# PROTEIN ENGINEERING OF A FUNGAL XYLANASE

**Dawn Elizabeth Stephens**

**Submitted in fulfilment of the requirements for the Degree of Doctor of Technology (Biotechnology), Department of Biotechnology, Faculty of Applied Sciences, Durban University of Technology, Durban, South Africa.**

**Supervisors:** **Prof. S. Singh (Ph. D)**
**Dr. K. Permaul (Ph. D)**

**Date submitted:** **February 2007**

## DECLARATION

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

_____

Dawn Elizabeth Stephens

**FINAL COPY APPROVED FOR SUBMISSION**

_____

Prof. S. Singh

_____

Dr. K. Permaul

“Men fear thought as they fear nothing else on earth -- more than ruin -- more even than death. Thought is subversive and revolutionary, destructive and terrible, thought is merciless to privilege, established institutions, and comfortable habit. Thought looks into the pit of hell and is not afraid. Thought is great and swift and free, the light of the world, and the chief glory of man.”

**Bertrand Russell , philosopher**

**(1872 - 1970)**

## ACKNOWLEDGEMENTS

I would like to pay tribute to the following:

The National Research Foundation, for the award of their scholarships which funded my post-graduate studies over the years;

Dr. Kugen Permaul and Prof. Suren Singh (Durban University of Technology) for their contributions;

Prof. Huimin Zhao and Dr. Frances Arnold (California Institute of Technology), for their suggestions and technical input;

The Inqaba Biotech team (Lufono, Oliver, Michael, Vanitha and Chantal), for their supremely professional assistance with DNA sequencing, technical support and service delivery;

Mr. Jack Turner (Diversa Corporation), for the kind donation of the commercial Luminase enzyme;

Dr. Karl Rumbold (Ghent University), who remains a splendid guide and an exceptional sounding board for research ideas;

Dr. Peter Biely (Slovak Academy of Sciences), for editorial input and technical advice;

Ms. Siphi Dlungwane and Mr. Elliot Ncayiyana (Durban University of Technology), for invaluable technical assistance and always being so obliging and helpful in the lab;

My Mum, Dad and family for their tremendous love, encouragement and prayers;

My darling nephews (Brian, Reglin, Brent, Andrei, Cody) and my delightful nieces (Regina, Caitlin), for always making me smile and helping me remember that life is truly about celebrating and appreciating the simple things;

My good friends, especially Naz, Ray, Kavi, Rachael and J: you make this journey through life so much more beautiful;

My late mentor and teacher, the Rev. Fr. Charles Langlois: Thank you for teaching me to appreciate that being different is a gift, not a curse;

Mostly all praise and honour be to the Almighty who works wonders in my life: You are my sanctuary in tempest and in tranquillity, in high tide and in low tide, in sunshine and in shadow, in popularity and in persecution. I acknowledge You always as the fount of all my strength and skill.

“When you get to the end of all the light you know and it's time to step into the darkness of the unknown, faith is knowing that one of two things shall happen: either you will be given something solid to stand on, or you will be taught how to fly.”

**Edward Teller, physicist**

**(1908 - 2003)**

# CONTENTS

**ABSTRACT**.....i

**LIST OF ABBREVIATIONS**.....ii

**LIST OF FIGURES** .....iii

**LIST OF TABLES**.....x

**CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW**.....1

**1.1 INTRODUCTION**.....1

**1.2 ENZYMES – THE CATALYSTS OF LIFE**.....3

**1.2.1 Biocatalysts with vision**.....3

**1.2.2 Stability – an industrial preference for enzymes**.....6

**1.3 PROTEIN ENGINEERING – THE ANSWER TO MODERN GENETICS**.....7

**1.4 RATIONAL DESIGN – THE LOGICAL APPROACH**.....9

**1.5 DIRECTED EVOLUTION – EMULATING MOTHER NATURE**.....12

**1.5.1 Requirements for successful directed evolution**.....14

**1.5.2 Generating sequence diversity**.....15

***1.5.2.1 Chemical mutagenesis and UV irradiation***.....16

***1.5.2.2 Mutator strains***.....16

***1.5.2.3 Degenerate oligonucleotides***.....17

***1.5.2.4 Ep-PCR***.....18

***1.5.2.5 Commercial random mutagenesis kits***.....20

***1.5.2.6 DNA recombination***.....21

**1.5.3 Screening for improved variants**.....25

**1.5.4 Commercial successes of directed evolution**.....**26**
**1.5.5 Future prospects for directed evolution**.....**27**

**1.6 XYLANASES**.....**29**
**1.6.1 Xylan: occurrence and structure**.....**29**
**1.6.2 Enzymatic hydrolysis of xylan**.....**30**
**1.6.3 Classification of xylanases**.....**31**
**1.6.4 Commercial application of xylanases**.....**33**
***1.6.4.1 Animal nutrition***.....**33**
***1.6.4.2 Baking***.....**33**
***1.6.4.3 Wines***.....**34**
***1.6.4.4 Surfactants***.....**34**
***1.6.4.5 Biobleaching***.....**35**
*1.6.4.5.1 Mechanism of biobleaching*.....**35**
*1.6.4.5.2 Characteristics of the ideal pulp xylanase*.....**36**
**1.6.5 Molecular biology of xylanases**.....**39**
***1.6.5.1 Cloning of xylanases***.....**39**
***1.6.5.2 Crystallization and bioinformatics of xylanases***.....**40**
***1.6.5.3 Protein engineering of xylanases***.....**41**
**1.6.6 *T. lanuginosus***.....**43**

**1.7 RATIONALE FOR THE PRESENT STUDY**.....**45**

**CHAPTER 2: RANDOM MUTAGENESIS OF *xynA***.....**47**

**2.1 INTRODUCTION**.....**47**

**2.2 MATERIALS AND METHODS**.....**50**
**2.2.1 Construction of plasmid pX4**.....**50**
***2.2.1.1 Plasmids***.....**50**
***2.2.1.2 Plasmid DNA isolation***.....**51**
***2.2.1.3 DNA quantification***.....**52**
***2.2.1.4 Preparation of λ DNA molecular weight marker***.....**52**

*2.2.1.5 Agarose gel electrophoresis*....52
*2.2.1.6 Primer design*....53
*2.2.1.7 PCR*....54
*2.2.1.8 Restriction analysis*....55
*2.2.1.9 Klenow treatment*....56
*2.2.1.10 Ligation*....56
*2.2.1.11 Preparation of SEM-competent cells*....56
*2.2.1.12 Transformation and screening*....57
**2.2.2 Random mutagenesis**....58
**2.2.3 Cloning of PCR products**....59
**2.2.4 Growth of mutants and enzyme extraction**....59
**2.2.5 Screening**....59
*2.2.5.1 Thermostability screening assay*....59
*2.2.5.2 Alkaline screening assay*....60
*2.2.5.3 Long-term thermal and alkaline stabilities*....60

**2.3 RESULTS**....62
**2.3.1 Construction of plasmid pX4**....62
*2.3.1.1 PCR*....62
*2.3.1.2 Screening transformants for xylanase production*....63
*2.3.1.3 PCR and restriction analysis of a few isolates*....64
**2.3.2 Random mutagenesis**....66
**2.3.3 Screening for xylanase mutants with improved stability**....67
*2.3.3.1 Short-term thermal stability*....67
*2.3.3.2 Screening for alkali-tolerant xylanase variants*....68
*2.3.3.3 Long-term thermal and alkaline stabilities*....70

**2.4 DISCUSSION**....72

**CHAPTER 3: DNA RECOMBINATION OF *xynA* VARIANTS.....79**

**3.1 INTRODUCTION.....79**

**3.2 MATERIALS AND METHODS.....82**

**3.2.1 Plasmids and plasmid isolation.....82**

**3.2.2 DNA quantification and agarose electrophoresis.....82**

**3.2.3 Optimization of StEP recombination PCR method.....82**

**3.2.4 Cloning of PCR products.....84**

**3.2.5 Growth of recombinants and enzyme extraction.....85**

**3.2.6 Short-term screening for thermo- and alkali-stable recombinants.....85**

**3.2.7 Long-term thermal and alkaline stabilities of recombinants.....85**

**3.2.8 Automated DNA sequencing and sequence analysis of *xynA* variants and recombinants.....86**

**3.3 RESULTS.....87**

**3.3.1 Optimization of StEP recombination PCR method.....87**

**3.3.2 Short-term thermal and alkaline screening of recombinants.....90**

**3.3.3 Long-term thermal and alkaline stabilities of recombinants.....91**

**3.3.4 Sequence analysis of *xynA* variants and recombinants.....93**

**3.4 DISCUSSION.....96**

**CHAPTER 4: BIOCHEMICAL CHARACTERIZATION OF XynA AND VARIANTS.....106**

**4.1 INTRODUCTION.....106**

**4.2 MATERIALS AND METHODS.....109**

**4.2.1 Growth of cultures and enzyme extraction.....109**

**4.2.2 Determination of pH and temperature optima.....110**

**4.2.3 pH and temperature stability.....110**

**4.2.4 First-order kinetics of characterized xylanases.....111**

**4.3 RESULTS**.....**112**
**4.3.1 pH and temperature optima**.....**112**
**4.3.2 pH and temperature stability**.....**115**
**4.3.3 First-order kinetics of characterized xylanases**.....**121**

**4.4 DISCUSSION**.....**123**

**CHAPTER 5: SCREENING FOR ACID-STABLE XynA VARIANTS**.....**130**

**5.1 INTRODUCTION**.....**130**

**5.2 MATERIALS AND METHODS**.....**132**
**5.2.1 Growth of *E. coli* clones and enzyme extraction**.....**132**
**5.2.2 Screening for acid-stable *xynA* variants**.....**132**
**5.2.3 Determination of pH and temperature optima**.....**133**
**5.2.4 Determination of xylanase stability over acidic pH range**.....**133**
**5.2.5 DNA sequence analysis of acid-stable *xynA* variants**.....**133**

**5.3 RESULTS**.....**134**
**5.3.1 Screening for acid-stable *xynA* variants**.....**134**
**5.3.2 Determination of pH and temperature optima**.....**136**
**5.3.3 Long-term xylanase stability over acidic pH range**.....**138**
**5.3.4 Sequence analysis of acid-stable mutants**.....**140**

**5.4 DISCUSSION**.....**141**

**CHAPTER 6: SUBSTRATE SPECIFICITY OF XynA VARIANTS**.....**146**

**6.1 INTRODUCTION**.....**146**

**6.2 MATERIALS AND METHODS**.....**149**
**6.2.1 Growth of *E. coli* clones and enzyme extraction**.....**149**
**6.2.2 Substrate specificity of XynA mutants**.....**149**
**6.2.3 Determination of pH and temperature optima**.....**149**

**6.2.4 DNA sequence analysis of selected mutants**.....**149**

**6.3 RESULTS**.....**150**

**6.3.1 Substrate specificity of XynA mutants**.....**150**

**6.3.2 Determination of pH and temperature optima**.....**152**

**6.3.3 Sequence analysis of selected mutants**.....**154**

**6.4 DISCUSSION**.....**155**

**CHAPTER 7: GENERAL DISCUSSION**.....**159**

**7.1 THE RESEARCH IN PERSPECTIVE**.....**159**

**7.2 FUTURE PROSPECTS FOR RESEARCH**.....**165**

**REFERENCES**.....**168**

## ABSTRACT

Protein engineering technologies, such as directed evolution and DNA recombination, are often used to modify enzymes on a genetic level for the creation of useful industrial catalysts. Pre-treatment of paper pulps with xylanases have been shown to decrease the amounts of toxic chlorine dioxide used to bleach pulp. This study was undertaken to improve the thermal and alkaline stabilities of the xylanase from the fungus *Thermomyces lanuginosus* using ep-PCR and DNA shuffling. An enzyme library consisting of 960 clones was analyzed in duplicate for stability at 80°C and pH 10. Mutant G41 retained 75% of its initial activity at 80°C for 90 min whilst mutant G53 retained 93% of its activity at pH 10 for 90 min. These mutants were then recombined using a DNA shuffling method called the staggered extension process (StEP) in order to assemble both these properties into a single xylanase. The products of the StEP reaction were cloned into *E. coli* and 516 recombinants were collectively obtained and tested for alkaline and thermal stability. Recombinants S340 and S325 displayed 60% activity at pH 10 and 54% and 85% activity at 80°C for 90 min respectively. DNA sequencing revealed parent mutants G53 and G41 to have one and four mutations respectively. The recombinants S340 and S325 exhibited three and five mutations respectively whilst other recombinants M28, M32 and M34 had between six and eight mutations. Most notable was the increase in arginine content of the highly stable enzymes as well as the disruption of the disulphide bridge in a less stable recombinant M28. The cloned XynA, its fungal counterpart, parents G41 and G53 as well as recombinants S325 and S340 were exposed along with a commercial xylanase, Luminase, to thermophilic and alkaline pulping conditions. At higher temperatures, G41 and S325 were quite stable. First order kinetics established that these were the most stable xylanases created in this study. The temperature and pH optima for the tested xylanases were subsequently established. In addition, the initial xylanase library was also screened for stability at acidic pH as well as altered activity on different xylan and cellulose substrates. Two mutants D57 and D63 were found to have better activity at pH 3 than wild-type XynA and DNA sequencing revealed that they both had mutations at the carboxyl ends of the gene. DNA sequence analysis of mutants D15 and G84 showed that alanine and valine substitutions at the amino and carboxyl ends of *xynA* were crucial for improved activity on the substrates tested.

## LIST OF ABBREVIATIONS

| | |
|---|---|
| 3-D | three dimensional |
| ASPD | artificially selected proteins and peptides |
| CAM | computer-assisted modelling |
| CBH | cellobiohydrolase |
| CMC | carboxymethyl cellulose |
| DES | diethylsulphate |
| dITP | deoxyinosine triphosphate |
| dNTPs | deoxynucleoside triphosphates |
| DNS | 3,5 dinitrosalicylic acid |
| DSC | differential scanning calorimetry |
| $E_a$ | activation energy of a reaction |
| EG | endoglucanase |
| EMS | ethylmethane sulphonate |
| ep-PCR | error-prone PCR |
| EX | endoxylanase |
| HPLC | high performance liquid chromatography |
| IPTG | isopropyl-β-D-thiogalactoside |
| k | enzyme reaction rate |
| LA PCR | long and accurate PCR |
| MMS | methylmethane sulphonate |
| $Mg^{2+}$ | magnesium |
| $Mn^{2+}$ | manganese |
| NTG | N-methyl-N´-nitro-N-nitrosoguanidine |
| pBSK | plasmid Bluescript $SK^-$ vector |
| PCR | polymerase chain reaction |
| p*I* | isoelectric point |
| RBB | Remazol Brilliant Blue |
| SDS-PAGE | sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| StEP | staggered extension process |
| $t_{50}$ | half-life |
| UV | ultraviolet |

# LIST OF FIGURES

**Fig. 1.1** Growth of the global enzyme market and its application sectors projected from 2002 to 2009 (Rajan, 2004).....4

**Fig. 1.2** A general scheme of rational protein design (Percival-Zhang *et al*., 2006).....10

**Fig. 1.3** A typical scheme for directed evolution (Percival-Zhang *et al*., 2006).....14

**Fig. 1.4** A simple description of DNA shuffling. Similar or related genes are mixed with the gene of interest. (1) Genes are digested separately with DNase1. (2) The fragments are denatured and mixed in differing ratios and allowed to form heterologous duplexes (preparative PCR). (3) dNTPs and *Taq* polymerase are added along with primers specific for gene ends. (4) PCR is carried out to amplify chimeric genes (Stemmer, 1994a; 1994b).....22

**Fig. 1.5** A schematic representation of the association between xylan, cellulose and lignin in a typical wood fibre (Techapun *et al*., 2003).....29

**Fig. 1.6** A hypothetical plant xylan structure showing different substituent groups with the sites of attack by the microbial xylanolytic system (Beg *et al*., 2001).....31

**Fig. 1.7** Ribbon representations of the main fold of the catalytic domains of the endoxylanases (EX) of family 10 and 11 (Davies and Henrissat, 1995).....32

**Fig. 1.8** A hypothetical structure of kraft pulp fibres. Xylanases are thought to promote pulp bleaching via the hydrolysis of re-precipitated xylan (Viikari *et al*., 1994).....36

**Fig. 1.9** Molecular structure of *xynA* from *T. lanuginosus* DSM 5826. (A) A perpendicular view of the structure of *xynA* after 3-D protein modelling. (B) Topology diagram depicting the arrangement of the β sheets and α-helix within *xynA* (Gruber *et al*., 1998).....44

**Fig. 2.1** A linear map of pX3 region showing location of the *xynA* gene from *T. lanuginosus* DSM 5826 drawn using BVTech Plasmid software (version 2.0). Restriction sites and locations on the polylinker of the plasmid used for cloning are shown and arrows indicate direction of DNA synthesis. pBSK MCS: vector multiple cloning sites.....50

**Fig. 2.2:** A linear map of pX4 region showing location of the *xynA* gene from *T. lanuginosus* DSM 5826 drawn using BVTech Plasmid software (version 2.0). Restriction sites and locations on the polylinker of the plasmid used for cloning are shown and arrows indicate direction of DNA synthesis. pBSK MCS: vector multiple cloning sites.....51

**Fig. 2.3** Gene sequence showing sense (E20) and antisense (X21) primer binding sites (grey highlights) on *xynA*. The actual recognition site within each primer is underlined and in boldface while the *xynA* start and stop codons are highlighted in black.....54

**Fig. 2.4** Agarose gel electrophoresis of pX3 after PCR with T3 and T7 primers (lane 2) and primers E20 and X21 (lane 3) against λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII (lane 1).....62

**Fig. 2.5** Recombinant *E. coli* colonies after transformation and incubation at 37°C for 16 h. Transformation mixtures were plated on LB medium supplemented with 0.4% RBB-xylan and ampicillin (100 μg/ml). The presence of a clear halo around the bacterial colony is indicative of xylanase production.....63

**Fig. 2.6** Agarose gel electrophoresis of 5 isolates after PCR with E20 and X21 primers. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: xylanase-producing clone 1; Lane 3: non-xylanase producing clone 2; Lane 4: non-xylanase producing clone 3; Lane 5: xylanase-producing clone 4; Lane 6: xylanase-producing clone 5.....65

**Fig. 2.7** Agarose gel electrophoresis of plasmids isolated from positive transformants subjected to restriction digestion. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: pBSK with *Eco*RI; Lane 3: 786 bp product obtained after PCR with E20 and X21; Lane 4: *Eco*RI; Lane 5: *Xba*I; Lane 6: *Eco*RI and *Xba*I; Lane 7: *Eco*RI and *Pst*I; Lane 8: *Eco*RI and *Xho*I; Lane 9: *Pst*I and *Xho*I; Lane 10: *Xho*I and *Xba*I; Lane 11: *Pst*I and *Xba*I.....65

**Fig. 2.8** Agarose gel electrophoresis of random mutagenesis of *xynA* after exposure to different mutagenic PCR conditions (A to H – refer to Table 2.1). Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: A (control amplified under normal PCR conditions); Lane 3: B (4.8 mM $Mg^{2+}$, 0.5 $Mn^{2+}$, 0.2 mM AG, 0.8 mM CT); Lane 4: C (1.5 mM $Mg^{2+}$, 0.04 mM AG, 0.2 mM CT); Lane 5: DI (2 mM $Mg^{2+}$, 0.04 mM $Mn^{2+}$, 0.02 mM AG, 0.2 mM CT); Lane 6: DII (2 mM $Mg^{2+}$, 0.02 mM AG, 0.2 mM CT, 0.04 mM dITP); Lane 7: Kit *Taq* restriction I (control); Lane 8: Kit *Taq* restriction IX (control); Lane 9: E (3.5 mM $Mg^{2+}$, 0.04 mM G, Diversify mix); Lane 10: F (3.5 mM $Mg^{2+}$, 0.64 mM $Mn^{2+}$, 0.04 mM G, Diversify mix); Lane 11: G (3.5 mM $Mg^{2+}$, 0.64 mM $Mn^{2+}$, 0.2 mM G, Diversify mix); Lane 12: H kit control (3.5 mM $Mg^{2+}$, 0.2 mM AGCT).....66

**Fig. 2.9** Screening plate showing zones of hydrolysis produced by crude xylanases created by mutagenesis of *xynA.* The wells contain crude xylanases from (1) XynA, (2) G41 and (3) G53. The plates were prepared according to the modified methods of Teather and Wood, 1982 and Béguin, 1990 and are described in section 2.2.5.2.....69

**Fig. 2.10** The effect of temperature on the stability of wild-type XynA and its mutant progeny at 80°C and pH 6.5. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....71

**Fig. 2.11** The effect of alkalinity on the stability of wild-type XynA and its mutant progeny at pH 10 and 60°C. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....71

**Fig. 3.1** Principle of the StEP recombination method. For simplicity, only one primer and single strands from two DNA templates (open and solid blocks) are shown. During priming, oligonucleotide primers anneal to the denatured templates. Short fragments are produced by brief polymerase-catalyzed primer extension. During subsequent random annealing-abbreviated extension cycles, fragments randomly prime the template (template switching) and extend further. The process is repeated for many cycles until the full-length genes are produced. The full-length chimeric genes can be amplified in a standard PCR (optional) (Zhao *et al.*, 1998; Zhao, 2004).....81

**Fig. 3.2** Optimization of StEP PCR reaction conditions as described in Table 3.1 at 94°C for 30 s and 42°C for 7 s for a total of 80 cycles. All reactions contained 1.5 mM $MgCl_2$, 1X PCR buffers and varying primer concentrations. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: 1.5 µM each T3 and T7 primers, 2 mM dNTPs, 2.5 U *Taq*; Lane 3: 0.5 µM each T3 and T7 primers, 2 mM dNTPs, 2.5 U *Taq*; Lane 4: 1.5 µM each T3 and T7 primers, 2 mM dNTPs, 2.5 U *Taq*; Lane 5: 0.5 µM each T3 and T7 primers, 2 mM dNTPs, 2.5 U *Taq*; Lane 6 (standard): 0.5 µM each T3 and T7 primers, 0.1 mM dNTPs, 1 U *Taq*.....88

**Fig. 3.3** Optimization of StEP PCR reaction conditions as described in Table 3.2 at 94°C for 30 s and 42°C for 5 s at the end of 20 and 80 cycles. All reactions contained 1.5 mM $MgCl_2$ and 1X PCR buffers. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: 1.5 µM each T3 and T7 primers, 1 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 3: 1.5 µM each T3 and T7 primers, 1 mM dNTPs, 2.5 U *Taq* after 80 cycles; Lane 4: 1.5 µM each T3 and T7 primers, 0.5 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 5: 1.5 µM each T3 and T7 primers, 0.5 mM dNTPs, 2.5 U *Taq* after 80 cycles; Lane 6: 1.5 µM each T3 and T7 primers, 0.25 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 7: 1.5 µM each T3 and T7 primers, 0.25 mM dNTPs, 2.5 U *Taq* after 80 cycles; Lane 8 (standard): 0.5 µM each T3 and T7 primers, 0.1 mM dNTPs, 1 U *Taq* after 20 cycles; Lane 9 (standard): 0.5 µM each T3 and T7 primers, 0.1 mM dNTPs, 1 U *Taq* after 80 cycles; Lane 10: 0.5 µM each T3 and T7 primers, 1 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 11: 0.5 µM each T3 and T7 primers, 1 mM dNTPs, 2.5 U *Taq* after 80 cycles; Lane 12: 0.5 µM each T3 and T7 primers, 0.25 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 13: 0.5 µM each T3 and T7 primers, 0.25 mM dNTPs, 2.5 U *Taq* after 80 cycles.....89

**Fig. 3.4** StEP PCR recombination of G41 and G53 at 94°C for 30 s and 42°C for 4 s at the end of 20 and 80 cycles. All reactions contained 10 ng total DNA, 1.5 mM $MgCl_2$, 0.5 µM each T3 and T7 primers, 1X PCR buffer and 1 U *Taq*. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: StEP reaction without mutagen after 20 cycles; Lane 3: StEP reaction without mutagen after 80 cycles; Lane 4: mutagenic StEP reaction with 640 µM $Mn^{2+}$ and 0.2 mM dGTP after 20 cycles; Lane 5: mutagenic StEP reaction with 640 µM $Mn^{2+}$ and 0.2 mM dGTP after 80 cycles.....90

**Fig. 3.5** The effect of temperature on the stability of wild-type XynA, its mutant progeny and recombinants at 80°C and pH 6.5. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....92

**Fig. 3.6** The effect of alkalinity on the stability of wild-type XynA, its mutant progeny and recombinants at pH 10 and 60°C. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....93

**Fig. 3.7** Amino acid sequence alignment of parent mutants (G41 and G53) and recombinants (S325, S340, M28, M32 and M34) with each other and with wild type *xynA*. The alignment was done by using the CLUSTALW (version 1.81) alignment program on the GenomeNet server (www.ebi.ac.uk/clustalw). 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; '.' indicates conservation within weaker groups of amino acids. The absence of an alignment character implies that an unrelated amino acid was substituted. Amino acid substitutions that differ from *xyn*A are indicated in red boldface. Catalytic residues that make up the active site of *xynA* are shown in italics.....95

**Fig. 4.1** The effect of pH on activity of XynA variants, fungal XynA and Luminase against cloned XynA from pH 2 – 12 at 50°C for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....113

**Fig. 4.2** The effect of temperature on the activity of XynA variants, fungal XynA and Luminase against cloned XynA from 40°C - 100°C at their optimum pH for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....114

**Fig. 4.3** The effect of temperature and alkalinity on the stability of wild-type XynA, its mutant progeny and recombinants at (A) 60°C, (B) 70°C, (C) 80°C and (D) 90°C at pH 8. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....117

**Fig. 4.4** The effect of temperature and alkalinity on the stability of wild-type XynA, its mutant progeny and recombinants at (A) 60°C, (B) 70°C, (C) 80°C and (D) 90°C at pH 9. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....118

**Fig. 4.5** The effect of temperature and alkalinity on the stability of wild-type XynA, its mutant progeny and recombinants at (A) 60°C, (B) 70°C, (C) 80°C and (D) 90°C at pH 10. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....119

**Fig. 4.6** The effect of acidity on the stability of wild-type XynA, its mutant progeny and recombinants at (A) pH 3, (B) pH 4 and (C) pH 5 at 60°C. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....120

**Fig. 5.1** Screening plate showing zones of hydrolysis produced by crude xylanases created by mutagenesis of *xynA*. The wells contained crude xylanases from (1) XynA, (2) D57 and (3) D63. The plates were prepared according to modified methods (Teather and Wood, 1982, Olaveson and Stokes, 1989, Béguin, 1990 and Xiong *et al*., 2004) and are described in section 5.2.2.....135

**Fig. 5.2** The effect of pH on the activity of D57 and D63 against wild-type XynA from pH 2 – 12 at 50°C for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....137

**Fig. 5.3** The effect of temperature on the activity of XynA, D57 and D63 against wild-type XynA from 40°C – 100°C at their optimum pH for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....137

**Fig. 5.4** The effect of acidity on the stability of XynA, D57 and D63 at 60°C and pH 3. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....138

**Fig. 5.5** The effect of acidity on the stability of XynA, D57 and D63 at 60°C and pH 4. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....139

**Fig. 5.6** The effect of acidity on the stability of XynA, D57 and D63 at 60°C and pH 5. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....139

**Fig. 5.7** Amino acid sequence alignment of mutants D57, D63 and wild type *xynA*. The alignment was done by using the CLUSTALW (version 1.81) alignment program on the GenomeNet server (www.ebi.ac.uk/clustalw). 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; '.' indicates conservation within weaker groups of amino acids. The absence of an alignment character implies that an unrelated amino acid was substituted. Amino acid substitutions that differ from *xyn*A are indicated in red boldface. Catalytic residues that make up the active site of *xynA* are shown in italics.....140

**Fig. 6.1** The effect of pH on the activity of D15 and G84 against wild-type XynA from pH 2 – 12 at 50°C for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....153

**Fig. 6.2** The effect of temperature on the activity of D15 and G84 against wild-type XynA from 40°C – 100°C at their optimum pH for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.....153

**Fig. 6.3** Amino acid sequence alignment of mutants D15, G84 and wild type *xynA*. The alignment was done by using the CLUSTALW (version 1.81) alignment program on the GenomeNet server (www.ebi.ac.uk/clustalw). 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; '.' indicates conservation within weaker groups of amino acids. The absence of an alignment character implies that an unrelated amino acid was substituted. Amino acid substitutions that differ from *xyn*A are indicated in red boldface. Catalytic residues that make up the active site of *xynA* are shown in italics.....154

# LIST OF TABLES


**Table 1.1** Examples of the industrial applications of some common enzymes.....5

**Table 1.2** Principles of different DNA shuffling techniques.....24

**Table 1.3** Summary of a few examples of directed evolution.....27

**Table 1.4** Characteristics of some commercial pulp bleaching xylanases.....38

**Table 1.5** Examples of xylanases mutated using SDM.....42

**Table 2.1** Mutagenic PCR conditions used for *xynA*.....58

**Table 2.2** Activity and stability of *xynA* mutant library at 80°C for 40 min.....68

**Table 2.3** Activity and stability of selected *xynA* mutants at pH 10, 60°C for 40 min.....69

**Table 3.1** StEP PCR carried out for 30 s at 94°C and 7 s at 42°C.....83

**Table 3.2** StEP PCR carried out under conditions of 30 s at 94°C and 5 s at 42°C.....83

**Table 3.3** Summary of amino acid changes in *xynA* mutants and recombinants.....94

**Table 4.1** Estimated $t_{50}$ in min, inactivation rates (k) and activation energies ($E_a$) of XynA, Luminase and xylanase variants at pH 8 for 180 min.....122

**Table 4.2** Estimated $t_{50}$ in min, inactivation rates (k) and activation energies ($E_a$) of XynA, Luminase and xylanase variants at pH 9 for 180 min.....122

**Table 4.3** Estimated $t_{50}$ in min, inactivation rates (k) and activation energies ($E_a$) of XynA, Luminase and xylanase variants at pH 10 for 120 min.....122

**Table 5.1** Screening of mutant xylanase library on 0.3% xylan agarose plates (pH 3).....135

**Table 6.1** Substrate specificity of mutant *xynA* library for different xylan and cellulose substrates.....151

**Table 6.2** Xylanase activity (nkat/ml) of wild type XynA, D15 and G84 on different xylan and cellulose substrates.....151

# CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

## 1.1 INTRODUCTION

In the field of chemical technology, where manufacture of a variety of products on a large scale has resulted in hazardous effluent and waste problems, the need for safer and "environment-friendly" technologies has become evident. This has resulted in scientists attempting to learn from natural processes, and therefore a new aspect of microbiology and biotechnology has rapidly begun to gain impetus. Microorganisms perform their myriads of biochemical reactions under ambient conditions with little or no toxic by-products. Therefore enzyme technologies, which began initially as an idea, have reached realistic proportions whereby they can be considered worthwhile and practical technologies of the future. A glimmer of hope to preserve the environment yet achieve the goals of chemical technology is gaining momentum in several frontier developments in this area (Kirk *et al*., 2002; Cherry and Fidantsef, 2003; Dalby, 2003).

At present, global production of chemically manufactured pulp stands at around 100 million tons per year (Khristova *et al*., 2006) and this figure is likely to increase despite the trend towards greater utilization of non-chemical methods and recycled paper. Political pressure from environmental groups and consumers has had a major impact on the industry in many western countries. Ironically, much of the environmental damage is caused by the same consumers having a preference for more attractive, bleached paper products. The chemical treatments used in chlorine-based bleaching procedures produce harmful chlorinated organic compounds that are often released into the environment. These chlorinated compounds are difficult to recycle and as a result, they are discarded – in some cases – directly into nearby lakes and rivers (Munkittrick *et al*., 1997). Additional pressure has fallen on the industry because of fears that some of these compounds (such as dioxin) may escape extraction steps and contaminate paper used for food packaging. The pulp and paper industry is faced with having to switch to "greener" pulping technologies. It is reluctant to do so because the chemical processes are cheap, effective and efficient; but they are also agents of pollution and public tolerance is waning (Turner *et al*., 1992; Roncero *et al*., 2005).

The cure for this industrial malady would be to use enzymes such as xylanases which can reduce the amounts of chlorine required for paper pulp bleaching. The positive effect of xylanase is generally attributed to the elimination of the xylans, thereby breaking the existing link between the cellulose and the lignin, and since the lignin is therefore more exposed, it is more readily eliminated in subsequent bleaching stages (Viikari *et al*., 1994; Techapun *et al*., 2003).

Kraft pulping requires conditions of high temperatures and alkaline pH and few xylanases, even those for extremophilic origin, can meet the challenges of such a process. To include the current commercially-available xylanases in the bleaching stages requires cooling and neutralization. The result is a costly alteration of the plant, an increased energy consumption and increased process time, which partly, explains why mill owners are reluctant to embrace this technology. The race is now on for finding new xylanases which can alleviate at least some of these problems (Techapun *et al*., 2003; Khandeparkar and Bhosle, 2006a).

For many years, the identification of new biocatalysts depended on labour-intensive screening of microbial cultures for the desired activities. Efforts to comb natural biodiversity for useful activities have been greatly facilitated by high-throughput screening technologies and by new methods for collecting genes from the environment and expressing them in recombinant organisms. These processes allow faster access to useful catalytic activities from organisms that cannot be cultured. But natural diversity cannot address all practical biocatalytic problems. Screening larger libraries of DNA or microbes may not be the fastest or most efficient route to obtaining a good catalyst. Some problems can be solved by immobilization or crystallization, which can stabilize weak protein structures. But many problems are best alleviated by engineering the catalyst itself (Arnold *et al*., 2001; Chatterjee and Yian, 2006).

Genetic and protein engineering (both rational and random) are modern techniques for the commercial production of enzymes for improved stability to high temperatures, extremes of pH, oxidizing agents and organic solvents. Cloning and expression of suitable genes from a thermophile into a suitable and faster-growing mesophilic host has also provided possibilities of producing the specific thermostable enzyme required

for a particular biotransformation process. Rational engineering is labour-intensive and requires much information about enzyme structure. Random mutagenesis or directed evolution offers the fastest and most effective means of improving biocatalysts provided the screening method is sensitive enough to detect the altered properties being screened for. Recombination, also termed 'sexual evolution', is when beneficial properties from different genes are combined into a single gene generating a protein with superior properties than its parent genes (Cedrone *et al*., 2000; Bacher *et al*., 2002; Haki and Rakshit, 2003).

For xylanases to be economically feasible for pulp application, they must be cellulase-free, retain stability from 60°C - 90°C at pH 8 - 10 for 3 - 5 h (Techapun *et al*., 2003). *T. lanuginosus* DSM 5826 produces high amounts of a cellulase-free xylanase, exhibits stability over a wide pH and temperature range, has been cloned into *E. coli* (Schlacher *et al*., 1996) and crystallized (Gruber *et al*., 1998). It is a potentially good candidate for protein engineering for pulp application. Thus the focus of this research study was to use the principles of directed evolution and recombination to genetically alter the xylanase from *T. lanuginosus* DSM 5826 to render it more temperature and alkaline stable for future pulp application. The genes of the most stable xylanase variants were sequenced and characterized according to existing pulp parameters. The initial xylanase library obtained was also screened for stability at acidic pH as well as altered substrate specificity on different xylan and cellulose substrates to determine if mutation had an effect (if any) on these properties.

## 1.2 ENZYMES – THE CATALYSTS OF LIFE

### 1.2.1 Biocatalysts with vision

The role of enzymes in many processes has been known for a long time. Their existence was associated with the history of ancient Greece where the Greeks used enzymes from microorganisms for baking, brewing, alcohol production and cheese making, etc. The industrial application of microbial enzymes in the western world started more than 100 years ago in 1894, with the patent of Takamine on the production of α-amylase from *Aspergillus oryzae* (cited by Techapun *et al*., 2003). The field of industrial enzymes is experiencing a surge in major research and

development initiatives, resulting in both the development of a number of new products and in improvement in the process and performance of several existing products. As both rational and random methods of enzyme improvement have evolved over the past 20 years, enzymes are being developed to play an ever-expanding role as catalysts in industrial processes. According to a report from Business Communications Company, Inc. (USA), the global market for industrial enzymes was estimated at $2 billion in 2004 and future projections indicate that the total industrial enzyme market is expected to reach nearly $2.4 billion in 2009 (Rajan, 2004).

The industrial enzyme market is divided into three application segments. The largest segment with 65% of sales is technical enzymes and includes enzymes in the detergent, starch, textile, leather, pulp and paper, and personal care industries. Food enzymes, the second largest segment comprising 25% of the market, includes enzymes employed in the dairy, brewing, wine and juice, fats and oils, and baking industries. Finally, feed enzymes, encompassing enzymes used in animal feeds, contributes approximately 10% of the market (McCoy, 2000; Kirk *et al*., 2002; Hibbert and Dalby, 2005). The projected growth of animal feed enzymes is expected to be almost 4% whilst growth of the technical enzymes will parallel the overall market. Overall, the enzymes in various food application sectors are expected to grow at a rate of about 3% (Rajan, 2004). These projections are reflected in Fig. 1.1.

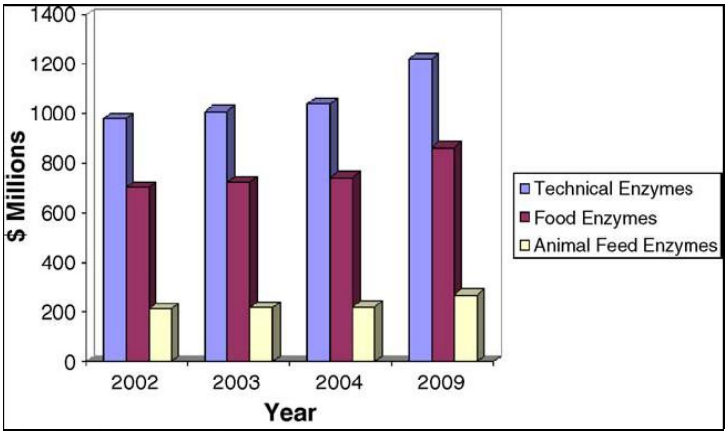


**Fig. 1.1** Growth of the global enzyme market and its application sectors projected from 2002 to 2009 (Rajan, 2004).

Enzymes offer substantial and increasingly important advantages over chemical catalysts in that they:

(i) are derived from renewable resources;

(ii) are biodegradable;

(iii) work under relatively mild reaction conditions of temperature and pH; and

(iv) tend to offer exquisite selectivity in both reactant and product stereochemistry.

As concerns about the environment mount with population pressures and the industrialization of the planet, the need to find more benign methods for the production of goods and provision of services leads to further interest in harnessing enzymes to replace hazardous traditional chemical-based methods (Arnold, 1996; Johannes and Zhao, 2006). Examples where enzymes have found application in industrial processes are shown in Table 1.1.

**Table 1.1** Examples of the industrial applications of some common enzymes

| Enzymes | Chemical replaced | Application | Reference |
|---|---|---|---|
| Protease | Phosphates | Laundry detergents | Gupta *et al*., 2002; Venugopal and Saramma, 2006 |
| Cellulase | Bleach | Fabric softening (e.g., denim) | Hongpattarakere, 2002 Miettinen-Oinonen *et al*., 2004 |
| Lipase | Chemical emulsifiers | Bread making | Pandey *et al*., 1999; Hasan *et al*., 2006 |
| Amylase | Chemical liquefying agents | Syrup and ethanol production | Rakshit, 1998; Asgher *et al*., 2007 |
| Xylanase | Chlorine and chlorine dioxide | Paper pulping | Khristova *et al*., 2006; Khandeparkar and Bhosle, 2006a |
| Transglutaminase | Heat; fat substitutes | Low fat meat products | Trespalacios and Pla, 2007 |
| Pectinase | Sodium hydroxide | Textile finishing | Van der Maarel *et al*., 2002 |
| Pectate lyase | Water and dew retting | Linen making | Akin *et al*., 2007 |

The list of enzyme applications grows bringing with it a list of significant social and environmental benefits. In all of these applications there is a strong demand for improving existing enzymes or for finding new ones, be it for optimizing an existing process, or for acquiring a marketable piece of intellectual property.

### 1.2.2 Stability – an industrial preference for enzymes

Although the advantages of replacing chemical treatments with enzymes are often compelling from a societal and ecological perspective, enzymes must compete economically with often entrenched and extremely inexpensive time-honoured chemical processes and the growth of the enzyme markets (Fig. 1.1) indicate that this is fast becoming reality.

The stability of enzymes and proteins, *in vitro*, particularly with regards to thermal stability remains a critical issue in industrial biotechnology. Thermostable enzymes are often preferred over their mesophilic counterparts since most industrial processes are often carried at high process temperatures. Some of these advantages include:

(i) reduced risk of contamination by common mesophiles;

(ii) improved bioavailability and solubility of organic compounds thereby facilitating efficient bioremediation (Becker, 1997);

(iii) higher reaction rates due to a decreased viscosity (Krahe *et al*., 1996); and

(iv) an increased diffusion coefficient of substrates and higher process yields due to improved solubility of substrates and products and favourable equilibrium displacement in endothermic reactions (Mozhaev, 1993; Kumar and Swati, 2001).

Traditionally, the selection of useful and robust enzymes to be incorporated into commercial applications has been done by screening for enzymes or microbes that have become adapted to extreme environmental conditions, such as geothermal vents and salt lakes (Niehaus *et al*., 1997). However, identification of new biocatalysts (for example, by screening of soil samples or strain collections by enrichment cultures) does not always yield suitable enzymes for a given synthetic problem. Metagenomic gene discovery represents another method for the detection of novel enzymes for biocatalytic applications. It is widely accepted that up to 95% of the microbes present in many environments are not readily culturable. Accordingly, traditional cultivation-based methods are insufficient for accessing the considerable reservoir of hidden natural diversity of the uncultured majority. To avoid the inherent loss of diversity related to cultivation, the collective genomes of all microorganisms present in a given habitat, the so-called metagenome (Handelsman *et al.*, 1998), can be used in a

cultivation-independent approach. It comprises the direct extraction of genomic DNA from environmental samples, its cloning into suitable expression vectors and subsequent screening of the constructed libraries (Cowan *et al.*, 2005; Lämmle *et al*., 2007).

In parallel, other techniques, such as enzyme immobilization, stabilizing additives and screening for improved microorganisms in chemostat cultures, have been developed (Harayama, 1998; Ó Fágáin, 2003). But, even these techniques only lend a slight degree of stability to the enzymes of interest. Another promising alternative in overcoming these limitations is to genetically engineer the already existing vast repertoire of enzymes with improved properties.

When natural enzymes are harnessed for industrial application, it is observed that they are often not well-suited to these tasks due to a variety of reasons such as poor substrate solubility and competing chemical reactions. Reflecting their participation in complex biochemical networks inside living cells, enzymes are often limited by their own substrates or products, either of which may severely limit the productivity of a biocatalytic process. Evolution is usually the reason: enzymes are optimized and often highly specialized for specific functions within the context of a living organism. Consequently, several strategies have been envisaged to modify the properties of natural enzymes (Bergquist *et al*., 2005).

## 1.3 PROTEIN ENGINEERING – THE ANSWER TO MODERN GENETICS

The advent of modern biotechnology, particularly large-scale fermentation and recombinant DNA technology, has not only made enzymes economically feasible, but in many cases it has been used to create enzymes that perform under decidedly unnatural conditions, such as in organic solvents. Commercial enzyme products were originally developed as single fermentation broths of naturally-occurring organisms. Often the resulting products were complex mixtures of secreted enzymes produced at relatively low yields (Arnold, 1996). Today, over 90% of industrial enzymes are produced recombinantly to maximize product purity and economy of production. Expression is carried out in fungal or bacterial hosts that have been modified to

remove unwanted side activities and to maximize expression of heterologous genes (Cherry and Fidantsef, 2003). In addition, a growing number of these heterologous genes have been modified to optimize the catalytic performance of enzymes they encode under application-relevant conditions. Protein engineering technologies involves creating new proteins by modifying existing ones and its importance in industry continues to grow as the number of applications of proteins expands, and the technology to efficiently discover proteins with useful properties is better able to address industrially-relevant problems (Rubingh, 1997).

Protein engineering design attempts to create a 'designer' molecule that is highly suited for a particular process using traditional methods of sequencing and site-directed mutagenesis (SDM). Among the motivations for protein design are the understanding of the relationship between sequence and structure, aiding structural and functional genomics, and creating proteins with desired functions. Protein engineering furthers the molecular biologist's knowledge of the relationship between sequence and structure in a very fundamental manner. Since protein design typically involves mutating sequences with known structure, it is possible to learn how specific changes in a sequence affect the structure of the protein (Freeland and Hurst, 1998; Bacher *et al*., 2002).

But what are the goals of engineering in the first place? A protein's role can be loosely divided into two aspects: folding and function, i.e., the intrinsic, structural aspects of the protein and its extrinsic interactions with its surroundings and both are targets for protein engineering (Doyle *et al*., 2003; Thirumilai *et al*., 2003; Eijsink *et al*., 2005; Otten and Quax, 2005).

The basic assumption of protein engineering is that a natural sequence can be modified to improve a certain function. This implies that:

(i) the protein is not already at an optimum for that function, otherwise it could not be improved;

(ii) the required sequence changes can be accommodated without disrupting the structure, otherwise it would not fold; and

(iii) the new sequence is not too different from the native sequence, otherwise it could not be detected or measured (Arnold *et al*., 2001; Eijsink *et al*., 2004).

There are two types of protein manipulation strategies in use for modern-day genetic engineering, *viz*., rational design and directed evolution.

## 1.4 RATIONAL DESIGN – THE LOGICAL APPROACH

Rational design was the earliest approach to protein engineering and is still the most widely used technique to introduce desired properties into a protein of interest. This approach has been made possible by progress in two techniques, *viz*., SDM and computer-assisted modelling (CAM) of the three-dimensional (3-D) structure of proteins. The strategy hinges on relating structure to function by replacing critical amino acid residues within the protein and monitoring the effects on the gross structure and electrostatic configurations of the protein (Fig. 1.2). Since the possibilities of replacement of amino acids are endless, computational design is used to trim the sequence space to areas of high interest, and then it's fewer possible resultant sequences can be verified experimentally (Blackburn, 2000; Voigt *et al*., 2001).

SDM is a powerful method that is commonly used in several areas of molecular biology and biochemistry. The importance of SDM goes beyond gene structure-function relationships because the technique enables mutant proteins with novel properties of value to be created as in protein engineering. During SDM, an amino acid can be changed into one or more other amino acids (site-saturation mutagenesis). Several protocols have been developed to mutate specific sequences in plasmid DNA. All employ oligonucleotide primers containing the desired mutations flanked by bases complementary to target sequences (Smith, 1985; Kunkel *et al*., 1987; Higuchi *et al*., 1988; Ho *et al*., 1989; Sarkar and Sommer, 1990; Ge and Rudolph, 1997; Angelaccio and Bonaccorsi di Patti, 2002). These methods require a well-designed strategy taking into account the length of the mutated fragment, the presence of (two) unique cleavage sites, and the purification of PCR intermediate products used as templates for further PCR reaction rounds (Ling and Robinson, 1997). Often SDM and CAM

are complementary techniques in the rational design of proteins (Rubingh, 1997; Bornscheuer and Pohl, 2001).

