IMPROVEMENT OF THERMOSTABILITY OF A FUNGAL XYLANASE USING ERROR-PRONE POLYMERASE CHAIN REACTION (EpPCR)

By

SARVESHNI PILLAY

Submitted in fulfillment for the Degree of

Master of Technology (Biotechnology)

Department of Biotechnology and Food Technology

Faculty of Applied Sciences

Durban University of Technology

Supervisors: **Dr. K. Permaul (Ph.D)**

Prof. S. Singh (Ph.D)

DECLARATION

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Technology, to the Durban University of Technology. It has not been submitted before for any degree or examination to any other University.

_____

S. PILLAY

February 2007

FINAL COPY APPROVED FOR SUBMISSION

_____

Prof. S. Singh

_____

Dr. K. Permaul

DEDICATION

IN LOVING MEMORY OF MY LATE DAD.

"The great tragedy of science- the slaying of a beautiful hypothesis by an ugly fact"

Thomas Huxley

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The author wishes to thank:

Prof. S. Singh and Dr. K. Permaul, for their invaluable guidance and assistance for the duration of this study;

Inqaba Biotec for sequencing of samples;

Colleagues and friends especially Dawn, Santosh, Predeshini and Prenaven in the Department of Biotechnology for their help and encouragement;

My mother for being a pillar of strength in my life;

My family especially my children, Varshen and Tharun, for their patience, support and understanding and

GOD Almighty, for His sovereignty over me.

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Interest in xylanases from different microbial sources has increased markedly in the past decade, in part because of the application of these enzymes in a number of industries, the main area being the pulp and paper industry. While conventional methods will continue to be applied to enzyme production from micro-organisms, the application of recombinant DNA techniques is beginning to reveal important information on the molecular basis and this knowledge is now being applied both in the laboratory and commercially. In this study, a directed evolution strategy was used to select an enzyme variant with high thermostability. This study describes the use of error-prone PCR to modify the xylanase gene from *Thermomyces lanuginosus* DSM 5826, rendering it tolerant to temperatures in excess of 80°C. Mutagenesis comprised of different concentrations of nucleotides and manganese ions. The variants were generated in iterative steps and subsequent screening for the best mutant was evaluated using RBB-xylan agar plates. The optimum temperature for the activity of xylanases amongst all the enzyme variants was 72°C whilst the temperature optimum for the wild type enzyme was 70°C. Long term thermostability screening was therefore carried out at 80°C and 90°C. The screen yielded a variant which had a 38% improvement in thermostability compared to the wild type xylanase from pX3 (the unmutated gene). Successive rounds of error-prone PCR were carried out and in each round the progeny mutant displayed better thermostability than the parent. The most stable variant exhibited 71% residual activity after 90 minutes at 80˚C. Sequence analysis revealed four single amino acid residue changes that possibly enhanced their thermostabilities. This *in vitro* enzyme evolution technique therefore served as an effective tool in improving the thermostable property of this xylanase which is an important requirement in industry and has considerable potential for many industrial applications.

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

While enzymes are a very useful complement to the increasing range of chemical catalysts for the synthesis of complex chiral molecules, the process conditions in an industrial catalytic reactor are rarely ideally suited to maintaining a highly active and long lasting enzyme. Natural enzymes have evolved over millions of years to operate in a specific environment and despite their wide diversity, natural and industrial environments are significantly different (Hibbert *et al.,* 2005). Thus, for numerous industrial processes a satisfactory enzyme cannot be readily found in nature, implying that there is a need for novel biocatalysts.

Although many complicated chemical reactions can be efficiently performed by biocatalysts, as mentioned, industrial conditions are usually different from those in nature with respect to substrate concentrations, sheering forces, temperature and organic solvents. These factors make the application of enzymes for industrial synthesis difficult, in particular, where multiple (rather than singular) traits need to be satisfied to create the ideal industrial enzyme catalyst. Most enzymes found in soil and water may display the desired activity, but are generally not suited for industrial use (Otten and Quax, 2005).

Nevertheless, there are very strong reasons for attempting to overcome these barriers, such as the exquisite chemo-, regio- and stereo-selective and specific properties of enzymes and their ability for effective catalysis under mild conditions (Hibbert *et al.*, 2005). One method of overcoming some of the barriers to implementing enzymes in industry is the application of directed evolution, in which the amino acid sequence of an enzyme is iteratively altered until the enzyme functions in the desired manner. Changing a protein by directed evolution opens the possibilities for moving towards a variety of required properties. The most popular targets for directed evolution to date have been activity, substrate specificity, thermal and oxidative stability, enantio-selectivity or enantio-specificity, pH range and tolerance to solvents. In many cases, directed evolution

of an enzyme has enabled industries to reduce the number of steps required in an equivalent chemical or biocatalytic synthesis (Hibbert *et al.*, 2005).

The industrial enzyme market is steadily growing due to improved production technologies, engineered enzyme properties and new application fields. The estimated value of world enzyme market is presently about 2.3 billion US dollars and it has been forecasted to grow to almost 3 billion US dollars by 2007 (Lesney, 2003). The applications are diverse, but the majority of bulk industrial enzymes are used in the detergent (37%), textile (12%), starch (11%), baking (19%) and animal feed (6%) industries which use about 75% of industrially produced enzymes. The remaining 25% include isomerases, ligases, lyases, oxidoreductases and transferases (Lesney, 2003). Research towards the use of biocatalysts is mainly driven by the desire of using sustainable technologies for the production of useful ingredients in a pharmaceutical and agrobiological context.

Directed evolution has developed rapidly to become a method of choice for protein engineers in order to create enzymes having desired properties for various processes. Over the last decade this technique has become a routine part of the molecular toolbox of every biochemist. This is emphasized by the increasing number of publications about the subject (Arnold *et al*., 2001; Zhao *et al*., 2002; and Dalby, 2003). In the 1990s, most publications were about new methods for mutagenesis and recombination (Arase *et al*., 1992; Giver *et al*., 1998 and Marrs *et al*., 1999). They described the evolution of a model enzyme to establish the new methods. In the past years, however, more research deals with enzymes that can be used in industry or research (Otten and Quax, 2005). Thus, new as well as extant enzymes have the capability to become seemingly whatever the biotechnologist chooses them to be through the use of protein engineering and directed evolution.

1.1 MICROBIAL XYLANASES

1.1.1 Occurrence and mechanistic action of xylanases

Xylan is the main carbohydrate found in the hemicellulosic fraction of plant tissues and accounts for one-third of all renewable organic carbon available on earth (Prade, 1996). Xylanase is the major component of an enzymic consortium, which acts in nature by depolymerizing xylan molecules into monomeric pentosan units that are used by bacterial and fungal populations as a primary carbon source. Xylanase producers have been isolated from all ecological niches where plant material is deposited (Prade, 1996). Numerous bacteria, yeasts and filamentous fungi can degrade xylan by producing a range of enzymes (Sunna and Antranikian, 1997). Endo-1,4-β-D-xylanase is regarded as the most important enzyme, but other debranching enzymes such as acetylxylan esterase and α-glucuronidase are thought to play a synergistic role in the hydrolysis process (Biely, 1993).

Xylanase from archaea are mainly intra-cellular or cell-associated (Niehaus *et al*., 1999). Although thermophilic filamentous fungi are not as extreme as thermophilic eubacteria or archaea, they typically excrete xylanase into the medium and produce several accessory xylanolytic enzymes which are necessary for debranching substituted xylan (Haltrich *et al*., 1996). Furthermore, the xylanase levels from fungal cultures are generally much higher than those from bacteria or yeast (Haltrich *et al*., 1996). This has been proven in a study carried out by Singh *et al*. (2000) where the thermophilic fungus *Thermomyces lanuginosus* SSBP has been reported to be one of the best xylanase producers. The enzymes from these filamentous fungi are of particular interest, from an industrial point of view, due to the fact that they excrete xylanases into the medium. Only a few thermophilic microorganisms are able to grow on xylan and secrete thermoactive xylanolytic enzymes (Maheshwari *et al*., 2000).

Xylan is composed of a backbone of 1,4-β-linked D-xylose units that may be substituted with acetic acid, arabinose and 4-O-methyl-D-glucoronic acid (Biely, 1993). β-D-

xylosidase hydrolyzes xylooligomers and xylobiose from the non-reducing end, liberating xylose (Rizzati and Sandrim, 2004). Hydrolysis of β-glycosidic linkages is specified by a catalytic reaction common to all glycanases, whereas substrate recognition, is specified by subsites that interact with adjacent glycosyl units. Some developments have shown that metabolic pathways can be transferred from one organism to another and proteins can be modified to gain conformational stability, suggesting that naturally occurring systems can be custom engineered to the situation in the fermentation tank (Prade, 1996).

1.1.2 Classification and structural description of xylanase families

Xylanases are classified into two major families (10 or F and 11 or G) of glycosyl hydrolases. Family 10 xylanases are larger, more complex and produce smaller oligosaccharides e.g., xylanases from *Cryptococcus albidus* and *Streptomyces lividans*. The family 10 catalytic domain is a cylindrical α/β barrel resembling a salad bowl, with the catalytic site at the narrower end, near the carboxyl terminus of the β-barrel. Family 11 xylanases are more specific for xylan. There are five crystal structures of low molecular weight endo-1,4-βxylanases belonging to family G/11, two xylanases from *Trichoderma reesei*, one from *Trichoderma harzianum*, one from *Bacillus circulans* (Torronen and Rouvinen, 1997) and one xylanase from *Thermomyces lanuginosus* (Gruber *et al*., 1998). The overall shape of the xylanase molecule resembles a partly-closed right hand which consists of two β-sheets and one α-helix (Fig.1). The two β–sheets and the α-helix form the "fingers" and the "palm" and the two loop regions form the "thumb" and "cord" (Fig.1).

However, variations in their functional properties, such as catalytic activity, substrate cleaving patterns, pH optima and thermostabilities, exist (Torronen and Rouvinen, 1997). Both families use ion pair catalytic mechanisms and both retain anomeric configuration following hydrolysis. The extensive sharing of structural features within β-glycanase families suggests that gene duplication and conversion events have occurred during xylanase evolution (Prade, 1996).

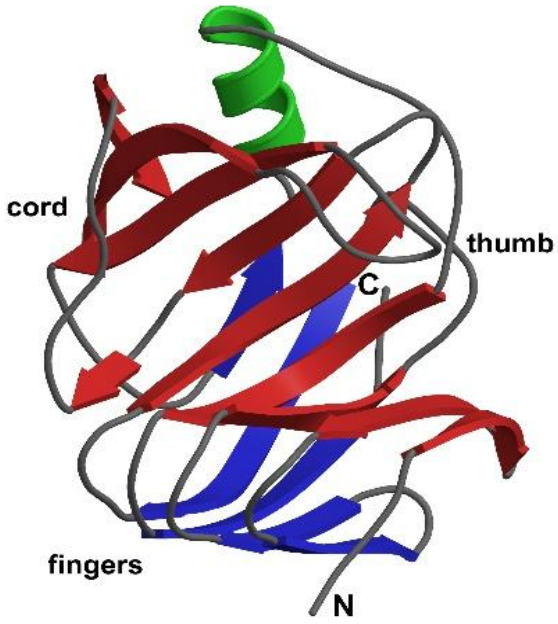


Fig. 1.1 Three-dimensional structure of endo 1,4-β-xylanase from *T. lanuginosus* (Gruber *et al*., 1998).

1.1.3 Properties of xylanase produced by *T. lanuginosus*

A thermostable xylanase from the filamentous fungus *T. lanuginosus* (DSM 5826) was purified (Schlacher *et al*., 1996). This enzyme has an apparent molecular weight of 24 to 26 kDa as determined by SDS polyacrylamide gel electrophoresis. cDNA and genomic DNA fragments coding for this enzyme were cloned and sequenced (Schlacher *et al*., 1996). The cDNA contains an open reading frame encoding a polypeptide of 225 amino acids and was functionally expressed in *E. coli* as a LacZ fusion protein. The family G/11 xylanases show clear amino acid homology and thus have a common fold. Comparison of the cDNA sequence with the genomic DNA sequence showed that the xylanase was encoded by two exons interrupted by an intron of 106 bp (Schlacher *et al*., 1996). It exhibited the highest activity at a pH of 7.0 and in the temperature range of 60-70°C (Cesar and Mrša, 1996).

Single-crystal x-ray diffraction revealed that the thermostability of *T. lanuginosus* was due to the presence of an extra disulfide bridge, as well as to an increase in the density of charged residues throughout the protein (Gruber *et al*., 1998). There is one disulfide bridge in the *T. lanuginosus* enzyme which does not exist in the majority of other family 11 xylanases. It connects the C-terminus of the β- strand B9 (residue 110) with the N-terminus of the α–helix (residue 154). All the other xylanases with known crystal structures have a hydrogen bond at the corresponding position. It was probably this unique structure of the enzyme that was responsible for its' intrinsic stability.

Puchart and co-workers (1999) reported very low amounts of debranching enzymes such as α-glucuronidase, acetylxylan esterase and α-L-arabinofuranosidase being produced by *T. lanuginosus* DSM 5826. There are no reports on the production of other types of hemicellulose-hydolyzing enzymes in *T. lanuginosus* and there is only a single report on the production of pectinolytic enzymes (Gomes *et al*., 1992).

1.1.4 Commercial application of xylanases - an overview

Interest in xylanases from different sources has increased markedly in the past decade, in part because of the application of these enzymes in a number of industries (Table 1.1). They have been used in the feed industry, to improve the digestibility of animal feeds and in the baking industry, to assist with dough preparation as a result of the reduction in viscosity of highly viscous cereal arabinoxylans (Collins *et al*. 2004). Enzymatic saccharification of xylan in agricultural and industrial wastes may contribute to a greater utilization of lignocellulosic biomass, resulting in useful chemicals and fuels.

The main focus area of xylanase research in this study is directed towards the pulp and paper industry. The use of xylanase in biobleaching of wood pulp produced in the sulphite (acid) and kraft (alkaline) processes has received much attention (Christov and Prior, 1993). Thermostable xylanases would be particularly useful in biobleaching,

directly after the cooking process of wood (Yu *et al*., 1987), which is usually conducted at temperatures above 45°C (Bierman, 1993). Xylanases produced by mesophilic fungi lack thermostability, thereby rendering their application to industry very limiting (Singh *et al.*, 2000). Despite an increased knowledge of microbial xylanolytic systems in the past few years, further studies are required to achieve a complete understanding of the mechanism of xylan degradation by micro-organisms and their enzymes. The enzyme system used by microbes for the metabolism of xylan is the most important tool for investigating the use of the second most abundant polysaccharide (xylan) in nature.

Table 1.1 Large scale applications of xylanase (Collins *et al*., 2004)

Industry	**Function**
Food: fruit and vegetable juices	Improves maceration and clarification
Baking: dough and bakery products	Improves elasticity and strength of the dough and improves bread texture
Feed: mono-gastric and ruminant feeds	Improves digestibility and nutritive value of poorly degradable feeds e.g. barley and wheat
Pulp and paper	Reduces chlorine consumption and toxic discharges Reduces use of mechanical pulping methods, hence reduces energy consumption
Starch	Reduces viscosity, improves gluten agglomeration and process efficiency
Textile	Reduces/replaces chemical retting methods

1.2 STRATEGIES FOR PROTEIN ENGINEERING

The importance of protein engineering in industry continues to grow as the number of applications of proteins expands, and the technology to discover proteins with useful properties are able to address industrially related problems.

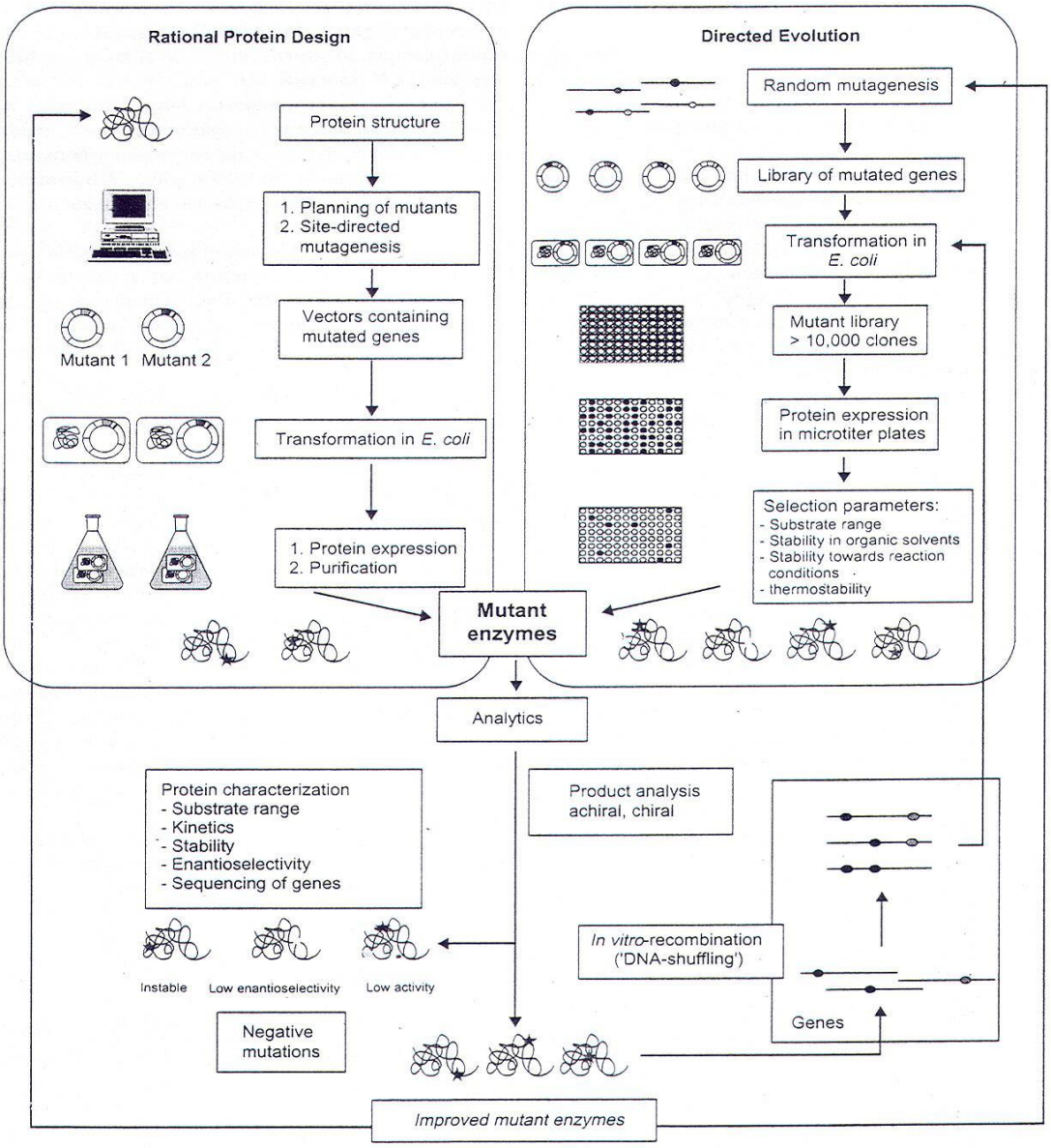


Fig. 1.2 Schematic diagram of protein engineering (Bornscheuer and Pohl, 2001)

1.2.1 Rational protein design

Rational design (Fig. 1.2) was the earliest approach to protein engineering and is still the most widely used way to introduce desired properties into a protein of interest. The strategy hinges on relating structure to function, frequently via molecular modeling techniques. It usually requires both the availability of the structure of the enzyme and

knowledge about the relationship between sequence, structure and mechanism/function, and is therefore very information-intensive (Bornscheuer and Pohl, 2001).

Rational design, relies on structural analysis and site-directed mutagenesis for engineering proteins with improved properties. Structure-based rational protein design has yielded some success, however, the approach has its drawbacks. Structural prediction and site-directed mutagenesis of target residues tend to be resource and labour-intensive, as well as time-consuming. Perhaps more importantly, the effects of multiple mutations on protein function are not easily predictable, even with structural data on hand (Tobin *et al*., 2000).

Advances in rational design depend on progress in structure determination, improved modeling procedures and significant new insight into structure-function relationships. A promising new technique for obtaining structural information on hard to crystallize proteins is the two-dimensional (2D) crystallization technique that utilizes metal ion co-ordination to surface histidines in order to crystallize proteins at interfaces (Rubingh, 1997).

1.2.2 Directed evolution: a powerful tool for protein engineering

Directed evolution (Fig. 1.2) also called evolutive biotechnology or molecular evolution, involves making random mutations in the gene encoding an enzyme of interest and does not require any structural or mechanistic information about the protein or about the reactions it performs (Jaeger and Reetz, 2000). Directed evolution is attracting increasing interest from industry as many questions on the relationship of protein properties to structure are still unanswered. In its simplest form, directed evolution is a method to optimise enzymes for desired properties through iterative rounds of mutagenesis (Benkovic *et al*., 2000).

Of all the novel methods for industrial enzyme discovery, directed evolution may be the most powerful. Directed evolution is a fast and inexpensive way of finding variants of existing enzymes that work better than naturally occurring enzymes under specific conditions. It mimics natural evolution in that it depends upon the selection of fitter 'individuals', in this case, enzymes, from a diverse population. Directed evolution is 'directed' in the sense that the researcher selects the variant enzymes that better meet some criterion via a series of step-wise improvements. It starts with a cloned gene for an enzyme that needs to be improved. Diversity is typically enhanced by *in vitro* mutagenesis of the isolated gene. The mutagenized gene, now a population of mutant variants, is cloned and expressed in a convenient host and many of the clones are assayed for the activity of the enzyme in question. The best of the variant clones is selected and it is used as a starting point for subsequent rounds of mutagenesis, screening and selection (Marrs *et al*., 1999).

