

**The relationship between structure and
thermostability of a nitrile hydratase from
Geobacillus pallidus RAPc8**

JENNIFER CAROLINE VAN WYK



A dissertation submitted in partial fulfilment of the requirement for the degree
Doctor of Philosophy

Department of Biotechnology
University of Western Cape

Supervisor: Dr. M. F. Sayed
Co-supervisor: Prof. D. A. Cowan

September 2008

**The relationship between structure and thermostability of a nitrile hydratase
from *Geobacillus pallidus* RAPc8**

Jennifer C. van Wyk

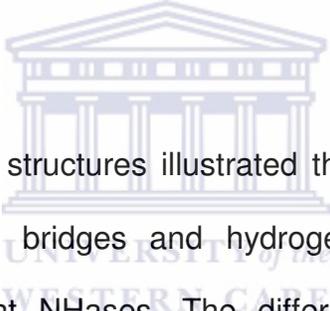
PhD thesis, Department of Biotechnology, University of Western Cape

ABSTRACT

Nitrile hydratases (NHases) are very important biocatalysts for the enzymatic conversion of nitriles to industrially important amides such as acrylamide and nicotinamide. An “ideal” NHase should fulfil several essential criteria including, high substrate conversion rates, being able to tolerate high substrate and product concentrations as well as being highly thermostable. The NHase used in the present study was isolated from *Geobacillus pallidus* RAPc8, a moderate thermophile. The primary aims of this study were to use random mutagenesis to engineer the *G. pallidus* RAPc8 NHase towards improved thermostability and then to use X-ray crystallography to investigate the molecular mechanism(s) involved in the enhanced thermostability.

Two randomly mutated libraries were constructed using MnCl₂ mediated error-prone PCR. The PCR reaction was performed using 0.05 mM and 0.10 mM MnCl₂ and a biased dNTP concentration. The hydroxamic acid assay was used to screen the randomly mutated libraries for NHase mutants with enhanced thermostability. Six mutants that exhibited thermostability-enhancing mutations were isolated from the randomly mutated libraries. The thermostabilised mutants contained between 3

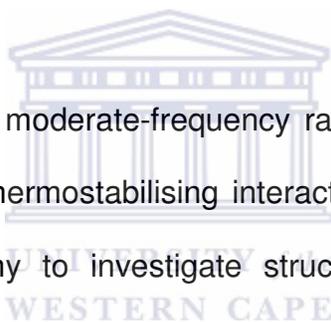
and 7 nucleotide changes per NHase operon. The wild-type and four thermostabilised mutant NHases (7D, 8C, 9C, 9E) were over-expressed, purified, crystallised and subjected to X-ray crystallography. The resolution of the diffraction data for all the mutant NHases were better than the 2.4Å previously obtained for the wild-type *G. pallidus* NHase. The best quality data was collected for mutant 9E, which diffracted to a resolution of 1.15Å. The high quality crystal structures allowed each thermostability-enhancing mutation to be viewed in detail. As most of the NHase mutants contained multiple mutations, the crystal structures were important in correlating the observed thermostabilisation with the structural effect of the mutations.



Analysis of the X-ray crystal structures illustrated the importance of electrostatic interactions, particularly salt bridges and hydrogen bonds in enhancing the thermostability of the mutant NHases. The difference in the free energy of activation of thermal unfolding ($\Delta\Delta G$) was used to compare the wild-type and mutant NHases thermostability. The most improved NHase, mutant 9C, was stabilised by both a buried inter-subunit salt bridge between α R169 and β D218 and an inter-helical hydrogen bond between β K43 and β K50. The stabilisation provided by these electrostatic interactions was 7.62 kJ/mol. Mutant 8C was primarily stabilised by the introduction of an electrostatic network consisting of a salt bridge between β E96 and α R28 and a hydrogen bond between β E96 and β E92. Also, an intra-helical salt bridge between α E192 and α K195 stabilised the helix consisting of α 190-196 in mutant 8C by shielding the helix backbone from solvation and

preventing co-operative unfolding of the α helix. However, mutant 8C was also destabilised by a mutation that disrupted a water-mediated hydrogen bond between β D167 and β K168 at the heterotetramer interface of the enzyme. Consequently, the net stabilisation energy provided as a result of stabilising and destabilising interactions was 6.16 kJ/mol. Mutant 7D was the only NHase mutant with only one possible thermostabilising mutation. This mutant was stabilised by 3.40 kJ/mol as the result of a water-mediated hydrogen bond between α S47 and β E33. Similarly, a water-mediated hydrogen bond between α S23 and β S103 provided a stabilisation energy of 4.27 kJ/mol to mutant 9E.

This project has shown that moderate-frequency randomly mutated libraries can yield mutants with multiple thermostabilising interactions. Also, the importance of utilising X-ray crystallography to investigate structure-function relationships in proteins has been illustrated.



DECLARATION

I declare that “*The relationship between structure and thermostability of a nitrile hydratase from Geobacillus pallidus*” is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete references.



Jennifer C. van Wyk

September 2008

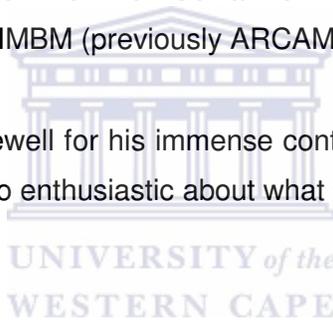
Signed:.....

ACKNOWLEDGEMENTS

Uiteindelik klaar!!!

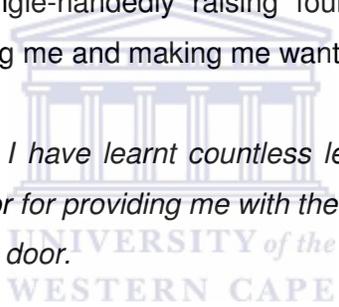
This PhD project was one of the most challenging endeavors I have completed to date and I would not have succeeded without the support of those around me. Therefore, I would like to express my sincere thanks to everyone who have contributed directly or indirectly to the completion of this project.

- Thank you to my supervisor Dr. Muhammed Sayed for giving me the opportunity to complete this project. Thank you for being so accommodating, patient and helpful
- Thank you to my co-supervisor Prof. Don Cowan for his support and patience and also for generously hosting me in IMBM (previously ARCAM) during my stay at UWC
- Thank you to Prof. Trevor Sewell for his immense contribution to the completion of this thesis. Thank you for being so enthusiastic about what you do and inspiring others to be the same.
- Thank you to Prof. Mike Danson for generously hosting me during my stay in Bath and for always being so kind and helpful. Also, thank you to his wife, Janet, for having such a beautiful spirit.
- The National Research Foundation and the University of the Western Cape supported this work. I would like to thank these organizations for their financial contribution to my studies. Also, thank you to the Royal Society (UK) for funding my visit to Bath University.
- Thank you to Dr. Heide Goodman who takes care of everyone who belongs to IMBM and ensures that everyone has what he or she needs to do good science.
- Thank you to William, Lucas, Quinton and Morne for very meaningful discussions that have kept me on the right course. Special thanks to William who came to my aid in my time of need and proofread part of this thesis.



- Special thank you to Rene for dancing into my life and taking me on an amazing adventure. Thank you for being such a loony dude and allowing me to be my loony self.
- Thanks to my three beloved “sisters” in crime Bronnie, Micks and Lynn. It is such a blessing and a privilege to have you in my life. Thank you for being such intelligent, strong, beautiful women; I truly admire all of you. Thanks for the laughter, the shared experiences, the evenings of dining out, movies, and of course cocktails ☺
- Last but certainly not least, I would like to thank my family for their patience and support (and hopefully understanding) while I spent yet another few years not earning my keep. Thank you to my parents who worked hard and sacrificed much to give their children the gift of education. Thank you to my mother for being so strong and brave and beautiful and for almost single-handedly raising four kids. Thanks to my younger siblings for always challenging me and making me want to do better.

Over the course of my studies I have learnt countless lessons that I will keep with me, always. Thank you to my Creator for providing me with the opportunity to walk on this earth and for always opening the next door.



“This thesis is dedicated to the memory of my father, Jan van Wyk, who always believed in me and taught me to believe in myself”

TABLE OF CONTENTS

Page title	Page number
Abstract.....	i
Declaration.....	iv
Acknowledgements.....	v
Table of contents.....	vii
Nomenclature.....	viii
List of tables.....	ix
List of figures.....	x
1. Literature review.....	1
2. Experimental procedures.....	36
3. Screen development and mutant library construction.....	50
4. Kinetic stability of thermostabilised mutant NHases.....	76
5. Structural evaluation of thermostabilised mutants.....	91
6. General discussion.....	126
References.....	143
Appendices.....	177



NOMENCLATURE

A	Activity
Å	angstrom
ΔG	energy of thermodynamic stabilisation
$\Delta\Delta G$	difference in activation energy of the of thermal unfolding
E_a	activation energy
k_d	rate of thermal unfolding
kJ/mol	kilojoule per mol
IPTG	isopropyl β -D-1-thiogalactopyranoside
μl	microlitre
mm	milimeter
mM	milimolar
M	molar
μM	micromolar
NHase	nitrile hydratase
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
SDM	Site-directed mutagenesis
SDS	sodium dodecyl sulphate
w/v	weight per volume



LIST OF TABLES

Table 1.1: Thermostability of mesophilic and thermophilic nitrile hydratases.....	19
Table 1.2: Thermal stability of purified and crude <i>Geobacillus pallidus</i> NHase.....	20
Table 2.1: Bacterial strains.....	37
Table 2.2: Primers used in this study.....	40
Table 3.1: Frequency of active enzymes for EP-PCR libraries.....	60
Table 3.2: Mutation frequency of Lib 0.05 MnCl ₂	69
Table 3.3: Mutation frequency of putative improved thermostable mutants.....	71
Table 3.4: Position of nucleotide (nt) and amino acid (aa) changes in randomly mutated NHases with improved thermostability compared to the wild-type.....	74
Table 4.1: Kinetic parameters for the thermal inactivation of wild-type and mutants NHases at 63°C.....	85
Table 4.2: Kinetic parameters for the thermal inactivation of wild-type and mutant NHases at 65°C.....	89
Table 5.1: X-ray data collection statistics for wild-type and mutant NHases.....	99
Table 5.2: Refinement statistics for wild-type and mutant NHases.....	101

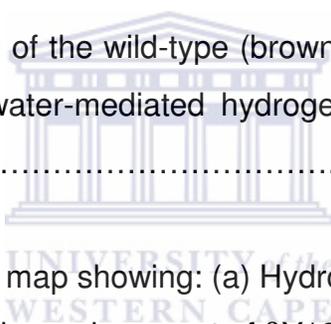
LIST OF SCHEMES AND FIGURES

Scheme 1.1: Enzymatic conversion of nitriles by nitrile hydratase and amidase.....	3
Scheme 3.1: Enzymatic conversion of nitriles to ammonia by nitrile hydratase and amidase.....	53
Scheme 3.2: Enzymatic conversion of nitrile to hydroxamic acid by NHase and amidase.....	55
Figure 1.1: Arrangement of the NHase operons from various microorganisms.....	8
Figure 1.2: Proposed mechanisms for the catalysis of nitriles by NHase.....	10
Figure 1.3: Detailed catalytic mechanism proposed by Mitra and Holz (2006) for nitrile hydratase.....	11
Figure 1.4: The general steps of directed evolution.....	29
Figure 2.1: Schematic of the pNH14K plasmid.....	39
Figure 2.2: The hanging drop vapour diffusion method.....	47
Figure 3.1: Effect of hydroxylamine on activity of NHase cell free extract (CFE).....	56
Figure 3.2: Effect of hydroxylamine on amidase activity.....	58

Figure 3.3: Effect of MnCl ₂ concentration on PCR product yield during error-prone PCR (EP-PCR).	59
Figure 3.4: Effect of thermal inactivation temperature on wild-type nitrile hydratase thermostability.....	63
Figure 3.5: The hydroxamic assay performed in microtitre plate format showing a positive result obtained from screening Lib 0.1 for NHases with improved thermostability.....	65
Figure 3.6: Using the hydroxamic assay to confirm the improved thermostability between the wild-type and mutant NHases.....	67
Figure 3.7: Distribution of nucleotide changes across the operon of thermostabilised NHases.....	74
Figure 4.1: Change in activation energy of a protein from the folded to the unfolded state.....	79
Figure 4.2: The effect of temperature on the extent of thermal inactivation on wild-type and mutant NHases.....	81
Figure 4.3: Thermal inactivation of wild-type and randomly mutated NHases.....	84
Figure 4.4: Comparison of thermal inactivation of wild-type, mutant 9C and site-directed mutant NHases.....	87
Figure 5.1: SDS PAGE analysis the NHase mutants at each step of the purification process.....	94

Figure 5.2: Crystals obtained for the wild-type (WT) and mutant NHases (7D, 8C, 9E, 9C) using the hanging drop vapour diffusion method.....	96
Figure 5.3: Electron density map showing the water-mediated hydrogen bond between α S47 and α E33 in mutant 7D.....	103
Figure 5.4: Three-dimensional structure of mutant 9C showing the position of W202 between α S47 and α E33 on the α 2 β 2 heterotetramer.....	104
Figure 5.5: Three-dimensional structure of mutant 9C showing the positions of α R28, β E96, β V167 and α V188 on the α 2 β 2 heterotetramer.....	105
Figure 5.6: Electron density maps showing (a) the single salt bridge between α SK195 and α E192 as a result of the M \rightarrow V (α 188) mutation and (b) the position of α K195 and α E192 in the wild-type NHase.....	107
Figure 5.7: Electron density map showing the double salt bridge between α R28 and β E96 and the hydrogen bond between β E96 and β E92 in mutant 8C.....	108
Figure 5.8: Electron density map showing the relative positions of β V167 and β K168 in mutant 8C.....	111
Figure 5.9: The NHase heterotetramer interface across the two-fold axis.....	112
Figure 5.10: Three-dimensional structure of mutant 9C showing the positions of the amino acid changes on the α 2 β 2 heterotetramer.....	113
Figure 5.11: Electron density map showing the double salt bridge between α R169 and β D218 in mutant 9C.....	114

Figure 5.12: Electron density map showing the hydrogen bond between β K43 and β K50 in mutant 9C.....	116
Figure 5.13: Electron density map showing β A150 and the surrounding amino acid residues in the local environment of mutant 9C.....	117
Figure 5.14: Three-dimensional structure of mutant 9E showing the position of the amino acid changes on the α 2 β 2 heterotetramer.....	118
Figure 5.15: Electron density map showing β L36 and the surrounding amino acids in the local environment.....	119
Figure 5.16: Superimposition of the wild-type (brown) and mutant (green) electron density maps showing the water-mediated hydrogen bond between α S103 and α S23.....	122
Figure 5.17: Electron density map showing: (a) Hydrogen bond formation of β Y127 with α S47 and α E57 and (b) the replacement of β Y127 with β N127 and filling of the resultant space with water.....	125



CHAPTER 1

LITERATURE REVIEW

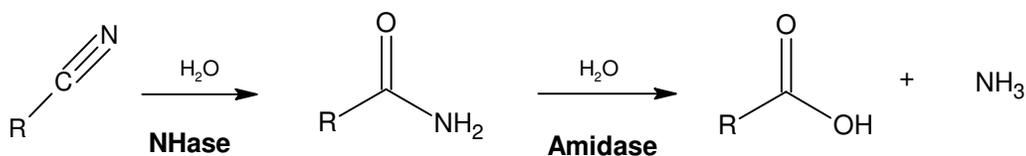
CONTENT

1.1. INTRODUCTION.....	2
1.2. DISTRIBUTION, ISOLATION AND BIOLOGICAL SIGNIFICANCE OF NHASES	4
1.3. MOLECULAR AND FUNCTIONAL CHARACTERISTICS OF NHASES.....	6
1.3.1. Molecular characteristics	6
1.3.2. Crystal structure and reaction mechanism	7
1.3.3. Functional expression in <i>E. coli</i>	12
1.3.4. Substrate specificity.....	13
1.4. APPLICATION OF NHASES.....	15
1.4.1. Industrial application.....	15
1.4.2. Bioremediation.....	16
1.5. THERMOSTABILITY OF NHASES.....	17
1.6. UNDERSTANDING PROTEIN THERMOSTABILITY	21
1.7. ENGINEERING PROTEIN THERMOSTABILITY.....	26
1.7.1. Rational design.....	26
1.7.2. Directed evolution.....	28
1.8. PROJECT OBJECTIVES	32
1.8.1. Rationale.....	32
1.8.2. Aim.....	34

1.1. INTRODUCTION

Since the discovery of nitrilases (EC 3.5.5.1) by Thimann and Mahadevan (1964a), it was generally believed that these enzymes were solely responsible for the enzymatic hydrolysis of nitriles. Nitrilase catalyses the conversion of nitriles to their corresponding carboxylic acid without the formation of an amide intermediate (Thimann and Mahadevan, 1964b). However, surprisingly, researchers reported the presence of amides during nitrilase purification (Hook and Robinson, 1964) and cultivation of microorganisms on aliphatic nitriles (DiGeronimo and Antoine, 1976). This “mystery” was later clarified when Asano *et al.* (1980) discovered nitrile hydratase (NHase; E.C.4.2.1.84) while investigating the microbial conversion of the toxic cyano-group of compounds such as acrylonitrile and acetonitrile.

NHase is one of the enzymes responsible for the bi-enzymatic conversion of nitriles to their corresponding carboxylic acids (Scheme 1.1). The NHase catalyses the conversion of the nitrile substrate to an amide that is further converted to a carboxylic acid by an amidase (E.C.3.5.1.4). The successful application of NHases for the industrial enzymatic production of commodity chemicals such as acrylamide, nicotinamide and 5-cyanovaleramide has generated widespread interest in the synthetic value of these enzymes (Nagasawa and Yamada, 1990; Thomas *et al.*, 2002). Moreover, the industrial production of acrylamide using NHase is the first successful example of the application of a bioconversion process for the production of a commodity chemical (Kobayashi and Shimizu, 1998).



Scheme 1.1: Enzymatic conversion of nitriles by nitrile hydratase and amidase.

Until recently, research has mainly focused on NHases isolated from mesophilic microorganisms, such as the enzymes used industrially for the production of acrylamide (Yamada and Kobayashi, 1996). These enzymes are inherently thermolabile and as a result biotransformations are conducted at low temperatures and require cooling facilities (Duran *et al.*, 1993; Nagasawa and Yamada, 1990; Takashima *et al.*, 1998). The discovery of thermostable NHases (Cramp and Cowan, 1999; Takashima *et al.*, 1998; Yamaki *et al.*, 1997) has broadened the scope of the potential application of these enzymes for organic synthesis. Consequently, the focus has shifted to the isolation, characterisation and exploitation of NHases from thermophilic microorganisms.

The aim of this chapter is to review recent literature pertaining to NHases, particularly thermophilic NHases. The scope of the project also requires a brief and general review of the literature regarding the understanding and engineering of protein thermostability.

1.2. DISTRIBUTION, ISOLATION AND BIOLOGICAL SIGNIFICANCE OF NHASES

Since the initial isolation of NHase from *Norcardia rhodochrous* LL-100-21 (Asano *et al.*, 1980), NHase activity has been found in microorganisms from various genera (e.g. *Agrobacterium*, *Arthrobacter*, *Bacillus*, *Brevibacterium*, *Rhodococcus*, *Pseudomonas*, *Pseudonocardia*) isolated from diverse ecosystems (e.g. shallow marine sediment, deep sea sediments, contaminated soils). Some microorganisms are known to use nitriles as their sole carbon and nitrogen sources (Watanabe *et al.*, 1987; Chapatwala *et al.*, 1990; Linardi *et al.*, 1996; Rezende *et al.*, 2003).

Microorganisms exhibiting NHase activity are readily isolated from environmental samples using selection and enrichment strategies in media containing a nitrile as sole carbon or nitrogen source (Watanabe *et al.*, 1987; Pereira *et al.*, 1998). Samples believed to contain thermophilic microbial isolates are also cultivated at high temperatures (40-80°C). Layh *et al.* (1997) showed that NHases generally have high activity against the nitrile substrate used during the enrichment procedure. Interestingly, this study also showed that the organic solvent used in the enrichment strategy could markedly affect the bioavailability of the substrate to the enzyme, where relatively polar solvents resulted in unsuccessful enrichments while less polar solvents resulted in successful enrichments.

Selection and enrichment strategies rely on the culturability of the microorganisms. Considering that a significant proportion (between 90 and 99%) of microorganisms

is unculturable, the development of culture-independent methods is proving to be increasingly useful in order to isolate novel enzymes (Cowan *et al.*, 2004). Molecular based approaches such as PCR amplification (Precigou *et al.*, 2001; Lourenco *et al.*, 2004) and metagenomic library screening (Liebeton and Eck, 2004) have been used to isolate novel NHases. However, the high sequence identity between NHases might still limit the application of PCR technologies in the search for novel enzymes.

The physiological role of nitrile-converting enzymes in microorganisms was elucidated from studies involving aldoxime metabolism. Aldoximes are intermediates that occur ubiquitously in the environment, particularly in the biosynthesis of plant natural products, and can be enzymatically converted to nitriles by aldoxime dehydratase (Kato *et al.*, 1999). Kato *et al.* (2000) investigated the relationship between aldoxime dehydratase and nitrile degrading enzymes in microorganisms and found that all active strains tested exhibited both aldoxime dehydration and nitrile degrading activities. The role of these enzymes was further corroborated when aldoxime hydratase genes was found to occur in the same gene cluster as nitrile degrading enzymes (Kato *et al.*, 2004; Xie *et al.*, 2003). In addition, Hashimoto *et al.* (2005) reported the occurrence of acetyl-CoA synthase in the aldoxime pathway gene cluster of *Pseudomonas chlororaphis* B23. The aldoxime-nitrile pathway (aldoxime → nitrile → amide → acid → acetyl-CoA) is thought to be important for microorganisms to metabolise aldoximes as a carbon source (Kato and Asano, 2006).

1.3. MOLECULAR AND FUNCTIONAL CHARACTERISTICS OF NHASES

1.3.1. Molecular characteristics

NHases are soluble metalloenzymes and are classified into two main groups based on their cofactor requirements. The iron (Fe) type NHases (Endo *et al.*, 2001; Huang *et al.*, 1997) have a non-heme iron at the catalytic centre while the cobalt (Co) type (Takashima *et al.*, 1998) have a non-corrinoid cobalt atom. An exception is the NHase from *Myrothecium verrucaria* that contains a zinc (Zn) ion in the catalytic centre (Maier-Greiner *et al.*, 1991).

Except for the NHase from *P. chlororaphis*, which is homotetrameric, all known NHases are heteromultimeric and consist of alpha (α) and beta (β) subunits, the sizes of which range from 23kDa to 30kDa (Cowan *et al.*, 1998). Although the subunits are similar in size there is no apparent homology between the α - and β subunits (Cameron *et al.*, 2005). The arrangement of the NHase operons from various microorganisms is shown in Figure 1.1. The α - and β subunits are encoded by two separate, adjacent open reading frames (ORFs). The sequence encoding the α -subunit is generally located upstream of that encoding the β -subunit except for *R. rhodochrous* J1, *Ps. Thermophila*, *Bacillus* sp. BR449, *Geobacillus pallidus* RAPc8 and *Bacillus smithii* where the α -subunit follows the β -subunit (Cowan *et al.*, 2003). All NHases exhibit high protein sequence identity and share a number of conserved regions; specifically they all share the cofactor binding motif CXLCS (known as the claw-setting motif) where X is serine in Fe-type enzymes and threonine in Co-type enzymes (Huang *et al.*, 1997).

For most NHase operons described to date, the amidase sequence is positioned approximately 100 bp upstream of the NHase genes (Cowan *et al.*, 2003). The exception is *R. rhodochrous* J1 where an amidase gene was found 1.9 kb downstream of the α -subunit of the L-NHase gene while no amidase gene was found close to the H-NHase gene (Kobayashi *et al.*, 1992). Rather, the H-NHase operon contained an insertion sequence, IS1164, upstream of the region encoding the α - and β -subunits that may explain the duplication of this operon in *R. rhodochrous* J1. Unlike their mesophilic homologs, amidases associated with the thermophilic *Bacillus* sp. BR449 and *G. pallidus* RAPc8 belong to the nitrilase family of amidases (Cowan *et al.*, 2003).

1.3.2. Crystal structure and reaction mechanism

The first two crystal structures of a NHase were reported for the Fe-type NHases from *Rhodococcus* sp. R312 (PDB accession number 1AHJ/2AHJ), to a resolution of 2.65Å (Huang *et al.*, 1997) and *Rhodococcus* sp. N-771, to a resolution of 1.7Å (Nagashima *et al.*, 1998). Since then, the crystal structures of three thermostable NHases have been described (*Pseudomonas thermophila*, PDB accession number 1IRE, Miyanaga *et al.*, 2001; *B. smithii* SC-J05, PDB accession number 1V29, Hourai *et al.*, 2003; *G. pallidus* RAPc8, Tsekoa, 2005).

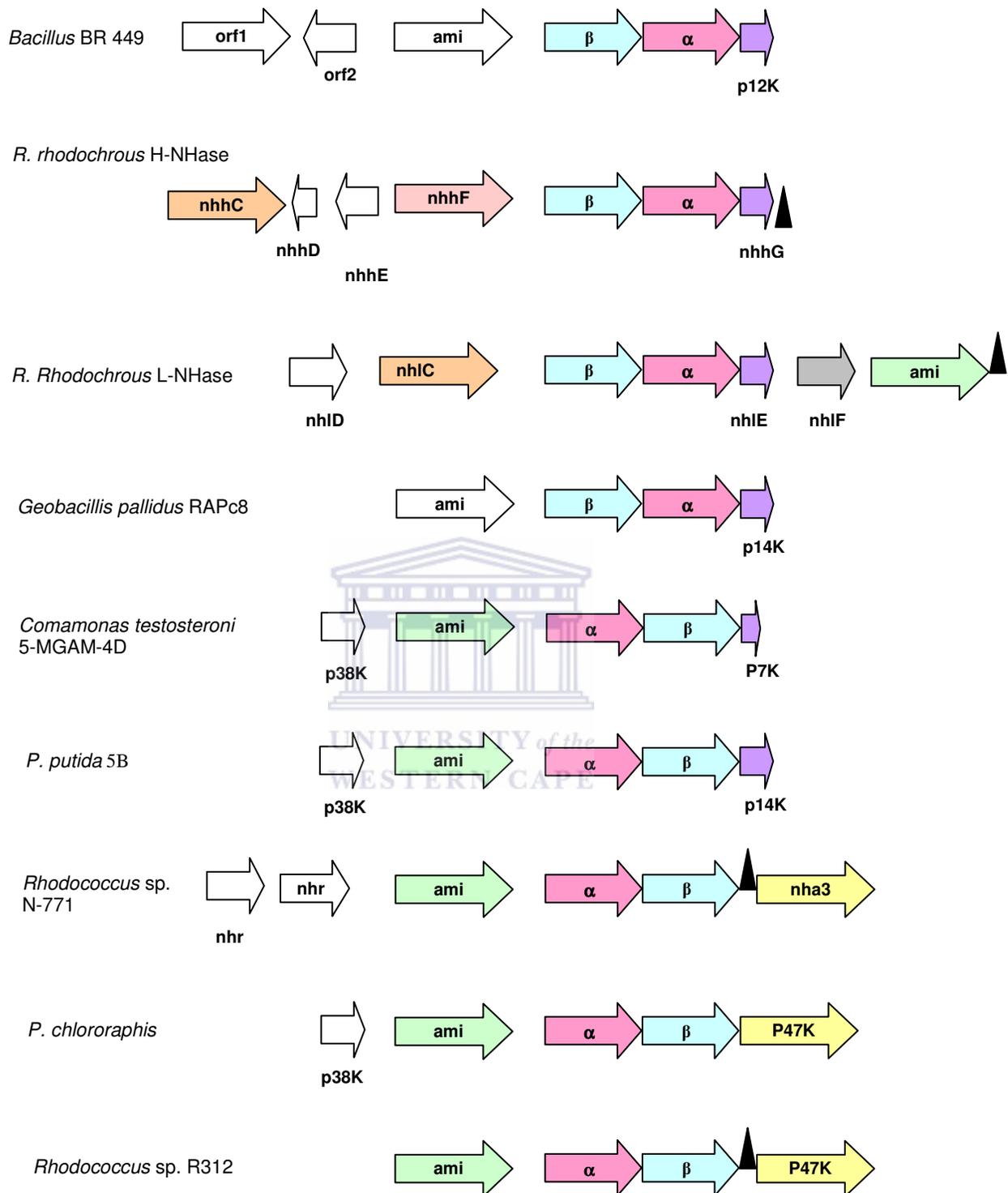
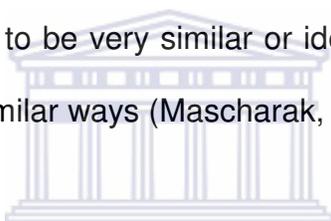


Figure 1.1: Arrangement of the NHase operons mainly consisting of amidase (*ami*), beta subunit (β), alpha subunit (α) and accessory protein (P14K or P47K) from various microorganisms. The arrows represent the direction of the genes; homologous genes are shown by the same colour shading.

The active centre of the *Rhodococcus* NHase is located in a cavity at the interface of the α - and β -subunits (Huang *et al.*, 1997). Three-dimensional analysis of the *Rhodococcus* sp. R312 NHase crystal structure revealed a novel iron centre that is formed by protein ligands from the α subunit only. While performing a structural alignment of the *G. pallidus* RAPc8 NHase with previously solved NHase structures, Tsekoa (2005) found that the co-factor binding motif is identical in all NHases including the Fe-type NHase from *Rhodococcus* sp. R312. Thus, it is believed that perhaps all NHases have very similar structures and thus a similar mechanism of catalysis. The coordination structures of the metal sites of the Fe- and Co-NHase are assumed to be very similar or identical, indicating that the two metal sites function in very similar ways (Mascharak, 2002).



Based on the crystal structure, Huang *et al.* (1997) proposed 3 possible catalytic mechanisms for the hydration of nitriles by NHase. All the mechanisms suggest that the metal ion acts as a Lewis acid as depicted in Figure 1.2. Huang *et al.* (1997) and others favoured the last mechanism (Fig. 1.2 c) that does not involve ligand exchange. This mechanism was supported by kinetic data from the two types of NHases, which illustrated that the Co-type NHase from *R. rhodochrous* sp. J1 and Fe-type NHase from *Brevibacterium* R312 are able to hydrolyze propionitrile at very similar rates under identical reaction conditions (Mascharak, 2002).

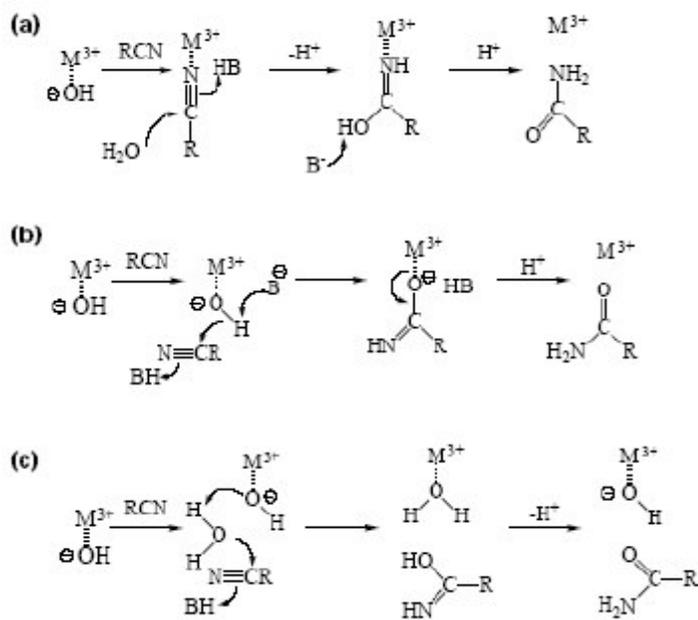
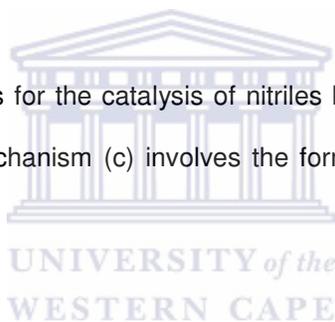


Figure 1.2: Proposed mechanisms for the catalysis of nitriles by NHase. Mechanisms (a) and (b) involve ligand exchange while mechanism (c) involves the formation of an outer sphere complex (Huang *et al.*, 1997)



Recently, Mitra and Holtz (2006) proposed a detailed catalytic mechanism for NHases by combining a comprehensive amount of data from a multitude of studies. They investigated the pH and temperature dependence of kinetic constants k_{cat} and K_m . This data was then combined with previously reported data using complementary methods such as X-ray crystallography, kinetic studies, spectroscopic studies, sequence comparisons and synthetic and theoretical modelling. The details of the model proposed by Mitra and Holz are shown in Figure 1.3. The nitrile substrate directly coordinates to the trivalent metal ion active site resulting in the displacement of the bound water molecule. Following binding of the substrate, an active site water molecule and an active site base are needed for

the reaction to continue (Kovacs, 2004). Sequence analysis of both Fe- and Co-type NHases revealed the presence of the conserved catalytic triad, α Ser112- β Tyr168- β Trp72. These residues are numbered according to the *P. thermophila* JCM 3095 NHase. Mitra and Holz proposed that the catalytic triad residues facilitate NHase activity where α Ser112 deprotonates β Tyr168, thereby providing the base that is needed for the reaction to proceed. Two protons are then transferred, one to the nitrile N-atom and the other to β Tyr168. After proton transfer, the imidate intermediate tautomerises to form an amide, which is subsequently displaced by a water molecule to yield the regenerated biocatalyst. This is the most comprehensive attempt to elucidate the mechanism of NHases.

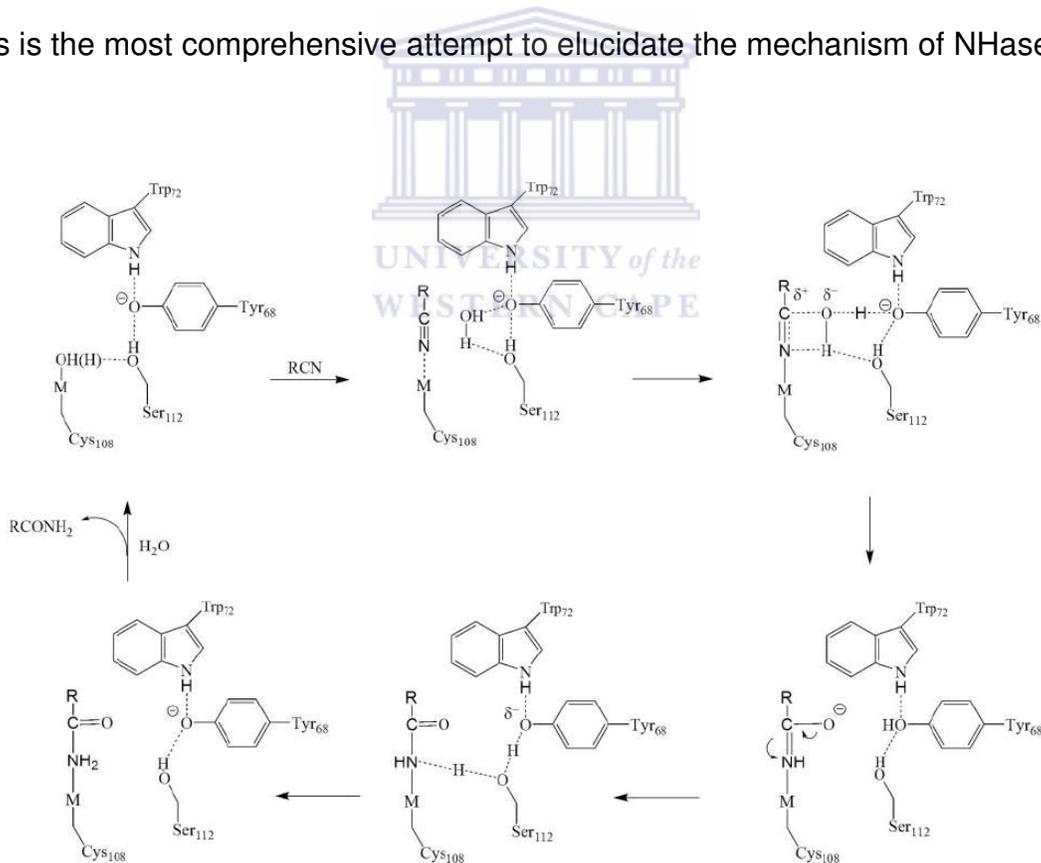


Figure 1.3: Detailed catalytic mechanism for nitrile hydratases as proposed by Mitra and Holz (2006).

1.3.3. Functional expression in *E. coli*

Early studies involving recombinant NHase expression in *E. coli* were mostly unsuccessful mainly due to production of insoluble and inactive inclusion bodies (Ikehata *et al.*, 1989; Kobayashi *et al.*, 1991; Wu *et al.*, 1997). Improvements in expression of soluble and active NHase were obtained by performing protein expression at 30°C or less (Endo *et al.*, 2001). Enhanced over-expression of NHase was achieved with the development of host-vector systems in *R. rhodochrous* (Hashimoto *et al.*, 1992) and the discovery of so-called “activator proteins” (Nishiyama *et al.*, 1991).

The genes encoding activator proteins are generally found in the same gene cluster but downstream of the NHase subunits (Figure 1.1). There is no homology between the activator proteins from Fe-type and Co-type NHase operons (Cameron *et al.*, 2005). Activator proteins associated with Fe-type NHases are 47 kDa proteins that are homologous to the ATP-dependent iron transporter MagA (Nishiyama *et al.*, 1991; Nojiri *et al.*, 1999) whereas those linked to Co-type NHase are small proteins (~14kDa) that are homologous to a portion of the β subunit (Cameron *et al.*, 2005). Petrillo *et al.* (2005) reported the presence of a putative accessory peptide, P7K, which is essential for optimum expression of the *C. testosteroni* NHase. Although this NHase was determined to be thermostable, to date no information is available regarding the co-factor requirement of this enzyme. All the thermostable NHases described thus far require cobalt as cofactor (Cameron *et al.*, 2005).

The role of activator protein in NHase expression remains unclear. Mutational analysis of the activator from *Rhodococcus* sp. N-771 indicated that it is essential for functional NHase expression (Lu *et al.*, 2003). Although the NHase from *Bacillus* sp. BR449 is seemingly unaffected by the co-expression of its associated P12K activator protein (Kim and Oriel, 2000) the 100% amino acid sequence identity between the P12K and P14K protein have led researchers to speculate that the P12K protein is truncated, thereby resulting in an inability to affect NHase expression (Cameron *et al.*, 2005).

Lu *et al.* (2003) identified a proposed metal binding domain CXCC in the sequence of the Fe-type NHase activator. However, no known metal binding motifs have been found in Co-type NHase activators (Cameron *et al.*, 2005). Various studies have reported the absence of the *p14k* transcript indicating that the P14K protein is either expressed at low levels or degraded immediately after the NHase folding (Cameron *et al.*, 2005; Wu *et al.*, 1997). Accordingly, Wu *et al.* (1997) postulated a chaperone or a catalytic role for activator proteins. Cameron *et al.* (2005) proposed that the P14K protein associated with the *G. pallidus* RAPc8 NHase acts as a subunit specific chaperone that mediates folding of the α subunit. Further studies to elucidate the role of activator proteins in NHase expression are in progress.

1.3.4. Substrate specificity

Previously, NHases were generally assumed to have aliphatic substrate specificity but lack aromatic specificity. However, the continued isolation and characterisation

of NHases with aromatic specificity (Meth-Cohn and Wang, 1995; Wieser *et al.*, 1998) and/or stereoselectivity (Fallon *et al.*, 1997; Wu *et al.*, 1997) have eliminated generalizations concerning the substrate specificity of this group of enzymes. For example, the NHase from *R. rhodochrous* AJ270 is able to hydrolyse a broad spectrum of saturated and unsaturated aliphatic nitriles, aromatic nitriles and heterocyclic nitriles (Meth-Cohn and Wang, 1997). Furthermore, whole cells of *R. rhodochrous* AJ270 have been demonstrated to be able to catalyse the stereoselective hydrolysis of α -substituted phenylacetonitriles (Wang *et al.*, 2000) and trans-3-aryl-2,2-dimethylcyclopropanecarbonitriles (Wang and Feng, 2002) to enantiopure amides and/or carboxylic acids. Enantioselective NHases have also been isolated from microorganisms such as *Pseudomonas putida* (Payne *et al.*, 1997), *Agrobacterium tumefaciens* (Bauer *et al.*, 1998), *R. equi* (Přepechalová, 2001) and *Rhodococcus* sp. CGMCC 0497 (Wu and Li, 2002).

The substrate specificities of three thermophilic NHases have been studied in detail. These are (*Geo*) *Bacillus pallidus* DAC 521 (Cowan *et al.*, 1998), *G. pallidus* RAPc8 (Pereira *et al.*, 1998; Cameron, 2003) and *Bacillus smithii* SC-J05-1 (Takashima *et al.*, 1998). Although no substrate specificity data are available for *Bacillus* sp. BR449 NHase, this enzyme is able to catalyse the conversion of acrylonitrile (Padmakumar and Oriel, 1999). The *Pseudonocardia thermophila* NHase may be specific for aliphatic nitriles, although this study only tested four nitrile substrates (Miyanaga *et al.*, 2004).

1.4. APPLICATION OF NHases

1.4.1. Industrial application

The chemical conversion of nitriles has several operational disadvantages. These include the use of harsh reaction conditions, such as strong acidic or basic solutions, elevated temperatures, and the formation of toxic byproducts and large amounts of salt (Banerjee *et al.*, 2002). Nitrile-degrading enzymes are extensively used as biocatalysts in both the chemical and pharmaceutical industries. The most successful industrial application of NHase is in the chemoenzymatic production of acrylamide (Thomas *et al.*, 2002; Zaks, 2001). Using a NHase from *Pseudomonas chlororaphis* B23, Mitsubishi Rayon Company (Kyoto, Japan; previously Nitto Chemical Company), produced 6,000 tons of acrylamide per year in the late 1980's (Nagasawa and Yamada, 1989). This was further improved to more than 30, 000 tons per year by using a NHase from *R. rhodochrous* J1 (Kobayashi *et al.*, 1992).

The biocatalytic manufacture of nicotinamide by Lonza Guangzhou Fine Chemicals is also commercially successful, where *R. rhodochrous* J1 cells was used to achieve 100% conversion of 3-cyanopyridine to the desired amide (Thomas *et al.*, 2002). Another successful commercial example is the regioselective conversion of adiponitrile to 5-cyanovaleramide (5-CVAM), which is an important intermediate in the synthesis of azafenidin (Milestone[®]), a DuPont herbicide (Hann *et al.*, 1999). Using *Pseudomonas chlororaphis* B23 as biocatalyst, 5-CVAM is produced with 93% yield, 96% selectivity and 99.5% reduction in catalytic waste (Thomas *et al.*, 2002). Recent studies have focused on using NHases for the chemo-, regio- and stereoselective synthesis of biologically active molecules (Mylerová and

Martínková, 2003). Studies on the enantioselective hydration of nitriles have focused on α -arylpropionic nitrile (Wu and Li, 2002). This substrate is very important in the synthesis of α -arylpropionic acids, compounds related to profens, which are non-steroidal anti-inflammatory drugs (Rieu *et al.*, 1986).

1.4.2. Bioremediation

Due to its widespread use as a synthetic starting material, as an extraction solvent and as a mobile phase solution in liquid chromatography, acetonitrile is often found in industrial and laboratory effluents (Kohyama *et al.*, 2006). Since acetonitrile is extremely toxic, it is important these compounds are removed from wastes before they are released into the environment. Several studies have attempted to develop microbial treatment methods for the detoxification of acrylonitrile-containing wastes (Wyatt and Knowles 1995; Battistel *et al.*, 1997; Kohyama *et al.*, 2006). This objective has proved to be difficult as such wastes often contain high levels of acrylonitrile and other toxic components that inhibit microbial growth (Wyatt and Knowles 1995; Battistel *et al.*, 1997). However, researchers have managed to overcome these limitations by using mixed microbial consortia.

Wyatt and Knowles (1995) isolated a number of microorganisms that were able to grow on the organic compounds associated with acetonitrile manufacture. The individual isolates were first acclimatised by exposure to a synthetic effluent containing increasing concentrations of the various components. The resultant stable mixed microbial population was able to degrade a highly toxic effluent

consisting of acrylonitrile, fumaronitrile, succinonitrile, acrylic acid, acrylamide, acrolein, cyanopyridine and maleimide. Immobilization of nitrile-degrading organisms eliminates the need for acclimitisation and allows the bacteria to be recycled (Battistel *et al.*, 1997).

