Structure, enzymology and genetic engineering of *Bacillus* sp. RAPc8 Nitrile Hydratase.

Tsepo L. Tsekoa

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Department of Biotechnology, University of the Western Cape.

Supervisor: Prof D.A. Cowan

November 2005

Abstract

Structure, enzymology and genetic engineering of *Bacillus* sp. RAPc8 Nitrile Hydratase.

Tsepo L. Tsekoa

Ph.D thesis, Department of Biotechnology, University of the Western Cape.

Microbial nitrile hydratases (NHases) are important industrial enzymes that catalyse the conversion of nitriles to the corresponding amides. A thermostable, cobalt-type *Bacillus* sp. RAPc8 NHase was previously cloned and expressed in *E. coli*. In this study, the primary aim was to determine the molecular structure of *Bacillus* sp. RAPc8 NHase.

The heterotetrameric enzyme was purified to near homogeneity using heat-purification, hydrophobic interaction chromatography and ion exchange chromatography. Purified NHase was crystallised using the hanging-drop vapour-diffusion method. Crystals produced in the presence of 30% PEG 400, 0.1M MES pH 6.5 and 0.1M magnesium chloride were selected for X-ray diffraction studies. These crystals diffracted well, with diffraction spots visible beyond 2.4Å, with little mosaicity. At 2.5Å, the data were 93% complete.

The crystal structure of *Bacillus* sp. RAPc8 NHase was solved *via* molecular replacement using the crystal structure of *Pseudonocardia thermophila* NHase as a search model. The final refined structure had good refinement statistics and geometry. The overall fold was very similar to that of previously determined NHase structures. *Bacillus* sp. RAPc8 NHase was most similar to *Bacillus smithii* NHase (0.355Å

r.m.s.d.) and least similar to *Rhodococcus* sp. R312 NHase (1.191Å r.m.s.d.). One cobalt atom per heterodimer was bound to a typical NHase metal-binding motif, with post-translationally modified cysteine residues among the ligands to the metal.

The substrate-binding and catalytic cavity of *Bacillus* sp. RAPc8 NHase was identified and described in detail. Surface representation of the structure revealed an extended, curved solvent accessible channel with access to bulk solvent from two locations in the heterodimer. The amino-acid residues forming the channel were identified and the geometric dimensions measured. Enzyme inhibition kinetics indicated that benzonitrile was a potent uncompetitive inhibitor of NHase. This information was used to aid the genetic engineering of aromatic substrate specificity into *Bacillus* sp. RAPc8 NHase. Site-directed mutants of NHase were prepared using the Quickchange mutagenesis procedure. Mutant βW76G showed a two to three fold decrease in benzonitrile inhibition compared with the wild-type. Analysis of the substrate channel of this mutant NHase showed an 11% increase in volume and a 20% increase in inner surface area compared to that of the wild-type NHase. Due to the lack of other significant differences between the two structures (an r.m.s.d. of only 0.101Å was observed), this difference was thought to be responsible for the decrease in benzonitrile inhibition. A structure-modelling based approach for assessing the likely structural differences that may result as a result of a specific mutation was suggested and tested. This approach may be of value in future mutagenesis work.

November 2005

Declaration

I declare that *Structure, enzymology and genetic engineering of Bacillus sp. RAPc8 Nitrile Hydratase* 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.

Tsepo L. Tsekoa November 2005

Signed: . . . . . .

Acknowledgements

First, I would like to thank Professor Don Cowan. I am very grateful and fortunate to have had the opportunity to work under his supervision and in his laboratory. I would also like to thank my co-supervisor, Dr Muhammed Sayed. I thank them both for gracefully teaching me. The contributions of the following other mentors must be gratefully acknowledged: Prof Trevor Sewell, Prof Mike Danson, Prof Sir Tom Blundell and Dr David Hough. The input of all of them has had an immense influence on me and on the successful completion of this thesis. It is humbling to be in awe of one's teachers.

I would also like to thank my friends, mentors and colleagues at the Advanced Research Centre for Applied Microbiology and the Department of Biotechnology. Heide, Rory, Koni, Majozi, Lisa, Manaks, Nazneen, Ntsiki, Nceby, Nthabi, Eloise, Eve, Jen, Antoinette, William, Ben, Gill, Jolanda, Estela, Lemese and Sam all contributed to making the time spent in the lab a lighter load. The contributions of Oz, Mikezozo and Bull should also be acknowledged. Heita daar!

I am most grateful to my family: 'Me, Ntate, Mongoli, Tumi, Marets'epile, Nthedi and Linda. Their continued love, support, motivation and unfailing patience is a real strength. Last (and certainly not least!), I would like to thank Mphos, whose amazing love and support was central to my time during the completion of this thesis.

This work was supported by grants from the National Research Foundation (SA) and the Royal Society of London (UK). I am grateful for this funding.

Table of Contents

List of Tables and Figures

Abbreviations

Å	angstrom
BSA	bovine serum albumin
Da	Dalton
g	gram or gravity
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k	kilo
kDa	kilodalton
K_M	Michaelis-Menten constant
L	litre
m	milli
M	molar
min	minute
NHase	Nitrile Hydratase
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
r.m.s.d	root mean square deviation
SDS	sodium dodecyl sulphate
sp.	specie
v/v	volume per volume
V_{max}	maximum velocity (rate of enzyme-catalysed reaction at infinite substrate concentration)
w/v	weight per volume
μ	micro

Chapter 1: Introduction

1.1 Introduction

Microbial nitrile metabolising enzymes have enjoyed significant interest from academic as well as industrial researchers. These enzymes act to degrade nitriles to amides or carboxylic acids *via* two distinct routes. In one pathway, nitrilase (E.C.3.5.5.1) catalyses direct hydrolysis of the nitrile to the corresponding carboxylic acid. In the second pathway, nitrile hydratase (NHase; E.C.4.2.1.84) catalyses hydration of the nitrile to the corresponding amide. The amide is then converted to a carboxylic acid by an amidase (E.C.3.5.1.4).

A major contributor to the widespread research interest in this group of enzymes is their already successful application as biocatalysts in large-scale commercial biotransformation of nitriles to higher value amides and carboxylic acids. The products of these biotransformations are important in the commodity chemical and pharmaceutical industries (see section 1.5).

A significant milestone in the chemical biotransformation field was the successful industrial-scale (6000 tons per year) conversion of acrylonitrile to acrylamide by the Nitto Chemical Company (presently part of Mitsubishi Rayon Company; Yokohama, Japan). This was achieved using a NHase from *Pseudomonas chlororaphis* B23 (Nagasawa and Yamada, 1989). This process has now been further improved to produce more than 30000 tons of acrylamide per year using *Rhodococcus rhodochrous* J1 NHase (Kobayashi *et al*., 1992).

A further motivation for work on these enzymes is their potential application in environmental bioremediation targeted at generally non-biodegradable nitrile industrial waste.

The benefits of biocatalysis over conventional methods are many and have been reviewed at length elsewhere (Zaks, 2001; Rasor and Voss, 2001). Such processes may be more environmentally friendly and allow syntheses with high yield and chemo-, regio- and/or stereo-selectivity under milder reaction conditions.

Until recently most attention has been directed at mesophilic nitrile metabolising organisms. Use of mesophile-derived enzymes in industrial biotransformation has limitations due to the lower stability of these proteins. In consequence, recent research has focussed on the identification, characterisation and exploitation of new nitrile-metabolising enzymes from thermophilic microorganisms. This chapter aims to review recent literature in this field with particular emphasis on the NHase enzymes.

1.2 Occurrence, distribution and isolation of nitrile metabolising enzymes

1.2.1 Biological significance of nitrile metabolism

The biological significance of nitrile metabolism in microorganisms has only recently become apparent. Historically NHases were presumed to have roles in nitrogen recovery and/or detoxification. However, it has been recently shown that many microbial strains metabolise aldoximes *via* the enzyme aldoxime dehydratase (Kato *et al*., 2000). Aldoximes are known as intermediates in the synthesis of natural products,

particularly in plants. Aldoxime dehydratase genes have also been found within the same gene cluster or operon as nitrile metabolising enzymes (NHase, nitrilase and amidase) (Kato *et al*., 2004; Kato *et al*., 1998; Xie *et al*., 2003). The role of this pathway in shunting aldoximes into central energy metabolism has been proposed. The existence of acyl-CoA synthetase within the same gene cluster in *Pseudomonas chlororaphis* B23 has substantiated this hypothesis (Hashimoto *et al*., 2005). The complete nitrile pathway (aldoxime → nitrile → amide → carboxylic acid → acyl CoA) may therefore be responsible for allowing microorganisms to use aldoximes as a carbon source.

1.2.2 Occurrence and distribution of NHases

NHases have been identified in a number of bacterial species from different genera (e.g., *Agrobacterium, Arthrobacter, Corynebacterium, Pseudomonas* and *Rhodococcus*) isolated from diverse ecosystems (e.g., thermal lake sediments, deep sea trenches and nitrile contaminated soils). The majority of these organisms are obligate mesophiles.

To date, only five thermostable NHases have been reported. The first was from an actinomycete, *Pseudonocardia thermophila* JCM 3095 (Yamaki *et al*., 1997). Subsequently, all other thermostable NHases have been from moderately thermophilic *Bacillus* species; *Bacillus smithii* SC-J05-1 (Takashima *et al*., 1998), *Bacillus* sp. RAPc8 (Pereira *et al*., 1998), *Bacillus pallidus* DAC 521 (Cramp and Cowan, 1999) and *Bacillus* sp. BR449 (Padmakumar and Oriel, 1999). Nitrile degrading activity has also been reported in a limited number of fungal genera (Harper, 1977; Cowan *et al*.,

1998) but exhaustive searches have not identified NHases in any member of the archaeal kingdom.

1.2.3 Isolation of NHases

Isolation of thermophilic nitrile metabolising enzymes has typically relied on "classical" selection or enrichment isolation strategies. Acrylonitrile or benzonitrile are typically used as single carbon/nitrogen sources and/or as substrates for primary or secondary screening procedures.

The importance of knowledge of factors that affect the production and activity of NHases in known microorganisms before employing these isolation strategies cannot be over-emphasised. For example, it was found that for *Bacillus smithii* SC-J05-1, control of the concentration of ammonium and glycerol in the medium was essential for optimisation of enzyme production (Takashima *et al*., 2000). Further, for NHases to be fully active, all required co-factors must be provided in the growth media or in the secondary-screening assay reaction. A coherent strategy must thus be employed to maximise the likelihood of finding a novel isolate.

Although these classical isolation approaches have been successful, they are not without limitations. Much of the bioresource available to provide new enzymes is not readily culturable. Culture-dependent isolation therefore captures only a small proportion of the true genomic diversity. Recently, studies aimed at isolation of new NHases have used metagenomic approaches as an alternative (see Cowan *et al*., 2005 for a review). Several new NHases have recently been isolated by PCR amplification of target genes from community DNA or libraries prepared from the metagenome

(Precigou *et al*., 2001; Lourenco *et al*., 2004 and Liebeton and Eck, 2004). Some of these NHases have been expressed successfully in an active form (Liebeton and Eck, 2004). However, it was observed that although novel, these new NHases were closely related to known NHase genes. This illustrates that PCR-based metagenomic approaches are not truly culture-independent, as the oligonucleotide primers used for amplification must be designed from conserved regions of already known genes.

1.3 Functional expression of NHase in E. coli

Heterologous expression of microbial NHases has been widely demonstrated in *E. coli*. The ability to clone NHases for expression in *E. coli* offers the opportunity to obtain large amounts of enzyme for characterisation. Expression in *E. coli* also allows the use of modern molecular biology techniques such as directed evolution and site directed mutagenesis to improve enzyme characteristics that may be important for biotechnology applications.

Over-expression of functional proteins in *E. coli* is well established and the various standard expression and expression optimisation approaches have been reviewed previously (e.g., Baneyx, 1999). Successful expression of NHase genes in *E. coli* has typically involved the use of high copy number plasmids under the control of an IPTG inducible promoter. Nishiyama *et al*. (1991) showed that the endogenous promoter of the NHase operon from *Pseudomonas chlororaphis* B23 is not functional in *E. coli*.

Early attempts to express cloned NHase in *E. coli* were of limited success due to formation of inactive and/or insoluble enzyme in the form of inclusion bodies (Kobayashi *et al*., 1991). Growth of cultures and induction of expression at 30-32˚C

instead of the optimal growth temperature of 37˚C had some success in reducing formation of inclusion bodies. It was also shown that the presence of the appropriate metal ion co-factor in the media was important for the generation of active recombinant NHase.

Co-expression of NHase genes with various activator proteins has been shown to improve the specific activity of the recombinant enzyme (Wu *et al*., 1997; Hashimoto *et al.*, 1994; Cameron *et al*., 2005). The genes for these activator proteins are generally found just downstream of genes encoding the NHase subunits within the same gene cluster, and vary in size and sequence. In *Pseudomonas putida* 5B the downstream open reading frame that is required for functional expression of the NHase gene encodes a protein of 14KDa (127 amino acids) and is termed P14K. This protein showed no significant homology to any sequences in protein databases (Wu *et al.*, 1997). More recently, Cameron *et al*. (2005) found that co-expression of NHase from *Bacillus* sp. RAPc8 with a homologous protein was essential for complete activity. This is in contrast to the requirements of *Bacillus* sp. BR449, where despite the presence of a gene encoding a protein of a similar size (12KDa), no accessory protein was required for expression of thermostable NHase in *E. coli* (Kim and Oriel, 2000). Since the NHase operon of these organisms is very similar, this finding was surprising and may possibly be due to mis-identification of a truncated downstream protein due to sequencing errors (Cameron R.A., personal communication).

The definitive role of these small activator proteins is still unclear. It has been suggested that they may be involved in post-translational modification of the active site. Other activator proteins that have been found to promote activity of NHase

activity in *E. coli* are implicated in incorporation of the metal ion co-factor into the enzyme (Endo, 2001) or into the host cell. *E. coli* lacks mechanisms for transport of cobalt ions into the cell. A new cobalt transporter in NHase-producing *Rhodococcus rhodochrous* J1 has been characterized (Komeda *et al.*, 1997). The gene for this 37KDa membrane protein is located within the NHase operon just downstream of the gene encoding the α subunit. Co-expression of this protein with NHase in *E. coli* and *R. rhodochrous* ATCC 12674 was conclusively shown to result in more active NHase being expressed. A putative chaperone-like role for P14K in the assembly of NHase has also been suggested (Cameron *et al.*, 2005). Further work aimed at confirming this hypothesis is presently in progress (Cameron R.A., unpublished results) but it is noteworthy that co-expression of *Comamonas testosteronii* NI1 NHase with the GroEL/GroES system from *E. coli* resulted in similar improvements in activity (Stevens *et al*., 2003).

1.4 Enzymology of NHases

1.4.1 Molecular characteristics

NHases are soluble enzymes and are classified into two groups based on their co-factor requirements. The iron-type (Fe-type) NHases (Endo *et al*., 2001; Huang *et al*., 1997) have a non-heme iron atom while the cobalt-type (Co-type) NHases (Cramp and Cowan, 1999); (Kim *et al*., 2001) have a non-corrinoid cobalt atom at the catalytic centre.

The minimal functional unit of most NHases is composed of two subunits; α and β, and contains one metal ion per αβ unit. However, at least two NHases so far

characterised have a homodimeric functional unit denoted 2α (*Agrobacterium tumefaciens* (Trott *et al*., 2001) and *Corynebacterium* sp. C5 (Yamamoto *et al*., 1992)) and one (*Rhodococcus equi* (Gilligan *et al*., 1993)) has a monomeric functional unit. The sizes of the various α and β subunits range from 23kDa to 30kDa.

The native oligomeric state of NHases varies widely between species. The native molecular weight ranges between 30kDa in *Rhodococcus equi* SHB-121 (Gilligan *et al*., 1993) and 500-530kDa in *Rhodococcus rhodochrous* J1 (Kobayashi *et al*., 1991). This has resulted in the classification of NHases as light NHase (L-NHase) and heavy NHase (H-NHase), in particular the enzymes from *Rhodococcus rhodochrous* J1. Table 1.1 summarises the molecular characteristics of bacterial NHases.

Table 1.1 Molecular characteristics of mesophilic and thermophilic NHases.

Organism	Native molecular weight (kDa)	Subunit composition	Size of α subunit (kDa)	Size of β subunit (kDa)	Metal Cofactor	Reference
Mesophilic						
Rhodococcus rhodochrous J1 (L-NHase)	100	2α2β	26	29	Co	
Rhodococcus rhodochrous J1 (H-NHase)	500-530	10α10β	26	29	Co	Kobayashi *et al*., 1991
Rhodococcus sp. N774	70	2α2β	28.5	29	Fe	Endo and Watanabe, 1989
Rhodococcus sp. N771	70	2α2β	27	27	Fe	Ikehata *et al*., 1989; Yamada and Kobayashi, 1996
Rhodococcus sp. R312	95				Fe	Nagasawa *et al*., 1986
Corynebacterium sp. C5	61	2α	27		Fe	Yamamoto *et al*., 1992
Arthrobacter sp. J-1	420		25	27		Asano *et al*., 1982
Thermophilic						
Pseudonocardia thermophila JCM 3095		2α2β	29	32	Co	Yamaki *et al*., 1997; Miyanaga *et al*., 2001
Bacillus sp. BR449			25	27	Co	Padmakumar and Oriel, 1999
Bacillus pallidus DAC 521	110	2α2β	27	29	Co	Cramp and Cowan, 1999
Bacillus smithii SC-J05-1	130	2α2β	26	29	Co	Takashima *et al.*, 1998
Bacillus sp. RAPc8	110	2α2β	28	29	Co	Pereira *et al*., 1998, Cameron *et al*., 2005

1.4.2 Structure and mechanism

Several NHase crystal structures have been solved (*Rhodococcus* sp. R312, PDB accession number 1AHJ/2AHJ, Huang *et al*., 1997; *Rhodococcus* sp. N-771, Nagashima *et al*., 1998; *Pseudomonas thermophila*, PDB accession number 1IRE, Miyanaga *et al*., 2001; *Bacillus smithii* SC-J05, PDB accession number 1V29, Hourai *et al*., 2003). The overall fold of all these enzymes is conserved.

The first detailed structure of a NHase, at 2.65Å resolution was reported in 1997 (Huang *et al*., 1997). This NHase belongs to the Fe-type subgroup. The α subunit of this enzyme consists of a long extended N-terminal arm and a globular C-terminal domain. The structure of this C-terminal globular domain is unusual and was described as a four layered α−β−β−α structure. The β subunit starts with a long loop of 30 residues that wraps around the α subunit in the αβ dimer. Following this is a five-helical domain that is followed by a β roll that contains six anti-parallel strands.

The α and β subunits of the *Rhodococcus* sp. R312 NHase form a tight dimer with a large interface area of 3800 $Å^2$. The iron centre is located in a large open cavity at this interface and all protein ligands to the metal ion were provided by the α subunit. The side-chains of cys110α, cys113α, cys115α and the main chain amide nitrogen atoms of residues ser114α and cys115α form the ligand sphere. The ligand sphere observed in the crystal structure was thus S_3N_2X with the five identified ligands occupying five vertices of an octahedron. The sixth position (denoted X) was proposed to be a hydroxide or a water molecule in the active form of the enzyme (Huang *et al*., 1997) and nitric oxide in the inactive form of the enzyme (Noguchi *et*

al., 1996). Guanidium groups from the conserved residues arg141β and arg56 form hydrogen bonds with sulphur atoms from cys115α and cys113α thus holding them in place. A crystal structure of another Fe-type NHase from *Rhodococcus* sp. N-771 (Nagashima *et al*., 1998) indicated that cys114α and cys112α (corresponding to cys115α and cys113α of *Rhodococcus* sp. R312) were post-translationally modified to cysteine-sulfenic acid and cysteine-sulfinic acid, respectively. In this enzyme, arg141β and arg56β form hydrogen bonds with oxygen atoms of these modified cysteine residues.

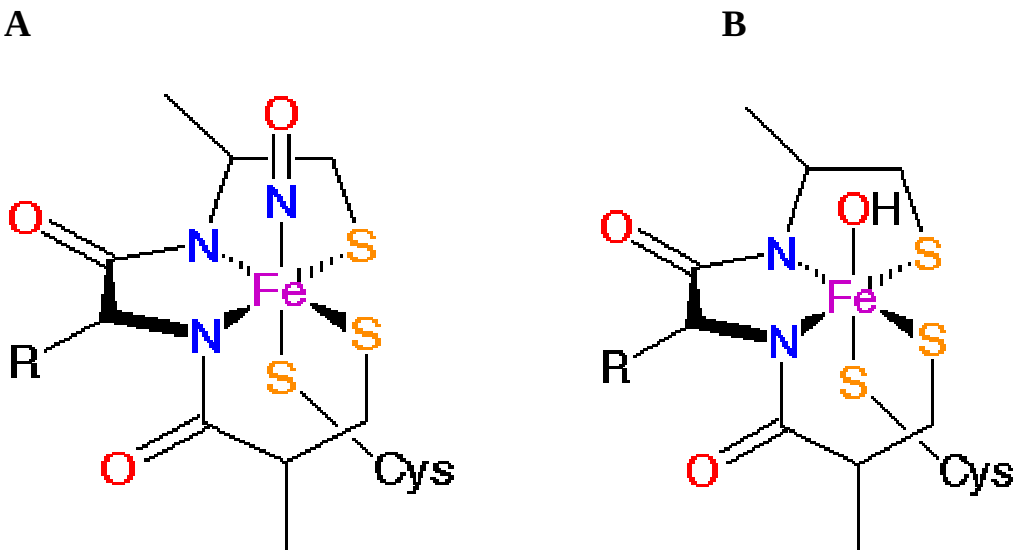


Figure 1.1 Metal-centre of Fe-type NHase showing claw-setting motif with nitric oxide in the inactive state (A) and hydroxide in the active (B) state.

It has been noted that the Val-Cys-Ser(Thr)-Leu-Cys-Ser-Cys sequence of the α subunit forming the metal binding motif (referred to as a 'claw-setting motif') is highly conserved among all known NHases (Endo *et al*., 2001; Huang *et al*., 1997). The Co-type NHases have a threonine in place of the serine found in Fe-type NHase.

This substitution is thought to be responsible for differing preferences for metal co-factor.

The open cavity occupied by the metal ion is accessible to the bulk solvent and has been proposed as the probable active site pocket of the enzyme. This view is supported by successful modelling of a substrate analogue (iodoacetonitrile) into the active site pocket of the crystal structure of *Rhodococcus* sp. R312 (Huang *et al*., 1997). Three different catalytic mechanisms have been proposed from this model (Figure 1.2). All mechanisms involve the metal ion acting as a Lewis acid to activate the nitrile for hydration. The mechanism that is thought most likely, due to the non-involvement of ligand exchange, involves indirect activation of the nitrile. Ligand exchange for trivalent metal ions is typically slow, making such mechanism less favoured. It has been demonstrated (Mascharak, 2002) that Co-type NHase from *R. rhodochrous* sp. J1 and Fe-type NHase from *Rhodococcus* sp. R312 hydrolyse propionitrile at almost identical rates under the same reaction conditions. Since ligand exchange kinetic characteristics of low-spin iron (III) and cobalt (III) are distinctly different, a similar reactivity of the two NHases suggests that the reaction mechanism of these enzymes does not involve ligand exchange. In the mechanism favoured by Huang *et al*., (1997), termed an outer sphere mechanism (Figure 1.2c), a hydroxide ion acts as a general base to activate a water molecule at the active site. This water molecule, in turn, attacks the nitrile. The schematic representation of these mechanisms is outlined below (Figure 1.2).

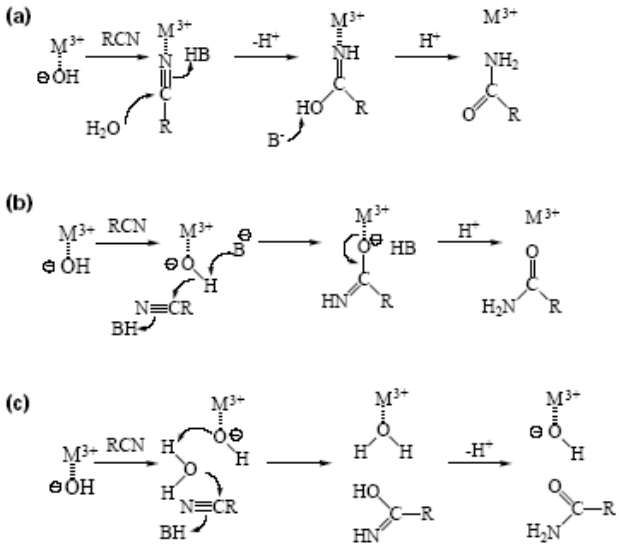


Figure 1.2 Proposed mechanisms for NHase catalysis (from Huang *et al*., 1997)

More recently, Miyanaga *et al*. (2001) reported the crystal structure of a Co-type NHase from the thermophilic *Pseudonocardia thermophila* JCM 3095. This structure (at a resolution of 1.8Å) exhibited high structural similarity with the mesophilic Fe-type NHase. However, a significant difference between the two enzymes was the presence of an additional interaction between the α and β subunits of the *P. thermophila* enzyme, provided by two additional α-helices. It has been suggested that this additional interaction may contribute to the thermostability of this enzyme.

The active centre/metal-binding site of the Co-type NHase has the same claw-setting motif as the Fe-type enzymes. The ligands to the cobalt ion are formed by sulphur atoms of residues cys108α, cys111α and cys113α, and the main chain amide nitrogen atoms of ser112α and cys113α. Electron density peaks indicating that cys111α and cys113α are posttranslationally oxidised to cysteine-sulfinic acid and cysteine-sulfenic acid respectively, have been identified (Miyanaga *et al*., 2001). As in the *Rhodococcus* sp. N-771 NHase (Nagashima *et al*., 1998) oxygen atoms from these

modified cysteine residues form hydrogen bonds with conserved arginine residues from the β subunit (arg52β and arg157β), thus stabilising the claw-setting motif.

1.4.3 Photoactivation of Fe-type NHase

Fe-type NHases from *Rhodococcus* sp. N771, *Rhodococcus* sp. N774, *Rhodococcus* sp. R312, *R. erythropolis* and *C. testosteroni* have been shown to possess a unique reactivity to light. The mechanism of photoreactivation in *Rhodococcus* sp. N771 has been recently and comprehensively reviewed (Endo *et al*., 2001).

Early physicochemical work implicated the iron centre in photoactivation (Tsujimura *et al*., 1997). The most convincing data to resolve the photoreactivation mechanism was provided by measuring difference Fourier-transform infra-red (FT-IR) spectra before and after photoactivation (Noguchi *et al.,* 1995). The spectra conclusively indicated the association of an endogenous NO molecule in the dark and the dissociation of the molecule upon exposure to light. Figure 1.3 shows a schematic representation of photoreactivation and light inactivation. Dissociation of NO upon light irradiation results in a conformational change at the active centre, facilitating substrate binding. The physiological significance of the photoactivation process is not known.

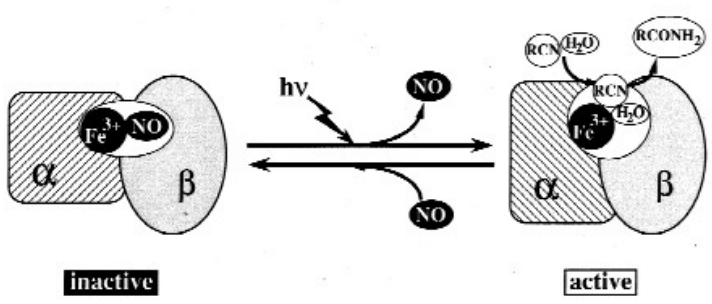


Figure 1.3 Schematic representation of photoactivation of Fe-type NHase (from Endo *et al*., 2001)

No reports of photoactivation of cobalt-containing NHases exist. Since all the thermostable NHases so far identified contain cobalt as the metal co-factor, photoactivation is currently restricted to mesophilic NHases.

1.4.4 Functional characteristics: Temperature and pH optima

There is wide variation in temperature optimum and thermostability among NHases from different microorganisms. The optimal temperature for activity of the purified enzymes is typically consistent with the optimal growth temperature of the source organism. The Co-type NHases generally exhibit higher thermostability than iron-type NHases (Payne *et al*., 1997), not surprisingly, since all the thermostable NHases so far isolated have cobalt as their metal co-factor. Recent publication of a detailed structure of a thermostable Co-type NHase from *Pseudonocardia thermophila* (Miyanaga *et al*., 2001) suggested that an additional α−β subunit-subunit interaction might contribute to the increased thermostability of this protein. Absence of detailed structures of other thermophile-derived and mesophile-derived NHases makes it impossible to determine whether the relatively high thermostability of mesophilic Co-type *Rhodococcus rhodochrous* J1 (H-NHase) NHase and other thermostable NHases is a result of a similar structural deviation from the structure of Fe-type NHase. It has been shown that thermal stability of crude thermophilic *Bacillus pallidus* NHase is substantially higher than that of the purified enzyme (Cowan *et al*. 1998). This indicated that the intracellular environment of this thermophile provided a significant level of extrinsic stabilisation.

Table 1.2 Thermal stability values of purified and crude *Bacillus pallidus* NHase (from 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

Optimal pHs of NHases so far characterised range between pH 6 (lower limit of optimal range for *Pseudomonas chlororaphis* NHase (Nagasawa *et al*., 1987)) and pH 10.5 (optimal pH for *Bacillus smithii* NHase (Takashima *et al*., 1998)). Known NHases are thus active between mildly acidic/neutral to very alkaline pH. Table 1.3 summarises the pH and temperature optima of selected mesophile-derived and thermophile-derived NHases.

Table 1.3 Temperature and pH optima of NHases.