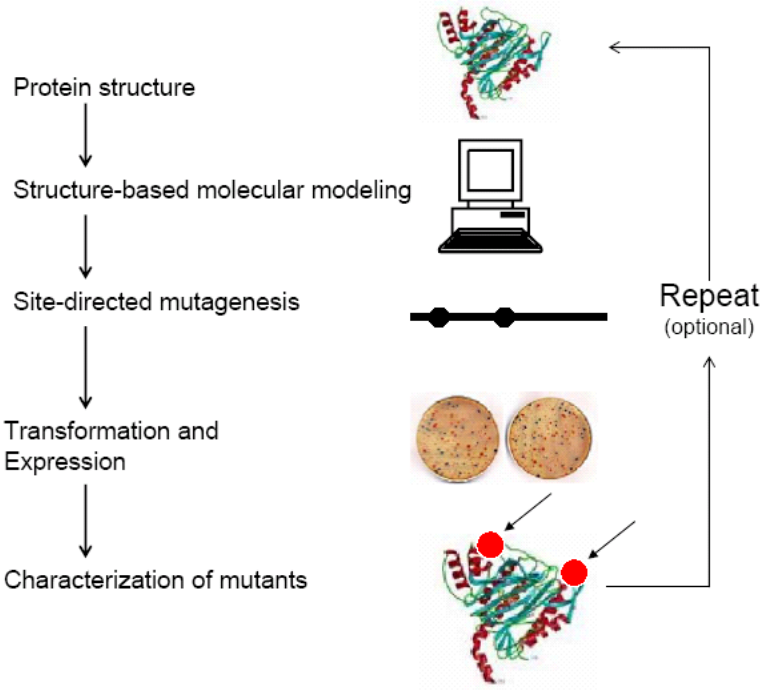


**Fig. 1.2** A general scheme of rational protein design (Percival-Zhang *et al*., 2006).

Rational design usually requires both the availability of the structure of the enzyme and knowledge about the relationships between sequence, structure and mechanism or function. Most importantly, it requires a hypothesis about the limiting step for the desired function. Thus, this method is therefore very information-intensive. In well characterized systems, the performance of rational engineering can be quite remarkable. As an example, the thermolysin-like protease (TLP) has been engineered with 8 point mutations for thermostability to resist boiling temperatures (Van den Burg *et al*., 1998). Remarkably, the mutant enzyme is as active at room temperature as the wild-type. Individual mutations were combined from an analysis of sequence differences to thermolysin, from an increase in the number of residues that lower the entropy of the unfolded state and from a designed disulphide bridge – tried and proven approaches to rational engineering. Depending on the purpose of the mutagenesis, amino acid substitutions are often selected by sequence comparison with homologous sequences. The results however, have to be carefully interpreted because minor sequence changes by a single point mutation may cause significant structural disturbance. Thus, comparison of the 3-D structures of mutant and wild-type enzymes

is necessary to ensure that a single mutation is really site-directed (Bornscheuer and Pohl, 2001).

As much as rational design is both labour- and information-intensive, rapid progress in resolving protein structures by nuclear magnetic resonance (NMR) spectroscopy instead of by X-ray diffraction of crystals, and the enormously increasing numbers of sequences stored in public databases have significantly eased access to data and structures. Using molecular modelling, it has been possible to predict how to increase the selectivity, activity and the stability of enzymes, even if there is no structural data available and the structure of a homologous enzyme is used as a model (Orrenius *et al*., 1998). The goals of molecular modelling are three-fold:

(i) to explain, on a molecular level, the known behaviour of an enzyme;

(ii) to suggest how to change the selectivity of a reaction by modification of substrate, enzyme or reaction conditions; and

(iii) to predict quantitatively, the stereoselectivity of an enzyme-catalyzed reaction (Kaslauskas, 2000).

That this process is becoming practical is evident from the successful *de novo* design of many proteins (Bolon *et al*., 2002). Two main benefits of rational design for evolutionary engineering can be identified:

(i) the possibility to construct stable structural frameworks for the display of combinatorial libraries; and

(ii) the generation of hypotheses that allow limiting the required size of the library, such as constraining diversity to spatially adjacent residues or conserving hydrophobicity profiles (Cedrone *et al*., 2000; Steipe, 2000; Arnold *et al*., 2001).

Unfortunately, this approach is only applicable to protein families in which the 3-D structure of at least one member protein has been resolved. Furthermore, many attempts to alter the properties of enzymes by this approach have failed because the introduced amino acid substitutions exerted unexpected influences on the structure and function of the target enzymes. In fact, successfully modified enzymes obtained by a different approach, *viz*., random mutagenesis, and subsequent intensive screening often possess substituted residues in regions that protein modelling could not predict

(Zhao and Arnold, 1999). In order to obtain improved enzymes for potential commercial benefit, more rapid methods of design are needed. One of these exciting alternative strategies is directed evolution.

## 1.5 DIRECTED EVOLUTION – EMULATING MOTHER NATURE

Evolution is a process in which the selection for specific traits is accomplished by applying environmental pressure. In nature, genetic diversity is obtained by time-consuming spontaneous mutations that occur during DNA replication or by recombination events. In a pool of genetically diverse organisms, the one most suited for survival (best suited for carrying out all relevant reactions at the specific habitat) is selected for, i.e., has a larger probability for passing its genetic material to the next generation. Thus, through recursive cycles of mutation, selection and amplification, new traits accumulate in a population of organisms. Those that provide an advantage under prevailing environmental conditions are passed from one generation to the next (Taylor *et al*., 2001).

Life can be found in almost any environment that has been explored on earth to date, including the depths of the ocean, volcanic and arctic regions or highly saline lakes. The microbes and other life forms that are capable of surviving such harsh environments have adapted their physiology by evolving enzymes, proteins and other characteristics that enable them to remain active and functional. Enzymes from such organisms can be candidate-enzymes to be applied in industrial processes that occur under conditions similar to that of their environmental surroundings (Vieille and Zeikus, 2001; Kirk *et al*., 2002; Straathof *et al*., 2002; Cherry and Fidantsef, 2003; Van den Burg, 2003; Robertson and Steer, 2004).

Since ancient times, man has exploited evolution in a directed way to produce plants and animals with useful characteristics. The cross-breeding of individuals with favourable traits successfully harnesses sexual recombination, one of the most powerful evolutionary strategies to generate new variants. From these crossings, progeny with improved features are chosen for additional breeding cycles, thus channelling the course for future development (Tobin *et al*., 2000; Dalby, 2003).

Proteins are extremely complex and are usually hovering on the precipice of instability. Because they are fairly well-designed to begin with, it is unfortunately much easier to damage one than it is to improve it. And even if the designing of one trait is successful (e.g., enhanced stability), it is often impossible to predict the cost to another (e.g., catalytic activity). All these significant hurdles, which merely reflect the dismally sparse understanding of protein structure-function, are bypassed with evolution (Arnold, 1998; Chatterjee and Yuan, 2006).

Through the pioneering work of Chen and Arnold (1993) and Stemmer (1994a; 1994b) powerful new tools were introduced in the mid-1990s that significantly enhanced existing methods of mutation and screening. These new approaches encouraged the broader acceptance of directed evolution to modify enzymes for application as biocatalysts. Although SDM of proteins contributed significantly to the understanding of protein structure and function, the technique was not designed as a tool for generating enzymes with specific characteristics in a random, selection-driven manner. Directed evolution combines two independent but powerful technologies, *viz*., methods for the generation of random mutagenic libraries and strategies for the screening or selection of variant enzymes that possess specific characteristics, especially increased catalytic activity, enhanced selectivity or improved stability (Tawfik, 2003; Turner, 2003). Point mutation and recombination are the two primary processes for the generation of genetic diversity. Iteration of a directed evolution cycle facilitates the accumulation of beneficial mutations and the elimination of deleterious ones as seen in Fig. 1.3.

Directed evolution differs from natural evolution in two key aspects. Firstly, natural evolution occurs under multiple and variable selection pressures, whereas directed evolution is accomplished under controlled selection pressure for predetermined functions. Secondly, in directed evolution, 'non-natural' functions, of practical use, can be obtained through the design of appropriate selection schemes, whereas natural evolution favours functions advantageous to the survival of the organism (Chatterjee and Yuan, 2006).

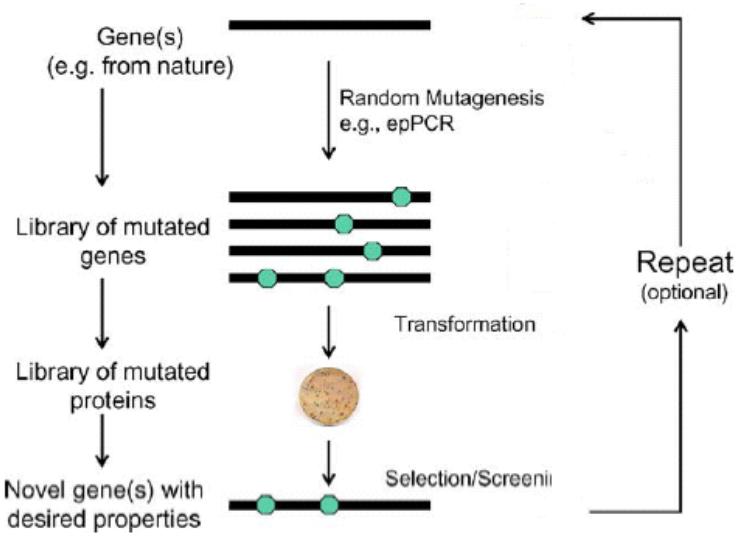


**Fig. 1.3** A typical scheme for directed evolution (Percival-Zhang *et al*., 2006).

### 1.5.1 Requirements for successful directed evolution

Successful directed evolution has four basic requirements:

(i) the desired function should be physically feasible;

(ii) the function should be biologically or evolutionary feasible (i.e., a mutational pathway exists to get from the start to end point through ever-improving variants);

(iii) it should be possible to make libraries of mutants complex enough to contain rare beneficial mutations (this usually means functional expression in a suitable microorganism); and

(iv) a rapid screen or selection that accurately reflects the desired function should be available (Kuchner and Arnold, 1997a; 1997b).

Ultimately, perhaps the second requirement - getting from the start to end point – is the most restrictive on directed evolution. Test tube evolution will not easily yield new protein folds, for example, not because those folds do not exist, but because there do not appear to be good evolutionary paths from one fold to another (Arnold, 1998; Arnold *et al*., 2001).

### 1.5.2 Generating sequence diversity

Mutations are often described as errors but, from the perspective of evolution, perhaps the most serious error is for a genome to be too "perfect". Fidelity of replication and repair is necessary to reproduce a genome but diversity among descendants may increase the chance that some will survive or even flourish, as the environment presents challenges and opportunities. Genomes that encode multiple mechanisms to generate genetic diversity may thus have a selective advantage (Caporale, 2000). Survival of descendants carries the genome forward in time.

Efforts to develop and optimize techniques for introducing point mutations randomly into DNA segments facilitated many of the first experiments in directed enzyme evolution. The quality of a mutant library is decisive for the success of a directed evolution experiment and many methods have been developed for generating diversity at the gene level. These methods differ significantly in the mutational spectra and mutational frequencies and are differently affected by the redundancy of the genetic code. A widely accepted way of assessing bias in mutational spectra is to analyze the ratio of transitions (interchanges between the pyrimidines cytosine and thymine or between the purines adenine and guanine) to transversions (interchanges between purines and pyrimidines). An ideal method would allow substituting every amino acid of a protein sequence by its 19 counterparts in a statistical manner. For a protein engineer developing a directed evolution strategy or performing a directed evolution experiment, it is important to know which amino acid substitutions can be generated on the protein level at positions that have been identified as important by sequence alignments or rational design studies of proteins (Woese, 1965; Myers *et al*., 1985; Vartanian *et al*., 1996; Patel *et al*., 2001; Di Guilio, 2005; Wong *et al*., 2006).

Random mutations have been introduced continuously along a stretch of DNA using error-prone PCR (ep-PCR), chemical mutagenesis, ultraviolet (UV) irradiation, mutator strains, and degenerate oligonucleotides concomitantly with DNA recombination (Singer and Kusmierek, 1982).

#### *1.5.2.1 Chemical mutagenesis and UV irradiation*

Generation of mutations by directly damaging DNA with chemical and physical agents has been used to dissect biological systems for many years. The basis of mutagenesis by UV irradiation, which generates primarily cyclobutane dimers and photo products at adjacent pyrimidine bases or alkylating agents, is that the damaged DNA is incorrectly replicated or repaired leading to mutation. While these methods were among the earliest used in evolutionary engineering, they have been largely superseded by the more modern techniques described below. Chemical mutagens used include sodium bisulphite (Shortle and Botstein, 1983), nitrous acid (causes oxidative deamination of particular bases), formic acid, hydrazine (Myers *et al*., 1985) or hydroxylamine (Busby *et al*., 1982; Grompe *et al*., 1989; Montandon *et al*., 1989). Newer alkylating agents currently in use include ethyl methane sulfonate (EMS), methyl methane sulfonate (MMS), diethylsulfate (DES), and N-Methyl-N'-nitro-N-nitrosoguanidine (NTG). These mutagens tend to prefer G-rich regions, reacting to form a variety of modified G residues, the result often being depurination (Singh *et al*., 1995; Wei *et al*., 1999; Vorobjeva *et al*., 2001). All the methods listed above do not yield highly mutagenized fragments because they target the entire cell and are not gene-specific. Furthermore, they cannot generate all the possible base substitutions. These chemical mutagens and UV irradiation also poses a great risk to the technicians working with them.

#### *1.5.2.2 Mutator strains*

The idea of relaxing the usually very high fidelity of DNA replication is also exploited in mutator strains. These bacterial strains have defects in one or several DNA repair pathways leading to a higher mutation rate. Alternatively, some strains may possess a modified polymerase with lower fidelity to generate sequence diversity. Genetic material that passes through these cells accumulates mutations at a vastly higher rate than usual. This is an effective and straightforward way of introducing mutations throughout a DNA construct (Nguyen and Daugherty, 2003).

However, in common with chemical and physical mutagens, the mutagenesis is indiscriminate. Thus the construct carrying the gene of interest as well as the gene

itself, and indeed the chromosomal DNA of the host cell, suffers mutation (Greener and Callahan, 1994). The process of mutagenesis using mutator strains can be quite slow as the level of mutagenesis is controlled by the length of the time the DNA spends in the strain. Constructing a library with mutagenesis levels of one or two nucleotide changes per gene can require multiple passages through the mutator strain. It is primarily this second disadvantage that has lead to the almost universal use of ep-PCR methods for the generation of diversity for directed evolution experiments (Nguyen and Daugherty, 2003).

Nevertheless, this method is chiefly attractive because it can obviate the labour of subcloning. The sequence which is subjected to mutations can be limited by locating the gene encoding the protein of interest on plasmids (Bornscheuer, 1998). Furthermore, the simplicity of mutator strains will appeal to groups entering the area of directed evolution, particularly those with less experience in molecular biology. They may also be the most appropriate methodology when a simple initial experiment is required to generate preliminary results. The *E. coli* XL1 Red strain, commercially available from Strategene, has been used in most experiments that utilize this strategy (Bornscheuer *et al*., 1999; Alexeeva *et al*., 2002).

The main disadvantage of the above 3 methods is their indiscriminate targeting of the entire genome (or at least an entire plasmid), and this makes it rather unlikely that the screen for function will be influenced by non-specific effects, like altered expression rates, or even the spontaneous modification of cellular enzymes to perform the task that is being screened for (Steipe, 2000).

#### *1.5.2.3 Degenerate oligonucleotides*

Another efficient approach for mutagenesis of DNA involves the synthesis of degenerate oligonucleotides by mixing nucleotides at each step. Accordingly, highly mutated synthetic oligonucleotide libraries can be prepared. Such a strategy is almost ideal in that it allows the researcher to:

(i) precisely define the mutagenized window;

(ii) mutagenize each nucleotide with a given probability; and

(iii) select the bases introduced at each position.

However, this method is limited to DNA fragments containing no more than 100 residues (Fromant *et al*., 1995).

The simplest and oldest approach is to use equimolar mixtures of all four nucleotides, for the codons that are to be changed. In procedures that involve the synthesis of a complementary strand, a bias for incorporating the original nucleotide will arise from the preferential hybridization of oligonucleotides that form larger Watson-Crick base pairs (Reidhar-Olson *et al*., 1991). This bias can be eliminated by reducing the concentration of the wild-type nucleotide during synthesis (Chirumamilla *et al*., 2001).

#### *1.5.2.4 Ep-PCR*

The polymerase chain reaction (PCR) has been employed extensively in the medical and biological sciences since its invention. It has caused a revolution in molecular biology (Bloch, 1991; Mullis *et al*., 1994). The crucial ingredient to the brilliance of PCR is the DNA polymerase, *Taq*, isolated originally from the thermophilic bacterium *Thermus aquaticus*. The fidelity of *Taq* DNA polymerase is influenced by different concentrations of additives, mostly bivalent metal ions, and deoxynucleoside triphosphates (dNTPs) along with the temperature program. Magnesium ($Mg^{2+}$) is the physiological co-factor of *Taq* DNA polymerase and the concentration of this bivalent cation is often varied in order to change the amplification specificity. The concentration of dNTPs and $Mg^{2+}$ were found to increase the mutation rate significantly. However, the former had a lower effect than the latter on mutation rate (Eckert and Kunkel, 1990). Beckmann *et al*. (1985) reported that the additional presence of non-physiological metal ions like manganese ($Mn^{2+}$) within a PCR reaction resulted in deficient *Taq* DNA polymerase activity. Leung *et al*. (1989), used this ground-breaking research to design the ep-PCR protocol; the first stride in the sphere of directed evolution. It remains perhaps the most important method of asexual evolution today.

This method yielded mutation frequencies higher than two percent and these were the first researchers to show that mutagenic rates can be controlled by varying the concentrations of $Mn^{2+}$ and dNTPs. Unfortunately, the presence of these poisoning

ions can considerably decrease the amount of PCR product obtained and this method exhibits a significant bias for transitions over transversions. The protocol of Cadwell and Joyce (1992) diminishes this bias but does not eliminate it. It was subsequently found that unbalanced concentrations of the four dNTPs result in more substitutions. Thus, an excess of CTP and TTP over ATP and GTP predictably leads to more transversions. By using $Mn^{2+}$ and unbalanced nucleotide concentrations, the overall mutagenesis becomes more random and given access to a higher sequence space. The effects of these parameters are different on every template and have to be established for each gene to be mutated (Shafikani *et al*., 1997).

Other ep-PCR protocols are based on the ability of *Taq* polymerase to incorporate universal bases like deoxyinosine triphosphate (Spee *et al*., 1993; Xu *et al*., 1999) or derivatives of nucleoside analogues like 8-oxo-dGTP (Pavlov *et al*., 1994). Such analogues can pair with different bases, thereby introducing all mutations possible. In the next PCR cycle, other natural nucleotides are added as a complement to the analogues. Thus, by using different additives in a PCR reaction, it becomes possible to choose a mutation type: either transitions or transversions.

Apart from its mutational bias, ep-PCR allows a series of mutations to be made, which differ in mutation frequency, simply by altering PCR conditions. Mutation frequency is an important parameter in directed evolution. More mutations increase the possibility of finding a totally new activity, but also disrupt the overall structure of the enzyme. A mutation frequency of 1 - 3 mutations per 1000 bp is commonly regarded as a good rate for finding improvements of the desired activity of an enzyme. Hypermutation is thought to result in too many non-functional mutants and is therefore hardly ever used. Ep-PCR is highly appropriate for most experiments in which the relationship between the protein structure and function is unclear. Another bottleneck when using this procedure is the often observed low ligation efficiency during PCR (Bornscheuer, 1998).

An important problem with ep-PCR and random mutagenesis of a gene is that mutations neutral to the function in question, but not necessarily neutral to other functions not part of the selection scheme, will not be selected against. Most of such mutations will be crossed out if the initial screening of a random library is followed

by a round of gene shuffling and further selection which, however, increases the necessary screening effort quite considerably (Leung *et al*., 1989).

#### *1.5.2.5 Commercial random mutagenesis kits*

Due to the rapid and timeous advancement of directed evolution, many biotechnological companies have simplified random mutagenesis for molecular biologists that make use of such techniques in a bid to improve the enzymes which they work with. A few of these kits are discussed below.

Diversify$^{TM}$ PCR Random Mutagenesis kit (CLONTECH) provides a range of buffer conditions that vary the concentrations of $Mn^{2+}$ and dGTP for performing random mutagenesis allowing the researcher to adjust the reaction for a desired error rate ranging from 2 - 8 mutations per 1000 bp. TITANIUM *Taq* is used in the kit. This system outperforms other PCR systems under the stress of ep-PCR conditions. The kit is reputed to have a wide mutational diversity to produce transition and transversion mutations.

The GeneMorph system (Strategene) is based on a highly error-prone *Taq* polymerase. This kit is straightforward to use and comes with detailed instructions and is therefore appealing to those entering the area. The level of mutagenesis is controlled by the concentration of template used and the number of serial amplification reactions performed.

QuikChange® Multi site-directed mutagenesis kit (Strategene) can introduce multiple point mutations simultaneously in a gene of interest and has a special enzyme blend containing *PfuTurbo*® DNA polymerase. The kit features a one-day, three-step procedure using a single mutagenic oligonucleotide to mutagenize each site in a double-stranded DNA template. It relies on the use of saturation mutagenesis for a more comprehensive study of structure-function relationships. This method accesses a larger faction of protein sequence space and it introduces all possible mutations at key sites or adjacent sites. *PfuTurbo*® DNA polymerase minimizes unwanted errors during mutant strand synthesis while delivering high mutagenesis efficiency.

Mutation Generation System (Finnzymes) is designed for rapid construction of insertion mutation libraries for any kind of DNA clones. The system employs the highly efficient transposition machinery of the bacteriophage Mu to generate a pool of 15 bp inserts. Some of the documented advantages include random insertions of five amino acids in all three reading fames; short in-frame insertions with no stop codons with PCR processes patented by Hoffman-La Roche.

GPS$^{TM}$–M Mutagenesis System (New England Biolabs) enables the insertion of a self-designed transposable element into the target DNA by *in vitro* transposition whilst LA PCR Mutagenesis kit (Takara Mirus Bio) is an improved system designed to introduce a series of site-directed mutations into long DNA fragments cloned in pUC or pUC-derived vectors containing the pUC multiple cloning site. This system takes advantage of LA (long and accurate) PCR technology by including the Takara LA Taq$^{TM}$ polymerase along with kit-supplied primers to provide an easy and efficient way to generate mutants in longer DNA fragments with high fidelity.

Although many elegant random mutagenesis kits are available on the market, the overall cost of producing an enzyme with industrially-relevant properties is what governs the final choice of kit for many researchers. What remains clear is that conventional mutagenesis with $Mn^{2+}$ is no longer the only option for generating improved biocatalysts.

#### *1.5.2.6 DNA recombination*

Accumulating point mutations is an effective fine-tuning mechanism, but nature also uses other means to create new molecular diversity on which evolution can act. One of those is recombination. Studies have shown that recombination is an extremely useful operation for laboratory evolution. In contrast to ep-PCR, sexual evolution starts from a pool of homologous parent genes. These can originate from asexual evolution, from related natural sequences, or from enzyme variants generated by rational design. The DNA shuffling method developed by Stemmer (1994a; 1994b) for random *in vitro* recombination represents a significant benchmark in the scientist's ability to mimic the natural mechanism of evolution since recombination allows the rapid accumulation of beneficial mutations identified in separate genes. Genes from

multiple parents and even from different species can be shuffled in a single step, operations that are forbidden in nature but may be very useful for rapid adaptation.

Screening protocols, for example, often produce multiple improved sequences. The beneficial mutations normally discarded when a single clone is chosen from these to parent the subsequent generation can be efficiently recombined by DNA shuffling. In addition, deleterious mutations are removed. Recombination is also a promising approach to combining two or more improved properties evolved separately (Kuchner and Arnold, 1997a; Bergquist *et al*., 2005). Thus DNA shuffling is used widely to generate highly improved biocatalysts, as well as ones with features not present in the parent enzymes and not known to occur in nature.

The principle of DNA or gene shuffling is based on the original procedure which involved partial DNase-I digestion of the gene followed by a recombination of fragments by PCR (Fig. 1.4). Again, an accumulation of positive and an elimination of negative mutations are desired. Disadvantages are that homologous genes are not always available, the DNase digestion for the generation of fragments of suitable size is difficult to control, and there are problems with ligation efficiency after reassembly PCR (Bornscheuer, 1998).

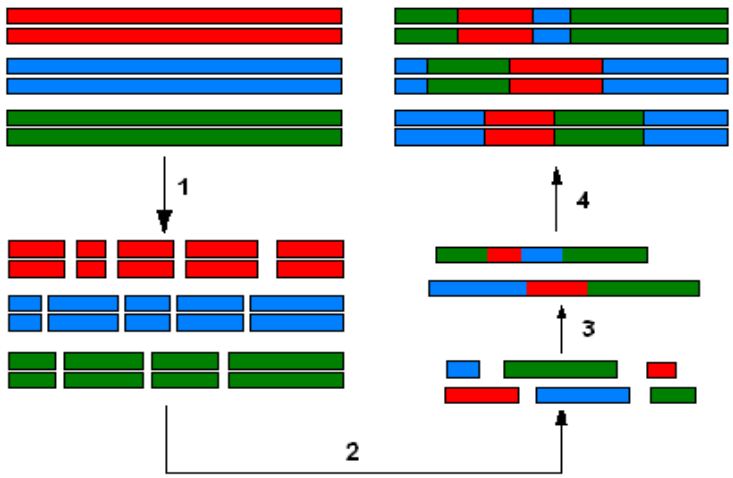


**Fig. 1.4** A simple description of DNA shuffling. Similar or related genes are mixed with the gene of interest. (1) Genes are digested separately with DNase1. (2) The fragments are denatured and mixed in differing ratios and allowed to form heterologous duplexes (preparative PCR). (3) dNTPs and *Taq* polymerase are added along with primers specific for gene ends. (4) PCR is carried out to amplify chimeric genes (Stemmer, 1994a; 1994b).

These disadvantages lead to the development of more sophisticated methods of shuffling such as the staggered extension process (StEP) (Zhao *et al*., 1998), exon shuffling (Gibbs *et al*., 2001; Kolkman and Stemmer, 2001) and recombination on truncated templates (RETT) (Lee *et al*., 2003), to name a few, and the list of new DNA shuffling technologies continues to grow. Some of the types of DNA shuffling are listed in Table 1.2.

**Table 1.2** Principles of different DNA shuffling techniques

| Group | Technique | Members | Reference |
|---|---|---|---|
| Shuffling | • Recombination of small fragments based on homology in the sequence between mutations that stem from all kinds of mutagenesis strategies of different family members<br>• Aims for high recombination but difficult to separate close mutations | Shuffling<br>Family shuffling<br>Family shuffling (ssDNA)<br>RPR<br>RETT<br>SCRATCHY | Stemmer, 1994a; 1994b<br>Crameri *et al*., 1998<br>Kikuchi *et al*., 2000<br>Shao *et al*., 1998<br>Lee *et al*., 2003<br>Lutz *et al*., 2001 |
| Full length parent shuffling | • Recombination of small fragments of different origin using one or more full length parent strands<br>• Higher recombination frequency but more elaborate | StEP<br>RACHITT | Zhao *et al*., 1998<br>Coco *et al*., 2001 |
| Single cross-over | • Recombination of non-homologous genes by ligating front and back of two different genes, selection of new genes based on size<br>• Recombination possible between low or non-homologous genes, but only one recombination point | ITCHY<br>SHIPREC<br>SCRATCHY | Ostermeier *et al*., 1999<br>Sieber *et al*., 2001<br>Lutz *et al*., 2001 |
| Domain swapping | • Recombination of structural, functional or less homologous parts of different family members<br>• More active enzymes in the resulting library, but only a few recombination points which are hard to find | Exon shuffling<br>DOGS | Gibbs *et al*., 2001<br>Kolkman and Stemmer, 2001 |
| *In vivo* recombination | • Recombination using the gap repair system of yeast or the recE/recT system of *E. coli*<br>• High yield, since no ligation necessary, but specialized vectors and multiple steps necessary | CLERY | Abecassis *et al*., 2000 |
| Synthetic shuffling | • Recombination of (un)known mutations in synthetic oligonucleotides<br>• Recombination of close mutations possible but expensive and good selection necessary | Single step shuffling | Stemmer *et al*., 1995 |

### 1.5.3 Screening for improved variants

The cornerstone of successful protein improvement is the correlation between the enzyme screen and the final application. In the best case, a screen is easily performed at small scale and with high-throughput, yet accurately reflects the stress of an industrial application. Creating effective screening protocols begins with a thorough biochemical understanding of what is required for the success or failure of an enzyme to perform adequately in an application, and often requires multiple rounds of application testing and screen modification. Advances in high-throughput screening include an interesting method utilizing substrate arrays for enzyme fingerprinting (Cohen *et al*., 2001; Reymond and Wahler, 2002).

The number of enzyme variants which can be generated by directed evolution grows exponentially with the size of the enzymes and the number of simultaneously exchanged amino acids. Even for a small protein of 200 amino acids, more than 9 billion possible variants can theoretically be generated simply by introducing 3 substitutions at the same time (Kuchner and Arnold, 1997b). These cannot be analyzed with common tedious and protracted methods such as high performance liquid chromatography (HPLC) or gas chromatography. Thus an efficient screening system is of extreme importance.

A selection can be based, for instance, on altered antibiotic resistance (Crameri *et al*., 1996) or growth on media lacking components essential for growth (Bornscheuer *et al*., 1998; MacBeath *et al*., 1998). Problems related to organic synthesis require assay systems which provide direct information about the enzyme properties. For example, the identification of a lipase variant with improved stereo-selectivity is possible by measuring the ratio of the rates of hydrolysis of enantiomerically pure (*R*)- or (*S*)-*ρ*-nitrophenol esters (Reetz *et al*., 1997). There are a myriad of other elegant screening technologies available based on colorimetry (Taylor *et al*., 1999) and fluorometry (Sundberg, 2000).

One of the chief problems of screening is that the enzyme library producing colonies cannot be assayed directly because the enzymes are, for example, produced intracellularly and a cell lysis is necessary. Additional prerequisites for the detection

of all variants are correct processing and folding to ensure the production of active enzymes. Furthermore, all enzyme variants should have high stability and activity under assay conditions. The directed evolution of enzymes is particularly dominated by biochemistry.

### 1.5.4 Commercial successes of directed evolution

Subtilisin (Bryan, 2000; Ness *et al*., 2002), hydantoin (May *et al*., 2000) and both Kannase® and Lipoprime® from Novozymes were some of the earliest published examples of enzymes that have been successfully evolved to benefit an industrial process or result in a commercial product. This number has grown exponentially over the last few years. The enzymes listed below have been developed using directed evolution, DNA shuffling and/or SDM. Some of the enzymes developed to date by Diversa Corporation for commercial use includes the Luminase brand of xylanases to enhance pulp bleaching; Purifine to increase the efficiency of vegetable oil processing; the vaccine Bayovac SRS to treat disease in farmed salmon and Phyzyme to enhance the nutritional value of animal feed and improve feed conversion efficiency (Diversa, 2006).

Iogen Corporation, one of Canada's leading biotechnology companies commercially produces Roxazyme G2 to improve and increase animal feed efficiency; the BioBrite® brand of enzymes used to bleach paper pulp; Denabraide® used in denim processing to substitute or enhance the performance of pumice stones which fade and soften fabric and Fibrilase® used to soften fabrics and remove or prevent fuzz on garments (Iogen, 2006).

Globally, Novozymes has almost a 60% share in the industrial enzyme market and has commercially produced over 100 types of enzymes and microorganisms and has not less than 700 different products currently out in the industrial enzyme market. Some of its products include PondPlus® and PondProtect® used to degrade organic pond waste; four enzymes including Vinozyme that facilitate maceration, clarification and maturation of grapes in wine; lecitase Ultra and Lipozyme TL IM to improve oil processing; eight enzymes including Lactozym involved in the processing of dairy products; six enzymes including Alcalase for ethanol bioconversion and enzymes

used in detergent formulations. The latter category has almost 50 patented products and sales from this category form the bulk of the revenue for Novozymes (Novozymes, 2006).

Approaches combining enzyme engineering techniques and screening in microplate systems that mimic the ideal conditions experienced in industrial biocatalytic processes are now being used to create such superior catalysts and a few examples of enzymes en route to commercialization are documented in Table 1.3.

**Table 1.3** Summary of a few examples of directed evolution*

| Enzyme | Mutations in evolved enzyme | | Enhancement factor | Method | References |
|---|---|---|---|---|---|
| | Total | Residue function | | | |
| 2-Hydroxybiphenyl-3-monoxygenase | 1 | substrate binding | 1.3 | Ep-PCR | Meyer *et al*., 2002) |
| Cytochrome P450 BM-3 | 11 | One substrate binding Five in active site lid | 2-108 | Ep-PCR/StEP | Glieder *et al*., 2002 |
| Benzoylformate decarboxylase | 1 | Active site gatekeeper and product release | 5 | Ep-PCR/SM | Lingen *et al*., 2002 |
| Glutathione-*S*-transferase | 1 | Helix flexibility and product release | 12.6 | DS – two parents | Broo *et al*., 2002 |
| Catalase | 3 | R47H accounts for activity change and interacts with catalytic haem centre R47W | 7.3 | DS/Ep-PCR | Ni *et al*., 2002 |
| Esterase | 8 | C-terminal domain | 10 | DS/Ep-PCR | Giver *et al*., 1998 |

* The examples shown all use established methods alone or in combination with other methods, such as error-prone PCR (Ep-PCR), DNA shuffling (DS), staggered extension process (StEP) or saturation mutagenesis (SM). The enhancement factor shown is relative to the wild-type enzyme.

### 1.5.5 Future prospects for directed evolution

The enzyme market and number of competitive enzyme-based processes is growing rapidly because of cheaper production methods, new application fields and new enzymes. The possibility to dramatically change enzyme properties by directed evolution and gene shuffling, and efficient methods to screen for new enzymes in the environment makes it feasible to use enzymes that are specifically tailored to their application and process conditions. Enzyme technology is close to a major breakthrough, due to many factors ranging from simple cost savings, the strongly

increasing demand for chiral chemicals, the trend towards sustainable industrial development (less waste, less carbon dioxide), and last, but not least, the opportunities created by emerging technologies (Schoemaker *et al*., 2003).

In directed evolution experiments, enzymes with altered properties are discovered frequently to have acquired these properties as a result of the alteration of the amino acids in such a way that would have been difficult to predict or even to justify using protein crystal structure alone. This highlights the limitation in the understanding of protein structures and of the ability to rationally design desired biocatalysts. However, directed evolution is also limited by its inability to exhaustively search the vast sequence space of proteins. Thus, the future developments of many enzyme biocatalysts might depend upon the successful amalgamation of rational design and directed evolution approaches (Arnold, 2001).

Directed evolution techniques are uncomplicated and have been successfully applied to the optimization of a range of protein activities including binding, stability and enzyme selectivity. The challenge now lies in pushing back the boundaries of what can be achieved using directed evolution. Tackling these challenges may require the construction of new types of libraries. A large toolkit of methods is available that makes possible the construction of novel and highly efficient libraries. These methods are necessarily more complex than ep-PCR, DNA shuffling and oligonucleotide-based mutagenesis, and are therefore unlikely to be the first choice of the general user (Arnold, 1996; 1998; 2001). However, they are likely to come into their own when the simpler methods have failed in the Holy Grail quest to tailor-make the perfect enzyme. They will also be valuable for those groups working to develop optimized and rational approaches to directed evolution. It is not, as yet, clear which of these new techniques will be most valuable or most popular. But, the ability to generate and combine such a wide range of sequence diversity is an important staging post en route to developing the full promise of the technology of directed evolution.

## 1.6 XYLANASES

### 1.6.1 Xylan: occurrence and structure

Plant biomass comprises on average 23% lignin, 40% cellulose and 33% hemicellulose by dry weight. Xylans or the hemicelluloses are situated between the lignin and the collection of cellulose fibres underneath. Consistent with their structural chemistry and side-group substitutions, the xylans seem to be interspersed, intertwined and covalently linked at various points with the overlying sheath of lignin, while producing a coat around the underlying strands of cellulose (Biely, 1985) via hydrogen bonding (Joseleau *et al*., 1992). The xylan layer with its covalent linkage to lignin and its non-covalent interaction with cellulose may be important in maintaining the integrity of cellulose *in situ* and in helping protect the fibres against degradation (Fig. 1.5) (Uffen, 1997).

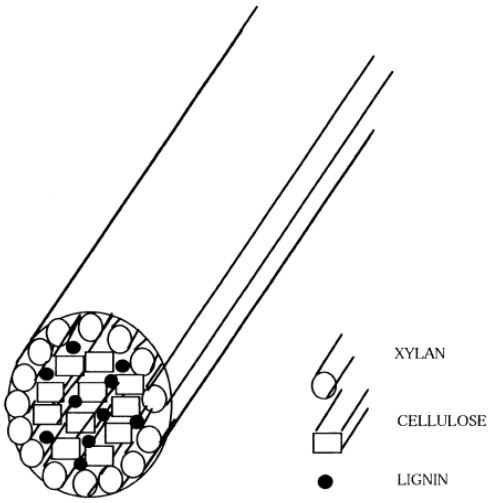


**Fig. 1.5** A schematic representation of the association between xylan, cellulose and lignin in a typical wood fibre (Techapun *et al*., 2003).

The term ‘hemicellulose’ was used to describe the fractions isolated or extracted from plant materials with dilute alkali (cited by Saha, 2003). Hemicelluloses are heterogeneous polymers of pentoses (xylose, arabinose), hexoses (mannose, glucose, galactose) and sugar acids. Unlike cellulose, which is an unbranched glucose polymer composed of D-glucose units linked by 1,4-β-D-glycosidic bonds, hemicelluloses are not chemically homogenous. Hardwood hemicelluloses contain mostly xylans, whereas softwood hemicelluloses contain mostly glucomannans (McMillan, 1993).

Xylans of many plant materials are heteropolysaccharides with homopolymeric backbone chains of 1,4-linked β-D-xylopyranose units. Besides xylose, xylans may contain arabinose, glucuronic acid or its 4-*O*-methyl ether, and acetic, ferulic and *ρ*-coumaric acids (Fig. 1.6). The frequency and composition of branches are dependent on the source of the xylan (Aspinall, 1980).

Xylan is the major hemicellulose in hardwood from angiosperms, but is less abundant in softwood from gymnosperms. Hardwood xylans are highly acetylated (e.g., birchwood xylan contains more than 1 mol of acetic acid per 2 mol of xylose). The presence of these acetyl groups is responsible for the partial solubility of xylan in water. These acetyl groups are readily removed when xylan is subjected to alkali extraction (Wong *et al*., 1988; Sunna and Antranikian, 1997). Xylans from softwood have a higher 4-*O*-methylglucuronic acid content than do hardwood xylans. They are unacetylated and have instead α-L-arabinofuranose units linked by α-1,3-glycosidic bonds at the C-3 position of the xylose (Puls and Schuseil, 1993).

### 1.6.2 Enzymatic hydrolysis of xylan

Due to the heterogeneity and complex chemical structure of plant xylan, its complete breakdown requires the action of a complex of several hydrolytic enzymes with diverse specificity and modes of action. Thus it is not surprising for xylan-degrading microorganisms to produce a variety of polymer-degrading proteins. Xylan hydrolysis usually requires a repertoire of hydrolytic enzymes: β-1,4-endoxylanase, β-xylosidase, α-L-arabinofuranosidase, α-glucuronidase, acetyl xylan esterase, ferulic and *ρ*-coumaric acid esterases (Fig. 1.6). All these enzymes act co-operatively to convert xylan into its constituent sugars. The presence of such a multifunctional xylanolytic system is quite widespread among fungi (Biely *et al*., 1986; Chadha *et al*., 2004), actinomycetes (Elegir *et al*., 1995) and bacteria (Collins *et al*., 2005).

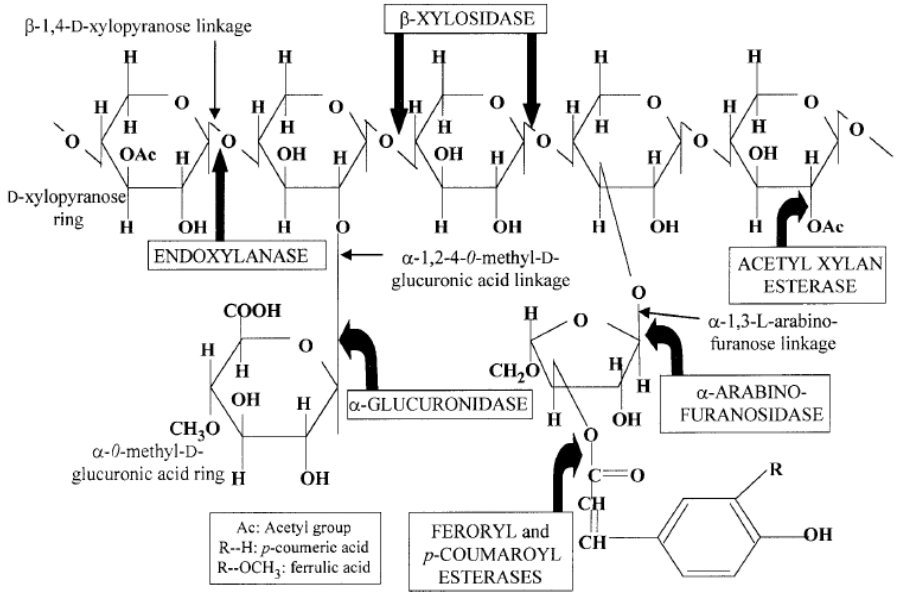


**Fig. 1.6** A hypothetical plant xylan structure showing different substituent groups with the sites of attack by the microbial xylanolytic system (Beg *et al*., 2001).

### 1.6.3 Classification of xylanases

Endoxylanases (EX) have been classified into families 5, 8, 10, 11 and 43 of glycoside hydrolases, based on hydrophobic cluster analysis and similarities in their amino acid sequences (Coutinho and Henrissat, 1999). Members of the two largest and best known families are the family F/10 and family G/11 xylanases (Jeffries, 1996). Sequence alignment suggests that xylanases evolved from two ancestral proteins and therefore can be grouped into two large families designated F and G. These families are analogous to glycohydrolase families 10 and 11 (Henrissat and Bairoch, 1993). A common feature of both enzyme families are their endo-acting characters that are demonstrated viscometrically and the double displacement mechanism of the hydrolysis of the glycosidic bond, which means that both types of enzymes are retaining glycanases (Withers, 2001).

F/10 xylanases consist of a cellulose binding domain and a catalytic domain connected by a flexible linker region (Fig. 1.7) and they generally have a higher molecular weight (Henrissat and Bairoch, 1993). They exhibit lower substrate specificity and hence greater catalytic versatility, because they are capable of cleaving glycosidic linkages in xylan main chains closer to side-chain substituents. As a result,

these EX release short xylooligosaccharides upon hydrolysis (Kolenová *et al*., 2006). Crystallization of xylanases from this family indicates that the tertiary fold is a typical 8-fold α/β barrel $(\alpha/\beta)_8$ resulting in a 'salad bowl' shape of the molecule (Derewenda *et al*., 1994). Recent studies show the presence of histidine in the active site of F/10 xylanases (Liu *et al*., 2004a). The substrate binding sites are shallow clefts. This feature, together with a possible greater conformational flexibility of the larger enzymes may account for the lower substrate specificity observed with members of this family. F/10 xylanases are more thermostable than their G/11 counterparts and this property is attributed to improved hydrophobic packing (Lo Leggio *et al*., 1999).

Members of xylanase family G/11 are commonly referred to as true xylanases because of their inability to hydrolyze cellulose (Jeffries, 1996). They consistently exhibit a low molecular weight and appear very small in native state. G/11 xylanases are highly specific and preferentially cleave in unsubstituted regions of the xylan backbone (Faulds *et al*., 2005) partly because they have deep clefts in the substrate binding sites (Fig. 1.7). These xylanases are well-packed and form mainly β-sheets (Törrönen *et al*., 1994). The catalytic groups are present in the cleft that accommodates a chain from five to seven xylopyranosyl residues. Thermostability in G/11 xylanases is attributed to the accumulation of several minor modifications such as the increase of charged residues and the stabilization of secondary structures (Hakulinen *et al*., 2003).

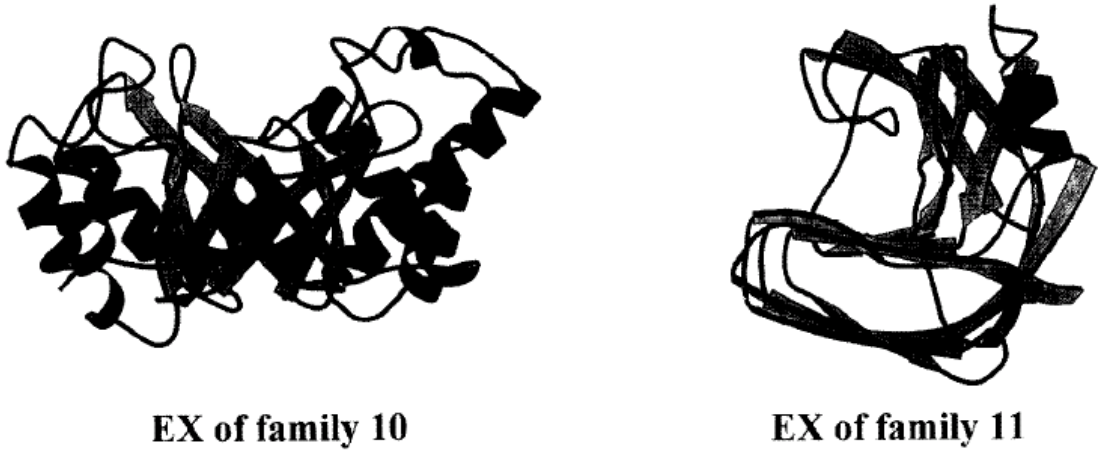


**Fig. 1.7** Ribbon representations of the main fold of the catalytic domains of the endoxylanases (EX) of family 10 and 11 (Davies and Henrissat, 1995).

### 1.6.4 Commercial application of xylanases

#### *1.6.4.1 Animal nutrition*

Xylanases have shown an immense potential for increasing the production useful products in a most economical fashion. Incorporation of xylanase into a rye-based diet of broiler chickens results in reduced intestinal viscosity, thus improving both their weight gain and their feed conversion efficiency (Bedford and Classen, 1992; Malathi and Devegowda, 2001).