Kohyama *et al.* (2006) successfully used nitrile-degrading microorganisms to construct a waste treatment process for the detoxification of high concentrations of acetonitrile. Resting cells of *Rhodococcus pyridinivorans* S85-2 and *Brevundimonas diminuta* AM10-C-1 were used in tandem for the conversion of acetonitrile and acetamide, respectively. Using this strategy, the conversion of 6M acetonitrile to acetic acid at a rate of greater than 90% in 10 hours was achieved. Recently, Li *et al.* (2007) used an enrichment strategy to develop a microbial consortium that was able to degrade organic nitriles. The mixed microbial culture was able to degrade acetonitrile, acrylonitrile and benzonitrile in a batch bioreactor under standard conditions (25°C, pH7.0).

1.5. THERMOSTABILITY OF NHASES

Microorganisms are divided into three main groups according to their temperature range of growth: psychrophiles (-5 to 20°C), mesophiles (15 to 40°C) and thermophiles (> 40°C). Thermophiles are further subdivided into moderate thermophiles with a maximum growth temperature of 60°C; extreme thermophiles

with a maximum growth temperature of 75°C and hyperthermophiles with a temperature range of between 80 and 121°C.

Only five NHases from moderate thermophiles have been described in detail. The first was isolated from the actinomycete, *Pseudonocardia thermophila* JCM 3095 (Yamaki *et al.*, 1997). Subsequently, two NHases were isolated from moderate thermophiles belonging to the genus *Bacillus* including *Bacillus smithii* (Takashima *et al.*, 1998) and *Bacillus* sp. BR449 (Padmakumar and Oriel, 1999). Another two moderately thermostable NHases were also isolated from the genus *Geobacillus* including, *G. pallidus* RAPc8 (Pereira *et al.*, 1998) and (*Geo*) *Bacillus pallidus* DAC 521 (Cramp and Cowan, 1999). The thermostability of selected mesophilic and thermophilic NHases is summarised in Table 1.1. The thermostabilities of most NHases are consistent with the environment from which they were isolated. However, Petrillo *et al.* (2005) have reported the isolation of a moderately thermostable NHase from *Comamonas testosteroni* 5-MGAM-4D, a mesophilic microorganism. This NHase retains 83% activity after 1 hour incubation at 50 °C.

Table 1.1: Thermostability of mesophilic and thermophilic nitrile hydratases

MICROORGANISM	TEMPERATURE (°C)	INCUBATION TIME	% RESIDUAL ACTIVITY	REFERENCE
Mesophilic				
<i>P. chlororaphis</i>	20	10 min	100	Nagasawa and Yamada, 1997
	35	10 min	53	
<i>P. putida</i>	50	10 min	0	Fallon <i>et al.</i> , 1997
<i>P. putida</i>	50	20 min	60	Payne <i>et al.</i> , 1997
<i>C. testoteroni</i>	50	30	90	Petrillo <i>et al.</i> , 2005
5-MGAM-4D	50	60	83	
<i>Rhodococcus</i> sp. N-774	30	30 min	100	Nagasawa and Yamada, 1995
<i>R. rhodochrous</i> J1 H-NHase	50	30 min	100	Nagasawa <i>et al.</i> , 1991
	60	1 h	50	Kobayashi <i>et al.</i> , 1992
<i>R. rhodochrous</i> J1 L-NHase	30	30 min	100	Wieser <i>et al.</i> , 1998
Thermophilic				
<i>Bacillus</i> sp. BR449	60	2 h	100	Padmakumar and Oriel, 1999
<i>G. pallidus</i> RAPc8	50	2.5 h	50	Pereira <i>et al.</i> , 1998
	60	16 min	50	
<i>B. pallidus</i> DAC521	50	51 min	50	Cramp and Cowan, 1999
	60	7 min	50	
<i>Bacillus smithii</i>	55	1.5 h	50	Takashima <i>et al.</i> , 1998
<i>Ps. thermophila</i>	50	2 h	100	Yamaki <i>et al.</i> , 1997
	60	2 h	90	

It has been noted that Co-type NHases are generally more thermostable than Fe-type NHases (Payne *et al.*, 1997) and that all thermostable NHases described to date have cobalt as a metal co-factor (Cameron *et al.*, 2005). The reason for this difference in enzyme thermostability remains unknown. Crude extracts of thermophilic *G. pallidus* RAPc8 NHase were significantly more thermostable than the purified enzyme (Table 1.2; Cowan *et al.* 1998) suggesting that the intracellular environment may be important in providing extrinsic stabilisation. Nakasako *et al.* (1999) showed that hydration water molecules play an important role in stabilising the tertiary and quaternary structure of the *Rhodococcus* sp. N-771 NHase. Hydration water molecules are thought to have a stabilising influence on protein structures by extending the networks of hydrogen bond formation (Meyer, 1992).

Table 1.2: Thermal stability of purified and crude *Geobacillus pallidus* NHase (Cowan *et al.*, 1998)

Temperature in °C	Activity half life (purified enzyme)	Activity half life (crude enzyme)
30	7.0 hours	120 hours
37	-	67 hours
50	-	4.5 hours
55	51 minutes	-
60	6.8 minutes	8.2 minutes

Miyanaga *et al.* (2001) suggested that increased subunit interactions of the Co-type NHase compared to the Fe-type may contribute to the increased thermostability of Co-type NHases. Increased subunit interaction is well-documented to be one of the factors involved in enhancing the stability of thermophilic proteins (Bjørk *et al.*, 2004; Hough and Danson, 1999; Jaenicke and

Böhm, 1998). Tsekoa (2005) noted that the most obvious structural difference between the thermolabile *Rhodococcus* sp. R312 Fe-type NHase and the thermostable *G. pallidus* RAPc8 Co-type NHase was the presence of an additional α helix comprising amino acids 116-126 of the β subunit. Tsekoa suggested that this extra helix may be a stabilising influence by providing an additional interaction between cognate dimers in the heterotetrameric structure. Increased helical content has been shown to contribute to protein thermostability (Kumar *et al.*, 2000).

1.6. UNDERSTANDING PROTEIN THERMOSTABILITY

Thermostable enzymes are becoming increasingly important as industrial biocatalysts (Cowan, 1996; Haki and Rakshit, 2003; Polizzi *et al.*, 2007). There are several economic advantages to performing a bioprocess at higher temperatures, including increased reaction rates, increased solubility of reactants and reduced microbial contamination (Van den Berg and Eijsink, 2002). The use of thermostable enzymes may also result in decreased enzyme turnover, which improves the economic feasibility of the industrial process (Eijsink *et al.*, 2004).

The potential exploitation of various thermophilic enzymes such as amylases, chitinases, xylanases, and proteases as industrial biocatalysts has been subject to extensive review (Haki and Rakshit, 2003; Turner *et al.*, 2007). The continued isolation and characterisation of thermophilic and hyperthermophilic enzymes and

their resultant application in industrial biocatalysis has generated considerable interest in understanding the molecular basis of protein stability (Sternier and Liebl, 2001; Van den Burg, 2003). In this regard, numerous studies have been conducted to identify and understand the mechanisms involved in conferring stability to thermostable proteins (reviewed by Vieille and Zeikus, 2001; Yano and Poulos, 2003). These studies have focused primarily on comparing sequence and structural data of thermophilic enzymes with their mesophilic counterparts.

The amino acid composition of a protein has long been considered an important factor in the thermostability of proteins (Argos *et al.*, 1979; Querol *et al.*, 1996; Fields, 2001). The first statistical comparison of the amino acid composition of mesophilic and thermophilic proteins revealed a preference toward substitutions such as Gly→Ala, Asp→Glu and Lys→Arg (Argos *et al.*, 1979). Rapid advances in bioinformatics and genomics has generated huge quantities of gene and genome sequence, which has allowed for large-scale statistical comparisons of mesophiles and thermophilic enzymes in order to determine the occurrence and relative distribution of particular amino acids or amino acid groups in these groups of proteins (Fields, 2001).

For example, Haney *et al.* (1999) compared the genome of one hyperthermophilic archaeon, *Methanococcus jannaschii*, with that of four mesophilic *Methanococcus* species. This allowed for the direct sequence alignment of 115 proteins. A comprehensive pairwise comparison of all 20 amino acids, which examined more than 7000 amino acid replacements, showed 26 pairs of significant amino acid

changes between the mesophilic and the thermophilic proteins. These changes were categorised into four main groups: a decrease in the number uncharged polar residues; an increase in the number of charged residues; an increase in the hydrophobicity of residues; and an increase in residue volume. Querol *et al.* (1996) analysed 195 different amino acid substitutions, reported to confer enhanced thermostability to proteins in order to relate protein conformational characteristics with thermostability.

It difficult to identify or predict amino acid changes that confer thermostability because even closely related proteins can achieve similar thermostability levels by different mechanisms: e.g., by different amino acid replacements (Yano and Poulos, 2003). Moreover, sequence differences between thermophilic and mesophilic proteins may not necessarily be the result of thermal adaptation (Acharya *et al.*, 2004). Also, despite the high sequence homology between thermophilic and mesophilic proteins, there is a significant variation in the overall amino acid distribution between these types of proteins (Kumar *et al.*, 2000). Therefore, the importance of amino acids in protein thermostability is probably more related to the distribution and interaction of amino acid residues in the protein structure rather than the amino acid composition of the protein (Veille and Zeikus, 2001). For example thermitase from *Thermoactinomyces vulgaris* and subtilisin BPN9 from *Bacillus amyloliquefaciens* are homologous and contain the same number of charged residues. However, the thermophilic thermitase contains eight more ionic interactions, which probably facilitates the increased thermostability of this enzyme (Teplyakov, 1990).

Depending on its conformational state within the protein, one amino acid can be associated with several different thermostabilising features. Therefore, structural analysis is necessary to elucidate the role of the amino acid in protein thermostability (Nakai *et al.*, 1988). It is widely accepted that protein stability is dependent on the structural properties rather than the sequential properties of amino acids (Pack and Yoo, 2004). In this regard, various studies have attempted to correlate the distribution of amino acids on the protein structure with thermostability. Kumar *et al.* (2000) performed a statistical analysis of various factors previously identified to enhance protein thermostability. The aim of this study was to identify systematic differences between thermophilic and mesophilic proteins across 18 different protein families. They found that the majority of thermophilic proteins have more salt bridges and side chain-side chain hydrogen bonds. In addition, Arg and Tyr are more frequent while Cys and Ser is less frequent in thermophilic proteins.

In another study, Pack and Yoo (2004) investigated the relationship between the three-dimensional structure and distribution of amino acid across 20 pairs of thermophilic and mesophilic proteins and found that protein thermostability depends on the residual structure state of the amino acid; i.e., whether the amino acid is buried or exposed. Specifically they found less partially buried Ser residues, less exposed and more buried Ala residues, more buried Glu residues and more exposed Arg residues. Although both these studies elucidated trends regarding the

contribution of amino acids to protein thermostability, no single factor showed a consistent trend across all protein families.

Several features have been identified to confer increased thermostability to proteins (Jaenicke and Bøhm 1998, Querol *et al.*, 1996). These are: increased hydrophobicity (Haney *et al.*, 1997), better packing, deletion or shortening of loops (Russell *et al.*, 1997), a reduced number of labile amino acids such as cysteine, asparagine and glutamine (Russell *et al.*, 1997; Yano *et al.*, 2003), increased proline content (Bogin *et al.*, 1998), increased electrostatic interactions by salt bridges or networks (Pappenberger *et al.*, 1997; Hakamada *et al.*, 2001; Torrez *et al.*, 2003), increased hydrogen bonding (Vogt *et al.*, 1997), increased aromatic interactions (Anderson *et al.*, 1993; Puchkaev *et al.*, 2003), an increase in metal binding capacity (Kataeva *et al.*, 2003), α -helix stabilization (Eijsink *et al.*, 1992; Forood *et al.*, 1993), higher packing efficiency (Li *et al.*, 1998) and increased oligomerisation (Salminen *et al.*, 1996).

Protein stability results from the complex balance of opposing forces (Jaenicke *et al.*, 1996) and thermophilic enzymes show only marginal increases in the free energy of stabilisation compared to their mesophilic homologs (Jaenicke and Bøhm, 2000). Although studies involving sequence and structural comparisons have elucidated various factors that contribute to increased thermostability there are still no general rules that govern the structural basis of enzyme thermostability (Hough and Danson, 1999; Song and Rhee, 2000). The only real consensus

among these studies are that stabilisation mechanisms involve all levels of the protein structural hierarchy from the amino acid sequence to the quaternary structure (Sterner and Liebl, 2001).

1.7. ENGINEERING PROTEIN THERMOSTABILITY

1.7.1. Rational design

Rational approaches used for thermostability engineering typically involve a comparison of either the amino acid sequence or the three-dimensional structure of the protein of interest with that of a more thermostable, homologous counterpart. Promising amino acid exchanges are then verified by replacement of these amino acids in the mesophilic homologue by site-directed mutagenesis (Eijsink *et al.*, 2004). By comparing the amino acid sequences of homologous enzymes of differing thermostabilities, Van den Burg *et al.* (1998) engineered a thermolysin-like protease from *Bacillus stearothermophilus* (TLP-ste) to be 8-fold more thermostable than the wild-type and to be functional at 100°C. The stabilising amino acid replacements were identified in a previous study where residues in the TLP-ste were replaced by the corresponding amino acid in thermophilic thermolysin (Veltman *et al.*, 1996).

The combined application of sequence homology and structural alignment has also proven successful for rational design. For example, Vetriani *et al.* (1998) used homology modeling and direct structural comparison to show that the hexameric glutamate dehydrogenase from the hyperthermophile *Pyrococcus furiosus*

contained a more extensive ion-pair network than the less stable enzyme variant from *Thermococcus litoralis*. Furthermore, by altering two amino acid residues, the less stable enzyme from *T. litoralis* was engineered to be four-fold more stable at 104°C than the wild-type enzyme. Using a similar strategy, Jiang *et al.* (2001) increased the stability and melting temperature (T_m) of the human Yes-associated protein WW domain by 10.5 kJ/mol and 28°C, respectively.

Although site-directed mutagenesis studies do not provide the protein engineer with a general strategy regarding the mechanisms of protein stability, the information gained can prove useful for the rational design of protein thermostability. For example, Reddy *et al.* (1998) analysed protein structural data from a set of 304 non-homologous protein structures in order to understand the effect of amino acid substitutions on the folding and stability of proteins. All amino acids substitutions studied were previously reported to affect protein stability. The effect of each mutation was correlated with the local structural environment of the amino acid. This information can potentially be used to indicate whether a region in the protein structure is unsuitable for engineering stability and also suggest stabilising point mutations for a given protein.

Although rational design has been instrumental in understanding the mechanism of protein unfolding and stability of some well-studied proteins such as thermolysin-like proteases (Imanaka *et al.*, 1986; Van den Berg *et al.*, 1998; Eijsink *et al.*, 2001), tetrameric malate dehydrogenases (Bjørk *et al.*, 2003; Bjørk *et al.*, 2004) and α -amylases (Declerck *et al.*, 2003; Machius *et al.*, 2003), this technique has

not provided a general strategy to improve the stability of proteins. Rational design is limited in its application, as it requires knowledge of both protein sequence and structure. Also, knowledge of the thermal inactivation process of the specific protein could be beneficial as many site-directed mutagenesis experiments aimed at altering protein stability fail because researchers do not target the areas on the protein that are critical for the inactivation process (Eijsink *et al.*, 2005). Due to the limitations of rational design, its successful application probably requires a complementary approach that combines this technique with other protein engineering strategies.

1.7.2. Directed evolution

Directed evolution or evolutionary protein engineering has emerged as a very useful method for the generation of enzymes with improved properties for industrial applications (Woodyer *et al.*, 2004). The general techniques of directed evolution mimic natural evolution processes such as random mutagenesis and sexual recombination to create genetic diversity (Arnold *et al.*, 2001). Unlike rational design, directed evolution does not require previous knowledge about enzyme structure or function (Woodyer *et al.*, 2004). Figure 1.4 illustrates the general steps involved in a directed evolution experiment (Tao and Cornish, 2002). The gene encoding the protein of interest is randomly mutated resulting in a library of mutant genes, each of which is cloned, expressed and then subjected to selection or screening to identify proteins with the desired property.

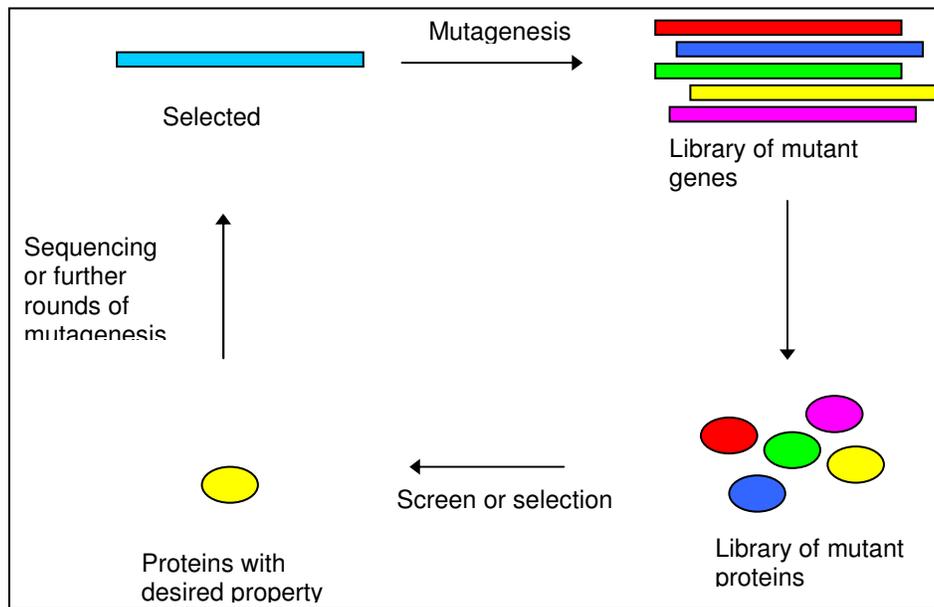


Figure 1.4: General steps of directed evolution (redrawn from Tao and Cornish, 2002)

Numerous methods have been used to generate genetic diversity (Bloom *et al.*, 2005; Neylon, 2004). Some are used to directly create sequence diversity, either by random point mutations or by introducing controlled insertions or deletions (Neylon, 2004). Random point mutations can be introduced by chemical mutagens, UV radiation, error-prone PCR (EP-PCR) and mutator strains (Woodyer *et al.*, 2004). Controlled randomisation involves the random insertion and deletion of sequences of defined length at random positions in the target sequence (Neylon, 2004). Methods that indirectly create diversity include recombination techniques such as DNA shuffling (Stemmer, 1994) and staggered extension process (Zhao *et al.*, 1998). Methods such as random mutagenesis, targeted mutagenesis and recombination remain fundamentally important strategies for directed evolution

experiments. However, continued research has resulted in interesting variations and modifications of these methods (Lutz and Patrick, 2004) or the development of new methods (Farinas *et al.*, 2001).

The simplest directed evolution strategy is to introduce random point mutations into the gene of interest (Eijsink *et al.*, 2005). Among the many random mutagenesis methods used, EP-PCR is by far the most common (Woodyer *et al.*, 2004). EP-PCR, first described by Leung *et al.* (1989) and later modified by Cadwell and Joyce (1992), is used to introduce random point mutations during each cycle of the PCR reaction. The aim is to reduce the fidelity of the polymerase without significantly affecting the level of PCR amplification. Due to its high inherent error frequency ($\sim 10^{-3}$ per nucleotide over 20-25 cycles of amplification), *Taq* polymerase is a suitable polymerase for EP-PCR studies (Cadwell and Joyce, 1994). However, this error frequency is insufficient to produce a diverse mutant library and the fidelity of the *Taq* polymerase is further reduced by the addition of manganese ions, an increased magnesium chloride concentration and by using a biased dNTP concentration in the PCR reaction (Cadwell and Joyce, 1992).

The success of a directed evolution experiment primarily depends on the availability of a screening or selection method to identify variants with the desired property (Zhao and Arnold, 1997). Selection and screening technologies limits the search for the desired variant as typically they cover no more than 10^8 and 10^4 variants, respectively (Arnold, 1998). It is very important that the selection or screening methodology is sensitive and specific for the desired property of the enzymes (i.e. 'You get what you screen for': You and Arnold, 1996). Consequently,

all screening and selection strategies aim to exploit the link between the gene, the enzyme it encodes and the product that results from enzyme activity (Aharoni *et al.*, 2005).

Optimum selection methodologies require that the activity of the target enzyme is linked to the growth and/or survival of the host microorganism (Bommarius *et al.*, 2006). Although the thermophilic microorganism *Thermus thermophilus* has been employed in selection studies at higher temperatures (Akunama *et al.*, 1998; Tamakoshi *et al.*, 2001), selection is rarely used to investigate protein stability (Bommarius *et al.*, 2006). Screening for protein stability usually involves measuring residual enzyme activity after exposure to denaturing conditions (Eijsink *et al.*, 2005). However, the development of high-throughput methods such as measuring the denaturation curves of proteins in 96-well plate format is in progress (Aucamp *et al.*, 2005).

The observation that naturally occurring thermophilic enzymes are not as catalytically active as their mesophilic homologs have led researchers to speculate that thermal adaption come at the cost of catalytic activity. Similarly, studies have suggested that an enzyme cannot be both thermostable and optimally catalytically active (Shoichet *et al.*, 1995). More recent studies have indicated that it is possible to enhance stability without sacrificing catalytic activity (Giver *et al.*, 1998; Zhao and Arnold, 1999). Indeed, Giver *et al.* (1998) and Strausberg *et al.* (2005) showed that both properties can be improved simultaneously, although this is extremely rare.

Directed evolution has been used successfully to enhance the thermostability of both mesophilic (Akanuma *et al.*, 1998; Giver *et al.*, 1998) and psychrophilic (Miyazaki and Arnold, 1999; Miyazaki *et al.*, 2000) enzymes. Sakaue and Kajiyama (2003) isolated several thermostable fructosyl-amino acid oxidase variants by using the mutator strain, *E. coli* XL1-Red. The most thermostable mutant contained 5 amino acid changes and was stable at 45°C while the wild-type was not stable above 37°C. Using a combination of EP-PCR and DNA shuffling, Niu *et al.* (2006) improved the optimum temperature of a lipase from *Rhizopus arrhizus* by more than 10°C. The increased thermostability of the lipase was attributed to only one of the three observed amino acid replacements. The directed evolution approach is useful in that it does not require knowledge about protein sequence and structure. Using this approach it is possible to stabilise a protein in ways that are not necessarily easy to rationalise. However, rational design can still provide valuable input as it limits the library size (one of the major constraints of the directed evolution approach) and hence increases the success of the experiment (Eijsink *et al.*, 2005). Therefore, a complementary approach that encompasses different techniques is likely to be more rewarding.

1.8. PROJECT OBJECTIVES

1.8.1. Rationale

One of the primary difficulties associated with the industrial application of biocatalysts is their inherent instability at high temperature (Yano and Poulos,

2003). The enzymatic industrial production of acrylamide is conducted at low temperature to minimise enzyme turnover (Nagasawa and Yamada, 1990). The isolation or engineering of highly thermostable NHases would increase the industrial value of these catalysts and aid in the application of NHases for the synthesis of high value chemicals. However, an “ideal” NHase would in addition to being highly thermostable, also fulfill other essential criteria such as high substrate conversion rates and being able to tolerate high substrate and product concentrations.

The NHase used in the present study was isolated from *Geobacillus pallidus* RAPc8, a moderate thermophile. Although the structure of a number of NHases, including the one used in the present study, has been elucidated, most studies have focused on investigating the reaction mechanism (Huang *et al.*, 1997), catalytic activity (Peplowski *et al.*, 2007; Takarada *et al.*, 2006) and substrate specificity (Tsekoa, 2005) of these enzymes. To date, no study has attempted to utilize the crystal structures to understand the molecular basis of NHase thermostability, either by investigating the relationship between protein sequence and structure or by engineering proteins with improved thermostability.

Considering our limited understanding of the thermostability of NHases and the number of different mechanisms associated with the thermostability of proteins in general, it is probably inappropriate to attempt to engineer a NHase with increased thermostability using rational design. Directed evolution has been used successfully to improve the thermostability of various enzymes such as esterases

(Giver *et al.*, 1998), fructosyl-amino acid oxidases (Sakaue and Kajiyama, 2003), maltogenic amylases (Kim *et al.*, 2003), and lipases (Acharya *et al.*, 2004). On this basis, directed evolution was chosen as suitable method to create a NHase with increased thermostability compared to the wild-type enzyme. This approach also offers a route to understanding the molecular basis of protein thermostability in this enzyme class.

1.8.2. Aim

The primary aim of this project was to use protein engineering coupled with X-ray crystallography to investigate the relationship between structure and thermostability of a nitrile hydratase isolated from *Geobacillus pallidus* RAPc8, a moderate thermophile. The integral components include:

- i. The generation a randomly mutated library of the nitrile hydratase gene using EP-PCR
- ii. The identification of a suitable enzyme activity assay that could be used to screen the library for a NHase with increased thermostability
- iii. The use of x-ray crystallography to investigate the relationship between enzyme structure and thermostability

CHAPTER 2

EXPERIMENTAL PROCEDURES

CONTENT

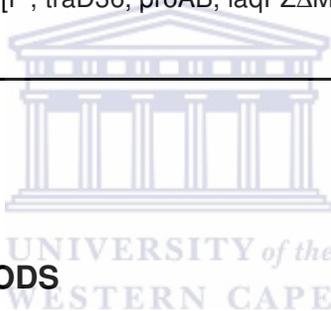
2.1. BACTERIAL STRAINS.....	37
2.2. ANALYTICAL METHODS	37
2.2.1. Protein determination	37
2.2.1.1. <i>Bradford assay</i>	37
2.2.1.2. <i>Fluorometric quantitation of protein</i>	37
2.2.2. SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)	38
2.3. RANDOM MUTAGENESIS USING ERROR-PRONE PCR (EP-PCR)	38
2.4. MUTANT LIBRARY CONSTRUCTION	39
2.5. SITE-DIRECTED MUTAGENESIS	41
2.6. OPTIMISATION OF THE HYDROXAMIC ACID ASSAY	41
2.7. SCREENING FOR IMPROVED THERMOSTABLE NHASE MUTANTS	43
2.7.1. Determination of thermal inactivation temperature	43
2.7.2. Cultivation and expression of mutant library in microtitre plates	43
2.7.3. Screening for improved thermostable variants	44
2.7.4. Kinetic stability analysis of improved thermostable variants	44
2.8. PROTEIN EXPRESSION AND PURIFICATION	45
2.8.1. Cell-free extract (CFE) preparation	45
2.8.2. Ammonium sulphate precipitation	46
2.8.3. Hydrophobic interaction chromatography (HIC)	46
2.8.4. Anion-exchange chromatography.....	46
2.9. STRUCTURE DETERMINATION USING X-RAY CRYSTALLOGRAPHY.....	47
2.9.1. Crystallisation	47
2.9.2. X-ray data collection and processing	48
2.9.3. Structure solution and refinement	49
2.9.4. Molecular graphics and structural analysis	49

2.1. BACTERIAL STRAINS

The *Escherichia coli* strains used in this study are listed in Table 2.1.

Table 2.1: *E. coli* strains

Bacterial strain	Genotype	Supplier
BL21 (DE3)	$F^- ompT hsdS_B(r_B^- m_B^-) gal dcm$ (DE3)	Novagen
BL21 (DE3) pLysS	$F^- ompT hsdS_B(r_B^- m_B^-) gal dcm$ (DE3) pLysS (Cam ^R)	Novagen
JM109 (DE3)	<i>endA1, recA1, gyrA96, thi, hsdR17</i> (r_k^-, m_k^+), <i>relA1, supE44</i> , $\Delta(lac-proAB)$, [F', traD36, proAB, laqI ^q Z Δ M15], λ (DE3)	Promega



2.2. ANALYTICAL METHODS

2.2.1. Protein determination

2.2.1.1. Bradford assay

Protein concentration measurements were performed using a modified version of the Bradford assay as described by Bradford (1976). Sigma reagents were used with bovine serum albumin as standard.

2.2.1.2. Fluorometric quantitation of protein

Protein concentration was also determined fluorometrically using the Quan-iT protein assay kit (Invitrogen) as described by the manufacturer.

2.2.2. SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Samples from all stages of protein purification were analysed with SDS-PAGE as previously described by Laemmli (1970). Typically, 12 or 15% acrylamide gels were prepared and samples were electrophoresed at a constant current of 15 mA using Hoefer SE 250 minigel electrophoresis unit. Gels were then stained with a solution containing Coomassie blue R-250 in 40% methanol and 10% acetic and destained in a 10% acetic acid solution. The rate of staining and destaining was increased by heating the solutions at medium setting in a 660-Watt microwave.

2.3. RANDOM MUTAGENESIS USING ERROR-PRONE PCR (EP-PCR)

EP-PCR was carried out using a modified version of the method described by Wintrobe *et al.* (2001). The pNH14K plasmid was used as template for the PCR reaction. The Nhop2 and pNH14k5'R primers contained *Nde* I and *Not* I restriction enzyme sites, respectively, to facilitate directional cloning of the mutant library. The reaction mixture contained: 10 mM Tris-HCl, 7 mM KCl, 0.01% (w/v) gelatin, 0.05-0.15 mM MnCl₂, 0.2 mM dATP, 0.2 mM dGTP, 1 mM dCTP, 1 mM dTTP, 50 pmol of each primer, 10 ng pNH14K. A controlled PCR that contained all the components of the PCR reaction except the template DNA was also performed. The PCR was conducted under the following conditions: an initial denaturation step at 98 °C for 30 seconds followed by 30 cycles of amplification using the following conditions: denaturation at 98 °C for 10 seconds, annealing at 65 °C for 30 seconds, extension at 72 °C for 4 minutes, final extension at 72 °C for 10 minutes.

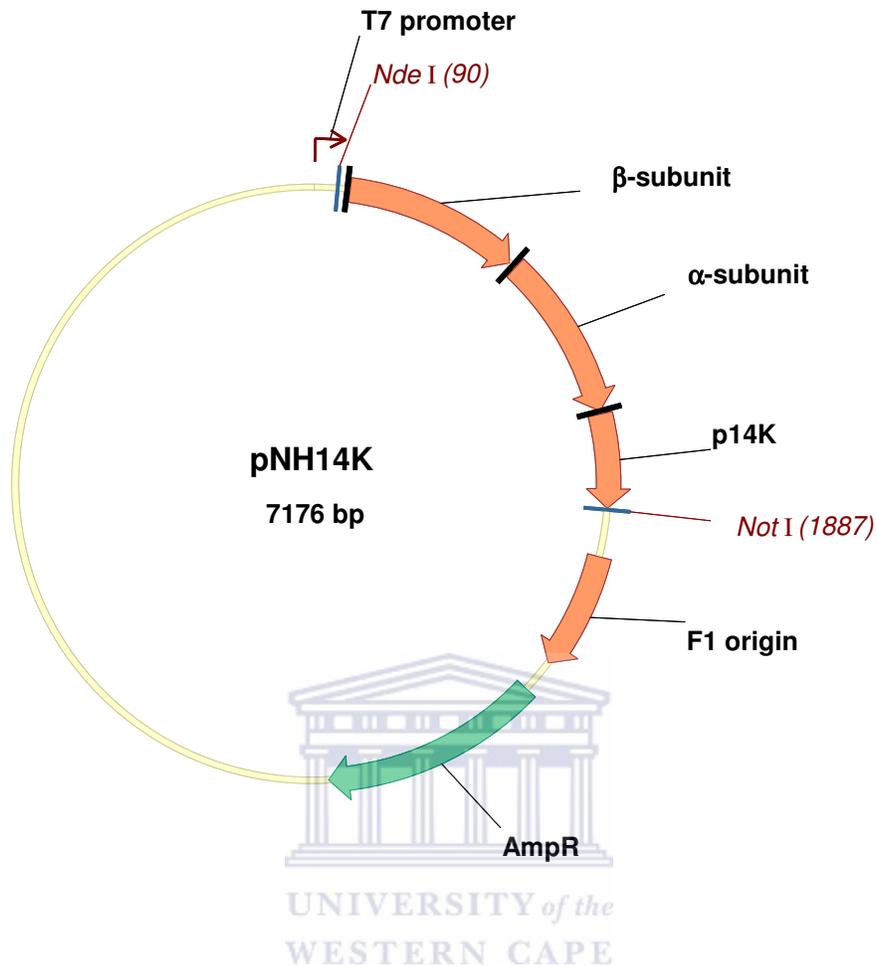


Figure 2.1: Schematic of the pNH14K vector generated using vector NTI Suite8 software (Infomax). The blue lines indicate restriction enzyme sites and black lines the position of ribosome-binding sites.

2.4. MUTANT LIBRARY CONSTRUCTION

The EP-PCR products were digested with *Dpn* I to remove template DNA and gel purified using the GFX PCR DNA purification kit (Amersham Biosciences). Purified products were digested overnight with the restriction enzymes *Nde* I and *Not* I in Buffer O containing 5 mM Tris-HCl (pH 7.5 at 37°C), 1 mM MgCl₂, 10 mM NaCl, 0.01 mg/ml BSA (Fermentas). The pNH14K plasmid was digested with the same

restriction enzymes and the 5.3 kb fragment (essentially pET21) gel-purified. The restriction enzymes were heat inactivated at 65°C for 20 minutes. The EP-PCR products were ligated with the 5.3 kb fragment of the pNH14K plasmid overnight at 16°C. The ligations were set up as follows: 2 µl vector, 3 µl insert, 4 µl of 5x ligation buffer (Fermentas), 0.5 µl T4 DNA ligase (Fermentas). The resulting constructions were transformed into freshly prepared JM109 DE3 competent cells using standard molecular biology techniques. A 50 µl aliquot of each library was plated onto Luria agar plates (containing per litre: 10.0 g NaCl, 5.0 g yeast extract, 10.0 g tryptone, 100 µg/ml ampicillin) in order to determine the transformation efficiency. The rest of the transformation of each library was inoculated into 5 ml Luria broth (containing per litre: 10.0 g NaCl, 5.0 g yeast extract, 10.0 g tryptone, 15.0 g bacteriological agar, 100 µg/ml ampicillin) and amplified for 45 minutes at 37°C. Each library was divided into fractions and glycerol stocks were prepared as described by Sambrook and Russell (2001).

Table 2.2: Primers used in this study

Random mutagenesis	
Nhop2	5'-GGGAATTCCATATGAACGGTATTCATGATGTTGG
NH14K5'R	5'-AAGGAAAAAAGCGGCCGCATTAATAAAAAACCTCATCTCC
Site-directed mutagenesis	
pNH9CalphaR505f	5'-Pho-ATCCGGGTATGGGACCGCAGTTCAGAAATTC
pNH9CalphaR505r	5'-Pho-TTCTACTGAATCTGGAAGATCTAATCCGAA

Sequences corresponding to restriction enzyme recognition sites are underlined and mutated residues are in bold.

2.5. SITE-DIRECTED MUTAGENESIS

Site-directed mutants were constructed using the Phusion™ High-fidelity DNA polymerase (Finnzymes). The method involved PCR amplification of the target vector with two phosphorylated primers. In each case the forward primer (designated f) carried the mutation while the reverse primer (designated r) was complementary to the vector sequence. The mutagenic primer was designed so that the mutation was in the middle of the primer with 10-15 perfectly matched nucleotides on either side. The reaction mixture contained: 10 µl 5 x Phusion HF buffer, 200 µM of each dNTP, 0.5 µM of each primer and 10 ng template DNA in a total volume of 50 µl. The PCR mixture was heated at 98°C for 30 seconds followed by 25 cycles of amplification using the following conditions: denaturation at 98°C for 10 seconds, annealing at 65°C for 30 seconds, extension at 72°C for 4 minutes, final extension at 72°C for 10 minutes. The PCR products were analysed using agarose gel electrophoresis and then circularized with DNA ligase. The ligation mixture contained: 1 µl PCR product, 0.5 µl DNA ligase, 4 µl 5 x ligation buffer in a total of 20 µl. A control ligation containing no ligase but all the other reagents was included in the analysis. DNA sequencing analysis provided by the University of Cape Town sequencing service confirmed the mutation.

2.6. OPTIMISATION OF THE HYDROXAMIC ACID ASSAY

NHase activity was determined using a modified version of the hydroxamic acid assay described by Fournand *et al.* (1998). The NHase cell free extract (CFE) was

prepared in 25 mM potassium phosphate buffer, pH 7.2. The assay involved two incubation steps, one for NHase, the second for the amidase. In the first reaction, 50 μ l of an appropriate dilution of the NHase CFE and 50 μ l 100 mM acrylonitrile were aliquoted into a microtitre plate. After a 30-minute incubation period at 37°C, 50 μ l 2 M hydroxylamine (prepared in 25 mM potassium phosphate buffer, pH 7.2) and 10 μ l amidase (prepared in 25 mM potassium phosphate buffer, pH 7.2) were added. This was followed by another 30 minute incubation step at 37°C. The reaction was terminated by the addition of 50 μ l of an acidic ferric chloride solution (0.1 M FeCl_3 in 1 M hydrochloric acid). This resulted in the formation of a red coloured compound that corresponded to the hydroxamic acid-Fe complex.

In order to determine the effect of hydroxylamine on amidase activity only the second step of the hydroxamic assay (as described above) that involved the amidase reaction was performed. The acetamide concentration was varied between 0-2000 mM. In order to determine the effect of hydroxylamine on NHase activity, the hydroxamic acid assay was performed as described above except that in addition to the NHase CFE and the nitrile substrate, the first reaction mixtures also contained between 0-100 mM hydroxylamine. The hydroxylamine concentration was adjusted to 500 mM in all the reaction mixtures before the amidase was added.

2.7. SCREENING FOR IMPROVED THERMOSTABLE NHASE MUTANTS

2.7.1. Determination of thermal inactivation temperature

The wild-type NHase was used to determine the temperature for future thermal inactivation experiments. The cell free extract (CFE) was prepared as described in section 2.8.1. The CFE was diluted in 25 mM potassium phosphate buffer and 50 μ l volumes aliquoted into eppendorf tubes. Thermal inactivation was performed at 55, 57 and 60°C over a total period of 30 minutes. For each temperature duplicate aliquots were removed at 5 minute intervals and placed immediately on ice. The aliquots were pipetted into 96-well microtitre plates and assayed as described in section 2.7.

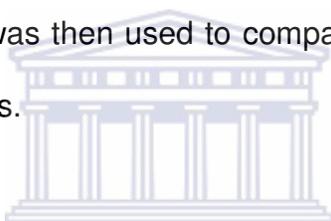
2.7.2. Cultivation and expression of mutant library in microtitre plates

The glycerol stocks of each library were plated onto Luria agar containing 100 μ M ampicillin. Single colonies were picked into 96-well microtitre plates containing 200 μ l Luria broth (LB: 1% tryptone, 0.5% yeast extract and 1% NaCl) and 100 μ M ampicillin. Microtitre plates were sealed with a Breathe-Easy membrane (Sigma) to ensure aerobic conditions. The cells were grown to saturation by overnight incubation at 37°C with vigorous aeration. Fresh medium (196 μ l) was inoculated with 2 μ l of overnight culture and incubated for 1.5 hours at 37°C with vigorous aeration. Protein expression induced with the addition of cobalt to a final concentration of 100 μ M at 1.5 hours and IPTG to a final concentration of 4 mM at 2 hours post-inoculation. Expression proceeded overnight at 37°C with vigorous

aeration. Cells were lysed in the plate by the addition of 10 μ l Pop Culture (Novagen) and the lysate used for thermostability screening.

2.7.3. Screening for improved thermostable variants

The lysate was diluted appropriately into 25 mM potassium buffer, pH 7.2 and 50 μ l volumes aliquoted into eppendorf tubes. Samples were partially inactivated by incubating eppendorf at 60°C for 10 minutes. After 10 minutes, duplicate aliquots were removed and placed immediately on ice. The aliquots were pipetted into 96-well microtitre plates and assayed at 37°C as described in section 2.6. The percentage residual activity was then used to compare the thermostabilities of the mutant and wild-type enzymes.



2.7.4. Kinetic stability analysis of improved thermostable variants

The modified hydroxamic acid assay was validated for linearity by using different concentrations of CFE under standard assay conditions as described in section 2.6. The protein sample was diluted in order for assay measurements to fall within the linear range for further experiments. Samples were assayed in triplicate for activity at 37°C while replica samples were first partially inactivated at a set temperature for a predefined period and then assayed for activity at 37°C. For kinetic stability determinations the percentage residual activity of each protein sample was determined as a function of time and the data fitted to the equation:

$$\ln (\% \text{Residual activity}) = k_d \times t$$

where k_d is the rate of thermal inactivation

2.8. PROTEIN EXPRESSION AND PURIFICATION

2.8.1. Cell-free extract (CFE) preparation

NHase was recombinantly expressed in either *E. coli* BL21 (DE3) (Novagen) or *E. coli* JM109 (DE3) (Promega). Cells were transformed with either expression vector pNH14K or the mutated pNH14K plasmids. Electrocompetent cells were prepared as described by Sambrook and Russell (2001) and transformed using a BioRad gene PulserTM. A 500 ml LB culture containing 100 µg/ml ampicillin (or carbenicillin), was grown at 37°C with shaking at 220 rpm to an optical density (at 600 nm) of 0.4, at which point expression was induced with 0.4 mM IPTG. Cobalt chloride was added to a final concentration of 0.1 mM, 30 minutes prior to induction. Cells were harvested by centrifugation 4 hours post induction and washed with 50 mM potassium phosphate buffer, pH 7.2.

The washed cell pellets containing expressed recombinant protein were stored overnight at -80°C. Cell pellets were thawed at 37°C, resuspended in 25 ml potassium phosphate buffer (pH 7.2) and disrupted by sonication (30s pulse, 30 s stop for 5 minutes) using a Bandelin Sonoplus HD2070 sonicator. The cell lysate was centrifuged at 9000 x g for 10 minutes and the supernatant collected. Heat-sensitive *E. coli* proteins were removed by incubating the cell free extract (CFE) at 55°C for 45 minutes. The CFE was centrifuged at 9000 x g for 10 minutes and the supernatant collected.

2.8.2. Ammonium sulphate precipitation

Solid ammonium sulphate was slowly added to the heat-treated sample to achieve 20% saturation followed by incubation on ice for 1 hour and centrifugation at 9000 x g for 10 minutes at 4 °C to remove precipitated proteins.

2.8.3. Hydrophobic interaction chromatography (HIC)

The supernatant from the ammonium sulphate precipitation step was loaded onto a HighLoad 16/10 Phenyl-Sepharose column (Amersham Biosciences) previously equilibrated with 50mM potassium phosphate buffer, pH 7.2 containing 1 M ammonium sulphate. Bound proteins were eluted with a linear gradient of decreasing ammonium sulphate concentration using 50 mM potassium phosphate buffer, pH 7.2 (5 column-volumes, 1 M - 0 M ammonium sulphate). Fractions containing NHase were pooled after confirmation with SDS-PAGE.

2.8.4. Anion-exchange chromatography

The pooled fractions from HIC were dialysed against 25 mM potassium phosphate buffer, pH 7.2 and loaded onto a HiLoad 26/10 Q-Sepharose column (Amersham Biosciences) equilibrated with the same buffer. Bound proteins were eluted with a linear gradient of increasing sodium chloride concentration using 25 mM potassium phosphate buffer, pH 7.2 containing 500 mM sodium chloride, (5 column volumes, 0M-500 mM sodium chloride). Fractions containing pure NHase were pooled after verification of homogeneity with SDS-PAGE.

2.9. STRUCTURE DETERMINATION USING X-RAY CRYSTALLOGRAPHY

2.9.1. Crystallisation

Prior to initial protein crystallisation, pooled fractions of pure NHase from Q-Sepharose chromatography were dialysed against 20 mM Tris-Cl, pH 7.2, filtered through a 0.22 μ M filter and concentrated to above 10 mg/ml using Vivaspin 15R 5000 molecular weight cut-off concentrator tubes (Sartorius AG).