Organism	Optimal Temperature (˚C)	pH optimum	Reference
Thermophilic			
Pseudonocardia thermophila JCM 3095	60		Yamaki *et al*. (1997)
Bacillus sp. RAPc8	60	7.0	Pereira *et al.*, 1998
Bacillus pallidus DAC 521	50	7.0-7.5	Cramp and Cowan, 1999
B. smithii SC-J05-1	40	10.5	Takashima *et al.*, 1998
Bacillus sp. BR449	55	7.5	Padmakumar and Oriel, 1999
Mesophilic			
Rhodococcus rhodochrous J1 (H-NHase)	35-40	6.0-8.5	Kobayashi *et al*., 1991; Komeda *et al*., 1996; Nagasawa *et al*., 1991
Rhodococcus sp. R312	35	7-8.5	Nagasawa *et al*., 1986
Agrobacterium tumefaciens	25	6.5-9.5	Bauer *et al*., 1994
Pseudomonas chlororaphis	20	6-7.5	Nagasawa *et al*., 1987

1.4.5 Substrate specificity

It has been often stated that "nitrilases are aromatic substrate specific and enantioselective" while "NHases are aliphatic substrate specific and lack enantioselectivity". With the identification and characterization of more NHases, some of which demonstrate aromatic substrate specificity and/or enantioselectivity, this view is no longer tenable. An overview of substrate specificity of mesophilic and thermophilic NHases is given in Table 1.4.

The substrate specificity of *Rhodococcus rhodochrous* AJ270 has been studied in detail (Meth-Cohn and Wang, 1995). While purified NHase was not used in this study, the authors were able to demonstrate that the predominant nitrile-metabolising system in this organism was the NHase/amidase system. The organism was able to hydrolyse a broad spectrum of saturated and unsaturated aliphatic nitriles, aromatic nitriles and heterocyclic nitriles. The study also showed that the enzyme was sensitive to substrate molecular geometry, particularly the presence of ortho-substituents in aromatic nitriles, and that a molecular diameter of no more than ~7 Å was the upper limit for effective nitrile hydrolysis.

Of the five known thermophilic NHases, the substrate specificities of three have been studied in detail; *Bacillus pallidus* DAC 521 (Cowan *et al*., 1998), *Bacillus* sp. RAPc8 (Pereira *et al*., 1998; Cameron, 2002) and *Bacillus smithii* SC-J05-1 (Takashima *et al*., 1998). No substrate specificity data are available for *Bacillus* sp. BR449 NHase although it has been shown that this enzyme was active against the commercially important substrate acrylonitrile (Padmakumar and Oriel, 1999). The NHase from

Pseudonocardia thermophila appeared to have preference for aliphatic nitriles (Miyanaga *et al*., 2004), although only four nitrile substrates were tested.

Table 1.4 Substrate specificities of NHases

Organism	Metal Cofactor	Substrate Specificity	Reference
Mesophilic			
Rhodococcus. erythropolis BL1		Wide: aliphatic, hetrocyclic and aromatic	Duran *et al*., 1993
Rhodococcus rhodochrous J1	Co	Wide: aliphatic saturated or unsaturated and aromatic nitriles	Nagasawa *et al*., 1991; Kobayashi *et al*., 1991
Corynebacterium sp. C5	Fe	Narrow: aliphatic	Yamamoto *et al*., 1992
Rhodococcus sp. R312	Fe	Wide: aliphatic, cyclic	Nagasawa *et al*., 1986
Rhodococcus sp. N774	Fe	Narrow: aliphatic	Yamada and Kobayashi, 1996
Thermophilic			
Bacillus pallidus DAC 521	Co	Narrow: aliphatic	Cowan *et al*., 1998
Bacillus sp. RAPc8	Co	Very wide: cyclic linear, dinitriles and branched aliphatics	Pereira *et al*., 1998
Bacillus smithii SC-J05-1	Co	Wide: linear aliphatics, branched aliphatics and dinitriles	Takashima *et al*., 1998

Detailed comparative specificities of selected thermophilic NHases are given in Table 1.5. While the range of nitriles studied for each of the NHases differs and different assay procedures may have been used, these data imply that *Bacillus* sp. RAPc8 NHase has the widest substrate specificity while *Bacillus smithii* and *Bacillus pallidus* NHases have relatively limited substrate specificities.

Table 1.5 Comparative specificities of thermophilic NHases. Values are relative activities, normalized to the maximum value. Data from: [1]: Pereira *et al*. (1998), Cameron (2002); [2]: Cramp (1998); [3]: Takashima *et al*. (1998).

Substrate	***Bacillus sp.* RAPc8[1]**	***Bacillus pallidus* DAC 521[2]**	***Bacillus smithii* SC-J05-1[3]**
Acetonitrile	93	60	100
Propionitrile	30	40	18
Butyronitrile	64	55	54
Valeronitrile	100	49	44
Acrylonitrile	62	79	72
Crotonitrile	38	100	4
Benzonitrile	0	0	0.2
3-Cyanopyridine	53	-	0.7
Cyclopentene-acetonitrile	58	0	-
Malononitrile	6	0	-
Glutaronitrile	30	0	-
Adiponitrile	27	0	1
Chloropropionitrile	-	-	7

1.5 Applications of nitrile-degrading enzymes

1.5.1 Biotransformations

As previously indicated, a major contributor to the widespread research interest in nitrile metabolising enzymes is their already successful application as biocatalysts in the industrial conversion of nitriles to higher value amides and carboxylic acids. Indeed, the use of *Rhodococcus rhodochrous* J1 in production of kiloton quantities of acrylamide per year by Mitsubishi Rayon (Kyoto, Japan) is one of the most impressive examples of the commercial success of industrial biocatalysis (Zaks, 2001). Another example of the industrial application of nitrile metabolising enzymes is the chemoenzymatic manufacture of nicotinamide by Lonza Guangzhou Fine Chemicals (Guangzhou, China) (Thomas *et al*., 2002). While there are only two commercially developed nitrile metabolising enzyme systems, there are numerous recent publications describing various other biotransformations that have potential for future exploitation. A selection of these is reviewed here.

The commercial-scale conversion of adiponitrile (ADN) to 5-cyanovaleramide (5-CVAM) using *Pseudomonas chlororaphis* B23 has been demonstrated and was close to large-scale industrial exploitation (Hann *et al*., 1999). 5-CVAM is an early intermediate in the production of the herbicide azafenidin (DuPont). Normally, 5-CVAM is chemically produced using Raney copper or manganese dioxide catalyst at high temperatures (typically around 130˚C). One of the main limitations of these processes was the co-production of significant quantities of adipamide.

The use of immobilised *P. chlororaphis* B23 introduced the desired regioselectivity, which resulted in minimal production of adipamide. A further limitation of the chemical catalysis process was the production of significant amounts of catalyst waste. In contrast, the *P. chlororaphis* process produced only 0.006 kg of catalyst waste per kg of 5-CVAM produced, with the solids content of the catalyst waste constituting only 7% (Hann *et al*., 1999). This new process is thus more 'environmentally friendly'.

D-phenylglycine amide is an intermediate in the industrial synthesis of β-lactam antibiotics. (Wegman *et al*., 2001) isolated a new *Rhodococcus* sp. (most closely related to the type strain of *R. globerulus*) that was able to carry out the biotransformation of a racemic mixture of phenylglycine nitrile to D-phenylglycine amide and L-phenylglycine. This organism harboured a nitrile metabolising system with a non-stereoselective nitrile hydratase and an extremely L-selective amidase. Organisms capable of aliphatic and aromatic nitrile biotransformations with similar stereochemistry have been reported previously (see the excellent review by Sugai *et al*., 1997 for example). This *Rhodococcus* sp. was novel due to its ability to retain very high activities in the presence of high substrate concentrations. The high activity minimised decomposition of phenylglycine nitrile to benzaldehyde and ammonia (Wegman *et al*., 2001).

2-arylpropanoic acids are an important class of anti-inflammatory pharmaceutical compounds (e.g., naproxen and ibuprofen are commercially important examples). The (*S*)-enantiomer of these agents has been shown to be much more active than the (*R*)-enantiomer. Chemical synthesis pathways that involve the resolution of isomers

by physical means can be costly and are thus commercially undesirable. One of the strategies being developed for the preparation of optically active 2-arylpropanoic acids is the enantioselective enzymatic hydrolysis of the corresponding nitriles. Effenberger and Böhme (1994) demonstrated the enantioselective hydrolysis of racemic (*R/S*) naproxen nitrile to (*S*)-naproxen using *Rhodococcus* sp. C3II cells. This strain constitutively expressed (*S*)-enantiomer specific nitrile hydratase and amidase. (Effenberger and Graef, 1998) showed that *Rhodococcus* sp. C3II and *R. erythropolis* MP50 were able to hydrolyse a variety of other substrates that led to the production of a variety of other 2-arylpropanoic acids. These two strains were found to be complementary with respect to regioselective biotransformation of dinitriles and diamides, where *Rhodococcus* sp. C3II selectively produced mononitrile and monoamide derivatives whereas *R. erythropolys* MP50 preferentially formed mononitrile, monoacid and monoamide derivatives (Effenberger and Graef, 1998). The enzymatic production of (*S*)-ibuprofen has also been demonstrated using *Rhodococcus* sp. AJ270 (Snell and Colby, 1998).

Since its isolation, *Rhodococcus* sp. AJ270 has proved to be a very robust and versatile biocatalyst, with very broad substrate specificity. For example, (Wang and Feng, 2000) reported the biotransformations of racemic 2-arylcyclopropanecarbonitriles leading to formation of enantiopure preparations of various pharmaceutically important cyclopropyl compounds. *Rhodococcus* sp. AJ270 was also successfully used to produce optically active L-arylglycines and D-arylglycine amides (Wang *et al*., 2001).

NHases have also been used in steroid biotransformations (Kaufmann *et al*., 1999). The progestin dienogest (17α-cyanomethyl-17β-hydroxy-estra-4,9-dien-3-one) derived from nortestosterone is a hormonal contraceptive. In order to develop this compound for other indications and to enhance potency, *Rhodococcus erythropolis* was used to transform the cyanomethyl group at position 17α to the amide and carboxylic acid derivative (Kaufmann *et al*., 1999). Within a 2-24h period, the dominant reaction was the aromatisation of ring A. Appearance of the amide derivatives was only found after prolonged fermentation (2-27 days). While this biotransformation was not particularly successful (and the amide derivatives not particularly useful), it demonstrates the potential of the nitrile hydratase/amidase system in producing a wide variety of useful compounds.

(*R*)-2-naphthylmethoxyacetic acid is a valuable chiral NMR reagent in the resolution of the absolute configurations of secondary alcohols. Typically, a racemate of this molecule is synthesized chemically and the enantiomers separated by chromatography. Recently, Kimura *et al*. (2002) reported an integrated chemo-enzymatic synthesis of (*R*)-2-naphthylmethoxyacetic acid. Among the steps employed was the hydration of a cyano group utilizing *Rhodococcus rhodochrous* IFO 15564.

1.6 Techniques in X-ray crystallography of biological macromolecules

X-ray crystallography is now the method of choice for elucidation of the molecular structures of biological macromolecules. Of the 33252 released structures in the Protein Data Bank (PDB), 28250 (85%) have been solved using X-ray crystallography (www.pdb.org). The technologies required to progress from cloning of a gene of interest to calculation of phases and refinement of the final structure are now accessible even to non-specialist crystallography laboratories, including research facilities on the African continent (Tsekoa *et al*., 2004; Kuhnert *et al*., 2005). This section aims to provide a concise review of the steps required for determination of a protein structure once x-ray diffraction data have been collected. More comprehensive reviews are available elsewhere (e.g., Carter and Sweet, 1997; Smyth and Martin, 2000; McRee, 1999). This section is essentially based on such reviews but includes computer programs and methods that have become widely used in the past few years. The reader is also referred to a very comprehensive book (McPherson, 1999) for a review of techniques and considerations for crystallisation of biological macromolecules. These are briefly discussed in section 3.3.

The various steps in a typical protein structure determination project after x-ray diffraction data have been collected are summarised in Figure 1.4. The data processing step can be sub-divided into several parts. First, the space group or crystal system (crystal packing symmetry) to which the crystals belong is determined. This process also involves the accurate determination of the unit cell dimensions of the crystal and the orientation of the crystal in the x-ray beam.

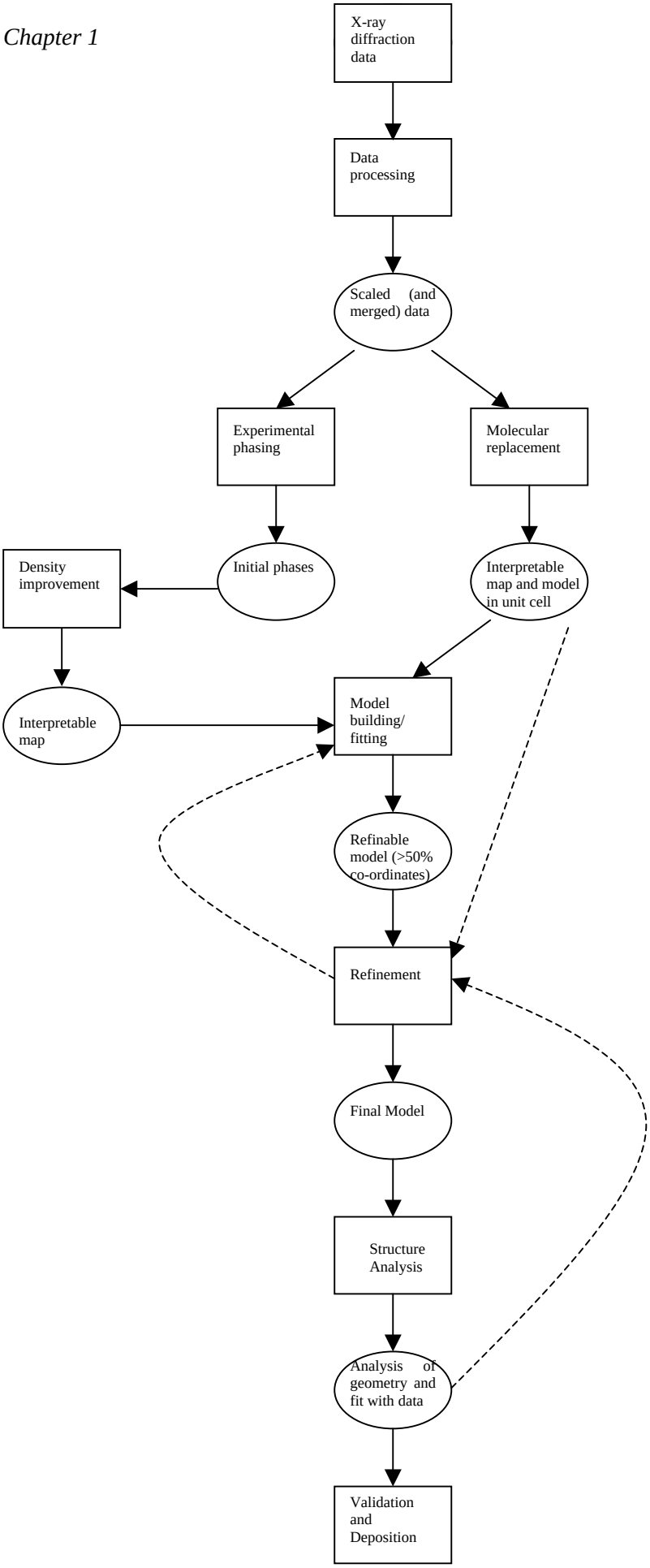


Figure 1.4 Stages of macromolecular structure determination.

Once this has been done, the data can be indexed. Simply put, indexing is the process in which all spots on all diffraction images within the data-set are assigned a Miller index described by three integers; *h, k* and *l.* The intensities of the spots are then measured and scaled to relate all images from the data-set with one another. All these procedures are carried out using computer programs with interactive intervention from the researcher. Programs that are currently in common use for these purposes are Mosflm and SCALA (CCP4, 1994), Crystal Clear (d*TREK) (Pflugrath *et al*., 1999) and HKL2000 (incorporating Denzo and Scalepack) (Minor *et al*., 2002).

The result of data-processing is a file describing a set of indexed intensities. The intensity of a given diffraction spot is a result of the amplitude and the relative phase of the diffracted waves. In the crystallographic experiment, amplitudes can be calculated (using programs such as Truncate from the CCP4 suite (CCP4, 1994)), but phases cannot be calculated directly. This characteristic of the crystallographic experiment has resulted in the notorious 'phase problem' in macromolecular crystallography. Once phases are calculated, structure factors can be calculated with relative ease. These can then be used to calculate a map of the distribution of electrons in the molecule (the electron density map) *via* Fourier transform. Along with successful production of crystals for analysis (and the associated preceding work), calculation of phases is regarded as one of the most significant hurdles in macromolecular structure determination.

The approach taken to calculate phases from crystallographic data depends on whether a structure of a closely related macromolecule is available. Increasingly often, an appropriate homologue is available for use as a search model. For this reason

the technique of molecular replacement (Rossmann, 2001) is now widely used. In molecular replacement, phase information from the search model is essentially "borrowed" for use in the new structure. The structure factors of the known homologue structure are subjected to inverse Fourier transform and the phases extracted. These phases are used as an estimate of the phases of the experimental data. For this process to be useful, the search model must be placed in the unit cell of the unknown structure in the correct position and orientation. The six-dimensional search for this correct position and orientation is carried out in two separate three-dimensional steps; a rotation function and a translation function. The observed amplitudes from the experiment and these newly estimated phases are then subjected to Fourier transform to calculate structure factors of the new structure. Commonly used computer programs for solution of structures by molecular replacement include Amore (Navaza, 2001), Molrep (Vagin and Teplyakov, 1997), EPMR (Kissinger *et al*., 1999) BEAST (Read, 2001) and Phaser (Storoni *et al*., 2004; McCoy *et al*., 2005).

In the absence of a structure of a homologue, experimental or *ab initio* phasing methods must be used. The most commonly used experimental phasing techniques are multiple isomorphous replacement (MIR) (Hengming, 1997) and multiwavelength anomalous dispersion (MAD) (Hendrickson, 1991). While this review will not attempt to cover these in detail, it is important to provide a brief overview.

Experimental phasing techniques require the presence of heavy atoms within the macromolecule. This derivatisation process is different in the two techniques. In MIR, heavy metal atoms are incorporated into the crystal by soaking or co-crystallisation in low-millimolar solutions of heavy-metal reagents. Mercury, platinum or gold are

common choices (Smyth and Martin, 2000). More recently, halides have also been used to derivatise crystals (Dauter and Dauter, 2001). In MAD, selenomethionine is genetically incorporated into the macromolecule in place of the normally occurring methionine residue (Hendrickson *et al*., 1990).

MIR relies on (near-) perfect isomorphism between the diffraction data from the native and derivative crystals. The positions of the heavy atoms can be determined based on the fact that major differences between the two data sets are not as a result of conformational changes or differences in the unit cell dimensions of the crystal but solely as a result of the incorporated heavy atoms. The determination and refinement of the positions of heavy atoms in the macromolecule are the starting point of phase calculation.

MAD became widely used relatively recently due to the increasing availability of appropriate synchrotron instrumentation (Ealick, 2000). Here, diffraction data is collected at several wavelengths near the absorption edge of the selenium atoms incorporated into the crystal. Phase information analogous to that obtained from MIR is then calculated from these data. Electron density maps that result from experimental phasing often require modification and improvement before the map can be interpreted accurately. Programs for experimental phasing include SOLVE/RESOLVE (Terwilliger and Berendzen, 1999) and SHARP/autoSHARP (Global Phasing Ltd).

Once an interpretable electron density map has been produced, the structure can be built in the form of co-ordinates describing a model of atoms and bonds that fit the

electron density map. The manual model-building or fitting process is carried out with the aid of computer programs such as O (Kleywegt and Jones, 1997), XtalView (McRee, 1999), coot (Emsley and Cowtan, 2004) and NOC (Institute of Biophysics, Chinese Academy of Sciences). Structures from high resolution data can even be built automatically by programs such as ARP/wARP (Lamzin and Perrakis, 2000). The structure is then refined to improve the fit between the original diffraction data and structure model.

In practice, the fitting and refinement steps are normally carried out cyclically. At the end of successful model-building and refinement, the result should be a model which sufficiently explains the experimental observations, while making physical, chemical and biological sense (Kleywegt and Jones, 1997). Success is measured primarily by calculation of an R-factor and a free R-factor (R_{free}) (Brunger, 1993). In simple terms, these values are a measure of the difference between calculated structure factors (from the model) and observed structure factors (from the original diffraction data). The R_{free} however, is calculated from a small portion of the data (typically 5 to 10%) that has been excluded from the refinement process (the test set). It is used as a more reliable cross-validation indicator than the R-factor (Kleywegt and Brunger, 1996). Some of the programs that are currently used for refinement are Refmac5 (Murshudov *et al.*, 1997), ARP/wARP (Lamzin and Perrakis, 2000) X-plor/CNX (Accelrys Ltd.).

Once a final refined model is produced, it has to be checked for reliability using other validation tools. PROCHECK (Laskowski *et al.*, 1993) appears to be the most widely cited program for this purpose. This program checks the geometric integrity of the structure and produces detailed Ramachandran plots as an output. Alternative

programs used for this purpose are WhatIf and RAMPAGE (Lovell *et al*., 2001). A fully refined, validated structure model is ready for comprehensive analysis and deposition into the database of 3 dimensional protein structures, the Protein Data Bank (PDB) (Berman *et al*., 2000; Berman *et al*., 2002).

Analysis of a crystal structure can involve any combination of a large number of analytical procedures (e.g., accessible surface area analysis, cavity analysis, analysis of topology, interfaces etc.) that require many programs. These programs, with very limited, specific functions will not be reviewed here. Invariably however, analysis of a structure requires use of molecular graphics programs, which aid the researcher in viewing and displaying the structure in different ways. All the programs designed for model-building (mentioned above) are effective for this purpose. The long list of currently used molecular graphics programs also includes Rasmol, Swiss-PDBviewer (Guex and Peitch, 1997), Cn3D (National Center for Biotechnology Information, USA), Chime (Elsevier MDL), PyMol (Delano Scientific LLC) and MolScript (Kraulis, 1991).

1.7 Aims

The primary aim of this study was to determine the molecular structure of *Bacillus* sp. RAPc8 NHase. This detailed structural information was to be used to aid in the engineering of aromatic substrate specificity into this enzyme. To achieve this aim the following specific objectives were pursued:

1. To purify recombinantly expressed *Bacillus* sp. RAPc8 NHase.

2. To produce diffraction quality crystals of this NHase for X-ray diffraction analysis.

3. To solve the high-resolution crystal structure of *Bacillus* sp. RAPc8 NHase.

4. To apply knowledge gained from the crystal structure to genetically engineer this enzyme to include aromatic nitriles among its substrates.

Reference List

Cramp,R.A. and Cowan,D.A. (1999) Molecular characterisation of a novel thermophilic nitrile hydratase. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1431:** 249-260.

Effenberger,F. and Graef,B.W. (1998) Chemo- and enantioselective hydrolysis of nitriles and acid amides, respectively, with resting cells of Rhodococcus sp. C3II and Rhodococcus erythropolis MP50. *Journal of Biotechnology* **60:** 165-174.

Hann,E.C., Eisenberg,A., Fager,S.K., Perkins,N.E., Gallagher,F.G., Cooper,S.M. et al. (1999) 5-cyanovaleramide production using immobilized Pseudomonas chlororaphis B23. *Bioorganic & Medicinal Chemistry* **7:** 2239-2245.

Hashimoto,Y., Hosaka,H., Oinuma,K., Goda,M., Higashibata,H., and Kobayashi,M. (2005) Nitrile pathway involving Acyl-CoA synthetase - Overall metabolic gene organization and purification and characterization of the enzyme. *Journal of Biological Chemistry* **280:** 8660-8667.

Kato,Y., Ooi,R., and Asano,Y. (2000) Distribution of aldoxime dehydratase in microorganisms. *Applied and Environmental Microbiology* **66:** 2290-2296.

Kato,Y., Ooi,R., and Asano,Y. (1998) Isolation and characterization of a bacterium possessing a novel aldoxime-dehydration activity and nitrile-degrading enzymes. *Archives of Microbiology* **170:** 85-90.

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.

Kaufmann,G., Dautzenberg,H., Henkel,H., Muller,G., Schafer,T., Undeutsch,B., and Oettel,M. (1999) Nitrile hydratase from Rhodococcus erythropolis: Metabolization of steroidal compounds with a nitrile group. *Steroids* **64:** 535-540.

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.

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 (BBA) - Gene Structure and Expression* **1129:** 23-33.

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.

Nagasawa,T. and Yamada,H. (1989) Microbial transformations of nitriles. *Trends in Biotechnology* **7:** 153-158.

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

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

Stevens,J.M., Rao Saroja,N., Jaouen,M., Belghazi,M., Schmitter,J.M., Mansuy,D. et al. (2003) Chaperone-assisted expression, purification, and characterization of recombinant nitrile hydratase NI1 from Comamonas testosteroni. *Protein Expression and Purification* **29:** 70-76.

Takashima,Y., Kawabe,T., and Mitsuda,S. (2000) Factors affecting the production of nitrile hydratase by thermophilic Bacillus smithii SC-J05-1. *Journal of Bioscience and Bioengineering* **89:** 282-284.

Takashima,Y., Yamaga,Y., and Mitsuda,S. (1998) Nitrile hydratase from a thermophilic Bacillus smithii. *Journal of Industrial Microbiology & Biotechnology* **20:** 220-226.

Thomas,S.M., DiCosimo,R., and Nagarajan,A. (2002) Biocatalysis: applications and potentials for the chemical industry. *Trends in Biotechnology* **20:** 238-242.

Tsujimura,M., Dohmae,N., Chijimatsu,M., Takio,K., Odaka,M., Yohda,M. et al. (1997) Photoreactive nitrile hydratase: Posttranslational modification of photoreactive site. *Journal of Inorganic Biochemistry* **67:** 335.

Wang,M.X. and Feng,G.Q. (2000) Enantioselective synthesis of chiral cyclopropane compounds through microbial transformations of trans-2-arylcyclopropanecarbonitriles. *Tetrahedron Letters* **41:** 6501-6505.

Wang,M.X., Li,J.J., Ji,G.J., and Li,J.S. (2001) Enantioselective biotransformations of racemic 2-aryl-3-methylbutyronitriles using Rhodococcus sp AJ270. *Journal of Molecular Catalysis B-Enzymatic* **14:** 77-83.

Wegman,M.A., Heinemann,U., van Rantwijk,F., Stolz,A., and Sheldon,R.A. (2001) Hydrolysis of D,L-phenylglycine nitrile by new bacterial cultures. *Journal of Molecular Catalysis B-Enzymatic* **11:** 249-253.

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.

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.

Yamamoto,K., Ueno,Y., Otsubo,K., Yamane,H., Komatsu,K.i., and Tani,Y. (1992) Efficient conversion of dinitrile to mononitrile-monocarboxylic acid by Corynebacterium sp. C5 cells during tranexamic acid synthesis. *Journal of Fermentation and Bioengineering* **73:** 125-129.

Chapter 2: Materials and Methods

2.1 Chemicals and reagents

Chemicals were supplied by Merck Chemicals and Laboratory Supplies, Sigma-Aldrich Chemical Company and Kimix Chemical and Laboratory Supplies. All chemicals used were of the highest grade available.

Oxoid Ltd and Biolabs supplied culture media.

DNA size markers, protein size markers and all DNA modifying enzymes (polymerases and restriction endonucleases) were purchased from Fermentas Life Sciences Ltd.

Oligonucleotide primers for polymerase chain reaction (PCR) were synthesized by Inqaba Biotech or Integrated DNA Technologies (IDT) Inc.

2.2 Buffers and solutions

Table 2.1 describes the compositions of routinely utilised buffers and solutions.

Table 2.1 Composition of buffers and solutions.

Buffer/Solution	Composition	pH
Agarose gel loading dye (6X)	0.25% (w/v) bromophenol blue 40% (w/v) sucrose	
Ammonia assay reagent A	0.59 M phenol 1 mM sodium nitroprusside	
Ammonia assay reagent B	0.11 M sodium hypochlorite 2 M sodium hydroxide	
1 M Potassium phosphate buffer	717 ml 1 M K_2HPO_4 283 ml 1 M KH_2PO_4	7.2
SDS-PAGE electrode buffer (10X)	0.25 M Tris HCl 2 M glycine 1% (w/v) SDS	8.3
SDS-PAGE Gel-loading Buffer (2X)	100 mM Tris-Cl 4% (w/v) SDS 0.2% (w/v) bromophenol blue 20% (v/v) glycerol 200 mM dithiothreitol	6.8
PAGE staining solution	0.2% (w/v) Coomassie Blue R 250 40% (v/v) methanol 10% (v/v) acetic acid	
PAGE destaining solution	40% (v/v) methanol 10% (v/v) acetic acid	
20X TAE buffer	2M Tris base 25mM EDTA (pH adjusted with glacial acetic acid)	8.3

2.3 Bacterial strains

Table 2.2 describes all bacterial strains used in this study.

Table 2.2 Bacterial strains used in this study.

Bacterial strain	Relevant Genotype	Supplier
Bacillus sp. RAPc8	NHase and amidase producer	ARCAM stock (Pereira, 1998)
E. coli BL21 (DE3)	*hsdS gal ompT* (λcI*ts*857 *ind*1 Sam7 *nin5 lac*UV5-T7 gene1)	Stratagene
E. coli BL21 pLysS	As above. [pLysS]	Stratagene

2.4 Plasmids

The expression plasmids used in this study are shown in Table 2.3

Table 2.3 Expression plasmids used in this study.

