André, C., Spangenberg, P., Rabiller, C. and Tellier, C. (2001b). Identification by saturation mutagenesis of a single residue involved in the α-galactosidase AgaB regioselectivity. Accepted in *Glycoconjugate journal.*

Dwek, R.A. (1996). Glycobiology: Toward understanding the function of sugars. *Chem. Rev.* **96:** 683-720.

Fourage, L. (2000). Characterisation of beta- glycosidase *ttgly* from *Thermus thermophilus*. Ph.D. thesis. University of Nantes

Fridjonsson, O., Watzlawick, H., Gehweiler, A., Rohrhirsch, T. and Mattes R. (1999). Cloning of the gene encoding a novel thermostable α-galactosidase from *Thermus brockianus* ITI360. *Appl Env Microbiol* **65** : 3955-3963.

Ganter, C., Bock, A., Buckel, P. and Mattes, R. (1988). Production of a thermostable, recombinant α-galactosidase suitable for raffinose elimination from sugar beet syrup. *J. Biotechnol.* **8:** 301-310.

Glottin-Fleury,A. (2000). Evolution dirigée de l'α-galactosidase AgaB901 de *Bacillus stearothermophilus* en vue d'améliorer son activité transférase. DEA. University of Nantes.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280:**309-316.

Henrissat, B. and Barioch (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293:**781-788.

Henrissat, B. and Davies, G. (1997). Structural and sequence -based classification of glycoside hydrolases. *Curr Opin Struct Biol* **7** : 637-644.

Hiroaki, I., Hiroshi, N., and Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96:**23-26.

Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R., and Stüber, D. (1988). Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent.*Bio/Technology* **6:** 1321-1325

Hoffmann., A. and Roeder, R. (1991). Purification of His-tagged proteins in nondenaturing conditions suggests a convenient method for protein interaction studies. *Nucl.Acids Res.* **19:** 6337-6338.

Hutchens, T.W., and Yip, T.-T. (1990). Differential interaction of peptides and protein surface structures with free metal ions and surface immobilized metal ions. *J. Chromatogr.* **500:** 531-542.

Ichikawa, Y., Look, G.C., and Wong, C.H. (1992). Enzyme-catalyzed oligosaccharide synthesis. *Anal. Biochem.* **202:** 215-238.

Ioanna, P.P. and Arnold, F.H. (2000). Designed evolution of enzymatic properties. *Curr. Opin. Biotech.* **11** : 325-330.

Jaeger, K.E. and Reetz, M.T. (2000). Directed evolution of enantioselective enzymes for organic chemistry. *Curr. Opin. Chem. Biol.* **4** : 68-73.

Joo, H., Lin, Z. and Arnold, F.H. (1999). Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. *Nature* **399** : 670-673.

Kirino H., Aoki, M., Aoshima, M., Hayashi Y., Ohba, M., Yamaghishi, A., Wakagi, T. and Oshima T. (1994). Hydrophobic interaction at the subunits interface contribute to the thermostability of 3-isopropylmalate deshydrogenase from an extreme thermophile, *Thermus thermophilus. Eur J Biochem* **220 :** 275-281.

Kuchner, K. and Arnold, F.H. (1997). Directed evolution of enzyme catalysts. *Trends biotechnol.* **15:** 523-530.

Leloir, L.F. (1971). Two decades of research on the biosynthesis of saccharides. *Science* **172:** 1299-1303.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227 :** 680-685.

Lineweaver, H and Burk D (1934). The determination of enzyme dissociation constants. *J Am Chem Soc* **56 :** 658- 666.

Lowe, J.B., Stoolman, LM., Nair, R.P., Larseden R.D., Berhend, T.L., and Marks, R.M. (1990). ELAM-1-dependent cell adhesion to vascular endothelium determined by a transfected human fucosyl transferase cDNA. *Cell* **63 :** 475-484.

Luonteri, E.M., Tenkanen, M., and Viikari, L. (1998). Substrate specificities of *Penicillium simplicissimum* α-galactosidases. *Enzyme Microb. Technol.* **22:**192-198.