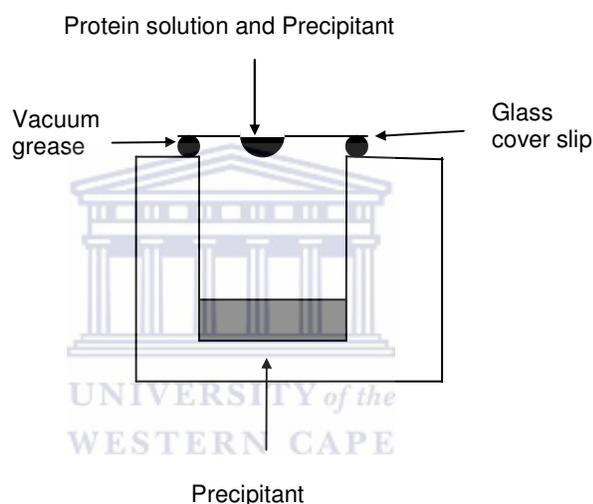


Figure 2.2: The hanging drop vapour diffusion method

Wild-type and mutant NHases were crystallised using the hanging-drop vapour diffusion method (Figure 2.2). Siliconized cover slips were used from Hampton research. Typically, 1 ml of crystallisation solution was pipetted into selected wells of a 24-well tissue culture plate (Linbro plates, ICN Biomedicals). The rim of the well was covered with a layer of vacuum grease. The hanging drops were mounted on siliconised cover slips (Hampton research) and contained 2 μ l protein and 2 μ l

reservoir solution. The cover slip was carefully inverted, placed on the top of the well and gently pressed around the rim to ensure good sealing. Diffraction quality crystals were grown at 22°C in a crystallisation solution that contained 30% PEG 400, 100 mM magnesium chloride, 100 mM MES (2[N-Morpholino]ethanesulfonic acid), pH 6.5 (10-40 mg/ml protein) as described by Tsekoa *et al.* (2004). Diffraction quality crystals were observed after one week under these conditions.

2.9.2. X-ray data collection and processing

Initial X-ray diffraction data from crystals of wild-type and the 7D mutant were collected at the in-house X-ray source of the Department of Biotechnology, University of the Western Cape. The equipment consisted of a Rigaku RUH3R copper-rotating anode X-ray source operated at 40 kV, 22 mA, a Rigaku R-axis IV+ image plate camera, an X-stream 2000 low temperature system and an AXCO PX50 glass capillary optic with a 0.1 mm focus. Crystals were mounted on a cryoloop (Hampton Research) and data were collected under cryo-conditions (in a nitrogen stream at a temperature of 100 K) with an oscillation angle of 0.5 degree, an exposure time of 15 min and a crystal-to-detector distance of 129 mm. It was not found to be necessary to add cryoprotectant to the crystals. The machine was operated with the program Crystal Clear (Pflugrath, 1999).

Subsequent data collection of wild-type and four randomly created NHase mutants was carried out by Prof. Trevor Sewell and Dr. Muhammed Sayed at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France on station BM14. Single crystals of NHase were flash cooled in liquid nitrogen and stored in a XT-20

(Taylor-Wharton) dry shipper dewar for transport to the ESRF. The X-ray diffraction patterns were recorded with an oscillation angle of 0.5 or 1 degree per image using a MAR 225 mm CCD based detector at a wavelength of 0.979Å, with a crystal-to-detector distance of 100 or 120 mm, and an exposure time of 10 seconds per image. The collected diffraction data was processed and refined using Crystal Clear (d*TREK) (Pflugrath, 1999). Solvent content and Matthews Coefficient were calculated using Matthews from the CCP4 suite of programs (CCP4, 1998).

2.9.3. Structure solution and refinement

The crystal structures of mutant NHases were solved by molecular replacement using the program PHASER (McCoy *et al.*, 2007). The wild-type NHase structure (PDB code: 2dpp) was used as a search model for this procedure. All structures were refined using Refmac5 from the CCP4 suite of programs (Murshudov, 1997) and model rebuilding was carried out using O (Jones *et al.*, 1991). The final model was validated using WHATCHECK (Hooft *et al.*, 1996). Prof. Trevor Sewell (University of Cape Town) performed all structure solution, refinement and validation procedures.

2.9.4. Molecular graphics and structural analysis

All structures were visualized using either PyMOL (Delano Scientific LLC) or the UCSF Chimera software package (<http://www.cgl.ucsf.edu/chimera>). Molecular graphics of all images were generated using UCSF Chimera.

CHAPTER 3

RANDOM MUTANT LIBRARY CONSTRUCTION AND SCREENING FOR IMPROVED THERMOSTABLE NHASES

CONTENT

3.1.	INTRODUCTION.....	51
3.2.	OPTIMISATION OF THE HYDROXAMIC ACID ASSAY	54
3.3.	RANDOM MUTANT LIBRARY CONSTRUCTION.....	58
3.3.1.	Library design and construction.....	58
3.3.2.	Frequency of active mutant enzymes.....	60
3.4.	THE RELATIONSHIP BETWEEN THERMOSTABILITY AND ACTIVITY	61
3.5.	SCREENING FOR NHASES WITH IMPROVED THERMOSTABILITY	63
3.6.	SEQUENCE ANALYSIS OF LIB 0.05	68
3.7.	SEQUENCE ANALYSIS OF THERMOSTABILISED NHASE MUTANTS	71

3.1. INTRODUCTION

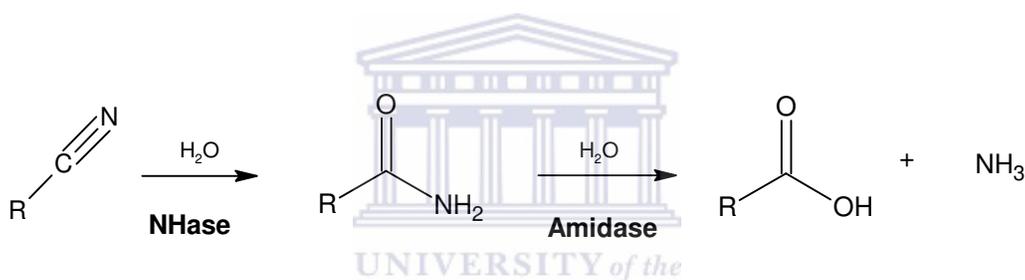
The quality of the mutant library is very important for the successful outcome of a random mutagenesis experiment (Wong *et al.*, 2006). Of the many random mutagenesis methods available, error-prone PCR (EP-PCR), first described by Leung *et al.* (1989) and later modified by Cadwell and Joyce (1992), is by far the simplest and most common method used to create genetic diversity. For the purpose of EP-PCR, *Taq* polymerase is used since it is inherently prone to introduce random errors into a PCR reaction. Error frequencies ranging between 2×10^{-4} (Eckert and Kunkel, 1991) and 7.2×10^{-5} (Ling *et al.*, 1991) errors per base pair have been reported for this enzyme group. However, this inherent error frequency is insufficient to generate a suitably diverse randomly mutated library and $MnCl_2$ is added to further reduce the fidelity of the polymerase.

The addition of $MnCl_2$ into the PCR reaction causes the *Taq* polymerase to increase the frequency with which random errors are introduced into the gene sequence. This process of misincorporation is further facilitated by biasing the dNTP concentration. The objective of EP-PCR is to generate mutant libraries with a low mutation frequency i.e. 2-5 base substitutions per gene, corresponding to approximately one amino acid residue change per mutated protein (Cadwell and Joyce, 1992). A low mutation rate is preferred for EP-PCR so that it possible to observe rare beneficial mutations. If the mutation rate is too high, rare beneficial point mutations are masked in the negative background and the discovery of (combinations of) novel mutations within the population of existing mutations is

severely inhibited (Moore *et al.*, 1997). Since the error frequency of the *Taq* polymerase essentially depends on the concentration of $MnCl_2$ and the number of cycles of the PCR reaction (Cadwell and Joyce, 1992; Shafikhani *et al.*, 1997), it is very important that these variables are carefully controlled.

The availability of a suitable selection or screening system is considered even more important than the method of library creation (Arnold and Volkov, 1999). A selection strategy involves linking the improved property to the survival of a suitable host organism (Bommarius *et al.*, 2006). The thermophilic bacterium, *Thermus thermophilus*, has been successfully used to select for thermostability at high temperatures (Tamakoshi *et al.*, 2001). However, screening for thermostability is preferred over selection strategies and typically involves colorimetrically measuring the residual enzyme activity following thermal inactivation at a predetermined temperature (Bommarius *et al.*, 2006). An effective screen will aim to exploit the link between the gene, the enzyme it encodes and the product that result from enzyme activity (Aharoni *et al.*, 2005). Since there is no simple assay to detect either the nitrile substrate or the amide product formed from NHase activity, high performance liquid chromatography (HPLC) or gas chromatography (GC) has been routinely used to measure these compounds (Takashima *et al.*, 1998; Cramp and Cowan, 1999; Kim and Oriel, 2000). However, in a combinatorial design experiment, the number of enzymes that need to be screened is often too large for conventional analytical assays such as HPLC and GC (Bommarius *et al.*, 2006).

NHase activity has also been determined colorimetrically using a modified version of the phenol-hypochlorite method described by Fawcett and Scott (1960). This assay indirectly estimates NHase activity by measuring the amount of ammonia released from the bi-enzymatic reaction. The nitrile substrate is first converted to the corresponding amide by NHase which is subsequently catalysed to the corresponding carboxylic acid by the amidase with the resultant release of ammonia (Scheme 3.1). The ammonia forms a coloured complex with phenol-hypochlorite that can be measured spectrophotometrically.

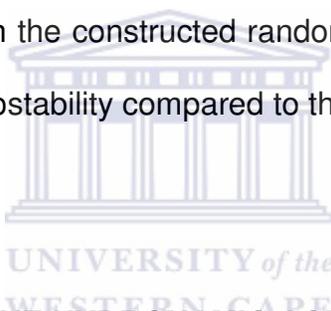


Scheme 3.1: Enzymatic conversion of nitriles to ammonia by nitrile hydratase and amidase

Although the phenol-hypochlorite method has been used previously to detect NHase activity (see, for example, Cameron *et al.*, 2005; Okamoto and Eltis, 2007), the assay was not necessarily suitable for use as a screening system to measure NHase activity during a random mutagenesis study. Since it is possible for microbial cells to produce endogenous ammonia irrespective of whether NHase is being expressed or not, it could potentially be problematic to generate an acceptable negative control with whole cells or cell extracts. Adhering to the first law of random mutagenesis; i.e. “You get what you screen for” (You and Arnold,

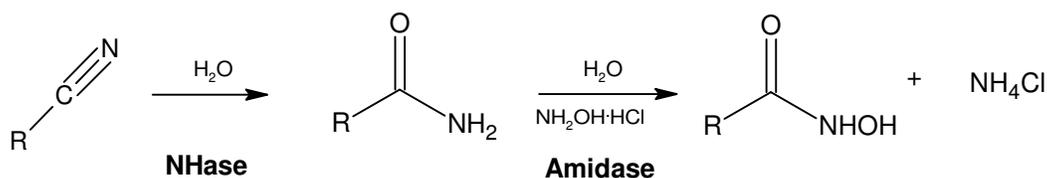
1996) and considering that the best screening strategy is always to directly measure the product that result from enzyme activity (Aharoni *et al.*, 2005), it was to identify and optimise an assay that would be more suitable to screen for NHase activity when using a random mutagenesis approach.

The aim of this part of the project was to use random mutagenesis techniques, specifically EP-PCR, to construct a NHase randomly mutated library. It was also necessary to identify and optimise a suitable assay that could be applied as a screening system to detect NHase activity in a microtitre plate format. This assay would then be used to screen the constructed randomly mutated library for NHase variants with improved thermostability compared to the wild-type.



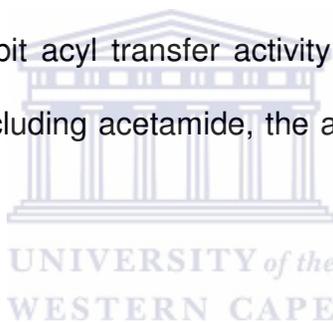
3.2. OPTIMISATION OF THE HYDROXAMIC ACID ASSAY

The hydroxamic acid assay is typically used to detect the acyl transfer activity of amidases by measuring the formation of hydroxamic acid that results from the transfer of the acyl group of the amide to hydroxylamine ($\text{NH}_2\text{OH}\cdot\text{HCl}$) (Fournand *et al.*, 1998). The hydroxamic acid assay can also indirectly measure NHase activity by including an additional incubation step of the NHase with the nitrile substrate as depicted in Scheme 3.2. This assay was preferred to the phenol-hypochlorite method as the formation of the final colorimetric product (hydroxamic acid), is dependent on the presence of the amide that only results from NHase activity.



Scheme 3.2: Enzymatic conversion of nitrile to hydroxamic acid by NHase and amidase

The amidase used in the present study was previously cloned from *Geobacillus pallidus* RAPc8 (Cameron *et al.*, 2005) and is part of the NHase operon. This amidase was shown to exhibit acyl transfer activity on various amide substrates (Makhongela *et al.*, 2007) including acetamide, the amide that results from NHase activity on acetonitrile.



In order to optimise the hydroxamic acid assay to measure NHase activity, it was necessary to ensure that the observed effect was only due to the relationship between hydroxylamine and the NHase. The first assay trials used a single incubation step, where hydroxylamine and amidase was added together with the NHase and acetonitrile. This showed that the NHase was inhibited by the hydroxylamine (data not shown). Further assay trials were conducted in order to investigate the extent of the inhibition of the NHase by hydroxylamine. The effect of hydroxylamine concentration on amidase activity was also determined.

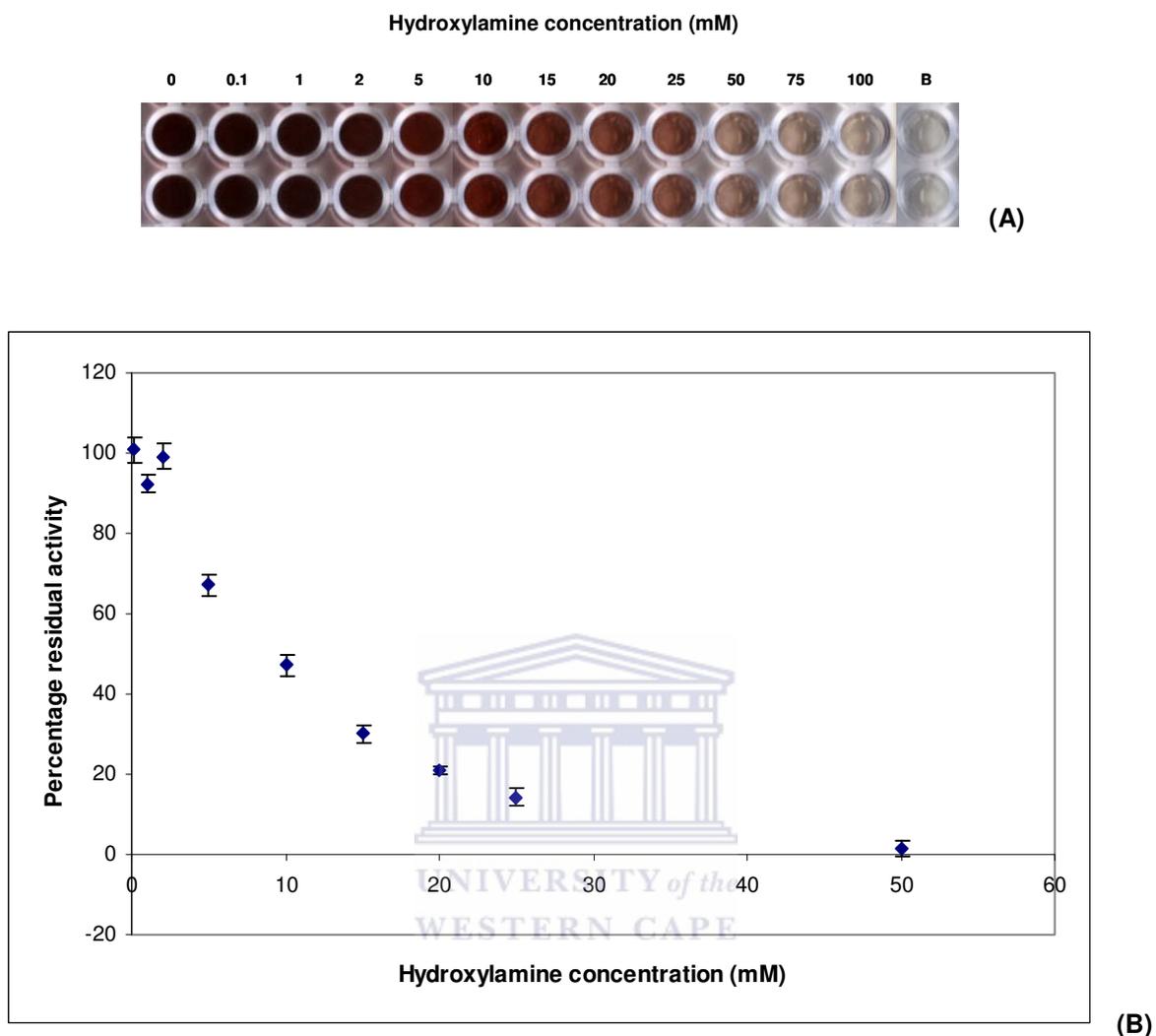
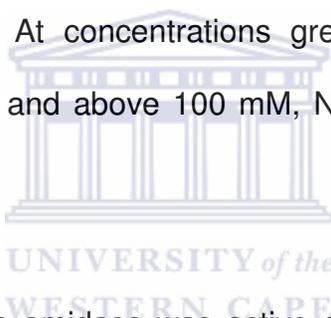


Figure 3.1: Effect of hydroxylamine on nitrile hydratase activity using cell free extract (CFE). (A) The observed effect of increasing hydroxylamine concentration on NHase activity. (B) The percentage residual activity for reaction mixtures containing increasing concentrations of hydroxylamine was determined relative to the reaction where the no hydroxylamine was present during the incubation of the NHase with the nitrile substrate. The error bars indicate the standard deviation between duplicate measurements.

The effect of hydroxylamine concentration on NHase activity was determined by the addition of increasing concentrations of hydroxylamine in the reaction mixture

during the preliminary incubation of the NHase with the nitrile substrate. The hydroxylamine concentrations in all the reaction mixtures was standardised to 500 mM for the second incubation step involving the amidase. This ensured that the effect observed was only due to the relationship between hydroxylamine and the NHase. The presence of hydroxylamine in the reaction mixture resulted in the inhibition of NHase activity. The effect was quite pronounced and could be observed visually (Figure 3.1a). The percentage residual NHase activity was determined relative to the reaction where no hydroxylamine was present. The percentage residual activity decreased with increasing hydroxylamine concentration (Figure 3.1b). At concentrations greater than 50 mM, very low NHase activity was detected and above 100 mM, NHase activity was completely inhibited.



In contrast to the NHase, the amidase was active at very high concentrations of hydroxylamine. Indeed, the amidase activity was optimal at a hydroxylamine concentration of 500 mM (Figure 3.2). Therefore, the concentration of hydroxylamine that is optimum for amidase activity effectively terminates NHase activity and the net result of the reaction is due to the relationship between the NHase and hydroxylamine. Consequently, the optimised hydroxamic acid assay was performed in two steps to assay for NHase activity. The first step involved the incubation of the NHase with 50 mM acetonitrile for 30 minutes at 37°C. The hydroxylamine was then added to a final concentration 500 mM and together with the amidase the reaction mixture was incubated 30 minute period at 37°C.

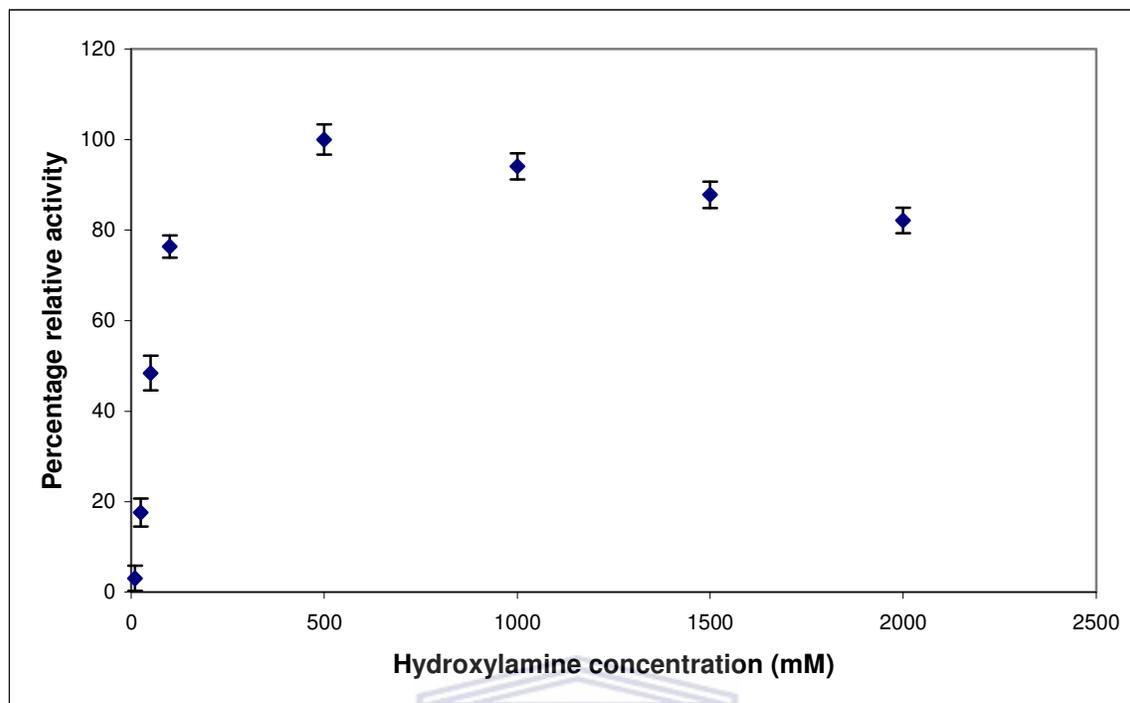
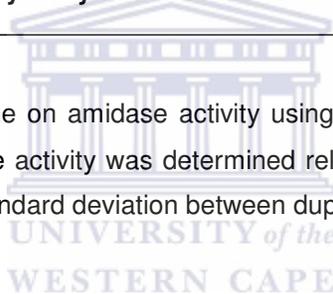


Figure 3.2: Effect of hydroxylamine on amidase activity using partially purified (heat-treated) cell free extract (CFE). The percentage activity was determined relative to the optimum activity at 500 mM. The error bars indicate the standard deviation between duplicate measurements.



3.3. RANDOM MUTANT LIBRARY CONSTRUCTION

3.3.1. Library design and construction

EP-PCR was performed as described in section 2.3. Initially, the EP-PCR reaction was attempted using 0.5 mM MnCl_2 . This MnCl_2 concentration was reported to generate low mutation frequency libraries i.e. 2-5 base substitutions per gene, corresponding to approximately one amino acid residue change per mutated protein (Cadwell and Joyce, 1992; Chen and Arnold, 1993; Daugherty *et al.*, 2000; Shafikhani *et al.*, 1997). However, no PCR product was obtained when the EP-PCR was performed with this concentration of MnCl_2 (data not shown). The fact

that the inhibition of the *Taq* polymerase was due to the $MnCl_2$ was confirmed by performing a $MnCl_2$ titration experiment.

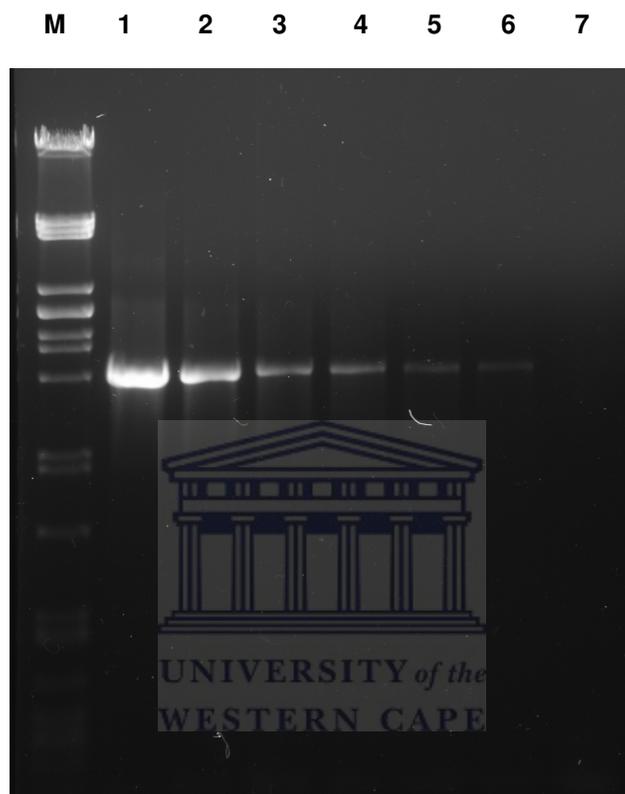


Figure 3.3: Effect of $MnCl_2$ concentration on PCR product yield during error-prone PCR (EP-PCR). M: Molecular weight marker (λ PstI); Lanes 1-7 $MnCl_2$ concentration; Lane 1: 0 mM, Lane 2: 0.05 mM, Lane 3: 0.1 mM, Lane 4: 0.15 mM, Lane 5: 0.2 mM, Lane 6: 0.25 mM, Lane 7: 0.5 mM

Figure 3.3 illustrates that the *Taq* polymerase used in the present study was inhibited by the presence of $MnCl_2$ in the PCR reaction as the amount of PCR product decreased with increasing $MnCl_2$ concentration. Even the lowest concentration of 0.05 mM had an effect on the PCR product yield. Since the sensitivity of *Taq* polymerase to the $MnCl_2$ concentrations was likely to affect the

error frequency of the library, it was decided to use a lower MnCl_2 concentration for the EP-PCR reaction. However, since it was impossible to know the extent to which the MnCl_2 concentration would affect the error frequency of this enzyme, it was decided to construct more than one library using different MnCl_2 concentrations. Consequently, two randomly mutated libraries were constructed using 0.05 and 0.10 mM MnCl_2 .

3.3.2. Frequency of active mutant enzymes

The two randomly mutated libraries were transformed into electrocompetent *E. coli* JM109 (DE3) cells. Each library was cultivated, expressed and lysed in 96-well microtitre plates. The wild-type NHase was used as a control. The cell lysates of the wild-type and mutant NHases were assayed for activity at 37 °C as described in section 2.6 of the experimental procedures.

Table 3.1: Frequency of active mutant enzymes for EP-PCR libraries

Library	MnCl_2 concentration (mM)	Percentage active enzymes
Lib 0.05	0.05	56±4.02
Lib 0.10	0.10	33±3.79

The percentage of active mutant enzymes was calculated using three different microtitre plates from each library. A little more than half (56%) of enzymes from

Lib 0.05 were found to be active whereas only 33% of the enzymes from Lib 0.1 were active (Table 3.1). Therefore, the percentage of active enzymes within the libraries decreased with increasing MnCl_2 concentration. This indicated that the number of inactive enzymes increased with increasing error frequency of the polymerase and hence the associated increase in the mutation frequency of the libraries.

3.4. THE RELATIONSHIP BETWEEN THERMOSTABILITY AND ACTIVITY

The potential industrial application of NHases requires that both the catalytic activity and the thermostability of the enzymes are optimised. It has been illustrated that thermophilic enzymes are generally less active at room temperature than their mesophilic counterparts (Daniel, 1996). This has led to suggestions that thermostability and activity are mutually exclusive and optimising the one property is detrimental to the other (Vieille and Zeikus, 2001). However, evidence from directed evolution studies suggests that thermostability and activity need not be mutually exclusive (Giver *et al.*, 1998). Furthermore, it is possible to evolve both these properties in an enzyme (Strausberg *et al.*, 2005).

The simplest way to ensure that thermostabilised enzymes are also catalytically active is to screen for thermostability based on catalytic activity (Kuchner and Arnold, 1997). This is done by performing the screen in two steps. First the initial activities of the mutant enzymes (A_i) are determined at the low temperature. The

enzymes are then subjected to partial thermal inactivation at an elevated temperature and then the residual activity (A_r) is again measured at the low temperature. The percentage residual activity (A_r/A_i) is used as an indication of the extent of the improvement in thermostability. By screening for both thermostability and activity it is possible to engineer enzymes that retain low temperature activity while the thermostability is enhanced (Giver *et al.*, 1998; Miyazaki *et al.*, 2000).

Since this study aims to engineer thermostabilised NHases it is also important that the catalytic activity of the enzymes is not compromised; i.e., the intricate balance between thermostability and activity must be maintained. The possibility that the observed lack of activity of randomly mutated enzymes was due to these enzymes having substantially reduced activity at 37°C was considered. The activity of the mutant NHase was determined at 37°C (the standard temperature used during this study) and 55°C (the optimum temperature of the wild-type) and found to be similar at both temperatures. Therefore, it was concluded that it would be feasible to perform a thermal inactivation step at an elevated temperature and then screen for residual activity at 37°C. By using the catalytic activity of the mutant NHases to screen for improved thermostable enzymes it was possible to ensure that enzymes isolated from the screening process would be more thermostable than the wild-type NHase as well as retain activity at 37°C.

3.5. SCREENING FOR NHASES WITH IMPROVED THERMOSTABILITY

The wild-type NHase was used to determine the screening temperature for thermal inactivation studies. The aim was to select a temperature at which the wild-type enzyme was partially inactivated following heat-inactivation for a set period of time.

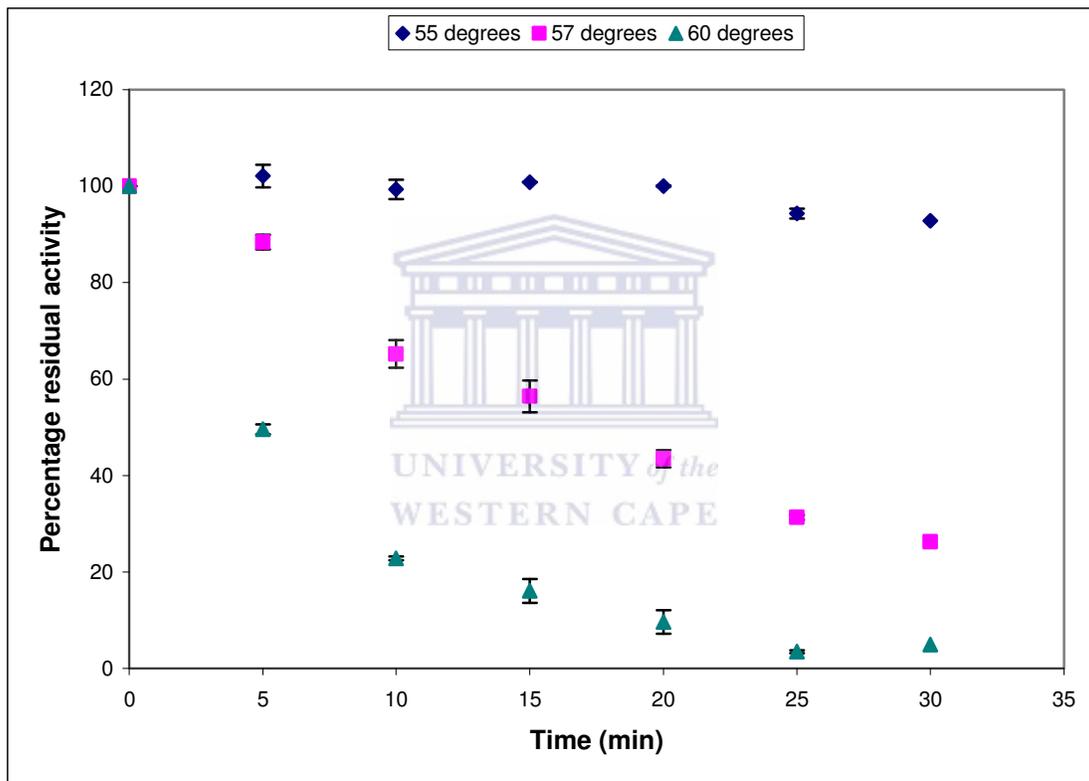


Figure 3.4: Effect of thermal inactivation temperature on wild-type NHase thermostability using cell free extract (CFE). The error bars indicate the standard deviation between duplicate measurements.

The wild-type NHase cell free extract (CFE) was subjected to partial thermal inactivation at 55, 57 and 60°C for five minute time intervals over a period of 30

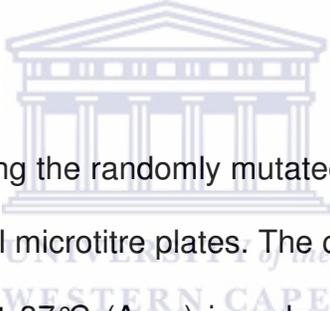
minutes. The percentage residual activity of NHase subjected to thermal inactivation at 60°C was determined relative to that at 37°C (i.e. no thermal inactivation) according to the equation:

$$\%RA = A_{(60^{\circ}\text{C})} / A_{(37^{\circ}\text{C})} * 100 \quad (\text{Equation 1})$$

Where %RA is the percentage residual activity, $A_{(60^{\circ}\text{C})}$ is the NHase activity at 37°C following thermal inactivation at 60°C and $A_{(37^{\circ}\text{C})}$ is the NHase activity at 37°C without a prior thermal inactivation step.

Figure 3.4 shows the percentage residual activity of the wild-type NHase following thermal inactivation at 55, 57 and 60°C. The enzyme was quite thermostable at 55°C and retained initial activity over the 30 minute incubation period. At 57°C the enzyme was partially inactivated. The thermal inactivation process proceeded moderately at this temperature and the NHase retained approximately 25% residual activity after 30 minutes. At 60°C thermal inactivation proceeded more rapidly than for the other two temperatures and approximately 20% residual activity was attained after only 10 minutes at this temperature. Consequently, it was decided to conduct thermal inactivation experiments of the mutant library at 60°C for 10 minutes. The wild-type was considerably inactivated at this temperature which was observed visually by the low intensity of the red coloured complex formed. This meant that it should be able to optically visualise an enzyme with improved thermostability compared to the wild-type.

Several attempts were made to perform the heat-inactivation in the microtitre plates using different incubators and ovens. However, none of the available incubators tested could facilitate uniform thermal inactivation across the microtitre plate. This was very important since it was clear for the wild-type that a 2-3°C difference in the thermal inactivation temperature resulted in significant differences in the degree of thermal inactivation (Figure 3.4). It was decided that since a large proportion of enzymes were inactive that only the active enzymes in each library should be subjected to thermal-inactivation studies. Therefore, these studies were conducted in a water bath as described in section 2.7.3 of the experimental procedures.



E. coli JM109 (DE3) harbouring the randomly mutated libraries were cultivated and expressed in duplicate 96-well microtitre plates. The cell lysates from one microtitre plate was directly assayed at 37°C ($A_{37^\circ\text{C}}$) in order to determine which enzymes were active. Mutant enzymes that displayed less than 10% of the parental activity were designated inactive and excluded from subsequent thermal inactivation analysis. The cell lysates of the selected randomly mutated enzymes were partially thermally inactivated at 60°C for 10 minutes and the NHase activity determined at 37°C. Figure 3.5 illustrates a typical positive result obtained from thermal inactivation analysis. The considerable extent of thermal inactivation of the wild-type NHase meant that a positive result could be observed visually by the increased intensity of the red colour of the improved mutant compared to the wild-

type. However, putative thermostable enzymes were selected based on a higher percentage residual activity compared to that of the wild-type NHase.

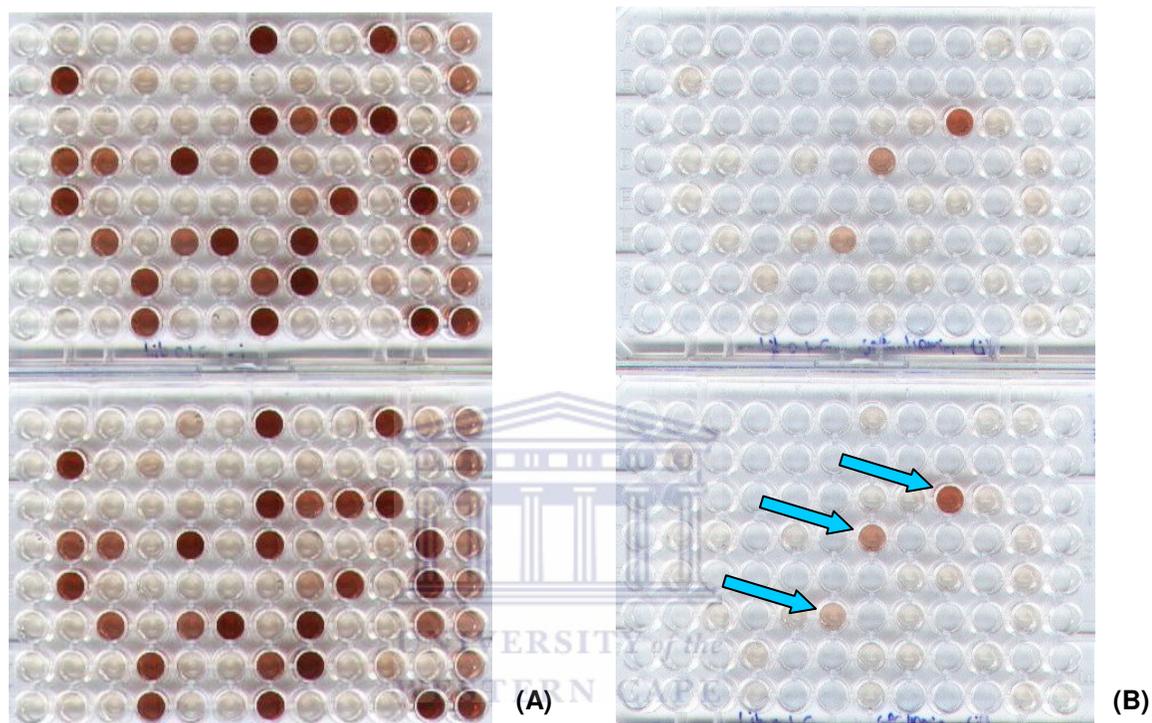


Figure 3.5: The hydroxamic assay performed in microtitre plate format showing a positive result obtained from screening Lib 0.1 for NHases with improved thermostability. (A) The activity of the enzymes was determined directly at 37°C. (B) The enzymes were then subjected to thermal inactivation at 60 degrees for 10 minutes and the activity measured at 37°C. Potential thermostabilised NHases are indicated with blue arrows.

The percentage residual activity of each enzyme was calculated according to Equation 1. Modified NHase enzymes with a percentage residual activity greater than 25% following thermal inactivation were considered to be more thermostable than the wild-type. Putative thermostable NHases were streaked from the master

plate, cultivated and the plasmid DNA extracted. The mutant plasmids were then transformed into electrocompetent *E coli* BL21 (DE3) and subjected to an identical procedure to the screening enzymes for improved thermostability to confirm that they were more thermostable than the wild-type NHase (Figure 3.6).

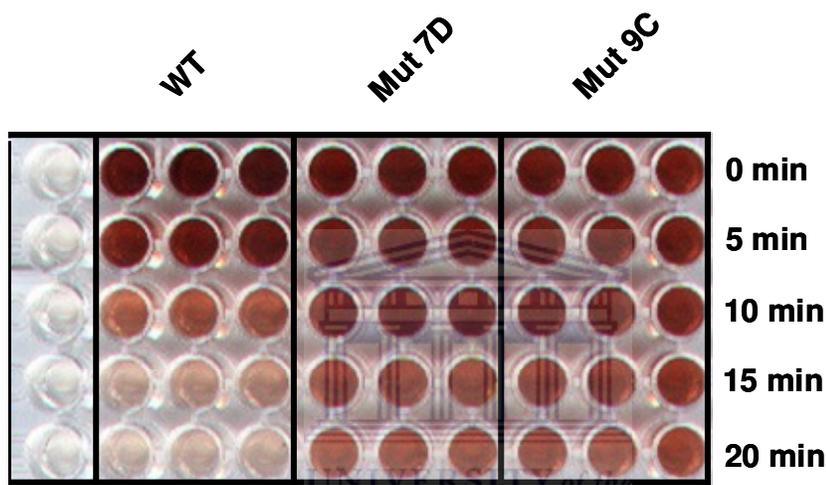


Figure 3.6: Using the hydroxamic assay to confirm the improved thermostability between the wild-type and mutant NHases

A total of approximately 800 suitably active modified NHases from each of the three libraries were screened for improved thermostability. In total, six putative positive mutant enzymes that exhibited increased thermostability compared to the wild-type were isolated. Four of the mutants were isolated from Lib 0.05 and two from Lib 0.1.

3.6. SEQUENCE ANALYSIS OF LIB 0.05

Since the mutation frequency of the library was considered a probable cause for the high percentage of inactive NHase mutants, a set of 25 randomly constructed NHase mutant genes from Lib 0.05 MnCl_2 was subjected to DNA sequence analysis in order to assess the mutation frequency of this library. The mutants were assessed for the number of mutations, the type of mutation and the distribution of the mutations across the genes of the NHase operon. The mutation frequency of Lib 0.05 MnCl_2 is summarised in Table 3.2.

Lib 0.05 was dominated by AT→GC transitions and AT→TA transversions (Table 3.2). This mutational bias was observed equally in the β - and α subunit as well as across the NHase operon (defined in this context to consist of the β - and α subunit as well as the P14K protein). The only inconsistency regarding the mutational bias was observed for the p14k transcript where 40% of the mutations were GC→CG transitions. However, this was mainly due to a nucleotide change at one position within the gene encoding the P14K protein. Interestingly, all of the randomly mutated genes sequenced contained the C→G transition at position 349 of the p14k transcript. The wild-type p14k transcript was sequenced to determine the validity of this mutation. Sequence analysis revealed that this mutation was genuine and not the result of a sequencing error that occurred when the wild-type gene was originally sequenced. This mutation probably occurred early in the EP-PCR and was subsequently amplified during later cycles resulting in a common mutation at this position.

Table 3.2: Mutation frequency of Lib 0.05 MnCl₂

	β subunit	α subunit	P14K protein	NHase operon
Base pairs sequenced	17 250	16 500	9 225	42 975
Total mutations	52	65	62	178
Mutation frequency, %	0.30	0.39	0.67	0.41
Ave. mutations per gene/operon	2.1	2.6	2.5	7.1
Nonsynonymous mutations, %	61.5	55.4	37.7	51.1
Synonymous mutations, %	28.8	40.0	52.5	41.0
Frameshift/nonsense mutations, %	9.6	4.6	9.8	7.9
Mutation types, %				
A→T, T→A	35	32	13	26
A→C, T→G	13	1	5	6
A→G, T→C	44	38	23	35
G→A, C→T	6	14	8	9
G→C, C→G	0	1	40	15
G→T, C→A	2	6	3	4
Frameshift	0	5	8	5

The mutation frequency was determined using DNA sequence data from a randomly selected set of 25 mutants.

The mutation frequency was determined for the NHase operon as a whole as well as for each separate gene encoding the α subunit, β subunit and P14K protein, respectively. The mutants NHases contained a total of 179 nucleotide mutations in the 43,005 base pairs of the NHase operon sequenced (Table 3.2). The overall mutation frequency was 0.41% per position across the NHase operon. However, the mutation frequency varied for the individual genes. The mutation frequency was 0.30% for the β subunit, 0.39% for the α subunit and 0.67% for the gene encoding the P14K protein. The variation in the mutation frequency of the genes in the NHase operon was not due to variations in the nucleotide composition of the genes as the %GC content was relatively similar for all three genes; i.e. 42.70% (β subunit), 41.16% (α subunit) and 42.66% (p14k gene). It should be noted that the highest mutation frequency occurred in the P14K protein and this was primarily due to the same C→G transition at position 349 of the p14k transcript.

The genes encoding the β subunit, α subunit and P14K protein contained an approximate average of 2.1, 2.6 and 2.5 mutations, respectively. The α subunit and P14K contained more base pair changes resulting in a higher mutation frequency and a higher average number of mutations per gene than the β subunit. However, more of the mutations in the β subunit (61.5%) were non-synonymous compared to the other two genes. This led to an average of 7.1 mutations per NHase operon. More than half (51%) of the mutations in the NHase operon were non-synonymous mutations while 41% were synonymous. This meant that each mutant NHase was postulated to contain an approximate average of 4 amino acid changes. Eight

percent of the mutations led to premature truncation of the genes due to frameshift or nonsense mutations. Most of these frameshift or nonsense mutations occurred in the β subunit (9.6%) and p14k transcript (9.8%).

3.7. SEQUENCE ANALYSIS OF THERMOSTABILISED NHASE MUTANTS

All genes encoding mutated NHases that exhibited improved thermostability compared to the wild-type enzyme were subjected to DNA sequence analysis. Each mutated gene was sequenced in duplicate to ensure that any observed change was due to random mutagenesis and not the result of sequencing errors. The mutation frequency of the putative improved thermostable mutants is summarised in Table 3.3.