1.2.2.1 *Requirements for successful directed evolution*

The main requirements for successful directed evolution are: the functional expression of the enzyme in a suitable microbial host, the availability of a screen (or selection) sensitive to the desired properties and identifying a workable evolution strategy (Kuchner and Arnold, 1997). The vast majority of possible evolutionary paths lead to poorer enzymes, the strategic challenge is to identify a path that will result in the improvement of the desired feature(s). The number of possible variants increases rapidly with the size of the enzyme and with the number of amino acids that are allowed to vary simultaneously. Even for a small protein, an impossibly large number of variants can be generated when multiple mutations are introduced (Kuchner and Arnold, 1997).

1.2.2.2 *Methods for directed evolution*

Mutations can be introduced into genes of interest by various methods. Random point mutagenesis by error-prone polymerase chain reaction (epPCR) involves a modified PCR protocol that uses variations in $MgCl_2$ or $MnCl_2$ concentrations to achieve an average

mutation level of 2-5 base substitutions per gene (Jaeger and Reetz, 2000). Another way to create mutant DNA is through DNA shuffling, as pioneered by Stemmer (1994) which mimics the process of natural recombination. This technique involves the recombinative exchange of small DNA fragments between two or more closely related genes, thereby leading to the creation of new genes (Jaeger and Reetz, 2000).

Zhao *et al*. (1998) developed two variations of recombinative DNA technology. The staggered extension process (StEP) involves PCR amplification of a gene using a set of primers and short time intervals for annealing and for polymerization (Zhao *et al*., 1998). Truncated oligomers are dissociated from the template and allowed to anneal randomly to different templates, thereby accomplishing recombination. Repetition of this process results in a 'staggered' extension of the DNA template during PCR (Jaeger and Reetz, 2000).

The random priming recombination (RPR) method uses random hexamer oligonucleotide primers for PCR that yield a collection of random oligomers containing additional point mutations introduced by the polymerase action (Shao *et al*., 1998). After removal of the full-length template, the oligomers can reassemble with other oligomers containing mutations and are finally amplified to yield complete genes, which are shuffled by a recombination-like process (Jaeger and Reetz, 2000). Libraries thus created are then usually assayed via high-throughput technologies to identify improved variants (Bornscheuer and Pohl, 2001).

1.3 WHY DIRECTED EVOLUTION?

When natural enzymes are recruited for industrial applications, from serving as catalysts in chemical synthesis to additives for laundry detergents, it has been discovered that they are often not well suited to these tasks. Due to poor substrate solubility, breakdown of unstable products or competing chemical reactions, the conditions for an enzyme reaction may be unsuitable for large-scale applications. In evolution, enzymes are optimised and

often highly specialized for specific biological functions within the context of a living organism. Biotechnology needs enzymes that are stable and active over long periods of time, are active in non-aqueous solvents and can accept different substrates (Arnold, 1998).

It is possible to produce new enzymes in recombinant organisms, altering the amino acid sequence and therefore the properties through appropriate modifications at the DNA level. This is hampered, however by ignorance of how the amino acid sequence affects every aspect of enzyme performance, from its ability to be expressed in a heterologous host to its catalytic activity in an industrial environment. Even if one trait is successfully designed such as enhanced stability, it is virtually impossible to predict the cost to another such as catalytic activity or expression level. All these hurdles are bypassed by directed evolution (Arnold, 1998).

Thus, the successful implementation of enzymes as industrial biocatalysts requires the availability of suitable enzymes with high activity, specificity and stability under process conditions (Zhao *et al.,* 2002). However, naturally occurring enzymes are often not optimized to fulfill these requirements. Within this context, directed evolution is very effective in closing these functional discrepancies (Zhao *et al.*, 2002).

Molecular diversity is typically created by various random mutagenesis and/or *in vitro* gene recombination methods (Kurtzman *et al.,* 2001). Functionally improved variants are first identified by a high-throughput screening or selection method and then used as the parents for the next round of evolution. As a result, the success of directed evolution experiments often depends on the choice of diversity-generation methods and the availability of screening/selection methods (Zhao *et al.,* 2002).

Depending on the choice of screening/selection methods, enzyme stability and activity could be simultaneously improved or one property is improved at the cost of another (Zhao *et al.,* 2002). By screening for initial activity and residual activity at an elevated temperature, both the thermostability and activity of mesophilic subtilisin E,

psychrophilic subtilisin S41 and mesophilic *p*-nitrobenzyl esterase were significantly increased using directed evolution strategies (Zhao *et al.*, 2002). By contrast, the evolved variants of a thermophilic indoleglycerol phosphate synthase using a genetic selection in *E. coli* exhibited increased activity and decreased thermostability (Merz *et al.*, 2000). Using a similar selection method, four variants of a thermophilic 3-isopropylmalate dehydrogenase with enhanced specific activity at low temperatures were obtained. Two of these variants maintained the wild type thermostability, whereas the other two variants exhibited decreased thermostability (Suzuki *et al.*, 2001). Clearly, with no direct selective pressure on both properties, either activity or thermostability could drift randomly in the fitness landscape (Zhao *et al.*, 2002).

The potential application of biocatalysts in industrial biocatalysis would be greatly expanded if enzymes could function in non-natural environments such as organic solvents (Zhao *et al.*, 2002). Using random mutagenesis, recombination and screening, Song and Rhee (2001) obtained three variants of phospholipase A_1 with enhanced stability and activity in organic solvents. Using a similar strategy, a variant of horseradish peroxidase with increased stability in the presence of hydrogen-peroxide, sodium dodecyl sulphate and salts were obtained (Morawski *et al.*, 2001).

The laboratory equivalent of Darwin's survival of the fittest is mutation and recombination accompanied by selection on the basis of desired functional change (Arnold and Volkov, 1999). A relatively simple algorithm of recursive random mutagenesis and recombination can deal with the many important protein engineering problems that fall in the 'tuning' category e.g., enhancing activity, altering substrate specificity, improving stability and a variety of other features already exhibited at some, low level by the parent enzyme (Arnold and Volkov, 1999).

Provided the desired enzyme is physically feasible, there appears to be a good chance that the accumulation of mutations identified in a well designed selection or screen will yield it. Once the goal has been identified and the enzyme expressed in a functional form in a suitable host, developing and validating the selection or screen becomes the most challenging part of the evolution experiment (Arnold and Volkov, 1999).

1.4 SUCCESSFUL IMPLEMENTATION OF DIRECTED EVOLUTION

It is clear that directed evolution has played a central role in improving commercially important enzymes and in finding new applications of proteins quite different from their natural function. From an industrial point of view, being able to rapidly identify such an enzyme is also very important (Otten and Quax, 2005).

Zhao and Arnold, (1999) converted a psychrophilic enzyme, subtilisin S41 from Antarctic *Bacillus* TA41, into its thermophilic counterpart. At 60°C, variant 3-2G7 has a half-life of 449 minutes, which is more than 500 fold greater than that of the wild type enzyme and considerably more than its mesophilic homologs, subtilisin BPN′ and SS11′, whose half-lives under the same conditions are only 28 and 34 minutes, respectively. Only eight amino acid substitutions were sufficient to create the new 'thermophilic' subtilisin E.

Applications for cold-adapted enzymes are appearing in food processing, laundry, environmental remediation and biosynthethic processes. Directed evolution was used to create a cold-adapted subtilisin BPN′ with twice the activity of the wild type enzyme at 10°C (Liu *et al.*, 1997).

Directed evolution has been used to probe the relationship between thermostability and catalytic activity during the evolution of a thermophilic version of the *p*-nitrobenzyl esterase from *B. subtilis* (Giver *et al.*, 1998). A 14°C increase in melting temperature (T_m) was achieved after four cycles of random mutagenesis and one of DNA shuffling. By screening simultaneously for thermostability and activity at 30°C, it was possible to create a variant that was both thermostable and retained its activity at lower temperatures. Catalytic activity increased with temperature, such that at high temperatures (near the T_m) the activity of the mutant was about twenty-fold than that of wild-type at low temperatures. This clearly demonstrates that thermostability and catalytic activity at low temperatures are not mutually exclusive: acquisition of thermostability need not come at the cost of catalytic activity.

The thermostability of 3-isopropylmalate dehydrogenase (IPMDH) from *B. subtilis* was improved by directed evolution using an *in vivo* selection. This dehydrogenase gene, when integrated into the chromosome of an auxotrophic strain of *Thermus thermophilus*, complements a defective *leuB* gene and supports cell growth in minimal media (Akanuma *et al.*, 1998). Spontaneous mutations were accumulated at lower temperatures and thermostable variants were then selected by growth at elevated temperatures. After three cycles, an IPMDH variant allowed bacterial growth at 70°C, a 14°C increase over the starting conditions. The increased thermostability of this mutant had only three amino acid substitutions (Akanuma *et al.*, 1998).

Flippase (FLP) recombinase from *Saccharomyces cerevisiae* is used in genetic experiments involving mutations such as chromosomal translocations, large deletions and site-specific integrations in order to manipulate DNA in a site-specific fashion (Arnold and Volkov, 1999). This yeast enzyme performs poorly at physiological temperatures in mammalian systems. Repeated cycles of error-prone PCR, DNA shuffling and screening were used to improve the performance of FLP recombinase (Buchholz *et al.*, 1998). The best variant showed a four-fold increase in recombinase activity at 37°C and a tenfold increase at 40°C. This improved enzyme did not start to denature until about 37°C, while the wild-type FLP recombinase already showed signs of denaturation at 20°C (Arnold and Volkov, 1999).

The lipase from *T. lanuginosus* was successfully evolved to improve its performance in laundry detergents (Buchholz *et al*., 1998). A combination of random mutagenesis of the entire gene and site-directed mutagenesis of selected regions of the protein followed by *in vivo* recombination in *S. cerevisiae*, produced a variant with the desired wash properties (Buchholz *et al*., 1998).

With directed evolution we now have the ability to tailor individual proteins as well as whole biosynthetic and biodegradation pathways for biotechnology applications (Hibbert *et al*., 2005). As is evident from the examples that are highlighted above, directed

evolution techniques have allowed generation of enzymes with enhanced characteristics and with new and completely novel substrate specificities.

Using the error-prone polymerase chain reaction (epPCR), saturation mutagenesis and screening, Jaeger and Reetz (2000) considerably increased the enantioselectivity of a *Pseudomonas aeruginosa* lipase towards 2-methyldecanoate from 2% to 90-93%. A particularly impressive example is the evolution of ampicillin-resistant activity from a functionally unrelated DNA fragment (Yano and Kagamiyama, 2001). The DNA fragment conferring very low ampicillin–resistant activity on *E. coli* was isolated from genomic DNA library of a hyperthermophilc *Pyrococcus furiosus* in which no β-lactamase activity has ever been detected. After 50 rounds of DNA shuffling and screening at increasing ampicillin concentrations, the ampicillin resistant activity of this clone on *E. coli* was significantly enhanced.

In 2005, Matsumura and Rowe obtained a clone from a library of β-galactosidase mutants that exhibited 100-fold improvement in catalytic activity. These authors optimized the performance of a plasmid-borne *E. coli lacZ* gene in two rounds of directed evolution. Firstly, its promoter was 'randomized' by whole-plasmid PCR. A genetically stable constitutive expression vector was isolated by *in vivo* genetic selection. Secondly, the entire plasmid was mutated randomly in a slightly mutagenic long polymerase chain reaction. Kichise *et al*. (2002) reported *Aeromonas caviae* FA440 synthase mutants with improved synthesis yields of the biodegradable polyester polyhydroxyalkanoate. Canada *et al*. (2002) improved the activity of toluene ortho-monooxygenase for 1-naphthol synthesis and degradation of chlorinated ethylene pollutants. Impressively, most of these experiments required only two to three rounds of mutagenesis and library sizes of 10^3 and 10^5 variants (Tao and Cornish, 2002).

Successes in the directed evolution approach suggest it may be more efficient than rational design. Directed evolution has become an increasingly important strategy in the development of enzymes, as it allows one to sidestep some of the difficult questions relating to the structural and functional properties of such proteins. It is also important to

remember that in order to engineer a commercially useful protein many properties need to be changed, and frequently these changes are interdependent (Lutz and Patrick, 2004).

1.5 GENERATING A MUTANT LIBRARY

Mimicking natural evolution, an initial parent gene is chosen and a diverse library of offspring genes is created through mutagenesis or recombination. A screen or selection is applied to the library and the mutants that exhibit the greatest improvement in the desired properties are chosen to become the parents to the next generation. This iterative search has generated large improvements in properties such as activity, specificity, stability and has been used to evolve systems of enzymes, such as that found in metabolic pathways (Schmidt-Dannert *et al.,* 2000). Directed evolution often discovers these improvements by making a few amino acid substitutions that collectively have an important functional effect. For this reason, the method generally requires a starting protein with some activity towards the desired reaction (Zhao and Arnold, 1999).

The laboratory approach to Darwinian evolution of biological macromolecules, often referred to as directed molecular evolution, presents a powerful means for exploring and altering enzyme substrate specificity and catalysis (Lutz and Patrick, 2004). In contrast to rational design, which generally concentrates on a small number of variants, evolutionary methods for protein engineering rely on the generation of vast molecular diversity by random mutagenesis and recombination. From among the thousands, millions or even billions of progeny, effective search engines such as genetic complementation and high-throughput screening identify candidates with improvements in a desired phenotype. Multiple iterations of this two-step process have led to the identification of numerous proteins with novel catalytic properties (Lutz and Patrick, 2004).

1.6 SCREENING AND SELECTION OF MUTANTS

Evolution is defined as the genotypic changes that occur in a population over time, and it is the result of mutation and selection. Natural selection is the process by which mutant genotypes become enriched in a population if they facilitate reproduction of the associated organism in a given environment. Thus, the environment provides the selection pressure that defines which members of a population are most fit. The experimentalist generates a pool, or library, of mutants and subjects them to conditions that favour the isolation or identification of individual members that possess the desired property (Holmberg *et al.*, 2005).

1.6.1 Degeneracy vs diversity

What is a 'library size?' Rather than giving a library size, it is much more informative to describe the library in terms of number of transformants, the fraction of the transformants that are 'meaningful' library members (as opposed to those members which, for example, did not receive the insert DNA) and the degeneracy of the library. The degeneracy of a library is the number of different members among the transformants (i.e. the number of independent clones). A distinction must therefore be made between degeneracy and diversity. Degeneracy describes the number of variants in a library, whereas diversity is a qualitative description of how much, on average, two randomly selected library members will differ. This point is emphasized by noting that two libraries can both have one million variants, and therefore the same degeneracy, but in one library the average difference between two randomly picked members may be three amino acids, whereas in the other library the average distance between two randomly picked members may be ten amino acids. Both libraries are equally degenerate but the latter is more diverse (Bosley and Ostermeier, 2005).

Thus, a directed evolution experiment begins with a library of mutant proteins. Mutations may be focused to a specific set of residues, or distributed throughout the entire sequence of the protein. Nucleotide mutations are introduced into the gene that

encodes the protein of interest by means of automated nucleic acid synthesis, chemical agents that covalently modify nucleotide bases, propagation in a bacterial mutator strain or error-prone PCR (Holmberg *et al.*, 2005).

1.6.2 Selection vs screening

One of the critical steps in directed evolution is the creation of a screen or selection that is rapid enough to process millions of mutants while accurately measuring a desired property (Zhao and Arnold, 1997). Screening by monitoring formation of the exact product of interest is desirable, but limits throughput. For this reason, the development of ingenious high-throughput methods is an active area of research (Bolon *et al.*, 2002).

Most selection schemes indirectly screen for catalysis, often through binding events. One of the difficulties of this is that the mutants that are discovered tend to bind more tightly to the substrates, but are not necessarily better catalysts (Bolon *et al.*, 2002). To overcome this limitation, catalytic turnover was coupled with fluorescence, so that high throughput could be achieved using fluorescence-activated cell sorter (Daugherty *et al.*, 2000). A fluorophore (F) and a quencher (Q) were tethered to a peptide substrate. Upon action of a protease, the substrate was cleaved, separating the F and Q moeities and disrupting the intra-molecular quenching. Fluorescence is thus linked to a turnover event. Because screening is often the limiting step, there have been several studies to optimize the mutation or recombination rate with respect to the number of mutants that can be screened (Voigt *et al.*, 2000).

A major challenge in directed evolution is to establish a screen or selection that is sensitive to the properties of interest (Arnold and Volkov, 1999). True selections in which only those clones carrying a desired improvement survive or grow faster are rare, usually highly specific to one problem and difficult to implement productively in directed evolution. When available and validated, however, they can deliver dramatic results (Martineau *et al.*, 1998). Genetic complementations, which provide a rapid way of identifying clones that retain at least some level of activity, have also been useful,

particularly for identifying the relatively few functional sequences in heavily mutated proteins (MacBeath *et al.,* 1998).

Selection is normally used for evolving intracellular enzymes, for which the required link between cell survival and enzyme activity is easier to implement (Arnold and Volkov, 1999). A study done by Naki *et al. (*1998) showed a simple way of linking cell growth to a secreted enzyme (i.e., subtilisin). Cells were grown in hollow fibres at an average density of one clone per fibre, using bovine serum albumin as the sole source of nitrogen. Cells secreting more subtilisin or more active enzyme should be able to obtain more nitrogen and grow faster. The hollow fibres provide tiny compartments to limit cross-feeding. One round of mutagenesis and enrichment produced a clone with five-fold increased protease production, and was found to arise from mutations outside the target gene (Arnold and Volkov, 1999).

Selection also requires that the target enzymatic activity does not interfere with cellular metabolism and that it can be distinguished from the background of all other cell reactions (Arnold and Volkov, 1999). *In vitro* translation systems may provide the basis for protein evolution in the absence of cells.

An interesting approach to selection for stabilized proteins links the protease resistance of the target protein to the infectivity of a filamentous phage (Sieber *et al*., 1998). The protein of interest is fused with the phage minor coat protein in such a way that the phage will infect cells only if the minor coat protein is intact and folded (Spada and Pluckthun, 1997). Selection of stabilized variants then consists of proteolytic treatment, infection and phage propagation. Application of one round of this selection to a library of RNase T1 mutants yielded several stabilized proteins. This method does not depend in any way on activity, and further screening of the stabilized variants would be necessary to retain the catalytic activity during evolution of thermostability (Arnold and Volkov, 1999).

A screen is required when the desired activity or feature cannot be linked easily to cell survival (Arnold and Volkov, 1999). Screens are more versatile than selections. The

throughput, however, is generally relatively low and inversely proportional to sensitivity to the desired property. A quantitative, colorimetric pH-based assay was developed for screening enantioselectivity of hydrolytic reactions (Janes *et al.*, 1998). Seventy-two commercial lipases and esterases were screened using this approach in order to identify the one with the best enantioselectivity toward solketal butyrate, an important building block in the synthesis of pharmaceuticals and biologically active compounds (Arnold and Volkov, 1999).

Thus, further advances in screening and selection technologies will reduce the time and cost of the experiments and will make it possible to tackle ever more difficult problems involving multiple enzymes, multi-component enzymes and the creation of new functional molecules (Arnold and Volkov, 1999).

1.7 TAILORING FUNDAMENTAL PROPERTIES OF ENZYMES

1.7.1 Thermostability

Interest in thermostable enzymes has increased dramatically as resistance to thermal inactivation has become a desirable property of the enzymes that are used in many industrial applications. Thermostable enzymes are generally defined as those with an optimum temperature above that of the maximum for growth of an organism or with exceptional stability above 50°C over an extended period of time (Gupta and Gupta, 1993)

Stability is an important parameter, which co-determines the economic feasibility of applying an enzyme in an industrial process. High stability is generally considered an economic advantage because of reduced enzyme turnover. In addition, stable enzymes permit the use of high process temperatures, which may have beneficial effects on reaction rates, reactant solubility and the risk of microbial contamination. Enzymes are increasingly used in feed where they may need to withstand hygienic heat treatments and harsh processes such as extrusion (Eijsink *et al.*, 2004).

There are now several examples of proteins, which have been stabilized by the introduction of numerous mutations with cumulative, small, stabilizing effects (Giver *et al.*, 1998). This may be taken to suggest that it is useful to use sequence statistics or statistical structural comparisons to discover 'rules' for protein stability. However, a clear conclusion to be drawn from other recent work is that very large stability differences in some cases are due to only one or very few point mutations (Williams *et al*., 1999). These latter observations indicate that statistical comparisons of the sequences and structures of proteins with varying stabilities may not be that useful.

In addition, the introduction of disulfide bonds, chemical cross-links and salt bridges has been widely used to increase stability, although not all disulfide bonds increase stability. Recent work by Clarke *et al. (*1995) on barnase by x-ray crystallography and hydrogen exchange shows that those disulfide bonds that do not improve stability disrupt the local structure. By comparing a number of glyceraldehyde phosphate dehydrogenase structures from organisms with different heat tolerance, they found a strong correlation between increased thermostability and the number of hydrogen bonds involving charged side-chains and neutral partners. On the other hand, Waldburger *et al*. (1995) found that for the Arc protein, a buried salt bridge between arginine and glutamic acid provided less stability than hydrophobic residues. Although not contradictory to Tanner *et al*. (1996) it does suggest that the contribution of salt bridges to stability remains unresolved (Rubingh, 1997).

Stephens *et al.* (2007) obtained three variants with enhanced thermostability using directed evolution. One of these variants 2B7-10, exhibited a five fold higher residual activity than the wild type XynA. It retained 71% of its activity after treatment at 80°C for 60 minutes and had a half-life of 215 minutes at 70°C, which is higher than that attained by XynA.