The use of fibrolytic enzymes for ruminant diets has attracted interest, following trials in which positive responses in nutrient digestion and animal performance was observed (Beauchemin *et al*., 1995; Schingoethe *et al*., 1999). In another study, alfalfa hay and corn silage were the feed substrates evaluated for animal nutrition using preparations of xylanases and cellulases. This study showed that the different modes of action of these exogenous enzymes attack different substrates and may partly explain enzyme-feed specificity. Alfalfa pretreated with the enzyme preparations was better digested by the ruminants indicating that the enzymes remove structural barriers that retard the microbial colonization of digestible fractions within the rumen. In corn silage however, the enzymes performed better when they interacted with ruminal enzymes and degraded the forage more rapidly (Colombatto *et al*., 2003).

Some commercial xylanases used in the feed industry today include Biofeed Beta and Biofeed Plus (Novozymes, Denmark), Sanzyme PX and Alpelase F (Sankyo, Japan), Enzeko xylanase (Enzyme Development, USA), Roxazyme G2 (Iogen Corp., Canada) and Phyzyme XP and Quantum phytases (Diversa Corp.)

#### *1.6.4.2 Baking*

Xylanases have also been shown to improve the quality of bread with an increase in specific bread volume. This is further enhanced when used in conjunction with amylase (Maat *et al*., 1992). The acid tolerant xylanase from *Aspergillus oryzae* has been used to commercially produce traditional Japanese foods such as sake (rice wine) and shoyu koji (soybeans and wheat). It was shown to efficiently degrade the

soybean and wheat cell walls thus improving the usage of raw materials and decreasing the amount of pressed cake left following the press-filtration of soy sauce mash which is difficult to dispose of (Kimura *et al*., 2002). Multifect xylanase (Genencor, Finland) and Enzeko xylanase (Enzyme Development, USA) are some of the older commercial xylanases used in the baking industry. Novozymes markets a new range of xylanases, sometimes in combination with other enzymes, for improved dough conditioning, especially in bread manufacture. Some of these commercial xylanase preparations include Celluclast BG, Fungamyl Super AX, Fungamyl Super MA and the Pentopan® brand of xylanases.

#### *1.6.4.3 Wines*

α-L Arabinofuranosidase and β-D-glucopyranosidase have been employed in food processing for aromatizing musts, wines and fruit juices (Spagna *et al*., 1998). The microflora of grapes is highly variable and can contribute to spoilage however; controlled growth of these indigenous species may in some circumstances enhance the chemical complexity and sensory quality of the wines. Xylanases produced by the microflora are purported to assist in degradation of the grape cell walls and in doing so, increase the amount of monoterpenyldiglycoside precursors in the must. These compounds produce aroma active compounds (Lambrechts and Pretorius, 2000; Strauss *et al*., 2001). Currently, no commercial enzyme preparations containing xylanases have been marketed as the potential use of xylanase in wines are still in the research and development phase. The closest commercial enzyme related to xylanase however, is Glucanex (Novozymes, Denmark), a glucanase from *Trichoderma harzianum*, which is used for wine clarification.

#### *1.6.4.4 Surfactants*

Alkyl glycosides are one of the most promising candidates for new surfactants. Commercially, they are produced from monomeric sugars such as glucose and a fatty alcohol. But the direct glycosylation using polysaccharide is much more feasible for their industrial production, because hydrolysis of polysaccharide and subsequent steps can be omitted. Thus, use of xylanase in this process provides a challenging opportunity. Xylanase from *Aureobasidium pullulans* has been used for direct

glycosylation of xylan, 1-octanol and 2-ethyl hexanol into octyl-β-D-xylobioside, xyloside and 2-ethylhexyl-β-D-xylobioside, respectively (Matsumura *et al*., 1999).

#### *1.6.4.5 Biobleaching*

Xylanases largely find application in xylanases is in the pre-bleaching of kraft pulps (Bajpai, 1999). Established pulping processes are relatively inefficient and environmentally costly due to the toxic by-products generated during chemical bleaching with chlorine. Thus, new pulping processes are needed. Enzymes continually play a role in such developments. Biobleaching involves using microorganisms and enzymes to bleach pulp. It relies on the ability of some microorganisms to depolymerize lignin directly and uses microorganisms or enzymes that attack hemicellulose which thus facilitate subsequent depolymerization (Techapun *et al*., 2003).

##### *1.6.4.5.1 Mechanism of biobleaching*

Xylanase attacks hemicellulose and alters the interface between the cellulose and lignin, thereby facilitating the removal of lignin-associated hemicellulosic fraction with minimal damage to the pulp (Fig. 1.8). This process is less drastic, less expensive and especially less toxic than conventional chemical treatment. Besides bleaching through lignin removal the use of xylanases helps increase pulp fibrillation, reduce beating times in the original pulp and increase the freeness in recycled fibres. Xylanase treatment can improve the chemical extraction of lignin from pulp. This consequently leads to a significant reduction in the amount of chemicals required for bleaching and in the levels of toxic chlorine compounds released into the environment.

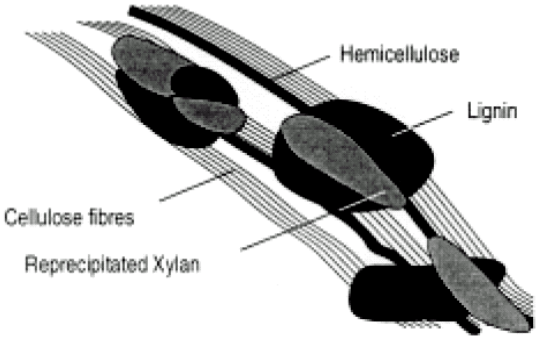


**Fig. 1.8** A hypothetical structure of kraft pulp fibres. Xylanases are thought to promote pulp bleaching via the hydrolysis of re-precipitated xylan (Viikari *et al*., 1994).

#### *1.6.4.5.2 Characteristics of the ideal pulp xylanase*

Since xylanases are used as bleaching additives in the processing of chemical or thermomechanical pulps, it would be desirable that the enzyme meets the requirements of the process rather than the process meeting the needs of the enzyme.

Several features determine which xylanases are effective for bleaching:

(i) the xylanase of interest must be cellulase-free to prevent reduced pulp quality;

(ii) high yields of enzyme must be obtained at very low cost;

(iii) it is desirable to have a low molecular weight xylanase that is able to penetrate the pulp fibre. An enzyme must have access to the substrate if it is going to release chromophoric groups and lignin;

(iv) it must have an alkaline pH optimum and isoelectric point since kraft pulp processes are highly alkaline (pH 8 - 10). Pulp fibres are negatively charged due to the presence of sugar acids, and if an enzyme has an alkaline pI, and hence a positive charge at the operational pH, it will bind more effectively;

(v) the xylanase must be stable at the temperature prevailing in the pulp stock (60 - 90°C) with incubation periods of 3 - 5 h. The combination of alkaline and thermal stability however, has been extremely difficult to obtain from

natural sources and has thus been the focus of genetic engineering efforts for decades; and

(vi) lastly, the xylanase should possess appropriate substrate specificity since the objective is to release the chromophores and extract residual lignin and not remove the bulk xylan. (Viikari *et al*., 1994; Jeffries, 1996; Beg *et al*., 2001; Techapun *et al*., 2003)

Current efforts are aimed at process optimization, simplification and reduced cost of applying xylanases in the pulp industry. The xylanases available for commercial use at present require a pH adjustment of the incoming pulp from pH 10 - 11 to pH 6 - 8 which is necessary for their optimal activity (Table 1.4). From an industrial point of view, it is simple to adjust the pH but difficult to control temperature due to the high expenses incurred for cooling. The ideal solution would therefore be to use xylanases with higher pH and temperature stability, which will make the operations simpler and economical. Thus future research studies are focussed on screening for thermostable xylanases at high alkaline pH and then developing process technologies for commercial-scale manufacture of such enzymes.

**Table 1.4** Characteristics of some commercial pulp bleaching xylanases

| Trade mark name | Distributor | Source | Operational enzyme parameters | Reference |
|---|---|---|---|---|
| Cartazyme | Clariant, UK | *Thermomonospora fusca* | 5% consistency pulp, 35 - 55°C, pH 3 - 5, 2 - 10 h | Bhardwaj *et al*., 1996 |
| Ecopulp X200 | ICI Forest Products, UK | *Trichoderma reesei* | 50 - 55°C, pH 5 - 6 | Vicuna *et al*., 1995 |
| Irgazyme 40 | Nalco-Genencor, Finland; Ciba Giegy, Switzerland, Genencor, USA | *Trichoderma longibrachiatum* | 10% consistency pulp, 50 - 55°C, pH 6.6 - 7.8 1 h, | Vicuna *et al*., 1995; Bhardwaj *et al*., 1996; Wong *et al*., 1996 |
| Resinase | Novozymes, Denmark | ND* | 4% consistency pulp, 40 - 50°C, pH 6.5 - 7 | Haltrich *et al*., 1996 |
| Pulpzyme | Novozymes, Denmark | *Bacillus* sp. V1-4 (isolated from kraft pulp, pH 11) | 10% consistency pulp, 50 - 70°C, pH 7 - 8, 3 h | Yang *et al*., 1995; Bhardwaj *et al*., 1996; Wong *et al*., 1996 |
| Bleachzyme | Biocon, India | ND* | 5% consistency pulp, 50°C, pH 7, 3 h | Bhardwaj *et al*., 1996; |
| Amano 90 | Amano Pharmaceuticals, Japan | *Aspergillus niger* | 5% consistency pulp, 50°C, pH 4.5 - 5, 3 h | Bhardwaj *et al*., 1996; |
| Biobrite | Iogen Corp, Canada | ND* | 5% consistency pulp, 55°C, pH 5 - 6, 3 h | Dhillon and Khanna, 2000 |
| Luminase | Diversa Corp, USA | DNA from soil sample from Kamchatka, Russia | 5% consistency pulp, 50°C, pH 6 - 8 1h | Diversa, 2004 |

*ND – not determined

### 1.6.5 Molecular biology of xylanases

Recent advances in molecular biology and genetic engineering over the last two decades have opened up the areas of application of gene cloning and recombinant DNA technology. Recombinant DNA techniques offer new opportunities for the construction of genetically modified microbial strains with selected enzyme machinery. To ensure the commercial utilization of hemicellulosic residues in the pulp and paper industries, the production of higher xylanase yields at low capital cost is required. In this respect, isolation and cloning of the xylanase gene represents an essential step in the engineering of the most efficient microorganism (Beg *et al*., 2001).

#### *1.6.5.1 Cloning of xylanases*

Although thermophiles can be good candidates in producing thermostable enzymes, it is often impractical because of low yield and the fact that high temperature fermentations may require specialized equipment (Sonnleitner and Fiechter, 1983). As a consequence, a molecular approach through genetic engineering becomes a good alternative towards bulk production economically via prokaryotic and eukaryotic systems. Expression of foreign proteins in prokaryotic systems is the most widely used approach to achieve high-level expression; both for fundamental studies and for commercial purposes (Georgiou, 1996). In addition, the fast growth rate and ease of cultivation technology of *E. coli* make it suitable for industrial application. However, there are two goals needed to be accomplished during gene expression, *viz*., high cell density and high-level gene expression (Lin *et al*., 2001). The expression vector and host are important issues for achieving maximal expression cloned genes. However, molecular cloning of a foreign gene does not ensure that the gene will be expressed successfully (Glick and Pasternak, 1994). The most difficult problems in bacterial expression are proteolytic degradation and the production of proteins that accumulate in misfolded forms, most often inclusion bodies (Georgiou, 1996).

Although yeasts have greater genetic complexity than bacteria and contain 3.5 times more DNA than *E. coli* cells, they have many technical advantages. Some of the properties that make yeast suitable for biotechnological studies include rapid growth,

dispersed cells, the ease of replica plating and mutant isolation, a well-defined genetic system, and most important, a highly versatile DNA transformation system. Being non-pathogenic, yeast can be handled with little precautions (La Grange *et al*., 1996).

Several studies have been made to clone and express xylanase in different bacterial and yeast hosts. Some examples of xylanases that were cloned in *E. coli* include *A. oryzae* (Kimura *et al*., 2002), *A. pullulans* var. *melanigenum* (Ohta *et al*., 2001), *Bacillus lyticus* (Srivastava and Mukherjee, 2001), *Clostridium thermocellum* (Fernandes *et al*., 1999) and *Caldocellum saccharolyticum* (Lüthi *et al*., 1990). Some xylanases such as that from *A. pullulans* var. *melanigenum* (Tanaka *et al*., 2004), *T. lanuginosus* IOC-4145 (Damaso *et al*., 2003) and *A. niger* (Berrin *et al*., 2000) have instead been cloned in *Pichia pastoris*. *Saccharomyces cerevisiae* has also served as a cloning host for the *A. pullulans* (Li and Ljungdahl, 1996) and *T. reesei* (La Grange *et al*., 1996) xylanases.

#### *1.6.5.2 Crystallization and bioinformatics of xylanases*

The identification of important active site residues has been sped up by the crystallization of numerous enzymes in the past years. The crystal structure of an enzyme usually gives a clue about the residues that are involved in catalysis, especially when the substrate or similar component is co-crystallized. Interactions between substrate and enzyme or between amino acids in the enzyme itself can be deduced from distances between the linked atoms. Unfortunately, crystal structures only show one possible state the enzyme can accept, which may imply that certain dynamic shapes of the enzyme will never be seen. This makes it difficult to fully comprehend the catalytic process from a single crystal structure and may lead to ignoring important residues. Nevertheless, the crystal structure of an enzyme or a close relative is a good start to search for residues that should be changed in order to modify the properties of an enzyme (Yano and Poulos, 2003; Johannes and Zhao, 2006).

Examples of xylanases whose crystal structures have been determined include *Streptomyces lividans* (Ducros *et al*., 2000), *Thermomonospora* sp. (George and Rao, 2001; George *et al*., 2001), *Chaetomium thermophilum*, *Nonomuraea flexuosa*

(Hakulinen *et al*., 2003), *Pseudoalteromonas haloplanktis* (Van Petegem *et al*., 2003) and *Thermoascus aurantiacus* (Vardakou *et al*., 2005).

Progress is also being made in the field of bioinformatics and there have been three publications in the last five years that focused on using sequence alignments to determine similarities in function and structure in the different xylanase families (Sapag *et al*., 2002; Liu *et al*., 2004a; 2004b).

#### *1.6.5.3 Protein engineering of xylanases*

Biotechnological applications of xylanases require a thermostable enzyme preparation with a wide pH and temperature range. Since the availability of the ideal enzyme preparation is limited, the application of protein engineering studies to xylanases has gained importance and has been previously reviewed (Stephens *et al*., 2004).

Protein engineering is also one of the principal means of examining the active site of an enzyme to identify the roles of specific residues on catalysis. Previously, the identification of the enzyme active site residues relied on the chemical modification of proteins. However, rapid developments in molecular biology have made it possible for the individual amino acids to be substituted by site-specific mutagenesis. Some recent examples of xylanases mutated using SDM are shown in Table 1.5.

Scientists have also used directed evolution technology and random mutagenesis to enhance the properties of xylanases (Arase *et al*., 1993; Chen *et al*., 2001; Palackal *et al*., 2004; Miyazaki *et al*., 2006; Xie *et al*., 2006; Stephens *et al*., 2007). Various studies have also focused on the shuffling of xylanases using different techniques with varying measures of success (Shibuya *et al*., 2000; Ahsan *et al*., 2001; Gibbs *et al*., 2001; Mesta *et al*., 2001; Nishimoto *et al*., 2002; Mangala *et al*., 2003; Bergquist *et al*., 2005).

**Table 1.5** Examples of xylanases mutated using SDM

| Xylanase source | Amino acid mutation(s) | Biochemical effect | Reference |
|---|---|---|---|
| *A. kawachi* | D37N | Shift in optima from pH 2 to pH 5 | Fushinobu *et al*., 1998 |
| *Streptomyces lividans* | W274A | $T_m$ increased by 4.1°C with 0.01 fold increase in activity | Roberge *et al*., 1999 |
| *Streptomyces* sp. S38 | T11Y | $T_m$ increased by 9°C | Georis *et al*., 2000 |
| *S. lividans* | I49A, I49C | 2.8 - 8 fold increase in thermal stability | Ebanks *et al*., 2000 |
| *Bacillus circulans* | N35D | Shifted pH optimum from 5.7 to 4.6; 20% increase in activity | Joshi *et al*., 2000 |
| *T. reesei* | S110C, N154C and Q162H or Q162Y | Increased half-life of xyn from 1 min to 14 min at 65°C | Turunen *et al*., 2001 |
| *T. reesei* | S186R, N67R, T26R, Q34R, N69R | pH optimum shifted to alkaline region but decreased thermal stability | Turunen *et al*., 2002 |
| *B. pumilis* A-30 | N71S | Shifted optimum 0.4 units to acidic range with 6.5 fold decrease in activity | Liu *et al*., 2002 |
| *C. thermocellum* | Transposed TSD* from N- to C-terminus | Temp optimum was decreased by 40°C with 2 fold loss in thermal stability | Shin *et al*., 2002 |
| *T. reesei* | T2C, T28C | Increase in half-life from 40 s to 20 min at 65°C | Fenel *et al*., 2004 |
| *A. pullulans* | D73N | Decreased activity by 81% and improved pH optimum from 2 to 4 | Tanaka *et al*., 2004 |
| *P. haloplanktis* | E78Q | Increased $T_m$ by 2.1 degrees with 250 fold drop in activity | Collins *et al*., 2005 |
| *Streptomyces* sp. S38 | E139K | Shifted pH optimum from 6 to 7.5 | Esteves *et al*., 2005 |
| *T. reesei* | T2C, T28C, K58R, N97R, F93W, H144K, I91D | Optimum was increased 4 - 5 °C at pH 9 | Fenel *et al*., 2006 |
| *A. niger* BCC14405 | T25R, T65R, S67R, S184R | Stability increased by 18 fold with no change in activity | Sriprang *et al*., 2006 |
| *Thermobacillus xylanilyticus* | S98C, N145C | 10 fold higher stability with 2 fold increase in activity | Paës and O'Donohue, 2006 |
| *B. stearothermophilius* No. 236 | S100C, N150C | Improvement in half-life from 7 to 25 min at 65°C and from 4 to 7 min at 70°C | Jeong *et al*., 2007 |

*TSD –thermostabilizing domain

### 1.6.6 *T. lanuginosus*

*T. lanuginosus* is a thermophilic Deuteromycete that thrives at temperatures up to 60°C, the upper temperature limit for growth of true fungi, and fails to grow below 20°C. The worldwide distribution of *T. lanuginosus* is a result of the common occurrence of self-heating masses of organic debris (Emerson, 1968) and the properties of the different strains have been reviewed (Singh *et al*., 2003).

*T. lanuginosus* has an exceptional ability to produce extremely high levels of thermostable cellulase-free hemicellulases, most notably of which is xylanase. These properties have elicited numerous groups to investigate the xylanase-producing ability of strains isolated from nature and obtained from culture collections. A notable variability in the production of xylanase by *T. lanuginosus* strains has been observed (Bennett *et al*., 1998; Puchart *et al*., 1999; Singh *et al*., 2000a; 2000b; 2000c). The xylanase from a high-producing member of the *T. lanuginosus* family, designated SSBP, has been produced using different types of fermentation and applied in the bleaching of paper pulp (Madlala *et al*., 2001; Bissoon *et al*., 2002; Christopher *et al*., 2005). The most well-researched *T. lanuginosus* strain researched to date however, has been the strain designated DSM 5826.

It was isolated by Dr I. Gomes from self-heated (65 – 70°C) jute stacks at the Jute Research Institute, Dhaka, Bangladesh, in 1988. The earliest studies on this strain focused on its ability to utilize cheap lignocellulosic wastes for xylanase production, most notably corn cobs (Gomes *et al*., 1989; 1992; 1993; Purkarthofer *et al*., 1993a; 1993b). The DSM 5826 xylanase is induced by xylose, xylan and xylobiose (Purkarthofer and Steiner, 1995). It has endo-acting character with a molecular weight of 26 kDa, pI of 4.1, optimal activities at pH 6.5 and 70°C and is stabilized by glycerol (Cesar and Mrša, 1996). Another study showed that its xylanase retained 80% of it activity at 70°C for 10 min when tested in the range of pH 4 - 12 (Lischnig *et al*., 1993). A separate study by Singh *et al*. (2000b), showed the xylanase from this strain to have a half-life ($t_{50}$) of 201 min at 70°C. The xylanase gene from this fungal strain was purified and successfully cloned into plasmid Bluescript (pBSK) and functionally expressed in *E. coli* as a LacZ-fusion protein. The *xyn*A-LacZ-fusion

protein was produced intracellularly with a tendency to form inclusion bodies (Schlacher *et al*., 1996).

The crystal structure analysis of this enzyme and a comparison of its structure with other family 11 xylanases were made by Gruber *et al*. (1998). It was found that the most prominent feature of the enzyme is a long cleft that spans the whole molecule and contains the active site. The structure is dominated by two heavily twisted β-sheets, designated sheets A and B (Fig. 1.9). Sheet A forms the outer surface of the enzyme and consists of five antiparallel strands. Its hydrophilic, solvent-accessible surface contains a large number of serine and threonine residues. Sheet B consists of nine mostly antiparallel strands. There is only one α-helix in the xylanase structure, which consists of ten residues, and is packed against the hydrophobic face of sheet B (Fig. 1.9B). The overall shape of the molecule resembles a right-hand with the two β-sheets and the α-helix forming the 'fingers' and the 'palm' and two loop regions forming the 'thumb' and a 'cord' (Fig. 1.9A). In the crystal structure, the thumb is well-ordered by interactions with sheet B and crystal contacts. Molecular dynamics simulations, however, indicate that it is one of the most flexible parts of the molecule.

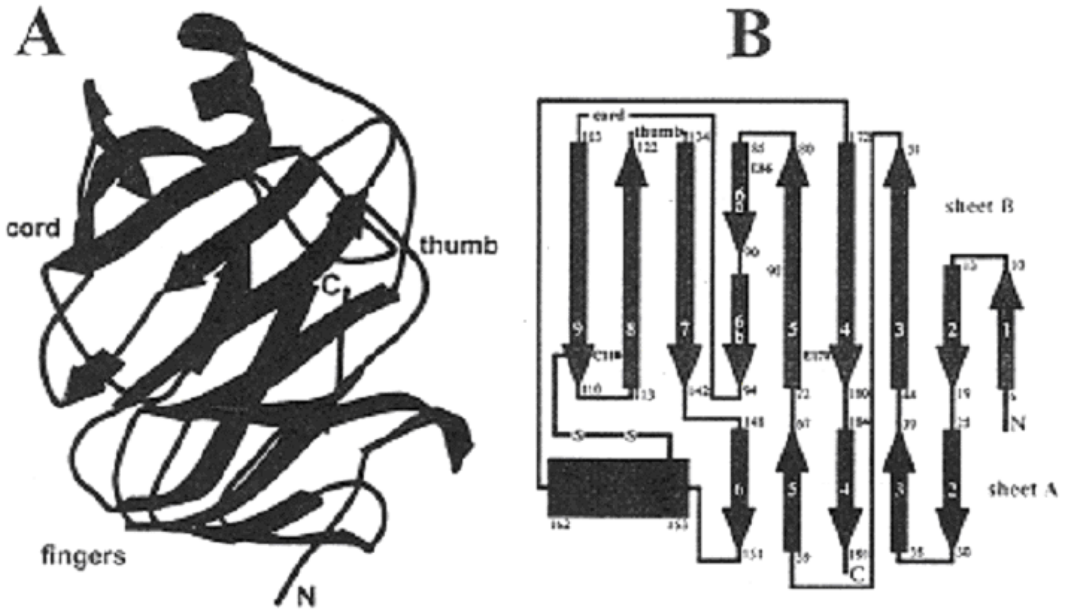


**Fig. 1.9** Molecular structure of *xynA* from *T. lanuginosus* DSM 5826. (A) A perpendicular view of the structure of *xynA* after 3-D protein modelling. (B) Topology diagram depicting the arrangement of the β sheets and α-helix within *xynA* (Gruber *et al*., 1998).

The thermostability of this xylanase was largely attributed to the presence of a disulphide bridge, an overall increase in the number of ion pairs and the presence of four salt bridges. The cloning and crystallization of this gene heralded the birth of a new approach to adapt it for the rigours of biotechnological processes: recombinant DNA technology.

## 1.7 RATIONALE FOR THE PRESENT STUDY

The implementation of xylanases, particularly in the pulp and paper industry, has many shortcomings that limit its application. The technology itself has tremendous potential, especially when the planet is facing the damning repercussions of centuries of pollution, some of which is contributed to by the pulping industry. Few natural xylanases are capable of withstanding the high temperatures and alkalinity encountered in the pulp and paper industry. In fact, pulp mills that utilize commercial xylanases today still need to cool and neutralize their pulp to ensure optimal functioning of these catalysts which presents significant downstream costs. With the dawn of an age where computational protein design and genetic engineering are moving to the forefront of technological innovation, the likelihood of conceiving a xylanase with the fitness to meet pulping process demands has become a reality.

Rational protein design is not an alien technology to the molecular biologist; it has been around for decades and has received a resurgence in popularity due to the breakthroughs and developments in bioinformatics and the sequence alignments of related proteins with known functions. However, what works brilliantly in theory may not match up to reality when the global structure of the protein in question is taken into account. Thus, random mutagenesis technologies such as directed evolution are more alluring to the protein engineer primarily because only functional proteins with the desired properties are selected for, provided that the screening technology is stringent enough. Finally, the possibility of using the age-old technique of recombination is an assured way of amalgamating many desirable properties into a single protein. By subjecting the xylanase gene to successive rounds of mutagenesis and selection of improved variants, it is envisaged that the gene might accrue constructive mutations that would pave the path for its eventual commercialization.

The premise for this study was that *T. lanuginosus* DSM 5826 is capable of withstanding alterations to its protein structure that would render it more tolerant to the thermophilic and alkaline biochemical stresses encountered during pulp bleaching.

The project was designed to initially reduce the size of the *xynA* cloning fragment and then proceed with random mutagenesis, which was accomplished using an assortment of ep-PCR techniques for creation of the functionally-fittest xylanase. The development of a suitable screening method was crucial to the success of the protein engineering strategy for the selection of functional alkaline and thermostable xylanase variants. The xylanases of mutants with promising stabilities were then further evaluated on a long-term basis. Two mutants individually displaying both these properties were selected and DNA shuffling was then used to combine the beneficial properties into a single xylanase. Recombinant xylanases displaying both thermal and alkaline stabilities were screened and selected for their tolerance to thermal and alkaline extremes. Parent and recombinant xylanases were subsequently sequenced to analyze beneficial mutations. The pH and temperature optima and stabilities as well as the kinetics of the successful parent and recombinant xylanases were then determined under operating pulp conditions. In addition, the original mutant library was screened for changes to other properties such as stability at acidic pH and altered specificity for different glucans and xylans. Mutants exhibiting notable changes in these properties were partially characterized and sequenced.

# CHAPTER 2: RANDOM MUTAGENESIS OF *xynA*

## 2.1 INTRODUCTION

Enzymes have been selected by natural evolution to sustain the development of living cells and organisms in their natural environment. These biological catalysts in general, are attractive for industrial purposes because they are efficient and selective in the chemistries they execute. However, natural enzymes do not necessarily have adequate properties, specificity, stability or catalytic activity to make them suitable for biotechnological applications. Several strategies have been envisaged to modify the properties of natural enzymes. SDM, guided by molecular modelling or hypotheses about natural evolution, has sometimes been successful in providing enzymes with novel expected properties (Eijsink *et al*., 1992; Cedrone *et al*., 2000; Gupta *et al*., 2002). Although protein chemists continue to elucidate the relationships between the sequence structure and function of proteins, the extensive knowledge that is necessary for the application of rational engineering approaches is available for only a small fraction of known enzymes. In view of this impediment, the alternative strategy of creating libraries of mutants and screening for interesting mutants, called directed evolution, is increasingly being followed in the absence of detailed structural data (Chirumamilla *et al*., 2001; Fernandez-Gacio *et al*., 2003).

Directed evolution has proved to be useful in enhancing enzyme performance in non-natural environments, as well as obtaining new features never required by nature, provided an efficient screening method can be found to channel the enzyme evolution towards the desired property (Arnold and Volkov, 1998). It mimics the process of Darwinian evolution in a test tube, combining random mutagenesis and, sometimes, recombination with screening or selection for enzyme variants that have the desired properties. A significant advantage of this approach over rational design methods is that neither structural information nor a mechanistic road map is required to guide the directed evolution experiment. It also allows the rapid and simultaneous exploration of sequence, structure and function space and thus provides a powerful tool to address biochemical questions involving protein stability and function (Kuchner and Arnold, 1997a). The success of this strategy depends on the size, quality and diversity of the

libraries and, crucially, on the sensitivity, efficiency and discriminatory power of the screening technique available (Fernandez-Gacio *et al*., 2003; Turner, 2003).

An important step in a directed evolution experiment is to efficiently explore sequence space through random mutagenesis. The sequence space of a truly randomized library is, by its nature, not limited by a pre-selection for functions under physiological conditions and allows scientists to adapt proteins to non-natural environments such as industrial processes. Among random mutagenesis methods, ep-PCR methods, based on the inaccurate amplification of genes, have been very successful and are generally employed in directed evolution experiments due to their simplicity and versatility (Wong *et al*., 2004). Ep-PCR methods can be divided into:

(i) methods that reduce the fidelity of the polymerase by unbalanced nucleotide concentrations and/or addition of inhibitors such as $Mn^{2+}$ (Cadwell and Joyce, 1992; 1994; Lin-Goerke *et al*., 1997);

(ii) methods that employ nucleotide analogues (Kuipers, 1996; Zaccolo *et al*., 1996);

(iii) methods that utilize 'mutagenic' polymerases (Cline and Hogrefe, 2000); and methods that employ a combination of the above strategies (Xu *et al*., 1999).

The primary intention of this study was to use the directed evolution strategies mentioned above to create modified enzymes useful for the pulp and paper industry, in an attempt to reduce the chlorinated organic by-products generated when pulp is bleached during paper manufacture. These by-products are harmful to the environment and 'greener' alternatives that result in their reduction or elimination are under investigation. One such technique is the use of xylanases as a prebleach treatment. Xylanase treatment is carried out after the initial alkali extraction of the pulp at high temperatures and thus thermostable and alkalophilic xylanases are highly desirable for such use (Raghukumar *et al*., 2004) and have been the subject of intensive research.

The initial part of the study involved the removal of the unnecessary adjuncts to *xynA*, which included the LacZ-fusion partner as well as redundant *T. lanuginosus* DNA from the original cloned fragment found in pX3. *xynA* was originally cloned from the

fungus by Schlacher *et al*. (1996), using the using λ Zap II cloning kit (Strategene), a strategy that cones larger chunks of DNA at a time. Newly designed primers were employed to trim off this excess DNA thus reducing the size of the cloned fragment. Once this was achieved, the main objective of this chapter was to create xylanase variants tolerant to alkaline and thermophilic conditions using different methods of ep-PCR. The screening step was two-fold: firstly, at the level of transformation to identify transformants producing a functional xylanase; and secondly, biochemical screens to test for stability under thermophilic and alkaline extremes.

## 2.2 MATERIALS AND METHODS

### 2.2.1 Construction of plasmid pX4

#### *2.2.1.1 Plasmids*

The original plasmid used in this study was designated pX3 (plasmid Bluescript containing *xynA*) as described by Schlacher *et al*. (1996) and the *E. coli* strain harbouring this plasmid was designated as X3. A linear plasmid map the region of pX3 containing *xynA* is depicted in Fig. 2.1. *xynA* is 678 bp in size and was cloned into a plasmid Bluescript SK$^-$ vector (pBSK) adjacent to the LacZ-fusion gene. The gene can therefore be amplified by PCR or sequenced using standard T3 (forward) and T7 (reverse) primers. The *xynA* gene-containing DNA fragment (1003 bp) is excised from the plasmid using the restriction enzymes *Xho*I and *Eco*RI, leaving behind a 2928 bp fragment that corresponds to the pBSK vector.

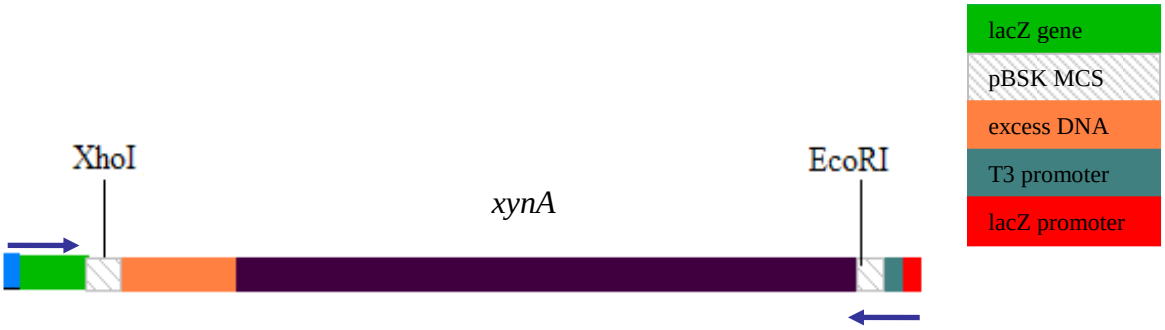


**Fig. 2.1** A linear map of pX3 region showing location of the *xynA* gene from *T. lanuginosus* DSM 5826 drawn using BVTech Plasmid software (version 2.0). Restriction sites and locations on the polylinker of the plasmid used for cloning are shown and arrows indicate direction of DNA synthesis. pBSK MCS: vector multiple cloning sites.

New primers were designed (section 2.2.1.6) to excise a 786 bp *xynA*-containing fragment from pX3 which was recloned into pBSK with *Xho*I and *Pst*I restriction enzymes. This new plasmid was designated as pX4 and the *E. coli* strain harbouring this plasmid was designated as X4. A linear map of the region containing *xynA* in pX4 is depicted in Fig. 2.2.

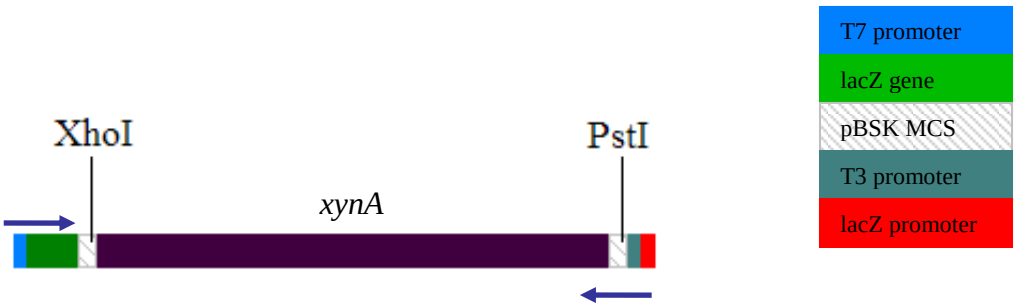


**Fig. 2.2:** A linear map of pX4 region showing location of the *xynA* gene from *T. lanuginosus* DSM 5826 drawn using BVTech Plasmid software (version 2.0). Restriction sites and locations on the polylinker of the plasmid used for cloning are shown and arrows indicate direction of DNA synthesis. pBSK MCS: vector multiple cloning sites.

#### *2.2.1.2 Plasmid DNA isolation*

The boiling lysis method of Sambrook *et al.* (1989) was used to isolate plasmid DNA and modifications to the original protocol are described. *E. coli* clones were inoculated in 5 ml LB medium containing 100 µg/ml ampicillin and grown for 12 - 16 h at 37°C. Cells were harvested by centrifugation of broth cultures at 5000 *g* for 5 min. The pellets were resuspended in 400 µl of STET-buffer [8% Sucrose, 5% Triton X-100, 50 mM EDTA (pH 8), 50 mM Tris-HCl (pH 8)] and transferred to 1.5 ml microcentrifuge tubes. Ten microlitres of lysozyme/RNase mix [10 mg/ml lysozyme and 1 mg/ml RNase in 50 mM Tris-HCl (pH 8)] was added and the mixture was incubated for 10 min at room temperature. The tube was transferred to a boiling water bath set at 95°C for 3 min. The tube and its contents were placed on ice for 5 min and then centrifuged in a microcentrifuge at 10 000 *g* for 10 min. The white pellet was removed with a sterile toothpick and discarded. Four hundred microlitres of ammonium acetate/isopropanol solution (1 vol 5 M ammonium acetate and 2 vols isopropanol) was added to the supernatant and centrifuged at 10 000 *g* for 10 min. The resulting pellet was washed with 700 µl 70% ethanol, without resuspension, and centrifuged at 10 000 *g* for 5 min. The pellets were air-dried for 10 - 15 min at room temperature. The DNA pellet was dissolved in 20 to 50 µl of sterile distilled water and stored at -20°C. For specialized applications like ep-PCR and DNA sequencing, where DNA of a higher purity was necessary, the FastPlasmid Mini kit (Eppendorf) was used according to the manufacturer's instructions.

#### *2.2.1.3 DNA quantification*

DNA concentration was estimated spectrophotometrically at 260 nm and calculated on the premise that an absorbance of 1 at $OD_{260\ nm}$ corresponds to 50 ng DNA/µl. Purity of the DNA sample was estimated using the absorbance ratio measured at both $OD_{260\ nm}$ and $OD_{280\ nm}$. A ratio greater than 1.7 is considered ideal purity for procedures like DNA sequencing.

#### *2.2.1.4 Preparation of λ DNA molecular weight marker*

An accurate estimation of the size of DNA fragments after agarose gel electrophoresis may determine the success or failure of many molecular biology procedures. Thus, phage λ DNA (Roche Molecular Biochemicals) was restricted with *Eco*RI and *Hin*dIII restriction enzymes to yield approximately 11 bands of known size. Forty seven microlitres of sterile distilled water and 10 µl restriction enzyme buffer B (Roche Molecular Biochemicals) were added to a sterile eppendorf, followed by the addition of 35 µl λ DNA (8.75 µg)and 4 µl each of *Eco*RI and *Hin*dIII.

The mixture was incubated at 37°C for 1 h and then placed at 65°C for 10 min to deactivate the enzymes. Ten microlitres of gel loading buffer (0.0375 g bromophenol blue, 4 g sucrose, 1.5 ml 10% SDS, 3 ml 0.5 M EDTA in a total volume of 15 ml) was added and then stored at 4°C. Aliquots were used as a molecular weight marker on all agarose gels.

#### *2.2.1.5 Agarose gel electrophoresis*

DNA molecules were separated on the basis of size using 0.8% agarose gels. The desired amount of agarose was placed in an Erlenmeyer flask together with the required amount of 1X TAE buffer, which was diluted from a 50X TAE stock (242 g Tris, 57.1 ml acetic acid, 100 ml of 0.5 M EDTA, pH 8). The contents of the flask were then microwaved for 1 min and poured into a casting tray with well combs and allowed to set. Gel loading buffer was added to the DNA samples in a ratio of 1:5 which was then loaded into the agarose gel wells. Samples were run alongside the λ DNA molecular weight marker (section 2.2.1.4) at 90 V for approximately 1 h. Gels

were then stained in ethidium bromide (0.05 mg/ml) for 20 min and destained in distilled water for a further 5 - 10 min. Stained gels were then viewed on a UV transilluminator and the band sizes compared to the DNA marker. The interaction of double-stranded DNA with ethidium bromide results in a strong, UV-excitable orange fluorescence, which shows the location of the DNA bands. Gel images were captured using a Scion CFW-1310M camera and Scion Image software (Scion Corporation).

#### *2.2.1.6 Primer design*

Two primers designated E20 and X21 were designed using the DNAMAN version 5.0 software package (Lynnon Biosoft) to remove the LacZ-fusion partner and excess *T. lanuginosus* DNA unnecessary for *xynA* synthesis from pX3 and thus reduce the size of the cloning fragment containing *xynA* from 1003 bp to 786 bp.

The first step of primer design involved the input of the *xynA* sequence as well as 60 bases on either side of gene into the program to determine the different restriction sites available for cloning without the restriction of *xynA* itself. The idea was to design primers with restriction sites for easier cloning. One pair was found to be most suitable: the 20-mer sense primer (E20) had an *Eco*RI restriction site and was located 43 bases from the *xynA* start codon whilst the 21-mer antisense primer (X21) had an *Xba*I restriction site almost 49 bases from the *xynA* stop codon. These primers, their *xynA* binding and restriction sites are shown in Fig. 2.3.

The primers were synthesized by Inqaba Biotechnical Industries. Lyophilized primers (70 nanomoles) were diluted with sterile double distilled water to a final stock concentration of 100 µM and stored at -20°C in 10 µl aliquots. Working stock concentrations of 10 µM were subsequently prepared as required.

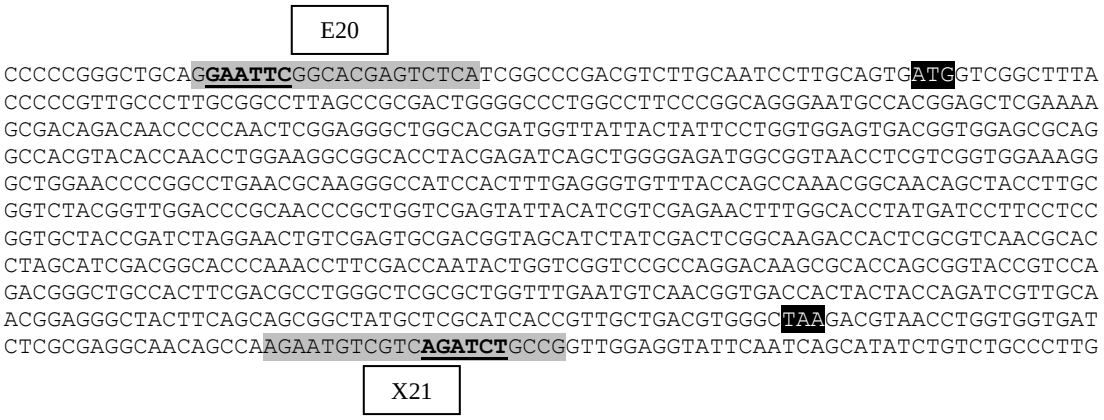


**Fig. 2.3** Gene sequence showing sense (E20) and antisense (X21) primer binding sites (grey highlights) on *xynA*. The actual recognition site within each primer is underlined and in boldface while the *xynA* start and stop codons are highlighted in black.

#### *2.2.1.7 PCR*

PCR of *xynA* in pX3 with newly designed primers E20 and X21 (section 2.2.1.6) was used to reduce the size of the cloning fragment in this study and is referred to hereafter as the test PCR. PCR of *xynA* in pX3 using standard primers T3 and T7 (Integrated DNA Technologies) served as the control.

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

Ten nanograms of template pX3 plasmid DNA, 0.5 µM of the primers, 0.1 mM dNTPs, 1.5 mM $MgCl_2$, 10X PCR buffer and 1U of *Taq* polymerase were used for control and test PCR reactions in a total volume of 50 µl.

Both sets of primers had different melting temperatures thus the control and test PCR reactions were carried out using PCR programs with different annealing temperatures; 42°C for control PCR with primers T3 and T7; and 52°C for test PCR with primers E20 and X21. The temperature program for amplification was as follows:

| | | |
|---|---|---|
| Denaturation | : | 1 min at 94°C |
| Primer annealing | : | 1 min at 42°C (52°C) |
| Primer extension | : | 2 min at 72°C |

Each PCR reaction comprised a total of 30 cycles and the reactions were performed using the PCR Genius thermal cycler (Techne).

PCR products were separated by agarose gel electrophoresis (section 2.2.1.5) to determine if the target DNA was successfully amplified. Upon successful amplification, the PCR bands were excised from the gel and recovered using the GFX PCR DNA and Gel Band Purification kit (Amersham) according to the manufacturer's instructions. A small aliquot of the purified PCR products was run on an agarose gel to ascertain successful elution from the purification column.

#### *2.2.1.8 Restriction analysis*

Restriction endonucleases were used for specific digestion of plasmid DNA for three different applications in this study. Firstly, restriction digestion of both the vector and the insert DNA with suitable restriction enzymes was used to create compatible sticky ends for ligation with DNA ligase. Secondly, recombinant plasmids were restricted with the same enzymes to determine if they contained the insert after transformation into *E. coli*. Such a restriction would yield two characteristic bands of different sizes: a large vector band and a significantly smaller insert band. Finally, restriction endonucleases were also used for preparation of the λ DNA molecular weight marker (section 2.2.1.4).

Standard protocols were followed for restriction analysis (Sambrook *et al*., 1989; 2001). For digestion, 0.1 volumes of the corresponding restriction buffers (10X) were added to the DNA solution. Mixtures were incubated with the restriction enzymes (Roche Molecular Biochemicals) at 37°C for 2 h or more. Restricted DNA was analyzed on 0.8% agarose gels.

The purified test PCR product (section 2.2.1.7) was restricted using *Eco*RI and *Eco*RV whilst the control PCR product was restricted using *Eco*RI and *Xho*I to create sticky ends compatible with the similarly digested pBSK vector to facilitate ligation between both the vector and insert. pBSK vector was digested separately for both conditions. Restriction reactions were purified from solution using the GFX PCR DNA and Gel Band Purification kit (Amersham) according to the manufacturer's

instructions. A small aliquot of the purified restriction products were run on an agarose gel to ascertain successful elution.

#### *2.2.1.9 Klenow treatment*

The Klenow enzyme (Fermentas) was used to fill in the recessed 3´-termini of test insert and vector restricted with *Eco*RI and *Eco*RV. Three micrograms of restricted test insert and vector was added to 10X Klenow reaction buffer, 2 mM dNTP mix, 1U Klenow fragment and topped up with sterile deionized water in a total volume of 20 µl. The reaction mixture was incubated at 37°C for 10 min and directly purified using the GFX PCR DNA and Gel Band Purification kit (Amersham) according to the manufacturer's instructions. A small aliquot of the purified filled-in products were run on an agarose gel to ascertain successful elution and then quantified (section 2.2.1.3).

#### *2.2.1.10 Ligation*

For the creation of recombinant test and control libraries, a molar vector:insert ratio of 1:3 was used to guarantee high ligation efficiency. The GFX-purified restricted PCR inserts and GFX-purified restricted pBSK vectors for both control and test PCR reactions were ligated using Rapid DNA Ligation kit (Roche Molecular Biochemicals) according to the manufacturer's instructions. The ligation mixtures were then transformed into SEM-competent *E. coli* host cells.

#### *2.2.1.11 Preparation of SEM-competent cells*

Host cells were made 'competent' or capable of taking up DNA from their surrounding environment, by exposing them to $Ca^{2+}$, which interacts with their cell envelopes. *E. coli* XL1 blue MRF' (Strategene) was used for all cloning procedures and were made competent using the Simple and Efficient Method prior to transformation (Ausubel *et al*., 1989).