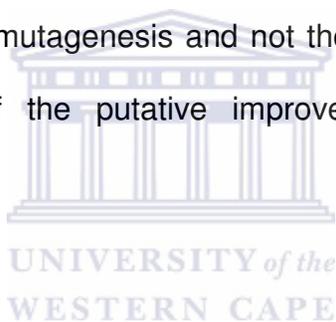


Table 3.3: Mutation frequency of putative thermostabilised mutants

	beta subunit	alpha subunit	P14K protein	NHase operon
Base pairs sequenced	4140	3960	2214	10 314
Total mutations	14	11	11	36
Mutation frequency, %	0.34	0.28	0.50	0.35
Ave. mutations per gene/operon	2.3	1.8	1.8	6.0
Synonymous mutations, %	21.4	36.4	54.5	36.1
Nonsynonymous mutations, %	78.6	63.6	45.5	63.9
Frameshift/nonsense mutations, %	0	0	0	0

The mutation frequency was 0.34% in the β -subunit compared to 0.28% and 0.50% in the α -subunit and p14K, respectively. Table 3.4 provides a summary of the specific nucleotide and amino acid changes that was observed in the β -subunit, α -subunit and P14K protein. As for Lib 0.05, the thermostabilised mutants were dominated by AT→TA transitions and AT→GC transversions. This was not surprising as most of the thermostabilised NHases were isolated from this randomly mutated library. Most (78.6%) of the mutations in the β -subunit were non-synonymous (Table 3.3) and as a result, all the mutants, excluding Mut 7D, had at least two and at most three amino acid changes in the β -subunit (Table 3.4). None of the improved thermostable enzymes contained mutations that resulted in premature truncation of the genes due to frameshift or nonsense mutations.

All the improved thermostable NHase mutants also contained the same GC→CG transition at position 349 of the *p14k* transcript as previously identified in Lib 0.05. This mutation occurred in mutants from both Lib 0.05 and Lib 0.1, thus, it is proposed that the source of this mutation was not from the early rounds of EP-PCR, but that this position might be prone to mutations. This base pair change resulted in the substitution of an Arg residue with a Gly residue. Moreover, mutant 9C, that was isolated from Lib 0.1 contained two consecutive base pair changes at positions 348 and 349 of the *p14K* transcript. It is very rare for consecutive base pair changes to occur during random mutagenesis. Daugherty *et al.* (2000) found only one adjacent nucleotide substitution for a library containing an average of 23

mutations per gene. This further supports the idea that this area on the *p14k* transcript is prone to mutations.

A total of 14 nucleotide changes occurred in the β -subunit compared to 11 in both the α -subunit and the *p14K* transcript (Table 3.3). The distribution of the nucleotide changes across the NHase operon for the thermostabilised mutants is shown in Figure 3.7. A minimum of three and a maximum of seven nucleotide changes were observed per NHase operon consisting of the α subunit, β subunit and P14K protein. This translated into a minimum of one and a maximum of five amino acid changes per NHase mutant. Mutant 7D contained only a single amino acid change and the Ser was obviously involved in the improved thermostability of this mutant. Since all the other mutants contained more than one amino acid change, it was more difficult to predict which amino acid residues are involved in conferring improved thermostability based on the sequence data alone. Some of the mutations did result in the introduction of charged amino acids such as Arg and Lys that could potentially facilitate increased electrostatic interactions; e.g., mutant 9C (Table 3.4). Sequence comparisons have shown that thermophilic proteins contain more Glu, Lys and Arg residues compared to mesophilic proteins (Das and Gerstein, 2000). This sequence comparison data can also be correlated with structural comparison data that show increased occurrences of electrostatic interactions in thermophilic proteins (Xiao and Honig, 1999).

Table 3.4: Position of nucleotide (nt) and amino acid (aa) changes in randomly mutated NHases with improved thermostability compared to the wild-type

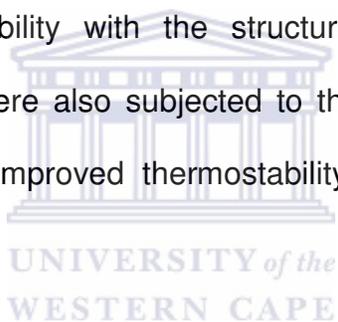
Library	Mutant	β subunit		α subunit		P14K		Total number nt changes	Total number aa changes
		nt changes	aa changes	nt changes	aa changes	nt changes	aa changes		
Lib 0.05	6B	A→T (321) G→A (637)	D→E (107) N→D (213)	A→G (14) A→G (109) T→G (285)	Q→R (5) I→V (37)	T→C (59) C→G (349)	V→A (20) R→G (117)	7	6
	7D	no mutations		T→G (140)	I→S (47)	A→G (54) C→G (349)	R→G (117)	3	2
	7G	C→G (421) A→G (506) T→C (687)	R→W (141) Y→C (169)	A→T (170) T→C (285)	E→V (57)	A→T (81) C→G (349)	R→G (117)	7	4
	8C	T→A (288) A→G (366) A→T (500) A→G (567)	D→E (96) D→V (167)	A→G (562)	M→V (188)	T→A (120) C→G (349)	R→G (117)	7	4
Lib 0.1	9C	T→A (128) A→G (448)	M→K (43) T→A (150)	A→G (66) A→C (505)	S→R (169)	A→G (348) C→G (349)	R→G (117)	6	4
	9E	T→C (106) T→C (308) T→A (379)	F→L (36) L→S (103) Y→N (127)	A→G (11) T→C (72)	D→G (4)	C→G (349)	R→G (117)	6	5

The position of the nucleotide or amino acid in the respective gene or protein sequence is indicated between brackets.



Figure 3.7: Distribution of nucleotide changes across the operon of thermostabilised NHases. Non-synonymous mutations are indicated in blue and synonymous mutations in pink.

However, depending on the conformation of the amino residue in the protein structure, a charged amino acid can be involved in different thermostabilising electrostatic interactions e.g. hydrogen bonds, salt bridges, charged-helix dipoles (Kumar *et al.*, 2000). Also, these stabilising interactions can only occur if there is a suitable amino acid in the surrounding environment with which the charged amino acid can interact. This highlights the need for structural data to be able to accurately identify the interactions that confer improved thermostability to the NHase mutants. Therefore, some of the mutants were subjected to structural analysis using X-ray crystallography in order to correlate the observed improvements in thermostability with the structural effects of the mutations (Chapter 5). The mutants were also subjected to thermal inactivation kinetics to assess the extent of their improved thermostability compared to the wild-type NHase (Chapter 4).



CHAPTER 4

KINETIC STABILITY OF THERMOSTABILISED MUTANT NHASES

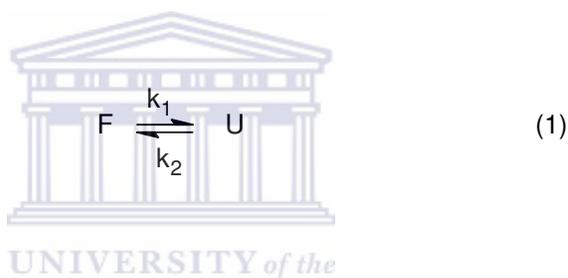
CONTENT

4.1.	INTRODUCTION.....	77
4.2.	PRELIMINARY ANALYSIS OF THERMOSTABILISED MUTANT NHASES.....	81
4.3.	KINETIC STABILITY OF RANDOMLY CREATED NHASE MUTANTS	83
4.4.	KINETIC STABILITY OF SITE-DIRECTED NHASE MUTANTS	86



4.1. INTRODUCTION

Enzyme thermostability incorporates both thermodynamic and kinetic stability factors (Vieille and Zeikus, 2001). These two parameters are used to describe different aspects of temperature-dependent protein unfolding. Thermodynamic stability is relevant if protein unfolding is reversible and can be described using the two-state model (1). According to the two-state model, proteins unfold from the native, folded state (F) to the reversibly unfolded state (U) without the formation of folding intermediates (Sternier and Liebl, 2001). The rate constants for unfolding and folding are given by k_1 and k_2 , respectively.



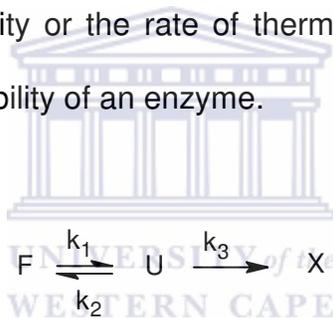
Consequently, thermodynamic stability is a measure of the difference in the free energies between the folded and unfolded states and is defined by the free energy of stabilisation, ΔG . The equilibrium constant $K_{eq} = (k_1/k_2) = [U]/[F]$ can be used to determine the ΔG according to the equation

$$\Delta G = -RT \ln K_{eq} \quad (2)$$

where R is the universal gas constant and T is the absolute temperature. The difference in thermodynamic stability of proteins is often expressed in terms of the melting temperature (T_m), i.e. the temperature at which 50% of the protein is unfolded (i.e. where, $\Delta G=0$ and $K_{eq}=1$). Methods used experimentally to measure

thermodynamic stability include differential scanning calorimetry, circular dichroism, tryptophan fluorescence or changes in absorbance (Polizzi *et al.*, 2007).

However, most proteins do not unfold reversibly and therefore it is impossible to measure their thermodynamic stability (Sternier and Liebl, 2001). Rather, most proteins undergo a transition state which involves either aggregation or the partial unfolding into scrambled structures which are kinetically stable (Vieille and Zeikus, 2001). Therefore, an additional step is incorporated into the model where X is the irreversible unfolded state and k_3 is the rate constant for the U to X transition (3). In such cases the kinetic stability or the rate of thermal inactivation is used as the determinant of the thermostability of an enzyme.



In contrast to thermodynamic stability, which is dependent on the ΔG , kinetic stability is dependent on the activation energy (E_a) of unfolding (ΔG^*) from F to U (Figure 4.1). The measurement of kinetic stability of proteins usually involves assaying for enzyme activity (Polizzi *et al.*, 2007) and is most commonly expressed as the overall observed deactivation constant ($k_{d,obs}$) for the irreversible unfolding from F to X where $k_d = (k_1 \cdot k_3) / (k_2 + k_3)$. However, since only U undergoes irreversible unfolding, the upper limit of k_d is determined by k_1 (Sternier and Liebl, 2001). This concept has important consequences for thermostable proteins because at high

temperatures, irreversible processes such as aggregation are expected to proceed quickly so that $k_3 \gg k_2$ and $k_d \sim k_1$.

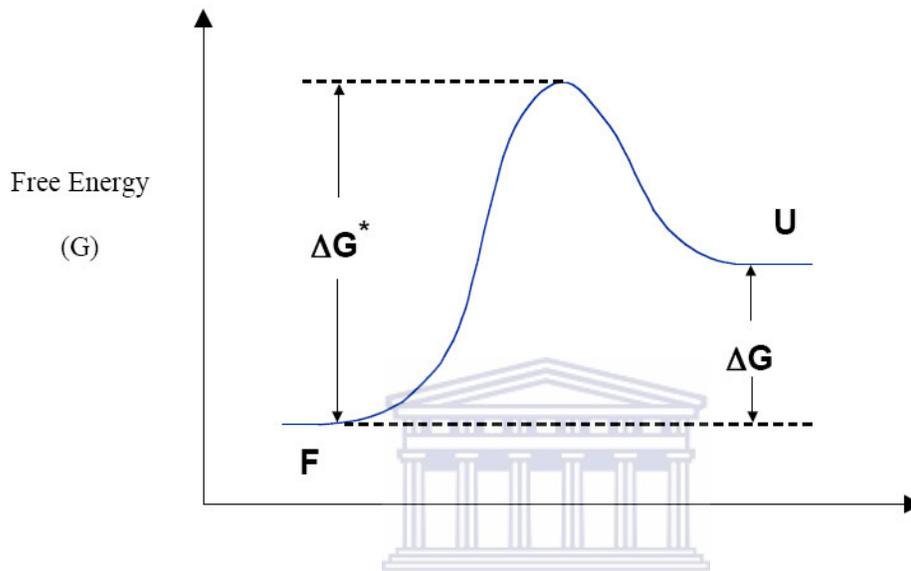


Figure 4.1: Change in activation energy of a protein from the folded to the unfolded state.

In industrial bioprocesses, kinetic stability (i.e., the rate of irreversible unfolding) is a key determinant of biocatalyst functionality (Eijsink *et al.*, 2005). Considering the potential industrial application of NHases, it was quite useful to use kinetic stability to assess and compare the thermostability of the wild-type and mutant NHases. Since random mutagenesis was used to engineer improved thermostability, any potential positive amino acid change was likely to have a local effect on the thermostability of the protein; i.e., to affect the partial unfolding step that occurs before the enzyme undergoes permanent denaturation.

The difference in difference kinetic stability between wild-type and mutant enzymes can be described by the difference in the free energy of activation of thermal unfolding ($\Delta\Delta G$). The $\Delta\Delta G$ can be derived from the Arrhenius equation by considering the relationship between the rate of thermal inactivation and the activation energy as follows:

$$k_d = Ae^{-(E_a/RT)} \quad (4)$$

Since ΔG^* is the activation energy of thermal unfolding it follows that:

$$k_{d,wt} = Ae^{-(\Delta G_{wt}/RT)} \quad \text{and} \quad k_{d,mut} = Ae^{-(\Delta G_{mut}/RT)}$$

$$\text{Then } \ln(k_{d,wt}) = \ln(A) - \frac{\Delta G_{wt}}{RT} \quad \text{and} \quad \ln(k_{d,mut}) = \ln(A) - \frac{\Delta G_{mut}}{RT}$$

$$\therefore \ln(k_{d,wt}) - \ln(k_{d,mut}) = -\frac{\Delta G_{mut}}{RT} + \frac{\Delta G_{wt}}{RT}$$

$$\therefore RT \cdot \ln\left(\frac{k_{d,wt}}{k_{d,mut}}\right) = \Delta G_{wt} - \Delta G_{mut} = \Delta\Delta G \quad (5)$$

4.2. PRELIMINARY ANALYSIS OF THERMOSTABILISED MUTANT NHASES

The enzymes isolated from screening the randomly mutated library were subjected to preliminary analysis to confirm whether they were more thermostable than the wild-type. Therefore, each thermostable mutant was subjected to thermal inactivation for a period 10 minutes at temperatures ranging from 60 and 70 °C. The residual enzyme activity was then determined, as an indicator of the thermostability of the enzymes. Figure 3.6 shows the percentage residual activity of the wild-type and mutant NHases as a function of temperature.

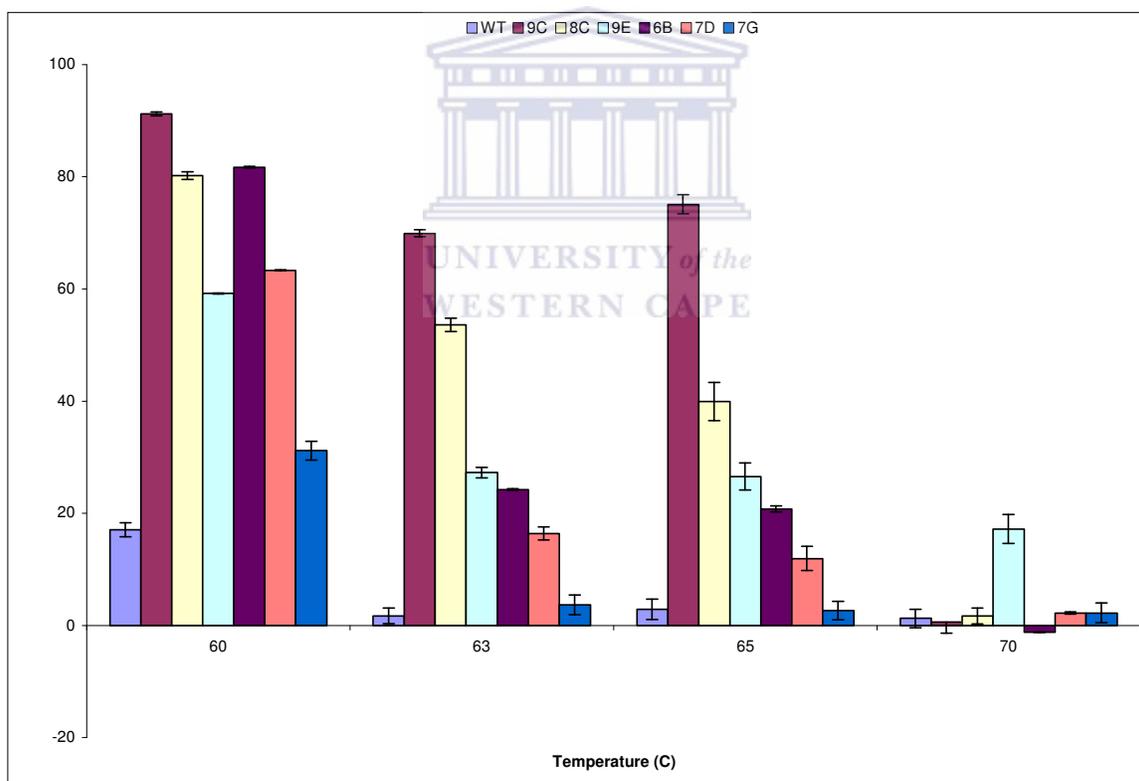


Figure 4.2: The effect of temperature on the extent of partial thermal inactivation on wild-type and mutant NHases. The enzyme samples were incubated in the absence of substrate for 10 minutes at the indicated temperatures. Residual activity was determined under standard assay conditions. The error bars indicate the standard deviation between triplicate experiments.

All of the mutant NHase isolated as a result of random mutagenesis were confirmed to be more thermostable than the wild-type. Small incremental changes in the thermostability of the mutants with respect to the temperature of unfolding were observed (i.e. between 3 and 5°C). However, the extent of the improvement in thermostability of the mutants compared to the wild-type at a specific temperature was quite substantial. Mutant 9C retained approximately 75% activity at temperatures up to 65°C. No activity was detected following thermal inactivation at 70°C (Fig. 4.2). Mutants 8C and 6B exhibited similar thermostabilities at 60°C. However, mutant 8C was more thermostable at 63 and 65°C, retaining approximately 54 and 40% residual activity following partial thermal treatment at these temperatures. Likewise, mutants 7D and 9E exhibited similar thermostabilities at 60°C where both retained approximately 60% residual activity. The percentage residual activity of mutant 7D declined to 17% and 11% after heat treatment at 63 and 65°C, respectively. In contrast, mutant 9E retained close to 30% activity up to 65°C. Mutant 9E was also the only mutant that retained some residual activity (17%) following thermal treatment at 70°C. Mutant 7G only exhibited slight improvement in thermostability at 60°C where it retained approximately 30% residual activity compared to 20% for the wild-type NHase.

Improving the thermostability of a protein is usually the result of the accumulation of thermostabilising amino acid substitutions where each individual substitution only marginally increases the unfolding temperature of a protein (Kuchner and Arnold, 1997). However, there is also experimental evidence that moderate and

high mutation frequency libraries yield more functionally improved isolates than low mutation frequency libraries (Daugherty *et al.*, 2000). This difference is likely due to mutants from moderate and high mutation frequency libraries having more than one mutation that contribute to the improved function. Since the mutants from this study were isolated from moderate frequency libraries, it was likely for a mutant to contain more than one thermostabilising mutation following only one round of random mutagenesis. This could explain the substantial thermostabilisation of the mutants compared to the wild-type. The variation of the thermostability of the mutants with temperature could be indicative of the different mechanisms involved in conferring improved thermostability. In addition, it could also be a reflection of differences in the relative importance of the location of the amino acid change in the protein structure. All the mutant enzymes were subjected to further kinetic analysis of thermal inactivation in order to more accurately compare the thermostability of wild-type and mutant NHases.

4.3. KINETIC STABILITY OF RANDOMLY CREATED NHASE MUTANTS

Determinations of kinetic stability were used to compare the thermostability of the mutant and wild-type NHases. Since some of the improved mutants were relatively stable at 60°C (Figure 4.2), thermal inactivation kinetics was performed at 63°C. The wild-type and mutant NHases exhibited first order thermal inactivation kinetics (Figure 4.3).

The linear regressions in Figure 4.3 represent fits to the equation

$$\ln(\% \text{residual activity}) = -k_d \times t \quad (6)$$

where k_d is the rate constant for the thermal inactivation and t is the time of incubation. The rate of thermal inactivation (k_d) (derived from the slope of the linear regressions in Figure 4.3) was used to calculate the $\Delta\Delta G$ for each thermostable mutant NHase relative to the wild-type according to Equation 5.

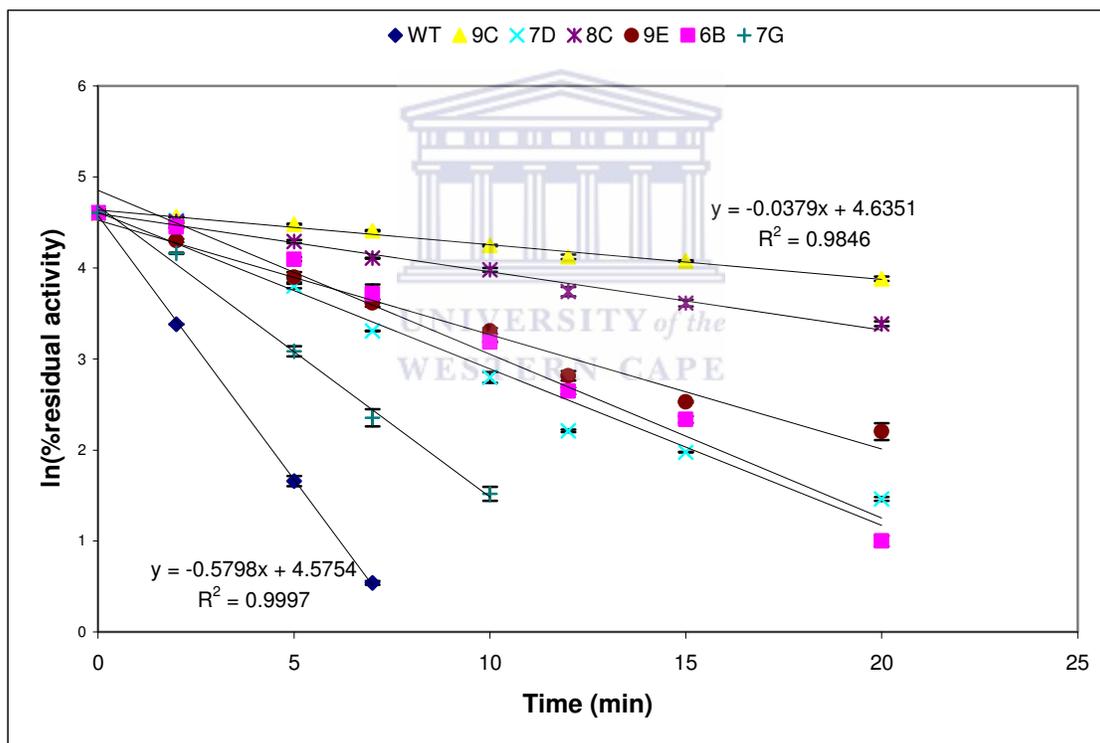


Figure 4.3: Thermal inactivation of wild-type and randomly constructed mutant NHases. The cell free extract (CFE) of each sample was incubated without substrate at 63°C for different periods of time. The percentage residual activity was determined under standard assay conditions. The error bars indicate the standard deviation between triplicate measurements. For clarity only the linear regression formulas and R^2 values of the wild-type and mutant 9C is displayed.

The kinetic stability parameters of the thermal inactivation of the wild-type and mutant NHases at 63°C are summarised in Table 4.1. The lower thermal inactivation rate constants of the mutants compared to the wild-type NHase indicated that the rates of thermal unfolding of all the mutants were slower than that of the wild-type NHase. Mutants 8C and 9C exhibited substantially slower unfolding kinetics, being more thermostable than the wild-type enzyme by 9 and 15 fold, respectively. The $\Delta\Delta G$ values for these mutants were 6.16 and 7.62, respectively. The rest of the mutants were between 2 and 5 fold more thermostable than the wild-type with an associated stabilisation energy of between 1.51 and 4.27 kJ/mol.



Table 4.1: Kinetic parameters for the thermal inactivation of wild-type and mutants NHases at 63°C.

	k_d (min^{-1})	$k_{d,wt}/k_{d,mut}$	$\Delta\Delta G$ (kJ/mol)
WT	0.579	-	-
6B	0.1707	3.4	3.42
7D	0.1719	3.3	3.40
7G	0.3372	1.7	1.51
8C	0.0639	9.1	6.16
9C	0.0379	15.3	7.62
9E	0.1256	4.6	4.27

4.4. KINETIC STABILITY OF SITE-DIRECTED NHASE MUTANTS

If multiple mutations occur in one generation it is difficult to differentiate between beneficial and non-beneficial mutations (Acharya *et al.*, 2004). Since all but one of the thermostable mutants contained more than one amino acid change it was decided to perform some preliminary rationale design experiments using site-directed mutagenesis (SDM) in order to ascertain which residue conferred improved thermostability. However, rationale design is a laborious and expensive process and therefore, this strategy was only implemented for one of the randomly created mutants. In order to select a multiple thermostabilised mutant for SDM, the amino acid residue changes in each mutant were assessed based on the type of amino acid change; i.e., amino acids which are more likely to enhance thermostability by the insertion of additional intra- or intermolecular bonds, as well as the location of the mutation i.e. the context of the amino acid within the secondary structure.

On the basis, mutant 9C was selected for SDM. This mutant contained 3 amino acid residue changes, two of which involved charged amino acids (Table 3.4). Charged amino acids are very important in electrostatic interactions such as salt bridges and hydrogen bonds (Vogt *et al.*, 1997). The S→R (α 169) mutation was especially interesting because the chemical and molecular properties of arginine are known to facilitate more stabilising interactions than other charged amino acids (Vieille and Zeikus, 2001). The S→R (α 169) mutation was also seen to occur on a loop (i.e., an area of undefined secondary structure) using molecular graphics

software. The shortening and stabilisation of loops is one of the mechanisms involved in conferring thermostability to enzymes (Kumar *et al.*, 2000).

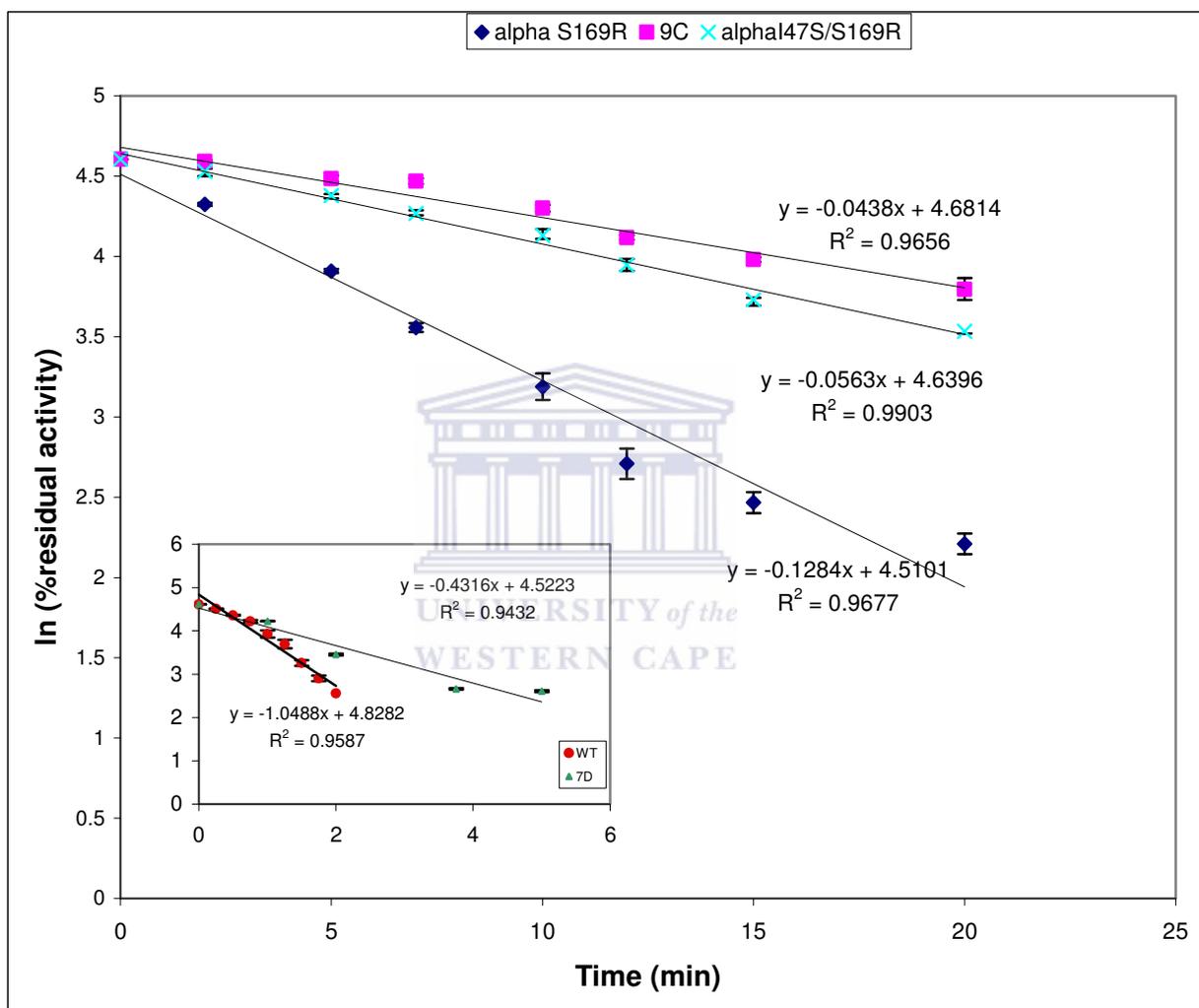


Figure 4.4: Comparison of thermal inactivation of wild-type (insert), mutants 7D (insert), mutant 9C and site-directed mutant NHases. The enzyme samples were incubated without substrate at 65°C for different periods of time. The residual activity determined under standard assay conditions. Error bars indicate the standard deviation between triplicate measurements.

The single site directed mutant, α S169R, was constructed using the wild-type NHase gene as template in order to determine if the α R169 residue conferred improved thermostability to mutant 9C. The thermal inactivation analysis was conducted at 65°C in order to clearly observe the difference in the temperature-dependent unfolding rates of mutant 9C and α S169R. The kinetic stability parameters of the thermal inactivation of mutant 9C and the SD mutant NHases at 65°C are summarised in Table 4.2. The linear regressions illustrate that the kinetics of thermal unfolding of the single mutant, α S169R was 3 fold faster than mutant 9C. The thermal inactivation rate constant for α S169R was 0.1284 compared to 0.0438 for mutant 9C. This indicated that the S→R (α 169) mutation was not solely responsible for the improved thermostability of mutant 9C.

Based on the DNA sequence data, the other amino acid change in mutant 9C that could potentially contribute to the improved thermostability of this mutant was the M→K (β 43) mutation. Since this is also a charged amino acid it is possible that it could also contribute to the thermostabilisation of mutant 9C by increased electrostatic interactions. Although it was possible to test this hypothesis by constructing the proposed site-directed mutant, it would be inefficient to use rational design to effectively reproduce results already obtained from random mutagenesis. The suitable alternative was to subject mutant 9C to crystal structure analysis and anticipate that this would provide detailed insight into the mechanisms involved in conferring improved thermostability to this NHase.

Table 4.2: Kinetic parameters for the thermal inactivation of wild-type and mutant NHases at 65°C.

	k_d (min^{-1})	$k_{d,wt}/k_{d,mut}$	$\Delta\Delta G$ (KJ/mol)	$k_{d,\alpha S169R}/k_{d,mut}$	$\Delta\Delta G$ (KJ/mol)
WT	1.0488	-	-	-	-
$\alpha S169R$	0.1284	8.2	5.9	-	-
Mut 9C	0.0438	23.9	8.9	2.9	3.0
Mut 7D	0.4313	2.4	2.5	-	-
$\alpha I47S/\alpha S169R$	0.0563	18.6	8.2	2.3	2.3

Once of the main expected outcomes of this project is that the structural analysis of the mutants will identify the key amino acid residues involved in the improved thermostabilisation. This would aid the future engineering of an “extremely” thermostable NHase. However, in order to combine all the thermostabilising mutations in one “extremely” thermostabilised mutant requires that the mutations are additive. The thermostabilisation effect of individual amino acids substitutions are likely to be independent and thus additive provided there is no interactions between the substituted residues (Bogin *et al.*, 2002; Zang *et al.*, 1995).

Since mutant 7D was the only mutant that contained only a single thermostabilising mutation, the double mutant, $\alpha I47S/\alpha S169R$ was constructed using the plasmid encoding mutant 7D as template. The double mutant was stabilised by 8.2 kJ/mol compared to the wild-type NHase whereas mutant 7D was stabilised by 2.5 kJ/mol and the site-directed mutant, $\alpha S169R$ was stabilised by 5.9 kJ/mol (Table 4.2). This indicated that the enhanced effect of kinetic unfolding as a result of these two

mutations was approximately additive. Also, the thermal inactivation rate constant of the site directed double mutant α I47S/ α S169R was in the same range as mutant 9C. Therefore, if all the mutations that confer improved thermostability to the randomly created NHase mutants could be identified and engineered into a single mutant, it is possible that this multiple mutant would be even more thermostable than the most thermostable isolate (mutant 9C). It is acknowledged that this conclusion ignores other effects such as conformational changes that could have a positive effect in one mutant but a negative effect when combined with another mutation.



CHAPTER 5

STRUCTURAL EVALUATION OF THERMOSTABILISED MUTANTS

CONTENT

5.1.	INTRODUCTION.....	92
5.2.	PROTEIN PURIFICATION AND CRYSTALLISATION.....	93
5.2.1.	Protein purification.....	93
5.2.2.	Crystallisation	95
5.3.	X-RAY CRYSTALLOGRAPHY.....	97
5.3.1.	Diffraction data collection	97
5.3.2.	Structure refinement and validation.....	100
5.4.	ANALYSIS OF CRYSTAL STRUCTURES OF IMPROVED MUTANTS	102
5.4.1.	Mutant 7D (α l47S).....	102
5.4.2.	Mutant 8C (α M188V / β D96E / β D167V).....	105
5.4.3.	Mutant 9C (α S169R / β M43K / β T150A)	113
5.4.4.	Mutant 9E (α D4G / β F36L / β L103S / β Y127N)	117

5.1. INTRODUCTION

X-ray crystallography is fast becoming the method of choice for determining the molecular structures of biological macromolecules. The objective of macromolecular X-ray crystallography is to use a protein crystal in order to obtain a three dimensional molecular structure of the protein (McPherson, 1999). The rate-limiting step is often the presence or availability of diffracting quality crystals that can be used to obtain the data on which the structural determination will be based. The production of crystals suitable for X-ray diffraction analysis is dependent on a high quality, homogeneous, soluble protein sample (Smyth and Martin, 2000). Once suitable crystals are available, the next step in protein structure determination is the actual diffraction experiment. In the course of the X-ray diffraction data collection experiment, a complete three-dimensional set of reflections should be recorded; these have to be phased and transformed into maps which has to be interpreted by the construction of an atomic model. The theory of X-ray crystallography and its application to macromolecular crystals are well established and widely published e.g. Blundell and Johnson, 1976; Mccree, 1993; Drenth, 1994.

In this chapter, the purification, crystallisation and structure determination of wild-type and improved thermostable mutant NHases are presented. The aim was to obtain three-dimensional structural models for all the mutants, which could then be used to correlate the experimentally observed improvement in thermostability with the structural effects of the mutations.

5.2. PROTEIN PURIFICATION AND CRYSTALLISATION

5.2.1. Protein purification

The *E. coli* BL21 (DE3)/pET21 expression system was used for the expression of recombinant NHases. This expression system resulted in the over-expression of more than adequate amounts of protein to proceed with a purification protocol.

The wild-type and all mutant NHases were purified under similar conditions to the wild-type as described by Tsekoa (2005). SDS-PAGE was used to visually assess the extent of each purification step. Figure 5.1 shows the SDS-PAGE analysis of purification process for one of the thermostable mutants. Since the wild-type NHase was known to be moderately thermostable (Pereira *et al.*, 1998) and the mutants exhibited improved thermostability over the wild-type, the first step in the purification process was the heat treatment of the crude protein sample at 55°C for 30 minutes. This step resulted in the removal of a large fraction of the *E. coli* thermolabile proteins (Figure 5.1a).

The protein sample was then subjected to ammonium sulphate precipitation at 25% w/v. Although this step only resulted in the removal of a small fraction of *E. coli* proteins, the resulting protein extract was in the appropriate buffer for hydrophobic interaction chromatography (HIC). The protein sample was then further purified to near homogeneity with HIC followed by anion exchange chromatography. HIC resulted in substantial purification of the NHase mutant and only one major contaminant was observed (Figure 5.1b).

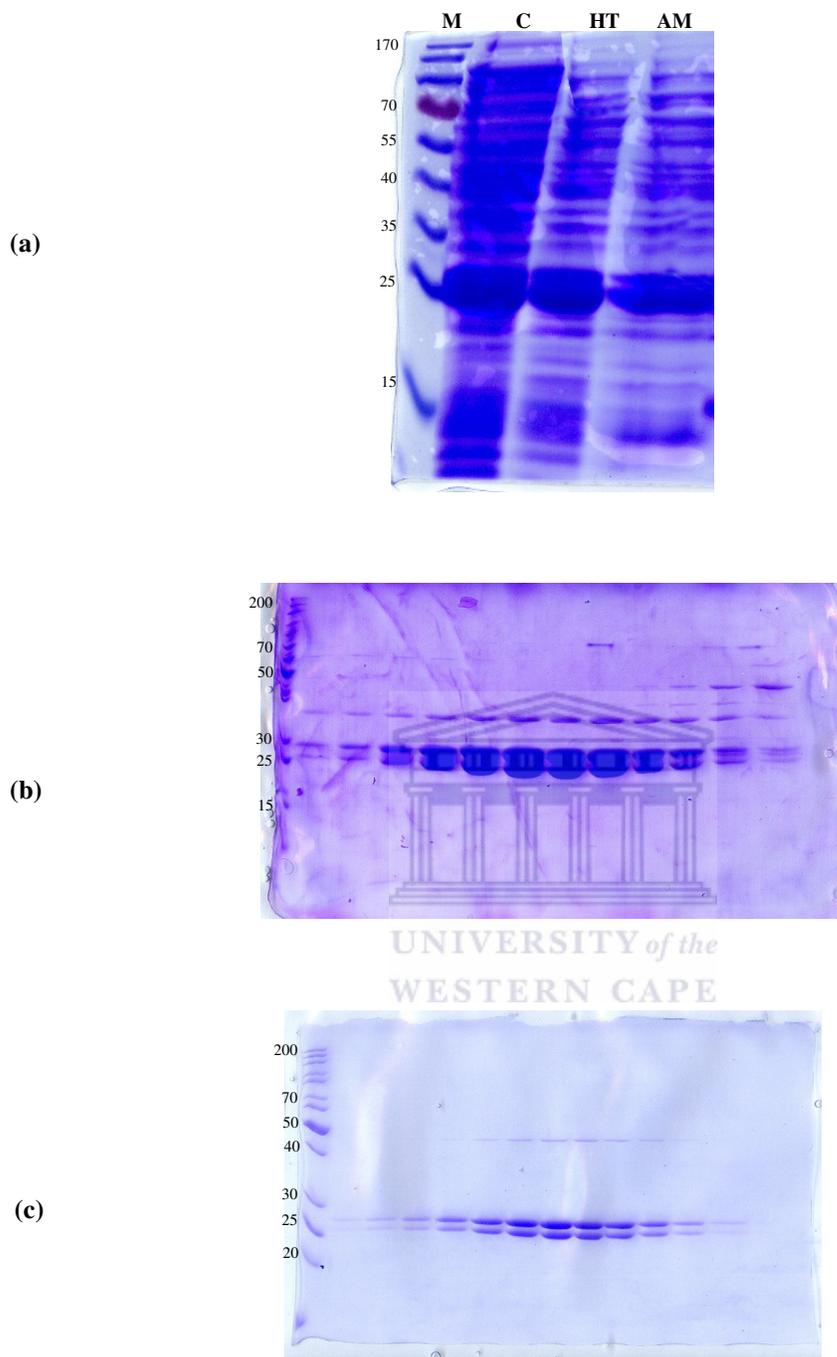


Figure 5.1: SDS PAGE analysis the NHase mutants at each step of the purification process. (a) Initial purification treatment of cell free extract where M=Molecular weight marker, C=Cell free extract, HT=Heat treatment, AM=ammonium sulphate precipitation. (b) Five microlitre volumes of each protein fractions obtained from phenyl-Sepharose hydrophobic interaction chromatography (HIC). (c) Five microlitre volumes of every second fraction obtained from Q-Sepharose chromatography.

Following anion exchange chromatography, the purified protein sample was assessed to be nearly homogeneous by SDS-PAGE (Figure 5.1c). All protein samples were treated identically and similar results were obtained.

5.2.2. Crystallisation

Prior to crystallisation, the purified NHases were dialysed against 20 mM Tris-HCl (pH 7.2), filtered through a 0.22 μm filter and concentrated to a range of between 10 mg/ml and 40 mg/ml. The hanging drop vapour diffusion method was used in order to obtain good quality crystals that were suitable for X-ray diffraction. Crystallisation of mutant NHases were set up as previously described by Tsekoa *et al.* (2004) and incubated at 22°C (section 2.9.1). The wild-type NHase was also crystallised under identical conditions as a control.

Crystals of good form and size were obtained for the both the wild-type and mutant NHase (Figure 5.2). Initial crystals were observed within three to four days following incubation at 22°C and reached maximum dimensions in approximately four weeks. The crystals used for x-ray diffraction were relatively large; i.e., approximately 0.3 mm in the largest dimension. Crystal size is a very important factor in X-ray crystallography, as the intensity of X-ray diffraction is proportional to the volume of the diffracting matter (McPherson, 1999). The acceptable size of a protein crystal used for a typical diffraction experiment is generally in the region of 0.2 - 0.5 mm (Smyth and Martin, 2000) although data from much smaller crystals can also be collected at synchrotron facilities and may produce better quality data.

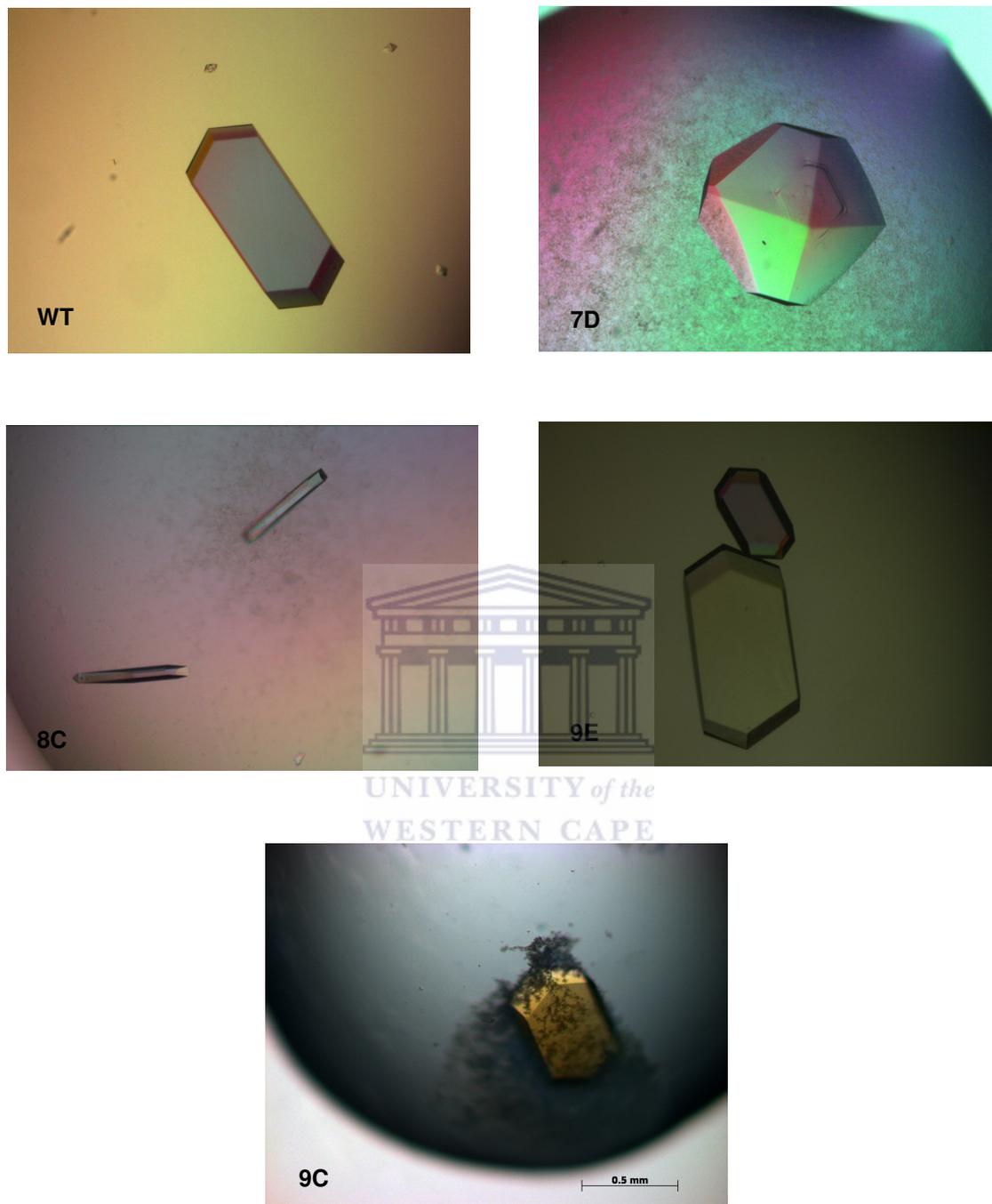


Figure 5.2: Crystals obtained for the wild-type (WT) and mutant NHases (7D, 8C, 9E, 9C) using the hanging drop vapour diffusion method.