1.7.2 Enzyme activity

Although improving the activity of an industrial enzyme is often a primary goal, it is also one of the most complex. This is partly because in many applications of enzymes the substrate is chemically complex and heterogenous. In a revealing study on the proteolysis of variants of a single chain monellin, the extent of proteolysis at a fixed time correlated with the free energy of protein unfolding, suggesting that in this case substrate unfolding may be rate limiting (Iijima and Moromoto, 1995). Many industrial enzymes, such as cellulases, amylases, lipases and even proteases in certain contexts, work on insoluble substrates. In this context, the rate of substrate turnover may be diffusion limited and controlled by enzyme mobility at the surface or by on/off enzyme desorption rates. These in turn are frequently related to the surface properties of the enzyme and conditions at the interface between enzyme and substrate (Brode *et al*., 1996).

1.7.3 Surface properties

Some enzymes, such as cellulases and xylanases, possess binding domains that aid in adsorption of the enzyme. Lipases are generally not active until adsorbed at an oil-water interface. These studies support the importance of charge interactions in adsorption. A number of protein engineering studies have focused on changing charges to affect surface properties (Brode *et al*., 1996).

An important issue, from both an industrial and academic view point, is the extent to which the stability, activity and surface properties are linked. Thus, engineering an enzyme for improved activity/stability might involve understanding something about the inter-relations of all three properties mentioned above (Rubingh, 1997).

1.8 CHOICE OF A SUITABLE PRODUCTION HOST

Recombinant DNA techniques used for obtaining and combining genes from a variety of sources and the possibility of expressing these genes in different host cells, have provided scientists in the biotechnology field with proteins in quantities previously impossible to obtain. *E. coli* has however, also been successfully used for the production of relatively complex proteins and progress over recent years has widened the use of this organism even further (Jonasson *et al*., 2002).

As a production host, *E. coli* has mainly been used for cost-efficient production of large amounts of proteins that are limited in size and have a relatively simple structure. Genetic-design approaches may be applied to influence the targeting of the gene product, i.e. whether the gene product is accumulated intra-cellularly in a soluble form or as inclusion bodies, or whether the product is secreted into the periplasm or even into the culture medium. The choice of host system, therefore, depends on many factors, such as, the size, structure and stability of the gene product and the requirements for post-translational modifications for biological activity. The necessary production yields, acceptable cost and quality specifications of the final product also have to be considered (Jonasson *et al*., 2002).

1.8.1 The expression vector

An *E. coli* expression vector should contain, apart from the gene of interest, an origin of replication, a gene that confers antibiotic resistance (or an alternative selectable marker), a promoter and a transcription terminator (Jonasson *et al*., 2002). The origin of replication determines the vector copy number, which could typically be in the range of 25-50 copies per cell if the expression vector is derived from the low-copy-number plasmid or between 150 and 200 copies per cell if derived from the high-copy-number plasmid. The copy number influences the plasmid stability, i.e., the maintenance of the plasmid within the cells during cell division. The gene coding for antibiotic resistance is necessary both for identifying transformants and to ensure antibiotic selective pressure,

that is, only cells that harbour an expression vector will divide, thus preventing plasmid loss (Jonasson *et al.,* 2002).

A number of strong promoters are available for high-level expression in *E. coli*. An important criterion of a promoter is its ability to be efficiently down-regulated under non-induced conditions. An early overproduction of the heterologous protein, due to a non-silent promoter, might impair cell growth. It is therefore desirable to be able to repress the promoter during a cell growth phase to achieve high cell densities, after which the high-rate protein production would be initiated by induction of the promoter. For laboratory-scale production, the isopropyl β-D-thiogalactopyranoside (IPTG) – inducible promoters, which are regulated by the product of the *lac* repressor, are widely used (Jonasson *et al*., 2002).

1.9 SCOPE OF THE PRESENT STUDY

Xylanases are hydrolases depolymerizing the plant cell wall component xylan, the second most abundant polysaccharide. The molecular structure and hydrolytic pattern of xylanases have been reported extensively. The use of these enzymes in various industries requires a complete understanding of their functional and genetic significance. Thus, while conventional methods will continue to be applied to enzyme production, the application of recombinant DNA techniques is beginning to reveal important information on the molecular basis of fungal enzyme production and this knowledge is now being applied both in the laboratory and commercially. There are several different techniques of directed evolution, a strategy of which is adopted in this study and could add to the wealth of knowledge. This study investigates the use of error-prone PCR to evolve a fungal xylanase.

In a previous study carried out by Stephens *et al*. (2007) a mutant with enhanced thermostability and activity was obtained. In this study, this mutant was used as a parent to generate progenies with improved characteristics. While conformational stability does

not provide a guarantee for activity it certainly is a prerequisite. According to Eijsink *et al*. (2004) the route to enzymes that are highly active at high temperatures always needs to start with stability engineering.

It is hypothesized that the xylanase gene from *T. lanuginosus* DSM 5826 could withstand numerous mutations without disrupting its actual conformation, and in doing so, the thermostability and/or activity of the enzyme could be improved. Hence, the objectives of this study were *viz.*,

i) To identify a workable evolution strategy,

ii) To functionally express the enzyme in a suitable microbial host which in this study was *E. coli*,

iii) To screen and select xylanase-producing mutants exhibiting improved thermostability/ activity,

iv) To verify that the thermostability of the enzyme variants were greater than those from the parental strains and

v) To sequence and analyse the xylanase gene from each of the mutants.

CHAPTER TWO

DIRECTED EVOLUTION OF XYLANASE VARIANTS

2.1 INTRODUCTION

The successful implementation of enzymes as industrial biocatalysts requires the availability of suitable enzymes with high activity, specificity and stability under process conditions. However, naturally occurring enzymes are often not optimized to fulfill these requirements and they must often be mutated to improve their performance (Cherry and Fidantsef, 2003).

Rational protein design is often not useful either, as the problems are complex and multifactorial, and there is seldom sufficient knowledge of the structure and/or function (Chen, 2001). In rational redesign, precise changes in amino acid sequence are preconceived, based on a detailed knowledge of protein structure and function and then introduced using site-directed mutagenesis. Numerous attempts at redesigning enzymes using this technique have failed (Chen, 2001). These failures resulted, to some extent, from an incomplete understanding of the underlying mechanisms required to enhance the desired enzyme properties (Chen, 2001).

Within this context, directed evolution is very effective in closing these functional discrepancies (Zhao *et al*., 2002). Directed evolution, mimicking the process of natural evolution, arguably offers the fastest and most effective means of creating improved mutants for industrial application (Cherry and Fidantsef, 2003). It generally begins with the creation of a library of mutated genes. Gene products that show improvement with respect to the desired property or set of properties are identified by selection or screening and the gene(s) encoding those enzymes are subjected to further cycles of mutation and screening in order to accumulate beneficial mutations. This evolution can involve few or

many generations, depending on how far one wish to progress and the effects of mutations observed in each generation (Kuchner and Arnold, 1997).

With recent advances in directed evolution even more choices are available for creating mutant libraries *viz.*, error-prone polymerase chain reaction (epPCR), combinatorial oligonucleotide mutagenesis, DNA shuffling, staggered extension process (StEP) recombination and *in vitro* recombination (Arnold and Georgiou, 2003; Neylon, 2004).

The first objective of this study was to identify a workable evolution strategy that would mutate the xylanase gene. This was achieved using epPCR, which was chosen as the directed evolution strategy. Random mutagenesis involved a modified PCR protocol that used variations in $MnCl_2$ and nucleotide base concentrations. By varying the concentration of $MnCl_2$ and nucleotides the *Taq* polymerase makes errors during the PCR reaction (Jaeger and Reetz, 2000).

The second objective was to clone the mutated gene library into a suitable microbial host. Cloning of the *xynA* gene library provided the basis for screening the library for mutants displaying desirable characteristics. The third objective was therefore to screen for improved xylanase variants. One of the critical steps in directed evolution is the creation of a screen or selection that is rapid enough to process millions of mutants while accurately measuring a desired property (Bolon *et al*., 2002).

Screening requires the evaluation of each member of a gene library in a laborious process of analysis. The mutants created in this investigation were chosen on the basis of their ability to hydrolyze xylan. Bacterial colonies producing functional xylanase were detected by a zone of clearing on the RBB-xylan plate which was used as an index of xylanolytic activity. Analysing the sequences of all the variants and the parent, pX3 was the final objective in this chapter. A multiple sequence alignment was performed with the *xynA* gene from the wild-type and all of the recombinants used and constructed in this study. The degree of homology between recombinants and wild-type *xynA* was compared to determine the positions of the amino acid substitutions in relation to the

overall structure of the xylanase gene. The properties of the enzyme variants were then tested and are highlighted in Chapter 3.

2.2 MATERIALS AND METHODS

2.2.1 Growth of plasmids and recombinant strains

The parent strain harbouring the plasmid pX3 was kindly provided by Prof. W. Steiner of the Technical University of Graz, Austria. This plasmid contained the *xynA* gene from *T. lanuginosus* DSM 5826 and was cloned and expressed in *E. coli* as a LacZ-fusion protein. The vector backbone which was plasmid Bluescript (pBSK), consisted of the β-lactamase (*bla*) gene which conferred ampicillin resistance to it. Plasmid Bluescript did not contain the *xynA* gene and served as the control for this study. pBSK and *E. coli* XL1 Blue MRF´ wee obtained from Stratagene. A variant 2B7-10, obtained by mutation of *xynA* in a previous study (Stephens *et al*., 2007) was used as starting material for the current study. All the recombinant strains as well as *E. coli* pX3 were grown on Luria Bertani (LB) agar (10 g/l bactopeptone, 5 g/l yeast extract, 5 g/l sodium chloride, 15 g/l agar) supplemented with 100 µg/ml ampicillin where necessary. Strains were incubated at 37°C overnight. All strains were subcultured once a month and stored at 4°C. Long term preservation was carried out by storage at -70˚C, with the addition of 15% glycerol as a cryoprotective agent to broth cultures.

2.2.2 Plasmid DNA isolation

In this study, the alkaline lysis technique of Birnboim and Doly (1979) was modified to provide adequate-quality plasmid DNA, in less than an hour. For specialized applications, where DNA-yield and DNA-quality was critical, the High Pure Plasmid DNA Isolation kit (Roche Molecular Biochemicals) was used, according to the manufacturer's instructions.

In accordance to the modified method, the cultures were grown in 5 ml LB medium containing 100 µg/ml ampicillin for 16 hours at 37˚C. Cells were harvested by centrifugation at 13 000 *g* for 1 minute. The pellets were resuspended in 100 µl of

Solution A (25 mM Tris-HCl, 50 mM glucose, 10 mM EDTA, pH 8). The cell suspension was lysed by the addition of 200 µl of freshly prepared Solution B (2 N NaOH, 10% SDS). The micro-centrifuge tube was immediately inverted repeatedly but gently, until the solution turned clear and became viscous. The cell membranes were disrupted by the anionic detergent and the chromosomal DNA was denatured by NaOH. One hundred and fifty microlitres of cold Solution C (3 M sodium acetate, pH 8) was added and gently mixed. The tube was placed on ice for 5 minutes and then centrifuged at 13 000 *g* for 15 minutes. Two volumes of cold absolute ethanol was added and incubated for 15 minutes at -70˚C. The precipitated DNA was recovered by centrifugation at 13 000 *g* for 15 minutes. The DNA pellet was washed once with 70% ethanol and centrifuged for a further 3 minutes. The pellet was air-dried and dissolved in 20 µl of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). The DNA was stored at -20˚C.

2.2.3 Determination of DNA

The purity and concentration of DNA was measured at both $OD_{260\ nm}$ and $OD_{280\ nm}$. Ideally pure DNA should have a 260/280 ratio of 1.8 and 2. DNA concentration was calculated on the premise that an absorbance of 1 at $OD_{260\ nm}$ corresponds to 50 ng DNA/µl.

2.2.4 Agarose gel electrophoresis

A 0.8% agarose gel was run in a 1X TAE buffer (242 g Tris-HCl, 57.1 ml acetic acid, 100 ml of 0.5 M EDTA, pH 8) at a constant voltage of 95 V. DNA samples were mixed with 0.2-0.3 volumes of sample buffer (0.0375 g bromophenol blue, 4 g sucrose, 1.5 ml 10% SDS, 3 ml of 0.5 M EDTA dissolved in 15 ml distilled water) before loading onto the gel. Phage λ DNA was restricted with *Eco*RI and *Hin*dIII to produce DNA fragments

of known molecular weights and this served as the standard marker (see section 2.2.5) for all agarose gels used in this study. Gels were stained with ethidium bromide (50 μg/ml) for 20 minutes, destained with distilled water and viewed on a UV transilluminator. Gel images were captured using a Scion digital imaging system.

2.2.5 Preparation of DNA molecular weight marker III

The molecular weight marker was prepared by mixing the following components: λ DNA (35 μl), *Eco*RI (10 U), *Hin*d III (10 U), Enzyme Buffer B (10 μl) and sterile distilled water (47 μl). The mixture was incubated at 37˚C for an hour and then incubated at 65˚C for 10 minutes to inactivate the enzymes. Ten microlitres of sample buffer was then added and the mixture was stored at 4˚C until further use.

2.2.6 Error-prone PCR

Error-prone PCR involved a modified PCR protocol that uses variations in $MnCl_2$ and nucleotide base concentrations. Five different mutagenic conditions were used. The standard PCR served as a control for all epPCR conditions and contained all PCR materials in amounts that facilitated normal amplification (Table 2.1).

The PCR mixtures for all three rounds of mutation were prepared according to Table 2.1. The variant 2B7-10 obtained from a previous study by Stephens *et al*. (2007) served as the template for the first round of mutagenesis. The first round of epPCR was performed, PCR products were ligated into pBSK, transformed into *E. coli*, mutants were selected and the best mutant from the first round was used as the template for the second round of epPCR. This pattern was followed through to the third round of random mutagenesis. The concentration of template DNA (10 ng) and the primers (200 ng) were kept constant in all the PCR reactions, the total volume of which was 50 μl.

The choice of annealing temperature is based on the primer sequences. Choosing the correct annealing temperature ensures that nonspecific annealing will be minimized and precise complimentarity will occur. The following primers (Whitehead Scientific) were used for the PCR reactions:

Forward primer, T3 (5'- ATTAACCCTCACTAAAGGGA-3')
Reverse primer, T7 (5'- TAATACGACTCACTATAGGG-3')

The forward primer is responsible for synthesis of the DNA in the forward direction whilst the reverse primer synthesizes DNA on the opposite DNA strand, thus accomplishing synthesis of the double-stranded DNA molecule.

Table 2.1 Mutagenic conditions used for epPCR

Conditions	**dNTP [mM]**				**MnC_{l2}**
	A	G	T	C	**[mM]**
STD	0.1	0.1	0.1	0.1	-
1	0.04	0.04	0.2	0.2	0.5
2	0.04	0.04	0.2	0.2	0.2
3	0.04	0.04	0.2	0.2	0.1
4	0.04	0.04	0.2	0.2	_
5	0.1	0.04	0.1	0.1	_

The PCR programme consisted of denaturation for 1 minute at 95°C, annealing for 1 minute at 42°C followed by extension for 1.5 minutes at 72°C (Stephens *et al*., 2007). The PCR reaction comprised a total of 30 cycles and the reactions were performed using a PCR Genius (Techne) thermal cycler. To verify that the target DNA was amplified, 6 µl of the PCR product was run on an agarose gel. The PCR product was purified using the GFX PCR DNA and Gel Band Purification Kit (Amersham).

2.2.7 Preparation of insert DNA and vector

Cutting DNA at particular nucleotide sequences is accomplished by restriction endonucleases that recognize and cleave double-stranded DNA at specific sequences. If the vector is cut with the same endonuclease as the insert, the sticky ends of the insert complements the sticky ends of the vector, making ligation easier. Thus, both pBSK (which was used as a vector in this study) and the purified PCR products were restricted with type II endonucleases *viz*., *Eco*R1 and *Xho*1, to create compatible sticky ends for ligation.

Restriction digestion was carried out using type II restriction endonucleases. For a standard restriction, a digestion procedure comprising of: DNA (approximately 100-300 ng), 1 µl *Eco*RI (10 U), 1 µl *Xho*I (10 U) and 2 µl Buffer B was carried out in a total volume of 20 µl, containing sterile distilled water. Mixtures were incubated with the restriction enzymes at 37°C for 3 hours. Restriction enzymes and buffers were purchased from Roche Molecular Biochemicals. The recovery of insert DNA and vector DNA was accomplished by gel extraction, using a GFX PCR DNA and Gel Band Purification Kit (Amersham).

2.2.8 Ligation

After quantification of the digested and purified target gene insert and plasmid vector, the insert was ligated to the vector. The vector and epPCR products were ligated at a ratio of 1:3 vector to insert. Ligase mixtures containing the vector and insert DNA were heated to 65°C for 5 minutes. After rapid chilling on ice, 1 × Ligase buffer and 1 U of T4 DNA ligase (Roche Diagnostics) was added, to a final volume of 10 µl and was incubated overnight at room temperature. The resulting ligation products were then ready for transfer into a host bacterial cell, by transformation (Sambrook *et al*., 1989).

2.2.9 Preparation of competent cells

Host cells are rendered artificially 'competent' by exposing early log phase cells to bivalent cations. In this study, *E. coli* XL1 blue MRF' (Stratagene,) was used for cloning. One millilitre of an overnight culture of *E. coli* XL1 blue was diluted to 30 ml with LB broth. The culture was incubated at 37°C until an optical density of 0.375 at 590 nm was obtained. The cells were pelleted at 5000 *g* for 10 minutes and the supernatant was discarded. The cells were resuspended in 10 ml ice cold 100 mM $CaCl_2$, recentrifuged and resuspended again in 10 ml cold 100 mM $CaCl_2$. Cells were incubated on ice for 20 minutes and then centrifuged. Competent cells were then resuspended in 2 ml $CaCl_2$ (100 mM), containing 10% glycerol (Ausubel *et al*., 1989). One hundred microlitres of the prepared cells were dispensed into eppendorf tubes and stored at -70°C.

2.2.10 Transformation

The transformation technique was used to transfer the plasmid construct containing the *xynA* gene insert to a bacterial host cell for amplification of the plasmid and inserted gene. In order to transform bacterial cells, the cells must be made 'competent' to take up DNA. This is accomplished by weakening its cell membrane. The cells were incubated in $CaCl_2$ followed by a mild heat shock when the temperature is sharply increased from 4°C on ice to 42°C. The cells were then quickly chilled and incubated for gene expression.

One microlitre of the ligation reaction mixture was added to the competent cells on ice and mixed gently. The mixture was incubated on ice for a further 20 minutes. It was then heat shocked for two minutes at 42°C, returned immediately to ice and incubated for 5 minutes. One milliliter of LB broth (section 2.2.1) was added to the eppendorf tubes and incubated for an hour at 37°C. Thereafter the transformation mixture was plated and incubated at 37°C overnight, in order to recover a gene library (Sambrook *et al*., 1989).

2.2.11 Screening mutant libraries for xylanolytic activity

Coupling of Remazol Brilliant Blue to xylan was achieved by following the protocol of Biely *et al*. (1985; 1988). The dye (RBB) was linked to birchwood xylan (Roth) so that β-xylanase producers could be distinguished from non-β-xylanase producers by a zone of clearing around the colonies. Transformants were directly plated onto 0.4% RBB-xylan plates (0.4% RBB-xylan, 10 g/l bacto-peptone, 5 g/l yeast extract, 5 g/l sodium chloride, 15 g/l agar) supplemented with 100 μg/ml ampicillin. Plates were incubated at 37°C overnight and were subsequently screened for β-xylanase production. A zone of clearing was used as an index of xylanolytic activity. This procedure served as a selection method as well as a screening method for mutants. Only those cells that had taken up a plasmid were able to grow on the selective media containing ampicillin and it further screened for recombinant plasmids and distinguished it from self-ligated plasmids by producing a zone of clearing around clones producing xylanase. Positive transformants were identified as one producing a zone of clearing around it.

2.2.12 Protein sequencing and analysis of gene sequences

Automated sequencing of cloned mutants was performed at Inqaba Biotec, Pretoria using the Genetic analysis system SCE2410 (SpectruMedix LLC). The sequencing reactions were performed using the BigDye version 3.1 Dye Terminator Cycle Sequencing Kit (Applied Biosystems). All reactions were performed according to the manufacturer's instructions.

The sequencing data was then analysed using DNAmend 1.01 (Piet Jonas and Björn Maul, University of Greifswald, Germany). After editing and assembling contiguous DNA sequences of all *xynA* variants and the wild type pX3, their DNA sequences were then translated into protein. Alignments of protein sequences to homologs were performed using CLUSTALW program (version 1.82) on the GenomeNet server.

2.3 RESULTS

2.3.1 Error-prone PCR

Before performing epPCR, different mutagenic conditions were tested to assess which mutation would produce the best gene library. From the agarose gel it was observed that the intensity of the bands varied with each condition used. Five different mutagenic conditions were used and correspond with lanes 3 to lane 7. The band in lane 2 was the brightest and was amplified under standard PCR conditions with no mutagenesis introduced to the DNA. This explains its high intensity. The bands present in lanes 3, 4 and 5 represent conditions 1, 2 and 3 respectively (Table 2.1) which produced very low intensity bands and increased as the concentration of Mn^{2+} decreased. These low intensity bands were expected due to sub-optimal conditions. Moreover, the epPCR products that were obtained with conditions 1, 2 and 3 produced very few positive transformants and were the reason for eliminating these conditions from further analysis.