*E. coli* XL1 Blue was cultured on LB-medium. A single colony was used to inoculate 5 ml sterile SOC medium (20 g/l tryptone, 5 g/l yeast extract, 40 mM glucose, 20 mM NaCl, 20 mM $MgCl_2$, 20 mM $MgSO_4$, 5 mM KCl), incubated at 37°C and shaken

overnight. One millilitre of this culture was used to inoculate 29 ml of fresh SOC medium and shaken at 37°C until it reached an $OD_{600\ nm}$ of 0.5. The culture was immediately placed on ice and kept cold for the duration of the procedure. The cells were pelleted at 5000 *g* for 10 min and the supernatant was discarded. The cells were resuspended in 10 ml cold 100 mM $CaCl_2$ and recentrifuged at the same speed and resuspended in 10 ml cold 100 mM $CaCl_2$. The entire mixture was incubated on ice for 20 min and then centrifuged. The competent cells were subsequently resuspended in 2 ml 100 mM $CaCl_2$ containing 10% glycerol. One hundred and fifty microlitres of the prepared competent cells were dispensed into eppendorfs, stored at 4°C overnight, and then transferred to -70°C, after freezing in liquid nitrogen. According to the protocol followed (Ausubel *et al*., 1989), it is postulated that SEM-competent cells are most efficient when prepared 24 h prior to transformation. The preparations were therefore incubated at 4°C overnight to enhance the effectiveness of the subsequent transformation procedure.

#### *2.2.1.12 Transformation and screening*

Two microlitres of the ligated DNA solutions (section 2.2.1.10) were added to one hundred and fifty microlitres of the SEM-competent cells and incubated on ice for 30 min and thereafter subjected to heat shock for 30 s at 42°C. Rapid addition of 800 µl of fresh SOC medium to each of the mixtures followed the heat shock procedure and they were then shaken at 37°C for 1 h. The transformation mixtures were plated on RBB-xylan-LB plates (0.4% RBB-xylan, 10 g/l bactopeptone, 5 g/l yeast extract, 5 g/l sodium chloride, 15 g/l technical agar) containing 100 µg/ml ampicillin and incubated at 37°C overnight. The dye, Remazol Brilliant Blue, was linked to birchwood xylan (Roth) for detection of xylanase producers during the transformation process. RBB-xylan was prepared according to the method of Biely *et al*. (1985; 1988).

This served as both a screening and selection method since only those cells that had taken up a fully functional plasmid were able to grow on the selective medium supplemented with ampicillin and it further screened for clones expressing xylanase from those that merely contained self-ligated vectors. Five isolates were initially subjected to PCR to determine presence of *xynA* and then restriction analysis to determine the set of restriction enzymes that would release the 786 bp fragment.

### 2.2.2 Random mutagenesis

PCR-based random mutagenesis generally involves performing a PCR reaction under conditions that reduce the fidelity of nucleotide incorporation. Such conditions include introduction of $Mn^{2+}$, sub- or super-optimal amounts of nucleotides and increasing amounts of $Mg^{2+}$ which impedes *Taq* DNA polymerase activity (You and Arnold, 1994; Wan *et al*., 1998; Melnikov and Youngman, 1999). A broad scope of conditions (Table 2.1) ranging from previously published research as well as the use of a commercial mutagenesis kit, Diversify Random Mutagenesis kit (CLONTECH), were investigated for mutagenesis of *xynA*.

**Table 2.1** Mutagenic PCR conditions used for *xynA*

| Condition | Concentration (mM) | | | PCR | | Reference |
|---|---|---|---|---|---|---|
| | $Mg^{2+}$ | $Mn^{2+}$ | dNTPs | Program | No. of cycles | |
| A (control) | 1.5 | - | 0.1 | 94°C – 1 min | 35 | - |
| B | 4.8 | 0.5 | 0.2 AG | 42°C – 1 min | | Matsumura and |
| | | | 0.8 CT | 72°C – 2 min | | Ellington, 2001 |
| C | 1.5 | - | 0.04 AG | | | Chen *et al*., 2001 |
| | | | 0.2 CT | | | |
| DI[†] | 2 | 0.04 | 0.02 AG | 94°C – 1 min | 20 | Xu *et al*., 1999 |
| | | | 0.2 CT | 46°C – 1 min | | |
| DII | | - | 0.02 AG | 72°C – 2 min | 30 | |
| | | | 0.2 CT | | | |
| | | | 0.04 dITP | | | |
| E[❖] | 3.5 | - | 0.04 G | 94°C – 1 min | 25 | Diversify kit manual |
| F[❖] | | 0.64 | 0.04 G | 46°C – 30 s | | |
| G[❖] | | 0.64 | 0.2 G | 68°C – 1 min | | |
| H[❖] (control) | | - | 0.2 AGCT | | | |

[†] 2 µl unpurified PCR product of DI was used as template for DII

[❖] Diversify Random Mutagenesis kit conditions. Special Diversify dNTP mix with unspecified concentrations was used for conditions E, F, G

The plasmid pX4, constructed in section 2.2.1, served as the template DNA for all mutagenic conditions tested. Standard T3 and T7 primers (0.5 µM) were used for all reactions and the total volume was always 50 µl. Five microlitres of each PCR product was analyzed by agarose gel electrophoresis to determine if the target DNA was successfully amplified. The remaining PCR products were recovered using the

GFX PCR DNA and Gel Band Purification kit (Amersham) and restricted with *Xho*I and *Pst*I at 37°C for 2 h to create sticky ends compatible with the pBSK vector to facilitate ligation between vector and insert. The restricted products were purified from the solution using the same kit, according to the manufacturer's instructions.

### 2.2.3 Cloning of PCR products

The purified and restricted PCR products and pBSK vector were quantified and ligated as previously described in sections 2.2.1.3 and 2.2.1.10. Ligated products were subsequently transformed into SEM-competent cells (sections 2.2.1.11 and 2.2.1.12).

### 2.2.4 Growth of mutants and enzyme extraction

All positive transformants, i.e., clones exhibiting xylanase activity, were picked with a sterile toothpick and initially streaked onto an LB plate containing 100 µg/ml ampicillin. Subsequently, a single colony from the overnight plate was inoculated into 5 ml LB broth containing 100 µg/ml ampicillin and incubated for 12 - 16 h at 37°C in a shaking incubator at 170 rpm. One hundred and fifty microlitres of these overnight cultures were placed in sterile 1.5 ml eppendorf tubes containing 30% sterile glycerol. These eppendorfs were incubated at -20°C and served as the master mutant library stock. The rest of the culture was centrifuged at 5000 *g* and the media discarded. The pellets were resuspended in 500 µl Bugbuster Protein Extraction Reagent (Novagen). The suspensions were shaken gently at 40 rpm in a shaking incubator at room temperature for 20 min to gently lyse *E. coli* cells and release the mutant xylanases. The lysates were then centrifuged at 15000 *g* at 4°C. The resulting pellets were discarded whilst the supernatants were stored at 4°C for further analysis.

### 2.2.5 Screening

#### *2.2.5.1 Thermostability screening assay*

The clear lysate obtained after enzyme extraction contained the crude enzyme and was used to test the thermostability of the xylanase variants. The protocol followed was a combination of the methods used by Giver *et al*. (1998) and Matsuura *et al*.

(1999). A temperature of 80°C was chosen for screening possible thermostable xylanase variants because XynA is documented to be stable up to 70°C (Schlacher *et al*., 1996). Prior to incubation at 80°C in a water bath, 0 min (untreated) samples were removed from the clear cell lysates and placed on ice. The crude enzymes were subsequently heated for 40 min, chilled on ice for 15 min and incubated for 30 min at room temperature to prevent low temperature denaturation of the enzymes. The samples were centrifuged and the supernatants assayed for residual xylanase activity (Bailey *et al*., 1992). Activities of the 0 min samples were considered as 100%, and activities of the 40 min incubation time were expressed as percentages of the untreated sample to determine the % residual activity after heat treatment. The wild type XynA served as the control.

#### *2.2.5.2 Alkaline screening assay*

Wells were punched (using an Ouchterlony well-maker) into plates containing 0.1% birchwood xylan (Roth) and 1% agarose made with 0.05 M glycine-NaOH buffer (pH 10). Ten microlitres of each crude enzyme was inoculated into the wells, with one well containing the control XynA on each screening plate. The plates were incubated at 60°C for 2 h and then stained with 1% Congo Red for 25 min. Excess dye was flushed off with 1 M NaCl for up to 1 h, until zones of hydrolysis were clearly visible (Teather and Wood, 1982; Béguin, 1990). Mutants displaying larger or more distinct zones than XynA were then grown and lysed as described in section 2.2.4. Supernatants were diluted in 0.05 M glycine-NaOH buffer (pH 10) and incubated in a 60°C water bath and residual activity of the enzymes were determined as previously described (section 2.2.5.1). A temperature of 60°C was used for detection of alkaline-stable mutants to effectively trim down the number of mutants screened during longer incubation periods.

#### *2.2.5.3 Long-term thermal and alkaline stabilities*

Mutants displaying more than 60% residual activity after heat and alkaline treatment in sections 2.2.5.1 and 2.2.5.2 were inoculated into 5 ml LB broth containing 100 µg/ml ampicillin and incubated overnight. One millilitre of this culture was used to inoculate 300 ml LB medium containing 100 µg/ml ampicillin and shaken at 37°C until the $OD_{600}$ of all flasks were 0.5 absorbance units. The cells were induced for xylanase production

by adding 1 mM IPTG to the flasks and incubated at 37°C for a further 24 h. Thereafter, the samples were centrifuged 5000 *g* and the media discarded. Each 50 ml pellet was resuspended in 2 ml of cold lysis solution.

Bugbuster Protein Extraction Reagent (Novagen) (used for large scale screening in section 2.2.4) is unstable for long-term incubation at high temperatures. Thus, the following lysis solution was designed to eliminate this problem. The lysis solution comprised 150 U DNase I (Roche Molecular Biochemicals), 2 g lysozyme (Roche Molecular Biochemicals) in 120 ml breaking buffer solution [6.80 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]. After autoclaving and cooling of the breaking buffer solution, 0.10 g/l phenylmethylsulfonylfluoride was added. The suspensions were left at 4°C overnight and then centrifuged at 15000 *g* and the supernatant lysates containing the enzymes were stored at 4°C until further use.

For determination of long-term thermal stability, samples were incubated at 80°C and samples were removed every 15 min for 90 min, incubated on ice then assayed for residual xylanase activity (section 2.2.5.1). Activities of each mutant were expressed as percentages of the 0 min sample.

For determination of long-term alkaline stability, samples were diluted in 0.05 M glycine-NaOH buffer (pH 10), incubated at 60°C and samples were removed every 15 min for 90 min, incubated on ice then assayed for residual activity (section 2.2.5.1). Activities of each mutant were expressed as percentages of the 0 min sample. All mutants were analyzed in duplicate.

## 2.3 RESULTS

### 2.3.1 Construction of plasmid pX4

#### *2.3.1.1 PCR*

After the original plasmid pX3 was isolated as described in section 2.2.1.1, it was subjected to PCR under standard conditions using T3 and T7 primers and using the newly designed primers E20 and X21, separately. A small aliquot of the PCR products were run on an agarose gel to determine the success and differences (if any) of each reaction and is shown in Fig. 2.4. Amplification of the standard 1003 bp product T3 and T7 primers is shown in lane 2, while PCR with E20 and X21 primers indicates a 786 bp product (lane 3) instead. The λ DNA molecular weight marker prepared in section 2.2.1.4 can be seen in lane 1.

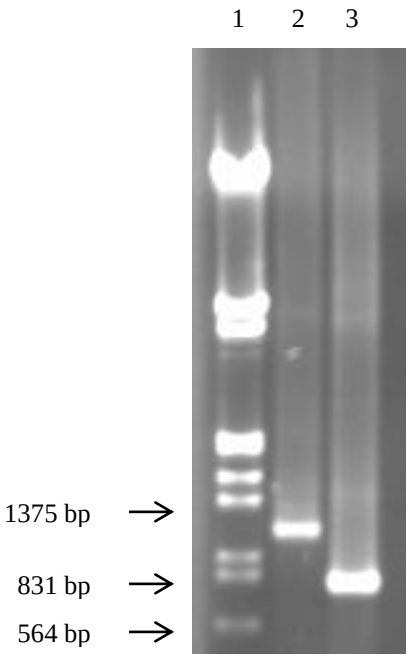


**Fig. 2.4** Agarose gel electrophoresis of pX3 after PCR with T3 and T7 primers (lane 2) and primers E20 and X21 (lane 3) against λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII (lane 1).

### *2.3.1.2 Screening transformants for xylanase production*

The standard and test PCR products obtained were cloned into *E. coli* and after plating on LB-RBB-xylan plates containing ampicillin, three positive and 112 negative colonies were obtained. Colonies producing xylanase hydrolyzed RBB-xylan into colourless degradation products which are observed as a halo around the bacterial colony (Fig. 2.5). Two positive colonies had almost the same zone size were very distinct (referred to as clones 1 and 4) whilst that of the third was very faint (clone 5).

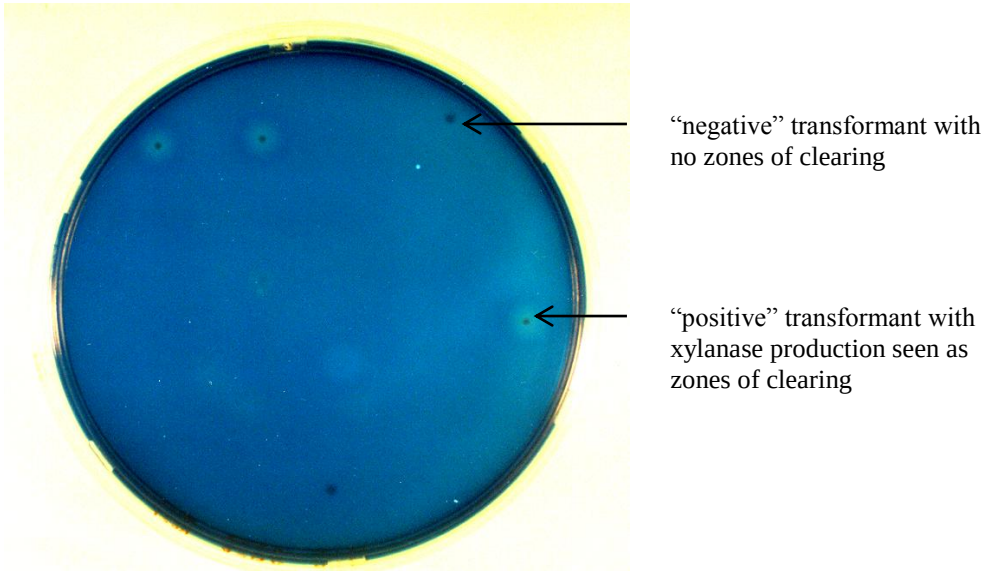


**Fig. 2.5** Recombinant *E. coli* colonies after transformation and incubation at 37°C for 16 h. Transformation mixtures were plated on LB medium supplemented with 0.4% RBB-xylan and ampicillin (100 μg/ml). The presence of a clear halo around the bacterial colony is indicative of xylanase production.

#### *2.3.1.3 PCR and restriction analysis of a few isolates*

The three clones exhibiting xylanase production (1, 4 and 5) along with two non-xylanase producers (clones 2 and 3) were selected for further analysis. They were grown in 5 ml LB supplemented with ampicillin (100 μg/ml) and the plasmids were isolated and subjected to PCR (section 2.2.1.7) to determine if both positive and negative xylanase producers were in possession of the *xynA* gene. Agarose gel electrophoresis of the PCR products revealed all clones to have the 786 bp fragment (Fig. 2.6). However, the brightness of the products differs for both positive and negative xylanase producers. Clone 1 had the brightest band (lane 2), followed by clones 4 and 5 (lanes 5 and 6), which are all positive for xylanase production. Clones 2 and 3, which did not produce xylanase, had fainter PCR products (lanes 3 and 4). Since clone 1 produced the brightest band, it was selected for all test restriction analysis.

Clone 1 was grown in 5 ml LB supplemented with ampicillin (100 μg/ml) and its plasmid was isolated and subjected to restriction analysis as described in section 2.2.1.8 using different combinations of restriction enzymes and is shown in Fig. 2.7. Linearization of pBSK occurs after digestion with *Eco*RI only (lane 2) and a small amount of PCR product of 786 bp is shown in lane 3 for comparison to show which restriction enzyme combination releases the fragment containing *xynA*. Lane 4 indicates the presence of covalently closed circular plasmid DNA uncut after treatment with *Eco*RI. This enzyme does not restrict the DNA since single bands are obtained in lanes 6, 7 and 8 when it is paired with *Xba*I, *Pst*I or *Xho*I. Restriction with *Xba*I linearizes the plasmid (Lane 5). Restriction of the plasmid with *Xba*I and *Pst*I (lane 11) also yields a single band since these two restriction enzymes have recognition sites only a few bases away from each other in the pBSK MCS and the second fragment produced is too small to be detected on the gel. However, when the plasmid is digested with combinations of *Pst*I and *Xho*I or *Xba*I and *Xho*I in lanes 9 and 10 respectively, the desired 786 bp *xynA* fragment is released leaving behind the 2928 bp pBSK vector band. Thus, either combination is suitable for releasing *xynA* from the vector. The combination in lane 9 was chosen for further cloning experiments.

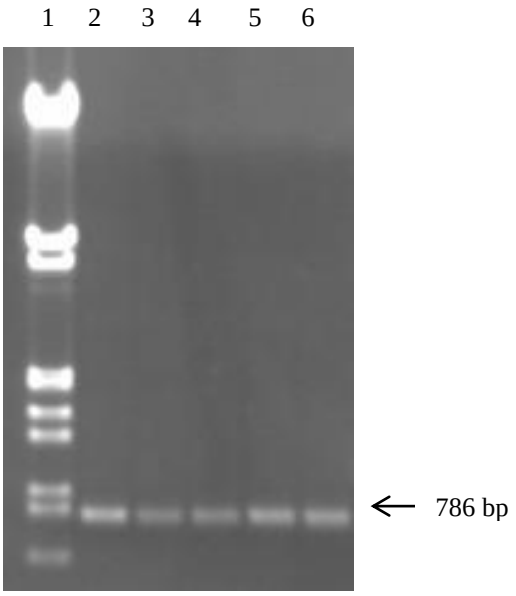


**Fig. 2.6** Agarose gel electrophoresis of 5 isolates after PCR with E20 and X21 primers. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: xylanase-producing clone 1; Lane 3: non-xylanase producing clone 2; Lane 4: non-xylanase producing clone 3; Lane 5: xylanase-producing clone 4; Lane 6: xylanase-producing clone 5.

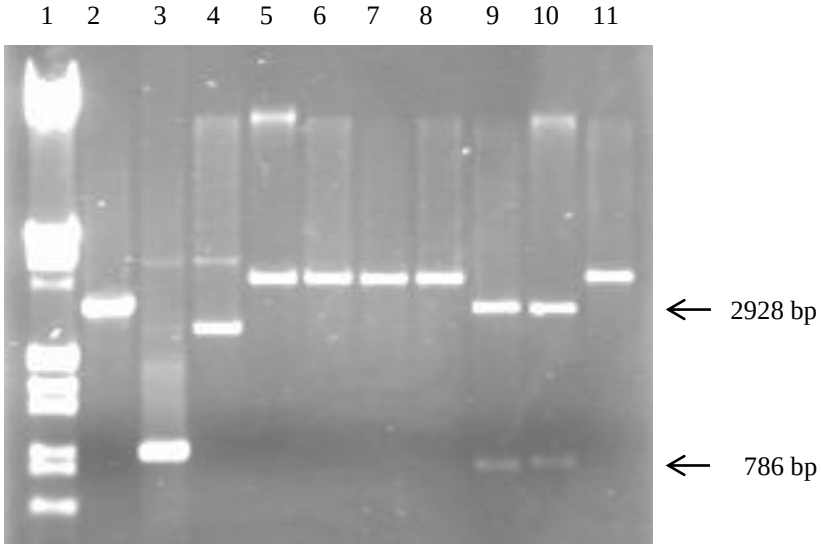


**Fig. 2.7** Agarose gel electrophoresis of plasmids isolated from positive transformants subjected to restriction digestion. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: pBSK with *Eco*RI; Lane 3: 786 bp product obtained after PCR with E20 and X21; Lane 4: *Eco*RI; Lane 5: *Xba*I; Lane 6: *Eco*RI and *Xba*I; Lane 7: *Eco*RI and *Pst*I; Lane 8: *Eco*RI and *Xho*I; Lane 9: *Pst*I and *Xho*I; Lane 10: *Xho*I and *Xba*I; Lane 11: *Pst*I and *Xba*I.

### 2.3.2 Random mutagenesis

Five microlitres of each PCR product corresponding to each of the mutagenesis conditions tested (Table 2.1) was separated by agarose gel electrophoresis (Fig. 2.8). A single band of 786 bp was obtained for all conditions except in lane 8, which was a kit *Taq* control and produced multiple bands as expected. Lanes 7 to 12 reflect mutagenic conditions generated with the use of the Diversify kit. Control reactions for the different PCR conditions are shown in lanes 2, 7, 8, and 12 and these bands are quite dense with the exception of lane 8, which is a mutagenesis reaction control. In contrast, the band in lane 9 is marginally lighter followed by the two lightest bands in lanes 10 and 11, which noticeably represents the most mutagenic kit reactions. Lane 3 which contains bands produced during reaction B is much brighter than the control in lane 2. Lane 5 contains the reaction for DI which contains high amounts of the universal base dITP and it serves as the precursor of DII (lane 6), which is much brighter.

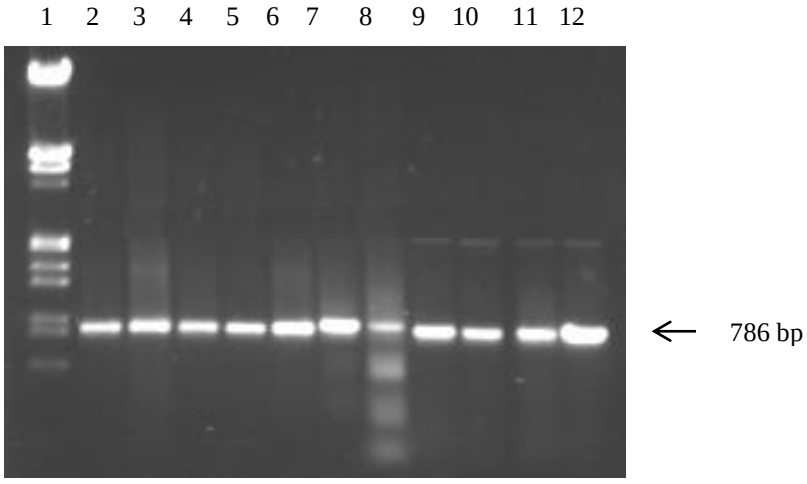


**Fig. 2.8** Agarose gel electrophoresis of random mutagenesis of *xynA* after exposure to different mutagenic PCR conditions (A to H – refer to Table 2.1). Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: A (control amplified under normal PCR conditions); Lane 3: B (4.8 mM $Mg^{2+}$, 0.5 $Mn^{2+}$, 0.2 mM AG, 0.8 mM CT); Lane 4: C (1.5 mM $Mg^{2+}$, 0.04 mM AG, 0.2 mM CT); Lane 5: DI (2 mM $Mg^{2+}$, 0.04 mM $Mn^{2+}$, 0.02 mM AG, 0.2 mM CT); Lane 6: DII (2 mM $Mg^{2+}$, 0.02 mM AG, 0.2 mM CT, 0.04 mM dITP); Lane 7: Kit *Taq* restriction I (control); Lane 8: Kit *Taq* restriction IX (control); Lane 9: E (3.5 mM $Mg^{2+}$, 0.04 mM G, Diversify mix); Lane 10: F (3.5 mM $Mg^{2+}$, 0.64 mM $Mn^{2+}$, 0.04 mM G, Diversify mix); Lane 11: G (3.5 mM $Mg^{2+}$, 0.64 mM $Mn^{2+}$, 0.2 mM G, Diversify mix); Lane 12: H kit control (3.5 mM $Mg^{2+}$, 0.2 mM AGCT).

### 2.3.3 Screening for xylanase mutants with improved stability

Nine hundred and sixty clones were collectively obtained after the cloning and transformation of the inserts obtained from the different mutagenic conditions by screening transformants for xylanase production on RBB-xylan as described in section 2.3.1.2. The clones were named according to the condition from which they were generated and were given a number for easy identification. The thermal and alkaline stabilities of XynA and its mutant progeny were determined separately and are shown as such.

#### *2.3.3.1 Short-term thermal stability*

The clones that underwent amplification under all control or standard conditions tested displayed similar thermal and alkaline stability profiles to XynA and are not reflected in these results. Of the 960 clones that were screened for improved thermal stability, only a small percentage displayed significant changes in this regard (Table 2.2). Condition E generated mutants with significant improvements in their catalytic activities although with a large majority remaining unchanged in terms of their stability. Conditions C and F were found to be the least effective mutagenic conditions since they exhibited a 75% and 64% loss in activity with a concomitant 63% and 50% reduction in thermal stability respectively. Condition D, however, generated the most thermostable mutants since 11% of the progeny retained more than 60% of their activity after heat treatment for 40 min. These mutants however, lost up to 90% of their catalytic activities.

**Table 2.2** Activity and stability of *xynA* mutant library at 80°C for 40 min

| Condition | No. of transformants | Activity (%) | | | Thermal stability (%) | | | |
|---|---|---|---|---|---|---|---|---|
| | | Enhanced | Reduced | Same | Enhanced | Marked increase♣ | Reduced | Same |
| B | 200 | 35 | 26 | 39 | 38 | 1.5 | 8 | 54 |
| C | 118 | 5 | 75 | 20 | 20 | 0 | 63 | 17 |
| D | 91 | 10 | 57 | 33 | 34 | 11 | 33 | 33 |
| E | 169 | 90 | 4 | 6 | 6 | 0 | 10 | 84 |
| F | 266 | 3 | 64 | 33 | 8 | 1.5 | 50 | 42 |
| G | 115 | 1 | 70 | 29 | 6 | 1.7 | 23 | 71 |

♣mutants with more than 60% remaining activity after incubation at 80°C for 40 min

#### *2.3.3.2 Screening for alkali-tolerant xylanase variants*

Alkaline screening was performed in 2 parts, *viz*., a plate assay and a liquid assay that tested stability at pH 10 for 40 min. The entire xylanase library was screened on pH 10 xylan plates and stained with Congo Red to determine zones of hydrolysis (Fig. 2.9). This screen effectively eliminated xylanases with inferior or equivalent stability to XynA. However, a highly active xylanase could produce a large zone of hydrolysis which could easily be misconstrued for superior stability. Thus, a liquid assay at pH 10 for 40 min would effectively curtail this problem.

XynA produces a large clear zone on the screening plate (Fig. 2.9). G53, on the other hand, produced a larger and clearer zone whereas G41 produced a weak, barely discernible zone. Only 62 mutant clones with significantly larger zones of hydrolysis than XynA were selected for further liquid biochemical testing (Table 2.3). It was observed that most of the 62 clones had similar activity to XynA, with only 24 having improvements in activity and 23 clones had more than 60% activity at pH 10 after 40 min at 60°C.

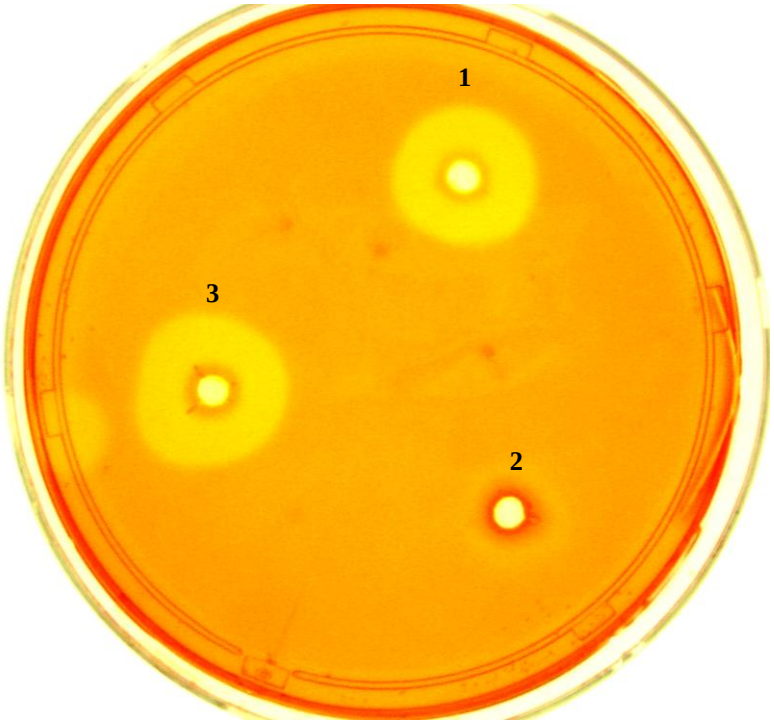


**Fig. 2.9** Screening plate showing zones of hydrolysis produced by crude xylanases created by mutagenesis of *xynA*. The wells contain crude xylanases from (1) XynA, (2) G41 and (3) G53. The plates were prepared according to the modified methods of Teather and Wood, 1982 and Béguin, 1990 and are described in section 2.2.5.2.

**Table 2.3** Activity and stability of selected *xynA* mutants at pH 10, 60°C for 40 min

| Condition | No. of clones screened | Enhanced activity | Alkaline stability | |
|---|---|---|---|---|
| | | | Enhanced | Marked increase♣ |
| B | 21 | 4 | 21 | 9 |
| C | 14 | 5 | 11 | 5 |
| D | 6 | 3 | 6 | 5 |
| E | 11 | 11 | 3 | 0 |
| F | 8 | 1 | 3 | 2 |
| G | 2 | 0 | 2 | 2 |

♣mutants with more than 60% remaining activity after incubation at pH 10, 60°C for 40 min

#### *2.3.3.3 Long-term thermal and alkaline stabilities*

Six of the ten mutants exposed to 80°C for 90 min displayed more than 60% remaining activity after 90 min and 3 of them are shown in Fig. 2.10. XynA lost almost 80% of its activity under these conditions. Mutant G53 displayed an even lower stability and retained only 10% of its activity. The most stable mutant generated by initial random mutagenesis of *xynA* was G41, which displayed 75% retention of its total activity after 90 min of heat treatment. Mutant D88 had similar thermal stability to G41 but its activity was lower whilst mutant B61 retained almost 65% of its activity. All thermostable mutants displayed poor activity.

The mutants fared better at pH 10 (Fig. 2.11). Six of the seven mutants tested retained more than 60% of their activity after 90 min at pH 10. XynA lost 70% of its activity while B144 and D34 retained 76% and 79% of their activity respectively. Mutant G41 lost nearly 60% of its activity under these extreme alkaline conditions. The most remarkable alkaline stable mutant in this study was, however, G53 which retained 93% of its activity after similar treatment. Most alkaline stable mutants did not exhibit a significant drop in catalytic activity.

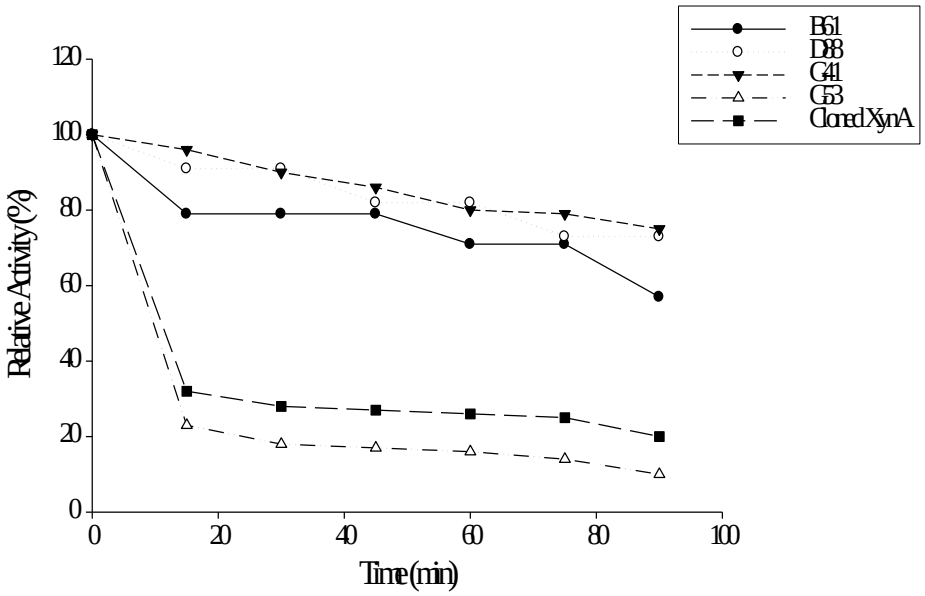


**Fig. 2.10** The effect of temperature on the stability of wild-type XynA and its mutant progeny at 80°C and pH 6.5. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

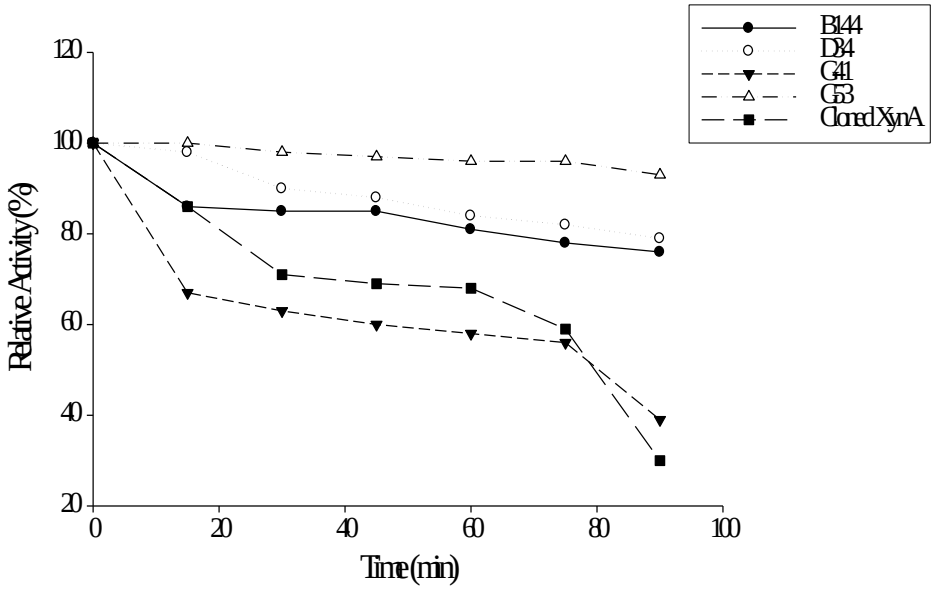


**Fig. 2.11** The effect of alkalinity on the stability of wild-type XynA and its mutant progeny at pH 10 and 60°C. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

## 2.4 DISCUSSION

This chapter outlines the numerous steps required for successful laboratory evolution of *xynA* for the creation of thermostable and alkaline stable enzyme variants. In 1996, this gene was originally cloned using the λ-ZapII cloning system (Strategene) by Schlacher *et al.* This system requires initial random fragmentation of the genome and subsequent cloning of the DNA fragments into the phage vector. Since it can take up large amounts of DNA, fewer clones are needed to create a cloned genomic library. This earlier research did not specifically target *xynA*, but rather a larger segment of the genomic DNA, which may not be involved in xylanase synthesis at all. An initial study (Stephens *et al*., 2007) revealed a reduced mutation rate of *xynA* with some mutations occurring within these unnecessary adjuncts of DNA instead. Shortening the cloned insert would therefore not only increase the mutational efficiency of an ep-PCR reaction, but it would also make sequencing easier since the fungal DNA fragment is now shorter.

Thus, it was decided to remove the LacZ-fusion partner upstream of *xynA* and most of the excess DNA upstream and downstream of *xynA*, thus reducing the size of the cloning and sequencing fragment, and then proceed with mutagenesis. This involved the design of primers E20 and X21 (section 2.2.1.6), which were used in a PCR reaction of the original plasmid pX3 (Fig. 2.4). The resultant shorter PCR product was then blunt-ended and cloned into pBSK. Clones were screened on LB-RBB-xylan plates (Fig. 2.5) and three positive xylanase-producing clones (1, 4 and 5) were found. A large number of clones did not produce a xylanase and two random colonies were selected for further analysis (clones 2 and 3). PCR of the plasmids of these positive and negative isolates revealed that they all possessed the 786 bp *xynA*-containing fragment (Fig. 2.6). However, the bands produced by the non-xylanase producers were much lighter compared with the xylanase-producing clones. Clone 1 had the brightest band and was used for all restriction analysis.

The primers designed in section 2.2.1.6 were designed to be cloned using *Eco*RI and *Xba*I. However, restriction analysis revealed that the *Eco*RI restriction site was no longer functional (Fig. 2.7). Restriction using different pairs of enzymes revealed that restriction of clone 1 with *Xho*I and either *Pst*I or *Xba*I produced the desired smaller

DNA fragment of 786 bp (Fig. 2.7). The plasmid from clone 1 was designated pX4 and is 3714 bp in size. This plasmid served as the template for the entire study and also as the unmutated control XynA.

Random mutagenesis is an important tool for exploring and optimizing protein properties. A large pool of gene variants is created, most commonly using ep-PCR (Leung *et al*., 1989; Cadwell and Joyce, 1994). After translation, the resulting protein library can be screened and mutants that outperform the wild-type are retained. In this study, a range of mutagenic conditions was explored for the random mutagenesis of *xynA* to create unique *xynA* variants. These included conditions recommended by a benchmark article on mutagenic PCR (Matsumura and Ellington, 2001), another that utilized a unique dITP and $Mn^{2+}$ method for mutagenesis of a cutinase gene from *F. solani pisi* (Xu *et al*., 1999), a condition with imbalanced dNTP concentrations used for mutation of *N. patriciarum* xylanase (Chen *et al*., 2001) and finally the use of a commercial mutagenesis kit, *viz*., Diversify Random Mutagenesis kit (CLONTECH).

With reference to mutagenesis reactions with imbalanced dNTPs, when the limiting dNTPs have been depleted, it is possible that misincorporation of one of the other bases is favoured. However, this misincorporation can be severely biased since it is widely known that *Taq* polymerase prefers certain types of base-pair changes over others, independent of the $Mn^{2+}$ concentration used. Mutations generated by this enzyme are dominated by AT→GC transitions and AT→TA transversions (Casson and Manser, 1995; Lin-Goerke *et al*., 1997; Shafikani *et al*., 1997). Consequently, all types of mutations are not equally represented and a high GC genetic content can further significantly reduce the average mutation frequency and library variation. Misincorporation is stimulated when inosine is present, since it is known to be incorporated by *Taq* polymerase, although with a four-fold reduced efficiency. In the subsequent PCR cycle, this would result in the incorporation of any of the three compatible nucleotides (Leung *et al*., 1989; Spee *et al*., 1993). Most of the conditions tested in this study served to address this deficiency.

A bright 786 bp band was produced after ep-PCR of *xynA* under all conditions tested (Fig. 2.8). Further examination of the bands revealed that products in lanes 4, 5, 10 and 11 were significantly less dense than the bands arising from the other conditions.

These lanes correspond to conditions C, DI, F and G respectively - the last three of which are extremely mutagenic. Lighter bands generally imply a high mutagenic rate in ep-PCR since the *Taq* polymerase does not work optimally. Therefore, it cannot form as much product compared to normal circumstances. This was an early indication that these conditions could generate novel mutant xylanases. Generally, with regards to transformation, as the degree of mutagenesis in ep-PCR reactions increases, the number of transformants obtained also decreases because less DNA is amplified during the ep-PCR reactions. Consequently fewer colonies are obtained and screened. However, with the advances in *in vivo* and restriction-free cloning, even these drawbacks are slowly being eradicated. Brighter bands, conversely, are indicative that the reaction largely proceeds as normal and fewer enzyme variants of unique capability can be expected. Thus, a large number of colonies need to be screened in order to obtain beneficial mutations.

Directed mutagenesis is a key method for the improvement of enzyme thermal tolerance as it allows the random selection of mutations that improve thermal tolerance. Its fundamental advantage over rational approaches is that it forces the enzyme to find the best solution for surviving exposure to heat. The solution might be increased thermostability, improved folding reversibility or a combination of the two and it is characteristic to a given enzyme and rarely predictable (Robertson and Steer, 2004). Because the fraction of proteins retaining function appears to decline exponentially with increasing numbers of amino acid substitutions, low mutational rates seek to create mutational diversity without destroying activity so that improved clones can be found (Drummond *et al*., 2005).

The most challenging aspect of directed evolution is to develop a screen suitable and stringent enough to detect suitable mutants. Probably the most universal assay type used in enzyme discovery is direct expression screening using substrates with bonds attacked by the enzyme class of interest. Because bond attack in substituted surrogates may not define the ultimate enzyme of interest, this method of screening usually requires tiered screening to first identify a group of enzyme candidates followed by more specific screening using low-throughput methods (Robertson and Steer, 2004).

Tiered screening was also used in this study and was three-fold. Transformation mixtures were plated on LB-RBB-xylan plates. Only clones capable of producing a fully functional xylanase would be able to break the linkage between the xylan and the RBB dye to produce colourless xylooligosaccharides which is visible as a zone of hydrolysis around the colony (Fig. 2.5). This served as the first screen - to select only xylanase-producing clones. Mutation sometimes produces clones that contain the gene of interest however, it is often so heavily mutated that an active, fully-functional protein is no longer produced (Malakauskas and Mayo, 1998; Drummond *et al*., 2005; Vanhercke *et al*., 2005). Also mutations are more likely to destroy enzyme activity rather than enhance it.

Nine hundred and sixty positive clones were obtained collectively for conditions B to G (Table 2.2). Normally in large-scale commercial mutagenesis studies, tens of thousands of clones are screened usually with the assistance of an automated robotic picker after cloning of the entire mutated PCR product in order to obtain the most novel and robust target protein. For the purposes of this study, the library size was kept small with only a fraction of the mutated products cloned. However, it is important to remember that the effective size of a library is not the number of mutants screened or the number usually reported, but rather the number of unique mutants screened (Drummond *et al*., 2005). Although only ninety-one clones were screened for condition D, almost 11% of these mutants retained more than 60% of their activity after heat treatment (Table 2.2), reinforcing that the mutational method has much more impact in directed evolution than merely the total number of clones screened.

The second aspect of the screen involved searching for xylanase variants tolerant to high temperatures. Directed mutagenesis is a key method for the improvement of enzyme thermal tolerance as it allows the random selection of mutations that improve thermal tolerance. Its fundamental advantage over rational approaches is that it forces the enzyme to find the best solution for surviving exposure to heat. The solution might be increased thermostability, improved folding reversibility or a combination of the two and it is characteristic to a given enzyme and barely predictable (Knappik and Plückthun, 1995; Yano and Poulos, 2003).

Thus, the mutant xylanase library generated through random mutagenesis was grown and the xylanases tested in a liquid assay before (0 min) and after (40 min) exposure to high temperature (80°C) and the results are summarized in Table 2.2. Conditions D and G generated the most number of thermostable mutants. However, their catalytic activities were severely compromised. Condition E, on the other hand, had excellent improvement in its substrate turnover rate but; stability of the variants was rather poor.

These findings echo that of previous work where improvements in thermal stability have deleteriously affected catalytic activity (Schoichet *et al*., 1995; Palackal *et al*., 2004; Fenel *et al*., 2006). It seems that reduced flexibility is a necessary consequence of thermostabilization, i.e., more stable proteins are less prone to have their structures perturbed by thermal fluctuations and therefore appear less flexible (Varley and Pain, 1991; Zavodsky *et al*., 1998; Finke *et al*., 2000). This change however, is counterproductive since an enzyme must be quite flexible in order to change its shape, degrade the substrate to its constituent monomers and then quickly return to its original conformation to convert more substrate (Arnold, 2001; Arnold *et al*., 2001; Wintrode *et al*., 2003). Thus, often thermal stability and catalysis work against each other and it is indeed rare, but not impossible, to improve both properties simultaneously in a single enzyme as was found in a landmark study conducted on the laboratory evolution of the ρ-nitrobenzyl esterase from *B. subtilis* which resulted in the co-evolution of both activity as well as thermostability in this enzyme (Giver *et al*., 1998). Other protein engineering studies also show that this trade-off can be eradicated under non-natural selective pressures and this is dependent on the gene of interest and its tolerance to mutation as well as the location of the mutation and structural change induced by the mutation concerned (Van den Burg *et al*., 1998; Williams *et al*., 1999; Miyazaki *et al*., 2001; Bjørk *et al*., 2003; Declerck *et al*., 2003).

It is theorized that the apparent trade-off behaviour does not reflect physical but rather evolutionary constraints. It is possible that natural selection actively avoids highly stable, highly active enzymes. An enzyme that is too stable, for example, might be resistant to degradation within the cell. For thermophilic organisms, high activity at moderate temperatures could translate into catalysts that are so active at elevated physiological temperatures that they impede cellular function. Another plausible

explanation is that as mutations accumulate during the course of evolution, unconstrained properties will drift downwards and this drift would tend to push enzyme function towards the biologically relevant minimum (Kiefhaber *et al*., 1991; Kaufmann, 1993; Arnold *et al*., 2001; Eijsink *et al*., 2005)

Screening for thermal tolerance is complicated because of the variability of enzyme concentration. Variants within a library may have variable expression levels and affect cell density and colony size, which can all influence the amount of each individual enzyme variant tested in the screen. As such, to eliminate false positives, the comparison of heated and non-heated samples is essential in the identification of improved variants (Zhang *et al*., 2003; Palackal *et al*., 2004). The comparison can be made in the primary screen or in a secondary screen where positive mutants identified in the primary screen are re-tested and compared under heated and non-heated conditions.

It is for this reason, that the most thermostable mutants were re-grown, their xylanases extracted and retested at the same temperature but subjected to a longer heat challenge of 90 min. XynA, although it has a $t_{50}$ of 201 min at 70°C (Singh *et al*., 2000b), loses almost 80% of its activity after 90 min at 80°C (Fig. 2.10). With the exception of G53, all the mutant xylanases tested have better thermal stability, albeit a lower activity, than XynA. Mutant G41 retained 75% of its relative activity under these conditions but its catalytic activity was low. Similar results were observed in a previous study which focussed on improving the stability of a *Bacillus* xylanase (Palackal *et al*., 2004). The most thermostable mutant Q11H displayed a notable reduction in enzymatic activity yet its melting temperature was increased from 64°C to 70°C. SDM of *T. reesei* xylanase II increased the $t_{50}$ to 100 min yet the activity was substantially reduced (Turunen *et al*., 2001). Interestingly, most papers on engineering thermostable xylanases measure stability after a heat exposure time of 5 - 15 min only. In this study, however, longer incubation times of 40 min, 90 min and finally 360 min was tested (chapter four).