5.3. X-RAY CRYSTALLOGRAPHY

5.3.1. Diffraction data collection

The crystals selected for X-ray diffraction data collection were approximately 0.3 mm in the largest dimension. The NHase crystals were previously shown to belong to the P4₁2₁2 space group with unit cell dimensions a=b=106.61 Å, c=83.23 Å and cell angles α=β=γ=90.0° (Tsekoea *et al.*, 2004). This information was important to devise a strategy that would enable the collection of a complete set of unique reflections.

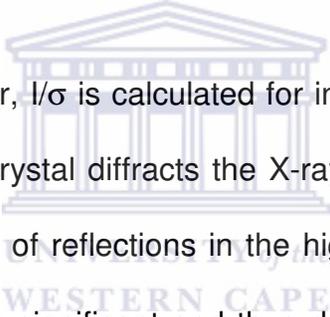
The collected diffraction images were reduced to a list of reflections with their hkl indices and intensities using the software package d*TREK (Pflugrath, 1999). The symmetry equivalent reflections were then merged together, to give an average intensity value to each reflection in a data set of purely unique reflections. These symmetry equivalent reflections give an indication of the quality of the data; a merging R-factor is computed to measure the internal agreement of the diffraction data.

$$R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_{i(hkl)} - \langle I_{(hkl)} \rangle|}{\sum_{hkl} \sum_i I_{i(hkl)}} \times 100\% \quad (1)$$

where $\langle I_{(hkl)} \rangle$ and $I_{i(hkl)}$ are the mean intensity and intensity of the i^{th} measurement.

Since the R_{merge} value increases with increasing resolution (because the signal becomes weaker), the processed data are usually divided in resolution shells.

Normally, only data with R_{merge} levels below 40% are used, thus limiting the resolution to a value corresponding to data with high quality merging statistics. The unique reflections lie in one portion of reciprocal space. Ideally 100% of these data need to be collected (referred to as the 'completeness' of the data). High redundancy, the ratio between the total number of observed reflections and the total number of unique reflections, can improve the accuracy of the data as a more accurate mean intensity of each reflection can be calculated. Higher symmetry crystals give more symmetry-related reflections over a certain rotation range than low symmetry ones and so give more highly redundant data.



The signal to noise parameter, I/σ is calculated for individual resolution shells and measures how strongly the crystal diffracts the X-ray beam. An average value of 3σ is required in the majority of reflections in the highest resolution shell. Values below 3σ are statistically less significant and thus define the maximum resolution limit of the data collection. The signal to noise parameter and R_{merge} are correlated and both will dictate the resolution cut-off. However, R_{merge} values are regarded as the more important of the two if the signal to noise ratio of the highest resolution shell starts to fall below 3σ but R_{merge} remains below 30%. In this case the resolution shell is generally kept for data processing.

The overall data statistics, given in Table 5.1, show that the data for the wild-type and mutant NHases are of very high quality, at least in terms of Resolution, R_{merge} , Redundancy and overall completeness. The wild-type NHase diffracted to a

resolution of 1.4Å. This was higher than the resolution of 2.4Å previously reported by Tsekoa *et al.* (2004) for the *G. pallidus* RAPc8 NHase. Moreover, the diffraction data for all the mutants was better than that previously obtained for the wild-type. The best diffraction data was obtained for mutant 9E, which diffracted to a resolution of 1.15Å.

Table 5.1: X-ray data collection statistics for wild-type and mutant NHases

Data set	WT	9e	8c	7d	9c
Wave length (Å)	0.979	0.979	0.979	0.979	0.979
Space group	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2
Unit cell Parameters	a = 105.44 b = 105.44 c = 83.04	a = 105.83 b = 105.83 c = 83.72	a = 106.18 b = 106.18 c = 83.12	a = 106.5 b = 106.5 c = 83.0	a = 106.13 b = 106.13 c = 82.88
	α=β=γ= 90.00°	α=β=γ= 90.00°	α=β=γ= 90.00°	α=β=γ= 90.00°	α=β=γ= 90.00°
Mosaicity	0.71	0.68	0.62	0.79	0.62
Resolution Range (Å) (outer shell)	31.16 - 1.40 (1.45 - 1.40)	27.86 - 1.15 (1.19 - 1.15)	44.74 - 1.47 (1.52 - 1.47)	31.90 - 1.8 (1.86 - 1.80)	44.69 - 1.45 (1.50 - 1.45)
Total observations	759392	1016398	797902	615278	310298
Total observations (unique)	92064	184357	81002	44724	83831
Completeness (%) (outer shell)	99.9 (100.0)	99.6 (100)	100	100	99.8 (99.9)
Redundancy (outer shell)	8.25 (8.17)	5.51 (4.87)	9.85 (9.71)	13.76 (13.66)	3.70 (3.63)
Signal-to-noise ratio (I/σ(I)) (outer shell)	11.8 (3.8)	11.4 (2.7)	14.8 (5.0)	17.5 (5.4)	13.5 (3.0)
R _{merge} (outer shell)	0.064 (0.347)	0.049 (0.398)	0.056 (0.315)	0.059 (0.336)	0.038 (0.321)
Reduced ChiSquared	0.97 (0.87)	0.95 (1.38)	0.94 (1.39)	0.98 (1.47)	0.99 (1.48)
Wilson plot average B- factor (Å ²)	14	9.75	14.5	11.46	17.05
Matthew's coefficient	2.26	2.28	2.29	2.30	2.28
Solvent content	45.51	46.13	46.32	46.57	46.17

All structures were subsequently solved by molecular replacement with PHASER (McCoy *et al.*, 2007) using the previously solved wild-type NHase structure as a search probe (data not shown). Similar molecules can form crystals with either the same or very different packing arrangements within the unit cell. These differences

arise even for molecules with identical shapes because the packing is determined not only by shape but by intermolecular interactions (charge, H-bond, hydrophobic effect) at the molecular surfaces. Knowledge of the molecular structure of one crystal can be used to derive the structure of the same or similar molecule in another (“target”) crystal. If the crystals are isomorphic (having the same space group and cell constants) the known molecule can be directly fitted into the target crystal; otherwise the technique of Molecular Replacement (MR) is used to conduct a search for the fit of the known molecule into the target crystal by comparing their X-ray diffraction patterns. MR is widely used as a method for macromolecular structure determination and has proved to be particularly successful in recent years (reviewed in: Rossmann, 1972; Tickle and Driessen, 1996).

5.3.2. Structure refinement and validation

The initial model of a protein structure obtained by MR or any other crystallographic method is likely to be crude and to contain errors, and therefore will need to be improved. This is done through crystallographic refinement, which involves adjusting an approximate atomic model of the protein in order to find agreement between the calculated structure factor amplitudes and the observed structure amplitudes. The success of structure refinement and validation is determined by the R-factor and free R-factor (R_{free}) (Brunger, 1993). These values are indicative of the difference between calculated and observed structure factors obtained from the model and from the original diffraction data, respectively. Typically, a small portion of the data (5 to 10%) are excluded from the refinement process (the test set) and used to calculate the R_{free} . This value is a more reliable

cross-validation indicator than the R-factor (Kleywegt and Brunger, 1996). Successive cycles of refinement and model rebuilding were done and data were included to a resolution shell where $1/\sigma(I)$ was greater than 2.0. The refinement statistics (R-factor and R_{free}) obtained for the wild-type and all mutant structures as shown in Table 5.2 falls within the known limits for a crystal structure solved at the resolution range mentioned in this study (Kleywegt and Brunger, 1996).

Table 5.2: Refinement statistics for wild-type and mutant NHases

Data set	WT	9e	8c	7d	9c
Number of Amino Acids	A10-A211, B1-B227	A10-A211, B1-B229	A10-A211, B1-B227	A10-A211, B1-B227	A8-A211, B1-B227
Protein Molecular mass (kDa)	51.15	51.15	51.15	51.15	51.15
Number of non-hydrogen atoms	3715	4179	3744	3692	3829
Number of water atoms	213	684	265	212	323
Number of reflections					
Working set	92064	157601	76797	42263	79484
Test set (%)	5.00	5.04	5.02	5.16	4.99
R-factor§(%)	20.8	18.0	20.0	21.0	20.6
$R_{\text{free}}\#$ (%)	22.7	19.83	21.8	23.8	23.0
Rms deviations from ideality					
Bond lengths (Å)	0.008	0.008	0.009	0.016	0.009
Bond angles (°)	1.27	1.20	1.23	1.59	1.32
Average B value (Å ²)	13.87	10.46	14.72	21.23	17.84

Validation of the refined wild-type and mutant structures with WHATCHECK (Hoofst *et al.*, 1996) confirmed that the root mean square (rms) deviations from ideal bond lengths and angles (Table 5.2) were within acceptable limits. Generally acceptable rms deviations for crystal structures from macromolecules are 4° for bond angles and 0.02Å for bond lengths (Laskowski *et al.*, 1993)

5.4. ANALYSIS OF CRYSTAL STRUCTURES OF IMPROVED MUTANTS

5.4.1. Mutant 7D (α 47S)

The only possible stabilising mutation evident in mutant 7D was I \rightarrow S (α 47). As a result of this amino acid change, a water-mediated hydrogen bond forms between α S47 and α E33. A detailed image of the specific interactions involved in the hydrogen bond formation is shown in Figure 5.3. α S47 interacts with w202 through the hydroxyl oxygen ($O\gamma$). The distance between these atoms are 2.51Å. α E33 interacts with the structured water via $O\epsilon$ 1 of the carboxyl side chain. Here, the bond distance is 2.61Å between $O\epsilon$ 1 and w202. Previous measurements of the difference in stability between mutant 7D and the wild-type NHase showed a stabilisation of ($\Delta\Delta G$) of 3.40 kJ/mol. This is slightly lower than the range of 4.18 to 9.20 kJ/mol that an average hydrogen bond contributes to protein stability (Myers and Pace, 1996; Takano *et al.*, 1999).

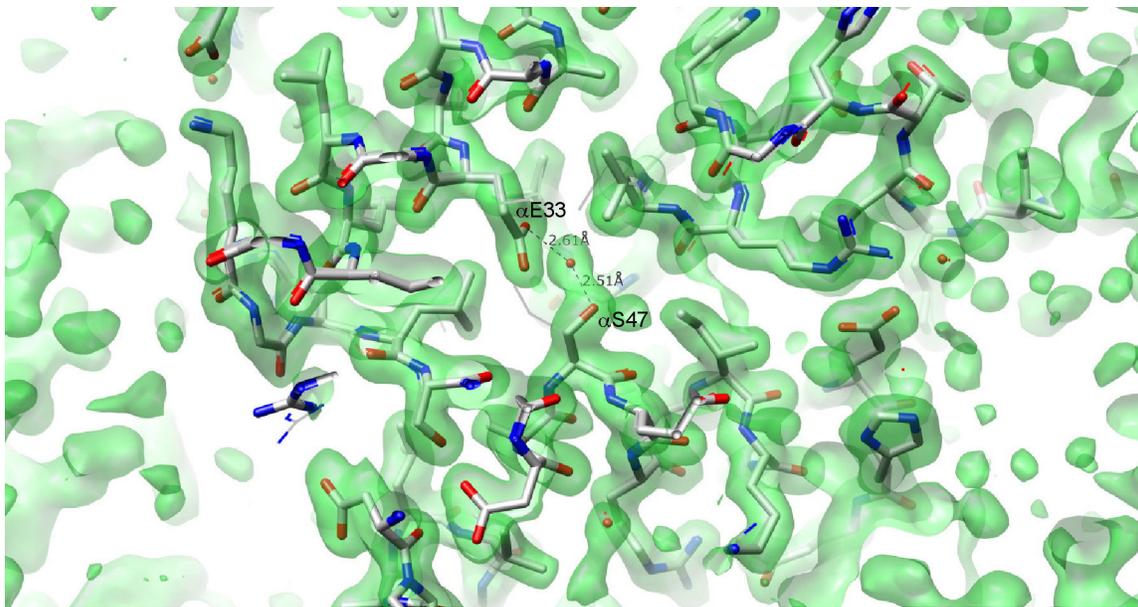


Figure 5.3: Electron density map showing the water-mediated hydrogen bond between α S47 and α E33 in mutant 7D. The electron density map was contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant. The amino acid residues are shown as sticks where the atoms are: C=white, O=red, N=blue, H₂O=red spheres.

An overview of the stabilising interaction between α S47 and α E33 in the context of the mutant 7D structure is shown in Figure 5.4. These two amino acid residues are found on the surface of the protein and are positioned on different α helices of the α subunit. The presence of polar residues (such as Ser) on protein surfaces are important for increasing hydrogen bonding networks that are implicated in protein thermostability (Zhou, 2002). The stabilising effect of polar residues is often mediated through water molecules localised on the protein surface (Vogt *et al.*, 1997). Structured or ordered waters are also considered important for protein stability (Levy and Onuchic, 2006). Together with Thr, Ser is known to be the best residue for interacting with water molecules (Mattos, 2002).

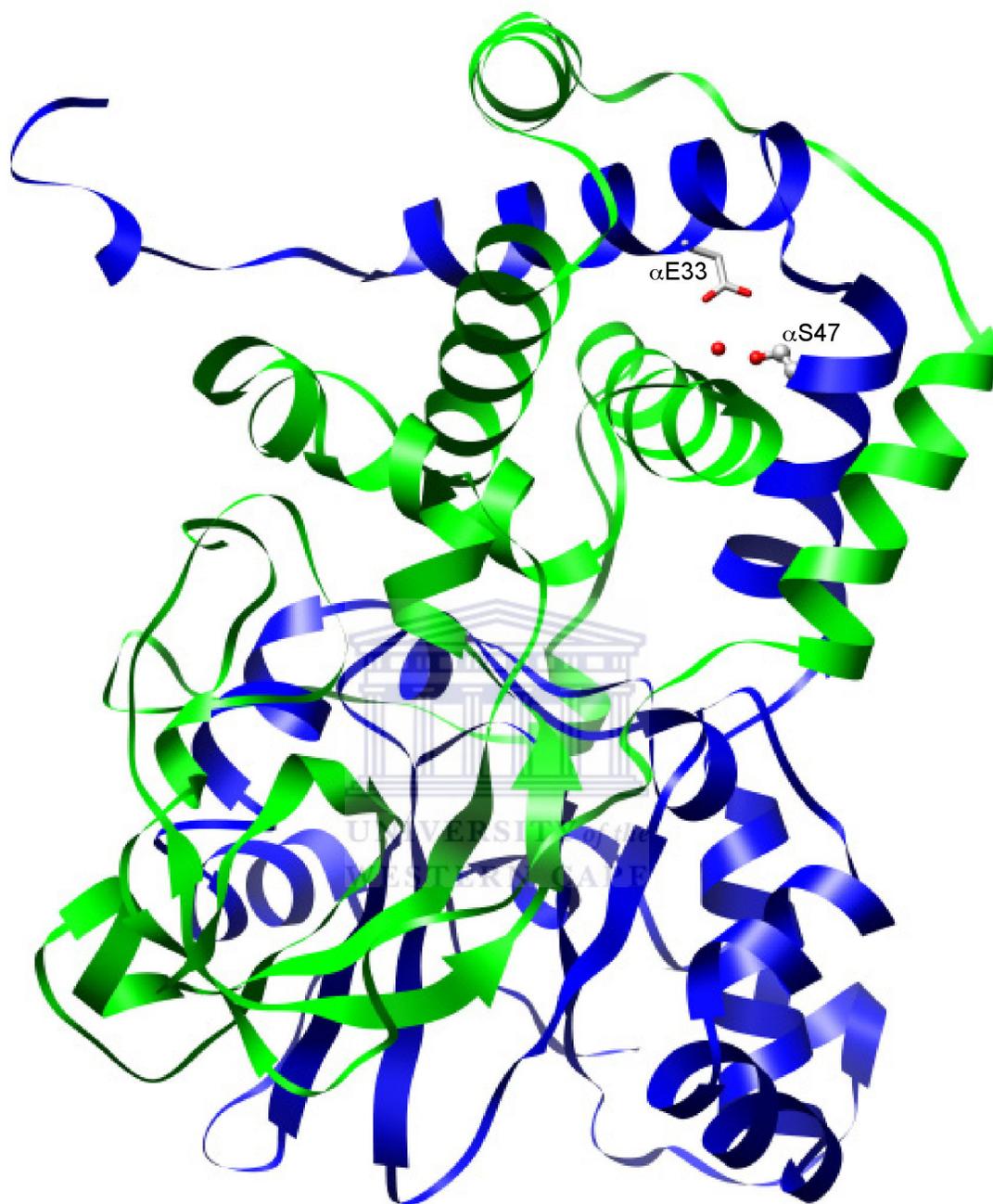


Figure 5.4: Three-dimensional structure of mutant 7D showing the position of W202 between α S47 and α E33 on the α β 2 heterotetramer. The α subunit is coloured blue and the β subunit green. The relevant mutations are shown in ball and stick where the atoms are: C=white, O=red, N=blue, H₂O=red spheres.

5.4.2. Mutant 8C (α M188V / β D96E / β D167V)

An overview of all the mutations observed in the mutant NHase structure is shown in Figure 5.5.

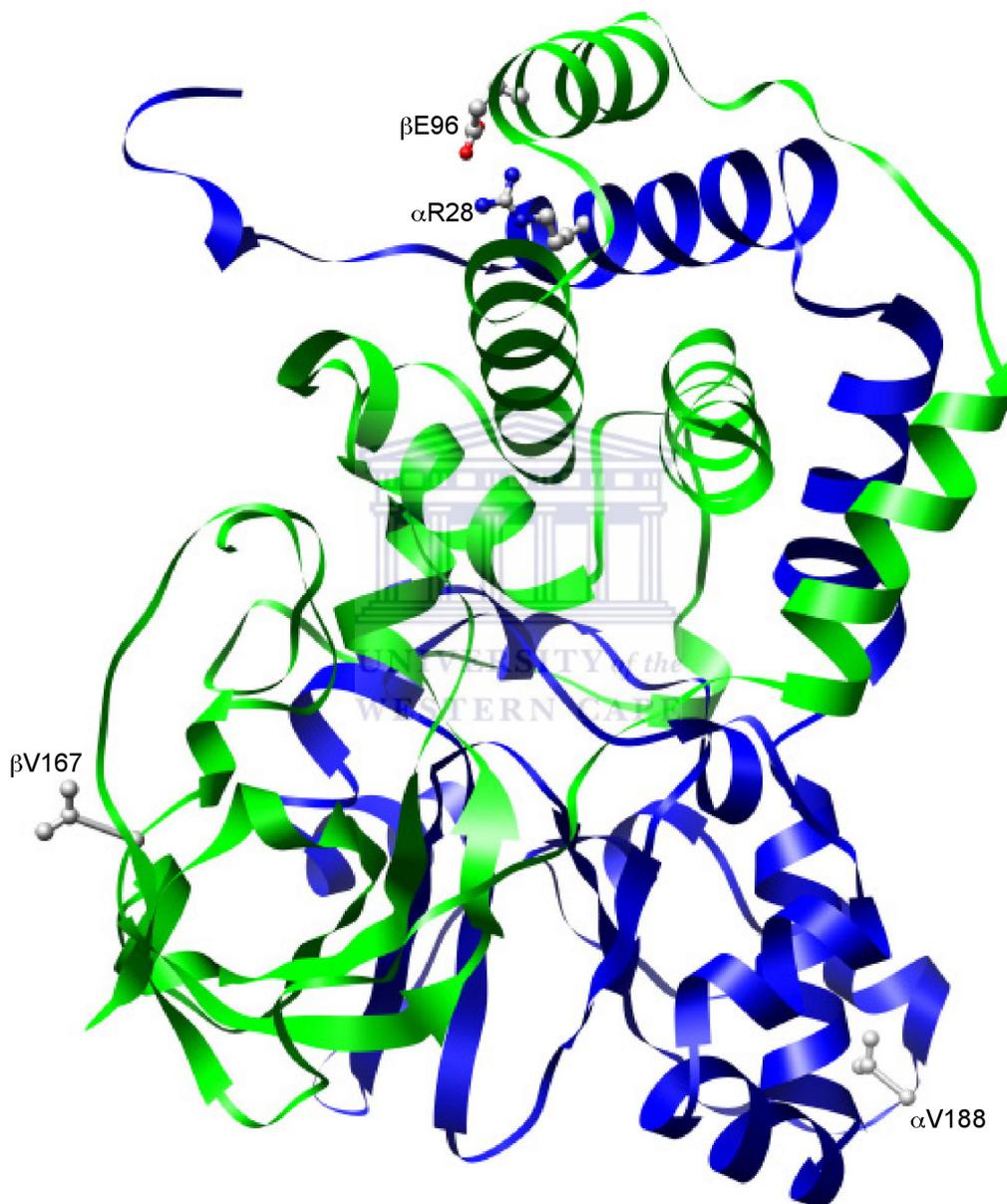
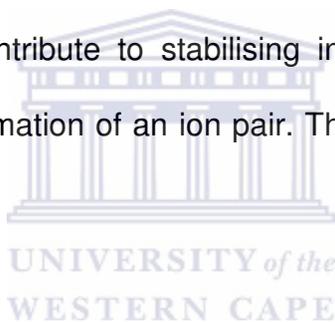


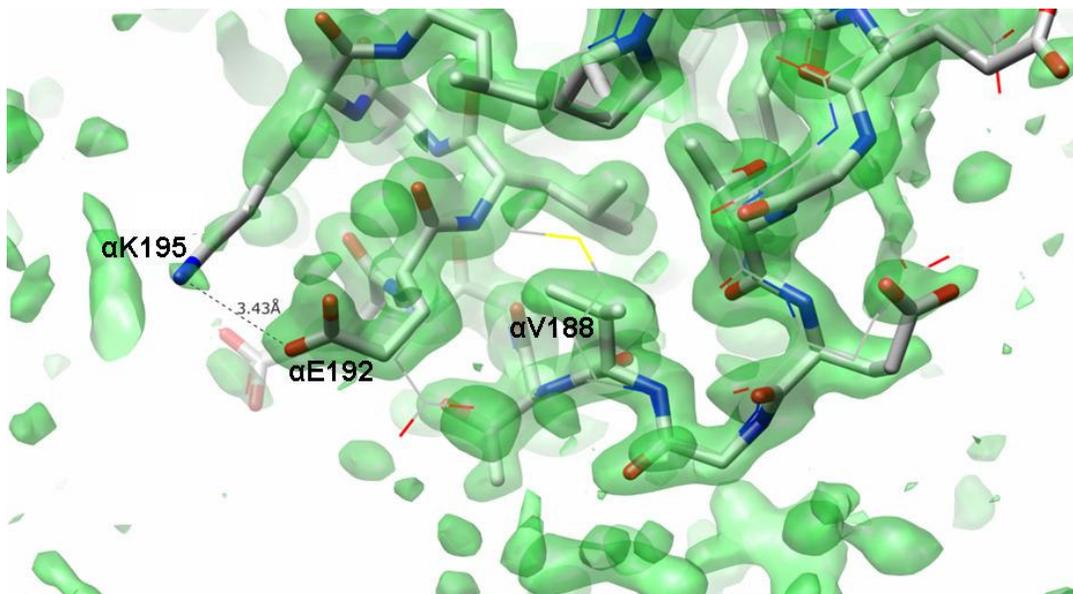
Figure 5.5: Three-dimensional structure of mutant 8C showing the positions of α R28, β E96, β V167 and α V188 on the α β β heterotetramer. The α subunit is coloured blue and the β subunit green. The relevant mutations are shown in ball and stick where the atoms are: C=white, O=red, N=blue

The M→V (α 188) mutation resulted in a 3D conformational shift of α E192 (Figure 5.6). The average atomic displacement parameters or the temperature factors (B values) at position α 188 were 26 and 29 for the wild-type and mutant respectively, suggesting that these residues are well-defined in the electron density maps. In a high-quality model the B factors reflect the mobility or flexibility of various parts of the molecule (Parthasarathy and Murthy, 2000). As a result of the M→V (α 188) amino acid change, the bond distance between the α E192 O ϵ 1 and the α K195 N ζ decreases from 7.59Å (Figure 5.6b) to 3.43Å (Figure 5.6a) and a single salt bridge forms between the respective carboxyl and amino side chains. Therefore, although α V188 does not directly contribute to stabilising interactions, the M→V (α 188) substitution mediates the formation of an ion pair. This is one of two ion pairs that occur in mutant 8C.



The suboptimal geometry of the residues involved in this ionic interaction prevents the formation of a double salt bridge. Since only one of the side chain nitrogen and oxygen atoms is within 4Å of each other, single ion pairs are also classified as nitrogen-oxygen (N-O) bridges (Kumar and Nussinov, 2002). The energetic contribution of N-O bridges to protein thermostability is generally less than that of geometrically optimal or double salt bridges such as the α R28- β E96 ion pair that also occurs in this mutant. This ionic interaction braces the helix consisting of residues α 190-196 and is involved in stabilising this helix. The stabilisation of helices by salt bridges is important for protein thermostability (Ghosh *et al.*, 2003; Marqusee and Sauer, 1994).

(a)



(b)

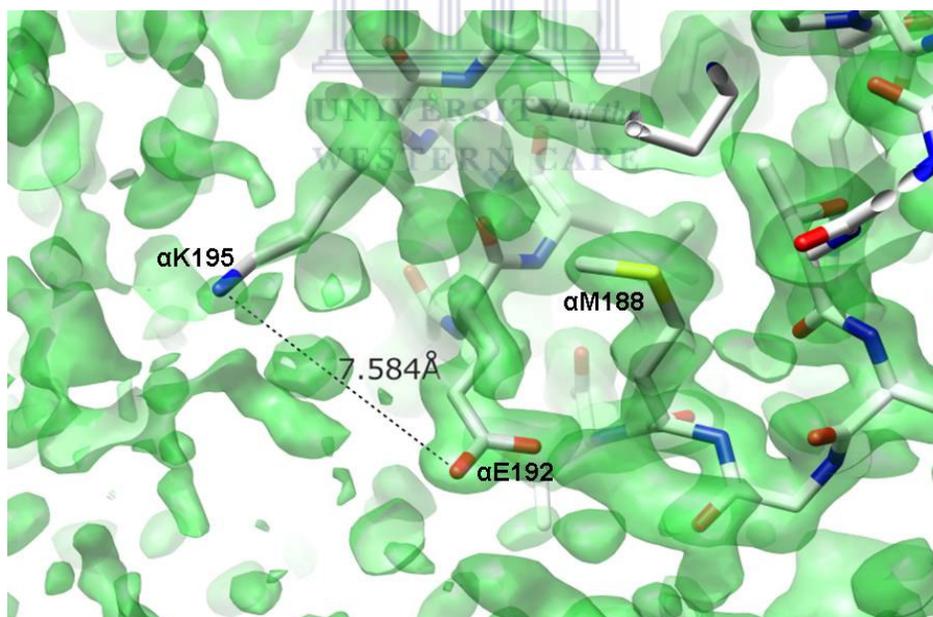


Figure 5.6: Electron density maps showing (a) the single salt bridge between α K195 and α E192 as a result of the M \rightarrow V (α 188) mutation and (b) the position of α K195 and α E192 in the wild-type NHase. The electron density map was contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant. The amino acid residues are shown as ball and stick where the atoms are: C=white, O=red, N=blue, S=yellow.

The D→E (β 96) mutation is thought to be one of the main contributors to the enhanced thermostability of mutant 8C. Specifically, a double salt bridge was formed between the negative carboxyl side chain of β E96 and the positive side chain of α R28 (Figure 5.7). The bond distance was 2.96Å and 3.02Å from the O ϵ 1 and O ϵ 2 of β E96 to the respective N η 2 and N η 1 of α R28. Most of the salt bridges (~92%) with this ion pair geometry are stabilising (Kumar and Nussinov, 2002).

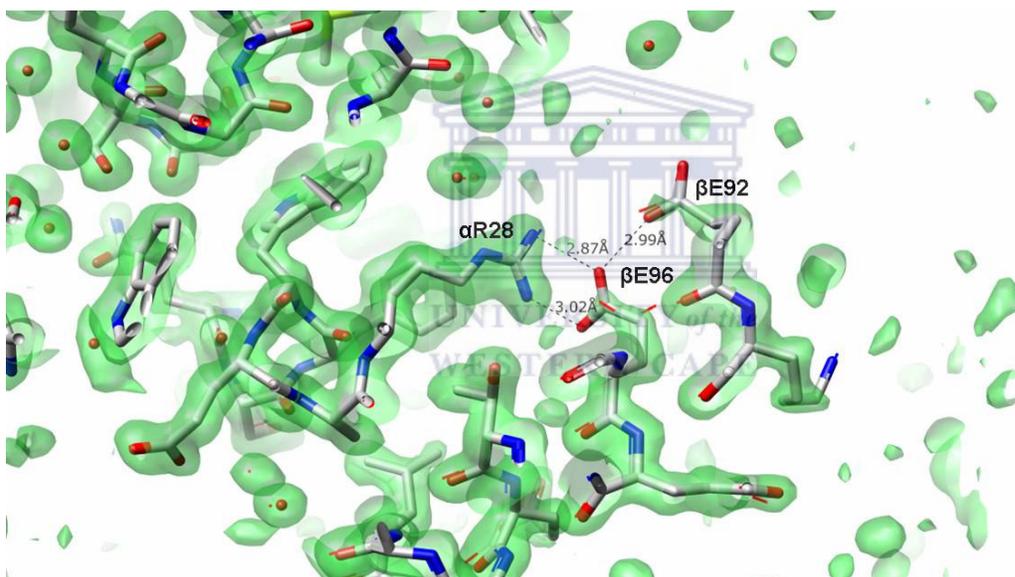
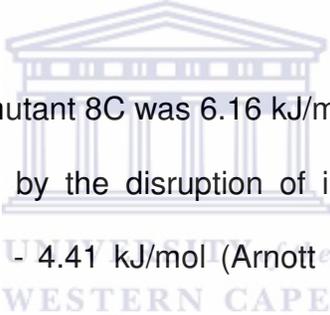


Figure 5.7: Electron density map showing the double salt bridge between α R28 and β E96 and the hydrogen bond between β E96 and β E92 in mutant 8C. The electron density map was contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant. The amino acid residues are shown as sticks in the mutant and lines in the wild-type where the atoms are: C=white, O=red, N=blue, H₂O=red spheres.

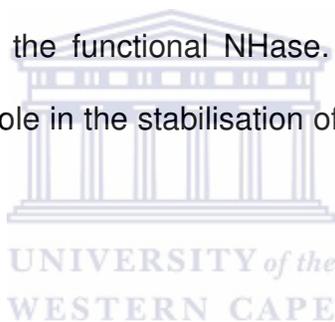
The O ϵ 2 of β E96 also forms a hydrogen bond with the O ϵ 1 of β E92 (2.99Å). The structure of the bonds between the three amino acid residues is typical of a triad where the interactions are often interdependent or coupled (Marqusee and Sauer, 1994). The stabilisation that is conferred by this electrostatic network is likely to be more than the stabilisation that could be conferred from the sum of the individual interactions. β E96 and α R28 occur on separate α helices of the respective β - and α subunits that are located almost perpendicular to each other (Figure 5.5). The slightly longer side chain of Glu compared to Asp favours the formation of the stabilising salt bridge.



The $\Delta\Delta G$ value obtained for mutant 8C was 6.16 kJ/mol. The stabilisation attributed to salt bridges as measured by the disruption of individual ion pairs has been reported to be between 3.31 - 4.41 kJ/mol (Arnott *et al.*, 2000) and 3.73 - 4.18 kJ/mol (Pappenberger *et al.*, 1997). The extent of the stabilisation of mutant 8C is less than expected considering the number and type of electrostatic interactions (both hydrogen bonding and salt bridges). Besides stabilising the individual helices, the salt bridge also facilitates interaction between the α - and β subunits.

One of the trends shown from structural comparisons of thermophilic and mesophilic homologs is that thermophilic proteins tend to be oligomeric (Bjork *et al.*, 2004). Furthermore, the additional interfaces are often characterised by intersubunit ionic networks that have been shown to contribute to the stability of the thermophilic proteins (Arnott *et al.*, 2000; Bjork *et al.*, 2004; Clantin *et al.*, 2001).

The D→V (β 167) mutation had minimal effect on the 3D crystal structure of mutant 8C and no significant perturbations of the local structural environment were observed (Figure 5.8). This residue was not involved in any additional interactions in the local environment that could explain the improved thermostability of this mutant. However, the D→V (β 167) mutation may be structurally important because in the wild-type NHase, a water-mediated hydrogen bond occurs between β D167 and β 'K168, which is positioned at the tetramer-forming interface of the enzyme (Figure 5.9). This favourable Asp-Lys inter-subunit interaction between the β subunits (β and β ') is disrupted in mutant 8C and could be involved in destabilising the heterotetrameric form of the functional NHase. Charged residues at protein interfaces play an important role in the stabilisation of oligomeric proteins (Jones *et al.*, 2000).



The importance in oligomerisation as a contributor to the thermostability of proteins, is widely recognised (Arnott *et al.*, 2000; Bjork *et al.*, 2004). Furthermore, interfacial protein-protein interactions play key roles in the stability of oligomeric proteins. Previously, Vetriani *et al.* (1998) showed that both inter trimer and intersubunit interactions are important for the thermostability of the glutamate dehydrogenases from the hyperthermophiles, *Pyrococcus furiosus* and *Thermococcus litoralis*. Similarly, Luke and Wittung-Stafshede (2006) reported that up to 80% of the stability of the oligomeric human chaperonin protein 10 is attributed to the interfacial interactions. Water is also important for the stability of

oligomeric proteins by mediating hydrogen bond formation at protein interfaces (Jiang *et al.*, 2005; Li and Lazaridus, 1996).

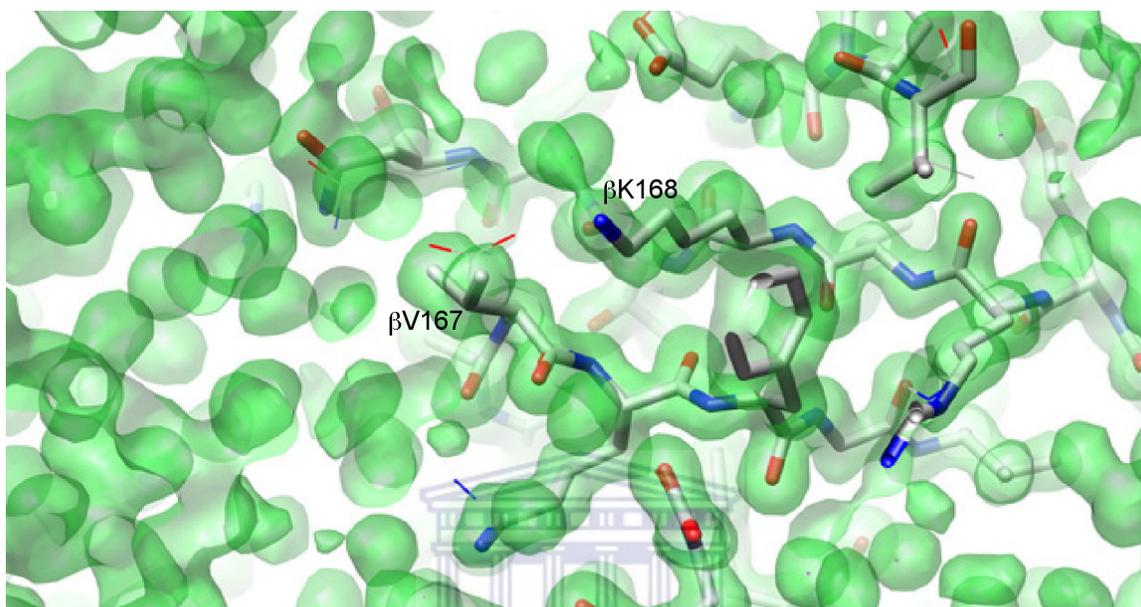


Figure 5.8: Electron density map showing the relative positions of β V167 and β K168 in mutant 8C. The electron density map was contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant. The amino acid residues are shown as sticks in the mutant and lines in the wild-type where the atoms are: C=white, O=red, N=blue.

The D \rightarrow V (β 167) mutation should have had a negative impact on the thermostability of the enzyme because it results in the disruption of a stabilising hydrogen bond at the heterotetramer interface. The disruption of a hydrogen bond at the protein-protein interface of an oligomeric enzyme can destabilise the enzyme by ~ 5.2 kJ/mol (Hiraga and Yutani, 1997). However, mutant 8C exhibited enhanced thermostability compared to the wild-type enzyme. This does not mean that the D \rightarrow V (β 167) mutation does not destabilise the enzyme. When comparing the number, type and position of favourable interactions in mutant 8C and 9C, it

would appear that 8C should be the more thermostable enzyme. However, the kinetic stability data have shown that 9C is 15 fold more thermostable than the wild-type NHase, whereas for 8C the improvement is 9 fold (Table 4.1). Also, the $\Delta\Delta G$ value of 6.16 kJ/mol was less than expected considering the contributions of both a hydrogen bond and two salt bridges to the thermostabilisation of this mutant.

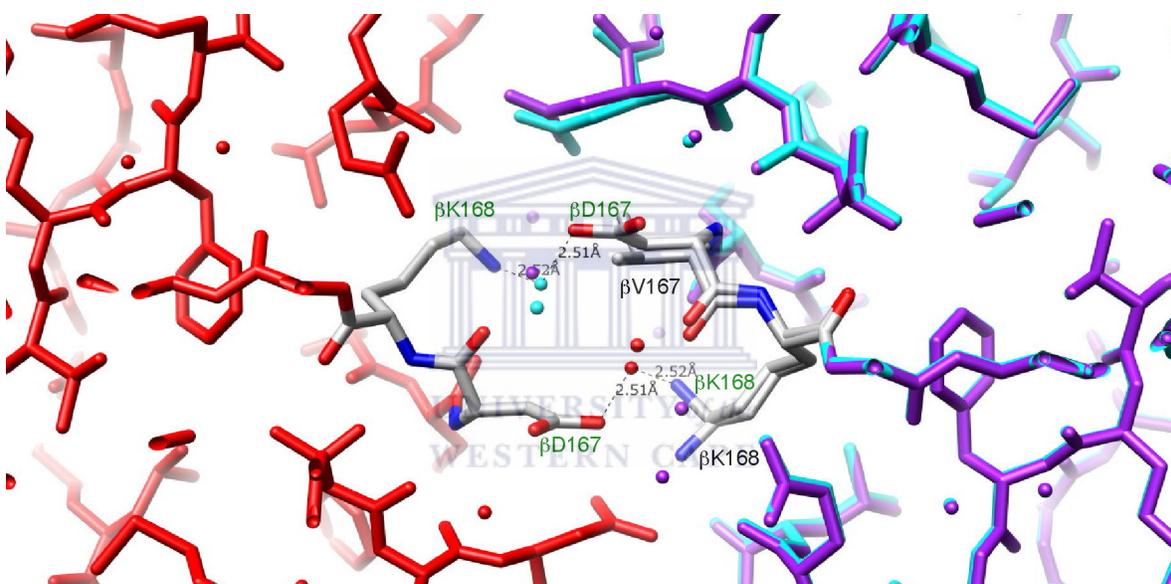


Figure 5.9: The NHase heterotetramer interface across the two-fold axis. The left hand side of the interface depicts the wild-type and is coloured in red. The right hand side shows the superimposition of the mutant and wild-type interface where the wild-type is coloured purple and the mutant cyan. The water molecules are coloured red, purple or cyan according to the position relative to the interface. The atoms of the amino acid residues of interest are coloured C=white, O=red, N=blue and are labelled in green for the wild-type and black for the mutant.

The discrepancy between the thermostability of 8C and 9C can probably be attributed to the D→V ($\beta 167$) mutation in 8C. If this hypothesis is correct, then the

D→V (β 167) substitution partially masks the stabilising effects of the electrostatic interactions that occur in this mutant. It would be interesting to evaluate the thermostability of 8C in the absence of this destabilising mutation.

5.4.3. Mutant 9C (α S169R / β M43K / β T150A)

An overview of the three mutations observed in mutant 9C is shown in Figure 5.11.

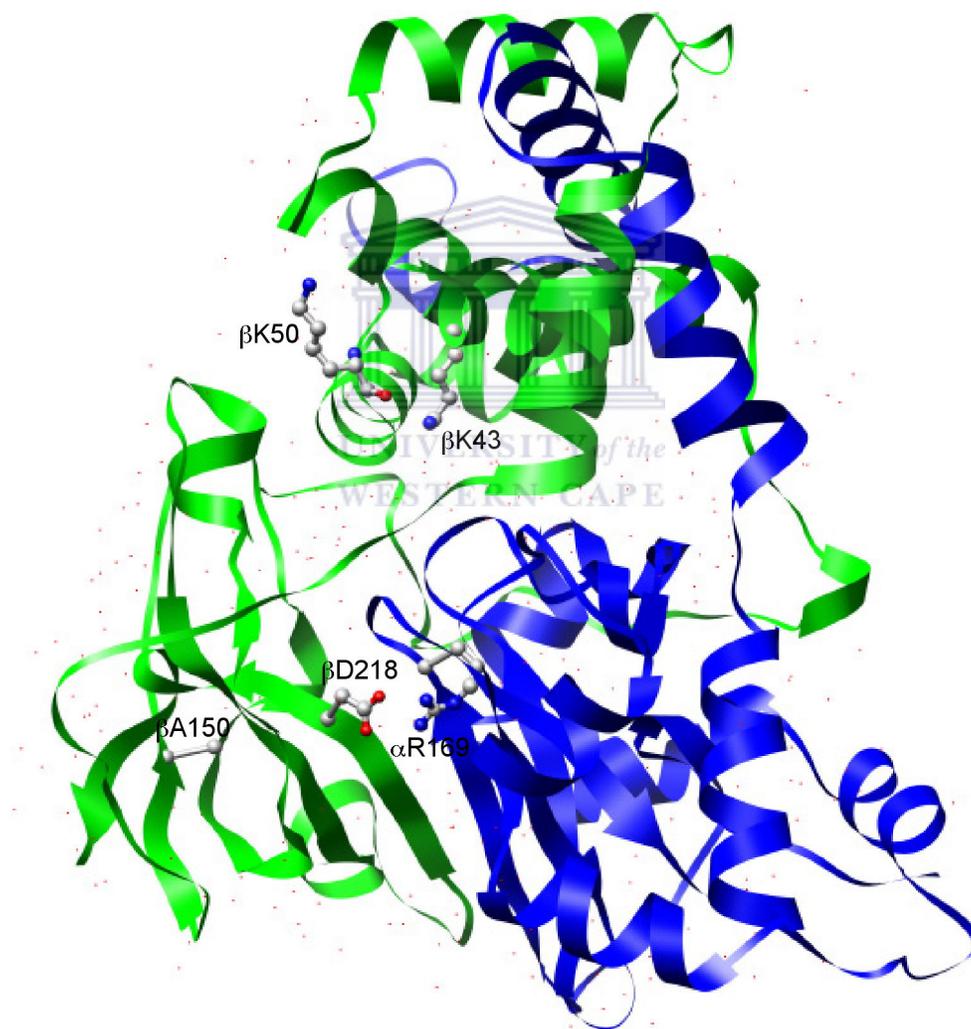


Figure 5.10: Three-dimensional structure of mutant 9C showing the amino acid changes on the α β 2 heterotetramer. The α subunit is coloured blue and the β subunit green. The relevant mutations are shown in ball and stick where the atoms are: C=white, O=red, N=blue

The S→R (α 169) mutation results in the formation of a double salt bridge between α R169 and β D218 (Figure 5.11). The bond distance is 2.75Å between the β D218 O δ 1 and the α R169 N η 2 and 3.45Å between the β D218 O δ 2 and the α R169 N η 2. The formation of this salt bridge also results in the movement of several other amino acid residues including α F175, α M176, β K210 and β D167. However, it was difficult to ascertain whether the movement of these amino acid residues contributed to the improved thermostability of mutant 9C. α R169 is located on a loop of the α subunit while β D218 is located on a β -sheet of the β subunit that is positioned parallel to the loop.