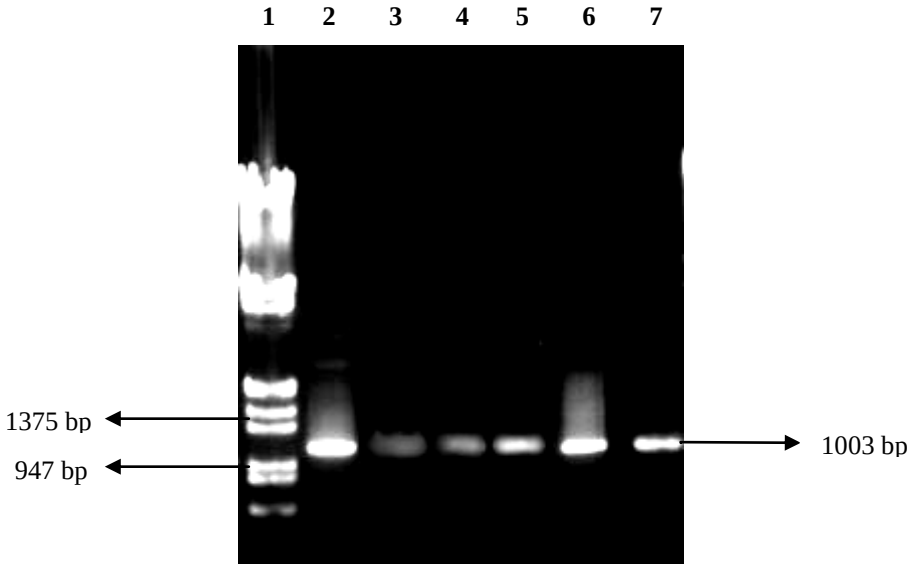


Fig. 2.1 Profiles of different mutagenic PCR conditions. Lane 1: DNA molecular weight marker III ; lane 2 (standard PCR conditions) ; lane 3 (0.5 mM Mn^{2+}) ; lane 4 (0.25 mM Mn^{2+}); lane 5 (0.1 mM Mn^{2+}) ; lane 6 (altered concentrations of dNTPs with no Mn^{2+}) ; lane 7 (altered a single dNTP (G) concentration without addition of Mn^{2+}) .

The band in lane 5 was visibly more intense than those in lanes 3 and 4, this was due to the fact that it contained the lowest concentration of Mn^{2+}. The brightest band was produced with condition 4 (lane 6) compared to those bands generated from the other mutagenic conditions (Fig 2.1). Also a full-length PCR product (1003 bp) was obtained under this condition. The band in lane 7 was just as intense as that found in lane 6 and this could be attributed to the fact that only one dNTP concentration that is dGTP, was altered. After having tested all the different mutagenic conditions, condition 4 which produced the highest amount of PCR products was used for subsequent rounds of mutagenesis.

EpPCR amplification of the three selected variants, with T3 and T7 primers produced a distinct banding pattern. It is evident from Fig. 2.2, that a single band of equal intensity and equivalent to the 1003 bp *xynA* gene was amplified in all of the reactions. This confirmed that all three mutants contained the *xynA* gene and was used as templates for epPCR. After different rounds of epPCR three promising mutants were identified *viz*., mutant 58, mutant 27 and mutant 47.

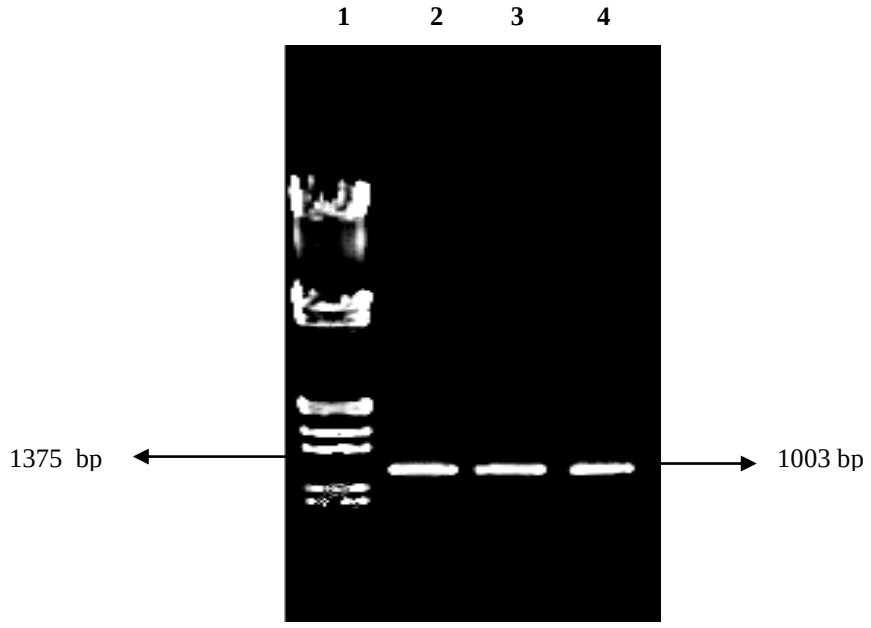


Fig. 2.2 PCR amplification of parental DNA for the three rounds of mutagenesis. Lane 1: DNA molecular weight marker III; lane 2: mutant 2B7-10; lane 3: mutant 58; lane 4: mutant 27.

Mutant 58 was selected after the first round of epPCR (using 2B7-10 as the template), and was used in the second round of epPCR as the template to obtain mutant 27 and this mutant was then used as the template in the third round of epPCR to obtain mutant 47. Mutants were selected on the basis of the fact that during thermostability assay they showed better thermostability compared to the parent. The mutant that was the most thermostable from each round of thermostability screening was used as the template for the subsequent round of mutagenesis.

2.3.2 Plasmid DNA isolation

Electrophoresis was carried out after performing plasmid isolation on all recombinant strains and the wild-type pX3 which revealed characteristic bands for all plasmids isolated (Fig. 2.3). Lanes 2-7 represent electrophoresis of uncleaved plasmid DNA.

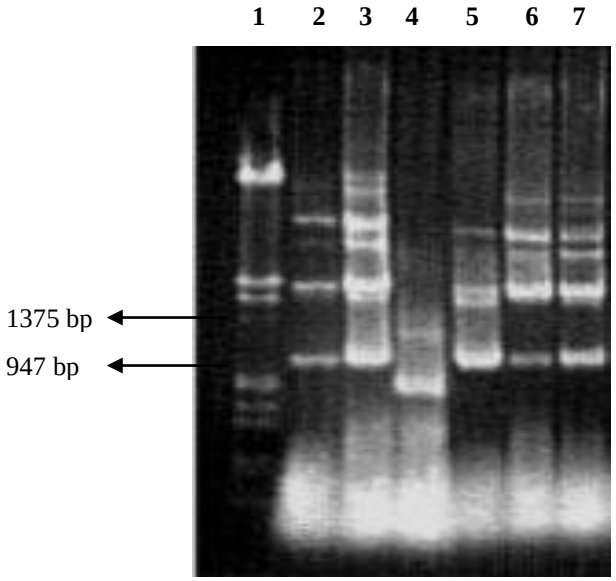


Fig. 2.3 Plasmid isolation using the alkaline lysis method. Lane 1: λ DNA molecular weight marker III; lane 2: mutant 2B7-10; lane 3: pX3; lane 4: pBSK; lane 5: mutant 58; lane 6: mutant 27 and lane 7: mutant 47.

Plasmid bluescript (pBSK) which does not possess the xylanase gene (lane 4) migrated the furthest because of its compact form and can be clearly seen (Fig. 2.3). Lanes 2 to 7 with the exception of lane 4 represents the three conformations of plasmids *viz*., superhelical, open circular and linear. However, since more DNA was loaded in some of these lanes, additional bands can be seen and these bands were brighter. These additional bands are multimers of the three plasmid conformations. The bands of pX3 and the bands found in lanes 5, 6 and 7 are identical since they are DNA of the same size. The smear at the bottom of the electrophoresis gel was contaminating RNA.

2.3.3 Restriction analysis of plasmids

Restriction profiles of all recombinant plasmids were verified against wild-type pX3 to ensure that the correct banding pattern was obtained. After restriction with *Eco*RI and *Xho*I two distinct fragments of molecular weights 2928 bp (vector) and 1003 bp (*xynA*) were produced (Fig. 2.4).

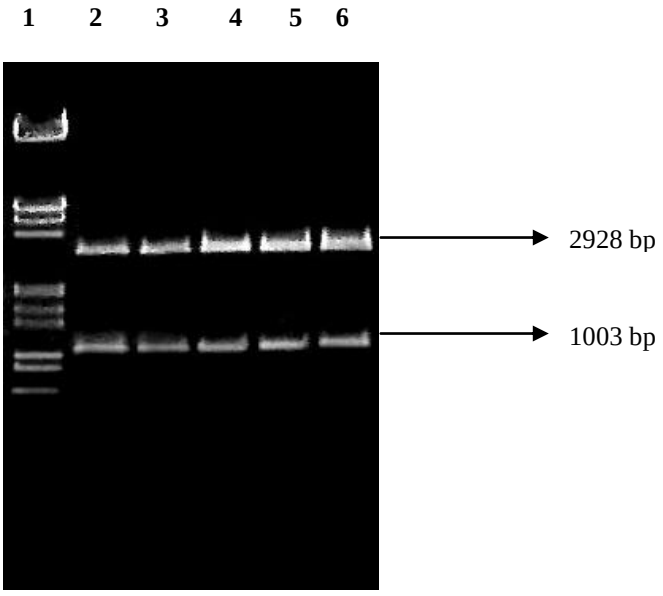


Fig. 2.4 Restriction of recombinant plasmids isolated from *E. coli*, showing the presence of the *xynA* **g**ene. Lane 1: λ DNA molecular weight marker III; lane 2: control pX3; lane 3: mutant 2B7-10; lane 4::mutant 58; lane 5: mutant 27 and lane 6: mutant 47.

All three variants *viz.*, 2B7-10, mutant 58 and mutant 27 were chosen as parents for each consecutive round of mutagenesis and were consistent in producing these fragments of known molecular weight. All of the variants contained the 1003 bp fragment, which confirms the presence of the insert.

2.3.4 Screening mutant libraries for β-xylanase production

Random mutagenesis conducted on the xylanase gene (*xynA*) of plasmid 2B7-10 produced a mutation library from which mutant forms of the wild-type xylanase were identified *viz.*, mutant 58 from the first round of screening, mutant 27 from the second round of screening and mutant 47 from the third round of screening. Mutants were identified by zones of clearing being formed around *E. coli* colonies that produce β-xylanase (Fig 2.5).

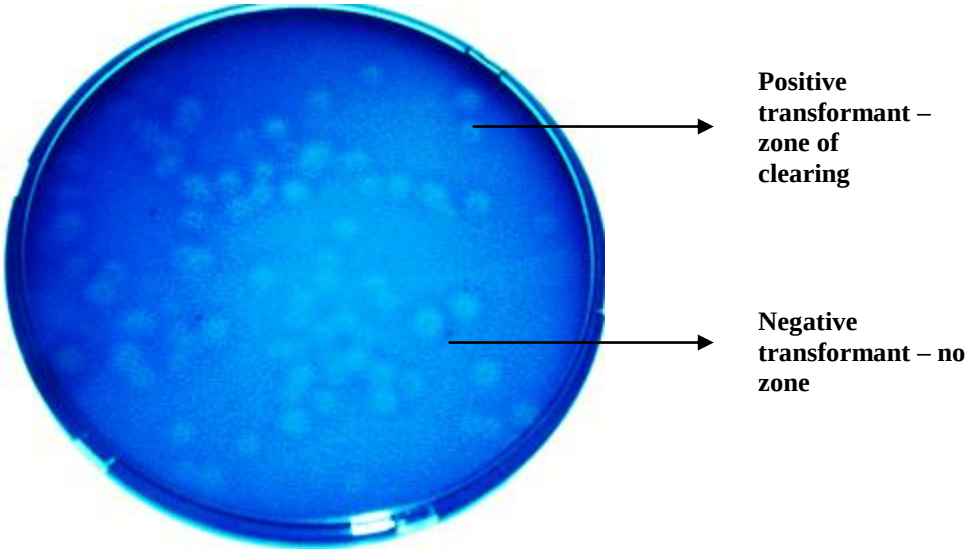


Fig. 2.5 RBB-xylan agar plate showing positive and negative transformants.

Xylanase hydrolyses RBB-xylan so that colourless xylooligosaccharides are produced, thereby making the identification process of transformants easier. Of the three rounds of epPCR that was carried out 73% of the colonies were positive transformants. In the first round of random mutagenesis, it was found that condition 4 (Table 2.1), which consisted of altered concentrations of dNTPs without the addition of Mn^{2+}, produced the largest number of positive transformants (approximately 200 colonies) compared to the other mutagenic conditions which produced positive transformants in the region of 20-90 colonies.

2.3.5 Protein sequence analysis

Three of the most promising mutants found in this study together with the parent mutant 2B7-10 (Stephens, 2007) and the wild type *xynA* were sequenced. The sequences of the remaining mutants that displayed poor thermostability were not analysed. After editing and translating their nucleotide bases to protein, their sequences were aligned to the protein sequence of *xynA*. The results of CLUSTALW similarity searches revealed that the degree of homology for mutant 2B7-10, mutant 58 (from the first round of epPCR) and mutant 27 (from the second round of epPCR) was 100%. This meant that the protein sequences of these mutants were identical to each other. When compared to the wild type *xynA* there was only one amino acid substitution (Fig. 2.6), which occurred at residue position 58 where tyrosine was substituted for phenylalanine. No other mutations have been observed for mutant 58 and mutant 27. On the basis of the fact that these mutants showed slight improvement in their thermostability compared to the parent it was assumed that these were actually enzyme variants when in fact they were not. Sequence analysis was therefore very important to establish whether a mutation occurred or not.

Mutant 47 however, which was obtained from the final round of epPCR had 98% similarity in comparison to the wild type *xynA*. As with all the other mutants, mutant 47 also had an amino acid substitution at residue position 58, whereby it changed from tyrosine (Y) to phenylalanine (F).

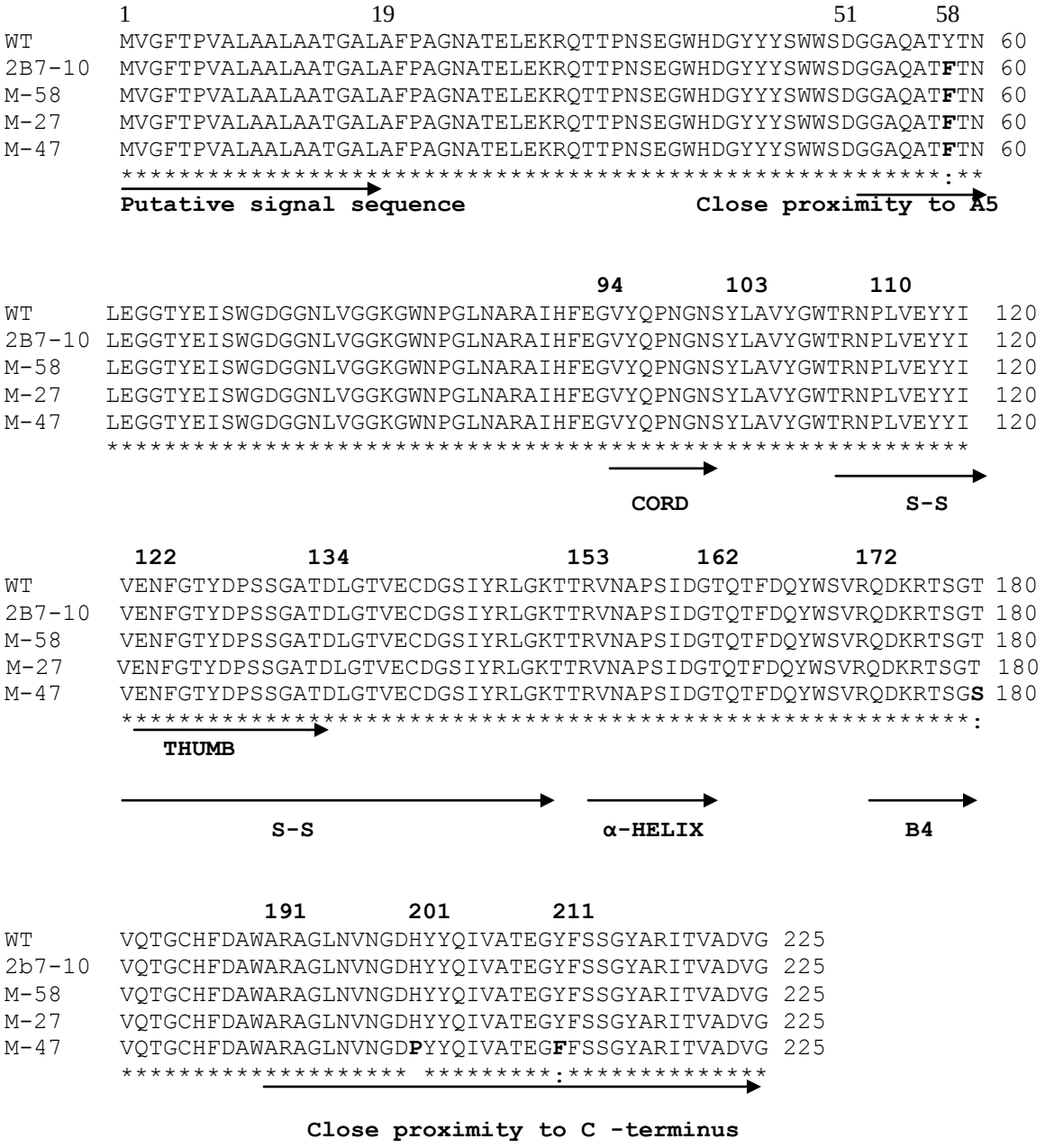


Fig. 2.6 Multiple protein sequence alignment of all recombinants used or constructed in this study. The mutants, 2B7-10, 58, 27 and 47 are aligned to the wild-type xyn*A*. Alignment sequence characters are indicated as follows: (*) indicates positions which have a single, fully conserved residue, (:) shows conservation within a strong group of amino acids. The absence of an alignment character implies that an unrelated amino acid was substituted. Amino acid substitutions that differ from the wild-type xyn*A* are indicated in bold face. Sequence alignment was done using CLUSTALW (version 1.82). For structural designations refer to Fig.1.2.

However, mutant 47 also had three additional mutations at positions 180, 201 and 211, respectively. At position 180, the amino acid changed from threonine (T) in the wild type to serine (S) in the mutant. At position 201, there was an unrelated amino acid substitution from histidine (H) in the wild type to proline (P) in mutant 47. There was conservation amongst a strong group of amino acids at residue 211 (Fig 2.6). At this position the amino acid tyrosine (Y) in the wild type was replaced with (F) phenylalanine. This occurred only in mutant 47. The regions between residue position 61 and residue position 179 were fully conserved among all mutant xylanases but the surrounding regions showed some variability (Fig. 2.6). Of interest is that for mutants 58 and 27 the mutations occurred in β-sheet A only. Whilst for mutant 47 only one mutation occurred in this region. The rest of the mutations occurred in close proximity to the C-terminus of the *xynA* structure and in beta sheet B.

2.4 DISCUSSION

An enzyme is defined by a unique sequence of amino acids, which in turn is dictated by the organism's DNA code (the gene) and assembled in the cell. This amino acid sequence determines how the chain folds, and ultimately how the enzyme functions. Thus, by modifying the amino acid sequence, we can alter the enzyme's function (Arnold, 1998). Enzyme sequences from mesophilic organisms have been compared to their homologs from thermophilic and hyperthermophilic microorganisms in an effort to identify the interactions responsible for conferring enhanced thermostability (Kinjo and Nishikawa, 2001). Numerous mutagenesis studies have also explored this phenomenon, however the mechanisms that mediate protein thermostability is still poorly understood (Farinas *et al.*, 2001).

Rules for engineering protein stability by rational design is protein specific, and any such design effort in combination with structure is both time-consuming and information intensive. A more simplistic approach to improving the thermal properties of an enzyme is directed evolution, which can contribute to the fast generation and identification of improved enzymes (Moore and Maranas, 2000). A key step in the directed evolution experimental cycle is the introduction of genetic diversity to the library. There are two basic ways for introducing diversity: epPCR and DNA recombination (Moore and Maranas, 2000). EpPCR was chosen as a method of preference for this enzyme evolution study.

It was hypothesized in this study that the thermostability and activity of fungal xylanase can be improved through directed evolution, using error-prone PCR. This was successfully achieved by introducing random mutations into protein-coding gene (*xynA*) by mutagenic PCR. Introducing these deliberate errors is crucial for successful epPCR (Moore and Maranas, 2000). Amplification with T3 (forward) and T7 (reverse) primers produced a distinct single band with a size of 1003 bp for all reactions tested (Fig. 2.2).

Since mutation is required for introducing gene diversity different mutagenic conditions were initially tested to determine which of these would provide a large gene library. As it is evident from the agarose gel, the intensity of the bands varied with each mutagenic condition used (Fig 2.1). A standard PCR was performed and used as a reference for all the mutagenic conditions that were tested.