Ly, H.D., and Withers, S.G. (1999). Mutagenesis of glycosidases. *Annu. Rev. Biochem.,* **68:** 487-522.

Magnusson S., Strokan V., Molne J., Nilsson K., Nilsson., Ryderberg L. and Breimer M. E. (2000). Blocking of anti-pig xenoantibodies by soluble Gal alpha 1-3Gal and Gal alpha 1-2Gal disacharides; studies in a pig kidney *in vitro* perfusion model. *Transpl Int* **13 :** 402-412.

Margolles-Clark, E., Tenkanen, M., Luonteri E. and Penttila, M. (1996). Three α-galactosidase genes of *Trichorderma reesei* cloned by expression in yeast. *Eur. J. Biochem.* **240 :**104-111.

Mathews, B.W. (1996). Structural and genetic analysis of the folding and function of T4 lysozyme. *FASEB J* **10 :** 35-41.

Maxwell, K., Mittermaier, A., Forman-Kay, J., Davidson A. (1999). A simple *in vivo* assay for increased protein solubility. *Protein Sci* **8 :** 1908-1911.

May, O., Nguyen, P.T., Arnold , F.H. (2000). Inverting enantioselectivity and increasing total activity of a key enzyme in a multi-enzyme synthesis creates a viable process for production of L-methionine. *Nat Biotechnol* **18 :** 317-320.

Menendez-Arias L., Argos P. (1989). Engineering protein thermal stability : Sequence statistics point to residue substitutions in α-helices. *J Mol Biol* **206 :** 397-406.

Miyazaki, K. and Arnold, F.H. (1999). Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. *J Mol Evol* **49 :** 716-720.

Moore, J.C., Jin, H.M., Kuchner, O., and Arnold, F.H. (1997). Strategies for the *in vitro* Evolution of Protein Function: Enzyme Evolution by Random Recombination of improved Sequences. *J. Mol. Biol.* **272:** 336-347.

Muller, S., Alving, K., Peter Katalinic, J., Zachara, N., Gooley, A.A. and Hanisch, F.G. (1999). High density O-glycosylation on tandem repeat peptide from secretory MUC1 of T47D breast cancer cells. *J Biol Chem* **274 :** 18165-18172.

Nilsson, K.G.I. (1987). A simple strategy for changing the regioselectivity of glycosidase-catalysed formation of disaccharides. *Carbohydr. Res.* **167:** 95-103.

Nilsson, KGI. (1997). Glycosidase-catalysed synthesis of di- and trisaccharide derivatives related to antigens involved in hyperacute rejection of xenotransplants. *Tetrahedron Lett.* **38 :** 133-136.

Nisole, A. (1999). Evolution in vitro de la régiosélectivité de l'α-galactosidase AgaB de *Bacillus stearothermophilus.* DEA. University of Nantes.

Petzelbauer, I., Reiter, A., Splechtna, B., Kosma, P. and Nidetzky, B. (2000). Transgalactosylation by thermostable beta-glycosidases from *Pyrococcus furiosus* and *Sulfolobus solfataricus.* Binding interactions of nucleophiles with the galactosylated enzyme intermediates make major contributions to the formation of new-glycosides during lactose conversion. *Eur. J. Biochem.* **267 :** 5055-5066.

Phillips ML., Nudelman E., Gaeta FC., Perez M., Singhal AK., Hakomori S., and Paulson JC. (1990). ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl- Lex. *Science.* **250 :** 1130-1132.

Porath, J., Carlsson, J., Olsson, I. and Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258:** 598-599.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Schmidt, R.R. (1986). New Methods for the synthesis of Glycosides and Oligosaccharides. Are There Alternatives to the Koenigs-Knorr Methods. *Angew. Chem.* Int.Ed.Engl. **25:** 212-236.

Scigelova M., Singh S.and Crout, D.H.G. (1999). Glycosidases- a great synthetic tool. *J. Molecular Catalysis B: Enzymatic.* **6:** 483-494.