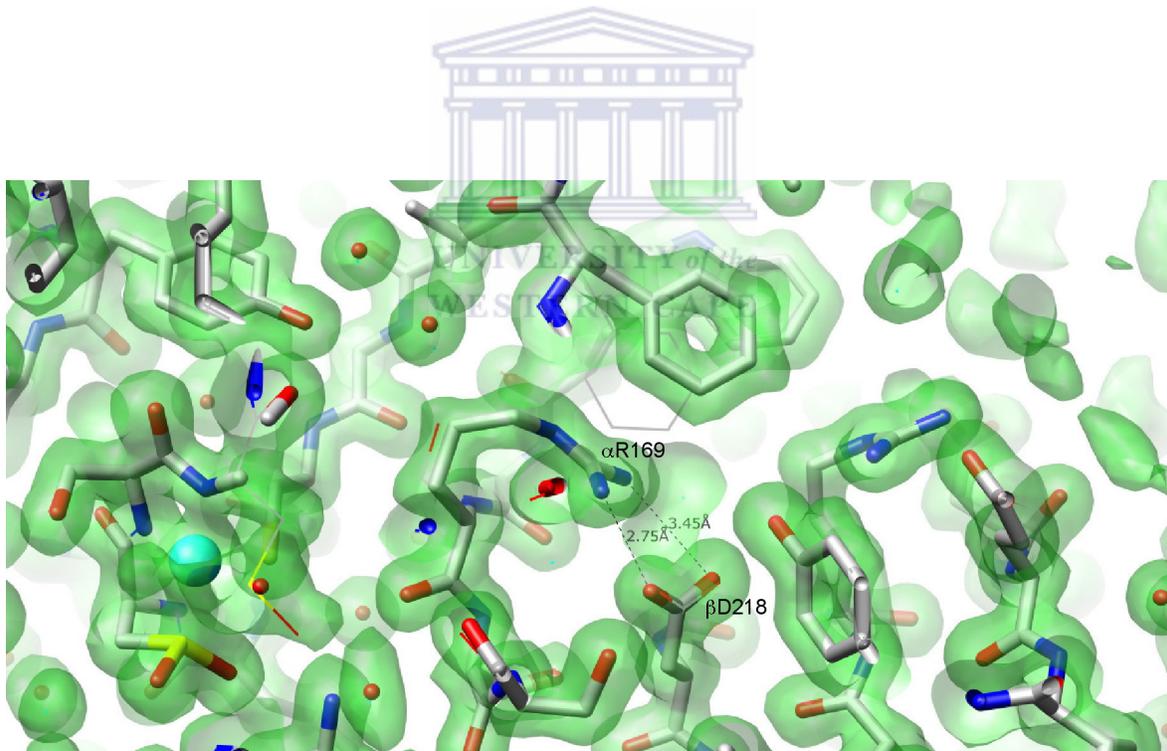
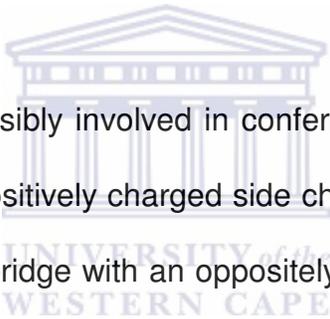


Figure 5.11: Electron density map showing the double salt bridge between α R169 and β D218 in mutant 9C. The electron density map was contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant. The amino acid residues are shown as sticks in the mutant and lines in the wild-type where the atoms are: C=white, O=red, N=blue, S=yellow, H₂O=red spheres, Co=cyan sphere.

In order to test the effect of the S→R (α 169) mutation on the thermostability of mutant 9C, a site-directed mutant, α S169R was constructed (see section 4.4). Thermal inactivation kinetic data showed that although α S169R was more thermostable than the wild-type, it was less thermostable than mutant 9C. Thus, the S→R (α 169) mutation is not solely responsible for the improved thermostability of this mutant. This observation is confirmed by the $\Delta\Delta G$ value of 7.62 kJ/mol for mutant 9C relative to the wild-type NHase (Table 4.1). The degree of the measured stabilisation is higher than is expected for the contribution of a salt bridge to the thermostability of this mutant.



The other residue that is possibly involved in conferring enhanced thermostability to mutant 9C is β K43. The positively charged side chain of lysine meant that β K43 could potentially form a salt bridge with an oppositely charged amino acid such as Glu or Asp. However, there is no negatively charged residue in the immediate environment of β K43 that could facilitate salt bridge formation. Instead, the amino side chain of β K43 interacts with the main-chain carbonyl oxygen of β K50 (Figure 5.12). This hydrogen bond can account for the $\Delta\Delta G$ value of 3.0 kJ/mol (Table 4.2) that is expected for the extra stabilising interaction. This degree of stabilisation is marginally less than the expected value of 4.18 kJ/mol for typical hydrogen bond. The reason for this difference is not clear. Charged residues are important for stabilising proteins; this is the result of both the formation of salt bridges and the formation of hydrogen bonds (Marqusee and Sauer, 1994; Pappenberger *et al.*, 1997).

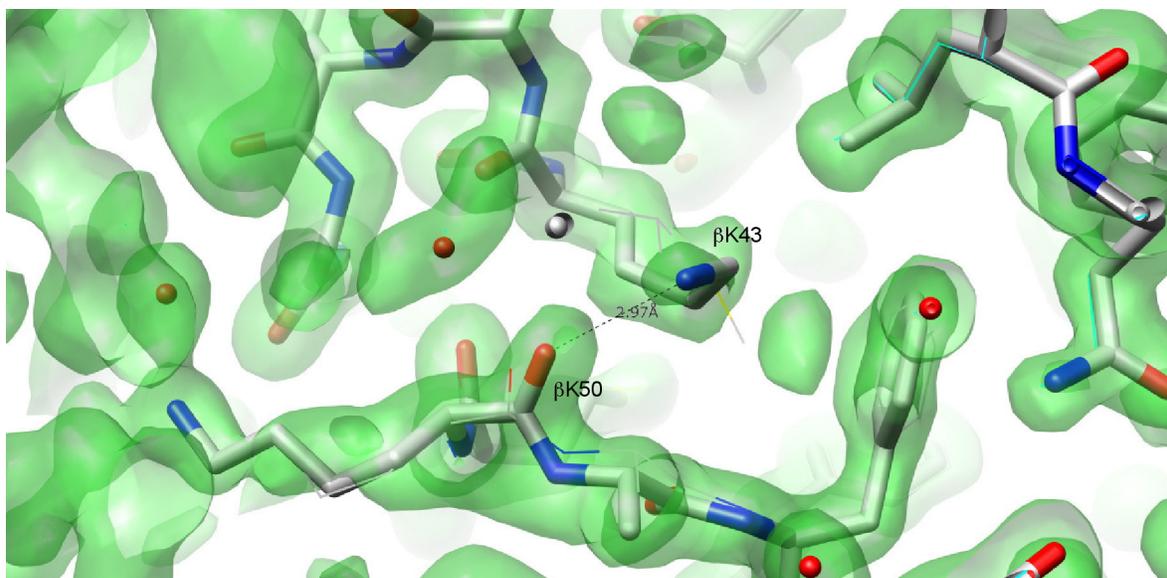


Figure 5.12: Electron density map showing the hydrogen bond between β K43 and β K50 in mutant 9C. The electron density map was contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant. The amino acid residues are shown as sticks where the atoms are: C=white, O=red, N=blue, H₂O=red spheres.

β A150 seems to effect the 3D conformation of mutant 9C. As a result of the T→A (β 150) mutation, β F161 and the whole region from α 142-148 move closer to β A150. β F161 is located on a loop while the region of α 142-148 comprises a loop between the ends of two α helices. This region is rich in ionic amino acids and contained one Glu and two Arg residues. It was difficult to ascertain whether the changes in the 3D conformation of mutant 9C as a result of this mutation had any effect on the thermostability of this mutant.

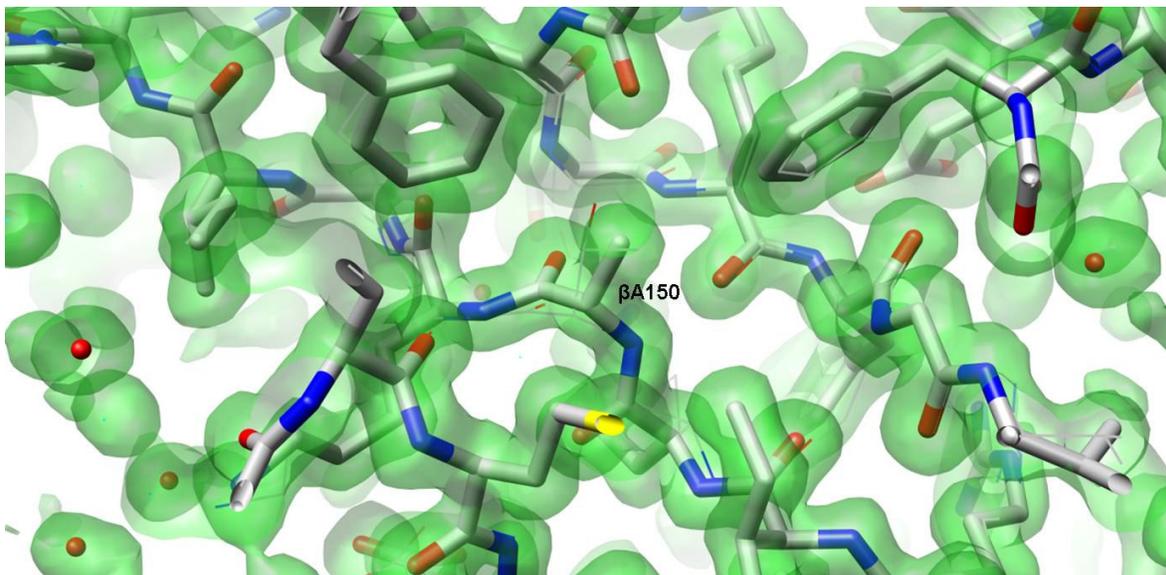
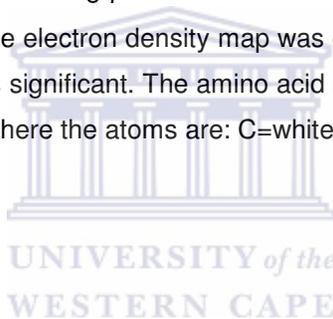


Figure 5.13: Electron density map showing β A150 and the surrounding amino acid residues in the local environment of mutant 9C. The electron density map was contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant. The amino acid residues are shown as sticks for the mutant and lines for the wild-type where the atoms are: C=white, O=red, N=blue, H₂O=red spheres.



5.4.4. Mutant 9E (α D4G / β F36L / β L103S / β Y127N)

An overview of the three mutations that were characterised in mutant 9E is shown in Figure 5.14. It was not possible to exactly characterise the D \rightarrow G (α 4) mutation, as α G4 is not well defined in the electron density map of this mutant probably due to the flexibility of the C-terminal region. It is impossible to speculate about the structural effect of this mutation on the thermostability of mutant 9E without any structural information for this region.

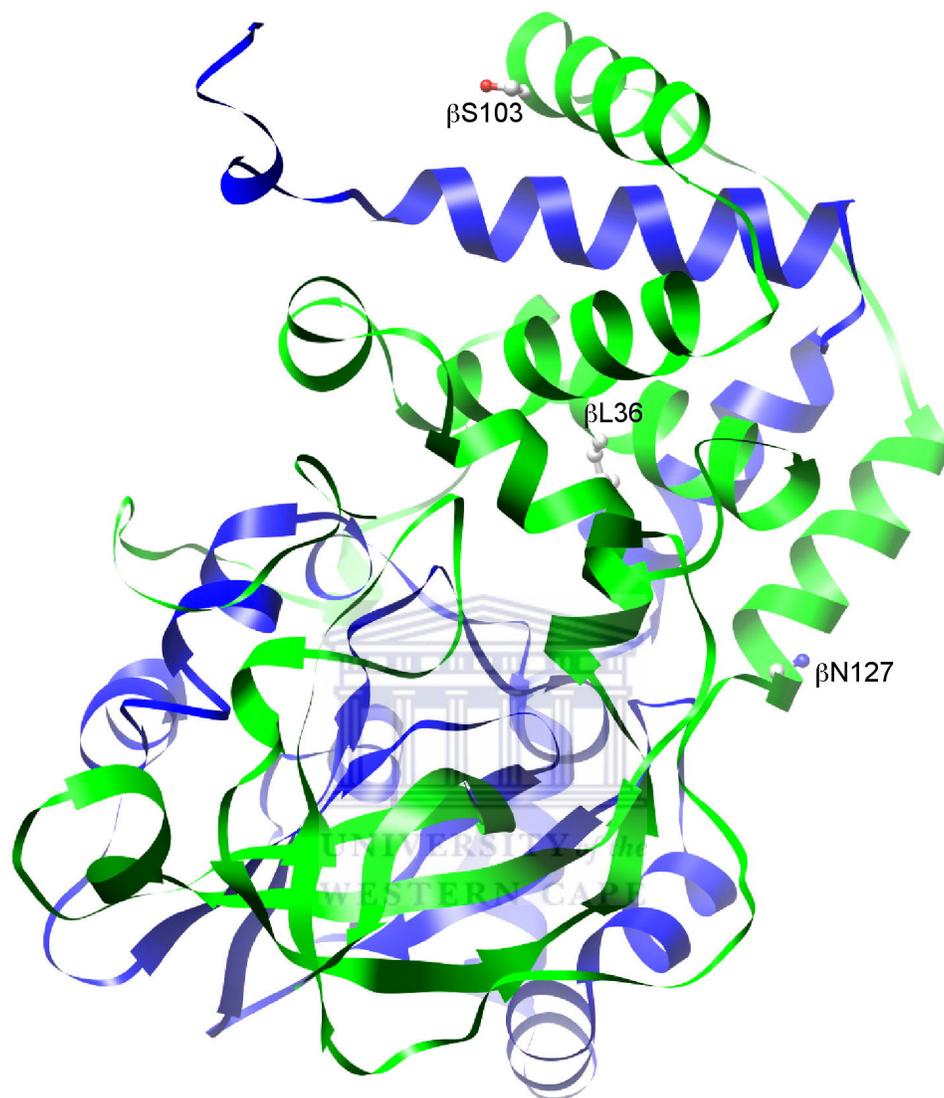


Figure 5.14: Three-dimensional structure of mutant 9E showing the amino acid changes on the $\alpha\beta$ 2 heterotetramer. The α subunit is coloured blue and the β subunit green. The relevant mutations are shown in ball and stick where the atoms are: C=white, O=red, N=blue

The F \rightarrow L (β 36) mutation does not seem to be involved in the improved thermostability of mutant 9E. There appear to be no significant 3D conformational

changes or movement of other amino acid residues in the local environment of β L36. Figure 5.15 illustrates that β L36 is in an almost identical position to the substituted β F36. The only observable difference between the two amino acids is the absence of the aromatic side chain in β L36. The environment surrounding the F \rightarrow L (β 36) mutation is quite rich in aromatic amino acid residues and in fact, forms part of the substrate channel leading to the NHase active site (Cameron, 2003; Tsekoa, 2005).

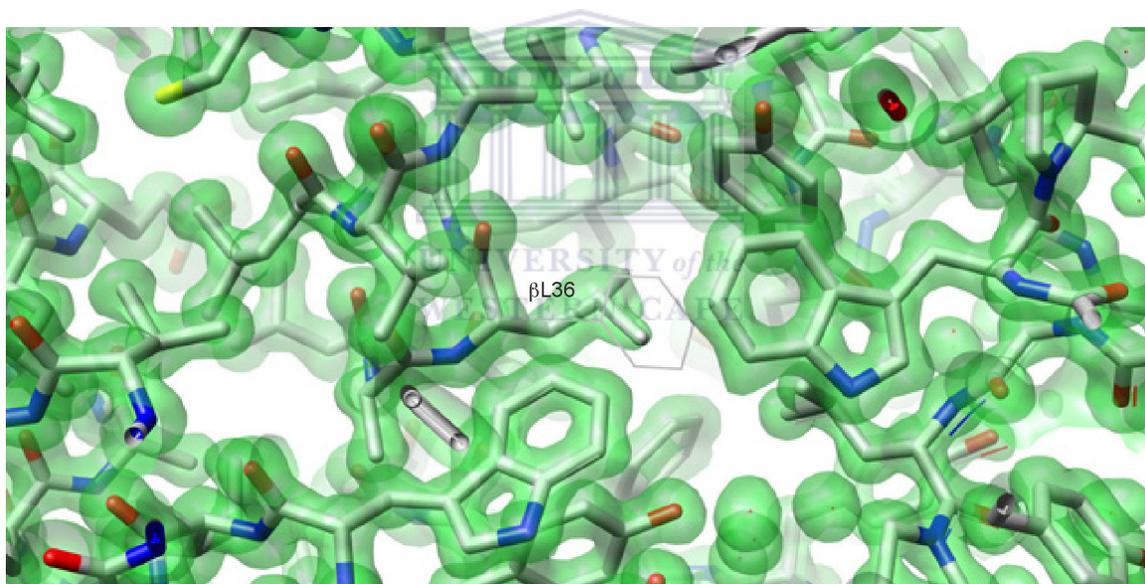
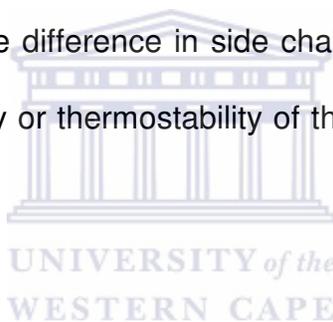


Figure 5.15: Electron density map showing β L36 and the surrounding amino acids in the local environment of mutant 9E. The electron density map was contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant. The amino acid residues are shown as sticks in the mutant and lines in the wild-type where the atoms are: C=white, O=red, N=blue.

Previously, Cameron (2003) created the site-directed mutant, β F36L, in an attempt to engineer homoaromatic substrate specificity into the *G. pallidus* RAPc8 NHase. The inability to catalyse homoaromatic substrates was thought to be related to the hydrophobicity of the substrate channel. Thus, β F36 was one of the amino acids chosen for SDM based on the hypothesis that it is located within the substrate channel in close proximity of the catalytic site cleft. Also, homology modelling with *Ps. thermophila* NHase showed that β F36 is highly flexible and could form potential multiple interactions with the homoaromatic substrate thereby preventing substrate recognition and catalysis. However, besides still being unable to catalyse homoaromatic substrates, the difference in side chains between Phe and Leu did not affect the catalytic activity or thermostability of the site-directed mutant, β F36L (Cameron, 2003).



Improved hydrophobic interactions at the protein interior is considered an important mechanism for protein thermostabilisation (Anderson *et al.*, 1993; Yutani *et al.*, 1987). This stabilisation effect is often associated with improved packing of the core and increased aromatic-aromatic interactions (Anderson *et al.*, 1993). The presence of additional aromatic clusters in thermophilic enzymes has been shown to be a key determinant of protein thermostability (Kannan and Vishveshwara, 2000). The stabilising effect of aromatic amino acids by increasing aromatic-aromatic interactions can stabilise an enzyme by up to 5.9 kJ/mol (Anderson *et al.*, 1993; Puchaev *et al.*, 2003, Serrano *et al.*, 1991).

In the context of the present study, the substitution of Phe with Leu results in decreased aromatic-aromatic interactions in the protein interior, which should destabilise mutant 9E. However, evidence from the site-directed mutant, β F36L indicated that the thermostability of this mutant was comparable to the wild-type (Cameron, 2003). Therefore, the decrease in regional hydrophobicity, as a result of the F \rightarrow L (β 36) mutation does not appear to have either a positive or a negative effect on the thermostability of mutant 9E.

The L \rightarrow S (β 103) mutation is thought to be primarily responsible for the improved thermostability of mutant 9E. The stabilising interaction results from the formation of a water-mediated hydrogen bond between β S103 and α S23 via the respective O γ atoms of their hydroxyl side chains (Figure 5.16). The superimposed electron density maps of the wild-type (brown) and 9E (green) shows that this structured water is absent in the wild-type. The bond distance was 2.82Å between the β S103 O γ and W139 and 3.07Å between W139 and the α S23 O γ . The L \rightarrow S (β 103) mutation occurs on the surface of mutant 9E (Figure 5.16) and once again illustrates the importance of increasing hydrogen bonding between polar surface residues and structured water molecules to mediate improved thermostability. The $\Delta\Delta G$ of mutant 9E indicates a stabilisation of 4.27 kJ/mol compared to the wild-type NHase (Table 4.1). This is within the range of 4.18 to 9.20 kJ/mol attributed to the stabilisation provided by a hydrogen bond (Myers and Pace, 1996; Takano *et al.*, 1999).

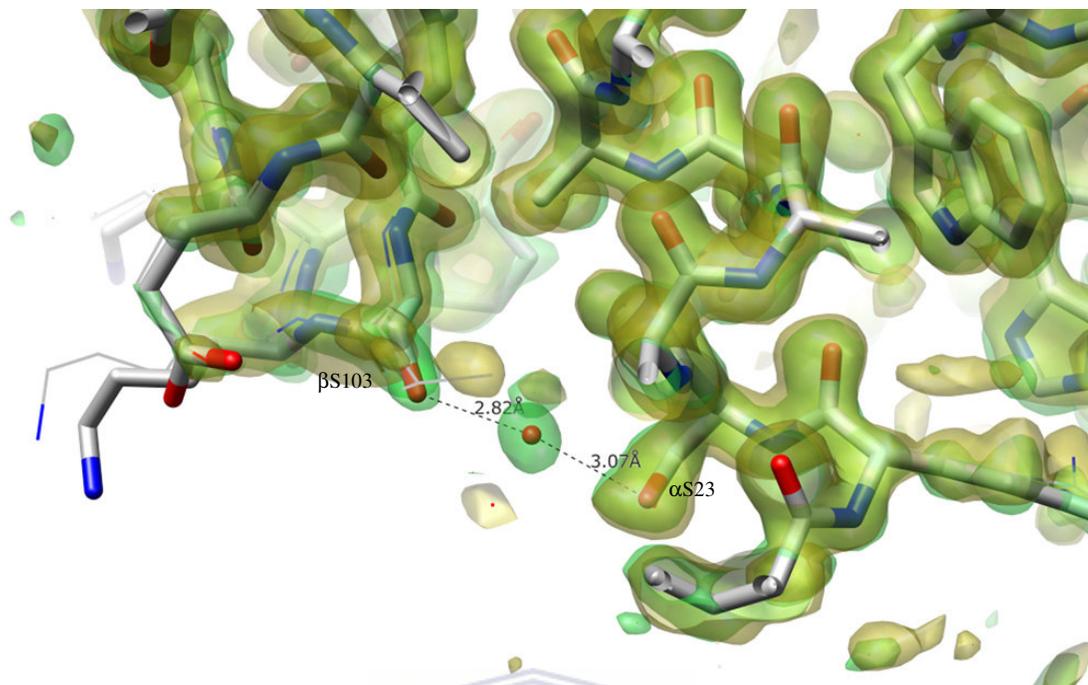


Figure 5.16: Superimposition of the wild-type (brown) and mutant 9E (green) electron density maps showing the water-mediated hydrogen bond between α S103 and α S23. The water molecule is only present in the density attributed to the mutant. The electron density maps were contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant. The amino acid residues are shown as sticks in the mutant and lines in the wild-type where the atoms are: C=white, O=red, N=blue, H₂O=red sphere.

However, the contribution of the L→S (β 103) mutation to the stabilisation of mutant 9E is slightly higher than for mutant 7D (3.40 kJ/mol). This is contradictory to experimental evidence that suggests that charged-neutral hydrogen bonds, such as mutant 7D, are more favourable than neutral-neutral hydrogen bonds, such as mutant 9E in water (Luo and Baldwin, 1997; Shan *et al.*, 1996). The difference between these mutants is that while β S103 and α S23 (mutant 9E) result in an inter-subunit hydrogen bond, the hydrogen bond between α S47 and α E33 (mutant

7D) occurs within the α subunit. This difference probably accounts for the difference in $\Delta\Delta G$ between these mutants as an inter-molecular hydrogen bond is likely to be more stabilising than an intra-molecular hydrogen bond.

The Y \rightarrow N (β 127) mutation involved the substitution of a polar amino acid with the same type of amino acid residue on the surface of mutant 9E. In the wild-type NHase, β Y127 forms hydrogen bonds with α R49 and α E57 via their respective N ζ (3.32Å) and O ϵ 1 (3.4Å) groups (Figure 5.18a). The Y \rightarrow N (β 127) substitution disrupts this intersubunit interaction and the cavity that is created by the replacement of the bulky Tyr with Asn is filled with several water molecules (Figure 5.17b). Tyr is more likely to be a hydrogen donor than a hydrogen acceptor because one of the lone pairs on the O η is thought to be partially delocalised within the aromatic ring. As a result, the Tyr O η generally forms only one hydrogen bond (Baker and Hubbard, 1984). The hydrogen bonds that are formed from this amino acid residue acting as a hydrogen acceptor tend to be weak (McDonald and Thornton, 1994). Therefore, it is possible that only the disruption of the hydrogen bond between β Y127 and α E57 was significant.

The disruption of intermolecular hydrogen bonds with Tyr has been shown to result in up to an 8.4 kJ/mol decrease in the stability of RNase Sa (Pace *et al.*, 2001). However, it is difficult to assess whether the disruption of the hydrogen bonds resulting from the Y \rightarrow N (β 127) mutation could destabilise mutant 9E. Most of the Tyr groups studied by Pace *et al.*, (2001) involved buried or partially buried

residues that were replaced by Phe. The stabilisation provided by buried polar residues involved in hydrogen bonding is more than that provided by surface residues and can be as much as 5.4 kJ/mol (Shirley *et al.*, 1992). β Y127 group is located on the enzyme surface and is replaced by Asn, which is capable of forming four hydrogen bonds. The Asn is surrounded by several water molecules that occupy the space vacated by the Tyr. Although the presence of these water molecules in the mutant was confirmed, they are not well-defined as shown by the lack of electron density in Figure 5.17b. Thus, it is difficult to determine whether there is any interaction between β N127 and the surrounding water molecules.

However, most polar groups in proteins are hydrogen bonded and surrounded by waters (Myers and Pace, 1996). Also, the occurrence of an unsatisfied polar group is rare (Fleming and Rose, 1995). Therefore, it is unlikely that β N127 does not hydrogen bond with the neighbouring water molecules. Also, hydrogen bond networks between water molecules are known to contribute to protein stability (Takano *et al.*, 1999). Thus, even if the Y \rightarrow N (β 127) mutation was destabilising this negative effect was countered by positive stabilisation forces that resulted from hydrogen bonding between β N127 and the water molecule(s), or hydrogen bonding between the water molecules that filled the cavity. Certainly, the $\Delta\Delta G$ of 4.27 kJ/mol corresponds to the hydrogen bond that results from the L \rightarrow S (β 103) mutation being responsible for stabilising mutant 9E.

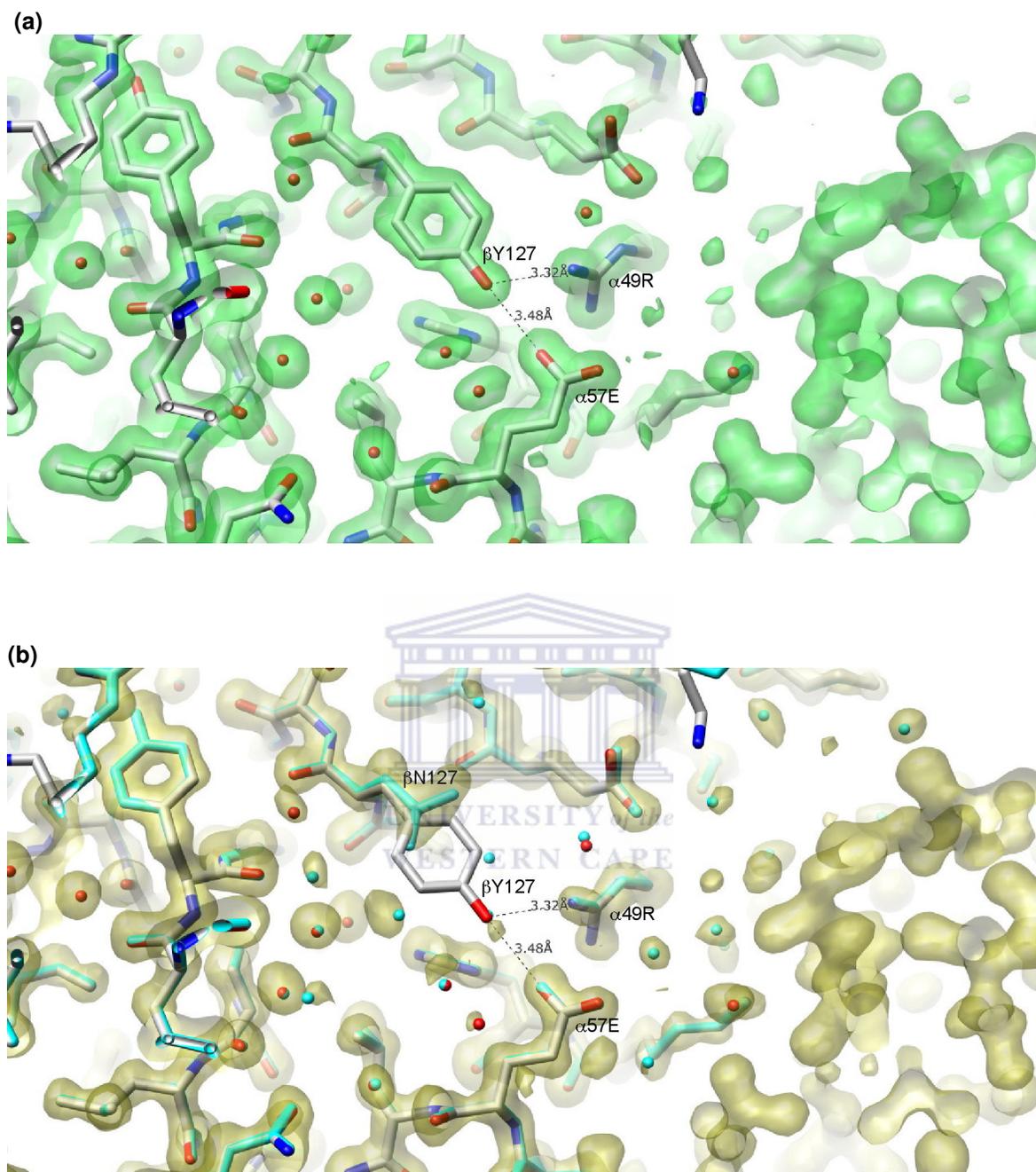


Figure 5.17: Electron density map showing: (a) Hydrogen bond formation of β Y127 with α R49 and α E57 in the wild-type NHase and (b) the replacement of β Y127 with β N127 in mutant 9E and filling of the resultant space with water. Both electron density maps were contoured at 1σ (1σ is one standard deviation) to ensure that the data is significant where: green=WT and brown=mutant. All the amino acid residues are shown as ball and stick. In the wild-type the atoms are: C=white, O=red, N=blue, H₂O=red spheres while in the mutant the both the amino acid residues and the water molecules are coloured cyan.

CHAPTER 6

GENERAL DISCUSSION AND CONCLUSION

CONTENT

6.1.	GENERAL DISCUSSION	127
6.1.1.	Library construction and evaluation.....	127
6.1.2.	The relationship between activity and thermostability	130
6.1.3.	Screening for improved thermostable mutants.....	133
6.1.4.	The structural determinants of improved NHase thermostability	134
6.1.4.1.	<i>Salt bridges</i>	134
6.1.4.2.	<i>Hydrogen bonding</i>	138
6.2.	CONCLUDING REMARKS.....	140



6.1. GENERAL DISCUSSION

The main focus of this study was to investigate the relationship between molecular structure and thermostability in a set of improved thermostable NHase mutants that were isolated from a randomly mutated library. The focus of this chapter is to discuss the experimental results already described in Chapters 3, 4 and 5 in the context of recent literature regarding the various aspects of this study.

6.1.1. Library construction and evaluation

During the present study, error-prone PCR (EP-PCR) was used to create a genetically diverse library since it is the simplest method to introduce random mutations into a gene sequence. The objective was to construct a low error frequency library that contained mostly beneficial mutations while the number of deleterious or neutral mutations was minimised. However, the *Taq* polymerase used in the present study was completely inhibited by a concentration of 0.5 mM MnCl_2 , which was previously reported to generate low mutation frequency libraries (Cadwell and Joyce, 1992). Similarly, Vanhercke *et al.* (2005) found that a MnCl_2 concentration higher than 0.25 mM resulted in significant reduction in the PCR product yield. A survey of the literature showed a number of studies where lower MnCl_2 concentrations were used: e.g. 0.2 mM (Suen *et al.*, 2004; Vanhercke *et al.*, 2005), 0.15 mM (Miyazaki and Arnold, 1999; Wintrode *et al.*, 2001), 0.1 mM (Nui *et al.*, 2006; Song and Rhee, 2000). However, with the exception of Vanhercke *et al.* (2005), the reason for lowering the MnCl_2 concentration was not stated.

Since the MnCl_2 is used to manipulate the efficiency of the *Taq* polymerase it is not unexpected that it would also affect the PCR yield. Indeed, the PCR yield and the error frequency are known to be dependent on the source of the polymerase and it is possible to create very different libraries under identical conditions by simply using *Taq* polymerases from different sources (Cirino *et al.*, 2003). Therefore, the relationship between PCR yield and error frequency is very important in deciding the MnCl_2 concentration for the mutagenic PCR reaction. However, experimental data is needed to establish such a relationship for each specific *Taq* polymerase. Since such data was not available for the *Taq* polymerase used in the present study, it was eventually decided to construct two mutant libraries using 0.05 and 0.10 mM MnCl_2 together with a biased dNTP concentration.

The frequency of active mutant enzymes is one determinant of the mutation frequency of an EP-PCR library. It is not uncommon to obtain a fraction of inactive enzymes during a random mutagenesis experiment. In fact, it is generally assumed that most randomly introduced mutations are deleterious and beneficial mutations are rare (Arnold, 1996). This is one of the reasons that the MnCl_2 concentration is carefully controlled during EP-PCR. In so doing, the *Taq* polymerase is able to introduce the required sequence diversity with minimal changes at the genetic level: i.e. a low error frequency and the resultant library contain a larger percentage of active enzymes. Shafikhani *et al.* (1997) reported that 58% and 29% of enzymes were active following one round of EP-PCR using 0.15 mM and 0.5 mM MnCl_2 , respectively. The observed frequency of active enzymes of 56% and 33% for Lib

0.05 and Lib 0.1, respectively, was lower than expected considering that the MnCl_2 concentrations were lower than that reported by Shafikhani and co-workers. Here, the low number of active enzymes compared to other studies was the result of a more profound effect of the MnCl_2 on the error-rate of the *Taq* polymerase due to the sensitivity of this enzyme to this compound.

Previous studies have shown that the frequency of active enzymes declines as the average number of mutations per gene increases (Daugherty *et al.*, 2000; Shafikhani *et al.*, 1997). The mutation frequency of Lib 0.05 was determined to assess whether the low frequency of active enzymes in this library could be related to the average number of mutations per gene. DNA sequence analysis of Lib 0.05 showed an average number of 7.1 nucleotide changes per *NHase* operon (Table 3.2). Thus, reducing the MnCl_2 concentration up to 10 fold (from 0.5 to 0.05 mM) was not enough to generate a low mutation frequency library (1-3 nucleotide changes per gene) but instead resulted in a moderate frequency library.

Although the desired result of constructing a low error frequency library was not achieved, the moderate frequency libraries were still useful. Recently, researchers have realised that moderate (up to 8 nucleotide changes) to high mutation frequency libraries (up to 30 nucleotide changes) can also be exploited to isolate variants with novel or improved functions. Increasing experimental evidence suggests that contrary to expectations higher error frequency libraries contain a surprising number of active mutant enzymes (Drummond *et al.*, 2005; Georgiou,

2001). Furthermore, such libraries have been demonstrated to lead to mutants with greater functional enhancements than low mutation frequency libraries (Daugherty *et al.*, 2000; Zacco and Gherardi, 1999). These moderate and high mutation frequency libraries are more diverse than low mutation frequency libraries and thus able to probe distant regions of protein sequence space that the latter cannot access. For the purpose of the present study, it was hoped that the moderate frequency libraries would provide NHase mutations that exhibited the required function; i.e., improved thermostability.

6.1.2. The relationship between activity and thermostability

An essential objective of thermostability engineering is to maintain the intricate balance between thermostability and activity. Therefore, one of the main concerns of this study was to ensure that the randomly introduced mutations did not affect the activity of the mutant NHases. In the past it has been suggested that thermostability and catalytic activity are mutually exclusive; i.e., that an improvement in one property is detrimental to the other (Vieille and Zeikus, 2001). This view is partly due to the observation that proteins isolated from thermophilic organisms are typically less active at room temperature than their mesophilic counterparts (Daniel, 1996; Jaenicke, 1991). Since mesophilic enzymes are believed to be derived from thermophilic enzymes it seemed likely that Nature has evolved efficient mesophilic biocatalysts at the cost of thermal stability (Volkl *et al.*, 1994; Vieille and Zeikus, 2001). There also seems to be a temperature-dependent, inverse relationship between catalytic activity and flexibility (Fields, 2001).

Thermostable enzymes are generally more rigid than their mesophilic counterparts (Sterner and Liebl, 2001). According to Zavodszky *et al.*, (1998) flexibility is a prerequisite for catalytic activity and the rigidity of thermostable enzymes seems to explain their reduced activity at low temperatures. However, this observation is probably enzyme specific. Conformational mobility is not required for catalytic activity for the *G. pallidus* NHase. Indeed, rigidity is required for the substrate to move through the narrow substrate channel to the active site (T. Sewell, Personal communication). Thermostable enzymes need to be more rigid to balance the effect of increasing conformational fluctuations that are associated with increasing temperatures (Vihinen, 1987).



The view that enhanced thermostability is detrimental to catalytic activity has been substantiated by comparative structure-function studies (e.g., Georlette *et al.*, 2003). A comparison of the 3D structure and activity of mesophilic and thermophilic DNA ligases showed that the latter exhibited both high thermostability and reduced low temperature activity. Similarly, there is evidence from site-directed mutagenesis studies which suggests that there is an inverse relationship between thermostability and catalytic activity (Zhang *et al.* 2005). However, there is also experimental evidence that illustrates that thermostability and activity need not be mutually exclusive and that it is possible to improve enzyme stability without affecting nor impeding activity (Van der Burg *et al.*, 1998; Giver *et al.*, 1998). There is also evidence from directed evolution studies that it is possible to simultaneously

engineer both these properties into an enzyme (Miyazaki *et al.*, 2000; Strausberg, 2005).

In the present study, the screening strategy was used to assess and ensure that the improved thermostable mutant NHases also retained low temperature catalytic activity (37°C). The screen was performed in two steps, whereby the activity of mutant NHases was first determined at 37°C. The NHase mutant libraries were then subjected to thermal inactivation at 60°C followed by residual NHase activity determination at 37°C in order to screen for improved thermostable enzymes; i.e., only mutant enzymes that retained activity at 37°C were subjected to thermostability screening. Linking the catalytic activity of mutant enzymes to the thermostability screen is the simplest way of ensuring that the balance between thermostability and activity was maintained. In this manner it is possible to isolate improved thermostable enzymes that also retain low temperature activity (Kuchner and Arnold, 1997). This study also illustrates that thermostability and activity need not be mutually exclusive since all the improved thermostable enzymes isolated showed levels of activity equivalent to that of the wild-type at 37°C. Clearly, the contrasting views on the relationship between thermostability and catalytic activity suggest that the mechanisms involved are not well-understood and require further investigation.

6.1.3. Screening for improved thermostable mutants

The hydroxamic acid method was identified as a suitable assay used to measure NHase activity. Since the formation of hydroxamic acid is dependent on the presence of an amide, this assay was the most direct method of colorimetrically determining NHase activity. However, the *G. pallidus* NHase was sensitive to hydroxylamine concentrations above 2 mM even though activity was observed up to ~50 mM. Various studies have reported that hydroxylamine is an inhibitor of NHase activity (Nagasawa *et al.*, 1991; Takashima *et al.*, 1998; Wieser *et al.*, 1998). Most of these studies measured the percentage residual NHase activity following pre-incubation with 1 mM hydroxylamine although some, such as Maier-Greiner *et al.* (1991), used higher hydroxylamine concentrations (20 mM).

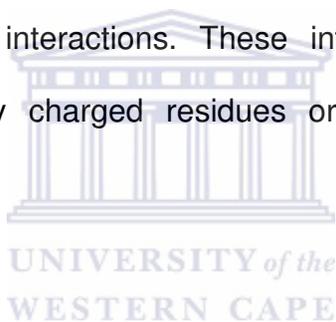
In using the hydroxamic acid assay to detect NHase activity, the inhibition of NHase by hydroxylamine was circumvented by exploiting the bi-enzymatic nature of the NHase / amidase system. This was not difficult since the amidase was able to tolerate high hydroxylamine concentrations and the activity of the amidase was optimal at a hydroxylamine concentration of 500 mM. Using this assay it was possible to isolate six NHase mutants that exhibited the required function; i.e., improved thermostability from both the EP-PCR libraries generated.

The kinetic stability of the improved mutants were assessed and compared to the wild-type by measuring the thermal inactivation rate; i.e., the rate of thermally-induced unfolding. The mutants were between 3 and 15 fold more thermostable

than the wild-type (Table 4.1) However, although a set of improved thermostable enzymes were isolated, the high number of amino acid changes per mutant enzyme (Table 3.2) made it difficult to assess which amino acids were involved in or conferred higher thermostability. Therefore, the most thermostable (according to kinetic stability data) mutants were subjected to x-ray crystallography.

6.1.4. The structural determinants of improved NHase thermostability

During this study, the most prominent stabilising features implicated in the enhanced thermostability of the randomly created mutant enzymes were the occurrence of electrostatic interactions. These interactions occurred as salt-bridges, between oppositely charged residues or through hydrogen bonding through a water molecule.



6.1.4.1. Salt bridges

The two most improved mutants, 8C and 9C, had mutations that facilitated the formation of stabilising salt bridges. Salt bridges have been extensively studied as one of the structural factors responsible for the difference in thermostability between thermophilic and mesophilic enzymes (Kumar *et al.*, 1999; Vieille and Zeikus, 2001). The generally observed trend is that thermophilic enzymes contain more ionic interactions and networks than their mesophilic counterparts (Karshikoff and Ladenstein, 2001; Kumar *et al.*, 2000; Robinson-Rechavi *et al.*, 2006). This corresponds with observations from genome sequence comparisons that thermophilic proteins tend to have a higher frequency of charged amino acids than

their mesophilic counterparts (Das and Gerstein, 2000; Haney *et al.*, 1999b; Liang *et al.*, 2005). Structure-based prediction of the positive contribution of salt bridges on the stability of thermostable enzymes has also been verified by site-directed mutagenesis studies (Arnott *et al.*, 2000; Pappenberger *et al.*, 1997; Vetriani *et al.*, 1998).