The crude xylanase mutants were then subjected to screening on alkaline xylan plates to select for alkaline stable variants (Fig. 2.9). This screen was less effective than the previous short-term thermostability screen because more mutants made it through to

the second stage of alkaline screening, and numerous mutants were screened even at pH 10 for 90 min. A plate assay could be misleading since a highly active enzyme could give a false positive result. Thus, it became necessary for the mutants selected from the plates to be subjected to a short-term liquid assay. Table 2.3 summarizes the short-term alkaline liquid assay screening of sixty-two mutants at pH 10. Thirty-nine percent of the mutants displayed better activity than XynA whilst 34% had more than 60% remaining activity after 40 min alkaline treatment. XynA retains 30% of its activity after 90 min at pH 10 whilst mutant G53 only lost only 7% activity under the same conditions making it the most alkaline stable mutant to be generated in this study (Fig. 2.11).This proves that the mutation did not improve overall activity of the enzyme, but rather only its stability at pH 10, which is a drawback of the plate screen used in this study. G41, on the other hand, has only 40% activity after treatment under identical conditions. A noteworthy observation was that the catalytic activity of the more alkaline stable variants was comparable to XynA; which is in complete contrast to the observations regarding thermal stability in this study. Previous studies on the alkaline stability of other xylanases, show similar findings (Chen *et al*., 2001; Turunen *et al*., 2002; Palackal *et al*., 2004). It seems that xylanases are much more sensitive to high temperatures than alkaline environmental conditions.

The findings in this chapter highlight that enzymatic properties often diverge from each other during evolution, with one property taking precedence over the other. After screening the entire mutant library only mutant G41 was sufficiently thermostable whilst mutant G53 was extremely alkali-tolerant. In contrast, G41 was found to have low alkaline stability at pH 10 and G53 displayed dismal thermal tolerance at 80°C. G53 had catalytic activity (197 nkat/ml) almost comparable to XynA (216 nkat/ml) whilst G41 had lower activity in orders of magnitude (22 nkat/ml). If random mutagenesis were to continue, the properties would diverge even further from each other on the "directed" evolutionary tree. The perfect way of preventing this would be to assemble these properties into a single xylanase using DNA shuffling or sexual recombination which was the objective of the subsequent chapter in this study.

# CHAPTER 3: DNA RECOMBINATION OF *xynA* VARIANTS

## 3.1 INTRODUCTION

In nature, radical changes in evolution originate from recombination. The variation produced by the crossovers between chromosomes of one's father and mother in the production of eggs and sperm, and the highly variable parts of antibodies are good examples of this process (Kurtzman *et al*., 2001). Homologous recombination is a ubiquitous process that plays an important role in species adaptation and survival. Its importance is illustrated by its duality of functions: increasing genetic diversity in populations by reshuffling genes and preserving genetic integrity by aiding in the repair of damaged genes (Stahl, 1987; Forrest, 1993). Natural *in vivo* recombination mechanisms, however, usually operate at low efficiencies, eliciting changes in gene structures or functions usually after ten or more generations (Holland, 1992; Kuchner and Arnold, 1997b).

Various approaches have been developed to mimic and accelerate nature's recombination strategy to direct the evolution of protein function and are summarized in Table 1.2 (section 1.5.2.6). *In vitro* recombination methods generally offer higher recombination efficiencies and greater experimental flexibility than *in vivo* approaches. While random mutagenesis methods introduce point mutations randomly into a single parent sequence to produce a library of progeny sequences, DNA recombination methods entail the block-wise exchange of genetic variations among multiple parent sequences created in the laboratory or existing in nature to produce a library of chimeric progeny sequences.

The key advantage of DNA recombination is its ability to accumulate beneficial mutations present in a population. It is also capable of simultaneously removing deleterious mutations which would otherwise accumulate in an asexual population in a process known to geneticists as Müller's ratchet (Müller, 1932). Such interchangeable processes may greatly accelerate the evolution of a protein or nucleic acid molecule of interest toward a specific function. It was demonstrated in computational simulation studies that DNA recombination plays a critical role in the

evolution of biological systems (Maynard Smith, 1988; Forrest, 1993; Moore *et al*., 1997).

The 'staggered extension process' (StEP) is a type of DNA family shuffling that is not only technically simple but conceptually novel as well. It was developed in 1998 by Zhao *et al.* and is based on template switching during polymerase-catalyzed primer extension. The method relies on high homology between parent DNA templates to create full-length progeny with multiple crossovers. As with any type of DNA shuffling, a high degree of recombination is important to get all possible combinations of mutations whilst homology is required for reconstruction of the full-length gene.

As illustrated in Fig. 3.1, the StEP method uses full-length genes as templates for the synthesis of chimeric progeny genes. It relies on repeated cycles of melting, annealing and extension to build up the full-length gene. However in StEP, fragments are added sequentially to the end of a growing strand. The growing strand is prevented from reaching its full length by keeping the extension time very short. This results in only partial elongation of a strand in any one extension step. The strand is then melted from its template and may anneal in the next step to a different template leading to a crossover. StEP is continued until full-length genes are formed (Kurtzman *et al*., 2001; Neylon, 2004; Zhao, 2004). It is interesting to note that StEP is, in some ways, reminiscent of the template-switching recombination mechanism that contributes to the evolutionary potential of retroviral populations (Hu *et al*., 1997).

The overall aim of this study was to use random mutagenesis to improve both the thermal and alkaline stabilities of *xynA* from *T. lanuginosus*. In the previous chapter, it was noted that both these traits appeared to evolve separately in the resulting mutant library, with the mutant xylanases displaying either good alkaline or thermal stabilities, but not both. The ideal way of combining these properties into a single xylanase would be the use of a DNA shuffling technology such as StEP to accomplish this. Both the thermostable mutant G41 and the alkaline stable mutant G53 were created through random chemical mutation of *xynA*. Both these mutants would thus have a high homology towards each other, which is a prerequisite for StEP recombination since it is needed for the reconstruction of the full-length gene, which will be translated into a protein that should be able to fold into a soluble and active

enzyme. Thus, the objectives of this phase of the study were to assemble both alkaline and thermostable properties of the parents G41 and G53 into a single xylanase using StEP recombination, to screen for recombinants that display these properties and finally to sequence the parents and recombinants to determine the amino acid substitutions responsible for the biochemical changes observed.

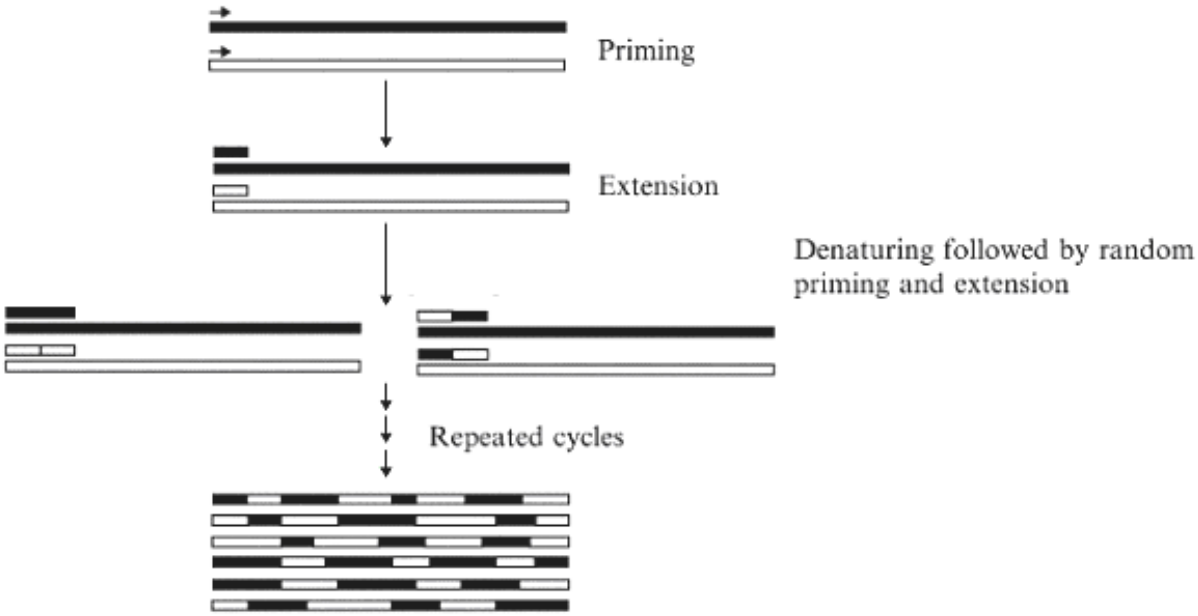


**Fig. 3.1** Principle of the StEP recombination method. For simplicity, only one primer and single strands from two DNA templates (open and solid blocks) are shown. During priming, oligonucleotide primers anneal to the denatured templates. Short fragments are produced by brief polymerase-catalyzed primer extension. During subsequent random annealing-abbreviated extension cycles, fragments randomly prime the template (template switching) and extend further. The process is repeated for many cycles until the full-length genes are produced. The full-length chimeric genes can be amplified in a standard PCR (optional) (Zhao *et al*., 1998; Zhao, 2004).

## 3.2 MATERIALS AND METHODS

### 3.2.1 Plasmids and plasmid isolation

The plasmid constructed in chapter 2, pX4, was initially used for optimization of StEP recombination method and served as the control for all subsequent recombination studies. For the final optimized StEP reactions, plasmids isolated from mutants G41 and G53 served as the templates.

A loopful of glycerol-preserved culture was removed aseptically and streaked onto fresh LB-ampicillin (100 μg/ml) agar plates and incubated at 37°C overnight. A single colony was used to inoculate 5 ml LB medium containing 100 μg/ml ampicillin and grown for 12 - 16 h at 37°C. Plasmids pX4, G41 and G53 were isolated from the *E. coli* clones harbouring them using the FastPlasmid Mini kit (Eppendorf), according to the manufacturer's instructions.

### 3.2.2 DNA quantification and agarose electrophoresis

DNA concentration was estimated spectrophotometrically as described in section 2.2.1.3. Preparation of λ DNA molecular weight marker and agarose electrophoresis for verification of recombination was carried out as described in sections 2.2.1.4 and 2.2.1.5, respectively.

### 3.2.3 Optimization of StEP recombination PCR method

For optimization of the recombination method, pX4 served as the template. Initially, the PCR was carried out as described in the original publication (Zhao *et al*., 1998). Ten nanograms of pX4, 1.5 μM each of T3 and T7 primers (IDT), 2 mM dNTP mix, 10X PCR buffer, 1.5 mM $MgCl_2$ and 2.5 U *Taq* polymerase were added together in a total volume of 50 μl.

The program for amplification was as follows:

| | | |
|---|---|---|
| Denaturation | : | 30 s at 94°C |
| Primer annealing | : | 7 s at 42°C |

Each PCR reaction comprised a total of 80 cycles and the reactions were performed using a PCR Genius thermal cycler (Techne). Samples were removed after every 20 cycles and analyzed on 0.8% agarose gels to determine amplification of the final product. The initial StEP reaction carried out according to the original method was unsuccessful.

To eliminate the short annealing time as a possible cause of unsuccessful PCR, a reaction containing the standard concentration of reactants (E) was carried out under the same PCR conditions. All parameters are outlined in Table 3.1. DNA (10 ng), PCR buffer (1X final concentration) and $Mg^{2+}$ concentrations (1.5 mM) were kept constant for all reactions. Primer concentrations were varied. Reaction pairs A and C and B and D were duplicate reactions carried out to eliminate experimental error. The PCR program outlined above was used.

**Table 3.1** StEP PCR carried out for 30 s at 94°C and 7 s at 42°C

| Reaction | Primers (μM) | dNTPs (mM) | *Taq* (U) |
|---|---|---|---|
| A | 1.5 | 2 | 2.5 |
| B | 0.5 | 2 | 2.5 |
| C | 1.5 | 2 | 2.5 |
| D | 0.5 | 2 | 2.5 |
| E | 0.5 | 0.1 | 1 |

Only reaction E underwent successful amplification after 80 cycles, which meant that the short annealing time was suitable. The next step focussed on varying dNTP and primer concentration and decreasing the annealing time to 5 s. These parameters are outlined in Table 3.2. Samples were removed after 20 and 80 cycles and analyzed on agarose gels to determine amplification.

**Table 3.2** StEP PCR carried out under conditions of 30 s at 94°C and 5 s at 42°C

| Reaction | Primers (μM) | dNTPs (mM) | *Taq* (U) |
|---|---|---|---|
| A | 1.5 | 1 | 2.5 |
| B | 1.5 | 0.5 | 2.5 |
| C | 1.5 | 0.25 | 2.5 |
| E (standard) | 0.5 | 0.1 | 1 |
| F | 0.5 | 1 | 2.5 |
| G | 0.5 | 0.25 | 2.5 |

Reaction E still gave the best results, so these conditions were chosen for the final StEP reactions of G41 and G53. Thermal stability was considered more important than alkaline stability, thus twice the amount of G41 plasmid was added to each reaction as compared with G53. The final concentration of DNA was still kept as 10 ng. One StEP reaction was made mutagenic by adding 640 μM $MnSO_4$, the concentration used in the Diversify Random Mutagenesis kit (CLONTECH). The second StEP reaction contained no mutagen. PCR buffers containing $Mg^{2+}$ were used; 0.5 μM each of T3 and T7 primer and 1U *Taq* was added. The annealing temperature was dropped further to 4 s and the products were analyzed on agarose gels after 20 and 80 cycles.

Upon successful amplification, the PCR bands were excised from the gel and recovered using the GFX PCR DNA and Gel Band Purification kit (Amersham), according to the manufacturer's instructions. Aliquots of the purified PCR products were run on an agarose gel to ascertain successful elution from the purification column.

Purified PCR products were restricted with *Xho*I and *Pst*I at 37°C for 2 h to create sticky ends compatible with the pBSK vector to facilitate ligation between the vector and insert. The restricted products were purified from solution using the GFX kit, according to the manufacturer's instructions.

### 3.2.4 Cloning of PCR products

The purified and restricted PCR products and pBSK vector were quantified and ligated as previously described (sections 2.2.1.3 and 2.2.1.10). Ligated products were subsequently transformed into SEM-competent cells onto LB plates supplemented with RBB-xylan and ampicillin (sections 2.2.1.11 and 2.2.1.12).

### 3.2.5 Growth of recombinants and enzyme extraction

Five hundred and thirty one recombinants were collectively obtained after transformation and plating onto LB-RBB-xylan. Recombinants were grown and preserved and their enzymes extracted as described in section 2.2.4. Recombinants arising from the normal StEP reaction were designated with the alphabet ‘S’ and those created via a mutagenic StEP reaction were designated with the alphabet ‘M’.

### 3.2.6 Short-term screening for thermo- and alkali-stable recombinants

Wild-type XynA served as the unmutated control for all screening assays and the parent mutants G41 and G53 were also used as controls. The recombinant xylanase lysates were analyzed for a duration of 40 min at 80°C for determination of thermal stability, whilst the lysates were initially screened on pH 10 plates and those with better zones than the control xylanases were subjected to a liquid assay for determination of alkaline activity as outlined in sections 2.2.5.1 and 2.2.5.2.

### 3.2.7 Long-term thermal and alkaline stabilities of recombinants

Recombinants possessing enzymes displaying more than 60% of residual activity after heat and alkaline treatment were grown in larger volumes, then lysed and analyzed over 90 min as described in section 2.2.5.

### 3.2.8 Automated DNA sequencing and sequence analysis of *xynA* variants and recombinants

DNA sequencing was carried out to determine the mutations that were responsible for the observed changes in stabilities of some xylanase variants. Automated sequencing of *xynA* variants were performed by Inqaba Biotechnical Industries (Pretoria, South Africa) using a Genetic Analysis System SCE2410 with 24 capillaries (SpectruMedix LLC). Chain terminating sequence reactions were performed using the BigDye® Terminator version 3.1 sequencing kit (Applied Biosystems). All sequencing reactions were performed according to the manufacturer's instructions and preliminary analysis were done using BaseSpectrum version 2.1.1 software (SpectruMedix).

After obtaining the raw DNA sequencing data, they were was initially processed using the Chromas Lite software package (Technelysium, version 2.0) downloaded from the company server (www.technelysium.com.au) and both DNA strands were edited to yield complete gene sequences. The DNA sequences were then translated into their protein counterparts using the Translate tool from the ExPASy website (www.expasy.org/tools/dna.html) and aligned to the wild type parent using the CLUSTALW (version 1.81) alignment program on the GenomeNet server (www.clustalw.genome.ad.jp).

## 3.3 RESULTS

### 3.3.1 Optimization of StEP recombination PCR method

When the StEP reaction was carried out according to the original protocol (Zhao *et al*., 1998), PCR products were not obtained. The PCR run was repeated with modifications to the primer concentration. The annealing time for a standard PCR reaction is 1 min but for the StEP reaction, a short annealing time of 7 s was used. To eliminate the annealing time as a possible reason for failure of the StEP method, a PCR reaction containing a standard concentration of reactants was carried out under StEP method conditions as outlined in Table 3.1. Following separation on 0.8% agarose (Fig. 3.2), it was found that only the standard reaction (lane 6) underwent amplification after 80 cycles. Thus, the reagents and PCR program were adequate, which meant the concentrations of the primers and dNTPs were not optimal.

Since a dense PCR product was obtained with an annealing time of 7 s, this time was further decreased to 5 s whilst the dNTP concentration was modified in concentrations ranging from 0.25 – 1 mM, whilst the primer concentrations were altered in the range of 0.5 – 1.5 μM as outlined in Table 1.2. Agarose electrophoresis showed that the standard reaction (lanes 8 and 9) still gave the best results without extra bands as seen in lanes 5, 7 and 13 after 80 cycles (Fig. 3.3). Conditions in lanes 2, 3, 10 and 11 yielded no visible PCR products. As expected, products obtained after 80 cycles were more intense than products obtained after 20 cycles as depicted in lanes 4, 5, 6, 7, 8, 9, 12 and 13. Products from the standard reactions in lanes 8 and 9 were less intense after both 20 and 80 cycles.

Final optimized StEP reactions were thus carried out using normal reactant concentrations and the parent DNA (G41:G53) was added in a ratio of 2:1 respectively. One reaction was made mutagenic by addition of $Mn^{2+}$. When comparing the density of the bands after 80 cycles, it can be observed that the band in lane 5, which represents the mutagenic StEP reaction, is lighter than that seen in lane 3 for the normal StEP reaction (Fig 3.4). However, the 20 cycle product in lane 4 is brighter than that in lane 2. The extra bands observed in all lanes represent the parent plasmid DNA.

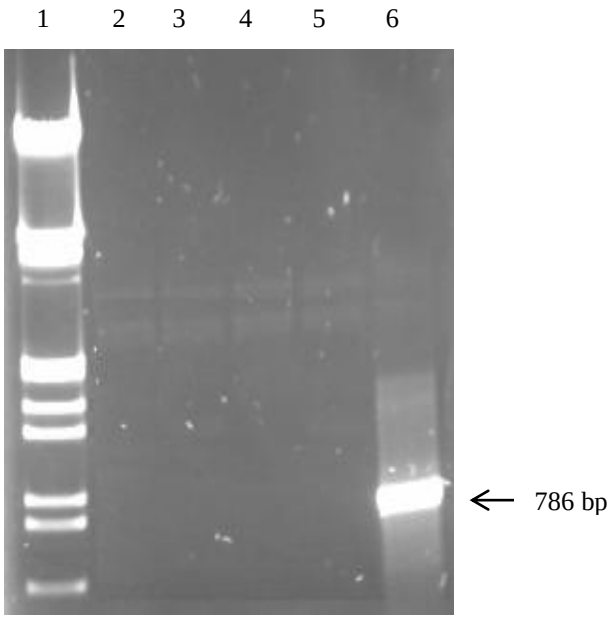


**Fig. 3.2** Optimization of StEP PCR reaction conditions as described in Table 3.1 at 94°C for 30 s and 42°C for 7 s for a total of 80 cycles. All reactions contained 1.5 mM $MgCl_2$, 1X PCR buffers and varying primer concentrations. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: 1.5 µM each T3 and T7 primers, 2 mM dNTPs, 2.5 U *Taq*; Lane 3: 0.5 µM each T3 and T7 primers, 2 mM dNTPs, 2.5 U *Taq*; Lane 4: 1.5 µM each T3 and T7 primers, 2 mM dNTPs, 2.5 U *Taq*; Lane 5: 0.5 µM each T3 and T7 primers, 2 mM dNTPs, 2.5 U *Taq*; Lane 6 (standard): 0.5 µM each T3 and T7 primers, 0.1 mM dNTPs, 1 U *Taq*.

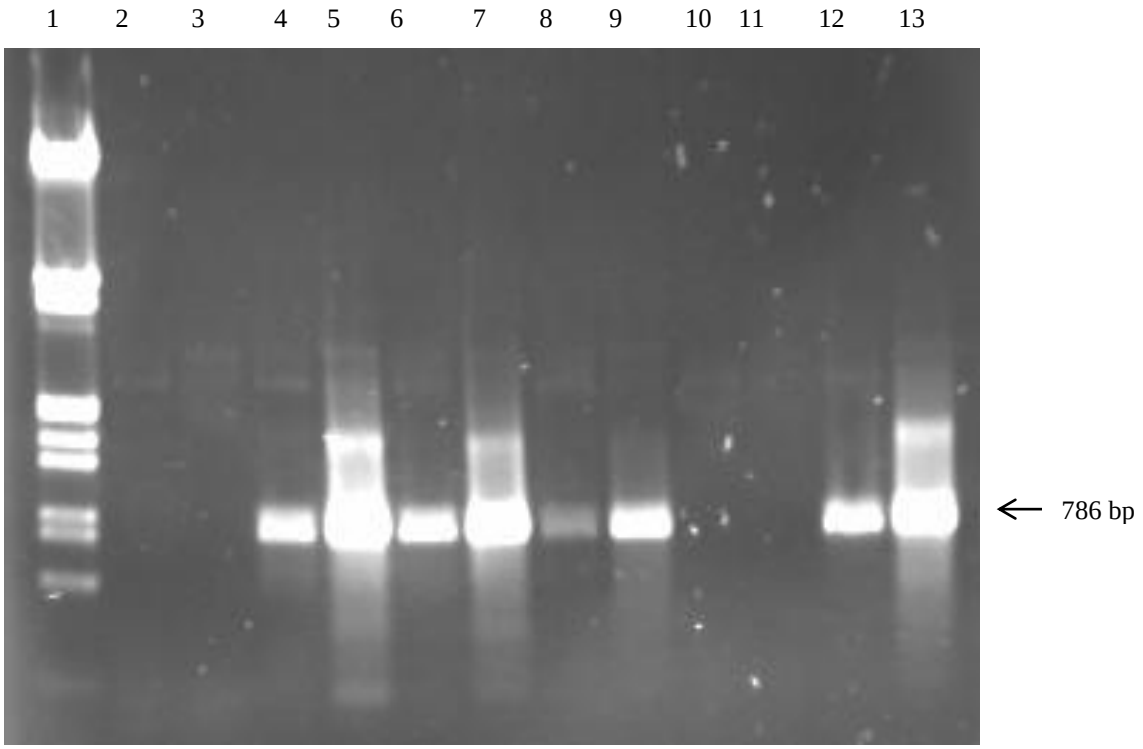


**Fig. 3.3** Optimization of StEP PCR reaction conditions as described in Table 3.2 at 94°C for 30 s and 42°C for 5 s at the end of 20 and 80 cycles. All reactions contained 1.5 mM $MgCl_2$ and 1X PCR buffers. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: 1.5 µM each T3 and T7 primers, 1 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 3: 1.5 µM each T3 and T7 primers, 1 mM dNTPs, 2.5 U *Taq* after 80 cycles; Lane 4: 1.5 µM each T3 and T7 primers, 0.5 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 5: 1.5 µM each T3 and T7 primers, 0.5 mM dNTPs, 2.5 U *Taq* after 80 cycles; Lane 6: 1.5 µM each T3 and T7 primers, 0.25 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 7: 1.5 µM each T3 and T7 primers, 0.25 mM dNTPs, 2.5 U *Taq* after 80 cycles; Lane 8 (standard): 0.5 µM each T3 and T7 primers, 0.1 mM dNTPs, 1 U *Taq* after 20 cycles; Lane 9 (standard): 0.5 µM each T3 and T7 primers, 0.1 mM dNTPs, 1 U *Taq* after 80 cycles; Lane 10: 0.5 µM each T3 and T7 primers, 1 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 11: 0.5 µM each T3 and T7 primers, 1 mM dNTPs, 2.5 U *Taq* after 80 cycles; Lane 12: 0.5 µM each T3 and T7 primers, 0.25 mM dNTPs, 2.5 U *Taq* after 20 cycles; Lane 13: 0.5 µM each T3 and T7 primers, 0.25 mM dNTPs, 2.5 U *Taq* after 80 cycles.

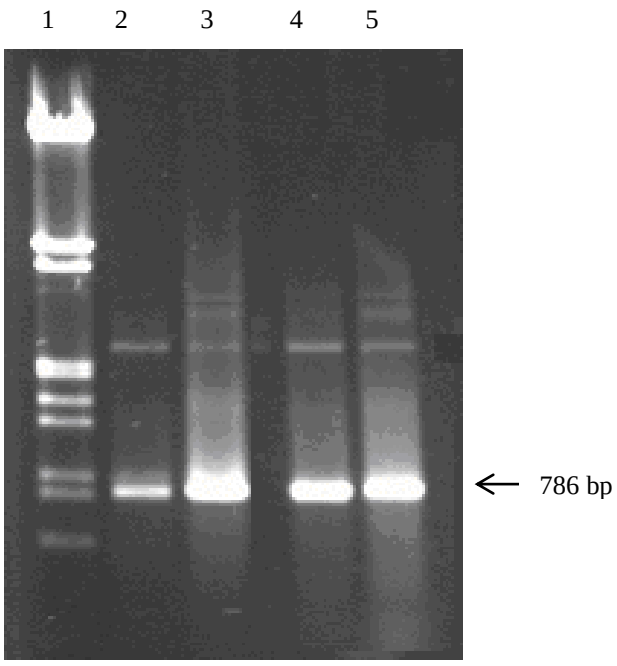


**Fig. 3.4** StEP PCR recombination of G41 and G53 at 94°C for 30 s and 42°C for 4 s at the end of 20 and 80 cycles. All reactions contained 10 ng total DNA, 1.5 mM $MgCl_2$, 0.5 µM each T3 and T7 primers, 1X PCR buffer and 1 U *Taq*. Lane 1: λ DNA molecular weight marker restricted with *Eco*RI and *Hin*dIII; Lane 2: StEP reaction without mutagen after 20 cycles; Lane 3: StEP reaction without mutagen after 80 cycles; Lane 4: mutagenic StEP reaction with 640 µM $Mn^{2+}$ and 0.2 mM dGTP after 20 cycles; Lane 5: mutagenic StEP reaction with 640 µM $Mn^{2+}$ and 0.2 mM dGTP after 80 cycles.

### 3.3.2 Short-term thermal and alkaline screening of recombinants

A total of 531 recombinants were screened for thermal and alkaline stability: 393 arose from the 'S' or normal StEP reaction whilst 138 were from the 'M' or mutagenic StEP reaction. Recombinants displaying greater stability than the thermostable parent G41 and with more stability than the alkaline stable parent G53 were considered to be successful.

It was observed that 63 of the 393 recombinants from the normal StEP recombination displayed more than 70% of their initial activities after being exposed to 80°C for 40 min as opposed to the 37 mutagenic StEP recombinants which have similar or better thermal stabilities than parent G41. Collectively, this represents 19% of the total xylanase variant library. This effectively meant that 81% of the recombinant library displayed poor thermal stability.

With respect to activity, 34 recombinants displayed similar activity to parent G53 (252 nkat/ml) and XynA (250 nkat/ml). Only 3 recombinants, *viz*., M6 (298 nkat/ml), M7 (283 nkat/ml) and M25 (308 nkat/ml), displayed marginally higher catalytic activities than G53. This means that most of the recombinants attained poor overall activities.

After screening on pH 10 agar plates, none of the clones produced better zones than G53. Consequently, xylanase activity was not quantitatively assessed at pH 10 and 60°C for 40 min. It was thus decided to screen only the thermostable recombinants to assess for an increase in alkaline stability at pH 10.

### 3.3.3 Long-term thermal and alkaline stabilities of recombinants

Twenty recombinants were quite stable at 80°C for 40 min and they were grown on a large scale and their xylanases extracted and tested for up to 90 min at the same temperature. Only the profiles of the 5 best recombinants are shown in Fig. 3.5. Thermostable parent G41 and alkaline stable parent G53, created by initial random mutation of *xynA*, and wild type XynA served as the controls for this experiment. One recombinant, S325, was found to have better thermal stability than thermostable parent G41 and retained 85% of its residual activity after 90 min of heat treatment. G41 retained 75% of its activity whilst G53 retained 10% of its initial activity after similar treatment. However, S325 had slightly lower catalytic activity (18.75 nkat/ml) than G41 (25.33 nkat/ml). Another recombinant S340 displayed intermediate stability between both parents G41 and G53 and displayed 54% residual activity with a marginally better activity (31.16 nkat/ml) than G41. Recombinants from the mutagenic StEP reaction included M28, M32 and M34. Recombinant M28 had 55% residual activity; M32 had 73% residual activity whilst M34 had 64% remaining activity. The initial activities of these recombinants, however, (25 nkat/ml, 23 nkat/ml and 21 nkat/ml, respectively) were relatively poor in comparison to G53 and XynA.

The lysates containing xylanases from the recombinant library were screened on 0.1% xylan plates (pH 10) as previously described (section 2.2.5.2). However, none of the recombinants produced larger or clearer zones than the alkaline stable parent G53 and were therefore not subjected to short-term alkaline screening at pH 10. The 20

thermostable recombinants, however, were subjected to alkaline testing to determine if these xylanases displayed any improvements in their ability to withstand long-term incubation at pH 10. The profiles of the 5 thermostable recombinants along with the control strains at pH 10 are depicted in Fig. 3.6. The immediate observation for this set of data is that G53 is still the most alkaline stable mutant to arise from this study, retaining 93% of its activity after incubation for 90 min at pH 10. However, what is interesting to note is that both recombinants S325 and S340 have 60% residual activity after similar treatment. This is a significant improvement on parent G41 which had only 40% remaining activity. The mutagenic StEP recombinants, however, varied slightly in terms of their stability at pH 10. The recombinants M28, M32 and M34 had 27%, 68% and 43% remaining activities, respectively, after 90 min at pH 10. These were good examples that highlighted the effect of mutation and recombination on G41, since M28 displayed inferior stability, M32 displayed superior stability and M34 displayed similar alkaline stability to this parent.

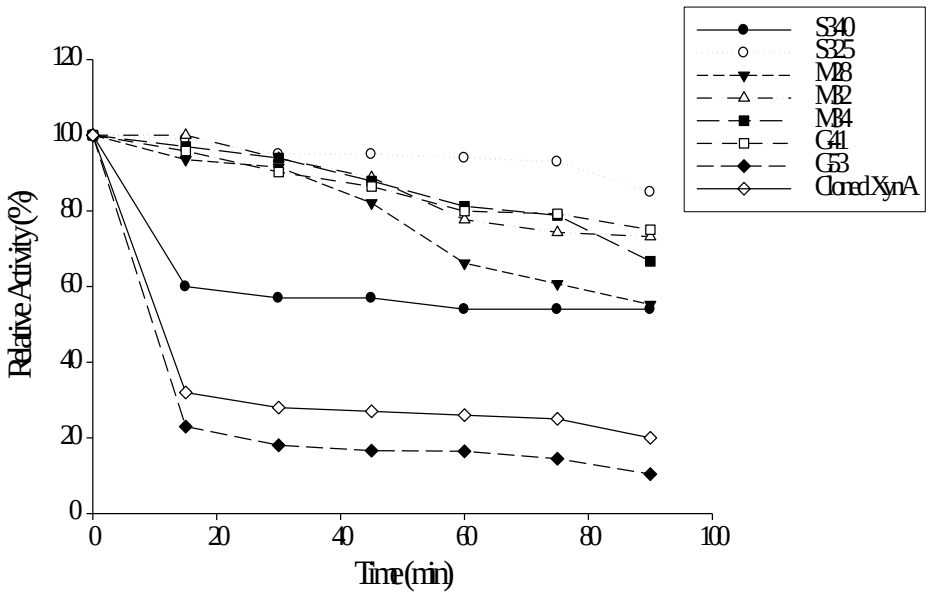


**Fig. 3.5** The effect of temperature on the stability of wild-type XynA, its mutant progeny and recombinants at 80°C and pH 6.5. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

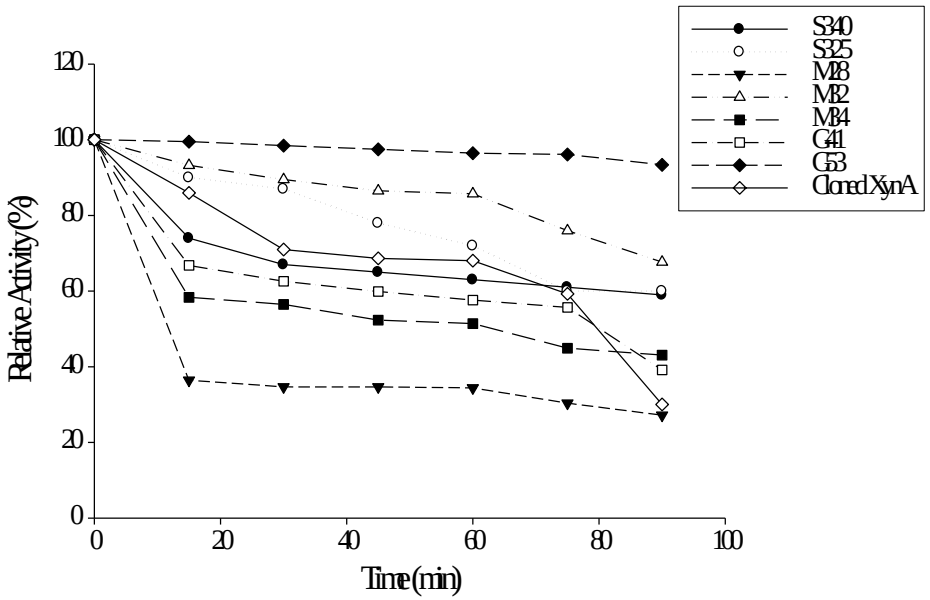


**Fig. 3.6** The effect of alkalinity on the stability of wild-type XynA, its mutant progeny and recombinants at pH 10 and 60°C. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

### 3.3.4 Sequence analysis of *xynA* variants and recombinants

The xylanase genes of the parent mutants (G41 and G53) and a few progeny recombinants (S325, S340, M28, M32, and M34) were sequenced to determine the amino acid substitutions that allowed them to exhibit the observed differences in stability under the tested conditions. The number and locations of the amino acid substitutions attained by the mutants and recombinants in this study as well as their sequence similarity to wild-type *xynA* are summarized in Table 3.3. The amino acid sequence alignment itself is reflected in Fig. 3.7. Parent G53 displays the greatest protein sequence similarity to *xynA* at 99%, parent G41 and S340 have a 98% similarity whilst S325 and M28 have 97% protein sequence identity whilst recombinants M32 and M34 have 96% similarity.

Alkali-stable mutant G53 has a single amino acid substitution (A54T) whilst thermostable mutant G41 has four such substitutions (K30E; W40R; T57A; K80R). Recombinant S340 has three mutations from parent G41 (W40R; T57A; K80R) while S325 had five mutations inherited from both parents (A54T; K30E; W40R; T57A; K80R). Mutagenic recombinant M28 inherited all four mutations from parent G41,

(K30E; W40R; T57A; K80R), as well as two new mutations (S178C; C185Y). Mutant M32 had a total of eight mutations; four from parent G41 (K30E; W40R; T57A; K80R) and four new ones (T26A; D128G; T177A; Q204R). Mutant M34 also attained four mutations from G41 (K30E; W40R; T57A; K80R) and four additional mutations (N24Y; Y45C; Q204R; F212L).

The protein sequence alignment (Fig. 3.7) revealed a total of fourteen mutations in all variants: eight are substituted for closely related amino acids; two are substituted for similar amino acids whilst four amino acid substitutions are unrelated to their predecessors. Seven substitutions are at the amino terminal end of the protein, which, for the most part, represent mutations inherited from the parents G41 and G53.

**Table 3.3** Summary of amino acid changes in *xynA* mutants and recombinants

| Mutant/Recombinant | % Similarity to *xynA* | Total no. of amino acid substitutions | Location of amino acid substitutions |
|---|---|---|---|
| G53 | 99 | 1 | A54T |
| G41 | 98 | 4 | K30E; W40R; T57A; K80R |
| S340 | 98 | 3 | W40R; T57A; K80R |
| S325 | 97 | 5 | K30E; W40R; A54T; T57A; K80R |
| M28 | 97 | 6 | K30E; W40R; T57A; K80R; S178C; C185Y |
| M32 | 96 | 8 | T26A; K30E; W40R; T57A; K80R; D128G; T177A; Q204R |
| M34 | 96 | 8 | N24Y; K30E; W40R; Y45C; T57A; K80R; Q204R; F212L |

```
xynA        MVGFTPVALAALAATGALAFPAGNATELEKRQTTPNSEGWHDGYYYSWWSDGGAQATYTN 60
G53         MVGFTPVALAALAATGALAFPAGNATELEKRQTTPNSEGWHDGYYYSWWSDGGTQATYTN 60
G41         MVGFTPVALAALAATGALAFPAGNATELEERQTTPNSEGRHDGYYYSWWSDGGAQAAYTN 60
S340        MVGFTPVALAALAATGALAFPAGNATELEKRQTTPNSEGRHDGYYYSWWSDGGAQAAYTN 60
S325        MVGFTPVALAALAATGALAFPAGNATELEERQTTPNSEGRHDGYYYSWWSDGGTQAAYTN 60
M28         MVGFTPVALAALAATGALAFPAGNATELEERQTTPNSEGRHDGYYYSWWSDGGAQAAYTN 60
M32         MVGFTPVALAALAATGALAFPAGNAAELEERQTTPNSEGRHDGYYYSWWSDGGAQAAYTN 60
M34         MVGFTPVALAALAATGALAFPAGYATELEERQTTPNSEGRHDGYCYSWWSDGGAQAAYTN 60
            ********************** *:***:********* **** ********:**:***

xynA        LEGGTYEISWGDGGNLVGGKGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
G53         LEGGTYEISWGDGGNLVGGKGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
G41         LEGGTYEISWGDGGNLVGGRGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
S340        LEGGTYEISWGDGGNLVGGRGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
S325        LEGGTYEISWGDGGNLVGGRGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
M28         LEGGTYEISWGDGGNLVGGRGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
M32         LEGGTYEISWGDGGNLVGGRGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
M34         LEGGTYEISWGDGGNLVGGRGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
            ********************:****************************************

xynA        VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
G53         VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
G41         VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
S340        VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
S325        VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
M28         VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTCGT 180
M32         VENFGTYGPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRASGT 180
M34         VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
            *******.**************************************************:.**

xynA        VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVG 225
G53         VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVG 225
G41         VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVG 225
S340        VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVG 225
S325        VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVG 225
M28         VQTGYHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVG 225
M32         VQTGCHFDAWARAGLNVNGDHYYRIVATEGYFSSGYARITVADVG 225
M34         VQTGCHFDAWARAGLNVNGDHYYRIVATEGYLSSGYARITVADVG 225
            **** ******************:*******:*************
```

**Fig. 3.7** Amino acid sequence alignment of parent mutants (G41 and G53) and recombinants (S325, S340, M28, M32 and M34) with each other and with wild type *xynA*. The alignment was done by using the CLUSTALW (version 1.81) alignment program on the GenomeNet server (www.ebi.ac.uk/clustalw). 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; '.' indicates conservation within weaker groups of amino acids. The absence of an alignment character implies that an unrelated amino acid was substituted. Amino acid substitutions that differ from *xyn*A are indicated in red boldface. Catalytic residues that make up the active site of *xynA* are shown in italics.

## 3.4 DISCUSSION

Genetic recombination is a key process for generating genetic diversity, making organisms adapt and evolve in nature and many *in vitro* recombination techniques such as DNA shuffling (Stemmer, 1994a; 1994b), RATCHITT (Coco *et al*., 2001), ITCHY (Ostermeier *et al*., 1999), SHIPREC (Sieber *et al*., 2001) and StEP (Zhao *et al*., 1998) have been developed to mimic and accelerate nature's recombination strategy.

In the previous chapter, random mutagenesis was used to evolve *xynA* to withstand conditions of high temperature and alkalinity. These properties however, were not concurrently displayed in a single enzyme, thus, the best thermostable mutant G41 and the most alkaline-stable mutant G53 were chosen as candidates for DNA shuffling to recombine these beneficial properties into a single xylanase. It could be argued at the outset that these mutations could potentially be generated by serial mutagenesis and screening. However, shuffling is a simpler and quicker strategy to accomplish the above objective. For example, the starting population of a library generated by mutagenic PCR typically contains almost 40% or more non-functional variants (Matsumura and Ellington, 2001), whereas most variants formed by DNA shuffling are functional. Thus, DNA shuffling allows a more streamlined exploration of sequence space and acquisition of novel protein phenotypes easily, as has indeed proven to be the case for a number of protein targets (Crameri *et al*., 1998; Ness *et al*., 1999; Joern *et al*., 2002) as well as in this study. The StEP method was chosen as the method of recombination in this study because unlike other recombination methods it can be performed using a pair of flanking primers in a single PCR tube and separation of parent templates from the recombined products is not necessary (Zhao *et al*., 1998).

The method required optimization before commencement of the actual recombination procedure. When the published procedure was followed, no amplification of the target DNA was observed, even after several attempts. The inclusion of a PCR reaction containing standard concentrations of reactants (Table 3.1), eliminated the short annealing time of 7 s as a possible reason for failure of the StEP reaction, since full-length product was obtained after 80 cycles (Fig. 3.2). Since the quality of the reagents and the PCR program were suitable, the only parameters requiring

optimization were dNTP and primer concentrations which were subsequently carried out using an even lower annealing time of 5 s (Table 3.2).

When this optimization was performed, it was noted that reactions A (Fig. 3.3 - lanes 2 and 3) and F (lanes 10 and 11) containing high concentrations of primer and dNTPs yielded no amplification products. There were products for reactions B, (lanes 4 and 5), C (lanes 6 and 7) and G (lanes 12 and 13) which all had lower than 1 mM dNTP concentration, implying that the concentration of this component in a StEP reaction is critical. Even though actions produced very bright PCR product bands after 80 cycles, they were unsuitable for the recombination as very bright PCR product bands were also produced after only 20 cycles, (lanes 4, 6 and 12). This is usually synonymous with a lower amount of template switching and higher amount of amplification, which is undesirable for StEP recombination. Generally, in directed evolution experiments, bright bands obtained after chemical mutagenesis, imply a smaller degree of mutation of the target DNA, which was also observed in this study and was discussed extensively in chapter two. The rule of thumb for any StEP reaction is: the faster the full length product appears in the extension reaction, the lower the recombination frequency due to the fewer number of template switching events (Zhao *et al*., 1998; Zhao, 2004).

The only reaction that could be construed as being possibly recombinogenic was the one containing the standard amount of reactants (lanes 8 and 9). For this reaction, the product is the least intense after 20 cycles (lane 8), which denotes conditions suitable for recombination to occur. Following the exact method from earlier publications need not necessarily work for other researchers, since there are many variables that need to be taken into account such as primer design, G-C content of the both the primers and DNA template, the annealing temperatures of the primers, thermal cycler etc. This anomaly is supported by a publication that focussed on the random mutagenesis and StEP recombination of a *B. thuringiensis* insecticidal crystal protein (Vanhercke *et al*., 2005). The researchers followed the original StEP method without any success, but obtained better results when they slightly modified their own standard method for use in the StEP reaction. Thus, in the current study, reactant concentrations identical to the standard amplification were used for the final StEP reaction between G41 and G53. The annealing time was reduced from 7 s to 5 s then

finally to 4 s (Prof. Huimin Zhao, personal communication). This was done to minimize the time spent in each StEP cycle i.e., prevent chain extension and proof-reading by *Taq* polymerase, thereby increasing the likelihood of recombination. The extension step is omitted during each stage of the initial StEP reaction since under ideal conditions, the polymerization rate for *Taq* polymerase is 750 - 1000 bp per second, which is sufficient for some extension to occur. Only partial extension is required for the recombinogenic template switching events to occur. An extension step may be added after the StEP reaction to increase the amount of available DNA for cloning, if insufficient full length product is obtained. This was, however, not required in the present study.

The StEP method aims at achieving higher recombination frequencies which is why the templates must be highly homologous. But, in doing so, fewer point mutations are introduced. Thus, the technology has a smaller probability of generating enzymes with totally new activities. Therefore, one StEP reaction was made mutagenic with the introduction of $Mn^{2+}$ and a higher concentration of dGTP, to increase the rate of mutation in the resultant recombinant library. The profiles of these reactions are seen in Fig. 3.4, with the final mutagenic StEP product in lane 5 less dense than the normal StEP reaction product in lane 3. This serves as an early indication that the clones generated using the products of the StEP reaction from lane 5 will most likely have a higher mutation rate and thus have a better chance at generating a more robust protein, since it explores more sequence space.

As much as StEP recombination has many advantages, it also has some drawbacks. The PCR conditions have to be strictly controlled in order to achieve staggered extension of primers by shortening or eliminating the polymerization time and lowering the reaction temperature. Failure to maintain a desirable range of temperatures during the PCR process may lead to non-specific annealing and formation of undesirable recombinants (Lee *et al*., 2003). StEP can also be more difficult for an inexperienced user to set up than conventional DNA shuffling, since full-length templates are included in the StEP reaction. This means that the production of full-length template may indicate simple amplification rather than recombination. Balancing the need for yield and recombination can be challenging, as it is not always straightforward to determine whether recombination has occurred, thus the

importance of analyzing the products obtained after 20 cycles. However, once optimized for a specific thermal cycler, primers and template, StEP can be easier to perform than DNA shuffling as fewer steps are involved (Neylon, 2004).

The StEP method has been used successfully to recombine gene variants created from random mutagenesis and naturally-occurring homologous genes that are approximately 80% identical. For example, it was used to increase the temperature optimum of subtilisin E by 18°C over that of the wild-type enzyme (Zhao and Arnold, 1999) and to broaden the substrate specificity of biphenyl dioxygenase (Bruhlmann and Chen, 1999). Other applications of StEP recombination include the alteration of the regioselectivity of a *Bacillus* α-galactosidase (Dion *et al*., 2001), thermostabilization of a cellulosomal endoglucanase (Murashima *et al*., 2002), the improvement of the enzyme activity of a fungal laccase (Bulter *et al*., 2003) and improvement of the thermal stability of *B. subtilis* xylanase (Miyazaki *et al*., 2006). From all research published regarding the use of the StEP method, the present study was the first that combined both ep-PCR and StEP recombination to improve proteins of interest.