After having tested condition 1, 2 and 3 (Fig 2.1) which produced very low intensity bands the transformants were plated out. It was found that there were very few positive transformants on the plate as a result of which, this condition was eliminated from further analysis. This decrease in the number of positive transformants yielded a decrease in the quantity of PCR products therefore the enzyme was inhibited to a large extent. If full-length products are produced but of a lower quantity, the yield of the insert after restriction with *Eco*R1 and *Xho*1 would also have been lower. This therefore, explains the low number of positive transformants. Secondly, those transformants that did grow under these conditions were probably hyper-mutated. This means that they had numerous mutations and were therefore not useful, as approximately one mutation per gene is normally desired. If too many mutations per gene occur it is difficult to predict the effect of each individual mutation.

In condition 5 only one dNTP (guanine) concentration was altered and although the negative transformants produced a full-length PCR product, the enzyme was not active when the transformants were plated out. There was no zone of clearing around the colonies on the RBB-xylan plates. In addition, the number of these negative transformants far outweighed the positive clones found on the RBB-xylan plates and were not considered for further analysis. This is indicative of the fact that slight changes in the PCR mixture may give rise to a large gene library but won't necessarily produce positive transformants. Rowe *et al*. (2003) also tested the efficiency of recursive random mutagenesis. The method that was used was identical to condition 5, in this study. They found that recursive random mutagenesis produced essentially asexual populations, within which beneficial mutations drove each other into extinction. A similar finding

was discovered in this study. Perhaps this clonal interference attributed to the fact that with this condition there were too many negative transformants.

The application of condition 4 of epPCR proved to be the best mutagenic condition as it generated the largest gene library from all of the conditions tested. By using altered concentration of dNTPs only and without the addition of Mn^{2+} (condition 4), the number of positive transformants found in the gene library were numerous. Of the several hundred colonies that were tested 73% of it was positive transformants. This can be expected since Mn^{2+} was not added hence, *Taq* polymerase would not be inhibited. The brightness of the band is also a good indication of the number of PCR products that can be obtained. By not adding Mn^{2+} to the PCR mix it offers the advantage of increasing the library size whilst still maintaining diversity amongst the encoded proteins. This condition was therefore chosen for the second and third rounds of mutagenesis as well.

Minipreparation of plasmid isolation (Fig. 2.3) was carried out for all isolates selected after epPCR in order to verify the presence and concentration of plasmid DNA. The miniprep procedure was adequate as small quantities of DNA were required for restriction digestion. After having performed restriction digestion on all the *xynA* variants, it was confirmed that all of the variants possessed the xylanase gene. This was verified by the agarose gel (Fig 2.4) showing the presence of the 1003 bp gene fragment. Wild type pX3 was used as a control in this procedure, to verify the presence of the vector and gene fragments in all of the recombinants. Restriction analysis was therefore an absolutely critical step in verifying the presence of the xylanase gene.

A major obstacle in the construction of random mutagenesis libraries occurs during the insertion of the gene fragment into the host vector. This typically relies upon a ligation step that requires considerable effort to optimize (Miyazaki and Takenouchi, 2002). Furthermore, many of the resulting plasmids contained no inserts or it contained multiple inserts. Attempts have been made to improve the efficiency and quality of ligations, though it is not clear if these improvements are broadly applicable or whether each new case requires optimization (Topcu, 2000).

An alternative approach is to eliminate the ligation step by the use of mutator strains such as *E. coli* XLI-red. These bacterial strains are deficient in DNA repair mechanisms and permit mutations throughout a plasmid of interest to accumulate (Greener *et al*., 1997). Such methods have been widely adapted to site-directed mutations (Miyakazi and Arnold, 1999) and more recently to create saturation site-directed random mutations (Braman, 2002) which both target a single codon for mutagenesis. These methods permit libraries to be reliably constructed within 3 days. Cloning and expression results in potentially active enzyme molecules being located either in the cell cytoplasm, in the periplasmic fluid or on the cell surface. This technique however, could not be used in this study as the mutations were performed randomly without targeting a specific codon.

A screening procedure was employed to isolate the numerous *E. coli* transformants containing sequences coding for the functional protein, xylanase. These clones represent a pool from which the best mutants had to be retrieved on the basis of detected enzyme activity. A critical phase to any directed evolution experiment therefore, is deciding how to search for variants with the desired properties. Only those recombinants that possessed a functional xylanase gene produced a zone of clearing which was used as an indication of xylanolytic activity (Fig. 2.5). This screening method therefore, helped to differentiate those transformants that had a functional xylanase gene from those that had no insert and did not produce a zone of clearing due to a non-functional xylanase.

The generation of large mutant libraries for *in vitro* enzyme evolution presents a challenge of effectively screening libraries of mutants on the basis of simultaneously assaying their biocatalytic activity (Cohen *et al*., 2001). A disadvantage of the screening method used in this study, was the fact that the enzyme variants had to be tested individually for the property of interest which was extremely time–consuming. Also, mutants generated by this screening method could not be assayed for thermostability simultaneously, as the enzymes were produced intracellularly and a cell lysis step was necessary before the mutants could be assayed for this property.

Perhaps alternative methods of screening should be looked at for future studies. Cohen *et al*. (2001) developed a new screening method for the 'one in a million' mutant in a single run. This method used immobilization of a single cell on a single bead followed by bead immobilization on a physical support. The method was coupled with digital imaging. A feasibility study was then carried out and it demonstrated the handling and screening of one million cell populations. Other methods that were investigated by Cohen *et al*. (2001) were single cells in multiwell plates and flow cytometry. However, these methods were not as successful as the immobilization method. Thus, incorporation of automated processing and digital imaging could lead to a substantial increase in the number of cells/colonies that could be screened in a short space of time.

While researchers can choose from a variety of methods to generate enzyme variants, it is still unclear which method is the best. The first law of directed evolution as proposed by Arnold (1998) 'you get what you screen for' stands true. Hence, the importance of establishing a screen or selection that reflects the desired result should not be under-estimated. One method of overcoming some of the barriers to implementing enzymes in industry is the application of directed evolution, in which the amino acid sequence of an enzyme is altered until the enzyme functions in a desired manner. The clones that were constructed in this investigation as well as the parent and the wild-type *xynA* were sequenced to determine whether the amino acid substitutions possibly contributed to an increase in themostability that was observed in all of the variants tested.

Multiple sequence alignment revealed the degree of similarity that is shared between all the recombinants and the wild type pX3. There was a common mutation for all the recombinants except the wild type at residue position 58 which is in close proximity to β-sheet A5. This sheet forms the "outer" surface of the enzyme and is hydrophilic in nature. A much lower similarity was shown with mutant 47 which had four mutational changes. The individual amino acid substitutions are reflected in Fig.2.6. Of interest is that at position 180 the amino acid threonine (T) was replaced with serine (S). This mutation occurred within the region of β-sheet B4, whilst it is sheet A that contains a large number of serine and threonine residues. At position 201, there was an unrelated

amino acid substitution *viz.*, from histidine (H) in the wild-type to proline (P) in mutant 47. This is an important finding as proline residues have been reported to contribute to improved thermostability (Gray *et.al*., 2001).

At residue position 211 the amino acid tyrosine (Y) was substituted with (F) phenylalanine (Fig. 2.6). This is noteworthy since an identical type of mutation occurred at residue position 58. Although the mutations were random, it seems that the first and the last mutations were identical. The fact that no mutations occurred in the region of the disulfide bridge could have conferred an increase in thermostability amongst the variants. Hence these factors were further investigated in Chapter 3.

Kumar *et al*. (2000) undertook a study whereby systematic differences among thermophilic and mesophilic proteins across the families were looked at. They observed that both thermophilic and mesophilic proteins have similar hydrophobicities, compactness, oligomeric states, polar and non-polar contribution to surface areas and main-chain and side chain hydrogen bonds. Insertions/ deletions and proline substitutions do not show consistent trends between the thermophilic and mesophilic members of the families. Additionally, comparisons of the sequences of the thermophile-mesophile homologous protein pairs indicate that arginine and tyrosine are significantly more frequent, while cysteine and serine are less frequent in thermophilic proteins.

Clones from mutant libraries were selected and screened until proteins with the desired properties were found. These clones then served as a starting point for iterative rounds of mutagenesis. EpPCR and subsequent screening yielded several mutants after each round, displaying xylanolytic activity. The findings of this study was in keeping with the main requirements of directed evolution *viz*., the functional expression of the enzyme in a suitable microbial host, screening and selecting recombinants that are sensitive to the desired property, iterative rounds of random mutagenesis and sequencing and analyzing the xylanase gene of each recombinant. All of these requirements were successfully achieved. Further research therefore focused on the evaluation of the thermostability of the progeny mutants in comparison to the parent, and is outlined in Chapter 3.

CHAPTER THREE

BIOCHEMICAL ANALYSIS OF THE XYLANASE FROM RECOMBINANT STRAINS

3.1 INTRODUCTION

In nature, evolution and creation of new functionalities is achieved by mutagenesis, recombination and natural selection of beneficial traits. Directed evolution strategies mimic this. Changing a protein by directed evolution opens the possibilities for screening for a variety of desirable properties (Hibbert *et al*., 2005). The most popular targets for directed enzyme evolution to date have been enzyme activity, substrate specificity, thermal stability, enantio-selectivity or enantio-specificity, pH range and tolerance to solvents. In many cases, directed evolution of an enzyme has enabled biocatalysis to reduce the number of steps required in an equivalent chemical synthesis (Hibbert *et al*., 2005).

Depending on the choice of selection and screening methods, enzyme stability and activity could be simultaneously improved or one property is improved at the cost of another (Zhao *et al*., 2002). Six generations of random mutagenesis, recombination and screening stabilized *Bacillus subtilis p*-nitrobenzyl esterase without comprising its catalytic activity at lower temperatures (Giver *et al*., 1998). By contrast, the evolved variants of a thermophilic indole-glycerol phosphate synthase exhibited increased catalytic values but decreased thermostability (Merz *et al*., 2000). Using a similar strategy, Suzuki *et al*. (2001) were able to obtain four variants of a thermophilic 3-isopropylmalate dehydrogenase with enhanced specific activities at low temperatures. Two of these variants maintained the wild type thermostability, whereas the other two variants exhibited decreased thermostability. Clearly, with no direct selective pressure on both properties, either activity or thermostability could drift randomly in the fitness landscape (Suzuki *et al*., 2001).

According to Giver *et al*. (1998) analysis of the stabilities and activities of large numbers of random mutants indicated that these properties are not inversely correlated. Mutations that increase thermostability while maintaining activity are very rare. Unless both properties are constrained (by natural selection or screening) the evolution of one by the accumulation of single amino acid substitutions typically comes at the expense of the other, regardless of whether the two properties are inversely correlated or not correlated at all (Giver *et al*., 1998).

Thermostable enzymes are gaining wide industrial and biotechnological interest due to the fact that their enzymes are better suited for harsh industrial processes (Zeikus *et al*., 1998). In recent years, the interest in xylanases has markedly increased due to its potential industrial uses, particularly in the pulp and paper industry. The wood used for the production of the pulp is treated at high temperatures, which implies that the enzymatic procedures require proteins exhibiting a high thermostability. Treatment with xylanase at elevated temperatures disrupts the cell wall structure and facilitates lignin removal in the various stages of bleaching (Jacques *et al*., 2000).

These enzymes have offered a major advance in the reduction of chlorine consumption in the bleaching process of kraft pulp, thus lowering environmental pollution by organic halogens (Erikson and Heitmann, 1998). The pulp and paper technology is one of the fastest growing industries and the use of the thermostable xylanases seems attractive since they provide global environmental benefits. However, scaling up of the enzyme to the levels required by industry remains to be seen (Erikson and Heitmann, 1998).

Xylanases for such a purpose must: (i) lack cellulolytic activity to avoid hydrolysis of the cellulose fibres; (ii) need to be of low molecular mass to facilitate their diffusion in the pulp fibres and (iii) high yields of the enzyme must be obtained at a very low cost (Niehaus *et al*., 1999). According to these authors, all commercially available xylanases can only partially fulfill these requirements, and the optimum temperature for the activity of most xylanases is reported to be between 50-60°C (Jacques *et al*., 2000). However,

some xylanases have been reported to exhibit higher thermostability and optimal activity ranging from 80 to 100°C (Morris *et al*., 1998).

Xylanases also have potential application in the food and feed industry, textile processes, the enzymatic saccharification of lignocellulosic materials and waste treatment (Castro *et al*., 1997). However, most of these processes are carried out at high temperatures, so that thermostable enzymes would certainly be an advantage. One extremely valuable advantage of conducting biotechnological processes at elevated temperatures is reducing the risk of contamination by common mesophiles (Haki and Rakshit, 2003). Other values of elevated process temperatures include higher reaction rates due to a decrease in viscosity and an increase in diffusion co-efficient of substrates and higher process yield due to increased solubility of substrates and products (Kumar *et al*., 2000). Such enzymes can also be used as models for the understanding of thermostability and thermo-activity, which is useful for protein engineering. The search for a thermophile with high yield of enzyme and the desired characteristics is still being pursued, hence the reason for this study.

Thermostable xylanases have been isolated from a number of bacterial and fungal strains. Members of the *Bacillus* species, *Streptomyces* species, *Thermoascus aurantiacus* and *Fusarium proliferatum* have been reported to produce xylanases which are active at temperatures between 50 and 80°C (Haki and Rakshit, 2003). While the *Dictyoglomus* sp. was described to produce xylanases operating at an optimum temperature of 90°C, a number of *Thermotogales* sp. were reported to secrete thermostable xylanases which can function at higher temperatures (Haki and Rakshit, 2003). Cellulase-free xylanases with an optimum temperature of 65°C from *Thermoactinomyces thalophilus* subgroup C (Kohilu *et al*., 2001) and *Clostridium abusonum* CFR-702 (Swaroopa and Krishna, 2000) were also reported. It was for these reasons that a temperature range of 80- 90°C was chosen, to screen for thermostability of the recombinant strains.

In Stephens *et al*., (2007) an enzyme variant 2B7-10 was obtained after two generations of error-prone PCR. This variant retained 71% of its activity after treatment at 80°C for

60 minutes and it had a half-life of 215 minutes at 70°C, which is higher than that attained by XynA. This variant showed increased activity and thermostability.

Many enzymes stabilized by protein engineering do have higher temperature optima, but there are now several examples in the literature of proteins showing an increase in stability without showing an increase in temperature optima (Bjork *et al*., 2003). Thus, some enzymes seem to have genuine temperature optima that are not always dictated by conformational stability (Daniel *et al*., 2001). Biocatalyst stability is a major concern in virtually all bioprocesses, because it may affect process economics at a number of levels. Poor biocatalyst stability will result in longer process operations (resulting from decreased catalytic efficiency), increased frequency of catalyst replacement and reduced product yields (Burton *et al*., 2002).

Thermostable enzymes are stable and active at temperatures which are even higher than the optimum temperatures for the growth of the microorganisms (Saboto *et al*., 1999). The relationship between thermal stability and the temperature optimum for activity is quite interesting. Relatively few studies have been carried out in this regard and the only available information is that thermophilic enzymes are more rigid proteins than their mesophilic counterparts.

The research reported in this chapter therefore focused on assessing temperature optima, activity and thermostability of the best enzyme variants obtained from each round of mutagenesis, the parent strain 2B7-10, as well as the wild type pX3 and the enzyme from *T. lanuginosus* DSM 5826.

3.2 MATERIALS AND METHODS

3.2.1 Growth of *Thermomyces lanuginosus* and recombinant strains

The fungal culture *T. lanuginosus* was obtained from the German Type Culture Collection (DSM 5826). *T. lanuginosus* DSM 5826 was grown on potato dextrose agar (PDA, Oxoid) at 50°C for 7 days. Plates were stored at 4°C until further use. For long term preservation, fungal cultures were freeze dried. All recombinant strains were cultured and maintained as outlined in section 2.2.1.

3.2.2 Induction of xylanase

A loopful of culture of all recombinants were streaked for single colonies on LB-ampicillin plates and incubated overnight at 37°C. Thereafter, 5 ml LB-ampicillin broth was inoculated with a single colony of the desired strain. A tube with medium only was retained as a control. The tubes were incubated at 37°C overnight, in a shaking incubator. One hundred microlitres of the overnight cultures was used to inoculate 500 ml flasks containing 100 ml LB-ampicillin medium. Duplicate flasks were incubated in a shaking incubator at 37°C until the OD_{600} of all the cultures were 0.5 absorbance units. Flasks were then inoculated with 0.1 mM IPTG and re-incubated at 37°C overnight after which the samples were extracted.

3.2.3 Enzyme extraction

Bacteria were concentrated by centrifugation at 5 000 × *g* for 5 minutes. The pellet was resuspended in 5 ml breaking buffer [6.8 g/l KH_2PO_4, 0.61 g/l $MgCl_2.6H_2O$, 0.77 g/l dithiothreitol, 0.37 g/l EDTA, pH 6.8, 0.1 g/l phenylmethylsulfonylfluoride (PMSF)] and immediately placed on ice. Whilst on ice, samples were sonicated at 5 000 Hz for 10 seconds per cycle with a total of three cycles per sample using a VirSonic sonicator (Virtis). Samples were then centrifuged at 10 000 × *g* for 45 minutes at 4°C. The supernatants were transferred to new eppendorf tubes and kept at 4°C for further analysis.

T. lanuginosus (DSM 5826) was grown in duplicate flasks containing 100 ml of medium (pH 6.5) comprising of: beechwood xylan (3.12 g), yeast extract (3.02 g) and potassium dihydrogen phosphate (0.5 g) as described by Purkarthofer *et al*., (1993). Flasks were inoculated with an agar block (1-2 cm^2) from a plate culture of *T. lanuginosus* and incubated at 50°C in a shaking incubator for 5 days. The culture medium was then centrifuged at 7 000 × *g* for 10 minutes followed by filtering of the supernatant using filter paper (Whatman no.1) and kept at 4°C until further analysis.

3.2.4 Determination of xylanase activity

Xylanase activity was assayed according to Bailey *et al*. (1992) by incubating the appropriately diluted enzyme solution at pH 6.5 and a temperature of 50˚C for 5 minutes by using a substrate solution of 1% (w/v) birchwood xylan (Roth). Reducing sugars were assayed by adding 3 ml DNS (2-hydroxy-3, 5 dinitrosalicylic acid) reagent, boiling for 5 minutes cooling and reading the absorbance at 540 nm. One unit of xylanase activity was defined as the release of one nmol of product per second.

3.2.5 Temperature optima and thermostability

Temperature optima of the variants was determined by diluting (1:2) the respective supernatants with buffer (pH 6.5) followed by incubation at temperatures from 50°C to 90°C at 10°C increments. Since all of the variants, wild type and the fungal enzyme were optimal between 70°C to 76°C, temperature optima was therefore further evaluated at 2°C increments between 70°C to 80°C. Samples were removed every 5 minutes and assayed for residual xylanase activity. For determination of thermostability of xylanase, crude enzyme preparations were diluted in 50 mM citrate buffer at pH 6.5 to reduce the concentrations of medium components which could influence the enzyme stability. The buffered enzyme solutions were incubated in a water bath at 80°C and 90°C, respectively. Samples were drawn at 15 minute intervals, immediately cooled on ice and the residual xylanase activity was determined (section 3.2.4).

3.3 RESULTS

Thermostable enzymes are generally defined as those with an optimum temperature above that of the maximum for growth of an organism or with exceptional stability above 50°C over an extended period of time (Gupta and Gupta, 1993). Before assessing the thermostability of the various recombinants obtained in this study, the temperature optima of these recombinants, the wild type and the fungal strain was determined.

The temperature optimum for xylanase activity was 70°C for all the recombinant strains, the wild type and the fungal strain. This was much higher than their respective growth temperature optima *viz*., for the bacterial strains it was 37°C whilst for the fungal strain it was 50°C. Since the temperature optima was identical for all the strains it was decided to further test their optima at two degree increments ranging from 70°C to 76°C. The temperature optima for all the recombinant strains tested were 72°C whilst the optima for the wild type pX3 was 70°C and for the fungal strain it was 76°C (Fig. 3.1).

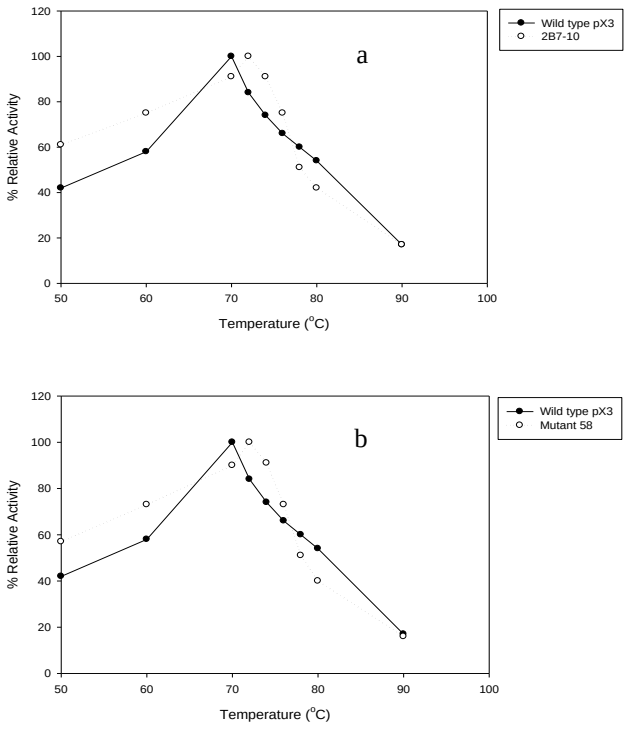
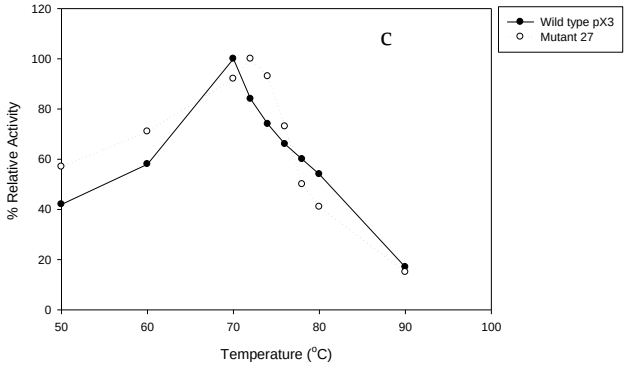
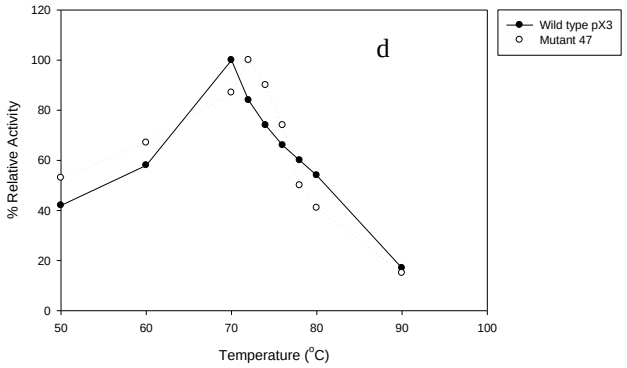


Fig. 3.1 Comparative temperature optima of the xylanases produced by the wild type pX3 to (a) Mutant 2B7-10 and (b) Mutant 58.