Mutant 8C has two salt bridges, both of which occur on the protein surface. There are numerous examples in the literature that illustrate that surface salt bridges contribute to protein thermostability (Takano *et al.*, 2000; Strop and Majo, 2000; Vetriani *et al.*, 1998). For example, Vetriani *et al.* (1998) demonstrated that surface inter-subunit ion pairs are important for the thermostability of the glutamate dehydrogenase from *Thermococcus litoralis*. The thermostability of the enzyme was four-fold improved over the wild-type by engineering two inter-subunit ion pairs onto the enzyme surface. The extent of the contribution of surface salt bridges tends to vary. Using mutational analysis, Pappenberger *et al.* (1997) showed that the disruption of a salt bridge on the surface of the D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophile, *Thermotoga maritima*, destabilised the enzyme by 3.73 kJ/mol for an R20N mutation and 4.18 kJ/mol for an R20A mutation. The salt bridge between β E96 and α R28 facilitate a stabilising interaction between the α - and β subunit. The geometry of this salt bridge is optimal and allows interaction between both side chain charged groups of each of the amino acids. Makhatadze *et al.* (2003) illustrated that depending on the context of the salt bridge; i.e., the interactions of the salt-bridge with the rest of the protein, the extent

of the contribution of a K11/E34 salt bridge to the thermostability of ubiquitin varied between 3.8 kJ/mol and 4.3 kJ/mol.

The location of the salt bridge on the secondary structure is also important. The salt bridge between α E192 and α K195 braces the helix consisting of residues α 190-196. This ion pair shields the hydrogen bonds along the helix backbone from water and prevents co-operative unfolding of the α helix. Since the helix backbone is shielded from hydration, the disruption of the backbone hydrogen bonds becomes energetically unfavourable (Garcia and Sanbonmatsu, 2002). In this regard, Marqusee and Sauer (1994) showed that an intrahelical salt bridge stabilised the N-terminal domain of λ repressor by 3.8 kJ/mol.

Mutant 8C is stabilised by 6.16 kJ/mol, which results in a 9 fold improvement in thermostability compared to the wild-type *G. pallidus* NHase. However, the extent of the stabilisation is lower than the expected contribution from the stabilising factors; i.e., the inter-subunit salt bridge between β E96 and α R28 in mutant 8C, which is part of a larger electrostatic network due to the hydrogen bond between β E96 and β E92 as well as the salt bridge between α E192 and α K195. This is probably because the thermostabilisation value is the net effect of both stabilising and destabilising interactions. The disruption of the water-mediated hydrogen bond between β E167 and β K168 probably accounts for the reduced stabilisation energy. Therefore, the stabilisation energy provided by the stabilising interaction in mutant

8C is probably 4.18-9.2 kJ/mol lower than the expected value, which is equivalent to the stabilisation provided by a hydrogen bond.

The salt bridge between α R169 and β D218 of mutant 9C occurs in the protein core. It has been suggested that the desolvation energy required to bury a salt bridge is greater than any potential stabilisation effects (Vieille and Zeikus, 2001). However, Kumar and Nussinov (1999) analysed 222 salt bridges from 36 monomeric proteins and found that buried salt bridges were more stabilising than surface salt bridges. Using nuclear magnetic resonance, Anderson *et al.* (1990) showed that an Asp-His salt bridge contributed 12.5-20.9 kJ/mol to the stability of T4 lysozyme. The contribution of buried salt bridges to thermostability is temperature dependent as the dielectric constant of water decreases from 80 at 25°C to 55 at 100°C (Kumar *et al.*, 2000). This means that the desolvation energy associated with burying a salt bridge becomes less at higher temperatures and could account for the larger number of buried salt bridges in thermophilic enzymes (Vieille and Zeikus, 2001). Merz *et al.* (1999) disrupted two surface salt bridges of the indoleglycerol phosphate synthase from *Thermotoga maritima* and found that the D184A mutants was destabilised by 0.5 kJ/mol while the R241A mutant was destabilised by 3.2 kJ/mol, indicating that salt bridges are still important for the stability of hyperthermophilic proteins.

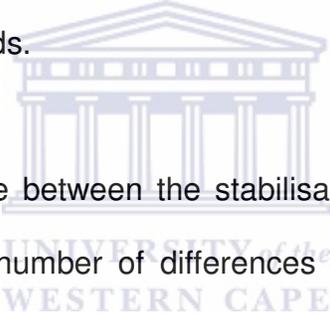
The present study suggested that a buried salt bridge partially contributed to the improved thermostability of mutant 9C. Mutant 9C is stabilised by 7.62 kJ/mol,

which resulted in a 15 fold improvement in the thermostability compared to the wild-type NHase. The positive effect of this mutation was confirmed by site-directed mutagenesis. However, the S→R (α 169) mutation was not solely responsible for the improvement in thermostability of this mutant. Mutant 9C is still 3 fold more stable than the α S169R site-directed mutant. However, the stabilisation effect of the buried salt bridge is significant.

6.1.4.2. Hydrogen bonding

Hydrogen bonding is another important type of weak electrostatic interaction that was observed to be associated with the improved thermostability of randomly created NHase mutants. Water-mediated hydrogen bonds were particularly important and were thought to be primarily responsible for the improved thermostability of mutants 7D and 9E. The evidence was especially convincing for mutant 7D, since this mutant only contained one amino acid change that could explain the improvement in thermostability. Both the hydrogen bonds in mutant 7D and mutant 9E resulted from the substitution of a hydrophobic amino acid, Iso or Leu, with a polar amino acid, Ser. The increased numbers of polar groups on the surfaces of thermophilic proteins facilitate the improved hydrogen bonding capacity across the surfaces of these enzymes (Zhou, 2002). Structured water molecules often enhance the hydrogen bonding capacity of polar groups on protein surfaces and are important for protein thermostability (Vogt *et al.*, 1997). In this regard, amino acids such as Ser and Thr are especially important since they are the best amino acids for interacting with water molecules (Mattos, 2002).

Mutant 7D is stabilised by a charged-neutral hydrogen bond between α S47 and α E33 while for mutant 9E the stabilisation results from a neutral-neutral hydrogen bond between β S103 and α S23. According to Tanner *et al.* (1996), charged-neutral hydrogen bonds are generally more stabilising than neutral-neutral hydrogen bonds. However, mutant 7D is stabilised by 3.40 kJ/mol, which is slightly below the range of between 4.18 to 9.20 kJ/mol that is attributed a single hydrogen bond (Myers and Pace, 1996). In contrast, mutant 9E is stabilised by 4.27 kJ/mol. According to Takano *et al.* (1999), water-mediated hydrogen bonds provide a stabilisation energy of 5.2 kJ/mol while the stabilisation energy is 8.5 kJ/mol for protein-protein hydrogen bonds.



The reason for the difference between the stabilisation on mutant 7D and 9E is unknown since there are a number of differences between these mutants. The difference between the stabilisation that the additional interactions provide can be due to the context of the hydrogen bonds. In mutant 7D the water-mediated hydrogen bond is thought to stabilise the α subunit while for mutant 9E, additional hydrogen bonds facilitate inter-subunit interactions. Perhaps the involvement of two Ser groups as opposed to one is also important since Ser residues are known to be the best amino acid for forming hydrogen bonds with water. Another possible reason for the difference could be that mutant 9E may have another mutation that provides marginal stabilisation. This could indicate that the Y \rightarrow N (β 127) mutation is significant either because of hydrogen bonding between the water molecules that fills the cavity that results from the substitution of the Tyr or by the improved

hydrogen bonding capacity provided by the Asn. However, these speculations could not be confirmed with the structural analysis of mutant 9E.

The importance of hydrogen bonds to protein thermostability was also illustrated by the D→V (β 167) mutation, which destabilised mutant 8C by disrupting a water-mediated hydrogen bond between β V167 and β K168 at the heterotetramer interface. Xu *et al.* (1997) analysed hydrogen bonds across 319 protein-protein interfaces and found that due to structural constraints between two proteins, hydrogen bond geometry is often not optimal and as a result such bonds either do not form or the bonds that do form are weak. The presence of water molecules at the interface facilitates bridging between the bonding atoms and thus mediates the formation of a stronger hydrogen bond across the protein-protein interface. Therefore, it is probable that the disruption of the water-mediated hydrogen bond partially masks the effect of the stabilising interactions in mutant 8C. Water-mediated hydrogen bonds at protein-protein interfaces are critical for protein thermostability (Jiang *et al.*, 2005; Spassov *et al.*, 1995)

6.2. CONCLUDING REMARKS

The generation of two moderate frequency libraries resulted in the isolation of six thermostabilised NHase mutants. Structural analysis of four of the mutants revealed that the two most improved mutants had more than one stabilising feature. This illustrates that moderate frequency libraries yields more functionally

improved mutants because these mutants contain multiple stabilising features. The diffraction data for all the mutations was better than that previously obtained for the wild-type *G. pallidus* NHase. The best diffraction data was obtained for mutant 9E at a resolution of 1.15Å. The high quality of crystal structures of the NHase mutants allowed each amino acid residue change and the neighbouring residues to be viewed clearly in the protein so that each thermostabilising feature could be assigned with confidence. This was very important as most of the mutants had several amino acid substitutions. The structural information was instrumental in identifying the water-mediated hydrogen bonds in 7D and 9E. Also, depending on its conformational state within the protein, one amino acid can be associated with several different thermostabilising features. For example, charged amino acids are associated with both hydrogen bonds and salt bridges. The structural analysis of the mutants was important to elucidate the exact role of the amino acid in the thermostabilisation of the NHase mutants.

Within its limited dataset, this study has identified electrostatic interactions as the key mechanism involved in conferring thermostability to randomly created NHase mutants. In particular, the importance of amino acid residue changes resulting in the formation of salt bridges and facilitating increased water-mediated hydrogen bonding was illustrated. Various reviews of the structural factors associated with protein thermostability have identified salt bridges and hydrogen bond networks as key components of protein thermostability (Sternier and Liebl, 2001; Li *et al.*, 2005). In this study, salt bridges and hydrogen bonds mediated their effect by increasing

intra-helical, inter-helical and inter-subunit interactions. Both surface and buried salt bridges were found to contribute to the improved thermostability of the mutants. This study also illustrated the importance of water-mediated hydrogen bonding between polar amino acids and between ion pairs to the thermostability of proteins. The structural information was also complemented with kinetic stability data, which allowed the mechanism of stabilisation to correlate with the stabilisation energy.

However, the structural analysis was still not conclusive for all residues. For example, β K150 had a significant effect on the 3D conformation of the local environment of this residue yet it is unclear whether this had any effect on the thermostability of mutant 9C. One way to do this would be to attempt some rationale design studies. The most effective strategy would be to employ rational design to eliminate probable destabilising mutations such as the D \rightarrow V (β 167) mutation at the heterotetramer interface of mutant 8C. Random recombination or DNA shuffling (Moore *et al.*, 1997; Stemmer, 1994) is proposed as a future strategy to further enhance the thermostability of the NHase biocatalyst. Using this strategy, it should be possible to construct a superior thermostable NHase mutant that contains multiple thermostability-enhancing mutations.

REFERENCES



UNIVERSITY *of the*
WESTERN CAPE

- Acharya, P., Rajakumara, E., Sankaranarayanan, R. and Rao, N. M.** 2004. Structural basis of selection and thermostability of laboratory evolved *Bacillus subtilis* lipase, *Journal of Molecular Biology*, **341**: 1271-1281
- Aharoni, A., Griffiths, A. D. and Tawfik, D. S.** 2005. High-throughput screens and selections of enzyme-encoding genes, *Current Opinion in Chemical Biology*, **9**: 210-216
- Akanuma, S. Yamagishi, A. Tanaka, N. and Oshima, T.** 1998. Serial increase in the thermostability of 3-isopropylmalate dehydrogenase from *Bacillus subtilis* by experimental evolution, *Protein Science*, **7**: 698-705
- Anderson, D. E., Hurley, J. H., Nicholson, H., Baase, W. A. and Matthews, B. W.** 1993. Hydrophobic core repacking and aromatic-aromatic interaction in the thermostable mutant of T4 lysozyme Ser177→Phe, *Protein Science*, **2**: 1285-1290
- Anderson, D. E., Bechtel, W. J. and Dahlquist, F. W.** 1990. pH-Induced denaturation of proteins: A salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme, *Biochemistry*, **29**: 2403-2408
- Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H., Frank, G. and Tratschin, J. D.** 1979. Thermal stability and protein structure, *Biochemistry*, **18**: 5698-5703.
- Arnold, F. H.** 1998. Design by directed evolution, *Accounts of Chemical Research*, **31**: 125-131
- Arnold, F. H.** 1996. Directed evolution: Creating biocatalysts for the future. *Chemical Engineering Science*, **51**: 5091-5102
- Arnold, F. H. and Volkov, A. A.** 1999. Directed evolution of biocatalysts, *Current Opinion in Chemical Biology*, **3**: 54-59

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

Arnott, M. A., Michael, R. A., Thompson, C. R., Hough, D. W. and Danson, M. J. 2000. Thermostability and thermoactivity of citrate synthases from the thermophilic and hyperthermophilic Archaea, *Thermoplasma acidophilum* and *Pyrococcus furiosus*, *Journal of Molecular Biology*, **304**: 657-668

Asano, Y., Tani, Y. and Yamada, H. 1980. A new enzyme "nitrile hydratase" which degrades acetonitrile in combination with amidase, *Agricultural and Biological Chemistry*, **44**: 2251-2252

Aucamp, J.P., Cosme, A. M., Lye, G. J. and Dalby P. A. 2005. High-throughput measurement of protein stability in microtiter plates, *Biotechnology and Bioengineering*, **89**: 599-607

Baker, E. N. and Hubbard, R. E. 1984. Hydrogen bonding in globular proteins, *Programmed Biophysical Molecular Biology*, **44**: 97-179

Banerjee, A., Sharma, R. and Banerjee, U. C. 2002. The nitrile-degrading enzymes: current status and future prospects, *Applied Microbiology and Biotechnology*, **60**: 33-44

Battistel, E., Benardi, A. and Maestri, P. 1997. Enzymatic contamination of aqueous polymer emulsions containing acrylonitrile, *Biotechnology Letters*, **19**: 131-134

Bauer, R., Knackmuss, H. -J. and Stolz, A. 1998. Enantioselective hydration of 2-arylpropionitriles by a NHase from *Agrobacterium tumefaciens* strain d3, *Applied Microbiology and Biotechnology*, **49**: 89-95

Bloom, J. D., Meyer, M. M., Meinhold, P., Otey, C. R., MacMillan, D. and Arnold, F. H. 2005. Evolving strategies for enzyme engineering, *Current Opinion in Structural Biology*, **15**: 447-452

Blundell, T. L. and Johnson, L. N. 1976. Protein crystallography, Academic Press, London

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

Björk, A., Dalhus, B., Mantzilas, D., Sirevåg, R. and Eijsink, V.G.H. 2004. Large improvement in the thermal stability of a tetrameric malate dehydrogenase by single point mutations at the dimer–dimer interface, *Journal of Molecular Biology*, **341**: 1215-1226

Bommarius, A. S., Broering, J. M., Chaparro-Riggers, J. F. and Polizzi, K. M. 2003. High-throughput screening for enhanced protein stability, *Current Opinion in Biotechnology*, **17**: 606-610

Bogin, O., Levin, I., Hacham, Y., Tel-Or, S., Peretz, M., Frolow, F. and Burstein, Y. 2002. Structural basis for the enhanced thermal stability of alcohol dehydrogenase mutants from the mesophilic bacterium *Clostridium beijerinckii*: contribution of salt bridging, *Protein Science*, **11**: 2561-2574

Bogin, O., Peretz, M., Hacham, Y., Korkhin, Y., Frolow, F., Kalb (Gilboa), A. J. and Burstein, Y. 1998. Enhanced thermal stability of *Clostridium beijerinckii* alcohol dehydrogenase after strategic substitution of amino acid residues with prolines from the homologous thermophilic *Thermoanaerobacter brockii* alcohol dehydrogenase, *Protein Science*, **7**: 1156-1163

- Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding, *Analytical Biochemistry*, **72**: 248-254
- Brunger, A. T.** 1993. Assessment of phase accuracy by cross validation: the free R value, Methods and Applications, *Acta Crystallographica D, Biological Crystallography*, **49**: 24-36
- Cadwell, R.C., and Joyce, G.F.** 1992. Randomization of genes by PCR mutagenesis, *PCR Methods and Applications*, **2**: 28-33
- Cadwell, R.C., and Joyce, G.F.** 1994. Mutagenic PCR mutagenesis, *PCR Methods and Applications*, **32**: 136-140
- Cameron, R. A.** 2003. Nitrile degrading enzymes from extreme environments, PhD thesis, University of London
- Cameron, R. A., Sayed, M. F. and Cowan, D. A.** 2005. Molecular analysis of nitrile catabolism operon of the thermophile *Bacillus pallidus* RAPc8, *Biochimica et Biophysica Acta*, **1725**: 35-46
- Chapatwala, K. D., Nawaz, M. S., Richardson, J. D. and Wolfham, J. H.** 1990. Isolation and characterization of acetonitrile utilizing bacteria. *Journal of Industrial Microbiology*, **5**: 65-70
- Chen, K. and Arnold, F. H.** 1993. Tuning the activity of an enzyme for unusual environments: Sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide, *Proceedings of the National Academy of Science USA*, **90**: 5618-5622

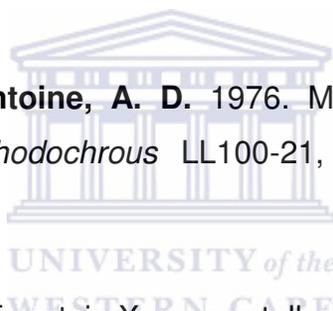
-
- Cirino, P. C., Mayer, K. M. and Umeno, D.** 2003. Generating mutant libraries using error-prone PCR. *Methods in Molecular Biology*, **231**: 3-9
- Clantin, B., Tricot, C., Lonhienne, T., Stalon, V. and Villeret, V.** 2001. Probing the role of oligomerisation in the high thermal stability of *Pyrococcus furiosus* ornithine carbamoyltransferase by site-specific mutants, *European Journal of Biochemistry*, **268**: 3937-2942
- Cowan, D.** 1996. Industrial enzyme technology, *Trends in Biotechnology*, **14**: 177-178
- Cowan, D. A., Arslanoglu, A., Burton, S. G., Baker, G. C., Cameron, R. A., Smith, J. J. and Meyer, Q.** 2004. Metagenomics, gene discovery and the ideal biocatalyst, *Biochemical Society Transactions*, **32**: 298-302
- Cowan, D. A., Cameron, R. A. and Tsekoa, T. L.** 2003. Comparative biology of mesophilic and thermophilic nitrile hydratases, *Advances in Applied Microbiology*, **52**: 123-158
- Cowan, D., Cramp, R., Pereira, R., Graham, D. and Almatawah, Q.** 1998. Biochemistry and biotechnology of mesophilic and thermophilic nitrile metabolising enzymes, *Extremophiles*, **2**: 207-216
- Cramp, R. A. and Cowan, D. A.** 1999. Molecular characterization of a novel thermophilic nitrile hydratase, *Biochimica et Biophysica Acta*, **143**: 249-260
- Daniel, R. M.** 1996. The upper limits of thermal stability, *Enzyme and Microbial Technology*, **19**: 74-79
- Das, R. and Gerstein, M.** 2000. The stability of thermophilic proteins: a study based on comprehensive genome comparison, *Functional and Integrative Genomics*, **1**: 76-88

Daugherty, P. S., Chen, G., Iverson, B. L. and Georgiou, G. 2000. Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies, *Proceedings of the National Academy of Science USA*, **97**: 2029-2034

Declerck, N., Machius, M., Joyet, P., Wiegand, G., Huber, R. and Gaillardin, C. 2003. Hyperstabilization of *Bacillus licheniformis* α -amylase and modulation of its stability over a 50°C temperature range. *Protein Engineering*, **16**: 287-293

Denisov V. P., Venu, K. Peters, J., Hörlein, H. D. and Halle, B. 1997. Orientational disorder and entropy of water in protein cavities. *Journal of Physical Chemistry B*, **101**: 9380-9389

DiGeronimo, M. J. and Antoine, A. D. 1976. Metabolism of acetonitrile and propionitrile by *Nocardia rhodochrous* LL100-21, *Applied and Environmental Microbiology*, **31**: 900-906



Drenth, J. 1994. Principles of protein X-ray crystallography, Springer-Verland, New York

Drummond, D. A., Iverson, B. L., Georgiou, G. and Arnold, F. H. 2005. Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins, *Journal of Molecular Biology*, **350**: 806-816

Duran, R., Nishiyama, M., Horinouchi, S and Beppu, T. 1993. Characterisation of nitrile hydratase genes cloned by DNA screening from *Rhodococcus erythropolis*, *Bioscience, Biotechnology and Biochemistry*, **57**: 1323-1328

Eijsink, V. G. H., Gåseidnes, S., Borchert, T. V. and van den Burg, B. 2005. Directed evolution of enzyme stability, *Biomolecular Engineering*, **22**: 21-30

- Eijsink, V. G. H., Bjørk, A., Gåseidnes, S., Sirevåg, R., Synstada, B., van den Burg, B. and Vriend, G.** 2004. Rational engineering of enzyme stability, *Journal of Biotechnology*, **113**: 105-120
- Eijsink, V. G. H., Vriend, G., Van den Burg, B., Van der Zee, J. R. and Venema, G.** 2002. Increasing the thermostability of a neutral protease by replacing positively charged amino acids in the N-terminal turn of α -helices, *Protein Engineering*, **5**: 165-170
- Eijsink, V. G. H., Vriend, G. and Van den Burg, B.** 2001. Engineering a hyperstable enzyme by manipulation of early steps in the unfolding process, *Biocatalysis Biotransformation*, **19**: 443-458
- Eckert, K. A. and Kunkel, T. A.** 1991. DNA polymerase fidelity and the polymerase chain reaction, *PCR Methods Application*, **1**: 17-24
- Endo, I., Nojiri, M., Tsujimura, M., Nakasako, M., Nagashima, S., Yohda, M., and Odaka, M.** 2001. Fe-type nitrile hydratase, *Journal of Inorganic Biochemistry*, **83**: 247 - 253
- Fallon, R. D., Stieglitz, B. and Turner, I. Jr.,** 1997. A *Pseudomonas putida* capable of stereoselective hydrolysis of nitriles, *Applied Microbiology and Biotechnology*, **47**: 156-161
- Farinas, E. T., Butler, T. and Arnold, F. H.** 2001. Directed enzyme evolution, *Current Opinion in Biotechnology*, **12**: 545-551
- Fawcett, J. and Scott, J.** 1960. A rapid and precise method for the determination of urea, *Journal of Clinical Pathology*, **13**: 156-159

- Fields, P. A.** 2001. Review: protein function at thermal extremes: balancing stability and flexibility, *Comparative Biochemistry and Physiology Part A*, **129**: 417-431
- Fleming, P. J. and Rose, G. D.** 2005. Do all backbone polar groups in proteins form hydrogen bonds? *Protein Science*, **14**: 1911-1917
- Forood, B., Feliciano, E. J. and Nambiar, K. P.** Stabilisation of α -helical structures in short peptides via end capping, *Proceedings of the National Academy of Science USA*, **90**: 838-842
- Fournand, D., Bigey, F. and Arnaud, A.** 1998. Acyl transfer activity of an amidase from *Rhodococcus* sp. strain R312: formation of a wide range of hydroxamic acids, *Applied and Environmental Microbiology*, **64**: 2844-2852
- Fukuchi, S. and Nishikawa, K.** 2001. Protein surface amino acid compositions distinctly differ between thermophilic and mesophilic bacteria. *Journal of Molecular Biology*, **309**: 835-843
- Garcia, A. E. and Sanbonmatsu, K. Y.** 2002. α -Helical stabilisation by side chain shielding of backbone hydrogen bonds, *Proceedings of the National Academy of Sciences USA*, **99**: 2782-2787
- Georgiou, G.** 2001. Analysis of large libraries of protein mutants using flow cytometry, *Advances in Protein Chemistry*, **55**: 293-315
- Georlette, D., Damien, B., Blaise, V., Depiereux, E., Uversky, V. N., Gerday, C. and Feller, G.** 2003. Structural and functional adaptations to extreme temperatures in psychrophilic, mesophilic and thermophilic DNA ligases. *Journal of Biological Chemistry*, **278**: 37015-37023

- Ghosh, T., Garde, S. and Garcia, A. E.** 2003. Role of backbone hydration and salt-bridge formation in stability of α -helix in solution, *Biophysical Journal*, **85**: 3187-3193
- Giver, L., Gershenson, A., Freskgard, P. -O. and Arnold, F. H.** 1998. Directed evolution of a thermostable esterase, *Proceedings of the National Academy of Science USA*, **95**: 12809-12813
- Graham, D., Pereira, R., Barfield, D. and Cowan, D. A.** 2000. Nitrile biotransformations using free and immobilized cells of thermophilic *Bacillus sp.*, *Enzyme and Microbial Technology*, **26**: 368-373
- Hakamada, Y., Hatada, Y., Ozawa, T., Ozaki, K., Kobayashi, T. and Ito, S.** 2001. Identification of thermostabilising residues in a *Bacillus* alkaline cellulase by construction of chimeras from mesophilic and thermostable enzymes and site-directed mutagenesis, *FEMS Microbiology Letters*, **195**: 67-72
- Haki, G. D. and Rakshit, S. K.** 2003. Developments of industrially important thermostable enzymes: a review, *Bioresource Technology*, **89**: 17-34
- Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R. and Olsen, G. J.** 1999. Thermal adaptation analysed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species, *Proceedings of the National Academy of Science USA*, **96**: 3578-3583
- Hann, E. C., Eisenberg, A., Fager, S. K., Perkins, N. E., Gallagher, F. G., Cooper, S. M., Gavagan, J. E., Stieglitz, B., Hennessey, S. M. and DiCosimo, R.** 1999. 5-Cyanovaleramide production using immobilised *Pseudomonas chlororaphis* B23, *Bioorganic and Medicinal Chemistry*, **7**: 2239-2245

Hashimoto, Y. Hosaka, H., Oinuma, K. I., Goda, M., Higashibata, H. and Kobayashi, M. 2005. Nitrile pathway involving Acyl-CoA synthase: overall metabolic gene organisation, and purification and characterisation of the enzyme, *Journal of Molecular Biology*, **280**: 8660-8667

Hashimoto, Y., Nishiyama, M., Yu, M. F., Watanabe, I., Horinouchi, S. and Beppu, T. 1992. Development of a host–vector system in a *Rhodococcus* strain and its use for expression of the cloned nitrile hydratase gene cluster, *Journal of General Microbiology*, **138**: 1003-1010

Hashimoto, Y., Nishiyama, M., Horinouchi, S. and Beppu, T. 1994. Nitrile hydratase gene from *Rhodococcus* sp. N-774 requirement for its downstream region for efficient expression, *Bioscience Biotechnology and Biochemistry*, **58**: 1859-1865

Hiraga, K. and Yutani, K. 1997. Roles of hydrogen bonding residues in the interaction between the α and β subunits in the tryptophan synthase complex, *Journal of Biological Chemistry*, **272**: 4935-4940

Hooft, R. W.W., Vriend, G., Sander, C. and Abola, E. E. 1996. Errors in protein structures, *Nature*, **381**: 272-272

Hook, R. H. and Robinson, W. G. 1964. Ricinine nitrilase. II. Purification and properties, *Journal of Biological Chemistry*, **239**: 4263-4267

Hourai, S., Miki, M., Takashima, Y., Mitsuda, S. and Yanagi, K. 2003. Crystal structure of nitrile hydratase from a thermophilic *Bacillus smithii*, *Biochemical and Biophysical Research Communications*, **312**: 340-345

Hough, D. W. and Danson, M. J. 1999. Extremozymes, *Current Opinion in Chemical Biology*, **3**: 39-46

- Huang, W., Jia, J., Cummings, J., Nelson, M., Schneider, G. and Lindqvist, Y.** 1997. Crystal structure of nitrile hydratase reveals a novel iron centre in a novel fold, *Structure*, **5**: 691-699
- Ikehata, O., Nishiyama, M., Horinouchi, S. and Beppu, T.** 1989. Primary structure of nitrile hydratase deduced from the nucleotide sequence of a *Rhodococcus* species and its expression in *Escherichia coli*, *European Journal of Biochemistry*, **181**: 563-570
- Imanaka, T., Shibasaki, M. and Takagi, M.** 1986. A new way of enhancing the thermostability of proteases, *Nature*, **324**: 695 - 697
- Jaenicke, R.** 2000. Stability and stabilisation of globular proteins in solution, *Journal of Biotechnology*, **79**: 193-203
- Jaenicke, R.** 1991. Protein stability and molecular adaptation to extreme conditions, *European Journal of Biochemistry*, **202**: 715-728
- Jaenicke, R. and Bøhm, G.** 1998. The stability of proteins in extreme environments, *Current Opinion in Structural Biology*, **8**: 738-748
- Jiang, X., Kowalski, J. and Kelly, J. W.** 2001. Increasing protein stability using a rational approach combining sequence homology and structural alignment: stabilising the WW domain, *Protein Science*, **10**: 1454-1465
- Jiang, L., Kuhlman, B., Kortemme, T. and Baker, D.** 2005. A “solvated rotamer” approach to modelling water-mediated hydrogen bonds at protein-protein interfaces, *PROTEINS: Structure, Function and Bioinformatics*, **58**: 893-904
- Jones, S., Marin, A. and Thornton, J. M.** 2000. Protein domain interfaces: characterisation and comparison with oligomeric protein interfaces, *Protein Engineering*, **13**: 77-82

- Jones, T. A., Zou, J. Y., Cowan, S. W. and Kjeldgaard, M.** 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallography A*, **47**: 110-119
- Kannan, N. and Vishveshwara, S.** 2000. Aromatic clusters: a determinant of thermal stability of thermophilic proteins, *Protein Engineering*, **13**: 753-761
- Karshikoff, A. and Ladenstein, R.** 2001. Ion pairs and the thermotolerance of proteins from hyperthermophiles: a “traffic rule” for hot roads, *Trends in Biochemical Sciences*, **26**: 550-556
- Kataeva, I. A., Uversky, V. N. and Ljungdahl, L. G.** 2003. Calcium and domain interactions contribute to the thermostability of domains of the multimodular cellobiohydrolase, CbhA, a subunit of the *Clostridium thermocellum* cellulosome, *Biochemistry Journal*, **372**: 151-161
- Kato, Y. and Asano, Y.** 2006. Molecular and enzymatic analysis of the “aldoxime-nitrile pathway” in the glutaronitrile degrader *Pseudomonas* sp. K-9, *Applied Microbial and Biotechnology*, **70**: 92-101
- Kato, Y., Ryoko, O. and Asano, Y.** 2000. Distribution of aldoxime dehydratase in microorganisms, *Applied and Environmental Microbiology*, **66**: 2290-2296
- Kato, Y., Tsuda, T. and Asano, Y.** 1999. Nitrile hydratase involved in aldoxime metabolism from *Rhodococcus* sp. Strain YH-3, *European Journal of Biochemistry*, **263**: 662-670
- Kato, Y., Yoshida, S., Xie, S. -X. and Asano, Y.** 2004. Aldoxime dehydratase co-existing with nitrile hydratase and amidase in the iron-type nitrile hydratase-producer *Rhodococcus* sp. N-771, *Journal of Bioscience and Bioengineering*, **97**: 250-259

- Kim J. -H., Choi, G. -S. Kim, S. -B. Kim, W. -H. Lee, J. -Y. Ryu, Y. -W. and Kim G. -J.** 2004. Enhanced thermostability and tolerance of high substrate concentration of an esterase by directed evolution. *Journal of Molecular Catalysis B: Enzymatic*, **27**: 169-175
- Kim, S. H., Padmakumar, R. and Oriel, P.** 2001. Cobalt activation of *Bacillus* BR449 thermostable nitrile hydratase expressed in *Escherichia coli*, *Applied Biochemistry and Biotechnology*, **91-3**: 597-603.
- Kim, S. and Oriel, P.** 2000. Cloning and expression of the nitrile hydratase and amidase genes from *Bacillus* sp. BR449 into *Escherichia coli*, *Enzyme and Microbial Technology*, **27**: 492-501.
- Kim, Y. -W., Choi, J. -H., Kim, J. -W., Park, C., Kim, J. -W., Cha, H., Lee, S. -B., Oh, B. -H., Moon, T. -W. and Park, K. -H.** 2003. Directed evolution of *Thermus* maltogenic amylase toward enhanced thermal resistance, *Applied and Environmental Microbiology*, **69**: 4866-4874
- Kleywegt, G. J. and Brunger, A. T.** 1996. Checking your imagination: applications of the free R value, *Structure*, **4**: 897-904
- Kobayashi, M., Nagasawa, T. and Yamada, H.** 1992. Enzymatic synthesis of acrylamide: a success story not yet over, *TRENDS in Biotechnology*, **10**: 402-408
- Kobayashi, M., Nishiyama, M., Nagasawa, T., Horinouchi, S., Beppu, T. and Yamada, H.** 1991. Cloning, nucleotide sequence and expression in *Escherichia coli* of two cobalt-containing nitrile hydratase genes from *Rhodococcus rhodochrous* J1, *Biochimica et Biophysica Acta*, **1129**: 23-33
- Kobayashi, M. and Shimizu, S.** 1998. Metalloenzyme nitrile hydratase: Structure, regulation, and application to biotechnology, *Nature Biotechnology*, **16**: 733-736

Kobayashi, M. and Shimizu, S. 2000. Nitrile hydrolases, *Current Opinion in Chemical Biology*, **4**: 95-102

Kohyama, E., Yoshimura, A., Aoshima, D., Yoshida, T. Kawamoto, H. and Nagasawa, T. 2006. Convenient treatment of acetonitrile-containing wastes using the tandem combination of nitrile hydratase and amidase-producing microorganisms, *Environmental Biotechnology*, **72**: 600-606

Komeda, H., Ishikawa, N. and Asano, Y. 2003. Enhancement of the thermostability and catalytic activity of D-stereospecific amino-acid amidase from *Ochrobactrum anthropi* SV3 by directed evolution, *Journal of Molecular Catalysis B: Enzymatic*, **21**: 283-290

Kovacs, J. A. 2004. Synthetic analogues of cysteinylated non-heme iron and non-corrinoid cobalt enzymes, *Chemistry Review*, **104**: 825-848

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

Kumar, S. and Nussinov, R. 2002. Relationship between ion pair geometries and electrostatic strengths in proteins. *Biophysical Journal*, **83**: 1595-1612

Kumar, S., Tsai, C. J. and Nussinov, R. 2000. Factors enhancing protein thermostability, *Protein Engineering*, **13**: 179-91

Kumar, S., Tsai, C. J. and Nussinov, R. 1999. Contribution of salt bridges toward protein thermostability, *Journal of Biomolecular Structure and Dynamics*, **11**: 79-85

Laemmli, U. K. 1970. Cleavage of structural proteins during assembly of the head of Bacteriophage T7, *Nature*, **227**: 680-685

Lasa, I. and Berenguer, J. 1993. Thermophilic enzymes and their biotechnological potential, *Microbiologia*, **9**: 77-89

Laskowski, R. A., Moss, D. S. and Thornton, J. M. 1993. Main-chain bond lengths and bond angles in protein structures, *Journal of Molecular Biology*, **231**: 1049-1067

Layh, N., Hirrlinger, B., Stolz, A. and Knackmuss, H. -J. 1997. Enrichment strategies for nitrile-hydrolysing bacteria, *Applied Microbiology and Biotechnology*, **47**: 668-674

Leung, D.W., Chen, E., and Goeddel, D.V. 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction, *Technique*, **1**: 11-15

Levy, Y. and Onuchic, J. N. 2006. Water mediation in protein folding and molecular recognition, *Annual Review of Biophysics and Biomolecular Structure*, **35**: 389-415

Li, W. T., Grayling, R. A., Sandman, K., Edmondson, S., Shriver, J. W. and Reeve, J. N. 1998. Thermodynamic stability of archaeal histones, *Biochemist*, **23**: 271-281

Li, Z. and Lazaridis, T. 2006. Water at biomolecular interfaces, *Physical Chemistry Chemical Physics*, **9**: 573-581

Li, T., Liu, J., Ohanja, D. -G. and Wong, F. -S. 2007. Biodegradation of organonitriles by adapted activated sludge consortium with acetonitrile-degrading microorganisms, *Water Research*, **41**: 3465-3473

- Liebeton, K. and Eck, J.** 2004. Identification and expression in *E. coli* of novel nitrile hydratases from the metagenome, *Engineering in Life Sciences*, **4**: 557-562
- Linardi, V. R., Dias, J. C. T. and Rosa, C. A.** 1996. Utilization of acetonitrile and other aliphatic nitriles by a *Candida famata* strain, *FEMS Microbiology Letters*, **144**: 67-71
- Ling, L. L., Keohavong, P., Dias, C. and Thilly, W. G.** 1991. Optimisation of the polymerase chain reaction with regard to fidelity: Modified T7, *Taq*, and Vent polymerases, *PCR Methods Application*, **1**: 63-69
- Lourenco, P. M., Almeida, T., Mendonca, D., Simoes, F. and Novo, C.** 2004. Searching for nitrile hydratase using the Consensus-Degenerate Hybrid Oligonucleotide Primers strategy, *Journal of Basic Microbiology*, **44**: 203-14.
- Lu, J., Zheng, Y., Yamagishi, H., Odaka, M., Tsujimura, M., Maeda, M. and Endo, I.** 2003. Motif CXCC in nitrile hydratase activator is critical for NHase biogenesis in vivo, *FEBS Letters*, **553**: 391-396
- Luke, K. and Wittung-Stafshede, P.** 2006. Folding and assembly pathways of co-chaperonin protein 10: origin of thermostability, *Archives of Biochemistry and Biophysics*, **456**: 8-18
- Luo and Baldwin,** 1997. Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water, *Biochemistry*, **36**: 8413-8421
- Lutz, S. and Patrick, W. M.** 2004. Novel methods for directed evolution of enzymes: quality, not quantity, *Current Opinion in Biotechnology*, **15**: 291-297

- Machius, M., Declerck, N., Huber, R. and Wiegand, G.** 2003. Kinetic stabilization of *Bacillus licheniformis* α -amylase through introduction of hydrophobic residues at the surface, *Journal of Biological Chemistry*, **278**: 11546-11553
- Maier-Greiner, U. H., Obermaier-Skrobranek, B. M. M., Estermaier, L. M., Kammerloher, W., Freund, C., Wulfing, C., Burkert, U. I., Breuer, M., Eulitz, M., Kufrevioglu, O. I. and Hartmann, G. R.** 1991. Isolation and properties of a nitrile hydratase from the soil fungus *Myrothecium verrucaria* that is highly specific for the fertilizer cyanamide and cloning of its gene, *Proceedings of the National Academy of Sciences USA*, **88**: 4260-4264
- Makhatadze, G. I., Loladze, V. V., Ermolenko, D. N., Chen, X and Thomas, S. T.** 2003. Contribution of surface salt bridges to protein stability: Guidelines for protein engineering, *Journal of Molecular Biology*, **327**: 1135-1148
- Makhongela, H. S., Glowacka, A. E., Agarkar, V. B., Sewell, B. T., Weber, B., Cameron, R. A. Cowan, D. A. and Burton, S. G.** 2007. A novel thermostable nitrilase superfamily amidase from *Geobacillus pallidus* showing acyl transfer activity, *Applied Microbiology and Biotechnology*, **75**: 801-811
- Marqusee, S. and Sauer, R. T.** 1994. Contribution of a hydrogen bond/salt bridge network to the stability of secondary and tertiary structure in λ repressor, *Protein Science*, **3**: 2217-2225
- Mascharak, P.** 2002. Structural and functional models of nitrile hydratase, *Coordination Chemistry Reviews*, **225**: 201-214
- Mattos, C.** 2002. Protein-water interactions in a dynamic world. *TRENDS in Biochemical Sciences*, **27**: 203-208

-
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J.** 2007. Phaser crystallographic software. *Journal of Applied Crystallography*, **40**: 658-674
- McDonald, I. K and Thornton, J. M.** 1994. Satisfying hydrogen bonding potential in proteins, *Journal of Molecular Biology*, **238**: 777-793
- McPherson, A.** 1999. *Crystallisation of biological macromolecules*, Cold Spring Harbor Press, New York
- McRee, D. E.** 1993. *Practical protein crystallography*, Academic Press: San Diego, California
- Merz, A. Knochel, T., Jansonius, J. N. and Kirschner, K.** 1999. The hyperthermostable indoleglycerol phosphate synthase from *Thermotoga maritima* is destabilised by mutational disruption of two solvent-exposed salt bridges, *Journal of Molecular Biology*, **288**: 753-763
- Meth-Cohn, O. and Wang, M. -X.** 1995. A powerful new nitrile hydratase for organic synthesis - aromatic and heteroaromatic nitrile hydrolyses - a rationalization, *Tetrahedron*, **36**: 9561-9564
- Meth-Cohn, O. and Wang, M.-X.** 1997. An in-depth study of the biotransformation of nitriles into amides and/or acids using *Rhodococcus rhodochrous* AJ270, *J. Chem. Soc., Perkin Trans.* **1**: 1099-1104
- Meyer, E.** 1992. Internal water molecules and H-bonding in biological macromolecules: A review of structural features with functional implications (1), *Protein Science*, **1**: 1543-1562

- Mitra, S. and Holz, R. C.** 2007. Unraveling the catalytic mechanism of nitrile hydratases. *Journal of Biological Chemistry*, **282**: 7397-7404
- Miyanaga, A., Fushinobu, S., Ito, K. and Wakagi, T.** 2001. Crystal structure of cobalt-containing nitrile hydratase, *Biochemical and Biophysical Research Communications*, **288**: 1169-1179
- Miyanaga, A., Fushinobu, S., Ito, K., Shoun, H. and Wakagi, T.** 2004. Mutational and structural analysis of cobalt-containing nitrile hydratase on substrate and metal binding, *European Journal of Biochemistry*, **271**: 429-438
- Miyazaki, K., Wintrode, P. L., Grayling, R. A., Rubingh, D. N. and Arnold, F. H.** 2000. Directed evolution study of temperature adaptation in a psychrophilic enzyme, *Journal of Molecular Biology*, **297**: 1015-1026
- Miyazaki, K. and Arnold, F. H.** 1999. Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function, *Journal of Molecular Evolution*, **49**: 716-720
- Moore, J. C., Jin, H. –M., Kuchner, O. and Arnold, F. H.** 1997. Strategies for the *in vitro* evolution of protein function: enzyme evolution by random recombination of improved sequences, *Journal of Molecular Biology*, **272**: 336-347
- Murshudov, G. N., Vagin, A. A. and Dodson, E. J.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica D, Biological Crystallography*, **53**: 240-255
- Myers, J. K. and Pace, C. N.** 1996. Hydrogen bonding stabilizes globular proteins, *Biophysical Journal*, **71**: 2033-2039

- Mylerová, V. and Martinková, L.** 2003. Synthetic Applications of Nitrile-Converting Enzymes, *Current Organic Chemistry*, **7**: 1-17
- Nagasawa, T. and Yamada, H.** 1990. Application of nitrile converting enzymes for the production of useful compounds, *Pure and Applied chemistry*, **62**: 1441-1444
- Nagashima, S., Nakasako, M., Dohmae, N., Tsujimura, M., Takio, K., Odaka, M., Yohda, M., Kamiya, N. and Endo, I.** 1998. Novel non-heme iron centre of nitrile hydratase with a claw setting of oxygen atoms, *Nature Structural Biology*, **5**: 347-351
- Nakai, K., Kidera, A. and Kanehisa, M.** 1988. Cluster analysis of amino acid indices for prediction of protein structure and function, *Protein Engineering*, **2**: 93-100
- Nakasako, M., Odaka, M., Yohda, M., Dohmae, N., Takio, K., Kamiya, N. and Endo, I.** 1999. Tertiary and quaternary structures of photoreactive Fe-type nitrile hydratase from *Rhodococcus* sp. N-771: Roles of hydration molecules in stabilising the structures and the structural origin of the substrate specificity of the enzyme, *Biochemistry*, **38**: 9887-9898
- Neylon, C.** 2004. Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution, *Nucleic Acids Research*, **32**: 1448-1459
- Nguyen, A. W. and Daugherty, P. S.** 2003. Production of randomly mutated plasmid libraries using mutator strains, *Methods in Molecular Biology*, **231**: 39-44
- Nishiyama, M., Horinouchi, S., Kobayashi, M., Nagasawa, T., Yamada, H. and T. Beppu, T.** 1991. Cloning and characterization of genes responsible for

metabolism of nitrile compounds from *Pseudomonas chlororaphis* B23, *Journal of Bacteriology*, **173**: 2465-2472

Niu, W. -N., Zhang, D. -W., Yu, M. -R. and Tan, T. -W. 2006. Improved thermostability of the optimum temperature of *Rhizopus arrhizus* lipase by directed evolution, *Journal of Molecular catalysis B: Enzymatic*, **43**: 33-39

Nojiri, M., Yohda, M., Odaka, M., Matsushita, Y., Tsujimura, M., Yoshida, T., Dohmae, N., Takio, K. and Endo, I. 1999. Functional expression of a nitrile hydratase in *Escherichia coli*: requirement of a nitrile hydratase from activator and post-translational modification of a ligand cysteine, *Journal of Biochemistry*, **125**: 696-704

Okamoto, S. and Eltis, L. D. 2007. Purification and characterisation of a novel hydratase from *Rhodococcus* sp. RHA1, *Molecular Microbiology*, **65**: 828-838

Pace, C. N., Horn, G., Hebert, E. J., Bechert, J., Shaw, K., Urbanikova, L., Scholtz, J. M. and Sevcik, J. 2001. Tyrosine hydrogen bonds make a large contribution to protein stability, *Journal of Molecular Biology*, **312**: 393-404

Pack, S. P. and Yoo, Y. J. 2004. Protein thermostability: structure-based difference of amino acid between thermophilic and mesophilic proteins, *Journal of Biotechnology*, **111**: 269-277

Padmakumar R. and Oriel P. (1999) Bioconversion of acrylonitrile to acrylamide using a thermostable nitrile hydratase, *Applied Biochemistry and Biotechnology*, **77-9**: 671-679

Pappenberger, G., Schurig, H. and Jaenicke, R. 1997. Disruption of an ionic network of leads to accelerated thermal denaturation of D-Glyceraldehyde-3-

phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*, *Journal of Molecular Biology*, **274**: 676-683

Parthasarathy, S. and Murthy, M. R. N. 2000. Protein thermal stability: insights from atomic displacement parameters (B values), *Protein Engineering*, **13**: 9-13

Payne, M. S., Wu, S., Fallon, R. D., Tudor, G., Stieglitz, B., Turner, I. M. Jr. and Nelson, M. J. 1997. A stereoselective cobalt-containing nitrile hydratase, *Biochemistry*, **36**: 5447-54

Peplowski, L., Kubiak, K. and Nowak, W. 2007. Insights into catalytic activity of industrial enzyme Co-nitrile hydratase. Docking studies of nitriles and amides, *Journal of Molecular Modeling*, **13**: 725-730

Pereira, R. A., Graham, D., Rainey, F. A., and Cowan, D. A. 1998. A novel thermostable nitrile hydratase, *Extremophiles*, **2**: 347-357

Petrillo, K. L., Wu, S., Hann, E. C., Cooling, F. B., Ben-Bassat, A., Gavagan, J. E., DiCosimo, R. and Payne, M. 2005. Over-expression in *Escherichia coli* of a thermally stable and region-selective nitrile hydratase from *Comamonas testosteroni* 5-MGAM-4D, *Applied Genetics and Molecular Biotechnology*, **67**: 664-670

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. 2004. UCSF Chimera - A visualisation system for exploratory research and analysis, *Journal of Computational Chemistry*, **25**: 1605-1612

Pflugrath, J. W. 1999. The finer things in X-ray diffraction data collection, *Acta Crystallographica D, Biological Crystallography*, **55**: 1718-1725

- Polizzi, K. M., Bommarius, A. S., Broering, J. M. and Chaparro-Riggers, J. F.** 2007. Stability of biocatalysts, *Current Opinion in Chemical Biology*, **11**: 220-225
- Precigou, S., Goulas, P. and Duran, R.** 2001. Rapid and specific identification of nitrile hydratase (NHase)-encoding genes in soil samples by polymerase chain reaction, *FEMS Microbiology Letters*, **204**: 155-161
- Prepechalova, I., Matnikova, L., Stolz, A., Ovesna, M., Bezouska, K., Kopecky, J. and Kren, V.** 2001. Purification and characterisation of the enantioselective nitrile hydratase from *Rhodococcus equi* A4, *Applied Microbiology and Biotechnology*, **55**: 150-156
- Puchkaev, A. V., Koo, L. S. and Ortiz de Montellano, P. R.** 2003. Aromatic stacking as a determinant of the thermal stability of CYP119 from *Sulfolobus solfataricus*, *Archives of Biochemistry and Biophysics*, **409**: 52-58
- Querol, E., Perez-Pons, J. A. and Mozo-Villarias A.** 1996. Analysis of protein conformational characteristics related to thermostability, *Protein Engineering*, **9**: 265-71
- Reddy, B. V. B., Datta, S. and Tiwari, S.** 1998. Use of propensities of amino acids to the local structural environments to understand effect of substitution mutations on protein stability, *Protein Engineering*, **11**: 1137-1145
- Rezende, R. P., Dias, J. C. T., Monteiro, A. S., Carraza, F. and Linardi, V. R.** 2003. The use of acetonitrile as the sole nitrogen and carbon source by *Geotrichum* sp. JR1, *Brazilian Journal of Microbiology*, **34**: 117-120
- Rhee, J. -K., Kim, D. -Y., Ahn, D. -G., Yun, J. -H., Jang, S. -W., Shin, H. -C., Cho, H. -S., Pan, J., -G. and Oh, J. -W.** 2006. Analysis of the thermostability determinants of hyperthermophilic esterase EstE1 based on its predicted three-

dimensional structure, *Applied and Environmental Microbiology*, **72**: 3021-3025

Rieu, J. -P., Boucherle, A., Cousse, H. and Gilbert Mouzin, G. 1986. Tetrahedron report number 205: Methods for the synthesis of antiinflammatory 2-aryl propionic acids, *Tetrahedron*, **42**: 4095-4131

Robinson-Rechavi, M., Alibés, A. and Godzik, A. 2006. Contribution of electrostatic interactions, compactness and quaternary structure to protein thermostability: Lessons from structural genomics of *Thermotoga maritima*, *Journal of Molecular Biology*, **356**: 547-557

Rossmann, M.G. (Ed.) 1972. The molecular replacement method. A collection of papers on the use of non-crystallographic symmetry, Gordon and Breach, Science Publishers, Inc., New York.