Short-term thermostability screening of the 531 recombinants for 40 min at 80°C revealed that 19% of these recombinants had 70% of their activity after heat treatment, whilst 81% experienced a loss in their stabilities. The latter amount was expected, since most variants in a library will be unable to tolerate amino acid changes to their gross 3-D structure and will thus exhibit poor stability at extreme conditions. However, the fact that 19% of these recombinants displayed improvements to their stability is reasonably high. This implies that a proper plate screen for thermostability is not really required, since the success rate of mutation and recombination is almost twenty percent, and thus manageable. Overall, these recombinants had poor activities and this highlights the fact that thermal stability may be preferred over that of catalytic activity, even after recombination. Unfortunately, none of the recombinants displayed better hydrolysis on pH 10 xylan agar than parent G53 and there was no need to subject the recombinant to short-term screening at pH 10 as previously done for the xylanase mutants in chapter two. However, it seems unlikely that variants having greater than 93% stability at pH 10 would be easily obtained.

Twenty recombinant strains containing variant xylanases were grown and lysed and their xylanases tested for stability at 80°C for 90 min and the stability of 5 such recombinants, their parents and control XynA are shown in Fig. 3.5. Recombinants S325 and M32 were highly thermostable displaying 85% and 73% residual activity after heat treatment. They also retained 60% and 68% of their activity after exposure to pH 10, respectively (Fig. 3.6). These recombinants now displayed intermediate pH stability with regards to their parents G41 and G53. Recombinant S340 had intermediate thermal stability at 54% (Fig. 3.5) and average alkaline stability as well at 60% (Fig. 3.6). Mutagenic recombinant M28 was inferior to its thermostable parent G41 in that its thermal stability was only 27% (Fig. 3.5) and its alkaline stability was 55% (Fig. 3.6). On the other hand, M34 had similar alkaline stability at 43% (Fig. 3.6) but its thermal stability was 64% (Fig. 3.5), which is slightly lower than G41. All these recombinants and XynA were more thermostable than G53, but none compared to its high alkaline stability of 93%.

Often during shuffling experiments, shuffled recombinants display a combination of the traits of both parents. It was found that the thermal stability of the *S. lividans* xylanase was greatly increased when it was shuffled with the highly stable *T. fusca* xylanase. All recombinants had better stability than the former but less stability than the latter (Shibuya *et al*., 2000). Two F/10 xylanases from *T. alba* and *S. olivaceoviridis* were shuffled. Xylanase from the latter organism was stable up to 50°C whilst *T. alba* xylanase is stable up to 70°C. One chimera was stable up to 60°C whilst another retained stability up to only 45°C (Ahsan *et al*., 2001). Two xylanases from *N. frontalis* were shuffled and although one recombinant was shown to have the pH and temperature optima of the parent, it displayed much higher activity (Mesta *et al*., 2001). Fusion of *C. stercorarium* xylanase with cellulase-free xylanase from *B. halodurans* resulted in the constructed chimera displaying an increased affinity for oatspelts xylan and acid-swollen cellulose (Mangala *et al*., 2003). These different studies which focussed on shuffling of xylanases show that the recombinants obtained can exhibit a vast array of properties.

One of the most profound changes seen with the thermostable mutant G41 and the recombinants is the trade-off they exhibit in terms of enhanced thermal stability but diminished catalytic activity. The debate on this commonly-observed compromise

between the two properties has been debated by scientists for decades. With directed evolution experiments, it is often seen that catalytic activity at moderate and low temperatures decreases with increasing thermostability in naturally-occurring enzymes (Jaenicke, 1981; Jaenicke and Bohm, 1998) and possible reasons for this trend have been addressed in chapter two.

It is common for significant conformational changes to occur in enzymes during catalysis (Gerstein *et al*., 1994). The conformational changes occurring in xylanase from *T. reesei* were studied on the basis of crystal structure (Havukainen *et al*., 1996). Structural changes may lead to the opening and closing of the active site, and this probably plays a role in the functioning of the enzyme. Therefore, amino acid substitutions which enhance the thermostability of an enzyme may often change its molecular structure, causing a reduction in the enzyme.

The mutants and recombinant generated in this study exhibited some biochemical differences in comparison to wild-type XynA. It was therefore necessary to sequence the genes from these variants to determine the genetic mutations that occurred in order for them to exhibit these changes in biochemistry. The sequenced variants displayed amino acid sequence similarities ranging from 96% to 99%, highlighting that *xynA* is quite tolerant to mutation (Table 3.3). Sequence alignment (Fig. 3.7) revealed that the mutations were scattered throughout the protein but concentrated in the regions forming β-sheet B with more than 50% of the total number of substitutions occurring in sheet A, which forms the hydrophilic, solvent-accessible part of *xynA* (Gruber *et al*., 1998). A previous study of G/11 xylanases indicated that this long β-sheet is responsible for thermostability, because it stabilizes the overall xylanase structure (Hakulinen *et al*., 2003).

Alkaline-stable parent G53 displayed a single mutation, *viz*., A54T, which may have increased the polarity of XynA and allowed it to exhibit great stability at pH 10 (Fig. 3.6). The substitution occurred at the enzyme surface, between the second strand of sheet B and the second strand of sheet A. SDM of the *T. reesei* endoglucanase where threonine replaced asparagine resulted in the shifting of the pH optimum to the alkaline range (Wang *et al*., 2005). A clear conclusion to be drawn from other work is

that very large stability differences, in some cases, are due to one or very few point mutations (Hasegawa *et al*., 1999; Williams *et al*., 1999; Sandgren *et al*., 2003).

Conversely, thermostable parent G41 has a sum total of 4 mutations *viz*., K30E, W40R, T57A and K80R, which cumulatively conferred to the enzyme considerable thermal stability, albeit with poor catalytic activity (Fig. 3.5). What is interesting about these mutations is that 2 of them are arginine substitutions. The guanidinium group of arginine is the most polar of all the common amino acid side chains found in proteins (Wolfenden, 1983) and is thus the residue most likely to be found on the surface in an aqueous environment. It is often involved in the formation of single or double hydrogen bonded salt bridges, usually with aspartic or glutamic acid side chains. Such salt bridges stabilize the tertiary and quaternary structures of numerous proteins (Riordan *et al*., 1977). A sequence and crystal structure alignment study showed that highly thermostable xylanases have an increased number of charged residues, especially arginine, resulting in enhanced polar interactions (Hakulinen *et al*., 2003). W40R occurs in the first strand of sheet B, which is on the surface of the protein. Important improvements of stability by mutagenesis of a single solvent-exposed residue have been reported (Strub *et al*., 2004). Interactions of non-polar residues with water present a thermodynamic disadvantage caused by the side-chain being more exposed to solvent in the native than in the denatured state. In addition, polar residues at the surface may provide additional hydrogen bonds with the solvent and then increase protein stability.

The K80R substitution occurs in the third strand of the B sheet and is in a shallow location of *xynA*. Lysine → arginine mutations in a number of enzymes lead to enhanced thermal stability. Arginine may be better able to interact with backbone carbonyl oxygens, providing added structural rigidity to the protein (Mrabet *et al*., 1992). Shallowly-buried arginines are likely to influence the overall electrostatic potential of the protein surface. Several studies indicate that there is a correlation between protein stability and the number of arginines on the protein surface. The comparison of mesophilic proteins and their thermophilic counterparts has revealed that thermophilic proteins have, on average, a higher arginine content on the protein surface (Argos *et al*., 1979; Vogt *et al*., 1997). At the same time, the amount of lysine was observed to decrease, thus increasing the arginine:lysine ratio. This result is

supported by experimental mutagenesis studies in which lysine → arginine mutations stabilized the protein structure (Cunningham and Wells, 1987; Mrabet *et al*., 1992). It has been proposed that an important reason for the increased stability is stronger hydrogen bonding of the large guanidinium group of arginine with nearby polar groups (Mrabet *et al*., 1992; Borders *et al*., 1994). Involvement in the ion pair networks is another possible factor responsible for the increased thermostability conferred by arginine (Mrabet *et al*., 1992). A similar effect was found with the SDM of *T. reesei* xylanase II when lysine → arginine mutations increased thermotolerance (Turunen *et al*., 2002). A change in polarity might also contribute to the thermostabilization of these variant xylanases by rearrangement of interactions such as those with hydrogen bonds and/or salt bridges (Murashima *et al*., 2002).

Genetic analysis of the recombinants proved that StEP recombination occurred since the recombinants displayed the amino acid substitutions from both parents G41 and G53. StEP recombinant S340 had three mutations (W40R, T57A, K80R) and retained 54% and 60% residual activity respectively at 80°C and pH 10. However, its catalytic activity of 31.16 nkat/ml is slightly higher than parent G41. Thus, the K30E substitution somehow impeded enzyme activity.

Recombinant S325 inherited four mutations from thermostable parent G41 and the single mutation from parent G53. S325 is the only sequenced recombinant to attain the A54T mutation from parent G53. This additional threonine allowed it to exhibit a higher stability at 80°C than G41. However, the four thermostable mutations reduced the alkaline pH stability to 60%.

The mutagenic StEP recombinants, M28, M32 and M34, had additional mutations due to the inclusion of mutagen in their recombination reactions. M28 has a very significant substitution - C185Y. This is the position of the second cysteine of the disulphide bridge, which is also located within the α-helix of the protein. This mutation and the resultant loss of the disulphide bridge could explain why this recombinant displayed poor stabilities of 55% and 27% at 80°C and pH 10 in comparison to its parents. Disulphide bridges located within the α-helices of proteins are known to contribute significantly to thermal stability (Daggett and Levitt, 1993; Wakarchuk *et al*., 1994a, 1994b; Turunen *et al*., 2001). The stabilizing effect of a

disulphide bridge is usually assumed to arise from its destabilizing effect on the unfolded state, by decreasing its conformational entropy (Creighton, 1988; Matsumura *et al*., 1989; Clarke *et al*., 2000). This is because the presence of such a covalent cross-link limits the motional freedom of the protein backbone in the unfolded state, thereby decreasing the entropy (Davoodi *et al*., 2007). But it cannot be the only means of stabilization since other mesophilic G/11 xylanases also have disulphide bridges and other xylanases displaying greater thermostability do not possess these structures.

The Q204R mutation in both M32 and M34 increased the arginine content of these variants and both had eight substitutions each. M32 retained 73% and 68% of its activity at 80°C and pH 10. Its alkaline stability was better than G41 but not G53. In M32, the D128G substitution occurred in the region between B sheet strand 6b and strand 9. This region seems to be of importance for stability especially at alkaline conditions in G/11 xylanases. The effect is probably due to a change in the ionization state at higher pH values, leading to a pH-induced stabilization of the enzyme (Fenel *et al*., 2006). Glycine has a very small R-group which facilitates better packing of this region, which in turn contributes to enhanced thermal stability through a reduction in the volume of the hydrophobic core of the protein (Hakulinen *et al*., 2003).

Mutations in evolved stable enzyme variants mostly concern residues located on the surfaces of the protein, while mutations in secondary structural elements are relatively rare. On the one hand, it is not surprising that mutations accumulate at the surface, because such mutations have a much greater chance of being structurally acceptable than mutations in the protein interior. In addition, the larger part of an enzyme surface will not be directly involved in catalytic activity, meaning that the chances of surface mutations affecting activity are small. However, a more recent report indicates that co-operative networks of surface or near-surface interactions are important for stability (Eijsink *et al*., 2005).

Only one mutation, F212L, was seen two residues away from the acid-base partner of the active site in variant M34. Mutations directly in the active site might not be the best ones. These mutations make large changes that can disrupt catalysis or drastically alter the shape of the binding pocket. Mutations just outside the active site that would

more subtly alter the substrate-binding site might be more effective (Morley and Kazlauskas, 2005).

Most random mutagenesis methods, such as ep-PCR or DNA shuffling, seek to mutate all the amino acids equally. These methods might contain bias for particular types of mutations (transitions over transversions) in ep-PCR (Fromant *et al*., 1995), or particular locations (regions of high sequence similarity) in DNA shuffling (Ness *et al*., 2002), but their goal is to mutate each amino acid or region with equal probability. Thus, mutation of the entire protein creates a greater number of distant mutations compared with mutations close to the active site. This occurs because there are more amino acids far way from the active site than near the active site with only a few amino acids (<10%) forming the substrate-binding site, and this explains why directed evolution experiments create distant mutations (Morley and Kazlauskas, 2005).

The ultimate goal of this project was to create a thermostable, alkaline xylanase suitable for pre-bleaching of paper pulp. Recombinants displaying both these properties were created using the StEP method and they were sequenced to determine their amino acid substitutions that afforded them these traits. S325 and S340 were true recombinants since they displayed properties of both parents G41 and G53 whilst M28, M32 and M34 were mutagenic recombinants that exhibited an array of genetic mutations due to the addition of $Mn^{2+}$. Only S340 and S325 were investigated further. Before application, these recombinant and parent xylanases needed to be characterized to determine the biochemical conditions under which they operate at their most optimal, which was the aim of the subsequent chapter.

# CHAPTER 4: BIOCHEMICAL CHARACTERIZATION OF XynA AND VARIANTS

## 4.1 INTRODUCTION

Modern experimental technologies are providing seemingly endless opportunities to generate massive amounts of genomic sequence, expression and functional data. Continuous advances and improvements have enabled proteome analysis to proceed with increased depth and efficiency. To capitalize on this enormous pool of information and in order to understand fundamental biological phenomena, it has become increasingly apparent that functional data for the gene products, in particular for enzymes, has either limited accessibility or is unavailable. Additionally, although enzyme structural information has been rapidly accumulated in databases, little effort has been invested towards systematic characterization of enzyme functions primarily because the exercise is rather costly and also the data itself is often dependent on the experimental conditions which vary across the world. Also, there exists a lack of substrates that are suitable for the demonstration of such functions.

In industry there is an increasing demand for biocatalysts because they can catalyze reactions with very high rates and have high reaction specificity and stereoselectivity. But prior to any such commercial application, adequate knowledge about the conditions that contribute to the optimal functioning of an enzyme is required, since catalysis is influenced by reaction conditions, particularly temperature and pH, which can have irreversible or reversible effects on the enzymes (Garrett and Grisham, 2005).

Enzyme kinetics is the investigation of enzyme performance by measuring how the initial rate of reactions is affected by varying experimental conditions such as temperature, pH and even substrate concentration (Brownie and Kernohan, 1999). Kinetic parameters and models are used for the development of bioprocesses that ensure maximum output with minimal economic expenditure and hazards. In theory, the higher the temperature, the faster a given chemical reaction will proceed. Quantitatively, the relationship between the rate a reaction proceeds (k) and its temperature is determined by the Arrhenius equation. At higher temperatures, the

probability that two molecules will collide is higher. This higher collision rate results in a higher kinetic energy, which has an effect on the activation energy of the reaction ($E_a$). The activation energy is the amount of energy required to ensure that a reaction occurs (Jencks, 1969; Price and Dwek, 1979; Bailey and Ollis, 1986).

Experimental studies have shown that introduction of specific amino acid changes through mutagenesis produces striking effects on the rates of the enzyme-catalyzed reactions even though the changes are far from the active site. These biocatalysts are highly flexible proteins capable of giving rise to conformations that facilitate the chemical transformation of the substrate by favouring orientations and conformations that provide a framework for optimal catalytic activity. The conservation of key amino acids throughout an enzyme from many species over the course of evolution hints at the operation of such a network (Benkovic, 2005).

For application of xylanases in pulping, xylanases must be cellulase-free, thermostable and alkalophilic. Due to the extremes of the pulp bleaching process, it is desirable to use enzymes that are stable under alkaline conditions (pH 8 - 10) and at high temperatures (60°C - 90°C) for incubation periods up to 3 – 5 h. The use of xylanases for delignification has been slowed down by the lack of large-scale availability of enzymes active at pHs above 8 and temperatures of around 60°C which are prevalent conditions in many bleaching processes (Techapun *et al*., 2003). Most commercial xylanases have not met the requirements of the pulp bleaching industry; pH adjustment between pH 5 and 8 and cooling of the pulp at temperatures between 40°C and 60°C has, so far, been necessary. An example of such a commercial xylanase is Luminase, produced by Diversa Corporation (USA), which was discovered in a soil sample from a terrestrial alkaline thermal vent in the Geyser Valley in the Kamchatka region of Russia and further subjected to SDM. It is capable of opening up wood fibres for bleaching and is recommended for use from 40°C - 70°C at pH 6 – 8 and at the time, was the latest commercial pulping xylanase released on the market.

The objectives of this chapter were: (1) to determine the pH and temperature for optimal activity of xylanase variants G41, G53, S325 and S345 against the cloned and fungal XynA and commercial enzyme Luminase; (2) to use these optima to determine

the long-term functioning of the variant xylanases at commercial pulping conditions; (3) to determine if mutation had altered their stabilities at acidic pH and; (4) to use first-order kinetics to determine the stability of the xylanases.

## 4.2 MATERIALS AND METHODS

### 4.2.1 Growth of cultures and enzyme extraction

The *E. coli* strains producing cloned XynA and variant xylanases G41, G53, S340 and S325 were grown in LB media containing ampicillin and their respective xylanases were extracted as described in section 2.2.4.

A modified protocol from Purkarthofer *et al*. (1993a) was used for the growth of fungal *T. lanuginosus* strain DSM 5826 producing XynA, in duplicate 500 ml Erlenmeyer flasks each containing 100 ml of growth medium. The composition of the growth medium comprised of 3.12% beechwood xylan (Sigma); 3.02% yeast extract (Oxoid); and 0.5% $KH_2PO_4$ dissolved in distilled water. The pH of the medium was adjusted to 6.5 with 2 M NaOH and then autoclaved. After the flasks were cooled, each flask containing the sterile medium was aseptically inoculated with a 2 $cm^3$ agar block cut from a fresh potato dextrose agar plate culture of the fungus. After 5 days of incubation on a shaker at 50°C at 150 rpm, the culture medium was centrifuged at 11 000 *g* for 10 min and the supernatant was filtered using Whatman no. 1 filter paper. The filtrate contained the crude fungal xylanase and was stored at 4°C until further analysis.

The commercial pulping xylanase, Luminase, used in this study was kindly donated by Dr Jay Short and Mr Jack Turner from Diversa Corporation (San Diego, USA). The commercial preparation is sold as a crude fermentation broth that is concentrated and preserved in glycerol. The enzyme was stored at 4°C.

### 4.2.2 Determination of pH and temperature optima

Crude xylanases were diluted in appropriate buffers over the pH range of 2 – 12 to determine pH optima of the respective xylanases. The buffer systems (0.05 M) used were citrate-NaOH (pH 2 – 6.5), Tris-HCl (pH 7 – 9) and glycine-NaOH (pH 10 – 12). Substrate (0.1% birchwood xylan, Roth) was prepared in buffers of the corresponding pH and assayed for xylanase activity as described in section 2.2.5.1.

Enzyme filtrates were diluted in their optimal pH buffers (pH 5 – 6.5) and assayed for xylanase activity from 40°C – 100°C for 5 min as described in section 2.2.5.1 for determination of the temperature for optimal activity for the xylanases.

### 4.2.3 pH and temperature stability

Xylanases were tested for stability in the alkaline pH range (pH 8 – 10) using 0.05 M Tris-HCl (pH 8 and 9) and 0.05 M glycine-NaOH (pH 10) from 60°C - 90°C for up to 180 min. Samples were removed every 15 min and treated as described in section 2.2.5.1.

Enzyme filtrates were diluted in 0.05 M citrate buffer (pH 3 - 5) and incubated at 60°C for 90 min to determine stability at acidic pH. Samples were taken every 15 min and treated as described in section 2.2.5.1.

### 4.2.4 First-order kinetics of characterized xylanases

The first-order constants of inactivation for each of the xylanases were determined by constructing Arrhenius plots. The constants were calculated from the semilogarithmic plots of the percentage of remaining xylanase activity over different time intervals at 60, 70, 80 and 90°C from pH 8 – 10. The Arrhenius equation (1) was used to calculate the activation energy of the reaction, where k is the reaction rate constant (first-order), A is the pre-exponential factor, R is the universal gas constant ($8.314 \times 10^{-3}$ kJ $mol^{-1}$ $K^{-1}$), T is the absolute temperature in Kelvin and $E_a$ is the activation energy for the reaction (Price and Dwek, 1979).

$$k = Ae^{(Ea/RT)} \qquad (1)$$

The half-lives ($t_{50}$) of the enzymes were calculated using equation 2, to determine the time frame in which the enzymes retained exactly 50% of their initial activities under the different conditions tested. For first-order reactions, the $t_{50}$ is independent of the enzyme concentration. If half of the amount of enzyme becomes inactivated, the ratio inside the log is always equal to ln 2, no matter what the initial concentration. The time needed for half of the reagent to react is called the half-life of the enzyme and is derived from the following equation:

$$\ln 2 = kt_{50}$$

$$t_{50} = \frac{\ln 2}{k}$$

$$t_{50} = \frac{0.693}{k} \qquad (2)$$

## 4.3 RESULTS

### 4.3.1 pH and temperature optima

The xylanases selected for further investigation displayed optima in the pH range of 5 – 6.5 (Fig. 4.1). The Luminase, fungal and cloned XynA enzymes and G53, exhibited optima of pH 6.5. G41 and S325 had optima of pH 6 whilst S340 had a pH optimum of 5. The xylanase variants G41, S325 and S340, also retained more than 60% and almost 40% of their initial activities at pH 2 and 12 respectively. These three enzymes had different profiles compared to the rest and also had a broader range of activity.

Luminase had a broad temperature optimum, ranging from 60°C - 80°C and retained almost 100% activity in this temperature range. However, its activity dropped sharply at temperatures above 80°C. Both the XynA controls and G53 exhibited optimal activity from 70°C - 78°C with an overall optimum of 70°C (Fig. 4.2). G41 and S325 had optima of 74°C whilst S340 had an optimum of 72°C.

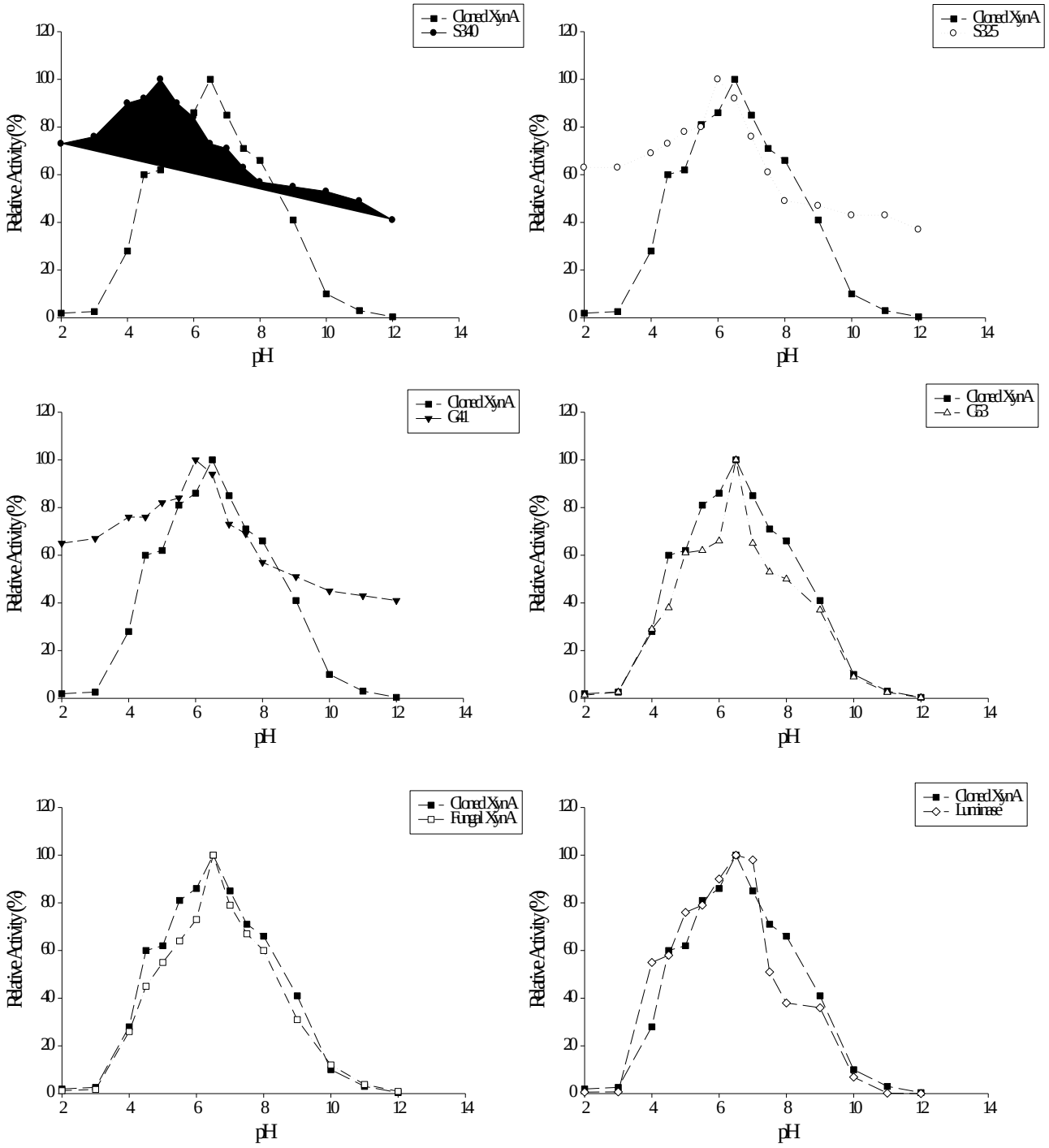


**Fig. 4.1** The effect of pH on activity of XynA variants, fungal XynA and Luminase against cloned XynA from pH 2 – 12 at 50°C for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

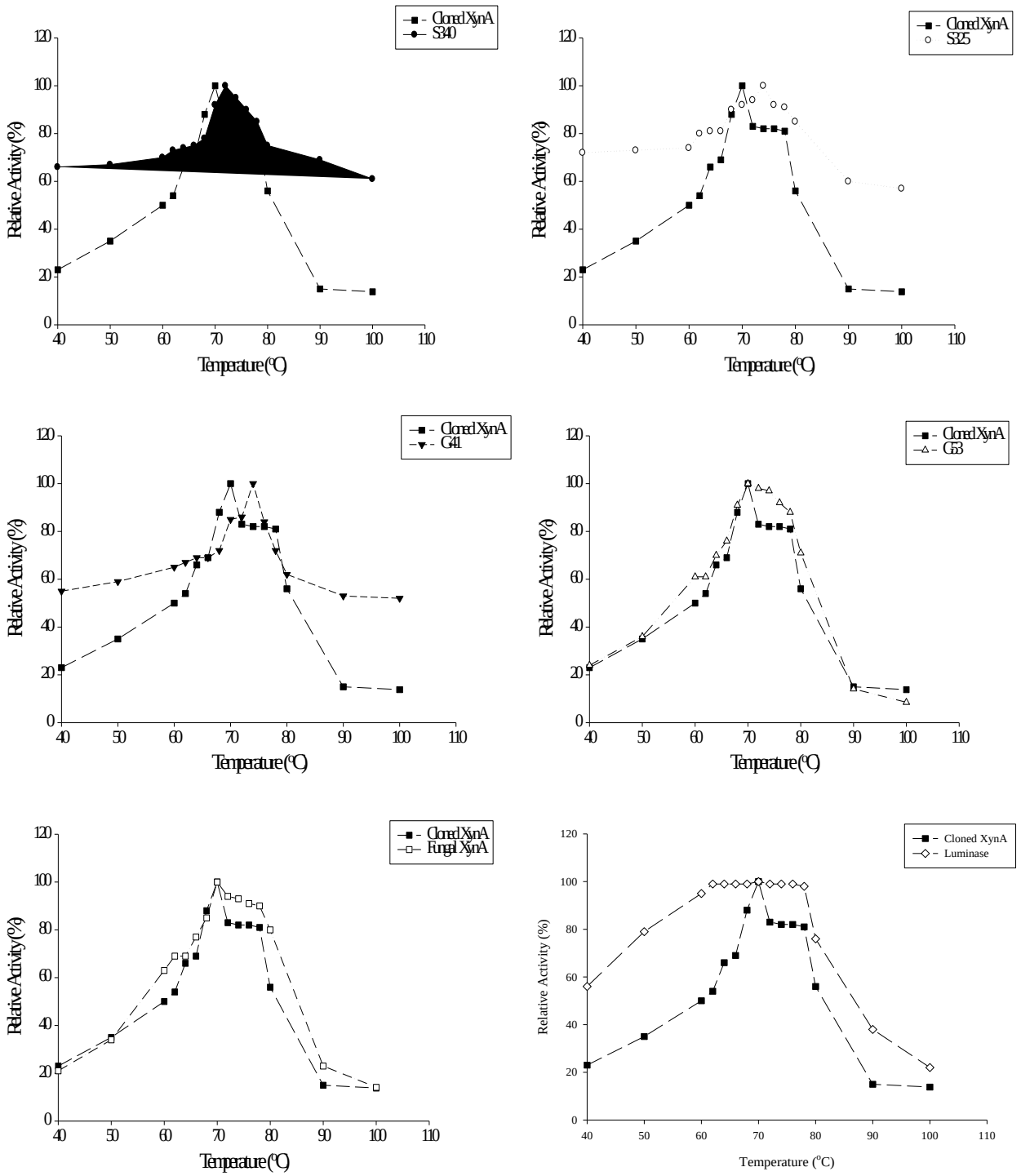


**Fig. 4.2** The effect of temperature on the activity of XynA variants, fungal XynA and Luminase against cloned XynA from 40°C - 100°C at their optimum pH for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

### 4.3.2 pH and temperature stability

Luminase was very stable at 60°C at pH 8 and 9, with no appreciable loss in activity after 180 min incubation (Fig. 4.3A and Fig. 4.4A). At pH 10, the enzyme was stable for approximately 80 min at 60°C and thereafter steadily lost stability (Fig. 4.5A). The loss in the stability of Luminase was great at temperatures higher than 60°C from pH 8 – 10 (Figs. 4.3 – Fig. 4.5).

The cloned and fungal XynA controls retained almost 50% activity when incubated at 60°C and pH 8 (Fig. 4.3A). This stability steadily dropped at pH 9 and both enzymes displayed approximately 40% residual activity (Fig. 4.4A) and this further decreased at pH 10 (Fig. 4.5A). At 70°C, their stabilities dropped to less than 15% at pH 8 and 9 (Figs. 4.3B and 4.4B) with less than 5% at pH 10 (Fig. 4.5B). At 80°C and 90°C at pH 8 (Figs. 4.3C and 4.3D), pH 9 (Figs. 4.4C and 4.4D) and pH 10 (Figs. 4.5C and 4.5D), these XynA controls have almost no activity after 180 min of incubation.

G53 had good stability when tested at 60°C retaining 82%, 81% and 76% of its activity at pH 8, 9 and 10 respectively (Figs. 4.3A, 4.4A and 4.5A). This stability at alkaline pH is severely compromised when the temperature is increased and G53 displayed less than 20% activity at 70°C from pH 8 – 10 (Figs. 4.3B, 4.4B and 4.5B), less than 3% activity at 80°C from pH 8 – 10 (Figs. 4.3C, 4.4C and 4.5C) and no activity at 90°C at the pH range tested (Figs. 4.3D, 4.4D and 4.5D).

At 60°C and pH 8, S340, S325 and G41 had 74%, 71% and 70% residual activity, respectively (Fig. 4.3A). At pH 9, this stability decreased to 61%, 64% and 70% (Fig. 4.4A) and further decreased to 59%, 60% and 66% at pH 10 (Fig. 4.5A). The remaining activities continued to decrease for these mutants at 70°C and 80°C from pH 8 – 10 (Figs. 4.3 - 4.5). What remained impressive was that they retained 50% activity at 90°C, pH 8 and 9 (Figs. 4.3D and 4.4D). At 90°C and pH 10, S325, S340 and G41 had 37%, 45% and 43% activity, respectively, making them the most stable xylanases in this study.

At pH 3, S340, S325 and G41 retained 51%, 78% and 87% residual activity after 45 min whilst Luminase had 40% activity (Fig. 4.6A). The other 3 enzymes had less than

25% activity at pH 3. At pH 4, all the xylanases displayed greater stability. S340 retained almost 100% activity for about 30 min and thereafter its stability decreased to less than 70% (Fig. 4.6B). At pH 5, all the xylanases tested retained more than 70% activity. Luminase was marginally better than the other xylanases (Fig. 4.6C).

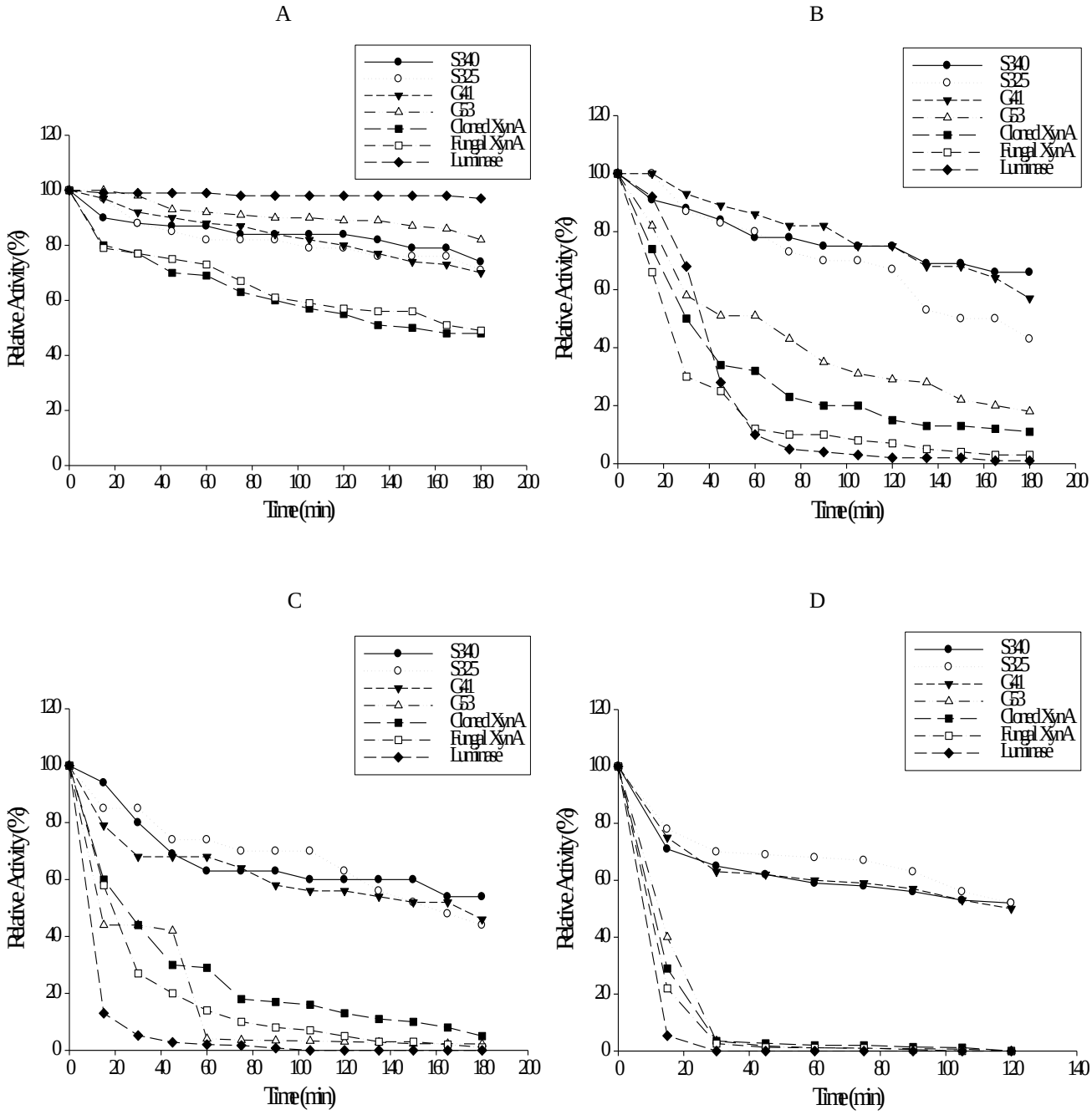


**Fig. 4.3** The effect of temperature and alkalinity on the stability of wild-type XynA, its mutant progeny and recombinants at (A) 60°C, (B) 70°C, (C) 80°C and (D) 90°C at pH 8. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

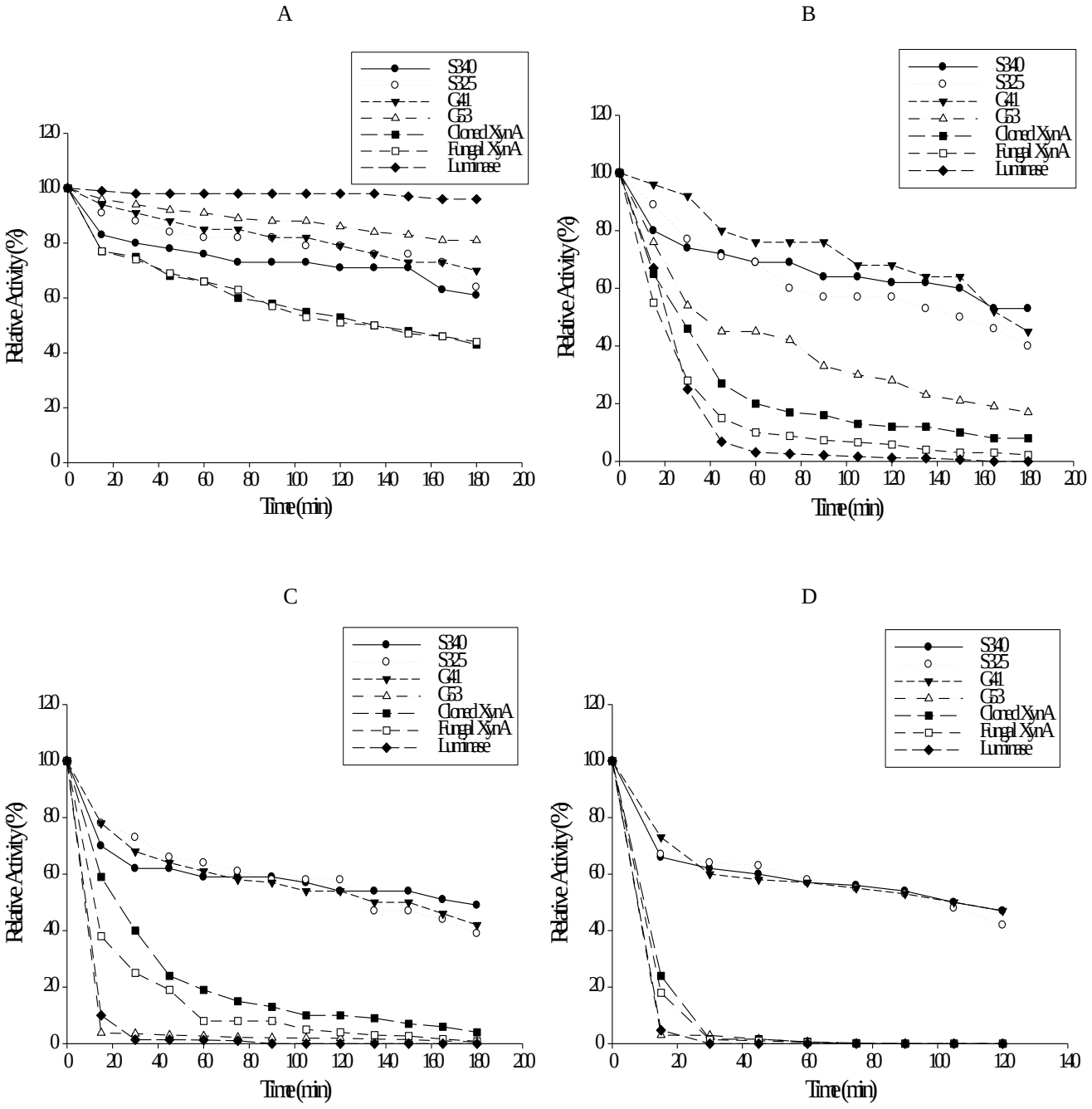


**Fig. 4.4** The effect of temperature and alkalinity on the stability of wild-type XynA, its mutant progeny and recombinants at (A) 60°C, (B) 70°C, (C) 80°C and (D) 90°C at pH 9. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

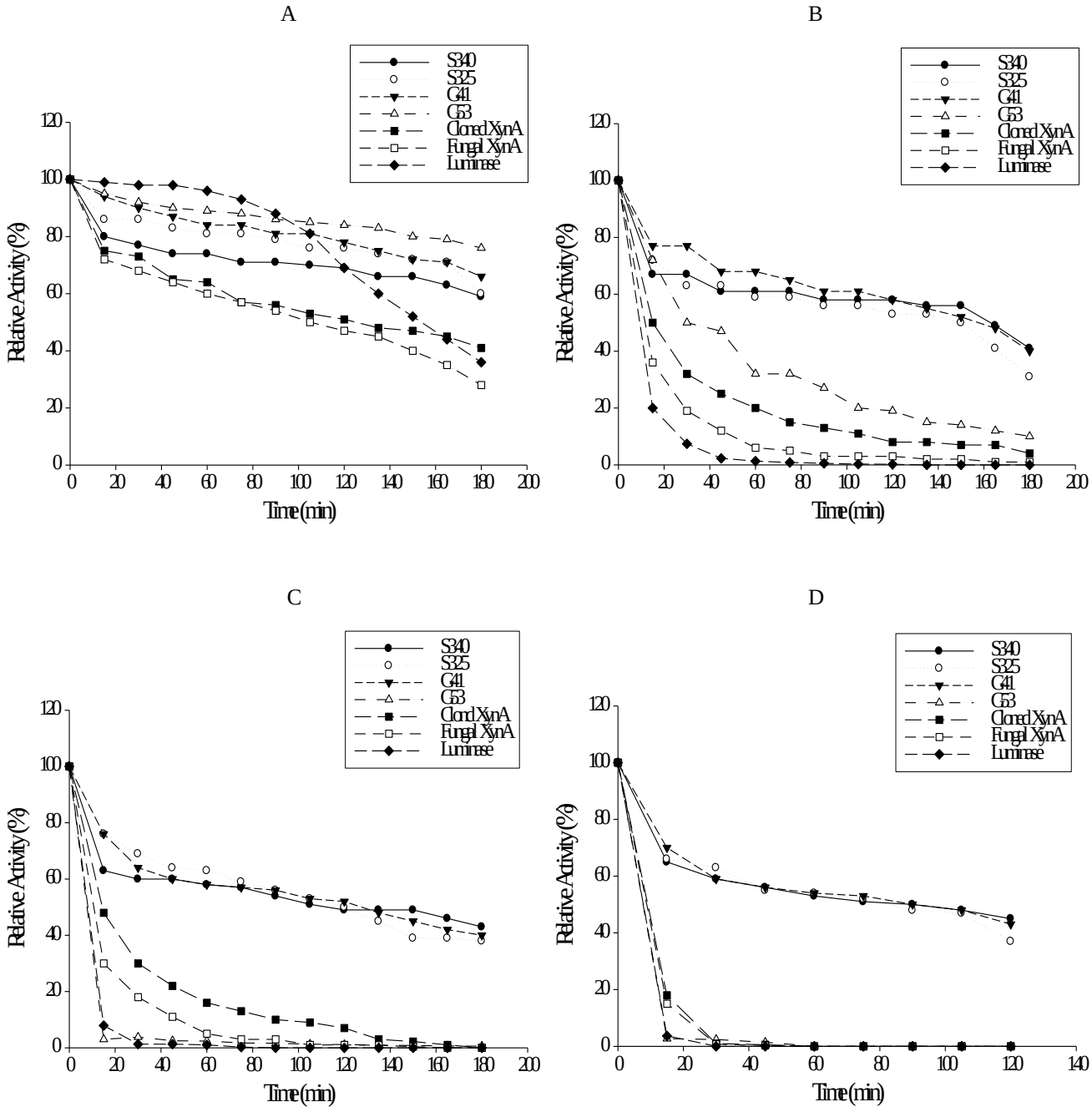


**Fig. 4.5** The effect of temperature and alkalinity on the stability of wild-type XynA, its mutant progeny and recombinants at (A) 60°C, (B) 70°C, (C) 80°C, and (D) 90°C at pH 10. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

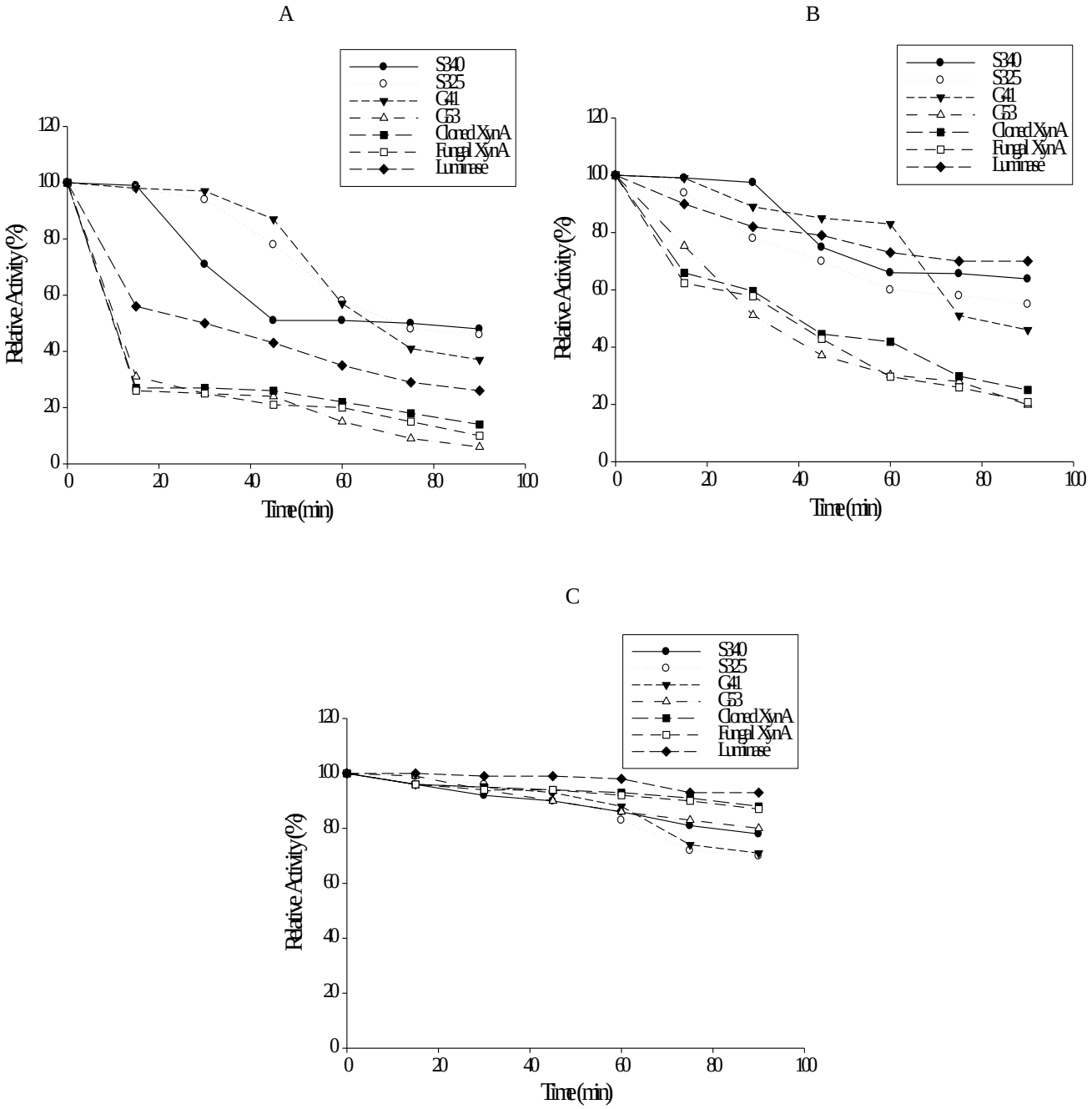


**Fig. 4.6** The effect of acidity on the stability of wild-type XynA, its mutant progeny and recombinants at (A) pH 3, (B) pH 4 and (C) pH 5 at 60°C. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

.