Plasmid	Description	Source
pET21a (+)	Vector for high-level protein expression in *E.coli.*	Novagen
pNH223	pET21a(+) derivative carrying amidase gene of *Bacillus* sp. RAPc8.	Cameron, 2002
pNH512	pET21a(+) derivative carrying NHase α and β subunit genes of *Bacillus* sp. RAPc8.	Cameron, 2002
pNH14K (also referred to as pNH461)	pET21a(+) derivative carrying NHase α and β subunit genes and P14K gene of *Bacillus* sp. RAPc8.	Cameron *et al*., 2005
pNH14K βF36L	NHase βF36L mutant of pNH14K	This work
pNH14K βF52G	NHase βF52G mutant of pNH14K	This work
pNH14K βF55L	NHase βF55L mutant of pNH14K	This work
pNH14K βY67A	NHase βY67A mutant of pNH14K	This work
pNH14K βW76G	NHase βW76G mutant of pNH14K	This work

2.5 Analytical Procedures

2.5.1 Spectrophotometry

Spectrophotometric analyses were carried out using a Biomate 3 spectrophotometer (Thermo Electronic Company) or a Cary 300 *Bio* spectrophotometer operated *via* WinUV software (Varian Ltd).

2.5.2 Determination of protein concentration

2.5.2.1 The Bradford assay

Bradford assays (Bradford, 1976) were performed using the Biorad Protein Assay (Biorad). Manufacturer's instructions were followed except that the procedure was scaled down to a 1ml reaction volume for convenience. Protein standards covering 0 to 1.4 mg/ml bovine serum albumin (BSA) were used. A typical standard curve is shown in Appendix 1.

2.5.2.2 Direct spectrophotometric determination

The concentration of pure *Bacillus* sp. RAPc8 NHase was determined by direct spectrophotometric analysis at a wavelength of 280 nm. The molar extinction coefficient (78090 M^{-1} cm^{-1}) of NHase was calculated based on the amino acid sequence using the PROTPARAM tool available on the World Wide Web (http://au.expasy.org/tools/protparam.html). This web server relies on the method of Gill and von Hippel, 1989.

2.5.3 Polyacrylamide gel electrophoresis (PAGE)

PAGE was carried out using an adaptation of the method of Laemmli (1970), on a Hoefer SE 250 minigel electrophoresis unit. Typically, 0.75 mm gels containing 12 or 15% acrylamide were prepared using a Hoefer SE 245 dual gel caster. For SDS-PAGE, samples were mixed with SDS-PAGE gel loading buffer and boiled briefly prior to loading. For native-PAGE, samples were mixed with the same buffer without SDS and DTT (Table 2.1). Components used for preparation of gels are listed in Tables 2.4 and 2.5. Native gels were prepared in the same manner except that they did not contain SDS. Gels were electrophoresed at a constant current of 50mA in running buffer. Gels were stained with PAGE staining solution; the process was speeded up by heating at medium setting in a 660-Watt microwave. Gels were then destained overnight using PAGE destaining solution.

Table 2.4 Preparation of 12% and 15% separating gels for PAGE. For preparation of native PAGE gels, SDS was excluded from the recipe.

	Volume (ml)	
	12% Gel	15% Gel
dd H_2O	6.6	4.6
Buffer (1.5 M Tris-HCl, pH 8.8)	5	5
20% (w/v) SDS	0.2	0.2
Monomer solution (30% acrylamide, 0.8% bis-acrylamide)	8	10
10% (w/v) ammonium persulfate	0.2	0.2
TEMED	0.01	0.01

Table 2.5 Preparation of stacking gels for PAGE. For preparation of native PAGE gels, SDS was excluded from the recipe.

	Volume (ml)
ddH_2O	2.8
0.5 M Tris-HCl	1.25
20% (w/v) SDS	0.05
Monomer solution (30% acrylamide, 0.8% bis-acrylamide)	0.85
10% (w/v) ammonium persulfate	0.05
TEMED	0.005

2.5.4 NHase activity assays

2.5.4.1 Ammonia detection assay

The ammonia detection assay is a calorimetric assay that quantifies the release of ammonia using a modification of the well-established phenol-hypochlorite detection method (Fawcett and Scott, 1960). The assay was carried out exactly as described previously (Cramp and Cowan, 1999). Partially purified amidase used in the assay was produced from *E. coli* BL21 (DE3) transformed with pNH223 as previously described (Cameron, 2002). The blue product formed was detected spectrophotometrically at 600nm absorbance and the results interpolated against a standard curve prepared using 0 to 2mM ammonium chloride. A typical standard curve is shown in Appendix 2.

2.5.4.2 Continuous spectrophotometric assay

A continous spectrophotometric assay was used for enzyme kinetic studies (section 2.9). Assays were carried out at 50˚C. A 1ml reaction containing 50mM potassium phosphate, pH7.2, and an appropriate concentration of nitrile dissolved in methanol was prepared in a quartz cuvette. The cuvette was incubated briefly at 50˚C to allow reaction componenets to equilibrate. The reaction was initiated by addition of an appropriate amount of enzyme. Details are described in section 2.9. Table 2.6 lists molar extinction coefficients and wavelengths used for detection of the products formed from each of the substrates.

Table 2.6 Molar extinction coefficients and wavelengths used for detection of amide products during assay for activity of specified nitriles.

Substrate	Wavelength for detection of product (nm)	Molar extinction coefficient of product ($mM^{-1}.cm^{-1}$)
Acrylonitrile	225	2.9
3-Cyanopyridine	235	3.2 (N.B.3-Cyanopyridine: 0.8)
Benzonitrile	242	5.5

2.6 Site-directed mutagenesis

Plasmid pNH14K was used as the template for mutagenesis. Site-directed mutants were generated using the polymerase chain reaction (PCR)-based QuickChange site-directed mutagenesis kit (Statagene) according to the manufacturer's recommendations.

The following cycling parameters were used for the amplification step:

95˚C for 30 seconds	1 Cycle
95˚C for 30 seconds 55˚C for 1 minute 68˚C for 12 minutes	16 Cycles

The primers used were as follows:

5`GATTGGGAAAGACTTGCGTT**A**GG**C**CTTGTAGCTGG3` (βF36L),

5`GATTGGGGATGAAGGC**CGG**TGATGAATTCAGGATCG3` (βF52G),

5`CTTTTGATGAA**C**TCAGGATCGGCATTG3` (βF55L),

5`GCTTATGCGTCCAGTGGAT**GG**TTTGACGTCGTCG3` (βY67A) and

5`CGTCGTATTATGGCCAT**G**GGATTGCAACCGTTG3` (βW76G).

Only the sense primers are shown in each case; the antisense primers were the reverse complement of the sequences shown. Nucleotides that resulted in mutation and/or new restriction sites are highlighted in bold and underlined. The presence of a mutation in the final product was first identified by restriction endonuclease digestion and then confirmed by DNA sequencing. Inqaba Biotech provided the DNA sequencing service.

2.7 Protein expression

Wild-type NHase was recombinantly expressed in *E. coli* BL21 (DE3) or *E. coli* BL21 (DE3) pLysS (Stratagene). Cells were transformed with either expression vector pNH14K or pNH512 (Table 2.3). Mutant NHases were expressed in the same

hosts transformed with mutated pNH14K constructs as listed in Table 2.3. Competent cells were prepared using the calcium chloride method and transformed using the heat-shock method (Sambrook and Russell, 2001).

An 800ml LB culture containing 50μg/ml ampicillin (or carbenicillin) and 30μg/ml chloramphenicol where appropriate, was grown at 37˚C with shaking at 220 rpm to an optical density (at 600nm) of 0.4, at which point expression was induced with 0.4mM IPTG. Cobalt chloride was added to a final concentration of 0.1mM, 15 to 30 minutes prior to induction. 4 hours after induction cells were harvested by centrifugation and washed with 50mM potassium phosphate buffer, pH 7.2.

2.8 Protein purification

All chromatographic procedures were performed on an ÄKTA FPLC liquid chromatography system (Amersham Biosciences) controlled *via* a Unicorn graphic user interface (Version 4.10). Unless otherwise indicated, default flow rates and pressure limits for each column were used. Samples from all stages of purification were analysed by SDS-PAGE (section 2.6.3). Gels containing 12% or 15% acrylamide were used. Samples were assayed for protein concentration using the Biorad Protein Assay kit (Biorad) with BSA as a standard (section 2.6.2.1) and for nitrile hydratase activity using the ammonia detection assay (section 2.6.4.1).

2.8.1 Preparation of cell-free extracts and heat purification

Washed cell pellets were resuspended in 25 ml potassium phosphate buffer, pH 7.2. Because of the presence of T7 lysozyme to facilitate lysis in the pLysS strain, cells were easily disrupted by one cycle of freezing at –80˚C and thawing at 37˚C followed by brief sonication on ice. Sonication was carried out on a Bandelin Sonoplus

HD2070 sonicator in cycles of 30s pulse, 30s stop for 8 minutes. The lysate was centrifuged at 5000 x g for 20 minutes and the supernatant collected. The relative thermostability of *Bacillus* sp. RAPc8 was exploited for the first step of purification. Heat treatment was used to precipitate heat-sensitive *E. coli* proteins. The cell free extract was heated at 55˚C for 45 minutes, centrifuged at 7000 x g for 20 minutes and the supernatant collected. Due to the lower thermostability of mutant NHases, crude extracts produced from strains producing these proteins were not subjected to this heat purification step.

2.8.2 Hydrophobic interaction chromatography (HIC)

Solid ammonium sulphate was added to the heat-treated sample to achieve 20% saturation (Appendix 3) and left on ice for 1 hour. Precipitated proteins were removed by centrifugation at 7000 x g for 30 minutes at 4˚C. The supernatant was loaded onto a HighLoad 16/10 Phenyl-Sepharose column (Amersham Biosciences) equilibrated with buffer containing 1.0M ammonium sulphate, 50mM potassium phosphate, pH 7.2. Bound proteins were eluted with a linear gradient of decreasing ammonium sulphate concentration generated with 50mM potassium phosphate buffer, pH 7.2 (5 column-volumes, 1.0M – 0M ammonium sulphate). Fractions containing NHase were pooled.

2.8.3 Ion exchange chromatography

Pooled fractions from HIC were dialysed against 25mM potassium phosphate buffer, pH 7.2 and loaded onto a HiPrep 16/10 Q-Sepharose FF column or 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 generated with buffer containing 500mM sodium chloride, 25mM

potassium phosphate, pH 7.2 (5 column volumes, 0M – 500mM sodium chloride). Fractions containing NHase were pooled.

Mutant NHases were further purified by chromatography on a MonoQ 5/50 (Amersham Biosciences) column. This step was also occasionally used for the wild type protein in the place of or after the step described above (see section 3.2.5). Pooled fractions from the previous step were dialysed against 25mM potassium phosphate buffer, pH7.2 and loaded onto a MonoQ 5/50 column. Bound proteins were eluted with a linear gradient of increasing sodium chloride concentration generated with buffer containing 500mM sodium chloride, 25mM potassium phosphate, pH 7.2 (20 column volumes, 0M – 500mM sodium chloride). Fractions containing NHase were pooled.

2.8.4 Size exclusion chromatography (Gel Filtration)

Prior to crystallisation experiments, pure NHase samples were subjected to gel filtration (section 3.2.4). Pure NHase was concentrated to 0.5ml and loaded on Sephacryl S300 gel filtration chromatography media packed in a XK26/100 column (Amersham Biosciences) according to manufacturer's guidelines. The column was equilibrated with 25mM potassium phosphate buffer, pH7.2 and run at a flow-rate of 0.8ml/minute for 1.5 column volumes.

2.9 Enzyme kinetics

All enzyme kinetic data were determined using the direct spectrophotometric assay described in section 2.5.4.2. For determination of V_{max} and K_M, assays were done in triplicate using 5mM, 10mM, 20mM, 30mM, 40mM and 50mM substrate. In each 1ml reaction, 0.6μg (according to spectrophotometric method, section 2.5.2.2)

enzyme was used. For determination of the inhibition kinetics of benzonitirile inhibition of NHase, assays were repeated in the presence of 100μM and 200μM benzonitrile. 0.6μg enzyme was used in each assay. Progress of reactions was monitored on a Cary 300 Bio spectrophotometer operated via WinUV software (Varian Inc.). Data were analysed using Enzpack (Biosoft).

2.10 Crystallisation experiments

2.10.1 Sample preparation

Prior to initial crystallisation, pooled fractions of pure NHase from Q-Sepharose chromatography were dialysed against 20 mM Tris, pH 7.2, filtered through a 0.22 micron filter and concentrated to 10 mg/ml (as determined by the direct spectrophotometric method described in section 2.6.2.2) using Amicon Centriprep or Whatman VectorSpin tubes.

2.10.2 Initial crystallisation trials

NHase was crystallised using the hanging-drop vapour diffusion method (see Figure 3.11). Initial crystallisation conditions were tested using Hampton Crystal Screen 1 (Hampton Research) and JBScreen 1 (JenaBioscience). The wild-type protein was used in all crystal trials. A total of 74 individual conditions were tested. Parallel screens at 20˚C and 4˚C were prepared. 24-well VDX plates (Hampton Research) were used. The hanging drop was mounted on square siliconised cover slips and contained 1μl protein and 1μl reservoir solution.

2.10.3 Optimisation of crystallisation

Once initial conditions were resolved, optimisation was carried out with drops containing 2μl protein and 2μl reservoir solution. Protein concentration, incubation temperature, precipitant concentration and reservoir solution pH were varied. Details are described and discussed in section 3.3.3.

Diffraction quality crystals were finally grown at 20˚C in 30% PEG 400, 100mM magnesium chloride, 100mM MES (2[N-Morpholino]ethanesulfonic acid), pH 6.5 (40 mg/ml protein). Mutant NHases were then crystallised under identical conditions.

2.11 Collection of X-ray diffraction data

X-ray diffraction data from crystals of wild-type and mutant NHases were collected at the in-house x-ray source at the Department of Biotechnology, University of the Western Cape. This equipment comprised 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. Data from crystals mounted on a cryoloop (Hampton Research) were collected under cryo-conditions (in a nitrogen stream at a temperature of 100 K) with a crystal-to-detector distance of 160 mm. It was not necessary to add cryoprotectant to the crystals (see Section 3.4). Data frames covering an oscillation angle of 0.5^0 per frame were collected for 10 minutes. The machine was operated *via* the program Crystal Clear (Pflugrath *et al*., 1999).

2.12 X-ray-data processing and phase calculation

X-ray data processing was performed with Crystal Clear (d*TREK) (Pflugrath *et al*., 1999) and HKL2000 (Minor *et al*., 2002). Solvent content and Matthews Coefficient were calculated using Matthews from the CCP4 suite of programs (CCP4, 1998.). Other routine procedures (e.g. file conversion etc.) were carried out using programs from the CCP4 suite. A comprehensive guide for the CCP4 suite is provided online at http://www.ccp4.ac.uk/docs.php.

The wild-type structure was solved by molecular replacement with EPMR (Kissinger *et al*., 1999) using the crystal structure of *Pseudonocardia thermophila* NHase (Miyanaga *et al*., 1997; PDB code: 1IRE) as a search probe. The most appropriate search probe from previously solved structures was identified using FUGUE (Shi *et al*., 2001). The mutant structures were solved using the wild-type *Bacillus* sp. RAPc8 structure as a search probe. The program Molrep was used.

2.13 Refinement and validation of crystal structures

Initial side-chain substitution prior to model-rebuilding in the wild-type structure was done using SCWRL accessible *via* a web server (http://www1.jcsg.org/scripts/prod/scwrl/serve.cgi). Manual model-rebuilding and addition of water molecules was performed in O (Kleywegt and Jones, 1997) and NOC (Institute of Biophysics, Chinese Academy of Sciences). All structures were refined using Refmac5 from the CCP4 suite of programs. The model-rebuilding and refinement procedures for the wild-type protein were carried out by Dr Ozlem Tastan-

Bishop under the supervision of Dr Muhammed Sayed, Department of Biotechnology, University of the Western Cape.

Validation of the crystal structures was performed using PROCHECK (Laskowski *et al*., 1993) and RAMPAGE (Lovell *et al*., 2003).

2.14 Molecular Graphics and structure analysis

Unless otherwise indicated, all molecular graphics operations were performed using PyMOL (Delano Scientific LLC). All images were ray-traced using PyMOL.

PDB format co-ordinate files were edited using PDBSet (CCP4 suite) or manually in WordPad (Microsoft Corporation) where appropriate.

Detailed analysis of the topology of the crystal structure of wild-type *Bacillus* sp. RAPc8 NHase was carried out using TOPS (Michalopoulos *et al*., 2004). TOPS was accessed at http://www.tops.leeds.ac.uk/.

Detailed analysis of the interface between cognate dimers in the structure of the NHase heterotetramer was performed using the Protein-Protein interface server (http://www.biochem.ucl.ac.uk/bsm/PP/server/).

Cavity analysis was carried out using CASTp (Liang *et al*., 1998) accessed on the World Wide Web (http://cast.engr.uic.edu/cast). Cavities were visualised using Chime (Elsevier MDL).

Site-directed mutant model structures were prepared using the *mutagenesis* tool in PyMOL.

Chapter 3: Purification, crystallisation and preliminary X-ray diffraction analysis of *Bacillus* sp. RAPc8 NHase

A crucial step in the process towards elucidating the structure of a protein by X-ray diffraction analysis is the efficient expression and purification of the protein of interest. While it is difficult to predict accurately all the factors that will affect production of crystals of adequate form and size for structural studies, the best practice (to increase the probability of producing good crystals of a new protein) is to start with a sample that is as homogeneous as possible.

3.1 Expression of active recombinant nitrile hydratase

Previous reports on expression of cobalt-containing NHases have shown that the presence of cobalt ions in the growth media is essential for production of active enzyme. Cameron (2002) expressed this nitrile hydratase with this premise in mind and demonstrated that addition of 0.1mM $CoCl_2$ to the media resulted in production of ~49 U/mg of nitrile hydratase activity in the crude extract. It was also shown (Kim *et al.,* 2001) that even higher specific activities of *Bacillus* sp. BR449 NHase could be achieved by incubating the apoenzyme in 5 μM $CoCl_2$ at 50˚C.

Cameron (2002) cloned the nitrile hydratase operon from *Bacillus* sp. RAPc8 and reported the construction of vectors for its expression in *E. coli.* Two of these expression constructs were acquired for use in the present study. pNH512 expressed the β and α subunits of nitrile hydratase, while pNH14K expressed these subunits

along with a small downstream protein termed P14K. The open reading frames were sub-cloned into pET21a for IPTG-inducible expression under control of the T7 RNA polymerase promoter. The expression host selected for use was *E.coli* BL21 (DE3), an effective protease-deficient expression strain lacking the *lon* protease and the OmpT protease.

Three different strategies for expressing highly active NHase were tested. In one, the α and β subunits of NHase were expressed in culture medium supplemented with 0.1 mM $CoCl_2$. In another strategy, the α and β subunits of NHase were expressed in the absence of $CoCl_2$ in the medium. The clarified extract containing the apoenzyme was then incubated in 5μM $CoCl_2$ for 45 minutes at 50˚C. In the third stategy the α and β subunits were co-expressed with P14K, a small protein whose gene was found downstream NHase in the same operon. The medium was supplemented with 0.1mM $CoCl_2$. It was found that optimal expression was achieved with the third strategy, with as much as 30-fold improvement in activity relative to NHase produced in the absence of P14K. The procedure described here was essentially a duplication and confirmation of work reported by Cameron *et al*. (2005).

Table 3.1 Treatment alternatives for activation of Bacillus sp. RAPc8 NHase

Treatment	**Activity (U/mg)**
0.1mM $CoCl_2$ in growth medium	50
Heat-treatment of apoenzyme in presence of 5μM $CoCl_2$	450
Co-expression with P14K with 0.1mM $CoCl_2$ in growth medium	1600

The requirement for small 'activator' proteins from flanking genes for heterologous expression of active NHases has been demonstrated in many other species (Nojiri *et al*., 1999, Petrillo *et al*., 2004). The results shown above for *Bacillus* sp. RAPc8 NHase are consistent with previous findings. The exact role played by these activator proteins is not known. Hypotheses for their function include the following:

1. Incorporation of metal co-factor into the active site.
2. Transport of metal co-factor into the cell (Komeda *et al*., 1997).
3. Chaperone-like activity during folding or assembly of NHase (Cameron *et al*., 2005).
4. Catalysis of the post-translational modification of cys residues of the metal-binding centre.

Further work aimed at elucidating the precise function of P14K in *Bacillus* sp. RAPc8 was in progress in parallel with this work and the results are reported in part elsewhere (Cameron *et al*., 2005).

3.2 Purification of Bacillus sp. *RAPc8 NHase*

NHase was previously purified from the native *Bacillus* sp. RAPc8 using a combination of hydrophobic interaction and ion exchange chromatographies (Pereira, 1998). Partial purification of heterologously expressed recombinant *Bacillus* sp. RAPc8 NHase was also achieved using a heat-treatment strategy (Cameron, 2002).

The various recombinant technologies available for simplified purification of recombinantly expressed protein would have allowed rapid purification of the NHase for the purpose of this study. These included expression of NHase as a fusion with GST followed by rapid purification on glutathione-Sepharose (i.e., the pGEX system) or expression with a hexa-histidine tag followed by purification by IMAC (immobilised metal affinity chromatography) on chromatography media incorporating chelated nickel (e.g., chelating Sepharose or NTA-agarose).

Here, techniques that did not require the incorporation of tags and/or additional amino acids were selected, in order that the crystal structure would closely resemble that produced by NHase from the native *Bacillus* organism.

3.2.1 Heat Purification

Earlier work demonstrated that *Bacillus* sp. RAPc8 NHase was moderately thermostable (Pereira, 1998). The recombinant enzyme was found to maintain maximal activity after incubation at 55˚C for up to 50 minutes (Cameron, 2002). This property and the fact that proteins produced by the mesophilic *E. coli* host are generally less tolerant to heat-treatment was exploited in the first step of purification.

Clarified crude extracts were subjected to heating at 55˚C for 45 minutes. Qualitative analysis by SDS-PAGE showed that this procedure was able to selectively precipitate a good proportion of the *E. coli* proteins (Figure. 3.1). The sample was also assayed for NHase activity after centrifugation. 107% of NHase activity was recovered in the heat-treated extract. The increase in activity relative to the clarified crude extract was possibly due to removal of inhibitory factors on heat treatment. This step gave a 1.8-fold purification.

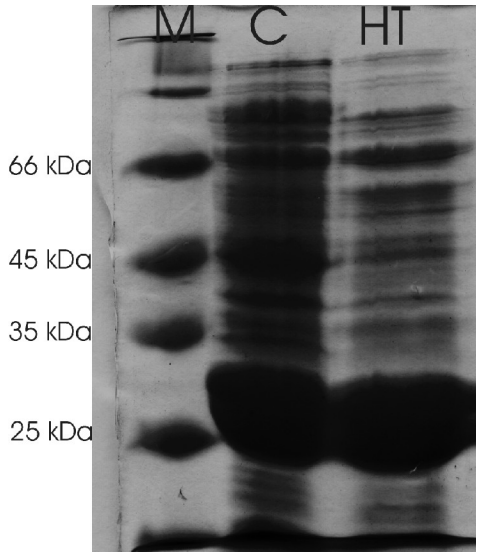


Figure 3.1 SDS-PAGE analysis of heat-treatment step of NHase purification. M: molecular weight marker, C: crude extract, HT: heat-treated extract.

3.2.2 Hydrophobic interaction chromatography (HIC)

Previous studies aimed at purification of NHases showed that these enzymes have a high affinity for HIC media. In the purification of recombinant *Bacillus* sp. BR449 NHase, which shares very high sequence identity with RAPc8 NHase, the use of phenyl-Sepharose HIC generated near-homogeneous enzyme (Kim *et al*., 2001). A similar technique was employed for purification of recombinant *Bacillus* RAPc8 NHase.

Prior to phenyl-Sepharose HIC the heat-treated extract was adjusted to a 1M (equivalent to 25% w/v saturation) ammonium sulphate. This procedure resulted in some purification as a small proportion of the proteins present in the extract were precipitated. It was difficult to assess either qualitatively or quantitatively the extent of this purification as the high concentration of ammonium sulphate present after treatment severely interfered with SDS-PAGE and the ammonia assay for NHase activity.

The heat-treated extract containing 1M ammonium sulphate was loaded onto a HiLoad 16/10 Phenyl-Sepharose HP column equilibrated with phosphate buffer containing 1M ammonium sulphate. Bound proteins were eluted with a 1-0M gradient of decreasing salt. The resulting chromatogram is shown in Figure 3.2.

Fractions exhibiting high absorbance at 280nm wavelength were analysed for presence of NHase by SDS-PAGE (Figure 3.3). NHase was eluted in fractions 15 to 23 at approximately 0.7-0.6M ammonium sulphate. Fractions with low background protein contamination were pooled, and assayed for NHase activity as outlined in section 2.5.4.1. The pooled fractions were also analysed by native PAGE showing the presence of one major contaminating protein (Figure 3.3, inset).

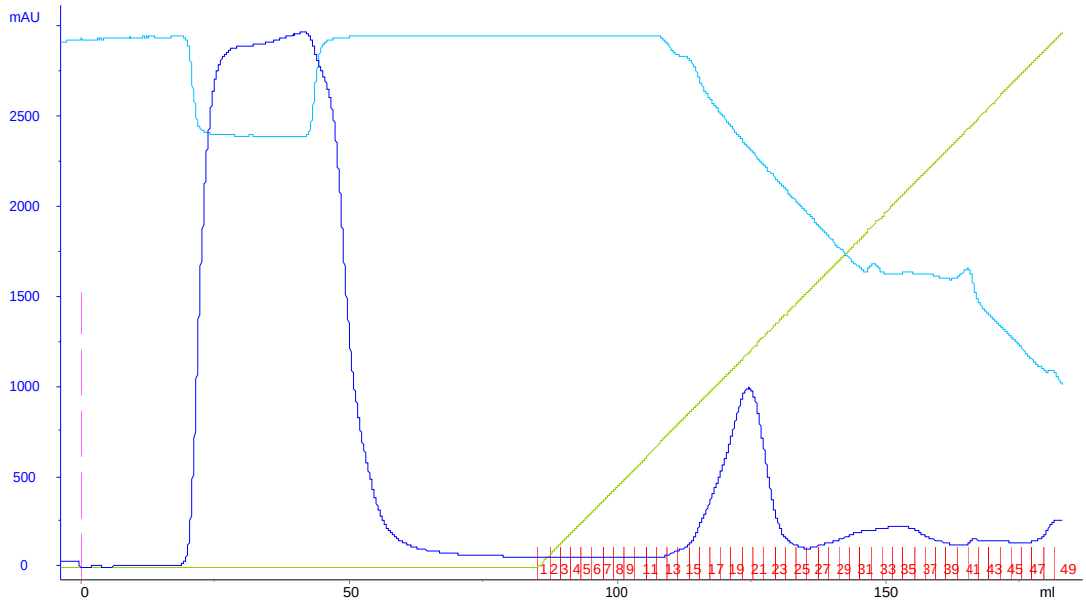


Figure 3.2 Chromatogram from phenyl-Sepharose hydrophobic interaction chromatography. A broad UV peak corresponding to the flow-through One distinct UV peak corresponding to fractions 15 –to 23 can be seen during gradient elution. — : UV, — : % Buffer B (0 to 100), — : Relative conductivity.

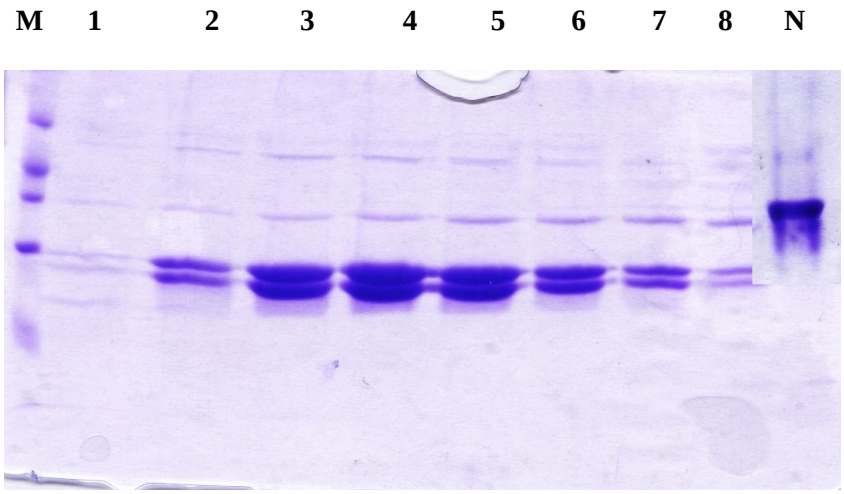


Figure 3.3 SDS-PAGE analysis of fractions from phenyl-Sepharose hydrophobic interaction chromatography (HIC). Lane M: molecular size marker, lanes 1 to 8: fractions 16 to 23. Panel N: Native-PAGE analysis of pooled fractions from HIC .

This step yielded 65% NHase activity with a 7.3-fold purification.

3.2.3 Anion exchange chromatography

Several reports on the purification of NHases have detailed the use of anion exchange chromatography. Pereira (1998), in particular, showed that use of Q-Sepharose (a strong anion exchanger) yielded a very high degree of purification for NHase from *Bacillus* sp. RAPc8.

Pooled, dialysed samples from HIC were loaded onto a HiPrep 16/10 Q-Sepharose FF column equilibrated with low salt buffer. Bound proteins were eluted with a 5 column-volume linear gradient of increasing sodium chloride concentration (0-500mM). The resulting chromatogram showed very small amounts of protein in the flow-through (unbound fraction), with one distinct peak eluting at approximately 150-250 mM sodium chloride (Figure 3.4).