Russell, R. J. M., Ferguson, J. M., Hough, D. W., Danson, M. J. and Taylor, G. L. 1997. *Biochemistry*, **36**: 9983-9994

Sakaue, R. and Kajiyama, N. 2003. Thermostabilization of bacterial fructosyl-amino acid oxidase by directed evolution, *Applied and Environmental Microbiology*, **69**: 139-145

Salminen, T., Teplyakov, A., Kankare, J., Cooperman, B. S., Lahti, R. and Goldman, A. 1996. An unusual route to thermostability disclosed by comparison of *Thermus thermophilus* and *Escherichia coli* inorganic pyrophosphatases, *Protein Science*, **5**: 1014-1025

Sambrook, J. and Russell, D. W. 2001. Molecular cloning: A laboratory manual, Cold Spring Harbour Laboratory Press, New York, USA

- Serrano, L., Bucroft, M. and Fersht, A. R.** 1991. Aromatic-aromatic interactions and protein stability. Investigation by double-mutant cycles. *Journal of Molecular Biology*. 218: 465-475
- Shafikhani, S., Siegel, R. A., Ferrari, E. and Schellenberger, V.** 1997. Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization. *BioTechniques*, **23**:, 304–310.
- Shan, S. O., Loh, S. and Herschlag, D.** 1996. The energetics of hydrogen bonds in model systems: implications for enzymatic catalysis, *Science*, **272**: 97-101
- Shirley, B. A., Stanssens, P., Hahn, U. and Pace, C. N.** 1992. Contribution of hydrogen bonding to the conformational stability of ribonuclease T1, *Biochemistry*, 31: 725-732
- Smyth M. S. and Martin, J. H. J.** 2000. X-ray crystallography, *Journal of Clinical Pathology: Molecular Pathology*, **53**: 8-14
- Song, J. K. and Rhee, J. S.** 2000. Simultaneous enhancement of thermostability and catalytic activity of Phospholipase A₁ by evolutionary molecular engineering, *Applied and Environmental Microbiology*, **66**: 890-894
- Spasov, V. Z., Karshikoff, A. D. and Ladenstein, R.** 1995. The optimisation of protein-solvent interactions: Thermostability and the role of hydrophobic and electrostatic interactions, *Protein Science*, **4**: 1516-1527
- Stemmer, W. P.** 1994. Rapid evolution of a protein in vitro by DNA shuffling, *Nature*, **370**: 389-391

- Strausberg, S. L., Ruan, B., Fisher, K. E., Alexander, P. A. and Bryan, P. N.** 2005. Directed coevolution of stability and catalytic activity in calcium-free subtilisin, *Biochemistry*, **44**: 3272-3279
- Strop, P and Majo, S. L.** 2000. Contribution of salt bridges to protein stability, *Biochemistry*, **39**: 1251-1255
- Suen, W. C., Zhang, N., Xiao, L., Madison, V. and Zaks, A.** 2004. Improved activity and thermostability of *Candida antarctica*, *Protein Engineering Design and Selection*, **17**: 133-140
- Takano, K., Yamagata, Y. Funahashi, J., Hioki, Y., Kuramitsu, S. and Yatani, K.** 1999. Contribution of intra- and intermolecular hydrogen bonds to the conformational stability of human lysozyme. *Biochemistry*, **38**: 12698-12708
- Takano, K., Tsuchimori, K., Yamagata, Y. and Yutani, K.** 2000. Contribution of salt bridges near the surface of a protein to the conformational stability. *Biochemistry*, **39**: 12375-12381
- Takarada, H., Kawano, Y., Hashimoto, K., Nakayama, H., Ueda, S., Yohda, M., Kamiya, N., Dohmae, N., Maeda, M. and Odaka, M.** 2006. Mutational study on α Gln90 of Fe-type nitrile hydratase from *Rhodococcus* sp. N771, *Bioscience, Biotechnology and Biochemistry*, **70**: 881-889
- Takashima, Y., Yamaga, Y. and Mitsuda, S.** 1998. Nitrile hydratase from a thermophilic *Bacillus smithii*, *Journal of Industrial Microbiology and Biotechnology*, **20**: 220-226
- Tamakoshi, M., Nakano, Y., Kakizawa, S., Yamagishi, A. and Oshima, T.** 2001. Selection of stabilized 3-isopropylmalate dehydrogenase of *Saccharomyces*

cerevisiae using the host-vector system of an extreme thermophile, *Thermus thermophilus*, *Extremophiles*, **5**:17-22

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

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

Teplyakov, A. V., Kuranova, I. P., Harutyunyan, E. H., Vainshtein, B., Frömmel, C., Höhne, W. E. and Wilson, K. S. 1990. Crystal structure of thermitase at 1.4Å resolution, *Journal of Molecular Biology*, **214**: 261-279

Tickle, I. J. and Driessen, H. P. C. 1996. Molecular replacement using known structural information, *Methods in Molecular Biology*, **56**: 173-203

Thimann, K. V. and Mahadevan, S. 1964a. Nitrilase: occurrence, preparation and general mode of action, *Arch Biochem Biophys*, **105**: 133-141

Thimann, K. V. and Mahadevan, S. 1964b. Nitrilase: substrate specificity and possible mode of action, *Arch Biochem Biophys*, **107**: 62-68

Thomas, S. M., DiCosimo, R. and Nagarajan, V. 2002. Biocatalysis: applications and potentials for the chemical industry, *Trends in Biotechnology*, **6**: 238-242

Torrez, M., Schultehenrich, M and Livesay, D. R. 2003. Conferring thermostability to mesophilic proteins through optimised electrostatic interactions, *Biophysical Journal*, **85**: 2845-2853

- Tsekoa, T.** 2005. Structure, enzymology and engineering of *Bacillus pallidus* RAPc8 nitrile hydratase, PhD Biotechnology thesis; University of the Western Cape
- Tsekoa, T., Sayed, M. F., Sewell, T. and Cowan, D. A.** 2004. Preliminary structural analysis of a thermostable nitrile hydratase, *South African Journal of Science*, **100**: 488-490
- Turner, P., Marno, G. and Nordberg Karlsson, E.** 2007. Potential and utilisation of thermophiles and thermostable enzymes in biorefining, *Microbial Cell Factories*, **6**: 9-32
- Van den Burg, B.** 2003. Extremophiles as a source for novel enzymes, *Current Opinion in Microbiology*, **6**: 213-218
- Van den Burg, B. and Eijsink, V. G. H.** 2002. Selection of mutations for increased protein stability, *Current Opinion in Biotechnology*, **13**: 333-337
- Van den Berg, B., Vriend, G., Veltman, O. R., Venema, G. and Eijsink, V. G. H.** 1998. Engineering an enzyme to resist boiling, *Proceedings of the National Academy of Science USA*, **95**: 2056-2060
- Vanhercke, T., Ampe, C., Tirry, L. and Denolf, P.** 2005. Reducing mutational bias in random protein libraries, *Analytical Biochemistry*, **339**: 9-14
- Veltman, O. R., Vriend, G., Middelhoven, P. J., Van den Burg, B., Venema, G. and Eijsink, V. G. H.** 1996. Analysis of the structural determinants of the stability of thermolysinlike proteases by molecular modeling and site-directed mutagenesis, *Protein Engineering*, **9**: 1181-1189
- Vetriani, C., Maeder, D. L., Tolliday, N., Yip, K. S. -P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H. and Robb, F. T.** 1998. Protein thermostability above

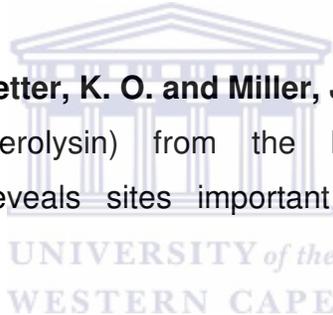
100°C: a key role for ionic interactions, *Proceedings of the National Academy of Science USA*, **95**: 12300-12305

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

Vihinen, M. 1987. Relationship of protein flexibility to thermostability, *Protein Engineering*, **1**: 477-480

Vogt, G., Woell, S. and Argos, P. 1997. Protein thermal stability, hydrogen bonds and ion pairs. *Journal of Molecular Biology*, **269**: 631-643

Volkl, P., Markiewicz, P., Stetter, K. O. and Miller, J. H. 1994. The sequence of a subtilisin-type protease (aerolysin) from the hyperthermophilic archaeum *Pyrobaculum aerophilum* reveals sites important for thermostability, *Protein Science*, **3**: 1329-1340



Wang, M. -X. and Feng, G. -Q. 2002. Nitrile biotransformation for highly enantioselective synthesis of 3-substituted 2,2-dimethylcyclopropanecarboxylic acids and amides, *Journal of Organic Chemistry*, **68**: 621-624

Wang, M. -X., Lu, G., Ji, G. -J., Huang, Z. -T., Meth-Cohn, O. and Colby, J. 2000. Enantioselective biotransformations of racemic α -substituted phenylacetone nitriles and phenylacetamides using *Rhodococcus* sp. AJ270, *Tetrahedron: Asymmetry*, **11**: 1123-1135

Watanabe, I., Satoh, Y., Enomoto, K, Seki, S. and Sakashita, K. 1987. Opticicultural conditions for cultivation of *Rhodococcus* sp N-774 and for conversion of acrylonitrile to acrylamide by resting cells, *Agriculture and Biological Chemistry*, **51**: 3201-3206

- Wieser, M., Takeuchi, K., Wada, Y., Yamada, H. and Nagasawa, T.** 1998. Low-molecular-mass nitrile hydratase from *Rhodococcus rhodochrous* J1: purification, substrate specificity and comparison with the analogous high-molecular-mass enzyme, *FEMS Microbiology Letters*, **169**: 17-22
- Wintrode, P. L., Miyazaki, K. and Arnold, F. H.** 2001. Patterns of adaptation in a laboratory evolved thermophilic enzyme, *Biochimica et Biophysica Acta*, **1549**: 1-8
- Wong, T. S., Roccatano, D. Zacharias, M and Schwaneberg, U.** 2005. A statistical analysis of random mutagenesis methods used for directed protein evolution, *Journal of Molecular Biology*, **355**: 858-871
- Woodyer, R., Chen, W. and Zhao, H.** 2004. Outrunning nature: Directed evolution of superior biocatalysts, *Journal of Chemical Education*, **81**: 126-133
- Wu, S., Fallon, R. D. and Payne M. S.** 1997. Over-production of stereoselective nitrile hydratase from *Pseudomonas putida* 5B in *Escherichia coli*: activity requires a novel downstream protein, *Applied Microbiology and Biotechnology*, **48**: 704-708
- Wu, Z. -L. and Li, Z. -Yi.** 2002. Enhancement of enzyme activity and enantioselectivity via cultivation in nitrile metabolism by *Rhodococcus* sp. CGMCC 0497, *Biotechnology and Applied Biochemistry*, **35**: 61-67
- Wyatt, J. M. and Knowles, C. J.** 1995. The development of a novel strategy for the microbial treatment of acrylonitrile effluents, *Biodegradation*, **6**: 93-107
- Xiao, L. and Honig, B.** 1999. Electrostatic contributions to the stability of hyperthermophilic proteins, *Journal of Molecular Biology*, **289**: 1435-1444

- Xie, S. -X., Kato, Y., Komeda, H., Yoshida, S. and Asano, Y.** 2003. A gene cluster responsible for alkylaldoxime metabolism coexisting with nitrile hydratase and amidase in *Rhodococcus globerulus* A-4, *Biochemistry*, **42**: 12056-12066
- Xu, D., Tsai, C. -J. and Nussinov, R.** 1997. Hydrogen bonds across protein-protein interfaces, *Protein Engineering*, **10**: 999-1012
- Xu, Z., Liu, Y., Yang, Y., Jiang, W., Arnold, E. and Ding, J.** 2003. Crystal structure of D-Hydantoinase from *Burkholderia pickettii* at a resolution of 2.7 angstroms: insights into the molecular basis of enzyme thermostability, *Journal of Bacteriology*, **185**: 4038-4049
- Yamada, H. and Kobayashi, M.** 1996. Nitrile hydratase and its application to industrial production of acrylamide, *Bioscience, Biotechnology and Biochemistry*, **60**: 1391-1400
- Yamaki, T., Oikawa, T., Ito, K. and Nakamura, T.** 1997. Cloning and sequencing of a nitrile hydratase gene from *Pseudonocardia thermophila* JCM3095, *Journal of Fermentation and Bioengineering*, **83**: 474-477
- Yano, J. K. and Poulos, T. L.** 2003. New understandings of thermostable and peizostable enzymes, *Current Opinion in Biotechnology*, **14**: 360-365
- You, L. and Arnold, F. H.** 1996. Directed evolution of subtilisin E in *Bacillus subtilis* to enhance total activity in aqueous dimethylformamide, *Protein Engineering*, **9**: 77-83
- Yu, H. -M., Shi, Y., Tian, Z. -L., Zhu, Y. -Q and Shen, Z. -Y.** 2006. An over expression and high efficient mutation system of a cobalt-containing nitrile hydratase, *Journal of Molecular Catalysis B: Enzymatic*, **43**: 80-85

- Zaccolo, M. and Gherardi, E.** 1999. The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase, *Journal of Molecular Biology*, **285**: 775-783
- Zaks, A.** 2001. Industrial biocatalysis, *Current Opinion in Chemical Biology*, **5**: 130-136
- Zhang, X. -J., Baase, W. A., Shoichet, B. K., Wilson, K. P. and Matthews, B. W.** 1995. Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive, *Protein Engineering*, **8**: 1017-1022
- Zavodszky, P., Kardos, J., Svingor, A. and Petsko, G.** 1998. Adjustment of conformational flexibility is a key even in the thermal adaptation of proteins, *Proceedings of the National Academy of Sciences USA*, **95**: 7406-7411
- Zhao, H. and Arnold, F. H.** 1997. Combinatorial protein design: strategies for screening protein libraries, *Current Opinion in Structural Biology*, **7**: 480-485
- Zhao, H. and Arnold, F. H.** 1999. Directed evolution converts subtilisin E into a functional equivalent of thermitase, *Protein Engineering*, **12**: 47-53
- Zhou, H. -X.** 2002. Toward the physical basis of thermophilic proteins: linking of enriched polar interactions and reduced heat capacity of unfolding, *Biophysical Journal*, **83**: 3126-3133
- Zhao, H., Giver, L., Shao, Z., Affholter, J. A. and Arnold, F. H.** 1998. Molecular evolution by staggered extension process (StEP) in vitro recombination, *Nature Biotechnology*, **16**: 258-261

Zhou, Z., Hashimoto, Y. and Kobayashi, M. 2005. Nitrile degradation by *Rhodococcus*: Useful microbial metabolism for industrial productions, *Actinomycetologica*, **19**: 18-26



APPENDICES



UNIVERSITY *of the*
WESTERN CAPE

Nucleotide alignment of wild-type and thermostabilised mutants α subunit

```

      10      20      30      40      50      60      70      80      90
alpha
21_6b_NHOP4  ATGACGATTGATCAAAAAAATACTAATATAGATCCAAAGATTTCCACATCATCATCCGCGTCCAACATCATTTTGGGAGGCCGTCGAAAA
70_7d_Nhop4  ATGACGATTGATCGAAAAAATACTAATATAGATCCAAAGATTTCCACATCATCATCCGCGTCCAACATCATTTTGGGAGGCCGTCGAAAA
13_7g_Nhop4  ATGACGATTGATCAAAAAAATACTAATATAGATCCAAAGATTTCCACATCATCATCCGCGTCCAACATCATTTTGGGAGGCCGTCGAAAA
71_8c_Nhop4  ATGACGATTGATCAAAAAAATACTAATATAGATCCAAAGATTTCCACATCATCATCCGCGTCCAACATCATTTTGGGAGGCCGTCGAAAA
22_Lib0_1C_9C_Nhop4 ATGACGATTGATCAAAAAAATACTAATATAGATCCAAAGATTTCCACATCATCATCCGCGTCCAACATCATTTTGGGAGGCCGTCGAAAA
16_9e_Nhop4  ATGACGATTGATCAAAAAAATACTAATATAGATCCAAAGATTTCCACATCATCATCCGCGTCCAACATCATTTTGGGAGGCCGTCGAAAA

      100     110     120     130     140     150     160     170     180
alpha
21_6b_NHOP4  GCTCTTGAATCCTTGTTGATTGAGAAAGGGCATCTTTCTCAGATGCTATTGAAAAGGTAATAAAAATTATGAGCATGAGCTGGGACCA
70_7d_Nhop4  GCTCTTGAATCCTTGTTGATTGAGAAAGGGCATCTTTCTCAGATGCTATTGAAAAGGTAATAAAAATTATGAGCATGAGCTGGGACCA
13_7g_Nhop4  GCTCTTGAATCCTTGTTGATTGAGAAAGGGCATCTTTCTCAGATGCTATTGAAAAGGTAATAAAAATTATGAGCATGAGCTGGGACCA
71_8c_Nhop4  GCTCTTGAATCCTTGTTGATTGAGAAAGGGCATCTTTCTCAGATGCTATTGAAAAGGTAATAAAAATTATGAGCATGAGCTGGGACCA
22_Lib0_1C_9C_Nhop4 GCTCTTGAATCCTTGTTGATTGAGAAAGGGCATCTTTCTCAGATGCTATTGAAAAGGTAATAAAAATTATGAGCATGAGCTGGGACCA
16_9e_Nhop4  GCTCTTGAATCCTTGTTGATTGAGAAAGGGCATCTTTCTCAGATGCTATTGAAAAGGTAATAAAAATTATGAGCATGAGCTGGGACCA

      190     200     210     220     230     240     250     260     270
alpha
21_6b_NHOP4  ATGAACGGAGCAAAGGTCGTAGCGAAGGCTTGGACTGATCCTGCTTTTAAAACAAAGATTGCTAGAAGATTAGAGACTGTATTAAGGGAG
70_7d_Nhop4  ATGAACGGAGCAAAGGTCGTAGCGAAGGCTTGGACTGATCCTGCTTTTAAAACAAAGATTGCTAGAAGATTAGAGACTGTATTAAGGGAG
13_7g_Nhop4  ATGAACGGAGCAAAGGTCGTAGCGAAGGCTTGGACTGATCCTGCTTTTAAAACAAAGATTGCTAGAAGATTAGAGACTGTATTAAGGGAG
71_8c_Nhop4  ATGAACGGAGCAAAGGTCGTAGCGAAGGCTTGGACTGATCCTGCTTTTAAAACAAAGATTGCTAGAAGATTAGAGACTGTATTAAGGGAG
22_Lib0_1C_9C_Nhop4 ATGAACGGAGCAAAGGTCGTAGCGAAGGCTTGGACTGATCCTGCTTTTAAAACAAAGATTGCTAGAAGATTAGAGACTGTATTAAGGGAG
16_9e_Nhop4  ATGAACGGAGCAAAGGTCGTAGCGAAGGCTTGGACTGATCCTGCTTTTAAAACAAAGATTGCTAGAAGATTAGAGACTGTATTAAGGGAG

      280     290     300     310     320     330     340     350     360
alpha
21_6b_NHOP4  CTAGGATACTATGGTTTACAGGGTGAGCATATCAGGGTAGTAGAAAAACGGATACGGTACACAAATGTTGTAGTCTGCACTTTATGTTCA
70_7d_Nhop4  CTAGGATACTATGGTTTACAGGGTGAGCATATCAGGGTAGTAGAAAAACGGATACGGTACACAAATGTTGTAGTCTGCACTTTATGTTCA
13_7g_Nhop4  CTAGGATACTATGGTTTACAGGGTGAGCATATCAGGGTAGTAGAAAAACGGATACGGTACACAAATGTTGTAGTCTGCACTTTATGTTCA
71_8c_Nhop4  CTAGGATACTATGGTTTACAGGGTGAGCATATCAGGGTAGTAGAAAAACGGATACGGTACACAAATGTTGTAGTCTGCACTTTATGTTCA
22_Lib0_1C_9C_Nhop4 CTAGGATACTATGGTTTACAGGGTGAGCATATCAGGGTAGTAGAAAAACGGATACGGTACACAAATGTTGTAGTCTGCACTTTATGTTCA
16_9e_Nhop4  CTAGGATACTATGGTTTACAGGGTGAGCATATCAGGGTAGTAGAAAAACGGATACGGTACACAAATGTTGTAGTCTGCACTTTATGTTCA

      370     380     390     400     410     420     430     440     450
alpha
21_6b_NHOP4  TGTACCCCTGGCCATTGCTTGGTTTACCGCCCTCATGGTACAAAGAAACCTGCTTATAGAGCTCGTGTCTGTAAGAGCCGAGACAAGTG
70_7d_Nhop4  TGTACCCCTGGCCATTGCTTGGTTTACCGCCCTCATGGTACAAAGAAACCTGCTTATAGAGCTCGTGTCTGTAAGAGCCGAGACAAGTG
13_7g_Nhop4  TGTACCCCTGGCCATTGCTTGGTTTACCGCCCTCATGGTACAAAGAAACCTGCTTATAGAGCTCGTGTCTGTAAGAGCCGAGACAAGTG
71_8c_Nhop4  TGTACCCCTGGCCATTGCTTGGTTTACCGCCCTCATGGTACAAAGAAACCTGCTTATAGAGCTCGTGTCTGTAAGAGCCGAGACAAGTG
22_Lib0_1C_9C_Nhop4 TGTACCCCTGGCCATTGCTTGGTTTACCGCCCTCATGGTACAAAGAAACCTGCTTATAGAGCTCGTGTCTGTAAGAGCCGAGACAAGTG
16_9e_Nhop4  TGTACCCCTGGCCATTGCTTGGTTTACCGCCCTCATGGTACAAAGAAACCTGCTTATAGAGCTCGTGTCTGTAAGAGCCGAGACAAGTG

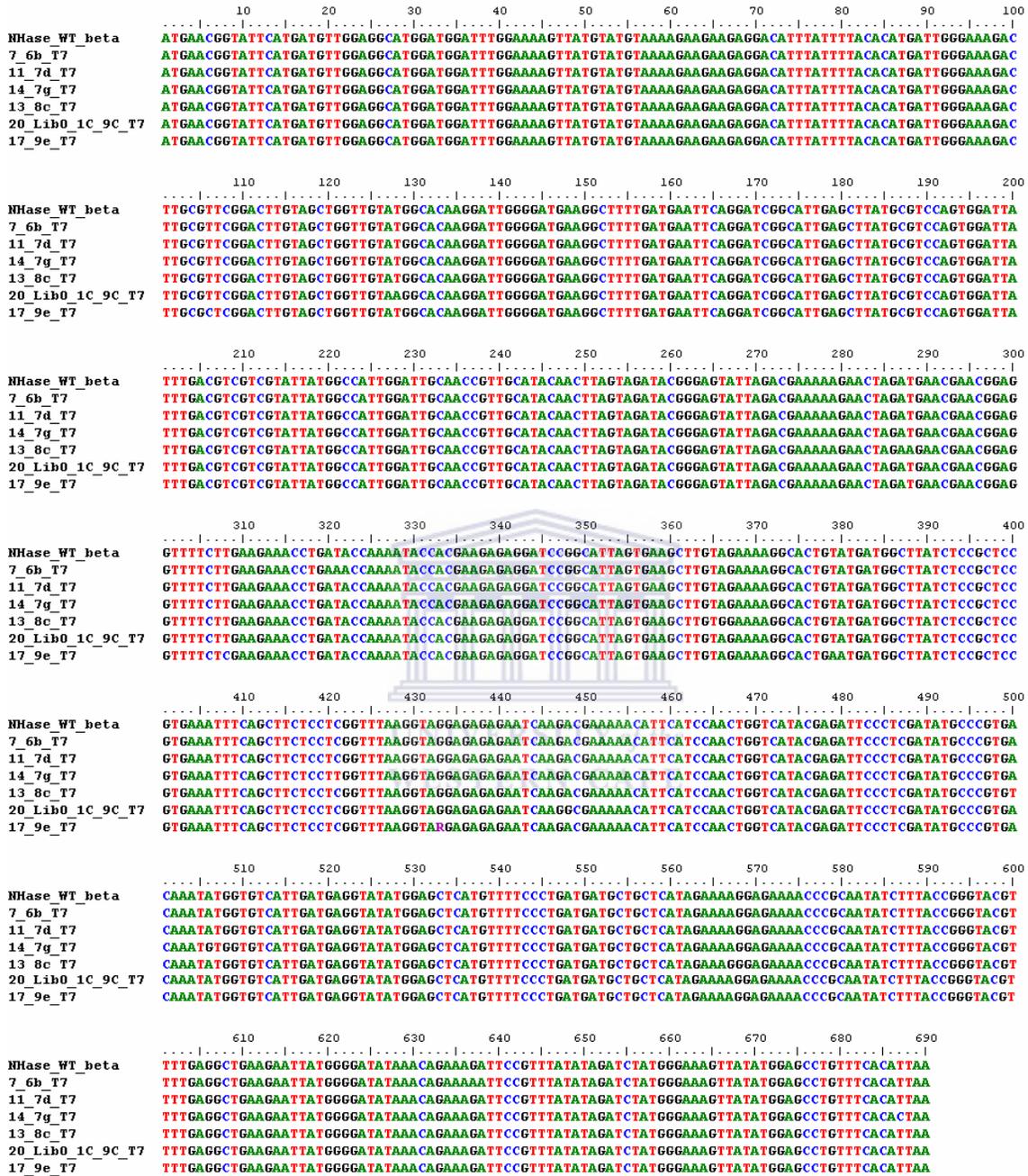
      460     470     480     490     500     510     520     530     540
alpha
21_6b_NHOP4  TTGAAAAGAAATTCGGATTAGATCTTCCAGATTCCAGTAGAAATCCGGGTATGGGACAGCAATTCCAGAAATTCGCTTTATGTTATTGCCGCAA
70_7d_Nhop4  TTGAAAAGAAATTCGGATTAGATCTTCCAGATTCCAGTAGAAATCCGGGTATGGGACAGCAATTCCAGAAATTCGCTTTATGTTATTGCCGCAA
13_7g_Nhop4  TTGAAAAGAAATTCGGATTAGATCTTCCAGATTCCAGTAGAAATCCGGGTATGGGACAGCAATTCCAGAAATTCGCTTTATGTTATTGCCGCAA
71_8c_Nhop4  TTGAAAAGAAATTCGGATTAGATCTTCCAGATTCCAGTAGAAATCCGGGTATGGGACAGCAATTCCAGAAATTCGCTTTATGTTATTGCCGCAA
22_Lib0_1C_9C_Nhop4 TTGAAAAGAAATTCGGATTAGATCTTCCAGATTCCAGTAGAAATCCGGGTATGGGACAGCAATTCCAGAAATTCGCTTTATGTTATTGCCGCAA
16_9e_Nhop4  TTGAAAAGAAATTCGGATTAGATCTTCCAGATTCCAGTAGAAATCCGGGTATGGGACAGCAATTCCAGAAATTCGCTTTATGTTATTGCCGCAA

      550     560     570     580     590     600     610     620     630
alpha
21_6b_NHOP4  AGACCTGAAAGGTAACGGAAGGAATGACGGAGGAGGACTTGCAAAACCTTGTACTCGAGACTCCATGATTTGGTGTCTGCTAAAAATAGAGCCG
70_7d_Nhop4  AGACCTGAAAGGTAACGGAAGGAATGACGGAGGAGGACTTGCAAAACCTTGTACTCGAGACTCCATGATTTGGTGTCTGCTAAAAATAGAGCCG
13_7g_Nhop4  AGACCTGAAAGGTAACGGAAGGAATGACGGAGGAGGACTTGCAAAACCTTGTACTCGAGACTCCATGATTTGGTGTCTGCTAAAAATAGAGCCG
71_8c_Nhop4  AGACCTGAAAGGTAACGGAAGGAATGACGGAGGAGGACTTGCAAAACCTTGTACTCGAGACTCCATGATTTGGTGTCTGCTAAAAATAGAGCCG
22_Lib0_1C_9C_Nhop4 AGACCTGAAAGGTAACGGAAGGAATGACGGAGGAGGACTTGCAAAACCTTGTACTCGAGACTCCATGATTTGGTGTCTGCTAAAAATAGAGCCG
16_9e_Nhop4  AGACCTGAAAGGTAACGGAAGGAATGACGGAGGAGGACTTGCAAAACCTTGTACTCGAGACTCCATGATTTGGTGTCTGCTAAAAATAGAGCCG

      640     650     660
alpha
21_6b_NHOP4  CCTAAAAGTTACGGTAGGTTAGGAGGAAAAATA
70_7d_Nhop4  CCTAAAAGTTACGGTAGGTTAGGAGGAAAAATA
13_7g_Nhop4  CCTAAAAGTTACGGTAGGTTAGGAGGAAAAATA
71_8c_Nhop4  CCTAAAAGTTACGGTAGGTTAGGAGGAAAAATA
22_Lib0_1C_9C_Nhop4 CCTAAAAGTTACGGTAGGTTAGGAGGAAAAATA
16_9e_Nhop4  CCTAAAAGTTACGGTAGGTTAGGAGGAAAAATA

```

Nucleotide alignment of wild-type and thermostabilised mutants β subunit



Nucleotide alignment of wild-type and thermostabilised mutants *p14k*

```

.....10.....20.....30.....40.....50.....60.....70.....80.....90.....100
P14K      ATGAAAAGTTGTGAGAATCAACCTAATGAATCATTGCTTTCGAAATATGCTGAAGAAATCGCACCTCCTAGAAAAACGGAGAGTTAGAAATCCAAAGAGC
14_6h_T7term ATGAAAAGTTGTGAGAATCAACCTAATGAATCATTGCTTTCGAAATATGCTGAAGAAATCGCACCTCCTAGAAAAACGGAGAGTTAGAAATCCAAAGAGC
12_7d_T7Term ATGAAAAGTTGTGAGAATCAACCTAATGAATCATTGCTTTCGAAATATGCTGAAGAAATCGCACCTCCTAGAAAAACGGAGAGTTAGAAATCCAAAGAGC
58_7g_T7Term ATGAAAAGTTGTGAGAATCAACCTAATGAATCATTGCTTTCGAAATATGCTGAAGAAATCGCACCTCCTAGAAAAACGGAGAGTTAGAAATCCAAAGAGC
14_8c_T7Term ATGAAAAGTTGTGAGAATCAACCTAATGAATCATTGCTTTCGAAATATGCTGAAGAAATCGCACCTCCTAGAAAAACGGAGAGTTAGAAATCCAAAGAGC
21_Lib0_1C_9C_T7Term ATGAAAAGTTGTGAGAATCAACCTAATGAATCATTGCTTTCGAAATATGCTGAAGAAATCGCACCTCCTAGAAAAACGGAGAGTTAGAAATCCAAAGAGC
59_9e_T7Term ATGAAAAGTTGTGAGAATCAACCTAATGAATCATTGCTTTCGAAATATGCTGAAGAAATCGCACCTCCTAGAAAAACGGAGAGTTAGAAATCCAAAGAGC

.....110.....120.....130.....140.....150.....160.....170.....180.....190.....200
P14K      CTTGGGAAAGACGCTCTTTTGGCATGACTCTTTCCTTGTACGAAGAAAAGCTGTATAGCTCTTGGGAGGATTTTCGATCCCCTTTGATTGAGGAGATCAA
14_6h_T7term CTTGGGAAAGACGCTCTTTTGGCATGACTCTTTCCTTGTACGAAGAAAAGCTGTATAGCTCTTGGGAGGATTTTCGATCCCCTTTGATTGAGGAGATCAA
12_7d_T7Term CTTGGGAAAGACGCTCTTTTGGCATGACTCTTTCCTTGTACGAAGAAAAGCTGTATAGCTCTTGGGAGGATTTTCGATCCCCTTTGATTGAGGAGATCAA
58_7g_T7Term CTTGGGAAAGACGCTCTTTTGGCATGACTCTTTCCTTGTACGAAGAAAAGCTGTATAGCTCTTGGGAGGATTTTCGATCCCCTTTGATTGAGGAGATCAA
14_8c_T7Term CTTGGGAAAGACGCTCTTTTGGCATGACTCTTTCCTTGTACGAAGAAAAGCTGTATAGCTCTTGGGAGGATTTTCGATCCCCTTTGATTGAGGAGATCAA
21_Lib0_1C_9C_T7Term CTTGGGAAAGACGCTCTTTTGGCATGACTCTTTCCTTGTACGAAGAAAAGCTGTATAGCTCTTGGGAGGATTTTCGATCCCCTTTGATTGAGGAGATCAA
59_9e_T7Term CTTGGGAAAGACGCTCTTTTGGCATGACTCTTTCCTTGTACGAAGAAAAGCTGTATAGCTCTTGGGAGGATTTTCGATCCCCTTTGATTGAGGAGATCAA

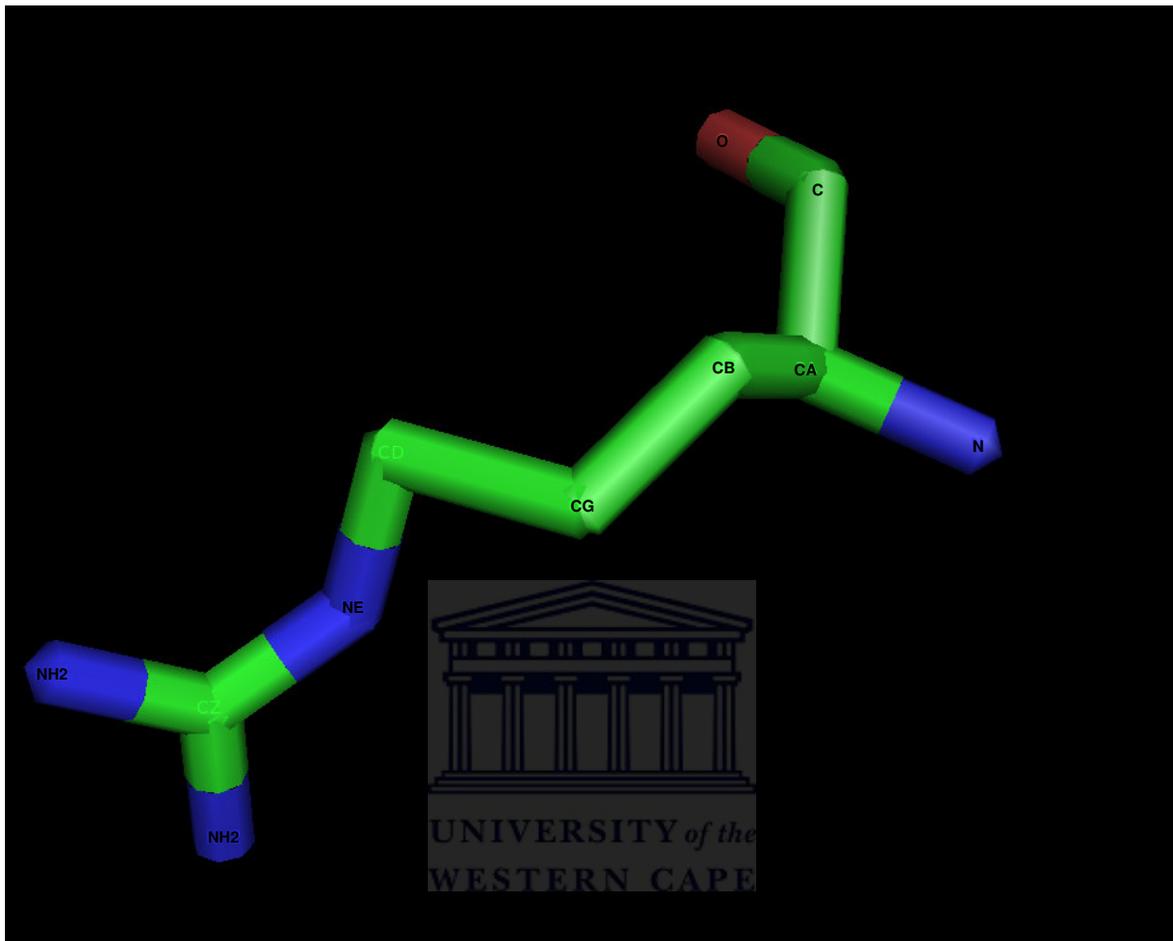
.....210.....220.....230.....240.....250.....260.....270.....280.....290.....300
P14K      GGGGTGGGAGACCGCGAAACAGAAAGGATAATTCGACTGGAACTACTATGAGCATTGGCTGGCCGCTTGGAAACGACTAGTAGTGGAAACAGGAATGTTA
14_6h_T7term GGGGTGGGAGACCGCGAAACAGAAAGGATAATTCGACTGGAACTACTATGAGCATTGGCTGGCCGCTTGGAAACGACTAGTAGTGGAAACAGGAATGTTA
12_7d_T7Term GGGGTGGGAGACCGCGAAACAGAAAGGATAATTCGACTGGAACTACTATGAGCATTGGCTGGCCGCTTGGAAACGACTAGTAGTGGAAACAGGAATGTTA
58_7g_T7Term GGGGTGGGAGACCGCGAAACAGAAAGGATAATTCGACTGGAACTACTATGAGCATTGGCTGGCCGCTTGGAAACGACTAGTAGTGGAAACAGGAATGTTA
14_8c_T7Term GGGGTGGGAGACCGCGAAACAGAAAGGATAATTCGACTGGAACTACTATGAGCATTGGCTGGCCGCTTGGAAACGACTAGTAGTGGAAACAGGAATGTTA
21_Lib0_1C_9C_T7Term GGGGTGGGAGACCGCGAAACAGAAAGGATAATTCGACTGGAACTACTATGAGCATTGGCTGGCCGCTTGGAAACGACTAGTAGTGGAAACAGGAATGTTA
59_9e_T7Term GGGGTGGGAGACCGCGAAACAGAAAGGATAATTCGACTGGAACTACTATGAGCATTGGCTGGCCGCTTGGAAACGACTAGTAGTGGAAACAGGAATGTTA

.....310.....320.....330.....340.....350.....360.....370
P14K      AATAAGCGTGATGTCGACACACGTACAAACGAATTTCTTACTGGAAAAGGAGATGAGGTTTTTTTATTAAT
14_6h_T7term AATAAGCGTGATGTCGACACACGTACAAACGAATTTCTTACTGGAAAAGGAGATGAGGTTTTTTTATTAAT
12_7d_T7Term AATAAGCGTGATGTCGACACACGTACAAACGAATTTCTTACTGGAAAAGGAGATGAGGTTTTTTTATTAAT
58_7g_T7Term AATAAGCGTGATGTCGACACACGTACAAACGAATTTCTTACTGGAAAAGGAGATGAGGTTTTTTTATTAAT
14_8c_T7Term AATAAGCGTGATGTCGACACACGTACAAACGAATTTCTTACTGGAAAAGGAGATGAGGTTTTTTTATTAAT
21_Lib0_1C_9C_T7Term AATAAGCGTGATGTCGACACACGTACAAACGAATTTCTTACTGGAAAAGGAGATGAGGTTTTTTTATTAAT
59_9e_T7Term AATAAGCGTGATGTCGACACACGTACAAACGAATTTCTTACTGGAAAAGGAGATGAGGTTTTTTTATTAAT

```



Notation used for description of amino acids



Arginine (R)