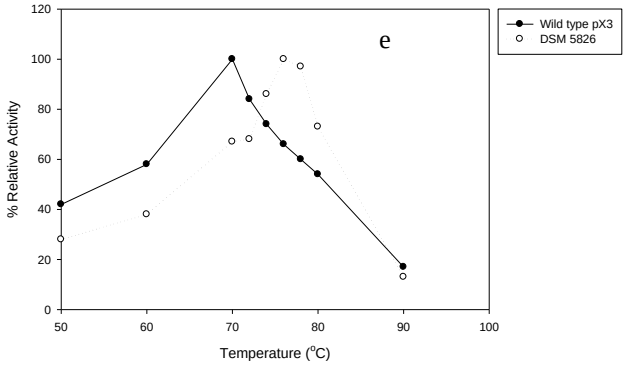


Fig. 3.2 Comparative temperature optima of the xylanases produced by the wild type pX3 to (c) Mutant 27, (d) Mutant 47 and (e) DSM 5826. Each point represents the mean of duplicate determinations.

Mutants were initially screened for improved thermal stability at 80°C for 90 minutes and the more promising mutants were then tested at 80°C for three hours. The parent clones and their progeny mutants were designated a number to differentiate them from each other. Three rounds of experiments were done in respect of short-term thermostability of mutants. Mutant 2B7-10 was used as the parent for the first round of error-prone PCR. This mutant was the product of two rounds of mutagenesis that was carried out in a previous study by Stephens *et al.* (2007).

Mutant 2B7-10 retained 52% of its activity at 80°C after 90 minutes as compared to its progeny mutant 58, which retained 57% of its activity under the same conditions (Fig. 3.3). Although mutant 13 and mutant 52 exhibited better thermostability for up to 45 minutes, after 90 minutes at 80°C both their activities dropped to 54%. Mutant 12 was marginally better than the parent with an activity of 55%. The wild type xylanase retained only 33% of its activity under the same conditions (Fig. 3.3).

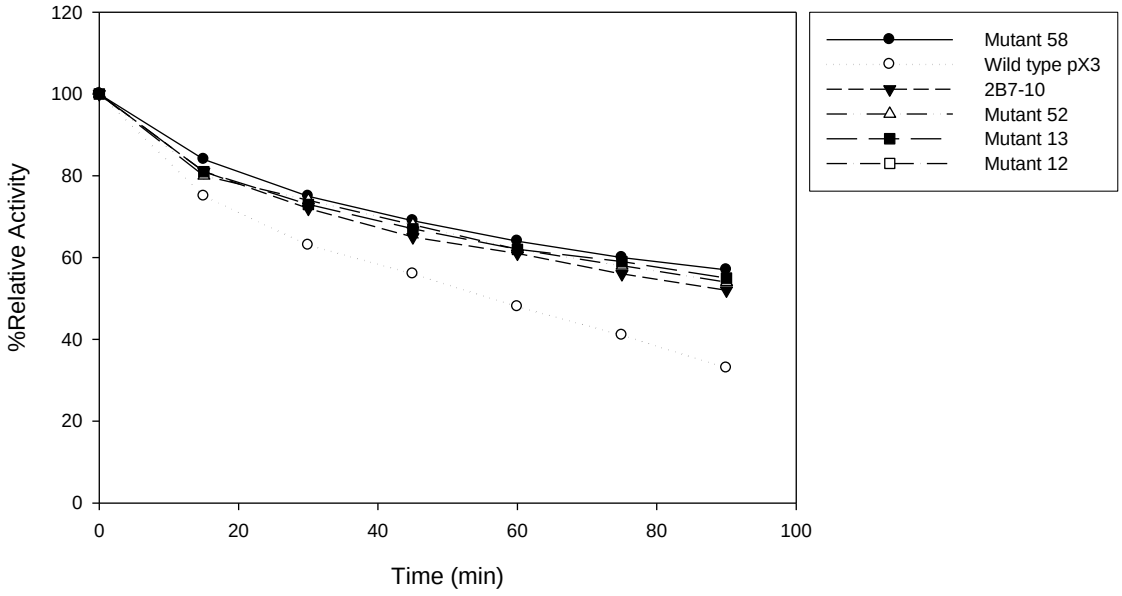


Fig. 3.3 Comparative thermostability screening of wild type XynA and the enzyme variants after the first round of random mutagenesis. The crude enzyme was incubated at 80°C for 90 minutes. Each point represents the mean of duplicate determinations.

Mutant 58 which exhibited the best thermostability from the first round of random mutagenesis was used as the parent for the successive round of mutagenesis. Its progeny, mutant 27 retained 60% of its relative activity at 80°C after 90 minutes (Fig. 3.4). This mutant was the most thermostable from all of the mutants tested, in the second round of mutagenesis. Mutant 105 was only slightly better than the parent with a relative activity of 58% whilst the two other enzyme variants *viz*., mutant 66 and mutant 28 performed poorly with relative activities of 55% and 53% respectively. Mutant 66 displayed 69% activity at 60 minutes which was slightly better than mutant 27 which showed 67% activity for the same period of time however, after an hour its activity dropped to 55%. Again, all of the mutants demonstrated improved stability in comparison to the wild type pX3 (Fig. 3.4).

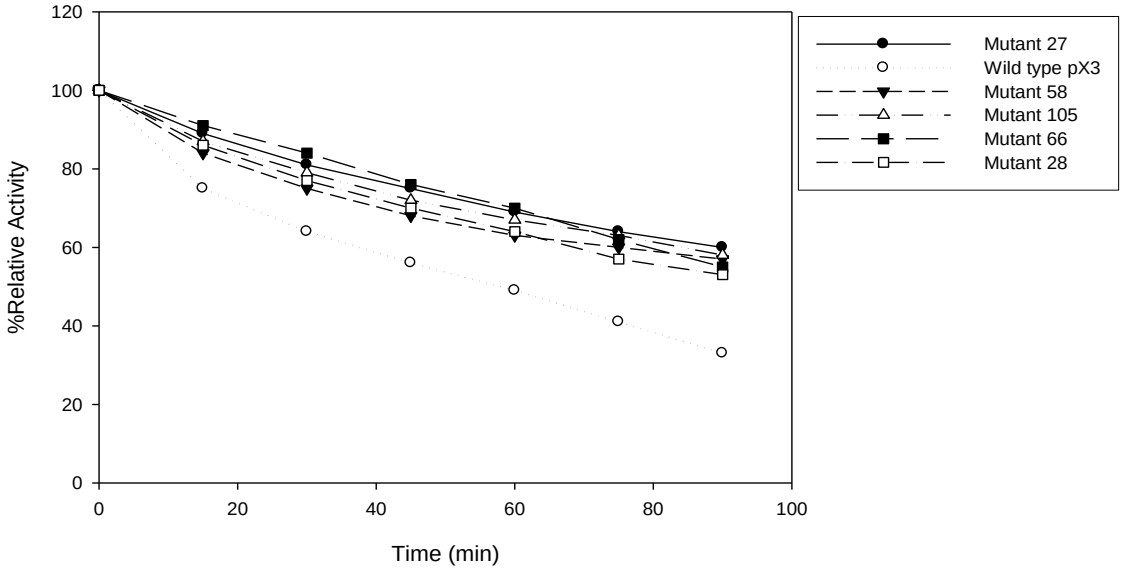


Fig. 3.4 Comparative thermostability screening of wild type XynA and the enzyme variants after the second round of random mutagenesis. The crude enzyme was incubated at 80°Cfor 90 minutes. Each point represents the mean of duplicate determinations.

For the final round of random mutagenesis, mutant 27 was used as the parent clone. Five promising mutants in addition to the parent and the wild type were evaluated in this round of experiments. Mutant 47, demonstrated exceptional relative activity of 71% at 80°C after 90 minutes, compared to the parent, mutant 27 which displayed 60% relative activity under the same conditions (Fig.3.5). The other mutants that were tested *viz.*, mutant 31, mutant 54, mutant 107 and mutant 116 showed slight differences in relative activity ranging from 61 to 65% compared to the parent.. The enzyme variants once again proved to be more stable than the wild type pX3 (Fig. 3.5).

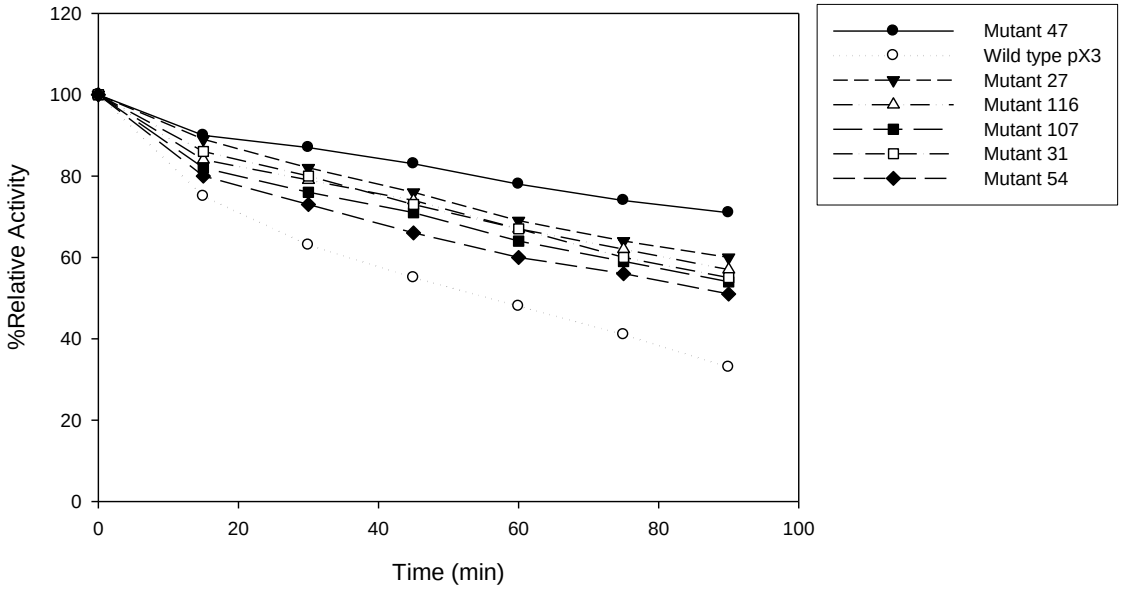


Fig. 3.5 **C**omparative thermostability screening of wild type XynA and the enzyme variants after the third round of random mutagenesis. The crude enzyme was incubated at 80°C for 90 minutes. Each point represents the mean of duplicate determinations.

The most improved enzyme variant from each round was tested for long term thermostability, at 80°C and 90°C, respectively for three hours. The wild type XynA served as the control and the crude enzyme from *T. lanuginosus* DSM 5826 was also evaluated in this set of experiments.

At 80°C, mutant 47 retained 40% relative activity after three hours whilst mutant 27 and mutant 58 displayed 31% and 30% relative activity, respectively under the same conditions (Fig. 3.6). A similar trend occurred with mutant 2B7-10 which exhibited 28% activity after three hours at 80°C. The wild type enzyme however, demonstrated very poor stability compared to the enzyme variants. It retained only 5% activity after three hours at 80°C. Of all the enzymes tested the xylanase from the DSM strain proved to be the most unstable. It displayed only 3% activity up to 45 minutes after which the enzyme was completely inactivated (Fig.3.6).

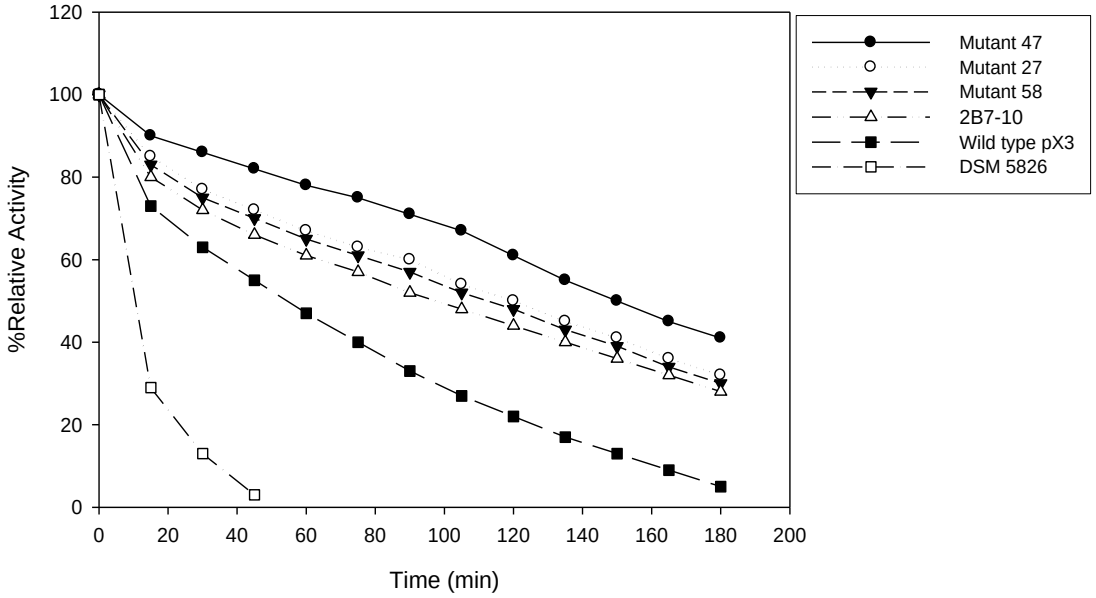


Fig. 3.6 Influence of temperature on the stability of wild type XynA, enzyme variants and xylanase from *T. lanuginosus* DSM 5826. Crude enzyme was incubated at 80°C for 3 hours. Each point represents the mean of duplicate determinations.

In terms of residual activity of xylanase amongst the recombinants, mutant 2B7-10 expressed the highest activity of 1280 nkat/ml at 80°C whilst mutant 47 displayed the lowest level, 867 nkat/ml at the same temperature (Appendix A). Mutant 27 and mutant 58 had 1117 nkat/ml and 1173 nkat/ml xylanase activity, respectively. In addition to its poor stability, the wild type pX3 expressed very low catalytic activity of 270 nkat/ml as

well. The fungal strain DSM 5826 however, expressed the highest amount of enzyme which was 20 600 nkat/ml (Appendix A). This is due to the fact that the xylanase is produced extracellularly in this fungal strain.

Table 3.1 Percentage improvement in long term thermostability at 80°C of enzyme variants vs wild type

Time (h)	Improvement in thermostability (%)				Residual activity (%)
	Mutant 47*	Mutant 27*	Mutant 58*	2B7-10*	Wild type pX3
1	32 (79)	20 (67)	18 (65)	15 (62)	(47)
2	38 (60)	27 (49)	26 (48)	22 (44)	(22)
3	26 (40)	17 (31)	17 (30)	14 (28)	(14)

* Residual activity values (%) for mutants are placed in parentheses

Mutant 47 showed remarkable improvement in terms of enzyme stability compared to the wild type. This is evident from Table 3.1 which clearly shows the individual percentage improvements of each enzyme variant in comparison to the wild type at specific times, during long term thermostability screening at 80C° for three hours. The greatest improvement of 38% was demonstrated with mutant 47 after two hours. The percentage improvement in thermostability amongst the other enzyme variants ranged from 14 to 27%. This clearly demonstrates that *in vitro* enzyme evolution can improve the property of an enzyme without detailed knowledge of the structure and function of the enzyme.

For the second set of experiments for long term thermostability, the crude enzyme of *T. lanuginosus* DSM 5826, the wild type and the recombinant strains were incubated at 90°C for three hours. The assessment of thermostability of the xylanase produced by the four recombinant strains showed that mutant 47 was the most thermostable, whereas the three other mutant strains proved to be less stable over the same period of time (Fig. 3.7). Mutant 47 retained 26% activity after three hours whilst mutant 27 and mutant 58 displayed 21% and 20% activity, respectively. Mutant 2B7-10 was marginally less thermostable having retained 19% activity after three hours (Fig. 3.7).

Both mutants 27 and 58 showed a greater degree of similarity to the parent mutant 2B7-10 as the thermostability trend that all of these mutants exhibited were very similar. The wild type was very unstable at this temperature as well. It retained only 2% of activity after three hours. The DSM strain however, completely lost all of its activity after 30 minutes at 90°C (Fig. 3.7). At this temperature, the enzyme activity of all mutant strains decreased with an increase in incubation time. However, it is evident that all of the progeny mutants displayed better thermostability than the parental strain.

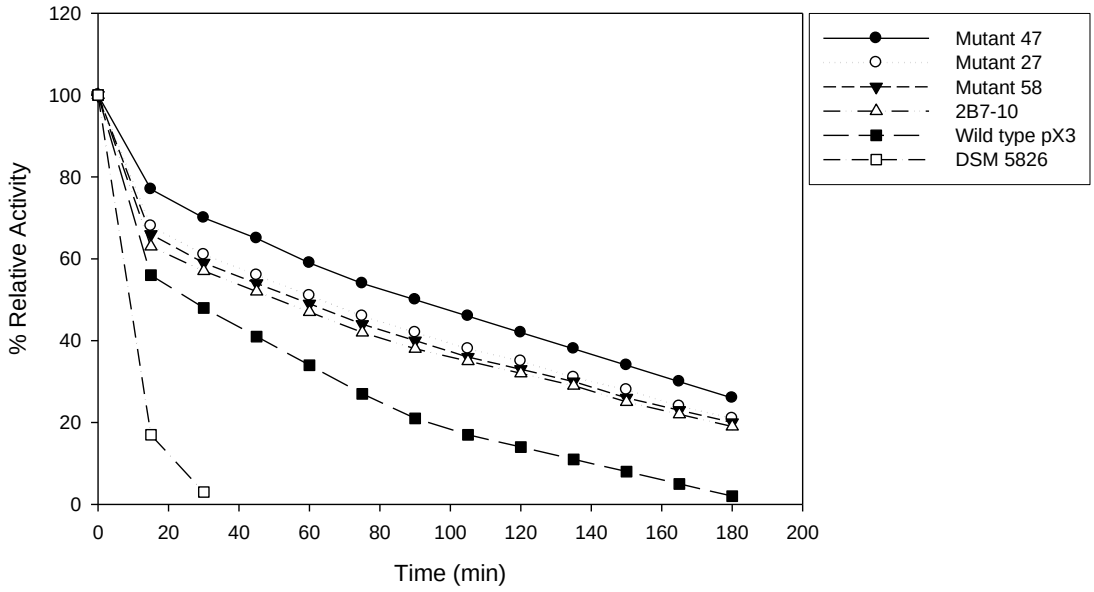


Fig. 3.7 Influence of temperature on the stability of wild type XynA,, enzyme variants and *T. lanuginosus* DSM 5826. Crude enzyme was incubated at 90°C for 3 hours. Each point represents the mean of duplicate determinations.

With regards to the residual activity at 90°C for three hours the fungal strain DSM 5826 exhibited the highest activity of 20 000 nkat/ml compared to the recombinant strains which had activities in the range of 1200 nkat/ml for mutant 2B7-10 and 840 nkat/ml for mutant 47 (Appendix B). The enzyme activity for mutant 27 and 58 was 1017 nkat/ml and 1147 nkat/ml, respectively. The wild type displayed very low enzyme activity of, 267 nkat/ml (Appendix B).

From Table.3.2 it is evident that all of the enzyme variants had improved stability compared to the wild type although some improvements were better than others. Mutant 47 showed the best improvement of 27% after two hours. At 60 minutes and at three hours the percentage improvement was identical. The percentage improvement for mutants 27, 58 and 2B7-10 ranged from 12% to 22%. Of interest is that the progeny mutants showed better improvement in stability than the parental strains.

Table 3.2 Percentage improvement in long term thermostability at 90°C of enzyme variants vs wild type

Time (h)	Improvement in thermostability (%)				Residual activity (%)
	Mutant 47	Mutant 27	Mutant 58	2B7-10	Wild type pX3
1	24 (59)	16 (51)	14 (49)	12 (47)	35
2	27 (41)	22 (36)	19 (33)	17 (31)	14
3	24 (27)	18 (21)	17 (20)	15 (18)	3

*Residual activity values (%) for mutants are placed in parentheses

3.4 DISCUSSION

In this enzyme evolution study, optimization of the thermostable property of the xylanase enzyme was investigated, due to its many potential commercial applications particularly in the pulp and paper industry. While conformational stability does not provide a guarantee for catalytic activity, it certainly is a prerequisite (Eijsink *et al*., 2004).