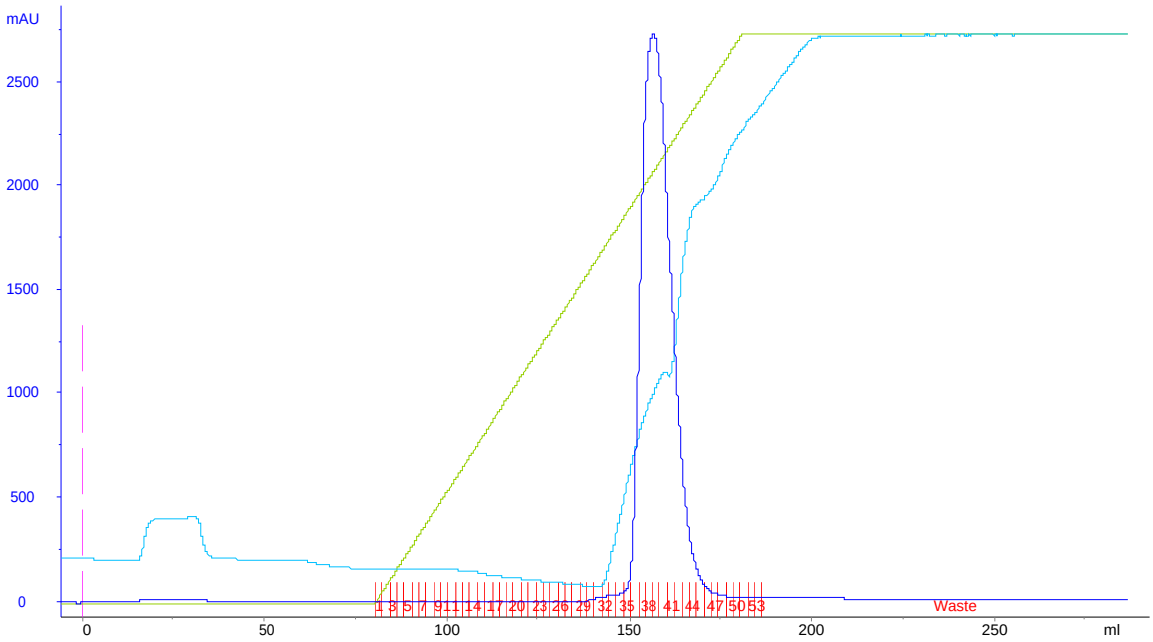


Figure 3.4 Chromatogram from Q-Sepharose ion exchange chromatography. One distinct UV peak corresponding to fractions 35 to 47 is present. No significant flow-through peak was observed. — : UV, — : % Buffer B (0 to 100), — : Relative conductivity.

Fractions from this peak were analysed by SDS-PAGE (Figure 3.5). Fractions 37 to 44 contained NHase that appeared almost electrophoretically pure. These fractions were pooled and dialysed into low ionic strength buffer. Assays for NHase activity revealed that this step resulted in 60.5 % yield and 8.9-fold purification.

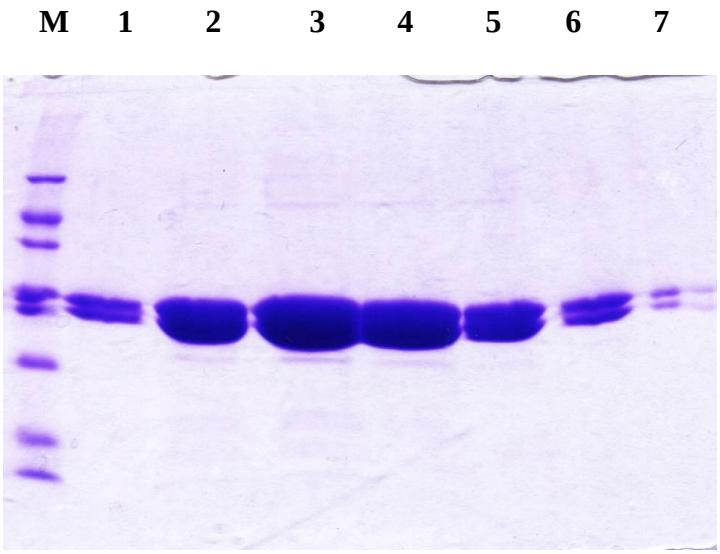


Figure 3.5 SDS-PAGE analysis of fractions from Q-Sepharose chromatography. Lane M: molecular size marker, lanes 1 to 7: fractions 35 to 47 at intervals of two.

3.2.4 Oligomeric state of pure NHase

The monodispersity of 'pure' protein destined for crystallisation must be assessed carefully prior to crystallisation trials. Analysis for electrophoretic purity by SDS-PAGE is not sufficient as the denaturing conditions of this technique preclude any useful information on the oligomeric state of the native protein.

Earlier characterisation of *Bacillus* sp. RAPc8 NHase showed that it exists as a heterotetramer of about 100 kDa molecular weight (Pereira, 1998). To determine the

oligomeric state of NHase after Q-Sepharose ion exchange chromatography, the sample was concentrated to 0.5 ml and loaded on Sephacryl S300 gel filtration chromatography matrix packed in a XK26/100 column.

This gel filtration procedure showed that the pure NHase sample contained only a single oligomeric state as indicated by the presence of one elution peak (Figure 3.6).

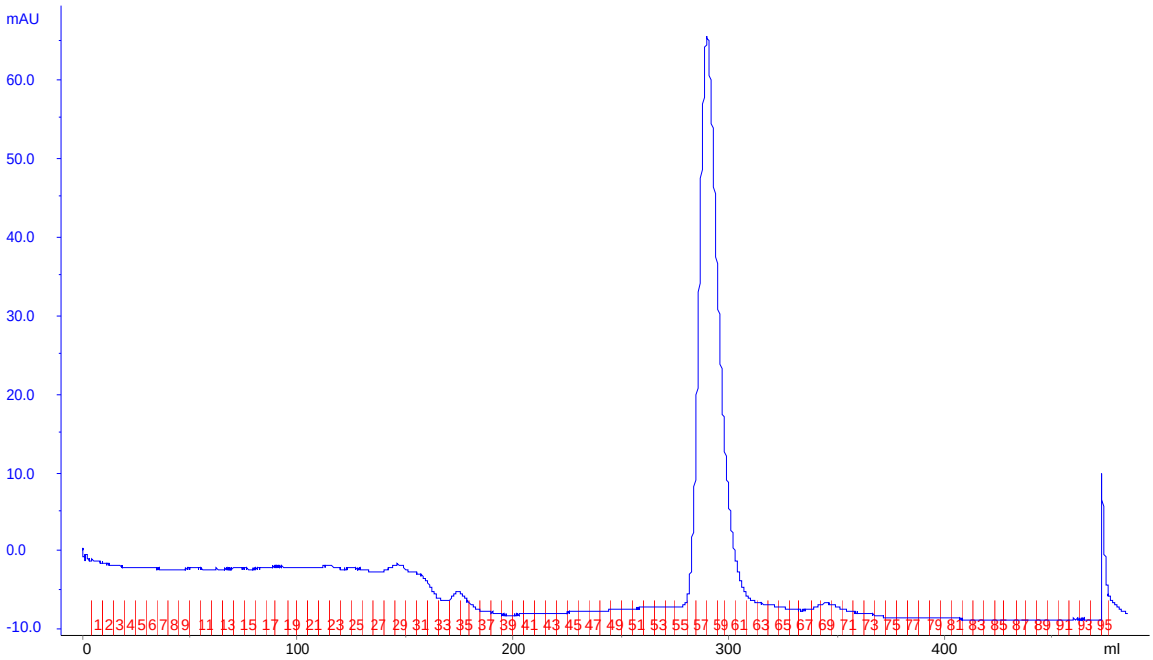


Figure 3.6 Chromatogram from Sephacryl S300 gel-filtration chromatography.

3.2.5 Alternative chromatography methods

Throughout this study, NHase was routinely purified using the procedures described above. Different chromatographic procedures were used as they became available in attempts to further optimize NHase purification.

NHase from HIC was loaded onto a pre-packed MonoQ 5/50 column and eluted with a gradient of increasing sodium chloride concentration. The resulting chromatogram

showed two prominent protein peaks eluting at around 200mM sodium chloride (Figure 3.7).

Analysis of the protein composition of these peaks by SDS-PAGE reveals bands corresponding to electrophoretically pure NHase in both peaks. Further analysis by native PAGE showed that NHase from both peaks was identical in size (~100kDa) (Figure 3.8). This observation removed the possibility that the differing interaction behaviour of the two pools of NHase with this particular column was due to presence of two oligomeric states.

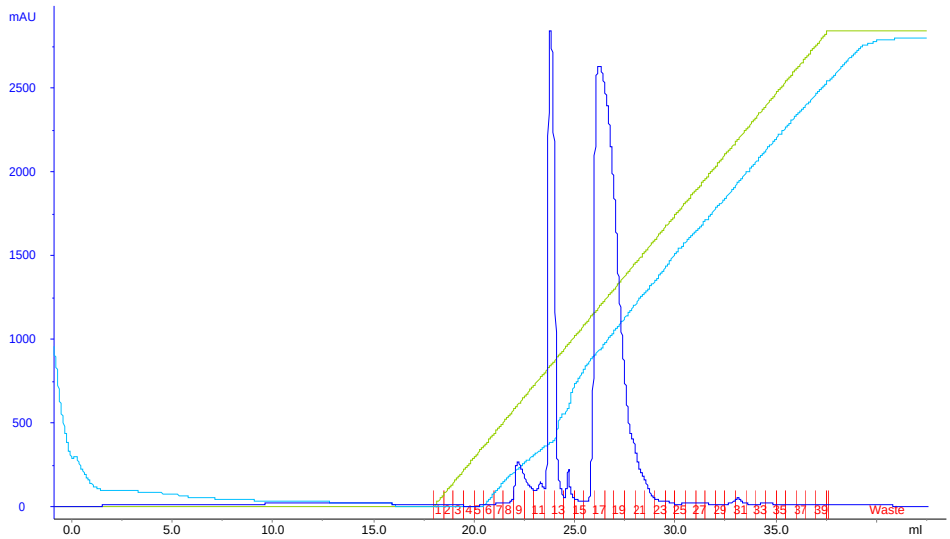


Figure 3.7 Chromatogram from MonoQ 5/50 ion exchange chromatography. Two prominent, distinct UV peaks (dark blue line) are present. — : UV, — : % Buffer B (0 to 100), — : Relative conductivity.

Some explanations for this behaviour can be proposed. One possibility can be deamination of a proportion of the protein sample during the heat-treatment step to give negatively-charged carboxyl groups that would inevitably affect the chromatography.

Figure 3.8 Native-PAGE analysis of pooled fractions from MonoQ 5/50 chromatography. Lane 1: pooled fractions from peak 1, lane 2: pooled fractions from peak 2.

3.2.6 Summary of NHase purification

Recombinant *Bacillus* sp. RAPc8 NHase was purified to near electrophoretic homogeneity from *E. coli* BL21 (DE3) in three steps. A summary of the purification steps and data is presented in Figure 3.9A. SDS-PAGE analysis of samples from each step is presented in Figure 3.9B. The purification table is shown in Table 3.2.

A

Purification (Fold)	Purification Step	Step Yield (%)
1	Crude cell free extract	100
	↓	
	Heat treatment for 45minutes at 55^0 C	106
1.8		
	↓	
7.3	Hydrophobic interaction chronmatography on phenyl-Sepharose	65
	↓	
8.9	Ionic exchange chromatography on Q-Sepharose	61

B

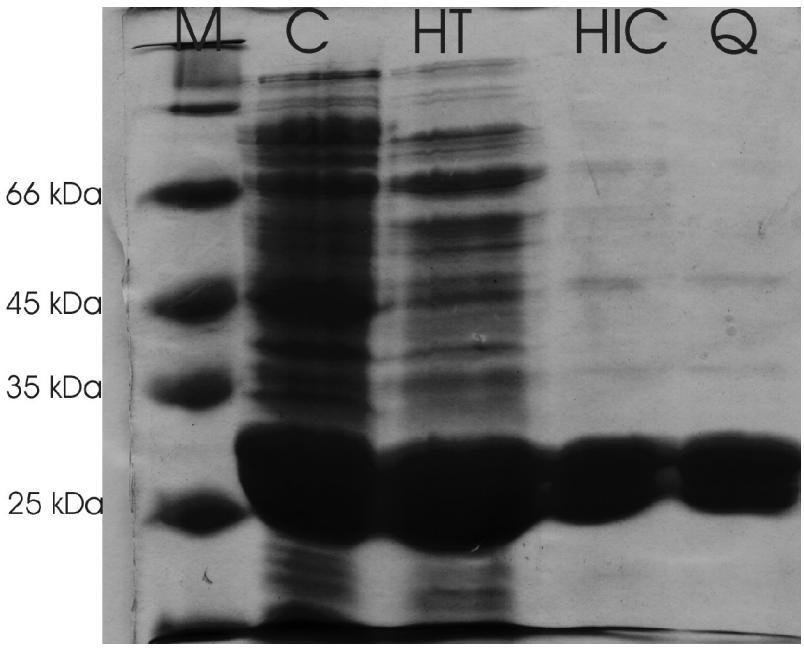


Figure 3.9 Summary of NHase purification (A) and SDS-PAGE analysis of all steps from the purification protocol (B). (M: molecular weight marker, C: crude extract, HT: heat-

treated extract, HIC: hydrophobic interaction chromatography, Q: ion exchange chromatography.)

Table 3.2 Purification table

Step	Total Protein (mg)	Total Activity (Units)	Specific Activity (Units/mg)	Purification (Fold)	Yield (%)
Crude Extract	275	330000	1200	1	100
Heat Treatment	158	351648	2220	1.8	106.6
Phenyl-Sepharose	24	213742	8839	7.3	64.8
Q-Sepharose	19	199716	10375	8.9	60.5

3.3 Crystallisation of Bacillus sp. RAPc8 NHase

3.3.1 The strategy for crystallisation screening

Every macromolecule has unique properties. This is true even for proteins that differ by just a few amino acids. While several NHases with significant homology to *Bacillus* sp. RAPc8 NHase were crystallised prior to this study, for the purpose of crystallisation this NHase was considered as a new protein. The many factors that may be varied to produce crystals of a suitable size and form for crystallography analysis require that a coherent screening strategy is employed.

Table 3.3 Factors affecting crystallization (from McPherson, 1999).

Physical	Chemical	Biochemical
Temperature/temperature variation	pH	Purity of the macromolecule/impurities
Surfaces	Precipitant type	Ligands, inhibitors, effectors
Methodology/approach to equilibrium	Precipitant concentration	Aggregation state of the protein
Gravity	Ionic strength	Posttranslational modifications
Pressure	Specific ions	Source of macromolecule
Time	Degree of supersaturation	Proteolysis/hydrolysis
Vibrations/sound/mechanical perturbations	Reductive/oxidative environment	Chemical modifications
Electrostatic/magnetic fields	Concentration of the macromolecule	Genetic modifications
Dielectric properties of the medium	Metal ions	Inherent symmetry of the macromolecule
Viscosity of the medium	Cross-linkers/polyions	Stability of the macromolecule
Rate of equilibration	Detergents/surfactants/amphiphiles	Isoelectric point
Homogeneous/heterogeneous nucleants	Non-macromolecular impurities	History of the sample

Crystallisation variables are classified in three groups: physical, chemical and biochemical. The biochemical factors (Table 3.3) are essentially the inherent properties of the protein and the method used to express, purify and store the protein. Care was taken to keep these properties as constant and/or reproducible as possible.

Physical factors are primarily a function of the physical environment/location in which crystals are grown and the nature of consumables used. Crystallisation screens were set-up and incubated in an air-conditioned room free from vibrations, noise and magnetic fields. Screens were typically performed at both 20˚C and 4˚C. In this study the highly cost-effective hanging drop vapour diffusion method (Figure 3.10) was used. Glass cover slips used were of high quality with siliconised surfaces to avoid wetting and drop spread.

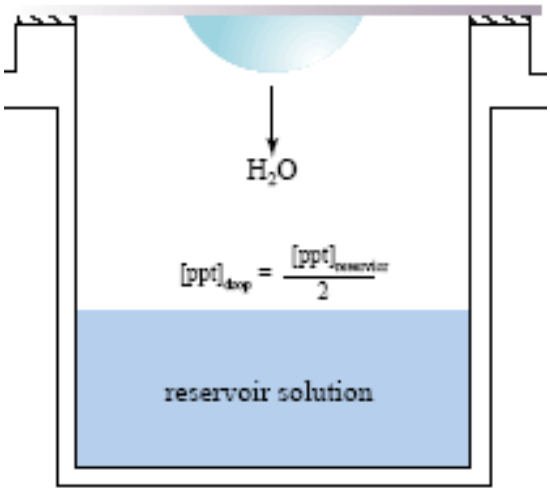


Figure 3.10 Process of vapour diffusion using the hanging drop approach (from Hampton Research crystallisation resource, Crystal Growth 101).

There were various approaches to screening available. Some can be described as systematic screens in that the investigator systematically varies each of the factors described above at fine intervals (i.e., each condition within the screen is closely related with another in some way). This approach to crystallisation screening is highly exhaustive/tedious and rapid success relies heavily on prior knowledge about the physical properties of the protein (e.g., solubility, pI etc.).

In this study, two commercial screening kits were used for initial crystallisation trials: Hampton Screen 1 (Hampton Research), first published by Jancarik and Kim (1991) and JBScreen (Classic) 1 (Jena Biosciences). These screens, comprising a total of 74 individual conditions, can be regarded as *shotgun* approaches to crystallisation screening in that they use a fairly large number of conditions that are coarsely distributed around the crystallisation condition space referred to as sparse matrix screens.

3.3.2 Initial crystallisation trials

Prior to initial crystallisation trials, NHase from Q-Sepharose chromatography was dialysed against 20mM Tris-Cl, pH 7.2, filtered through a 0.22-μm filter and concentrated to 10 mg/ml (as determined by absorbance at 280 nm). Parallel screens at 20˚C and 4˚C were prepared. 24-well VDX plates (Hampton Research) were used. The hanging drops, mounted on square siliconised cover slips, contained 1 μl protein and 1μl reservoir solution. This procedure yielded some form of crystal in four different conditions (Table 3.4) at both temperatures. (The absence of operational photographic equipment in the laboratory at this stage meant that photographs of these promising crystals could not be taken.)

The remaining conditions yielded various forms of precipitate or failed to show any visible change.

Table 3.4 Initial screen conditions yielding crystals of NHase. HS1: Hampton Screen 1 (Hampton Research), JB1: JB Screen 1 (JenaBiosciences).

Screen (condition number)	Composition	Observation
HS1 (9)	100mM sodium citrate (pH 5.6), 30% PEG 4000, 200mM ammonium citrate	Thin needles
HS1 (15)	100 mM sodium cacodylate (pH 6.5), 30% PEG 8000, 200mM ammonium sulphate	Thin needles
HS1 (38)	100mM HEPES (pH 7.5), 1.4M sodium citrate	Small three-dimensional crystals
JB1 (B4)	100mM MES (pH 6.5), 30% PEG 400, 100mM magnesium chloride	Small three-dimensional crystals

3.3.3 Optimisation of crystallisation

Of the four conditions that yielded crystals, two were selected for further optimisation experiments. HS1 (38) and JB1 (B4) were selected on the basis of the promising (aesthetic) form of the crystals at this early stage of crystallisation trials. A finer systematic grid screen was designed around each condition. The composition of each grid screen used is shown in Figure 3.11.

A

[NaCit] / **pH**	1 M	1. 1M	1.2M	1.3 M	1.4 M	1.5 M
6.5						
7.0						
7.5						
8.0						

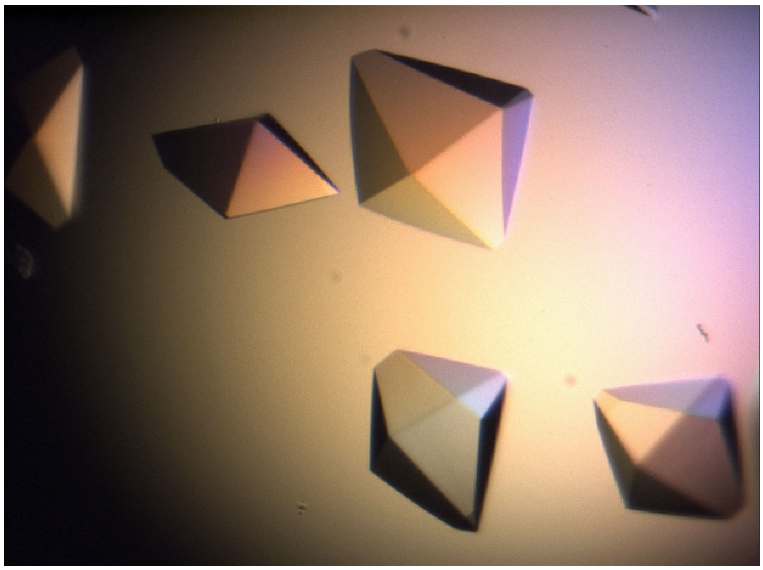
B

[PEG400] / **pH**	15%	20%	22.5%	25%	27.5%	30%
6.5						
7.0						
7.5						
8.0						

Figure 3.11 Composition and set-up of fine optimisation screen around the conditions that yielded promising crystals. A: Grid screen around HS1 (38), B: Grid screen around JB1 (B4).

Each of these screens was repeated with three different protein concentrations: 10 mg/ml, 20 mg/ml and 40 mg/ml at 20˚C. A 4µl drop volume was used (2µl reservoir: 2µl protein) in an attempt to encourage growth of larger crystals. Through this optimisation procedure, crystals of very good form and size were produced (Figure 3.12).

A



B

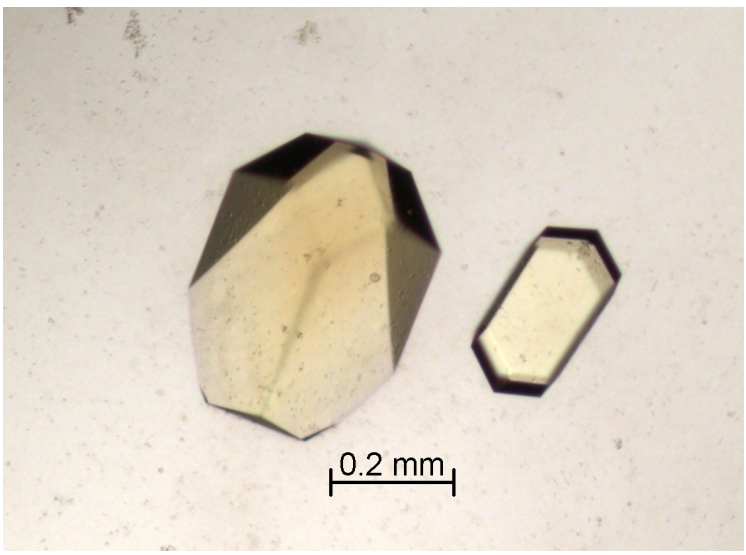


Figure 3.12 Crystals of NHase produced in final crystallisation optimisation experiments. A: 1.1M sodium citrate, 100mM HEPES, pH 7.5 and B: 30% PEG 400, 100mM MgCl, 100mM MES, pH 6.5.

Crystals of different form/external morphology were produced in each of the conditions. Table 3.5 shows the exact conditions and crystal forms resulting from the successful crystallisation optimisation experiments.

Table 3.5 Observations made during the final crystallisation optimisation experiments

Conditions	Approximate size of largest crystals (mm)	Time to reaching maximum dimensions
100mM HEPES (pH 7.5), 1.1M sodium citrate, 20 mg/ml NHase	0.1 x 0.1 x 0.2	~7 days
100mM MES (pH 6.5), 30% PEG 400, 100mM magnesium chloride, 40 mg/ml NHase	0.2 x 0.2 x 0.3	~ 4 weeks

3.4 X-ray diffraction data

3.4.1 Preliminary diffraction experiments

Crystals produced in the presence of PEG 400 as described above were selected for structural studies. The larger size of crystal produced in these conditions encouraged this choice.

In the initial diffraction test, a single, large crystal was removed from the hanging drop using a cryoloop (Teng, 1990) of appropriate size. The crystal was then mounted on the X-ray diffraction apparatus at the in-house source at the University of the Western Cape (see section 2.11) and exposed to the beam for 10 minutes. It was found that these crystals diffracted strongly with spots visible beyond 2.4 Å (Figure 3.13).

Several crystals were tested for diffraction. Visual examination of the resulting diffraction images showed that most of the crystals diffracted to a similar resolution with equally well-defined spots (as an initial, crude estimation of relative mosaicity). A diffraction image collected at 90^0 orientation demonstrated similar results.

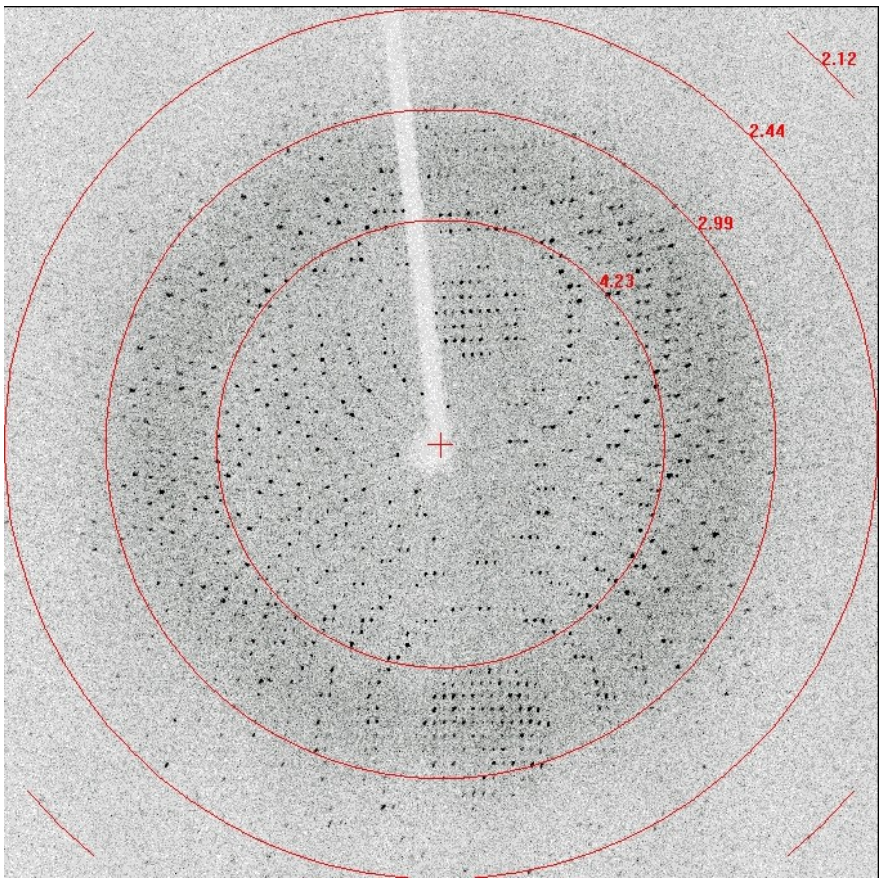


Figure 3.13 X-ray diffraction image from NHase crystals. Spots were visible to beyond 2.4 Å. Resolution circle values in Å.

3.4.2 X-ray diffraction data collection

The crystal finally selected for collection of a complete data-set had dimensions of approximately 0.2mm x 0.1mm x 0.1mm. Diffraction data were collected under cryoconditions (at a temperature of 100K). The use of cryoconditions has many advantages including decreased damage to the crystal on prolonged exposure to X-rays and the potential for better quality data due to decreased thermal motion (or disorder) of the molecules within the crystal (Garman, 1999). The use of cryoprotectants was also tested. A small amount of the mother liquor was collected with a cryoloop (Teng, 1990), placed in the path of the cryostream and monitored. This procedure indicated that the high percentage (30%) of PEG 400 present in the crystallisation mother liquor was sufficient to prevent ice formation. This is consistent with the diffuse shadow (the scattering ring) seen on a typical diffraction image (see Figure 3.13). The equally shallow slopes on the high and low-resolution ends of the image are characteristic of a well-cryoprotected solution (Garman and Mitchell, 1996).

The first diffraction image collected was used to estimate the unit cell dimensions and space group of the crystal *via* the data-collection and processing program CrystalClear (Pflugrath, 1999). The crystal was preliminarily found to belong to the primitive tetragonal system (P4) with unit cell dimensions a = 106.61Å, b = 106.61Å, c = 83.23Å, alpha = beta = gamma = 90^0. This information was used by CrystalClear to devise a data collection strategy that would enable collection of a complete set of

unique reflections. While this suggested strategy was taken into account, generous availability of beam-time at the University of the Western Cape in-house X-ray source provided the possibility of collecting ample data to allow for maximal completeness. A total of 319 images, each covering an oscillation angle of 0.5^0 were collected at a crystal-to-detector distance of 160mm. These data were processed using Crystal Clear, and HKL 2000 (incorporating Xdisp, Denzo and Scalepack; Minor *et al*., 2002).

A summary of the data-collection statistics and other crystallographic information is shown in Table 3.6. Based on the estimated molecular weight of the NHase heterodimer (51 kDa), the calculated unit cell dimensions would allow a single heterodimer per asymmetric unit with 46.5% solvent content and a Matthews coefficient of 2.3 $Å^3$ Da^{-1} (Matthews, 1968).

Table 3.6 X-ray data collection statistics and crystallographic information.