### 4.3.3 First-order kinetics of characterized xylanases

Using the Arrhenius equation (section 4.2.4), it was possible to determine the k, $t_{50}$ and $E_a$ values for all characterized xylanases at 60 - 90°C from pH 8 – 10 (Tables 4.1 - 4.3). Under all conditions tested, the cloned and fungal XynA had the shortest $t_{50}$ tested. Luminase displayed remarkable $t_{50}$ of 385 and 346.50 min at pH 8 and 9, respectively. However, this was substantially diminished as the treatment conditions became progressively more alkaline and thermophilic. On the other hand, its activation energies were very high under all conditions tested. Both the cloned and fungal XynA displayed almost identical kinetic profiles under all conditions tested. G53 however, had very good $t_{50}$ of 46.20, 40.76 and 36.47 min respectively from pH 8 – 10 at 60°C in comparison with the other variant xylanases. Overall, however, G41 and S340 had the best $t_{50}$ and the lowest $E_a$ values at temperatures greater than 60°C. Conversely, Luminase had high reaction rates under the parameters tested from pH 8 – 9.

**Table 4.1** Estimated $t_{50}$ in min, inactivation rates (k) and activation energies ($E_a$) of XynA, Luminase and xylanase variants at pH 8 for 180 min

| Xylanases | Temperature (°C) | | | | | | | | $E_a$ (kJ/mol) |
|---|---|---|---|---|---|---|---|---|---|
| | 60°C | | 70°C | | 80°C | | 90°C | | |
| | k | $t_{50}$ | k | $t_{50}$ | k | $t_{50}$ | k | $t_{50}$ | |
| S340 | 0.017 | 40.76 | 0.032 | 21.66 | 0.045 | 15.40 | 0.064 | 10.83 | 26.07 |
| S325 | 0.024 | 28.88 | 0.060 | 11.55 | 0.064 | 10.83 | 0.069 | 10.04 | 29.24 |
| G41 | 0.027 | 25.67 | 0.044 | 15.75 | 0.050 | 13.86 | 0.068 | 10.19 | 27.53 |
| G53 | 0.015 | 46.20 | 0.134 | 5.17 | 0.319 | 2.17 | 0.392 | 1.77 | 94.38 |
| Cloned XynA | 0.056 | 12.38 | 0.177 | 3.92 | 0.213 | 3.25 | 0.452 | 1.53 | 65.48 |
| Fungal XynA | 0.051 | 13.59 | 0.278 | 2.49 | 0.328 | 2.11 | 0.442 | 1.57 | 52.19 |
| Luminase | 0.0018 | 385 | 0.217 | 3.19 | 0.295 | 2.35 | 0.372 | 1.86 | 142.84 |

**Table 4.2** Estimated $t_{50}$ in min, inactivation rates (k) and activation energies ($E_a$) of XynA, Luminase and xylanase variants at pH 9 for 180 min

| Xylanases | Temperature (°C) | | | | | | | | $E_a$ (kJ/mol) |
|---|---|---|---|---|---|---|---|---|---|
| | 60°C | | 70°C | | 80°C | | 90°C | | |
| | k | $t_{50}$ | k | $t_{50}$ | k | $t_{50}$ | k | $t_{50}$ | |
| S340 | 0.029 | 23.90 | 0.039 | 17.77 | 0.046 | 15.07 | 0.070 | 9.90 | 28.17 |
| S325 | 0.026 | 26.65 | 0.063 | 11.00 | 0.065 | 10.66 | 0.083 | 8.35 | 34.86 |
| G41 | 0.028 | 24.75 | 0.056 | 12.38 | 0.057 | 12.16 | 0.074 | 9.36 | 30.71 |
| G53 | 0.017 | 40.76 | 0.135 | 5.13 | 0.223 | 3.11 | 0.405 | 1.71 | 102.25 |
| Cloned XynA | 0.060 | 11.55 | 0.195 | 3.55 | 0.233 | 2.97 | 0.489 | 1.42 | 65.70 |
| Fungal XynA | 0.060 | 11.55 | 0.284 | 2.44 | 0.337 | 2.06 | 0.463 | 1.50 | 63.57 |
| Luminase | 0.002 | 346.50 | 0.231 | 3.00 | 0.381 | 1.82 | 0.385 | 1.80 | 151.92 |

**Table 4.3** Estimated $t_{50}$ in min, inactivation rates (k) and activation energies ($E_a$) of XynA, Luminase and xylanase variants at pH 10 for 120 min

| Xylanases | Temperature (°C) | | | | | | | | $E_a$ (kJ/mol) |
|---|---|---|---|---|---|---|---|---|---|
| | 60°C | | 70°C | | 80°C | | 90°C | | |
| | k | $t_{50}$ | k | $t_{50}$ | k | $t_{50}$ | k | $t_{50}$ | |
| S340 | 0.030 | 23.10 | 0.044 | 15.75 | 0.047 | 14.74 | 0.076 | 9.12 | 43.21 |
| S325 | 0.029 | 23.90 | 0.064 | 10.83 | 0.072 | 9.63 | 0.093 | 7.45 | 36.70 |
| G41 | 0.030 | 23.10 | 0.057 | 12.16 | 0.059 | 11.75 | 0.082 | 8.45 | 30.98 |
| G53 | 0.019 | 36.47 | 0.180 | 3.85 | 0.237 | 2.92 | 0.653 | 1.06 | 124.34 |
| Cloned XynA | 0.061 | 11.36 | 0.224 | 3.09 | 0.359 | 1.93 | 0.502 | 1.38 | 65.48 |
| Fungal XynA | 0.083 | 8.35 | 0.340 | 2.04 | 0.372 | 1.86 | 0.497 | 1.39 | 71.56 |
| Luminase | 0.082 | 8.45 | 0.286 | 2.42 | 0.319 | 2.17 | 0.418 | 1.66 | 160.68 |

## 4.4 DISCUSSION

In directed evolution studies, once enzyme variants display better stabilities than their parent counterparts, it makes them attractive for commercial application. It is important to not only sequence them to determine the amino acid substitutions that foster this stability, but it is also vital to fully characterize them on a biochemical level. This biochemical information is crucial since it will not only complement the genetic data, but it will also provide useful information that will allow bioprocess engineers to tailor the industrial process for future commercial applications.

Enzyme-substrate recognition and the catalytic events that ensue are greatly dependent on pH. An enzyme possesses an array of ionisable side chains and prosthetic groups that not only determine its secondary and tertiary structure, but may also be intimately involved in its active site. The charges on these groups will vary, according to their acid dissociation constants, with the pH of the environment. This will affect the total net charge of the enzymes and the distribution of charge on their exterior surfaces, in addition to the reactivity of the catalytically-active groups. Furthermore, the substrate itself often has ionizing groups, and one or another of the ionic forms may preferentially interact with the enzyme (Garrett and Grisham, 2005). When the medium of an enzyme reaction is made more acidic, the hydrogen ion concentration of the solution increases. This increase results in a congruent increase in competition for any metal cationic binding sites on the enzyme effectively reducing the bound metal cation concentration. Increasing the alkalinity of the reaction solution increases the hydroxyl ion concentration which competes against the enzymes' ligands for divalent and trivalent cations, causing their conversion to hydroxides and their complete removal from the enzyme (Chaplin, 2004).

Enzymes are usually active as catalysts within a narrow range of pH, with exceptions, and usually the enzyme is most active at a specific pH called the optimum pH. This was determined for Luminase, cloned and fungal XynA, G53, G41, S325 and S340 (Fig. 4.1) and all displayed typical bell-shaped pH profiles when tested in the range of pH 2 – 12 at 50°C. Luminase, cloned and fungal XynA and G53 all exhibited optimal activity at pH 6.5 whilst G41 and S325 showed pH optima at pH 6. S325 showed a shift of 2.5 pH units to the acidic range with an optimum of pH 5. Thus, the amino

acid substitutions described for these mutants in chapter 3 had a definite effect on the pH profile of these enzymes. Bell-shaped pH profiles are a phenomenon typical for most enzymes. It seems that the enzyme is unstable at pH values away from the optimum and at these pH values the enzyme starts to lose the tertiary structure necessary for conformation of the active site. Another plausible explanation, in the majority of cases, is that the effects of pH on the ionization of acidic and basic groups at the active centre of the enzyme are being observed (Wynn, 1979).

The variation of activity with pH, within a range of 2 – 3 units each side of the optimal pH is a reversible process. Extremes of pH will, however, cause a time- and temperature-dependent, essentially irreversible denaturation. In alkaline solutions, there may be a partial destruction of cysteine residues due to base catalyzed β-elimination reactions, whereas, in acid solutions, hydrolysis of the labile peptide bonds, sometimes found next to aspartic acid residues, may occur (Wynn, 1979; Hulbert and Preston, 2001; Chaplin, 2004).

Luminase had a temperature optimum of 70°C but exhibited minor changes in activity when exposed to temperatures of 62 - 78°C, which resulted in a very broad temperature optimum, as seen with the plateau formed in this range (Fig. 4.2). The cloned and fungal XynA and G53 had optima of 70°C whilst S340 had an optimum of 72°C. G41 and S325 exhibited optimal activity at 74°C. Clearly, genetic mutation has resulted in these variants shifting to the right towards the thermostable range.

Although these xylanase variants exhibited properties more stable than either XynA or the Luminase controls in this study, it was important to evaluate their capabilities under simulated pulping conditions that were both thermophilic and alkaline from 60°C - 90°C at pH 8 – 10. These parameters were combined and carried out at 60°C, first at pH 8, then pH 9 and finally pH 10. The temperature was raised by 10°C increments till 90°C, and these results are reflected in Figs. 4.3 – 4.5.

From these graphs, it is evident that Luminase is a thermophilic enzyme and exhibited remarkable stability at 60°C at pH 8 and 9 and retained almost all of its activity after 180 min of incubation (Figs. 4.3A and 4.4A). Even at pH 10, it still had almost 70% of its initial activity after 120 min of treatment (Fig. 4.5A). However, as soon as the

enzyme was exposed to higher temperatures of 70°C - 90°C, it was rapidly denatured, in most instances, within a mere 20 min of incubation (Figs. 4.3 – 4.5). The enzyme is relatively new in the pulping market and a database search did not reveal any publications where pulping data on the xylanase's capability is available. Presumably, it has been applied purely on a commercial level. The enzyme's selling point however, seems not to be in its stability at 60°C from pH 8 – 9, but rather the fact that it is produced in a cheap fermentation broth, heavily concentrated and preserved in glycerol. Since the heat denaturation of enzymes is primarily due to the proteins' interactions with the aqueous environment, enzymes are more stable in concentrated rather than dilute solutions. Thus, even if the Luminase preparation is exposed to high temperatures, it would effectively be able to perform bleaching of paper pulp because it is so highly concentrated. Luminase II was released for commercial use in early 2006. It still retains the alkaline characteristics of its predecessor but, it has been engineered using SDM to be more thermophilic than the initial version.

The cloned and fungal XynA controls exhibited similar profiles, with a few inconsistencies. Since it was not possible to purify all the tested xylanases to homogeneity, unknown factors in the crude enzyme preparations such as ingredients of culture media and *E. coli* proteins or cell metabolites may have interacted with the xylanase and in doing so, affected enzyme stability as previously noted (Kim *et al*., 2002). Although these control xylanases are reported to have good stability at 20°C, pH 8 (Gomes *et al*., 1993) and 70°C, pH 6.5 (Singh *et al*., 2000b), in this study it was observed that this stability was lost when the pH of the surrounding medium was made more alkaline and the temperature was concurrently increased. At 60°C (pH 8) they retained 50% of activity after 180 min of incubation (Fig. 4.3A), and at pH 9 they retained 45% activity (Fig. 4.4A). Their stability steadily dropped as the conditions were made more thermophilic (70°C - 90°C) and alkaline (pH 10).

G53 was quite stable at 60°C from pH 8 – 10 (Figs. 4.3A, 4.4A and 4.5A). Under these conditions, it outperformed all the variant xylanases created in this study. However, the enzyme was not thermostable because of its inability to tolerate exposure to simultaneous heat (70°C – 90°C) and alkaline treatment (pH 8 – 10) (Figs. 4.3 - 4.5).

S340, S325 and G41 were considerably stable under all conditions tested and seemingly satisfied the stability requirements for commercial application. The fact that there remained any activity of these xylanases at all even after 120 min exposure to 90°C, is quite fascinating since both the cloned and wild type XynA from 70°C (pH 10) onwards exhibited almost no activity. Regardless of the pH of incubation, G53 was inactivated at 120 min at 80°C. At pH 3, S340, S325 and G41 retained 51%, 78% and 87% residual activity after 45 min whilst Luminase has 40% activity (Fig. 4.6A). The 3 xylanase variants exhibited lower pH optima and this presumably enhanced their stability at pH 3, compared to the other enzymes tested. However, as the pH increased, the stability of the other xylanases increased as well (Figs. 4.6B and 4.6C).

These preliminary biochemical characterization results seem impressive and in some instances, similar, when placed in context with the capabilities of other recombinant xylanases studied to date. Engineering of a disulphide bridge into the *T. reesei* endo-xylanase II resulted in the increase of its $t_{50}$ from 40 s to 20 min at 65°C and from 10 s to 6 min at 70°C at pH 5 (Fenel *et al*., 2004). A more recent study was conducted into this mutant with the intention of improving alkaline stability (Fenel *et al*., 2006). At pH 9, the construct showed an increase of 5°C in the $t_{50}$ value when studied over 10 min. Directed evolution of a xylanase isolated from a genomic DNA library resulted in improvement of the stability of the enzyme at 80°C (pH 6) after 20 min (Palackal *et al*., 2004). Another directed evolution study on *N. patriciarum* xylanase resulted in its mutant having 60% activity after 60 min at 60°C (pH 6) (Chen *et al*., 2001). Introduction of a disulphide bridge into the thermostable xylanase of *T. xylanilyticus* resulted in the improvement of its stability to 75% after 180 min at 60°C and pH 5.8 (Paës and O'Donohue, 2006).

Some natural xylanases with novel pH and temperature stabilities have also been documented. *Bacillus* sp. AR-009 produces 2 xylanases, one of which retained 54% and 67% of its activity after 60 min incubation at 65°C (Gessesse, 1998). Another *Bacillus* sp. NCIM 59 strain produces a xylanase stable at pH 10 with a $t_{50}$ of 24 h at 25°C (Nath and Rao, 2001a; 2001b). Recent examples include *Enterobacter* sp. MTCC 5112 which retained a remarkable 90% of its activity after 40 min at 90°C (Khandeparkar and Bhosle, 2006b) and 2 xylanases isolated from *Myceliophthora* sp. IMI 387099 that are stable at pH 9.2 at 60°C for 120 min (Chadha *et al*., 2004).

These results reinforced the theoretical concept that the catalytic action of an enzyme is influenced by physico-chemical conditions, in particular, those of pH and temperature. Changes in these conditions can have irreversible or reversible effects on the enzyme. Enzyme kinetics provides valuable information about the mechanism of an enzyme-catalyzed reaction and allows further insight into the likely properties of an enzyme in a simulated environment of its commercial application (Price and Dwek, 1979; Brownie and Kernohan, 1999). In the previous two chapters, thermostability and alkalinity were studied independently of each other and since a prebleaching pulp reaction combines both these variables, it was important to characterize their kinetics on this level as well. All kinetic data are reflected as a function of the 3 different pH conditions (pH 8 – 10) studied from 60°C – 90°C and are reflected in Tables 4.1 – 4.3.

First-order kinetics dictates that the k value will increase with temperature to a certain point until it denatures. Before a reaction can take place, it is necessary for the reactants to overcome the 'energy barrier' which is called the $E_a$. The greater this value, the more heat is required to 'energize' the enzymes for the reaction to proceed (Wynn, 1979). All 7 xylanases follow first-order deactivation and Arrhenius kinetics since the k value and $E_a$ increased with increasing temperature. Luminase had tremendous stability at pH 8 – 9 at 60°C, which was not only observed in the previous graphs (Figs. 4.3A and 4.4A) but was echoed in its superb $t_{50}$ values at of 385 and 346.50 min, respectively (Tables 4.1 and 4.2). At this temperature, this xylanase had the lowest k value but overall when its stability was examined at pH 8 from 60°C - 90°C, it was noted that Luminase was not very stable at high temperatures since its $E_a$ (142.84 kJ/mol) is the highest. Temperature seemed to be more critical for its stability, since it managed to retain 70% activity after 120 min at pH 10 (60°C). Its $E_a$ further increased to 151.92 and 160.68 kJ/mol at pH 9 and 10 respectively (Tables 4.2 and 4.3). Generally, $E_a$ and k are directly proportional to each other and increases in both values indicate poor catalyst stability under the tested conditions (Chaplin, 2004).

The cloned and fungal XynA controls were not remarkably stable under the combined challenge of extreme alkaline and thermal conditions used in this study. G53, on the other hand, was the most stable xylanase variant generated in this study at 60°C. Its $t_{50}$ values of 46.20, 40.76 and 36.47 min at pH 8, 9 and 10 respectively bear testimony to

this (Tables 4.1 – 4.3). However, this alkaline stability was lost as soon as the heat challenge was increased, resulting in k and $E_a$ values even higher than the control XynA. Therefore, in terms of temperature stability, G53 is not a thermostable xylanase since at temperatures greater than 60°C from pH 8 – 10, it almost always displayed the highest k and $E_a$ and the lowest $t_{50}$ values; in many ways, it proved to be even less stable than the control strains.

S340, S325 and G41 displayed better k and $E_a$ values when tested under these extreme conditions. At pH 8, S340 always had the best $t_{50}$ value from 60 - 90°C and the margin of difference in $t_{50}$ values between S340, S325 and G41 was large at 60°C but the margin became smaller as the temperature increased, until they displayed almost identical $t_{50}$ values at 90°C (Table 4.1). As the pH increased, the margin of difference between them was relatively less for each progressive rise in temperature (Tables 4.2 and 4.3). A noteworthy observation was that at pH 8 and 9, S340 was the most stable variant since its $E_a$ values were the lowest (Tables 4.1 and 4.2) but at pH 10, from these 3 variants, its $E_a$ value was the highest, even though it still technically had the best $t_{50}$ values. G41 was the next best stable variant closely followed by S325.

Thus, it seems that the mutations for these variant xylanases outlined in chapter 3 in some way, influenced their unfolding kinetics by slowing down denaturation and thus increasing thermostability and crucially, also improving alkaline stability. Previous studies conducted on *Bacillus* NCL 87-6-10 and *B. agaradhaerens* G/11 xylanases postulated that the enhanced thermal stability of these enzymes inadvertently also helped them attain a higher alkaline stability (Poon *et al*., 2003; Balakrishnan *et al*., 2006). Due to better solubility of xylan under alkaline conditions, alkaline-active xylanases may also find other potential applications in addition to pulp bleaching. For example, in waste management programs xylanases can be used to hydrolyze xylan in industrial municipal waste.

G41 and S340 were found to be the most stable xylanases generated in this study since their stabilities under combined stresses of extreme heat and pH were consistent for the range of conditions tested and they almost always exhibited the lowest $E_a$ and k values, which are the hallmarks of stable catalysts. Although the variants generated in this study had much lower activities, they exhibited significantly higher stabilities

than Luminase under the tested conditions and this represents the first step towards possible commercialization of a product with novel properties.

Biochemical characterization revealed that the mutations analyzed at a genetic level in chapter 3, not only altered the pH and temperature for optimal activity of the variant xylanases but, also rendered them more tolerant of combined thermophilic and alkaline conditions. The logical step after biochemical characterization would be to apply these xylanase variants to a pulping reaction and evaluate them against a commercial xylanase such as Luminase. However, low expression of the recombinant xylanases was a huge stumbling block for potential commercial application since larger quantities of purified xylanases are required. Cloning the xylanases into a His-tag vector and purification using affinity columns would certainly aid in future studies. For application tests and determination of commercial viability of the xylanases, high-level expression, preferably in a eukaryotic host, is necessary.

In this study thus far, only the thermal and alkaline properties of the mutant library were investigated since they are the chief requirements for the success of any pulping reaction. However, mutation could have altered other properties of the xylanase library, such as tolerance to acidic conditions and altered substrate preference. These latter properties not were investigated thus far and there existed the possibility that xylanases exhibiting these traits could be found in the original library and the next chapters examine these two aspects.

# CHAPTER 5: SCREENING FOR ACID-STABLE XynA VARIANTS

## 5.1 INTRODUCTION

The uses of alkaline thermostable xylanases have not only been well-documented in literature, but have been expounded extensively in this study as well. Most of the F/10 and G/11 xylanases have preferences for pH optima ranging from pH 5 – 9. Most mutational studies have even focussed on improving this alkalophilicity, as well as the thermophilicity of xylanases. However, few research groups have executed studies specifically for improvement of xylanase stability at acidic pH, possibly because of the difficulty in achieving this end.

Surprisingly, there exist many avenues of biotechnological application for acid-tolerant xylanases. Xylanase I from *A. niger* is produced at an industrial scale by the Gist-Brocades Company in Delft, Netherlands where it is used as a supplement in chicken feed and plays an important role in effective feeding and in the reduction of animal excrement (Krengel and Dijkstra, 1996). Acidic xylanases also have potential use in the wine industry where they interact with grape precursor compounds and play a role in varietal aroma of the final wine product (Strauss *et al*., 2001).

In Japan, the acid-tolerant properties of aspergilli are extensively exploited in the food and beverage industry. *A. kawachii* is used for making 'schochu', a Japanese traditional spirit, and produces cellulolytic enzymes that contribute to the digestion of barley. The pH of schochu mash is very low (about pH 3) because of the citric acid production by *A. kawachii*. Nevertheless, the enzymes act efficiently without inactivation (Fushinobu *et al*., 1998). *A. oryzae* is used for production of traditional foods such as 'shoyu' (soy sauce) and 'sake' (a potent rice wine). Analysis of the soy sauce mash, which consists of a mixture of soybeans and wheat, revealed that the fungus produces copious amounts of cell wall-degrading enzymes such as cellulase and xylanase during fermentation (Kimura *et al*., 2002).

Another highly attractive use of acid-stable xylanases is found in the pulping industries that utilize the bisulphite cooking process, which is performed under acidic conditions of pH 3 - 5 with a cooking liquor consisting of mainly bisulphite and

without excess $SO_2$. The acidity of the process effectively hydrolyzes and dissolves lignin. Another advantage is that ordinary sulphite woods, as well as pine and different hardwoods, are successfully digested. The process is not as successful with raw materials such as straw, bagasse and bamboo (Christov *et al*., 2000). The presence of certain metal cations in the pulp, such as $Fe^{2+}$ and $Mg^{2+}$, causes unwanted degradation of cellulose and removal of cations is an essential part of bleaching. The metals can be removed by chelating agents such as EDTA or by acid washing at low pH. The disadvantage of using EDTA is its slow natural degradation. In order to carry out simultaneous metal removal and bleach boosting, enzymes with optimal functioning at the acid washing pH would be beneficial. A further advantage of using an acidic xylanase is the avoidance of pH adjustment steps prior to ozone or chlorine dioxide bleaching stages, which are carried out in the acidic pH range (Tenkanen *et al*., 1997.

The effects of mutation can allow a protein to become multi-faceted and exhibit changes not only in global structure but in biochemical behaviour as well. The previous chapters focussed on using ep-PCR and recombination to obtain thermo- and alkaline-stable *xynA* variants. However, it was also possible that mutation could have expanded the biochemistry of a few of the *xynA* variants to include stability at acidic pH. Thus, the objectives of this chapter were: (1) to develop a screening method to detect acid-stable *xynA* variants from the original mutant library; and (2) to characterize potential acid-stable candidates and determine their stability at acidic pH.

## 5.2 MATERIALS AND METHODS

### 5.2.1 Growth of *E. coli* clones and enzyme extraction

The original mutant library consisting of 960 clones (chapter 2) were grown and their enzymes extracted as described in section 2.2.4. The *E. coli* clone carrying the wild-type XynA served as the unmutated control for all screening assays.

### 5.2.2 Screening for acid-stable *xynA* variants

The acidified media was prepared by modification of a protocol devised for screening of acidophilic algae by Olaveson and Stokes (1989). Double strength birchwood xylan (0.3%) was made up in 150 ml 0.05 M citrate buffer (pH 2.9) and autoclaved separately from 3% agarose dissolved in 150 ml warm distilled water. After autoclaving, both solutions were cooled to 50°C, carefully swirled in the flask then poured into sterile petri dishes to form a thick layer. After the medium had solidified, evenly-sized wells were punched into the agar using an Ouchterlony well maker.

The actual acid screening method was adapted for use in this study from the original protocols by Xiong *et al*. (2004) and Teather and Wood, (1982). Twenty microlitres of each crude extract of the 960 xylanases tested was inoculated onto pH 3 xylan plates prepared above and then incubated at 60°C for 4 h. XynA served as the control and was included on each screening plate for reference. The plates were then stained with 0.2% NaOH for 30 min. After the 0.2% NaOH was decanted, 1% Congo Red was added to the plate and incubated at room temperature for 30 – 90 min depending on the development of a clear zone of hydrolysis. Finally, the Congo Red was decanted and 1 M NaCl was added for 5 – 10 min to wash off excess dye. Enzyme variants that produced zones which were larger and clearer than that of XynA were chosen to undergo a liquid enzyme stability assay.

Culture lysates were diluted in 0.05 M citrate buffer (pH 3) and incubated in a 60°C water bath and residual activity of the enzymes were determined as previously described (section 2.2.5.1).

### 5.2.3 Determination of pH and temperature optima

The effect of pH on enzyme activity of the best acid-stable mutants were determined in a range of buffers at various pH values at 50°C, as outlined in section 4.2.2. Similarly, the effect of temperature on the enzyme activity of the most acid-stable mutants was determined at their optimal pH as outlined in section 4.2.2.

### 5.2.4 Determination of xylanase stability over acidic pH range

Enzyme filtrates from the acid-stable mutants were diluted in 0.05 M citrate buffer (pH 3 – 5) and incubated at 60°C for 90 min. Samples were removed every 15 min and assayed for residual xylanase activity as described in section 2.2.5.1.

### 5.2.5 DNA sequence analysis of acid-stable *xynA* variants

DNA sequencing and analysis of the acid-stable enzyme variants were performed as described in section 3.2.8 to determine what mutations contributed to the differences observed at acidic pH.

## 5.3 RESULTS

### 5.3.1 Screening for acid-stable *xynA* variants

Acid screening was performed using two assays, *viz*., a plate assay and a liquid assay that tested stability at pH 3 for 40 min. The entire xylanase library of 960 clones was screened on pH 3 xylan plates and stained with NaOH and Congo Red to determine the sizes of the zones of hydrolysis. XynA produced a weak zone of hydrolysis on xylan plates at pH 3. Enzymes that exhibited similar or poorer xylan degradation at this pH were not considered for further screening. Two variants, *viz.*, D57 and D63 produced distinct xylan hydrolysis on the acidic plates (Fig. 5.1). Mutant D57 produced a larger and clearer zone (1.8 cm) than either XynA (1 cm) or mutant D63 (1.5 cm) and this was an early indication of its improved activity at pH 3. From the entire library screened at pH 3, 150 of the enzymes displayed better activity than XynA (Table 5.1).

After incubation of these enzyme variants at pH 3 for 40 min at 60°C, the 2 best mutants, *viz*., D57 and D63, retained 61% and 58% residual activity, respectively, in comparison with XynA which had 19% residual activity.

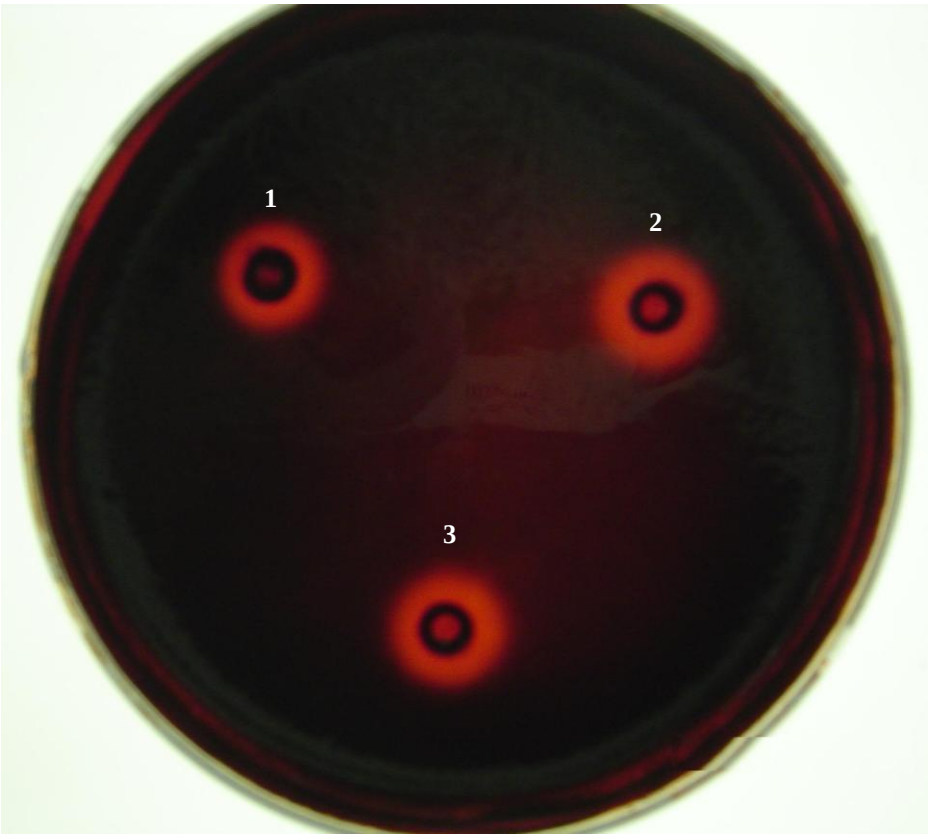


**Fig. 5.1** Screening plate showing zones of hydrolysis produced by crude xylanases created by mutagenesis of *xynA*. The wells contained crude xylanases from (1) XynA, (2) D57 and (3) D63. The plates were prepared according to modified methods (Teather and Wood, 1982, Olaveson and Stokes, 1989, Béguin, 1990 and Xiong *et al*., 2004) and are described in section 5.2.2.

**Table 5.1** Screening of mutant xylanase library on 0.3% xylan agarose plates (pH 3)

| Condition[♣] | Number of mutants screened | Number of mutants producing zones larger than the parent |
|---|---|---|
| B | 200 | 32 |
| C | 119 | 21 |
| D | 91 | 27 |
| E | 169 | 11 |
| F | 266 | 42 |
| G | 115 | 17 |

[♣]Alphabets refer to the different mutagenic conditions used in this study (see Table 2.1)

### 5.3.2 Determination of pH and temperature optima

Xylanases from D57 and D63 exhibited typical bell-shaped profiles during determination of their pH optima (Fig. 5.2). However, they exhibited wider pH profiles than XynA. Mutant D57 had the same pH optimum as XynA at pH 6.5 with the exception that it had almost 99% activity as compared to XynA which had 67% activity at pH 6. The optimum of mutant D63 moved slightly towards the acidic range and had a new optimum at pH 6.

Analysis of the effect of temperature on xylanase activity showed D63 to have a temperature optimum of 70°C, which is the same as that of XynA (Fig. 5.3). Both D63 and XynA displayed broader operating temperature ranges and were seen to have wide pH curves. D57, on the other hand displayed optimal activity at 73°C and was shown to have a narrower temperature profile with a slight shift towards greater thermotolerance.

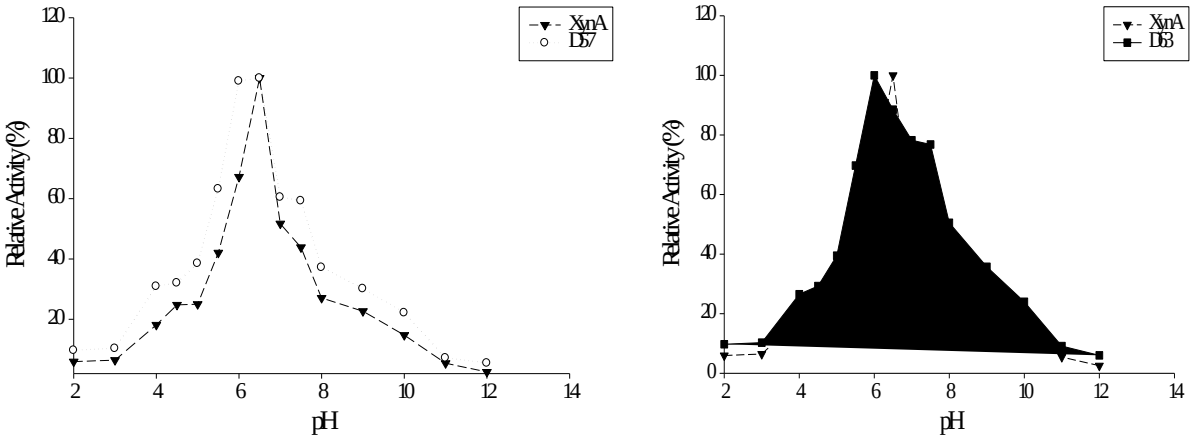


**Fig. 5.2** The effect of pH on the activity of D57 and D63 against wild-type XynA from pH 2 – 12 at 50°C for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

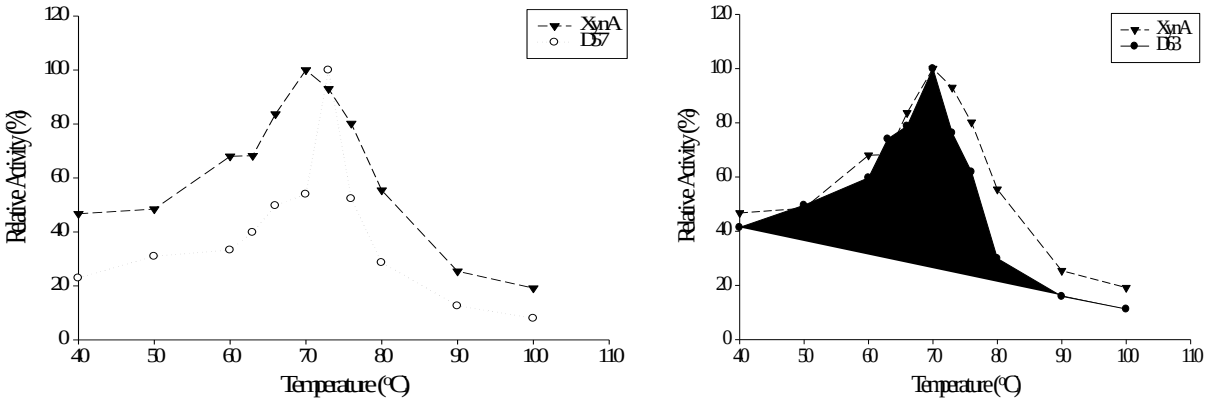


**Fig. 5.3** The effect of temperature on the activity of XynA, D57 and D63 against wild-type XynA from 40°C – 100°C at their optimum pH for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

### 5.3.3 Long-term xylanase stability over acidic pH range

Only two mutants, *viz*., D57 and D63, were tested for 90 min from pH 3 - 5 to determine the long-term effects of acidity on enzyme stability. Both these mutants were more stable than XynA for almost 60 min at pH 3 (Fig. 5.4). D57 and D63 had better stabilities than XynA at 57%, 48% and 26% respectively. However, after 90 min at pH 3, all three enzymes attained almost 15% residual activity. At pH 4 the stabilities of XynA, D57 and D63 attained residual activities of 25%, 37% and 24% respectively after 90 min (Fig. 5.5). At pH 5, the stabilities of XynA and D57 increased to 88% and 85% whilst D63 became more stable with a final residual activity of 59% (Fig. 5.6).

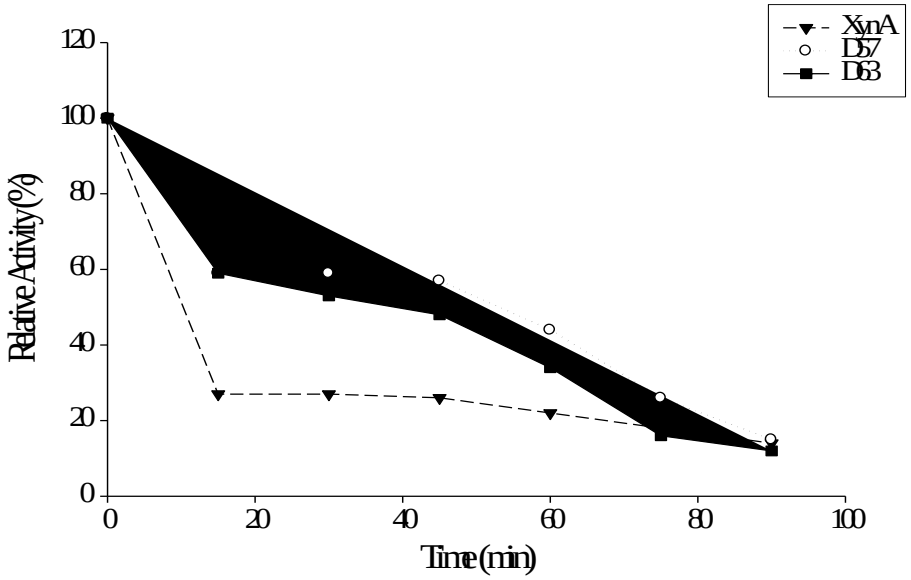


**Fig. 5.4** The effect of acidity on the stability of XynA, D57 and D63 at 60°C and pH 3. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

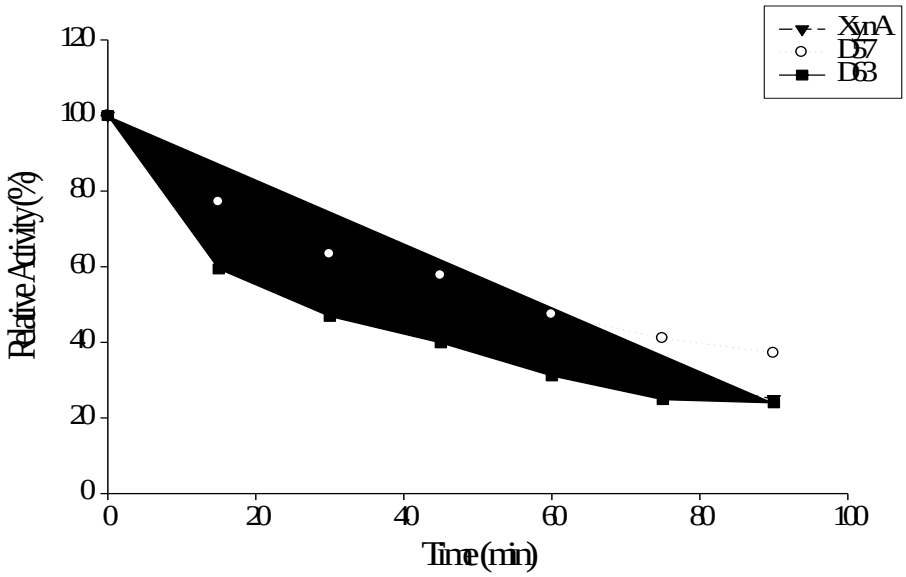


**Fig. 5.5** The effect of acidity on the stability of XynA, D57 and D63 at 60°C and pH 4. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

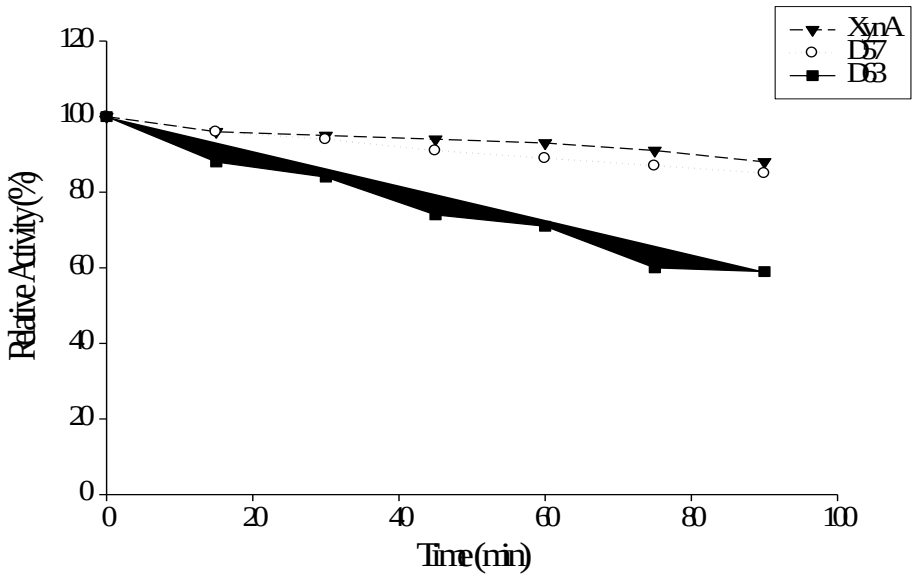


**Fig. 5.6** The effect of acidity on the stability of XynA, D57 and D63 at 60°C and pH 5. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

### 5.3.4 Sequence analysis of acid-stable mutants

The xylanase genes of mutants D57 and D63 were sequenced to determine the amino acid changes that rendered them slightly more stable than *xynA* at pH 3. Both enzymes had 99.5% amino acid sequence similarity to *xynA* with a single substitution each. Both substitutions occurred at the carboxyl termini of *xynA*. For D57, it was D223E, and for D63 it was T208A, which resides directly adjacent to the catalytic glutamic acid that forms part of the active site (Fig. 5.7).

```
xynA      MVGFTPVALAALAATGALAFPAGNATELEKRQTTPNSEGWHDGYYYSWWSDGGAQATYTN 60
D63       MVGFTPVALAALAATGALAFPAGNATELEKRQTTPNSEGWHDGYYYSWWSDGGAQATYTN 60
D57       MVGFTPVALAALAATGALAFPAGNATELEKRQTTPNSEGWHDGYYYSWWSDGGAQATYTN 60
          ************************************************************

xynA      LEGGTYEISWGDGGNLVGGKGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
D63       LEGGTYEISWGDGGNLVGGKGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
D57       LEGGTYEISWGDGGNLVGGKGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
          ************************************************************

xynA      VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
D63       VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
D57       VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
          ************************************************************

xynA      VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVG 225
D63       VQTGCHFDAWARAGLNVNGDHYYQIVAAEGYFSSGYARITVADVG 225
D57       VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVAEVG 225
          ***************************:***************:**
```

**Fig. 5.7** Amino acid sequence alignment of mutants D57, D63 and wild type *xynA*. The alignment was done by using the CLUSTALW (version 1.81) alignment program on the GenomeNet server (www.ebi.ac.uk/clustalw). 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; '.' indicates conservation within weaker groups of amino acids. The absence of an alignment character implies that an unrelated amino acid was substituted. Amino acid substitutions that differ from *xyn*A are indicated in red boldface. Catalytic residues that make up the active site of *xynA* are shown in italics.

## 5.4 DISCUSSION

Directed evolution has rapidly developed to become a method of choice for protein engineers in order to create enzymes which have enviable properties for all kinds of processes. In fact, every directed evolution experiment is an individual search for the most superior enzyme for the desired reaction, taking into account codon similarity and bias. In order to be able to tackle each question as efficiently as possible, it is convenient to have some general tools at hand. One of the possibilities is to maintain a huge variety of similar enzymes in the freezer and screen them for the desired activity when necessary. Having some general directed evolution methods at hand and readily adjustable screening or selection methods is also vital (Steipe, 2000; Otten and Quax, 2005).

During directed evolution, scientists are often hampered by ignorance of how amino acid sequence affects every aspect of enzyme performance; even if one trait is successfully designed, it is often impossible to predict the cost to another (Arnold and Volkov, 1998). Alkaline-stable xylanases are far more commercially-lucrative and have accordingly, been the subject of much scientific scrutiny. Acid-stable xylanases nonetheless, also have potential in the wine, feed, food and pulping industries. It is much more difficult, though to engineer enzymes to have tolerance at the lower end of the pH range if they do not have some meagre tolerance of acidity. The majority of xylanases reported to date are optimally active in the mildly acidic to neutral pH range but few can tolerate conditions of lower than pH 4 (Gessesse, 1998). Thus, the objective of this chapter was to determine if the original *xynA* mutant library, which was created primarily for evolving thermostable and alkaline stable enzymes, contained any variants that were tolerant of acidic conditions.

Development of a suitable plate screening method sensitive enough to detect acid-tolerant xylanases was a challenging task. Initially, problems were experienced with the actual preparation of the plates, which did not solidify because of the low pH. Separately autoclaving the agarose from the acidic components solved this problem by preventing breakdown of the agarose at low pH. For detection of zones of hydrolysis, Congo Red was flooded onto a plate containing the enzyme after an appropriate incubation time. This resulted in the plate immediately turning blue-black

without any visible clearings. A literature search revealed the cause of this observation: Congo Red serves as an indicator in acidic media and turns it blue. Subsequently, a method used by Xiong *et al*. (2004), originally intended to make polyacrylamide gels less acidic prior to staining, was adapted to suit the acid screen in this study. It was found that soaking the pre-incubated plate with 0.2% NaOH for 30 min prior to Congo Red staining was sufficient enough to detect zones of hydrolysis (Fig. 5.1). These zones were made more distinct by first washing off with 1M NaCl, then leaving for 5 – 10 min to remove unbound dye. However, plates had to be examined immediately as the plates turned completely black after 10 min.