The xylanase gene incorporating the secretion signal sequence of *T. lanuginosus* (DSM 5826) was expressed in *E. coli* as a *LacZ*-fusion protein (Schlacher *et al*., 1996). Even though the secretion signal is present, *E. coli* does not recognize it and therefore localizes the enzyme intracellularly. The enzymes of all the recombinants were produced intracellularly with a tendency to inclusion body formation. Consequently, the recombinant strains produced much lower levels of enzyme compared to the DSM strain. The highest residual activity of xylanase was obtained from mutant 2B7-10 which produced 1280 nkat/ml when induced with 0.1 mM isopropyl thiogalactosidase as opposed to 20 600 nkat/ml of xylanase produced by DSM 5826 strain. This low extracellular enzyme activity of xylanase from recombinants can be attributed to low level expression by bacterial promoters, the absence of a post-translational modification system in *E. coli*, non-cleavage of the signal sequence to yield the mature protein and the intracellular accumulation of the recombinant xylanases (Rizzatti *et al*., 2004).

What was most interestng in this study is that, with an increase in stability there was a decrease in activity. This occurred among all the enzyme variants that were screened during long term thermostability assay at both 80°C and 90°C. This phenomenon of high stability and low activity was also observed by Giver *et al*. (1998) who stated that mutations that increase thermostability while maintaining activity are rare. Unless both of the properties *viz*., stability and activity are constrained the evolution of one, by mutagenesis comes at the expense of the other.

This has led to the suggestion that there is a trade–off between high stability and high activity, possibly because rigidification of enzymes is needed to attain high stability, whereas conformational flexibility is beneficial for catalysis (Jaenicke and Bohm, 1998). Clearly, the putative inverse correlation between stability and activity is much more complex than previously suggested. One explanation may be that activity requires local flexibility, whereas stability may be determined by flexibility in regions of the protein that do not affect local flexibility of the active site (Eijsink *et al*., 2004).

The xylanase from mutant 47 which was obtained from the final round of error-prone PCR displayed the highest thermostability. There was a 38% improvement in stability of mutant 47 compared to the xylanase produced by the wild type pX3 strain (Table 3.1). This was significant as it was the highest percentage increase in terms of stability that was obtained in this study. This stability trend occurred for both the temperatures (80°C and 90°C) that were tested. From Table 3.1 and Table 3.2 it is evident that all of the progeny mutants demonstrated improved stability when compared to the wild type pX3. Random mutagenesis using epPCR was therefore a good method choice in improving the thermostable properties of the enzyme.

During long term thermostability screening, the xylanase produced by wild type pX3 displayed poor stability whilst the enzyme from the DSM strain was completely inactive after 45 minutes at 80°C and after 30 minutes at 90°C. This is evident of the fact, that the enzyme from pX3 as well as from the fungal strain DSM 5826 cannot withstand high temperatures of 80°C and 90°C for prolonged periods.

It seems evident that thermal adaptation cannot be attributed to any specific characteristic, but results from various changes which contribute to the stability of the organism in an adaptive manner. This might be the result of many intracellular adjustments, such as membrane lipids composition, structures of proteins and protein synthesis machinery, or cytoplasmic thermostabilizing substances, such as polyamines (Rizzatti *et al*., 2004).

Another interesting issue is the relationship between thermostability and the temperature optimum for activity. It is well known that enzymes isolated from thermophilic organisms tend to be as active at their (high) temperature optima as their counterparts from mesophilic organisms at their temperature optima (Jaenicke and Bohm, 1998). This was clearly evident from the temperature optima studies that were carried out in this study. The temperature optimum for xylanase activity in crude extracts was 72°C for all recombinant strains selected which is much higher than their growth temperature optima (Eijsink *et al*., 2004). The same occurred for the wild type pX3 and DSM 5826 strain, whereby the temperature optimum for xylanase activity was 70°C and 76°C respectively, whilst the growth temperature optimum was 37°C and 50°C for each of them respectively (Eijsink *et al*., 2004).

Thus, some enzymes like the fungal strain seem to have genuine temperature optima that are not always dictated by conformational stability (Daniel *et al*., 2001). This explains its high temperature optima of 76°C whilst it was not very stable for a prolonged period of time. The mutants on the other hand, had both a higher temperature optima than pX3 and were more stable than the wild type pX3. Perhaps this is a clear indication that enzymes that are stabilized by protein engineering do have higher temperature optima (Eijsink *et al*., 2004).

In this investigation, sequence analysis was carried out for two important reasons. Firstly, in order to verify whether a mutation did in fact occur, and secondly, to identify those amino acid substitutions that could possibly have contributed to the remarkable thermostability of the enzyme variants. What is noteworthy is that mutant 47 which had the most number of mutations was also found to have the highest thermostability amongst mutants that were evaluated from each of the rounds of mutagenesis. This was significant because these amino acid substitutions could possibly have attributed to the increase in thermostability that was observed during enzyme assays. Several factors that could have conferred thermostability include extra disulfide bridges, electrostatic interactions, hydrophobic interactions, extra networks of hydrogen bonds, protection or shortening of N- and C- terminus and extra proline or glycine substitution sites (Gruber *et*

al., 1998). Perhaps a single or a combination of several of the above factors may have been responsible for the thermostability of mutant 47, which requires an in-depth knowledge of structure and function and goes beyond the scope of this study. Moreover, large changes in thermostability are usually the result of cumulative small contributions of a number of mutations (Gray *et al*., 2001). These results should be useful for future mutagenesis studies aimed at improving the thermostability of this important class of enzymes.

According to Cherry and Fidantsef (2003), epPCR, cassette mutagenesis and saturation mutagenesis all work with varying degrees of success to randomly mutate a target gene, but are very dependent on screening of the resulting diversity to identify variants with improved fitness. It was unfortunate that the screening strategy that was employed in this study could not differentiate between mutations that were neutral or beneficial. Thus, mutant 58 and mutant 27 were redundant enzymes as their sequences were identical to the parent mutant 2B7-10 and their relative activities which ranged from 52% to 60% were slightly different from each other, under the same conditions. These slight differences in thermostability, does not indicate that these are actually enzyme variants. The screening method for identifying useful mutations should ensure that the expected small enhancements brought about mainly by single mutations can be measured. The importance of an efficient screening system therefore, cannot be underestimated. The basis for variation of thermostability among, mutant 58 and mutant 27 investigated in this study requires further elucidation. Sequence analysis is therefore crucial after each consecutive round of random mutagenesis in order to determine whether a mutation occurred or not and what the effects of each mutation was.

In a study carried out by Wakarchuck *et al.* (1994), attempts to enhance the thermostability of *B. circulans* xylanase by introducing disulfide bridges via site-directed mutagenesis increased the thermostability of the enzyme by as much as 15˚C. While it is beyond doubt that the disulfide bridge enhances the thermostability of the *T. lanuginosus* xylanase, other factors also contribute to this property of the enzyme (Gruber *et al*., 1998). This seems to confirm the opinion that the thermophilic attributes, including

improved enzyme thermostability, result from a series of modifications of pre-existing mesophilic functions, in order to optimize survival at higher temperatures (Kinjo and Nishikawa, 2001). All properties cited may be important for improving the properties of this important xylanolytic complex.

In addition, there are conflicting reports in literature with respect to how rigid hyperstable enzymes really are (Jaenicke, 2000; D' Amico *et al*., 2003). The wealth of reports, Strop and Mayo, (2000); Tollinger *et al*., (2003) and Gray et al., (2001) on stability engineering underlines one clear principle: one cannot generalize the effect of a certain type of mutation on stability. The effect of every single point mutation should always be viewed in the light of its structural context (Strop and Mayo, 2000; Tollinger *et al*., 2003).

Thermostable proteins have been studied extensively to gain insight into the adaptive mechanisms for achieving thermostability. Enzyme sequences from mesophilic organisms have been compared to their homologs from thermophilic and hyperthermophilic microorganisms in efforts to identify the interactions responsible for conferring enhanced thermostability (Kinjo and Nishikawa, 2001). Numerous mutagenesis studies have also explored this phenomenon, however, the mechanisms that mediate protein thermostabilization is still poorly understood. Nevertheless, stability remains an important parameter, which co-determines the economic feasibility of applying an enzyme in an industrial process (Farinas *et al*., 2001).

One of the most remarkable features of biological catalysts is that they are capable of evolution. This study demonstrated that directed evolution using epPCR can be used to generate enzymes that display improved thermostability. This was evident with all of the progeny mutants and the parental strains that were investigated in this study (Table 3.1 and Table 3.2). Continued growth of the industrial enzyme market is dependent on technological innovation to improve performance and reduce cost. Thus, enhancement of the thermostability of the fungal xylanase using random mutagenesis represents a valuable tool for both industrial and research applications. The future of this field is fascinating and boundless!

CHAPTER FOUR

GENERAL CONCLUSIONS AND RECOMMENDATIONS

Enzymes are the catalytic cornerstone of metabolism, and as such are the focus of intense worldwide research. The latter half of the twentieth century saw an unprecedented expansion in our knowledge of the use of microorganisms, their metabolic products, and enzymes in a broad area of research and their potential industrial applications. Only in the past two decades however, have microbial enzymes been used commercially in the pulp and paper industry (Beg *et al*., 2001). Xylanases belong to the large group of glycosyl hydrolases, and catalyze the degradation of xylan, the most abundant hemicellulose synthesized in the biosphere (Jeffries, 1996).

Proteins from thermophilic organisms have proven to be extremely useful for high temperature applications (Flores and Ellington, 2002). High levels of cellulase-free thermostable xylanases are known to be produced by the deuteromycete *T. lanuginosus* DSM 5826 (Purkarthofer *et al*., 1993). Since this enzyme is known to be thermostable, its 3D structure was compared to known structures of mesophilic family 11 xylanases in view of possible factors conferring thermostability. It was found that the thermostability is due to the presence of an extra disulfide bridge, as well as to an increase in the density of charged residues throughout the protein (Gruber *et al*., 1998). This study focused on the xylanase enzyme from *T. lanuginosus* DSM 5826 and variants of this enzyme derived by directed evolution. After several rounds of error-prone PCR, plate screening and thermostability screening, an enzyme variant that displayed 71% relative activity at 80°C for 90 minutes was obtained. It is thought that perhaps the presence of the disulfide bridge in the protein structure contributed to the increase in thermostability that was observed in this variant.

Mutant 47 displayed the highest thermostability at both, 80°C and 90°C. At 80°C, the enzyme variant retained 613 nkat/ml of enzyme activity after 90 minutes whilst the wild type retained 90 nkat/ml of enzyme activity under the same conditions (Appendix A). At

90°C a similar trend was observed. Whilst the enzyme variant retained 413 nkat/ml of activity, only 57 nkat/ml of activity remained in the wild type pX3 (Appendix B). An overall improvement of 38% increase in thermostability was evident in mutant 47 in comparison to pX3.

One of the critical steps in directed evolution is the creation of a screen or selection that is rapid enough to process millions of mutants while accurately measuring a desired property (Zhao and Arnold, 1997). The mutant libraries that were created in this study were screened and promising mutants were assayed to identify improved variants. The screening method employed in this investigation was effective. The active enzymes from mutants that possess a functional xylanase gene created a visible halo or a clear zone surrounding the bacterial colony. Only mutants that produced this clear zone were tested further, those mutants that did not produce a clear zone were discarded. For this reason, the method generally required a starting protein with some activity towards the desired reaction. Therefore, the primary screening is most effective in reducing the workload as well as time taken to test all positive strains. This methodology can be applied to other enzymes to create thermostable variants. A first screen is preferable to reduce the number of samples in the second screen.

Screening by monitoring formation of the exact product of interest is desirable, but limits throughput. Future research in directed evolution should include development of large-scale screening methods, so that efficient searches of large mutant libraries can be performed. According to Burton *et al.* (2002), many of the necessary molecular and screening technologies are in place and there is evidence of an evolution from mono-functional to multi-functional screening. The full implementation of the latter, is the key to the successful acquisition of ‘new’ designer biocatalysts for truly optimized bioprocesses. With the dye-linked agar plate method that was used in this study, only enzyme variants that showed some zones of clearing were tested further, hence those variants that did not exhibit zones could have displayed thermostable properties and yielded a different result. This can be overcome by coupling the screening method used with a digital imaging system or by using flow cytometry. The implementation of a

multi-functional screening whereby activity as well as stability could be tested simultaneously would therefore be of great benefit as it would be less laborious and less time-consuming.

Ideally, enzymes undergoing directed evolution should be directed into the extra-cellular medium, however, in this study the *xynA* gene was functionally expressed in *E. coli* as a *lac*Z fusion protein. Cloning and expression resulted in potentially active enzyme molecules being located in the cell cytoplasm or in the periplasmic fluid. Eukaryotic proteins expressed intracellularly in *E. coli* are frequently sequestered into insoluble inclusion bodies. The intermolecular association of hydrophobic domains during folding is believed to play a role in the formation of inclusion bodies (Khan *et al*., 2002).

Although the progenies greatly resembled their parents, after many generations new features developed such that the descendent variant 47 is quite different from its ancestor, wild type pX3. A similar observation was also made by Frances Arnold in 1998, where variants that were obtained in her study were quite different from their parents. In terms of enzyme activity at 80°C, all the recombinant strains produced the same or lower levels of enzyme compared to the parent (Appendix A). Mutant 2B7-10 (the parent) produced the highest activity of 1280 nkat/ml whilst its progeny, mutant 58 produced 1173 nkat/ml. Like wise, mutant 27 (the parent) produced 1117 nkat/ml of enzyme whilst mutant 47 (the progeny) displayed the lowest activity of 867 nkat/ml. A similar trend occurred at 90°C for the same period of time for all the recombinant strains (Appendix B). The wild type had the least amount of enzyme activity at both temperatures and for the same period of time. It was also very unstable compared to the recombinant strains. However, the xylanase from *T. lanuginosus* DSM 5826 was completely inactive after 30 minutes at 80°C and after 45 minutes at 90°C but it produced the highest amount of xylanase 20600 nkat/ml at 80°C and 20000 nkat/ml at 90°C (Appendix A and Appendix B). What was interesting was the fact that a temperature optimum of 72°C for xylanase activity was observed in all recombinant strains tested in this investigation. There was a clear shift to the right in terms of temperature optima of xylanase activity in all of the enzyme variants in comparison to the wild type enzyme.

While there were variants that were more stable and variants that were more active than the wild type pX3, none of the clones selected from the mutant libraries generated in this study, were both more stable and more active than the wild type. According to Giver *et al*. (1998), mutations that improve both properties simultaneously are rare. Since most mutations are neutral or deleterious, an increase in any one property is likely to come at the expense of another if the experiment, does not constrain both properties (Giver *et al*., 1998). In this investigation it was found that by selecting only thermostable mutants to parent subsequent generations, the available evolutionary paths to highly active mutants were significantly restricted. Because single mutations that improve two or more properties are rare, it should be more efficient in the future to direct the evolution of multiple independent properties by recombining those properties that evolved separately (Kuchner and Arnold, 1997).

Sequence analysis revealed that with all the recombinant strains selected from each round of mutagenesis, a mutation at residue position 58 took place. This mutation was inherited from the parent 2B7-10. The amino acid substitution changed from tyrosine to phenylalanine. No other differences were observed for mutant 58 and mutant 27 whilst for mutant 47 three other amino acid substitutions occurred. At residue position 180, the amino acid changed from threonine to serine, at position 201, histidine was substituted for proline and in position 211, tyrosine was substituted for phenylalanine. The substitution of the four amino acids in residue positions 58, 180 201 and 211 contributed to the improved thermostability of the enzyme variant.

In this evolutionary study, this phenomenon needs to be investigated further before any conclusions are drawn about the increased thermostability that was obtained in mutant 47. In order to test this, it is possible for future studies to use site-directed mutagenesis to create enzymes having each or all of the four mutations identified in this study and to verify whether the properties are the same as that produced by mutant 47. A library of such mutations should then be created. Therefore, the properties of mutant 47 can possibly be enhanced by adding other beneficial amino acids and should probably be looked at for future prospects. The sequencing, structure and mutagenesis information

accumulated in the last 20 years confirmed that no single mechanism is responsible for the remarkable stability of thermophilic proteins (Vieille and Zeikus, 2001).

Failure to enhance both the properties of stabiliy and activity might result, to some extent, from an incomplete understanding of the underlying mechanisms required to enhance the desired enzyme properties. Thus, future research should include a combination of computational modeling, bioinformatics and genetic tools for creating smart libraries, allowing a protein engineer to eliminate redundant enzyme variants before constructing libraries that are most likely to yield the desired properties.

The advantage of directed evolution is that this strategy can be employed even when limited information is available on an enzyme's structure or catalytic mechanism. Since the vast majority of proteins remain largely uncharacterized, this marks a huge advantage for the evolutionary methods. The disadvantage of this method however, is that it requires confirmation of the mutation by sequencing and then extraction of the enzyme from the mutants, following each round of random mutagenesis. Such an approach is expensive and time-consuming and might be impractical for multiple cycles of mutagenesis. Another disadvantage lies in the fact that directed evolution has a codon bias towards a certain group/s of amino acids.

Therefore, the results obtained in this study should be useful for future mutagenesis studies aimed at improving the thermostability of this important class of enzymes. With the availability of a thermostable enzyme, a number of new applications in the future are likely. Although believed to provide tremendous economical benefits, production of a thermostable enzyme to the level required by industry still remains a challenge. However, the important significance of the enzyme produced in this study by directed evolution, lies in the fact that the increase in its thermostable property may be useful in the pulp and paper industry. Future work therefore, entails improving the activity of mutant 47 using a combination of molecular strategies and thereafter testing the potential application of this enzyme in an industrial process.

REFERENCES

Akanuma, S., Yamagishi, A., Tanaka, N. and Oshima, T. 1998. Serial increase in the thermal stabilty of 3-isopropylmalate dehydrogenase from *Bacillus subtilis* by experimental evolution. *Protein Science* **7**: 698-705.

Arase, A., Yomo, T., Urabe, I., Hata, Y., Katsube, Y. and Okada, H. 1993. Stabilization of xylanase by random mutagenesis. *Federation of European Biochemical Societies* **316**: 123-127.

Arnold, F.H. 1998. Design by directed evolution. *Accounts of Chemical Research* **31**: 125-131.

Arnold, F.H. and Volkov, A.A. 1999. Directed evolution of biocatalysts. *Current Opinion in Chemical Biology* **3**: 54-59.

Arnold, F.H., Wintrode, P.L., Miyazaki, K. and Gershenson, A. 2001. How enzymes adapt: lessons from directed evolution. *TRENDS in Biochemical Sciences* **26**: 100-106.

Arnold, F.H. and Gergiou, G. 2003. Directed evolution library construction: methods and protocols. Humana Press Inc., Totowa New Jersey.

Bailey, M.J., Biely, P. and Poutanen, K. 1992. Interlaboratory testing of methods for assay of xylanase activity. *Journal of Biotechnology* **23**: 257-270.

Beg,Q.K., Kapoor, M., Mahajan, L. and Hoondal, G.S. 2001. Microbial xylanases and their industrial applications: a review. *Applied Microbiology and Biotechnology* **56**: 326-338.

Benkovic, S.J., Perham, R., and Aebersold, R. 2000. Protein technologies and commercial enzymes. *Current Opinions in Biotechnology* **11**: 317-318.

Biely, P., Mislovicova, D. and Toman, R. 1985. Soluble chromogenic substrates for the assay of endo-β-1,4-xylanases and endo-β-1,4-glucanases. *Analytical Biochemistry* **144**: 142-146.

Biely, P., Mislovicova, D. and Toman, R. 1988. Remazol brilliant blue xylan: a soluble chromogenic substrate for xylanases. *Methods in Enzymology* **160**: 536-541.

Biely, P. 1993**.** Biochemical aspects of the production of microbial hemicellulases. In: Hemicellulose and Hemicellulases. pp. 29-51. Coughlan, M.P. and Hazelwood, G.P. (eds.). Portland Press, London.

Bierman, D.J. 1993. Essentials of pulping and papermaking. pp. 48-59. Academic Press, New York.

Birnboim, **H**.**C. and Doly**, **J.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acid Reviews* **7**: 1513-1523.

Bjork, A., Dalhus, B., Mantzilas, D., Eijsink, V.G.H. and Sirevag, R. 2003. Stabilization of a tetrameric malate dehydrogenase by introduction of a disulfide bridge at the dimer-dimer interface. *Journal of Molecular Biology* **334**: 811-821.

Bolon, D.N., Voigt, C.A. and Mayo, S.L. 2002. *De novo* design of biocatalysts. *Current Opinion in Chemical Biology* **6**: 125-129.

Bosley, A.D. and Ostermeier, M. 2005. Mathematical expressions useful in the construction, description and evaluation of protein libraries. *Biomolecular Engineering* **22**: 57-61.

Bornscheuer, U.T. and Pohl, M. 2001. Improved biocatalysts by directed evolution and rational protein design. *Current Opinion in Chemical Biology* **5**: 137-143.

Braman J. 2002. *In vitro* mutagenesis protocols: methods in molecular biology. Humana Press Inc., Totowa New Jersey.

Brode, P.F., Erwin, C.R., Rauch, D.S., Barnett, B.L., Armpriester, J.M., Wang, E.S.F. and Rubingh, D.N. 1996. Subtilisin BPN2 variants: increased hydrolytic activity on surface-bound substrates via decreased surface activity. *Biochemistry* **35**: 3162-3169.