Data Collection Statistics	
Wavelength	1.54 Å
Space Group	$P4_12_12$ or $P4_32_12$
Cell dimensions	a = b = 106.61 Å, c = 83.23 Å
Cell angles	alpha = beta = gamma = 90^0
Resolution range	20 - 2.5 Å
Number of unique reflections	16152
Completeness	93 %
Last Shell (2.59 – 2.52)	77 %
R_{merge} (Last Shell)	0.091 (0.391)

These statistics indicated that the data set obtained was of adequate quality for progression to structure solution. The ambiguity in space group is discussed in Section 4.1. It may be worthwhile to note that these spacegroups do not appear in any of the previously crystallised NHases.

Chapter 4: Crystal structure of *Bacillus* sp. RAPc8 Nitrile Hydratase

4.1 Solution of the crystal structure

Once soluble, pure target protein is produced, the single most significant obstacle for the crystallographer is the 'phase problem' that results from the fact that phase information is not directly obtainable from the crystallographic experiment. However, several NHase structures have been previously determined using X-ray crystallographic methods. Availability of these structures allowed determination/estimation of initial phases by the now widely used method of molecular replacement. The first step was to identify the best possible search model out of the three NHase structures whose co-ordinates were available for use at the time (*Rhodococcus* sp. R312, PDB accession number 1AHJ/2AHJ; (Huang *et al*., 1997) and *Ps. thermophila,* PDB accession number 1IRE; (Miyanaga *et al*., 2001)).

The amino acid sequence of *Bacillus* sp. RAPc8 NHase was submitted to the Fugue server at the University of Cambridge (Shi *et al*., 2001), which performed a sequence-structure comparison with all proteins within HOMSTRAD (Mizukuchi *et al*., 1998) and identified structural homologues. The results of this procedure indicated that *Bacillus* sp. RAPc8 NHase was most likely to fold in a form most closely resembling that of the NHase from *P. thermophila* (PDB accession number: 1IRE). The Z-scores achieved were 42.75 for the α subunit and 35.15 for the β subunit indicating a very high likelihood. This structure was thus used as a search model for molecular replacement using the program EPMR (Kissinger *et al*., 1999).

At this stage, the space group of the NHase crystals from which crystallographic data were collected could not be unambiguously assigned (see Table 3.6). The crystals were found to belong to either $P4_1212$ or $P4_3212$. It was thus necessary to carry out the molecular replacement procedure with data processed in both space groups to remove the ambiguity. In the respective outputs from EPMR, $P4_1212$ gave a correlation coefficient of 35% and an R-factor of 56.2% while $P4_3212$ gave a correlation of 13% and an R-factor of 60%. The better statistics achieved with the $P4_1212$ data-set indicated that the crystal is more likely to belong to this space group. The solution produced in $P4_1212$ was therefore used in subsequent procedures.

4.2 Refinement and validation of the structure

Inspection of the electron density map from this first solution showed regions that were well resolved. This was particularly the case for regions that had the highest homology to the structure used as a search probe. Use of SCWRL (http://www1.jcsg.org/scripts/prod/scwrl/serve.cgi) for side-chain substitution resulted in a very significant improvement in phase information. The resulting map revealed density for regions that were invisible in the first solution. This rapid and convenient procedure resulted in a model with an R-factor of 29% and an R_{free} of 37%.

Subsequent model rebuilding was done using O (Kleywegt and Jones, 1997) and refinement was done using Refmac 5 from the CCP4 suite of programs (CCP4, 1994). These procedures were carried out by Dr Ozlem Tastan-Bishop (under supervision of Dr Muhammed F Sayed, Department of Biotechnology, University of the Western

Cape) and Table 4.1 presents the refinement statistics reported for the model handed over for analysis in this study.

Table 4.1 Refinement statistics

R-factor	20 %
R_{free}	27 %
Mean B-factor	45.82
Main chain	44.77
Side chain atoms and water molecules	46.73
Number of water molecules	210
Number of cobalt atoms	1

The model-rebuilding process is prone to a certain degree of subjectivity on the part of the crystallographer. Before the last reported refinement model described above was analysed in the context of the known functional properties of *Bacillus* sp. RAPc8 NHase, it was essential to gain full confidence in its reliability.

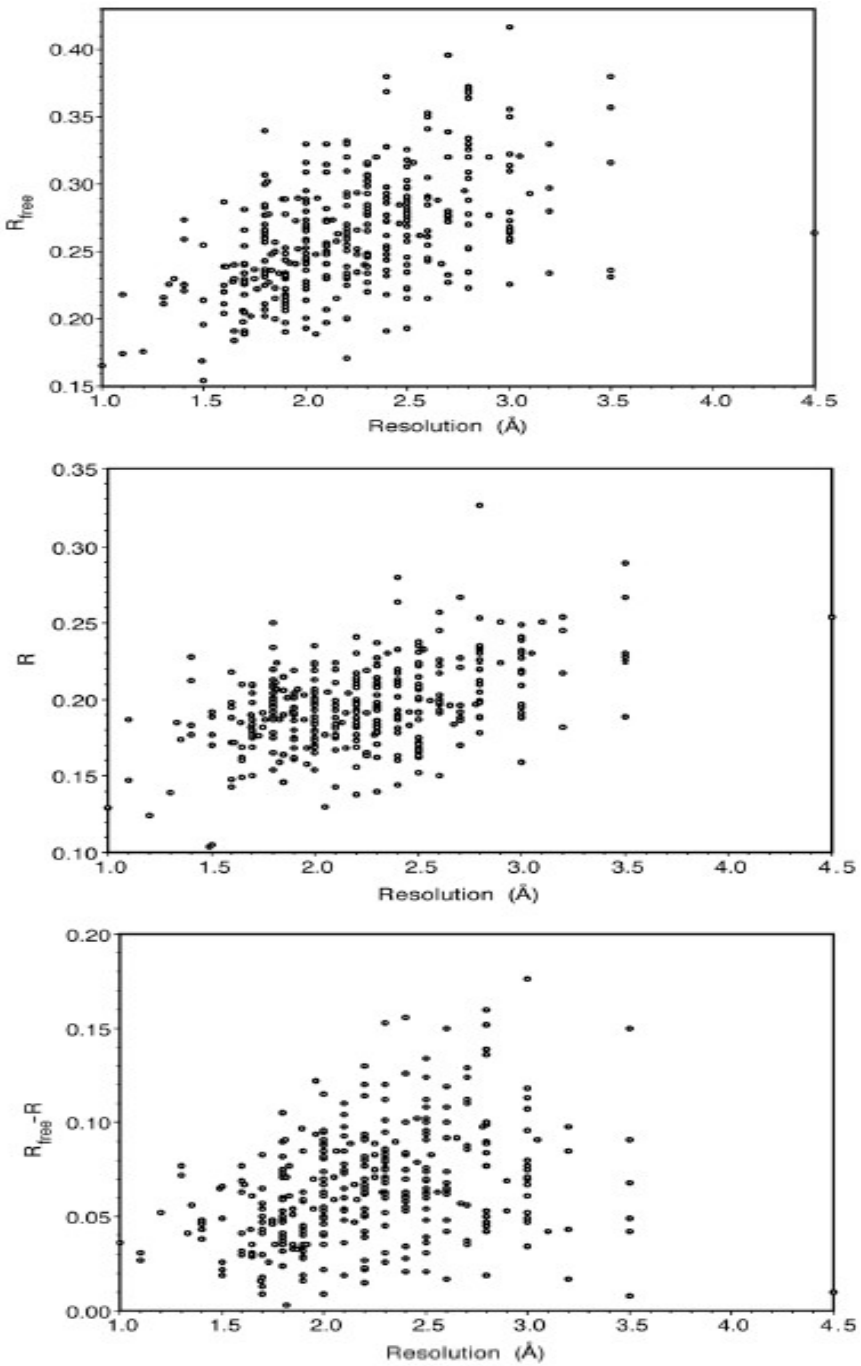


Figure 4.1 Plots of R_{free}, R-factor and R_{free} – R-factor as a function of resolution for structures in the Protein Data Bank, May 1996 release. (Taken from Kleywegt and Brunger, 1996)

This validation procedure (eloquently referred to as "checking your imagination" by Kleywegt and Brunger (1996)) is an essential part of any structure determination project.

The refinement statistics (particularly the R-factor and the R_{free}) reported above were scrutinised against known limits for a crystal structure of the resolution reported here (2.5Å) in mind. Figure 4.1 shows the range within which structures deposited in the Protein Data Bank typically fall (Kleywegt and Brunger, 1996) and shows that the R_{free}, R-factor and R free – R-factor achieved in this study (7%) are within these limits.

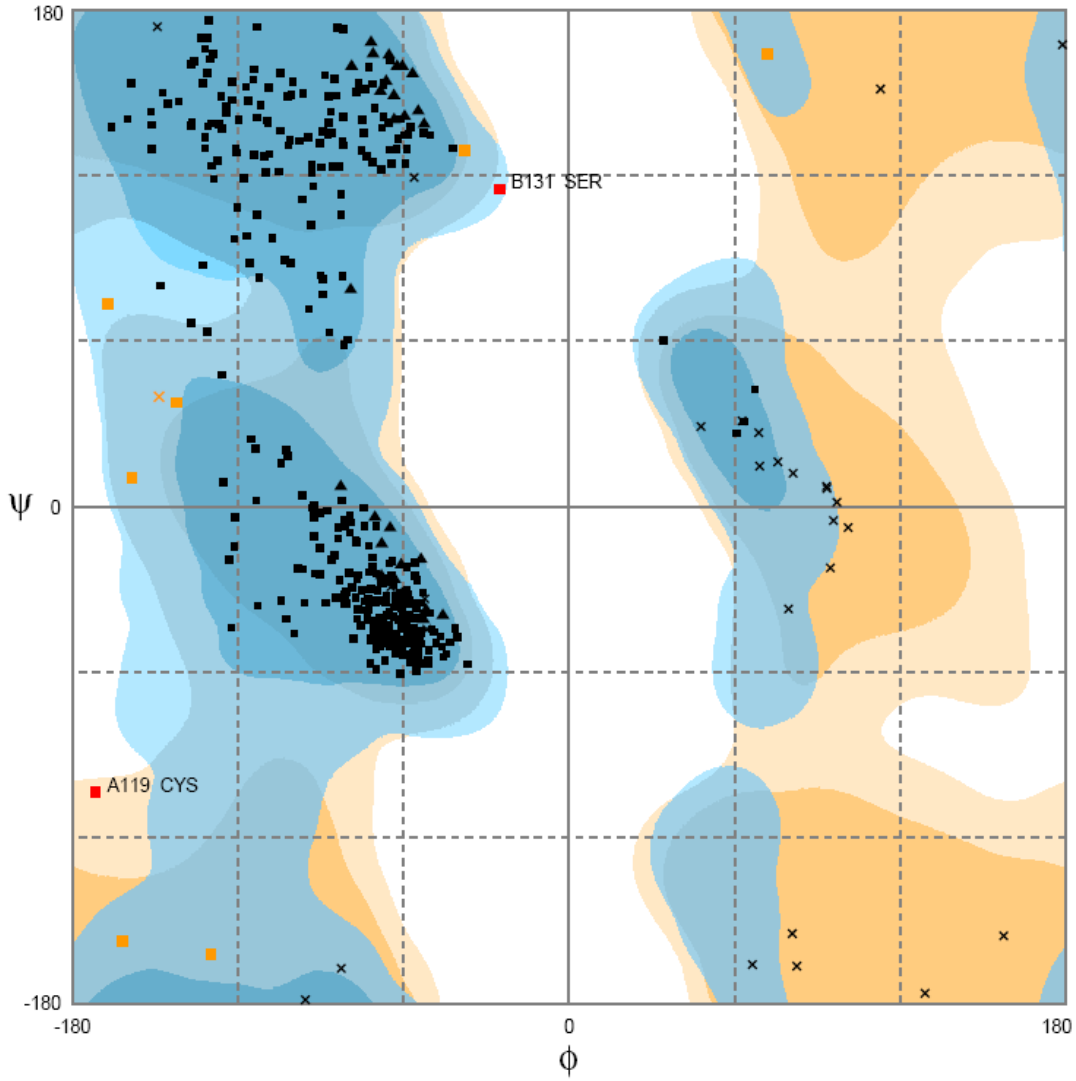
This good R-factor value (20%) indicated that there was a reasonably good correlation between the original crystallographic data and the final NHase structure model. The reasonable R_{free} (27%) further demonstrated that the model produced could accurately predict the diffraction data for data that would have been excluded from the refinement process (the test set).

Whereas the above analysis relied on a cross-validation technique, the remainder of the structure validation procedures conducted in this study were carried out with computational tools that are all based on prior knowledge of correct features or properties of high-resolution X-ray structures of previously solved macromolecules and small molecules.

First and arguably most importantly, the geometric soundness of the final structure model was analysed with PROCHECK (Laskowski *et al*., 1993) and RAMPAGE

(Lovell *et al*., 2003). Both of these protocols were used because they define favoured and disallowed regions for phi (ϕ) and psi (ψ) angles along the Cα chain of the polypeptide in slightly different ways. The stringency of the methods also differs. In the resulting outputs (Figure 4.2), the NHase structure was found to have satisfactory geometry. Interestingly, among the very few problematic amino-acid residues found in both outputs was the post-translationally modified cysteine 119 from the α-subunit (see section 4.3). Inspection of the electron density map showed that this region had a perfect fit to the electron density. The equivalent modified residue in the structure of *Rhodococcus* sp. R312 NHase was also found to be problematic in a cursory geometric analysis of this structure. It is therefore possible that the conformation of the metal-binding claw-setting motif unique to NHases forces ψ and ϕ angles of amino-acids forming this motif into rare or strained conformations.

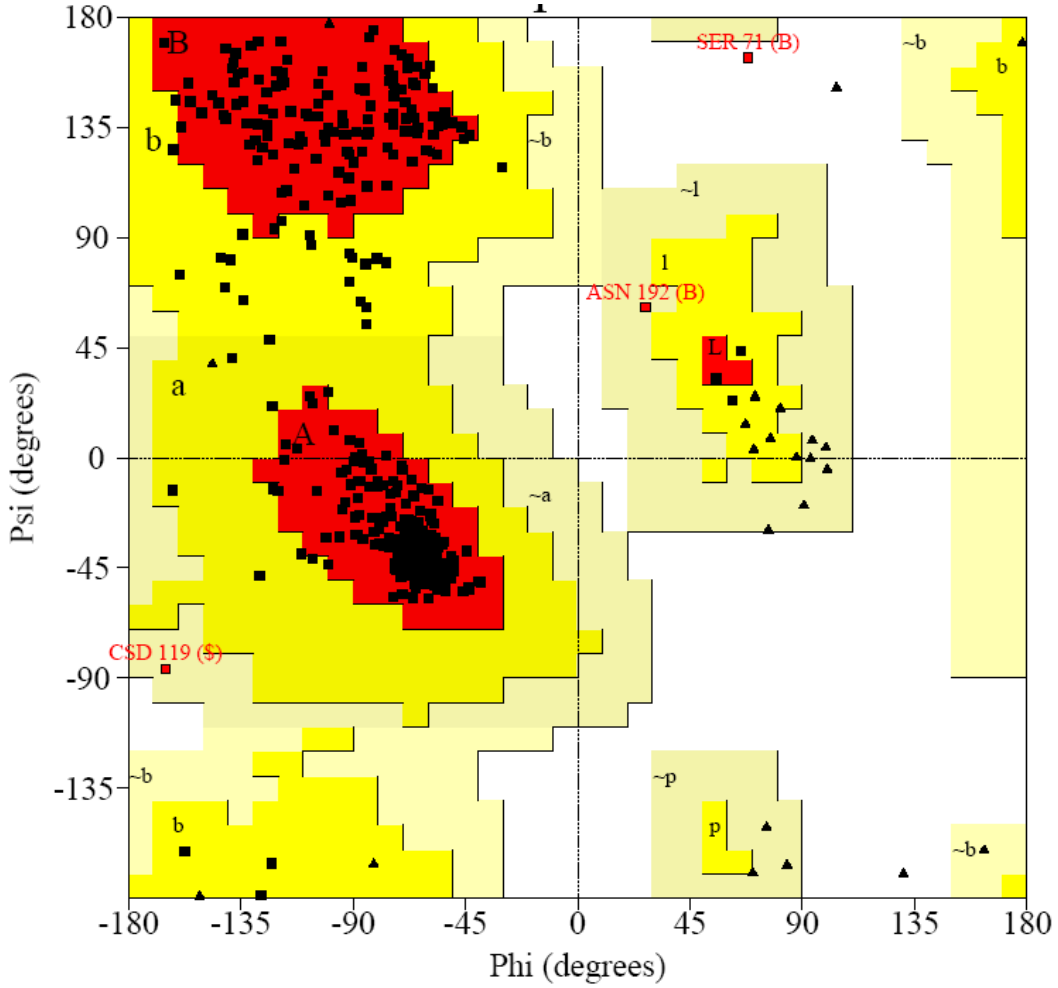
As a final assessment of the geometric integrity of the final refinement model, average bond lengths and bond angles within the structure were analysed. The r.m.s.d. (root mean square deviation) from ideal was found to be 2^0 for bond angles and 0.02 Å for bond lengths. Accepted deviations within crystal structures of macromolecules are 4^0 and 0.02 Å for bond angles and bond lengths respectively (Laskowski *et al*., 1993b).



Number of residues in favoured region (~98.0% expected)	: 415 (97.6%)
Number of residues in allowed region (~2.0% expected)	: 8 (1.9%)
Number of residues in outlier region	: 2 (0.5%)

Figure 4.2A Ramachandran plot of NHase structure output from RAMPAGE.

(continues next page)



Plot statistics

Residues in most favoured regions [A,B,L]	330	89.7%
Residues in additional allowed regions [a,b,l,p]	35	9.5%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.5%
Residues in disallowed regions	1	0.3%
Number of non-glycine and non-proline residues	368	100.0%
Number of end-residues (excl. Gly and Pro)	207	
Number of glycine residues (shown as triangles)	30	
Number of proline residues	28	
Total number of residues	633	

Figure 4.2B Ramachandran plot of NHase output from PROCHECK.

4.3 Description of the structure

4.3.1 The αβ heterodimer

The α subunit of *Bacillus* sp. RAPc8 has a protruding N-terminal arm comprising three α-helices separated by a loop from a large globular domain. This globular domain consists of nine α-helices and four β-strands (Figure 4.3 and Figure 4.5A).

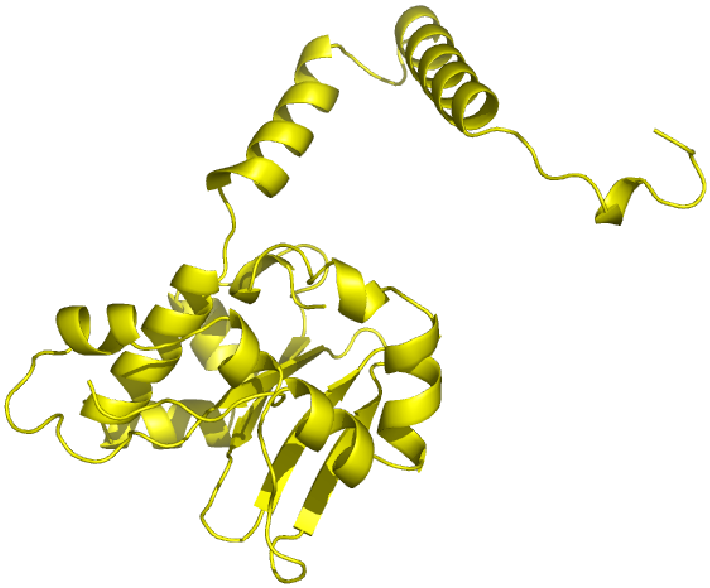


Figure 4.3 Cartoon representation of the α subunit.

The β subunit has two distinct domains. The N-terminus begins with a long loop interrupted by a very short α-helix. This loop leads to a domain that comprises seven α-helices (the α-helix domain). The second domain is comprised of two small α-helices, two small β-strands and a prominent anti-parallel β-sheet formed by four

strands (the β-sheet domain) (Figure 4.4 and Figure 4.5 B). Table 4.2 describes detailed description of the topology of the α and β subunits of NHase.

In the heterodimer (Figure 4.6) the long N-terminal arm of the α subunit wraps around the α-helix domain of the β subunit. The globular domain of the α subunit is nestled in the groove-like region between the α-helix domain and the β-sheet domain of the β subunit. This globular domain interacts with amino acids from both domains of the β subunit. The long N-terminal loop of the β subunit wraps around this globular domain.

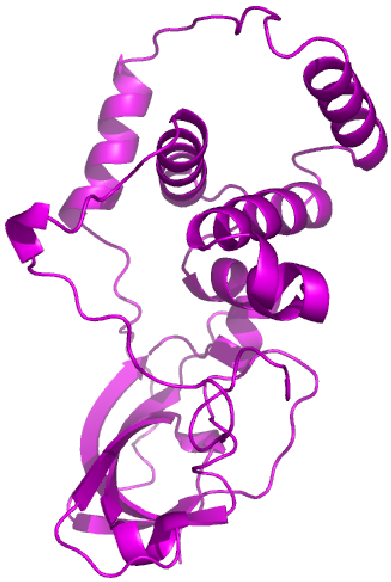


Figure 4.4 Cartoon representation of the β subunit.

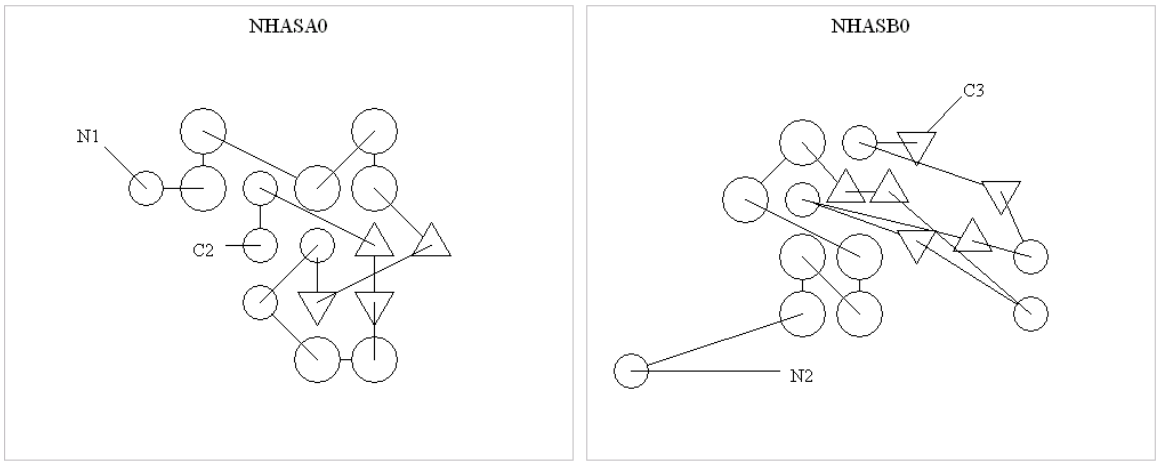


Figure 4.5 Schematic representation of topology of NHase α and β subunits. Circles represent α-helices while triangles represent β-strands. The figure was generated using TOPS (Michalopoulos *et al*., 2004).

These relatively intricate interactions between the two subunits are consistent with the view that dimer formation in NHases is a dynamic process rather than a simple docking process (Nagashima *et al*., 1998) and may provide support for the hypothesis that the P14K protein found in the *Bacillus* sp. RAPc8 NHase operon may be involved in assembly of the dimeric structure. Further work aimed at elucidating the precise function of P14K in *Bacillus* sp. RAPc8 is in progress (Cameron RA, unpublished results) and will be reported elsewhere.

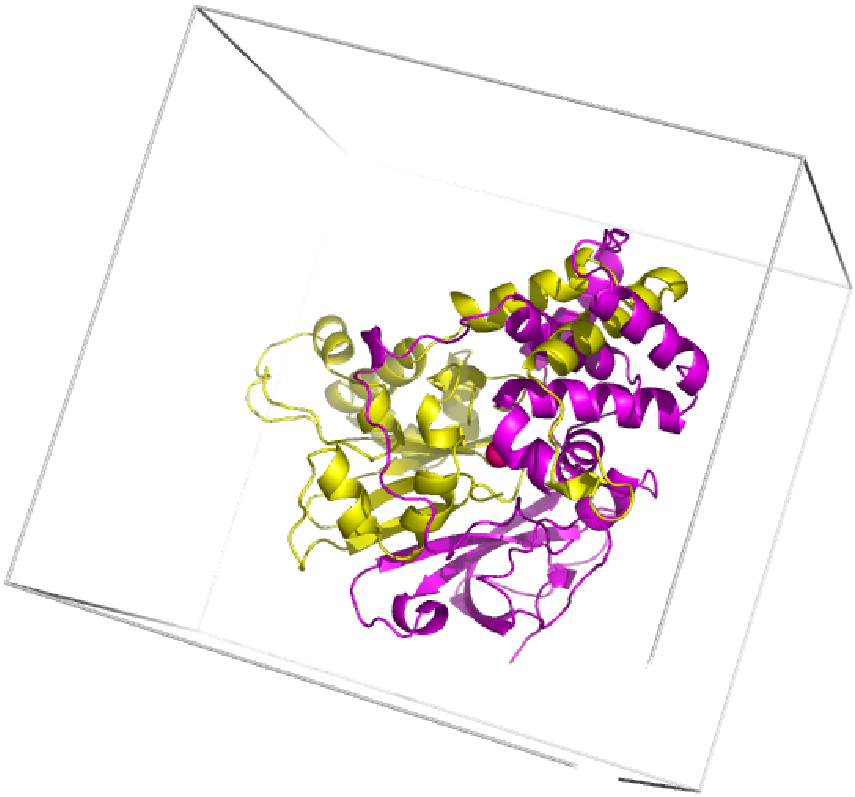


Figure 4.6 Cartoon representation of the αβ heterodimer of *Bacillus* sp. RAPc8 NHase. The heterodimer was found within each asymmetric unit. Unit cell axes are shown with white lines. (β subunit in purple, α subunit in yellow, cobalt ion – red sphere.)

Table 4.2 Detailed description of NHase topology. The amino-acid regions representing each secondary structural element were deduced from the TOPS output (depicted in Figure 4.5).

α subunit		**β subunit**	
Secondary structure	**Residues**	**Secondary structure**	**Residues**
α helix	15-17	α helix	21-23
α helix	22-38	α helix	30-43
α helix	44-55	α helix	52-60
α helix	61-72	α helix	64-69
α helix	74-82	α helix	72-86
α helix	84-90	α helix	92-104
β sheet	101-105	α helix	116-128
β sheet	110-115	β sheet	133-134
α helix	124-127	β sheet	147-150
α helix	132-135	α helix	163-165
α helix	137-142	β sheet	169-179
α helix	147-153	α helix	182-185
β sheet	163-168	β sheet	194-202
β sheet	174-178	α helix	203-207
α helix	190-194	β sheet	213-220
α helix	200-226	α helix	221-223
		β sheet	224-226

4.3.2 The $\alpha_2\beta_2$ heterotetramer

Characterisation of *Bacillus* sp. RAPc8 NHase revealed that the native protein existed in the form of an $\alpha_2\beta_2$ heterotetramer (Pereira *et al*., 1998). These experiments, including gel filtration and native gel electrophoresis suggested that this heterotetrameric form represents the sole *in vivo* functional unit. Although the crystal structure described here indicates that the αβ heterodimer exists within each asymmetric unit (Figure 4.6), it was nonetheless important to inspect the crystal structure for the presence of interactions that may have formed a heterotetramer. This would facilitate analysis of the structure in the context of its biological situation.

Symmetry-related elements within 10Å of the core molecule were generated and displayed using the *symexp* command in PyMOL (DeLano Scientific LLC). Inspection of the resulting output revealed an interaction between two αβ heterodimers in very close proximity to one another (Figure 4.7).

The αβ heterodimers interact along a two-fold crystallographic axis. On visual inspection, the interaction was apparently loose and was formed mainly by amino acids of the β subunit of each dimer. A precise description of the interactions forming this interface (Table 4.3) was produced using the Protein-Protein interface server (http://www.biochem.ucl.ac.uk/bsm/PP/server/). .

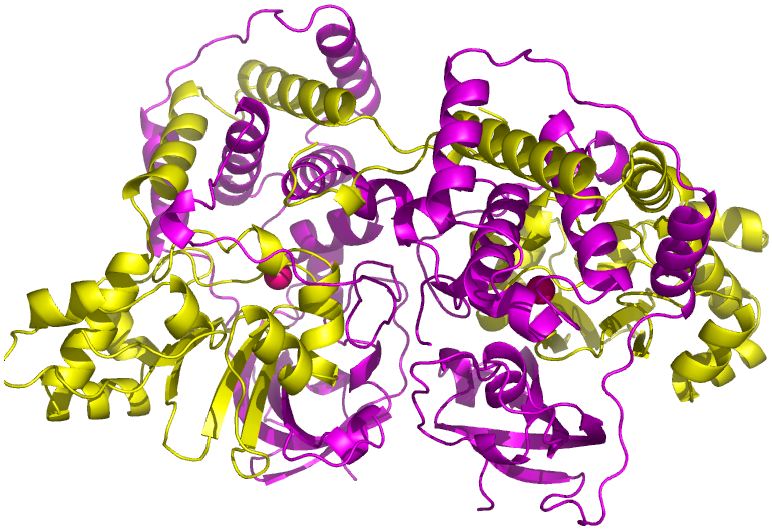


Figure 4.7 Cartoon representation of the $\alpha_2\beta_2$ heterotetramer. The interface between two heterodimers is formed primarily by amino acids of the β subunit (magenta).

Table 4.3 Detailed description of the interface between cognate dimers in the heterotetrameric structure of *Bacillus* sp. RAPc8.