This type of solid-phase screening relies on product solubilisation following an enzymatic reaction that gives rise to a zone of clearing. Congo Red screening generally precludes any quantitative comparison of enzyme activities. The method has inherent flaws that can influence zone size. These include enzyme size, migration of enzymes from the producing host as well as variations in growth rate between colonies. To minimize the effects of these variables, the same volume of lysate and an appropriate control was included on each screening plate.

Of the 960 clones screened, only 150 displayed better zones than XynA (Table 5.1). This was expected since only a small percentage of mutations are beneficial. All clones produced a zone; most with equal intensity to the control. XynA produced a clear, albeit small zone on the pH 3 agar, whilst the other two mutants produced better zones than XynA with differing clarities (Fig. 5.1). This screen was thus effective since it eliminated most of the undesirable xylanase variants. Enzyme evaluation of the 150 mutants for 40 min at pH 3 (60°C) showed that all but two mutants *viz*., D57 and D63, had poor stability at acid pH. These mutants displayed more than 50% residual activity under these conditions. Interestingly, upon examination of their alkaline and temperature stabilities, these mutants have less than 20% activity after 40 min at 80°C and their zones on pH 10 agarose were similar to XynA, so much so, that they were not even subjected to a liquid assay at pH 10 (refer to chapter two). Mutants D57 and D63 were then subjected to further pH and temperature profiling.

In determining their pH optima, it was noted that both displayed a typical bell-shaped pH activity profile. This profile shows loss of activity in the direction of increasing

acidity, attributed to protonation of the nucleophile partner, whilst loss of activity on the direction of increasing basicity is thought to be due to the dissociation of a proton from the proton donor (Lawson *et al*., 1997; Withers, 2001). Mutant D57 had an identical pH optimum to XynA of pH 6.5, whilst D63 had an optimum of pH 6 (Fig. 5.2). What is interesting, however, is that these two enzyme variants exhibited a wider pH curve than XynA, which displayed a narrow optimal range especially from pH 5 - 7. Conversely, D63 and XynA were found to have optimal activity at 70°C whilst D57 had optimal activity at 73°C (Fig. 5.3). XynA and D63 had slightly wider optimal temperature activity ranges but D57 had a narrower temperature curve with a shift towards increasing thermotolerance.

Xylanases with acidic pH optima also occur in nature. *A. niger* xylanase A has a low optimum pH (pH 4) and exhibits significant stability in acidic environments (Liu *et al*., 2006). Xylanase I from *A. niger* has a low optimum of pH 3, which is unusual for F/10 xylanases. These xylanases, in general, are most active at a pH of about 5 (Krengel and Dijkstra, 1996). Xylanase C from *A. kawachii* is another example of an acid-stable xylanase with an optimum of pH 2 and it is quite stable at pH 1 (Fushinobu *et al*., 1998). *A. pullulans* has an optimum of pH 2 (Ohta *et al*., 2001; Tanaka *et al*., 2004).

Long-term incubation at pH 3 - 5 of these enzymes along with the control XynA for 90 min at 60°C showed no distinct differences from XynA, since their remaining activities after this time was approximately the same under all conditions of treatment (Figs. 5.4 – 5.6). What is slightly out of the ordinary is that at pH 3, for about 60 min, both D57 and D63 have slightly better stabilities than XynA (Fig. 5.4). At pH 4 and 5, however, D63 had lesser stability than XynA (Figs. 5.5 and 5.6).

Sequence analysis of these mutants showed them to have a single amino acid substitution each: D57 had a D223E substitution near the C-terminus, located on the outer surface of the protein and D63 had a T208A mutation, which is adjacent to one of the catalytic glutamic acid residues in *xynA* (Fig. 5.7). Threonine is a bulky amino acid and the alanine substitution could have structural implications near the active site which accounts for the lower thermal stability of D63. Somehow, this substitution allowed the enzyme to function better at pH 3 but, it remains significant because at

pH 4 and 5, its stability was still lower than XynA, showing that mutations near the active site can compromise stability and activity (Morley and Kazlauskas, 2005). SDM studies on acidophilic adaptation of xylanases have highlighted the importance of residues located in the vicinity of these glutamic acids, especially in the thumb and at the edge of the cleft (Fushinobu *et al*., 1998; Joshi *et al*., 2000; Esteves *et al*., 2004). Currently, there are no publications based on the use of random mutagenesis to obtain acidophilic xylanases, presumably because such mutations are most effective when targeted at specific regions of the protein known to be involved in pH stability.

In D57, the D223E substitution also lends the enzyme good stability for 60 min at pH 3. A decrease in the number of aspartic acid and lysine residues was found to influence the pH character of enzymes in different ways, depending on the location of the substitution (Wang *et al*., 2005).

A few SDM studies regarding the acidic character of xylanases have been conducted. Xylanase C from *A. kawachii* was engineered to have a D37N mutation. This resulted in the shifting of its pH optimum from pH 2 to pH 5 with a 15% decrease in activity (Fushinobu *et al*., 1998). The N35D substitution in the *B. circulans* xylanase shifted the optimum from pH 5.7 to pH 4.6 with an almost 20% increase in activity (Joshi *et al*., 2000). Several mutants were constructed and it was shown that the glutamic acid at position 157 and aspartic acid at position 73 play a crucial role in stability at acid pH of *A. pullulans* xylanase (Tanaka *et al*., 2004). SDM was also recently used to shift the pH profile of *A. niger* phytase from pH 5.5 to 3.5 to enhance its effectiveness as an animal feed additive (Kim *et al*., 2006).

On the basis of a structural comparison between xylanases exhibiting different pH dependences, the acidic residues at the edge of the cleft and some aromatic residues forming subsites on the enzyme surface were suggested to be crucial structural differences. It was found that the negative charge of the active site may be important for molecular recognition and substrate specificity, because xylan exhibits a certain degree of substitution, especially of negatively-charged glucuronic acid substituents (Törrönen and Rouvinen, 1997; Fushinobu *et al*., 1998). Three-dimensional modelling of these *xynA* variants would undoubtedly uncover the biochemical basis of these enzymes' increased stability at pH 3 and form the foundation for future mutagenesis

studies for improvement of the xylanases' acid stability. The stability of these mutants at pH 3 could be further enhanced by further mutagenesis, SDM and shuffling.

# CHAPTER 6: SUBSTRATE SPECIFICITY OF XynA VARIANTS

## 6.1 INTRODUCTION

Despite the impressive ability of natural enzymes to catalyze a broad array of reactions and utilize diverse substrates, alteration of substrate specificity of an existing enzyme is vital. It provides not only efficient catalysts with designed substrate specificity but also valuable information on the mechanism of substrate recognition. Modifying enzyme substrate specificity represents a critical frontier for adapting enzymes to industrially desirable reactions. A number of directed evolution studies have demonstrated that substrate preference can be evolved (Yano *et al*., 1998; Chirumamilla *et al*., 2001; Kaufmann and Schmidt-Dannert, 2001; Kittur *et al*., 2003; Paës and O'Donohue, 2006). Substrate specificity is the preference that an enzyme manifests for one substrate over competing substrates. Implicit in this definition is the understanding that specificity is dependent both on substrate binding and on the utilization of the bonding energy for catalytic turnover. The binding of a substrate is necessary, but not sufficient, for catalysis (Harris and Craik, 1998).

The two main xylanase families differ from each other in substrate specificity. F/10 xylanases have a greater catalytic versatility or lower substrate specificity and are inclined to form oligosaccharides with a low degree of polymerization. Many of the xylanases from this family are able to degrade cellulose. G/11 xylanases are cellulase-free, tend to be more specific for xylan, producing larger oligosaccharides, and do not accommodate a modular substrate binding domain (Moers *et al*., 2003). In many directed evolution studies conducted thus far, the grafting of a new catalytic activity of interest upon a known enzyme's active site remains rare and xylanases are no exception to this rule. But, this study essentially focussed on the effects of mutation on G/11 xylanase and mutations that somehow broaden the xylanases' catalytic activity which could have implications for this group of enzymes.

The main xylan-containing substrates used for current research are birchwood, beechwood and oatspelts xylans. Birchwood (Roth) is a hardwood xylan consisting of 90% xylose, 1% arabinose, 1.5% glucose and 7.5% anhydrouronic acid. It is highly acetylated, which is responsible for its partial solubility in water. Beechwood xylan

(Sigma) is also a hardwood xylan that consists of a backbone of β-1,4-linked D-xylopyranose residues, with side chains of 4-*O*-methylglucuronic acid attached to the C-2 position of xylose and *O*-acetyl groups at C-2 or C-3 positions (Freixo and de Pinho, 2002). It has 90% xylose and 10% 4-*O*-methyl-D-glucuronic acid residues. Oatspelts xylan (Sigma) is a highly unsubstituted, largely insoluble xylan made up of 75% xylose, 15% glucose and 10% arabinose residues. These substrates contain a large amount of xylose, which makes them ideal substrates to standardize the activities of various xylanases (Li *et al*., 2000).

Cellulases, which are capable of hydrolyzing β-1,4-glycosidic bonds in cellulose, have been broadly divided into two types: endoglucanases (EG), which hydrolyze bonds internally in cellulose chains and cellobiohydrolases (CBH) that act preferentially on chain ends, cleaving off cellobiose as the main product. EGs have traditionally been recognized by their ability to readily degrade soluble substituted cellulose derivatives such as carboxymethyl cellulose (CMC), and the term CMCase has also been used to describe them. EG and CBH enzymes show considerable synergism in their simultaneous action on crystalline cellulose substrates (Henrissat *et al*., 1985; Bailey *et al*., 1993; Irwin *et al*., 1993; Medve *et al*., 1998). CMC (Sigma) is a derivative of cellulose formed by its reaction with alkali and chloroacetic acid. It has β-1,4-D-glucopyranose units which are primarily hydrolyzed by cellulases.

Lichenan from Icelandic moss (*Cetraria islandica*) is essentially a glucose polysaccharide made up of a predominantly linear β-(1→3) linked cellotriosyl units. The proportion of cellotriosyl units is higher in moss than in cereal β-glucans (Tosh *et al*., 2004). Lichenan is suggested to be primarily a structural element of the fungal wall, with important functions in thallus-water relations and enzymes specific for its degradation are called lichenases (Honeger and Haish, 2001). Most G/11 xylanases do not efficiently hydrolyze lichenan, whilst F/10 xylanases can hydrolyze this polymer to a degree. Interestingly, lichenases strictly hydrolyse β-(1→4)-linkages (like cellulases), but they typically have no activity against true β-(1→4)-glucans such as CMC (Grishutin *et al*., 2006).

The xylanase used in this study is a G/11 xylanase and therefore more specific in its hydrolytic functions. However, it is possible that the random mutagenesis experiments

conducted earlier in this study could have broadened the selectivity of the variant xylanases for different xylan-containing and other similar substrates. Thus, the objectives of this chapter were: (1) to screen the mutant xylanase library on different xylan substrates and cellulose substrates; and (2) to characterize and sequence mutants displaying changes in substrate specificity on the various substrates.

## 6.2 MATERIALS AND METHODS

### 6.2.1 Growth of *E. coli* clones and enzyme extraction

Eight hundred and ninety clones from the original mutant library (chapter 2) were grown and their enzymes extracted as described in section 2.2.4. The *E. coli* clone carrying the wild-type XynA served as the unmutated control for all screening assays. All xylanases were assayed in duplicate.

### 6.2.2 Substrate specificity of XynA mutants

The substrates used for determining specificity were birchwood xylan (Roth), beechwood xylan (Sigma), oatspelts xylan (Sigma), lichenan from *C. islandica* (Sigma) and CMC (Sigma). Each substrate (0.1%) was prepared in sodium citrate buffer (0.05 M, pH 6.5) by boiling and moderate heating and then clarified by centrifugation before use (Bailey *et al*., 1992). Crude xylanases from each mutant were diluted and their activities evaluated on the respective substrates as described in section 2.2.5.1. Forty one mutants that displayed improved activity on these substrates were re-grown in 300 ml LB medium as previously outlined in section 2.2.5.3 and reassayed using the same substrates.

### 6.2.3 Determination of pH and temperature optima

The effect of pH on enzyme activity of the best mutants were determined in a range of buffers from pH 2 – 12 at 50°C, as outlined in section 4.2.2. Similarly, the effect of temperature on enzyme activity of the best acid-stable mutants was determined at their optimal pH as outlined in section 4.2.2.

### 6.2.4 DNA sequence analysis of selected mutants

Sequence analysis of the best mutants was performed as described in section 3.2.8.

## 6.3 RESULTS

### 6.3.1 Substrate specificity of XynA mutants

Most mutants, like XynA, had the best catalytic activity on birchwood xylan, followed by beechwood and then oatspelts xylan (Table 6.1). Enzyme assays are very subjective and this experimental error was taken into account by assigning a 200 nkat/ml difference before describing an enzyme of having more, less or similar activity from the parent on a specific substrate. From the table, it is evident that most of the mutants did not display a marked improvement in their ability to hydrolyze these xylans. Mutagenic condition G, almost always had the most number of mutants with diminished activity on the different substrates but what remains interesting is that this condition, along with conditions C and E generated a few mutants with minor improvements in their affinity for lichenan. All mutants did not display any significant change in the ability of XynA to hydrolyze CMC. A summary of these findings are outlined in Table 6.1 and depicts the number of mutants that had improved, reduced or similar activities to XynA on these substrates.

Forty one mutants displayed high activities on birchwood, beechwood and oatspelts xylan in the preliminary analysis and were subsequently re-grown and assayed using the same substrates. Most of the mutants had activities comparable to XynA which had 2963 nkat/ml, 2254 nkat/ml, 2178 nkat/ml, 3 nkat/ml and 10 nkat/ml on birchwood, beechwood and oatspelts xylans, lichenan and CMC respectively. However, mutant D15 had slightly improved activities of 3222 nkat/ml and 3175 nkat/ml on birchwood and beechwood xylans, respectively. It displayed lower activities of 1792 nkat/ml, 8 nkat/ml and 2 nkat/ml on oatspelts xylan, lichenan and CMC, respectively. Mutant G84 was very interesting in that it had an almost similar activity on birchwood xylan to XynA, yet its activities on beechwood (1496 nkat/ml) and oatspelts xylan (1505 nkat/ml) were lower (Table. 6.2).

**Table 6.1** Substrate specificity of mutant *xynA* library for different xylan and cellulose substrates[♣]

| Condition[*] | No. of clones screened | Substrates | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | BIRCHWD X | | | BEECHWD X | | | OATSPELTS X | | | LICHENAN X | | | CMC | | |
| | | ↑ | ↓ | = | ↑ | ↓ | = | ↑ | ↓ | = | ↑ | ↓ | = | ↑ | ↓ | = |
| B | 131 | 15 | 15 | 70 | 0 | 35 | 65 | 18 | 24 | 58 | 0 | 35 | 65 | 0 | 63 | 37 |
| C | 118 | 16 | 21 | 63 | 27 | 18 | 55 | 8 | 74 | 18 | 3 | 81 | 16 | 0 | 81 | 19 |
| D | 91 | 21 | 37 | 42 | 22 | 32 | 46 | 31 | 52 | 17 | 0 | 97 | 3 | 0 | 82 | 18 |
| E | 169 | 53 | 18 | 29 | 15 | 60 | 25 | 36 | 43 | 21 | 2 | 89 | 9 | 0 | 88 | 12 |
| F | 266 | 36 | 24 | 40 | 3 | 69 | 28 | 24 | 33 | 43 | 0 | 73 | 27 | 0 | 45 | 56 |
| G | 115 | 23 | 47 | 30 | 1 | 76 | 23 | 19 | 53 | 28 | 5 | 53 | 44 | 0 | 50 | 50 |

[♣] All values are tabulated as percentages of the no. of mutants screened per mutagenic condition and are reflected as activity values (with a 200 nkat/ml margin of experimental error) higher than (↑), lower than (↓) or equivalent (=) to the control XynA.

[*]Alphabets refer to the different mutagenic conditions used in this study (see Table 2.1)

**Table 6.2** Xylanase activity (nkat/ml) of wild type XynA, D15 and G84 on different xylan and cellulose substrates

| Xylanases | Substrates | | | | |
|---|---|---|---|---|---|
| | BIRCHWD X | BEECHWD X | OATSPELTS X | LICHENAN X | CMC |
| XynA | 2963 | 2254 | 2178 | 3 | 10 |
| D15 | 3222 | 3175 | 1792 | 8 | 2 |
| G84 | 2361 | 1496 | 1505 | 0 | 5 |

### 6.3.2 Determination of pH and temperature optima

Since mutants D15 and G84 displayed some differences in activity on these substrates, their pH and temperature optima were determined. All tested strains had optimal activity at pH 6.5 (Fig. 6.1). However, the pH profile of mutant D15 was slightly broader at all pH values tested, particularly from pH 5 – 9 than that for the other two strains. From pH 5.5 – 9, G84 had a similar pH profile to XynA however; from pH 2 – 3 it was more active than both XynA and D15.

XynA, D15 and G84 all displayed optimal activity at 70°C (Fig. 6.2). Nonetheless, some slight changes to the temperature curves of the mutants were observed. Mutant D15 exhibited a narrower temperature profile than either XynA or G84, particularly from 60 - 80°C. The profile of G84, on the other hand was marginally broader than both XynA and D15 especially from 80°C - 100°C.

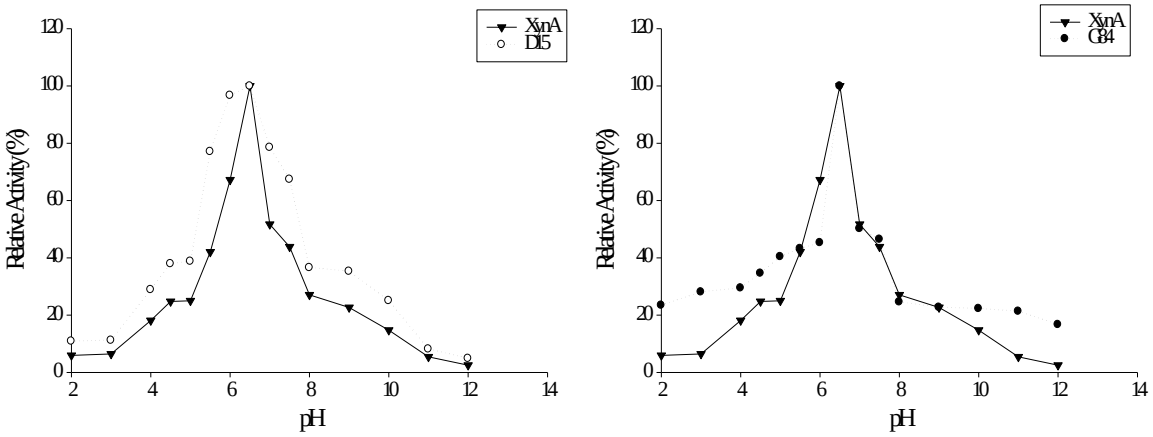


**Fig. 6.1** The effect of pH on the activity of D15 and G84 against wild-type XynA from pH 2 – 12 at 50°C for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

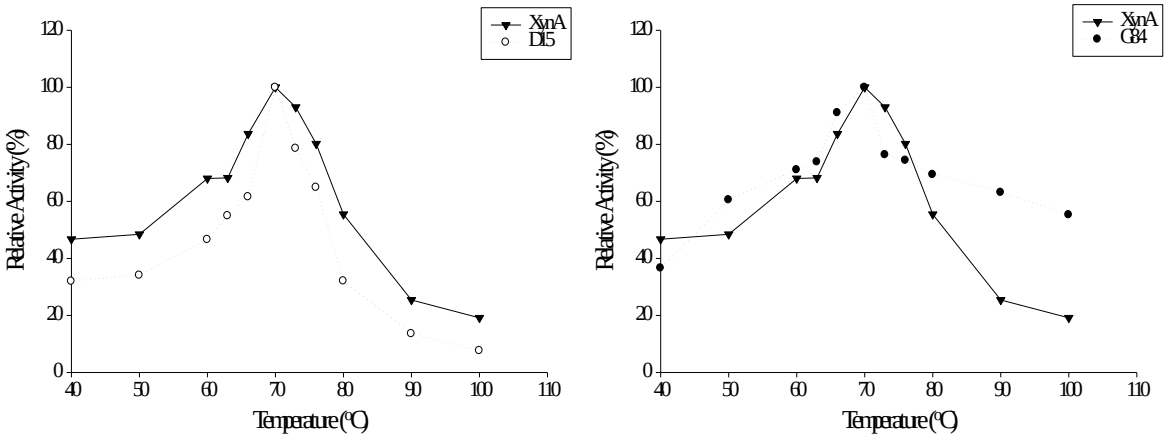


**Fig. 6.2** The effect of temperature on the activity of D15 and G84 against wild-type XynA from 40°C – 100°C at their optimum pH for 5 min. The remaining activities were expressed as percentages of the original activities with each point representative of duplicate determinations.

### 6.3.3 Sequence analysis of selected mutants

The xylanase genes of D15 and G84 were sequenced to determine the amino acid changes that allowed them to exhibit the observed differences when assayed on the different substrates. Both sequences displayed a 99.5 % amino acid sequence identity to *xynA* (Fig. 6.4). Mutant G84 had an alanine → valine substitution at position 10 and mutant D15 had a valine → alanine substitution at position 224.

```
xynA      MVGFTPVALAALAATGALAFPAGNATELEKRQTTPNSEGWHDGYYYSWWSDGGAQATYTN 60
G84       MVGFTPVALVALAATGALAFPAGNATELEKRQTTPNSEGWHDGYYYSWWSDGGAQATYTN 60
D15       MVGFTPVALAALAATGALAFPAGNATELEKRQTTPNSEGWHDGYYYSWWSDGGAQATYTN 60
          *********.**************************************************

xynA      LEGGTYEISWGDGGNLVGGKGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
G84       LEGGTYEISWGDGGNLVGGKGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
D15       LEGGTYEISWGDGGNLVGGKGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYI 120
          ************************************************************

xynA      VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
G84       VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
D15       VENFGTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGT 180
          ************************************************************

xynA      VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVG 225
G84       VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVG 225
D15       VQTGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADAG 225
          *******************************************.*
```

**Fig. 6.3** Amino acid sequence alignment of mutants D15, G84 and wild type *xynA*. The alignment was done by using the CLUSTALW (version 1.81) alignment program on the GenomeNet server (www.ebi.ac.uk/clustalw). 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; '.' indicates conservation within weaker groups of amino acids. The absence of an alignment character implies that an unrelated amino acid was substituted. Amino acid substitutions that differ from *xyn*A are indicated in red boldface. Catalytic residues that make up the active site of *xynA* are shown in italics.

## 6.4 DISCUSSION

Natural enzymes are typically quite specific for their intended substrates, and it is often necessary to modify or broaden their substrate selectivity for industrial or biotechnological applications. The relationship between enzyme structure and selectivity is complex and modifying the substrate selectivity of an enzyme is a bigger hurdle than improving the activity of an enzyme in different environments (Tao and Cornish, 2002).

The objective behind this aspect of the study was to determine if mutation had an effect (if any) on the ability of the mutant xylanase library to better hydrolyze certain substrates. Most of the variants hydrolyzed birchwood xylan best, followed by beechwood then oatspelts, which is the trend followed by XynA as well (Table 6.1). This echoes the oft observed finding that a significant proportion of the variants in directed evolution experiments maintain wild-type activity. From these results, it is also noted that subjecting a gene to increasing mutagenic conditions need not necessarily markedly improve its ability to hydrolyze certain substrates better than XynA; it can however, allow for slight improvements that can be enhanced in a stochastic fashion. It must be remembered that enzyme assays can be influenced by many factors, such as experimental error, variance in manufacture of different batches of substrates as well as actual substrate preparation itself. This was the reason for allowing a margin of difference of at least 200 nkat/ml when designating a specific clone as having either improved, lower or similar activities on the xylans tested. The use of substrate and enzyme blanks can help but because single colonies of *E. coli* clones rather actual protein measurements were used as standardization tools, it was necessary to regrow and exhaustively assay the clones displaying "improved" activities on the tested substrates.

From the original 890 clones tested for altered substrate specificity, 41 mutants with better activity than XynA were selected, re-grown and assayed on the same substrates and two were found to behave differently from XynA (Table 6.2). Mutant D15 not only had better activity on birchwood xylan but showed a 1.4 fold improvement in its ability to hydrolyze beechwood xylan. However, its activity on oatspelts xylan dropped 1.2 fold. It also displayed a marginal improvement in its ability to degrade

lichenan. Mutant G84 on the other hand had roughly the same amount of activity on birchwood xylan as XynA but its specificity for beechwood and oatspelts xylans was diminished 1.5 and 1.4 fold respectively. The differences in the substrate specificity of these mutants are probably due to the differences in the composition of the polysaccharides. In the case of birchwood xylan, the backbone is decorated with more acidic 4-*O*-methyl-α-D-glucuronopyranosyl residues. Neutral α-L-arabinofuranosyl residues are common side chains in oatspelts xylan (Kittur *et al*., 2003).

In determining the effect of pH on the activity of the xylanases of these mutants, it was noted that whilst all exhibited optimal activity at pH 6.5, their profiles differed (Fig. 6.1). XynA and G84 have distinctly narrow profiles from pH 5 - 7.5 whilst D15 had a wider pH profile in the same range. Conversely, mutant G84 showed a wider range of functioning from 60 - 80°C (Fig. 6.2) than XynA and D15 even though 70°C was the temperature at which they all exhibited optimal activity.

Sequence analysis of these mutants showed them to differ from *xynA* in a single amino acid each (Fig. 6.3). D15 had a substitution at the carboxyl terminus of the protein (V224A) whilst G84 had a mutation (A10V) at the amino terminal end of *xynA*, which is actually the signal sequence of the protein and may have a marginal effect (if any), particularly if this region is cleaved off during secretion of the final product in the native fungus. What is noteworthy is that both these mutations involve only valine and alanine and occur on opposite ends of the protein and are far from the active site. DNA shuffling experiments sometimes showed that amino acid changes distant from the active site can affect substrate specificity (Zhang *et al*., 1997; Yano *et al*., 1998). Such changes work by altering the orientation of the active-site residues or the conformational dynamics of the entire protein (Benkovic and Hammes-Schiffer, 2003; James and Tawfik, 2003) and are therefore difficult to anticipate. A previous study showed that mutations which alter substrate specificity occur less frequently than those that increase protein expression or solubility. Also, mutations distant from the active site somehow increased overall activity without really altering substrate specificity (Parikh and Matsumura, 2005). It is thus also important to take these suggestions into consideration as well.

In mutant D15, the V224A substitution accounts for improvements in the enzyme's ability to hydrolyze birchwood and beechwood but it decreased affinity for oatspelts xylan. The alanine substitution is a largely neutral one occurring at a non-critical part of the enzyme and was found to play a role in substrate recognition in *Cellovibrio* F/10 xylanase (Xie *et al*., 2006). In mutant G84, the A10V mutation caused it to exhibit a lower specificity for beechwood and oatspelts xylans, which coincidentally have less xylose residues than birchwood xylan. This mutation somehow favours catalysis with a more substituted xylan like birchwood. It seems surprising that a single substitution would influence substrate binding and catalysis but it has been previously noted and reviewed (Jestin and Vichier-Guerre, 2005). It seems that these substitutions are more often than not mutations that change the Van de Waals forces operating in proteins and more often than not occur in non-critical regions of proteins as in this study. It was also noted that mutation somehow modifies other unmeasured physico-chemical properties, which in turn influence the enzyme's progression through the cell wall network and its interactions with various cell wall components (Paës and O'Donohue, 2006).

Prior to conducting this investigation, it was known that it would be ineffectual to hope to acquire a xylanase variant with markedly different substrate selectivity on the substrates tested after just a single round of mutagenesis. This is primarily because directed evolution experiments to obtain enzymes that catalyze fundamentally different reactions have not yet been demonstrated. Drastic changes in protein function require considerable conformational rearrangements not likely to occur in libraries of highly homologous sequences, let alone libraries obtained after a single round of ep-PCR. It is difficult for mutations to coax a particular enzyme to perform a fundamentally different chemistry since in nature, most enzymes may have been selected so as to have narrow enzyme substrate specificities to ensure that only one specific molecule or class of molecules will allow for catalysis to occur (Georgiou and DeWitt, 1999; Jestin and Vichier-Guerre, 2005.

There are some excellent examples of how the substrate specificity of enzymes was changed using protein engineering. Incremental truncation and shuffling resulted in the ability of a few mutants of the phospholipase from a strain of *Serratia* to not only degrade phosphoglycerides but also behave like lipases and degrade triacylglycerols

(Song *et al*., 2002). Dioxygenase enzymes like toluene dioxygenase are involved in the catabolism of aromatic hydrocarbons by soil microorganisms and the enzyme from *Pseudomonas putida* was successfully evolved within the laboratory to degrade the heterocyclic substrate 4-picoline (Sakamoto *et al*., 2001). What is important to note from these examples is that the substrate specificities of these enzymes was broadened to accept other substrates and this involved minor changes in biochemistry. The majority of examples to date have obtained improvement in activity towards substrates that are poorly accepted by the wild type enzyme. Obtaining entirely new substrate specificity is a challenge that will require further innovations into the exploration of sequence space.

This is an indication that directed enzyme evolution is probably the most efficient way approach to broadening enzyme substrate specificity. Such experiments reveal extensive relationships between sequences and catalytic activities and define a strategy for identifying amino acid substitutions associated with improvements in catalytic activity or shifts towards extended specificities. They further represent a straightforward strategy for identifying mutations that are sufficient to broaden the biochemical activities of a polypeptide chain. Such mutations should be of great use in the fields of biocatalysis and synthetic chemistry in general.

# CHAPTER 7: GENERAL DISCUSSION

## 7.1 THE RESEARCH IN PERSPECTIVE

In the face of growing environmental awareness at governmental and regulatory levels, various industries are increasingly evaluating biotechnology as a technically and commercially sound tool to improve overall sustainability. However, few natural biocatalysts are capable of meeting the requirements of modern biotechnological processes. So, to increase their utility, it is necessary to adapt their properties accordingly through *in vitro* evolution technologies such as directed evolution and DNA shuffling. These techniques mimic natural evolution in nature and seek to explore new avenues of sequence space of enzymes, thus enhancing the function of these mutated enzymes in non-natural environments (Kaufmann and Macready, 1995; Jaeger and Reetz, 2000).

It has been demonstrated that pre-treatment of kraft pulp with xylanase prior to bleaching can facilitate lignin removal by bleaching chemicals, thereby reducing the demand for elemental chlorine and improving final paper brightness (Viikari *et al*., 1994). To satisfy the specific conditions of the bleaching process at the pulp and paper mills, xylanases that are active and stable at elevated temperatures and alkaline pH are needed. Few natural xylanases are capable of meeting these demands and thus the application of *in vitro* molecular evolution techniques to enhance the properties of existing xylanases are highly attractive.

Thus, this study was undertaken to improve both the alkaline and thermal stabilities of the xylanase, *xynA*, from *T. lanuginosus* DSM 5826, originally cloned into *E. coli* as a fusion protein (Schlacher *et al*., 1996) and later crystallized to elucidate its 3-D structure (Gruber *et al*., 1998). The first phase of this study, prior to mutagenesis, was to remove the fusion partner from the xylanase, effectively reducing the size of the fungal DNA fragment from 1003 bp to 786 bp. This allowed beneficial mutations to occur more frequently in the *xynA* gene rather than the adjacent sequences. The new 3714 bp plasmid, pX4, served as the template for the entire study and also as the source for the unmutated control XynA enzyme.

This study focussed on the effects of mutation on the thermal, alkaline and acid stabilities of *xynA* as well as altered substrate specificity for xylan and cellulose substrates. Ep-PCR was used to randomly mutate *xynA* for improved performance under these conditions. Many methods were chosen to increase the repertoire of novel xylanase variants generated and are outlined in Table 2.1. Different conditions were used to enhance the probability of obtaining mutants with beneficial properties and also to increase the frequency of mutations. After screening of the xylanase variant library for improved properties, it was observed that the thermostable mutants displayed poor catalytic activities, possibly as they needed to become more rigid to tolerate thermal extremes. The xylanase variants were invariably more sensitive to temperature than pH fluctuations in their surroundings. After analyzing the results obtained, it is evident that mutation conditions D and G consistently generated mutants with better stabilities than XynA.

Condition D not only generated 11% of the thermostable mutants in total but, also produced mutants stable at pH 3 (D57 and D63), as well a mutant that had improved activities on all substrates tested except cellulose (D15). Condition D was based on the protocol developed by Xu *et al*. (1999) and involved the sequential use of dITP and $Mn^{2+}$ in a two-step ep-PCR reaction. The success of using inosine in an ep-PCR reaction is that it is a "universal" base and can base pair equally well with any of the three nucleotide bases (Cubero *et al*., 2001). The method was developed to compensate for the bias normally observed during traditional ep-PCR with $Mn^{2+}$, and in this study generated clones with different mutations and stabilities. Condition G, on the other hand, generated mutants that were thermostable (G41), alkali-stable (G53) and displayed altered substrate specificity (G84). This condition was the most mutagenic condition from the Diversify Random Mutagenesis kit (CLONTECH) used and was based on excess dGTP, imbalanced dNTP concentrations and the addition of $Mn^{2+}$.

One can use a plethora of methods to randomly mutate a target gene of interest but what is crucial for any protein engineering strategy is the use of a screen sensitive enough to detect the altered protein character (Salazar and Sun, 2003). This was the most arduous task in this study, especially with regards to detecting thermal tolerance and altered substrate specificity. For these experiments, it was necessary to extract the

enzymes of the entire mutant library and laboriously expose each filtrate to different temperatures and substrates and eliminate the undesirable clones sequentially.

Challenging screens for thermal tolerance have resulted in the identification of enzyme variants with substantially improved thermostability (Kim *et al*., 2003; Garrett *et al*., 2004; Palackal *et al*., 2004). In these studies, the improvement in thermostability was measured using differential scanning calorimetry (DSC) rather than by the measurement of relative activity with increasing temperature. In general, there are two main approaches that have been used for high-throughput screening for variants with increased residual activity. The first approach makes use of agar plates or membranes and has been used to improve the thermal tolerance of many enzymes including amino acid oxidase (Sakaue and Kajiyama, 2003) amylase (Kim *et al*., 2003), glucose dehydrogenase (Baik *et al.*, 2003), N-carbonyl-D-amino acid amidohydrolase (Oh *et al*., 2002) and glucanase (Murashima *et al*., 2002).

In the second approach, the assays are performed in microtitre plates. Microtitre plate assays can be highly quantitative and they allow kinetic measurement of enzyme activities. In most cases, turnover is quantified by absorbance or fluorescence using model substrates such as dye-linked polysaccharides, ρ-nitrophenyl or methylumbelliferyl substrate derivatives (Salazar *et al*., 2003; Zhang *et al*., 2003; Garrett *et al*., 2004) or colorimetric indicators (Gray *et al*., 2001; Sriprapundh *et al*., 2003) Another alternative method is based on phage display (Sieber *et al*., 1998; Martin *et al*., 2001). This method is based on the use of filamentous phages and on assessment in the form of proteolytic susceptibility. The premise for this method is that decreased proteolytic susceptibility relates to increased thermal stability but this method has many limitations since the method depends on viable phages that are susceptible to very high temperatures.

In this study, plate screens were developed for detection of mutant xylanases tolerant of acid or alkaline extremes. Developing a plate screen for alkaline pH was simple and is well documented in literature, but the screen itself was not as stringent in selecting the most stable xylanase at pH 10. The acid screen took longer to develop, since to date, no screen is available in literature for the selection of acid-tolerant xylanases, presumably because not much research is currently conducted in this area. The acid

plate screen was quite capable of selecting the acid-stable variants from those that were similar or inferior to XynA. In all instances, it was necessary to regrow the selected mutants, extract and detect enzyme activity under physiological conditions.

A lack of resources impeded the development of a more flexible high-throughput xylanase assay screening system using the 96-well format. All publications based on either SDM, directed evolution or recombination of xylanases used the traditional clear zone (halo) plate assay Congo Red staining method (Teather and Wood, 1982 and Béguin, 1990) or the RBB-xylan plate screening method used in this study. This highlights the challenges facing research in academic as opposed to commercial environments, where the use of robotics makes screening much simpler. Having such a system in place would have undoubtedly not only result in economy of labour and materials, but also the screening of a larger library which is crucial in the quest to obtain the most novel variant. Collaboration between four research institutes in Japan resulted in the development of a system for the easy detection of thermotolerant *B. subtilis* xylanase variants created using directed evolution (Miyazaki *et al*., 2006). The researchers modified the traditional dinitrosalicylic (DNS) method for reducing sugar detection (Miller, 1959) to a 96-well format and used the BugBuster Protein Extraction Reagent (Novagen) for mild extraction of the enzymes from *E. coli*. However, an expensive microplate centrifuge, microplate reader and thermotolerant microplates were required.

It was noted from screening the initial library for thermostable and alkaline-stable variants, that they displayed improvements in only one of these properties and not both. Thus, the two best mutants, *viz*., G41 and G53, were chosen as candidates for recombination to combine these beneficial mutations into a single xylanase. Recombination is one of the fastest ways of creating more functional enzyme variants. There are many methods of recombination and in this study, the StEP method was chosen and the reasons for this choice are documented in chapter 3. Two types of StEP reactions were carried out: one normal recombination and the second was made additionally mutagenic using ep-PCR condition G. This is the first time that the StEP reaction was ever made mutagenic; all other published work refers to it merely as a vehicle of recombination between highly homologous genes that can sometimes

generate point mutations (Bruhlmann and Chen, 1999; Zhao and Arnold, 1999; Dion *et al*., 2001; Murashima *et al*., 2002; Bulter *et al*., 2003; Miyazaki *et al*., 2006).

Most of the 531 recombinants screened had low activities and were more thermostable than either G53 or XynA. One recombinant in particular, S325, had better thermal stability than the thermostable parent G41. It, along with S340, attained intermediate alkaline stability. Recombinants created using mutagenic StEP displayed better thermal stability than G53 and XynA but had very poor activities. Only three recombinants in total had better activities than G53 and XynA.

The evolved parental variants (G41 and G53) and recombinants (S325 and S340) were subjected, along with XynA controls and the commercial Luminase, to long term stability testing at combined thermal and alkaline parameters similar to that carried out in pulp mills. Mutation resulted in the slight shifting of the pH optima of the xylanases towards the acid region whilst the temperature optima slightly shifted towards the thermophilic region. Kinetic data indicated that G41 and S340 were the best thermal and alkaline stable mutants generated in this study because of their low activation energies and reaction rates.

A huge discrepancy exists between the amount of xylanase produced by the fungal and cloned forms of the enzyme. The original fungal host secretes almost 30 000 nkat/ml of the xylanase into its culture medium (Purkarthofer *et al*., 1993a; 1993b). This is in stark contrast with the recombinant *E. coli* strain containing the cloned *xynA*, which produces xylanase with maximal activity of 4000 nkat/ml. The reasons for variability in enzyme expression between bacterial and fungal hosts include: the use of stronger promoters in fungi than bacteria; target proteins are commonly secreted intracellularly in inclusion bodies in bacteria and require lysis whereas fungal systems generally secrete their enzymes and the lack of post-translational modification systems in bacteria which leads to inactive or incomplete proteins, to name a few (Baneyx, 1999; Fernandez, 1999; Rai and Padh, 2001). Another reason is that fungal systems may contain a battery of enzymes that assist in substrate degradation, e.g., *T. lanuginosus* culture supernatant contains other enzymes involved in xylan degradation, even though they are produced at low levels. These include β-xylosidases that degrade the short xylo-oligosaccharides produced by the xylanases and debranching enzymes like α-

glucuronidases. The cumulative and synergistic action of all these enzymes would therefore lead to larger amounts of reducing sugars being produced compared with a culture with only β-xylanase activity (La Grange *et al*., 1996).

Sequence analysis revealed mutations in alkaline and thermostable xylanase variants were scattered throughout the genes, away from the active site of the enzyme and towards the surface of the proteins. The highly thermostable enzymes (G41 and S325) had an increase in arginine content. Arginine residues are charged and are highly involved in forming salt bridges across the global structures of proteins. Investigations into the role of electrostatic interactions in systemic protein flexibility have shown that salt bridges and their networks observed in protein crystal structures may easily break and reform in solution (Kumar and Nussinov, 2000; 2001). Several studies have shown that an increase in arginine content on the surface of proteins can increase their thermostability (Riordan *et al*., 1977; Argos *et al*., 1979; Vogt *et al*., 1997; Turunen *et al*., 2002). The disruption of the disulphide bridge in mutant M28 may be responsible for its low stability under all parameters tested.

Sequencing of the mutants that had better stability than XynA for 1 h at pH 3, *viz*., D57 and D63, showed that both variants had one mutation each at the carboxyl end of the gene. This region contains one of the two catalytic glutamic acid residues of the acid-base nucleophile that forms the active site and previous SDM studies have shown that mutation near these regions are thought to influence the acid character of xylanases (Fushinobu *et al*., 1998; Joshi *et al*., 2000; Esteves *et al*., 2004). This is the first report of an acidophilic xylanase obtained using ep-PCR: all others have been obtained using SDM.

Determining altered substrate specificity of the xylanase variants on different xylan and cellulosic substrates was difficult without standardizing protein concentration. Sequencing of mutants G84 and D15, which showed altered activities on the different substrates, revealed the importance of valine and alanine in substrate hydrolysis.

Directed evolution studies have frequently concluded that beneficial mutations do not necessarily occur within the active site of the enzyme (Arnold, 2001). However, the contribution of each mutation observed in the final enzyme variant is rarely

determined. Also, directed evolution is an iterative process that rarely identifies multiple synergistic mutations. Consequently, it remains unclear whether certain enzyme properties are more likely to arise from subjecting entire genes to a low intensity of mutagenesis or targeted regions to a high intensity (Dalby, 2003).

## 7.2 FUTURE PROSPECTS FOR RESEARCH

One of the most challenging aspects for research into the evolution of xylanases is undoubtedly improving the activity of thermostable enzymes. A variety of prokaryotic and eukaryotic organisms have been used as expression hosts in attempts to improve the production of xylanases for biotechnological applications. However, heterologous gene expression in *E. coli* and *S. cerevisiae* is often limited by differences in the expression systems from the native organism. Different codon usage, missing chaperones and post-translational modifications such as disulphide bridges and glycosylation can cause low expression levels and misfolded proteins that are degraded or driven into inclusion bodies (Bulter *et al*., 2003). These microorganisms however, still remain excellent hosts particularly in the context of protein engineering and gene manipulation. Once stable variants of the enzyme of interest are found, they could be linked to genes that secrete proteins into the surrounding medium or better still, be homologously recombined back into the original fungal host in which the gene coding for the native enzyme has been knocked out (Walsh and Bergquist, 1997). Cloning enzymes with improved properties after directed evolution into both alternative yeast hosts like *P. pastoris* and homologous recombination back into native *T. lanuginosus* fungus should be investigated to improve overall enzyme production.

Before commercialization, an accurate measurement of an enzyme's specific activity is required since pulp bleaching depends on enzyme dosage. Luminase is mass-produced in a yeast expression system and sold for commercial use. SDS-PAGE analysis of Luminase (results not shown) indicated that although the enzyme is impure with six other proteins visible in its profile, the xylanase itself is available in large volumes in the preparation, which compensates for this slight impurity. In future, the most promising xylanase variants should be ideally recloned into a His-Tag vector, purified using suitable columns and possibly further concentrated to obtain a dosage comparable to the commercial xylanase for further application.

Another avenue of interest would be the set up of an automated screening method to detect novel xylanases. This would undoubtedly pave the way towards obtaining the fittest catalyst in a library within a relatively small time frame. The method should ideally be one that could be used, with minor modifications, to detect almost any kind of altered physiological property of interest.

Recombining more parents with diverse traits using the StEP method could create multi-faceted xylanases. Also, elegant and sophisticated new methods of recombination are continuously being developed and these need to be applied to the existing *xynA* mutant library. More time needs to be spent on using SDM and 3-D modelling studies as tools of mutation and catalyst enhancement since knowledge-based mutagenesis saves time, money and human resources. Recombination and random mutagenesis are also vital since they increase the amount of sequence space explored. However, understanding the physiological role (if any), of an amino acid is not easy.

It is important that as the number of applications of directed evolution increases, the growing amount of data is used to learn something about enzyme structure/function relationships. To some extent, the compilation of a database of artificially selected proteins and peptides (ASPD) (Valuev *et al*., 2002) will aid this challenging prospect. A database could, in future, provide a knowledge base for the effects of certain mutations in relation to structural information, thus aiding decisions on mutagenic strategy. Such an effort will require both functional and structural data for the countless enzyme variants being created. Researchers are increasingly using bioinformatics to design more stable proteins that can predict whether certain mutations are feasible or not for a given protein. The 3-D structure of a related xylanase, however, must be available since the programs rely on electrostatic calculations for predictions. Some of these programs include the disulphide bond program, Disulfide by Design (Dombkowski, 2003), used to successfully improve the stability of *B. stearothermophilus* xylanase (Jeong *et al*., 2007), Insight II Molecular modeling software (Accelrys Software Inc.) for improvement of thermal stability of *A. niger* BCC 14405 xylanase (Sriprang *et al*., 2006) and Swiss-PdbViewer (http://www.expasy.ch/spdbv/) (Guex and Peitsch, 1997) for improvement of stability of *T. reesei* xylanase (Turunen *et al*., 2001; Fenel *et al*., 2004).

Most research on genetic engineering of xylanases focuses on tolerance to thermal and alkaline extremes. However, there exists a huge market for application of acid-stable xylanase particularly in the feed, traditional food and wine industries. There should be further investigation into using random mutagenesis methods to obtaining acid-stable xylanases, rather than follow the traditional route of using SDM studies. It would be also be more accurate to use 3-D modelling and binding affinity studies to determine exactly how the dynamics of the enzyme has changed after mutation to render it more active on certain substrates or more stable at physiological extremes. The availability of larger sequence and structural data on proteins would undoubtedly lead to better understanding of protein stability.

The number of enzyme properties that need to be altered to suit industrial process conditions and perhaps a reluctance of chemists to use enzymes may be hampering enzyme commercialization (Schoemaker *et al*., 2003). The disciplines of process engineering, biocatalytic engineering and directed enzyme evolution need to be brought together with the aim of reducing the development time for industrial enzymes. As the computational tools and understanding of enzyme mechanism through dynamics and quantum effects continue to advance, directed evolution may be brought closer to the ultimate goal of rationally designing improved or novel enzyme functions.

Continued growth of the industrial enzyme market is dependent on technological innovation – the identification and characterization of new enzymes from natural sources, the modification of these enzymes for optimal performance in selected applications, and high level expression of the enzyme. Biocatalytic technologies will ultimately gain universal acceptance when enzymes are perceived to be robust, specific and inexpensive. Genomics-based gene discovery from novel biotopes and the broad use of technologies for accelerated laboratory evolution promise to revolutionize industrial catalysis. As high-throughput screening technologies, ever more intelligent assay design and process-conscious research are brought to bear, the utility and economics of enzymes should mandate the use of these mild, selective catalysts.