Shibuya, H., Kobayashi H., G., Park G., Komatsu T., Sato T., Kim, W.-S., Yoshida, S., Kaneko, R., Nagasaki, H., Yoshida, S., Kasamo, K., and Kusakabe I. (1995). Purification and some properties of α-galactosidase from Penicillium pupurogenum. *Biosci. Biotechnol. Biochem.* **59:**2333-2335.

Smith, P.K., Krohn, R.I., Hermasan, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.I., Goeke, N.M., Olson, B.J. and Klenk, D.C. (1985). Measurement of protein using Bicinchoninic acid. *Anal. Biochem.* **150 :** 76-85.

Somers, W., Tang J., Shaw G., and Camphausen R. (2000). Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E- selectin bound to Sle (X) and PSGL-1. Cell **103 :** 467-479.

Spangenberg, P., Chiffoleau-Giraud, V., André, C., Dion, M. and Rabiller, C. (1999). Probing the transferase activity of glycosidases by means of in situ NMR spectroscopy. *Tetrahedron: Asymmetry* **10:** 2905-2912.

Spangenberg, P. (2000). α-galactosidases et Synthèse Regioselective d'Oligosaccharides. *Thèse de Doctorat*, University of Nantes.

Stemmer, W.P.C. (1994). Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* **370:** 389-390.

Stroud, M.R., Levery S.B., Martensson, S., Salyan, M.E., Clausen, H. and Hakomori, S. (1994). Human tumour -associated Le(a)-Le(x) hybrid carbohydrate antigen IV3 (Gal beta 1-3[Fuc alpha 1-4] Glc NAc) III3 FucnLc4 defined by monoclonal antibody 43-9F : enzymatic synthesis, structural characterization and comparative reactivity with various antibodies. *Biochemistry.* **33 :** 10672-10680.

Stutz, A.E. (1996). A novel approach for clarifying the reaction mechanism of retaining glycoside hydrolases. *Angew. Chem. Int. Ed. Engl.* **35:** 1926-1928.

Toone, E.J., Simon, E.S., Bednarski, M.D. and Whitesides, G.M. (1989). Enzyme-catalysed synthesis of carbohydrates. *Tetrahedron* **45:**5365-5422.

Vaughan , H.A., Loveland B.E. and Sandrin M.S. (1994). Gal alpha (1→3) Gal is the major xenoepitope on pig endothelial cells recognized by naturally occurring cytotoxic human antibodies. *Transplantation* **58 :** 879-882.

Vetere, A. and Paoletti, S. (1996). Complete synthesis of 3-sialyl-N-acetyllactosamine by regioselective transglycosylation. *FEBS Lett.* **399 :** 203-206.

Wang, C., Eujeini, M., Turano, C. and Giartosio, A. (1996). Influence of carbohydrate moiety on the stability of glycoproteins. *Biochemistry* **35 :** 7299-7307.

Wang, Q., Withers S.G. (1995). Substrate- assisted catalysis in glycosidases. *J. Am. Chem. Soc.* **117:**10137-10138.

Watanabe, K., Chishrio, K., Kimatura, K. and Suzuki Y. (1991). Proline residues responsible for thermostability occur with high frequency in loop regions of an extremely stable oligo-1,6-glucosidase from *Bacillus thermoglucosidiacus* KP1006. *J Biol Chem* **266 :** 24287-24294.

Withers, S., Warren, R.A.J., Street, I.P., Rupitz, K., Kempton, J.B.and Aebersold, R. (1990). Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of "retaining" glycosidase.*J. Am. Chem. Soc.* **112:** 5887-5889.

Withers, S.G. (1995). Enzymatic cleavage of glycosides: how does it happen? *Pure& Appl. Chem.* **67:**1673-1682.

Yoon, J.H. and Ajisaka, K. (1996). The synthesis of galactopyranosyl derivatives with β-galactosidases of different origins. *Carbohydr. Res.* **292 :** 153-163.

Zhao, H., Giver, L., Shao, Z., Affholter, J.A. and Arnold, F.H. (1998). Molecular evolution by staggered extension process (StEP) *in vitro* recombination. *Nature Biotechnology* **16 :** 258-261.