Protein Interface Parameter	Value
Interface Accessible Surface Area	1827.32 $Å^2$
% Interface Accessible Surface Area	9.90
Planarity	5.18
Length & Breadth	41.47 & 42.18
Length/Breadth Ratio	0.89
Interface Residue Segments	7
% Polar Atoms in Interface	38.18
% Non-Polar Atoms in Interface	61.80
Secondary Structure	Alpha
Hydrogen Bonds	10
Salt Bridges	0
Disulphide Bonds	0
Gap Volume	34.99 Å
Bridging Water Molecules	7

4.3.3 The active site region of NHase

4.3.3.1 The metal-binding centre

Prior knowledge of the requirement of *Bacillus* sp. RAPc8 NHase for cobalt chloride in the medium for optimal activity meant that it was anticipated from the onset that the crystal structure would show evidence of cobalt in the enzyme. Inspection of the difference electron density map during the late stages of refinement showed clear spherical electron density in a cavity at the interface between the α and β subunits of the structure. A cobalt atom with an occupancy of 1 was built into this density (Tastan-Bishop O., personal communication). This finding in itself was not conclusive evidence that *all* heterodimers contain one cobalt atom. Further occupancy refinement and analysis with ICP (inductive coupled plasma) will give definitive verification.

The cobalt-binding site was a typical *claw-setting motif* unique to NHases as described by Huang *et al*. (1997). The co-ordination to the cobalt was octahedral with amino acids αCys119, αSer120, αCys121 forming ligands to the cobalt atom on vertices of the octahedron. All NHase crystal structures reported to date except that of the *B. smithii* NHase (Hourai *et al*., 2003) indicated that two of the cysteine side chains were post-translationally modified. Inspection of the electron density in this region demonstrated that, consistent with the bulk of previous reports, αCys119 and αCys121 in *Bacillus* sp. RAPc8 were oxidized to cysteine sulfinic and cysteine sulfenic acid respectively (Figure 4.8). Presumably, during recombinant expression in *E. coli* this modification was achieved spontaneously *via* an autocatalytic mechanism or was effected in some way by P14K.

The exceptional finding of Hourai *et al*. (2004) that the NHase of *B. smithii* did not exhibit post-translational modification of cysteine residues is surprising, especially in view of the results of the current study. These two enzymes share very high sequence homology (88 % amino acid identity in the α subunit and 85 % amino acid identity in the β subunit) and would be expected to present similar molecular properties. Further, it is shown (section 4.3.4) that the overall crystal structures of these enzymes are very similar. Since these modifications are thought to be essential for activity of NHase (Nojiri *et al*., 1999), the apparent lack of modification in the *B. smithii* NHase may have been due to errors or omissions during solution of this structure or during the refinement process. The suggestion that the failure to identify the post-translational modification was due to poor resolution is unlikely as *Bacillus* sp. RAPc8 NHase was solved to the same resolution.

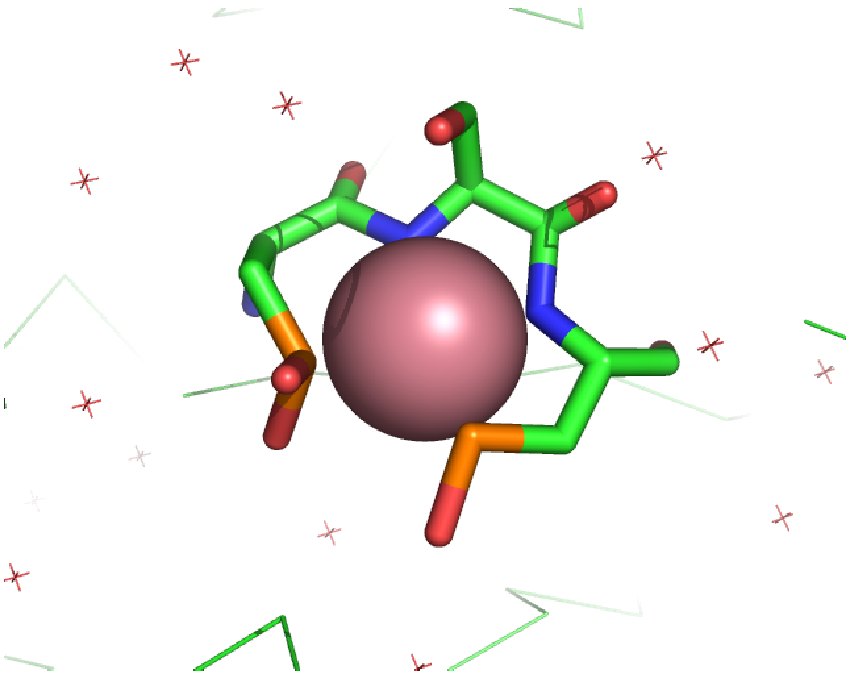


Figure 4.8 Metal Binding Centre showing 'claw-setting motif'.

4.3.3.2 The substrate channel

During the analysis of biologically relevant features in *Bacillus* sp. RAPc8 NHase, the molecular surface of the structure was rendered and visualized in PyMOL (DeLano Scientific). This procedure was deliberately carried out with the molecule in the 'dry state' (i.e., all water molecules removed) so that the surface observed is that contributed only by amino acid residues.

Inspection of this surface representation revealed an extended, curved solvent accessible channel with access to bulk solvent from two locations in the heterodimer (Figure 4.9). This channel lies at the interface between the α and β subunits of NHase and traverses past the metal binding site. This channel was putatively identified as the substrate entry channel to the active site. While this procedure gave the possible structure of the substrate channel, it was not possible to produce a more precise description from the images created using standard molecular graphics software.

Detailed analysis of cavities and/or channels within the structure was thus carried out using CASTp (Liang *et al*., 1998). An initial search for pockets was carried out with a probe radius of 1.5 Å (a water molecule has a molecular radius of approximately 1.4 Å). Using this approach, a total of 57 pockets were identified. Many were ignored for the purposes of this analysis (e.g., those that were on the surface of the structure and those that did not intrude deep into the interior of the protein). Among the larger pockets found, pocket 55 and 56 appeared to overlap with the region identified in Figure 4.9. These two pockets together (Figure 4.10) form "an extended, curved solvent accessible channel with access to bulk solvent from two locations in the heterodimer" as described earlier in this chapter.

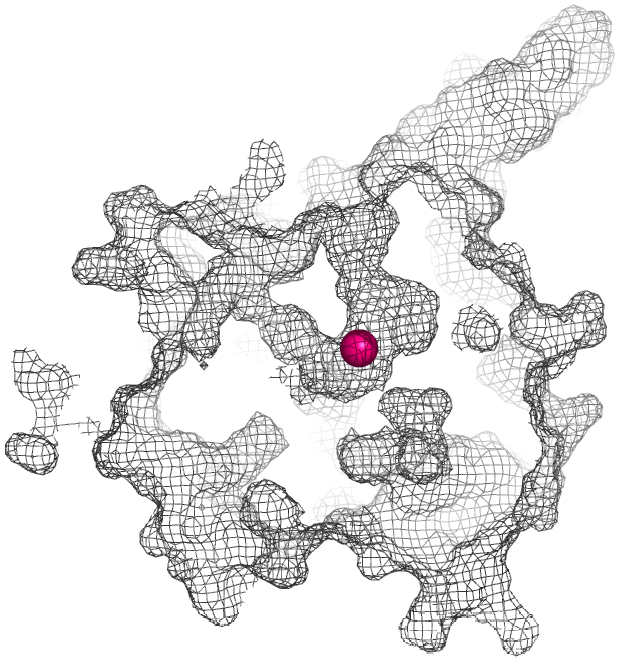


Figure 4.9 Mesh representation of surface in dry state (with waters removed) clipped to show internal cavities. A continuous channel traversing past the cobalt binding site/active site can be seen.

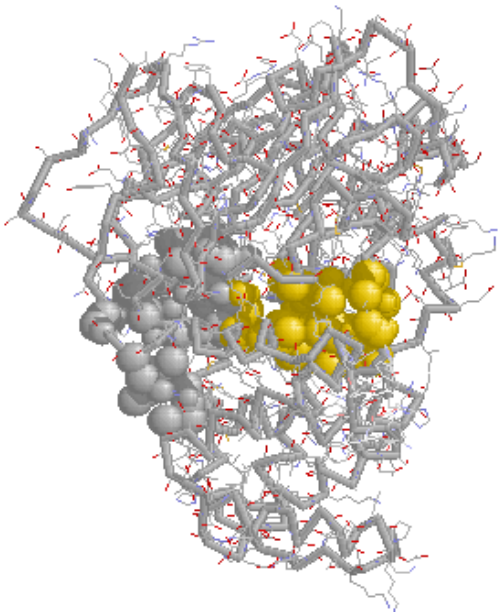


Figure 4.10 Two major pockets/cavities within *Bacillus* sp. RAPc8 NHase identified using a probe radius of 1.5 Å. These pockets overlap with each other to form an extended channel that traverses the cobalt centre. Yellow spheres: pocket 55, grey spheres: pocket 56. (Figure generated using CASTp (Liang *et al*., 1998) and Chime.)

Geometric dimensions of these cavities and the amino acid residues that contribute to the lining of the channel are listed in Table 4.4.

Table 4.4 Geometric dimensions and amino acid residues that contributed atoms to the lining of channel.

Pocket	Volume/Å^3	Inner surface area/Å^2	Amino acids
55	481	542	αGLN97, αLEU118, αCSD119, αSER120, αCEA121, αTYR122, αTRP124, αLEU129, αPRO130, αLYS135, αGLU136, αPRO137, αARG140 βHIS5, βPHE36, βPHE52, βPHE55, βARG56, βILE59, βTYR67, βLEU68, βSER71, βTYR72, βTYR73,βHIS75, βTRP76
56	494	506	αGLN97, αGLU99, αCEA121, αSER171, αGLU172 βGLY46, βGLY48, βMET49, βLYS50, βALA51, βPHE52, βASP53, βGLU54, βARG56, βTHR87, βLEU130, βPRO132, βARG134, βARG160, βVAL180, βPRO182, βALA185, βARG188, βGLY190, βGLU191, βTRP220

This complete description of the channel clearly localised it at the interface between the α and β subunits. The presence of the modified cysteine residues (that form the cobalt-binding claw-setting motif at the active site) among the residues that are

identified as part of the channel provides some confidence in the designation as the substrate channel. Further structural and functional analyses that focus on this substrate channel in the context of the substrate specificity of *Bacillus* sp. RAPc8 are described and discussed in Chapter 5.

4.3.4 Comparison with other NHase structures

Several other NHase structures have been determined experimentally. All these structures exhibit a conserved overall fold. The structure of *Bacillus* sp. RAPc8 as described here clearly exhibited overall features found in these other NHases. It was necessary however, to analyse these similarities in a more quantitative manner.

Structure alignments of previously solved NHase structures with that of *Bacillus* sp. RAPc8 NHase were produced. As expected, all NHase structures aligned very well and showed little deviation from that of *Bacillus* sp. RAPc8 NHase (Figure 4.11). The metal binding motif was essentially identical in all the NHases, including that of the iron-type *Rhodococcus* sp. R312 NHase. This observation is consistent with a study that demonstrated that *Rhodococcus* sp. N-771 iron type NHase was able to bind cobalt in place of iron. This cobalt-substituted iron-type NHase exhibited many-fold higher activity than enzyme produced in the complete absence of metal (Nojiri *et al*., 2000).

A detailed interpretation of the graphic output shown in Figure 4.11 is presented in Table 4.5. The relative similarities (r.m.s.d's) of these structures were consistent with the relative amino acid sequence identities/similarities previously observed (i.e., even in amino-acid sequence, *Rhodococcus* sp. R312 was the least similar while *Bacillus smithii* was the most similar).

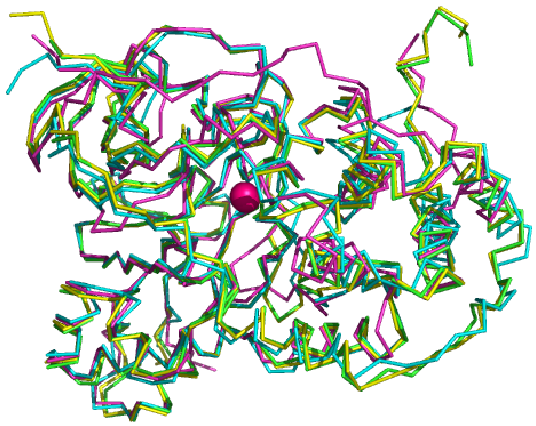


Figure 4.11 Ribbon representation of superimposition of NHase crystal structures available in the Protein Data Bank with *Bacillus* sp. RAPc8 NHase. Green: *Bacillus* sp. RAPc8, magenta: *Rhodococcus* sp. R312, blue: *Pseudonocardia thermophila*, green: *Bacillus smithii* .

Table 4.5 Detailed description of structure alignment of NHases in the PDB with *Bacillus* sp. RAPc8 NHase. (*r.m.s.d, root mean square deviation)

Source of NHase	PDB accession no.	No. of Cα atoms superimposed	r.m.s.d*
Rhodococcus sp. R312	1ahj	336	1.191
Pseudonocardia thermophila	1ire	377	1.066
Bacillus smithii	1v29	399	0.355

The higher r.m.s.d of the *Rhodococcus* NHase structure was also consistent with the fact that of the four enzymes compared, this enzyme alone is classified as an iron-type NHase. This enzyme also has the lowest thermostability of the group. Apart from differences in the loops and termini, the single most prominent difference between

this NHase and the more thermostable *Bacillus* sp. RAPc8 NHase was the presence of an additional α helix comprising amino acids 116-126 of the β subunit of *Bacillus* sp. RAPc8 NHase (Figure 4.12).

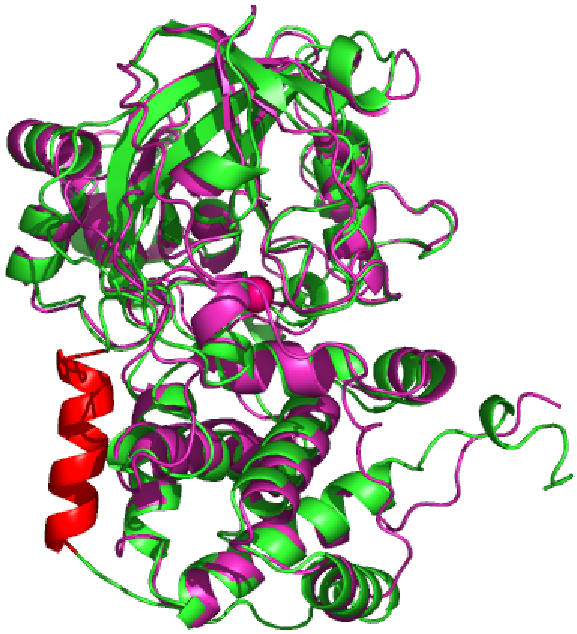


Figure 4.12 Cartoon representation of structure alignment between *Rhodococcus* sp. R312 and *Bacillus* sp. RAPc8 NHase crystal structures. Green: *Bacillus* sp. RAPc8, magenta: *Rhodococcus* sp. R312. The extra helix insertion in *Bacillus* sp. RAPc8 NHase is coloured in red.

Increased helical content in proteins has been linked with an increase in thermostability (Querol *et al*., 1996; Kumar *et al*., 2000). This additional helix in *Bacillus* sp. RAPc8 NHase may thus contribute to the additional thermostability of this enzyme. In the heterotetramer structure, this helix was involved in an additional interaction between cognate dimers. This interaction may further contribute to thermostability.

Reference List

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.

Miyanaga,A., Fushinobu,S., Ito,K., and Wakagi,T. (2001) Crystal structure of cobalt-containing nitrile hydratase. *Biochemical and Biophysical Research Communications* **288:** 1169-1174.

Chapter 5: Engineering aromatic substrate specificity in *Bacillus* sp. RAPc8 NHase

5.1 Background and strategy

Desirable features for NHases destined for use in industrial biotransformations include the lack of product inhibition, increased stability under harsh conditions and high activity on a wide range of substrates. The issue of substrate specificity was selected for investigation in this study.

Bacillus sp. RAPc8 NHase has been shown to have broad substrate specificity, converting nitrile substrates of various groups including linear and branched aliphatics, heteroaromatics and di-nitriles (see Tables 1.4 and 1.5). However, the enzyme does not catalyse conversion of homoaromatic nitriles such as benzonitrile and benzyl cyanide (Cameron, 2002; Pereira *et al*., 1998).

In this chapter, work aimed at genetically engineering aromatic substrate specificity into *Bacillus* sp. RAPc8 NHase is described and discussed. The great potential of aromatic products of NHase biotransformations as pharmaceuticals has been highlighted in section 1.5.1 of this thesis. A NHase that has the capacity to produce these compounds would thus be highly desirable.

The observation that *Bacillus* sp. RAPc8 NHase is active on heteroaromatics (3-cyanopyridine), yet was unable to convert homoaromatics (benzonitrile) formed the foundation of a working hypothesis. Cameron (2002) suggested that, since these

substrates are superficially equal in bulk, the lack of activity was not due to a small substrate-binding cavity as was suggested by Nagashima *et al*. (1998) for the *Rhodococcus* sp. N-771 NHase. This hypothesis further suggested that since benzonitrile appeared to be a potent inhibitor of NHase activity, the lack of activity was due to aromatic-aromatic or hydrophobic interactions between active-site residues and the substrate, trapping the aromatic nitrile substrate in a non-catalytic orientation.

An earlier study (Cameron, 2002) based the design of mutant NHases on this hypothesis and a crude homology model of *Bacillus* sp. RAPc8 NHase. At this stage, the enzyme studied was produced in the absence of P14K and was only marginally active. A similar study was conducted here. In this work however, fully active enzyme produced by co-expression with P14K was used (see Table 3.1). In addition, the crystal structure described in Chapter 4 was used as a core component of a typical protein engineering cycle (Figure 5.1).

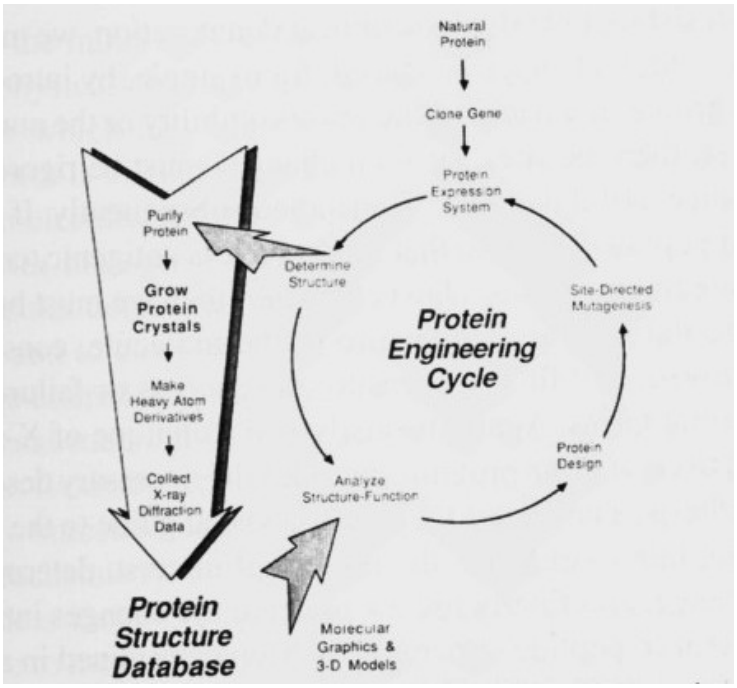


Figure 5.1 The protein engineering cycle (from McPherson *et al.*, 1999).

5.2 Structural enzymology of benzonitrile inhibition in Bacillus sp. RAPc8 NHase

5.2.1 Enzyme kinetics of inhibition

The study by Cameron (2002) estimated a K_i for inhibition of NHase by benzonitrile. This work was preliminary in that non-standard methods for determining inhibition kinetics were used. Here, different methods were used.

Standard kinetics of conversion of acrylonitrile (a good, industrially important aliphatic nitrile substrate) were determined using a continuous assay as described in section 2.9. The same assays were repeated in the presence of 100μM and 200μM benzonitrile. These data are shown here using a Lineweaver-Burke plot for easy visualisation (Figure 5.2).

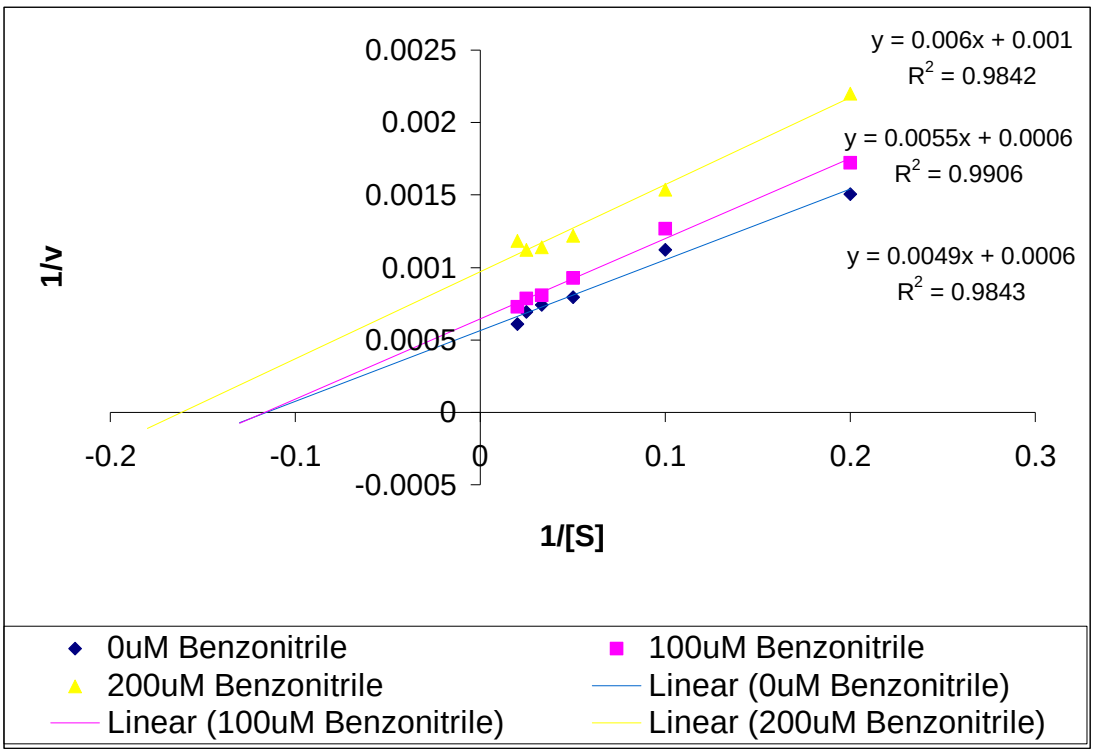


Figure 5.2 Lineweaver-Burke plot of kinetic data of wild-type NHase in the presence and absence of benzonitrile.

For more reliable, state-of-the art analysis, the program Enzpack (Biosoft) was used. Table 5.1 shows the resulting outputs from the different methods of fitting the Michaelis-Menten equation. The data were analysed using Lineweaver-Burke, Hanes, Eadie-Hofstee and the direct linear plot of Eisenthal and Cornish-Bowden (1974). The non-linear regression method (Wilkinson) of analysis was also used.

Table 5.1 Kinetic constants of wild-type *Bacillus* sp. RAPc8 NHase with acrylonitrile as substrate in the presence and absence of benzonitrile.

[Benzonitrile] (μM)		Lineweaver-Burke	Hanes	Eadie-Hofstee	Direct Linear	Wilkinson
0	**K_M (mM)**	8.64	11.0	9.17	10.1	10.4
	V_{max} (U/mg)	1768	1920	1819	1832	1882
100	**K_M (mM)**	8.57	9.84	8.99	10.2	9.71
	V_{max} (U/mg)	1554	1617	1579	1629	1617
200	**K_M (mM)**	6.19	4.24	5.82	5.29	5.43
	V_{max} (U/mg)	1031	955	1014	1016	1000

The limitations and merits of each method of calculation have been discussed by Eisenthal and Cornish-Bowden (1974) and Atkins and Nimmo (1975). The values obtained from the Direct-Linear and Wilkinson methods of analysis were taken as the most reliable. Statistical analysis of these data is shown in Table 5.2 (Wilkinson) and Table 5.3 (Direct Linear).

Table 5.2 Kinetic constants of wild-type NHase in the presence and absence of benzonitrile as calculated using non-linear regression analysis (Wilkinson method).

[Benzonitrile] (μM)	K_M (mM)	Standard error	V_{max} (U/mg)	Standard error	V_{max}/K_M range
0	10.4	1.74	1882	96	147-228
100	9.71	0.863	1617	42	149-200
200	5.43	0.956	1000	39	150-232

Table 5.3 Kinetic constants of wild-type NHase in the presence and absence of benzonitrile as calculated using Direct Linear Plot analysis.

[Benzonitrile] (μM)	K_M (mM)	68%Confidence limits	V_{max} (U/mg)	68% Confidence limits	V_{max}/K_M range
0	10.1	8.12-12.8	1832	1743-2059	136-254
100	10.2	8.51-11.4	1629	1541-1693	135-199
200	5.29	3.7-6.35	1016	931-1064	147-288

The K_M determined here in the absence of benzonitrile compares well with previously determined values for native RAPc8 NHase (11mM; Pereira, 1998) and recombinant NHase (11.9mM; Cameron, 2002). The differences observed between the values for V_{max} are a result of the fact that the NHase preparations characterised in the respective studies were produced in different ways.

The data in the presence of 100μM and 200μM benzonitrile were analysed to determine the type and extent of inhibition by benzonitrile. Previous reports suggested that benzonitrile was a competitive inhibitor with an inhibition constant (K_i) of 2.25mM (Cameron, 2002). The data obtained here show a significant contrast to this previous work. Comparison of data in the presence of 100μM benzonitrile with data in the absence of inhibitor indicated a small difference in the K_M and V_{max} values. Comparison of data in the presence of 200μM benzonitrile with data in the absence of inhibitior was more conclusive, showing a more significant drop in both V_{max} and K_M on addition of 200μM benzonitrile. The values of V_{max}/K_M in all instances (0μM, 100μM and 200μM benzonitrile) were essentially the same (considering the error limits; Tables 5.2 and 5.3). These observed changes in K_M and V_{max} with maintenance of a constant V_{max} to K_M ratio are characteristic of uncompetitive inhibition, where inhibitor binds only to the enzyme-substrate complex. The apparent inhibition constant (K_{ies}) was found to be 227μM (calculated using Enzpack). The very large difference between this inhibition constant and the K_M for acrylonitrile indicated that the strength of inhibition by benzonitrile was very high.

5.2.2 Structural biology of substrate and inhibitor binding

The observation that benzonitrile is an uncompetitive inhibitor of *Bacillus* sp. RAPc8 indicated that this nitrile binds only to the enzyme-substrate complex. This suggests that conformational changes that occur upon substrate-binding result in a binding site for benzonitrile being exposed. Substrate/substrate analog binding in NHases generally does not result in widespread conformational changes. For example, in the crystal structures of *Ps. thermophila* NHase solved in the absence and presence of

butyric acid, the Cα atom root mean square deviation (r.m.s.d) between the two is only 0.067Å. The overall r.m.s.d. (i.e., all atoms) is 0.272Å (Miyanaga *et al*., 2003), indicating that the predominant changes upon substrate-binding were at the side-chain level, probably within the substrate channel. In the active and inactive forms of *Rhodococcus* sp. N-771/R312 NHase (these two enzymes share 100% sequence identity), again large conformational changes were not detectible. The (all atom) r.m.s.d between substrate-channel-forming atoms however, was just short of 0.4Å (Huang *et al*., 1997; Nagashima *et al*., 1998). It is thus clear that conformational changes in NHase upon substrate binding are common but are limited to the substrate channel. These subtle conformational changes at active sites are a regular feature of most enzyme systems (Gutteridge and Thornton, 2005). Normal mode analysis of *Bacillus* sp. RAPc8 NHase is in progress and preliminary findings demonstrate that the substrate channel of this enzyme also shows a degree of flexibility that would allow conformational changes upon substrate binding (O. Tastan-Bishop, unpublished results).

It can be suggested therefore that the (inhibitor) binding site for benzonitrile that is exposed or created upon substrate binding in *Bacillus* sp. RAPc8 is probably located within the substrate-binding channel but away from the actual active site. Exhaustive co-crystallisation and crystal soaking experiments aimed at producing complex structures that provide a more definite location for benzonitrile binding are in progress. In a structural alignment between *Bacillus* sp. RAPc8 NHase and the butyric acid-bound NHase from *P. thermophila* (Miyanaga *et al*., 2004; PDB code: 1UGP), some of the residues that lie deep in the channel (at the interface between pocket 55 and 56; section 4.3.3.2) overlap exactly with the substrate-binding region of

P. thermophila NHase (Figure 5.3). Some of the residues and water molecules forming polar contacts with the substrate are conserved between the two enzymes. This region can therefore be designated as the substrate-binding/active site with reasonable confidence.