Buchholz, F. Angrand, P.O. and Stewart, A.F. 1998. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nature Biotechnology* **16**: 657-662.

Burton, S.G., Cowan, D.A. and Woodley, J.M. 2002. The search for ideal biocatalyst. *Nature Biotechnology* **20**: 37-45.

Castro, L.P.M., Trejo-Aguilar, B.A. and Osorio, G.A. 1997. Thermostable xylanases produced at 37°C and 45°C by a thermo-tolerant *Aspergillus* strain. *FEMS Microbiological Letters* **146**: 97-102.

Canada, K.A., Iwashita, S., Shim, H. and Wood, T.K. 2002. Directed evolution of toluene ortho-monooxygenase for enhanced 1-naphthol synthesis and chlorinated ethane degradation. *Journal of Bacteriology* **184**: 344-349.

Cesar, T. and Mrša, V. 1996. Purification and properties of the xylanase produced by *Thermomyces lanuginosus*. *Enzyme and Microbial Technology* **19**: 289-296.

Chen, R. 2001. Enzyme engineering: rational redesign versus directed evolution. *Trends in Biotechnology* **19**: 13-14.

Cherry, J.R. and Fidantsef, A.L. 2003. Directed evolution of industrial enzymes: an update. *Current Opinion in Biotechnology* **14**: 438-443.

Christov, L.P and Prior, B.A. 1993. Esterases of xylan degrading micro-organisms: Production, properties and significance. *Enzyme and Microbial Technology* **15**: 460-473.

Clarke, J., Hendrik, K. and Fersht, A.R. 1995. Disulfide mutants of barnase 1: changes in stability and structure assessed by biophysical methods and x-ray crystallography. *Journal of Molecular Biology* **253**: 493-504.

Cohen, N., Abramov, S., Dror, Y and Freeman, A. 2001. In vitro enzyme evolution: the screening challenge of isolating the one in million. *Trends in Biotechnology* **19**: 507-510.

Collins, T., Gerday, C. and Feller, G. 2004. Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiological Reviews* 1-3.

D'Amico, S., Marx, J.C., Gerday, C. and Feller, G. 2003. Activity-stability relationships in extremophilic enzymes. *Journal of Biological Chemistry* **278**: 7891-7896.

Dalby, P.A. 2003. Optimising enzyme function by directed evolution. *Current Opinion in Structural Biology* **13**: 1-6.

Daniel, R.M., Danson, M.J. and Eisenthal,R. 2001. The temperature optima of enzymes: a new perspective on an old phenomenon. *Trends in Biochemical Science* **26**: 223-225.

Daugherty, P.S., Iverson, B.L. and Georgiou, G. 2000. Flow cytometric screening of cell-based libraries. *Journal of Immunological Methods* **243**: 211-227.

Eijsink, V.G.H., Bjork, A., Gaseidnes, S., Sirevag, R., Synstad, B., Van den Burg, B. and Vriend, G. 2004. Rational engineering of enzyme stability. *Journal of Biotechnology* **113**: 105-120

Eriksson, L. and Hietmann, J. 1998. Applications of enzyme technology in the paper industry. In: Proceedings of the Technical Association of the Pulp and Paper Industry. p.3. Montreal.

Farinas, E.T., Bulter, T. and Arnold, F.H. 2001. Directed enzyme evolution. *Current Opinion in Biotechnology* **12**: 545-551.

Flores, H. and Ellington, A.D. 2002. Increasing the thermal stability of an Oligomeric protein, Beta-glucuronidase. *Journal of Molecular Biology* **315**: 325-337.

Giver, L., Gershenson, A., Freskgard, P.O. and Arnold, F.H. 1998. Directed evolution of a thermostable esterase. *Proceedings of the National Academy of Sciences, USA.* **95**: 12809-12813.

Gomes, I., Raajit, K.S., Mohluddin, G. and Hoq, M.M. 1992. Isolation and characterization of cellulase-free pectinolytic and hemicellulolytic thermophilic fungi. *World Journal of Microbial Biotechnology* **8**: 559-562.

Gray, K.A., Richardson, T.H., Kretz, K., Short, J.M., Bartnek, F., Knowles, R., Kan., L., Swanson, P.E. and Robertson, D.E. 2001. Rapid evolution of reversible denaturation and elevated melting temperature in a microbial haloalkane dehalogenase. *Advanced Synthesis and Catalysis* **343**: 607-617.

Greener, A., Callahan, M. and Jerpseth, B. 1997. An efficient random mutagenesis technique using an *E. coli* mutator strain. *Molecular Biotechnology* **7**: 189-195.

Gruber, K., Klintschar, G., Hayn, M., Schlacher, A., Steiner, W. and Kratky, C. 1998. Thermophilic xylanase from *Thermomyces lanuginosus*: High resolution X-ray structure and modeling studies. *Biochemistry* **37**: 13475-13485.

Gupta, S and Gupta, M.N. 1993. Mechanisms of irreversible thermo-inactivation and medium engineering. In: Thermostability of enzymes. pp. 114-122. Gupta, M.N. (ed.) Narosa Publishing House, New Delhi.

Haki , G. D. and Rakshit, S.K. 2003. Developments in industrially important enzymes: a review. *Bioresource Technology* **89**: 17-34.

Haltrich,D., Nidetzky, B., Kulbe,K.D., Steiner,W. and Zupancic,S. 1996. Production of fungal xylanases. *Bioresource Technology* **58**: 137-161.

Hibbert, E.G., Baganz, F., Hailes, H.C., Ward, J.M., Lye, G.J., Woodley, J.M. and Dalby, P.A. 2005. Directed evolution of biocatalytic processes. *Biomolecular Engineering* **22**: 11-19.

Holmberg, R.C., Henry, A.A. and Romesberg, F.E. 2005. Directed evolution of novel polymerases. *Biomolecular Engineering* **22**: 39-49.

Iijima, H. and Morimoto, K. 1995. Controlling susceptibility against protease digestion. *Annals of the New York Academy of Sciences* **750**: 62-65.

Jacques, G., Frederic, D.L., Joste, L.B., Viviane, B., Bart, D., Fabrizio, G., Benoit, G. and Jean-Marie, F. 2000. An additional aromatic interaction improves the thermostability and thermophilicity of a mesophilic family 11 xylanase: structural basis and molecular study. *Protein Science* **9**: 466-475.

Jaeger. K. and Reetz, M.T. 2000. Directed evolution of enantioselective enzymes for organic chemistry. *Current Opinion in Chemical Biology* **4**: 68-73.

Jaenicke, R. 2000. Do ultastable proteins from hyperthermophiles have high or low conformational rigidity? *Proceedings of the National Academy of Sciences,* USA **97**: 2962-2964.

Jaenicke, R. and Bohm, G. 1998. The stability of proteins in extreme environments. *Current Opinion in Structural Biology* **8**: 738-748.

Janes, L.E., Lowendahl, A.C. and Kazlauskas, R.J. 1998. Quantitative screening of hydrolase libraries using pH indicators: identifying active and enantioselective hydrolases. *Chemistry-A European Journal* **4**: 2324-2331.

Jonasson, P., Liljeqvist, S., Nygren, P. and Stahl, S. 2002. Genetic design for facilitated production and recovery of recombinant proteins in *Escherichia coli. Biotechnology of Applied Biochemistry* **35**: 91-105.

Kichise, T., Taguchi, S. and Doi, Y. 2002. Enhanced accumulation and changed monomer composition in polhdroxyalkanoate (PHA) coployester by *in vitro* evolution of *Aeromonas caviae* PHA synthase. *Applied and Environmental Microbiology* **68**: 2411-2419.

Kinjo, A.R. and Nishikawa, K. 2001. Comparison of energy components of proteins from thermophilic and mesophilic organisms. *European Biophysics Journal* **30**: 378-384.

Khan, S.N., Akter, M.Z. and Sims, P.F.G. 2002. Cloning and expression analysis of two endo-1,4 -β-xylanase genes from *Phanerochaete chrysosporium. Plant Tissue Culture* **12**: 57-68.

Kohilu, U., Nigam,P., Singh, D. and Chaudhary,K. 2001. Thermostable, alkaliphilic and cellulose free xylanases production by *Thermoactinomyces thalophilus* subgroup C. *Enzyme and Microbial Technology* **28**: 606-610.

Kuchner, O. and Arnold, F.H. 1997. Directed evolution of enzyme catalysts. *Trends in Biotechnology* **15**: 523-530.

Kumar, S., Tsai, C. and Nussinov, R. 2000. Factors enhancing protein thermostability. *Protein Engineering* **13**: 179-191.

Kurtzman, A.L., Govindarajan, S., Vahle,K., Jones, J.T., Heinrichs, V. and Patten, P.A. 2001. Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins. *Current Opinion in Biotechnology* **12**: 361-370.

Lesney, M.S. 2003. Biocatalysis on board. Today's Chemist at work. pp 20-23.

Liu, D.R., Magliery, T.J., Pastrnak, M. and Schultz, P.G. 1997. Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins *in vivo*. *Proceedings of the National Academy of Sciences, USA* **94**: 10092-10097.

Lutz, S and Patrick, W.M. 2004. Novel methods for directed evolution of enzymes: quality, not quantity. *Current Opinion in Biotechnology* **15**: 291-297.

MacBeath, G., Kast, P.and Hilvert, D. 1998. Probing enzyme quaternary structure by combinatorial mutagenesis and selection. *Protein Science* **7**: 1757-1767.

Maheshwari, R., Bharadwaj, G. and Bhat, M.K. 2000. Thermophilic fungus: their physiology and enzymes. *Microbiology of Molecular Biology Review* **4**: 461-488.

Marrs, B., Delagrave, S. and Murphy, D. 1999. Novel approaches for discovering industrial enzymes. *Current Opinion* in *Microbiology* **2**: 241-245.

Martineau, P., Jones, P. and Winter, G. 1998. Expression of an antibody fragment at high levels in the bacterial cytoplasm. *Journal of Molecular Biology* **280**: 117-127.

Matsumura, I. and Rowe, L.A. 2005. Whole plasmid mutagenic PCR for directed protein evolution. *Biomolecular Engineering* **22**: 73-79.

Merz, A., Yee, M.C., Szadkowski, H., Pappenberger, G., Crameri, A., Stemmer, W.P., Yanofsky, C. and Kirschner, K. 2000. Improving the catalytic activity of a thermophilic enzyme at low temperatures. *Biochemistry* **39**: 880-889.

Miyazaki, K. and Arnold, F.H. 1999. Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. *Journal of Molecular Evolution* **49**: 716-720.

Miyazaki, K. and Takenouchi, M. 2002**.** Creating random mutagenesis libraries using megaprimer PCR of whole plasmid. *BioTechnique***s 33:** 1033-1038.

Moore,G.L., and Maranas,C.D. 2000. Modeling DNA mutation and recombination for directed evolution experiments. *Journal of Theoretical Biology* **205**: 483-503.

Morawski, B., Quan, S. and Arnold, F.H. 2001. Functional expression and stabilization of horseradish peroxidase by directed evolution in *Sacchoromyces cerevisiae*. *Biotechnology and Bioengineering* **76**: 99-107.

Morris, D., Gibbs, M., Chin, C., Koh, H., Wong, K., Allison, R., Nelson, P. and Bergquist, P. 1998. Cloning of the xynB gene from *Dictyoglomus thermophilum* Rt 46B.1 and action of the gene products on kraft pulp. *Applied and Environmental Microbiology* **64**: 1759-1765.

Naki, D., Paech, C., Ganshaw, G. and Schellenberger, V. 1998. Selection of a subtilisin-hyperproducing *Bacillus* in highly structured environment. *Applied Microbiology* and *Biotechnology* **49**: 290-294.

Neylon, C. 2004. Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acid Research* **32**: 1448-1459.

Niehaus, F., Bertoldo, C., Kahler, M. and Antranikian,G. 1999. Extremeophiles as a source of novel enzymes for industrial application. *Applied Microbiology and Biotechnology* **51**: 711-729.

Otten, L.G. and Quax, W.J. 2005. Directed evolution: selecting today's biocatalysts. *Biomolecular Engineering* **22**: 1-9.

Prade,R.A. 1996. Xylanases: from biology to biotechnology. *Biotechnology and Genetic Engineering Review* **13**: 101-131.

Puchart, V., Katapodis, P., Biely, P., Kremnicky, L., Christakopoulos, P., Vranska, M., Kekos, D., Macris, B.J. and Bhat, M.K. 1999. Production of xylanases, mannases, and pectinases by the thermophilic fungus *Thermomyces lanuginosus*. *Enzyme and Microbial Technology* **24**: 355-361.

Purkarthofer, H., Sinner, M. and Steiner, W. 1993. Cellulase-free xylanase from *Thermomyces lanuginosus*: optimization of production in submerged and solid-state culture. *Enzyme and Microbial Technology* **15**: 677-680.

Rizzati, A.C.S. and Sandrim, V.C. 2004. Influence of temperature on the properties of the xylanolytic enzymes of the thermotolerant fungus *Aspergillus phoenicis*. *Journal of Industrial Microbiology and Biotechnology* **31**: 88-93.

Rowe, L.A., Geddie, M.L., Alexander, O.B. and Matsumura. I. 2003. A comparison of directed approaches using the β-glucuronidase model system. *Journal of Molecular Biology* **332**: 851-860.

Rubingh, N. D. 1997. Protein engineering from a bioindustrial point of view. *Current Opinion in Biotechnology* **8**: 417-422.

Saboto, D., Nucci, R., Rossi, M., Gryczynski, I., Gryczynski, Z. and Lakowicz, J. 1999. The β-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus:* enzyme activity and conformational dynamics at temperatures above 100˚C. *Annals of the New York Academy of Sciences* **672**: 137-141.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular cloning: a laboratory manual. 2nd edition. Cold Spring Harbor Laboratory, Cold Spring, New York.

Schimdt-Dannert, C., Umeno, D. and Arnold, F.H. 2000. Molecular breeding of carotenoid biosynthetic pathways. *Nature Biotechnology* **18**: 750-753.

Schlacher, A., Holzmann, K., Hayn, M., Steiner, W. and Schwab, H. 1996. Cloning and characterization of the gene for the thermostable xylanase XynA from *Thermomyces lanuginosus. Journal of Biotechnology* **49**: 211-218.

Shao, Z., Zhao, H., Giver, L. and Arnold, F.H. 1998. Random priming in vitro recombination: an effective tool for directed evolution. *Nucleic Acids Research* **26**: 681-683.

Sieber, V., Pluckthun, A. and Schmid, F.X. 1998. Selecting proteins with improved stability by a phage-based method. *Nature Biotechnology* **16**: 955-960.

Singh, S., Pillay, B. and Prior, B.A. 2000. Thermal stability of β-xylanase produced by different *Thermomyces Lanuginosus* strains. *Enzyme and Microbial Technology* **26**: 502-508.

Song, J.K and Rhee, J.S. 2001. Enhancement of stability and activity of phospholipase A_1 in organic solvents by directed evolution. International Journal of *Biochemistry, Biophysics and Molecular Biology* **1547**: 370-378.

Spada, S. and Pluckthun, A. 1997. Selectively infective phage (SIP) technology: a novel method for *in vivo* selection of interacting protein-ligand pairs. *Nature Medicine* **3**: 694-696.

Stemmer,W.P.C. 1994. Rapid evolution of a protein *in vitro* by DNA shuffling. ***Nature*** **370**: 389-391**.**

Stephens, D.E., **Rumbold, K., Permaul, K., Prior, B.A. and Singh, S.** 2007. Directed evolution of the thermostable xylanase from *Thermomyces lanuginosus*. *Journal of Biotechnology* **3**:348-354**.**

Strop, P. and Mayo, S.L. 2000. Contribution of surface salt bridges to protein stability. *Biochemistry* **39**: 1251-1255.

Sunna, A. and Antranikian, G. 1997. Xylanolytic enzymes from fungi and bacteria. *Critical Reviews in Biotechnology* **17**: 39-67.

Suzuki, T., Yasugi, M., Arisaka, F., Yamagishi, A. and Oshima, T. 2001. Adaptation of a thermophilic enzyme, 3-isopropylmalate dehydrogenase, to low temperatures. *Protein Engineering* **14**: 85-91.

Swarroopa, D. and Krishna, N. 2000. Production of thermostable cellulose-free xylanase by *Clostridium absonum* CFR-702. *Proceedings in Biochemistry* **36**: 355-362.

Tanner, J.J., Hecht, R.M. and Krause, K.L. 1996. Determinants of enzyme thermostability observed in the molecular structure of *Thermus acquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5*A* resolution. *Biochemistry* **35**: 2597-2609.

Tao, H. and Cornish, V.W. 2002. Milestones in directed enzyme evolution. *Current Opinion in Chemical Biology* **6**: 858-864.

Tobin, M.B., Gustafsson, C. and Huisman, G.W. 2000. Directed evolution: the 'rational' basis for 'irrational' design. *Current Opinion in Structural Biology* **10**: 421-427.

Tollinger, M., Crowhurst, K.A., Kay, L.E. and Forman-Kay, J.D. 2003. Site-specific contributions to the pH-dependence of protein stability. *Proceedings of the National Academy of Sciences,* USA **100**: 4545-4550.

Topcu, Z. 2000. Densitometric quantification of DNA topoisomers in ethidium bromide-stained agarose gels and chemiluminescence-detected X ray films. *Acta Biochimica Polonica* **47**: 841-846.

Torronen, A. and Rouvinen, J. 1997. Structural and functional properties of low molecular weight endo-1,4-β-xylanases. *Journal of Biotechnology* **57**: 137-149.

Vieille, C. and Zeikus, G.J. 2001. Hyperthermophilic Enzymes: Sources, uses and molecular mechanisms for thermostability. *Microbiology and Molecular Biology Reviews* **65**: 1-43.

Voigt, C.A., Kauffman, S. and Wang, Z.G. 2000. Rational evolutionary design: the theory of *in vitro* protein evolution. *Advances in Protein Chemistry* **55**: 79-160.

Wakarchuk, W.W., Sung, W.L., Campbell, R.L., Cunningham, A., Watson, D.C. and Yaguchi, M. 1994. Thermostabilization of the *Bacillus circulans* xylanase. *Protein Engineering* **7**: 1379-1386.

Waldburger, C.D., Schildbach, J.F. and Sauer, R.T. 1995. Are buried salt bridges important for protein stability and conformational specificity? *Nature Structural and Molecular Biology* **2**: 122-128.

Williams, J.C., Zeelen, J.P., Neubauer, G., Vriend,G., Backmann, J., Michels, P.A.M., Lambeir, A.M. and Wierenga, R.K. 1999. Structural and mutagenesis studies of *Leishmania* triphosphate isomerase: A point mutation can convert a mesophilic enzyme into a superstable enzyme without losing catalytic power. *Protein Engineering* **12**: 243-250.

Yano, T. and Kagamiyama, H. 2001. Directed evolution of ampicillin-resistant activity from a functionally unrelated DNA fragment: A laboratory model of molecular evolution. *Proceedings of the National Academy of Sciences, USA* **98**: 903-907.

Yu, E.K.C., Tan, L.U.L., Maria, K.H., Chan, L.D. and Saddler, N.J. 1987. Production of a thermostable xylanases by a thermophilic fungus, *Thermoascus aurantiacus. Enzyme and Microbial Technology* **9**: 16-24.

Zeikus, J., Vielle, C. and Savachenko, A. 1998. Thermozymes: biotechnology and structure-function relationships. *Extremophiles*. **2**: 179-183.

Zhao, H. and Arnold, **F.H.** 1997. Combinatorial protein design: strategies for screening protein libraries. *Current Opinion in Structural Biology* **7**: 480-485.

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.

Zhao, H. and Arnold, F.H. 1999. Directed evolution converts subtilisin E into a functional equivalent of thermitase. *Protein Engineering* **12**: 47-52.

Zhao, H., Chockalingam, K. and Chen, Z. 2002. Directed evolution of enzymes and pathways for industrial biocatalysis. *Current Opinion in Biotechnology* **13**: 104-110.

APPENDIX

A: Long term thermostability screening of enzyme variants, the wild type and *T. lanuginosus* DSM 5826 at 80˚C for three hours. The average residual activities of xylanase are expressed as nkat/ ml.

TIME	Mutant 47	Mutant 27	Mutant 58	Mutant 2B7-10	Wild type pX3	DSM 5826
0	867	1117	1173	1280	270	20600
15	787	950	973	1027	197	6200
30	747	850	880	920	170	2400
45	707	800	827	840	150	600
60	680	750	760	787	127	-
75	653	700	720	720	107	-
90	613	667	667	667	90	-
105	573	600	613	613	73	-
120	520	550	560	560	60	-
135	480	500	507	507	47	-
150	440	450	453	467	36	-
165	387	400	400	413	23	-
180	347	350	347	360	13	-

B: Long term thermostability screening of enzyme variants, the wild type and *T. lanuginosus* DSM 5826 at 90˚C for three hours. The average residual activities of xylanase are expressed as nkat/ ml.

TIME	Mutant 47	Mutant 27	Mutant 58	Mutant 2B7-10	Wild type pX3	DSM 5826
0	840	1017	1147	1200	267	20000
15	653	683	747	760	150	3400
30	587	617	680	680	130	600
45	547	567	613	627	110	-
60	493	517	560	560	93	-
75	453	467	507	507	73	-
90	413	433	453	453	57	-
105	387	383	413	413	47	
120	347	367	373	373	37	-
135	320	317	347	333	29	
150	293	283	293	293	23	-
165	253	250	267	253	13	
180	227	217	227	213	7	-