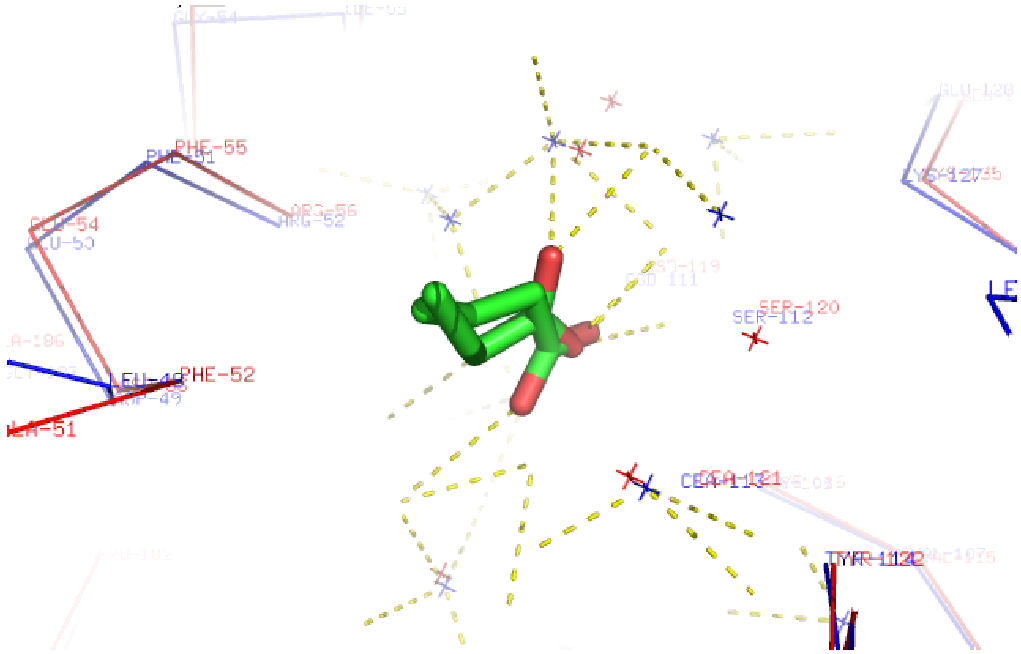


Figure 5.3 Superposition of *Bacillus* sp. RAPc8 NHase structure with the crystal structure of *Pseudonocardia thermophila* NHase with bound butyric acid. (Red: *Bacillus* sp. RAPc8, Blue: *Pseudonocardia thermophila*).

The amino-acid residues selected as targets for site directed mutagenesis were located in this region. In keeping with the working hypothesis described above, all amino-acid residues selected had aromatic side chains. Of the 52 residues forming the substrate channel, 11 had aromatic side-chains. Five of these eleven aromatic amino acid residues were selected for initial site-directed mutagenesis.

5.3 Construction, purification and crystallisation of mutant NHases

Mutants βF36L, βF52G, βF55L, βY67A and βW76G were constructed using the Quickchange Site Directed Mutagenesis Kit (Stratagene) as described in section 2.6. The resulting constructs were sequenced to confirm the presence of the new codons. All mutant constructs were successfully prepared.

Mutant *Bacillus* sp. RAPc8 NHases were expressed and purified essentially as described for the wild-type protein except that the heat-treatment step was not carried out (section 2.8). This was due to the fact that some of the mutations described here result in decreased protein thermostability (Cameron, 2003). Mutant NHases were successfully crystallised using the hanging-drop vapour diffusion method in exactly the same conditions as described for the wild-type protein. At the time of preparation of this thesis, refinement of the structures was in progress. Of all the mutants constructed, βW76G was the only one characterised to completion. Time constraints precluded the complete characterisation of the other mutant NHases.

5.4 The βW76G mutant NHase

In the structure of wild-type *Bacillus* sp. RAPc8 NHase, amino acid residue βW76 lies in helix 5 of the β subunit. In a surface representation of the structure of the wild-type NHase, the location of this residue within the substrate channel can clearly be seen (Figure 5.4). According to the CASTp analysis described in section 4.3.3.2, atoms NE1, CH2 and CZ2 (see Appendix 5 for the notation used for atoms in

tryptophan residues) from this residue directly contribute to the lining of the substrate channel.

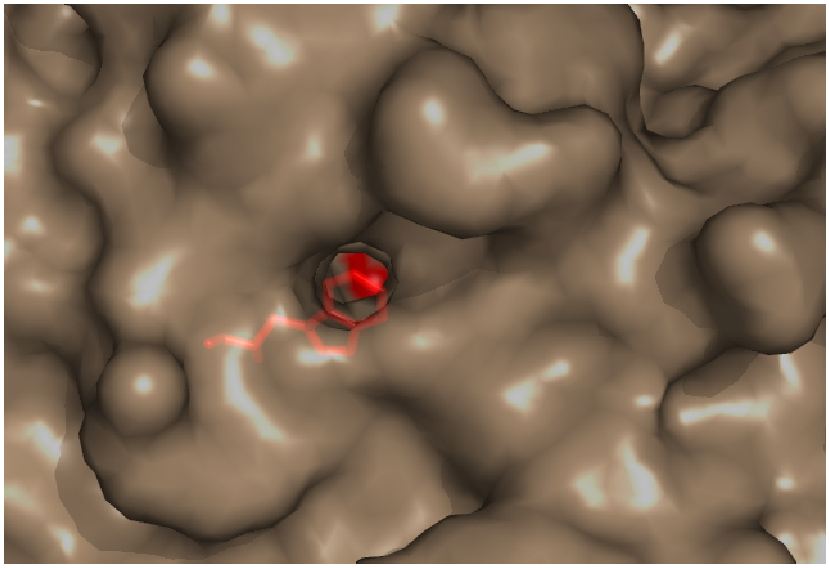


Figure 5.4 Surface representation of wild-type *Bacillus* sp. RAPc8 NHase showing location of residue βW76 (Red).

Mutation of this tryptophan to glysine did not result in detectable levels of activity on benzonitrile. This mutation did, however, result in decreased inhibition by benzonitrile. Enzyme kinetic characterisation of the mutant protein was carried out in a similar manner to the wild-type protein. These data are presented using a Lineweaver-Burke plot (Figure 5.5). Kinetic data calculated using non-linear regression (Wilkinson) and Direct Linear analysis are presented in Table 5.4 and Table 5.5, respectively.

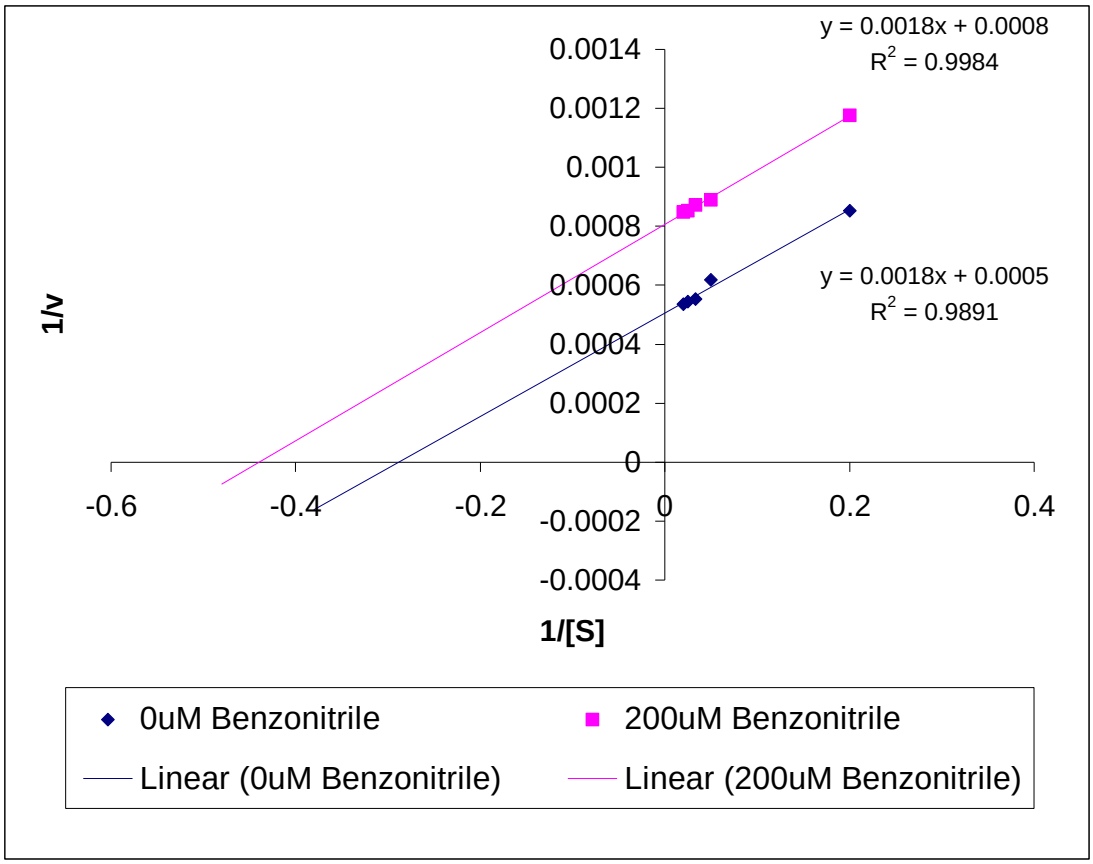


Figure 5.5 Lineweaver-Burke plot of kinetic data of NHase βW76G mutant in the presence and absence of benzonitrile.

Table 5.4 Kinetic constants of NHase βW76G mutant in the presence and absence of benzonitrile as calculated using non-linear regression analysis (Wilkinson method).

[Benzonitrile] (μM)	K_M(mM)	Standard error	V_{max}(U/mg)	Standard error	V_{max}/K_M range
0	4.05	0.555	2007	53	420-589
200	2.29	0.468	1461	110	428-862

Table 5.5 Kinetic constants of NHase βW76G mutant in the presence and absence of benzonitrile as calculated using Direct Linear Plot analysis.

[Benzonitrile] (μM)	K_M (mM)	68%Confidence limits	V_{max} (U/mg)	68% Confidence limits	V_{max}/K_M range
0	3.87	3.52-5.29	2011	1996-2076	377-589
200	2.49	1.02-4.22	1402	1223-1553	289-1522

The K_{ies} for benzonitrile inhibition of the βW76G mutant was found to be 536μM. This corresponds to more than two-fold less potent inhibition compared with the wild-type NHase. This finding indicated a likely role for residue βW76 in benzonitrile binding.

The crystal structure of this mutant NHase was solved by molecular replacement using the wild-type NHase structure as a search model. The program Molrep (CCP4 suite) was used for this procedure. The first solution had an R-factor of 33.2% and a correlation coefficient of 73.2%. Inspection of the electron density map at the site of mutagenesis (i.e., position β76) clearly revealed the substitution of glycine for tryptophan. The mutated residue was successfully fitted into this electron density (Figure 5.6).

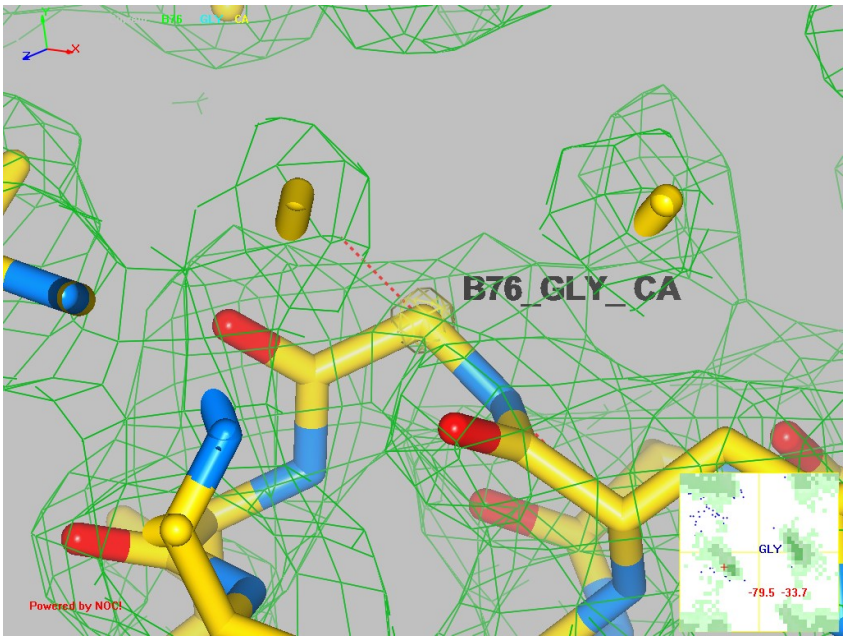


Figure 5.6 Stick representation of NHase βW76G crystal structure showing electron density in the region of βG76. (Figure generated using NOC, Institute of Biophysics, Chinese Academy of Sciences)

The structure was refined until an R-factor of 22.4% and an R_{free} of 26.7% was achieved. The integrity of the geometry of the structure was assessed using RAMPAGE (Lovell *et al.*, 2001). Figure 5.7 shows the output from RAMPAGE. From these validation procedures it was concluded that the structure was suitable for analysis.

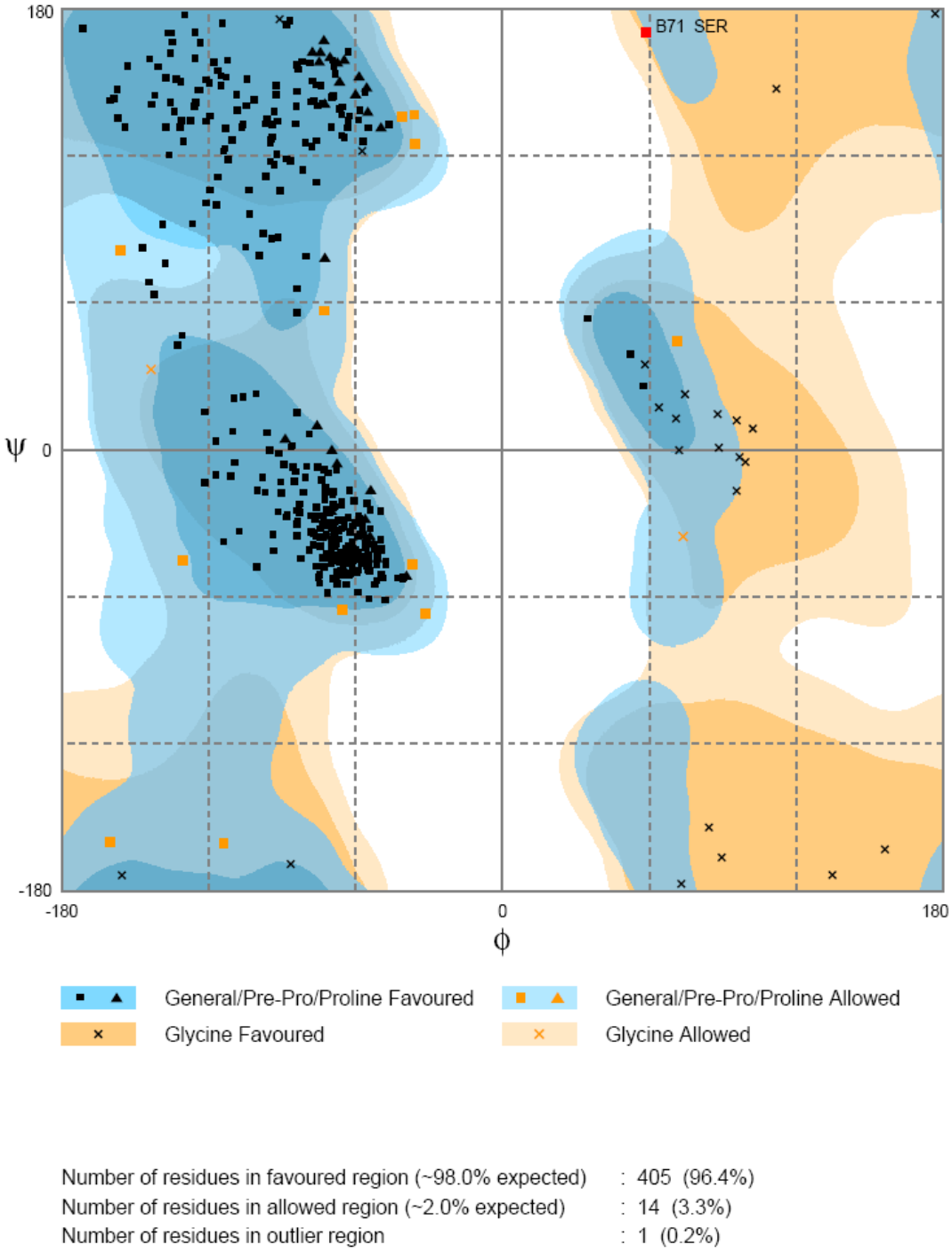


Figure 5.7 Ramachandran plot of βW76G NHase structure.

No widespread conformational differences were observed in a comparison of this mutant structure with that of the wild-type protein (an r.m.s.d. of 0.101Å was observed, with 416 out of 429 C-α atoms superimposed).

Analysis of the cavities within the crystal structure of this mutant NHase was perfomed using CASTp in a similar manner to the wild-type protein. This analysis revealed a total of 52 pockets or cavities. Among these, the two largest pockets overlapped with the region identified as the substrate channel in section 4.3.3.2. The pockets formed a complete channel that traverses across the molecule through the cobalt centre (Figure 5.8).

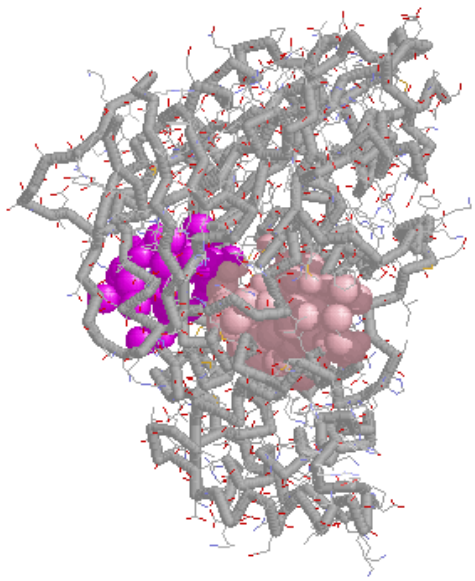


Figure 5.8 Two major pockets/cavities within βW76G mutant NHase. Pink spheres: pocket 51, magenta spheres: pocket 52. (Figure generated using CASTp and Chime.)

Table 5.6 Comparison of geometric dimensions of the cavities forming the substrate channel of wild-type and βW76G NHase.

Wild-type NHase			**βW76G NHase**		
Pocket	**Volume (Å^3)**	**Inner surface area (Å^2)**	**Pocket**	**Volume (Å^3)**	**Inner surface area (Å^2)**
55	481	542	**51**	739	849
56	494	506	**52**	343	405
Total	975	1048	**Total**	1082	1254

Comparison of the geometric dimensions of these (substrate-channel-forming) pockets in βW76G with those from the wild-type NHase revealed some differences (Table 5.6). The mutation resulted in an 11% increase in volume and a 20% increase in inner surface area of the substrate channel. This increase in cavity dimensions was the single most dramatic difference between the wild-type and βW76G NHases. It can be suggested that this difference is the likely cause for the observed decrease in inhibition.

5.5 Possibilities for new NHase mutants

In the heterotetrameric structure of wild-type *Bacillus* sp. RAPc8 NHase, access to bulk solvent from the pocket 55 end of the channel is blocked by the N-terminal arm of the cognate dimer (Figure 5.9). The pocket 56 end, is however accessible (Figure 5.10). It is therefore more likely that substrate and inhibitor enter the channel through the entrance on the pocket 56 end of the channel than the entrance on the pocket 56 end. This would mean that since benzonitrile binds only after substrate binding, the site of inhibition is adjacent to the substrate-binding site in the direction of the entrance into pocket 56. Further amino acid targets for site-directed mutagenesis should therefore be located in this region of the substrate channel.

Although both this study and the work of Cameron (2002) are incomplete, it is clear that mutation of aromatic side-chains so far studied does not result in any activity on benzonitrile. It can be concluded therefore that the ability of *Bacillus* sp. RAPc8 NHase to accept 3-cyanopyridine and not benzonitrile as a substrate is probably not exclusively due to aromatic-aromatic trapping. However, the hypothesis can still be explored and is strengthened by the work of Pochapsky and Gopen (1992) and Pereira de Araujo *et al*. (1999).

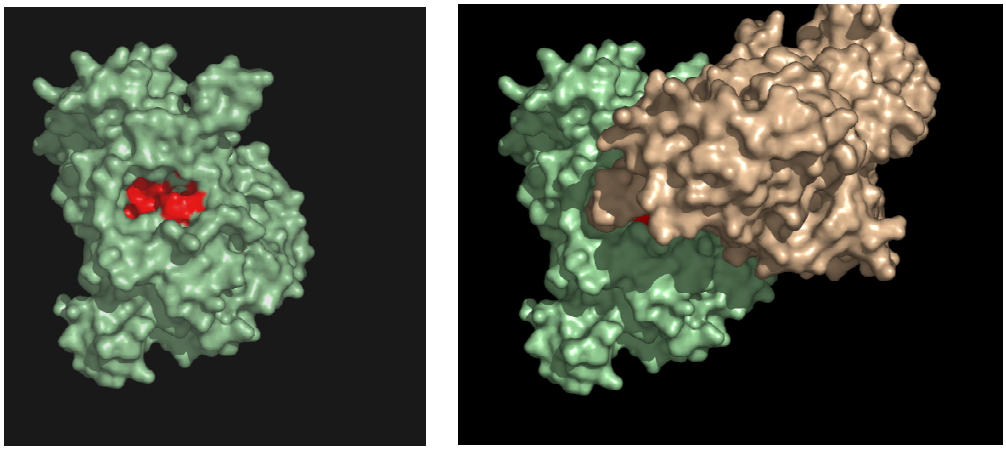


Figure 5.9 Surface representation of wild-type *Bacillus* sp. RAPc8 NHase heterodimer (left) and heterotetramer (right). The entrance to the channel through pocket 55 is blocked by an N-terminal arm in the tetramer. Channel-forming residues are coloured in red.

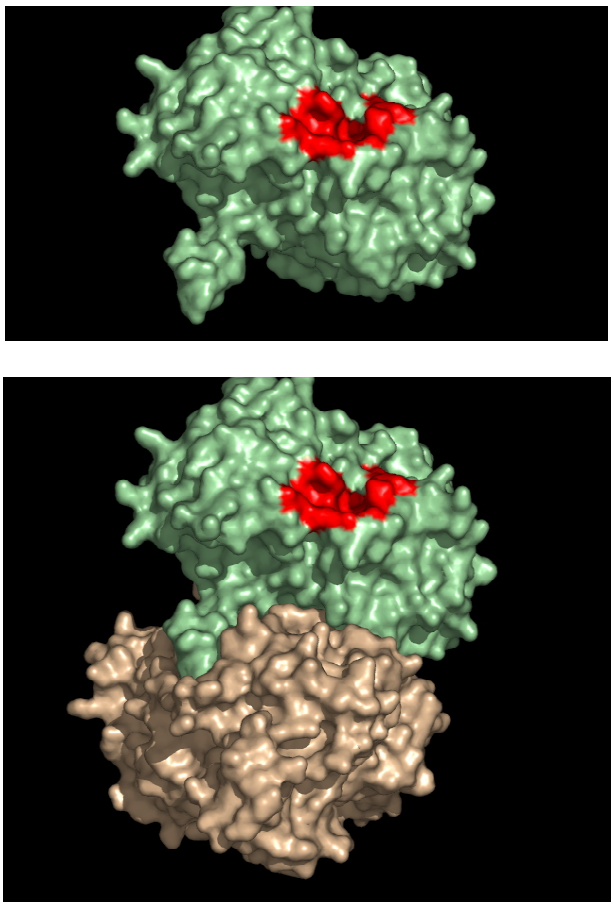


Figure 5.10 Surface representation of wild-type *Bacillus* sp. RAPc8 NHase heterodimer (top) and heterotetramer (bottom). The entrance to the channel through pocket 56 is accessible in the tetramer. Channel-forming residues are coloured in red.

These studies showed that interactions between aromatic-aromatic entities (or side-chain mimics) are remarkably strong at all temperatures. The phenomena described in these studies make the possibility of trapping of benzonitrile by aromatic side-chains possible.

Further work will not focus changes only on aromatic side chains. The strategy employed will attempt to increase the volume of the substrate channel in various regions. One of the consequences of this may be the removal of the possibility of aromatic-aromatic trapping through pi-cloud-pi-cloud association between aromatic side-chains and benzonitrile. The following section describes and discusses the prospect of using *in silico*-modelled structures of *Bacillus* sp. RAPc8 NHase to predict the likely increase in cavity dimensions as a result of specific mutations.

It is also anticipated that solution of structures of *Bacillus* sp. RAPc8 NHase in complex with substrates and/or inhibitors will yield more precise targets for mutagenesis.

An alternative strategy will be to produce an NHase mutant with a truncation at the N-terminal arm. In this way, the entrance to the substrate channel that is apparently blocked in the wild-type protein would potentially be accessible. Entrance of benzonitrile through this opening might then be possible. It may be necessary to use this approach in conjunction with further site-directed mutagenesis of the truncated NHase.

It is clear from this discussion that the so-called rational approach to engineering enzyme substrate specificity is fraught with a high degree of uncertainty even where the work is approached logically. Locating the relevant residues for mutagenesis is often not straightforward. For this reason random approaches are now sometimes used. A study that aims to produce a library of *Bacillus* sp. RAPc8 NHase mutants by directed evolution using error prone PCR is currently underway (van Wyk J., unpublished results). This library will be screened for activity on benzonitrile and the residues responsible for the observed changes may be identified.

5.6 In silico modelling as a tool for the design of new mutants

Mutation of amino-acid residue βW76 to glycine did not result in any widespread conformational changes in the structure of NHase. It is proposed that the observed differences between the substrate channel of the wild-type protein and this mutant were primarily as a result of the space occupied by the side-chain.

In order to assess the potential use of modelled NHase structures to predict changes in substrate channel dimensions, two *in silico* models of the structure of *Bacillus* sp. RAPc8 NHase were prepared as outlined in section 2.13. One model was a βW76G mutant prepared from the experimentally solved wild-type NHase structure and the other was a wild-type structure prepared from the experimentally solved βW76G mutant. The procedure used assumes that the amino-acid substitution in the modelled structure does not result in widespread conformational changes and that the conformation of the new side chain is the same as that of the original one.

The cavities within each of the modelled structures were analysed using CASTp in the same manner as the experimentally solved crystal structures. The dimensions of the substrate channel in each of the structures were compared with the dimensions of the substrate channel of the corresponding solved structure (Table 5.7).

Table 5.7 Comparison of substrate channel dimensions of experimentally solved NHase structures with modelled NHase structures.

Stucture	Volume ($Å^3$)	Inner surface area ($Å^2$)
X-ray WT	975	1048
Model WT	968	1139
X-ray W76G	1082	1254
Model W76G	1180	1254

The cavity-dimensions of the modelled structures were very similar to or identical with those of the X-ray structures. This indicated that modelling of structures in future mutagenesis studies would be a useful predictive tool for the design of new mutants. Further research efforts would thus be directed to NHase mutants that show the desired cavity dimensions *in silico*.

Chapter 6: General Discussion

The primary aim of this study was to determine the molecular structure of *Bacillus* sp. RAPc8 NHase. The detailed structural information was used as a central component of a protein engineering cycle aimed at engineering aromatic substrate specificity into this enzyme.

A crucial step in the process of elucidating the structure of a protein by X-ray diffraction analysis is the efficient expression and purification of the protein of interest. Chapter 3 describes the purification, crystallisation and preliminary x-ray diffraction analysis of NHase. Recombinant *Bacillus* sp. RAPc8 NHase was expressed in *E. coli* BL21 and purified in three steps; heat treatment, phenyl-Sepharose (hydrophobic interaction) chromatography and Q-Sepharose (ion exchange) chromatography. On completion of the purification protocol, the NHase produced was close to electrophoretic purity and was of a single oligomeric form as determined by gel filtration. A purification factor of 8.9-fold was achieved with 61% yield. The various steps used here to produce protein of sufficient purity for crystal trials have previously been used to (partially) purify other NHases for various purposes (e.g., Kim *et al*., 2001).

Purified NHase was subjected to crystallisation trials to determine appropriate conditions for crystallisation. The use of commercially available sparse matrix screens for the search of these conditions proved to be a very efficient strategy. Diffraction quality crystals were finally grown at 20˚C in 30% PEG 400, 100mM magnesium chloride, 100mM MES (2[N-Morpholino]ethanesulfonic acid), pH 6.5 (40mg/ml

NHase). This composition was very similar to the composition of the conditions used to crystallise *Bacillus smithii* NHase (Hourai *et al*., 2003).

The crystals finally used for collection of a complete set of X-ray diffraction data were approximately 0.2 x 0.2 x 0.3 mm in size. These crystals diffracted well with diffraction spots visible beyond 2.4Å, with little mosaicity. At 2.5Å, the data were 93% complete. The data collection statistics indicated that the data set obtained was of adequate quality for progression to structure solution.

Once soluble, pure target protein is produced, the single most significant obstacle for the crystallographer is the 'phase problem' that results from the fact that phase information is not directly obtainable from the crystallographic experiment. Availability of the structure of a homologue allowed determination/estimation of initial phases by the now widely used method of molecular replacement. The crystal structure of *Pseudonocardia thermophila* NHase was used as a search model in this procedure. The program EPMR was used for estimation of phases and the best model had a correlation coefficient of 35% and an R-factor of 56.2%. The structure was refined until an R-factor of 22% and an R_{free} of 27% was achieved. The final structure had very good geometry, with very few residues found in outlier regions of the Ramachandran plot. One of the outliers was the post-translationally modified cys119 from the metal-binding site of α subunit. Since the fit of the model to the electron density map in this region was very good, it is possible that the conformation of the metal-binding claw-setting motif unique to NHases forces ψ and ϕ angles of amino-acids forming this motif into rare or strained conformations.

The refined structure was analysed in detail. The overall topology was similar to other known NHase crystal structures. The metal binding motif was essentially identical to that found in other NHases, including that of the iron-type *Rhodococcus* sp. R312 NHase (Huang *et al*., 1997). This observation is consistent with a study that demonstrated that *Rhodococcus* sp. N-771 iron type NHase was able to bind cobalt in place of iron. This cobalt-substituted iron-type NHase exhibited several-fold higher activity than enzyme produced in the complete absence of metal (Nojiri *et al*., 2000).

All NHase crystal structures reported to date except that of the *B. smithii* NHase (Hourai *et al*., 2004) indicated that two of the cysteine side chains were post-translationally modified. The electron density map of *Bacillus* sp. RAPc8 NHase showed that αCys119 and αCys121 were, as expected, oxidized to cysteine sulfinic and cysteine sulfenic acid, respectively. The apparent lack of modification in the *B. smithii* NHase may have been due to errors or omissions during solution of this structure or during the refinement process.

The substrate-binding and catalytic cavity of RAPc8 NHase was identified and described in detail. Surface representation of the structure revealed an extended, curved solvent accessible channel with access to bulk solvent from two locations in the heterodimer. The amino-acid residues forming the channel were identified and the geometric dimensions measured. This information was used to aid in the attempted engineering of aromatic substrate specificity into *Bacillus* sp. RAPc8 NHase. The strategy followed was a typical protein engineering cycle, with the high resolution crystal structure used as a core component. A previous hypothesis stated that since benzonitrile appeared to be a potent inhibitor of NHase activity, the lack of activity

was due to aromatic-aromatic or hydrophobic interactions between active-site residues and the substrate, trapping the aromatic nitrile substrate in a non-catalytic orientation (Cameron, 2002). This hypothesis was tested.

Enzyme inhibition kinetics indicated that benzonitrile was a potent uncompetitive inhibitor of NHase. From these data, the substrate channel structure and through superimposition of the structure of RAPc8 NHase with that of substrate-bound NHase from another organism, initial amino-acid residues in the substrate channel were selected as targets for mutagenesis. In keeping with the working hypothesis described above, all amino-acid residues selected had aromatic side chains. Of all the mutants constructed, βW76G was the only one characterised to completion. Time constraints precluded the complete characterisation of the other mutant NHases.

Enzyme kinetics of this mutant were determined and the crystal structure solved. Although this mutation did not result in activity on benzonitrile, it was observed that the strength of inhibition was some two to three fold lower. This finding indicated a likely role for residue βW76 in benzonitrile binding. Analysis of the substrate channel of this mutant NHase showed an 11% increase in volume and a 20% increase in inner surface area compared to that of the wild-type NHase. Due to the lack of other significant differences between the two structures (an r.m.s.d of only 0.101Å was observed), this difference is thought to be responsible for the decrease in benzonitrile inhibition.

Since other aromatic-side-chain-mutants were not studied in detail, conclusions regarding the working hypothesis tested in this study cannot be made with confidence.

The observed decrease in benzonitrile inhibition that was associated with an increase in cavity dimensions led to the conclusion that, in any future mutagenesis studies, the primary target may be to increase the geometric dimensions of the substrate channel of *Bacillus* sp. RAPc8 NHase in specific regions. A structure-modelling based approach for assessing the likely structural differences that may result as a result of a specific mutation was suggested and tested. This approach may be of value in future work.

Although the crystal structure described here indicates that the αβ heterodimer exists within each asymmetric unit, it was nonetheless important to inspect the crystal structure for the presence of interactions that may have formed a heterotetramer. *Bacillus* sp. RAPc8 NHase has previously been shown to exist as a heterotetramer (Pereira, 1998). Analysis of the accessible surface structure of the tetrameric form indicated that the opening of one end of the channel may be blocked by an N-terminal protrusion from the cognate dimer. This observation led to the suggestion that further targets for mutagenesis should be selected from the accessible end of the channel. It is anticipated that solution of structures of *Bacillus* sp. RAPc8 NHase in complex with substrates and/or inhibitors will yield more precise targets for mutagenesis.

Other studies of NHase substrate specificity have also suggested that differences in amino-acid composition of the substrate channel result in the observed differences in substrate preference. For example, conservation of amino-acid residues βL48, βF51 and βW72 between NHases from *Pseudonocardia thermophila* and *Rhodococcus rhodochrous* J1 (light NHase) has been linked to the similar substrate preference profile of these enzymes (Miyanaga *et al*., 2003). These two enzymes are also unable

to catalyse the conversion of benzonitrile. In the RAPc8 NHase structure, the analogous residues are βF52, βF55 and βW76, respectively. The present study has already implicated βW76 in binding of benzonitrile and has reported the selection of the remaining two residues as targets for mutagenesis based on the hydrophobic trapping hypothesis. It is possible that mutation of these residues to less bulky residues in combination (i.e., a triple mutant) might produce the desired aromatic specificity.

In the absence of high-resolution structures of all NHases whose substrate specificities have been studied in detail, it is difficult to reconcile primary structural differences with differences at the three-dimensional level. Although tedious, the rational approach to engineering substrate specificity in NHases is likely to remain the most suitable approach for the foreseeable future. However, other approaches, which may include directed evolution, should also be pursued.

References

Atkins G.L., and Nimmo I.A. (1975) A comparison of seven methods for fitting the Michaelis-Menten equation. *Biochemical Journal*. **149:** 775-7.

Baneyx F. (1999) Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology*. **10:** 411-21.

Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., and Bourne P.E. (2000) The Protein Data Bank. *Nucleic Acids Research*. **28:** 235-42.

Berman H.M., Battistuz T., Bhat T.N., Bluhm W.F., Bourne P.E., Burkhardt K., Feng Z., Gilliland G.L., Iype L., Jain S., Fagan P., Marvin J., Padilla D., Ravichandran V., Schneider B., Thanki N., Weissig H., Westbrook J.D., and Zardecki C. (2002) The Protein Data Bank. *Acta Crystallographica D, Biological Crystallography*. **58:** 899-907.

Bradford MM. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry.* **72:** 248-54.

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.

Cameron R. (2002). Nitrile degrading enzymes from extreme environments. *Ph.D. thesis*, University of London, London, UK.

Cameron R.A., Sayed M., and Cowan D.A. (2005) Molecular analysis of the nitrile catabolism operon of the thermophile *Bacillus pallidus* RAPc8. *Biochimica Biophysica Acta.* **1725:** 35-46.

Carter Jr. C.W. and Sweet R.M. (eds) (1997) Macromolecular Crystallography. *Methods in Enzymology*. **276**

Collaborative Computational Project, Number 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallographica D, Biological Crystallography*. **50:** 760-3.

Cowan D., Cramp R., Pereira R., Graham D., and Almatawah Q. (1998) Biochemistry and biotechnology of mesophilic and thermophilic nitrile metabolizing enzymes. *Extremophiles.* **2:** 207-216.

Cowan D.A., Cameron R.A., and Tsekoa T.L. (2003) Comparative biology of mesophilic and thermophillic nitrile hydratases. *Advances in applied microbiology*. **52:** 123-153.

Cramp R.A. and Cowan D.A. (1999) Molecular characterisation of a novel thermophilic nitrile hydratase. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*. **1431:** 249-260.

Cramp R. (1997) Novel nitrile degrading enzymes in thermophilic bacteria. *Ph.D. thesis*, University of London, London, UK.

Dauter Z., and Dauter M. (2001) Entering a new phase: using solvent halide ions in protein structure determination. *Structure (Cambridge)*. **9:** R21-6.

Duran R., Nishiyama M., Horinouchi S. and Beppu T. (1993) Characterization of nitrile hydratase genes cloned by DNA screening from *Rhodococcus erythropolis*. *Bioscience Biotechnology and Biochemistry*. **57:** 1323-8.

Ealick S.E. (2000) Advances in multiple wavelength anomalous diffraction crystallography. *Current Opinion in Chemical Biology*. **4:** 495-9.

Effenberger F. and Graef B.W. (1998) Chemo- and enantioselective hydrolysis of nitriles and acid amides, respectively, with resting cells of *Rhodococcus* sp. C3II and *Rhodococcus erythropolis* MP50. *Journal of Biotechnology*. **60:** 165-174.

Effenberger F. and Bohme J. (1994) Enzyme-catalysed enantioselective hydrolysis of racemic naproxen nitrile. *Bioorganic and Medicinal Chemistry*. **2:** 715-21.

Eisenthal R., and Cornish-Bowden A. (1974) The direct linear plot. A new graphical procedure for estimating enzyme kinetic parameters. *Biochemical Journal*. **139:** 715-20.

Emsley P. and Cowtan K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallographica D, Biological Crystallography*. **60:** 2126-32.

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.

Fawcett J., and Scott J. (1960) A rapid and precise method for the determination of urea. *Journal of Clinical Pathology*. **13:** 156-159.

Garman E. (1999) Cool data: quantity AND quality. *Acta Crystallographica D, Biological Crystallography*. **55:** 1641-53.

Garman E.F. and Mitchell E. P. (1996). Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions. *Journal of Applied Crystallography.* **29:** 584-587.

Gill S.C. and von Hippel P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry*. **182:** 319-326.

Gilligan T., Yamada H., and Nagasawa T. (1993) Production of S-(+)-2-phenylpropionic acid from (R,S)-2-phenylpropionitrile by the combination of nitrile hydratase and stereoselective amidase in *Rhodococcus equi* TG328. *Applied Microbiology and Biotechnology*. **39:** 720-5.

Gutteridge A., and Thornton J. (2005) Conformational changes observed in enzyme crystal structures upon substrate binding. *Journal of Molecular Biology*. **346:** 21-8.

Guex N., and Peitsch M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. **18:** 2714-23.

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 immobilized *Pseudomonas chlororaphis* B23. *Bioorganic and Medicinal Chemistry*. **7:** 2239-45.

Harper D.B. (1977) Fungal degradation of aromatic nitriles. Enzymology of C-N cleavage by *Fusarium solani*. *Biochemical Journal.* **167**: 685-92.

Hashimoto Y., Hosaka H., Oinuma K., Goda M., Higashibata H., and Kobayashi M. (2005) Nitrile pathway involving Acyl-CoA synthetase - Overall metabolic gene organization and purification and characterization of the enzyme. *Journal of Biological Chemistry.* **280:** 8660-8667.

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 Biochemistry*. **58:** 1859-65.

Hendrickson W.A. (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science.* **254:** 51-58.

Hendrickson W.A., Horton J.R., and LeMaster D.M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO Journal.* **9:** 1665-1672.

Hengming K. (1997) Overview of isomorphous replacement phasing. *Methods in Enzymology*. **276:** 448–61.

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:** 340-5.

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.

Jancarik J., and Kim S.H. (1991) Sparse Matrix Sampling: a screening method for crystallization of proteins. *Journal of Applied Crystallography.* **24:** 409-411.

Kato Y., Ooi R., and Asano Y. (2000) Distribution of aldoxime dehydratase in microorganisms. *Applied and Environmental Microbiology.* **66:** 2290-2296.

Kato Y., Ooi R., and Asano Y. (1998) Isolation and characterization of a bacterium possessing a novel aldoxime-dehydration activity and nitrile-degrading enzymes. *Archives of Microbiology.* **170:** 85-90.

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.

Kaufmann G., Dautzenberg H., Henkel H., Muller G., Schafer T., Undeutsch B., and Oettel M. (1999) Nitrile hydratase from *Rhodococcus erythropolis*: Metabolization of steroidal compounds with a nitrile group. *Steroids.* **64:** 535-540.

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.

Kimura M., Kuboki A., and Sugai T. (2002) Chemo-enzymatic synthesis of enantiomerically pure (R)-2-naphthylmethoxyacetic acid. *Tetrahedron: Asymmetry*. **13:** 1059-1068.

Kissinger C.R., Gehlhaar D.K. and Fogel D.B. (1999). Rapid automated molecular replacement by evolutionary search. *Acta Crystallographica D, Biological Crystallography.* **55:** 484-491.

Kleywegt G.J., and Jones T.A. (1997) Model-building and refinement practice. *Methods in Enzymology*. **277:** 208-230.

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 (BBA) - Gene Structure and Expression.* **1129:** 23-33.

Komeda H., Kobayashi M., and Shimizu S. (1997) A novel transporter involved in cobalt uptake. *Proceedings of the National Academy of Science, U S A.* **94:** 36-41.

Kraulis P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography*. **24:** 946-950.

Kuhnert D.C., Sayed Y., Mosebi S., Sayed M., Sewell T., and Dirr H.W. (2005) Tertiary interactions stabilise the C-terminal region of human glutathione transferase A1-1: a crystallographic and calorimetric study. *Journal of Molecular Biology*. **349:** 825-38.

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

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

Lamzin V.S., and Perrakis A. (2000) Current state of automated crystallographic data analysis. *Nature Structural Biology*. **7:** 978-81.

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:** 1049-67.

Laskowski R.A., MacArthur M.W., Moss D.S., and Thornton J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*. **26:** 283-291.

Liang J., Edelsbrunner H., and Woodward C. (1998) Anatomy of Protein Pockets and Cavities: Measurement of Binding Site Geometry and Implications for Ligand Design. *Protein Science*. **7:** 1884-1897

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.

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.

Lovell S.C., Davis I.W., Arendall W.B. 3rd, de Bakker P.I., Word J.M., Prisant M.G., Richardson J.S., and Richardson D.C. (2003) Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins.* **50:** 437-50.

Mascharak P. (2002) Structural and functional models of nitrile hydratase. *Co-ordination Chemistry Reviews*. **225:** 201-214

Matthews B.W. (1968). Solvent content of protein crystals. *Journal of Molecular Biology.* **33:** 491-497.

McCoy A.J., Grosse-Kunstleve R.W., Storoni L.C., and Read R.J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallographica D, Biological Crystallography.* **61:** 458-64.

McPherson A. (1999) Crystallization of biological Macromolecules. Cold Spring Harbor Laboatory Press, New York, USA.

McRee D.E. (1999) Practical protein crystallography. Academic Press, San Diego, USA.

Meth-Cohn O., and Wang M-X. (1995) A powerful new nitrile hydratase for organic synthesis - aromatic and heteroaromatic nitrile hydrolyses - a rationalisation, *Tetrahedron Letters*. **36:** 9561-9564.

Minor W., Cymborowski M., and Otwinowski Z. (2002) Automatic system for crystallographic data collection and analysis. *Acta Physica Polonica A.* **101:** 613-619.

Michalopoulos I., Torrance G.M., Gilbert D.R., and Westhead D.R. (2004) TOPS: an enhanced database of protein structural topology. *Nucleic Acids Research.* **32:** D251-4.

Miyanaga A., Fushinobu S., Ito K., and Wakagi T. (2001) Crystal structure of cobalt-containing nitrile hydratase. *Biochemical and Biophysical Research Communications.* **288:** 1169-74.

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:** 429-38.

Mizuguchi K., Deane C.M., Blundell T.L. and Overington J.P. (1998) HOMSTRAD: a database of protein structure alignments for homologous families. *Protein Science*. **7:** 2469-71.

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:** 240-55.

Nagasawa T., Ryuno K., and Yamada H. (1986) Nitrile hydratase of Brevibacterium R312--purification and characterization. *Biochemical Biophysical Research Communications.* **139:** 1305-12.

Nagasawa T. and Yamada H. (1989) Microbial transformations of nitriles. *Trends in Biotechnology.* **7:** 153-158.

Nagasawa T., Nanba H., Ryuno K., Takeuchi K. and Yamada H. (1987) Nitrile hydratase of *Pseudomonas chlororaphis* B23. Purification and characterization. *European Journal of Biochemistry.* **162:** 691-8.

Nojiri M., Yohda M., Odaka M., Matsushita Y., Tsujimura M., and Yoshida T. (1999) Functional expression of nitrile hydratase in *Escherichia coli*: Requirement of a nitrile hydratase activator and post-translational modification of a ligand cysteine. *Journal of Biochemistry.* **125:** 696-704.

Nagasawa T., Takeuchi K., and Yamada H. (1991) Characterization of a new cobalt-containing nitrile hydratase purified from urea-induced cells of *Rhodococcus rhodochrous* J1. *European Journal of Biochemistry*. **196**: 581-9.

Nagashima S., Nakasako M., Dohmae N., Tsujimura M., Takio K., Odaka M., Yohda M., Kamiya N., and Endo I. (1998) Novel non-heme iron center of nitrile hydratase with a claw setting of oxygen atoms. *Nature Structural Biology*. **5:** 347-51.

Navaza J. (2001) Implementation of molecular replacement in AmoRe. *Acta Crystallographica D, Biological Crystallography.* **57:** 1367-1372.

Nishiyama M., Horinouchi S., Kobayashi M., Nagasawa T., Yamada H. and Beppu T. (1991) Cloning and characterization of genes responsible for metabolism of nitrile compounds from *Pseudomonas chlororaphis* B23. *Journal of Bacteriology*. **173:** 2465-72.

Noguchi T., Honda J., Nagamune T., Sasabe H., Inoue Y., and Endo I. (1995) Photosensitive nitrile hydratase intrinsically possesses nitric oxide bound to the non-heme iron center: evidence by Fourier transform infrared spectroscopy. *FEBS Letters*. **358:** 9-12.

Noguchi T., Hoshino M., Tsujimura M., Odaka M., Inoue Y., and Endo I. (1996) Resonance Raman evidence that photodissociation of nitric oxide from the non-heme iron center activates nitrile hydratase from *Rhodococcus* sp. N-771. *Biochemistry.* **35:** 16777-81.

Nojiri M., Nakayama H., Odaka M., Yohda M., Takio K., and Endo I. (2000) Cobalt-substituted Fe-type nitrile hydratase of *Rhodococcus* sp. N-771. *FEBS Letters.* **465:** 173-7.

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

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

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

Pereira R. (1998) A novel nitrile hydratase and amidase from a thermophilic Bacillus isolate. *Ph.D. thesis*, University of London, London, UK.

Pereira de Araùjo A.F. (1999) Folding protein models with a simple hydrophobic energy function: The fundamental importance of monomer inside/outside segregation. *Proceedings of the National Academy of Science, U S A.* **96:** 12482–12487

Petrillo K., Wu S., Hann E., Cooling F., Ben-Bassat A., and Gavagan J. (2004) Over-expression in Escherichia coli of a thermally stable and regio-selective nitrile hydratase from *Comamonas testosteroni* 5-MGAM-4D. *Applied Microbiology and Biotechnology*. **55:** 345-67.

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

Pochapsky T.C., and Gopen Q. (1992) A chromatographic approach to the determination of relative free energies of interaction between hydrophobic and amphiphilic amino acid side chains. *Protein Science*. **1:** 786-95.

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.

Querol E., Perez-Pons J.A., and Mozo-Villarias A. (1996) Analysis of protein conformational characteristics related to thermostability. *Protein Engineering*. **9:** 265-71.

Rasor J.P., and Voss E. (2001) Enzyme-catalyzed processes inpharmaceutical industry. *Applied catalysis A: General*. **221:** 145-158.

Read RJ. (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallographica D, Biological Crystallography*. **57:** 1373-82. Eratum in: *Acta Crystallographica D, Biological Crystallography*. (2003) **59:** 404.

Rossmann M.G. (2001) Molecular replacement - historical background. *Acta Crystallographica D, Biological Crystallography.* **57:** 1360-1366.

Shi J., Blundell T.L., and Mizuguchi K. (2001) FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *Journal of Molecular Biology*. **310:** 243-57.

Sambrook J., and Russell D.W. (2001) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboatory Press, New York, USA.

Smyth MS, Martin JH. (2000) x Ray crystallography. *Molecular Pathology*. **53:** 8-14.

Snell D., and Colby J. (1999) Enantioselective hydrolysis of racemic ibuprofen amide to **S**-(+)-ibuprofen by *Rhodococcus* AJ270. *Enzyme and Microbial Technology*. **24:** 160-163.

Stevens J.M., Rao Saroja N., Jaouen M., Belghazi M., Schmitter J.M., and Mansuy D. (2003) Chaperone-assisted expression, purification, and characterization of recombinant nitrile hydratase NI1 from *Comamonas testosteroni*. *Protein Expression and Purification.* **29:** 70-76.

Storoni L.C., McCoy A.J., and Read R.J. (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallographica D, Biological Crystallography*. **60:** 432-8.

Sugai T., Yamazaki T., Yokoyama, M., and Ohta H. (1997) Biocatalysis in organic synthesis: The use of nitrile- and amide-hydrolyzing microorganisms *Bioscience Biotechnology and Biochemistry.* **61:** 1419-1427.

Takashima Y., Kawabe T., and Mitsuda S. (2000) Factors affecting the production of nitrile hydratase by thermophilic *Bacillus smithii* SC-J05-1. *Journal of Bioscience and Bioengineering.* **89:** 282-284.

Takashima Y., Yamaga Y., and Mitsuda S. (1998) Nitrile hydratase from a thermophilic *Bacillus smithii*. *Journal of Industrial Microbiology & Biotechnology*. **20:** 220-226.

Teng T.Y. (1990) Mounting crystals for macromolecular crystallography in a free standing thin film. *Journal of Applied, Crystallography*. **23:** 387-391.

Terwilliger T., and Berendzen J. (1999) Automated MAD and MIR structure solution. *Acta Crystallographica D, Biological Crystallography.* **55:** 849-861.

Thomas S.M., DiCosimo R., and Nagarajan A. (2002) Biocatalysis: applications and potentials for the chemical industry. *Trends in Biotechnology.* **20:** 238-242.

Trott S., Bauer R., Knackmuss H.J., and Stolz A. (2001) Genetic and biochemical characterization of an enantioselective amidase from *Agrobacterium tumefaciens* strain d3. *Microbiology.* **147:** 1815-24.

Tsekoa, T.L., Sayed, M.F., Cameron, R.A., Sewell B.T. and Cowan D.A. (2004) Purification, crystallization and preliminary X-ray diffraction analysis of thermostable nitrile hydratase. *South African Journal of Science.* **100:** 488-490.

Tsujimura M., Dohmae N., Chijimatsu M., Takio K., Odaka M., and Yohda M. (1997) Photoreactive nitrile hydratase: Posttranslational modification of photoreactive site. *Journal of Inorganic Biochemistry.* **67:** 335.

Vagin A., and Teplyakov A. (2002) An approach to multi-copy search in molecular replacement. *Acta Crystallographica D, Biological Crystallography*. **56:** 1622-4.

Wang M.X. and Feng G.Q. (2000) Enantioselective synthesis of chiral cyclopropane compounds through microbial transformations of trans-2-arylcyclopropanecarbonitriles. *Tetrahedron Letters.* **41:** 6501-6505.

Wang M.X., Li J.J., Ji G.J., and Li J.S. (2001) Enantioselective biotransformations of racemic 2-aryl-3-methylbutyronitriles using *Rhodococcus* sp AJ270. *Journal of Molecular Catalysis B-Enzymatic.* **14:** 77-83.

Wegman M.A., Heinemann U., van Rantwijk F., Stolz A., and Sheldon R.A. (2001) Hydrolysis of D,L-phenylglycine nitrile by new bacterial cultures. *Journal of Molecular Catalysis B-Enzymatic.* **11:** 249-253.

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.

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.

Yamada H., Kobayashi M. (1996) Nitrile hydratase and its application to industrial production of acrylamide. *Bioscience Biotechnology and Biochemistry*. **60:** 1391-400.

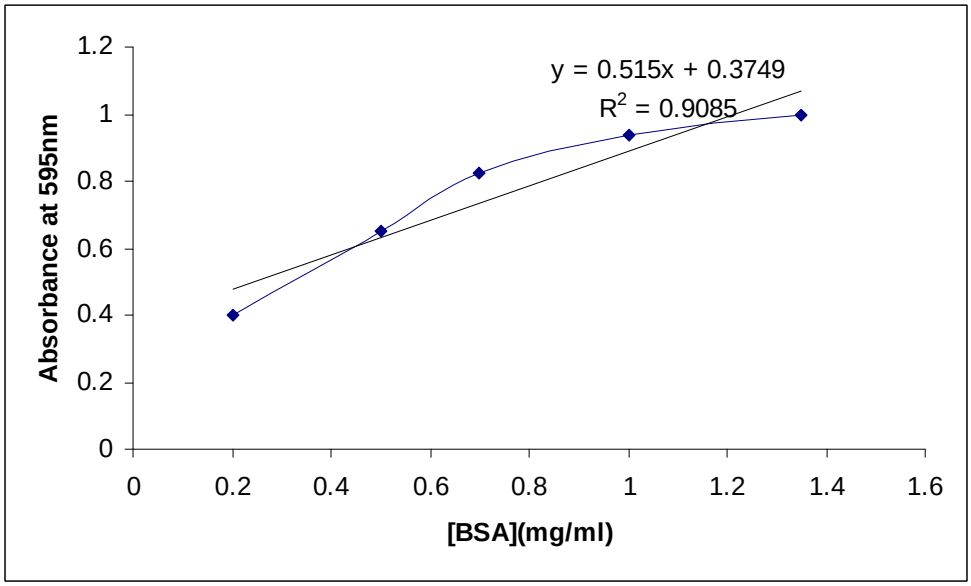
Yamaki T., Oikawa T., Ito K., and Nakamura T. (1997) Cloning and sequencing of a nitrile hydratase gene from *Pseudonocardia thermophila* JCM3095. *Journal of Fermentation and Bioengineering.* **83:** 474-477.

Yamamoto K., Ueno Y., Otsubo K., Yamane H., Komatsu K., and Tani Y. (1992) Efficient conversion of dinitrile to mononitrile-monocarboxylic acid by *Corynebacterium* sp. C5 cells during tranexamic acid synthesis. *Journal of Fermentation and Bioengineering.* **73:** 125-129.

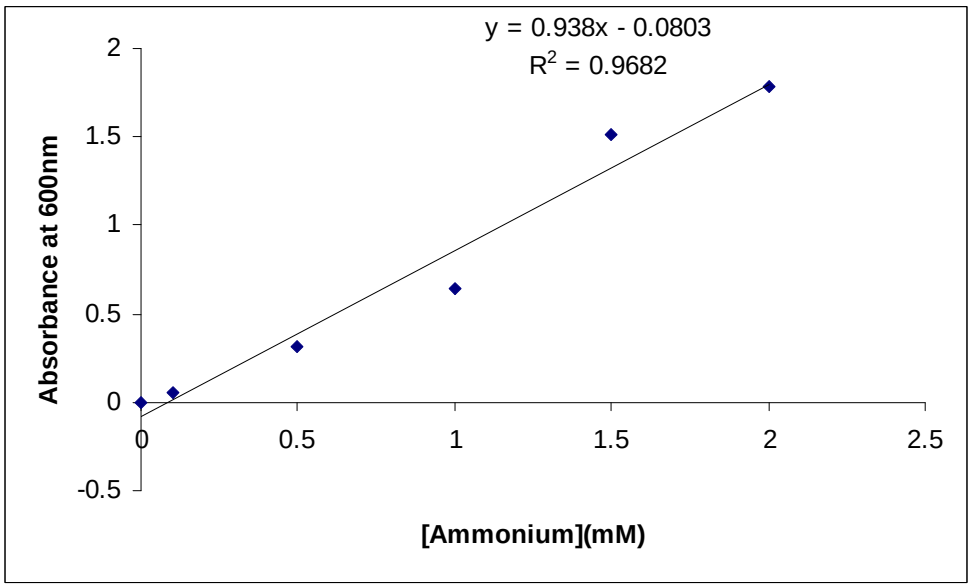
Zaks A. (2001) Industrial biocatalysis. *Current Opinion in Chemical Biology*. **5:** 130-6.

Appendices

Appendix 1 *Biorad protein assay standard curve*



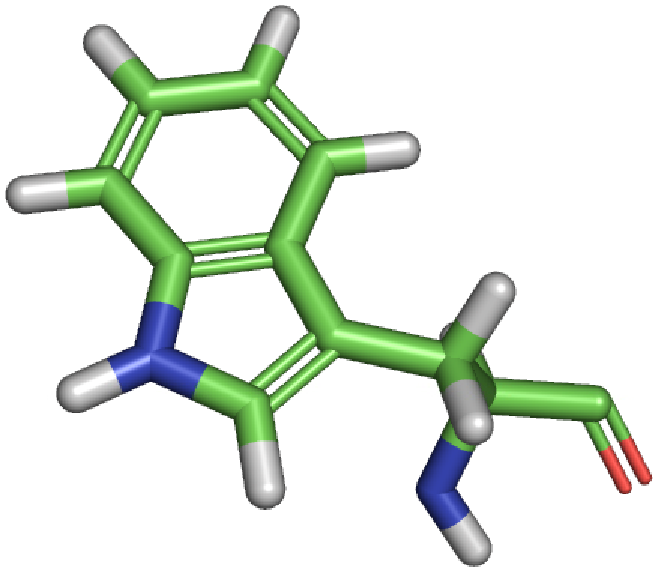
Appendix 2 *Ammonia assay standard curve*



Appendix 3 *Ammonium sulphate fractionation*

	Final concentration of ammonium sulphate, % saturation																
	10	20	25	30	33	35	40	45	50	55	60	65	70	75	80	90	100
	Grams solid ammonium sulphate to be added to 1L. of solution																
0	56	114	144	176	196	209	243	277	313	351	390	430	472	516	561	662	767
10		57	86	118	137	150	183	216	251	288	326	365	406	449	494	592	694
20			29	59	78	91	123	155	189	225	262	300	340	382	424	520	619
25				30	49	61	93	125	158	193	230	267	307	348	390	485	583
30					19	30	62	94	127	162	198	235	273	314	356	449	546
33						12	43	74	107	142	177	214	252	292	333	426	522
35							31	63	94	129	164	200	238	278	319	411	506
40								31	63	97	132	168	205	245	285	375	469
45									32	65	99	134	171	210	250	339	431
50										33	66	101	137	176	214	302	392
55											33	67	103	141	179	264	353
60												34	69	105	143	227	314
65													34	70	107	190	275
70														35	72	153	237
75															36	115	198
80																77	157
90																	79

Appendix 5 *Notation used for description of atoms in amino acids*



Tryptophan